

具有瞬态动态的神经网络

摘要： 我们研究了以不同的半稳定活动模式的时间序列为特征的动力系统，因为它们在皮层神经活动模式中能够被观察到。我们提出并讨论了一种通用机制，允许吸引子网络与特定的邻接瞬态网络之间连接，这是严格耗散的。具有瞬态的动态系统在自动调节其工作点时仍具有功能——避免长时间停滞或漂移到快速波动的状态。我们在连续时间神经网络模型中显示，用于在线学习的单个本地更新规则同时允许 (i) 通过无监督的 Hebbian 型学习进行信息存储，(ii) 自适应调整工作点，以及 (iii) 抑制突触生长失控。我们给出了仿真结果；讨论了时间反转对称性和链接对称性的自发发放。

1、引言

动力系统通常根据其长期行为进行分类，例如，可能表现为混沌或规则[1]。特别引人注意的是吸引子，环和极限环，因为它们决定了在各自吸引区内开始的所有轨道的命运。

吸引子状态在递归神经网络理论中起着核心作用：能够根据部分初始信息进行泛化和重建完整的记忆[2]。每当网络被认为是一个包含自治信息处理系统的功能子单元时，即一个自治认知系统[3]，循环神经网络中的吸引子状态便会面对基本的功能二分结构。一旦轨迹在吸引子处闭合，信息处理就会基本停止。对于神经网络的技术应用而言，“手动”重启系统是一个可行的选择，但不适用于自主运行的认知系统。

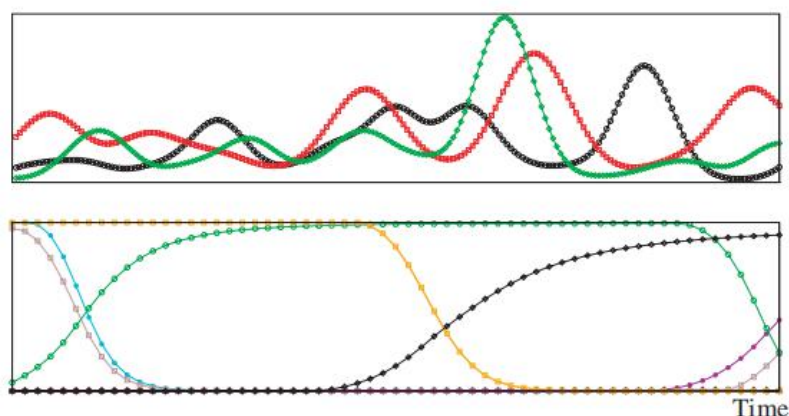


图 1 波动（顶部）和瞬态动力学（底部）的图示。

解决这个难题的一个方法是只考虑没有吸引子状态的动力学系统，即具有一种持续不断的“波动动力学”，如图 1 所示，在严格意义上的动力学系统理论意义上可能是混乱的。然而，问题是决策过程。如果没有定义良好的状态（这些状态持续某些最短的时间），则系统将没有确定的信息承载状态，可以基于该状态

来生成输出信号。有趣的是，在这种情况下，已经观察到了皮层神经活动中准平稳模式的迹象[4-6]。这些准平稳状态可以使用多元时间序列分析进行分析，表明大脑活动的自组织模式[7]。脑电图记录的研究已经根据大脑状态进行了解释，显示出经过吸引子序列的非周期性进化状态，这些吸引子在访问时支持记忆的过程[8]。这些发现表明，如图 1 所示，“瞬态动力学”对于皮层发射模式可能很重要。

从动力学系统理论的角度来看，有可能将瞬态视为轨道接近吸引子时的明确定义的周期。对于瞬态吸引子或吸引子残迹，我们在这里表示相空间中的一个点，当调整进入动力学系统演化方程的某些参数时，该点可以连续变成稳定的吸引子。动力学在吸引子的破坏附近变慢，并且在动力学变量集合中出现了定义明确的瞬态。瞬态动力学的概念在概念上与混沌迭代有关（请参见[9]和其中的参考文献），该术语用于表征动力学系统，在该动力学系统中，混沌高维轨道在一定时期内间歇性地靠近低维吸引子。对于动态迭代的出现，由于动态相互作用或噪声引起的不稳定性是必需的。

在论证了瞬态动力学对于广泛的现实世界动力学系统而言很重要之后，问题是如何以可控制的方式并以适用于各种启动系统的方式产生这种动力学行为。我们对神经网络感兴趣，而该神经网络会根据有意义的时间序列状态（任意接近预定义的吸引子）生成瞬态动力学。

我们将在这里采用的方法是从原始的吸引子神经网络开始，然后通过辅助局部变量耦合来将稳定吸引子集转换为暂态吸引子，我们将其称为“储层”，受长时标控制。我们注意到，在离散时间和相位耦合振荡器[10]，以及“Winnerless competition”的背景下，针对以“latching transitions” [11, 12]为语言处理目标的网络研究了相关问题[13-15]。能够生成后续状态时间序列的神经网络的其他示例是具有时间依赖性非对称突触强度[16]或动态阈值[17]的神经网络。我们还注意到，已经研究了自发波动动力学的发生[18]，特别是与基础网络几何形状有关的[19]。

神经网络的固有任务是学习并适应传入的刺激。对于自适应神经网络，这意味着对其动力学特性的连续修改。因此，如果不采取预防措施，学习过程可能会使网络脱离其预期的工作状态，即瞬态动力学状态。在这里，我们将表明可以制定局部学习规则，通过不断优化其自身的工作点来使系统保持在适当的动态状态。具体来说，让我们用 \bar{t} 表示准稳态瞬态的平均持续时间，用 Δt 表示从一种准稳态过渡到另一种稳态所需的典型时间。然后将动态工作点定义为比率 $\Delta t / \bar{t}$ 。

这些时间尺度 \bar{t} 和 Δt 是皮质神经元网络的结果，它是由各个神经元的属性（基本上与时间无关）以及突触强度（它们是受 Hebbian 型学习影响的缓慢动态

变量)产生的[20]。然后得出结论,神经间突触强度的修改具有双重功能:一方面它们参与记忆存储任务[20],另一方面它们需要将工作点保持在最佳状态。在这里,我们证明了这种双重功能可以在广义神经网络模型内实现。我们显示,当将 Hebbian 学习规则重新定义为优化过程时,可以获得工作点优化,从而导致导致单个神经元的突触集之间的竞争。最终的学习规则与优化内存存储容量的规则密切相关[21]。

2、模型

2.1 集合编码

具有稀疏编码的神经网络,即平均触发率低,具有非常大的存储器存储容量[22]。稀疏的编码在极端情况下会导致“Winner take all”的配置,单个单元只能对一个内存进行编码。但是,在此限制下,存储容量会再次降低,并且与网络大小成线性比例关系,就像原始的 Hopfield 模型[23]一样。在这里,我们选择“透明编码”的中间情况。就图论而言,一个集合是一个完全互连的子图,如图2所示的7顶点网络。集合编码对应于“Winner take all”的设置。获胜集合的所有成员彼此相互激发,同时将所有异常神经元的活动抑制为零。

