

Capacidade crítica de armazenamento de uma rede neural com conectividade espaço-temporal estruturada

Karim Aquere Filho, R.M.C. de Almeida

Laboratório de Modelos Teóricos e Computacionais, Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051 - 91501-970, Porto Alegre, RS, Brasil
e-mail: karim@if.ufrgs.br
e-mail: rita@if.ufrgs.br

J. A. Quillfeldt

Laboratório de Psicobiologia e Neurocomputação, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, 91501-970, Porto Alegre, RS, Brasil
quillfe@vortex.ufrgs.br

In this paper we present a new neural network model with structured temporal and finite connectivity capable of storing and recovering temporal sequences or cycles of information using a Hebbian-like rule. We perform a series of numerical simulations to determine the critical capability the system using the Hamming distance for cycles and observe that this model presents a better critical capability for long-cycles than short-cycles.

I. INTRODUÇÃO

No cérebro humano, um estímulo sensorial isolado, ou uma curta seqüência de estímulos, pode dar origem à lembrança de uma série de memórias conectadas temporalmente. São impressionantes os efeitos tais como a lembrança de uma música, que pode ser provocada por um estímulo bastante simples, e não necessariamente relacionado à música em si, o que indica algum tipo de correlação entre estados subseqüentes da rede. Algo similar ocorre na lembrança de uma seqüência de números ou do próprio alfabeto.

Nesse sentido, o armazenamento, o aprendizado e a recuperação de seqüências temporais de memórias formam um aspecto fundamental da nossa capacidade mental, já que nos permitem compreender o curso dos fatos e ações cotidianas que, de outra forma, consistiriam de simples figuras sem significado nem movimento.

Além disso, seqüências periódicas de impulsos nervosos são de fundamental importância para o controle de funções motoras do corpo, tais como as batidas do coração e o ritmo respiratório, que ocorrem com grande regularidade (cerca de três bilhões de vezes durante a vida média de uma pessoa). Sob esta ótica, tentar explicar como as Redes Neurais podem sustentar essa atividade periódica durante um longo período de tempo e estudar seu comportamento sob condições específicas pode fornecer informações importantes para a compreensão desses processos.

Existe uma série de modelos de redes neurais que possuem uma estrutura espacial e/ou temporal capaz de recuperar seqüências temporais. Dentre estas, destacamos o

modelo do Trion [1] e o trabalho de Ji e colaboradores [2], que introduzem redes unidimensionais com condições de contorno periódicas, onde neurônios formais estão conectados espaço-temporalmente com um número finito de neurônios. Utilizando técnicas numéricas que permitem monitorar e analisar a dinâmica destas redes verifica-se que estes tipos de estrutura evoluem para padrões cíclicos de ativação.

Neste trabalho apresentaremos um modelo de redes neurais com conectividade espaço-temporal finita e estruturada, mostrando resultados obtidos através de simulações computacionais que indicam que uma rede com este tipo de estrutura apresenta uma capacidade crítica de armazenamento que privilegia ciclos de informações em detrimento de padrões isolados. Nas seções seguintes apresentamos o novo modelo, descrevemos as simulações numéricas para obter a capacidade máxima de armazenamento α_c , apresentamos os resultados obtidos e finalmente discutimos os resultados e concluímos.

II. O MODELO

Consideramos uma rede unidimensional com condições de contorno periódicas composta por unidades S_i , $i = 1, \dots, N$, onde $S_i = \pm 1$. O estado da rede é definido por $\vec{S} = (S_1, \dots, S_N)$ e evolui no tempo de forma síncrona, governado pela equação dinâmica

$$S_i(t+1) = f[h_i(t)], \quad (1)$$

onde f é a função transmissão, que deve ser escolhida adequadamente, e $h_i(t)$ é o campo local atuando sobre o sítio S_i no tempo t , definido como

$$h_i(t) = \sum_{t'=1}^{t_{\max}} \sum_{\langle j \rangle} W_{ij}^{t'} S_j(t-t'), \quad (2)$$

onde $\langle j \rangle \in [i - \frac{k}{2}, i + \frac{k}{2}]$ e define a vizinhança de um dado sítio S_i , k é a conectividade do sistema (consideramos k assumindo somente valores pares, assegurando a validade da equação 2), $W_{ij}^{t'}$ é a matriz sináptica que descreve a contribuição do estado do sistema no tempo $t - t'$

ao campo local no tempo t e t_{\max} é o número máximo de passos anteriores ao tempo t que contribuem para o campo local.

