Análise de modelo de Hopfield
com topologia de rede complexa

Fabiano Berardo de Sousa

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Análise de modelo de Hopfield com topologia de rede complexa

Fabiano Berardo de Sousa

***Orientador:* Prof. Dr. Zhao Liang**

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências - Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

USP – São Carlos
Janeiro de 2014

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

S725a Sousa, Fabiano Berardo de
 Análise de modelo de Hopfield com topologia de
 rede complexa / Fabiano Berardo de Sousa;
 orientador Liang Zhao. -- São Carlos, 2013.
 78 p.

 Dissertação (Mestrado - Programa de Pós-Graduação
 em Ciências de Computação e Matemática
 Computacional) -- Instituto de Ciências Matemáticas
 e de Computação, Universidade de São Paulo, 2013.

 1. Redes Neurais Artificiais. 2. modelo de
 Hopfield. 3. Redes Neurais Caóticas. 4. Redes
 Complexas. 5. Sistemas Dinâmicos. I. Zhao, Liang,
 orient. II. Título.

À minha família.

Agradecimentos

Agradeço ao meu orientador Zhao Liang pela oportunidade e pela ótima orientação, à minha família pelo apoio emocional, à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo apoio financeiro.

Resumo

Redes neurais biológicas contêm bilhões de células (neurônios) agrupadas em regiões espacial e funcionalmente distintas. Elas também apresentam comportamentos complexos, tais como dinâmicas periódicas e caóticas. Na área da Inteligência Artificial, pesquisas mostram que *Redes Neurais Caóticas*, isto é, modelos de Redes Neurais Artificiais que operam com dinâmicas complexas, são mais eficientes do que modelos tradicionais no que diz respeito a evitar memórias espúrias. Inspirado pelo fato de que o córtex cerebral contém agrupamentos de células e motivado pela eficiência no uso de dinâmicas complexas, este projeto de pesquisa investiga o comportamento dinâmico de um modelo de Rede Neural Artificial Recorrente, como o de Hopfield, porém com a topologia sináptica reorganizada a ponto de originar agrupamentos de neurônios, tal como acontece em uma Rede Complexa quando esta apresenta uma estrutura de comunidades. O modelo de treinamento tradicional de Hopfield também é alterado para uma regra de aprendizado que posta os padrões em ciclos, gerando uma matriz de pesos assimétrica. Resultados indicam que o modelo proposto oscila entre comportamentos periódicos e caóticos, dependendo do grau de fragmentação das sinapses. Com baixo grau de fragmentação, a rede opera com dinâmica periódica, como consequência da regra de treinamento utilizada. Dinâmicas caóticas parecem surgir quando existe um alto grau de fragmentação. Mostra-se, também, que é possível obter caoticidade em uma topologia adequadamente modular, ou seja, como uma estrutura de comunidades válida. Desta forma, este projeto de pesquisa provê uma metodologia alternativa para se construir um modelo de Rede Neural Artificial que realiza tarefas de reconhecimento de padrões, explorando dinâmicas complexas por meio de uma estrutura de conexões que se mostra mais similar à topologia existente no cérebro.

Palavras-chave: redes neurais artificiais, modelo de hopfield, redes neurais caóticas, redes complexas, sistemas dinâmicos.

Abstract

Biological neural networks contain billions of neurons divided in spatial and functional clusters to perform different tasks. It also operates with complex dynamics such as periodic and chaotic ones. It has been shown that Chaotic Neural Networks are more efficient than conventional recurrent neural networks in avoiding spurious memory. Inspired by the fact that the cerebral cortex has specific groups of cells and motivated by the efficiency of complex behaviors, in this document we investigate the dynamics of a recurrent neural network, as the Hopfield one, but with neurons coupled in such a way to form a complex network community structure. Also, we generate an asymmetric weight matrix placing pattern cycles during learning. Our study shows that the network can operate with periodic and chaotic dynamics, depending on the degree of the connection's fragmentation. For low fragmentation degree, the network operates with periodic dynamic due to the employed learning rule. Chaotic behavior seems to rise for a high fragmentation degree. We also show that the neural network can hold both chaotic dynamic and a high value of modularity measure at the same time, indicating an acceptable community structure. These findings provide an alternative way to design dynamical neural networks to perform pattern recognition tasks exploiting periodic and chaotic dynamics by using a more similar topology to the topology of the brain.

Keywords: artificial neural networks, hopfield model, chaotic neural networks, complex networks, dynamical systems.

Sumário

Lista de Figuras	xi
1 Introdução	1
1.1 Contextualização	1
1.2 Objetivos e Motivações	2
1.3 Organização do Documento	3
2 Tópicos Fundamentais	5
2.1 Sistemas Dinâmicos	5
2.1.1 Dinâmica Estacionária	7
2.1.2 Dinâmica Periódica	8
2.1.3 Dinâmica Caótica	10
2.1.4 Estabilidade de Pontos Fixos em Sistemas Dinâmicos Não-Lineares	11
2.1.5 Análise do Comportamento de Sistemas Dinâmicos	14
2.2 Algumas Características das Redes Neurais Biológicas	23
2.2.1 A Estrutura Básica do Cérebro	23
2.2.2 A Neurodinâmica	25
2.2.3 O Aprendizado Hebbiano	27
2.3 O Modelo de Hopfield	28
2.4 Redes Complexas	32
3 Extensões de Redes Recorrentes	39
3.1 Redes Celulares	40
3.2 Rede Hopfield com Topologia Livre de Escala	43
3.3 Redes Neurais Caóticas	43

4	O Modelo de Rede Recorrente com Estrutura de Comunidades	53
4.1	Metodologia	53
4.1.1	Geração da Topologia de Rede	54
4.1.2	Regra de Aprendizado	55
4.1.3	Regra de Propagação	56
4.1.4	Análise do Comportamento Dinâmico	57
4.2	Análise do Modelo	57
4.2.1	Cálculo do Número de Conexões	57
4.2.2	Relação entre Parâmetros na Condição de Bifurcação	58
4.3	Experimentos e Resultados	60
4.3.1	Verificação do Comportamento Periódico	60
4.3.2	Análise do Surgimento do Comportamento Caótico	63
4.3.3	Análise do Comportamento Dinâmico no Contexto de Comunidades	63
5	Considerações Finais	69
5.1	Conclusões	69
5.2	Contribuições	70
5.3	Trabalhos Futuros	71
	Bibliografia	73

Lista de Figuras

2.1	Representação da dinâmica estacionária do mapa logístico da Equação 2.4 quando $a = 2,75$ e $x_0 = 0,50$	8
2.2	Representação da dinâmica periódica do mapa logístico da Equação 2.4 com $a = 3,20$ e $x_0 = 0,50$. Nota-se uma dinâmica de período 2.	9
2.3	Representação da dinâmica periódica do mapa logístico da Equação 2.4 com $a = 3,50$ e $x_0 = 0,50$. Neste caso, há uma dinâmica de período 4.	9
2.4	Representação da dinâmica caótica do mapa logístico da Equação 2.4 com $a = 3,95$ e $x_0 = 0,50$. Nota-se um comportamento errático não-periódico.	11
2.5	Diagrama de bifurcação que apresenta o comportamento dinâmico do mapa logístico da Equação 2.4 em relação aos valores assumidos pelo coeficiente de bifurcação a	15
2.6	Expoentes de Lyapunov λ do mapa logístico segundo a Equação 2.18 em relação aos valores assumidos pelo coeficiente de bifurcação a	16
2.7	Expoentes de Hurst H do mapa logístico da Equação 2.4 em relação aos valores do coeficiente de bifurcação a . O espectro exhibe dinâmicas periódicas (valores próximos a 0) e caóticos (valores próximos a 0,5). Nota-se que não existe expoente para os valores de a em que o mapa apresenta uma dinâmica estacionária. . .	19
2.8	Coeficientes de Fourier c da série temporal com dinâmica estacionária exibida pela Figura 2.1. A curva constante salienta um comportamento estacionário. . .	21
2.9	Coeficientes de Fourier c da série temporal com dinâmica periódica de período 2 exibida pela Figura 2.2. O pico ao centro indica um comportamento periódico. . .	22
2.10	Coeficientes de Fourier c da série temporal com dinâmica periódica de período 4 exibida pela Figura 2.3. Notam-se 3 picos sobressalentes, indicando que há um relacionamento entre o número de picos e o número de períodos da série. . . .	22

2.11	Coeficientes de Fourier c da série temporal com dinâmica caótica exibida pela Figura 2.4. A curva irregular reforça a indicação de um comportamento caótico.	22
2.12	Mapa do córtex cerebral evidenciando suas diversas regiões estruturais e funcionais. As áreas consideradas mais importantes são: 1, 2 e 3: córtex somatossensorial; 4. córtex motor; 6. córtex pré-motor; 17, 18 e 19: córtex visual; 41 e 42: córtex auditivo. Fonte: Haykin 1999, p. 10.	25
2.13	Exemplo de uma Rede Complexa: rede de relações entre membros do clube de karatê Zachary [Zachary 1977].	33
2.14	Tipos de Redes Complexas: (a) Rede Aleatória, (b) Rede Pequeno-Mundo, (c) Rede Livre de Escala.	35
2.15	Estrutura de comunidades da rede de relações entre membros do clube de karatê exibida na Figura 2.13. 5 comunidades são observadas.	36
3.1	Expoentes de Lyapunov para a Equação 3.16 pela variação do parâmetro $\theta(t) = a$, com $k = 0,5$, $\epsilon = 0,015$ e $\alpha = 1,0$. Valores positivos de λ indicam dinâmica caótica; valores negativos indicam outras dinâmicas; valores iguais a zero indicam pontos de bifurcação. Fonte: Adachi e Aihara 1997.	46
3.2	Diagrama de bifurcação da Equação 3.16. Gráfico do estado interno u do neurônio pela variação do parâmetro $\theta(t) = a$, com $k = 0,5$, $\epsilon = 0,015$ e $\alpha = 1,0$. Fonte: Adachi e Aihara 1997.	46
3.3	Base utilizada na análise do modelo proposto. Padrões de (a) a (d) são armazenados e o padrão (e) utilizado para recuperação associativa dos padrões armazenados. Fonte: Adachi e Aihara 1997.	48
3.4	Demonstração da dinâmica estacionária para $\alpha = k_f = k_r = 0$. O padrão visto em $t = 0$ converge para o padrão armazenado (c) da Figura 3.3. Fonte: Adachi e Aihara 1997.	49
3.5	Saídas da rede durante o ciclo de recuperação de padrão para $\alpha = 8,0$, $k_f = 0,15$ e $k_r = 0,5$. Nota-se um comportamento periódico ou quase-periódico, no qual a rede, algumas vezes, aproxima-se do padrão (c) da Figura 3.3, porém não o mantém estável. Fonte: Adachi e Aihara 1997.	49
3.6	Saídas da rede durante o ciclo de recuperação de padrão para $\alpha = 10,0$, $k_f = 0,2$ e $k_r = 0,9$. Nota-se uma trajetória semelhante a de uma itinerância caótica, na qual cada um dos quatro padrões, ora ou outra, é visitado pela rede. Fonte: Adachi e Aihara 1997.	51

4.1	Exemplo da topologia do modelo de RNA proposto para $N = 100$, $C = 4$, $P_{in} = 1,00$ e $P_{out} = 0,10$. As comunidades contêm 25 neurônios cada. A Medida de Modularidade para esta topologia é $Q = 0,452569$	55
4.2	Organização de 10 padrões ($M = 10$) em 2 ciclos ($K = 2$). Cada ciclo contém 5 diferentes padrões ($L = 5$).	56
4.3	Curva da relação entre P_{out} e C , dada pela Equação 4.10.	59
4.4	Base de padrões: imagens binárias de 10×10 pixels. (a) mostra os padrões a serem armazenados. O padrão em (b) é utilizado na fase de propagação.	60
4.5	Evolução da rede para o padrão apresentado na Figura 4.4(b). Os padrões da Figura 4.4(a) são organizados em um ciclo ($K = 1$ e $L = 4$). Nota-se um comportamento periódico de período 4, no qual a rede converge para o ciclo único.	61
4.6	Série temporal do estado interno u referente à evolução exibida pela Figura 4.5. Reforça-se a observação de um comportamento periódico de período 4.	61
4.7	Evolução da rede para o padrão apresentado na Figura 4.4(b). Os padrões da Figura 4.4(a) são organizados em 2 ciclo, 2 padrões em cada ciclo ($K = 2$ e $L = 2$). Desta vez, nota-se um comportamento periódico de período 2, no qual a rede converge para o primeiro dos dois ciclos.	62
4.8	Série temporal do estado interno u referente à evolução exibida na Figura 4.7. Reforça-se a observação de um comportamento periódico de período 2.	62
4.9	Análise do comportamento dinâmico da rede considerando uma base de 30 padrões divididos em 5 ciclos, 6 padrões por ciclo ($K = 5$ e $L = 6$). Em (a) é exibido o diagrama de bifurcação do estado interno u em relação a P_{in} , para $C = 1$ e $P_{out} = 1$. (b) mostra o espectro dos expoentes de Hurst H para a mesma evolução. Nota-se uma região de bifurcação para P_{in} em torno de 0,20.	64
4.10	Comportamento da rede para $P_{in} = 0,30$, $C = 1$ e $P_{out} = 1$. (a) mostra uma série temporal periódica de período 6 que é ratificado pelos coeficientes de Fourier c em (b). O expoente de Hurst para a série é $H = 0,060113$, valor que também indica um comportamento periódico.	65
4.11	Análise do comportamento da rede para $P_{in} = 0,10$, $C = 1$ e $P_{out} = 1$. Em (a) é exibida a série temporal e em (b) os relacionados coeficientes de Fourier. O expoente de Hurst da série é $H = 0,493295$. Este dados sugerem um comportamento caótico.	66
4.12	Variação do valor crítico de P_{out} em relação ao número C de comunidades. Nota-se que não há alteração significativa no valor de P_{out} para $C \geq 50$	66

- 4.13 Comparação entre as curvas dadas pela Figura 4.12 e pela Equação 4.10 quando $N = 400$, $p^{(c)} = 0,20$ e $\theta = 0,202$. Observa-se que o resultado experimental é compatível com o resultado teórico. 67
- 4.14 Variação da Medida de Modularidade Q pelo número de C , considerando $P_{out} = 0,05$ e $P_{in} = 1$. (a) mostra a curva completa e (b) magnifica a região na qual $Q \geq 0,3$. É indicado que a rede possui uma estrutura de comunidade válida para $2 \leq C \leq 29$ 67
- 4.15 Comportamento dinâmico da rede para $C = 20$, $P_{out} = 0,05$ e $P_{in} = 1$. (a) mostra a série temporal e (b) os coeficientes de Fourier. Para esta série, a Modularidade é $Q = 0,387818$ e o expoente de Hurst é $H = 0,587774$. Este resultado indica que a rede pode comportar-se caoticamente e, ao mesmo tempo, exibir uma estrutura de comunidades válida. 68

Introdução

Este capítulo discute, na Seção 1.1, o contexto de pesquisa no qual este projeto se insere. A Seção 1.2 disserta sobre as motivações e objetivos gerais. A organização do restante deste documento é comentada na Seção 1.3.

1.1 Contextualização

O cérebro é considerado o principal órgão responsável pelo processamento da informação em animais dotados de um sistema nervoso central. Pode ser encarado como um processador de sinais que explora um paralelismo massivo por meio de um número muito vasto de elementos processadores chamados neurônios. Estas células nervosas conectam-se umas as outras por meio de substâncias químicas (os neurotransmissores) formando as sinapses neurais, as quais permitem a continuidade do sinal elétrico que trafega pela célula. A existência, a quantidade e o tipo das sinapses é o que dá ao cérebro sua capacidade de armazenar, processar, recuperar e, principalmente, de generalizar informações (aprendizado) [McClelland e Rumelhart 1986].

Inspiradas na estrutura do cérebro humano, as Redes Neurais Artificiais (RNAs) representam uma tentativa computacional para explorar os mecanismos de processamento e memória presentes nos sistemas neurais biológicos. Tratam-se, portanto, de sistemas de processamento paralelo e distribuído compostos por unidades de processamento simples (neurônios artificiais) que calculam determinadas funções matemáticas, normalmente não-lineares, chamadas funções de ativação. Tais unidades são dispostas em uma ou mais camadas e interligadas por um certo número de conexões (sinapses), normalmente unidirecionais. Na maioria dos modelos, as conexões estão associadas a pesos, os quais servem para armazenar o conhecimento adquirido e ponderar a entrada recebida por cada neurônio da rede [Haykin 1999].

Tal como o cérebro, RNAs são capazes de: (i) aprender por meio de exemplos, (ii) generalizar

a informação armazenada e (iii) reconhecer padrões, tornando-se importante ferramenta na área de aprendizado de máquina. Há muitos modelos diferentes de RNAs, cada qual baseado em diferentes metodologias e aplicado para fins específicos. Um modelo que particularmente interessa a este projeto é o modelo proposto por Hopfield em 1982, por tratar-se de uma Rede Neural Recorrente que funciona como uma memória associativa, como o cérebro, e que opera com um comportamento dinâmico estacionário, conforme discutido na Seção 2.3. Por outro lado, tem-se demonstrado que o cérebro dificilmente opera com dinâmicas estacionárias, mas sim complexas (como periódicas e caóticas). Neste sentido, as chamadas *Redes Neurais Caóticas* surgem como modelos recorrentes alternativos e com um comportamento dinâmico mais similar ao funcionamento do cérebro. Além disso, tem-se demonstrado que o uso de dinâmicas complexas pode ajudar a superar algumas das limitações de modelos recorrentes tradicionais como, por exemplo, a existência de memórias ditas indesejáveis, ou *espúrias*.

Além do comportamento dinâmico complexo, as redes biológicas também apresentam uma topologia extremamente complexa e difícil de ser reproduzida por modelos artificiais. Há um número muito grande de neurônios e sinapses que permanecem em constante evolução e se organizam em blocos funcionais especializados, cada qual responsável por uma atividade específica. Neste sentido, uma área de pesquisa que atualmente fornece o ferramental teórico e prático para lidar com tais estruturas de grande porte é a das chamadas *Redes Complexas*.

Na área da ciência da computação, tem-se visto um número cada vez maior de pesquisas relacionadas ao estudo das redes complexas, dado ao aumento do poder computacional nas últimas décadas e, principalmente, à sua relevância multidisciplinar. Lidando com quantidades muito vastas de unidades e conexões, as redes complexas modelam perfeitamente vários sistemas dinâmicos reais (como redes metabólicas, redes sociais, redes de citações, redes de computadores - como a Internet - e, neste caso particular de interesse, as redes neurais biológicas) sendo sua teoria aplicada, seja para estudar e resolver problemas específicos da área, seja para extrair informações relevantes destes sistemas [Albert e Barabási 2002], [Boccaletti et al. 2006], [Newman 2003].

1.2 Objetivos e Motivações

Inspirado pelo fato de que o córtex cerebral apresenta dinâmica e topologia complexas, este projeto analisa o comportamento de um modelo de Hopfield porém com seus neurônios e conexões reestruturados de forma a apresentarem a estrutura topológica de comunidades de uma rede complexa. Desta forma, é proposto um novo modelo de Rede Neural Caótica cuja dinâmica caótica surge como consequência da estrutura de comunidades. Trata-se, portanto de

um modelo inédito cuja dinâmica e estrutura topológica são mais similares ao cérebro do que outros modelos de RNAs.

Assim sendo, os seguintes objetivos são definidos:

- Gerar um modelo de RNA recorrente, baseado no modelo de Hopfield, mas que apresente uma estrutura de comunidades válida.
- Analisar o comportamento dinâmico do modelo proposto em diferentes configurações da estrutura de comunidades.
- Avaliar o grau de modularidade nas diferentes configurações da estrutura de comunidades utilizadas no item anterior.

Tais objetivos são revisitados e aprofundados no Capítulo 4.

