计算机科学 2001Vol. 28№. 7

# 量子神经网络*)

Quantum Neural Networks

解光军[1,2] 庄镇泉[1]

(中国科技大学电子科学与技术系 合肥 230026)[1](合肥工业大学应用物理系 合肥 230009)[2]

**Abstract** In recent years, the researches on combination of quantum theory and neural networks have attracted much attention. This paper reviews the development and status about this field. Some quantum neural networks (QNN) models are discussed, the applications and prospects are also given, which show that QNN have great competence and potential in the computational intelligence field.

**Keywords** Quantum neural networks, Quantum computing, Quantum learning, Quantum computational intelligence

## 1 引言

计算的实质是一个物理过程，要受到物理规律的支配。现代的计算技术建立在经典物理学基础之上，而量子计算(Quantum Computing)则是建立于量子理论的原理之上。目前，量子计算之所以获得广泛瞩目，是由于它能极大地提高计算的效率，具有沟通与大脑和意识关系的潜力，以及当计算机技术发展到纳米量级时必须考虑到量子效应等原因。

量子计算过程具有随机性和不确定性，而神经网络技术是模仿人脑的工作机理，其工作过程本身也具有复杂的非线性动力学特征，显而易见，二者之间存在着许多相似之处，那么我们自然就会设想：量子理论与神经网络技术能否结合起来以产生一种新的计算范式？理论上讲它应该更加智能化，具有更有效的学习和泛化能力。

1989 年英国 Oxford 大学的 Penrose 教授研究了量子理论与人脑意识的关系问题，并指出，解决量子测量问题是最终解决意识问题的先决条件；1994 年美国 Aizona 大学的 Hameroff 教授则认为，在神经元内骨骼细胞的微管(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；1996 年 Perus 指出，量子波函数的坍缩(Collapse)十分类似于人脑记忆中的神经模式重构现象；而 Narayanan 则认为以上观点对于认知科学的基础以及意识/人脑在计算意义上的解释都提出了深刻的质疑。他们这些富有创见性的研究和讨论为量子理论与神经网络的结合提供了有益的支持。

事实上，确实已有少数先行者尝试着将量子理论引入神经网络领域，并提出了诸如量子联想、并行学习、经验分析等新概念，开创了量子神经网络(Quantum Neural Networks)的新的交叉学科，但是到目前为止，这些工作还处在理论研究水平，且相关报道较少。我们认为，就量子理论与神经网络技术结合的形式而言，大致有两种：一种是在神经网络的结构或训练过程中引入量子理论，从而提高神经网络的学习和推广能力；另一种则是直接借用量子理论中的某些原理或概念，来指导设计神经网络拓扑结构或训练算法。

本文首先介绍量子理论中的一些原理和基本概念，接着讨论量子计算和量子学习问题，然后结合当前的研究动态分析几种量子神经网络模型及其应用，最后给出结论。

## 2 量子理论的基本原理及概念

量子理论的核心在于揭示原子级、亚原子级微观粒子(如电子、光子等)运动规律，不过，量子理论中的一些基本原理和概念具有非直观的(Counterintuitive)特性，与经典物理理论存在着很大的区别。

### 2.1 不确定性原理、双缝实验及不同的量子观点

Heisenberg 不确定性原理认为，一个粒子的位置和动量不可能同时被测量，这是因为在测量过程中外界必定对该粒子注入能量，比如说由测量激光发射出的光子打在粒子上，光子的能量会被粒子吸收，并将其

*)中国科学大技量子通讯与量子计算开放研究实验室资助项目，**解光军** 讲师，博士生，主要研究方向：量子信息处理，计算智能方法，**庄镇泉** 教授，博导，主要研究方向：神经网络、智能信息处理等。

送入其它的运行轨道，这样，倘若测量时能够确定该粒子的位置，那么此时粒子的动量已被测量环境所改变。

另外，在量子力学中有一个著名的 Thomas Young 双缝实验。由光子枪发射出光子，通过屏幕上的两个窄缝打在接收屏上。如果控制两个窄缝的开、关状况，就会得到一个奇特的现象：当只有一个缝开启时，光子会穿过它，在接受屏上得到一个分离的概率分布模式（靠近窄缝一侧概率大）；但是当两个缝同时开启时，接受屏上所得到的并非两个模式的简单叠加，而是更为复杂的模式，这个现象根本无法用经典物理理论来解释。

