

研究工作

不同图样间联想的神经网络模型

刘 军 梁明理

(武汉大学 物理系)

提 要

本文讨论了 Hopfield 模型和其恢复特性。提出了不同图样间联想模型,推导了主态和辅态为初态时该模型的演化结果。对于 30 个随机图样、400 个神经元组成的神经网络,给出了计算机模拟计算结果。并且讨论了不同情况下 Hopfield 模型和图样间联想模型的计算结果。

一、导 言

1982 年, J. J. Hopfield 提出了神经网络模型^[1]。这种由大量的两态神经元对称联接而组成的网络, 显示出很多集体运算的性质^[1-3]。尤其是它具有按内容寻址贮存记忆的特性, 即它能从部分信息通过一种算法, 而联想出全部信息。现在, 我们在 Hopfield 模型的基础上, 提出一记忆间彼此联系的模型, 即为了记忆某一态(我们称之为主态)而引入另一态(我们称之为辅态)和主态相联系, 产生一新的突触强度矩阵 $\{T_{ij}^{new}\}$, 系统的动力学保证了从主态附近或从辅态附近都能演化出主态来, 而且对于 Hopfield 模型中的一些问题, 也能给予一定的解决。

二、Hopfield 模型和其恢复特性

考虑 N 个神经元组成的神经网络, 每一神经元 i 有两种状态, 用变量 S_i 描述为:

$$S_i = \begin{cases} +1 & (\text{firing}) \\ -1 & (\text{not firing}) \end{cases} \quad (1)$$

该网络的状态可用一 N 维矢量来描述

$$\mathbf{S} = (S_1, S_2, \dots, S_N) \quad (2)$$

显然, 该系统的状态可以代表某一物理事件, 或某一图样 (pattern), 或某一记忆。对 M 个随机图样 $\mathbf{S}^{(1)}, \mathbf{S}^{(2)}, \dots, \mathbf{S}^{(M)}$ 。其中

$$\mathbf{S}^{(k)} = (S_1^{(k)}, S_2^{(k)}, \dots, S_N^{(k)}) \quad (3)$$

的每一分量 $S_i^{(k)}$ ($k = 1, 2, \dots, M; i = 1, 2, \dots, N$), 我们以相同的几率随机地取为 $+1$ 或 -1 。显然, 图样 $\mathbf{S}^{(m)}$ 和 $\mathbf{S}^{(n)}$ 之间有赝正交^[4]

$$\begin{aligned} 1/N \sum_{i=1}^N S_i^{(m)} S_i^{(n)} &= \delta_{mn} \\ &\pm (1 - \delta_{mn}) O(1/\sqrt{N}) \end{aligned} \quad (4)$$

网络的 N 个神经元之间由突触联系。神经元 j 到 i 的突触强度记为 T_{ij} , 为了保证这 M 个图样对应的状态在状态空间中稳定, 且能吸引状态空间的流, 从而满足作为按内容寻址的记忆的条件^[1], 设置 T_{ij} 如下:

$$\begin{cases} T_{ij} = \sum_{k=1}^M S_i^{(k)} S_j^{(k)} & (i \neq j) \\ T_{ii} = 0 \end{cases} \quad (5)$$

这样神经元 i 的输入为:

$$V_i = \sum_{j=1}^N T_{ij} S_j \quad (6)$$

神经元将照下列规则改变其状态

$$S_i = \text{sgn}(V_i - U_i) \quad (7)$$

式中 U_i 为神经元的阈值, 令 $U_i = 0$ 。为简单起见, 认为所有的神经元状态的改变都同时进行^[4], 这样神经元状态的演化将以分离的时间步骤 $t = 0, 1, 2, \dots$ 进行, 而且满足

$$S_i(t+1) = \text{sgn} \left(\sum_{j=1}^N T_{ij} S_j(t) \right) \quad (8)$$

如果初态为 S , 即 $S = S(0)$ 。那么在 (8) 式所表述的算法作用下, 演化到 S' 。如果 S' 态就是某一记忆态 $S^{(k)}$ 的话, 就认为 $S^{(k)}$ 被联想出来。为了定量地描述 S' 态和 $S^{(k)}$ 态的相近程度, 定义态 S' 和态 $S^{(k)}$ 的重叠为:

$$q(S', S^{(k)}) = \frac{1}{N} \sum_{i=1}^N S'_i \cdot S_i^{(k)} \quad (9)$$

$$d = (1 - q)/2$$

式中 d 叫 Hamming 距离。显然 d 代表了 $S', S^{(k)}$ 之间相异部分的大小。如果 $d = 0$, 那么 $S^{(k)}$ 态被完全恢复。

