

## 摘 要

本文将人工神经网络理论与量子理论相结合,研究一种全新的计算模式——量子神经计算,它具有良好的理论和应用前景,属于交叉学科的前沿课题,目前国际上对此领域的研究尚处在起步阶段。本文的研究将量子计算的理论和方法引入了传统的人工神经网络理论,这有助于提升神经网络信息处理的能力,也丰富了量子计算的研究领域,同时还有利于理解量子理论与人工神经网络理论各自在更深层次上的含义。

本文力图从信息处理的角度出发,进行有关量子神经计算的理论研究和量子神经网络模型及算法的设计,并结合一些仿真实验,探索其在模式分类、记忆等方面的应用。论文主要围绕着两个主题展开研究:一是神经计算与量子计算的结合,二是量子神经网络模型、算法的构造和检验。

总结全文,主要工作与贡献包括:

- (1) 对相关的量子理论和量子计算原理进行了系统介绍,综述了当前国际上对于量子神经计算的研究现状及水平;
- (2) 对比研究量子理论和神经网络理论,以及二者在数学形式和动力学特征上存在的许多相似之处,据此提出用量子系统来实现神经计算的可能性,并推导出基于量子系统的 Hebb 学习规则,为神经计算与量子计算的结合提供了一定的理论基础;
- (3) 从不同的量子观点、坍缩原理、路径积分以及量子实验等方面出发,提出神经计算量子推广的几种基本方法;
- (4) 利用量子理论中的基本原理和观点,在量子理论波动力学的基础之上建立了量子神经元模型,并针对不同的逻辑函数探讨了其中叠加权值的形式及其相关含义,同时利用量子理论中的多宇宙观点,以多 BP 网络的结构设计出多宇宙量子神经网络模型,应用于模式分类;
- (5) 借鉴量子理论或量子计算中的某些概念、实验、效应等来设计神经计算模型和算法,分别提出了一种能够进行数值计算的多狭缝-多栅板结构的量子神经网络模型,以及基于量子通用门组单元的神经网络模型,并设计了一些仿真实验来评价上述模型的计算性能;

- (6) 研究了量子神经网络的学习问题, 借用量子隐形传态的结构提出基于量子纠缠的学习模型, 用它来解决信息的智能化传递、有效信息的发掘等问题, 在量子算法的理论框架下, 结合经典 Hamming 网络的学习原理, 设计了一种用于竞争学习的量子算法。

本文的主要创新点有:

- (1) 利用量子理论中的多宇宙观点, 以多 BP 网络的结构设计出多宇宙量子神经网络模型, 它包含了量子态的叠加、纠缠和坍缩等量子过程, 在训练时借鉴 Grover 量子算法提出了全新的搜索策略, 并针对不同的求解问题设计了不同的坍缩算法, 解决了多宇宙的测量问题。分别对单层及两层量子神经网络作了大量的数值实验, 仿真结果显示, 该模型与经典 BP 网络相比, 在保证相近泛化能力的前提下大大提高了训练速度;
- (2) 在指出 Chrisley 理论模型缺陷的基础上, 对于量子双缝干涉实验中的参量赋予了全新的物理定义, 提出了多狭缝-多栅板结构的神经网络模型, 该模型与以往的模型相比, 可以方便地实现数值计算。利用改进的模拟退火算法实现了类似二值异或运算的模拟仿真, 同时给出相应的理论推导, 从两个侧面验证了结果的正确性, 为线性不可分问题提供了一个单层网络的解决方案;
- (3) 提出一种基于量子通用门组单元的神经网络模型, 它利用量子计算中的一位相移门和两位受控非门所构成的通用门组单元作为网络的激活函数, 利用改进的复值 BP 算法进行训练, 应用结果显示, 该网络有较好的计算性能, 此外, 该模型还具有易于物理实现的潜力;
- (4) 对照经典 Hamming 网络提出了一种用于竞争学习的量子算法, 该算法含有两个幺正算子, 分别用来实现模式的匹配和竞争学习, 可以有效地实现模式的分类或记忆, 此举拓宽了量子算法的应用领域。

本文得到国家自然科学基金项目(60171029)和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的共同资助。

## Research on Quantum Neural Computation and Its Models

### Abstract

A novel computational paradigm called quantum neural computation (QNC) is studied in this dissertation combined with artificial neural networks (ANN) theory and quantum theory. Although nowadays the research on this field is at the start of a race, the combination has great potential and good foreground, it is of important theory significance and practical significance. In this dissertation the theory and methods of quantum computation are introduced into classical artificial neural networks, the information processing ability of ANN will be prompted, at the same time, this research enriches the fields of quantum computation, and also contributes to understand the deep meanings of both quantum theory and ANN theory.

Based on the view of information processing, this dissertation tries to do theoretical research on quantum neural computation and to design several models and algorithms of quantum neural networks (QNN), and also to discuss QNN's applications for patterns recognition and patterns memory. In brief, there are two themes in this dissertation, one is the combination of neural computation and quantum computation, another is the conformation and inspection of quantum neural networks models and algorithms.

Overall the main content of this dissertation includes: (1) introduction of basic quantum theory and quantum computation, summary of research status and level on QNC in the world; (2) theoretical prove on the possibility of neural computation in the quantum system; (3) basic methods of the extension of neural computation to quantum domain; (4) discussion of quantum neuron model and the model of many universes QNN with its collapse algorithms; (5) the QNN model based on double-slit interference experiment and the QNN model based on universal quantum gates cell; (6) the learning model of quantum entangled neural network based on teleportation and the quantum competitive learning algorithm.

The main contributions of this dissertation are:

(1) The QNN model based on many universes interpretation is proposed, it contains quantum processes such as states superposition, entanglement and collapse, specially in

the model's training it uses the strategy inspired by Grover algorithm, with different collapse algorithms for different questions. A lot of simulating experiments are done for single-layer and two-layer neural networks, the results show the training rate of this model is faster than classical BP networks under the precondition of close generalizing ability.

(2) To overcome the shortcoming of Chrisley's model, this dissertation endows new physical definitions to those parameters in quantum double-slit interference experiment, a many-slit many-bar QNN model is presented, the excellence of the model is that it is easy to implement numerical computing. With the improved simulated annealing algorithm, this computational model realizes similar two-value XOR logic operation, at the same time such operation is also proved by corresponding physical deductions. These results indicate that the linearly non-separable problem can be solved with single-layer network.

(3) A QNN model based on universal quantum gates cell is presented in this dissertation, in this model a 1-qubit phase-shift gate and a 2-qubit controlled-NOT construct a universal cell, which acts as the active function of neural networks. With a improved complex-BP training algorithm, the QNN model is applied to solve two practical financial problems, the simulating results show this model is superior to classic BP and RBF networks at learning rate and precision. Furthermore, this QNN model has the potential to be implemented with quantum hardware.

(4) Compared with classical Hamming nets a quantum algorithm is designed for competitive learning, in this algorithm there are two main quantum operator, one for patterns matching, and another for competitive learning. The simulating result shows this quantum algorithm is available for patterns classification or patterns memory. This study develops the application fields of quantum algorithms.

This work is financially supported by the national natural science foundation (No. 60171029) and the cooperation item foundation provided by the Laboratory of Quantum Communication & Quantum Computation of USTC.

## 图表目录

图 2-1 量子隐形传态模型示意图	15
图 2-2 量子异或门(受控非门)符号	17
图 2-3 量子 Toffoli 门符号	18
图 3-1 五原子量子点系统	36
图 3-2 量子系统实现神经计算过程示意图	40
图 3-3 双缝干涉实验示意图	40
图 4-1 逻辑求反函数的量子权值分布	46
图 4-2 逻辑异或函数的量子权值分布	47
图 4-3 两个模式“01”和“11”分别以虚线和实线权值进行训练	51
图 4-4 测试输入时神经元 A 和 B 的输入值分别为“1”和“0”	51
图 4-5 两层网络的 16 种网络类型	64
图 5-1 光子双缝干涉实验示意图	67
图 5-2 Chrisley 非叠加态量子神经计算模型	68
图 5-3 基于双缝干涉实验的 QNN 模型	69
图 5-4 QNN 利用 MSA 算法学习二值逻辑函数的模拟结果	71
图 5-5 量子神经元示意图	74
图 5-6 三层前馈量子神经网络结构示意图	74
图 5-7 非平稳时间序列的预测结果比较图	77
图 6-1 量子隐形传态单元 II 模型	80
图 6-2 含有三个量子隐形传态单元的量子纠缠神经网络模型	81
图 6-3 Hamming 神经网络的拓扑结构	82
图 6-4 量子学习算法实现过程示意图	83

	表 1-1 几种典型量子神经网络的设计要点-----	3
	表 2-1 Bob 的解码变换矩阵-----	16
	表 3-1 神经计算概念及其量子对照物-----	38
C	表 4-1 单层网络实验一的测试环境一-----	57
	表 4-2 单层网络实验一的测试环境一下的实验结果-----	57
	表 4-3 单层网络实验一的测试环境二-----	57
	表 4-4 单层网络实验一的测试环境二下的实验结果-----	57
	表 4-5 单层网络实验一的测试环境三-----	58
	表 4-6 单层网络实验一的测试环境三下的实验结果-----	58
	表 4-7 单层网络实验一的测试环境四-----	58
C	表 4-8 单层网络实验一的测试环境四下的实验结果-----	58
	表 4-9 单层网络实验一的测试环境五-----	58
	表 4-10 单层网络实验一的测试环境五下的实验结果-----	58
	表 4-11 单层网络实验一的测试环境六-----	58
	表 4-12 单层网络实验一的测试环境六下的实验结果-----	58
C	表 4-13 单层网络实验一的平均测试环境-----	59
	表 4-14 单层网络实验一的平均实验结果 (激活函数 tansig) -----	59
	表 4-15 单层网络实验一的平均实验结果 (激活函数 purelin) -----	59
	表 4-16 单层网络实验二的训练环境一-----	59
	表 4-17 单层网络实验二的训练环境一下的实验结果-----	59
	表 4-18 单层网络实验二的训练环境二-----	59
C	表 4-19 单层网络实验二的训练环境二下的实验结果-----	60
	表 4-20 单层网络实验二的训练环境三-----	60
	表 4-21 单层网络实验二的训练环境三下的实验结果-----	60

表 4-22 单层网络实验二的训练环境四	60
表 4-23 单层网络实验二的训练环境四下的实验结果	60
表 4-24 单层网络实验二的训练环境五	60
表 4-25 单层网络实验二的训练环境五下的实验结果	60
表 4-26 单层网络实验二的训练环境六	60
表 4-27 单层网络实验二的训练环境六下的实验结果	61
表 4-28 单层网络实验二的平均实验结果	61
表 4-29 单层网络实验二平均实验环境	61
表 4-30 单层网络实验二在坍塌方式二下的平均实验结果	61
表 4-31 单层网络实验三平均实验环境	62
表 4-32 单层网络实验三的平均实验结果	62
表 4-33 单层网络实验四平均实验环境	62
表 4-34 单层网络实验四的平均实验结果	62
表 4-35 两层量子神经网络实验一训练环境	64
表 4-36 两层量子神经网络实验一实验结果 (使用坍塌方式一)	65
表 4-37 两层量子神经网络实验一实验结果 (使用坍塌方式二)	65
表 5-1 类似 XOR 逻辑真值表	72
表 5-2 例 1 的训练样本集	76
表 5-3 例 1 的训练结果	76
表 5-4 例 2 实验数据结果比较	77

# 第一章 绪 论

## 1.1 论文的选题背景和意义

人工神经网络 (Artificial Neural Networks) 是对人脑工作机理的简单模仿, 它建立于简化的神经元模型和学习规则的基础之上, 却由此产生了许多计算上的优势, 并且拥有了许多非常成功的应用<sup>[1-3]</sup>。不过自 1982 年 Hopfield 神经网络提出以来, 人工神经网络领域除了大量的应用研究之外, 从本质上讲一直没有较大的理论突破, 随着应用的深入推广和实际问题的日益涌现, 神经计算 (Neural Computation) 的局限与不足逐渐显现出来, 这些本质上的缺陷使得人们对传统神经网络理论的进一步发展提出了强烈的要求<sup>[4]</sup>。根据学科发展的一般规律我们可以推知, 神经网络技术与新的理论相结合, 形成新的交叉学科就有可能产生质的飞跃。于是, 近年来便出现了神经网络与其它理论相结合的实践<sup>[5-8]</sup>, 其中神经网络与量子理论的结合是一个崭新的、极富前景和创意的尝试。

量子理论自诞生 100 多年以来已取得了巨大的成功, 但之前它与计算机科学、信息理论一直是作为不同的学科并行发展, 几乎很少有人注意到它们之间有联系和交叉的可能性, 这一局面一直维持到 1982 年, Beniof 和 Feynman 发现了将量子力学系统用于推理计算的可能<sup>[9, 10]</sup>, 接着 1985 年 Deutsch 提出第一个量子计算模型<sup>[11]</sup>, 由此, 量子计算 (Quantum Computation) 迅速成为一门引人入胜的新学科, 特别是近年来由 Shor (1994 年) 提出的大数质因式分解量子算法<sup>[12]</sup>和 Grover (1996 年) 提出的无序数据库量子查询算法<sup>[13]</sup>, 他们对传统算法实施了近乎指数级的改进, 更是极大地推动了该领域的发展。随后, 便出现了以量子计算机、量子通讯以及量子编码为代表的量子信息科学, 使人们越来越清楚地认识到: 信息理论、计算机科学和量子物理学之间存在着密切而又深刻的内在联系<sup>[14]</sup>。

量子计算与传统意义上的计算有着质的不同, 它的特点主要体现在量子态的叠加 (Superposition)、干涉 (Interference) 和纠缠 (Entanglement) 上, 许多计算上的优势诸如量子并行性 (Quantum Parallelism)、量子非局域性 (Quantum Non-locality)、量子隐形传态 (Quantum Teleportation) 等皆是由此而产生<sup>[15-18]</sup>。量子计算所表现出的惊人潜力和异乎寻常的特征源于对传统计算进行的量子改造, 而神经计算是对生物行为以信息处理方式的模拟, 那么, 若将神经计算推广到量子领域就有可能利用到量子计算的巨大威力, 同时, 神经网络的动力学特征

与量子系统之间存在着许多相似之处，由此可以推断：神经网络与量子理论相结合将会构建出一种新的量子神经计算（Quantum Neural Computation）模式。从理论上分析，这种结合具有很大的潜力和良好的前景，它将有助于理解人脑和意识的本质，有助于求解 NP-hard 问题，同时也有助于理解神经网络理论和量子理论本身，更重要的是它开辟了一种新的思路，能够创造出一种全新的神经计算模型——量子神经网络（Quantum Neural Networks）。

因此，在人们开始寻找新的计算模式的今天，量子神经计算无疑是值得关注的领域，其中，量子理论与神经网络的结合有着内在的合理性和很强的创新性。本文尝试对这一全新的计算模式进行探讨，通过理论分析和模型、算法研究，讨论该模式的计算性能，以及它在模式分类、记忆等方面的应用，从而深入理解量子理论与神经网络理论的本质联系和各自在更深层次上的含义，以及量子神经网络给传统人工神经网络和智能信息处理领域所带来的创新意义。

## 1.2 研究历史及现状

近年来，国际上已有少数研究者在量子神经计算领域展开了诸如量子联想、并行学习、经验分析等问题的讨论，虽然目前还只是处于理论分析阶段，但是他们构筑了人工神经网络中量子计算的基础。

在 1989 年，英国 Oxford 大学的 Penrose 教授最早讨论了量子理论与人脑意识之间的关系问题，他指出：解决量子测量问题是最终解决人脑意识问题的先决条件，而连接量子现象和广义相对论的新型物理学可以解释一些脑力活动，如理解、感知、意识等<sup>[19, 20]</sup>；美国 Arizona 大学的 Hameroff 教授则认为，上述方法不再是研究由神经元所组成的网络结构，而是着重分析细胞内部的细微结构，例如在神经元内骨骼支架的微管（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<sup>[21]</sup>；最近斯洛文尼亚 Ljubljana 大学的 Perus 博士指出：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<sup>[22]</sup>。诸如此类的讨论还有许多<sup>[23, 24]</sup>，虽然目前神经科学界尚无法确认人脑中是否确实存在有量子效应，但是，用量子理论来解释大脑现象（即所谓的量子思维或量子意识）的确富有创见和一定的合理性。

而真正将神经计算与量子理论结合起来的是美国 Louisiana 州立大学的 Kak 教授，他在 1995 年发表的“*On Quantum Neural Computing*”一文中首次提出量子神经计算的概念，开创了该领域研究的先河<sup>[25]</sup>；同年英国 Sussex 大学的 Chrisley 博士提出了量子学习（Quantum Learning）的概念，他结合量子双缝干涉实验装置给

出了非叠加态量子神经网络模型和类 BP 的学习算法<sup>[26]</sup>, 此模型在理论上可以实现复杂的学习功能, 但遗憾的是它并没有利用量子计算的超并行性, 因此仅仅是一个粗糙的理论模型, 且缺乏应用的潜力; 1996 年美国 Wichita 州立大学 Behrman 博士领导的研究小组利用量子点 (Quantum Dot) 开发了量子神经网络的时间和空间物理模型, 他们结合 Feynman 路径积分 (Path Integrals) 原理初步实现了简单的逻辑学习功能, 此法具有很好的研究前景<sup>[27]</sup>; 1997 年美国 Brigham Young 大学的 Ventura 博士和 Martinez 教授给出了具有量子力学特性的人工神经元模型的框架, 为利用量子理论分析神经网络结构提供了可参考的方法<sup>[28]</sup>; 随后, 他们于 1998 年又提出了量子联想记忆 (Quantum Associative Memory) 网络模型, 并从数学上证明了在记忆容量和回忆速度方面该模型远远优于经典的 Hopfield 网络<sup>[29]</sup>, 但其中的网络结构和关键算法尚有待于进一步改进; 1998 年英国 Exeter 大学的 Menneer 博士构造了多感知机模型, 借助量子理论中的态叠加和测量坍缩原理, 初步实现了量子衍生神经网络 (Quantum-inspired Neural Networks) 的训练过程, 数据实验结果证实, 对于分类问题该网络要比传统的神经网络有效<sup>[30]</sup>, 此模型的构造思想很有新意, 但其中的坍缩算法过于简单, 且无法明确说明该网络优越的内在原因; 此外, 巴西 Brasilia 大学的李伟钢 (Li Weigang) 博士在 1998 年发表了关于量子并行自组织映射 (QuSOM) 的文章, 他在并行环境下对传统 Kohonen-SOM 算法进行了改进, 并给出未来使用量子逻辑门实现的方案<sup>[31]</sup>; 2001 年美国 Los Alamos 国家实验室的 Ezhov 教授讨论了量子神经计算中的干涉和纠缠问题<sup>[32]</sup>; 2002 年加拿大 Lethbridge 大学的 Andrecut 博士探讨了有关量子感知机的模型问题<sup>[33]</sup>, 此外, 还有其它一些相关的研究<sup>[34-36]</sup>。总之, 最近特别是近几年来量子神经计算领域的研究日趋活跃, 正因为此, 在 2000 年 3 月召开的第四届国际计算智能和神经科学会议上专门开设了“量子计算与神经量子信息处理”的专题会议, 会后还出版了论文专辑<sup>[37]</sup>, 此举引起了国际理论界的广泛关注。

下面我们选择其中几种典型的量子神经计算模型加以简单的评述。首先列出它们的设计要点, 见下表 1-1。

表 1-1 几种典型量子神经网络的设计要点

模型	神经元	连接	变换	网络类型	动力学特性
Perus	量子神经元	Green 函数	线性	时域	坍缩
Chrisley	狭缝位置	经典连接权	非线性	多层 BP	非叠加
Behrman	时间片段	光子作用	非线性	时、空域	路径积分
Goertzel	经典神经元	量子连接	非线性	经典	路径积分
Menneer	经典神经元	经典连接权	非线性	多感知机	多宇宙
Ventura	量子位	态纠缠	非线性	多感知机	非么正变换

1995年 Chrisley 提出了量子学习的概念,他利用量子双缝干涉实验装置与传统 BP 神经网络结构的相似性,给出了非叠加态量子神经网络模型,此模型具有前馈的拓扑结构:一个粒子 G 首先入射栅板 B,之后是一个光敏接受盘 P,栅板上有许多狭缝,其中上面部分组成输入区,下面部分组成权值区,调节狭缝的不同位置便可以在接受盘上产生不同的干涉模式,这样便实现了两组映射:一个是由输入(如特征串、图象、查询等)到栅板狭缝结构的映射(I),另一个是由干涉模式到输出(如分类、存储数据等)的映射(O)。进一步地,假设输入为  $\bar{x}$ , 输出为  $\bar{a}$ , 权值为  $\bar{\omega}$ , 干涉模式为  $\bar{p}$ , 这样可得  $S(\bar{x}, \bar{\omega}) = \bar{p}$ , 此处  $S(\cdot)$  为 sigmoid 函数, 而  $\bar{a} = O(S(\bar{x}, \bar{\omega}))$ , 再定义系统的误差函数为  $E = \sum (\bar{d} - \bar{a})^2$ , 其中  $\bar{d}$  为期望输出, 这样, 该系统就可以学习任意一个映射, 其训练算法类似于经典的 BP 算法。

此模型的开发极富创意,致使该结构随后又被多次提及<sup>[38, 39]</sup>。Chrisley 模型不同于传统 BP 网络之处在于它的权值和输入可以同时调节,能够实现更为复杂的学习功能,但遗憾的是它并没有利用量子计算的超并行性,因此仅仅是一个粗糙的理论模型,另外,该模型的数值实现相当困难,不易于实际应用。

1997年 Ventura 和 Martinez 给出了具有量子力学特性的人工神经元模型的研究框架,他们针对传统人工神经网络模型中最简单的感知机来研究其量子对照物。在他们所建立的量子神经元模型当中,原有的单个连接权被一个波函数  $\psi(\omega, t)$  所取代,其基态为经典模型的权矢。于是,感知机的权矢就被许多权矢的量子叠加所代替。他们假设  $\psi$  对时间恒定,便将寻找  $\psi$  的问题等价于量子力学中求解刚性箱体问题,但随后对于物理解的形式及其含义的解释比较含糊。不过这些却给我们提供了一种全新的研究方法,使我们可以在量子理论的层面上研究神经计算的意义,这将有助于获取前所未有的新结论。

1998年 Ventura 和 Martinez 又提出量子联想记忆模型。我们知道,传统的人工神经网络(如 Hopfield 网络)允许关联模式响应,但是其主要缺点在于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,它可存储的模式数仅为  $m \leq kn$ , 一般地  $0.15 \leq k \leq 0.5$ , 而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  $n$  量子位(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 上进行多项式级别的基本

操作, 操作过程中的关键算子<sup>[40]</sup>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,  $m \leq p \leq 1$ 。

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) 响应模式: 采用 Grover 量子查询算法, 其基本思想是将期望基态的相位反转, 然后以平均幅度反转所有的基态, 执行反复操作直到所期望基态的概率幅度趋近于 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, 且概率接近于 1。

定义  $\hat{i}_\phi$  为恒等矩阵, 除了  $i_{\phi\phi} = -1$  之外, 它用来反转基态  $|\phi\rangle$  的相位, 再定义  $\hat{W} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ , 称为 Walsh 变换, 还有算子  $\hat{G} = -\hat{W}\hat{i}_\phi\hat{W}$ , 用来以平均幅度反转所有的量子态。

(3) 量子联想记忆: 将上述两个算法结合起来, 定义  $\hat{P}$  算子完成 (1) 中所描述的存储算法, 那么记忆一组模式就可表示为  $|\psi\rangle = \hat{P}|0\rangle$ , 其中  $|\psi\rangle$  是基态的量子叠加态, 而每个基态代表一个模式。

现在假定已知一个模式的  $(n-k)$  位, 希望回忆起模式的全部, 利用 (2) 中的算法来回忆模式, 即  $|\psi\rangle = \hat{G}\hat{i}_\phi\hat{G}|\psi\rangle$ , 随后  $|\psi\rangle = \hat{G}\hat{i}_\phi|\psi\rangle$ , 重复  $O(\sqrt{2^n})$  次, 用  $\hat{i}_\phi$  反转所有代表已存储模式态的相位,  $\hat{i}_\phi$  反转那些与已知  $(n-k)$  位匹配态的相位。于是, 拥有  $(2n+1)$  Qubit 的量子联想记忆算法便可以在  $O(mn)$  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。

Ventura 和 Martinez 从数学上证明<sup>[41]</sup>, 在记忆容量和回忆速度上该模型与传统

的 Hopfield 网络相比有指数级的提高。但是，上述量子联想记忆网络的结构和算法仍有待于改进，具体地，可以借鉴神经计算的分布式结构来改善其存储结构，在模式响应时应结合模式自身的特点对 Grover 查询算法进行适当的改进<sup>[42]</sup>。

1998 年 Menneer 等人构造了多感知机模型，借助量子理论中的态叠加和坍缩原理，初步实现了量子衍生神经网络的训练过程。通过数据实验证实，对于分类问题来说该网络要比传统神经网络更为有效。

通常，在训练一个传统的神经网络时需要反复学习模式集，然而这种模式的重复出现并无多少人类学习的生理基础。因此 Menneer 等人认为，训练集中的模式犹如一个粒子，它在不同的宇宙中被大量分立的感知机（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狭缝，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立的宇宙，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，并将其推广到所有的输入模式中，结果的叠加权矢量就构成所谓的量子衍生波函数 (QuIWF)，它将坍缩到实际的输入模式上，具体的坍缩依赖于输入模式和坍缩的方式<sup>[43]</sup>。其中定义泛化规则为：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。

实验的初步结果表明，该模型在训练时权值改变的次数要比传统神经网络减少近 50%，却无泛化能力的损失，此外，它还具有消除灾变性失忆 (Catastrophic Forgetting) 的能力，这是因为一个网络仅训练一个模式，在训练过程中模式之间不发生相互干扰。

此模型的构造思想很有新意，它借助态叠加、坍缩等概念将传统的感知机与量子计算结合起来，提供了构造广义量子神经网络的一般方法。但是，其中的坍缩算法过于简单，与传统的模板匹配没有多少区别，因此无法体现出量子神经计算在信息处理上的优势。

除了上面讨论的量子神经计算模型之外，值得一提的还有 Perus 的量子联想网络<sup>[44]</sup>，Chrisley 的本体论量子模型<sup>[45]</sup>等等，但是在所有这些已有的模型当中，大多数只停留在理论分析水平，涉及到具体计算模型比较少，而能够设计出数值仿真的模型就更少，这也正是本文所期望加以改进和创新的原因。

虽然目前量子神经计算的研究还处于萌芽阶段，但通过初步的探讨和理论分析可以总结得出，与传统的神经计算相比较，量子神经计算及其模型至少在以下几

个方面具有明显的优势<sup>[46]</sup>：(1) 指数级的记忆容量和回忆速度；(2) 小的规模和简洁的网络拓扑结构，因此理论上有很好的稳定性和有效性；(3) 快速学习能力和一次学习能力；(4) 消除灾变性失忆的能力；(5) 高的信息处理速度等。可以预见，随着量子神经计算研究的不断深入，各种计算模型将不断出现，它的计算性能会进一步被揭示出来。

### 1.3 本文的主要研究内容

本文要解决的主要问题和研究的主要内容可以归纳如下：

- (1) 神经计算与量子计算的结合是构成量子神经计算的基础，因此从理论上探讨二者结合的可能性成为首要问题；
- (2) 如何调和神经网络的非线性特征与量子系统么正变换的线性特征是建立量子神经计算网络所必须解决的问题，本文将从量子理论的原理、观点、概念、实验方法等方面为其寻找合理的解决途径；
- (3) 针对不同的量子势函数，将传统意义的神经元进行量子化改造，建立具有量子力学特性的神经元模型，它是组成广义量子神经网络的基本单元，它的数学描述和物理含义对其自身以及整个网络的性能至关重要；
- (4) 量子理论中的多宇宙观点特别适合于构造用于分类任务的神经计算模型，本文将研究利用多 BP 网络构建多宇宙神经网络结构，并引入 Grover 搜索机制，提出新的坍塌算法，从而实现模式分类；
- (5) 利用量子理论中实际或假想的实验来构造计算模型，一方面可以充分利用其中隐含的量子计算优势，另一方面也有助于用实际的量子系统来实现神经计算；
- (6) 借鉴量子算法的结构思想，直接构造量子学习算子，用于量子系统的竞争学习，进一步拓宽量子算法的研究领域；
- (7) 基于量子神经网络的应用研究，一方面可以检验量子神经网络的有效性，另一方面还可以充分发挥量子神经计算的特点，研究其快速有效解决实际问题的能力。

通过上述工作，本文所预期达到的目的是：

- (1) 从理论上初步证明在量子系统实现神经计算的可能性；

- (2) 利用量子理论中的基本原理和观点、概念、方法、技术等设计出若干种计算模型及其算法，使之融合量子计算与神经计算二者的优点；
- (3) 基于上述模型探讨它们求解实际问题的计算能力，并与传统神经网络方法进行比较，以此来评价量子神经计算的性能。