我们注意到,集合的数量可能非常多。为了说明,让我们考虑一个具有 N 个顶点和链接概率 p 的随机 Erdős-Renyi 图。包含 Z 个顶点的集合总数由下式统计得出:

$$\binom{N}{Z} p^{Z(Z-1)/2} (1-p^Z)^{N-Z}, \quad (1)$$

其中 $p^{Z(Z-1)/2}$ 是图中的 Z 个顶点由 $Z(Z-1)/2$ 个边相互连接的概率,而最后一项是 $N-Z$ 中的每个顶点没有同时连接到集合的所有 Z 点的概率。

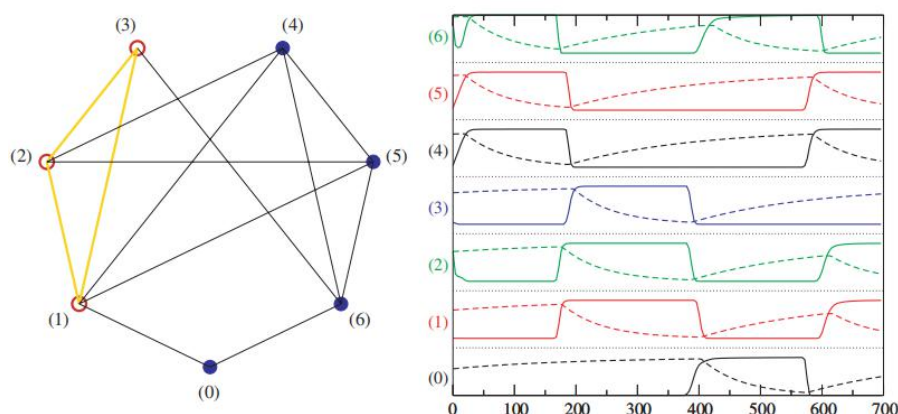


图2 一个小型的7顶点网络的几何和模拟结果。左面板: $w_{ij} > 0$ 的链接,包含六个集合 (0, 1), (0, 6), (3, 6), (1,2,3) (突出显示), (4, 5, 6) 和 (1, 2, 4, 5)。右图: 活动 $x_i(t)$ (实线) 和相应的储层 $\phi_i(t)$ (虚线) 的瞬态动力学作为时间的函数,从 (4, 5、

6), 到 (1、2、3), 再到 (0, 6), 最后到 (1、2、4、5)。有关参数值, 请参见第 2.6 节。

正如我们将在下面进一步讨论的那样, 具有集合编码的网络特别适合瞬态动态, 并且在生物学上是可行的。广泛的感觉预处理发生在大脑的各个皮质区域 [20], 导致单个神经元或小细胞集合体对特征和物体的表示。在此框架中, 此处考虑的有效神经网络的一个顶点(即一个神经中心)对应于这样的小单元格集合, 并且对应于一个内存的稳定表示形式, 通过将预处理算法从感觉输入流中提取的有限特征集绑定在一起来实现。

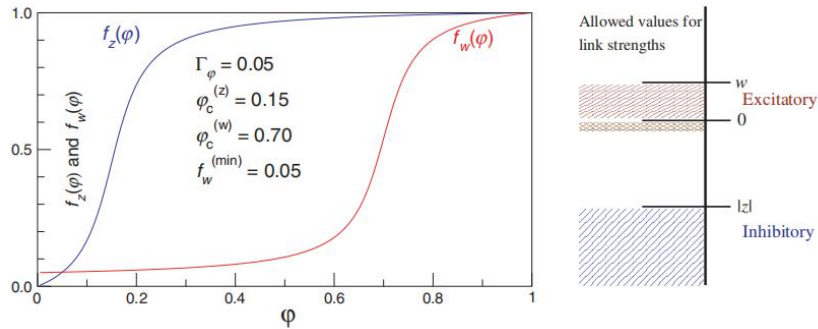


图 3 左图: 储层函数 $f_{z/w}(\phi)$ 的图示, 呈 S 形, 参见方程式 (3), 参见脚注 1, 具有相应的转折点 $\phi_c^{(f/z)}$, 宽度 Γ_ϕ , 和最小值 $f_z^{(\min)} = 0$ 。右图: 抑制链接 $z_{ij} < -|z|$ 和活动兴奋链接 $0 < w_{ij} < w$ 的突触强度分布, 导致集合编码。请注意, 由于优化过程 (11), w 不是严格的上限。零以下的阴影区域与非活动 w_{ij} 有关, 请参见公式 (5) 和 (11)。

2.2 连续时间动态

为了研究神经网络中瞬态动力学的可能机制, 我们考虑 i, \dots, N 个人工神经元, 其编码率 $x_i(t)$ 和连续时间 t 。让我们对最后一点进行简短评论。人工神经网络领域的大多数研究都涉及离散时间 $t=0, 1, 2, \dots$ [20] 的情况。但是, 正如引言部分所述, 我们对展现自主生成的动态行为的网络感兴趣, 因为它们通常发生在完整的自主认知系统的环境中。因此, 我们对具有与认知系统的其他组件的交互兼容的更新规则的网络感兴趣。离散时间更新不适用于这种情况, 因为最终的动态特性 (i) 取决于同步更新与异步更新的选择, 并且 (ii) 当有效的递归循环由于与自主认知的其他组件耦合系统而受到强烈影响时。因此, 我们在这里考虑并研究具有连续时间的模型。

2.3 神经网络模型

我们用 $x_i(t)$ 表示编码活动水平的状态变量, 并假定是连续变量 $x_i(t) \in [0, 1]$ 。此外, 我们为每个顶点引入一个变量 $\phi_i(t) \in [0, 1]$, 称为“储层”, 它用作疲劳记忆, 有助于自生瞬态时间序列。我们考虑以下一组微分方程:

$$\dot{x}_i = (1 - x_i)\Theta(r_i)r_i + x_i\Theta(-r_i)r_i, \quad (2)$$

$$r_i = \sum_{j=1}^N [f_w(\varphi_i)\Theta(w_{ij})w_{i,j} + z_{i,j}f_z(\varphi_j)]x_j, \quad (3)$$

$$\dot{\varphi}_i = \Gamma_\varphi^+(1 - \varphi_i)(1 - x_i/x_c)\Theta(x_c - x_i) - \Gamma_\varphi^-\varphi_i\Theta(x_i - x_c), \quad (4)$$

$$z_{ij} = -|z|\Theta(-w_{ij}). \quad (5)$$

现在我们讨论经过适当修改的 Lotka - Volterra 方程 (2)–(5) 的一些性质:

1、**规范化**。方程 (2) 和 (4) 中的 $x_i(t)$, $\varphi_i(t) \in [0,1]$, 方程 (2)–(4) 中的 $x_i, (1-x_i), \varphi_i, (1-\varphi_i)$ 归一化, 对应于各自的生长和耗竭过程。 $\Theta(r)$ 是 Heaviside-step 函数: $\Theta(r < 0) = 0$ 和 $\Theta(r > 0) = 1$ 。

