

量子系统实现神经计算的理论分析¹

解光军 * ** 李 斌 * 庄镇泉 *

*(中国科技大学电子科学与技术系 合肥 230026)

** (合肥工业大学应用物理系 合肥 230009)

摘 要 该文通过量子系统的演化与人脑信息处理过程的比较研究,发现量子理论与神经网络理论中存在有许多相对应的数学表述,利用其中有关的公式提出一个广义的基于量子系统的 Hebb 学习规则,并结合量子点系统设计了一个神经计算的理论模型。

关键词 量子信息处理, 神经计算, 量子系统

中图分类号 TN-052

1 量子系统与人脑信息处理

人工神经网络 (ANN) 是建立在极其简化的神经元模型基础之上,随着信息量与信息复杂度的增加,它已无法适应,这就需要寻找更基本、更自然的系统来加以改进。目前已有一些研究成果表明,人脑处理信息的过程可能与量子现象有关(即量子意识或量子思维),比如英国 Oxford 大学的 Penrose 教授在研究了量子理论与人脑意识的关系后指出,解决量子测量问题是最终解决意识问题的先决条件,他在 1989 年出版的著作 "The Emperor's New Mind" 中提到,曾有这样这样一个生理实验,当一个光子撞击蟾蜍的视网膜时,有时足以激发起一个神经冲动,但是对于人类而言该现象似乎被噪声所抑制,不过 Penrose 仍然认为,确实有事实表明在人体内的一些细胞对单个量子敏感,因此大脑中存在量子力学效应的可能性依然可以维持^[1]; 1994 年美国 Arizona 大学的 Hameroff 教授指出,在神经元内细胞支架的微管 (Cytoskeletal Microtubule) 之中或周围,意识是作为一个宏观量子态由量子级事件相干的一个临界级突现 (Emerge) 出来的^[2]; 1996 年 Perus 博士认为,量子波函数的坍塌 (Collapse) 十分类似于人脑记忆中的神经模式重构现象^[3],等等。

通过比较分析,我们认为量子系统的演化与人脑信息处理过程之间的关系为 (1) 量子系统是所有物理过程的微观基础,同样也应该是生物和心理过程的基础; (2) 量子系统的概念超出了物质中有关粒子、波、相互作用以及场的划分,它综合表现为一种不可分的并行分布式处理系统,是用来描述人脑作用整体性的比较合适的候选者; (3) 人脑作用源于复杂的系统动力学,并非传统人工神经网络所能完全描述,而量子系统具有与生物神经网络相似的动力学特征,因此将 ANN 与量子理论结合起来会更好地模拟人脑的信息处理过程。

2 量子理论和神经网络理论之间的对应关系

量子理论与神经网络理论在数学表述上有许多相似之处,这里,我们从信息处理的角度出发,来分析这些相似关系,并给出其中隐含的深刻的物理意义^[4,5]。

2.1 矢量表示

神经状态矢量与量子波函数的比较:

定义神经态矢为 $q(r, t)$, 神经模式为 v , 则神经态矢可表示为神经模式的叠加, 即

$$q(r, t) = \sum_{k=1}^p c_k(t) v_k(r) \quad (1)$$

¹ 2001-09-03 收到, 2002-06-24 改回

国家自然科学基金 (60171029)、中国科技大学量子通讯与量子计算开放研究实验室合作项目资助

其中 p 为分量的总数。

相应地, 量子波函数也可以表示为量子本征波函数的叠加, 即

$$\psi(\mathbf{r}, t) = \sum_{k=1}^p c_k(t) \varphi_k(\mathbf{r}) \quad (2)$$

其中 $\varphi_k(\mathbf{r})$ 为本征波函数。

对于式中的概率系数 $c_k(t)$, 在数学上可表示为标量积, 即投影:

$$c_k(t) = \langle v_k, q \rangle = \iint v_k(\mathbf{r})^* q(\mathbf{r}, t) d\mathbf{r} dt \quad (3)$$

$$c_k(t) = \langle \varphi_k, \psi \rangle = \iint \varphi_k(\mathbf{r})^* \psi(\mathbf{r}, t) d\mathbf{r} dt \quad (4)$$

2.2 时空积分

神经信号的时空积分为

$$q(\mathbf{r}_2, t_2) = \iint J(\mathbf{r}_1, \mathbf{r}_2, t_1, t_2) q(\mathbf{r}_1, t_1) d\mathbf{r}_1 dt_1 \quad (5)$$