## 1.4 论文安排

基于上述内容，本文的结构安排如下：

首先对文中所用到的必要的量子知识加以简单的介绍；然后从认知科学、神经科学及计算科学的角度讨论神经计算与量子计算二者的对应关系和结合的可能性，并给出相应的理论推导，由此引出量子神经计算这一全新的计算模式；接下来是文章的重点，将按照量子理论的系统结构，依次设计出几种计算模型及相应的学习算法，并通过理论论证和实例分析综合探讨它们的计算性能；最后总结全文。

## 第二章 量子理论与量子计算原理简介

只有当人们的视野转向微观世界时，才有可能发现微观粒子运动的规律。当进入 19 世纪末，人们才首次窥见一个完全不同于宏观世界的崭新世界即微观世界及其所展现出的至今仍令人迷惑的行为，这些应归功于当时的三大科学发现：即 X 射线、放射性元素铀和镭、基本粒子电子。当时，面对如此对象人们尚无法解释，正如 1900 年 4 月英国物理学家 W. Thomson 发表的一篇名为“19 世纪热和光的动力学理论上空的乌云”的文章<sup>[47]</sup>，文中指出经典物理学中存在着两个缺陷：一是电磁学理论中的关于光的波动学说，二是关于能量均分的 Maxwell-Boltzmann 理论，主要表现为热学理论中能量均分定则在气体比热及热辐射能谱的理论解释与实验结果不符，尤以黑体辐射理论中出现的“紫外灾难”最为突出。实际上，正是这两点导致了 20 世纪物理学的两大革命，前者导致了狭义相对论，后者则导致了量子力学的产生。

1900 年 12 月 14 日 M. Planck 提出了划时代的能量量子化假说，其关于“量子”的概念打破了一切自然过程都是连续的经典理论，第一次揭示了微观世界的奇异本质；继他之后 A. Einstein 发展了量子概念，首次提出“光量子”的概念，并以此成功地解释了光电效应；1913 年 N. Bohr 将量子概念应用于氢原子结构的分析中，他融合了光谱学和量子论，对量子理论的发展起到了推动作用；1924 年 L. de Broglie 提出了波粒二象性假说，在理论上将物质粒子与物质波联系起来；1925 年 W. Heisenberg、M. Born、Jordan、P. M. Dirac 等人发展了量子力学的矩阵表述形式；1926 年 E. Schrödinger 在 de Broglie 物质波的基础上提出了著名的波动力学形式—Schrödinger 方程；同年 Dirac 将矩阵力学与波动力学统一起来，最终完成了非相对论量子力学的创建工作<sup>[48]</sup>。

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）的运动规律，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。随着科学技术的发展，量子理论正在不断地得以完善，但是由于研究对象的特殊性，加之认知方式和技术手段等的限制，迄今还有一些基本的量子问题无法圆满解决，因此可以这么说：对于量子世界，“人们发现了它的部分规律，但却没有发现它本身”。

量子计算是 20 世纪的一大奇迹，它的概念直到 20 年前才出现。量子计算成功地利用了量子理论中态的叠加、纠缠、干涉、坍缩等概念，使计算取得了质的突

破,使得一些传统意义上的 NP-hard 问题在量子计算层面上有了多项式级求解的可能性,同时,量子计算及量子理论已经与计算机科学、信息科学等相互交叉,产生了诸如量子通信、量子计算机、量子信息处理等新的研究领域,为传统理论的发展提供了全新的思路和研究方法。

本文所关心的是那些有可能被应用于信息处理领域的量子知识,因此本章并不过多地去论述量子理论与量子计算,而是对其中的一些备用内容加以阐述,以期为后续章节提供必要的量子基础,从而使得文章结构更加系统、合理。

## 2.1 量子理论中的基本概念

### 2.1.1 不同的量子观点

为了合理解释波函数 $\psi$ 与粒子之间的联系,Born把散射波振幅绝对值的平方看成散射粒子偏转通过立体角中的几率,进而将波函数绝对值的平方设想为表示在某点附近的空间区域内遇见粒子的几率,即认为波函数是一种几率波而非物理实体;之后 Heisenberg 提出了关于微观运动规律的不确定性关系;同年 Bohr 阐述他的互补原理。至此,由不确定性关系、互补原理和 Born 的统计解释形成了量子力学中正统的 Copenhagen 量子观点。

另一方面,量子的实在观点是以粒子的经典连续运动为基础的,其中被广泛讨论的有 D. Bohm 的隐变数理论<sup>[49]</sup>和 H. Everett 的多宇宙观点<sup>[50]</sup>。

隐变数理论假设粒子与波函数同时存在,其中波函数被直接当作是存在于数学置配空间中的物理场,并假设它满足 Schrödinger 方程,而粒子则是由波函数导引进行经典连续运动,此外还假设粒子与某种未知背景系统之间存在随机相互作用,即需要假设背景随机扰动的存在,因此隐变数理论与量子力学具有不同的预测,尽管差别很小。该理论的主要缺陷在于以下几点:(1)缺乏合理的相对论协变形式的扩展;(2)存在所谓的 Born 规则问题;(3)量子非定域性问题;(4)位置测量无法显示 Bohm 粒子的真实位置;(5)空波问题;(6)作用-反作用原理问题;(7)基态的稳定性问题等等。实际上,在 Bohm 的理论中,不仅粒子的径迹是不可观察的,它所假设的另一物理实在— $\psi$ 场更是不可探测的隐变数,因为对于单个粒子的物理测量一般只产生一个关于粒子性质的确定的结果,而根本测不到任何 $\psi$ 的性质,这样隐变数理论所提供给我们的两种物理实在—粒子径迹和 $\psi$ 场都是不可

观测的。

多宇宙观点认为存在一些平行的完全的宇宙，它们在某种确定的意义上在相同的时间和空间中存在着，就某种意义而言，它们与我们共享同样的时间和同样的空间。不同的平行宇宙是一个公共的物理实体的一部分而关联在一起，物理实在就是纠缠在一起的所有宇宙的集合。在理论上，多平行宇宙是与波函数的各个投影分支相联系的，当宇宙面临一种量子选择时，它就分裂成两个不同的宇宙。该理论提供了一个完全决定性的理论，当考虑到环境引发解耦时，多宇宙解释的实验预测将等同于正统统计解释的预测，而当屏蔽掉环境引发解耦后，多宇宙解释就与正统解释及修正的量子动力学理论具有不同的实验预测。该理论的主要缺陷在于：(1) 量子自杀问题；(2) 几率解释问题，这是其根本问题，即关于确定性测量结果的几率解释如何来源于一个决定论的波函数演化，这个波函数包容了对应于所有可能测量结果的状态；(3) 随机性问题即未解释量子耦合规则；(4) 优选基问题；(5) 公理化问题；(6) 主观世界问题；(7) 内部观点与外部观点的不一致性等等。

实际上，隐变数理论与多宇宙观点都没有为波函数这一物理实体及其演化规律—Schrödinger 方程提供任何物理解释，在隐变数理论中波函数被直接当作是存在于数学置配空间中的物理场，而在多宇宙观点中波函数同样被直接看作是一种勿需任何解释的物理实体，因此二者实际上并没有为隐藏在波函数背后的神秘的微观实在提供一幅清晰的物理图象。

### 2.1.2 线性叠加与态叠加原理

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，波函数的形式表示为  $\psi(\vec{r}, t) = Ae^{i(\vec{k}\cdot\vec{r} - \omega t)} = Ae^{\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et)}$ ，显然，如此的波函数不再是经典意义上的，它体现了量子的波粒二象性。线性叠加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  $|\phi_i\rangle$  为基态，那么一个量子态就可以表示为  $|\psi\rangle = \sum_i c_i |\phi_i\rangle$ ，也就是说量子态  $|\psi\rangle$  是所有基态  $|\phi_i\rangle$  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，这一性质就称为态叠加原理。

### 2.1.3 相干 (Coherence)、消相干 (Decoherence) 以及测量原理

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消相干或坍缩, 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  $|\phi_i\rangle$  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, 因此由振幅  $c_i$  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。

### 2.1.4 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  $\hat{A}|\phi_i\rangle = a_i|\phi_i\rangle$ , 其中  $a_i$  为本征值, 方程的解  $|\phi_i\rangle$  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), 即满足  $\hat{A}^\dagger \hat{A} = \hat{A} \hat{A}^\dagger = \hat{I}$ , 其中  $\hat{I}$  为单位算子,  $\hat{A}^\dagger$  为算子  $\hat{A}$  的复共轭变换。

### 2.1.5 干涉

干涉是一种常见的现象, 它是由于相位的关系而产生的波的幅度的增强或减

弱，该现象适用于各种经典的波动力学对象，从水波到光波，同样由微观粒子的双缝干涉实验结果可知，它对于量子力学中的几率波也同样适用。

### 2.1.6 纠缠态及量子关联现象

这是量子理论中最能体现其非直观特性的现象。所谓纠缠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纠缠形式，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：

$$|\varphi_1\rangle = c_1|\alpha_1\rangle + c_2|\alpha_2\rangle; \quad |\varphi_2\rangle = c'_1|\beta_1\rangle + c'_2|\beta_2\rangle$$

当两个子系统相互独立时，由它们组成的大系统的态是它们态的张量积，即  $|\varphi\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle$ ；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，这样的态就称为纠缠态。

我们假设一个量子体系处在  $|\varphi\rangle = c_1|\alpha_1\rangle|\beta_1\rangle + c_2|\alpha_2\rangle|\beta_2\rangle$  的纠缠态中，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，此过程称为量子关联现象。

### 2.1.7 量子隐形传态

量子隐形传态用接近于物理规律的思想可解释为：先提取原物的所有信息，并将这些信息传送到接收地点，然后依据这些信息，选取与构成原物完全相同的基本单元（如原子），制造出原物的完美的复制品。但遗憾的是，这种方法违背了量子力学的不确定关系，因为不确定关系认为，不可能对一个物体（量子体系）的所有物理量都进行精确测量，因而提取一个物体的所有信息是不可能的。当然，人们还可以提出另一种不必对物体进行力学量测量的可能方案，因为按照量子力学的理论，物体的全部信息可由其量子态给出，换句话说，只要我们能精确地复制物体的量子态，并使远处相同类型的原子集合处于与原物完全相同的量子态上，则这种隐形传态就能成为现实，但是可惜的是，完全精确地复制量子态也是不可

能的，它违背了量子力学的另一条基本定理，即量子不可克隆定理<sup>[51]</sup>，也就是说，通用的量子复制机是不存在的。

直到 1993 年，Bennett 等人发表了一篇题为“经由经典和 EPR 通道传送未知量子态”的开创性文章<sup>[52]</sup>，重新点燃了人们研究隐形传态的兴趣，并引发一系列富有成果的研究。Bennett 等人提出的方案，其基本思想是为实现传送某个物体的未知量子态，可将原物的信息分成经典信息和量子信息两个部分，它们分别经由经典信道和量子信道传送给接收者。经典信息是发送者对原物进行某种测量而获得的，量子信息是发送者在测量中未提取的其余信息。接受者在获得这两种信息之后，就可以制造出原物的完美的复制品。在这个过程中，原物并未被传给接受者，它始终留在发送者处，被传送的仅仅是原物的量子态，发送者甚至可以对这个量子态一无所知，而接受者是将别的物质单元（如粒子）变换成为处于与原物完全相同的量子态，原物的量子态在发送者进行测量及提取经典信息时已遭破坏，因此，这是一种量子态的隐形传送，最终恢复原物量子态的粒子也可以不必与原物同类，只要它们满足相同的量子代数即可。由于经典信息对量子态的隐形传送是必不可少的（否则将违背量子不可克隆定理），而经典信息传递速度不可能快于光速，因此，量子隐形传态也不会违背相对论的光速最大原理。

为了更好地说明上述问题，首先需解释的是 EPR 效应，它是 1935 年由 Einstein、Podolsky 和 Rosen 三人提出的一个被人们称为 EPR 佯谬的著名假想实验，其本意是企图证明量子力学是不完备的<sup>[53]</sup>，然而后来的实验事实并不支持 EPR 的观点。这个佯谬的起源在于局域性（Locality）的经典概念，它将一个客观物理实在与每个粒子及其力学变量相联系，但是在量子力学中，只要体系不处于某个力学量算符的本征值，则该动力学变量并不确切地具有一个值，除非对它进行测量，因此，在某种意义上讲，测量产生物理实在。按照量子力学理论，EPR 粒子对作为一个量子系统处于如下的量子态（俗称为 EPR 态）：

$$|\varphi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \quad (2-1)$$

这实际上就是一种量子的最大纠缠态，不管两个粒子相距多远，它们都处于这种相互关联状态，这就是量子力学的非局域性效应（Non-local Effect）。许多实验证实，非局域性效应是量子力学的基本特性。虽然 Einstein 等人对量子力学理论的质疑被否定，但是上述非局域性效应却是他们根据量子力学原理在 EPR 实验中揭示出来的，因此又称之为 EPR 效应。

量子隐形传态的基本原理可描述如下：

- (1) Alice (发送者)、Bob (接收者) 和一个 EPR 源, 再加上一个经典信道与一个量子信道便构成一个量子隐形传态单元模型 (见图 2-1);
- (2) Alice 和 Bob 希望能够相互通讯, 由 EPR 源发送给每个人一个纠缠粒子, 两个粒子构成一个 EPR 对, 即  $|\varphi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ ;
- (3) Alice 试图通过经典信道和量子信道将她的量子态  $|\phi\rangle = a|0\rangle + b|1\rangle$  传送给 Bob, 她首先对她的量子态和纠缠态求直积, 得到一个初始态, 即

$$\begin{aligned} |\phi\rangle \otimes |\varphi_0\rangle &= \frac{1}{\sqrt{2}} [a|0\rangle \otimes (|00\rangle + |11\rangle) + b|1\rangle \otimes (|00\rangle + |11\rangle)] \\ &= \frac{1}{\sqrt{2}} (a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle) \end{aligned} \quad (2-2)$$

- (4) Alice 应用  $C_{not} \otimes I$  和  $H \otimes I \otimes I$  变换去直乘初始态, 其中  $C_{not}$  为受控非门 (Controlled-NOT),  $I$  为单位变换矩阵,  $H$  为 Hadamard 变换矩阵, 即

$$\begin{aligned} &(H \otimes I \otimes I)(C_{not} \otimes I)(|\phi\rangle \otimes |\varphi_0\rangle) \\ &= (H \otimes I \otimes I)(C_{not} \otimes I) \frac{1}{\sqrt{2}} (a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle) \\ &= (H \otimes I \otimes I) \frac{1}{\sqrt{2}} (a|000\rangle + a|011\rangle + b|110\rangle + b|101\rangle) \\ &= \frac{1}{2} [ |00\rangle(a|0\rangle + b|1\rangle) + |01\rangle(a|1\rangle + b|0\rangle) + |10\rangle(a|0\rangle - b|1\rangle) + |11\rangle(a|1\rangle - b|0\rangle) ] \end{aligned} \quad (2-3)$$

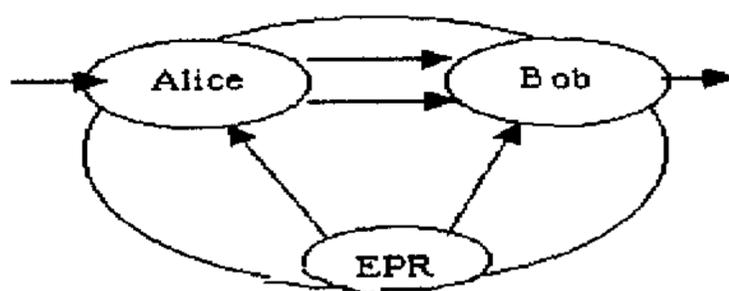


图 2-1 量子隐形传态模型示意图

- (5) 这样, Alice 控制着前两位, Bob 控制着最后一位, 并且 Alice 对前两位四种态 ( $|00\rangle$ 、 $|01\rangle$ 、 $|10\rangle$  和  $|11\rangle$ ) 的测量概率相等, 都等于  $1/4$ , 她将测量结果作为经典信息传送给 Bob;
- (6) 依据 Alice 的测量结果, Bob 所控制的位的态分别对应为  $a|0\rangle + b|1\rangle$ 、 $a|1\rangle + b|0\rangle$ 、

$a|0\rangle - b|1\rangle$  和  $a|1\rangle - b|0\rangle$ ，它们是由基态（0 或 1）和相态（+或-）联合组成的；

(7) Bob 收到 Alice 的信息后将他的一个 EPR 粒子解码，最终获得 Alice 原有态  $|\phi\rangle = a|0\rangle + b|1\rangle$ ，其间他利用了相应的解码变换矩阵，见表 2-1，其中  $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

为单位变换矩阵， $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  为负单位变换矩阵， $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  为移相矩阵，

$$Y = ZX = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

表 2-1 Bob 的解码变换矩阵

前两位可能态	后一位可能态	解码变换矩阵
00	$a 0\rangle + b 1\rangle$	I
01	$a 1\rangle + b 0\rangle$	X
10	$a 0\rangle - b 1\rangle$	Z
11	$a 1\rangle - b 0\rangle$	Y

## 2.2 量子门与量子门组网络

### 2.2.1 量子比特

我们知道，经典计算机的存贮单元是比特，它只有两种状态，要么为 0，要么为 1。而量子计算机最基本的存贮单元则是量子比特，它是任何一个有二维 Hilbert 态空间的量子体系，它的态空间有两个基，记为  $|0\rangle$  和  $|1\rangle$ 。与经典计算机中的比特不同的是，量子比特的状态可以为任意归一化的态  $|\psi\rangle = a|1\rangle + b|0\rangle$ ，其中  $|a|^2 + |b|^2 = 1$  满足归一化条件，所以，一个量子比特所包含的信息要比经典的比特多。各量子比特的状态可以是相互独立的，也可以是相互纠缠的。

### 2.2.2 量子门<sup>[54]</sup> (Quantum Gates)

对量子位的态进行一系列的么正变换可以实现某些逻辑功能，变换所起的作用相当于逻辑门所起的作用，我们称在一定时间间隔内实现逻辑变换的量子装置为量子门。按照其作用的量子位数目的不同可分为一位门、两位门和三位门等。

在量子计算中常用的一位门有相移 (Phase Shift) 门、非 (NOT) 门、Hadamard

门等。

相移门对一位量子位的态进行相移变换，表示为  $\phi = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}$ ，即  $\phi: |0\rangle \rightarrow |0\rangle, |1\rangle \rightarrow e^{i\theta}|1\rangle$ ，虽然它只是使量子态转动一个角度，没有逻辑意义，但在量子计算中会和 Hadamard 门结合起来实现一个量子位的任意么正变换，因此具有重要作用。

量子“非”门对一个量子位的态进行“非”变换，表示为  $X = |0\rangle\langle 1| + |1\rangle\langle 0| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ，即  $X: |0\rangle \rightarrow |1\rangle, |1\rangle \rightarrow |0\rangle$ ，它不仅是么正变换，而且还是厄米 (Hermitian) 变换，既是可逆的，又是可测量的。

Hadamard 门对一个量子位的态进行 Hadamard 变换 (或称 Walsh 变换)，表示为  $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ ，即  $H: |0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |1\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ ，相当于将  $|0\rangle$  态顺时针旋转  $45^\circ$ ，将  $|1\rangle$  态逆时针旋转  $135^\circ$ 。

两位门中最常用的是受控非门，又称量子“异或”门，它有两个输入端  $|x\rangle$  和  $|y\rangle$ ，各有一个量子位，即构成两位量子系统，对它们进行变换的算子应该是  $4 \times 4$  的么正矩阵，表示为

$$C_{not} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \text{ 即 } C_{not}: |00\rangle \rightarrow |00\rangle, |01\rangle \rightarrow |01\rangle, |10\rangle \rightarrow |11\rangle, |11\rangle \rightarrow |10\rangle$$

同时，量子“异或”门是么正矩阵，也是厄米矩阵，用图可以表述如下。

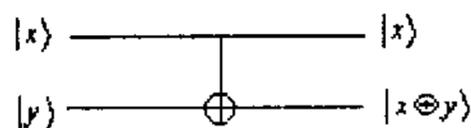


图 2-2 量子异或门 (受控非门) 符号

三位门有 Toffoli 门，又称为量子“与”门或“控控非”门。它有三个输入端  $|x\rangle$ 、 $|y\rangle$  和  $|z\rangle$ ，只有当  $|x\rangle$  和  $|y\rangle$  同时为“1”时， $|z\rangle$  才变为相反的状态，表示为

$$CC_{\text{not}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

即  $CC_{\text{not}}: |000\rangle \rightarrow |000\rangle, |001\rangle \rightarrow |001\rangle, |010\rangle \rightarrow |010\rangle, |011\rangle \rightarrow |011\rangle, |100\rangle \rightarrow |100\rangle, |101\rangle \rightarrow |101\rangle, |110\rangle \rightarrow |110\rangle, |111\rangle \rightarrow |110\rangle$ . Toffoli 曾证明该量子门对经

典计算是通用的<sup>[55]</sup>, 用图可以表述如下, 显然若令  $|z\rangle$  为  $|0\rangle$  Toffoli 门就实现了“与”门的功能。



图 2-3 量子 Toffoli 门符号

### 2.2.3 量子门组网络

迄今已经提出的量子计算机模型主要有 Turing 机、量子门组网络模型和量子细胞自动机模型。已有证明, 量子 Turing 机可以用多项式大小的量子门组网络或用量子门组网络多项式大小的时间花费模拟, 另外由于量子门组网络模型与经典计算机存在着更直接的对比和联系, 因此是目前讨论的最多的量子计算机模型, 它是经典计算机门组网络结构的量子推广。所谓量子门组网络是由多个量子门组成的, 而这些量子门的操作在时间上保持同步。例如, 可以用一个  $C_{\text{not}}$  门和一个量子“与”门组成量子半加器, 还可以由它们构成量子全加器, 并由此推广再构成多位量子加法器, 等等。

在量子计算中存在着通用门组, 1989 年 Deutsch 将经典 Toffoli 门推广到量子情况<sup>[56]</sup>, 得到了 Deutsch 门, 它是一个三位控制-控制-R 门, 当且仅当前两个量子位都处在态  $|1\rangle$  时, 才对第三位态施加一个 R 变换, 即  $R = -iR_x(\theta) = -i \begin{pmatrix} \cos \theta/2 & i \sin \theta/2 \\ i \sin \theta/2 & \cos \theta/2 \end{pmatrix}$ ,

可以证明 Deutsch 门可以实现经典 Toffoli 门的作用。1995 年 Divincenzo 证明 Deutsch 门可以用两位门组来实现<sup>[57]</sup>, Deutsch 又进一步证明<sup>[58]</sup>, 差不多任意两位门或 n 位

门对量子计算都是通用门组。之后, Barenco 等人又证明<sup>[59]</sup>, 通用量子门还可以由经典多位门和量子一位门构成, 特别是经典两位  $C_{\text{not}}$  门和一位量子门就构成了计算的通用门组。

## 2.3 量子算法

### 2.3.1 概述

某些计算复杂度很高的问题往往不能有效地在经典计算机上计算, 但在量子计算机上有可能降低其复杂度, 前提是要有高效的量子算法。量子并行计算虽然依次变换就可以产生所有  $x$  的  $f(x)$ , 然而测量时只能测得一个结果, 而不能选择所需的结果。量子算法的核心思想就是利用量子态的干涉特性, 使所需的结果增强, 同时使不需要的结果减弱, 从而使得所需的结果在测量时就会以很高的概率出现。

由于应用背景和设计难度的限制, 迄今已提出的量子算法并不多, 其中比较有代表性的有四个: Deutsch 算法、Simon 算法、Shor 算法和 Grover 算法。下面简单加以论述。

### 2.3.2 Deutsch 算法<sup>[60]</sup>

Deutsch 问题是至今所知的第一个和最简单的量子算法之一, 然而它却阐明了量子计算的一些重要的性质。问题的描述是: 仅取最简单的一个输入位  $x$  和一个输出位 (未知函数)  $f(x)$ , 因为每个  $f(0)$  和  $f(1)$  有两个可能值, 我们共有四个可能的函数值, 即:  $f_1(0) = f_1(1) = 0$ ;  $f_2(0) = f_2(1) = 1$ ;  $f_3(0) = 0$  和  $f_3(1) = 1$ ;  $f_4(0) = 1$  和  $f_4(1) = 0$ 。目的是要判断  $f(x)$  是恒定的 (即  $f(0) = f(1)$ , 如  $f_1$  或  $f_2$ ) 还是均衡的 (即  $f(0) \neq f(1)$ , 如  $f_3$  或  $f_4$ )。直观上, 最好的经典策略是对输入 0 和 1 计算  $f$ , 然后比较结果, 即需要函数  $f$  的两次演算。但是, 量子算法仅需要一次演算就能够得到结果。算法步骤如下, 需要两个量子位:

- (1) 第一个量子位初始化在态  $|0\rangle$ , 且第二个量子位初始化在态  $|1\rangle$ , 总的态是  $|01\rangle$ ;
- (2) 对于每个量子位作用 Hadamard 变换, 就使得态成为  $\frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle)$ ;

(3) 通过定义作用在基矢上的两位门  $U_f: |i, j\rangle \rightarrow |i, j \oplus f(i)\rangle$ , 在上述迭加上计算函数  $f$ , 其中  $i, j = 0, 1$ , 且  $\oplus$  表示模去 2 相加;

(4) 最后, 再一次对每位进行 Hadamard 变换。容易验证, 两个量子位的终态是:  $|01\rangle$  (如  $f = f_1$ ),  $-|01\rangle$  (如  $f = f_2$ ),  $|11\rangle$  (如  $f = f_3$ ),  $-|11\rangle$  (如  $f = f_4$ )。

所以, 最终的对于第一个量子位的测量将展现函数是否恒定或均衡。从上述算法中可见, 量子力学的叠加和线性特性是关键所在, 当  $U_f$  作用于态  $\frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle)$  上, 它叠加的四个态上的  $f$  被同时有效地计算出来, 同时量子干涉也是重要的, 因为最后的两个 Hadamard 变换导致叠加中不同部分相干涉得到有用的终态。因此, 量子力学的规律起到了决定性的作用。

### 2.3.3 Simon 算法<sup>[61]</sup>

设有一个量子黑盒, 它计算函数  $f: \{0, 1\}^n \rightarrow \{0, 1\}^n$ , 假设函数  $f$  是  $2 \rightarrow 1$  的, 并且  $f$  对某一个  $a \in \{0, 1\}^n$  满足:  $f(x) = f(y)$ , 当且仅当  $y = x \oplus a$ , 就是说如果认为  $x$  在  $\{0, 1\}^n$  中取值,  $a$  就是函数  $f$  的周期。Simon 问题就是知道  $f$  求函数  $f$  的周期。

在经典上这是一个难题, 因为对于  $n$  位  $a$  需要取输入从  $|0\rangle$  开始运行量子黑盒最多  $2^n$  此, 直到碰巧对两次输入的  $x, y$  有  $f(x) = f(y)$ , 才能从结果中得到  $a$ 。但使用量子算法只需  $n$  次运行就可求出  $a$ 。算法步骤如下, 需要有两个  $n$  位寄存器:

(1) 用  $H^{(n)}$  作用到处在  $|0\rangle^n$  态的第一寄存器上, 制备出态  $\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle$ , 并保持第二寄

存器仍为  $|0\rangle^n$ ;

(2) 以上述两个寄存器所在的态作为量子黑盒的输入态, 其输出将是两个寄存器的纠缠态, 即  $\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle |f(x)\rangle$ ;

(3) 测量第二寄存器, 结果将从  $f(x)$  的  $2^n$  可能项中选出一个, 每一个将等概

率地出现, 假设结果为  $f(x_0)$ , 由于仅  $x_0$  和  $x_0+a$  才被  $f$  映射为  $f(x_0)$ , 所以测量在第一寄存器中制备了态  $\frac{1}{\sqrt{2}}(|x_0\rangle + |x_0 \oplus a\rangle)$ ;

(4) 为了得到  $a$ , 对第一寄存器再施加  $H^{(n)}$  作用, 得到  $\frac{1}{2^{\frac{n-1}{2}}} \sum_{a \cdot y=0} (-1)^{a \cdot y} |y\rangle$ ;

(5) 测量上述态, 并随机选出一个  $|y\rangle$ , 满足条件  $a \cdot y = 0$ , 并重复上面的算法, 每次找到一个满足条件的  $y$  值;

(6) 将  $n$  个线性独立的  $|y_1, y_2, \dots, y_n\rangle$  组成方程组, 求解就可以得到  $a$ 。

这样, 考虑到有可能两次计算最后测量得到相同的  $y$ , 重复运行量子黑盒的次数仍是  $n$  的多项式数量级。可见, 量子算法相对经典算法可以得到指数级的加速。

### 2.3.4 Shor 算法<sup>[12]</sup>

Shor 算法用于求解大数的质因式分解, 所谓质因式分解可描述为: 若寻找某个数  $N$  的质数因子, 等于要找到  $N$  的最小因子  $r$ , 使得  $a^r = 1 \pmod{N}$ , 其中  $a$  被选择为  $N$  的互质数, 即除了 1 之外,  $a$  和  $N$  没有公约数。换言之, 就是要确定函数  $a^r \pmod{N}$  的周期。