2、**突触强度**。突触强度分为兴奋性贡献 w_{ij} 和抑制性贡献 z_{ij} , 其中 w_{ij} 是主要变量: 抑制作用 z_{ij} 仅在链接不是兴奋性时才存在 (5)。在整篇文章中, 我们都使用 $z \equiv -1$, 即 $|z|=1$, 然后定义了时间开发的逆参考单位。

3、**Winners-take-all 网络**。等式 (2) 和 (3) 描述了在没有通过 $f_{z/w}(\varphi)$ 耦合到储层的情况下, 具有集合编码的 Winners-take-all 神经网络。系统朝着由吸引点 $w_{p_i, p_i} > 0 (i, j = 1, \dots, Z)$ 连接的 Z 个顶点 (p_1, \dots, p_Z) 组成的下一个吸引子放松。

4、**储层功能**。储层函数 $f_{z/w}(\varphi) \in [0,1]$ 控制活动水平 x_i 和储层水平 φ_i 之间的相互作用。可以将它们选择为 S 形的阶梯函数, 具有合适的宽度 Γ_φ 和拐点 $\varphi_c^{(w/z)}$, 请参见图 3。

5、**储层动力学**。获胜集合的储层内容缓慢耗尽, 请参见方程式 (4) 和图 2, 并且仅在给定顶点的活动水平 x_i 降至 x_c 以下时才恢复, 这定义了 $x_i > x_c$ 时处于活动状态的顶点。在储层生长过程中出现的因数 $(1 - x_i/x_c)$, 请参见 (4), 用于稳定两个后续记忆状态之间的过渡。当给定 i 的活动水平 x_i 降到 x_c 以下, 无法立即重新激活; 由于 (4) 中的 $(1 - x_i/x_c)$, 对于 $x_i \approx x_c$, 储层无法再次充满。

6、**时间尺度的分离**。当 Γ_φ^+ 远小于活动兴奋限制的强度, 即典型的 $w_{ij} > 0$ 时, 会导致时间尺度的分离, 从而导致瞬态动态。获胜集合的储层一旦耗尽, 它会失败, 通过 $f_w(\varphi)$ 抑制了其抑制其他部位和相互内部激励的能力。

7、**缺少固定解**。对于方程 (2) 和 (4), $\dot{x} = 0 = \dot{\varphi}$ ($i=1, \dots, N$) 时没有平稳解, 只要 $\Gamma_\varphi^+ > 0$ 不消失并且对于任何非平凡的耦合函数 $f_{z/w}(\varphi) \in [0,1]$ 。

当将活动与储层分离时, 通过将 $f_{z/w}(\varphi)$ 设置为 1, 对于属于/不属于获胜集合的顶点, 可以得到 $x_i = 1/0$ 和 $\varphi_i = 0/1$ 的稳定吸引子, 比较图 4。

在图 2 中, 说明了由等式 (2) - (5) 得出的瞬态动力学。图 2 中显示的是在没有外部感官信号的情况下自主动力学的的数据。我们将在下面进一步讨论外部刺激的作用。在图 2 中, 我们仅显示了一个非常小的网络的数据, 该网络包含七

个顶点，可以很容易地用图形表示。我们还对包含数千个顶点的超大型网络进行了广泛的仿真，并发现了稳定的瞬态动态。

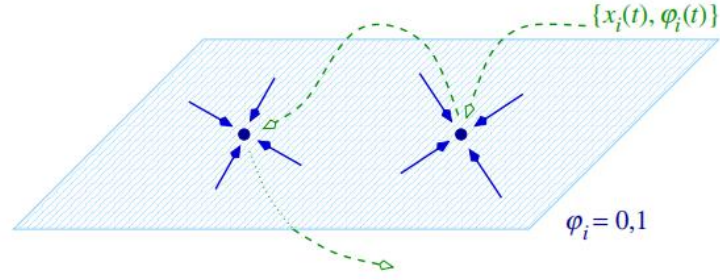


图4 原始网络的吸引子，即通过设置 $f_{z/w}(\varphi) \equiv 1$ 关闭与储层的耦合时，对应于 $x_i = 1/0$ 和 $\varphi_i = 0/1$ （适用于获胜集合的个体/其他成员）。与储层 φ_i 的有限耦合导致了 $\{x_i(t), \varphi_i(t)\}$ 轨道，这些轨道在短时间内被吸引子吸引，并在较长时间内被排斥。这是由于时间尺度的分离，因为储层 $\varphi_i(t)$ 的时间演化发生在比主要动态变量 $x_i(t)$ 的时间尺度上慢得多的时间尺度上。

2.4 储层的作用

这里讨论的动力系统首先代表了自上而下的认知系统方法，并且不打算与皮层结构一一对应。但是，该设置受到生物学类比的启发，我们可以确定方程（2）描述的人工神经网络的位置 i ，而不是单个神经元，而是神经集合或神经中心。因此，储层变量 $\varphi_i(t)$ 可以解释为在连续活动的神经组件中发生的有效疲劳过程。已经提出[24]，用于根据不同且可识别的对象绑定异类感官信息的神经编码在本质上可能是暂时的。在经过实验研究的这种时间编码假设内[25]，当定义相同的对象时，神经程序集同相触发，即同步，而对不同的对象进行编码时，则异步。通常，对象与记忆之间存在密切的关系。因此，一种引人入胜的可能性是用时域编码理论的同步激发神经元来识别在本方法中研究的瞬态网络的存储器。获胜的特征是高水位，然后对应于时间编码范式和储层耗竭时间 $1/\Gamma_\varphi^-$ 将对应于对象结合神经元的退相干时间。

2.5 耗散动力学

观察到的数值和动态鲁棒性的原因可以追溯到其松弛性质。对于较短的时间尺度，我们可以考虑储层变量 φ_i 近似恒定，并且系统松弛到下一个集合吸引子中。一旦接近瞬态吸引子， x_i 就基本上是恒定的，即接近于 $1/0$ ，并且储层逐渐耗尽。动态性对噪声具有鲁棒性，因为波动仅影响两个松弛过程的细节，而不影响其整体行为。确切地说，我们注意到相空间相对于储层变量收缩，即

$$\sum_i \frac{\partial \dot{\varphi}_i}{\partial \varphi_i} = - \sum_i [\Gamma_\varphi^+ (1 - x_i/x_c) \Theta(x_c - x_i) + \Gamma_\varphi^- \Theta(x_i - x_c)] \leq 0, \quad \forall x_i \in [0, 1],$$

这里我们使用了 (4)。我们注意到对链接矩阵的对角贡献消失, $z_{ii} = 0 = w_{ii}$, $\partial r_i / x_i = 0$ 。因此, 相空间也就活动而言收缩,

$$\sum_i \frac{\partial \dot{x}_i}{\partial x_i} = \sum_i [\Theta(-r_i) - \Theta(r_i)] r_i \leq 0,$$

这里我们使用 (2)。因此, 在没有学习和外部干扰的情况下, 该系统是严格耗散的, 从而导致观察到的数值鲁棒行为。

2.6 严格的瞬态动力学

图 2 所示的自生瞬态动力学在 $x_i(t)$ 中表现出特征性的平稳期, 因为已将较小的值用于储层的枯竭和增长率, $\Gamma_\phi^- = 0.005$ 和 $\Gamma_\phi^+ = 0.005$ 。图 2 所示的仿真是针对所有非零激励互连使用 $w_{ij} = 0.12$ 进行的。