1.3 Organização do Documento

Este documento está organizado da seguinte forma: no Capítulo 2, apresenta-se uma revisão bibliográfica sobre conceitos fortemente relacionados com este projeto, como Sistemas Dinâmicos, funcionamento básico do cérebro e neurodinâmica, o modelo de Hopfield e Redes Complexas. No Capítulo 3, abordam-se alguns trabalhos de pesquisa, como Redes Celulares, Rede de Hopfield com topologia de Rede Livre de Escala e Redes Neurais Caóticas, os quais, de forma direta ou indireta, relacionam-se com o modelo de Hopfield e com este projeto. O Capítulo 4 discute a metodologia e os resultados obtidos. Por fim, as conclusões e contribuições são dissertadas no Capítulo 5.

Tópicos Fundamentais

São apresentados neste capítulo os principais conceitos relacionados com o tema de pesquisa deste trabalho. O capítulo inicia-se abordando conceitos básicos sobre a teoria de Sistemas Dinâmicos na Seção 2.1. Na Seção 2.2 são descritas algumas características das redes neurais biológicas, as quais embasam as motivações deste projeto de pesquisa. O modelo de RNA de Hopfield, proposto em 1982, é discutido na Seção 2.3. O capítulo encerra-se com a abordagem, ainda que em caráter introdutório, dos conceitos e características das Redes Complexas na Seção 2.4.

2.1 Sistemas Dinâmicos

Sistemas Dinâmicos são sistemas que evoluem no tempo e essa evolução pode ser descrita por modelos matemáticos baseados em equações diferenciais (para sistemas cujo tempo é considerado contínuo) ou equações de diferenças (para sistemas cujo tempo é considerado discreto) [Ott 1993]. Sistemas dinâmicos contínuos são também chamados de *fluxo* e sistemas dinâmicos discretos de *mapa*. Tais modelos não se preocupam em obter soluções explícitas para as equações, mas sim em tentar prever o comportamento do sistema a longo prazo, medindo propriedades como: estabilidade mediante perturbações, ocorrência de mudanças (bifurcações), entre outras.

A teoria de sistemas dinâmicos teve início a partir da invenção do Cálculo Diferencial, por Newton e Leibnitz, utilizado principalmente para interpretar problemas relacionados com a mecânica celeste, como o problema dos três corpos [Strogatz 1989]. Hoje em dia, as equações diferenciais são ferramenta indispensável para estudar e descrever os mais diversos fenômenos que evoluem no tempo, sejam físicos, econômicos, sociais, biológicos, entre outros, tais como a dinâmica cerebral (também chamada de *Neurodinâmica*), cujo primeiro trabalho conhecido

remete a Nicholas Rashevsk em 1938, de acordo com Haykin 1999.

A característica fundamental de um sistema dinâmico é a retroalimentação (recorrência, ou *feedback*), ou seja, o estado do sistema em um dado instante de tempo depende do estado do sistema no instante de tempo anterior (configurando uma espécie de memória). Matematicamente, um sistema dinâmico de tempo discreto pode ser descrito pela equação de diferenças como segue:

$$x(t+1) = f(x(t)), \quad (2.1)$$

onde t é o tempo (discreto), x é a variável independente que denota o estado do sistema e f é a função que descreve como o sistema evolui no tempo. Por outro lado, se x assumir valores contínuos, o sistema é descrito pelas seguintes equações diferenciais:

$$\begin{cases} \frac{dx_1}{dt} = f_1(x_1(t), x_2(t), \dots, x_n(t)) \\ \frac{dx_2}{dt} = f_2(x_1(t), x_2(t), \dots, x_n(t)) \\ \vdots \\ \frac{dx_n}{dt} = f_n(x_1(t), x_2(t), \dots, x_n(t)), \end{cases} \quad (2.2)$$

onde n é a dimensão do sistema. Em formato vetorial, o mesmo sistema pode ser representado como:

$$\begin{cases} \dot{x} = f(x(t)) \\ x(t_0) = x_0, \end{cases} \quad (2.3)$$

onde x é o vetor de estados, $\dot{x}$ significa dx/dt e x_0 é o estado inicial.

Para qualquer estado inicial x_0 , o sistema alcançará um estado futuro $x(t, t_0, x_0)$, para $t > 0$. O caminho gerado pela mudança contínua de estado no espaço de estados é conhecido como *órbita* ou *trajetória*. Se f é uma função diferenciável, então existe uma única solução para a Equação 2.3 dada uma condição inicial x_0 [Ott 1993]. De acordo com o comportamento da função f , um sistema dinâmico pode ser classificado como linear (se f for linear), ou não-linear (se f for não-linear). Vale a pena mencionar que observa-se na natureza que a maioria dos fenômenos, ou leis, comportam-se de maneira não-linear.

Em relação à análise a longo prazo, um dos principais objetos de interesse da teoria de sistemas dinâmicos é o estudo do comportamento assintótico do sistema, ou seja, o comportamento quando $t \rightarrow \infty$ [Devaney 1989]. Neste sentido, um sistema dinâmico pode apresentar diversos tipos de *atratores*, ou seja, um ponto, ou uma região no espaço de estados para onde as trajetórias convergem a partir de determinadas condições iniciais. Os atratores mais comuns são: estacionário (ou *ponto fixo*), ciclo limite (ou *periódico*), quase-periódico e caótico. O período

de tempo que o sistema leva até atingir a convergência para um atrator é normalmente descrita como *fase de transiência* (ou simplesmente *transiente*).

As próximas subseções discutem mais aprofundadamente alguns destes atratores. Como resultado de simulações computacionais, estas subseções apresentam séries temporais obtidas para o *mapa logístico*, dado pela Equação 2.4. Trata-se de um mapa unidimensional e não-linear que emergiu na década de 70 para cálculos demográficos e que, atualmente, é largamente utilizado como fonte para demonstração de comportamentos dinâmicos, pois, pela variação de apenas um parâmetro, é possível obter dinâmicas distintas e facilmente observáveis.

$$x(t+1) = ax(t)(1-x(t)), \quad (2.4)$$

onde t é o tempo (discreto), x é a variável de estado e a é dito *parâmetro de bifurcação*. De acordo com o valor de a , pode-se observar as diversas dinâmicas abordadas nas subseções a seguir.

2.1.1 Dinâmica Estacionária

Um *ponto fixo* (também chamando de *ponto de equilíbrio* para sistemas de tempo contínuo) é um estado do espaço de estados que, uma vez atingido, o sistema nele permanece ao decorrer do tempo. Em outras palavras, ao atingir-se um ponto fixo, o sistema retorna sempre o mesmo valor. Matematicamente, um dado ponto x^* é considerado um ponto fixo se:

$$f(x^*) = x^*. \quad (2.5)$$

No caso do sistema dinâmico logístico apresentado na Equação 2.4, o ponto fixo representa uma solução constante dada por:

$$x(t+1) = x(t). \quad (2.6)$$

Na Figura 2.1 é ilustrada a série temporal (ou seja, o comportamento da trajetória ao longo do tempo) do mapa logístico, considerando-se o valor do parâmetro de bifurcação $a = 2,75$ e condição inicial $x_0 = 0,50$. Nota-se que quando $t \rightarrow \infty$, o valor de $x(t+1)$ permanece constante em 0,636364, evidenciando que o sistema converge para um ponto atrator.

Um ponto fixo é dito *estável* se soluções que têm início em uma certa vizinhança próxima deste, permanecem nesta mesma vizinhança conforme o tempo prossegue. De maneira mais formal, definições sobre a estabilidade de pontos fixos são dadas a seguir.

Definição: Um ponto fixo x^* é *estável* se, para toda vizinhança N_ϵ de x^* , existe uma

Figura 2.1: Representação da dinâmica estacionária do mapa logístico da Equação 2.4 quando $a = 2,75$ e $x_0 = 0,50$.

vizinhança N_δ em N_ϵ , tal que toda solução $x(t)$, com $x(0)$ em N_δ , é definida e permanece em N_ϵ para todo $t > 0$.

Definição: Se um ponto fixo x^* é estável e se toda solução $x(t)$ tende a x^* (ou seja: $\lim_{t \rightarrow \infty} x(t) = x^*$), então x^* é dito *assintoticamente estável*.

Definição: Se existe ao menos uma vizinhança N_ϵ de x^* tal que, para toda vizinhança N_δ em N_ϵ , existe ao menos uma solução $x(t)$, com início $x(0)$ em N_δ , que não pertence a N_ϵ , então x^* é *instável*.

Um critério para determinação da estabilidade de pontos fixos pode ser estabelecido pelas seguintes condições:

(a) Se $|f'(x^*)| < 1$, então x^* é assintoticamente estável.

(b) Se $|f'(x^*)| > 1$, então x^* é instável.

(observa-se que no caso $|f'(x^*)| = 1$, maiores análises são necessárias).

Uma análise mais profunda sobre a estabilidade de pontos fixos, considerando-se o caso de sistemas não-lineares, é apresentada na Subseção 2.1.4.

2.1.2 Dinâmica Periódica

Tem-se um atrator do tipo periódico quando o comportamento do sistema se repete ao longo do tempo, ou seja, há uma repetição sucessiva de valores quando $t \rightarrow \infty$. O número de iterações entre as repetições é chamado de *período*. Um ciclo de período p ocorre se:

$$\begin{cases} x(t+p) = x(t) \\ x(t+j) \neq x(t), \quad j = 1, 2, \dots, p-1. \end{cases} \quad (2.7)$$

Outra forma de se representar um comportamento periódico é:

$$f^p(x) = x, \quad (2.8)$$

onde f^p representa a p -ésima iteração da função f , ou seja: $f(f(f(\dots)))$ p vezes. Neste caso, diz-se que há uma dinâmica de *período* p .

Considerando um valor de $a = 3,20$ para mapa logístico dado pela Equação 2.4, existe um comportamento periódico de período 2, o qual é ilustrado pela série temporal da Figura 2.2. Já a Figura 2.3 ilustra o comportamento periódico de período 4 quando $a = 3,50$.

Figura 2.2: Representação da dinâmica periódica do mapa logístico da Equação 2.4 com $a = 3,20$ e $x_0 = 0,50$. Nota-se uma dinâmica de período 2.

Figura 2.3: Representação da dinâmica periódica do mapa logístico da Equação 2.4 com $a = 3,50$ e $x_0 = 0,50$. Neste caso, há uma dinâmica de período 4.

A estabilidade de atratores periódicos é definida em função da análise da estabilidade dos

pontos que compõem a trajetória dentro do ciclo. Um critério para a determinação da estabilidade de ciclos periódicos pode ser estabelecido como uma extensão do critério utilizado para pontos fixos. Seja $\Omega = \{x(0), x(1), \dots, x(p-1)\}$ um ciclo periódico de período p . Então:

(a) Se $|f'(x(0)) \cdot f'(x(1)) \cdot \dots \cdot f'(x(p-1))| < 1$, então Ω é assintoticamente estável.

(b) Se $|f'(x(0)) \cdot f'(x(1)) \cdot \dots \cdot f'(x(p-1))| > 1$, então Ω é instável.

(da mesma forma observada para o critério da estabilidade de pontos fixos, caso $|f'(x(0)) \cdot f'(x(1)) \cdot \dots \cdot f'(x(p-1))| = 1$, uma análise mais detalhada do sistema será necessária).

Existe também um tipo de dinâmica definida como *quase-periódica*, na qual o sistema exibe uma dinâmica recorrente, porém com comportamento imprevisível, isto é, impossível de ser medido. O mapa logístico da Equação 2.4 não apresenta esse tipo de comportamento. Quando um sistema não apresenta recorrência para dado valor de parâmetro, tem-se uma dinâmica dita ser *não-periódica*. Dinâmicas caóticas são, por definição, não-periódicas, conforme apresentado na Subseção a seguir.

2.1.3 Dinâmica Caótica

Um dos primeiros relatos sobre dinâmica caótica em equações diferenciais remete aos trabalhos do matemático, filósofo e pesquisador francês Henry Poincaré, no século XIX. A demonstração de tal fenômeno, entretanto, era impraticável na época, dada à complexidade dos cálculos de uma trajetória do sistema. Com o advento da computação digital, o assunto reemergiu graças a Edward Lorez que observou o mesmo fenômeno em simulações computacionais de eventos meteorológicos. Dadas condições iniciais muito próximas, o sistema apresentava, aparentemente, o mesmo resultado nas primeiras iterações, porém resultados cada vez mais divergentes a cada iteração seguinte. Foi Lorez que popularizou a famosa ideia do *efeito borboleta*, no qual o simples bater de asas de uma borboleta em um canto do mundo pode provocar uma tempestade no canto oposto, em analogia às consequências de perturbações nas condições iniciais do sistema.

O *caos* pode ser definido como uma dinâmica limitada, não-periódica em um sistema determinístico com grande sensibilidade a condição inicial [Kaplan e Glass 1995]. *Limitada* pois a variável de estado permanece sempre restrita a um domínio fechado, nunca se aproximando de $\pm\infty$. *Não-periódica* pois um estado nunca é repetido mais de uma vez. *Determinístico* pois o sistema não é governado por termos randômicos, mas por uma lei de evolução bem definida. *Grande sensibilidade a condição inicial*, pois pontos inicialmente bem próximos terão

trajetórias totalmente distintas com o decorrer do tempo. Esta última propriedade é um dos aspectos essenciais da Teoria do Caos.

Particularmente, o caos descreve um tipo complexo de dinâmica presente em sistemas dinâmicos não-lineares. E este comportamento caótico provê aos pesquisadores um extraordinário problema de previsibilidade, no qual previsões a curto prazo são muito difíceis e, a longo prazo, impossíveis, dada à divergência exponencial de trajetórias próximas. Logo, o comportamento de um sistema caótico a longo prazo é indeterminado, a não ser que se conheça o valor exato da condição inicial. Um exemplo de dinâmica caótica pode ser observado na Figura 2.4 que apresenta a dinâmica do mapa logístico dado pela Equação 2.4 com valor de $a = 3,95$.

Figura 2.4: Representação da dinâmica caótica do mapa logístico da Equação 2.4 com $a = 3,95$ e $x_0 = 0,50$. Nota-se um comportamento errático não-periódico.

2.1.4 Estabilidade de Pontos Fixos em Sistemas Dinâmicos Não-Lineares

Esta Subseção visa aprofundar a análise da estabilidade de pontos fixos, considerando-se o caso particular de sistemas dinâmicos não-lineares. Uma das principais ferramentas utilizadas para esta análise são os métodos *Direto* e *Indireto* propostos pelo matemático e engenheiro russo Aleksandr Mikhailovich Lyapunov. Tais abordagens exigem uma melhor formalização do conceito de estabilidade em relação ao conteúdo apresentado na Subseção 2.1.1. Assim, seja x^* um ponto fixo, então:

- (a) x^* é *estável*, no sentido de Lyapunov, se, dados $\epsilon > 0$ e $t_0 \in [0, \infty)$, existe um $\delta(\epsilon, t_0) > 0$, tal que $\|x(t_0) - x^*\| < \delta$ implica que $\|f^n(x(t_0)) - x^*\| < \epsilon$, para todo $t > t_0$.
- (b) x^* é *assintoticamente estável*, no sentido de Lyapunov, se, além de ser estável, existe $\eta(t_0) > 0$, tal que $\|x(t_0) - x^*\| < \eta$ implica que $\lim_{t \rightarrow \infty} x(t) = x^*$ (comportamento

chamado de *convergência*).

(c) x^* é *instável* se ele não for estável.

A seguir, são descritos, respectivamente, o *Método Indireto* e o *Método Direto* de Lyapunov, considerando-se as definições de estabilidade conforme apresentadas nos itens (a), (b) e (c) acima.

O Método Indireto de Lyapunov (Linearização)

Este método basicamente opera transformações matemáticas na função f não-linear do sistema dinâmico a fim de que sua estabilidade seja analisada da mesma forma que seria feito caso f fosse linear. Para tanto, considere o sistema dinâmico autônomo e não-linear:

$$\dot{x} = f(x), \quad (2.9)$$

supondo f uma função continuamente diferenciável. Considerando-se que o sistema possui um ponto fixo x^* e expandindo f na série de Taylor em torno de x^* , tem-se:

$$f(x) = f(x^*) + Df(x^*)(x - x^*) + g(x), \quad (2.10)$$

onde $g(x)$ representa os termos de ordem superior da expansão e $Df(x^*)$ é a matriz Jacobiana que contém as derivadas parciais de f em x^* , calculada como segue:

$$Df(x^*) = \left. \frac{\partial f(x)}{\partial x} \right|_{x=x^*} \quad (2.11)$$

Assim, a Equação 2.9 pode ser reescrita como:

$$\dot{x} = f(x^*) + Df(x^*)(x - x^*) + g(x). \quad (2.12)$$

Aplicando a mudança de variável $y = x - x^*$, tem-se que $\dot{y} = \dot{x}$ (como consequência da derivação aplicada em ambos os lados da Equação 2.12). Fazendo ainda $f(x^*) = 0$, a Equação 2.12 fica:

$$\dot{y} = Df(x^*)y + g(y + x^*), \quad (2.13)$$

para a qual x^* representa a origem do sistema. Considerando o comportamento das soluções apenas em uma pequena vizinhança em torno de x^* , pode-se desprezar os termos de ordem superior $g(y + x^*)$, obtendo um sistema linear da forma:

$$\dot{y} = Ay, \quad (2.14)$$

onde A é a matriz Jacobina $Df(x^*)$. É possível demonstrar que, por estar-se considerando uma vizinhança muito próxima ao ponto de origem x^* , as trajetórias das solução deste sistema comportam-se da mesma forma que no sistema não-linear equivalente [Hartman 1964]. Assim, o critério para avaliação da estabilidade de um sistema dinâmico não-linear, segundo o método indireto de Lyapunov, pode ser estabelecido como segue:

(a) x^* é *assintoticamente estável*, se todos os autovalores de A tiverem parte real negativa, ou seja, $\text{Re } \lambda_i < 0$, para todo i .

(b) x^* é *instável* se um ou mais autovalores de A tiverem parte real positiva, ou seja, $\text{Re } \lambda_i > 0$, para algum i .

Nota-se que este critério nada conclui para o caso de autovalores nulos ($\text{Re } \lambda_i = 0$). Uma forma de se contornar este problema foi encontrada pelo próprio Lyapunov no método descrito a seguir.

O Método Direto de Lyapunov

Este método de Lyapunov permite generalizar a noção de norma em $\mathbb{R}^n$ para superar a limitação imposta em seu método indireto. Lyapunov demonstrou que é possível utilizar outras funções para se verificar o decrescimento de $|x(t) - x^*|$, quando $x(t)$ é uma solução em uma vizinhança próxima de x^* . Tais funções são conhecidas como *funções de Lyapunov*. Considere, novamente, o sistema dinâmico autônomo não-linear dado pela Equação 2.9 e x^* como sendo um ponto fixo e também a origem do sistema 2.9. Seja V uma função de Lyapunov, diferenciável, cuja derivada é dada por:

$$\dot{V}(x) = \nabla V(x) \cdot \dot{x}, \quad (2.15)$$

definida para cada x ao longo das soluções da Equação 2.9.

Definição: V é dita ser *definida positiva* no espaço das soluções do sistema se (i) $V(x^*) = 0$ e (ii) $V(x) > 0$ para $x \neq x^*$.

Definição: $\dot{V}$ é dita ser *semidefinida negativa* se $\dot{V} \leq 0$.

Definição: $\dot{V}$ é dita ser *definida negativa* se $\dot{V} < 0$.

Com base nas definições acima, é possível estabelecer o critério para análise de sistemas dinâmicos não-lineares, segundo o método direto de Lyapunov:

- (a) x^* é *estável* se, em uma vizinhança próxima de x^* , existe uma função *definida positiva* V , tal que sua derivada $\dot{V}$ é *semidefinida negativa*.
- (b) x^* é *assintoticamente estável* se, em uma vizinhança próxima de x^* , existe uma função *definida positiva* V , tal que sua derivada $\dot{V}$ é *definida negativa*.
- (c) x^* é *instável* se não for estável.