经典算法分解一个大数需要指数级的时间, 而在量子计算机上应用 Shor 算法则可以多项式时间高效运算。例如 1994 年曾有人就 RSA-129 问题 (一种当前被公认为是最安全、经典计算机不能破译的公钥密码系统) 用 1600 台经典计算机进行因式分解, 结果需要 8 个多月才能完成, 并进一步推算, 若将问题的位数增加到 1000 位则需要  $10^{25}$  年, 而在量子计算机上使用 Shor 量子算法只需运行几百万步即可完成。该算法的主要特征和步骤可表示如下:

(1) 一个待分解数为  $N$ , 取两个具有  $k$  个量子位的量子寄存器 (其中  $k$  近似为  $\log_2 N$ ), 并制备第一个寄存器, 使其处在从 0 到  $2^k-1$  连续的自然数的等权迭加中, 留下第二个寄存器处在 0 态, 即  $|\psi\rangle = \frac{1}{\sqrt{2^k}} \sum_{n=0}^{2^k-1} |n\rangle |0\rangle$ , 将  $n$  转换为二进制位链, 即为寄存器的一般状态;

(2) 在第二个寄存器中计算函数  $a^n \pmod{N}$ , 结果为  $|\psi_1\rangle = \frac{1}{\sqrt{2^k}} \sum_{n=0}^{2^k-1} |n\rangle |a^n \pmod{N}\rangle$ ;

(3) 对第二个寄存器做投影测量  $|u\rangle\langle u| = |a^m \pmod N\rangle\langle a^m \pmod N|$ ，可得

$|\psi_2\rangle = \frac{1}{\sqrt{2^k}} \sum_{j=0}^{2^k-1} |jr+l\rangle|u\rangle$ ，其中略去了归一化因子，并且在以后的步骤中第二个寄存器将不再使用；

(4) 为提取在第一个寄存器中包含的周期  $r$ ，对它进行分立 Fourier 变换 (DFT)，

即  $U_{DFT}|jr+l\rangle = \frac{1}{\sqrt{2^k}} \sum_{y=0}^{2^k-1} \exp\{2\pi i(jr+l)y/2^k\}|y\rangle$ ，根据正交条件可知，仅当  $y=mM$

( $m=0, 1, \dots$ ) 时， $\sum_{j=0}^{M-1} \exp\{2\pi i j y / M\} = M$ ，否则为零，那么当  $2^k/r$  是整数时，

可得终态为  $|\psi_3\rangle = \frac{1}{\sqrt{r}} \sum_{m=0}^{r-1} \exp(2\pi i m / r) |m2^k / r\rangle$ 。但是当  $2^k/r$  不是整数时，需要进行更为仔细的分析；

(5) 在  $y=m2^k/r$  基进行测量，其中  $m$  是一个整数，所以一旦获得了特定的  $y$ ，就必须解方程  $y/2^k=m/r$ ，假定  $m$  和  $r$  没有除 1 以外的公约数，通过把  $y/2^k$  约化到一个不可约的分数得到  $r$ ，那么就很容易根据因式化算法推断  $N$  的因式，倘若  $m$  和  $r$  有公约数，那么该算法失败，整个算法需从头重复。

由此可见，一般而言 Shor 算法是概率性的，即通过运行以上的量子算法，我们得到  $N$  的因式或许有时不是正确的答案，不过可以通过反乘因式检查结果是否正确，若不正确，可以从头重复 Shor 算法，直到得到正确的答案为止。已有证明，即使 Shor 算法中存在随机的要素，但它仍是一个有效的算法，在整体上它是多项式时间的。

### 2.3.5 Grover 算法<sup>[13]</sup>

在经典计算机中，从具有  $N$  个条目的无序数据库中找到一个特定的条目通常需要进行  $N/2$  步的搜索，但 Grover 量子算法却可以在量子计算机中只用  $O(\sqrt{N})$  步即可完成同样的搜索，尽管不象 Shor 算法那样有指数级的提高，但可以证明，随着数据大小和集成度的增加，对于时间的节省还是有很大意义的。例如，使用 Grover 量子算法可以仅用 2 亿步取代经典计算机的大约  $3.5 \times 10^{16}$  步，破译 56 位 DES 数据编码标准（即一种被广泛用于保护银行间和其他方面金融事务的标准）。

上述的搜索问题可形式化地表示为：考虑到分别被标记为  $S_0, S_1, \dots, S_N$  的  $N$  ( $=2^L$ ) 个不同状态和一个条件  $C_j$ ，假定仅有一个态  $S_j$  满足该条件，目的是当最小

化条件  $C_j$  的演化数目时确定  $S_j$ 。解该问题的 Grover 算法可归纳如下:

(1) 设置单个量子寄存器, 使它处在所有计算态的等权迭加之中, 记为

$$|\psi(\theta)\rangle = \sin\theta|j\rangle + \frac{\cos\theta}{\sqrt{N-1}} \sum_{i \neq j} |i\rangle, \text{ 其中 } j \text{ 是要搜索的元素 } t=S_j \text{ 的指标;}$$

(2) 仅翻转迭加中一个特定元素的态  $|j\rangle$  的符号, 其他不变, 接着做 DFT, 改变

除  $|0\rangle$  之外的所有分量的符号, 最后做 Fourier 变换反演回来, 所有这些运算

完成了如下变换的精细的干涉效应, 即

$$U_G|\psi(\theta)\rangle = U_{DFT}^{-1}(2|0\rangle\langle 0| - I)U_{DFT}(I - 2|j\rangle\langle j|)|\psi(\theta)\rangle = |\psi(\theta + \phi)\rangle, \text{ 其中 } I \text{ 是单位矩阵,}$$

$\sin\phi = 2\sqrt{N-1}/N$ , 此时特定元素的系数稍微比所有其他元素大一些;

(3) 运用  $U_G$  算子  $m$  次, 在此  $m \approx (\pi/4)\sqrt{N}$ , 缓慢的转动引起  $\theta$  非常接近于  $\pi/2$ ,

使得量子态几乎精确地等于  $|j\rangle$ , 此时测量的误差率只有  $O(1/M)$ 。

此外, Grover 也已经将他的思想推广到其他问题, 证明了均值或中值的估计问题能够用类似的技术求解。

## 2.4 量子计算的性能分析

经典计算机科学中有专门的可计算性和算法复杂性理论, 在该理论中, 称一个特殊指定的指令序列为算法, 算法的难易程度由算法复杂性来衡量, 而算法复杂性取决于执行这个算法解决具体问题所消耗的物理资源的多少。一般地, 限制计算机能力的物理资源有两类: 即时间和空间。按照经典算法复杂性理论, 一个问题的大小可以用一个整数  $n$  表示,  $n$  是指定这个问题需要输入的信息量的度量, 若一个问题的大小是  $n$ , 解这一问题的最好算法需要的时间(或计算步数)为  $T(n)$ , 如果当  $n$  增大时,  $T(n)$  的增加始终不比  $n$  的一个多项式函数增加得更快, 就称这一算法是容易的, 在算法复杂性分类中就属于 P 类问题; 如果  $T(n)$  随  $n$  指数增大, 就说算法是难的, 属于 NP 类问题。一个 NP 类问题通常说成是“不可解的”, 这是指当具体问题增大时, 所需的物理资源是如此之多, 以至对任何实际计算设备都是难以承受的。上面的分类不依赖于具体计算设备, 这一点由 Church 定理保证: 任何一个物理计算装置都可以被经典 Turing 机以多项式资源耗费来模拟。

在经典计算理论中还阐明了某些问题是 NP 完备的概念。对于一个 NP 完备问

题,如果能找到一个可以在多项式时间内解决它的算法,即把它转化为 P 类问题,则其它任何 NP 类问题都可转化为 P 类问题。虽然未加证明,经典计算机科学家相信,所有的 NP 类问题似乎都是 NP 完备的。

量子计算具有超出经典计算的能力,是通过它能以多项式时间求解某些经典计算中非 P 类问题说明的。量子计算服从量子力学规律,它可以支持新类型的量子算法。既然 Shor 已经找到了大数质因式分解的快速算法,可以在多项式时间内解决在经典计算机上需要指数时间的问题,这就在量子计算机上把一个 NP 类问题转化为 P 类问题(虽然没有证明大数质因式分解是 NP 类问题,但一直没找到解这一问题经典机器的多项式算法,于是很多人相信它是 NP 类的)。这样,按照上述经典算法复杂性理论,就有可能把经典计算理论中的某些 NP 类问题在量子计算上化成易解的 P 类问题。这样经典计算复杂性理论分类在量子计算上将失去绝对性,原来的以经典 Turing 机为基础的计算复杂性理论就需要重新考虑。但是直到最近,仍不能肯定这种推理是否正确,即是否所有的 NP 类问题都能在量子计算上化成 P 类问题,因为 Shor 算法毕竟只是解决了这类问题中一个。

量子计算之所以能通过一次运算就给出问题的答案,是因为它不仅能运算和,而且它能对和的任意叠加态进行运算,然后从中取出有关函数的整体信息,这就是所谓的量子并行性。我们可取初态作用一次,这样,有关函数的某些整体信息就荷载在这个态中,因为实际上这个态包含了所有函数值的信息。通过适当的变换和测量,从上面态中取出所要的整体信息。这样,经典算法需要逐个求出所有函数值的问题在量子算法中只运算一次(操作)就能解决,从而使得很多不能用经典算法解决的问题有可能用量子算法来解决。量子算法的并行性利用得越充分,量子算法的优越性也越突出。

## 2.5 量子计算机

### 2.5.1 量子计算机概念的产生和发展

以计算机科学为核心的信息论和相对论、量子论一起,被称为 20 世纪科学的三大发现。近 20 年来,信息论和量子论结合产生了量子信息理论。量子计算机作为量子信息处理的工具,就成为近几年来人们十分关注的问题。

量子计算和量子计算机的概念起源于著名物理学家 Feynman。1982 年, Feynman 在研究物理系统的计算机模拟时,论证了用经典计算机模拟量子力学系统,随输入(粒子数、自由度)的增大,计算资源(时间和空间)的消耗将指数

增大, 并由此启发了用量子力学性质工作的计算机(即量子计算机)有可能避免这一困难。1985年, Deutsch 提出了第一个量子计算机的设计蓝图—量子计算机的网络模型, 定义了量子 Turing 机, 并预言了量子计算机的潜在能力。在上世纪 90 年代初, Deutsch、Berthianme 等人就寻找量子计算机可以比经典计算机更有效求解问题的量子算法。1994 年, Shor 设计了一个具体的量子算法, 可以在量子计算机上以输入的多项式时间分解大数质因式, 由于分解大数质因式在经典算法复杂性理论中被认为是个“难解问题”, 现在广泛使用的公钥密码系统 RSA 就是以这个问题为基础的, Shor 算法的发现, 使量子计算机研究有了实用背景, 因此也获得了新的推动力。因此, 1994 年以后, 量子计算机的研究出现了迅猛发展的势头, 并从最初仅是学术上感兴趣的对象, 变成了对计算机科学、密码学、通讯技术以及国家安全和商业应用都有潜在重大影响的领域。

### 2.5.2 量子计算机的物理原理和模型结构

任何计算装置都是一个物理系统。量子计算机是根据物理系统的量子力学性质、规律执行计算任务的机器。这里信息以量子态叠加、干涉的形式存贮, 计算过程就是对初始态按照算法要求执行的一系列么正演化, 最后测量末态输出计算结果。

已经提出的量子计算机模型之一是量子门组网络模型, 其中要处理的信息载体是量子位, 而信息处理则由量子门实现, 量子门组网络则是由量子门通过“线路”连结构成。

在量子计算机中, 充当信息存储单元的物理介质是一个双态量子系统, 比如一个双能级系统或具有两个自旋态的电子、质子等。为了叙述方便, 通常把自旋  $1/2$  的电子作为量子位。量子位与经典位不同就在于它可以同时处在两个量子态的叠加态之中, 而若干个量子位的有序集合构成一个量子寄存器。一个二位量子寄存器, 有 4 个独立的态矢, 张起四维矢量空间, 寄存器可以存在这 4 个态的叠加态中, 因此可以同时包含这 4 个态的信息。类似地, 三位量子寄存器的态矢张起 8 维矢量空间。对量子位进行的最简单的么正操作称为基本量子门, 任何复杂的么正操作都可通过量子门的“通用集合”组合实现, 而一个量子门的“通用集合”可以由少数几个基本量子门组成。由于可以制备经典上不同态的叠加态, 量子计算机对叠加态的演化, 就是对其中各个叠加成分的演化, 即同时沿着经典上互不相同的路径计算, 这就是所谓的量子并行性, 量子计算机具有超出经典计算机信息处理能力, 就源于它的这种高度并行计算。

### 2.5.3 量子计算机的实现及困难

量子计算机的网络模型把作用到多个量子位上复杂的么正操作归结为最基本的么正操作, 这些基本么正操作可以由通用量子门组来实现。因此, 量子计算机的实现首先必须解决量子位和量子门组的物理实现问题。

迄今为止已经提出的实现技术有“离子阱技术”、“核磁共振技术”、“高光学腔技术”等<sup>[62~66]</sup>。已经发现用激光束操纵冷冻在其空阱中的离子串有希望实现少量量子位数百个逻辑门操作。这种技术的主要困难是离子串的冷却和由于噪声引起的“脱散”。而量子计算机实现中最主要的困难是量子计算机系统和环境的相互作用引起计算态的“脱散”问题。“脱散”使计算机量子态衰变为经典态, 而丢失掉计算中的量子信息。一方面量子计算机依赖于量子位之间的关联态叠加而具有大规模并行计算能力, 另一方面量子位和环境之间的关联又破坏了量子计算的进行。

1996年, Shor、Steane建立了量子纠错的理论和方法, Shor、Kitaev等人还提出了量子容错计算的思想, 即只要量子位和运算操作中的出错率低于某一阈值, 就可以实现任意精度的量子计算。这些理论和技术, 使人看到了克服“脱散”困难, 最后实现量子计算机的新希望。现在可以说制造量子计算机在原理上已经没有根本性的困难了, 但是真正造出一个实际运行的称得上计算机的设备, 还需要解决许多技术上的难题。

## 2.6 小结

总之, 量子理论与信息论、计算机科学的结合产生了量子计算, 量子计算机系统态的叠加、干涉和纠缠等量子力学性质为计算提供了新的物理资源, 这种资源的开发和应用正引起信息理论和计算机科学的变革。

本章的论述为后文提供了一些预备知识, 如不同的量子观点、量子态的特性、量子计算的本质以及量子算法的结构等等, 同时, 上述讨论也为研究量子神经计算模式、建立量子神经计算模型提供了素材和借鉴。

## 第三章 神经计算与量子计算的结合

自从 1982 年 Hopfield 神经网络提出以来, 人工神经网络得到了很大的发展, 并在各个学科领域进行广泛的应用研究, 但从本质上讲一直没有较大的突破, 随着实际问题的日益涌现, 神经计算技术面临着很大的挑战。根据学科发展的一般规律可以推知, 与新的理论相结合, 形成全新的交叉学科就有可能产生新的飞跃。近年来, 国际上出现了将神经网络与量子理论相结合的实践, 从理论上分析, 这种结合有着很大的潜力和良好的前景, 对它的深入研究将有助于理解人脑和意识的本质, 也有助于理解人工神经网络理论和量子理论本身, 更主要的是它开辟了一种新的思路, 能够产生一种全新的计算模式——量子神经计算。

将神经计算与量子计算结合起来的基本动机主要有三: (1) 有关量子思维的讨论, 用量子的眼光来重新审视人脑思维, 会得到一些全新的结论, 也有一定的合理性, 目前这一讨论在认知学界广为盛行; (2) 由于量子计算与传统计算存在着质的不同, 它包含有许多特殊的性质, 是目前科学界研究的热点之一, 将它引入神经计算可能会产生一些意想不到的结果; (3) 随着计算机硬件的发展, 当集成电路的线宽继续缩小下去, Moore 定律将会失效, 门和连线仅包含几个原子, 量子效应自然就会出现并且发挥重要的作用, 那时就不得不涉及到量子计算, 同样道理, 神经计算的量子推广也是其发展的必然趋势, 因为二者之间存在着某些本质上的联系。

本章将首先指出传统神经计算的主要缺陷, 接着讨论量子思维, 然后从理论上分析神经计算与量子计算结合的可能性, 并给出量子系统中 Hebb 学习规则的初步推导, 最后从总体上阐述神经计算量子推广的基本方法。

### 3.1 传统神经计算的主要缺陷

传统的神经计算通过对人脑工作机理的简单模仿, 建立在简化的神经元模型和学习规则的基础之上, 但由此却产生了许多计算上的优势, 如并行、分布式信息处理方式, 以及从输入到输出的非线性映射, 并且这些特点都拥有了许多非常成功的应用。但是, 随着研究的深入和普及, 特别是待处理信息量的急速膨胀和信息类型的多样化, 神经计算的局限与不足逐渐显现出来, 这主要表现在:

- (1) 传统意义上的学习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) 神经网络的记忆容量有限;
- (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现象, 等等。

这些本质上的缺陷使得人们对传统神经网络理论的进一步发展提出强烈的要求, 于是便出现了神经网络理论与其它理论相结合的尝试, 以期待产生新的计算模式。

## 3.2 量子思维

### 3.2.1 量子理论与人脑意识

“物质决定意识, 意识反作用于物质”, 这一著名论断构成了现代唯物辩证法的基础, 这里所说的“反作用”可以解释为人类有意识地去改造客观世界的活动, 那么进一步地, 意识改造客观世界的作用机理是什么呢? 也就是说, 人脑是如何产生意识而意识又是如何控制人体的行动呢?

目前, 许多科学家都坚信现代生物学关于人脑的结构和功能的观点, 即人的大脑是由非常多的(约  $10^{11}$  个)神经元构成的, 神经元之间依靠能够不断自动生成的树突和轴突来通讯, 信号从一个神经元传到另一个神经元时必须越过树突与轴突之间很窄的缝隙(约五千分之一毫米), 即突触间隙。一个神经信号就会使突触内外的离子浓度(神经元内外有配比不同的  $K^+$ 、 $Na^+$ )发生变化, 并由此使另一突触内外的离子浓度发生变化, 从而产生电脉冲传给另一个神经元<sup>[67,68]</sup>。因此, 神经传导的一个重要特点是其信号完全是“全有或全无”的现象, 即信号强度不变化, 要么有要么没有。从某种意义上讲, 人脑可以看作是一台经典意义上的复杂的电脑, 其中意识就好比是软件, 大脑就是硬件, 意识作用于大脑的过程就如同软件作用于硬件的过程。由此自然就会认为, 不管头脑的“线路细节”如何, 似乎总是可以用电脑的功能来模拟? 但是事实却并非如此。

我们知道, 只要外界条件不变, 电脑是根据程序一步步执行人们赋予的命令的。对电脑本身(包括硬件和软件)而言, 有几个确定的输入就会有几个确定的输出, 这种确定性不会改变。但对于人脑, 如果我们能够清楚地知道人脑的一切细节, 那我们就能预测出人脑的输出吗? 电脑只不过是一台自动机而已, 但我们绝不会认为自己是一台自动机, 因为我们能感受到自我意识, 而且能够作出决断。由此可见, 意识对人脑的作用过程绝不能简单地类比为软件对硬件的作用过程。

当然也许有人会认为，人脑非常复杂，它的行为与电脑不同，而且意识仅仅是人脑的机能，不能脱离人脑而单独讨论。这个观点仍局限于传统的生物学观点，按照这一观点，假如从理论上我们完全了解了大脑的结构，那么就能预测出其行为，这样一来，大脑似乎就成了具有确定性的东西了，那又如何能容得下我们自己的自由意志呢？

### 3.2.2 量子效应在人脑意识中的作用

量子理论也许能更好地解释人类意识的本质。在量子理论中，一切都是不确定的。这里所说的“不确定”，与经典意义中的“随机性”有着本质上的区别。在经典理论中，“随机性”是由于观测手段的精确度以及观测者所掌握的物理规律的限制而产生的。例如，抛一枚硬币，它落地后是正面朝上还是背面朝上是“随机”的。但在量子理论中，“不确定性”不是受观测手段等外界因素限制的，而是由物质的本性—波粒二象性所决定的，即不能同时精确确定一些物理量（如两个非共轭算符描述的物理量）从而准确测定客体的量子行为。考虑了量子效应后，我们就能跳出以往“决定论”的怪圈。

我们可以猜测，人类意识的不确定性与量子效应有关，所谓的“灵感”、“顿悟”可能就是由量子的不确定性产生的。由于人脑非常复杂，而且又构成了一个紧密的“网络”，所以哪怕是一点很微小的扰动，也能产生巨大的影响。前面说过，如果把人脑比作经典电脑，它是不可能产生自我意识的。那么，是否就是量子效应才使我们有了自我意识，才能感觉到自我的存在，并能对事物作出有意识的反应呢？

在量子理论中，有一个著名的“薛定谔猫”的佯谬：想象在一个封闭的容器（没有任何向内或向外的影响能通过容器壁）中有一只猫和一台能被放射性原子衰变触发的仪器。如果仪器被触发而打碎装有氰化物的瓶子，猫即被毒死，如果该事件没发生，则猫将继续活着。如果把这只猫描述为一个量子系统，那么按量子理论，当没有观察者去看容器时，整个系统应处于混合状态，即猫应处于“死活”的叠加态：

$$|\text{猫}\rangle = 1/\sqrt{2}|\text{活}\rangle + 1/\sqrt{2}|\text{死}\rangle \quad (3-1)$$

一旦有观察者去看时，猫的态矢量就坍缩成其中的一种选择：要么是活的，要么是死的。但是猫自己能感觉到自己处于“死活”的叠加态吗？如果把猫换成人，而他事后还活着，他会说曾经有一段时间他自己处于“死活”的叠加态吗？

显然不会。

从这个例子中，我们已能大致领略到量子效应对人类意识的作用过程。我们不妨再回到人脑与电脑的关系上来。事实上，人脑的不同部分具有独立运算的功能，但这仅仅是生物上的“并行”。也许我们可以猜测：人类思维过程可能与量子并行有关。按照科学家的估计，如果要制造出一台和人脑具有相同功能的经典电脑的话，这台电脑大约有一座摩天大楼那么高，它的耗电量相当于一座中型发电厂的发电量，而如果用量子电脑的并行计算，则体积可以大大缩小。那么，如果意识的状态在某种形式上和量子态相同，则人脑就可以和量子电脑类比，而且这远比与经典电脑类比合适。由此可以设想：如果能够模仿人脑制造出一台量子电脑，也许能够使它具有人类的意识，从而就能够揭示意识的本质。目前，人们已经从理论上尝试着利用量子并行计算来模拟人脑的某些功能（如学习功能），并取得了一定突破。

也许有人会认为人脑的功能并不是由于量子并行，而是由于神经元的并行，因为人脑中的神经元有  $10^{11}$  个，完全可以进行复杂的并行处理。这种看法也可能是正确的，但不能否认的是，神经元中的量子并行也是可能的。有一个典型的例子，一种叫大头金龟子的生物，它的身体内只有几个神经元，在它每天从洞里出来觅食的过程中，可以利用太阳光的偏振记录下自己所走的距离以及每次转弯的角度，然后经过复杂的三角积分运算，可以在返回时取最短的路线。按传统的生物学观点可以推知，仅仅几个神经元是不可能实现如此复杂的并行计算的，而在大头金龟子身上确实出现了这个“奇迹”，或许是因为它的神经元本身就是一台量子计算机。

### 3.2.3 对量子效应奇异性的初步探索

量子效应表现出种种违背人类常识的现象，并导致了很多悖论。量子力学的奠基人之一狄拉克曾指出：“……非常可能，从现在的量子力学到将来的相对论量子力学的修改，会象从玻尔轨道理论到目前量子力学的那种修改一样激烈。”也许在量子理论不断完善的过程中，关于人类意识的本质的问题将得到圆满的解决。

目前在认知学界，有许多有关量子思维（或称量子意识、量子大脑）的讨论，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就曾指出，连接量子现象和广义相对论的新式物理学可以解释一些脑力活动，如理解、感知、意识等，而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<sup>[19]</sup>，曾有这样一个生理实验，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,因此大脑中存在着量子力学效应的可能性依然可以维持;1994 年美国 Arizona 大学的 Hameroff 教授指出,上述方法不再研究由神经元所组成的网络,而是着重分析细胞内部的细微结构,如神经元内骨骼支架的微管,其中在神经元内细胞支架的微管之中或周围,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出来的<sup>[21]</sup>;1996 年斯洛文尼亚 Ljubljana 大学的 Perus 博士认为,量子波函数的坍缩十分类似于人脑记忆中的神经模式重构现象<sup>[22]</sup>,等等。虽然目前神经科学界尚无法确认人脑中是否确实存在有量子效应,但是,用量子理论来解释大脑现象的确富有创见和一定的合理性。

### 3.3 神经计算与量子计算结合的可能性分析

#### 3.3.1 量子系统与人脑信息处理

人工神经网络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),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之一;(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模拟人脑的信息处理过程。

#### 3.3.2 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存在着许多相似的地方,这里,我们仅从信息处理的角度出发,来分析这些相似关系,并给出其中隐含着的深刻的物理意义<sup>[69, 70]</sup>。

##### (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  $q(\bar{r}, t)$ ，神经模式为  $\bar{v}$ ，则神经态矢可表示为神经模式的叠加，即

$$q(\bar{r}, t) = \sum_{k=1}^P c_k(t) v_k(\bar{r}) \quad (3-2)$$

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

$$\psi(\bar{r}, t) = \sum_{k=1}^P c_k(t) \phi_k(\bar{r}) \quad (3-3)$$

其中， $\phi_k(\bar{r})$  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：

$$c_k(t) = \langle \bar{v}_k, \bar{q} \rangle = \iint v_k(\bar{r})^* q(\bar{r}, t) d\bar{r} dt \quad (3-4)$$

$$c_k(t) = \langle \phi_k, \psi \rangle = \iint \phi_k(\bar{r})^* \psi(\bar{r}, t) d\bar{r} dt \quad (3-5)$$

## (2) 时空积分

神经信号的时空积分为

$$q(\bar{r}_2, t_2) = \iint J(\bar{r}_1, \bar{r}_2, t_1, t_2) q(\bar{r}_1, t_1) d\bar{r}_1 dt_1 \quad (3-6)$$

其中  $J$  是单个突触连接的强度。

而在非相对论量子力学中同样有

$$\psi(\bar{r}_2, t_2) = \iint G(\bar{r}_1, \bar{r}_2, t_1, t_2) \psi(\bar{r}_1, t_1) d\bar{r}_1 dt_1 \quad (3-7)$$

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。

## (3) Green 函数与模式相关编码

我们还可以将上面两个方程 (3-6)、(3-7) 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

$$J(\vec{r}_1, \vec{r}_2, t_1, t_2) = \sum_{k=1}^P v_k(\vec{r}_1, t_1) v_k(\vec{r}_2, t_2) \quad (3-8)$$

Green 函数作为单个量子模式  $\varphi_k$  的相关系数, 有

$$G(\vec{r}_1, \vec{r}_2) = \sum_{k=1}^P \varphi_k^*(\vec{r}_1) \varphi_k(\vec{r}_2) \quad (3-9)$$

在相对论中 G 可转变为

$$S(\vec{r}_1, t_1, \vec{r}_2, t_2) = -i \sum_{k=1}^P \sum_{j=1}^2 \varphi_k^j(\vec{r}_1, t_1) \varphi_k^j(\vec{r}_2, t_2) \quad \text{when}(t_2 > t_1) \quad (3-10)$$

$$S(\vec{r}_1, t_1, \vec{r}_2, t_2) = i \sum_{k=1}^P \sum_{j=3}^4 \varphi_k^j(\vec{r}_1, t_1) \varphi_k^j(\vec{r}_2, t_2) \quad \text{when}(t_2 < t_1) \quad (3-11)$$

显然, 上述两个方程 (3-8) 与 (3-9) 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。

#### (4) 神经模式重构与量子坍缩

神经模式重构可表示为

$$q(\vec{r}, t) = \sum_{k=1}^P c_k(t) v_k(\vec{r}) \Rightarrow q(\vec{r}, t_0) = v_{k_0}(\vec{r}) \quad (3-12)$$

而波函数坍缩可表示为

$$\psi(\vec{r}, t) = \sum_{k=1}^P c_k(t) \varphi_k(\vec{r}) \Rightarrow \psi(\vec{r}, t_0) = \varphi_{k_0}(\vec{r}) \quad (3-13)$$

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。

#### (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

$$\dot{q}(\vec{r}, t) = \sum_{k=1}^P \lambda_k c_k(t) v_k(\vec{r}) \quad (3-14)$$

$$\dot{\psi}(\bar{r}, t) = -\frac{i}{\hbar} \sum_{k=1}^p E_k c_k(t) \varphi_k(\bar{r}, t) \quad (3-15)$$

其中  $\lambda_k$  为神经模式  $\bar{v}_k$  的注意度参数, 即矩阵  $J$  的对应于本征矢量  $\bar{v}_k$  的本征值; 而  $E_k$  具有与  $\lambda_k$  类似的作用。