Neste trabalho consideramos os casos onde o número de conexões se dá em apenas dois passos anteriores no tempo de forma estruturada (figura II). Notamos que tomando $k = N - 1$ temos o caso de conectividade total, como no modelo de Hopfield [3] e tomando $t_{\max} = 1$ recaímos no modelo de Little [4], que também apresenta ciclos durante sua evolução.

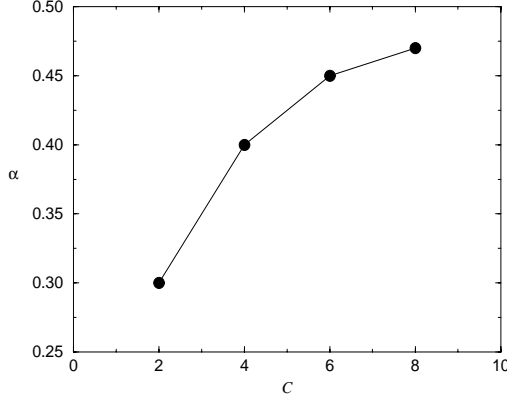


FIG. 1. Diagrama espacial-temporal das conexões de um sítio S_i da rede com seus k vizinhos de forma estruturada, recebendo influência de dois passos de tempo anteriores. As linhas cheias representam as conexões da rede no passo de tempo $(t - 1)$ associados a W_{ij}^1 e as linhas tracejadas as conexões com o passo de tempo $(t - 2)$ associados a W_{ij}^2 .

As matrizes sinápticas são construídas utilizando a informação contida em um conjunto de padrões, vistos aqui como ciclos compostos por um conjunto de C configurações, escolhidos para serem armazenados via uma regra tipo-Hebb. Consideramos P padrões de C passos temporais cada um, dados por $\xi^{\mu\tau}$, onde $\mu = 1, \dots, P$ é o número de padrões ou ciclos armazenados e $\tau = 1, \dots, C$ é o número de passos temporais de cada padrão. Os $\xi^{\mu\tau}$ são formados por CN valores sorteados aleatoriamente com igual probabilidade. Ou seja, sorteamos o valor para o estado de cada neurônio i , para cada passo de tempo τ , e cada padrão μ com igual probabilidade de ser ± 1 .

A regra de aprendizado é definida, dentro do espírito da regra de Hebb, como

$$W_{ij}^{t'} = \frac{1}{k} \sum_{\mu=1}^P \sum_{\tau=1}^C \xi_i^{\mu,\tau} \xi_j^{\mu,(\tau-t')}, \quad (3)$$

válida para i e j dentro da região de interação, onde $(\tau - t') \in [1, C]$, isto é, se $(\tau - t') < 1$ então $(\tau - t') \rightarrow (\tau - t') + C$. Vale salientar que esta forma de construir as sinapses é especialmente indicada para a recuperação de um número inteiro qualquer de ciclos, diferente dos modelos tradicionais de redes capaz de recuperar seqüências [5–11]. que armazenam apenas uma

seqüência de cada vez.

No nosso caso, construímos as sinapses através da regra 3 de forma que cada configuração da rede evolua para um determinado atrator cíclico previamente escolhido.

Visando explorar a possibilidade da realização de simulações numéricas sobre diversas configurações da rede evoluindo a partir de diferentes condições iniciais definimos a chamada Distância de Hamming para Ciclos.

III. A DISTÂNCIA DE HAMMING PARA CICLOS

Para podermos realizar uma análise mais detalhada do sistema é necessária a definição de uma expressão análoga à distância de Hamming, passível de realizar uma comparação entre duas configurações da rede. Esta expressão deve ser capaz de medir a diferença entre um dado padrão composto de C configurações e um conjunto de C configurações sucessivas da rede e recebe o nome de distância de Hamming para ciclos, \mathcal{H}^μ .