设 t 时刻, 态 $S(t)$ 和 $S^{(k)}$ 的重叠为 $q(t)$ 。现在来推导 $q(t+1)$ 和 $q(t)$ 的关系, 即记忆的恢复特性。

显然, $d(S(t+1), S^{(k)})$ 将由 $S_i^{(k)} \cdot S_i(t+1) < 0$ 那部分给出。那么, 考虑到 (8) 式和 (5) 式:

$$\begin{aligned} S_i^{(k)} \cdot S_i(t+1) &= \text{sgn} \left[S_i^{(k)} S_i^{(k)} \right. \\ &\quad \left. \sum_j \left(S_j^{(k)} S_j(t) + \sum_{m \neq k} S_j^{(k)} S_j^{(m)} S_j^{(m)} S_j(t) \right) \right] \\ &= \text{sgn} [Nq(S^{(k)}, S(t)) + R_i] \end{aligned} \quad (10)$$

忽略各态之间的相互关系, R_i 是一随机变量, 而且

$$\langle R_i \rangle = 0 \quad \Delta^2 = \langle R_i^2 \rangle = N \cdot M \quad (11)$$

对于 N 很大时, R_i 是一高斯分布

$$P(R) = (1/\sqrt{2\pi\Delta}) \exp(-R^2/2\Delta^2) \quad (12)$$

这样 $d(S(t+1), S^{(k)})$ 将由 $Nq(S^{(k)}, S(t)) + R_i < 0$ 的那部分给出。从而得到

$$q(t+1) = F \left(\frac{N \cdot q(t)}{\Delta} \right) \quad (13)$$

其中

$$F(x) = \sqrt{2/\pi} \int_0^x \exp(-y^2/2) dy$$

$$= \text{erf}(x/\sqrt{2})$$

当系统的状态稳定时有:

$$q(t+1) = q(t) \quad (14)$$

令稳定的重叠为 q^* , 它由下式给出

$$q^* = F(\sqrt{N/M} q^*) \quad (15)$$

最后恢复的错误为

$$d^* = (1 - q^*)/2 \quad (16)$$

这样得到了恢复过程中重叠的逆推关系 (13) 式和最后稳定点的 q^* 和 d^* (15), (16) 式。

三、图样间联想模型

对于 M 个随机图样, 由 Hopfield 模型, 有 (5) 式设定的突触强度:

$$\begin{cases} T_{ij} = \sum_{k=1}^M S_i^{(k)} S_j^{(k)} & (i \neq j) \\ T_{ii} = 0 \end{cases}$$

如果为了记忆 $S^{(k_1)}$ 态而引入一 $S^{(k_2)}$ 态和它发生联系。我们称 $S^{(k_1)}$ 态叫主态, $S^{(k_2)}$ 态叫辅态。那么新的突触强度定义为:

$$T_{ij}^{new} = T_{ij}^{old} + S_i^{(k_1)} \cdot S_j^{(k_2)} - S_i^{(k_2)} S_j^{(k_1)} \quad (17)$$

式中右边的第一项为原来的 Hopfield 模型的 T_{ij} , 第二项表明了 $S^{(k_1)}$ 态和 $S^{(k_2)}$ 态的联系, 第三项说明了为了从辅态 $S^{(k_2)}$ 联想出 $S^{(k_1)}$ 态, 而对辅态原始意义的抑制。

这样当初态设置为 $S^{(k_1)}$ 时, 即

$$S(0) = S^{(k_1)} \quad (18)$$

由 (8) 式和 (17) 式得:

$$\begin{aligned} S_i(1) &= \text{sgn} \left(\sum_{j=1}^N T_{ij}^{new} S_j(0) \right) \\ &= \text{sgn} \left[\left(\sum_{j=1}^N \sum_{k=1}^M S_i^{(k)} S_j^{(k)} S_j^{(k_1)} \right. \right. \\ &\quad \left. \left. + \sum_{j=1}^N (S_i^{(k_1)} S_j^{(k_2)} S_j^{(k_1)} - S_i^{(k_2)} S_j^{(k_2)} S_j^{(k_1)}) \right) \right] \\ &= \text{sgn} \left[S_i^{(k_1)} \sum_{j=1}^N S_j^{(k_1)} S_j^{(k_1)} \right. \\ &\quad \left. + S_i^{(k_1)} \sum_{j=1}^N S_j^{(k_1)} S_j^{(k_2)} \right] \end{aligned}$$

$$- S_i^{(k_2)} \sum_{j=1}^N S_j^{(k_2)} S_j^{(k_1)} \Big] \\ S_i(1) = \text{sgn}(NS_i^{(k_1)}) \quad (19)$$