对上述奇异现象的不同解释就产生了不同的量子观点，主要的有："波粒二象性"观点（Born），认为量子同时具有波动性和粒子性的特征；"多宇宙"观点（Everett），认为量子叠加态实际上是宇宙的叠加，一个宇宙里有一个可观测的值；"本体论"观点（Bohm），认为量子实际上是粒子，它的波动行为可由量子势引起的额外作用力来解释。

### 2.2 线性叠加和态叠加原理

在量子力学中使用 Hilbert 空间中的波函数来描述微观粒子的运动状态，波函数的形式为 $\Psi(\vec{r},t)=Ae^{i(\vec{k}\cdot\vec{r}-\omega t)}=Ae^{\frac{i}{\hbar}(\vec{p}\cdot\vec{r}-Et)}$，显然，如此的波函数不再是经典意义上的波函数，它体现了量子的波粒二象性。线性叠加（Liner Superposition）的概念与矢量的线性组合有关，由于量子态可用波函数 $|\Psi\rangle$ 来描述，而 Hilbert 空间是以一组 $|\varphi_i\rangle$ 为基态，那么一个量子态就可以表示为 $|\Psi\rangle=\sum_i c_i|\varphi_i\rangle$，也就是说量子态 $|\Psi\rangle$ 是所有基态 $|\varphi_i\rangle$ 的一个线性叠加，其中系数 $c_i$ 为复数。这样，从某种意义上讲该量子态同时存在于所有基态之中。同样，由于量子系统的完备性，那么由基态组合所得到的任一量子态（满足归一化条件）也是 Hilbert 空间中的一个矢量，这一性质被称为态叠加原理。

### 2.3 相干、去相干及测量原理

相干和去相干是与线性叠加的概念紧密相关的。如果一个量子系统处于其基态的线性叠加之中，那么就称此量子系统是相干的；但是当一个相干的系统以某种方式与它所处的环境发生相互作用（即测量）时，线性叠加就会被破坏，由此所引起的相干的损失就称为去相干或坍缩，它由波函数 $\Psi$ 控制，系数 $c_i$ 称为概率振幅，由 $|c_i|^2$ 给出量子态 $|\Psi\rangle$ 坍缩到态 $|\varphi_i\rangle$ 的概率。因为波函数 $\Psi$ 描述了一个真实的物理系统，它必定完全坍缩到一个基态，所以由振幅 $c_i$ 决定的概率加起来的和一定是等于 1，这个约束条件可以表示为 $\sum_i |c_i|^2=1$。例如，对于最简单的自旋-1/2 系统，它是一个两态系统，其基态可描述为 $|\uparrow\rangle$（自旋向上）和 $|\downarrow\rangle$（自旋向下），在此系统中，波函数 $\Psi$ 是有关两值的一个分布，而相干态 $|\Psi\rangle$ 是 $|\uparrow\rangle$ 和 $|\downarrow\rangle$ 的线性叠加，比如说可能是 $|\Psi\rangle=\frac{2}{\sqrt{5}}|\uparrow\rangle+\frac{1}{\sqrt{5}}|\downarrow\rangle$，只要系统保持其量子相干性，它就不能说是处在自旋向上状态或是自旋向下状态，应该是同时处在两种状态之中，但是当此系统发生去相干时，一个可能的结果是以 $(2/\sqrt{5})^2=0.8$ 的概率坍缩到 $|\uparrow\rangle$ 态。

上述的两态自旋-1/2 系统可作为量子计算的基本单元，定义为量子位（Qubit），两个态可重新命名为 $|0\rangle$ 和 $|1\rangle$。

### 2.4 算子（Operator）

在 Hilbert 空间中常使用算子来描述一个波函数变化到另一个波函数，它们是用字符加上标来表示的（如 $\hat{A}$），代表作用于矢量的矩阵。使用了算子以后，便可以将本征方程写作 $\hat{A}|\varphi_i\rangle=a_i|\varphi_i\rangle$，其中 $a_i$ 为本征值，方程的解 $|\varphi_i\rangle$ 为本征态，可用来构成一个 Hilbert 空间的基态。