其中 J 是单个突触连接的强度。在非相对论量子力学中同样有

$$\psi(\mathbf{r}_2, t_2) = \iint G(\mathbf{r}_1, \mathbf{r}_2, t_1, t_2) \psi(\mathbf{r}_1, t_1) d\mathbf{r}_1 dt_1 \quad (6)$$

其中 G 构成一个量子系统的 Green 函数或传播子, 它表示为一个矩阵, 描述了整个系统由初态到终态的并行分布式演化过程, 而系统自身变换到一个新的状态则表现为内部量子点间的大量的相互作用, 这种转换过程从信息论上来理解, 实际上就是联想。

2.3 Green 函数与模式相关编码

我们还可以将 (5), (6) 式看作是相关函数, 其中单个突触的信号传送由 Hebb 学习规则决定, 即

$$J(\mathbf{r}_1, \mathbf{r}_2, t_1, t_2) = \sum_{k=1}^p v_k(\mathbf{r}_1, t_1) v_k(\mathbf{r}_2, t_2) \quad (7)$$

Green 函数作为单个量子模式 φ_k 的相关系数, 有

$$G(\mathbf{r}_1, \mathbf{r}_2) = \sum_{k=1}^p \varphi_k^*(\mathbf{r}_1) \varphi_k(\mathbf{r}_2) \quad (8)$$

在相对论中 G 可转变为

$$S(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = -i \sum_{k=1}^p \sum_{j=1}^2 \varphi_k^j(\mathbf{r}_1, t_1)^* \varphi_k^j(\mathbf{r}_2, t_2), \quad t_2 > t_1 \quad (9)$$

$$S(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = i \sum_{k=1}^p \sum_{j=3}^4 \varphi_k^j(\mathbf{r}_1, t_1)^* \varphi_k^j(\mathbf{r}_2, t_2), \quad t_2 < t_1 \quad (10)$$

显然, (7) 与 (8) 式有着本质的不同, 后者包含有虚数部分, 这表明量子系统是不同于神经网络的复杂类型。

在神经网络中, 模式间的相关关系对于记忆至关重要, 而在量子力学中就对应于波函数不同部分之间的相位差, 由它们控制着几率幅度分布的时间演化, 因此改变波函数间的相位关系就类似于神经网络中的学习。

2.4 神经模式重构与量子坍缩

神经模式重构可表示为

$$q(\mathbf{r}, t) = \sum_{k=1}^p c_k(t) v_k(\mathbf{r}) \Rightarrow q(\mathbf{r}, t_0) = v_{k0}(\mathbf{r}) \quad (11)$$

而波函数坍缩可表示为

$$\psi(\mathbf{r}, t) = \sum_{k=1}^p c_k(t) \varphi_k(\mathbf{r}) \Rightarrow \psi(\mathbf{r}, t_0) = \varphi_{k0}(\mathbf{r}) \quad (12)$$

这是认知过程的重要特征之一, 它们皆是系统受到环境影响的结果, 环境偏向于选择那些最相似于环境状态的模式。

2.5 状态函数动力学

在量子理论与神经网络理论中存在着类似的动力学表示, 它们分别为

$$\dot{q}(\mathbf{r}, t) = \sum_{k=1}^p \lambda_k c_k(t) v_k(\mathbf{r}) \quad (13)$$

$$\dot{\psi}(\mathbf{r}, t) = -\frac{i}{\hbar} \sum_{k=1}^p E_k c_k(t) \varphi_k(\mathbf{r}, t) \quad (14)$$

其中 λ_k 为神经模式 v_k 的注意度参数, 即矩阵 J 的对应于本征矢量 v_k 的本征值; 而 E_k 具有与 λ_k 类似的作用。

3 基于量子系统的 Hebb 学习规则

由上面的对比分析可知, 量子系统的演化与神经网络系统的动力学在理论上有许多可以类比的地方, 因此我们利用量子理论中的有关公式, 并结合神经网络理论中的一些原理, 提出一个广义的基于量子系统的 Hebb 学习规则, 该算法可以实现量子系统中的神经计算。

假设 $|x\rangle = \alpha|0\rangle + \beta|1\rangle$, 其中 $|0\rangle$ 和 $|1\rangle$ 态可以表示为光的不同极化状态或电子的不同自旋方向等。