显然有 $S_i(1) = S_i^{(k_1)}$ 即通过一步就能演化到 $S^{(k_1)}$ 态。如果 $S(0)$ 是带有噪声的 $S^{(k_1)}$ 态时,它能通过数步而演化到 $S^{(k_1)}$ 态。当然对一定的 $\alpha = M/N$, 噪声的大小必须在一限度之内。

同样地在 $S^{(k_2)}$ 态为初态时即 $S(0) = S^{(k_2)}$, 有

$$\begin{aligned} S_i(1) &= \text{sgn} \left(\sum_{j=1}^N T_{ij}^{\text{new}} S_j(0) \right) \\ &= \text{sgn} \left(\sum_{j=1}^N T_{ij}^{\text{new}} S_j^{(k_2)} \right) \\ &= \text{sgn} \left[\left(\sum_{j=1}^N \sum_{k=1}^M S_i^{(k)} S_j^{(k)} S_j^{(k_2)} \right) \right. \\ &\quad \left. + \sum_{j=1}^N (S_i^{(k_1)} S_j^{(k_2)} S_j^{(k_2)}) \right. \\ &\quad \left. - \sum_{j=1}^N S_i^{(k_2)} S_j^{(k_2)} S_j^{(k_2)} \right) \Big] \\ &= \text{sgn}(NS_i^{(k_2)} + NS_i^{(k_1)} - NS_i^{(k_2)}) \\ &= \text{sgn}(NS_i^{(k_1)}) = S_i^{(k_1)} \end{aligned}$$

所以同样地演化到 $S^{(k_1)}$ 态。即初态为辅态时,通过联想而将主态恢复出来。

四、数字模拟结果和讨论

我们用计算机对 $N = 400$, $M = 30$ 的系统进行了模拟计算。30 个图样随机地选取,其中包括字符“A”的图样。我们用 400 个神经元的态来表示 20×20 图形矩阵,当 $S_i = 1$ 时给出“*”, $S_i = -1$ 时为空白。

1. 当 $k_1 = k_2$ 时。由 (17) 式

$$\begin{aligned} T_{ij}^{\text{new}} &= T_{ij}^{\text{old}} + S_i^{(k_1)} S_j^{(k_2)} - S_i^{(k_2)} S_j^{(k_2)} \\ &= T_{ij}^{\text{old}} \end{aligned}$$

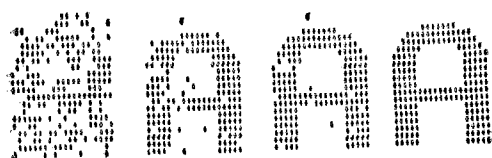


图1 Hopfield 模型字符“A”的恢复

即退回到 Hopfield 模型,满足第二节中讨论的恢复特性。对 $\alpha = 30/400$, $d(0) = 32\%$, 经过 4 步演化出记忆态。如图 1。

2. 当 $k_1 \neq k_2$ 时。不论初态为主态还是辅态,都将演化出主态来(图 2)。在 Hopfield 模

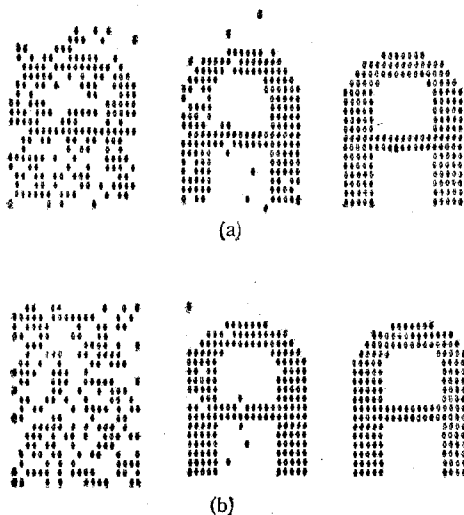


图2 主态辅态联想模型字符“A”的恢复

主态为字符“A”,辅态为任一随机图样

(a) 初态为带有 30% 噪声的主态,最后演化出主态字符“A”; (b) 初态为带有 30% 噪声的辅态,最后联想出主态字符“A”