在量子公式中，所有的性质皆由算子表示，其本征态是相关性质 Hilbert 空间的基态，其本征值是该性质所允许的量子值。另外，量子力学中的算子必须是线性的，并且描述一个态时间演变的算子必定是幺正的（Unitary），即 $\hat{A}^{+}\hat{A}=\hat{A}\hat{A}^{+}=\hat{I}$，其中 $\hat{I}$ 为单位算子，$\hat{A}^{+}$ 为算子 $\hat{A}$ 的复共轭变换。

### 2.5 缠绕态（Entangled State）及量子关联现象

这是量子理论中最能体现其非直观特性的现象。所谓缠绕态是指发生相互作用的两个子系统中所存在的一些态，它们不能表示为两个子系统态的张量积，而是表达为子系统中态的某种缠绕形式，例如有两个双态量子系统 $A$ 和 $B$，其中 $A$ 的两个状态记为 $|\alpha_1\rangle$ 和 $|\alpha_2\rangle$，$B$ 的状态记为 $|\beta_1\rangle$ 和 $|\beta_2\rangle$，一般地 $A$ 和 $B$ 都处于叠加态，即：

$$|\varphi_1\rangle=c_1|\alpha_1\rangle+c_2|\alpha_2\rangle$$

$$|\varphi_2\rangle=c'_1|\beta_1\rangle+c'_2|\beta_2\rangle$$

当两个子系统相互独立时，由它们组成的大系统的态是其态的张量积，即 $|\varphi\rangle=|\varphi_1\rangle|\varphi_2\rangle$；但是若两个子系统发生作用时，系统的自由度将受到一些限制，便存在一些态，如 $|\varphi\rangle=c_1|\alpha_1\rangle|\beta_1\rangle+c_2|\alpha_2\rangle|\beta_2\rangle$ 等，它们并不能表示为两个子系统态的张量积，这样的态就称为缠绕态。

假设一个量子体系处在 $|\varphi\rangle=c_1|\alpha_1\rangle|\beta_1\rangle+c_2|\alpha_2\rangle|\beta_2\rangle$ 的缠绕态中，如果对它测量，所得到的结果是 $A$ 处在 $|\alpha_1\rangle$ 态，那么 $B$ 无可避免地应处在 $|\beta_1\rangle$ 态之中，这个过程不需要任何时间，也就是说，$B$ 的演变

与对 $A$ 的测量同时发生，信息的传递是瞬时的，此过程称为隐形传态（Teleportation）或量子关联现象。

## 3 量子计算与量子学习

### 3.1 量子计算原理及其特点

量子计算是建立于量子理论之上的计算理论，其中态叠加原理是量子态之所以能用来进行并行运算的关键所在，也是量子计算机与经典计算机在计算原理上的根本区别。近年来在计算机和认知科学领域中对于量子计算的兴趣越来越浓厚，其直接原因是Deutsch和Shor指出被计算机学家认为是NP-难或NP-完全问题可以用量子计算机来解决，因为量子并行计算的特征，对于某类问题（如大数的因式分解），量子计算机存在多项式级算法，而经典计算机却需要指数级算法。在基于量子力学理论的新式计算机中，量子力学的基态就是计算机的逻辑状态，所谓计算就是这些态对其他态的幺正映射，算符用来定义其中的相互作用。这种量子计算的特点是通过态的叠加允许若干问题同时求解。

Deutsch认为任何一个物理系统皆为一个计算机，它应具备三个可观察的特征：即输入、输出和中断。在经典计算机系统中，输入的分离本征值对应着系统的期望输入，输出亦然，系统计算函数 $f$，一旦系统的输入值设定为输入的本征值 $x$，系统将中断，而处在与 $f(x)$ 相应输出的本征值状态中。另外，对于任何一个给定的计算系统，我们可以构建一个多维输入空间，每个输入作为一个分离的维，对输出同样进行处理，于是计算机便可看作是从其中之一个空间到另一个空间的转换矩阵。量子计算机与经典计算机的不同在于：经典计算机计算的是稀疏矩阵，在每行每列上具有精确的赋值（1或0），而量子计算机并无此限制，量子计算的终态是经典输出的一个线性叠加。