Considerando um ciclo de C configurações $\xi^{\mu,\tau}$, onde $\tau = 1, \dots, C$, verificamos que existem C possíveis fases para tal ciclo, isto é, para um intervalo de tempo de C passos temporais a rede pode apresentar cronologicamente os estados $\xi^{\mu,1}, \xi^{\mu,2}, \dots, \xi^{\mu,C}$, ou $\xi^{\mu,2}, \xi^{\mu,3}, \dots, \xi^{\mu,C}, \xi^{\mu,1}$, ou $\xi^{\mu,3}, \xi^{\mu,4}, \dots, \xi^{\mu,C}, \xi^{\mu,1}, \xi^{\mu,2}$, até finalmente $\xi^{\mu,C}, \xi^{\mu,1}, \dots, \xi^{\mu,C-1}$. Assim para definir a distância entre um padrão de tamanho C e uma sucessão de C estados da rede $\vec{S}^{(t-c)}, \vec{S}^{(t-c-1)}, \dots, \vec{S}^{(t-1)}, \vec{S}^t$, precisamos levar em conta as C fases possíveis deste padrão. Então definimos primeiramente a distância de Hamming entre uma sucessão de estados e um padrão em uma dada fase : $\mathcal{H}^{\mu,\dagger}$.

$$\mathcal{H}^{\mu,\dagger} = \frac{\infty}{CN} \sum_{\tau=t}^{C-\infty} \sum_{\lambda=\infty}^N (\infty - \delta(\mathcal{S}_i(\sqcup - \tau), \xi_i^{\mu,(C-\tau+\lambda)})) \quad (4)$$

onde

$$\delta(S_i(t - \tau), \xi_i^{\mu,C-\tau+a}) = \begin{cases} 1 & \text{se } S_i(t - \tau) = \xi_i^{\mu,C-\tau+a} \\ 0 & \text{se } S_i(t - \tau) \neq \xi_i^{\mu,C-\tau+a}, \end{cases} \quad (5)$$

com $0 \leq a \leq C - 1$ e lembrando que $\xi^{\mu,\lambda} = \xi^{\mu,\lambda+C}$ para λ inteiro. E finalmente tomamos a menor distância calculada para todas as fases, isto é:

$$\mathcal{H}^\mu = \mathbb{I} \setminus \downarrow [\mathcal{H}^{\mu,\dagger}] \quad (6)$$

Vale ressaltar que a equação 4 realiza uma comparação sítio a sítio entre dois conjuntos de CN neurônios, apontando apenas se dois sítios estão ou não no mesmo estado, não existindo distância entre dois estados maior que 1, mesmo no caso de três estados. Os resultados numéricos são obtidos essencialmente pelo monitoramento de \mathcal{H}^μ para diferentes protocolos tomando o valor médio da

distância de Hamming para ciclos dada por $\langle\{\mathcal{H}^\mu\}\rangle$, onde $\langle\dots\rangle$ indica uma média da distância de Hamming para ciclos sobre diversas configurações da rede e $\{\dots\}$ indica a média sobre várias evoluções partindo de diferentes condições iniciais.

Além disso observamos que, sob condições controladas, a rede é estável e determinamos a forma matemática da capacidade de armazenamento, definida como $\alpha = \frac{PC}{k}$, onde PC define a quantidade de informação armazenada na rede, ou seja o número de sítios de cada padrão, P, multiplicado pelo tamanho do ciclo C.

Já a capacidade crítica de armazenamento de informação por conexão sináptica, é definida como

$$\alpha_c = \frac{(P.C)_{max}}{k} \quad (7)$$

onde $(P.C)_{max}$ indica a quantidade máxima de informação para que ocorra uma recuperação satisfatória de um determinado padrão.

IV. RESULTADOS

Simulamos numericamente a média sobre diversas configurações da distância de Hamming para ciclos em função de α .

As figuras 2,3,4 e 5 mostram $\langle\{\mathcal{H}^\mu\}\rangle$ em função de α para $N = 1000$, para diferentes valores de C e valores de k crescentes. Observa-se que à medida que k aumenta, a função $\langle\{\mathcal{H}^\mu\}\rangle(\alpha)$ tende a uma transição de segunda ordem entre as fases de reconhecimento e não-reconhecimento.

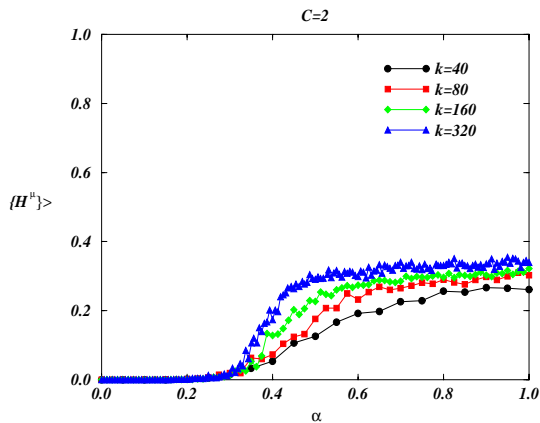


FIG. 2. Simulações com $N = 1000$, mostrando distância de Hamming para ciclos $\langle\{\mathcal{H}^\mu\}\rangle$, em função de α , para $C = 2$, onde $\alpha_c \simeq 0.30$.