Desta forma, Lyapunov define um método que estende seu método anterior. Ainda assim, o critério apresenta uma condição suficiente, mas não necessária para se determinar a estabilidade de um ponto fixo, uma vez que afirma a existência de uma função V , mas não sugere como encontrá-la. Abstrai-se que a função de Lyapunov pode ser encarada como uma generalização do conceito de energia de um sistema mecânico dissipativo, ou seja, sistemas em que a energia sempre decresce até alcançar o equilíbrio. Esta analogia é a base para a construção de muitos modelos recorrentes de RNAs, como o modelo de Hopfield, conforme discutido na Seção 2.3.

2.1.5 Análise do Comportamento de Sistemas Dinâmicos

Conforme visto anteriormente, além de pontos fixos, um sistema dinâmico pode apresentar comportamentos mais complexos como periodicidade e caoticidade. Esta subseção disserta sobre algumas ferramentas matemáticas que retornam indícios sobre o comportamento dinâmico de um sistema. São discutidos: *Diagramas de Bifurcação*, *Expoentes de Lyapunov*, *Expoentes de Hurst* e *Coefficientes de Fourier*.

Diagrama de Bifurcação

Nas subseções de 2.1.1 a 2.1.3 são apresentadas as séries temporais para o sistema logístico dado pela Equação 2.4, onde várias dinâmicas são observadas de acordo com o valor do parâmetro a . Percebe-se, portanto, que ao passo que o valor de a aumenta, o sistema transita por entre diversos comportamentos dinâmicos. Para $a < 3$, por exemplo, o sistema apresenta regime de convergência a um ponto fixo. Com $a > 3$, observam-se dinâmicas periódicas de períodos 2 e 4 e dinâmicas bastante complexas, que inclui trajetórias periódicas de altos períodos até, enfim, atingir órbitas caóticas para a próximo de 4.

Um das maneiras de se analisar a dinâmica global do sistema é descobrir para quais valores de parâmetros o sistema altera seu comportamento. Tais valores são conhecidos como

pontos de bifurcação. Um *diagrama de bifurcação* é um gráfico que resume o comportamento do sistema de acordo com os valores de um determinado parâmetro. Para cada valor do parâmetro, todos os pontos obtidos durante a iteração do sistema são colocados no gráfico, após a fase de transiência (isto é, após o sistema convergir para um ponto atrator), permitindo visualizar os pontos de bifurcação e as diferentes dinâmicas por estes separados. O diagrama de bifurcação para o mapa logístico da Equação 2.4 é apresentado na Figura 2.5.

Figura 2.5: Diagrama de bifurcação que apresenta o comportamento dinâmico do mapa logístico da Equação 2.4 em relação aos valores assumidos pelo coeficiente de bifurcação a .

Um sistema dinâmico que não apresenta pontos de bifurcação é dito ser *estruturalmente estável*, pois seu comportamento não é significativamente alterado quando seus parâmetros são perturbados.

Expoentes de Lyapunov

Outra maneira de se analisar o comportamento dinâmico de um sistema, principalmente em relação à sua caoticidade, é por meio dos *expoentes de Lyapunov* que, basicamente, medem o grau de divergência de trajetórias inicialmente próximas [Ott 1993]. O expoente de Lyapunov é definido como segue:

$$\lambda_i = \lim_{k \rightarrow \infty} \frac{1}{k} \ln |D_k(x_i)u_i|, \quad (2.16)$$

onde k é o número de iterações em que o sistema é observado, λ_i é o i -ésimo expoente de Lyapunov do sistema N -dimensional, D_k é um vetor composto pelos N autovalores da matriz Jacobiana J e u são os autovetores associados aos autovalores do vetor D_k . A matriz J pode

ser calculada como:

$$J(x) = \prod_{t=t_0}^k J(x(t)), \quad (2.17)$$

onde $J(x(t))$ é a matriz Jacobiana do sistema no instante t .

Os valores dos expoentes λ_i indicam a condição dinâmica do sistema para dados valores de parâmetros. Considera-se que, se ao menos um dos N expoentes de Lyapunov for positivo, o sistema pode ser considerado caótico. Se todos os expoentes forem negativos, o sistema é dito ter trajetórias convergentes a pontos fixos ou dinâmicas periódicas. Um valor nulo representa uma dinâmica quase-periódica ou ainda uma mudança significativa no comportamento do sistema (isto é, um ponto de bifurcação). Apesar de existirem contra-exemplos a esta regra, como verificado em Leonov e Kuznetsov 2007, a condição de caoticidade dada por ao menos um expoente positivo é aceitável até que se prove o contrário. Assim, o expoente que realmente interessa para avaliação da dinâmica do sistema normalmente é o expoente de valor máximo.

No caso do sistema logístico representado pela Equação 2.4, o cálculo do máximo expoente de Lyapunov para dado valor do parâmetro a é dado por:

$$\lambda = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{t=t_0}^k \ln |f'(x(t))|, \quad (2.18)$$

onde $f'(x(t)) = a(1 - 2x(t))$ é a primeira derivada da Equação 2.4. A representação gráfica dos expoentes de Lyapunov do mapa logístico pode ser conferida na Figura 2.6.

Figura 2.6: Expoentes de Lyapunov λ do mapa logístico segundo a Equação 2.18 em relação aos valores assumidos pelo coeficiente de bifurcação a .

Expoentes de Hurst

Os Expoentes de Hurst [Hurst 1951] foram propostos por Harold Edwin Hurst no ramo de hidrologia, porém ganharam espaço em outras áreas, como, por exemplo, em economia utilizados para analisar tendências em curvas financeiras. Trata-se de uma ferramenta que avalia o grau de autocorrelação de uma série temporal. Em outras palavras, um expoente de Hurst H é um número real entre 0 e 1 onde:

- $0 \leq H < 0,5$ indica autocorrelação negativa, ou seja, quando mais próximo de 0, maior a probabilidade de que uma subida na curva seja imediatamente seguida, a longo prazo, por uma descida. Assim sendo, pode-se afirmar que o grau de autocorrelação negativa reflete o grau de periodicidade da série temporal (quanto mais próximo de zero for H , menor será o período da série periódica).
- $0,5 < H \leq 1$ indica autocorrelação positiva, ou seja, quando mais próximo de 1, maior a probabilidade de que uma subida na curva seja imediatamente seguida por outra subida, ou que uma descida seja imediatamente seguida por outra descida a longo prazo.
- quando mais próximo de 0,5 for H , maior será a imprevisibilidade da autocorrelação, isto é, tornar-se impossível prever o comportamento da série a longo prazo. Pode-se considerar, portanto, que, partindo de valores próximos a 0,5, quanto maior o valor do expoente, menos periódica é a série, ou seja, o período p cresce até que se torne indefinido. Neste ponto, a série passa a exibir um comportamento não-periódico o qual é forte indício de uma dinâmica caótica. Nota-se que o expoente não é calculado para séries com comportamento estacionário.

Uma das formas de se calcular os expoentes de Hurst é pelo método de análise do *rescaled range* da série temporal após t observações. Data uma série temporal T de tamanho N , inicialmente, geram-se L componentes T_i , sendo:

- T_1 composto por uma subsérie de tamanho $n = N$ (ou seja, a própria série T);
- T_2 composto por duas subséries de tamanho $n = N/2$ (resultado da divisão de T em duas partes);
- T_3 composto por quatro subséries de tamanho $n = N/4$ (resultado da divisão de T em quatro partes);
- T_4 composto por oito subséries de tamanho $n = N/8$ (resultado da divisão de T em oito partes).

- O processo de decomposição segue até atingir um componente T_L , para o qual o tamanho das subséries é $n = 8$.

Para cada componente T_i , calcula-se a média dos chamados *rescaled ranges* RS . Um RS é calculado para cada uma das subséries que compõem a componente T_i da seguinte forma:

$$RS = \frac{R(n)}{S(n)}, \quad (2.19)$$

sendo S o desvio padrão e R o chamado *range* da subsérie. Considerando uma subsérie composta por n elementos $x_1, x_2, \dots, x_n$, o algoritmo para o cálculo do RS é dado como segue:

1. calcula-se a média m :

$$m = \frac{1}{n} \sum_{i=1}^n x_i$$

2. gera-se o chamado *ajuste da média* Y :

$$Y_i = x_i - m, \quad i = 1, 2, \dots, n$$

3. calcula-se a chamada *série de desvio cumulativo* Z :

$$Z_i = \sum_{j=1}^i Y_j, \quad i = 1, 2, \dots, n$$

4. calcula-se o *range* R :

$$R(n) = \max(Z_1, Z_2, \dots, Z_n) - \min(Z_1, Z_2, \dots, Z_n)$$

5. finalmente, calcula-se o desvio padrão S :

$$S(n) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - m)^2}$$

Após calculados os *rescaled range* médios para cada componente T_i , gera-se uma tabela de L linhas, onde cada linha contém o $\log_2$ do RS médio e o $\log_2$ do tamanho n das subséries referentes à componente T_i . Com essa tabela log em mãos, aplica-se uma regressão linear a fim de se encontrar a reta que melhor se ajusta aos pontos. O expoente de Hurst H é dado pelo valor coeficiente angular da reta gerada.

A Figura 2.7 exibe o espectro de Hurst do sistema logístico dado pela Equação 2.4 em relação à variação do parâmetro a . Nota-se que os expoentes mantêm-se iguais a (ou muito próximos de) 0 para $0,30 < a < 3,60$ indicando um comportamento periódico. Nota-se, também, que não há valor de expoente para quando o sistema exibe um comportamento estacionário. A partir de $a \sim 3,60$, os expoentes começam a se afastar de zero e mantêm-se em direção a 0,50 enquanto a cresce, indicando um possível comportamento caótico.

Figura 2.7: Expoentes de Hurst H do mapa logístico da Equação 2.4 em relação aos valores do coeficiente de bifurcação a . O espectro exhibe dinâmicas periódicas (valores próximos a 0) e caóticos (valores próximos a 0,5). Nota-se que não existe expoente para os valores de a em que o mapa apresenta uma dinâmica estacionária.

Coeficientes de Fourier

Os trabalhos do matemático francês Jean Baptiste Joseph Fourier mostram que uma função periódica pode ser aproximada pela soma de funções senóides e cossenóides, chamadas de *harmônicas*, cada qual adequadamente regulada por um coeficiente [Boashash 2003]. Assim, dada uma função periódica $f(x)$ de período p , sua expansão na chamada *série de Fourier*, considerando um conjunto de N valores, é feita da seguinte forma:

$$f_N(x) \sim \frac{a_0}{2} + \sum_{n=1}^N [a_n \cos(2\pi nx/p) + b_n \sin(2\pi nx/p)], \quad (2.20)$$

onde a_0 , a_n e b_n são os ditos *coeficientes de Fourier* e servem para ajustar o peso de cada uma das harmônicas da série. Tais coeficientes são calculados como:

$$\begin{aligned} a_0 &= \frac{2}{p} \int_{x_0}^{x_0+p} f(x) dx \\ a_n &= \frac{2}{p} \int_{x_0}^{x_0+p} f(x) \cos(2\pi nx/p) dx \\ b_n &= \frac{2}{p} \int_{x_0}^{x_0+p} f(x) \sin(2\pi nx/p) dx \end{aligned}$$

com $n = 1, 2, \dots, N$. A Equação 2.22 pode ainda ser reescrita em notação complexa de Euler como segue:

$$f_N(x) \sim \sum_{n=-N}^N c_n e^{2\pi i n x/p} \quad (2.21)$$

onde os coeficientes c_n são calculados pela Equação abaixo.

$$c_n = \frac{1}{p} \int_{x_0}^{x_0+p} f(x) e^{-2\pi i n x/p} dx \quad (2.22)$$

Os coeficientes c_n podem ser repensados como uma função F que depende de n . Quando x é uma variável de tempo, n passa a representar valores de frequência (se x é *segundos*, n é *Hertz*). Neste caso, F é chamada de *Transformada de Fourier*: uma função inversível que retorna coeficientes relacionados a valores de frequência de uma série temporal. Assim, a transformada faz uma conversão de valores em um *domínio de tempo* para valores em um *domínio de frequência*. A Equação 2.23 define uma transformação de Fourier considerando valores contínuos. $F(w)$ conterá informações de amplitude e fase referentes à frequência w . A transformada inversa é dada pela Equação 2.24.

$$F(w) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x w} dx \quad (2.23)$$

$$f(x) = \int_{-\infty}^{\infty} F(w) e^{2\pi i w x} dx \quad (2.24)$$

Considerando valores discretos, a *Transformada Discreta de Fourier* pode ser aplicada segundo a Equação que segue, considerando uma série temporal com N elementos $\{x_1, x_2, \dots, x_N\}$.

$$F_k = \frac{1}{N} \sum_{n=1}^N x_n e^{-2\pi i k n/N} \quad (2.25)$$

onde $k = 1, 2, \dots, N$, gerando a chamada *série de coeficientes de Fourier* $\{F_1, F_2, \dots, F_N\}$.

Vislumbrando uma forma computacionalmente mais eficiente para o cálculo os coeficientes de Fourier sobre valores discretos, Cooley e Tukey 1965 desenvolveram um algoritmo conhecido como *Transformada Rápida de Fourier*, cuja ideia é calcular a Transformada Discreta de Fourier aplicando-se o *método dos desdobramentos sucessivos*. Dada a transformada discreta F sobre uma série temporal $\{x_0, x_2, \dots, x_{N-1}\}$, tem-se que:

$$\begin{aligned}
F_k &= \sum_{n=0}^{N-1} x_n e^{-2\pi i n k / N} = \\
&= \sum_{n=0}^{N/2-1} x_{(2n)} e^{-2\pi i (2n) k / N} + \sum_{n=0}^{N/2-1} x_{(2n+1)} e^{-2\pi i (2n+1) k / N} = \\
&= \sum_{n=0}^{N/2-1} x_n^{par} e^{-2\pi i n k / (N/2)} + e^{-2\pi i k / N} \sum_{n=0}^{N/2-1} x_n^{impar} e^{-2\pi i n k / (N/2)}
\end{aligned}$$

A análise sobre a dinâmica da série temporal pode ser feita construindo-se a curva que contém as amplitudes dos coeficientes de Fourier. Tal curva é conhecida como *espectro de Fourier*. Séries periódicas geram espectros de Fourier com um número de *picos* bem definido. O espectro de séries não-periódicas pode ser considerado como uma curva “irregular”. Para exemplificar a aplicação dos Coeficientes de Fourier, as figuras a seguir exibem os espectros de Fourier das séries temporais exibidas pelas Figuras de 2.8 a 2.11, respectivamente, as quais representam os diversos comportamentos dinâmicos assumidos pelo mapa logístico da Equação 2.4, em relação aos valores do parâmetro a . Nota-se que quando a série temporal possui um período p , seu espectro contém $p - 1$ picos. Quando o comportamento da série é caótico, o espectro é irregular.

Figura 2.8: Coeficientes de Fourier c da série temporal com dinâmica estacionária exibida pela Figura 2.1. A curva constante salienta um comportamento estacionário.

Esta Subseção apresentou quatro maneiras para se verificar o comportamento de um sistema dinâmico. Conforme já dito, o tais comportamentos podem ser observados em vários sistemas da natureza, incluindo redes neurais biológicas. Alguns aspectos interessantes do cérebro, incluindo análises sobre sua dinâmica, serão discutidos a seguir.

Figura 2.9: Coeficientes de Fourier c da série temporal com dinâmica periódica de período 2 exibida pela Figura 2.2. O pico ao centro indica um comportamento periódico.

Figura 2.10: Coeficientes de Fourier c da série temporal com dinâmica periódica de período 4 exibida pela Figura 2.3. Notam-se 3 picos sobressalentes, indicando que há um relacionamento entre o número de picos e o número de períodos da série.

Figura 2.11: Coeficientes de Fourier c da série temporal com dinâmica caótica exibida pela Figura 2.4. A curva irregular reforça a indicação de um comportamento caótico.

2.2 Algumas Características das Redes Neurais Biológicas

Esta Subseção aborda os principais conceitos de redes neurais biológicas que fundamentam as pesquisas relacionadas às RNAs. A estrutura básica do cérebro, alguns aspectos de neurodinâmica, o aprendizado Hebbiano e o modelo de Hopfield são descritos a seguir.

2.2.1 A Estrutura Básica do Cérebro

O cérebro humano é um órgão extremamente complexo no qual se observa a ocorrência de vários fenômenos físicos e a compreensão de muitos aspectos de seu funcionamento ainda permanece obscura. Esta Subseção pretende abordar apenas algumas características do dinamismo e dos mecanismos computacionais utilizados por este órgão, descobertos por estudos na área da neurociência, e que se mostram relevantes para fornecer o embasamento necessário ao leitor para abstrair adequadamente as motivações e os objetivos deste projeto de pesquisa.

Basicamente, a rede neural biológica pode ser compreendida como um processador de informações composta por unidades de processamento básico chamadas *neurônios* que se conectam uns aos outros por meio das chamadas *sinapses neurais*, formando uma vasta e complexa rede. Para se ter uma ideia de tal complexidade, estima-se que existam mais de 100 bilhões de neurônios no cérebro humano (sendo estes de milhares de tipos diferentes) e mais de 100 trilhões de sinapses neurais (um neurônio pode vir a formar mais de 10 mil sinapses ao mesmo tempo). A compreensão de como esta rede opera e processa informação ganhou força a partir dos estudos do médico e histologista espanhol Santiago Ramón y Cajál no início do século XX, quando introduziu-se a ideia de que os neurônios são a estrutura constituinte do cérebro [Haykin 1999].

Os neurônios são células compostas por basicamente três estruturas: (i) dendritos - prolongamentos curtos que recebem os estímulos vindos de outros neurônios, (ii) corpo celular - que inclui o núcleo e componentes básicos da célula, e (iii) axônios - prolongamentos longos que propagam o estímulo recebido às outras células nervosas. Desta forma, neurônios operam como processadores muito simples, cuja função é, portanto, propagar (ou não) um estímulo nervoso (um pulso elétrico) mediante um potencial de ativação. Foi Ramón y Cajál que pela primeira vez discutiu a unidirecionalidade da propagação do pulso no chamado *Princípio da Polarização Dinâmica*, o qual afirma que sinais elétricos fluem em apenas uma direção: da recepção do neurônio (normalmente dendritos e corpo celular) para a zona de disparo do axônio [Haykin 1999]. Ao chegar na zona de disparo, o pulso pode ser propagado para outra célula nervosa por meio da sinapse neural.

Também foi Ramón y Cajál que apresentou o *Princípio da Especificidade Conectiva*, o

qual postula que as células nervosas não se conectam indiscriminadamente, nem formam redes aleatórias. Observa-se no cérebro basicamente dois tipos de sinapses neurais: elétricas e químicas. As primeiras são menos comuns e ocorrem em algumas poucas áreas do cérebro. A propagação do sinal elétrico ocorre por meio de canais de junções, ou seja, uma conexão citoplasmática entre a célula *pré-sináptica* (que enviará o pulso) e a célula *pós-sináptica* (que receberá o pulso). Muito mais abundantes são as sinapses químicas, nas quais não há contato direto entre as membranas das células pré e pós-sinápticas. A transmissão do estímulo se dá através dos chamados *neurotransmissores*. Estima-se que haja mais de 100 tipos distintos de neurotransmissores, cada qual exigindo uma morfologia única das vesículas que armazenam a substância nos axônios, assim como dos receptores químicos nos dendritos. O tipo de neurotransmissor utilizado na sinapse pode tanto estimular quanto inibir a propagação do sinal elétrico para a célula vizinha, agindo como um potencial de ativação.

Um modelo elaborado para explicar o mecanismo de origem e propagação do pulso elétrico ao longo de uma célula nervosa foi proposto em Hodgkin e Huxley 1952. O *modelo de Hodgkin-Huxley* elucida matematicamente a chamada *bomba de sódio-potássio*: mediante a um estímulo que atinge o potencial de ação, a membrana celular torna-se permeável à entrada de íons de sódio permitindo que o interior, repleto de íons de potássio, torne-se mais positivamente carregado em relação ao exterior da célula. A diferença de potencial elétrico gerada no local do estímulo (*despolarização*) estimula a seção subjacente da membrana tornando-a também permeável ao sódio. Desta forma, tem-se um efeito cascata ao longo da membrana e uma corrente elétrica é gerada. Alguns milisegundos após aplicação do estímulo, um processo de difusão ativa bombeia íons de sódio para o meio extracelular e íons de potássio para o meio intracelular, fazendo com que o local previamente despolarizado retorne ao estado de repouso (*repolarização*). Por este trabalho, Hodgkin e Huxley ganharam o prêmio nobel na área de fisiologia ou medicina em 1963.