### 3.2 量子学习

Hopfield指出，任何物理的动力学系统都可以看作是一个神经网络结构，其中系统中的某些固定点作为"记忆"，它可以通过内容编址的方式联想回忆起来。据此思路，量子的神经网络学习算法不仅可以寻求更有力的计算体系，而且有助于我们深入理解大脑、意识以及量子理论本身。

是否存在大脑对量子效应敏感的证据呢？曾有这样一个实验表明，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，但对人类而言该现象似乎被噪声过滤所抑制，不过Penrose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维持。

Mead曾指出，时至今日计算机在硬件上的革新也只是集中在尺寸上：尺寸越小性能越好、尺寸越小速度越快。同时他又指出，虽然人脑在尺寸和速度方面都不及当前的计算机，但其性能却远远优越，所以他认为与其寻求又小又快的硬件，不如去探讨一下人脑的构造，而神经网络算法正是这种新型计算方式研究的一部分。比较一下Deutsch的量子计算机设想，他仅仅考虑了尺寸和速度，并没有特别讨论任何非传统的算法，量子计算的优越只是在于并行计算，而他提到的并行计算也只是经典意义上的概念。一个在量子空间和时间尺度上的计算机，其性能肯定优于当代计算机，但Mead的疑问依然存在。所以，为什么不同时寻求既在尺度上又在计算形式上的共同改善呢？这正是我们讨论量子学习问题的原因。

## 4. 量子神经网络模型

现有的人工神经网络理论与人脑相比有许多缺陷，特别是：(1)传统意义上的学习是一种序列式的处理模式，在信息量大的情况下处理速度过慢，不符合人脑实时反应、大容量作业的特征；(2)人脑将记忆和回忆信息作为学习的一部分，但对神经网络而言，训练的样本并不被存储，只是用来改变连接权值，之后立即被遗忘；(3)神经网络需要反复训练，而人脑具有一次学习的能力；(4)神经网络在接受新的信息时会发生灾变性失忆（Catastrophic Forgetting）现象等。因此，为进一步发展神经网络理论必须引入新的理论、新的思想。量子理论与神经网络的结合便是一个有益的尝试。

下面我们结合最近的研究成果，讨论几种量子神经网络模型。

### 4.1 量子神经元模型（Quantum Neuron Model）

我们知道，在经典人工神经网络模型中最简单的是感知机（Perceptron），它以 $n$ 个二进制数集 $\{i_j\}$ 作为输入值，还定义了连接权矢量 $W=(\omega_1,\omega_2,\cdots,\omega_n)^T$，阈值 $\theta$ 和激活函数 $f$：

$$f=\begin{cases}1，若\sum_{j=1}^{n}\omega_j i_j>\theta \\ -1，其它\end{cases} \tag{1}$$

该模型无法解决线性不可分问题，但鉴于其形式和性质的简单，首先就研究它的量子对照物，即量子神经元模型。在量子神经元模型中，同样取 $\{i_j\}$ 作为输入，单个连接权矢被一个波函数 $\Psi(\omega,t)$ 所取代，其余部分皆与经典模型类似。这里的波函数处于Hilbert空间，其基态为经典模型的权矢，$\Psi(\omega,t)$ 代表了权矢空间中所有可能权矢的概率幅度（广义上为复数），在任意时刻t满足归一化条件，即

$$\int_{-\infty}^{+\infty}|\Psi|^2\mathrm{d}\omega = 1 \tag{2}$$

于是，感知机的权矢就被许多权矢的量子叠加所代替，不过，一旦当它与环境发生作用时将立即坍缩到其中之一的经典权矢之上，且概率为$|\Psi|^2$。

现在考察一个单输入单输出的双极性函数，用来对输入值求反(NOT)。为方便起见，将权值限定在$-\pi\leqslant\omega_i\leqslant\pi$区间之内，这样为了使一个神经元能够学习该函数，就必须找到$\Psi(\omega,t)$和$\theta$，再假定$\Psi$时间恒定，只与$\omega$有关，则波函数$\Psi(\omega,t)$蜕变为一维矢量。考虑到权值的限制，这时寻找$\Psi$就等价于量子力学中求解一维刚性箱体问题，其解的形式为：

$$\Psi(\omega_0)=A\sin(\frac{n\pi}{\alpha}\omega_0) \tag{3}$$

其中$A$为归一化常数，可由(2)式求得；$n=1,2,3,\cdots$；$\omega_0$为$W$的单个元素；$\alpha$是箱体的宽度，通常取$2\pi$，也可以更小。