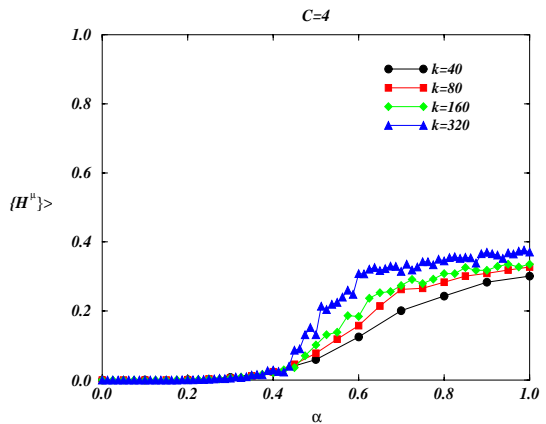


FIG. 3. Simulações com $N = 1000$, mostrando distância de Hamming para ciclos $\langle\{\mathcal{H}^\mu\}\rangle$, em função de α para $C = 4$, onde $\alpha_c \simeq 0.40$.

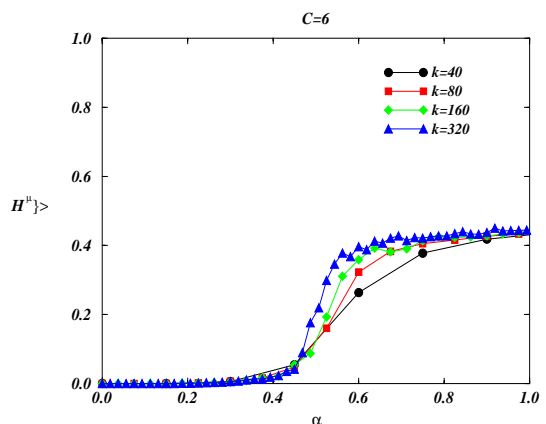


FIG. 4. Simulações com $N = 1000$, mostrando distância de Hamming para ciclos $\langle\{\mathcal{H}^\mu\}\rangle$, em função de α para $C = 6$, onde $\alpha_c \simeq 0.45$.

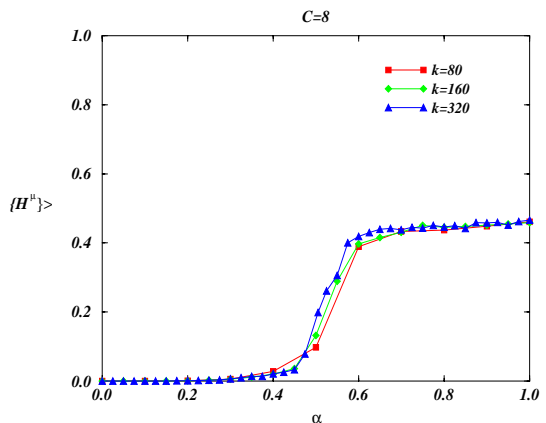


FIG. 5. Simulações com $N = 1000$, mostrando distância de Hamming para ciclos $\langle\{\mathcal{H}^\mu\}\rangle$, em função de α para $C = 8$, onde $\alpha_c \simeq 0.47$.

A figura IV mostra valores de α_c em função de C . As simulações indicam um crescimento de α_c em função de C , mostrando ser vantajoso, do ponto de vista da quantidade de informação armazenada, utilizar seqüências de informações temporalmente relacionadas em detrimento de padrões isolados ($C=1$, como no caso do modelo de Hopfield). Ainda na figura ?? observa-se que fazendo uma extrapolação para os valores de C , obtemos um limite superior para $\alpha_c \approx 0.20$ no caso de três estados e $\alpha_c \approx 0.50$, no caso binário, o que é esperado, pois, caso contrário poderíamos ter uma capacidade de armazenamento de informação infinita.

