

57-60

计算机科学1998 Vol. 25 No. 1

一种混沌神经网络和动态联想记忆

A Chaotic Neural Network and Dynamical Associative Memory

冯久超¹邱玉辉²

TP18

(西南师范大学物理系¹ 计算机科学系² 重庆400715)

摘 要 Based on Aihara's neuron model, a chaotic neural network model is constructed in this paper. The network is used to associate Chinese characters after the dynamical characteristic of the model is analyzed. Simulation result shows that the network can realize the dynamical associative memory of the embedded patterns and distinguish from similar patterns.

关键词 Chaos, Chaotic neural networks, Associative memory

1. 引言

混沌现象是非线性确定性系统中的一种内在随机过程的表现,普遍存在于自然界及人类社会中^[1]. 表征这种运动的基本特征量是李雅普诺夫指数(Lyapunov exponent). 当前,人们已经确认在人类脑电图(EEG)中存在混沌现象^[2],因为通过嵌入定

理在高维空间中对 EEG 时间序列的分析和计算表明了 EEG 具有正的李雅普诺夫特征指数. 不过,遗憾的是到目前为止科学家们还不能全面解释混沌在脑信息处理中的作用,但人脑在其长期进化中形成的完善性及混沌现象的普遍性使科学家们相信混沌在神经活动中起着重要作用,并充分认识到进行大量的理论和应用研究的必要性.^[3]

8. 冗余性. 能提高可靠性,鲁棒性.

9. 复杂性. 包括基因型的结构复杂性,解码复杂性,计算时-空复杂性(基因解码,适应值,再生等).

其中满意的特性有:完全性,可测性,复杂性. 但是以上特性有时是矛盾的.

编码应该具有完全性,否则对满足条件的神经网络结构可能无法编码. 满足条件的神经网络结构是否存在,若存在,是一个还是多个. 若有多个结构,它们之间可能差别极大. 具有很强描述能力的编码方法,一般比较复杂. 另外,如何判断是否完全,以及完善程度也有困难.

封闭性不满足时,将产生孤立的结点,从输入到输出无路径. 但在实际中,希望编码方法仅允许正确的网络,即所有编码对应着合理的网络. 同时仅正确的网络可以编码. 这就阻止了不正确的网络参与进化循环,省去了无用训练. 因此群体数和迭代数减少了,加快了进化学习时间,使问题易处理.

当一种编码允许无效的染色体通过遗传算子产生,修补它们不总是最好的. 修补结果是重要的遗传

物质失去了,降低了搜索的有效性.

编码方法应包括两方面:(1)编码过程,将神经网络映射成相应的编码表示.(2)解码过程,将编码表示映射成相应的神经网络. 不但要考虑编码过程,而且同时要考虑解码过程. 对两个过程同时评价. 最好两个转换过程都很方便,并且是一一对应的.

结束语 编码表示是进化神经网络的第一步,也是关键和核心的一步,它决定着整个进化过程的效率,及其进化神经网络的性能. 研究表明[Stephen, G. R. et al. 1995]:

1. 编码方法影响最后所生成的进化神经网络性能.

2. 编码串变太时,进化算法性能下降,伴随着进化生成的神经网络性能下降.

3. 对于小问题,网络中仅有少数结点,强编码表示机制优于其它两种表示机制.

4. 当问题大时,弱编码表示机制开始变得好了.

5. 文法编码方法相对而言不受问题大小的影响,用该编码方法进化生成的神经网络,总是比标准结构的神经网络要好.(参考文献共12篇,略)

冯久超 副教授,主要从事神经网络、信号和图像处理等研究工作. 邱玉辉 教授,校长.