为了更好理解该量子神经元模型，应注意$\Psi$是一给定权矢的概率幅度，在一个量子神经元中其权矢存在于所有可能的经典权矢的相干叠加之中，当叠加权矢与环境作用时(例如遭遇到一个实际输入)，它必定会去相干到其中之一的基态上，其发生概率为$|\Psi|^2$。注意，在此简单情况中，不论任何权矢在去相干时被选择，它必定将导致正确的输出。

我们再来考察更复杂一些的二值输入异或(XOR)问题，它是线性不可分的。与上面的推导思想相似，可得二维权空间问题的解等价于二维刚性箱体问题，其解的形式为：

$$\Psi(\omega_0,\omega_1)=A\sin(\frac{n_{\omega_0}\pi}{\alpha}\omega_0)\sin(\frac{n_{\omega_1}\pi}{\alpha}\omega_1) \tag{4}$$

其中变量和常量的含义与(3)式相同，注意这里有两个不同的$n$，每一个对应于一个不同的权值。显然，由于量子神经元保持了权的相干叠加，因此它能够解决线性不可分问题。

当然，上述例子和相应的解都比较简单，但即使在最普遍的情况下，其解常等价于一个$n$维刚性箱体问题，这样$\Psi$就有上面的通用形式，但倘若换成其它物理模型，如非刚性箱体、简谐振荡器、氢原子模型等，解的形式就要发生相应的变化。另外，在方程中尚未考虑时间变量，还有一个重要问题是训练神经元，它必须通过改变波函数来实现，而波函数又是由 Schrodinger 方程所决定，即

$$\nabla^2\Psi=\frac{2m}{h^2}[U-E]\Psi \tag{5}$$

这里与时间无关，其中可改变的项只有势函数$U$，所以$\Psi$可随改变$U$而改变。

4.2 **量子衍生神经网络(Quantum-inspired Neural Networks)**

在神经网络的设计、开发和实现过程中这种方法使用了量子理论的“多宇宙”观点。

我们知道，在训练一个传统的神经网络时需要反复学习模式集，直到网络对每个模式达到合适的输出为止，不过这种模式的重复出现(有时是成百上千次的)并无人类学习的生物基础。由量子“多宇宙”概念衍生出的神经网络方法则认为，训练集中的模式犹如一个粒子，在不同的宇宙中被大量分离的神经网络(可能是同种类型也可能不是)所处理，就类似于一个光子同时穿过许多窄缝，而训练集中的每个模式仅在自身宇宙中被处理。这样，实际上就需要许多神经网络同时进行训练，网络个数应等于训练模式数，每个网络与其相关的训练模式存在于一个分离的宇宙，一旦每个网络在其宇宙中训练成功，就计算这些网络的量子叠加态，并由此推导出一个量子衍生神经网络(QuINN)，把它推广到所有输入模式中，而结果的叠加权矢量称为量子衍生波函数(QuIWF)，它坍缩到实际输入模式，具体坍缩依赖于输入模式和坍缩的方式。

表 1

| | 权值 | | | | 训练输入 | | | |
|---|---|---|---|---|---|---|---|---|
| Net | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 1 | a1 | b1 | c1 | d1 | 1 | 1 | 1 | 1 |
| 2 | a2 | b2 | c2 | d2 | 1 | 0 | 1 | 0 |
| 3 | a3 | b3 | c3 | d3 | 1 | 0 | 0 | 1 |