考察一个感知机 $y = f\left(\sum_{j=1}^n \omega_j x_j\right)$, 其中 Hebb 学习规则可以表示为

$$\omega_j(t+1) = \omega_j(t) + \eta(d - y)x_j \quad (15)$$

式中 η 为学习率, d 为期望输出。

在量子系统中, 我们使用量子力学符号将感知机改写为 $|y_i\rangle = \hat{F} \sum_{j=1}^n \omega_i^j |x_j\rangle$, 其中 $\hat{F}$ 是一个未知算子, 它可由量子门网络来实现。这时, 学习过程就转变为从初态到终态的演化过程, 即 $|y\rangle_m \Rightarrow |d\rangle_m$, 而 $|d_i\rangle = \hat{F} \sum_{j=1}^m \tilde{\omega}_i^j |x_j\rangle$ 。于是调节权值就变成寻找某个合适的集合 $\{\tilde{\omega}_i^j\}$, 因此它可以设定为一个量子算子, 即

$$H = H_d + H_y = E \left\{ \sum_{j=1}^m |d_j\rangle \langle d^j| + \sum_{j=1}^m |y_j\rangle \langle y^j| \right\} \quad (16)$$

式中 E 为常数, 代表系统的本征值或能量; 么正性保证了概率守恒。

定义 $|d_i\rangle = (U_{H_d+H_y})_i^k |y_k\rangle$, 则 $\tilde{\omega}_i^j = \sum_{k=1}^m (U_{H_d+H_y})_i^k \omega_k^j$. 再定义 $|d_i\rangle, |y_i\rangle$ 以正交基表示, 即 $\langle d|y\rangle = a \rightarrow 1$, 于是权值矩阵就变换 $\omega_i^j = |y_i\rangle\langle x^j|$, $\tilde{\omega}_i^j = |d_i\rangle\langle x^j|$, 可得 Hebb 学习规则为

$$\delta\omega_i^j = \eta(|d_i\rangle - |y_i\rangle)\langle x^j| = -iE \sum_k \{(H_d + H_y)_i^k\} |y_k\rangle\langle x^j| = -iE(|d_i\rangle\langle d|y\rangle + |y_i\rangle)\langle x^j| \quad (17)$$

特殊地, 取 $\langle d|d\rangle = 1$, 则有 $\langle d|y\rangle = a$, 可得

$$\delta\omega_i^j = -iE(|d_i\rangle a + |y_i\rangle)\langle x^j| \quad (18)$$

下面, 我们简单证明一下该算法的收敛性:

在最简单的情况下, 即 $\hat{F} = 1$ 为单位算子时, t 时刻网络的输出为 $|y(t)\rangle = \sum_{j=1}^n \omega_j(t)|x_j\rangle$.

学习规则变为

$$\omega_j(t+1) = \omega_j(t) + \eta(|d\rangle - |y(t)\rangle)\langle x_j| \quad (19)$$

于是得

$$\| |d\rangle - |y(t+1)\rangle \|^2 = \left\| |d\rangle - \sum_{j=1}^n \omega_j(t+1)|x_j\rangle \right\|^2 = (1 - n\eta)^2 \| |d\rangle - |y(t)\rangle \|^2 \quad (20)$$

可见, 当 $0 < \eta < (1/n)$ 且归一化输入态 $\langle x_j|x_j\rangle = 1$ 时, 网络的输出值最终收敛于期望值 $|d\rangle$, 同时, 我们还可以从 (17) 式出发, 利用类似于证明 Grover 算法收敛性的方法来证明, 上述广义的量子 Hebb 规则在收敛性上要优于非量子算法 [6].

4 量子点神经计算模型的实例分析

目前已有报道, 利用量子点可以初步实现量子计算, 并且也有人提出了量子点细胞自动机 (QCA) 的模型 [7], 在这里, 我们取一个五原子量子点系统来分析一下在该系统中实现神经计算的可能性.

制备量子点系统 (即量子点阵列) 的方法主要有外延生长和再结晶法等, 其原理是: 在衬底上淀积几组相邻的异质原子孤岛, 其中单个量子点之间的距离保证足够的近, 使得量子点上的过剩电子有可能通过隧穿效应 (Tunneling Effect) 穿越孤岛, 在不考虑外界势能的作用时, 在五原子分子中电子就会优先占据方格的对角位置, 从而产生一个双倍退化的基态, 这就是分子的极化, 共有两种状态, 记为 $|\pm 1\rangle$, 在 (21) 式中就是位置变量 $x(t)$. 我们将每种状态作为一个量子比特 (Qubit), 如图 1 所示 [8].