Encarando as células nervosas como unidades simples de processamento, um neurônio isoladamente processa informação a um tempo de ciclo de aproximadamente 10^{-3} s. Por outro lado, a largura de banda que se observa na rede neural biológica em sua totalidade é de 10^{14} bits/seg [Russell e Norvig 2010], p. 12. Essa alta capacidade computacional do cérebro pode ser explicada pelo fato de que o processamento global é explorado de forma massivamente paralela. Além disso, a informação é armazenada de forma distribuída por entre a enorme quantidade de neurônios e conexões. Outra característica importante é que o cérebro possui áreas especializadas para tratar eventos específicos (agrupamentos funcionais localizados). Diferentes atividades são tratadas em diferentes porções do córtex cerebral, em baixo ou alto nível da sua organização topológica. As principais áreas do córtex cerebral são ilustradas na Figura 2.12.

Figura 2.12: Mapa do córtex cerebral evidenciando suas diversas regiões estruturais e funcionais. As áreas consideradas mais importantes são: 1, 2 e 3: córtex somatosensorial; 4: córtex motor; 6: córtex pré-motor; 17, 18 e 19: córtex visual; 41 e 42: córtex auditivo. Fonte: Haykin 1999, p. 10.

Na tentativa de se reproduzir tal eficiência computacional, as RNAs inspiram-se nos fundamentos conceituais das redes neurais biológicas com a intensão de gerar modelos simplificados dos mesmos, absorvendo muitos dos mecanismos até agora descobertos. Além dos já mencionados, como processamento paralelo e armazenamento distribuído, observa-se também a capacidade da rede recuperar uma informação armazenada partindo-se de apenas parte ou de um conjunto distorcido da informação original (diferentemente do método convencional de memória endereçável, no qual a informação é recuperada via acesso direto a um endereço de memória). Alguns modelos de RNAs (como o de Hopfield) implementam esta capacidade, a qual dá-se o nome de *memória associativa* e seu funcionamento exige que a rede exiba atratores dinâmicos. Uma discussão sobre alguns pontos do comportamento dinâmico do cérebro é apresentada na Subseção a seguir.

2.2.2 A Neurodinâmica

A *neurodinâmica* é um campo de pesquisa que enxerga as redes neurais biológicas como sistemas dinâmicos não-lineares de alta dimensionalidade, com ênfase em problemas de estabilidade [Haykin 1999], ou seja, a análise de possíveis configurações, dadas condições iniciais, para

as quais o sistema encontra-se estável ou instável, no sentido de Lyapunov (conforme apresentado na Subseção 2.1.4). Porém, por mais de 25 anos, muitos trabalhos de pesquisa analisam a dinâmica cerebral e comprovam a existência dos demais comportamentos discutidos na Seção 2.1: não somente estacionária, mas o cérebro também apresenta, principalmente, dinâmicas periódicas e caóticas, alterando esse comportamento de acordo com estímulos recebidos do ambiente.

Os pesquisadores Skarda e Freeman [Skarda e Freeman 1987], por exemplo, demonstraram a presença de pontos de bifurcação em redes neurais biológicas durante o condicionamento de coelhos ao uso de determinados odores. Medindo a atividade cerebral destes animais por meio da técnica de eletroencefalograma (EEG), demonstrou-se, experimentalmente, que os sinais de atividade cerebral destes animais apresentavam uma dinâmica caótica quando não sofriam estimulação (não inalavam nenhum odor específico). Porém, alternavam para uma dinâmica quase periódica quando um odor conhecido era inalado, evidenciando que o cérebro realiza alternância entre comportamentos dinâmicos.

Além do trabalho de Skarda e Freeman 1987, Babloyantz et al. 1985 mediram o EEG de um cérebro humano durante o período de sono e observaram a presença do atrator caótico durante o sono profundo (REM), ou quando o sonhador era perturbado por um ruído qualquer no ambiente. Já Aihara e Matsumoto 1986 identificaram dinâmicas periódicas, quase-periódicas e caóticas nas atividade de potencial de membranas dos axônios de lulas gigantes, demonstrado que comportamentos dinâmicos estão presentes no cérebro também em nível celular.

Comprova-se, portanto, que redes neurais biológicas possuem comportamento dinâmico, apresentando diversos tipos de atratores, desde periódicos a até caóticos. RNAs que operam dinamicamente são conhecidas como redes *recorrentes* e são modeladas como sistemas dinâmicos não-lineares de alta dimensionalidade, utilizando, normalmente, atratores estacionários para armazenamento e busca de informação, cuja estabilidade é mapeada pelo método direto de Lyapunov. Nota-se, porém, que não é comum observar convergência a pontos fixos na dinâmica cerebral - fato que compromete a plausibilidade biológica de muitos modelos recorrentes (como o modelo de Hopfield, abordado na Seção 2.3).

Além dos aspectos dinâmicos que redes neurais apresentam durante o processo de busca e recuperação de informação, outro mecanismo que é objeto de largo interesse, tantos para neurocientistas, quanto para cientistas da computação, é o método, ou regra, pela qual a rede adquire conhecimento (aprendizado), ainda obscuro em muitos de seus aspectos. Uma das principais regras observadas no cérebro e aplicada à RNAs é o aprendizado Hebbiano, discutido na Subseção seguinte.

2.2.3 O Aprendizado Hebbiano

O cérebro possui a capacidade de se adaptar ao ambiente a sua volta. A esta capacidade dá-se o nome de *plasticidade neural*. Em um sistema neural adulto, a plasticidade se faz presente por meio de dois mecanismos: (i) criação de novas conexões sinápticas entre neurônios e (ii) modificação de sinapses já existentes [Haykin 1999]. É esta constante reconfiguração sináptica que dá ao cérebro sua capacidade de aprender e de generalizar a informação aprendida. O conhecimento emerge, não de células isoladas, mas da rede como um todo, de acordo com a configuração de unidades e conexões que a rede apresenta.

Observando tal fenômeno, o neurofisiologista Donald Hebb teorizou, de forma eficiente, como a plasticidade do sistema nervoso pode ser subentendida em nível celular. O *Postulado de Hebb* é assim descrito em Hebb 1949, p. 69:

“When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.”

Em outras palavras, o postulado afirma que quando dois neurônios são simultaneamente (sincronamente) ativados, a ligação sináptica entre eles é, de alguma forma, fortalecida. Caso contrário é inibida. Esta ideia por ser descrita matematicamente por meio de uma regra de atualização de pesos sinápticos, também conhecida como *Hipótese de Hebb* e dada pela seguinte equação:

$$\Delta W_{ij} = \eta x_i x_j, \quad (2.26)$$

onde W_{ij} é o valor do peso sináptico do neurônio j para o neurônio i , x_j é o sinal pré-sináptico (valor de saída do neurônio j), x_i é o sinal pós-sináptico (valor de saída do neurônio i) e η é uma constante positiva que determina a *taxa de aprendizado*.

Esta capacidade de reforço, presente nas sinapses Hebbianas, utiliza: (i) mecanismo dependente de tempo (sincronia entre sinais pré e pós-sinápticos), (ii) localidade (a informação que trafega pela sinapse é utilizada como fator de modificação sináptica local), (iii) interatividade (a mudança na sinapse depende de ambos os sinais pré e pós-sinápticos) e (iv) correlação (a mudança só ocorre pela conjunção dos sinais pré e pós-sinápticos).

Trata-se, portanto, de uma regra de aprendizado *não-supervisionado*, pois leva em conta apenas os valores dos sinais pré e pós-sinápticos (não há informação sobre a saída desejada como, por exemplo, ocorre na regra Delta presente no algoritmo de aprendizado *Error Back-Propagation* [Bryson e Ho 1969], [Rumelhart et al. 1986]). O aprendizado de Hebb é um dos

mais famosos e utilizados por modelos artificiais de redes neurais, um dos quais é o modelo de Hopfield, conforme será visto a seguir.

2.3 O Modelo de Hopfield

Um dos primeiros modelos de RNAs propostos foi o *Perceptron* [Rosenblatt 1958]. Composto por somente um neurônio do tipo MCP [McCulloch e Pitts 1943], ele era usado para classificar apenas padrões linearmente separáveis. Visando superar esta limitação, surgiu a rede *Perceptron Multicamadas* (MLP) com algoritmo de aprendizado *Error Back-Propagation* [Bryson e Ho 1969], [Rumelhart et al. 1986], o qual é capaz de classificar qualquer tipo de padrão utilizando, pelo menos, três camadas. Porém, o tempo gasto para o treinamento e a dificuldade em se encontrar a configuração ideal dos parâmetros da rede tornaram-se um obstáculo à sua aplicação. Outros modelos surgiram, ou para aprimorar a eficiência computacional, ou para se tentar resolver problemas mais complexos, ou para refletir fenômenos mais similares ao comportamento do cérebro. Entre os novos conceitos estão as redes *recorrentes* (ou *associativas*), cuja resposta obtida é devolvida como entrada para o próximo ciclo de processamento, provendo um sistema dinâmico não-linear de alta dimensionalidade.

Pouco interesse havia no estudo de redes neurais recorrentes, dada a dificuldade em se promover treinamento e análise. Em 1982, o físico J. J. Hopfield, no trabalho Hopfield 1982, apresentou um modelo recorrente que faz analogia a um sistema físico real. Cada estado da rede pode ser associado a um valor de energia que decresce de maneira monotônica a medida que uma trajetória é descrita no espaço de estados em direção a um ponto fixo. Estes pontos fixos, obtidos durante o treinamento, correspondem a pontos estáveis de energia mínima análogos aos pontos atratores estacionários de um sistema dinâmico não-linear, uma vez que a estabilidade é garantida pelo método direto de Lyapunov, conforme discutido na Subseção 2.1.4.

A rede de Hopfield é organizada em uma única camada, na qual os neurônios são globalmente conectados (cada neurônio forma sinapses com todos os demais, menos com ele mesmo) de forma recorrente, ou seja, a saída de um neurônio será o valor de entrada dos demais neurônios ao longo da evolução do sistema. Assim, uma dada resposta irá depender sempre do estado do sistema no intervalo de tempo anterior, o que confere à rede uma espécie de memória. Uma informação (ou padrão de memória) é armazenada de forma distribuída e pode ser recuperada com base no conteúdo total ou parcial da mesma, isto é, o modelo funciona como uma memória associativa, como também ocorre no cérebro.

A rede pode ser concebida para trabalhar com valores discretos ou contínuos. Considerando-se a versão discreta do modelo, os neurônios da rede de Hopfield são do tipo MCP: modelo de

neurônio que basicamente realiza uma soma ponderada das entradas e seus respectivos pesos, produzindo uma saída u_i , chamada de *estado interno* do neurônio i . Ao estado interno é aplicada uma função f , não-linear, chamada de *função de ativação*, fabricando a saída final x_i do neurônio i .

O tipo de neurônio usado no modelo de Hopfield, exerce uma atividade *tudo-ou-nada*, ou seja: ou dispara, ou não dispara, dependendo do valor alcançado pelo estado interno u_i . Portanto, a saída x_i é binária: um valor para representar a ativação e outro para representar a não-ativação. Este cálculo é realizado pela função de ativação f , que introduz a não-linearidade, provocando a conversão do estado interno em um valor binário. No caso do modelo discreto de Hopfield, $x_i \in \{-1, 1\}$ e f é a função *sinhal*, dada por:

$$f = \begin{cases} 1 & \text{se } u_i > 0; \\ -1 & \text{se } u_i < 0. \end{cases} \quad (2.27)$$

Outro fator que influencia na ativação de um neurônio do tipo MCP é o chamado *limiar de ativação* (ou *threshold*) θ que compõe a soma ponderada na computação do neurônio. Seu significado é o de estabelecer um valor limite para o estado interno u . Se u atingir, ou ultrapassar este limite, a função de ativação provocará o disparo do neurônio. Caso contrário, o neurônio não dispara. Muitas vezes o *threshold* é também chamado de *bias* b . Apesar de ambos no modelo matemático representarem a mesma entidade, cabe ressaltar que, conceitualmente, possuem significado distinto: o objetivo do *bias*, além de atuar como um ajuste para a fronteira de decisão em um problema de classificação linear, é também o de simular a possibilidade biológica de um neurônio disparar mesmo que não receba estímulos externos. Desta forma, o *bias* b terá sempre valor unitário positivo - uma entrada que sempre está ativada e que será ponderada por um valor de peso sináptico. Assim, enquanto b atua como um fator de incremento na computação do neurônio, θ exerce o papel oposto:

$$b = -\theta. \quad (2.28)$$

No modelo de Hopfield, o *threshold* (ou *bias*) possui valor nulo para garantir a convergência do modelo discreto durante o processo de evolução dos padrões. A Equação 2.29 descreve a regra de evolução pelas quais passam as órbitas dos neurônios até convergirem para um ponto atrator estacionário.

$$\begin{cases} x_i(t+1) = f(u_i(t)), \\ u_i(t) = \sum_{j=1}^N W_{ij}x_j(t) \quad j \neq i \text{ e } i = 1, \dots, N, \end{cases} \quad (2.29)$$

onde N é o número de neurônios da rede, u_i é o estado interno do i -ésimo neurônio, W_{ij} codifica

a eficiência sináptica (ou peso) entre os neurônios i e j , $x_i \in \{1, -1\}$ é a saída, ou resposta, emitida pelo i -ésimo neurônio e f é a função de ativação, dada pela função sinal representada pela Equação 2.27.

O treinamento da rede, cuja função é a atualização dos pesos sinápticos de forma a garantir a existência de pontos atratores, é feito em um único ciclo iterativo e seguindo as regras do aprendizado de Hebb (abordado na Subseção 2.2.3), matematicamente descrito pela Equação 2.26. Considerando o armazenamento de M padrões de memória $\xi^p, p = 1, 2, \dots, M$, sendo cada entrada da forma $\xi = [\xi_1, \dots, \xi_n]^T$ e o valor de saída de cada $\xi_i \in \{-1, 1\}, i = 1, 2, \dots, N$, a regra de aprendizado pode ser reescrita em formato matricial como:

$$W = \frac{1}{N} \sum_{p=1}^M \xi^p (\xi^p)^T - MI, \quad (2.30)$$

onde W representa a matriz de pesos da rede e I é a matriz Identidade. Esta regra de aprendizado calcula a autocorrelação dos padrões ξ gerando uma matriz de pesos simétrica. Isto garante que os padrões sejam armazenados como pontos fixos assintoticamente estáveis da Equação 2.29, pois cada padrão representará um ponto de energia mínima local, segundo os critérios de estabilidade derivados do método direto de Lyapunov.

A recuperação de um item de memória consiste em se fixar a condição inicial com a informação (possivelmente parcial) do item de memória a ser recuperado e, a partir de então, o sistema irá evoluir por entre estados intermediários, os quais são alcançados por meio da atualização gradual dos neurônios (descrita pela Equação 2.29), até convergir para o estado correspondente ao padrão de memória ξ^p que representa a informação a ser recuperada. A eficiência do reconhecimento do padrão pode ser medida através de um *overlap* entre o padrão recuperado ξ'^p e o padrão original ξ^p , por meio da equação:

$$O = \frac{1}{N} \sum_{i=1}^N \xi_i^p \xi_i'^p, \quad (2.31)$$

onde N é o número de elementos dos padrões (que corresponde ao número de neurônios da rede) e ξ_i^p representa o i -ésimo elemento do padrão ξ^p . Este somatório retorna o cálculo da distância de *Hamming*, ou seja, o número de elementos diferentes entre ambos os padrões ξ^p e ξ'^p .

Para a demonstração de que o modelo converge para um atrator estacionário armazenado, parte-se do conceito de energia da rede (no sentido de Lyapunov), definida como:

$$E(x) = -\frac{1}{2} \sum_i \sum_{j, j \neq i} W_{ij} x_i x_j \quad (2.32)$$

Primeiramente, deve-se encontrar a variação discreta da energia $E(x)$ ao longo das soluções da Equação 2.29. Para isto, deve-se considerar as seguintes características do modelo:

1. As conexões são simétricas ($W_{ij} = W_{ji}$).
2. Os neurônios não são conectados a si mesmos ($W_{ij} = 0$, quando $i = j$).
3. A atualização do estado interno dos neurônios é feita de forma assíncrona (ou seja, apenas um neurônio altera seu estado interno por vez e, uma vez atualizado, o valor de saída deste neurônio é prontamente utilizado para atualizar o estado interno dos demais neurônios durante uma iteração do processo cíclico de busca e recuperação de padrão - uma iteração é completada quando todos os neurônios da rede têm seus estados internos atualizados).

Desta forma, a variação da energia $E(x)$ é dada por:

$$\begin{aligned} \Delta E(x) &= -\frac{1}{2} \sum_i \sum_{j, j \neq i} W_{ij} x_i(t+1) x_j(t+1) + \frac{1}{2} \sum_i \sum_{j, j \neq i} W_{ij} x_i(t) x_j(t) \\ &= (-x_i(t+1) + x_i(t)) \sum_{j, j \neq i} W_{ij} x_j(t) \\ &= (-x_i(t+1) + x_i(t))(x_i(t)) \\ &= -(x_i(t+1))^2 + x_i(t) x_i(t+1) \\ &\leq 0, \end{aligned}$$

onde a primeira passagem considera as características (1) e (2) do modelo e a segunda passagem considera a Equação 2.29.

Como consequência do resultado, pode-se afirmar que: (i) $\Delta E(x)$ é limitada; (ii) $\Delta E(x)$ sempre decresce; (iii) $\Delta E(x) = 0$ somente nos pontos fixos. Disto, conclui-se que $x(t)$ não pode oscilar e só existem pontos fixos assintóticos estáveis para a Equação 2.29. Mais detalhes desta demonstração podem ser encontrados em Rojas 1996, p. 350.

Assim sendo, a rede neural proposta por Hopfield apresenta-se como um modelo cujos mecanismos de memória se diferem de outros modelos tradicionais e aproxima-se de um sistema neural biológico no que diz respeito ao armazenamento da informação (de maneira distribuída) e à recuperação da informação (por meio da evolução de estados no tempo e de forma associativa). Por outro lado, o modelo é baseado na convergência para um atrator estacionário, sendo

este tipo de dinâmica pouco comum no cérebro (como discutido na Subseção 2.2.2). Além disso, como outros modelos tradicionais de RNAs, Hopfield possui uma estrutura topológica simplificada quando comparada a de uma rede biológica, a qual apresenta grupos geométricos funcionais bem definidos, com ligações entre um número muito vasto de unidades neuronais. Uma teoria que contempla e modela muito bem estas propriedades é das chamadas *Redes Complexas*, tema abordado a seguir.

2.4 Redes Complexas

Redes Complexas é uma área multidisciplinar de grande interesse cuja pesquisa iniciou-se em meados de 1930 no ramo da sociologia. Atualmente, a fim de resolver problemas específicos de diversas áreas, pode ser utilizada para representar e estudar diversos aspectos do mundo real, tais como a estrutura física de grandes redes de computadores (como a Internet), relações sociais entre indivíduos e grupos de indivíduos, redes neurais biológicas, redes metabólicas, cadeias alimentares, entre outros [Albert e Barabási 2002], [Boccaletti et al. 2006], [Newman 2003]. Ou seja, tais redes podem modelar sistemas reais e dinâmicos com eficiência.