我们可以通过一个简单的算例来测试该方法的可行性，同时使用传统的神经网络来作为参照。假设有3个网络，每个网络有4个输入、1个输出及相应的4个连接权，分别将所有的连接权标注为a1～d1，a2～d2，a3～d3。表1给出网络已训练过的输入模式和相应权值。训练之后，单个网络的权值结合起来得到最终QuINN的叠加权值(即QuIWF)，如QuINN中由输入单元1到输出单元的连接(即link1)具有a1、a2、a3叠加组合推导出的量子连接权，在这里使用一个简单的概率矩阵来推导QuIWF。于是，每个输入位(1/0)作为所有模式的一个亚粒子，QuINN连接作为物体，单个连接权作为态，这样当一个亚粒子与一个物体发生作用时它将进入一个叠加态之中，这取决于物体和亚粒子，即亚粒子1与QuINN连接1作用时它将进入态a1和a2的叠加之中，因此当1输入QuINN单元1时，与link1的QuIWF有关的概率矩阵对a1、a2取正值，对a3取负值，所谓正值表示接受粒子进入该态，负值则排斥。这里定义泛化规则如下：当一个输入模式的所有或大多数位被某个宇宙的态所接受时，它将在此宇宙中被处理。例如，见表1，当输入模式为1010时，它被宇宙2(即Net2)完全接受，所以正确的输出只发生在该宇宙中；而当输入为1001时，宇宙3比其它两

个接受的态更多，因此测试模式就选取此宇宙来进行处理。

当然，上面的叠加方法是由"多宇宙"的 Fermion 观点衍生出来的，它认为单个粒子的波函数相互之间不发生重叠，当坍缩发生时某个存在的宇宙(即训练过的网络)就是结果；而 Boson 观点则认为波函数可以相互交叠，相应的宇宙相互融合，并且可以坍缩到一个在训练过程中并不独立存在的网络之中。

实验的初步结果表明，QuINN 在训练时权值改变的次数比传统神经网络的减少近 50%，而泛化能力相同，因此利用"多宇宙"观点来设计和训练神经网络，在某些场合会提供更快的训练速度，同时又无泛化能力的损失。此外，量子衍生神经网络具有消除灾变性失忆的潜力，这是因为一个网络仅训练一个模式，因此在训练过程中模式之间不发生相互干扰。与此同时，QuINN 还为线性不可分问题提供了一个单层网络的解决方案，实验表明，传统的单层神经网络不能学习某些训练集，而量子网络则可以。

### 4.3 量子并行 SOM 模型

量子并行 SOM 模型是针对并行计算环境，通过对传统的 Kohonen-SOM 模型进行改进后得到的，此模型用来模拟人脑的一次学习和记忆功能，其结构如图 1 所示，在训练过程中输入、输出层神经元之间的每个连接分别作为一个独立的处理器(共有 $M\times P$ 个)，这样每个矩阵的所有元素便可以同时进行计算，极大地提高了权值更新的能力。

与传统 SOM 算法比较该模型主要有以下改进：(1)输入和输出层中神经元以及它们之间的连接的数目等于输入信号所有元素数($M$)与数据可能的分类模板数($P$)的乘积；(2)在一对输入和输出神经元之间只有一个连接，每个连接可看作是一个独立工作的处理器；(3)权值更新被视为一定次序的矩阵相乘，它适合于并行处理。

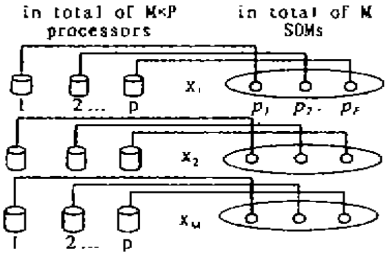



图 1　量子并行 SOM 模型结构

假设有一个信号序列 $x^t=\{x(1),x(2),\cdots,x(M)\}$，为了一次输入全部数据，并行 SOM 的前触突神经元数至少有 $M$ 个，如果输入数据可能分类成 $P$ 个模板，那么神经网络结构就变成二维阵列，分别具有 $M\times P$ 个前触突和后触突神经元，其中前触突表示为 $X=(x_1,x_2,\cdots,x_p)$，且 $x_1=x_2=\cdots=x_p=x$，后触突神经元用 $Y$ 表示，其中元素为 $y(i,k)$，$i=1,2,\cdots,M$，$k=1,2,\cdots,P$，每一个输入层神经元 $x_k(i)$ 仅与输出层神经元 $y(i,k)$ 相连接，权值为 $\omega^t(i,k)\in W^t$，这里 $t$ 是当前操作时间，且 $t=0,1,2,\cdots,T$。以下是并行 SOM 算法竞争权值更新的主要步骤：

(1)一次学习：网络允许对所有的输入数据只学习一次，这样全部的输入信号 $X$ 在这一步被读入，其中 $X=(x,x,\cdots,x)_{M\times P}$，$x^t=\{x(1),x(2),\cdots,x(M)\}$；