### 3.3.3 基于量子系统的 Hebb 学习规则<sup>[71]</sup>

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

$$y = f\left(\sum_{j=1}^n \omega_j x_j\right) \quad (3-16)$$

其中 Hebb 学习规则可以表示为

$$\omega_j(t+1) = \omega_j(t) + \eta(d - y)x_j \quad (3-17)$$

式中  $\eta$  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

$$|y\rangle = \hat{F} \sum_{j=1}^n \omega_j |x_j\rangle \quad (3-18)$$

其中  $\hat{F}$  是一个未知算子, 它可由量子神经网络来实现。这时, 学习过程就转变为从初态到终态的演化过程, 即  $|y\rangle_n \Rightarrow |d\rangle_n$ , 而

$$|d\rangle = \hat{F} \sum_{j=1}^n \tilde{\omega}_j |x_j\rangle \quad (3-19)$$

于是调节权值就变成寻找某个合适的集合  $\{\tilde{\omega}_j\}$ , 因此它可以设定为一个量子

算子, 即

$$H = H_d + H_y = E \left\{ \sum_{j=1}^n |d_j\rangle\langle d_j| + \sum_{j=1}^n |y_j\rangle\langle y_j| \right\} \quad (3-20)$$

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  $|d_i\rangle = (U_{H_d+H_y})^t |y_i\rangle$ , 则

$$\tilde{\omega}_i^j = \sum_{k=1}^n (U_{H_d+H_y})_i^k \omega_k^j \quad (3-21)$$

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为

$\omega_i^j = |y_i\rangle\langle x^j|$ ,  $\tilde{\omega}_i^j = |d_i\rangle\langle x^j|$ , 可得 Hebb 学习规则为

$$\delta\omega_i^j = \eta(|d_i\rangle - |y_i\rangle)\langle x^j| = -iE \sum_k (H_d + H_y)_i^k |y_k\rangle\langle x^j| = -iE(|d_i\rangle\langle d|y\rangle + |y_i\rangle)\langle x^j| \quad (3-22)$$

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

$$\delta\omega_i^j = -iE(|d_i\rangle a + |y_i\rangle)\langle x^j| \quad (3-23)$$

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

$$|y(t)\rangle = \sum_{j=1}^n \omega_j(t) |x_j\rangle \quad (3-24)$$

学习规则变为

$$\omega_j(t+1) = \omega_j(t) + \eta(|d\rangle - |y(t)\rangle)\langle x_j| \quad (3-25)$$

于是得

$$\| |d\rangle - |y(t+1)\rangle \|^2 = \left\| |d\rangle - \sum_{j=1}^n \omega_j(t+1) |x_j\rangle \right\|^2 = (1 - n\eta)^2 \| |d\rangle - |y(t)\rangle \|^2 \quad (3-26)$$

可见, 当  $0 < \eta < \frac{1}{n}$  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 (3-22) 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<sup>[72]</sup>。

### 3.3.4 量子点系统实现神经计算的实例分析

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<sup>[73, 74]</sup>, 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<sup>[75]</sup>, 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 $|\pm\rangle$ , 在方程 (3-27) 中就是位置变量 $x(t)$ 。我们将每种状态作为一个量子比特, 如图 3-1 所示<sup>[76]</sup>。

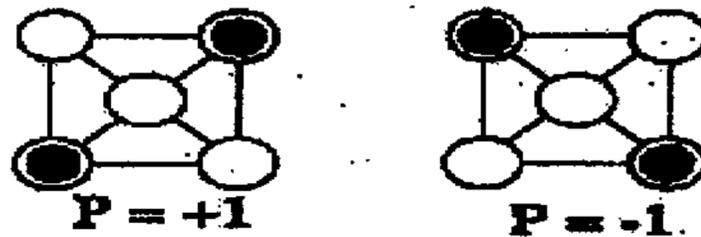


图 3-1 五原子量子点系统

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<sup>[77]</sup>形式将量子点系统的状态演化过程表示为

$$\begin{aligned}
 |\Psi(x_T, T)\rangle &= G(x_T, T; x_0, 0) |\Psi(x_0, 0)\rangle \\
 &= \int_{(x_0, 0)}^{(x_T, T)} D[x(t)] \exp\left(\frac{i}{\hbar} \int_0^T d\tau \left[ \frac{1}{2} m \dot{x}^2 - V(x) \right]\right) |\Psi(x_0, 0)\rangle \\
 &= \lim_{N \rightarrow \infty} \int_{x_1, 0}^{(x_{T-1}, T)} dx_1 \dots dx_N \left( \frac{m}{2\pi i \hbar \Delta t} \right)^{(N-1)/2} \exp\left(\frac{i \Delta t}{\hbar} \sum_{j=0}^{N-1} \left[ \frac{m}{2} \left( \frac{x_{j+1} - x_j}{\Delta t} \right)^2 - V(x_j) \right]\right) |\Psi(x_0, 0)\rangle
 \end{aligned} \tag{3-27}$$

其中,  $|\Psi(x_0, 0)\rangle$  是输入态, 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; 公式第二行是

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；第三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  $j\Delta t$  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），于是方程 (3-27) 就可简化为

$$|\Psi(\sigma_s(N\Delta t), T)\rangle = \sum_{\{\sigma_s(j\Delta t)\}} \exp\left(\frac{i}{\hbar} \sum_j [K\sigma_s(j\Delta t) + \epsilon(j\Delta t)\sigma_s(j\Delta t)] I(\sigma_s(t))\right) |\Psi(\sigma_s(0), 0)\rangle \quad (3-28)$$

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

进一步展开 (3-28) 式，就可以得到以下三个可调整的参数：(1) 耦合强度  $\{\lambda_s\}$ ；(2) 每个时间片段  $j$  时刻的电场强度值  $\{\epsilon_j\}$ ；(3) 振荡频率  $\{\omega_s\}$ 。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

$$Output = Th(|\langle + | \Psi(\sigma_s(N\Delta t), T) \rangle|^2) \quad (3-29)$$

其中  $Th()$  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。

## 3.4 神经计算的量子推广

### 3.4.1 概述

将神经计算推广到量子领域，首要的任务是建立神经计算概念在量子领域的对照物<sup>[39]</sup>，如表 3-1 所示。具体的解释如下：

表 3-1 神经计算概念及其量子对照物

量子对照物	神经计算概念
波函数	神经元
量子态	输入、输出
态叠加（相干）	连接权
测量（消相干或坍缩）	趋向吸引子的演化
态纠缠	学习规则
幺正变换	增益函数

- (1) 波函数与神经元：将神经计算的基本单元（神经元）转变为量子波函数，它具有复数的形式，通过处理可以使它同样满足神经元简化模型的条件，比如说具有阈值、激活函数等，即  $f(g(W) + \theta) \Rightarrow \psi(\phi(|W\rangle) + |\theta\rangle)$ ，不过此时量子神经元拥有了更丰富的含义和性质，从而可能简化网络结构；
- (2) 量子态与神经网络的输入、输出：在量子系统中可以将网络的输入和输出看成是量子态的初态和终态；
- (3) 态叠加与连接权：连接权是神经网络实现学习和计算的关键部件，也是信息分布式存储的宿体，在神经计算中它们可以通过学习进行调整、修改，在量子领域中我们可以使用量子态的叠加形式来表示权值，即

$\vec{w} = (w_1, w_2, \dots, w_n) \Rightarrow |\vec{w}\rangle = |w_1 w_2 \dots w_n\rangle$ , 这样就可以借助量子的超并行性实现更为快捷的计算;

- (4) 结果测量与趋向吸引子的演化: 量子态的测量过程是一个非么正变换过程, 它表示结果的输出和量子态演化的终止, 这就类似于人工神经网络中某个状态趋向其吸引子的演化过程, 如 Hopfield 神经网络中趋向某个稳态或伪稳态;
- (5) 态纠缠与学习规则: 神经网络通过学习规则对连接权实施调整, 而在量子计算过程中计算的进行和中间结果的传递都要利用态纠缠来实现, 即  $U_f: |x\rangle|y\rangle \rightarrow |x\rangle|f(x)\rangle$ , 从而体现计算的效率<sup>[78]</sup>;
- (6) 么正变换与增益函数: 在量子计算中计算过程就是量子态转化的过程, 它是么正性 (即线性) 的, 结合测量的非么正过程来实现计算, 这就相当于神经计算中的增益函数。

当然, 上述的联系并非一成不变, 但建立如此的对应关系是构成量子神经计算的前提, 并且合理的定义更有利于计算模型和算法的设计与实现。

### 3.4.2 量子推广的方法

有了上述的对应关系, 接下来就是如何在此基础上进行推广, 以建立合理、有效的量子神经计算模型 (即量子神经网络)。通过比较分析两种计算模式的动力学过程, 可以发现, 量子推广的关键问题在于如何解决神经计算的非线性映射和量子态的么正 (线性) 变换之间的矛盾, 只有调和了这个矛盾之后, 才能够实现量子神经计算。

以下是我们提出的几种不同的解决方案:

- (1) 依据 Copenhagen 学派的观点 (波粒二象性, Wave-Particle Duality), 它认为微观粒子同时具有波动性和粒子性, 其状态用波函数来表示, 它满足 Schrödinger 方程。在量子系统中的演化算子是么正的, 但在结果的测量时 (即坍缩) 所发生的消相干却是一个非么正过程, 可以把它视作量子系统趋向某个吸引子的非线性演化。

据此, 最直接的方法是利用波函数表示, 将传统神经计算引入量子系统, 同时设计量子学习算法和非么正 (测量) 算子, 从而实现输入空间到输出空间的非

线性映射，其具体实现过程如图 3-2 所示。

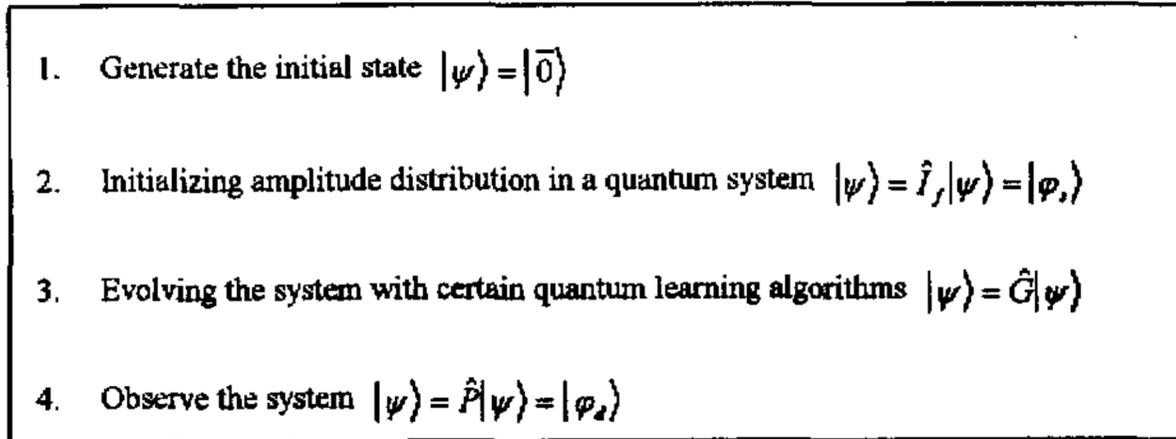


图 3-2 量子系统实现神经计算过程示意图

其中，算子  $\hat{I}_r$ 、 $\hat{G}$ 、 $\hat{P}$  依次是初始化算子、演化算子和测量算子，量子态  $|\varphi_s\rangle$  和  $|\varphi_e\rangle$  分别为初态和终态。另外，需要特别指出的是，我们在这里借用了计算机程序的表述习惯，等式从左到右描述了量子态的演化过程，因此等号两边符号相同的量子态并非表示同一个量子态。上述过程中存在的关键问题是初态的制备和量子学习算法的设计。

在本文的第六章中将利用上述的理论框架来设计量子竞争学习算法和相应的量子算子。

(2) 可以利用著名的双缝干涉实验来实现神经计算，参见图 3-3。

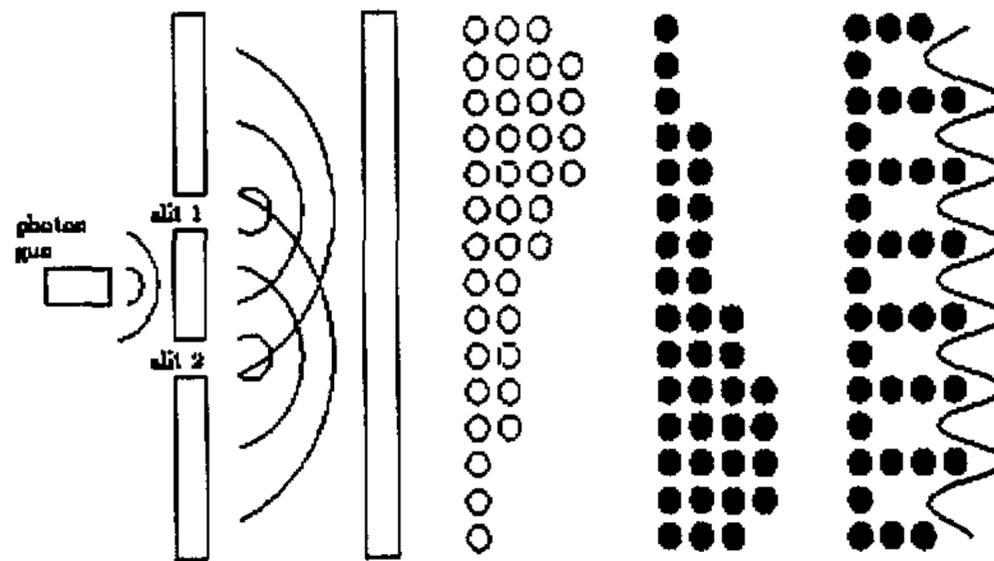


图 3-3 双缝干涉实验示意图

上图中从左至右依次为：光子枪（量子发射源）、双缝光栅、接收屏、只开启缝 1 时所得到的接收条纹、只开启缝 2 时所得到的接收条纹以及双缝同时开启时所得到的干涉条纹，我们注意到，双缝干涉条纹并非前面两个条纹的简单叠加（即

并非波的叠加), 而是发生了量子干涉现象, 这充分体现了光(量)子的波粒二象性。由于该实验模型的结构与前馈神经网络的拓扑结构类似, 因此有许多研究者据此构造出相应的量子神经网络, 如 Chrisley 的非叠加态模型<sup>[26]</sup>, 但是他们所提出的模型都缺乏数值计算的可行性, 在实际应用中难以实现。

我们进一步地考察上述结构, 若选择干涉条纹的适当特征位置作为网络的输出, 便可以得到非常有趣的结论: 例如, 取干涉条纹的中心点下一行(即条纹增强处)作为输出, 它表明当两个输入值同为 0 时, 输出为 0, 当单个输入值为 1 或两个同为 1 时, 输出为 1(令强度阈值为两个点), 即完成“逻辑或”运算; 同样, 若取中心点(即条纹完全抵消处)作为输出, 它表明当两个输入值同为 1 或 0 时, 输出为 0, 当单个输入值为 1 时, 输出为 1(同样令强度阈值为两个点), 即完成“逻辑异或(XOR)”运算, 更有意义的是, 上述“逻辑异或”运算的完成仅仅使用了单层网络结构, 这是传统单层前馈神经网络所不能实现的。

下面的问题是如何选择合适的物理参量, 使得基于双缝干涉实验结构的量子神经计算模型具有实际的可计算性。本文将在第五章提出一个全新的计算模型, 它从设计输入量着手, 以介质的介电常数矢量作为网络的输入, 以粒子实际运动轨迹作为输出, 从而形成多狭缝-多栅板的拓扑结构。该模型可以方便地进行数值计算, 我们利用该模型的单层结构实现了类似于 XOR 的逻辑运算。

(3) 利用 Feynman 路径积分<sup>[77]</sup>的时、空模型公式, 因为在其公式中, 即

$$|\psi(t)\rangle = \sum_{\text{all-paths}} e^{i \int_0^t \left[ \frac{m\dot{x}^2(\tau)}{2} - V(x(\tau)) \right] d\tau} \quad (3-30)$$

体现了非线性, 这是由于式中存在势函数  $V(x)$  和指数函数的缘故。

依此观点, 在量子神经网络中, 信息是沿着由许多支路所组成的相互纠结的宽信道传递的。假设有一个量子神经网络, 它含有  $N$  个量子神经元, 相互连接的突触权值在 0 到 1 之间, 为了简化模型, 我们假定该神经元是二值的, 并且在任一时刻某个神经元要么处于激活状态要么处于抑制状态。考察  $t$  时刻, 令  $q(l, t) = 1$  表示第  $l$  个量子神经元在  $t$  时刻被激活, 下一步, 第  $l$  个神经元将把信息传递给与之相连的另一个神经元, 虽然与该神经元存在有多个连接神经元, 但激活神经元将只选择其中一个, 假设为第  $J$  个, 选择概率为它们之间的连接权值。

下面简要讨论一下如何求解上述概率。假设信息到达第  $J$  个神经元的时刻为

$t+r$ , 于是概率可由公式  $P[q(J, t+r)=1, q(I, t)=1]$  给出, 在经典物理学中使用概率论的标准规则就能计算出来, 而在量子物理中则需利用 Feynman 路径积分来计算, 即概率取决于  $t$  时刻到  $t+r$  时刻从第  $I$  个神经元到第  $J$  个神经元所有可能的路径的和。令每条路径  $x$  具有的能量为  $H(x)$ , 则 Feynman 和  $F[q(J, t+r)=1|q(I, t)=1]$  为所有的  $\exp[\frac{i\hbar}{H(x)}]$  之和, 其中  $\hbar$  为 Planck 常数, 那么可得选择概率为

$$P[q(J, t+r)=1, q(I, t)=1] = \frac{F[q(J, t+r)=1|q(I, t)=1]}{F[q(J, t+r)=1|q(I, t)=1] + F[q(J, t+r)=0|q(I, t)=1]} \quad (3-31)$$

(4) 利用 Everett 的多宇宙观点, 它认为波函数始终满足 Schrödinger 方程, 测量的结果是将观测值分为互不察觉的若干部分, 而每个部分 (即宇宙) 只可能有一个结果。

此观点特别适用于模式识别, 以感知机为例<sup>[30, 43]</sup>, 假设用一个感知机只记忆一种模式, 则有多少种模式就有多少个感知机, 与传统的方法 (即仅用单个感知机记忆多种模式) 相比, 它有如下优点: 由于模式之间不发生干涉, 因此训练速度快, 泛化能力强, 无伪记忆模式。

此法的关键是设计合适的坍塌算法, 来实现测量时多宇宙的坍塌。Menneer 等人提出的量子衍生神经网络在模拟测量时只是简单采用模式匹配的方法, 忽视了多宇宙之间可能存在的纠缠。本文将在第四章重新设计多宇宙模型, 以 BP 网络为例, 引入 Grover 量子搜索策略, 针对不同的求解对象设计出相应的坍塌算法。

(5) 利用量子门组的硬件特征, 设计基于门组结构的神经网络。由于量子计算可以通过量子门组来实现, 因此量子门中实际上就蕴涵了量子计算的一些特征, 如量子并行性, 若能将这些量子门组合起来构成神经网络, 那么待量子门制作技术成熟之时, 就极易用硬件来实现量子神经计算。本文将在第五章讨论基于量子通用门组单元的神神经计算模型。

### 3.5 小结

本章首先分析了传统神经计算所存在的一些问题, 然后从认知科学、神经科学及计算科学的角度讨论了神经计算与量子计算之间的对应关系以及二者结合的可能性, 提出了基于量子系统的 Hebb 学习规则, 并结合量子点系统实例讨论了如何实现神经计算。

虽然目前对于量子神经计算的界定尚无公论，文中就作者的理解，提出了几种可能实现的途径，紧接着，本文随后的章节将详细讨论这些实现的途径和方法，提出几种量子神经计算模型及其相应的学习算法。

## 第四章 量子神经计算模型 (1)

### —— 基于量子理论基本原理的量子神经计算模型

量子神经计算模型是建立在神经计算与量子计算的结合之上, 依据上一章所讨论的量子推广方法, 我们将这种结合分为两大类: 一类是利用量子理论的基本原理和观点, 从较根本的层次来建立量子神经网络模型; 另一类则是借鉴量子理论或量子计算中的某些概念、实验、效应等来设计新的模型、算法。本章所讨论的两个模型分别建立于量子理论波动力学原理和量子理论多宇宙观点的基础之上, 试图构造出基本的量子神经元, 以及用于模式分类的量子神经网络通用模型。

#### 4.1 量子神经元模型 (Quantum Neuron Model)

量子计算是利用微观量子级效应来执行计算任务, 它已在某些领域取得了比经典计算有指数级提高的效果。我们知道, 在经典人工神经网络模型中选择最佳的权值组合是实现神经计算的实质, 但这是一项相当费时的工作。于是, 我们就会很自然地设想: 能否构建量子意义上的神经元——即量子神经元模型, 利用量子计算的超并行性来求解针对某个问题的神经网络的权值组合?

接下来, 我们借助文献[28]的研究方法, 在量子理论的层面上研究经典神经网络的意义。首先构造一个经典神经网络的量子对照物, 然后对其量子态权值的性质进行分析, 结合实例讨论求解该量子态的方法及其对应的经典含义, 最后将此问题进一步加以推广。

由于在经典人工神经网络模型中最简单的是感知机, 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  $W = (\omega_1, \omega_2, \dots, \omega_n)^T$ , 阈值  $\theta$  和激活函数

$$f = \begin{cases} 1, & \text{if } \sum_{j=1}^n \omega_j i_j \geq \theta \\ -1, & \text{otherwise} \end{cases} \quad (4-1)$$

该模型无法解决线性不可分问题, 但鉴于它形式和性质的简单, 我们首先就研究它的量子对照物, 构建其相应的量子神经元模型。

在新的量子神经元模型中, 我们同样取  $\{i_j\}$  作为输入, 单个连接权值被一个波函数  $\psi(\omega, i)$  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值, 称为经典权矢,  $\psi(\omega, t)$  代表了权矢空间中所有可能权值的概率幅度 (广义上为复数), 且在任意时刻  $t$  满足归一化条件, 即

$$\int_{-\infty}^{\infty} |\psi|^2 d\omega = 1 \quad (4-2)$$

于是, 经典感知机的权值被许多权矢的量子叠加态所代替, 但是一旦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, 其坍缩概率为  $|\psi|^2$ 。

既然我们已经将权值的概念融合到波函数之中, 那么求解神经网络的权值组合就等同于求解量子态方程的本征波函数。我们知道, 在量子力学中波函数是由 Schrödinger 方程所决定, 即

$$\nabla^2 \psi = \frac{2m}{\hbar^2} [U - E] \psi \quad (4-3)$$

简单地, 假设这里方程与时间无关, 那么其中可改变的项只有势函数  $U$ , 所以  $\psi$  可由改变  $U$  而改变。

最直接的想法是: 将经典神经网络求解的问题转化为一定形式的 Schrödinger 方程, 然后通过解该方程来计算出波函数  $\psi$  的值, 再将波函数化解为一个一个的权值组合, 这样就完成了经典神经网络的训练 (学习) 过程。但是, 就目前我们的理解, 这种算法存在以下几个难点: (1) 现有量子理论所支持的 Schrödinger 方程是一个线性方程, 它无法实现神经网络的非线性映射; (2) 量子么正算子也是线性的, 在非线性引入之前尚无法设计合适的么正矩阵; (3) 即使量子方程含有非线性因素, 对于复杂的神经网络学习问题也很难构造出相应的方程及训练算法。因此, 现阶段我们所作的是将上述问题简化, 即在已知经典权矢的基础上, 考虑如何求解它们的线性组合, 而现有的 Schrödinger 方程可以很方便地用来计算此值。于是, 我们本节的重点就放在如何将计算结果与经典神经网络意义的沟通上, 以及如何合理地解释求解量子方程所得到的结果上。

如上所述, 选取 Schrödinger 方程中合适的势函数  $U$ , 就可以求解相应的波函数  $\psi$ 。在已知可解的势函数中有刚性箱体模型, 还有非刚性箱体、简谐振荡器、氢原子模型等等<sup>[79]</sup>, 选取不同的势函数就会得到不同的解的形式。这里, 为了讨论的方便, 我们选取表达形式简单的无限深势阱的势函数, 其势函数可以表示为

$$U = \begin{cases} 0, & |x| < a \\ \infty, & |x| \geq a \end{cases} \quad (4-4)$$

其中,  $a$  是势阱的宽度。

现在考察一个单输入、输出函数, 它用来对输入值求反 (NOT)。如果我们将经典权值限定在  $-a \leq \omega_j \leq a$  区间之内, 再假定  $\psi$  与时间无关, 只与  $\omega$  有关, 那么波函数  $\psi(\omega, t)$  就蜕变为—维矢量, 这样考虑到权值的限制, 通过神经元的学习, 找到相应的  $\psi(\omega, t)$  就等价于量子力学中求解—维无限深势阱问题, 由量子力学的知识可知, 其解的形式为:

$$\psi(\omega_0) = A \sin\left(\frac{n\pi}{a} \omega_0\right) \quad (4-5)$$

其中  $A$  为归—化常数, 可由 (4-2) 式求得;  $n=1, 2, 3, \dots$ , 为量子数;  $\omega_0$  为  $\psi$  的单个元素, 即经典权矢; 通常势阱宽度  $a$  取  $2\pi$ 。

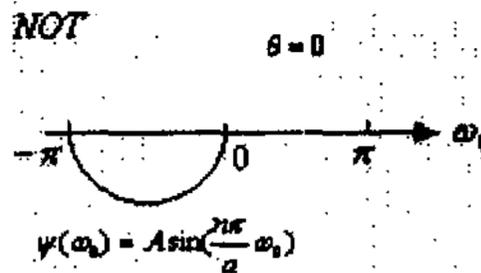


图 4-1 逻辑求反函数的量子权值分布

上图所示当阈值  $\theta=0$ , 量子数  $n=2$  时函数 NOT 的权值分布, 显然当  $\omega_0 = -\frac{\pi}{2}$  时解的概率最大。由量子方程得出的结果说明了经典感知机权值分布的有效性, 不过此处由于是单个输入、输出, 每次只有一个权值, 因此不存在权值的组合问题。

下面再来考察较为复杂的二值输入 XOR 问题。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无限深势阱问题, 其解的形式为:

$$\psi(\omega_0, \omega_1) = A \sin\left(\frac{n_0\pi}{a} \omega_0\right) \sin\left(\frac{n_1\pi}{a} \omega_1\right) \quad (4-6)$$

其中变量和常量的含义与 (4-5) 式相同, 注意这里有两个不同的量子数  $n$ , 每一

个对应于一个不同的经典权值。

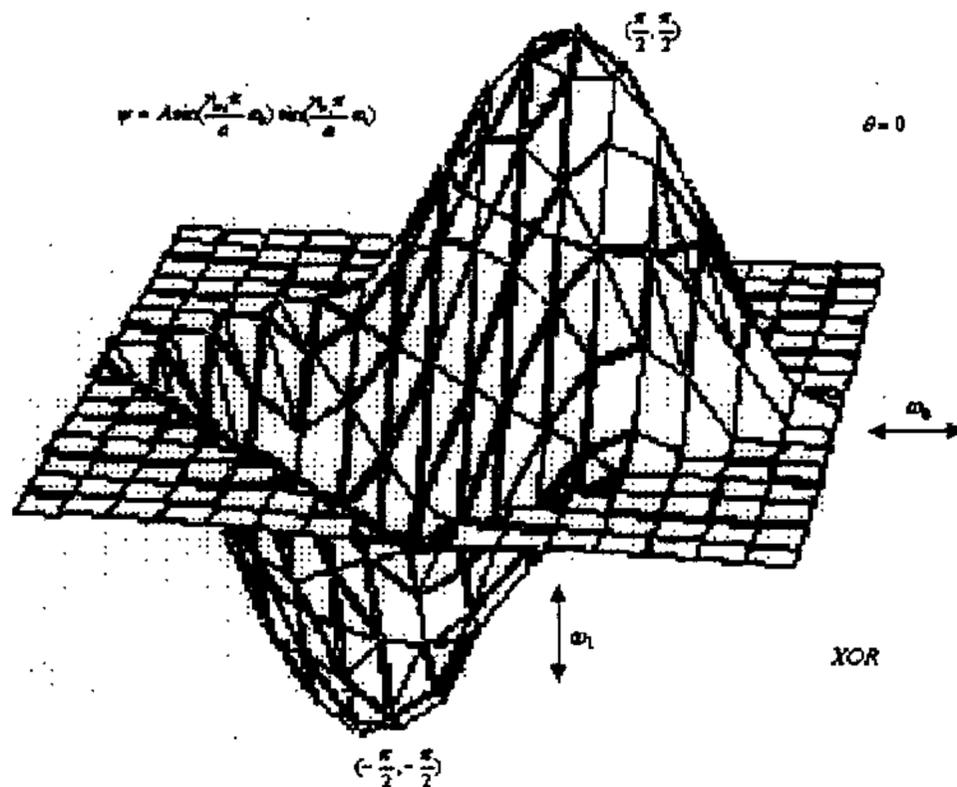


图 4-2 逻辑异或函数的量子权值分布

同样，我们可以画出二维波函数的图形，由上图可见，当权矢组合取  $(\frac{\pi}{2}, \frac{\pi}{2})$  和  $(-\frac{\pi}{2}, -\frac{\pi}{2})$  时解的概率最大，而处于反对角线上的权矢组合近似无解。