混沌神经网络是以非线性的数学模型为依据,探索大脑这一复杂的非线性系统在混沌状态下的动态行为和动态信息处理能力,以及潜在的工程应用。本文在介绍了 Aihara 神经元模型^[4]以后,建立了一种混沌神经网络模型,并对该模型的动力学特征进行了分析。将这一网络模型用于汉字的动态联想记忆是本文的主要目的。最后,文末作了一些必要的讨论。

2. 混沌神经网络的建模

2.1 单个神经元模型

Aihara 神经元模型是离散时间的非线性差分方程,即

$$z(t+1)=f\{\beta(t)-\alpha\sum_{d=0}^{\infty}k^d g[z(t-d)]-\varphi\} \quad (1)$$

其中 $z(t)$ 是离散时刻 t 时神经元在 $[0,1]$ 之间的输出值; $\beta(t)$ 是外部激励; $f(\cdot)$ 是连续的输出函数; g 是描写神经元输出与不应性大小之间的关系函数; α 和 k 分别是刚性尺度参量(refractory scaling parameter)和不应性衰减率; φ 是阈值。

定义神经元的内部状态 $y(t+1)$ 为:

$$y(t+1)=\beta(t)-\alpha\sum_{d=0}^{\infty}k^d g[z(t-d)]-\varphi \quad (2)$$

若记 $\theta(t)=\beta(t)-k\beta(t-1)-\varphi\cdot(1-k)$,则由方程(1)和(2)可以导出 Aihara 模型的简化方程:

$$z(t+1)=f[y(t+1)] \quad (3)$$

$$y(t+1)=ky(t)-\alpha g[f(y(t))]+\theta(t) \quad (4)$$

2.2 网络模型

考虑到由其它内部神经元的反馈输入及外部输入后,由 N 个 Aihara 神经元组成的网络模型为:

$$y_i(t+1)=\sum_{j=1}^M u_{ij}\sum_{d=0}^{\infty}k_e^d \beta_j(t-d)+\sum_{j=1}^N w_{ij}\sum_{d=0}^{\infty}k_f^d z_j(t-d)-\alpha\sum_{d=0}^{\infty}k_r^d g[z_i(t-d)]-\theta_i \quad (5)$$

$$z_i(t+1)=f[y_i(t+1)] \quad (6)$$

式中 $y_i(t)$ 和 $z_i(t)$ 是第 i 个神经元的内部状态和输出; $u_{ij}(t)$ 和 $w_{ij}(t)$ 分别是第 i 个神经元与外部输入及与第 j 个神经元之间的连接权重; k_e 和 k_f 分别是外部输入、反馈输入的衰减参量; k_r 是记忆参量。若将 $y_i(t)$ 分解成三个状态变量 $y_i^{(1)}(t)$, $y_i^{(2)}(t)$, $y_i^{(3)}(t)$, 即满足:

$$y_i^{(1)}(t+1)=\sum_{j=1}^M u_{ij}\beta_j(t)+k_e y_i^{(1)}(t) \quad (7)$$

$$y_i^{(2)}(t+1)=\sum_{j=1}^N w_{ij}z_j(t)+k_f y_i^{(2)}(t) \quad (8)$$

$$y_i^{(3)}(t+1)=-\alpha g[z_i(t)]+k_r y_i^{(3)}(t)-\theta_i(1-k_r) \quad (9)$$

• 58 •

$$\text{其中: } y_i^{(1)}(t+1)=\sum_{j=1}^M u_{ij}\sum_{d=0}^{\infty}k_e^d \beta_j(t-d)$$

$$y_i^{(2)}(t+1)=\sum_{j=1}^N w_{ij}\sum_{d=0}^{\infty}k_f^d z_j(t-d)$$

$$y_i^{(3)}(t+1)=-\alpha\sum_{d=0}^{\infty}k_r^d g[z_i(t-d)]-\theta_i$$

分别是和外部输入及内部神经元的反馈输入相应的内部状态,以及记忆项。这样方程(6)变成:

$$z_i(t+1)=f[y_i^{(1)}(t+1)+y_i^{(2)}(t+1)+y_i^{(3)}(t+1)] \quad (10)$$

方程组(7)-(10)描述了神经网络的动力学行为,由方程(10)可以计算出网络的输出值。

为便于将上面的网络用于动态联想记忆,本文取 $g(x)=x$, $f(x)=1/[1+\exp^{-x/\epsilon}]$, 式中 ϵ 为温度参数;并考虑有两个状态变量 $y_i^{(2)}$ 和 $y_i^{(3)}$ 的简化方程:

$$z_i(t+1)=f[y_i^{(2)}(t+1)+y_i^{(3)}(t+1)] \quad (11)$$

$$y_i^{(2)}(t+1)=\sum_{j=1}^N w_{ij}z_j(t)+k_f y_i^{(2)}(t) \quad (12)$$

$$y_i^{(3)}(t+1)=-\alpha z_i(t)+k_r y_i^{(3)}(t)+c_i \quad (13)$$

式中 c_i 表示第 i 个单元的阈值与固定的外部输入之和。方程组(11)-(13)是在外部激励为常数的情况下方程组(7)-(10)的简化方程组,这时 $y_i^{(1)}(t)$ 包含在时间常数 c_i 之中,这就是本文要研究的网络。很明显,当 $k_f=k_r=\alpha=0$ 时,该网络就转化为 Hopfield 自联想记忆网络^[5]。

3. 网络的动力学特征分析

正的 Lyapunov 特征指数是混沌运动的标志^[1],因此,下面分析和计算本文所研究网络的 Lyapunov 特征指数谱。

考虑由 16×16 个神经元组成的网络,每个神经元有两个内部状态,用变量 $y_i^{(2)}$ 和 $y_i^{(3)}$ 来表示,这样整个网络有 512 个内部状态变量,定义为:

$$Y(t)=(y_1^{(2)}, y_2^{(2)}, \dots, y_{256}^{(2)}, y_1^{(3)}, y_2^{(3)}, \dots, y_{256}^{(3)}) \quad (14)$$

我们可以将整个网络视为 512 维状态空间中的非线性动力学系统,即

$$Y(t+1)=F(Y(t)) \quad (15)$$

若记 $Y(t)$ 的无穷小偏差为 $\delta Y(t)$, 则有

$$\delta Y(t+1)=DF[Y(t)]\delta Y(t) \quad (16)$$

式中 DF 是 F 在 $Y(t)$ 处的雅可比矩阵。为了计算 Lyapunov 特征指数谱,考虑在 512 微切空间中的基矢 u_i ($i=1, 2, 3, \dots, 512$) 的变换:

$$e_i(t+1)=DF[u_i(t)] \quad (17)$$

式中 $e_i(t+1)$ 是非正交的变换矢量集。利用 Gram-

Schmidt 正交变换过程,可以形成正交化基矢 $u_i(t+1)$:

$$e'_i(t+1) = e_i(t+1) - \sum_{j=1}^{i-1} \langle e_i(t+1), u_j(t+1) \rangle u_j(t+1) \quad (18)$$

$$u_i(t+1) = \frac{e'_i(t+1)}{\|e'_i(t+1)\|} \quad (19)$$

其中 $\langle \cdot \rangle$ 表示内积, $\| \cdot \|$ 代表矢量范数。重复迭代 (17)-(19), 就可以计算出 Lyapunov 指数谱 $\lambda_i (i=1, 2, 3, \dots, 512)$:

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=1}^t \log \|e'_i(s)\| \quad (20)$$

假如 $\text{Max}\{\lambda_i, i=1, 2, \dots, 512\} > 0$, 该网络就表现出混沌行为。

图1是网络参数为神经元数 $N=256, c_i=2.0 (i=1, 2, 3, \dots, 256), \alpha=10.0, k_i=0.2, k_r=0.99$, 及 $\epsilon=0.015$ 的情况下, 最大 Lyapunov 指数随时间演化的曲线, 由此可知本文所研究的网络是混沌神经网络。下节将仿真该网络在混沌状态下对汉字的动态联想记忆。