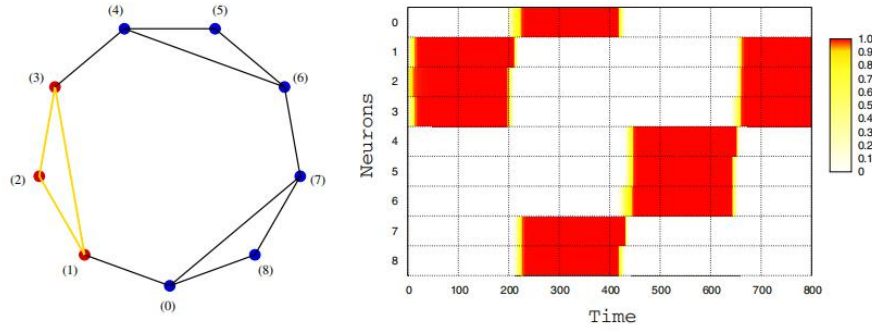


图 5 具有对称激励链接 $w_{ij} = w_{ji}$ 的循环 9 顶点网络的几何和模拟结果。左: $w_{ij} > 0$ 的链接, 包含六个集合 ((0,1), (1,2,3) (突出显示), (3,4), (4,5,6), (6,7) 和 (7,8,0)。右: 周期瞬态动力学 (1,2,3) (7,8,0) (4,5,6) 的活动 $x_i(t)$ 与时间的关系, 参数值请参见 2.6。“旋转”的两个方向 (顺时针/逆时针) 都是动态可能且稳定的, 实际方向取决于动态初始条件。

如果存在一组控制参数, 我们可以将一个动态系统定义为“严格的瞬态动力学”, 从而可以将瞬态转换为稳定的吸引子。式 (2) - (5) 满足了这一要求, 对于 $\Gamma_\phi^- \rightarrow 0$, 图 2 中观察到的稳态平稳期的平均持续时间 t 发散了。

或者, 通过为 Γ_ϕ^- 和 Γ_ϕ^+ 选择合适的值, 可以调节瞬态动力学的“速度”, 这是应用程序的重要考虑因素。对于工作的认知系统 (例如大脑), 仅在确定状态并对其进行操作所需的最短时间范围内, 过渡状态就保持稳定就足够了。更长的时间只会是“浪费时间”。

2.7 通用性

我们注意到, 在此提出的用于产生稳定的瞬态动力学的机制是通用的, 因为它可以应用于瞬时状态下的广泛的动力学系统, 即由吸引子和循环确定。

从物理上讲, 我们在这里提出的机制是通过与其他局部慢变量 ϕ_i 耦合, 将吸引网络的相空间 x_i 嵌入更大的空间 $\{x_i(t), \phi_i(t)\}$ 。由于新变量允许系统通过局部逃

逸过程进行相应的逃逸,因此新变量允许系统逃逸原始吸引子 $\{x_i = 1/0, \varphi_j = 0/1\}$ 的储层 $\varphi_i(t)$ 。请注意,嵌入是通过方程式 (3) 中的函数 $f_{z/w}(\varphi)$ 进行的,并且即使在等式 (3) 中通过设置 $f_{z/w}(\varphi) \rightarrow 1$ 来关闭耦合,储层变量仍会保持跟从的动力学 (4)。

图 4 说明了这种机制。局部性是该机制起作用的重要因素。否则,轨迹将不会再次接近任何吸引子,即原始吸引子被所有具有相似强度的排斥器所排斥,从而产生图 1 所示类型的波动动力。

2.8 周期和时间反转对称

图 2 和 5 所示的系统非常小,并且瞬态动态很快就陷入了吸引子破坏的循环中,因为在各个模拟中都没有考虑传入的感觉信号。对于包含大量顶点的网络,吸引子的数量可能会非常大,从而导致周期长度增加。我们对一个包含 713 个集合编码的内存的 100 个顶点的网络进行了仿真。即使对于包含多达 4400 个瞬态的瞬态序列,我们也没有发现任何循环行为。我们注意到,一旦给定的集合第二次稳定下来,系统就不一定会回溯其自身的轨迹,这在任何有限的系统中都需要发生,原因是当给定的给定等级时,储层的分布通常是不同的集合第二次被重访。

我们注意到,时间重复的对称性是“自发地”破坏的,因为这种重复性的瞬态动态类型

$$(\text{clique } A) \rightarrow (\text{clique } B) \rightarrow (\text{clique } A) \rightarrow (\text{clique } B) \rightarrow \dots$$

通常不会出现。原因很简单。当第一个集合停用后,其相应的储层高度需要一定时间才能再次充满,请比较图 2。然而,时间反转对称性将在极限 $\Gamma_{\phi}^{-} \ll \Gamma_{\phi}^{+}$ 中恢复,即当储层的补充比消耗的快得多时。

2.9 可重现的序列生成

动物需要产生多种活动的神经活动序列,例如对于运动或周期性的内部肌肉收缩,心跳就是最好的情况。这些序列需要是一系列明确定义的发射模式,可用于控制执行器,即肌肉。然后出现一个问题,在这种情况下,动力学系统会产生可重复定义的活动模式序列,即活动状态的受控时间序列[26, 27]。

在这种情况下,有两点值得注意。

1、方程 (2) - (4) 描述的动力学对于随时间变化可能会或不会变化的随机选择的链接矩阵 w_{ij} 都适用。具体而言,可以选择特定的方向,以便引起特定的一系列瞬态,图 5 中给出了一个示例。从原理上讲,该网络能够可靠地生成大量不同的序列瞬态。对于网络顶点和链接 w_{ij} 的几何排列,人们发现了遍及系统的瞬态波。

2、在第3节中，我们将讨论如何从外部老师向网络呈现的培训模式中学习适当的 w_{ij} 。我们将在第3节中集中讨论对个人记忆的训练和学习，适当的训练模式序列也可以用于学习记忆的时间序列。

3、自主在线学习

外部刺激 $\{b_i^{(ext)}(t)\}$ 影响各个神经中枢的活动 $x_i(t)$ 。这对应于各自的增长率 r_i 的变化，

$$r_i \rightarrow r_i + f_z(\varphi_i)b_i^{(ext)}(t), \quad (6)$$

比较方程(3)，其中 $f_z(\varphi_i)$ 是适当的耦合函数，具体取决于当地储层水位 φ_i 。当外部刺激的作用很强时，即当 $f_z b_i^{(ext)}$ 强烈时，通常会导致相应神经中枢 i 的激活 $x_i \rightarrow 1$ 。持续有效的刺激不会传达新的信息，另一方面，它会导致习惯化，从而减少对系统的影响。持续存在的强大刺激会导致所涉及的神经中枢的高活动水平 x_i 延长，从而通过(4)导致相应储层水平的耗竭，时间是由储层逆耗竭率 $1/T-\phi$ 给出的。然后，通过(6)中的耦合函数 $f_z(\varphi_i)$ 来调节，因为对于 $\varphi_i \rightarrow 0$ ， $f_z(\varphi_i)$ 变得非常小，比较图3。因此，(6)中包含的习惯化的影响使系统将其转向注意其他竞争性刺激，其中新颖的刺激更有可能影响正在进行的瞬态动力学。