对于 XOR 问题，当输入为 (1, -1) 和 (-1, 1) 时可以为上述任意一个权矢组合所满足，但对于输入为 (1, 1) 和 (-1, -1) 而言却并非如此，于是我们可以设计合适的量子算子，当输入为 (1, 1) 时使得对于正值权矢所对应的概率幅度减少，对于负值权矢所对应的概率幅度增加，同理，当输入为 (-1, -1) 时情况正好相反，这样就可以有效地保证计算的正确性。因此，经过量子化改造，再加上量子态叠加的特性，理论上说单个量子神经元就可以实现 XOR 学习。

当然，上述例子和相应的解都比较简单。我们还可以考虑将其换成其它物理模型，并考虑时间变量的影响，那么解的形式就要发生相应的变化，通过分析这些公式我们便可以深入理解量子神经元所包含的物理意义。

## 4.2 多宇宙量子神经网络

### 4.2.1 概述

在量子理论中有三个重要的概念，即：态叠加、态纠缠和坍缩，而在设计和

构造量子神经网络中用到的也往往就是这三个概念。利用量子态的叠加现象可以使经典问题的并行处理成为可能；同时，叠加现象本身又和坍缩紧紧联系在一起，因为为了获取叠加中的信息，量子态必须发生坍缩；而纠缠现象又和叠加现象紧密相关。因此，总的来说，量子叠加现象就成为量子理论诸多方面中的一个最主要的环节。

正如第二章和第三章中所讨论的那样，量子理论中有一种多宇宙观点<sup>[50]</sup>，该观点认为存在有一些平行的完全的宇宙，它们在某种确定的意义上在相同的时间和空间中存在着，就某种意义而言，它们与我们共享着同样的时间和同样的空间。不同的平行宇宙作为一个公共的物理实体的一部分而关联在一起，物理实在就是纠缠在一起的所有宇宙的集合。在理论上，多平行宇宙是与波函数的各个投影分支相联系的，当宇宙面临一种量子选择时，它就分裂成两个不同的宇宙。该理论提供了一个完全决定性的理论，考虑到环境引发解耦时，多宇宙观点的实验预测将等同于正统统计解释的预测。可见，多宇宙观点强调了量子态的叠加性，并由此产生了并行的量子信息处理机制，特别适合于解决模式分类之类的问题。

同时，从某种程度上讲，人工神经网络的激活过程可以看作是各种模式的叠加，虽然神经网络并没有具体地包含多种状态，但由于其在训练过程中连接权值的改变响应于不同的输入模式，一个新输入模式的激活会受到不同模式的混合影响，因此，所有模式的学习都将对网络的训练产生影响，也就可以称其为所有输入模式的叠加。

于是，我们会很自然地联想到利用量子态的叠加现象将多宇宙观点引入到神经网络的结构设计之中。实际中，我们应用一个叠加状态的网络，使用现成的学习规则，将网络的输入视为一个微粒，网络部分作为一个实体，每组权重作为一个状态，这样就产生了一个微粒和一个实体相互作用从而进入一个特殊状态的量子过程。为了使神经网络的过程具有量子效应，最关键的问题就是如何构造态叠加。

本节的目标是通过模拟实验的方法来寻找最有效的网络结构以及网络训练和测试时坍缩的方法，并给出量子神经网络的结果和相应的经典神经网络的结果，用以评价不同模型的计算性能。

#### 4.2.2 量子神经网络的体系设计

量子神经网络的体系结构必须根据规则和用系统的方法来处理。如同各种类型的神经网络一样，量子神经网络也是由两个步骤组成的，即训练和测试。在一

些模拟的过程中, 这些训练的方法会有所变化。在网络的测试阶段包含了其坍塌过程。不同坍塌方法的选取取决于量子神经网络的体系结构和训练方法。

在对于整个量子神经网络进行设计时, 我们首先要有以下关键性的设定:

### C (1) 网络中叠加现象的组成

量子的叠加现象可以在一个人工神经网络的许多部分得以体现: 神经元、神经元之间的连接以及整个神经网络。由于网络本身是由神经元及神经元之间的连接组成的, 每一条连接就是权重的一个叠加, 所以网络的叠加可以是权重集合的叠加, 也就相当于所有连接的叠加。而为了达到单个神经元的叠加状态必须也要通过改变神经元的输入和输出, 使一个神经元区别于其他神经元的方法就是赋予其不同的权值。由上可知, 这三种方式在实践中效果是等同的, 都可以归结为权重的叠加。

C 因此, 为了形成一个量子神经网络的结构体系必须使用权重的叠加, 输入本身处于一个叠加状态, 它们会坍塌到不同的宇宙完全取决于用来处理它们的权值。这样, 我们可以构造不同的子网络, 每个子网络是同质的, 但是却由不同的权值组成, 各个子网络组合在一起形成整个量子神经网络的叠加状态。

### (2) 每个子网络仅有一个训练模式

C 选择一个网络只对应一个训练模式的原因有很多, 其中最主要的一个原因是: 一个网络拥有一个训练模式, 可使我们使用的方法真实地影响到最终的结果, 同时, 仅有一个模式, 也可以使模式之间的学习不会相互干涉。因此, 学习是迅速的、无隐藏的。另一个原因是, 由于在经典的神经网络训练中, 一个模式的学习记忆会被另一个模式的学习所削弱, 所以, 使用一个模式学习要快速的多。此外, 由于没有其它模式的干扰和替代, 训练模式可以被有效地存贮在部分网络中。这样, 从某种程度上来说, 特殊的例子将被存贮起来。而经典的神经网络却无法存贮一个固定的输入。

### (3) 网络是否在重叠中训练

C 子网络可以作为一个叠加现象的各个状态来建立, 也可以作为多个经典网络来建立。如果这些网络是作为一个叠加存在的, 它们可以并行训练, 但叠加现象必须在整个训练过程中保持, 此种网络称为叠加训练子网络, 即训练是通过叠加现象来完成的。如果网络的叠加发生坍塌, 这些子网络的信息将会丢失。对于另一种训练方式来说, 子网络为经典网络, 在训练中叠加现象不会并行产生, 网络可

以存储信息, 这种训练方式产生的网络称为存储性子网络。

#### (4) 坍缩的方法

量子神经网络的坍缩有两种方法, 一种是特殊性坍缩, 即对于每一个测试输入来说, 叠加态都会坍缩到一个子网络上, 不同的测试输入会坍缩到特定的子网络上。在这种情况下, 剩余的子网络都将在坍缩过程中丢失。我们可使用上述的存储性子网络来实现对整个子网络的信息保存, 使之在下一次测试中仍然有效。

另一种方法是一般性坍缩, 即对于所有的测试模式, 坍缩到的子网络都是固定的。在这种情况下, 坍缩到的子网络对于以后的所有输入测试都将有效, 所以不需保存所有子网络的信息。因此, 可以使用上述的叠加训练子网络, 而无需在测试时发生的坍缩现象会遗失子网络的信息。

本文中采用特殊性坍缩方法。

#### (5) 处于叠加状态的量子神经网络发生坍缩的时间

因为网络处于叠加状态中, 我们可能会应用所有的子网络来处理一个输入, 所以在一个真正的量子神经网络中需要和使用一个单独的组建网络相同的处理时间。每一个子网络在处理输入时其概率会得到演化, 因此坍缩发生在模式被处理以后。

### 4.2.3 网络的训练

从上文的讨论中可知, 各子网络可以有两种训练方式: 一是子网络作为整个的网络叠加进行训练, 二是子网络单独作为经典网络进行训练, 在训练结束后组合在一起构成叠加。具体的选择哪种方式训练将取决于测试的坍缩方式。

量子神经网络是由子网络的一系列权值构成, 每个连接都对应着一组叠加的权值, 同一组的权值以某种方式相互纠缠, 当处于叠加状态的一个连接坍缩到某一子网络的权值上时, 所有的连接都会坍缩到该子网络的权值上。当然, 对于多层量子神经网络, 衡量的标准可能会变为权值的子部分, 在那种情况下, 纠缠仅仅发生在一个量子部分。

### 4.2.4 网络的测试

量子神经网络的坍缩是指权值的每个叠加状态退化为一个单独的确定的权值。如上所述, 权值的叠加状态相互纠缠从而使坍缩发生在一个子网络的一组权

值上。当然，在一般性坍缩实验中，通过对叠加中的权值可以计算出一组新的权值。

下面给出我们所设计的坍缩测试的主要思想，如图 4-3 所示。在训练过程中，训练模式“11”与实线权值相关联，训练模式“01”则与虚线权值相关联，测试时，将这两个训练模式分别和测试模式进行比较，当与实线权值相关联的训练模式与测试模式最接近时，实线权值的相位将会得到相应调整，使得坍缩到这个宇宙的概率增加。

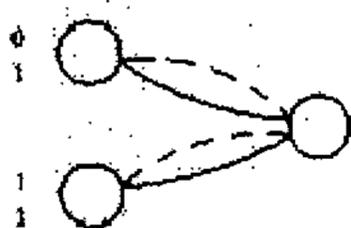


图 4-3 两个模式“01”和“11”分别  
以虚线和实线权值进行训练

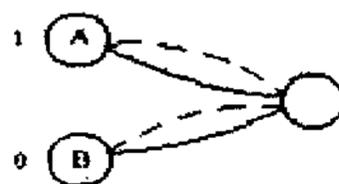


图 4-4 测试输入时神经元 A 和 B  
的输入值分别为“1”和“0”

具体地，如图 4-4 所示，当第一个输入 Qubit 为 1 时，输入神经元 A 可能会具有两个状态，一个是与训练模式的 Qubit“1”相关联，另一个是和训练模式的 Qubit“0”相关联，将测试 Qubit 和每个宇宙中的训练 Qubit 进行比较，在实线的宇宙中，测试 Qubit 和训练 Qubit 一致，所以其相位进行调整，此过程类似于 Grover 量子搜索算法<sup>[13]</sup>。而在虚线的宇宙中，其相位保持不变。同理，当第二个训练 Qubit 为“0”时，神经元 B 的两个状态都与 Qubit“1”相关联，由于测试 Qubit“0”与两个宇宙中所训练的 Qubit 都不同，因此虚线和实线宇宙的相位或者说坍缩的概率都将保持不变。最后，测试输入将会坍缩到概率大的一个宇宙中。

#### 4.2.5 量子神经网络的量子过程分析

量子神经网络的叠加形成以后，每个测试模式都与某个子网络的权值对应起来，每个子网络有其自身的相关训练模式，它们相互纠缠，处于同一个宇宙之中。测试模式与训练模式的相似度评估将会使模式相同的某个子网络的相位得到相应增加，因此处于该子网络的坍缩概率值也有所增加。

(1) 叠加：一个叠加或者说是波函数描述了一个实体所处的状态，典型的叠加会包含有该物体多于一个的状态值。在量子神经网络中的叠加即为网络的叠加，或者说是系列权值的叠加，这种叠加描述了网络连接中可能的权值。

(2) 坍缩: 坍缩是对于一个叠加或者说波函数而言的, 即波函数所描述的实体的状态值退化到了某一个状态值。在一个量子神经网络中, 权值的叠加在测试中发生坍缩, 各种可能的子网络最后都会归结为由某一个网络为输入模式所提供的输出结果。

(3) 纠缠: 纠缠发生在同一个宇宙内的权值之中, 或者说发生在那些取自同一个子网络中的权值之间。这种纠缠确保了权值的坍缩会集中在同一个子网络里。纠缠也同样会发生在训练时输入模式和参与训练此模式的网络之间, 这种纠缠保证了比较训练输入模式和测试模式后, 一个恰当的子网络将会提供对应测试输入的输出。

#### 4.2.6 仿真实验的设计和实验方法

真正的量子神经网络依赖于量子技术的实现, 因为它需要在一个量子计算机上来实现, 使用的是量子寄存器, 从而产生真正意义上的量子态叠加。由于目前的硬件条件不允许, 因此我们的实验研究仅仅是独立于量子技术的经典仿真。仿真所运用的是量子理论, 它可以独立于一个真正的量子体系存在, 在经典计算机上实现。当然这种基于量子理论的概率的特点不是很好实现, 但模拟和实际的量子神经网络有充分的相似性, 以确保其将来可被扩展到量子硬件上完成。

运行仿真实验的一个突出优点是可以进一步研究量子神经网络是如何运作的。通过模拟我们可以探讨的问题有: 量子神经网络和经典神经网络的特征性差别、算法问题的新型解决方法、量子过程的调查, 以及设计和实现量子算法。

##### (1) 概率因子的模拟

在一个量子神经网络中, 坍缩的最终状态是由叠加态中的概率因子所决定的, 为了使量子神经网络的实验能够在经典的硬件设施上实现, 概率因子的模拟是一个重要环节。

当子网络以概率因子来选择时, 需要有大量的输入测试集合才能实际的观测到整个量子神经网络是如何运作的。每一次输入测试 Qubit 通过网络就代表了一个测试模式, 但是当每个测试模式仅一次从网络中通过时, 某个测试模式可能会坍缩到一个概率较小而不是概率最大的子网络, 这样其处理的结果就会产生偏差。

为了使实验更完整、全面, 时间的限制无法忽视。我们不可能有足够的时间对所有实验中的所有网络都进行大量的输入测试来确定每个子网络的坍缩概率。在这些实验中, 我们设定所有的叠加状态都会以其最大的概率坍缩到某一子

网络上,这样就使得整个过程变得比较确定,因此我们就以系数代替了概率。

如果有两个子网络都有最大概率时,我们就将此量子网络的坍缩直接定义为最先接触到那个网络,这样通过每次选择第一个发现的网络,我们可以有效地在较短的时间内选择一个最好的子网络,避免了随机选择所花费的时间。

### (2) 坍缩的模拟

如前所述,为了计算出各个子网络的概率值,坍缩会发生在所有的子网络对测试模式进行处理之后。但在实际实验中,量子神经网络只能通过一个顺序的结构体系来模拟,这样势必会增加每个子网络的处理时间。因此,仅使用一个子网络来处理测试模式将会使效率提高。我们可以事先计算好每个子网络的概率,然后将最有可能的子网络选出,来直接处理测试模式,这样的处理仍然是和量子理论一致,因为即使所有的子网络都用来处理输入,结果也只能从某一个宇宙中获得。

### (3) 结构体系的选择

对于量子神经网络结构体系的选择与经典神经网络大致相似,这取决于实验的具体内容,即如果训练模式的目标输出已知,可采用监督型学习体系,选用反馈网络就比较合适;如果训练模式的目标输出未知,则可采用非监督型学习体系,此时可选用竞争学习网络。

量子神经网络的结构体系与相应的经典网络的区别在于:前者有多份整个网络的拷贝,或者是网络一部分的拷贝,每一份拷贝都以适合该网络体系结构的学习规则来训练一个模式,而网络中可能有一部分为经典网络,另一部分为量子网络,网络的经典部分(非重复部分)用于处理所有的模式,这些部分的激活来自于所有的其前的重复部分,而其权值的改变又取决于所有其后的重复部分的反馈。

### (4) 实验环境设定

实际仿真过程中,我们将给出一些规定来确定一个量子神经网络所需要的神经元的数目,这些规定是针对一个子网络部分的神经元数目的,即一个子网络部分的神经元是恒定的。这些规定也同样适用于那些并没有应用网络部分的多份拷贝仅仅是经典的网络部分,换言之,无论这个神经元是否为一个量子子部分,其所使用的神经元的数目是相同的,只不过是若为量子的部分,每个子部分都会有一份相同数目的神经元。

量子神经网络的输入神经元的数目由输入模式值的长度来决定,输出神经元

数目则由输出模式决定,这一点与经典神经网络类似。对于一个经典的神经网络,并没有什么明确的规则来确定隐层神经元的数目,一般的经验表明,隐层的神经元数目不应当超过输入模式的节点数,数目过多训练时间会增加,网络解决问题的能力有所下降,还会导致重复运算。反之,隐层的神经元数目太少时,会使模式的学习无法正常的进行,需要训练的次数多,训练精度也不高。这些规则同样实用于量子神经网络。因此,对于一些小型的简单例子,经典神经网络和量子神经网络中的隐层的神经元数目保持一致,而在一些较大型复杂的例子里,量子神经网络中所需的隐层神经元数目要少于经典网络。

#### (5) 网络参数设定

量子神经网络的学习率因子与经典网络一致,通常设为 0.1,测试的误差范围也与经典网络保持一致,当输出值为“0”或“1”时,误差定为 0.5,取于两者之间,而训练时允许的误差范围要比经典网络略低,经典网络取 0.1,量子网络则取 0.05。

### 4.2.7 实验结果及训练环境

神经网络的实验结果包含有两个值:一个是模式训练完成的时间,我们用训练循环的次数来衡量,量子神经网络的训练时间是所有的子网络的训练时间的总和;另一个是运行精度,定义为给出正确的输出的测试模式与测试模式总集的比值。这里,测试模式的集合中不再包括已经参加训练的模式。

对于每一个网络类型,都会在不同的训练环境下重复做多个网络训练,这样才能保证结果没有随机性。在下面的实验中,对于每一个网络,我们会更改训练集的规模大小及其在全集中的选取(测试集的规模及选取与之互补),以及网络的权值在初始化时的随机选择,以此来构成网络的训练环境。例如,全集中共有 16 个模式,训练集合分别选取 4、5、6、7 个模式,即 4 个不同规模,而同一个规模下,我们在 16 个模式中选取四个不同训练模式组合,并在每次的训练集合训练时,取 3 次不同的权值初始化,这样就构成了  $4 \times 4 \times 3$  个训练环境,使网络在 48 个不同的环境下进行训练。在某些实验中,由于时间的原因,选取的环境值可能或多或少,这都因实验而异。实验的两个结果取网络在所有的环境下训练结果的平均值。

### 4.2.8 其它规定

通常情况下,网络的学习都会在输出误差处于允许的误差范围之内后正常结

束, 但当某些特殊情况发生时, 我们也作了如下规定:

#### (1) 网络总的学习次数

有时对于一个训练集, 网络不能有效的学习, 这里有很多的原因, 例如没有足够的隐层节点数去容纳信息, 或者在训练集中出现了不协调的训练模式, 这时, 网络往往无法达到较小的误差将其收敛。在这些实验中, 我们对网络的最大学习次数作了一个硬性的规定。当然, 对于不同的实验, 由于实验的特性和复杂度不同, 学习次数的限制也会有所改变。当网络学习达到最大次数时, 学习停止。

#### (2) 误差恒定

另一个决定一个网络是否继续学习的方法就是记录在训练中目标输出和实际输出之间的误差不再减少的次数。如果网络的误差最小值无法超越的次数达到了一个事先给定值时, 即网络已不能将学习效果进一步提高, 则停止学习。当然这个给定值的设定不能过小, 因为网络学习在某个局部阶段其误差变化不大, 会有所波动, 甚至呈上升趋势, 但总的趋势仍然是下降的, 若给定值过小, 往往会使网络不能达到最佳的学习效果。我们一般把此值定为最大训练次数的十分之一。

### 4.2.9 设计小结

在经典神经网络设计中, 一个神经网络的设计由任务的特性而定。我们为这个任务选择一个有最大可能的效果的网络来作为参照的经典神经网络, 输入和输出的神经元数目由该输入输出模式的长度决定, 如果输出神经元的个数较多, 可采用某种编码方式, 或者在输出层后再加一神经元层, 将输出神经元的激活转化为后一层的输出。对于输入神经元也可利用同样的处理方法。

针对不同的任务, 隐层神经元的数目会有所变动, 我们用实验的方法找出能达到最好的结果的神经元数目。设计量子神经网络与经典神经网络相对应, 取与经典神经网络有相同的体系结构, 且每一个子网络具有同样的神经元数目。但同时考虑到作为参照的经典神经网络的设计有更多的自由, 两者比较的可靠性也会因此而受到影响。对于较复杂的任务, 使用不同的隐层神经元的数目往往会达到更好的效果, 这样, 可以对经典网络和量子网络采用一些不同的体系结构, 以两者最佳的效果作为相互比较的双方。

### 4.2.10 单层反馈网络的实验模型

我们选用的是 BP 网络, 其激活函数为  $\text{tansig}$  或  $\text{purelin}$ , 每一个输入模式与子

网络一一对应。各子网络中选择一个来处理测试输入，它对应的训练模式与输入模式比较，模式最接近的那个获胜，即测试输入模式将坍缩到获胜子网络进行处理。

设共有  $x$  个子网络组成了整个网络的叠加，单层量子神经网络所采用的主要算法如下：

*BEGIN*

(1) 从整个模式集合中选取  $x$  个模式作为训练集合；

(2) for  $i=1$  to  $x$

对每个子网络进行初始化；

end;

(3) for  $i=1$  to  $x$

每个子网络用其对应的训练模式进行训练，直到输出误差达到允许的范围；

end;

(4) 计算  $x$  个子网络总的训练时间；

(5) 将剩下的模式作为测试模式；

利用特定的坍缩规则决定每一个测试模式将坍缩到的宇宙；

(6) 用坍缩到的子网络对测试模式进行测试，结果在误差允许范围内的作为有正确输出的模式；

计算识别率；

(7) 对于同一训练集，至(2)重复，将网络初始化；

(8) 至(1)选择不同的训练集重复至固定次数；

*END*

为了使实验排除一些干扰因素，我们多次重复实验，取平均结果作为最后的比较对象，使实验具有统计意义。

### 实验一

计数实验，取输入为 4 个 Qubit，对输入中“1”的个数进行计数，当“1”的个数大于 3 时，输出即为“1”，否则输出为“0”。在总的模式集合中共有  $2^4=16$  种模式，从中分别选取  $x$  ( $x=4, 5, 6, 7, 8, 9, 10$ ) 个模式作为训练模式，剩余的  $16-x$  个模式作为测试模式。当选定固定个数的训练模式后，网络重复作 3 次实验，每次重新赋不同的网络初值，这样就构成了整个网络的训练环境。实验结果主要记录两方面：一为识别率，即测试集中测试结果与实际结果的差别在误差允许范围之内的模式与总的测试模式的比例；二为训练时间，即整个网络训练所有的训练模式的次数。对于量子神经网络，由于条件的限制，无法作并行的运算，

所以量子网络的实现是将各子网络的训练串行进行的, 整个训练时间平均被每个子网络分配以作为与经典网络比较的对象。这正是利用了量子的并行性, 取得的良好效果。我们设定不同的训练环境, 也就是为了排除一些干扰因素, 使实验具有统计意义, 所以最后进行比较的结果为不同训练环境的平均结果。

当进行测试时, 经典网络仅有一个网络模型, 所有的测试集都经过这个模型得出结果, 而对于量子网络, 它由多个子网络构成, 每个子网络同质, 但各个权值完全不同。对于单层网络, 共有两组权值, 即输入层与输出层之间的权值和输出层神经元的偏差权值。当测试模式坍塌到一个网络的一组权值上时, 它必将坍塌到该网络的另一组权值上 (多层网络将有其特有的坍塌方式, 稍后讨论)。我们定义单层量子神经网络的坍塌规则为:

(1) 有  $x$  个训练模式  $P_i (i=1,2,\dots,x)$ , 分别与  $x$  个宇宙相关联, 对于测试模式  $P$ , 计算出  $M_i = |P - P_i|$ :

$M_c = \min(M_i)$ :

(2) 其中  $c$  即为测试模式坍塌到的宇宙下标;

(3) 若同时有多个  $M_i$  取到最小值, 规定取第一个能取到最小值的那个宇宙。

定义了训练规则和坍塌规则后, 利用 Matlab 工具箱编程分别实现两个网络, 其实验结果如下表所示:

表 4-1 单层网络实验一的训练环境一

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
4	12	20	3	0.1

表 4-2 单层网络实验一训练环境一下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	47.6	49.3
训练时间	146.7	36.8

表 4-3 单层网络实验一的训练环境二

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
5	11	20	3	0.1

表 4-4 单层网络实验一训练环境二下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	50.2	49.1
训练时间	173.7	38.6

表 4-5 单层网络实验一的训练环境三

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
6	10	20	3	0.1

表 4-6 单层网络实验一训练环境三下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	53.0	52.7
训练时间	196.8	37.5

表 4-7 单层网络实验一的训练环境四

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
7	9	20	3	0.1

表 4-8 单层网络实验一训练环境四下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	50.2	49.7
训练时间	220.9	40.9

表 4-9 单层网络实验一的训练环境五

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
8	8	20	3	0.1

表 4-10 单层网络实验一训练环境五下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	50.1	49.8
训练时间	248.2	39.7

表 4-11 单层网络实验一的训练环境六

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
9	7	20	3	0.1

表 4-12 单层网络实验一训练环境六下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	56.7	55.2
训练时间	297.1	42.0

表 4-13 单层网络实验一的平均训练环境

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
4-9	7-12	20	3	0.1

表 4-14 单层网络实验一的平均实验结果 (激活函数 tansig)

	经典神经网络	量子神经网络
识别率 (%)	51.3	50.2
训练时间	231.9	39.3

由以上的数据可以看出, 在识别率大致相同的情况下, 量子神经网络的训练时间大幅度降低, 这正体现出量子神经网络的优越性。同时, 我们又将激活函数 tansig 改为 purelin, 在相同的实验环境下, 得到的实验结果如下:

表 4-15 单层网络实验一的平均实验结果 (激活函数为 purelin)

	经典神经网络	量子神经网络
识别率 (%)	50.6	50.5
训练时间	92.0	19.6

## 实验二

实验中输入模式为 4-qubit 值, 输出模式为 1-qubit 值, 当输入模式中含奇数个“1”, 则输出为“1”, 否则为“0”。我们采用的坍塌方式仍然与上文所述一致, 网络结构也与上文保持一致。实验结果如下表:

表 4-16 单层网络实验二的训练环境一

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
4	12	20	3	0.05

表 4-17 单层网络实验二训练环境一下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	41.9	36.5
训练时间	276.2	49.6

表 4-18 单层网络实验二的训练环境二

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
5	11	20	3	0.1

表 4-19 单层网络实验二的训练环境二下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	41.7	30.9
训练时间	283.2	48.3

表 4-20 单层网络实验二的训练环境三

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
6	10	20	3	0.1

表 4-21 单层网络实验二训练环境三下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	40.3	27.5
训练时间	296.1	47.2

表 4-22 单层网络实验二的训练环境四

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
7	9	20	3	0.1

表 4-23 单层网络实验二的训练环境四下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	41.8	35.4
训练时间	312.4	52.3

表 4-24 单层网络实验二的训练环境五

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
8	8	20	3	0.1

表 4-25 单层网络实验二的训练环境五下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	40.2	26.5
训练时间	324.8	47.3

表 4-26 单层网络实验二的训练环境六

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
9	7	20	3	0.1

表 4-27 单层网络实验二训练环境六下的实验结果

	经典神经网络	量子神经网络
识别率 (%)	42.7	27.1
训练时间	339.4	48.7

表 4-28 单层网络实验二的平均实验结果

	经典神经网络	量子神经网络
识别率 (%)	41.5	28.4
训练时间	305.5	48.8

从实验结果来看, 显然这种坍塌方式并没有取得好的效果。虽然其训练时间有显著的减少, 但识别率与经典网络差别较大。因此我们提出了一种新的坍塌方式将其改进。

### 实验二的另一种坍塌方式

我们采用一种新的方法作为此例的坍塌规则:

(1) 共有  $x$  个训练模式  $P_i (i = 1, 2, \dots, x)$ , 分别与  $x$  个宇宙相关联, 对于测试模式  $P$ , 计算出  $M_i = |P - P_i|$ ;

(2) 计算出  $S_i = \begin{cases} 0 & \text{当 } P \text{ 与 } P_i \text{ 的距离差为偶数时;} \\ 1000 & \text{当 } P \text{ 与 } P_i \text{ 的距离差为奇数时} \end{cases}$ ;

(3)  $M_c = \min(M_i + S_i)$ ; 其中  $c$  即为测试模式坍塌到的宇宙下标;

(4) 若同时有多个  $M_c$  取到最小值, 规定取第一个能取到最小值的那个宇宙。

使用这种坍塌规则, 我们重新进行上述实验, 平均结果记录如下:

表 4-29 单层网络实验二平均实验环境

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
4-9	7-12	20	3	0.05

表 4-30 单层网络实验二在坍塌方式二下的平均实验结果

	经典神经网络	量子神经网络
识别率 (%)	41.5	67.7
训练时间	305.3	47.4

显然, 改进后的坍塌方式使得整个量子网络的性能有了显著的提高, 不仅训练时间上有明显差别, 在识别率上也更优于传统的神经网络。

### 实验三

在完成了上述实验的基础上, 我们再利用其它的一些实验来检验这个量子神

神经网络结构体系的特性。将输出的 Qubit 数从一个增加到多个, 但整个网络仍然为单层 BP 网络。

实验三描述为: 输入 Qubit 数为 3, 输出 Qubit 数为 3, 且输入等于输出。总的实验集合有  $2^3=8$  个模式。实验结果如下表:

表 4-31 单层网络实验三平均实验环境

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
3-5	3-5	20	3	0.05

表 4-32 单层网络实验三平均实验结果

	经典神经网络	量子神经网络
识别率 (%)	25.3	13.7
训练时间	599.4	151.6

#### 实验四

实验四描述为: 输入输出均为 3 个 Qubit, 输出为输入的倒置, 实验集合也为 8 个模式。实验结果如下表:

表 4-33 单层网络实验四平均实验环境

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
3-5	3-5	20	3	0.05

表 4-34 单层网络实验四的平均实验结果

	经典神经网络	量子神经网络
识别率 (%)	25.8	14.1
训练时间	640.8	154.7

从上面的实验和结果可知, 这种结构体系对于单层网络来说一般只适用于把实验结果分成两类的两分型实验模型, 同时, 对于不同的实验, 可以根据其本身固有的特性, 采用特定的坍塌规则, 这样往往能取得更好的效果。

#### 4.2.11 两层量子神经网络

两层神经网络同样运用了上述的基础理论, 与单层网络的设计大体上一致, 但由于存在隐层神经元的关系导致了一些与单层网络不同的地方, 以下我们对其特有的网络性质及一些设定作进一步的讨论。

对于一个两层的反馈网络来说共有 5 个主要组成部分: 输入层到隐层的权值、

隐层到输出层的权值、输入神经元、输出神经元和隐层神经元。输入神经元的函数通常简单、不加改变，只是把输入值传递到输入层和隐层之间的权值。如果只有一个输入，输入神经元将无法作任何变换，因此叠加现象无法产生。为了获得隐层或输出神经元的多宇宙现象，神经元的处理过程必须改变，唯一的用来改变一个反馈神经元的计算办法就是改变对此神经元的权值的偏差。因此，权值偏差的多重发生被应用于实现量子隐层或输出神经元的叠加状态。

在训练过程中，类似于单层网络那样，一个子网络对应于一个训练方式，每个子网络是由四个不同部分的权重组成的：即输入层到隐层之间的权重、隐层到输出层之间的权重、隐层的偏差权重和输出的偏差权重。在这里，我们用四个符号来代替以上四个不同部分的权值，分别记为：IH、HO、HB 和 OB。当处理一个输入时，每个部分都完全独立于其他部分。当叠加的量子神经网络形成时，每个子网络的部分都不会与网络形成纠缠。这也就意味着当一个量子部分坍缩到一个固定子网络的相应部分时，其他的量子部分并不需要也坍缩到相同子网络的相应部分。例如：一个测试模式可能会坍缩到某个子网络的 IH 部分和另一个子网络的 HO 部分，这样，网络的设计也就更加复杂了。

在网络中叠加状态就在 IH、HO、HB 及 OB 这四个部分中产生，由于每一个部分的权值都可以单独的作为叠加状态中的一部分而存在，所以在网络中经典部分和量子部分的组合将会有 16 种不同的情况。除了四个部分皆为经典部分的经典网络以外，其它的 15 种网络都可视为量子神经网络。图 4-5 给出了这 16 种类型网络的示意图。为了方便起见，我们将 IH、HO、HB 及 OB 四个部分的类型用两类字母标识，经典部分记作“c”(classical)，量子部分则记为“q”(quantum)，四个部分按 IH—HO—HB—OB 的顺序表示。这样网络类型“qqcc”即表示输入层到隐层和隐层到输出层之间的权值都为量子化的，而隐层的偏差权值和输出层的偏差权值为经典意义下的。

由于两层网络隐层的特殊性，我们除了采用单层网络所运用的那种坍缩方法以外，还设计了另一种坍缩方式：

(1) 共有  $x$  个训练模式  $P_i (i=1,2,\dots,x)$ ，分别与  $x$  个子网络即  $x$  个宇宙相关。隐层的激活值  $A_i$  的表达式为：

$$A_i = f(IH_i \cdot P_i + HB_i), \text{ 其中 } IH_i \text{ 和 } HB_i \text{ 为第 } i \text{ 个子网络的输入至隐层的权值和隐层的偏差权值。当前}$$

的测试模式为  $P$ ， $P$  对于各子网络的隐层激活  $a_i = f(IH_i \cdot P + HB_i)$ ；

(2) 取  $M_i = |A_i - a_i|$  为两者之间的距离差；

(3)  $M_c = \min(M_i)$ ; 其中  $c$  即为测试模式坍塌到的宇宙下标;

若同时有多个  $M_i$  取到最小值, 规定取第一个能取到最小值的那个宇宙。

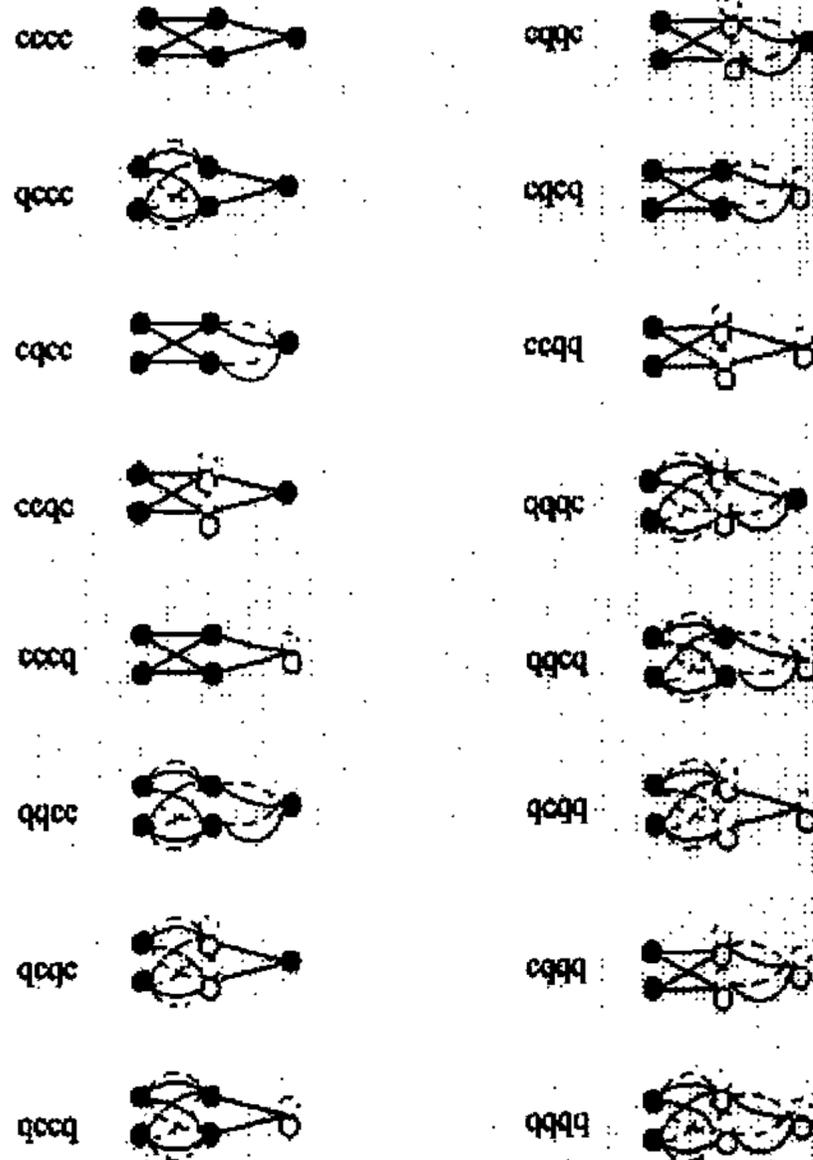


图 4-5 两层网络的 16 种网络类型

### 实验五

对于以上两种坍塌方式, 我们设计的实验五如下。实验描述为: 整数 2-15 用 4 位二进制数表示, 即输入 4 个 Qubit, 输出为 1 个 Qubit。若此整数为素数, 则输出为 1, 否则输出为 0。隐层节点取两个, 实验结果如下所示。

表 4-35 两层量子神经网络实验一训练环境

训练模式数	测试模式数	选取不同训练模式次数	赋不同初值数	训练误差
4	10	20	3	0.1

表 4-36 两层量子神经网络实验一实验结果 (使用坍塌方式一)

网络类型	识别率 (%)	训练时间
cccc	57.3	2473.2
cccq	57.7	666.8
ccqc	59.0	850.1
ccqq	59.0	502.7
cqcc	59.0	3615.7
cqqc	53.7	611.4
cqqq	56.3	597.3
qccc	60.0	449.0
qccq	51.3	1439.8
qcqc	58.3	2244.7
qcqq	56.3	747.2
qqcc	53.7	422.6
qqcq	57.3	1799.4
qqqc	54.0	549.7
qqqq	55.0	707.6
qqqq	57.7	412.3

表 4-37 两层量子神经网络实验一实验结果 (使用坍塌方式二)

网络类型	识别率 (%)	训练时间
cccc	57.3	2473.2
cccq	54.3	713.3
ccqc	52.7	835.5
ccqq	55.7	459.6
cqcc	54.0	1550.3
cqqc	53.7	660.7
cqqq	51.3	679.6
qccc	54.7	432.3
qccq	51.0	1686.5
qcqc	47.3	603.2
qcqq	50.3	702.1
qqcc	49.7	462.8
qqcq	55.0	1380.0
qqqc	50.3	624.8
qqqq	54.0	705.2
qqqq	49.3	419.3

从表 4-36 和表 4-37 中明显可以看出, 两层量子神经网络与经典神经网络结果相比较, 量子神经网络的费时更短, 并且在某些网络类型上, 识别率也更优于经典网络。

#### 4.2.12 总结及展望

本节所描述的量子神经网络是将经典的人工神经网络与量子理论中的多宇宙概念结合形成的，这是一个新型的网络结构。此外，我们通过设计一系列的仿真实验来检验这种结构与经典神经网络相比存在着哪些优缺点。

从对比结果来看，量子神经网络的确有着广阔的发展前途，不论是在单层网络中，还是在两层网络中，量子神经网络在保证识别率相近的前提下，比经典神经网络大幅度减少了训练时间，并使得量子的并行处理成为可能，甚至在某些类型的量子网络中，网络的识别率还要优于经典神经网络。今后，我们还可以把这种优势推广到多层网络，使之有更大的应用范围。

在上述实验中还发现，量子网络可以有多种坍缩规则，针对特定实验的特点，可采用特殊的坍缩规则，这样将会取得意想不到的效果。因此，我们在坍缩规则的选取方面，还可以作进一步的研究，除了可以实行硬竞争的方式，还可以尝试软竞争，即在训练的过程中不断的比较训练效果，动态地改变坍缩的几率，真正实现 Grover 量子算法的搜索策略。

目前，我们所处理的实验数据类型均为二进制数据，实验的类型也仅仅限于分类问题。今后可以将实验的类型加以扩展，把数据类型转变为多元化的，如整型、实型等以适应不同的实验要求。

## 第五章 量子神经计算模型 (2)

### —— 基于量子实验和量子门组概念的量子神经计算模型

本章将结合量子理论中的实验方法、量子门组概念来构造量子神经网络模型，同时还设计了一些仿真实验来评价模型的计算性能。这些模型与以往的模型相比，更具有实用价值。

## 5.1 基于量子双缝干涉实验的神经网络

### 5.1.1 双缝干涉实验

著名物理学家 Feynman 曾说过：“双缝实验包含了量子力学中的所有神秘”。

双缝干涉是量子理论中一个著名的实验，它能最直观地展现微观粒子的波粒二相性，同时它也是量子力学中最为离奇、在概念上最难理解的现象之一。该实验的示意图如图 5-1 所示，入射源是微观粒子（如电子、光子等），这里以光子为例。

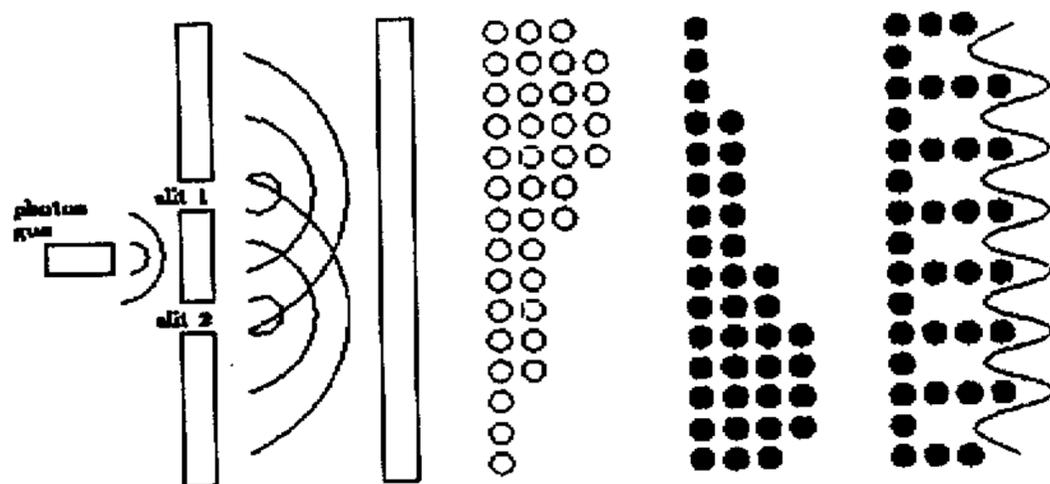


图 5-1 光子双缝衍射实验示意图

控制左侧的入射光子枪，使光子几乎一个一个地经过双缝，然后打在探测底板上，起初底板上出现一些点子的分布，随着时间的延长点子会越来越多，最后在底板上会形成一个有规律的衍射花纹（见图的最右侧）。若实验中只开启其中一个狭缝，则会得到中间两个对称的花纹，但当双缝同时开启时，最终的衍射花纹并非是它们的简单叠加，这就表明光子不仅具有粒子性，而且同时还具有波动性。

依据量子理论<sup>[80]</sup>，可以得到底板上点子的强度分布为

$$I(x, y) = |\psi(x, y)|^2 = \psi^*(x, y)\psi(x, y) = 4A^2 \cos^2\left(\frac{kd}{2D}x\right) \quad (5-1)$$

其中, 几率幅波函数为

$$\psi(x, y) = 2A \exp\left(ik\left(D + \frac{(d/2)^2 + x^2 + y^2}{2D}\right) \cos\left(\frac{kd}{2D}x\right)\right) \quad (5-2)$$

式中,  $A$  是光子源在底板上产生的振幅,  $D$  是栅板到底板的距离,  $k$  为波矢,  $d$  为双缝间距,  $x$ 、 $y$  是底板直角坐标轴。

### 5.1.2 Chrisley 非叠加型量子神经计算模型

鉴于上述实验装置的结构与前馈神经网络模型十分相似, Chrisley 提出了基于非叠加态的量子神经计算模型, 并给出了相应的学习算法<sup>[26]</sup>。Chrisley 模型如图 5-2 所示, 粒子束  $G$  入射栅板  $B$ , 之后是一个光敏接受盘  $P$ , 在栅板  $B$  上有许多狭缝, 其中上部分组成输入区, 下部分组成权值区, 调节狭缝的结构便可以在接受盘  $P$  上得到不同的干涉模式, 这样便实现了两组映射, 即: 一个是由输入 (如特征串、图象、查询等) 到栅板  $B$  上狭缝结构的映射 ( $I$ ), 另一个是由干涉模式到输出 (如分类、存储数据等) 的映射 ( $O$ )。

假设输入为  $\bar{x}$ , 输出为  $\bar{a}$ , 权值为  $\bar{w}$ , 干涉模式为  $\bar{p}$ , 这样可得  $S(\bar{x}, \bar{w}) = \bar{p}$ , 此处  $S(\bullet)$  为 sigmoid 函数, 而  $\bar{a} = O(S(\bar{x}, \bar{w}))$ 。再定义系统的误差函数为  $E = \sum (\bar{a} - \bar{a})^2$ , 其中  $\bar{a}$  为期望输出, 然后利用类似于经典 BP 算法进行训练, 理论上就能够学习任意的函数映射关系。于是, 利用上述神经网络结构便实现了量子学习。

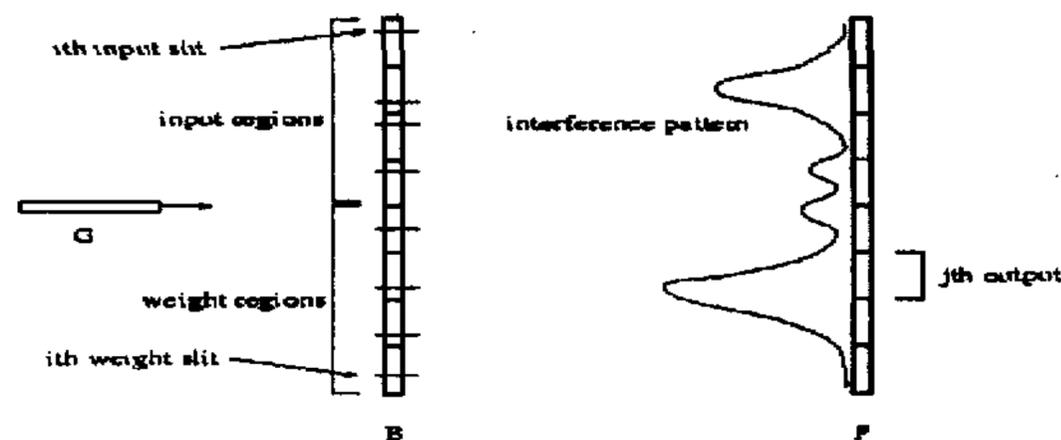


图 5-2 Chrisley 非叠加态量子神经计算模型

Chrisley 模型巧妙地将量子实验的结构与神经计算的方式结合起来, 改造了传统的神经网络模型, 并从中产生一些新的特点, 主要表现在: (1) 权值更新。它

不同于传统神经网络中每个权值仅调节与之相联的节点的连接关系, 在这里, 每改变权值区中的任何一个权值 (即权值区狭缝的结构), 将会影响到所有输入节点的连接; (2) 输出。该模型的输出是几率幅的叠加, 本质上是复值的, 它增加了网络的信息容量; (3) 物理实现。它具有很直观的物理意义, 易于物理系统实现。但是, 该模型也存在着以下几点问题: (1) 其前端未考虑量子态的叠加, 本意是为了避免相干态的维持问题, 但现实中量子态的叠加是客观存在的, 仅仅利用上述理论框架, 将会失去由量子态叠加而产生的计算优势, 如量子超并行、量子纠缠等等; (2) 虽然 Chrisley 给出了学习算法的大致框架, 但很难实现, 因为在该模型中很难设计出神经网络的量子对照物, 比如说输入量的表示, 由于粒子源是单一的, 它对于每个狭缝应该是等同的, 仅靠狭缝的不同位置不宜表征不同的输入量, 若采取粒子的入射角度或相位来表示, 就会导致计算复杂度的极大提高甚至无解。上述两点严重地制约了 Chrisley 模型的应用。

### 5.1.3 新的计算模型

我们所提出的新模型也是建立在双缝干涉实验基础之上, 并将它推广到多狭缝-多栅板结构。考虑到 Chrisley 模型的缺陷, 我们从网络的输入量着手, 通过设计量子实验来寻找合适的神经网络对照物。由量子理论可知, 当微观粒子进入不同介质后其运动轨迹会发生改变, 我们就选取不同介质的介电常数矢量  $\vec{n} = (n_1, n_2, \dots, n_m)$  作为网络的输入量 (其中  $m-1$  为栅板的个数), 以粒子穿越栅板到达探测器的实际运动轨迹  $Tr \langle S \rightarrow s_1^i \rightarrow \dots \rightarrow s_2^j \rightarrow \dots \rightarrow s_{m-1}^k \rightarrow D \rangle$  作为网络的权值 (实际上它最终由狭缝的位置  $s_j^i$  和栅板间距  $h_j$  所决定), 取探测器所测得点的强度 (即几率)  $P(\vec{n})$  作为网络的输出, 以此来构造一个全新的量子神经网络模型, 参见图 5-3。为了准确理解此模型的意义, 我们必须将其几何构形与神经网络上的含义区分开来。

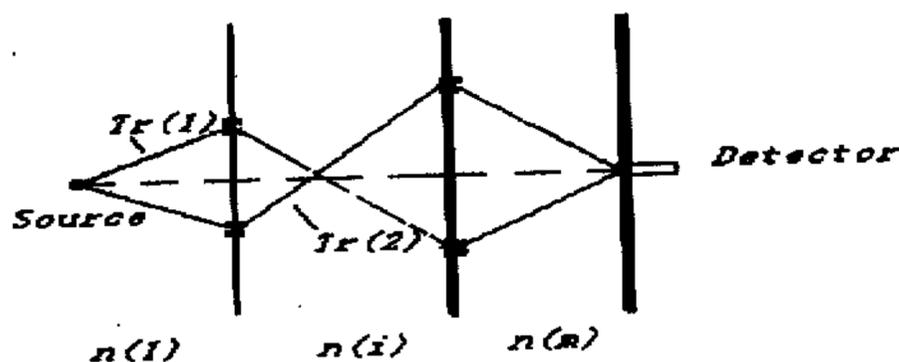


图 5-3 基于双缝干涉实验的 QNN 模型

不同于传统神经网络中每个权值仅调节与之相联的节点的连接关系, 在这里, 每改变权值区中的任何一个权值 (即权值区狭缝的结构), 将会影响到所有输入节点的连接; (2) 输出。该模型的输出是几率幅的叠加, 本质上是复值的, 它增加了网络的信息容量; (3) 物理实现。它具有很直观的物理意义, 易于物理系统实现。但是, 该模型也存在着以下几点问题: (1) 其前端未考虑量子态的叠加, 本意是为了避免相干态的维持问题, 但现实中量子态的叠加是客观存在的, 仅仅利用上述理论框架, 将会失去由量子态叠加而产生的计算优势, 如量子超并行、量子纠缠等等; (2) 虽然 Chrisley 给出了学习算法的大致框架, 但很难实现, 因为在该模型中很难设计出神经网络的量子对照物, 比如说输入量的表示, 由于粒子源是单一的, 它对于每个狭缝应该是等同的, 仅靠狭缝的不同位置不宜表征不同的输入量, 若采取粒子的入射角度或相位来表示, 就会导致计算复杂度的极大提高甚至无解。上述两点严重地制约了 Chrisley 模型的应用。

### 5.1.3 新的计算模型

我们所提出的新模型也是建立在双缝干涉实验基础之上, 并将它推广到多狭缝-多栅板结构。考虑到 Chrisley 模型的缺陷, 我们从网络的输入量着手, 通过设计量子实验来寻找合适的神经网络对照物。由量子理论可知, 当微观粒子进入不同介质后其运动轨迹会发生改变, 我们就选取不同介质的介电常数矢量  $\vec{n} = (n_1, n_2, \dots, n_m)$  作为网络的输入量 (其中  $m-1$  为栅板的个数), 以粒子穿越栅板到达探测器的实际运动轨迹  $Tr \langle S \rightarrow s_1^i \rightarrow \dots \rightarrow s_2^j \rightarrow \dots \rightarrow s_{m-1}^k \rightarrow D \rangle$  作为网络的权值 (实际上它最终由狭缝的位置  $s_j^i$  和栅板间距  $h_j$  所决定), 取探测器所测得点的强度 (即几率)  $P(\vec{n})$  作为网络的输出, 以此来构造一个全新的量子神经网络模型, 参见图 5-3。为了准确理解此模型的意义, 我们必须将其几何构形与神经网络上的含义区分开来。

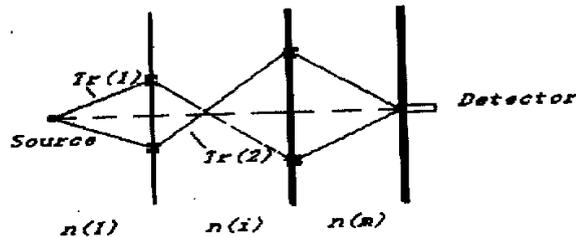


图 5-3 基于双缝干涉实验的 QNN 模型

下面我们利用有关的量子理论知识来推导此模型的动力学表达式。

令  $l(s_j^i, s_{j+1}^{i+1})$  为连接第  $j$  个栅板上狭缝  $s_j^i$  与第  $j+1$  个栅板上狭缝  $s_{j+1}^{i+1}$  之间的路径长度, 其中  $b_j = 1, 2, \dots, N_j$ , 是第  $j$  个栅板上的狭缝, 共有  $N_j$  个;  $r_j^i$  是狭缝  $s_j^i$  距离中心轴的距离。可以得到,

$$l(s_j^i, s_{j+1}^{i+1}) = \sqrt{h_j^2 + (r_j^i - r_{j+1}^{i+1})^2} \quad (5-3)$$

这里考虑了微观粒子的粒子性。

那么, 粒子沿轨迹 1 到达底板所测得的几率幅波函数为

$$\psi(Tr_1) = \frac{1}{\prod_{j=1}^n l(s_{j-1}^{j-1}, s_j^j)} \times \exp\left(ik \sum_{j=1}^n n_j l(s_{j-1}^{j-1}, s_j^j)\right) \quad (5-4)$$

此处有  $s_0^0 = S, s_n^n = D$ 。这里考虑了微观粒子的波动性。

我们再定义

$$\theta = \ln\left(\prod_{j=1}^n l(s_{j-1}^{j-1}, s_j^j)\right) \quad (5-5)$$

$$\omega_j = kl(s_{j-1}^{j-1}, s_j^j) \quad (5-6)$$

则 (5-4) 式便可改写为

$$\psi(Tr_1) = \exp\left(i \sum_{j=1}^n n_j \omega_j - \theta\right) = \exp(i\bar{W}\bar{n} - \theta) \quad (5-7)$$

最终可得神经网络的输出 (即强度) 为

$$P(\bar{n}) = |\psi(Tr_1) + \psi(Tr_2) + \dots|^2 \quad (5-8)$$

可见, 方程 (5-7) 具有与经典神经元极其相似的数学表示形式, 但又有所区别。

鉴于其广义上均为复值形式, 我们利用改进的模拟退火算法<sup>[81]</sup> (MSA) 来训练该网络的权值, 具体算法表述如下:

```

Begin
  initialize     $\bar{\omega} = \bar{\omega}_0, T(0) = T_0, \tau_{\max}, \varepsilon$ 
  set           $\tau = 0$ 
  while         $\Delta E > \varepsilon$ 
    begin
      while     $\tau < \tau_{\max}$ 
        begin
           $\tau = \tau + 1, \omega_i = \omega_i + \Delta \omega_i, \Delta E = E(\bar{\omega} + \Delta \bar{\omega}) - E(\bar{\omega})$ 
           $prob = \min(1, e^{-\frac{\Delta E}{T}})$ 
          if     $prob > random(0,1)$ 
            then update ( $\omega_i$ )
          end
           $T(t) = \alpha T(t)$ 
        end
      end
    output     $\omega_i$ 
  End

```

考虑到复值函数的计算复杂度，我们仅以 3 狭缝-1 栅板结构为例（即  $N_j = 3, m = 2$ ），训练上述 QNN 学习二值逻辑函数  $AND(n_1, n_2)$  和  $XOR(n_1, n_2)$ ，其模拟结果如图 5-4 所示，其中，输入值范围取  $n_1, n_2 \in [1, 5/3]$ ，输出强度阈值取 0.6。

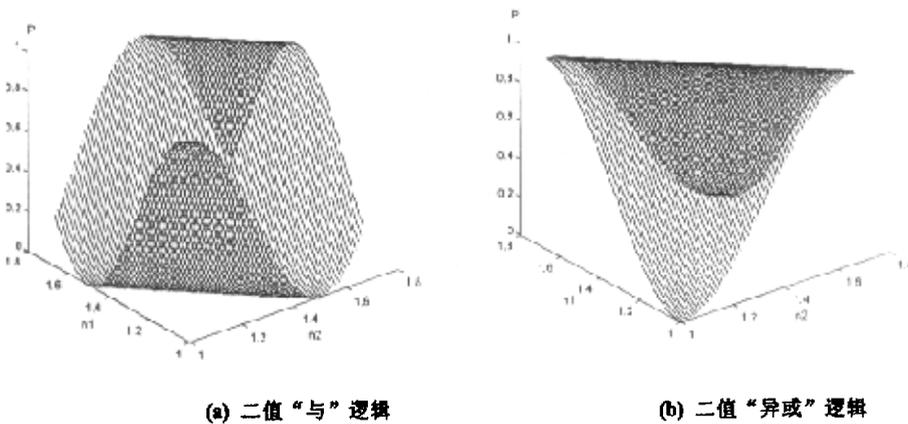


图 5-4 QNN 利用 MSA 算法学习二值逻辑函数的模拟结果

特别地，对于二值 XOR 函数，考察其最简单的 QNN 模型（即  $N_j = 2, m = 2$ ）。我们可以直接从物理公式出发进行推导，由于微观粒子运动的两条轨迹分别为

$$\psi(T_1) = \frac{1}{l_1^2} \exp(ik(n_1 l_1^1 + n_2 l_1^2)), \quad \psi(T_2) = \frac{1}{l_2^2} \exp(ik(n_1 l_2^1 + n_2 l_2^2)) \quad (5-9)$$