(2)权值初始化：对初始权值矢量 $W^0$ 选取随机值，这里唯一的限制是 $W^0(i,k)=W^0(i+1,k)$ 且 $W^0(i,k)\neq W^0(i,k+1)$，其中 $i=1,2,\cdots,M$，$k=1,2,\cdots,P$，并且希望使初始权值的模较小；

(3)相似性匹配：重复第(3)、(4)、(5)、(6)步 $T$ 次，令 $W^t=V$，其中 $V$ 是权值矩阵与权值转换矩阵的乘积。首先有 $W^t=W^0$，然后计算所有的欧氏距离 $D^t=\|X-W^t\|$，在第 $t$ 次利用最小距离标准寻找最佳匹配(即获胜者)，在并行环境可以利用 Grover 算法来查询，即 $d^t(i,k_{\min})=\min\{d^t(i,1)\cdots,d^t(i,p)\}$ 它是 $D^t$ 的第 $i$ 行的最小欧氏距离；

(4)同步更新：神经元之间的每个连接作为一个处理器，利用下面的公式来更新权值矩阵，即

$$\begin{cases}\omega^{t+1}(i,k)=\omega^t(i,k)+\eta(t)[x(i,k)-\omega^t(i,k)], & \text{if } k\in A_{k\min}\\ \omega^{t+1}(i,k)=\omega^t(i,k) & otherwise\end{cases}$$

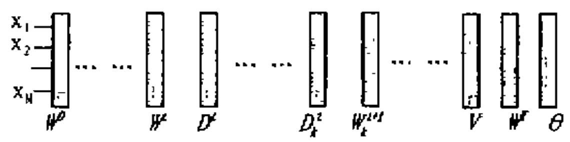
其中 $\eta(t)$ 是学习率，$A_{k\min}$ 是以获胜神经元 $k_{\min}$ 为中心的邻近函数，它们在学习过程中同时进行动态调整以达到最优结果；

(5)停止条件：依照求解问题的需要设定一个最小误差矩阵 $\varepsilon$，若 $W^{t+1}-W^t<\varepsilon$ 条件满足则转向第(4)步；

(6)矩阵重组：用权值转换矩阵 $Q$ 与权值矩阵 $W^{t+1}$ 相乘得到一个新的矩阵 $V$，这里 $QQ^{-1}=I$；

(7)登记：保存权值矩阵 $W^{t+1}$ 并停止。

该算法可以利用量子计算方法实现，即 QuSOM 算法，流程如下：




如果利用量子门实现上述算法，文献证明所需 qubit 数只有 $\log_2 M\times P$ 个，这样 QuSOM 算法在保证同样的收敛性的条件下，可以大大降低计算的复杂度。

### 4.4 量子联想记忆

传统的人工神经网络(如 Hopfield 网络)允许关联模式响应,但是其主要缺点是存储的容量受到限制,例如要存储一个长度为 n 的模式需要 n 个神经元的网络,可存储的模式数为 $m\leqslant kn$,一般地 $0.15\leqslant k\leqslant 0.5$。利用量子联想记忆可以极大地扩展记忆的容量。具体地,在量子联想记忆中使用了两个关键的量子算法,一个用于模式存储,另一个则用于模式响应。

(1)存储模式:设计一个量子算法用来构造一个 n Qubit 的相干态,它表示一个有 m 个模式的集合。该算法仅在若干 Qubit 上进行了多项式级别的基本操作,操作过程中的关键算子为:

$$\hat{S}^{p}=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & \sqrt{\frac{p-1}{p}} & \frac{-1}{\sqrt{p}}\\ 0 & 0 & \frac{1}{\sqrt{p}} & \sqrt{\frac{p-1}{p}}\end{bmatrix}$$