Como o nome sugere, refere-se a uma rede representada na forma de um grafo e cuja organização e estrutura topológica fogem do trivial por estarem sujeitas a constantes alterações devido à configuração dinâmica do sistema. Um exemplo simples pode ser observado na Figura 2.13 que retrata uma rede de relações entre membros do clube de karatê Zachary [Zachary 1977]. No caso geral, as redes complexas são composta por um conjunto muito vasto de vértices e arestas e possuem propriedades topográficas peculiares quando comparadas a grafos convencionais. Algumas destas propriedades, que refletem o dinamismo deste tipo de estrutura, são:

- coeficiente de aglomeração: representa a chance de dois vértices estarem conectados considerando o fato de ambos estarem ligados a um terceiro nó em comum (exemplo: se Fulano é amigo de Beltrano e Beltrano é amigo de Ciclano, então Fulano e Ciclano tem grandes chances de serem amigos);
- distribuição de graus: trata-se de uma função de distribuição de probabilidade que indica a chance de um determinado vértice ter grau fixo, ou seja, o número de arestas que se conectam a este vértice permanecer o mesmo durante a evolução da rede;
- resistência: indica a capacidade da rede manter sua funcionalidade estável após a remoção de alguns vértices ou arestas;

- misturas de padrões: capacidade de uma rede em apresentar grupos de vértices cada qual representando diferentes tipos de objetos (exemplo: cadeias alimentares nas quais os vértices podem representar plantas, animais herbívoros, ou animais carnívoros);
- correlação de graus: indica se as arestas em uma rede interligam vértices com graus parecidos (utilizada para investigar a conexão entre vértices de diferentes tipos).

Figura 2.13: Exemplo de uma Rede Complexa: rede de relações entre membros do clube de karatê Zachary [Zachary 1977].

A forma como vértices e arestas se correlacionam define, basicamente, o tipo da rede. Desta forma, existem três principais categorias de redes complexas, descritas a seguir:

- **Redes Aleatórias:** É o primeiro e o mais simples modelo de rede complexa. Proposto por Erdős e Rényi 1961, tal modelo considera que arestas são adicionadas de forma aleatória entre um número fixo de N vértices. O grau esperado de cada vértice é dado pela equação:

$$\langle k \rangle = p(N - 1), \quad (2.33)$$

onde p é a probabilidade de um dado vértice se ligar a outro vértice qualquer, segundo uma distribuição de Poisson. Desta forma, pode-se dizer que todos os vértices terão,

aproximadamente, a mesma quantidade de conexões, como também, as mesmas chances de receberem novas ligações, tornando a construção da rede um processo aleatório [Barabási e Albert 1999]. A estrutura de uma rede aleatória pode ser observada na Figura 2.14(a);

- **Redes Pequeno-Mundo:** Watts e Strogatz 1998 propuseram um modelo que, partindo-se de uma rede regular (rede em que todos os vértices possuem o mesmo grau e cada vértice liga-se a k vizinhos mais próximos em cada direção), algumas arestas são retiradas e reinseridas de maneira aleatória e com probabilidade p . Como consequência, a rede regular passa a apresentar algumas ligações entre vértices mais distantes. Estas ligações funcionam como “atalhos” que provocam a redução da distância média da rede. O fenômeno de *pequeno-mundo* é observado em redes sociais e linguísticas, nas quais a maioria dos vértices se conecta a outros através do *caminho mínimo* (distância geodésica formada pelo menor número de arestas que conecta o vértice origem ao vértice destino) e uma minoria de “atalhos” que depende do valor de p . O comprimento l do caminho mínimo médio entre pares de vértices de uma rede não direcionada é dado pela equação:

$$l = \frac{1}{n(n+1)/2} \sum_{i \geq j} d_{ij}, \quad (2.34)$$

onde n é o número de vértices e d_{ij} é a distância geodésica entre os vértices i e j . Um exemplo de rede pequeno-mundo é apresentado na Figura 2.14(b);

- **Redes Livres de Escala:** Este modelo de rede baseia-se em características bem definidas presentes na dinâmica de sua formação. A principal delas é a de um determinado vértice conectar-se, preferencialmente, a vértices que possuem um número maior de conexões (maior grau). Barabási e Albert 1999 observaram que tal propriedade está presente em diversos sistemas reais, tais como a Internet, redes de metabolismos, redes de citações de artigos científicos, redes neurais biológicas, entre outros. Assim sendo, os autores demonstraram que a distribuição do grau de diversas redes complexas obedece à *Lei da Potência*, que é dada pela equação:

$$P(k) \sim k^{-\lambda}, \quad (2.35)$$

onde k é o número de ligações de um vértice aleatório e λ é o expoente de escala. Em outras palavras, este tipo de rede conterà muitos vértices com baixo grau de conexões e

poucos vértices com um grau elevado de conexões (denominados *hubs*), conforme pode ser observado na Figura 2.14(c).

Figura 2.14: Tipos de Redes Complexas: (a) Rede Aleatória, (b) Rede Pequeno-Mundo, (c) Rede Livre de Escala.

A Estrutura de Comunidades

Outra característica não menos importante das redes complexas é que estas podem apresentar estrutura de comunidades (tal como agrupamentos vistos em grafos simples), na qual existem muitas conexões entre elementos de um mesmo grupo e poucas conexões entre elementos de grupos distintos, conforme é observado em redes reais como a World Wide Web, redes de citações, redes de transporte, redes de telefonia, redes sociais, entre outras. Um exemplo de comunidades em uma rede complexa é ilustrado na Figura 2.15.

Uma vez que estas redes podem modelar sistemas reais, a detecção destes determinados agrupamentos nestes sistemas pode levar a descoberta de características específicas e à análise da organização e evolução dinâmica dos mesmos. Algoritmos para detecção de comunidades são, na maior parte dos casos, baseados em técnicas utilizadas em aprendizado não-supervisionado [Jain e Dubes 1988], [Radicchi et al. 2004]. Estes algoritmos necessitam de uma métrica para avaliar a qualidade de uma determinada partição da rede (ou seja, se a estrutura de comunidades encontrada pode ou não ser considerada relevante). Uma das métricas mais utilizadas é a *Medida de Modularidade* proposta em Newman e Girvan 2004. A modularidade Q foi originalmente definida como:

$$Q = \sum_i (e_{ii} - a_i^2), \quad (2.36)$$

Figura 2.15: Estrutura de comunidades da rede de relações entre membros do clube de karatê exibida na Figura 2.13. 5 comunidades são observadas.

onde e_{ii} representa a fração das arestas da rede que estão inseridas dentro da comunidade i , $a_i = \sum_j e_{ij}$, onde e_{ij} é a fração das arestas que conectam vértices da comunidade i a vértices da comunidade j . Em redes cujas arestas são inseridas sem a preocupação de se saber quais as comunidades dos vértices que estão sendo conectados, pode-se considerar $e_{ij} = a_i a_j$. Assim, a_i^2 é considerada a fração dos vértices que são inseridos na comunidade i com probabilidade aleatória.

Em outras palavras, a função mede o quanto a partição encontrada se aproxima (ou se afasta) de uma formação aleatória. Valores próximos de 0 indicam baixa probabilidade da rede estar dividida em comunidades relevantes, visto que tais agrupamentos não diferem da casualidade de sua formação. Assim, quanto mais o resultado for positivo e distante de 0, mais relevante será a estrutura de comunidades encontrada. Segundo os autores, $Q \geq 0,3$ já reflete uma estrutura de comunidades significativa.

Inspirados na novidade computacional do conceito de modularidade, vários outros trabalhos surgiram utilizando, ou propondo melhorias em sua definição e aplicação [Newman 2004a], [Newman 2004b], [Clauset 2005], [Duch e Arenas 2005], [Muff et al. 2005], [Massen e Doye 2005], [Newman 2006b], [Newman 2006a], [Reichardt e Bornholdt 2006], [Newman 2008], [Newman 2011]. Uma versão mais precisa desta medida foi apresentada em Clauset et al. 2004, como

segue:

$$Q = \frac{1}{2e} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2e} \right] \delta(c_i, c_j), \quad (2.37)$$

onde e é o número de arestas da rede, A_{ij} é o elemento da linha i e coluna j da matriz de adjacência A (ou seja, A_{ij} retorna 1 caso exista conexão entre os vértices i e j e 0 caso contrário), k_i é o grau do vértice i (isto é, o número de vértices com os quais o vértice i está conectado), k_j é o grau do vértice j , c_i é o componente (comunidade) do vértice i , c_j é o componente (comunidade) do vértice j . A função δ retorna 1 caso i e j pertençam ao mesmo componente (comunidade) e 0 caso contrário. O somatório percorre todos os pares de vértices i, j da rede. Nota-se que esta fórmula privilegia a fração de arestas que estão inseridas em uma mesma comunidade. Mais uma vez, $Q \geq 0,3$ indica uma estrutura de comunidades válida. No caso das comunidades observadas na Figura 2.15, $Q = 0,3532216$.

Este projeto de pesquisa não pretende conceber ou implementar algoritmos de detecção de comunidades, mas utiliza a modularidade definida na Equação 2.37 para validar as partições formadas na rede neural proposta, uma vez que as sinapses originais de Hopfield serão quebradas para se gerar a estrutura de comunidades. No capítulo a seguir, serão apresentados alguns trabalhos que, assim como este, estendem a arquitetura original de Hopfield visando obter, ou maior similaridade com algum aspecto do funcionamento do cérebro, ou maior eficiência e precisão em relação ao modelo original.

Extensões de Redes Recorrentes

Este capítulo versa sobre projetos de pesquisa que estendem modelos dinâmicos de RNAs, ou para aferirem maior desempenho e precisão, ou para oferecerem maior grau de similaridade com certas características do funcionamento das redes neurais biológicas. Atualmente, as tendências na área de RNAs mapeiam-se, basicamente, em duas vertentes: (i) estudo e modelagem de diferentes tipos de unidades neuronais artificiais, algumas com a intenção de atingir maior plausibilidade em relação a um neurônio biológico e (ii) estudo e modelagem de diferentes regras de aprendizado e topologias, algumas visando maior semelhança à estrutura organizacional do cérebro.

No primeiro caso, modelos de redes neurais são implementados segundo características que esboçam um comportamento dinâmico, e até caótico, observadas em células nervosas, levando em conta, entre outras, propriedades como intensidade e atraso do pulso elétrico ao percorrer a membrana celular. Redes neurais que se baseiam em tal modelagem neuronal são normalmente conhecidas como *Redes Pulsadas* (do inglês *Spiking Neural Networks*) [Maass 1997], [Maass e Bishop 2001]. Estes modelos não são foco de aprofundamento neste trabalho.

Este projeto de pesquisa contempla o segundo cenário, ou seja, a elaboração de modelos de RNAs cuja proposta significativa se concentra no aperfeiçoamento da estrutura topológica. Assim sendo, este capítulo apresenta algumas ideias de trabalhos correlacionados, que, de forma geral, referem-se à aperfeiçoamentos de modelos de redes recorrentes, mais especificamente do modelo original de Hopfield. Embora muitos existam, três são evidenciados devido a sua relevância: a Seção 3.1 descreve, ainda que superficialmente, o modelo de Rede Neural Celular [Chua e Yang 1988] que, ao estender o modelo clássico de Hopfield, promove acoplamentos locais (ao invés de globais). A Seção 3.2 discute as ideias apresentadas em Stauffer et al. 2003, onde neurônios são adicionados ao modelo original de Hopfield para formar uma Rede Livre de Escala. Por fim, modelos de Redes Neurais Caóticas, que introduzem dinâmicas complexas à

modelos recorrentes tradicionais, são investigados na Seção 3.3.

3.1 Redes Celulares

Proposto em 1988 por Chua e Yang no artigo Chua e Yang 1988, este modelo de RNA estende o modelo original proposto por Hopfield, visando eliminar algumas de suas limitações, assim como fornecer-lhe maior similaridade no aspecto de processamento massivamente paralelo do cérebro. Algumas das limitações do modelo de Hopfield são: (i) a possível formação de estados indesejáveis (*espúrios*) na rede e (ii) a capacidade de armazenamento experimentalmente limitada em $0,15N$ padrões (onde N é o número de neurônios da rede). Aponta-se, também, que o fato de neurônios serem acoplados globalmente e conectados de forma simétrica enfraquecem o modelo no que tange à similaridade com relação à estrutura topológica do cérebro.

Neste sentido, as Redes Neurais Celulares (RNCs) alteram a topologia de Hopfield para tentar explorar o paralelismo massivo, passível de implementação em hardware, restringindo as conexões de um neurônio a uma certa vizinhança do mesmo, ao invés da arquitetura de conexões globalmente acopladas. Esta configuração gera uma rede reticulada com acoplamentos locais. O nome da rede é uma referencia aos *Autômatos Celulares* (sistemas dinâmicos discretos localmente acoplados) cuja semelhança está em relação à conectividade local entre os neurônios.

Cada unidade da rede é chamada de *célula*. Dada uma RNC com estrutura bidimensional contendo $M \times N$ células, a vizinhança de uma célula $C(i, j)$ (célula na posição de linha i e coluna j) é definida como:

$$N_r(i, j) = \{C(k, l) : \max(|k - i|, |l - j|) \leq r\}, \quad 1 \leq k \leq M, \quad 1 \leq l \leq N, \quad (3.1)$$

onde r é chamado de *raio* da vizinhança em torno da célula. Assim, de acordo com o valor de r , pode-se definir vários níveis de vizinhança para uma mesma célula. Definida a vizinhança, a lei que rege o comportamento dinâmico para busca e recuperação de informação é dada pelas seguintes equações:

$$\begin{cases} y_{kl} = f(x_{kl}) \\ \frac{dx_{ij}(t)}{dt} = -x_{ij}(t) + \sum_{(k,l) \in N_r(i,j)} A_{ij,kl} y_{kl}(t) + \sum_{(k,l) \in N_r(i,j)} B_{ij,kl} u_{kl}(t) + I(t), \end{cases} \quad (3.2)$$

para $1 \leq i \leq M$ e $1 \leq j \leq N$, onde $x_{ij} : \mathbb{R} \rightarrow \mathbb{R}$ é o estado interno da célula $C(i, j)$, $u_{ij} \in \mathbb{R}$ representa um estímulo externo, $y_{ij} \in [-1, 1]$ é a saída da célula, I é o limiar de ativação (*bias*, ou *threshold*), A é a matriz de pesos de retro-alimentação, que trata as conexões das entradas y_k

provenientes de células vizinhas, e B é a matriz de pesos de controle, que pondera os estímulos externos u_k . Esta equação ainda pode ser simplificada e reescrita em forma vetorial, como segue.

$$\begin{cases} y = f(x) \\ \dot{x} = -x + Ty + I, \end{cases} \quad (3.3)$$

onde T representa uma matriz esparsa de interconexões, referindo-se ao tratamento de conexões de retro-alimentações (elimina-se, portanto, a influência de estímulos externos). A função de ativação f é definida como a função *Saturação*:

$$f(x_i) = \begin{cases} -1, & \text{se } x_i < -1; \\ x_i, & \text{se } -1 \leq x_i \leq 1; \\ 1, & \text{se } x_i > 1, \end{cases} \quad (3.4)$$

Entre as aplicações de RNCs, a área de Processamento de Imagens recebe especial destaque, pelo fato de que células da retina apresentam um padrão de interconexão local [Werblin et al. 1995]. Assim, estas redes tornam-se adequadas para modelarem certas funções como detecção de movimento, extração de contornos e suavização de imagens. Porém, RNCs também podem funcionar como memórias associativas, sendo treinadas para o reconhecimento de padrão baseado em uma versão distorcida do mesmo. Para tanto, a fase de treinamento deve ajustar os pesos sinápticos da rede, a fim de que padrões de memória sejam armazenados como pontos fixos assintoticamente estáveis da Equação 3.3. Existem vários métodos propostos na literatura da área. Citam-se: (i) baseado em matriz pseudo-inversa [Grassi 2001], (ii) baseado em decomposição por valor singular (SVD) [Liu e Michel 1994], (iii) baseado em Perceptrons [Liu e Lu 1997], (iv) baseado em LMIs [Park et al. 2001] e (v) baseado em aprendizado Hebbiano [Szolgay et al. 1997]. A seguir, será apresentado o método baseado no modelo de aprendizado de Hebb proposto em Szolgay et al. 1997.

A ideia do método é garantir pontos de equilíbrio dentro de uma *região de saturação*. Estimando-se o valor de uma constante E que define tal região, estima-se também a capacidade de recuperação da informação armazenada. Considerando o armazenamento de p padrões de memória em uma RNC $M \times N$, onde $n = M \times N$ é o número total de células da rede, escreve-se a equação da dinâmica da célula como segue:

$$\begin{cases} y^p = f(x^p) \\ x^p = Ty^p + I, \quad p = 1, \dots, m, \end{cases} \quad (3.5)$$

onde $y^p \in \{-1, 1\}$ representa os p padrões a serem armazenados e $x^p \in C^n(y^p)$ (com $C^n(y) =$

$\{x \in \mathbb{R}^n : x_i, y_i > 1, i = 1, \dots, n\} \forall y \in B^n$) os pontos fixos associados. Considerando o domínio de x^p pertencente ao conjunto $C^p(y^p)$, conforme definido acima, esta equação pode ser reescrita como um sistema de desigualdades da forma:

$$x_i^p y_i^p > 1, \quad i = 1, \dots, n \quad e \quad p = 1, \dots, m, \quad (3.6)$$

onde x_i^p é o i -ésimo elemento de x^p e é calculado como:

$$x_i^p = T_i y^p + I_i, \quad i = 1, \dots, n \quad e \quad p = 1, \dots, m, \quad (3.7)$$

em que T_i refere-se à i -ésima linha da matriz I e I_i o i -ésimo elemento do vetor de *thresholds*. Com a hipótese de que x_i^p existem dentro de uma região de saturação limitada por uma constante E (pois estes são pontos de equilíbrio), tem-se que:

$$(T_i y^p) y_i^p + I_i y_i^p > E > 1, \quad i = 1, \dots, n \quad e \quad p = 1, \dots, m. \quad (3.8)$$

Considerando que as conexões da RNC são locais, haverá elementos T_i em T iguais a zero. Da mesma forma, elementos y_i^p em y_p que são multiplicados pelos elementos nulos de T_i também serão nulos. Sendo T_i^* um novo vetor contendo apenas elementos não-nulos de T_i e sendo y^{*p} o vetor obtido a partir da eliminação dos elementos nulos de y^p , a Inequação 3.8 fica:

$$(T_i^* y^{*p}) y_i^p + I_i y_i^p > E > 1, \quad i = 1, \dots, n \quad e \quad p = 1, \dots, m. \quad (3.9)$$

Sendo:

$$\begin{cases} T_i^* = [T_i^*, I_i] \\ y^{*p} = [y^{*p} \cdot y_i^p, y_i^p], \end{cases} \quad (3.10)$$

a Equação 3.9 pode ser reescrita como:

$$T_i^* y^{*p} > E > 1, \quad i = 1, \dots, n \quad e \quad p = 1, \dots, m. \quad (3.11)$$

Considerando a definição da Hipótese de Hebb (ver Capítulo 2, Subseção 2.2.3), tem-se que:

$$T_i^* = \sum_{p=1}^m w_i^p (y_i^{-P})^T, \quad i = 1, \dots, n, \quad (3.12)$$

onde w_i^p representa um “fator de ganho”(ou taxa de aprendizado). Assim, os pontos de equilíbrio podem ser encontrados pela Equação 3.11 utilizando-se o valor de T_i^* encontrado na Equação 3.12 (pois serão calculadas as correlações entre os padrões de memória).

Desta forma, o modelo de RNCs estende o modelo de Hopfield em todas as suas funcionalidades, agregando-lhe novas características e capacidades. Um outro trabalho que também investiga possíveis modificações na estrutura topológica original do modelo de Hopfield como forma de lhe proporcionar maior plausibilidade biológica é discutido na seção seguinte.

3.2 Rede Hopfield com Topologia Livre de Escala

O trabalho de Stauffer et al. 2003 discute os resultados obtidos em tarefas de reconhecimento de padrão de uma Rede Neural de Hopfield na qual N novos neurônios são adicionados à m neurônios que formam uma rede de Hopfield, a fim de se gerar a arquitetura de uma *Rede Livre de Escala* [Barabási e Albert 1999], ou seja, neurônios são adicionados de forma que a distribuição de graus da rede segue a *Lei da Potência* (neurônios “ricos” ficam mais “ricos”). A justificativa é a de que observa-se a arquitetura livre de escala na topologia cerebral.