设  $r = r_1 + r_2$ ,  $h_1 = h_2 = h$ , 则可得量子神经网络的输出为

$$P(n_1, n_2) = |\psi(Tr_1) + \psi(Tr_2)|^2 \cong \frac{1}{(r^2 + h^2)^2} |1 + \cos k(n_1 + n_2)(Tr_1 - Tr_2)|^2 \quad (5-10)$$

令  $A = \frac{1}{(r^2 + h^2)^2}$ ,  $Tr_1 = Tr_2 + \frac{3}{4}\lambda$ ,  $n_1 = 1$ ,  $n_2 = 5/3$ , 于是有

$$\begin{aligned} P(n_1, n_1) &= A|1 + \cos 3\pi|^2 = 0; & P(n_1, n_2) &= A|1 + \cos 4\pi|^2 = 4A; \\ P(n_2, n_1) &= A|1 + \cos 4\pi|^2 = 4A; & P(n_2, n_2) &= A|1 + \cos 5\pi|^2 = 0 \end{aligned} \quad (5-11)$$

再令  $A = 1/4$ , 就可以得到表 5-1 所示的结果。

表 5-1 类似 XOR 逻辑真值表

Output	Input1	Input2
0	1	1
1	1	5/3
1	5/3	1
0	5/3	5/3

可见, 单层结构的量子神经网络可以完成类似于 XOR 的逻辑运算, 这在传统神经网络中是无法实现的。另外, 从上述推导过程中也可以看到, 正是由于 QNN 中包含了量子态波函数的几率幅叠加, 在输出层出现了二次交叉乘积项, 提高了网络的信息容量, 使它具备了量子计算的优越性。

我们利用量子双缝干涉实验的结构构造了一个全新的 QNN 模型, 它具有学习布尔逻辑函数的功能, 特别是可以用单层网络结构来实现类似异或逻辑运算, 使得 QNN 具备了量子波函数的一些性质, 突破了传统神经网络的局限性。实际上, 该网络还可以看作是利用神经网络来调节量子实验装置的参数, 从而实现量子学习的功能, 因此它具有物理实现的潜力, 能够作为量子系统实现神经计算的一个范例, 这为探索神经网络与量子理论的结合提供了一个很好的途径。

## 5.2 基于量子通用门组单元的神经网络

### 5.2.1 概述

我们知道, 神经网络可以实现从输入空间到输出空间的非线性映射, 在实际应用中可以拟合多变量函数, 完成回归分析或模式识别、系统辨识等任务。在神

设  $r = r_1 + r_2$ ,  $h_1 = h_2 = h$ , 则可得量子神经网络的输出为

$$P(n_1, n_2) = |\psi(T_{r_1}) + \psi(T_{r_2})|^2 \cong \frac{1}{(r^2 + h^2)^2} |1 + \cos k(n_1 + n_2)(T_{r_1} - T_{r_2})|^2 \quad (5-10)$$

令  $A = \frac{1}{(r^2 + h^2)^2}$ ,  $T_{r_1} = T_{r_2} + \frac{3}{4}\lambda$ ,  $n_1 = 1$ ,  $n_2 = 5/3$ , 于是有

$$\begin{aligned} P(n_1, n_1) &= A|1 + \cos 3\pi|^2 = 0; & P(n_1, n_2) &= A|1 + \cos 4\pi|^2 = 4A; \\ P(n_2, n_1) &= A|1 + \cos 4\pi|^2 = 4A; & P(n_2, n_2) &= A|1 + \cos 5\pi|^2 = 0 \end{aligned} \quad (5-11)$$

再令  $A = 1/4$ , 就可以得到表 5-1 所示的结果。

表 5-1 类似 XOR 逻辑真值表

Output	Input1	Input2
0	1	1
1	1	5/3
1	5/3	1
0	5/3	5/3

可见, 单层结构的量子神经网络可以完成类似于 XOR 的逻辑运算, 这在传统神经网络中是无法实现的。另外, 从上述推导过程中也可以看到, 正是由于 QNN 中包含了量子态波函数的几率幅叠加, 在输出层出现了二次交叉乘积项, 提高了网络的信息容量, 使它具备了量子计算的优越性。

我们利用量子双缝干涉实验的结构构造了一个全新的 QNN 模型, 它具有学习布尔逻辑函数的功能, 特别是可以用单层网络结构来实现类似异或逻辑运算, 使得 QNN 具备了量子波函数的一些性质, 突破了传统神经网络的局限性。实际上, 该网络还可以看作是利用神经网络来调节量子实验装置的参数, 从而实现量子学习的功能, 因此它具有物理实现的潜力, 能够作为量子系统实现神经计算的一个范例, 这为探索神经网络与量子理论的结合提供了一个很好的途径。

## 5.2 基于量子通用门组单元的神经网络

### 5.2.1 概述

我们知道, 神经网络可以实现从输入空间到输出空间的非线性映射, 在实际应用中可以拟合多变量函数, 完成回归分析或模式识别、系统辨识等任务。在神

神经网络中,最常用的和研究最多的就是前馈型网络,如BP网络、RBF网络等,用数学公式可以将它们统一表示为:

$$y \approx \sum_{n=1}^N a_n g_n(\varphi(x, b_n)) \quad (5-12)$$

进一步地,有

$$\varphi(x, b_n) = \sum_{m=1}^d \omega_m \psi(x, c_m) \quad (5-13)$$

其中,  $a_n, b_n, \omega_m, c_m$  为可调权值,  $g_n(\cdot)$  为激活函数(或称基函数),  $\varphi(\cdot)$  为内部函数。网络的性能由激活函数和内部函数的种类和数目决定,常见的BP、RBF网络的激活函数分别具有如下特征:  $(x_i - c_m)$  和  $(x_i - c_m)^2$ 。

由此可见,改变激活函数的形式将会得到新的计算模型,随之产生新的计算特性。考虑到近年来有关量子计算性能的研究日益深入,我们设想:能否使用某种变换,将量子计算的概念引入到传统的神经计算模型之中,从而改变它的计算特征?研究这种新的计算模型将有助于结合二者的优势,得到更好的性能。

### 5.2.2 量子门和量子通用门组

在第二章中我们曾经介绍过有关量子门和量子门组网络的预备知识,量子门是物理实现量子计算的基础,从某种意义上说,量子门就代表了量子计算,它包含了量子计算的特点(如量子并行性)。

如前所述,在量子门组中存在着所谓的通用门组,并且已有证明,它可以由两类量子门构成,即一位的相移门和两位的受控非门。鉴于此,我们就考虑用此通用门组组成基本单元,以它们作为神经网络的激活函数,来构成新的计算模型,并考察它的计算性能。

这两类门对量子态的操作有其各自特点,为了便于应用我们首先将其推广到广义范畴,并用数学公式表述如下<sup>[82]</sup>:

$$\text{量子态: } f(\theta) = e^{i\theta} = \cos\theta + i\sin\theta$$

$$\text{相移门: } f(\theta + \theta') = f(\theta) \cdot f(\theta') = e^{i(\theta + \theta')}$$

$$\text{受控-U门: } f\left(\frac{\pi}{2}\kappa - \theta\right) = \begin{cases} \sin\theta + i\cos\theta & (\kappa = 1) \\ \cos\theta + i\sin\theta & (\kappa = 0) \\ \text{else} & (0 < \kappa < 1) \end{cases}$$

这样，实际上得到了用复数表示的量子态及量子通用门组形式。这里，我们将受控非门推广成了受控-U 门。

### 5.2.3 量子神经元及其网络

接下来，我们就构造基于量子通用门组单元的量子神经元，其结构如下图所示。在这样的一个神经元中包括了两大类量子基本门，它们完成对输入量子态的复合操作，实质上是对量子态的幅度和相位分别进行处理。为了便于推广，我们使用了多位态叠加输入和输出，并引入了相位转移系数、阈值系数、相位控制因子等可调整网络参数。

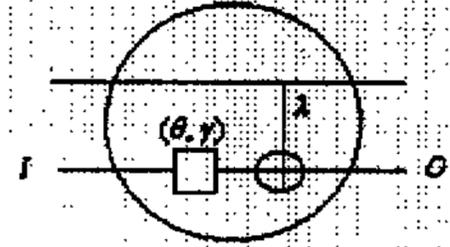


图 5-5 量子神经元示意图

具体地，分别为输入  $I$ ，输出  $O$ ，相位转移系数  $\theta$ ，阈值系数  $\gamma$  和相位漂移因子  $\lambda$ 。

用数学公式表述该量子神经元激活函数为

$$u_k = \sum_I f(\theta_{I,k}) \cdot I_I - f(\gamma_k); \quad y_k = \frac{\pi}{2} \cdot g(\lambda_k) - \arg(u_k); \quad O_k = f(y_k) \quad (5-14)$$

其中， $g(x) = \frac{1}{1+e^{-x}}$ 。

有了上述的神经元，我们借助传统前馈神经网络模型的结构就可以建立起多层量子神经网络模型，如图 5-6 所示。

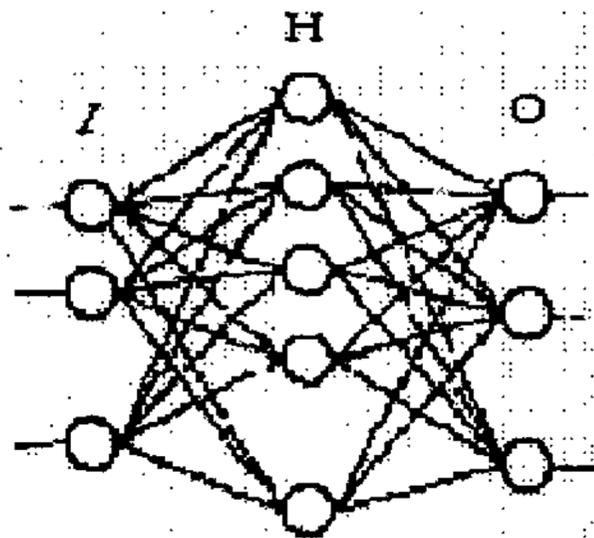


图 5-6 三层前馈量子神经网络结构示意图

该量子神经网络模型结构具有以下几个特征：(1) 复数形式；(2) 类似传统

BP 网络; (3) 应用量子门进行计算, 具有量子硬件实现的潜力。在此网络中, 信息处理的过程可以描述为: 输入量子态 (以态叠加的形式) 由输入端 I 输入, 经由相移门和受控 U 门组成的通用门组单元处理, 其相位及幅度均受到调制, 然后进入隐层 H 混合处理, 并由输出层 O 输出量子态, 最后对其进行测量后便可得到最终的结果。

### 5.2.4 训练算法

我们取改进的复值 BP 算法<sup>[83, 84]</sup> (Complex-BP) 作为上述量子神经网络的训练算法, 首先定义误差函数为

$$E = \frac{1}{2} \sum_p (t_p - \text{Measure}_p)(t_p - \text{Measure}_p)^* \quad (5-15)$$

其中星号表示共轭复数。

训练过程表示如下:

$$\begin{aligned} \theta^{new} &= \theta^{old} - \eta \cdot \frac{\partial E}{\partial \theta^{old}}; \\ \gamma^{new} &= \gamma^{old} - \eta \cdot \frac{\partial E}{\partial \gamma^{old}}; \\ \lambda^{new} &= \lambda^{old} - \eta \cdot \frac{\partial E}{\partial \lambda^{old}} \end{aligned} \quad (5-16)$$

其中  $\eta$  为学习率, 通常取 0.0 - 1.0。

### 5.2.5 仿真实验

我们将上述量子神经网络模型应用于金融数据分析。

例 1: 由公司的财务状况判断其破产风险<sup>[85]</sup>。具体地, 通过两个有代表性的财务指标 (即流动比率、债务/资产比率) 预测某公司的经营情况, 是否会导致其破产。这实质上是一个分类问题。此类问题由于没有明确的分类标准且不是线性可分的, 因此常用学习 (训练) 样本集的方法发现其中的规律, 然后进行泛化推广。

作为考察上述量子神经网络计算性能的例子, 我们选取部分公司样本, 分别用经典 BP 网络和量子神经网络进行处理, 然后比较二者结果, 从而得出结论。针对本例, 量子神经网络要优于经典 BP 网络, 它在训练速率和精度上都有了明显的提高。例 1 的训练样本集及训练结果分别见表 5-2 和表 5-3。

表 5-2 例 1 的训练样本集

序号 Number	流动比率 Current ratio	债务/资产比率 Debt/assets ratio (%)	是否破产 Yes/No
1	3.6	60	No
2	3.0	20	No
3	3.0	60	No
4	3.0	76	Yes
5	2.8	44	No
6	2.6	56	Yes
7	2.6	68	Yes
8	2.4	40	Yes
9	2.4	60	No
10	2.2	28	No
11	2.0	40	No
12	2.0	48	No
13	1.8	60	Yes
14	1.6	20	No
15	1.6	44	Yes
16	1.2	44	Yes
17	1.0	24	No
18	1.0	32	Yes

表 5-3 例 1 的训练结果

	量子神经网络	经典 BP 网络
网络结构	2-7-1	2-8-1
训练次数	342	654
训练误差	0.00069	0.013

例 2: 汇率预测问题, 即时间序列的拟合。在经济领域中的时间序列大多数是非平稳的, 这里我们以美元对新加坡元的汇率<sup>[86]</sup>为例, 应用本文所提出的量子神经网络进行预测, 并把预测结果与经典 RBF 网络和文献[86]中所得的结果进行比较。

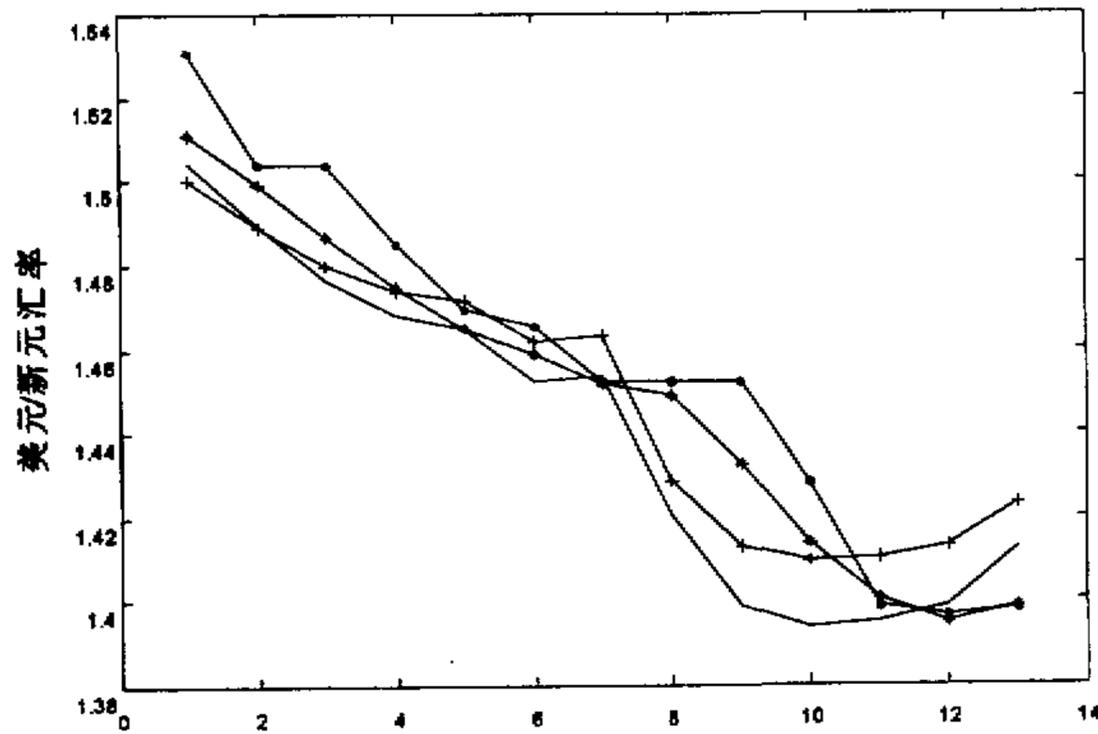


图 5-7 非平稳时间序列的预测结果比较图

实际汇率以月计, 从 1990 年 1 月到 1995 年 8 月共 68 个数据, 我们将前 55 个作为训练集, 后 13 个作为测试集。实验结果如图 5-7 所示, 其中实心线表示实际值, “+号线”、“\*号线”和“.”号线”分别代表量子神经网络、经典 RBF 神经网络以及文献[86]采用模糊神经网络所得到的预测值, 具体数值见下表。经比较可见, 量子神经网络在预测精度方面具有显著的优越性。

表 5-4 例 2 实验数据结果比较

测试集 (Mon./yr)	实际值	量子神经网络预测值	经典 RBF 预测值	文献[5]预测值
08/94	1.5043	1.5015	1.5111	1.5309
09/94	1.4891	1.4961	1.4992	1.5040
10/94	1.4763	1.4784	1.4863	1.5040
11/94	1.4681	1.4734	1.4744	1.4847
12/94	1.4651	1.4723	1.4648	1.4693
01/95	1.4523	1.4613	1.4588	1.4653
02/95	1.4538	1.4611	1.4517	1.4523
03/95	1.4203	1.4273	1.4490	1.4523

---

04/95	1.3985	1.4119	1.4326	1.4523
05/95	1.3937	1.4072	1.4135	1.4281
06/95	1.3949	1.4098	1.4005	1.3984
07/95	1.3984	1.4120	1.3946	1.3963
08/95	1.4125	1.4204	1.3985	1.3979
误差平方和 (SSE)		0.0012	0.0030	0.0075

---

本节将量子门组作为神经网络的激活函数,从而实现神经计算,所提出的模型的计算性能要优于传统的 BP 网络和 RBF 网络,究其原因是因为该模型已经扩展到了复数域,复数形式的数据在训练时分为实部和虚部同时进行,大大节省了计算时间,同时由于量子门组单元中含有可微细调节的相位参数,使得学习精度也有所提高。更重要的是,由于该模型可以通过量子门组来实现,那么待量子门的制作技术成熟之时,就极易用硬件来实现量子神经计算。因此,该模型具有广泛的应用前景。

## 第六章 量子神经网络的学习模型

神经计算的核心是学习，本章将讨论量子神经网络的学习问题。针对量子纠缠及量子演化的特征，分别提出基于隐形传态的量子纠缠网络以及量子竞争学习网络模型，详细分析其中的学习原理和过程。

### 6.1 基于隐形传态的量子纠缠神经网络

#### 6.1.1 概述

第二章我们曾用较多的篇幅讨论了量子纠缠态的特性及量子隐形传态的基本原理和实现步骤，值得注意的是，在上述信息传递过程中只是讨论了如何在量子系统中实现量子态信息的传递，而没有涉及到任何有关学习的内容。在这里，我们将尝试基于量子隐形传态的结构，把传统神经计算的学习功能引入其中，试图把它与量子纠缠的某些特性结合起来，开发出新颖的信息处理方式。

#### 6.1.2 基于量子隐形传态的纠缠神经网络模型

由于量子纠缠的奇异性，在数学上还无法准确地描述<sup>[87]</sup>，我们仅讨论它在信息处理中的简单理论模型，至于仿真计算暂时尚未涉及。

设计的目的是：在量子系统中实现学习、计算和信息传递等功能。该模型的设计要点有三：（1）量子纠缠态的定义；（2）参与隐形传态的各个量子态的表示；（3）学习功能的模拟。

由于基于 EPR 效应的量子隐形传态模型内部能够完成量子计算和量子态信息的无损传递，我们就将上述的由 Alice、Bob、EPR 对组成的模型作为基本单元，参见图 6-1，它主要完成信息的传送功能，然后将许多这样的单元组成一个大的网络，对于某一个基本单元来说，就要考虑单元之间的信息传递问题，于是，一个单元同时既拥有内部信息又拥有外部信息，这些信息以一定形式的组合形成知识源，在定义了态纠缠和各量子态的表示之后，核心问题就是如何利用此网络结构以及传送方式从中获取知识。在这里可以把信息处理过程分为两部分，一是在内部进行信息传递，其间完全利用了量子纠缠的作用，在结果测量阶段可以引入智能坍塌机制，二是接受外部信息，即基本单元之间的信息传递，在此可以引入某种规则（如 Hebb 规则）实现学习功能。具体地，量子纠缠神经网络的理论模型及

中间的信息处理过程可详细描述如下:

- (1) Alice (发送者)、Bob (接收者) 和一个 EPR 源, 再加上一个经典信道与一个量子信道便构成一个量子隐形传态单元模型 (见图 6-1);
- (2) Alice 和 Bob 希望能够相互通讯, 由 EPR 源传送给每个人一个纠缠粒子, 两个粒子构成一个 EPR 对, 即  $|\varphi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ ;
- (3) Alice 试图通过经典信道和量子信道将她的量子态  $|\phi\rangle = a|0\rangle + b|1\rangle$  传送给 Bob, 她首先对她的量子态和纠缠态求直积, 得到一个初始态, 即

$$\begin{aligned} |\phi\rangle \otimes |\varphi_0\rangle &= \frac{1}{\sqrt{2}} [a|0\rangle \otimes (|00\rangle + |11\rangle) + b|1\rangle \otimes (|00\rangle + |11\rangle)] \\ &= \frac{1}{\sqrt{2}} (a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle) \end{aligned} \quad (6-1)$$

- (4) Alice 应用  $C_{\text{not}} \otimes I$  和  $H \otimes I \otimes I$  变换去直乘初始态, 其中  $C_{\text{not}}$  为受控非门 (Controlled-NOT),  $I$  为单位变换矩阵,  $H$  为 Hadamard 变换矩阵, 即

$$\begin{aligned} &(H \otimes I \otimes I)(C_{\text{not}} \otimes I)(|\phi\rangle \otimes |\varphi_0\rangle) \\ &= (H \otimes I \otimes I)(C_{\text{not}} \otimes I) \frac{1}{\sqrt{2}} (a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle) \\ &= (H \otimes I \otimes I) \frac{1}{\sqrt{2}} (a|000\rangle + a|011\rangle + b|110\rangle + b|101\rangle) \\ &= \frac{1}{2} [|00\rangle(a|0\rangle + b|1\rangle) + |01\rangle(a|1\rangle + b|0\rangle) + |10\rangle(a|0\rangle - b|1\rangle) + |11\rangle(a|1\rangle - b|0\rangle)] \end{aligned} \quad (6-2)$$

- (5) 这样, Alice 控制着前两位, Bob 控制着最后一位, 根据单元的内部及外部信息 Alice 对前两位量子态进行测量, 并将测量结果作为经典信息传送给 Bob;

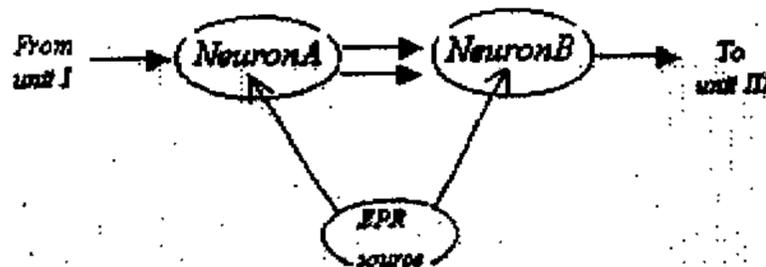


图 6-1 量子隐形传态单元 II 模型

- (6) 依据 Alice 的测量结果, Bob 所控制的位的态分别对应为  $a|0\rangle + b|1\rangle$ 、 $a|1\rangle + b|0\rangle$ 、

$a|0\rangle - b|1\rangle$  和  $a|1\rangle - b|0\rangle$ ，它们是由基态（0 或 1）和相态（+或-）联合组成的；

- (7) Bob 收到 Alice 的信息后将他的一个 EPR 粒子解码，最终获得 Alice 原有态  $|\phi\rangle = a|0\rangle + b|1\rangle$ ，其间他利用了相应的解码变换矩阵。

### 6.1.3 学习机理讨论

在上述过程中，与第二章所不同的地方在于第（5）步，Alice 在测量前两位时综合考虑了单元内、外部信息，我们这里用决策因子 D 来表示，之后她依据 D 测量量子态时还应用了类似 Grover 算法，使得期望得到的量子态在坍塌时的概率增大，并抑制不期望得到的量子态的概率。

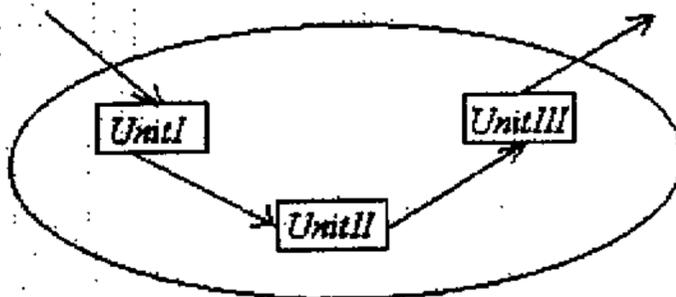


图 6-2 含有三个量子隐形传态单元的量子纠缠神经网络模型

在单元内部模型完全按类似于经典隐形传态的工作步骤来执行信息的传递过程，但值得注意的是，由于存在单元之间的连接（参见图 6-2），这样在下一个单元的处理过程中不仅有其内部的信息还有上一个单元所传递来的信息，这一部分信息的传递就要通过学习来获得，因为单元之间存在连接权，它调节了单元之间信息传递的强度，即决策因子 D。我们可以采用最简单的 Hebb 学习规则，假定上一个单元传出的信息对下一个单元的内部信息是同相增强的，即量子态中相位相同的其概率幅度将得到增强。

### 6.1.4 待解决的问题

由于纠缠态尚无广义的数学表述，我们对实际意义的特征描述在用量子态表示时只能按思维假定来规定纠缠，这一点极大地限制了它的应用。但是，一旦规定好处于纠缠态的特征，在上述神经网络处理过程中就会具有量子计算的一些优势，如量子并行，它就符合人类无重复学习的特点。另外，适合模型的学习规则仍需进一步挖掘，还有多单元之间（含反馈）的连接讨论等等。本节提出的纠缠神经网络模型只是一个粗糙的理论模型，有待于进一步完善。

## 6.2 量子竞争学习网络

量子计算利用了量子理论中有关量子态的叠加、纠缠和干涉等特性,通过量子并行计算使得某些在经典计算机上计算复杂度很高的问题有可能降低其复杂度(如 NP-hard 问题等),这其中的关键就是设计有效的量子算法。但是,目前已知的量子算法不是很多,究其原因可能是,一方面算法设计的难度较大,另一方面与其应用领域的狭窄有关。因此,寻找新的应用领域和应用对象就成为拓展量子计算的一个重要研究方向。

竞争学习是人工神经网络中的一种重要的学习方式,它模仿了人脑中神经工作的机制,可以实现模式的分类和联想记忆等功能。Hamming 神经网络是结构简单的一种竞争学习模型<sup>[88]</sup>。下面简单介绍它的网络结构和竞争学习实现过程。

图 6-3 所示为 Hamming 网络的拓扑结构,它是一个两层的前馈网络,由下层的匹配子网络和上层的竞争子网络共同组成,其中,下层计算输入模式与存储模式之间的 Hamming 距离,上层则依据距离最小准则寻找与输入模式最匹配的某个存储模式。如果将存储模式作为类别,那么就可以实现模式分类,类似地,也可以实现模式的联想记忆。

假设输入模式的维数为  $m$ ,已存储模式的样本数为  $n$ ,即网络容量。其算法过程可描述如下:

### 第一层:模式匹配

假设输入模式为  $X=(x_1, x_2, \dots, x_m)$ ,网络中已存储的模式为  $P^j=(p_1^j, 0 \leq i \leq m, 0 \leq j \leq n)$ ,计算  $X$  与每个  $P$  之间的 Hamming 距离,即

$$HD(X, P^j) = \frac{m - X \cdot P^j}{2} \quad (0 \leq j \leq n) \quad (6-3)$$

其中,  $X \cdot P^j$  是指两个模式的内积。简单地,若模式以二进制数表示,那么 Hamming 距离就定义为两个模式矢量中取不同值(即相反值)的分量的数目,它用来度量两个模式的差异,值越小表示两个模式越接近,反之则表示越远离。于是可得,两个相同模式间的 Hamming 距离为零,Hamming 距离的最大值为  $m$ 。