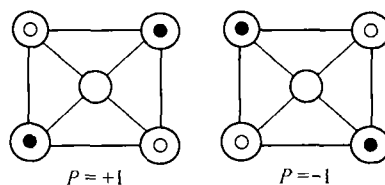


图 1 五原子量子点系统

下面, 我们利用该量子系统的时间演化方程与神经网络中神经元输出方程在数学形式上的相似性, 将量子点阵列转换为一个时间轴上的神经网络, 再引入合适的学习规则, 就可以实现神经计算.

具体地, 采用 Feynman 路径积分^[9]形式将量子点系统的状态演化过程表示为

$$\begin{aligned}
 |\Psi(x_T, T)\rangle &= G(x_T, T; x_0, 0) |\Psi(x_0, 0)\rangle \\
 &= \int_{(x_0, 0)}^{(x_T, T)} D[x(t)] \exp \left(\frac{i}{\hbar} \int_0^T d\tau \left[\frac{1}{2} m \dot{x}^2 - V(x) \right] \right) |\Psi(x_0, 0)\rangle \\
 &= \lim_{N \rightarrow \infty} \int_{(x_0, 0)}^{(x_{T-1}-x_T, T)} dx_1 \cdots dx_N \left(\frac{m}{2\pi i \hbar \Delta t} \right)^{(N-1)/2} \\
 &\quad \times \exp \left(\frac{i \Delta t}{\hbar} \sum_{j=0}^N \left[\frac{m}{2} \left(\frac{x_{j+1} - x_j}{\Delta t} \right)^2 - V(x_j) \right] \right) |\Psi(x_0, 0)\rangle \quad (21)
 \end{aligned}$$

其中 $|\Psi(x_0, 0)\rangle$ 是输入态, 表示系统在时刻 0 处于位置 x_0 的状态, 为量子点系统的初始态; $|\Psi(x_T, T)\rangle$ 是输出态, 为系统在时刻 T 的状态; G 为 Green 函数, 表示系统在时间轴上的前向演化; (21) 式第 2 行是用 Feynman 路径积分形式来表示 G , 它有无穷条路径, 每条路径的权值由复指数函数给出; 第 3 行是取时间段的求和表示, N 趋于无穷大. 此时, 若取 N 为有限值, 即满足准连续条件时, 那么在时间 $i\Delta t$ 上的 N 个中间态就可以被看作是 N 个量子神经元, 而神经计算中的非线性就隐含在方程的动量项 $(x_{j+1} - x_j)^2$ 和指数函数项之中, N 个神经元可能的输出状态则取决于最终的测量值, 它们可以通过改变势函数 $V(x)$ 来调节.

于是, 该量子系统的实时演化就变成了一个循环时间的量子神经计算模型, 我们用固定的初始电子态作为输入, 输出则由电子态在 T 时间后的测量值决定, 这样, 在原理上神经网络的节点就是在连续时间片段上的分子瞬时状态, 这些节点通过内部的相互作用与衬底发生局部的声学模式的间接作用, 而这些模式可以优先被光学激发, 因此可通过改变外界条件来控制, 把激发数理解为网络的训练权值.

除了调整势函数 $V(x)$ 外, 还可以将该系统耦合到外界环境中去, 从而获得一个外加的可训练的非线性参数. 具体地, 选择环境为一个高斯集合, 即具有二次 Hamilton 或正态分布特征, 在模型中它用量子点分子中的电子与衬底格子间的声子耦合来表示, 在物理上此类耦合很弱, 可以近似地表示为线性 (例如 GaAs 材料就满足此条件, 它的电子-声子耦合系数仅为 0.08, 远小于 1), 于是 (21) 式就可简化为

$$|\Psi(\sigma_z(N\Delta t), T)\rangle = \sum_{\{\sigma_z(j\Delta t)\}} \exp \left(\frac{i}{\hbar} \sum_j [K \sigma_x(j\Delta t) + \epsilon(j\Delta t) \sigma_z(j\Delta t)] \right) I(\sigma_z(t)) |\Psi(\sigma_z(0), 0)\rangle \quad (22)$$

其中 K 为隧穿概率; σ_z 为极化和; σ_x 为 Pauli 矩阵; $\epsilon(t)$ 为时变电场; $I(\sigma_z(t))$ 表示声子的总体效应.