现在，我们提供了一组学习规则，使系统可以在动态活动的正常阶段动态获取新模式。具有不同学习时期和性能的替代模型网络在神经网络的技术应用中得到了广泛的应用，但是在我们的连续活动的认知系统中，它并不是我们感兴趣的内容。

3.1 短期和长期的突触可塑性

对于具有瞬态动力学的神经网络来说，选择合适的突触可塑性有两个基本考虑因素。学习是一个非常缓慢的过程，没有短期记忆。训练模式需要一遍又一遍地呈现给网络，直到引起大量的突触变化为止[20]。短期记忆可以稳定外部模式，从而使系统有时间巩固长期的突触可塑性，因此可以大大加快学习过程。

使用稀疏编码的系统基于强大的抑制背景，即平均抑制链接强度 $|z|$ 。远大于平均兴奋性连接强度 \bar{w} ，

$$|z| \gg \bar{w}.$$

显然，渐进式学习只有在主要影响兴奋性联系时才有效：大参数的微小变化不会导致新的短暂吸引子，也不会实质性地影响认知动力学。

然后得出结论，将突触可塑性分为两部分很方便，

$$w_{ij} = w_{ij}(t) \equiv w_{ij}^S(t) + w_{ij}^L(t), \quad (7)$$

其中 $W^{S/L}$ 分别对应于短期和长期的突触可塑性。

3.2 负基线

根据上面的讨论，等式 (5) $z_{ij} = -|z| \Theta(-w_{ij})$ 表示抑制性链接强度为零或 $-|z|$ ，但在学习期间不会直接改变。因此，我们可以考虑两种“激励性联系优势”。

1、活动：活动 w_{ij} 为正，并通过公式 (5) 强制 $z_{ij} = 0$ 。

2、不活动：不活动的 w_{ij} 稍微为负，我们默认使用 $w_{ij} = W_L^{(\min)} < 0$ 。它通过方程式 (5) 对相同的链接强制执行 $z_{ij} = -|z|$ ，并且对动力学没有贡献，因为激励链接在 (3) 中输入为 $\theta(w_{ij})$ 。

当 w_{ij} 在学习过程中获得正值时，相应的抑制性链接 z_{ij} 通过等式 (5) 断开，而兴奋性链接 w_{ij} 确定等式 (3) 中相应项值。在整个模拟过程中，我们使用了一个小的负基线 $W_L^{(\min)} = -0.01$ 。

3.3 短期记忆动态

短期突触可塑性具有最大可能值 W (max) 是合理的。因此，我们考虑以下 Hebbian 型学习规则：

$$\dot{w}_{ij}^S(t) = \Gamma_S^+ (W_S^{(\max)} - w_{ij}^S) f_z(\varphi_i) f_z(\varphi_j) \Theta(x_i - x_c) \Theta(x_j - x_c) - \Gamma_S^- w_{ij}^S, \quad (8)$$

当突触前和突触后神经中枢都活跃时， $\dot{w}_{ij}^S(t)$ 以速率 Γ_S^+ 迅速增加，即当它们各自的活动高于 x_c 时。否则，其以速率 Γ_S^- 衰减至零。耦合函数 $f_z(\varphi_i)$ 阻止了短期记忆的长时间自我激活。当突触前和突触后中心的活动时间足够长，以至耗尽其各自的储层水平时，短期记忆将通过 $f_z(\varphi_i)$ 关闭。

在图 6 中，我们给出了一些选定的 $\dot{w}_{ij}^S(t)$ 的时间演化，用于使用图 2 中所示的网络进行仿真。在三种情况下，将激活短期记忆。

1、当现有的集合（即长期存储器 w_{ij}^L 中编码的集合）被激活时，如图 6 所示数据的 (0, 1) 的情况一样，相应的内部内部 w_{ij}^S 也被激活。这种行为是一个副作用，因为对于此处选择的参数值，短期链接强度的幅度明显小于长期链接强度的幅度。

2、在瞬态动力学过程中，当前活跃的集合与随后的活跃的集合存在一定的重叠。在这短时间内，激活了连接这两个集合的突触的短期可塑性 w_{ij}^S 。一个示例是图 6 中所示的仿真链接 (2, 4)。

3、当外部刺激作用在两个没有通过兴奋性长期记忆链接 w_{ij}^L 连接的位点上时，短期可塑性 w_{ij}^S 产生质的差异。暂时稳定相应的链接和相应的链接既可以单独成为新的集合 (i, j)，也可以成为已经存在的扩大集合的一部分。一个例子是图 6 中所示的仿真的链接 (3, 6)，但是请注意，如果不随后转移到长期记忆中，一旦引起外部刺激消失，这些新状态将以速率 Γ_S^- 消失。。

最后一点是最重要的一点，因为它可以暂时稳定感觉输入流中存在的新模式。

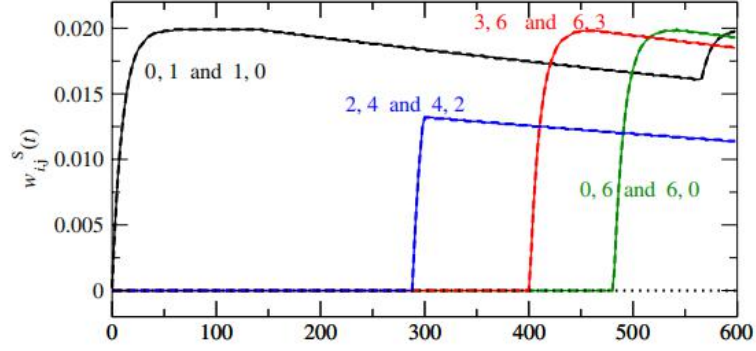


图 6 对于某些选定的连接 w_{ij}^S 和图 2 所示的网络，短期内存的时间演化，瞬态为 $(0, 1) \rightarrow (4, 5, 6) \rightarrow (1, 2, 3) \rightarrow (3, 6) \rightarrow (0, 6) \rightarrow (0, 1)$ 。位置 (3) 和 (6) 处的外部刺激作用时间 $t \in [400, 410]$ ，强度 $b^{(\text{ext})} = 3.6$ 。

3.4 长期记忆动力学

信息处理动力学系统只有在某些情况下保持其动力学特性时才能保留其功能。他们需要调节自己的工作点。对于此处讨论的具有瞬态动态特性的系统类型，其工作点如“简介”部分所述，定义为系统从一种准平稳状态过渡到另一种稳态所需的时间 Δt 。各个准平稳状态的长度 t ，由 $1/\Gamma_\phi^-$ 给出。

神经网络中的认知信息处理发生在短到中间时间尺度。为了使这些过程正常运行，需要对体内的整体突触可塑性（即长期记忆环节的平均强度）进行稳态调节。增长率 r_i 的平均幅度，请参见等式 (3)，确定完成从一个获胜集合到下一个过渡状态的过渡所需的时间 Δt 。因此，它构成了一个中心量调节系统的工作点，因为 $\bar{r} \sim 1/\Gamma_\phi^-$ 是固定的，所以储层枯竭率 Γ_ϕ^- 不会受到学习过程的影响，而学习过程只影响神经突触间的强度。