O modelo utiliza as mesmas equações originais para aprendizado (Equação 2.30 e recuperação de informação (Equação 2.29), porém restringindo as conexões sinápticas de um neurônio aos vizinhos dentro da topologia livre de escala. Os autores mostram que quanto maior o número de neurônios inseridos na rede original de Hopfield, menor será a eficiência do reconhecimento de padrão. Por outro lado, a taxa de reconhecimento de padrão aumenta quando o número m de neurônios que formam o núcleo da rede (neurônios que integram a rede de Hopfield original) aumenta.

Segundo os autores, ainda não está claro se a eficiência do reconhecimento resulta apenas do uso de altos valores médios para m ou se a adição de novos “ingredientes” também influenciariam o resultado. Mais importante é que este trabalho de pesquisa mesclou duas grandes áreas: Redes Neurais Artificiais e Redes Complexas ao analisar a qualidade do reconhecimento de padrão de um modelo de Hopfield onde neurônios são adicionados para formar uma topologia que também se observa no cérebro.

3.3 Redes Neurais Caóticas

Uma forma interessante de se estudar a neurodinâmica cerebral é gerar modelos artificiais que reproduzam, ao menos em parte, as complexas dinâmicas observadas no funcionamento do córtex. As *Redes Neurais Caóticas* são modelos de RNA recorrentes que provocam a ocorrência de atratores dinâmicos diversos (não somente estacionários) e possíveis pontos de bifurcação, tal como ocorre no cérebro, conforme discutido no capítulo anterior, Subseção 2.2.2.

Várias são as formas de se provocar dinâmicas (especialmente caóticas) em um modelo de

rede recorrente, como o de Hopfield. Algumas delas, utilizadas de forma isolada ou combinadas com outras, são: (i) uso de neurônios caóticos [Adachi e Aihara 1997], [Matykiewicz 2004], [Zhao et al. 2006], uso de diferentes regras de aprendizado [Nara et al. 1993], uso de diferentes funções de ativação [Zheng et al. 2010], [Huang e Huang 2011], [Mathias e Rech 2012], alteração da estrutura topológica [Kaslik e Balint 2009], [Marichal et al. 2012], entre outras.

A maioria dos modelos de Redes Neurais Caóticas opera em duas fases: uma caótica e outra não-caótica. Inicialmente, a rede opera na fase caótica, na qual os neurônios exibem trajetórias *itinerantes* que vagam pelo espaço de estados e visitam proximamente todos os padrões armazenados. Este comportamento, intrínseco a um sistema caótico, é conhecido como *itinerância caótica* [Kaneko e Tsuda 2003]. Quando as órbitas se aproximam suficientemente de um padrão desejado (de acordo com alguma medida de similaridade adotada), a rede altera seu comportamento para a fase não-caótica, onde ela se comporta como uma memória associativa convencional, fazendo com que as trajetórias convirjam para o ponto fixo referente ao padrão desejado.

Além da simulação dos aspectos dinâmicos do cérebro, outra motivação para pesquisas nesta área é o possível aumento na eficiência em tarefas de reconhecimento de padrão, pois a dinâmica estacionária, utilizada em modelos recorrentes convencionais, pode apresentar pontos indesejáveis (espúrios) de memória e levar ao reconhecimento impreciso de padrões. O uso e a combinação de atratores diversos, principalmente o caótico, proporciona uma forma para se evitar estes pontos espúrios, uma vez que a manifestação da itinerância caótica impede que as órbitas dos neurônios sejam capturadas por um mínimo local.

Tem havido grande interesse na pesquisa de modelos de RNAs caóticas, sendo estas utilizadas para os mais variados fins, como: (i) simular a dinâmica cerebral [Zhao et al. 2006], [Tokuda et al. 2010], (ii) aprimorar tarefas de reconhecimento associativo de padrões [Zhao et al. 2006], [Kaslik e Balint 2009], (iii) desenvolver novos algoritmos de aprendizado [Lysetskiy e Zurad 2004], entre outros. A seguir, é abordada a pesquisa de Adachi e Aihara 1997, um dos trabalhos seminais da área, e que investiga o aparecimento de comportamentos dinâmicos alternativos pelo uso de um modelo de neurônio caótico. A descrição deste trabalho neste documento possui a intenção de ilustrar a análise das dinâmicas e itinerância caóticas em um modelo de Rede Neural Caótica.

Rede Neural Caótica com Modelo de Neurônio Caótico

O trabalho de Adachi e Aihara 1997 apresenta a análise de um modelo de RNA com memória associativa (como o modelo de Hopfield), porém composto por um modelo de neurônio, dito caótico, cuja auto-recorrência produz dinâmicas alternativas ao do atrator estacionário utilizado

em modelos recorrentes tradicionais. Tal análise não considera eficiência ou precisão da rede, mas, pelo contrário, requer que o processo de recuperação de padrão fuja do ponto de equilíbrio a fim de simular alguns aspectos de associatividade presentes no modelo real biológico. Esta Subseção não pretende detalhar todo o processo de análise realizado pelos autores, apenas apresentar pontos que enfatizam a possibilidade de que demais dinâmicas possam estar presentes em uma rede recorrente, partindo-se de um modelo tradicional.

O modelo do neurônio caótico proposto é dado pela seguinte equação:

$$x(t) = f[u(t)], \quad (3.13)$$

onde $x(t) \in [0, 1]$ é a saída do neurônio no tempo discreto t ($t = 0, 1, 2, \dots$), f representa a função de ativação e $u(t) \in \mathbb{R}$ é o estado interno do neurônio no tempo discreto t , que é calculado da seguinte forma:

$$u(t+1) = A(t) - \alpha \sum_{d=0}^t k^d g[x(t-d)] - \theta, \quad (3.14)$$

onde $A(t)$ é o chamando *estímulo externo* no tempo discreto t , g é a chamada *função de obstinação* (pois preserva as $t-d$ saídas anteriores do neurônio), α é o parâmetro de escala da obstinação, k é o parâmetro de decaimento da obstinação e θ é o limiar de ativação (*threshold*). Substituindo 3.14 na Equação 3.13, tem-se a equação do neurônio caótico como segue:

$$x(t+1) = f \left\{ A(t) - \alpha \sum_{d=0}^t k^d g[x(t-d)] - \theta \right\}, \quad (3.15)$$

na qual, nota-se, diferentemente de modelos matemáticos tradicionais, que este é um neurônio cujo estado atual depende dos estados anteriores do mesmo. Esta característica de auto-recorrência do próprio neurônio permite o desdobramento de dinâmicas, desde periódicas à caóticas, de acordo com valores escolhidos para os parâmetros. Para demonstrar a ocorrência destas dinâmicas, aos autores calculam os expoentes de Lyapunov dos estados internos u , porém considerando-se a diferença reduzida da Equação 3.14, dada por:

$$u(t+1) = ku(t) - \alpha g\{f[u(t)]\} + \theta(t), \quad (3.16)$$

onde $\theta(t) \equiv A(t) - kA(t-1) - \theta(1-k)$. Considera-se, ainda, $g(x) = x$, $\theta(t) = a$ (constante) e f a função logística dada por $f(x) = 1/\{1 + \exp(-u/\epsilon)\}$, com parâmetro de declividade ϵ . Os expoentes de Lyapunov para a Equação 3.16 são calculados da seguinte forma:

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=t_0+1}^{t_0+n} \ln \left| \frac{du(t+1)}{du(t)} \right|, \quad (3.17)$$

onde n é o número de iterações. Na Figura 3.1, é apresentado o conjunto dos expoentes de Lyapunov pela variação dos valores do parâmetro $\theta(t) = a$. Observa-se a ocorrência de dinâmicas caóticas (quando $\lambda > 0$) e demais dinâmicas (quando $\lambda < 0$). O diagrama de bifurcação é ilustrado na Figura 3.2.

Figura 3.1: Expoentes de Lyapunov para a Equação 3.16 pela variação do parâmetro $\theta(t) = a$, com $k = 0,5$, $\epsilon = 0,015$ e $\alpha = 1,0$. Valores positivos de λ indicam dinâmica caótica; valores negativos indicam outras dinâmicas; valores iguais a zero indicam pontos de bifurcação. Fonte: Adachi e Aihara 1997.

Figura 3.2: Diagrama de bifurcação da Equação 3.16. Gráfico do estado interno u do neurônio pela variação do parâmetro $\theta(t) = a$, com $k = 0,5$, $\epsilon = 0,015$ e $\alpha = 1,0$. Fonte: Adachi e Aihara 1997.

A partir da concepção do neurônio caótico, constrói-se uma rede neural caótica, na qual a

dinâmica no i -ésimo neurônio integrado à rede é descrita como:

$$x_i(t+1) = f[\xi_i(t+1) + \eta_i(t+1) + \zeta_i(t+1)], \quad (3.18)$$

onde ξ_i , η_i e ζ_i representam, respectivamente: (i) os termos de estado interno referente aos estímulos externos, (ii) entradas de retro-alimentação dos neurônios que constituem a rede e (iii) a *obstinação*, ou seja, a auto-recorrência do i -ésimo neurônio, levando em conta seus próprios estados anteriores. Os três componentes ξ_i , η_i e ζ_i são calculados como segue:

$$\xi_i(t+1) = \sum_{j=1}^X V_{ij} \sum_{d=0}^t k_e^d A_j(t-d), \quad (3.19)$$

$$\eta_i(t+1) = \sum_{j=1}^N W_{ij} \sum_{d=0}^t k_f^d x_j(t-d), \quad (3.20)$$

$$\zeta_i(t+1) = -\alpha \sum_{d=0}^t k_r^d g[x_i(t-d)] - \theta_i, \quad (3.21)$$

onde V_{ij} representa o peso sináptico entre o neurônio i e a j -ésima entrada externa (de um total de X entradas externas), W_{ij} representa o peso sináptico entre os neurônios i e j (de um total de N neurônios), k_e é o parâmetro de decaimento dos estímulos externos, k_f é o parâmetro de decaimento das entradas de retro-alimentação e k_r é o parâmetro de decaimento da obstinação.

De acordo com os valores dos parâmetros, a rede pode apresentar, desde um comportamento clássico de modelos recorrentes compostos pelo neurônio MCP [McCulloch e Pitts 1943] (por exemplo, com $\alpha = k_e = k_f = k_r = 0$), até redes com comportamento dinâmico caótico (para $\alpha \neq 0$; $k_e \neq 0$; $k_f \neq 0$; $k_r \neq 0$). Em outras palavras, o modelo apresentado pelos autores é uma extensão natural dos modelos de RNAs dinâmicos convencionais de tempo discreto. Isto quer dizer, que o modelo pode funcionar como uma memória associativa, tal como o modelo de Hopfield, porém incorporando dinâmicas caóticas.

Para analisar as características fundamentais do modelo proposto, os autores constroem uma RNA composta por $N = 100$ neurônios conectados recorrentemente, visando a recuperação de padrões de forma associativa. A dinâmica dada pela Equação 3.18 é simplificada, eliminando-se a componente ξ_i (que opera sobre estímulos externos). Portanto, a saída do i -ésimo neurônio da rede é calculada como sendo:

$$x_i(t+1) = f[\eta_i(t+1) + \zeta_i(t+1)], \quad (3.22)$$

preservando o cálculo das componentes η_i e ζ_i conforme apresentado nas equações 3.20 e 3.21,

respectivamente. O treinamento é feito utilizando-se uma matriz simétrica auto-associativa como segue:

$$W_{ij} = \sum_{p=1}^M (\xi_i^p - \bar{\xi})(\xi_j^p - \bar{\xi}), \quad (3.23)$$

onde ξ_i^p é o i -ésimo elemento do padrão p , ξ_j^p é o j -ésimo elemento do padrão p e $\bar{\xi}$ representa a média dos valores espaciais dos M padrões armazenados. Na figura 3.3 é ilustrada a base de padrões utilizada. Os quatro primeiros, de (a) a (d), são utilizados no processo de armazenamento. O quinto padrão (e) é utilizado para recuperação associativa de um dos quatro padrões armazenados.

Figura 3.3: Base utilizada na análise do modelo proposto. Padrões de (a) a (d) são armazenados e o padrão (e) utilizado para recuperação associativa dos padrões armazenados. Fonte: Adachi e Aihara 1997.

De acordo com os valores utilizados para os parâmetros α , k_f e k_r pode-se obter diferentes comportamentos dinâmicos. Na Figura 3.4, observa-se a completa convergência do padrão exibido em $t = 0$ para o padrão armazenado (c) da Figura 3.3, demonstrando a dinâmica estacionária para $\alpha_i = k_f = k_r = 0$. Assim, com valores nulos para os parâmetros, a rede comporta-se como um modelo de memória associativa tradicional. Porém, ao passo que o grau de obstinação aumenta (incremento de k_r), a rede passa a não mais apresentar um ponto de equilíbrio estável. As Figuras 3.5 e 3.6 ilustram o processo de recuperação do padrão distorcido em $t = 0$. Na primeira, com um grau baixo de obstinação ($\alpha = 8,0$, $k_f = 0,15$ e $k_r = 0,5$), nota-se que a rede, algumas vezes, aproxima-se do padrão armazenado (c), mas não o mantém estável, como em um comportamento periódico, ou quase-periódico. Na segunda figura, já com um alto grau de obstinação ($\alpha = 10,0$, $k_f = 0,2$ e $k_r = 0,9$), nota-se que o processo de recuperação de informação se aproxima, ora ou outra, de cada um dos padrões armazenados. A evolução descrita na Figura 3.6 é um exemplo de itinerância caótica.

Análises mais profundas presentes no artigo deduzem que tais trajetórias itinerantes e transitórias movem-se não-periodicamente sem nunca estacionarem em um ponto de equilíbrio. Porém, após vários ciclos de iteração, convergem para um atrator periódico atingindo uma certa estabilidade. Isto demonstra que a introdução do fator de obstinação no modelo do

Figura 3.4: Demonstração da dinâmica estacionária para $\alpha = k_f = k_r = 0$. O padrão visto em $t = 0$ converge para o padrão armazenado (c) da Figura 3.3. Fonte: Adachi e Aihara 1997.

Figura 3.5: Saídas da rede durante o ciclo de recuperação de padrão para $\alpha = 8, 0$, $k_f = 0, 15$ e $k_r = 0, 5$. Nota-se um comportamento periódico ou quase-periódico, no qual a rede, algumas vezes, aproxima-se do padrão (c) da Figura 3.3, porém não o mantém estável. Fonte: Adachi e Aihara 1997.

neurônio torna o processo de busca e recuperação de padrão muito mais dinâmico do que em modelos recorrentes tradicionais. Tal dinamismo mostrou-se similar ao comportamento do sistema olfativo biológico [Skarda e Freeman 1987], no qual dinâmicas caóticas exercem um papel fundamental, não somente para o acesso a um padrão de memória armazenado, mas também para classificar e aprender novos padrões.

Figura 3.6: Saídas da rede durante o ciclo de recuperação de padrão para $\alpha = 10, 0$, $k_f = 0, 2$ e $k_r = 0, 9$. Nota-se uma trajetória semelhante a de uma itinerância caótica, na qual cada um dos quatro padrões, ora ou outra, é visitado pela rede. Fonte: Adachi e Aihara 1997.

O Modelo de Rede Recorrente com Estrutura de Comunidades

Tal como o trabalho de Stauffer et al. 2003, este projeto altera a estrutura sináptica padrão do modelo de Hopfield a fim de gerar uma topologia de Rede Complexa. Porém, enquanto Stauffer et al. 2003 analisa a eficiência de uma rede neural livre de escala, esta pesquisa gera estrutura de comunidades de Rede Complexa, cuja inspiração é a existência de agrupamentos espaciais e funcionais observados no córtex cerebral. Os resultados aqui apresentados sugerem que é possível a emergência de dinâmicas complexas (especificamente a caótica) quando a rede exibe uma topologia de comunidades e a utilização de dinâmicas complexas pode contribuir para melhorar a eficiência e precisão de modelos recorrentes, conforme discutido no trabalho de Zhao et al. 2006.

Assim, este projeto de pesquisa se mostra relevante ao propor um novo modelo de Rede Neural Caótica contendo uma estrutura topológica mais similar a do cérebro. A Seção 4.1 aborda a metodologia utilizada para implementação e análise do modelo de RNA proposto. A Seção seguinte, 4.2, apresenta algumas análises matemáticas que basicamente abordam consequências da fragmentação das conexões sinápticas. Por fim, os experimentos, simulações e resultados são discutidos na Seção 4.3. Tais resultados geraram um artigo científico aceito para publicação no *IEEE International Joint Conference on Neural Networks* (IJCNN) de 2013 (ver Seção 5.2).

4.1 Metodologia

Esta Seção descreve os métodos e técnicas empregados neste projeto de pesquisa. A Subseção 4.1.1 apresenta a etapa de passos seguidos para a geração da topologia de rede pro-

posta. A seguir, a Subseção 4.1.2, discute a regra de aprendizado utilizada que posta os padrões em ciclos gerando uma matriz assimétrica de pesos. A regra de propagação de padrão é abordada na Subseção 4.1.3. Por fim, a Subseção 4.1.4 disserta sobre as métricas utilizadas para a avaliação do comportamento dinâmico do modelo.

4.1.1 Geração da Topologia de Rede

Para gerar a RNA com estrutura de comunidades, a topologia convencional de uma rede de Hopfield (formada por N neurônios e número de conexões $e = N^2 - N$) é fragmentada levando-se em conta três parâmetros:

- C : Número de comunidades desejadas;
- P_{in} : Valor entre 0 e 1 que representa a probabilidade de um neurônio conectar-se a outro da mesma comunidade (conexões *intra-comunidades*);
- P_{out} : Valor entre 0 e 1 que representa a probabilidade de um neurônio conectar-se a outro de uma comunidade diferente (conexões *inter-comunidades*).

Após definidos estes parâmetros, os neurônios são rotulados em comunidades, como segue: $Comunidade_do_neuronio_i = i \% C$, onde $i = 1, 2, \dots, N$ e $\%$ representa a operação *mod* (nota-se, portanto, que as comunidades terão, no geral, o mesmo número de neurônios - obviamente que uma das comunidades conterá um diferente número de neurônios dependendo da relação entre N e C). Em seguida, uma matriz de adjacência A é construída aplicando-se as probabilidades P_{in} e P_{out} na ligação entre os neurônios: para cada par de neurônios (i, j) , com $i \neq j$, gera-se um número aleatório Z entre 0 e 1. Se $Z \leq P$ ($P = P_{in}$ se i e j pertencem à mesma comunidade; $P = P_{out}$ caso contrário), então $A_{ij} = 1$. Se $Z > P$, $A_{ij} = 0$. Assim, a matriz A guarda a topologia da nova RNA, que, de agora em diante, será referenciada como $RNA_{(Com)}$.

Devido à natureza estocástica na formação das sinapses, nota-se que a matriz A pode variar de conteúdo para o mesmo conjunto de valores de parâmetros, incluindo a possibilidade de originar uma topologia com grupos desconexos (fato que prejudicaria a convergência do modelo a um atrator). Por esta razão, força-se conexão entre os neurônios i e $i + 1$ (e entre o último e o primeiro) gerando, assim, uma sempre existente estrutura de ciclo, mesmo que os valores de P_{in} e P_{out} sejam nulos. Com a matriz A em mãos, a análise da estrutura de comunidades é feita por meio da *Medida de Modularidade* Q (Equação 2.37), como apresentado no Capítulo 2, Seção 2.4. A Figura 4.1 ilustra um exemplo da $RNA_{(Com)}$ composta por 100 neurônios ($N = 100$) e parametrizada para originar uma estrutura com 4 comunidades ($C = 4$).

Figura 4.1: Exemplo da topologia do modelo de RNA proposto para $N = 100$, $C = 4$, $P_{in} = 1,00$ e $P_{out} = 0,10$. As comunidades contêm 25 neurônios cada. A Medida de Modularidade para esta topologia é $Q = 0,452569$.