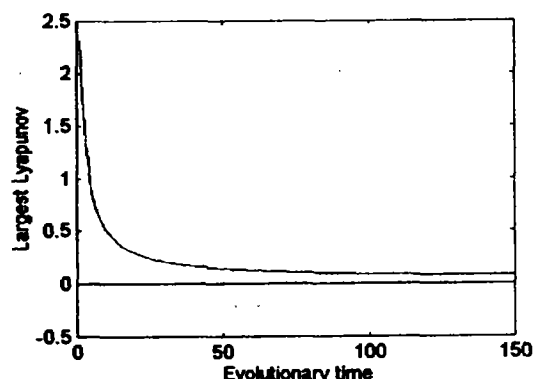


图1 最大 Lyapunov 指数随时间演化的曲线。网络参数为: 神经元数 $N=256, c_i=2.0 (i=1, 2, 3, \dots, 256), \alpha=10.0, k_i=0.2, k_r=0.99, \epsilon=0.015$ 。

4. 动态联想记忆

将上面的混沌神经网络(网络参数同上)用于联想记忆四个 16×16 点阵汉字, “余”, “余”, “代”, “伐”。在仿真过程中, 对应这四个已存储模式的对称性相关矩阵为:

$$w_{ij} = \sum_{r=1}^4 P_i^{(r)} P_j^{(r)}, \quad (21)$$

式中 $P_i^{(r)}$ 和 $P_j^{(r)}$ 分别代表第 r 个已存模式的第 i 个分量和第 j 个共轭分量, 其共轭分量满足:

$$\sum_{i=1}^{256} P_i^{(r)} P_j^{(r)} = \delta_{rj}$$

$$P_i^{(r)} = \sum_{j=1}^4 h_{ij} P_j^{(r)}, h_{ij} = (4 \sum_{r=1}^4 P_i^{(r)} P_j^{(r)})^{-1} \quad (22)$$

图2显示了在初始化模式($t=0$)的激励下, 该网络的联想记忆过程。这个演化过程说明尽管目标模式有一定的相似性(如本例中的“代”和“伐”, “余”和“余”), 该网络也能恢复记忆出原有的嵌入模式, 且能区分相近的模式。为了比较, 本文也仿真了相应的 Hopfield 网络的自联想记忆过程, 如图3所示。从 $t=3$ 开始, 网络收敛到一个目标模式“伐”。由此可见, 对应的 Hopfield 联想记忆网络在记忆恢复时, 最终输出模式是静态和唯一的。

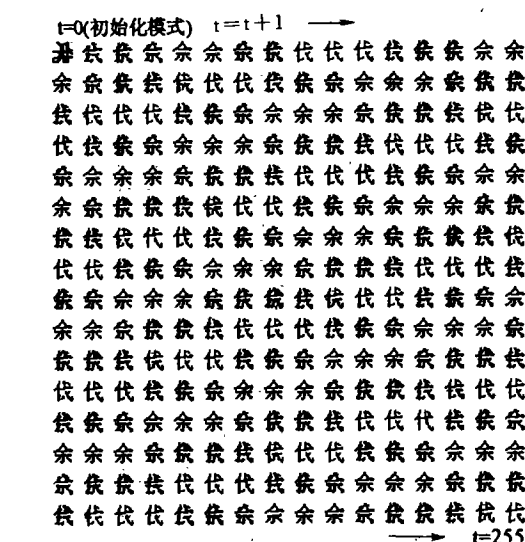


图2 混沌神经网络对嵌入模式“代”, “伐”, “余”, “余”的联想记忆过程

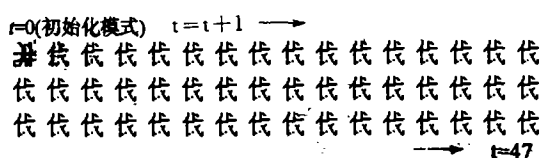


图3 Hopfield 联想记忆网络对嵌入模式“代”, “伐”, “余”, “余”的联想记忆过程

5. 讨论

本文构造的混沌神经网络是离散时间的非线性全连接网络, 在适当的参数条件下, 整个网络呈现出混沌行为。由于混沌吸引子是相空间中无数不稳定周期轨道的集合, 若将该混沌吸引子作为一个记忆