其中,$m\leqslant p\leqslant 1$。这实际上是条件变换的一组算子,其中与被存储的每个模式有关的 $\hat{S}^{p}$ 算子各不相同。

(2)完成模式:采用 Grover 量子查询算法,其基本思想是将期望基态的相反转,然后反转所有平均幅度的基态,反复操作直到期望基态的幅度趋近于 1,同时使得返回原有量的期望态的幅度减少。该过程的周期为 $\frac{\pi}{4}\sqrt{2^{n}}$,在 $O(\sqrt{2^{n}})$ 次查询之后,系统便可以在期望的态中被测定,且幅度接近于 1。我们定义 $\hat{I}_{\varphi}$ 为恒等矩阵,除了 $\iota_{\varphi\varphi}=-1$ 之外,它用来反转基态 $|\varphi>$ 的相;定义 $\hat{W}=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$,称为 Walsh 变换;还定义 $\hat{G}=-\hat{W}\hat{I}_{0}\hat{W}$,用来反转平均幅度的态。

(3)量子联想:将上述两个算法结合起来便构成量子联想记忆算法,定义 $\hat{P}$ 算子完成(1)中所描述的记忆模式算法,那么记忆一组模式就可表示为 $|\Psi>=\hat{P}|0>$,其中 $|\Psi>$ 是基态的量子叠加态,而每个基态代表一个模式。现在我们假定已知一个模式的 $n-k$ 位,希望回忆起模式的全部,利用改进的 Grover 算法来回忆模式,即 $|\Psi>=\hat{G}\hat{I}_{p}\hat{G}\hat{I}_{\tau}|\Psi>$,随后 $|\Psi>=\hat{G}\hat{I}_{\tau}|\Psi>$,重复 $O(\sqrt{2^{n}})$ 次,用 $\hat{I}_{p}$ 反转所有代表已存储模式的态的相,$\hat{I}_{\tau}$ 反转那些与已知 $n-k$ 位匹配的态的相。于是,拥有 $2n+1$ Qubit 的量子联想记忆算法可以在 $O(mn)$ 时间内存储 $2^{n}$ 个模式,并且在 $O(\sqrt{2^{n}})$ 时间内回忆起一个模式。

**讨论** 量子理论与人工神经网络理论之间有着本质的联系,因此可以尝试把它们结合起来,以产生新的研究课题即量子神经网络,它将有助于我们更深入地理解量子理论和神经网络理论本身。本文首先讨论了量子计算和量子学习问题,然后结合目前该领域的最新研究成果,详细探讨了其中有代表性的几种网络模型及其学习算法。

通过上述的量子神经网络我们可以发现,它们在结构和学习方面与传统神经网络相比有许多不同,而性能往往优于传统网络,并且大多数模型与量子并行计算有关,由此也可以得出一些构造量子网络的经验结论:

(1)将待求解的问题转换分解为一些小的子问题;

(2)每一个子问题可以相对独立地并行求解;

(3)问题最终的解决必须有一个类似于量子测量的过程,在此过程中有一个坍缩的概率幅度,而这个概率对网络的泛化性能至关重要。

*致谢:感谢美国 Penn 州立大学 Dan Ventura 博士和 Wichilta 州立大学 Elizabeth C Behrman 博士提供的相关资料。*

## 参考文献


1 Penrose R. The Emperor's New Mind. Oxford University Press,1989

2 Penrose R. Shadows of the Mind. Oxford University Press,1994

3 Deutsch D. Quantum Computaional Networks. Proceedings of the Royal Society,London A,1989,425:73～90

4 Shor P W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarthms on a Quantum Computer. SIAM Journal of Computing,1997,26(5):1484～1509

5 Kak S C. On Quantum Neural Computing. Information Sciences,1995,83:143～160

6 Menneer T,Narayanan A. Quantum-inspired Neural Networks:[Technical Report R329]. Department of Computer Science,University of Exeter,UK,1995

7 Chrisley R. Quantum Learning In,Pylkkanen P,Pylkko P,eds. New Directions in Cognitive Science. Proceedings of the International Symposium,Saariselda,1995,Lapland,Finland,77-89,Helsinki Finnish Association of Artificial Intelligence

8 Perus M. Neuro-Quantum Parallelism in Brain-Mind and Computers. Informatica,1996,20:173～183

9 Behrman E,et al. A Quantum Dot Neural Network. In:Proc. of the Workshop on Physics of Computation,1996,22～24

10 Ventua D. Artificial Associative Memory Using Quantum Processes. In:Proc. of the Int'Conf on Computational Intelligence and Neuroscience,1998,2:218～221

11 Weigang L. A Study of Parallel Self-Organizing Map. E-Print:http://xxx.lanl.gov/quant-ph/9808025,1998