4.1.2 Regra de Aprendizado

Uma vez definida e gerada a topologia, dá-se início ao treinamento dos padrões. Para isto, é utilizada a regra de aprendizado que posta os padrões em ciclos, proposta em Mori et al. 1989. Trata-se de uma regra para redes recorrentes e que exige uma propagação síncrona dos padrões (isto é, diferentemente da evolução assíncrona de Hebb, cujo valor de saída de um neurônio é imediatamente atualizado para uso dentro de uma mesma iteração, na evolução síncrona os neurônios terão seus valores de saída atualizados apenas após a evolução de todos os demais neurônios e antes do início da iteração seguinte). Por meio de uma hetero-correlação entre os padrões, uma matriz assimétrica de pesos W é gerada conforme mostra a Equação 4.1, na qual K é o número de ciclos e L é o número de padrões em cada ciclo ($K \times L = M$ é o total de padrões a serem armazenados). A Figura 4.2 ilustra a organização de 10 padrões divididos em 2 ciclos.

$$W = \frac{1}{N} \sum_{\mu=1}^K \sum_{v=1}^L \xi^{\mu, \{v+1\} \% L} (\xi^{\mu, v})^T \quad (4.1)$$

onde $\xi^{\mu,v} = (\xi_1^{\mu,v}, \xi_2^{\mu,v}, \dots, \xi_N^{\mu,v})$, com $\xi_i^{\mu,v} \in \{-1, 1\}$, representa o v -ésimo padrão do ciclo μ . % é o operador *mod*. Para $L = 1$ tem-se o comportamento da regra de autocorrelação convencional.

Figura 4.2: Organização de 10 padrões ($M = 10$) em 2 ciclos ($K = 2$). Cada ciclo contém 5 diferentes padrões ($L = 5$).

A razão para a alteração da regra de aprendizado padrão de Hopfield se faz pelo fato de que a matriz simétrica de pesos gerada pela autocorrelação entre os padrões (como feito no aprendizado Hebbiano) provê apenas dinâmicas estacionárias ou periódicas de, no máximo, período 2 [Goles-Chacc et al. 1985], [Marcus e Westervelt 1989]. Já a regra da Equação 4.1 gera uma matriz assimétrica por dividir os padrões em ciclos. Considerando que cada ciclo contém L padrões, o sistema evolui de forma naturalmente periódica, recuperando padrões com uma dinâmica periódica de período L . Além disso, foi demonstrado que uma matriz assimétrica de pesos sinápticos facilita a emergência de dinâmicas caóticas dependendo da estrutura topológica que rede apresenta [Nara et al. 1993].

4.1.3 Regra de Propagação

A fase de propagação é aquela na qual padrões distorcidos ξ' são apresentados à rede e neurônios evoluem de acordo com uma regra de evolução. Ao final do processo, a rede gera como resposta um padrão ξ que pode ou não corresponder fielmente a um dos padrões previamente armazenados, dependendo da precisão do sistema. A regra aqui utilizada exige uma evolução síncrona, conforme discutido na Subseção anterior, e é dada pela Equação 4.2 a seguir.

$$\begin{cases} x_i(t+1) = f(u_i(t)), \\ u_i(t) = \sum_{j=1}^N A_{ij} W_{ij} x_j(t) \quad j \neq i \text{ e } i = 1, \dots, N. \end{cases} \quad (4.2)$$

na qual x_i é a saída do neurônio i , u_i é o estado interno do neurônio i , N é o número total de neurônios, A_{ij} é o elemento (i, j) da matriz de adjacência A ($A_{ij} = 1$ se existe conexão entre os neurônios i e j e $A_{ij} = 0$ caso contrário) e f é a função de ativação *sinál*, a mesma dada pela Equação 2.27, conforme apresentado no Modelo de Hopfield, Seção 2.3. Por meio da geração

da matriz A (que codifica a topologia da rede), pode-se estimular diferentes comportamentos dinâmicos durante a evolução dos neurônios, de periódicos à caóticos, conforme discutido na Seção 4.3.

4.1.4 Análise do Comportamento Dinâmico

Para analisar o comportamento dinâmico do modelo proposto foram utilizados diagramas de bifurcação, coeficientes de Fourier das séries temporais advindas da propagação de padrões e, também, os expoentes de Hurst, calculados por meio do método de *rescaled range* (ver Subseção 2.1.5). Os expoentes de Hurst foram preferidos aos expoentes de Lyapunov (tradicionalmente empregados para avaliação da dinâmica de sistemas) pois o modelo de $RNA_{(Com)}$ proposto baseia-se em uma regra de propagação que assume valores discretos e que emprega uma função de ativação descontínua, fato que impossibilita o cálculo da matriz Jacobiana D , conforme apresentado na Subseção 2.1.5.

4.2 Análise do Modelo

Algumas análises matemáticas consideradas relevantes são descritas nesta Seção. A Subseção 4.2.1 apresenta uma estimativa para o número total de conexões da rede quando esta se encontra em um contexto de comunidades. Um estudo sobre o comportamento dos parâmetros utilizados para construção da rede em situação de caos é abordada na Subseção 4.2.2.

4.2.1 Cálculo do Número de Conexões

É importante observar que a $RNA_{(Com)}$ gerada conterá o mesmo número N de neurônios da rede original de Hopfield. O número de conexões e , porém, estará implícito na matriz A , mas pode ser estimado analiticamente como segue.

- Número de conexões intra-comunidades e_{intra} : considerando $P_{in} = 1$, um neurônio faz $(N/C) - 1$ conexões em sua comunidade i . Assim, todos os neurônios da comunidade i somarão $[(N/C) - 1](N/C)$ conexões. Portanto, para todas as comunidades da rede tem-se $[(N/C)^2 - (N/C)]C = (N^2/C) - N$ conexões. Considerando-se, então, uma probabilidade $P_{in} < 1$, o número de conexões intra-comunidades pode ser estimado como sendo:

$$e_{intra} = \left(\frac{N^2}{C} - N\right)P_{in} \quad (4.3)$$

- Número de conexões inter-comunidades e_{inter} : considerando $P_{out} = 1$, um neurônio faz $N - (N/C)$ conexões com neurônios de comunidades diferentes da sua. Todos os neurônios da rede somarão, portanto, $[N - (N/C)]N = (N^2) - (N^2/C)$ conexões. Pode-se estimar, então, o número total de conexões inter-comunidades para uma dada probabilidade $P_{out} < 1$ como:

$$e_{inter} = (N^2 - \frac{N^2}{C})P_{out} \quad (4.4)$$

- Número total de conexões e : o número total de conexões da $RNA_{(Com)}$ é dado por $e = e_{intra} + e_{inter}$. Assim, e pode ser estimado pela Equação 4.5 a seguir.

$$e = (\frac{N^2}{C} - N)P_{in} + (N^2 - \frac{N^2}{C})P_{out} \quad (4.5)$$

4.2.2 Relação entre Parâmetros na Condição de Bifurcação

Considera-se que a RNA_{Com} proposta altera seu comportamento dinâmico quando alcança determinado grau de fragmentação de suas conexões. Porém, dada à natureza estocástica da matriz A , o ponto de bifurcação torna-se, na realidade, uma *região de bifurcação*, pois diferentes resultados podem ser obtidos para a mesma configuração dos parâmetros P_{in} , P_{out} e C no contexto de um grau de fragmentação crítico.

Assim sendo, definindo uma probabilidade $p^{(c)}$ que corresponde à *probabilidade crítica* na qual existe uma *região de bifurcação*, o cálculo do número crítico de conexões (isto é, o número de conexões para a qual o sistema encontra-se na região de bifurcação) pode ser definido como:

$$e^{(c)} = (N^2 - N)p^{(c)} \quad (4.6)$$

É interessante analisar a relação entre os parâmetros no caso em que a rede encontram-se na região de bifurcação. Para isto, considera-se que os neurônios da rede mantêm uma estrutura de comunidades, isto é: $C > 1$ e $P_{out} < 1$. Na região de bifurcação, o número total de conexões e deve corresponder ao número crítico de conexões e^c , isto é: $e = e^{(c)}$. Desta forma, igualando as Equações 4.5 e 4.6, tem-se que:

$$(\frac{N^2}{C} - N)P_{in} + (N^2 - \frac{N^2}{C})P_{out} = (N^2 - N)p^{(c)} \quad (4.7)$$

No contexto da Equação 4.7, que evidencia uma situação de bifurcação, intenciona-se estudar o relacionamento entre os parâmetros C , P_{in} e P_{out} . No entanto, para simplificar a análise, considera-se também que a rede apresenta uma estrutura de comunidades válida e com um

grau máximo de modularidade, ou seja, faz-se $P_{in} = 1$ e mantém-se variáveis os valores de C e P_{out} . Nesta situação, a Equação 4.7 pode ser expandida da seguinte forma:

$$\begin{aligned}
 \frac{N^2}{C} - N + (N^2 - \frac{N^2}{C})P_{out} &= (N^2 - N)p^{(c)} \\
 \frac{N^2}{C} - N + P_{out}N^2 - P_{out}\frac{N^2}{C} &= p^{(c)}N^2 - p^{(c)}N \\
 N^2(\frac{1}{C} - \frac{P_{out}}{C} + P_{out}) - N &= p^{(c)}N^2 - p^{(c)}N \\
 N^2(\frac{1}{C} - \frac{P_{out}}{C} + P_{out}) &= p^{(c)}N^2 - p^{(c)}N + N \\
 \frac{1}{C} - \frac{P_{out}}{C} + P_{out} &= \frac{[p^{(c)}N^2 + (1-p^{(c)})N]}{N^2} \\
 \frac{1}{C} - \frac{P_{out}}{C} + P_{out} &= p^{(c)} + \frac{1-p^{(c)}}{N}
 \end{aligned} \tag{4.8}$$

Fazendo $p^{(c)} + \frac{1-p^{(c)}}{N} = \theta$:

$$\begin{aligned}
 \frac{1}{C} - \frac{P_{out}}{C} + P_{out} &= \theta \\
 \frac{1-P_{out}}{C} + P_{out} - \theta &= 0 \\
 \frac{1-P_{out}+CP_{out}-C\theta}{C} &= 0 \\
 1 - P_{out} + CP_{out} - C\theta &= 0 \\
 P_{out}(C - 1) &= C\theta - 1
 \end{aligned} \tag{4.9}$$

Tem-se, então, a relação:

$$P_{out} = \frac{\theta C - 1}{C - 1} \tag{4.10}$$

onde $\theta = p^{(c)} + (1 - p^{(c)})/N$. Esta equação traça uma curva hiperbólica na qual P_{out} converge assintoticamente para θ quando C cresce, como pode ser observado na Figura 4.3.

Figura 4.3: Curva da relação entre P_{out} e C , dada pela Equação 4.10.

A curva da Figura 4.3 indica que, em relação à caoticidade, a partir de um certo número de comunidades, o sistema atinge saturação em termos do valor crítico de P_{out} (relacionado a um número mínimo de conexões necessárias para que o sistema inicie um comportamento caótico). Em outras palavras, a partir de certo valor de C , torna-se irrelevante aumentar o número de comunidades, pois o grau crítico de fragmentação da rede tende a um valor ideal de $\theta\%$. Este fato é comprovado experimentalmente na Seção a seguir.

4.3 Experimentos e Resultados

Nesta Seção serão descritos as simulações computacionais e os resultados por estas obtidos. Inicialmente o comportamento periódico que surge como resultado da regra de aprendizado é verificado na Subseção 4.3.1. Em seguida, analisa-se o grau de fragmentação necessário para que a rede passe a apresentar um comportamento caótico (descobrimento da *região de bifurcação*) na Subseção 4.3.2. Por fim, na Subseção 4.3.3 a dinâmica caótica é analisada no contexto de comunidades propriamente dito.

4.3.1 Verificação do Comportamento Periódico

O primeiro experimento aqui descrito é o de averiguar o comportamento naturalmente periódico que surge como resultado da aplicação da regra de aprendizado dos ciclos limitantes. Para tanto, utiliza-se a base de padrões ilustrada pela Figura 4.4; 4.4(a) apresenta os padrões utilizados pela regra de aprendizado - nota-se, portanto, um número $M = 4$ de padrões, cada um consistindo de uma imagem binária de 10×10 pixels ($N = 100$). 4.4(b) é o padrão utilizado pela regra de propagação. Considera-se, também, neste caso, uma topologia de conexões que não apresenta fragmentações, isto é, $C = P_{in} = P_{out} = 1$. Desta forma, a rede apresenta a regra de propagação de Hopfield, fazendo com que o padrão 4.4(b) convirja para um dos ciclos.

Figura 4.4: Base de padrões: imagens binárias de 10×10 pixels. (a) mostra os padrões a serem armazenados. O padrão em (b) é utilizado na fase de propagação.

A Figura 4.5 apresenta o caso em que os 4 padrões da Figura 4.4(a) são organizados em apenas um ciclo ($K = 1$ e $L = 4$). Após um período de transiência de 488 iterações (de um

total de 1000 iterações para a evolução do padrão da Figura 4.4(b)), o sistema retorna, em sequência, os 4 padrões armazenados. Nota-se, portanto, que a rede converge para o ciclo único e apresenta um comportamento periódico de período 4. A Figura 4.6 mostra a série temporal da propagação do padrão 4.4(b), reforçando a dinâmica periódica.

Figura 4.5: Evolução da rede para o padrão apresentado na Figura 4.4(b). Os padrões da Figura 4.4(a) são organizados em um ciclo ($K = 1$ e $L = 4$). Nota-se um comportamento periódico de período 4, no qual a rede converge para o ciclo único.

Figura 4.6: Série temporal do estado interno u referente à evolução exibida pela Figura 4.5. Reforça-se a observação de um comportamento periódico de período 4.

Em um segundo caso, os mesmos 4 padrões da Figura 4.4(a) são organizados em 2 ciclos, 2 padrões em cada ciclo ($K = 2$ e $L = 2$). O primeiro ciclo contém os dois primeiros padrões

da Figura 4.4(a) enquanto que o segundo ciclo contém os dois últimos. Nota-se na Figura 4.7 que a rede continua a apresentar um comportamento periódico, porém, agora de período 2, onde o padrão 4.4(b) evolui para ser aprisionado no primeiro ciclo. Esta dinâmica periódica de período 2 é enfatizada pela série temporal da Figura 4.8.

Figura 4.7: Evolução da rede para o padrão apresentado na Figura 4.4(b). Os padrões da Figura 4.4(a) são organizados em 2 ciclo, 2 padrões em cada ciclo ($K = 2$ e $L = 2$). Desta vez, nota-se um comportamento periódico de período 2, no qual a rede converge para o primeiro dos dois ciclos.

Figura 4.8: Série temporal do estado interno u referente à evolução exibida na Figura 4.7. Reforça-se a observação de um comportamento periódico de período 2.

4.3.2 Análise do Surgimento do Comportamento Caótico

A próxima simulação visa analisar a caoticidade do sistema de acordo com o grau de fragmentação das conexões sinápticas. Para tanto, utiliza-se uma base com 30 padrões ($M = 30$) de 20×20 pixels cada ($N = 400$) gerados aleatoriamente porém de forma a manterem-se proximamente ortogonais uns aos outros. Esta alteração da base visa estabelecer uma proporção $\alpha = N/M = 3/40$. Esta proporção é empiricamente aceita como sendo a que mais facilita o aparecimento de pontos espúrios [Mori et al. 1989]. Assim, acredita-se que esta proporção também facilita o surgimento de dinâmicas complexas.

Para a etapa de treinamento, os 30 padrões são divididos em 5 ciclos ($K = 5$), com 6 padrões em cada ciclo ($L = 6$). Assim, o comportamento dinâmico natural da rede será, portanto, periódico de período 6, devido à regra de treinamento. O primeiro padrão gerado também é utilizado na etapa de propagação, isto é, a rede evolui para um padrão não distorcido - isto tem a intenção de promover maior estabilidade do comportamento dinâmico durante a evolução do padrão.

Para analisar a porcentagem de fragmentação necessária para iniciar um comportamento caótico, considera-se uma rede com apenas uma comunidade ($C = P_{out} = 1$), na qual as conexões são fragmentadas pela variação dos valores de P_{in} (P_{in} inicia em 1 e decresce para 0). A Figura 4.9(a) exibe o diagrama de bifurcação em relação aos valores de P_{in} . A Figura 4.9(b) mostra o espectro de Hurst para o mesmo caso. Nota-se que a região de bifurcação do sistema está em torno de $P_{in} = 0,20$. Isto significa que aproximadamente 80% das conexões devem ser descartadas para que a dinâmica periódica natural do sistema ceda lugar a um comportamento caótico.

As Figuras 4.10 e 4.11 investigam o comportamento do sistema ao redor da região de bifurcação. A Figura 4.10(a) mostra a série temporal da evolução da rede para $P_{in} = 0,30$ e a Figura 4.10(b) exibe os respectivos coeficientes de Fourier da série da Figura 4.10(a). Nota-se um comportamento periódico de período 6. Já as Figuras 4.11(a) e 4.11(b) exibem, respectivamente, a série temporal e os coeficientes de Fourier para $P_{in} = 0,10$. Neste caso, tais resultados sugerem a existência de um comportamento caótico.

4.3.3 Análise do Comportamento Dinâmico no Contexto de Comunidades

Os resultados da Subseção anterior indicam uma porcentagem crítica (ou seja, a porcentagem necessária para que a rede altere seu comportamento dinâmico de periódico para caótico) de $p^{(c)} \sim 20\%$ do número de conexões totais da rede. Resta, portanto, a investigação do com-

Figura 4.9: Análise do comportamento dinâmico da rede considerando uma base de 30 padrões divididos em 5 ciclos, 6 padrões por ciclo ($K = 5$ e $L = 6$). Em (a) é exibido o diagrama de bifurcação do estado interno u em relação a P_{in} , para $C = 1$ e $P_{out} = 1$. (b) mostra o espectro dos expoentes de Hurst H para a mesma evolução. Nota-se uma região de bifurcação para P_{in} em torno de 0,20.

portamento da rede no contexto de comunidades, ou seja, para $C > 1$. Para tanto, considera-se o experimento no qual o valor de P_{out} varia em relação a um dado valor de C . A fim de garantir uma estrutura de comunidades válida, faz-se $P_{in} = 1$, ou seja, as comunidades internamente são totalmente conectadas a fim de assegurar um alto valor da Medida de Modularidade Q . A Figura 4.12 mostra a curva entre número de comunidades C e os valores críticos de P_{out} , isto é, valores de P_{out} para os quais a rede encontra-se na região de bifurcação, porém na fronteira do lado caótico. Um valor de P_{out} é considerado estar na fronteira caótica da região de bifurcação quando o expoente de Hurst H da série temporal obtida para um dado valor de C excede um certo limiar, aqui definido como 0,3.

Figura 4.10: Comportamento da rede para $P_{in} = 0,30$, $C = 1$ e $P_{out} = 1$. (a) mostra uma série temporal periódica de período 6 que é ratificada pelos coeficientes de Fourier c em (b). O expoente de Hurst para a série é $H = 0,060113$, valor que também indica um comportamento periódico.

A análise da curva exibida na Figura 4.12 sugere que a rede exibe um comportamento caótico para $C \geq 5$. Para $C = 5$ o P_{out} crítico é 0,03 (cerca de 97% de conexões descartadas) e cresce até alcançar um estado de saturação ao redor de 0,20 quando $C \geq 50$. Este comportamento assintótico é analiticamente comprovado pela Equação 4.10, conforme demonstrado na Subseção 4.2.2. A Figura 4.13 compara a curva exibida na Figura 4.12 com a curva teórica gerada pela Equação 4.10, quando $N = 400$, $p^{(c)} = 0,20$ e, consequentemente, $\theta = 0,202$. Assim, mostra-se que os resultados experimentais são compatíveis com os resultados teóricos.