由于突触后储层进函数 $f_w(\varphi_i)$ 的时间相关性，增长率 $r_i(t)$ 非常强烈地与时间相关，请参见方程式 (3)。有效传入突触信号强度

$$\tilde{r}_i = \sum_j [w_{i,j} + z_{i,j} f_z(\varphi_j)] x_j, \quad (9)$$

独立于突触后储层 φ_i ，是一个更方便的局部控制参数。当所有位置的有效传入信号平均具有可比较的幅度 $r^{(\text{opt})}$ 时，认知系统的工作点是最佳的，

$$\tilde{r}_i \rightarrow r^{(\text{opt})}. \quad (10)$$

长期记忆有两个任务：通过无监督学习来提取和编码外部刺激方程 (6) 中存在的模式，并将动态系统的工作点保持在所需范围内。可以通过一个本地学习规则来完成这两项任务，

$$\dot{w}_{ij}^L(t) = \Gamma_L^{(\text{opt})} \Delta \tilde{r}_i \left[\left(w_{ij}^L - W_L^{(\text{min})} \right) \Theta(-\Delta \tilde{r}_i) + \Theta(\Delta \tilde{r}_i) \right] \Theta(x_i - x_c) \Theta(x_j - x_c), \quad (11)$$

$$-\Gamma_L^- d(w_{ij}^L) \Theta(x_i - x_c) \Theta(x_c - x_j), \quad (12)$$

现在，我们评论这些演化方程的一些性质：

1、**Hebbian 学习**。学习规则 (11) 是属于 Hebbian 类型。仅当突触前和突触后神经元处于活动状态时，即当它们各自的活动水平高于阈值 x_c 时，才进行学习。函数 $d(w_{ij}^L)$ 确定遗忘对实际突触强度的功能依赖性；为了简单起见，我们使用 $d(w_{ij}^L) = \theta(w_{ij}^L)w_{ij}^L$ 。

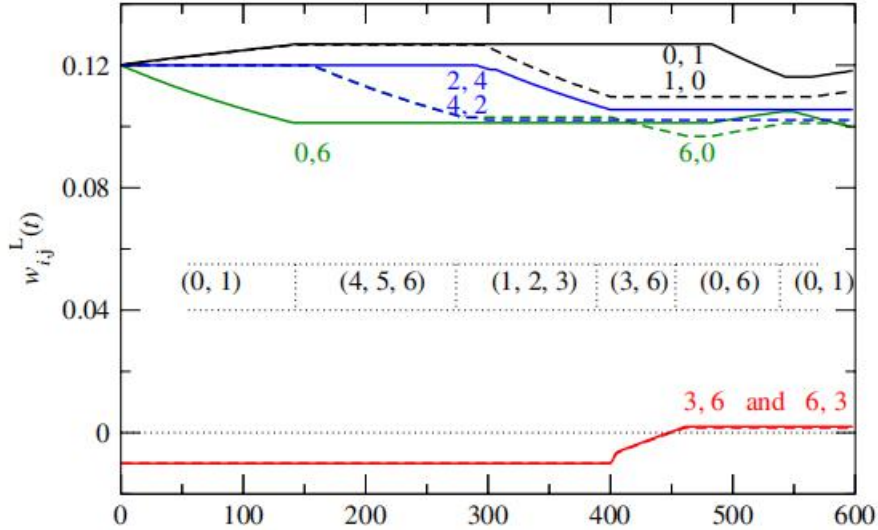


图 7.某些选定项的长期记忆的时间演变。连接 w_{ij}^L 和图 2 所示的网络，没有连接 (3, 6)。瞬态为 $(0, 1) \rightarrow (4, 5, 6) \rightarrow (1, 2, 3) \rightarrow (3, 6) \rightarrow (0, 6) \rightarrow (0, 1)$ 。位置 (3) 和 (6) 处的外部刺激作用于 $t \in [400, 410]$ ，强度为 $b^{(ext)} = 3.6$ 。刺激模式 (3, 6) 已被系统学习，在学习间隔 $t \in [400, 460]$ 中 $w_{3,6}$ 和 $w_{6,3}$ 变为正。由于短期记忆的激活，学习间隔明显长于裸露的刺激长度。在没有外部刺激的情况下，某些 w_{ij}^L 的衰减是由于遗忘 (11) 所致，通常这是一个非常微弱的效应。

2、**突触竞争**。当有效输入信号 \tilde{r}_i 相对于最佳值 $r^{(opt)}$ 弱/强时，活动连接被增强/减弱，其中 W_L^{min} 是 w_{ij} 的最小值。比较图 3 和图 7，基线 W_L^{min} 略为负。然后，以输入突触之间的时间竞争的形式进行 Hebbian 型学习，通常，活动的输入链接平均会获得强度。

3、**快速学习新模式**。在图 7 中，显示了某些选定 w_{ij}^L 的时间演变。网络学习了一种简单的输入模式。在此仿真中，学习参数 $\Gamma_L^{(opt)}$ 已设置为一个很大的值，使得学习可以一步完成（快速学习）。

4、**抑制失控的突触生长**。当神经网络反复受到相同或相似的外部刺激时，无监督学习通常会导致所涉及的突触强度不受控制地增长。这种现象被称为“突触生长失控”，也可能在具有连续自我生成活动的网络中发生，这种情况是一遍又一遍自动生成类似的活动模式。所提出的链接动力学抑制了两种突触失控生长 (11)。

5、基线为负。注意， $w_{ij} = w_{ij}^S + w_{ij}^L$ 作为 $\theta(w_{ij})$ 进入演化方程 (3)。我们因此可以区分活动配置 ($w_{ij} > 0$) 和非活动配置 ($w_{ij} < 0$)，比较图 3。输入 (11) 的负基线 $w_L^{(\min)} < 0$ 然后可以消除正向链接，并提供防止小的障碍。

在瞬态中，对于获胜集合，所有顶点的 $x_i \rightarrow 1$ ，而集合外顶点 $x_j \rightarrow 0$ ，会导致

$$\tilde{r}_i \approx \sum_j w_{i,j}, \quad j \in \text{active sites},$$

因此，工作点优化规则 (10) $\tilde{r}_i \rightarrow r^{(opt)}$ 等效于本地归一化条件，该条件使活动的传入连接强度之和为恒定值，即与顶点无关。该规则与 Bienenstock 等[28]提出的自我调节皮质神经元平均放电速率的机制密切相关。

3.5 在线学习

我们在这里考虑的神经网络是连续活动的，与是否有通过等式 (6) 的感觉输入无关，原因是进化方程 (2) 和 (4) 生成了一个永无止境的瞬态时间序列。这是本研究的主要假设，连续和自发的神经活动是建模整体脑部活动或发展自主认知系统的必要条件[3]。