进一步展开 (22) 式, 就可以得到以下 3 个可调整的参数: (1) 耦合强度 $\{\lambda_k\}$; (2) 每个时间片段 j 时刻的电场强度值 $\{\epsilon_j\}$; (3) 振荡频率 $\{w_k\}$. 相应地, 它们可以分别通过以下物理手段来实现控制: (1) 光学激发多声子; (2) 改变外电场强度; (3) 调节激发声子频率.

具体实现过程如下: 制备量子点分子阵列, 通过调节外电场强度初始化输入态, 并测定其极性, 然后使该量子系统自然演化, 在 T 时刻以某一个分子为对象, 测定其极性作为输出, 并以一定阈值加以限制, 给出实际输出值, 即

$$\text{Output} = \text{Th}(|\langle +|\Psi(\sigma_z(N\Delta t), T)\rangle| > |\langle -|\Psi(\sigma_z(N\Delta t), T)\rangle|) \quad (23)$$

其中 $\text{Th}(\)$ 为阈值函数.

在网络训练过程中可以采用上述的广义 Hebb 学习规则, 从而实现神经计算.

5 结 论

本文通过比较分析量子系统演化与人脑信息处理过程之间的关系,发现量子理论和神经网络理论中存在许多相似的数学表示,从而认为量子系统有可能实现神经计算。结合有关公式,我们提出了一个广义的基于量子系统的 Hebb 学习规则,并初步证明其收敛性能。最后,利用五原子量子点阵列,我们构造了一个量子神经计算模型,并进行相应的理论分析。研究表明,选择合适的量子系统能够实现神经计算的功能。

参 考 文 献

- [1] R. Penrose, *The Emperor's New Mind*, New York, Oxford University Press, 1989, 432–466.
- [2] S. R. Hameroff, Quantum coherence in microtubules: 'An intra-neuronal substrate for emergent consciousness?' *J. of Consciousness Studies*, 1994, 1, 91–118.
- [3] M. Perus, Neuro-quantum parallelism in brain-mind and computers, *Informatica*, 1996, 20, 173–183.
- [4] 曾谨言, 量子力学, 第三版, 北京, 科学出版社, 2000, 23–56, 270–295.
- [5] H. Haken, Synergetic Computers for Pattern Recognition, and Their Control by Attention Parameter, In *Neurocomputers and Attention II, Connectionism and Neurocomputers*, V. I. Kryukov, A. Holden (Eds.), UK, Manchester University Press, 1991, 551–556.
- [6] E. Fahri, S. Gutman, Analog analogue of a digital quantum computation, *Phys. Rev. A.*, 1998, 57, 2403–2407.
- [7] J. C. Lusth, B. Dixon, A characterization of important algorithms for quantum-dot cellular automata, *Information Science*, 1999, 113, 193–204.
- [8] E. Behrman, *et al.*, Simulation of quantum neural networks, *Information Science*, 2000, 128, 257–269.
- [9] R. P. Feynman, A. R. Hibbs, *Quantum Mechanics and Path Integrals*, New York, McGraw-Hill, 1965, 149–236.

THEORETICAL ANALYSIS ON THE IMPLEMENTATION OF NEURAL COMPUTATION IN QUANTUM SYSTEMS

Xie Guangjun* ** Li Bin* Zhuang Zhenquan*

*(*Department of Electronic Science and Technology of USTC, Hefei 230026, China*)

**(*Department of Applied Physics, Hefei University of Technology, Hefei 230009, China*)

Abstract Through the comparison between quantum systems and information processing in human brain it can be found that there are many similar mathematical expressions both in quantum theory and neural networks, with some formula a generalized Hebb learning algorithm is proposed based on quantum system, and its convergence performance is also certified. Finally, a quantum neural computational model in a quantum dot system is designed, after analysing it results in that this model can realize the learning of simple logic.

Key words Quantum information processing, Neural computation, Quantum systems

解光军: 男, 1970 年生, 讲师, 研究方向为量子信息处理、计算智能方法.

李 斌: 男, 1970 年生, 讲师, 研究方向为数据挖掘、神经网络.

庄镇泉: 男, 1938 年生, 教授, 研究方向为智能信息处理、计算机辅助设计.