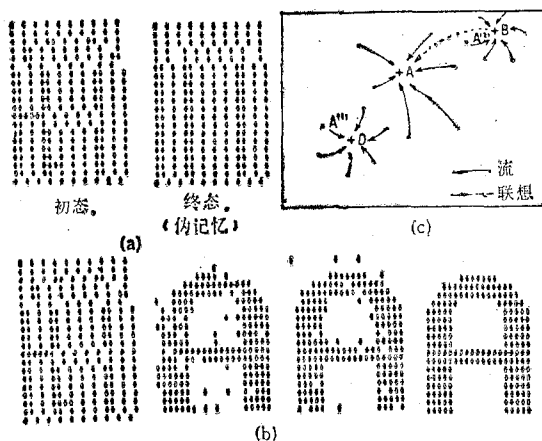


图3 噪声较大时字符“A”的恢复

(a) 带有 $d(0) = 50\%$ 噪声的字符“A”为初态,在 Hopfield 模型中最后输出一“伪记忆”; (b) 带有 $d(0) = 50\%$ 噪声的字符“A”为初态,它和辅态的重叠为 $q = 0.615$ 通过联想演化出主态字符“A”来 (c) 主态 A 和辅态 B。 $A^{(1)}$ 为带有较大噪声的 A 态,它流向一伪记忆 D 态。但如果带有噪声的主态落在辅态 B 附近如 $A^{(2)}$, 就能联想出主态来。图中“+”为吸引子,“·”为初态

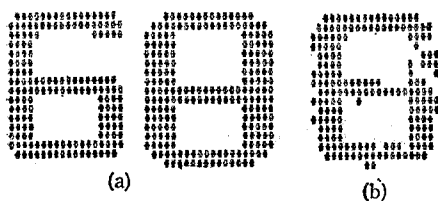
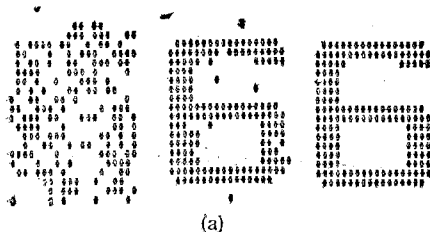
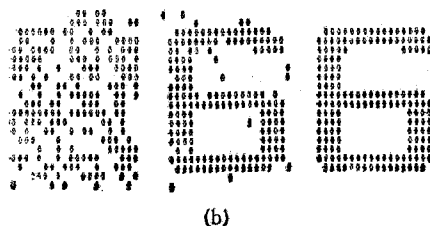


图4 Hopfield 模型中两相近记忆产生混乱

(a) 两相近记忆态字符“6”和字符“8”，Hamming 距离为 $d = 0.23$ (b) 演化的结果给出一混乱态



(a)



(b)

图5 两相近记忆态间联想主态字符“6”的恢复

(a) 带有 37% 噪声的主态为初态，经 2 步演化出纯主态字符“6” (b) 带有 37% 噪声的辅态为初态，经 2 步演化出纯主态字符“6”

型中，对一定的 $\alpha = M/N$ ，当 $d(0)$ 太大而越过“相边界”时^[2]，那么演化的结果将输出一“伪记忆”(图 3 a)。在我们的模型中一般情况也相似。但如果某一带有较大噪声 ($d(0) \geq 50\%$) 的主态正好落在和它相联系的辅态附近，那么通过联想能演化出主态(图 3 b)。这一结果可通过图 3 c 加以理解。

3. 在 Hopfield 模型中，当存在两个记忆态非常接近时，就会产生混乱^[1]，这时这两个态都不能恢复出来(图 4)。但如果让这两态相联系，就能演化出主态来(图 5)。因为图样间联想算法，实际上对辅态进行了抑制，从而解除了由于两态接近时产生的混乱，使主态得以恢复。

参 考 文 献

- [1] Hopfield, J. J.: *Proc. Natl. Acad. Sci. USA*, 1982, 79, 2554.
- [2] Tank, D. W., Hopfield, J. J.: *IEEE Trans.*, 1986, CAS-33, 533.
- [3] Hopfield, J. J., Tank, D. W.: *Biol. Cybern.*, 1985, 52, 141.
- [4] Kinzel, W.: *Z. Phys. B-Condensed Matter*, 1985, 60, 205.
- [5] Shinomoto, S.: *Biol. Cybern.*, 1987, 57, 197.

[本文于 1988 年 7 月 20 日收到]

(上接第 111 页)

- 1986, 157, 199.
- [14] Bialkowska-Hobrzanska, H.: *J. Clin. Microbiol.*, 1987, 25, 338.
- [15] Pinkel, D. et al.: *Proc. Natl. Acad. Sci. USA*, 1986, 83, 2934.
- [16] Rayburn, A. L. and Gill, B. S.: *J. Hered.*, 1985, 76, 78.
- [17] Meselidis, L. et al.: *J. Cell Biol.*, 1982, 95, 619.

- [18] Smith, G. H. et al.: *Anal. Biochem.*, 1986, 156, 17.
- [19] Khan, A. M. and Wright, P. J.: 1987, *J. Virol. Methods*, 15, 121.
- [20] Kulski, J. K. and Noval, M.: 1985, *Arch. Virol.*, 83, 3.

[本文于 1987 年 12 月 15 日收到]