突触可塑性的演化方程，即短期记忆的 (8) 和长期记忆的 (11)，是动力学系统的一部分，即它们确定了在任何时候，无论是否通过 (6) 向网络提供了外部刺激的 w_{ij}^S 和 w_{ij}^L 的时间演化。因此，对于一个连续活跃的神经网络，通常需要满足突触可塑性的演化方程两个条件。

(a) 在训练条件下，即通过 (6) 将输入模式呈现给系统时，系统应能够相应地修改突触链接强度，以使训练模式以新记忆的形式存储，即代表吸引子并导致准平稳状态的集合。

(b) 在没有输入的情况下，持续的瞬态动力学将通过 (8) 和 (11) 不断导致突触修饰。这些修改可能不会引起质变，例如现有内存的自动破坏或伪造新内存的自发生成。新的记忆应专门通过外部刺激训练获得。

在以下部分中，我们将提供仿真以调查这些问题。我们发现，由于方程 (11) 中长期突触可塑性的优化原理，本研究中制定的演化方程同时满足上述条件 (a) 和 (b)。

4、模拟

我们通过不断学习，对多达数千个顶点的系统进行了广泛的网络动态模拟。我们发现，即使存在由方程式 (8) 和 (11) 支配的连续在线学习，动力学也能保持长期稳定，表现出获胜集合的半规则序列，如图 2 所示。工作点是自适应调节的在模拟中没有观察到长时间的停滞或困滞状态；都没有发生快速或不受控制的振荡周期。

最终，在没有外部信号的情况下，具有有限数量的顶点 N 和集合的有限数量的任何系统最终都会陷入一系列的瞬态状态。对 N 为 20-100 的系统的初步研究导致周期平均跨越网络编码的所有集合的有限部分。这是一个显着的结果，因为比较方程式 (1)，因为存储在网络中的集合总数很容易比顶点 N 本身的数量大几个数量级。自治网络的循环行为的详细研究将在其他地方介绍。

表 1. 具有 N 个顶点和 N_{links} 个兴奋性链接以及包含 2 个, ..., 6 个顶点的 N_2, \dots, N_6 集合的系统的学习结果。 N_{tot} 是要学习的内存总数。 N_1 和 N_{par} 表示完全/部分学习的记忆的数量。

N	N_{links}	N_2	N_3	N_4	N_5	N_6	N_{tot}	N_1	N_{par}
20	104	1	10	42	11	1	65	60	3
100	901	26	563	122	2	0	713	704	7

4.1 学习新的记忆

网络通过 $r_i \rightarrow r_i + f_w(\phi_i)b_i^{(\text{ext})}(t)$ 外部呈现给系统的训练模式 $\{p_1, \dots, p_Z\}$ ，以及激活短路功能由网络获知对应模式内链接的长期记忆。在图 6 和 7 中，我们提供了一个案例研究。正在进行的内部瞬态动态在时间 $t=400$ 处被外部信号中断，该信号激活了短期记忆，请参见图 6。请注意，只要给定链接变为活动状态，短期记忆就会被外部刺激和内部激活，即当突触前和突触后位点同时活跃时。但是，内部激活不会导致内部生成虚假的内存，因为内部激活的链接无论如何都属于一个或多个已经存在的派系。

在图 7 中，我们展示了各个长期突触修饰 w_{ij}^L 的时间发展。此处选择的学习参数可实现快速学习。暂时保留在短期存储器中的，与外部信号相对应的模式将进一步存储，即在过渡到下一个集合之前，相应的 w_{ij}^L 变为正值。对于实际应用，较小的学习速率可能更适合，因为它们可以避免学习由环境噪声生成的寄生信号。

在表 1 中，我们提供了从头开始学习具有 $N_2=0$ 和 $N=100$ 的两个网络的结果。初始网络仅包含两个相连的集合，以实现非平凡的初始瞬态动态。所有其他链接都是抑制性的。通过使用与图 6 和 7 相同的参数，通过训练和存储外部呈现的模式来进行学习几乎是完美的。如我们所测试的，学习率可以在很宽的范围内选择。在这里，对于 100 个顶点的网络，到 $t=5 \cdot 10^4$ 时，培训阶段已经完成。回到 3.5 节的讨论，我们得出结论，网络满足制定的条件 (a)，能够有效地存储培训模式作为集合形式的吸引者。

4.2 连接不对称

我们注意到，通过工作点优化方程式 (11) 进行的 Hebbian 学习导致链接矩阵中不对称性的自发生成，即 $w_{ij}^L \neq w_{ji}^L$ ，因为突触可塑性取决于突触后的增长率。

在图8中,对于两个模拟,我们给出了表1的100个顶点的网络在时间 $t=5 \times 10^4$ 时所有正 w_{ij}^L 的连接不对称 $w_{ij}^L - w_{ji}^L$ 的分布。图8所示的分布是特定实现稳态分布,即它们在较大范围的总仿真时间内没有明显变化。

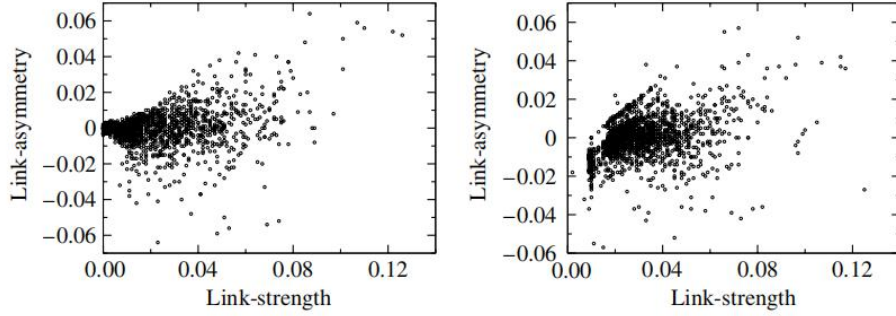


图8.在时间 $t=5 \times 10^5$ 时有713个集合的100个顶点的网络的正 w_{ij}^L 的连接不对称 $w_{ij}^L - w_{ji}^L$, 对应于大约4500个瞬态。左面板:从头开始学习。训练在 $t \approx 5 \times 10^4$ 时结束。右图:从 $w_{i,j} \rightarrow 0.12$ 开始,属于一个或多个集合的所有链接。

(i) 在第一次仿真中,网络是从头开始学习的。将713个训练模式的集合提交给网络 $t \in [0, 5 \times 10^4]$ 。之后,对于 $t \in [5 \times 10^4, 5 \times 10^5]$,系统自由演化。在时间 $t = 5 \times 10^5$ 总共产生了3958个瞬态,但是由于进行中的突触优化,方程式(8)和(11),系统尚未陷入瞬态循环。在 $t = 5 \times 10^5$ 时还剩下661个团体,因为链接竞争导致了对一些很少使用的链接的压制。

(ii) 在第二个模拟中,对于所有的内部链接,均已在 $t=0$ 处手动设置了均匀且对称的起始激励链接 $w_{ij}^L \rightarrow 0.12$ 。与N相同的N=100网络在没有外部刺激的情况下进行了仿真。尽管链接强度分布已从最初的均匀分布到平稳分布进行了大幅重组,但在 $t = 5 \times 10^5$ 时,所有713个队列仍然存在。在模拟过程中,总共生成了4123个瞬态,而系统不会进入周期。