60-63

BP 网络的权值诱导与层次训练算法

Weight Value Induction of BP Network and Corresponding Hierarchical Training Algorithm

胡建军 黄安贻 张仲甫

TP18

(武汉汽车工业大学机电工程学院 武汉430070)

摘 要 According to the difficulties of the physical realization of BP network, this paper presents a valuable weight induction training algorithm to induce and adapt the network weight values to the required range of electronical circuits and parts. As an example, a BP network with two hidden layers on 3×3 TC recognition problem is trained with this algorithm, and a brief comment is presented.

关键词 BP network, Training algorithm, Physical realization, Optimization algorithm

BP 神经网络以其结构简单、工作状态稳定等特点,成为当前应用最广的一种神经网络模型,应用范围包括识别分类、非线性映射、复杂系统仿真等方面。但是,BP 网络也存在局部极小、收敛速度慢、网络结构选取困难等缺点。对此,已提出了多种改进算

法^[3]。然而,事实上,BP 网络的硬件实现是其从实验室走向应用的关键环节,怎样得到一组既能达到预定功能又便于硬件实现的网络结构与权值,就成为 BP 网络研究的一个重要方面。本文针对这个问题,提出了 BP 网络的权值诱导的概念以及相应的一种

单位,先将目标模式(允许有相近的模式)全部嵌入到此混沌吸引子中,在记忆恢复阶段,网络便不会收敛于某一特定的目标模式,即记忆恢复过程是动态的,但能遍历目标模式(混沌遍历性)。这一点有一定的生理依据,因为生物信息处理也是动态的,它必须依据环境的变化而变化。作为对比,本文也仿真了 Hopfield 联想记忆网络的记忆恢复过程,它就不能再现全部目标模式,恢复过程是静态和唯一的,缺乏灵活性。因此,该混沌神经网络有助于克服在利用不动点进行信息处理中常见的网络陷入局部最小的困难局面。也应该看到,这种混沌神经网络没有一种最终的期望输出模式,若要获得期望的目标模式,一方面可以通过引入模式识别层,进行判断和识别;另一方面也可以通过混沌控制,使整个网络收敛于某一周期轨道(即某一固定点,对应于某一目标模式),或者研究新的学习算法(如自组织学习和竞争等)。此外,扩大网络的记忆容量也是值得研究的。所有这些问题都需要进一步深入研究。

总之,探索混沌神经网络在动态行为下的信息处理过程和能力是面向未来的研究课题,随着人们对神经系统的逐步理解,今后的研究也会渐渐深入,并将丰富神经网络的理论和应用。

参考文献

- [1] Yorke J., Grebogi C., The impact of chaos on science and society, United Nations University press, 1993
- [2] Masahiro N., Masaki O., On the chaos region of the modified Nagumo-Sato model, J. of the Physical Society of Japan, 61(4), 1992
- [3] 韩明虎、余英林, 基于耦合的混沌神经网络建模方法, 通信学报, 16(2), 1995
- [4] Aihara K. et al., Chaotic neural networks, Physics Letters, 144 A(6,7) 1990
- [5] Hopfield J. J., Neural networks and physical systems with emergent collective computation abilities, Proc. of National Academy of Sciences, USA, 1982, 79

胡建军 硕士生,研究方向:智能 CAD/CAM。黄安贻 博士生,研究方向:机械制造与自动化。张仲甫 教授,博士生导师,研究方向:机械制造与自动化。