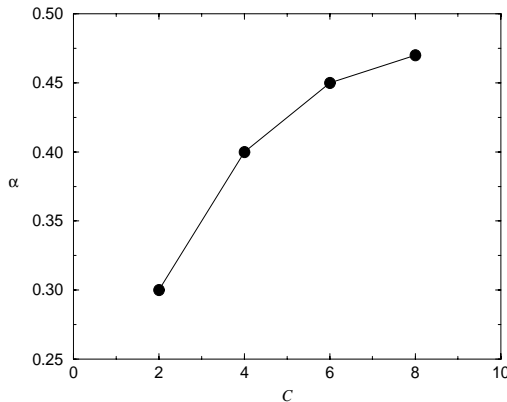


FIG. 6. α_c em função de C para o caso binário (a) e para o caso de três estados (b), mostrando que a medida que aumenta C , aumenta α_c , indicando ser vantajoso armazenar informação em forma de ciclos do que em forma de padrões isolados.

V. CONCLUSÕES

O nosso trabalho procura introduzir alguns aspectos estruturais e funcionais importantes do ponto de vista do funcionamento do sistema nervoso, dando ênfase à Regra de Aprendizado de Hebb, uma das hipóteses mais difundidas atualmente no que diz respeito ao armazenamento de memória nos mamíferos, e ao Princípio Organizacional de Mountcastle [12], que procura estabelecer a estrutura para o processamento de informação no encéfalo.

É necessário salientar que apesar do nosso modelo também estar afastado de sua contrapartida biológica, algumas características como conectividade finita e estruturada são condizentes com a realidade do sistema nervoso, pois é evidente que os neurônios mantêm sinapses com um número limitado de outros neurônios de forma relativamente estruturada, o que significa que em alguns casos os neurônios estão conectados com os seus vizinhos mais próximos, sendo isto é verdadeiro para várias regiões, em especial no córtex dos mamíferos [12,13].

O aspecto mais importante deste trabalho é o fato da estrutura do modelo permitir o aprendizado e a recuperação de ciclos de informações de forma mais eficiente

do que em padrões isolados. O que implica que uma determinada quantidade de informação pode ser melhor recuperada caso seja armazenada em seqüência temporal. Observamos que, numa determinada região, quanto maior é o ciclo armazenado, maior a capacidade crítica de armazenamento do sistema. O que de certa forma está de acordo com o observado para a memória humana, que tem maior facilidade de recuperar informações associadas em seqüência do que várias informações descorrelacionadas, como por exemplo ao evocar uma melodia aprendida ou um lista aleatória de palavras ou números. A própria conectividade temporal tem sua interpretação biológica, já que a atividade dos neurônios dependem do estado de seus vizinhos em instantes de tempo anteriores, para poder disparar ou não, nos instantes seguintes. Além disso, vale lembrar, a atualização dos estados no nosso modelo é feita em paralelo, o que também se verifica no ser humano, onde vários agrupamentos celulares trabalham em concerto para executar uma determinada tarefa.

-
- [1] SHAW, G. L, SILVERMAN, D. J., PEARSON, J. C. **Proc. Natl. Acad. Sci. USA** **82**, 2364, 1985.
 - [2] JI, D., HU, B., CHEN, T. **Physica A** **229**, 147, 1995.
 - [3] HOPFIELD, J. J. **Proc. Natl. Acad. Sci. USA.**, **79**, 2554, 1982
 - [4] LITTLE, W. A. **Mathematical Biosciences**, **19**, 101, 1974.
 - [5] KLEINFELD, D., SOMPOLINSKY, H. **Methods in neuronal modeling: from synapses to networks**, 195 Cambridge: MIT Press, 1989.
 - [6] AMIT, D. **Proceedings of the National Academy of Sciences USA**, **85**, 2141, 1988.
 - [7] HORN, D., USHER, M. **Phys. Rev. A** **40**, 1036, 1989.
 - [8] KURTEN, K.E. **J. Physique** **51**, 1585, 1990.
 - [9] CANNING, A., GARDNER, E. **J. Phys. A** **21**, 3275, 1988.
 - [10] KURTEN, K.E. **Phys. Lett. A** **129**, 157, 1988.
 - [11] DERRIDA, B. **J. Phys. A** **22**, 2069, 1989.
 - [12] MOUNTCASTLE, V. B., EDELMAN G. M. **The Mindful Brain: Cortical Organization and the Group-Selective Theory of Higher Brain Function** Cambridge: MIT Press, 1978.
 - [13] PURVES, D., RIDDLE, D. R., LAMANTIA, A. S. **Trends in Neuroscience**, **15**, 362, 1992.