Conclui-se, portanto, que há indícios relevantes de que a rede, no contexto de comunidades, altera seu comportamento de periódico para caótico quando a rede apresenta uma certa porcentagem crítica na fragmentação das conexões que ligam comunidades distintas. Ainda resta investigar se a rede pode ao mesmo tempo comportar-se caoticamente e apresentar uma estrutura válida de comunidades. Para tanto, investiga-se a qualidade da estrutura de comunidades da rede aplicando-se a medida de modularidade dada pela Equação 2.37. Para simplificar o experimento, supõe-se que a rede será tanto mais modular quanto maior for o número de conexões entre neurônios da mesma comunidade e menor for o número de conexões entre neurônios de comunidades distintas. Desta forma, a Figura 4.14 mostra a relação entre a Medida de Modu-

Figura 4.11: Análise do comportamento da rede para $P_{in} = 0,10$, $C = 1$ e $P_{out} = 1$. Em (a) é exibida a série temporal e em (b) os relacionados coeficientes de Fourier. O expoente de Hurst da série é $H = 0,493295$. Este dados sugerem um comportamento caótico.

Figura 4.12: Variação do valor crítico de P_{out} em relação ao número C de comunidades. Nota-se que não há alteração significativa no valor de P_{out} para $C \geq 50$.

laridade Q e o número de comunidades C para $P_{in} = 1$ e $P_{out} = 0,05$, ainda considerando a base de 30 padrões 20×20 ($N = 400$). A curva mostra uma Medida de Modularidade válida ($Q \geq 0,3$) para $2 \leq C \leq 29$.

Figura 4.13: Comparação entre as curvas dadas pela Figura 4.12 e pela Equação 4.10 quando $N = 400$, $p^{(c)} = 0, 20$, e $\theta = 0, 202$. Observa-se que o resultado experimental é compatível com o resultado teórico.

Figura 4.14: Variação da Medida de Modularidade Q pelo número de C , considerando $P_{out} = 0,05$ e $P_{in} = 1$. (a) mostra a curva completa e (b) magnifica a região na qual $Q \geq 0,3$. É indicado que a rede possui uma estrutura de comunidade válida para $2 \leq C \leq 29$.

Comparando os resultados apresentados nas Figuras 4.12 e 4.14, conclui-se que a rede pode comportar-se caoticamente e com alto grau de modularidade para $5 \leq C \leq 29$, considerando $P_{in} = 1$ e $P_{out} = 0,05$. Este resultado é reforçado pela Figura 4.15 que exhibe o comportamento

dinâmico da rede para $C = 20$ sugerindo uma dinâmica caótica.

Figura 4.15: Comportamento dinâmico da rede para $C = 20$, $P_{out} = 0,05$ e $P_{in} = 1$. (a) mostra a série temporal e (b) os coeficientes de Fourier. Para esta série, a Modularidade é $Q = 0,387818$ e o expoente de Hurst é $H = 0,587774$. Este resultado indica que a rede pode comportar-se caoticamente e, ao mesmo tempo, exibir uma estrutura de comunidades válida.

Considerações Finais

Há tempos que muitos projetos de pesquisa tentam compreender e assimilar a complexa funcionalidade das redes neurais biológicas. Modelos artificiais surgiram com propostas de reproduzir aspectos simples da funcionalidade do cérebro, baseados em um padrão de aprendizado conexionista. Demonstrou-se que o córtex cerebral exibe comportamentos dinâmicos complexos. Modelos artificiais, como o de Hopfield, operam com comportamento dinâmico, mas com um dinamismo estacionário sob uma estrutura topológica simples. Variações do modelo de Hopfield surgiram, ou modificando o tipo de neurônio, ou a regra de aprendizado, ou a topologia das conexões sinápticas. Neste sentido, as Redes Neurais Caóticas operam com dinâmicas complexas, porém ainda com uma estrutura topológica simples, quando comparadas à das redes neurais biológicas. Além do dinamismo complexo, o cérebro contém grupos especializados de células, que podem ser estudados como as estruturas de comunidades das Redes Complexas.

Desta forma o projeto de pesquisa aqui descrito analisou o comportamento dinâmico de um novo modelo de Rede Neural Caótica que opera com dinâmicas periódicas e caóticas sob uma arquitetura de comunidades, sendo assim mais similar à estrutura de uma rede neural biológica. A Seção 5.1 disserta sobre as conclusões da pesquisa. As principais contribuições científicas são discutidas na Seção 5.2. O capítulo é encerrado na Seção 5.3, sugerindo alguns tópicos a serem investigados em uma próxima etapa desta pesquisa.

5.1 Conclusões

Os resultados exibidos na Subseção 4.3.1 mostram que, em uma topologia sem fragmentações, a rede neural proposta opera com dinâmica periódica como consequência natural da regra de aprendizado utilizada. Para altos graus de fragmentação, há fortes indícios de que a rede altera

seu comportamento para uma dinâmica caótica em duas situações:

- utilizando-se um baixo valor de probabilidade de ocorrerem ligações entre neurônios da mesma comunidade em uma topologia com apenas uma comunidade, conforme exibido na Subseção 4.3.2. O diagrama de bifurcação e o espectro dos expoentes de Hurst indicam a região de bifurcação, reforçada pela análise dos coeficientes de Fourier em pontos ao redor da bifurcação.
- utilizando-se um baixo valor de probabilidade de ocorrerem ligações entre neurônios de comunidades distintas em uma topologia com considerável número de comunidades. Os gráficos da Subseção 4.3.3 mostram que o uso de comunidades provoca aparecimento da região de bifurcação, porém o aumento do número de comunidades, a partir de um valor crítico, não necessariamente provoca o aumento da porcentagem de fragmentação necessária para surgir a região de bifurcação. Assim, torna-se irrelevante utilizar um número muito elevado de comunidades. Mostrou-se, também, que para um intervalo específico do número de comunidades, a rede comporta-se caoticamente e exibe um alto valor da medida de modularidade, indicando que pode existir caos em uma estrutura de comunidades válida.

5.2 Contribuições

Este projeto de mestrado desenvolveu uma pesquisa inédita, propondo um modelo de Rede Neural Caótica que possui topologia mais similar à topologia observada em uma rede neural biológica. Se modelos tradicionais de RNA de tipos *Feedforwards* ou recorrentes, formam redes totalmente conectadas, ou se Redes Neurais Celulares apresentam topologia em retículos, este trabalho fornece uma arquitetura alternativa cujos neurônios são agrupados em comunidades, similarmente ao que ocorre no córtex cerebral. Além disso, também como no cérebro, o modelo pode operar com dinâmicas periódicas e caóticas em tarefas de reconhecimento de padrões.

Diferentemente da maioria dos modelos de Redes Neurais Caóticas, onde modelos de neurônios caóticos são empregados para facilitar a ocorrência do caos, a rede aqui proposta utilizou um modelo simples de neurônio (MCP) e o comportamento caótico foi derivado, tão somente, da fragmentação da estrutura sináptica da rede - algo que, comparativamente, é muito mais difícil de ser alcançado. Este fato, por si só, já indicaria uma ótima contribuição para a grande área de pesquisa. Todavia, além disso, os resultados foram expressivos e, portanto, fornecem garantia de aplicabilidade da metodologia aqui desenvolvida.

Outra importante contribuição científica desta pesquisa foi a elaboração, publicação e apresentação de um artigo científico que contempla os principais resultados aqui descritos [Sousa e Zhao 2013]. O artigo foi aceito no *IEEE International Joint Conference on Neural Networks* (IJCNN), de 2013, realizado na cidade de Dallas, EUA. Trata-se de um dos principais eventos da área de RNA, categorizado como A2 pelo Qualis CAPES e com Índice-H 46 em 2012. Uma apresentação oral foi realizada na referida conferência, durante a qual o artigo foi um dos premiados com o *Travel Grant Award*.

5.3 Trabalhos Futuros

Uma vez estabelecidas as garantias de que a rede opera com dinâmicas complexas por meio de uma arquitetura de comunidades, conforme indicam os resultados obtidos neste projeto de pesquisa, a próxima etapa óbvia é o desenvolvimento uma aplicação para reconhecimento de padrões que utiliza a metodologia aqui apresentadas. Para tanto, faz-se necessária a adequação da rede à uma base de dados específica e uma medida de similaridade adequada para contemplar o objetivo da aplicação. A medida de similaridade deve ser utilizada para provocar a alteração da dinâmica da rede a partir de uma fase caótica para uma fase não-caótica (periódica). Além disso, deve-se imaginar uma maneira efetiva para se dividir os neurônios em comunidades, visto que, conforme mencionado na Seção 5.1 acima, a rede não precisa, necessariamente, conter comunidades para que apresente um comportamento caótico.

Em relação ao processo de análise da dinâmica da rede, algumas complementações aos resultados aqui apresentados seriam relevantes. A curva dos expoentes de Lyapunov, por exemplo, poderia reforçar a indicação do comportamento caótico quando a rede apresenta um alto grau de fragmentação. Para isto, deve-se substituir a função de ativação por outra contínua (como uma função sigmoide configurada para se aproximar do comportamento da função sinal). Outra tarefa é constatar o processo de itinerância caótica exibindo a evolução de padrões durante a fase de propagação. Para isto, deve-se construir uma base de padrões visivelmente diferenciáveis e em quantidade suficiente para provocar a dinâmica caótica. Porém, deve-se ter em mente que a regra de aprendizado utilizada, para que funcione adequadamente, exige padrões proximamente ortogonais uns aos outros e pouca variabilidade na proporção $\alpha = M/N$ (normalmente em torno de 3/40).

Referências Bibliográficas

- [Adachi e Aihara 1997] Adachi, M. e Aihara, K. (1997). Associative dynamics in a chaotic neural network. *Neural Networks*, 10:83–89.
- [Aihara e Matsumoto 1986] Aihara, K. e Matsumoto, G. (1986). Chaotic oscillations and bifurcations in squid giant axons. *Chaos*, pages 257–269.
- [Albert e Barabási 2002] Albert, R. e Barabási, A. L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics*, 74:47–97.
- [Babloyantz et al. 1985] Babloyantz, A., Salazar, J. M., e Nicolis, C. (1985). Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A*, 111(3):152–156.
- [Barabási e Albert 1999] Barabási, A. L. e Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.
- [Boashash 2003] Boashash, B., editor (2003). *Time-frequency signal analysis and processing: A comprehensive reference*. Oxford: Elsevier Science.
- [Boccaletti et al. 2006] Boccaletti, S., Latora, V., Moreno, Y., Chaves, M., e Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308.
- [Bryson e Ho 1969] Bryson, A. E. e Ho, Y. C. (1969). *Applied optimal control*, volume 1. Blaisdell, New York.
- [Chua e Yang 1988] Chua, L. O. e Yang, L. (1988). Cellular neural networks: Theory. *IEEE Transactions on Systems and Circuits*, 35(10):1257–1272.
- [Clauset 2005] Clauset, A. (2005). Finding local community structure in networks. *Physical Review E*, 72(2):026132.

- [Clauset et al. 2004] Clauset, A., Newman, M. E. J., e Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, 79(6):066111.
- [Cooley e Tukey 1965] Cooley, J. W. e Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier Series. *Mathematics Computation*, 19:297–301.
- [Devaney 1989] Devaney, R. L. (1989). *An introduction to chaotic dynamical systems*. Addison-Wesley, Redwood City, California.
- [Duch e Arenas 2005] Duch, J. e Arenas, A. (2005). Community detection complex networks using extremal optimization. *Physical Review E*, 72(2):027104.
- [Erdős e Rényi 1961] Erdős, P. e Rényi, A. (1961). On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1-2):261–267.
- [Goles-Chacc et al. 1985] Goles-Chacc, E., Fogelman-Soulie, F., e Pellegrin, D. (1985). Decreasing energy functions as a tool for studying threshold networks. *Discrete Applied Mathematics*, 12(3):261–277.
- [Grassi 2001] Grassi, G. (2001). On discrete-time cellular neural networks for associative memories. *IEEE Transactions on Circuits and Systems*, 48(1):107–111.
- [Hartman 1964] Hartman, P. (1964). *Ordinary differential equations*. Wiley, New York.
- [Haykin 1999] Haykin, S. (1999). *Neural networks: A comprehensive foundation*. Prentice Hall, New York, 2nd edition.
- [Hebb 1949] Hebb, D. O. (1949). *The organization of behavior*. Wiley, New York.
- [Hodgkin e Huxley 1952] Hodgkin, A. e Huxley, A. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544.
- [Hopfield 1982] Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the USA*, 79(8):2554–2558.
- [Huang e Huang 2011] Huang, W. Z. e Huang, Y. (2011). Chaos, bifurcations and robustness of a class of Hopfield neural networks. *International Journal of Bifurcation and Chaos*, 21(3):885–985.

- [Hurst 1951] Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799.
- [Jain e Dubes 1988] Jain, A. K. e Dubes, R. C. (1988). *Algorithms for clustering data*. Prentice Hall.
- [Kaneko e Tsuda 2003] Kaneko, K. e Tsuda, I. (2003). Chaotic itinerancy. *Chaos*, 13(3):926–936.
- [Kaplan e Glass 1995] Kaplan, D. e Glass, L. (1995). *Understanding nonlinear dynamics*. Springer-Verlag, New York.
- [Kaslik e Balint 2009] Kaslik, E. e Balint, S. (2009). Complex and chaotic dynamics in a discrete-time-delayed Hopfield neural network with ring architecture. *Neural Networks*, 22:1411–1418.
- [Leonov e Kuznetsov 2007] Leonov, G. A. e Kuznetsov, N. V. (2007). Time-varying linearization and the Perron effects. *International Journal of Bifurcation and Chaos*, 17(4):1079–1107.
- [Liu e Lu 1997] Liu, D. e Lu, Z. (1997). A new synthesis approach for feedback neural networks based on the perceptron training algorithm. *IEEE Transactions on Neural Networks*, 8(6):1468–1482.
- [Liu e Michel 1994] Liu, D. e Michel, A. N. (1994). Sparsely interconnected neural networks for associative memories with applications to cellular neural networks. *IEEE Transactions on Circuits and Systems*, 41:295–307.
- [Lysetskiy e Zurad 2004] Lysetskiy, M. e Zurad, J. M. (2004). Bifurcating neuron: computation and learning. *Neural Networks*, 17:225–232.
- [Maass 1997] Maass, W. (1997). Networks of spiking neurons: The third generation of neural network models. *Neural Networks*, 10:1659–1671.
- [Maass e Bishop 2001] Maass, W. e Bishop, C. M. (2001). *Pulsed neural networks*. The MIT Press.
- [Marcus e Westervelt 1989] Marcus, C. M. e Westervelt, R. M. (1989). Dynamics of iterated-map neural networks. *Physical Review A*, 40(1):501–504.
- [Marichal et al. 2012] Marichal, R. L., González, E. J., e Marichal, G. N. (2012). Hopf bifurcation stability in Hopfield neural networks. *Neural Networks*, 36:51–58.

- [Massen e Doye 2005] Massen, C. P. e Doye, J. P. K. (2005). Identifying communities within energy landscapes. *Physical Review E*, 71:046101.
- [Mathias e Rech 2012] Mathias, A. C. e Rech, P. C. (2012). Hopfield neural network: The hyperbolic tangent and the piecewise-linear activation functions. *Neural Networks*, 34:42–45.
- [Matykiewicz 2004] Matykiewicz, P. (2004). Chaotic itinerancy for patterns separation. *Lecture Notes in Computer Science*, 3070:235–240.
- [McClelland e Rumelhart 1986] McClelland, J. L. e Rumelhart, D. E. (1986). *Parallel distributed processing - explorations in the microstructure of cognition: Foundations*, volume 1. The MIT Press.
- [McCulloch e Pitts 1943] McCulloch, W. S. e Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 52(1-2):99–115.
- [Mori et al. 1989] Mori, Y., Davis, P., e Nara, S. (1989). Pattern retrieval in an asymmetric neural network with embedded limit cycles. *Journal of Physics A*, 22:L525–L532.
- [Muff et al. 2005] Muff, S., Rao, F., e Caflisch, A. (2005). Local modularity measure for network clusterizations. *Physical Review E*, 72(5):056107.
- [Nara et al. 1993] Nara, S., Davis, P., e Totsuji, H. (1993). Memory search using complex dynamics in a recurrent neural network model. *Neural Networks*, 6:963–973.
- [Newman 2003] Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45:167–256.
- [Newman 2004a] Newman, M. E. J. (2004a). Detecting community structure in networks. *The European Physical Journal B*, 38(2):321–330.
- [Newman 2004b] Newman, M. E. J. (2004b). Fast algorithm for detecting community structure in networks. *Physical Review E*, 69(6):066133.
- [Newman 2006a] Newman, M. E. J. (2006a). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, 74(3):036104.
- [Newman 2006b] Newman, M. E. J. (2006b). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582.
- [Newman 2008] Newman, M. E. J. (2008). Community structure in directed networks. *Physical Review Letters*, 100:118703.

- [Newman 2011] Newman, M. E. J. (2011). Communities, modules and large-scale structure in networks. *Nature Physics*, 8:25–31.
- [Newman e Girvan 2004] Newman, M. E. J. e Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2):026113.
- [Ott 1993] Ott, E. (1993). *Chaos in dynamical systems*. Cambridge University Press, New York.
- [Park et al. 2001] Park, J., Kim, H. Y., e Lee, S. W. (2001). A synthesis procedure for associative memories based on space-varying cellular neural networks. *Neural Networks*, 14:107–113.
- [Radicchi et al. 2004] Radicchi, C., Castellano, F., Cecconi, F., Loreto, V., e Parisi, D. (2004). Defining and identifying communities in networks. *Proceedings of the National Academy of Science USA*, 101(9):2658–2663.
- [Reichardt e Bornholdt 2006] Reichardt, J. e Bornholdt, S. (2006). Statistical mechanics of community detection. *Physical Review E*, 74:016110.
- [Rojas 1996] Rojas, R. (1996). *Neural networks: A systematic introduction*. Springer-Verlag, Berlin.
- [Rosenblatt 1958] Rosenblatt, F. (1958). The Perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65:386–408.
- [Rumelhart et al. 1986] Rumelhart, D. E., Hinton, G. E., e Williams, R. J. (1986). Learning representations of back-propagation errors. *Nature (London)*, 323:533–536.
- [Russell e Norvig 2010] Russell, S. e Norvig, P. (2010). *Artificial intelligence: A modern approach*. Prentice Hall, New Jersey, 3rd edition.
- [Skarda e Freeman 1987] Skarda, C. A. e Freeman, W. J. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, 10(2):161–195.
- [Sousa e Zhao 2013] Sousa, F. B. e Zhao, L. (2013). Investigation of complex dynamic in a recurrent neural network with network community structure and asymmetric weight matrix. *Proceedings of International Joint Conference on Neural Networks*, 1:999–1005.
- [Stauffer et al. 2003] Stauffer, D., Aharony, A., Costa, L. F., e Adler, J. (2003). Efficient hopfield pattern recognition on a scale-free neural network. *The European Physical Journal B*, 32:395–399.

- [Strogatz 1989] Strogatz, S. H. (1989). *Nonlinear dynamics and chaos: With applications to physics, chemistry, and engineering*. Addison-Wesley, Reading, Massachusetts.
- [Szolgay et al. 1997] Szolgay, P., Szatmári, I., e László, K. (1997). A fast fixed point learning method to implement associative memory on cnns. *IEEE Transactions on Circuits and Systems*, 44(4):362–366.
- [Tokuda et al. 2010] Tokuda, I. T., Han, C. E., Aihara, K., Kawato, M., e Schweighofer, N. (2010). The role of chaotic resonance in cerebellar learning. *Neural Networks*, 23:836–842.
- [Watts e Strogatz 1998] Watts, D. J. e Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393:440–442.
- [Werblin et al. 1995] Werblin, F., Roska, T., e Chua, L. (1995). The analogic cellular neural network as a bionic eye. *Int. J. Circuit Theory and Applications*, 23:541–569.
- [Zachary 1977] Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33:452–473.
- [Zhao et al. 2006] Zhao, L., Cáceres, J. C. G., Damiance-Jr., A. P. G., e Szu, H. (2006). Chaotic dynamics for multi-valey content addressable memory. *Neurocomputing*, 69:1628–1636.
- [Zheng et al. 2010] Zheng, P. S., Tang, W. S., e Zhang, J. X. (2010). A new chaotic Hopfield network with piecewise linear activation function. *Chinese Physics B*, 19(3):030514.