对于这两种模拟,所有演化方程,即活动和储层的演化方程(2)和(4),短期记忆的方程(8),长期记忆的方程(11)和(12)都确定了时间 $t \in [0, 5 \times 10^4]$ 的动态。(i)与(ii)之间的差异是通过外部刺激,方程式(6)进行训练来确定记忆的方式,如(i)所述,或通过人工进行(ii)进行。

比较图8所示的两个链接分布,我们注意到总体相似性,这是连续作用的工作点优化的结果。对于较小的链接强度,主要区别在于,因为这两个模拟都是从相反的极端(消失/强烈的初始激励链接)开始的。链接分布取决于参数的详细信息如图8所示。对于图8所示的结果,我们出于说明目的使用了 $\Gamma_L = 0.1$,这对于调节弱健忘的参数来说是一个非常大的值。我们还用另一个极端 $\Gamma_L = 0$ 进行了仿真,发现连接不对称分布更加分散。

回到 3.5 节中的讨论，我们得出的结论是网络满足了制定的条件 (b)，因为在随后的自由进化过程中，基本上没有破坏训练状态下获得的记忆或自发创建了虚假的新记忆。

5、结论

我们研究了有关具有自主生成的瞬态动力学的神经网络的一系列问题。我们提出了一种通用方法，可以将初始吸引子网络转换为能够生成无限时间序列的瞬态的网络。最终的动力学系统具有严格收缩的相空间，瞬态与原始网络的吸引子之间具有一对一的绝热对应关系。

然后，我们讨论了稳态的问题，即系统需要自适应地调节其自身的工作点。我们为无监督的 Hebbian 型学习制定了一个简单的学习规则，从而解决了稳态问题。我们在这里注意到，该规则（等式 (11)）类似于为优化离散时间神经网络的整体存储容量而显示的学习规则[22]。

我们研究了使用集合编码的连续时间神经网络模型，并表明该模型非常适合于在广泛的学习条件下结合正在进行的动态学习来研究瞬态动力学。新记忆的快速在线学习和慢速学习都与网络自行生成的瞬态动态兼容。最后，我们转向对瞬态动力学的解释。对随后激活的团体的典型时间序列（如图 2 所示）进行检查后发现，集合的序列不是随机的。每个集合都通过兴奋性联系与其前身联系在一起；据说它们是“关联”连接的[29]。因此，随后的活跃集合的顺序可以被视为“联想思想过程”[29]。但是，仍需要研究此类过程在认知信息处理中的可能用途。

6、参考文献

- [1] Katok A, Hasselblatt B and Mendoza L 1995 Introduction to the Modern Theory of Dynamical Systems(Cambridge: Cambridge University Press)
- [2] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities Proc. Natl Acad. Sci. USA 79 2554
- [3] Gros C 2007 Autonomous dynamics in neural networks: the dHAN concept and associative thought processes Cooperative Behaviour in Neural Systems (Ninth Granada Lectures) ed P L Garrido, J Marro and J J Torres AIP Conf. Proc. 887 129–138 (Preprint q-bio.NC/0703002)
- [4] Abeles M et al 1995 Cortical activity flips among quasi-stationary states Proc. Natl Acad. Sci. USA 92 8616
- [5] Kenet T, Bibitchkov D, Tsodyks M, Grinvald A and Arieli A 2003 Spontaneously emerging cortical representations of visual attributes Nature 425 954
- [6] Ringach D L 2003 States of mind Nature 425 912

- [7] Hutt A and Riedel H 2003 Analysis and modelling of quasi-stationary multivariate time series and their application to middle latency auditory evoked potentials *Physica D* 177 203
- [8] Freeman W J 2003 Evidence from human scalp electroencephalograms of global chaotic itinerancy *Chaos* 13 1067
- [9] Kaneko K and Tsuda I 2003 Chaotic itinerancy *Chaos* 13 926
- [10] Timme M, Wolf F and Geisel T 2002 Prevalence of unstable attractors in networks of pulse-coupled oscillators *Phys. Rev. Lett.* 89 154105
- [11] Treves A 2005 Frontal latching networks: a possible neural basis for infinite recursion *Cogn. Neuropsych.* 22 276
- [12] Kropff E and Treves A 2006 The complexity of latching transitions in large scale cortical networks *Nat.Comput.* doi:10.1007/s11047-006-9019-3
- [13] Rabinovich M, Volkovskii A, Lecanda P, Huerta R, Abarbanel H D I and Laurent G 2001 Dynamical encoding by networks of competing neuron groups: winnerless competition *Phys. Rev. Lett.* 87 068102
- [14] Seliger P, Tsimring L S and Rabinovich M I 2003 Dynamics-based sequential memory: winnerless competition of patterns *Phys. Rev. E* 67 011905
- [15] Rabinovich M I, Varona P, Selverston A I and Abarbanel H D I 2006 Dynamical principles in neuroscience *Rev. Mod. Phys.* 78 1213
- [16] Sompolinsky H and Kanter I 1986 Temporal association in asymmetric neural networks *Phys. Rev. Lett.* 57 2861
- [17] Horn D and Usher M 1989 Neural networks with dynamical thresholds *Phys. Rev. A* 40 1036
- [18] Metzler R, Kinzel W, Ein-Dor L and Kanter I 2001 Generation of unpredictable time series by a neural network *Phys. Rev. E* 63 056126
- [19] Paula D R, Araujo A D, Andrade J S, Herrmann H J and Gallas J A C 2006 Periodic neural activity induced by network complexity *Phys. Rev. E* 74 017102
- [20] Arbib M A 2002 *The Handbook of Brain Theory and Neural Networks* (Cambridge, MA: MIT Press)
- [21] Chechik G, Meilijson I and Ruppin E 2001 Effective neuronal learning with ineffective Hebbian learning rules *Neural Comput.* 13 817
- [22] Okada M 1996 Notions of associative memory and sparse coding *Neural Netw.* 9 1429
- [23] Amit D J, Gutfreund H and Sompolinsky H 1985 Storing infinite numbers of patterns in a spin-glass model of neural networks *Phys. Rev. Lett.* 55 1530

- [24] von der Malsburg C and Schneider W 1986 A neural cocktail-party processor Biol. Cybern. 54 29
- [25] Gray C M, König P, Engel A K and Singer W 1989 Oscillatory responses in cat visual cortex exhibit incolumnar synchronization which reflects global stimulus properties Nature 338 334
- [26] Huerta R and Rabinovich M 2004 Reproducible sequence generation in random neural ensembles Phys. Rev.Lett. 93 238104
- [27] Rabinovich M I, Huerta R, Varona P and Afraimovich V S 2006 Generation and reshaping of sequences in neural systems Biol. Cybern. 95 519
- [28] Bienenstock E L, Cooper L N and Munro P W 1982 Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex J. Neurosci. 2 32
- [29] Gros C 2005 Self-sustained thought processes in a Dense Associative Network (Springer Lecture Notes in Artificial Intelligence (KI2005) vol 3698) p 375 (Preprint q-bio.NC/0508032)