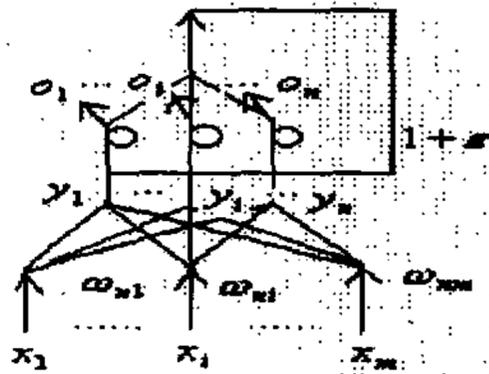


图 6-3 Hamming 神经网络的拓扑结构

## 第二层：竞争学习

依据 Hamming 距离最小准则，即  $y_j = X \cdot P^j$  最大，经如下迭代可得到分类结果，

$$y_j(t+1) = f(y_j(t) - \varepsilon \sum_{k \neq j} y_k(t)) \quad (6-4)$$

其中， $\varepsilon < \frac{1}{n}$ 。

## 6.2.1 量子学习算法的描述

我们知道，量子算法利用了量子态的干涉特性，使所需的结果增强，同时使不必要的结果减弱，这样所需的结果在测量时就会以较高的概率出现。在整个计算过程中量子算法实际上是一系列幺正算子的连续操作，它描述了量子态由初态变化到终态的演化过程。类似地，我们认为量子学习算法也应具有相似的特征。不失一般性，我们利用量子理论中的一些概念将上述过程描述如下：

量子理论中的 Copenhagen 学派观点认为，微观粒子同时具有波动性和粒子性（即波粒二象性），其状态用波函数来表示且满足 Schrödinger 方程，在量子系统中的演化算子是幺正的，但在结果的测量即坍缩（Collapse）时所发生的消相干却是一个非幺正过程，可以把它视作量子系统趋向某个吸引子的非线性演化。据此观点，最直接的描述方法就是利用波函数形式，将传统神经计算引入量子系统，同时设计量子学习算法和非幺正（测量）算子，从而实现输入空间到输出空间的非线性映射，其具体实现过程如图 6-4 所示。

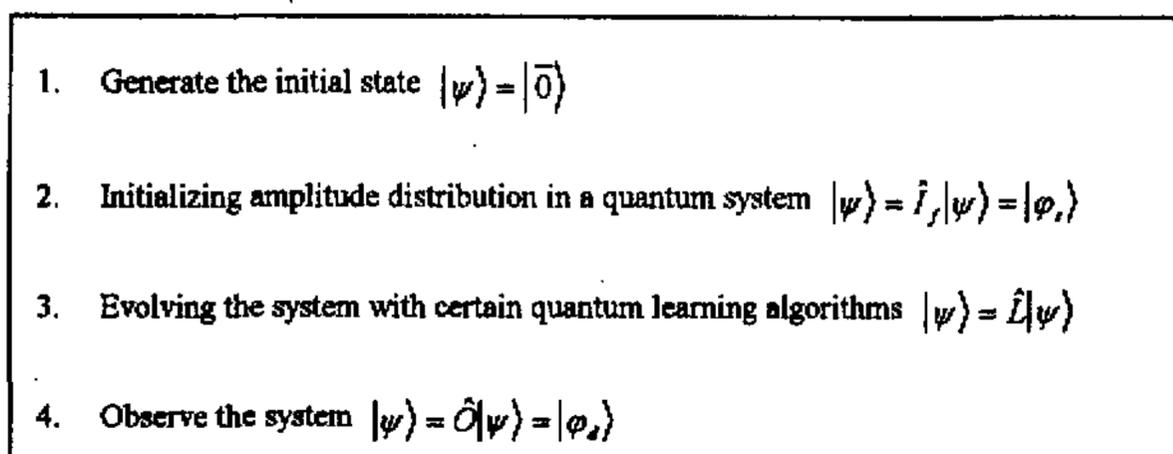


图 6-4 量子学习算法实现过程示意图

其中，算子  $\hat{I}_1$ 、 $\hat{L}$ 、 $\hat{O}$  依次是初始化算子、学习算子和测量算子，量子态  $|\varphi_1\rangle$  和  $|\varphi_s\rangle$  分别为初态和终态，同时注意，等号两边符号相同的量子态并非表示同一个量子

态, 从左到右描述了量子态的演化过程。上述过程中存在的关键问题是初态的制备和量子学习算法的设计。

### 6.2.2 量子竞争学习算法 (QCLA)

具体地, 我们通过设计若干算子来构造 Hamming 神经网络的量子对照物, 从而提出一种量子竞争学习算法, 它用于模式分类和联想记忆。具体过程描述如下:

#### STEP1 量子态的初始化

$$|\psi\rangle = |\bar{0}\rangle \quad (6-5)$$

$$|\bar{0}\rangle = |\psi\rangle = \frac{1}{\sqrt{n}} \sum_{0 \leq j < n} |P^j\rangle \quad (6-6)$$

首先将所有的量子位清零, 然后利用 Ventura-Martinez 算法<sup>[29]</sup>将待存储的模式存入寄存器中作为初态。

#### STEP2 量子态的编码

本文采用矢量编码的方法, 例如可将三个量子位的量子态 (如 $|101\rangle$ ) 用矢量编码为 $(00000100)^T$ 。

#### STEP3 竞争学习

前一个子矩阵 (6-7) 用来计算 Hamming 距离和模式匹配, 其中,  $i^m$  和  $j^m$  是指行、列的前  $m$  位,  $a$  是模式匹配阈值, 对角线上的 1 表示相应的模式可以参与竞争, 而非对角线上的 -1 是为防止伪匹配所添加的惩罚系数, 它们皆由阈值大小所决定; 后一个子矩阵 (6-8) 用来实现竞争学习, 它定义已存储的模式为合法。

$$\hat{C} = c_{ij} = \begin{cases} 1 & \text{if } i = j, HD(X, i^m) \leq m - a \\ -1 & \text{if } HD(X, i^m) \leq m - a, HD(X, j^m) \leq m - a \\ & \text{and } HD(X, i^m) < HD(X, j^m) \\ 0 & \text{otherwise} \end{cases} \quad (6-7)$$

$$\hat{R} = r_{ij} = \begin{cases} 1 & \text{if } i = j, HD(P, i^m) = 0 \\ 0 & \text{otherwise} \end{cases} \quad (6-8)$$

**STEP4 模式分类 (联想记忆)**

$$|\psi\rangle = \hat{R}\hat{C}|\psi\rangle \xrightarrow{\text{collapse}} |P\rangle \quad (6-9)$$

对终态进行测量，最终实现模式分类或联想记忆。

**6.2.3 实例分析**

现举例讨论上述 QCLA 流程：

已知存储模式为： $P = \begin{cases} 00 \rightarrow 1 \\ 01 \rightarrow 0 \\ 11 \rightarrow 0 \end{cases}$ ，对测试模式  $X = 10$  进行分类。

(1) 初始化量子态，并用 Ventura-Martinez 算法将已知模式存储在 3-qubit 量子系统中，即

$$\hat{S}|\bar{0}\rangle = |\psi\rangle = \frac{1}{\sqrt{3}}|001\rangle + \frac{1}{\sqrt{3}}|010\rangle + \frac{1}{\sqrt{3}}|110\rangle \quad (6-10)$$

(2) 将初态编码为矢量形式，即

$$|\psi\rangle = (0, 1, 1, 0, 0, 0, 1, 0)^T \quad (6-11)$$

(3) 已知  $m=2$ ，设  $a=1$ ，依据测试模式  $X = 10$  来计算两个竞争学习子矩阵，可得

$$\hat{c} = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 & 1 \end{bmatrix} \quad (6-12)$$

$$\hat{R} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (6-13)$$

$$\text{计算 } |\psi\rangle = \hat{R}\hat{C}|\psi\rangle = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad (6-14)$$

(4) 可得终态为

$$|\psi\rangle = \frac{1}{\sqrt{2}}|001\rangle + \frac{1}{\sqrt{2}}|110\rangle \quad (6-15)$$

经测量得

$$|P\rangle = \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|0\rangle \quad (6-16)$$

即测试模式  $x=10$  归为 0 类和 1 类的概率均为 50%。这与传统 Hamming 网络计算的结果一致，初步证明了该量子算法的有效性。

#### 6.2.4 讨论

量子计算具有许多独特的计算性能，将它与现有的计算技术相结合可以拓宽其应用领域，并可能发现新的量子算法。由于神经计算和量子计算同属于一类非精确性的、概率性的计算方法，因此它们的结合有良好的前景。

本节尝试将二者联系起来，通过设计若干量子算子将传统的 Hamming 神经网络引申到量子领域，从而提出一种用于模式分类和联想记忆的量子竞争学习算法。不过，由于量子神经计算是一种完全不同于传统方法的计算模式，对它的把握和应用因为没有经典的对照物而有一定难度，另外，文中所提的算法还比较简单，有必要进一步加以完善。

## 第七章 全文总结

### 7.1 本文所做的主要工作

量子理论与神经网络的结合有着内在的合理性和很强的创新性, 本文在已有文献调研的基础上, 系统探讨了量子神经计算这一全新计算模式, 并利用有关量子理论知识设计出几种量子神经网络模型以及相应的学习算法, 研究了它们在模式分类、记忆等方面的应用。本文的研究发展了传统人工神经网络理论, 在其中引入了量子计算的理论和方法, 这将有助于提升神经网络信息处理的能力, 同时也丰富了量子计算的研究领域, 有助于深入理解量子理论与神经网络理论的本质联系和各自在更深层次的含义。

总结全文, 本文研究的主要内容和关键技术包括:

- (1) 量子系统实现神经计算可能性的初步论证;
- (2) 神经计算量子推广的基本方法;
- (3) 具有量子力学特性的神经元模型讨论;
- (4) 多宇宙量子神经网络模型及其坍缩算法;
- (5) 基于双缝干涉实验的量子神经网络模型及学习算法;
- (6) 基于量子通用门组单元的神经网络模型及其算法;
- (7) 基于隐形传态的量子纠缠神经网络的理论模型;
- (8) 量子竞争学习网络。

### 7.2 论文的创新之处

量子神经计算是人工神经网络理论与量子理论相结合的产物, 它兼有神经计算和量子计算的优势, 其计算模型是一种全新的神经网络模型, 它具有与传统神经网络不同的优异性能, 对它的进一步研究对于计算机科学、信息科学、量子物理学、认知科学甚至复杂性科学都将有所推动。因此, 这是一个很有价值和发展前景的研究方向。

本文的研究成果具有以下几个方面的贡献:

- (1) 量子理论中的多宇宙观点特别适合于构造用于分类任务的神经网络, 本文利用多 BP 网络结构设计出多宇宙量子神经网络模型, 它包含了量子态的叠加、纠缠和坍缩等量子过程, 在训练时利用了 Grover 量子搜索策略, 针对不同问题设计了不同的坍缩算法, 使期望输出量子态的相位逐步增大, 同时抑制其它可能的输出状态, 从而实现模式分类功能。本文对单层及两层量子神经网络作了大量的数值实验, 仿真结果显示, 该类模型与经典 BP 网络相比, 训练速度大大提高, 而泛化能力基本持平;
- (2) 本文在指出 Chrisley 理论模型的缺陷的基础上, 对于量子双缝干涉实验给出全新的物理定义, 提出了基于该实验的多狭缝-多栅板网络模型, 该模型为线性不可分问题提供了一个单层网络的解决方案。本文利用改进的模拟退火算法实现了类似二值异或运算的模拟仿真, 同时对所得结果给出物理公式推导, 从两个侧面证明了结果的正确性;
- (3) 由于前馈网络的激活函数的性能在很大程度上决定了神经网络的计算性能, 本文提出基于量子通用门组单元的神经网络模型, 它利用量子计算中的一位相移门和两位受控非门所构成的通用门组单元作为神经网络的激活函数, 将经典的前馈网络推广到复数域, 利用改进的复值 BP 算法进行训练, 应用结果显示, 该网络与传统 BP 网络和 RBF 网络相比, 具有更好的拟合能力和更快的学习速率。此外, 该模型还具有易于物理实现的潜力;
- (4) 本文利用量子算法的理论框架, 结合经典 Hamming 网络的学习原理, 提出了一种用于竞争学习的量子算法, 该算法含有两个么正算子, 分别实现模式的匹配和竞争学习, 初步的实验表明, 此算法对模式的分类或记忆有效, 它拓宽了量子算法的应用领域。

### 7.3 进一步的研究

目前有关量子神经计算的研究刚刚起步, 还处于初始阶段, 本文虽然在一些方面取得了一定的成果, 但随着研究的深入, 我们将会发现该领域十分宽广, 还有许多课题有待于探索。基于当前的认识, 作者认为有以下问题值得进一步研究:

- (1) 量子系统实现神经计算的进一步论证, 广义量子学习算法的提出;
- (2) 利用 Feynman 路径积分设计量子神经网络, 该理论似乎更符合神经网络的拓扑结构, 由于路径积分在理论上与量子矩阵力学和波动力学等价, 因此

也应该同样有效，文中曾给出部分讨论，该法设计的难点在于学习的实现；

- (3) 可以考虑将经典神经网络的激活函数量子化<sup>[89]</sup>，即针对不同输入作分段输出，这样对于划分、分类问题可能会取得更精确的效果；
- (4) 量子门组计算网络的应用更能体现出量子计算的性质，并具有硬件实现的潜力，因此基于量子门的幺正算子设计是量子神经计算的一个重要方向；
- (5) 量子理论中能否引入些许非线性<sup>[90]</sup>？如有可能，将极大地改善量子神经网络的设计环境，更容易沟通神经计算与量子计算之间的联系。有无非线性在量子理论研究领域中是一个有争议、但又很有意义的课题，是否能从量子神经计算这一侧面取得一些进展？

## 参 考 文 献

- [1] Kohonen T. The Self-organizing Map. Proc. of the IEEE, 1990, 78: 1464-1480
- [2] Lapedes A. and Farber R. "Nonlinear signal processing using neural networks: Prediction and system modeling", LA-VR-87-2662, Los Alamos National Laboratory, 1987
- [3] Pao Y. H. Adaptive Pattern Recognition and Neural Networks. Mass.: Addison-Wesley Publishing Co., 1989
- [4] Bailey D. and Thompson D. M. "Developing Neural Network Applications". AI Expert, 1990, 9: 33-41
- [5] Gupta M.M. "Fuzzy neural networks: Theory and applications." Proceedings of SPIE, 1994, 2353: 303-325
- [6] Stefan P. et al. "Feature Extraction from Wavelet Coefficients for Pattern Recognition Tasks". IEEE Trans. Pattern Analysis and Machine Intelligence, 1993, 21(1): 83-88
- [7] Yeten B. and Umrah F. G. "The Use of Fractal Geostatistics and Artificial Neural Networks for Carbonate Reservoir Characterization". Transport in Porous Media, 2000, 41: 173-195
- [8] Levin E. et al. "A Statistical Approach to Learning and Generalization in Layered Neural Networks". Proc. of the IEEE, 1990, 78(10): 4568-1573
- [9] Beniof P. "Quantum Mechanical Hamiltonian Model of Turing Machines". J. Stat. Phys., 1982, 29: 515-546
- [10] Feynman R. P. "Simulating Physics with Computers". Int. J. Theo. Phys., 1982, 21: 467-470
- [11] Deutsch D. "Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer". Proc. R. Soc., London A, 1985, 400: 97-117
- [12] Shor P. W. "Algorithms for Quantum Computation Discrete logarithm and Factoring". Proc. of the 35th Annual Sym. on Foundations of Computer Science, 1994, 124-134
- [13] Grover L. K. "A Fast Quantum Mechanical Algorithm for Database Search". Proceedings of the 28th Annual ACM Symposium on the Theory of Computing, 1996, 212-219, ACM, New York
- [14] Bennett C. H. et al. "Quantum Information and Computation". Nature, 2000, 404: 247-255
- [15] Grover, L. K. "Quantum Computation". Proceedings of Twelfth International Conference On VLSI Design, 1999, 548-553
- [16] Narayanan A. "Quantum Computing for Beginners". Proceedings of the 1999 Congress on

- Evolutionary Computation, 1999, 3: 1999-2238
- [17] Shor P. W. "Fault-tolerant Quantum Computation". Proceedings of 37th Annual Symposium on Foundations of Computer Science, 1996, 56-65
- [18] DiVincenzo D. P. "Principles of Quantum Computing". Proceedings of IEEE International Solid-State Circuits Conference, 1995, 312-313
- [19] Penrose R. *The Emperor's New Mind*. Oxford University Press, 1989
- [20] Penrose R. *Shadows of the Mind*. Oxford University Press, 1994
- [21] Hameroff R. "Quantum Coherence in Microtubules: An Intra-neuronal substrate for Emergent Consciousness"? *J. of Consciousness Studies*, 1994, 1: 91-118
- [22] Perus M. "Neuro-Quantum Parallelism in Brain-Mind and Computers". *Informatica*, 1996, 20: 173-183
- [23] Mershin A., Nanopoulos D.V., Skoulakis E.M.C. "Quantum Brain"? E-Print <http://xxx.lanl.gov/lanl-e-print-quant-ph/0007088>
- [24] Pribram H. "Quantum Holography: Is It Relevant to Brain Function"? *Information Science*, 1999, 115: 97-102
- [25] Kak S C. "On Quantum Neural Computing". *Information Sciences*, 1995, 83: 143-160
- [26] Chrisley R. "Quantum Learning". In Pylkkänen P. and Pylkkö P. Editors, *New Directions in Cognitive Science: Proceedings of the International Symposium, Saariselka, 4-9 August 1995, Lapland, Finland, 77-89*, Helsinki. Finnish Association of Artificial Intelligence
- [27] Behrman E. et al. "A Quantum Dot Neural Network". Proceedings of the Workshop on Physics of Computation, 1996, 22-24
- [28] Ventura D. and Martinez T. R. "An Artificial Neuron with Quantum Mechanical Properties". Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms, 1997, 482-485
- [29] Ventura D. and Martinez T. R. "Quantum Associative Memory". *Information Sciences*, 2000, 124: 273-296
- [30] Narayanan A. and Menneer T. "Quantum Artificial Neural Network Architectures and Components". *Information Sciences*, 2000, 128(3-4): 231-255
- [31] Li W. "A Study of Parallel Self-Organizing Map". E-Print <http://xxx.lanl.gov/lanl-e-print-quant-ph/9808025>, 1998
- [32] Ezhov A. A., Berman G. P. "Role of Interference and Entanglement in Quantum Neural Processing". SPIE's International Sym. on Microelectronics and Micro-Electro-Mechanical

- Systems, 2001, 4591-4605
- [33] Andrecut M., Ali M. K. "A Quantum Perceptron", *International Journal of Modern Physics B*, 2002, 16(2): 75-88
- [34] Chandrashekar V. G., Behrman E. C., Steck J. E. "Physical Realizations of A Temporal Quantum Neural Computer". *Proceedings of International Joint Conference on Neural Networks*, 2001,2: 1571-1576
- [35] Matsuda S. "Quantum Neurons And Their Fluctuation". *Proceedings of 1993 International Joint Conference on Neural Networks*, 1993, 2: 1610-1613, Nagoya.
- [36] Behrman E. C., Steck J. E., Skinner S. R. "A spatial quantum neural computer." *Proceedings of International Joint Conference on Neural Networks*, 1999, 2: 874-877
- [37] Ventura D. and Kak S. C. "Quantum computing and neural information processing". *Information Sciences*, 2000,128(3/4): 147-148
- [38] Ezhov A. A. "Spurious Memory, Single-class and Quantum neural Networks". In: Pylkkänen P. and Pylkkö P. Editors, *Brain Mind and Physics*, IOS Press, Amsterdam, 1997, 126-139
- [39] Ezhov A. A. and Ventura D. "Quantum Neural Networks". In *Future Directions for Intelligent Systems and Information Science* (Ed. N. Kasabov), Physica-Verlag, 2000
- [40] Ventura D. "Artificial Associative Memory Using Quantum Processes". *Proceedings of the International Conference on Computational Intelligence and Neuroscience*, 1998,2:218-221
- [41] Ventura D. and Martinez T. "Quantum Associative Memory with Exponential Capacity". *Proceedings of IEEE International Joint Conference on Neural Networks*, 1998,1: 509-513
- [42] Ezhov A. A., Nifanova A. V., Ventura D. "Quantum Associative Memory with Distributed Queries". *Information Sciences*, 128(3/4): 271-293
- [43] Menneer T. and Narayanan A. "Quantum-inspired Neural Networks". *Technical Report R329*, Department of Computer Science, University of Exeter, UK, 1995
- [44] Perus M. and Ecimovic P. "Memory and Pattern Recognition in Associative Neural Networks". *International Journal of Applied Science and Computation*, 1998,4: 283-310
- [45] Chrisley R L. "Bohmian Quantum Neural Networks". <http://www.cogs.susx.ac.uk/users/ronc>, 2000
- [46] Ventura D. "Quantum Computational Intelligence: Answers and Questions". *IEEE Intelligent System*, 1999, July/August,14-16
- [47] 周世勋. *量子力学*. 北京: 高等教育出版社, 1979

- [48] Bohm D. *Quantum Theory*. New York: Prentice Hall, 1951
- [49] Bohm D. "A Suggested Interpretation of Quantum Theory in Terms of Hidden Variable". *Phys. Rev.*, 1952, 85: 1-2
- [50] Everett H. "Relative State Formulation of Quantum Mechanics". *Reviews of Modern Physics*, 1957, 29: 454-462
- [51] 曾谨言. *量子力学* (第三版). 北京: 科学出版社, 2000
- [52] Bennett C H., Brassard G., Crepeau C., et al. "Teleporting An Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels". *Phys. Rev. Lett.*, 1993, 70: 1895-1899
- [53] Einstein A. et al. "Can Quantum Mechanical Description of Physical Reality be Considered Complete"? *Phys. Rev.*, 1935, 47: 56-68
- [54] Griffiths D.J. *Introduction to Quantum Mechanics*. Prentice-Hall, Englewood Cliffs, New Jersey, 1995
- [55] Toffoli T. "Reversible Computing". In *Automata Languages and Programming*, J. W. de Bakker and J. van Leeuwen editors, Springer-Verlag, 1980, New York
- [56] Deutsch D. "Quantum Computational Networks". *Proceedings of the Royal Society, London A*, 1989, 425: 73-90
- [57] Divincenzo D. P. "Two-bit Gates Are Universal for Quantum Computation". *Phys. Rev. A*. 1995, 50: 1015-1019
- [58] Deutsch D. et al. "Universality in Quantum Computation". *Proc. R. Soc. London A*, 1995, 52: 3457-3462
- [59] Barenco A. et al. "Elementary Gates for Quantum Computation". *Phys. Rev. A*. 1995, 52: 3457-3461
- [60] Deutsch D. and Jozsa R. *Proc. R. Soc. London A*, 1992, 439: 553-560
- [61] Simon D. R. "On the Power of Quantum Computation". *Proceedings of 35th Annual Symposium on Foundations of Computer Science*, 1994, 116-123
- [62] Weitz M. "Towards Controlling Larger Quantum Systems: from Laser Cooling to Quantum Computing". *IEEE Journal of Quantum Electronics*, 2000, 36(12): 1346-1357
- [63] Zhou X., Habif J. L., Herr A.M., et al. "A Tipping Pulse Scheme for A rf-SQUID Qubit". *IEEE Transactions on Applied Superconductivity*, 2001, 11(1): 1018-1021
- [64] Sanders G. D., Kim K. W., Holton, W. C. "An Optically Driven Quantum Dot Quantum Computer". *Proceedings of 57th Annual Device Research Conference Digest*, 1999, 68-69

- [65] Monroe C., Itano W. M., Kielpinski D., et al. "Quantum Computing with Trapped Ions". Proceedings of Quantum Electronics and Laser Science Conference, 1999, 4-8
- [66] Kurtsiefer C. and DeVoe R. G. "Recent Experimental Results in Using  $^{137}\text{Ba}$  in A Miniaturized Ion Trap for Quantum Computation". Proceedings of Quantum Electronics Conference, 1998, 125-128
- [67] Shepherd G. M. *Neurobiology*. Oxford University Press, Oxford, 1988
- [68] Pribram K. H. "The Deep and Surface of Memory and Conscious Learning: toward A 21st Century Model". In Solso R. L. (Eds.) *Mind and Brain Science in The 21<sup>st</sup> Century*, MIT Press, Cambridge, 1997
- [69] Ohanian H.C. *Principles of Quantum Mechanics*. Prentice-Hall, Englewood Cliffs, New Jersey, 1990
- [70] Haken, H. "Synergetic Computers for Pattern Recognition, and Their Control by Attention Parameter". In *Neurocomputers and Attention II: Connectionism and Neurocomputers*, V.I. Kryukov and A. Holden (Eds.), Manchester University Press, UK, 1991, 551-556
- [71] Morel B. "Biologically Plausible Learning Rules for Neural Networks and Quantum Computing". *Neurocomputing*, 2000, 32-33: 921-926
- [72] Fahri E. and Gutman S. "Analog Analogue of A Digital Quantum Computation". *Phys. Rev. A*, 1998, 57: 2403-2407
- [73] Lusth J C. and Dixon B. "A Characterization of Important Algorithms for Quantum-Dot Cellular Automata". *Information Science*, 1999, 113: 193-204
- [74] Lent C. S., Tougaw P. D., Porod W. "Quantum Cellular Automata: the Physics of Computing with Arrays of Quantum Dot Molecules". Proceedings of Workshop on Physics and Computation, 1994, 5-13
- [75] Blick R. H. and Lorenz H. "Possible Definition of Quantum Bits in Coupled Quantum Dots". Proceedings of IEEE International Symposium on Circuits and Systems, 2000, 2: 245-248
- [76] Behrman E. C., Nash L. R., Steck J. E., et al. "Simulations of Quantum Neural Networks". *Information Sciences*, 2000, 128: 257-269
- [77] Feynman R. P. and Hibbs A R. *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill, 1965
- [78] Ventura D. "On the Utility of Entanglement in Quantum Neural Computing". Proceedings of International Joint Conference on Neural Networks, 2001, 2: 1565-1570
- [79] Liboff R.L. *Introductory Quantum Mechanics*. Addison-Wesley, Massachusetts, 1992

- [80] Loudon R. *The Quantum Theory of Light* (2<sup>nd</sup> Ed.). Oxford: Oxford University Press, 1983
- [81] 庄镇泉, 王煦法, 王东生. *神经网络与神经计算机*. 北京: 科学出版社, 1992
- [82] 松井伸之, 高井真人, 西村治彦. “量子画像ニューロンに基づく量子回路対応ネットワークモデル”, 电子情报通信学会论文志 A, 1998, J81-A (12): 1687-1692
- [83] T. Nitta. "An Extension of Back-Propagation Algorithm to Complex Numbers", *Neural Networks*, 1997,10(8): 1391-1416
- [84] George M. G. and Cris K. "Complex Domain Backpropagation". *IEEE Trans. on Circuit and Systems-II*, 1992, 30(5): 330-334
- [85] Robert M. et al. "Classification with Holographic Neural Networks". In Vito L. P., Branko S. and Giuseppe V. (Eds.) *Frontier Decision Support Concepts*. John Wiley and Sons, 1994, 94
- [86] Xiang Li, Cheng-Leong and Robert Gray. "An Intelligent Business Forecaster for Strategic Business Planning". *Journal of Forecasting*, 1999, 18: 181-204
- [87] Weinfurter H., Bouwmeester D., Daniell M., et al. "Quantum Communication and Entanglement". *Proceedings of the 2000 IEEE International Symposium on Circuits and Systems*, 2000, 2: 236-239, Geneva
- [88] 袁曾任. *人工神经网络及其应用*. 北京: 清华大学出版社, 1999
- [89] Purushothaman G. and Karayiannis N. B. "Quantum Neural Networks (QNNs): Inherently Fuzzy Feedforward Neural Networks". *IEEE Transactions on Neural Networks*, 1997, 8(3): 679-693
- [90] Abrams D. and Lloyd S. "Nonlinear Quantum Mechanics Implies Polynomial-time Solution for NP-complete and # P Problems". *Phys. Rev. Lett.*, 1998, 81(18): 3992-3995

## 致 谢

首先感谢我的导师庄镇泉教授，是他将我引入“量子神经计算”这一神奇而有趣的研究领域，使我大开眼界，并有机会接触到学科的前沿，给我带来了科研上的乐趣。此外，三年学习期间导师所表现出的治学态度和处世风格使我获益匪浅，并将终生受用。

感谢中国科大量子通讯与量子计算开放实验室的郭光灿教授、叶淮老师、王涛老师，计算机科学系的陈国良教授、孙广中同学，感谢他们给予我的教诲和无私帮助，并感谢他们给我提供了许多宝贵的学习机会。

感谢国外同仁 M. Perus 博士、A. Ezhov 教授、D. Ventura 博士、Li Weigang 博士、E. Behrman 博士、T. Martinez 教授和 M. Andrecut 博士有益的讨论和建议，以及所提供的相关文献和跨国界的友谊。

感谢同实验室的李斌、李海鹰、高清维、李军、顾理、狄晓龙、叶学义、张云超等同学平日里的帮助和有益的讨论，以及所提供的和谐的氛围。

感谢中国科大信息学院冯焕清院长、电子科学与技术系郭立主任、刘同怀书记、朱领娣老师，合肥工大理学院何晓雄院长、微电子学教研室刘声雷主任及理学院教务办公室的工作人员，感谢他们的理解和帮助。

最后，感谢我的全家人，特别是妻子在生活上的关心和精神上的支持。

## 在读期间所参加的科研项目汇总

- 1 国家自然科学基金项目：“量子神经网络模型与算法研究”（60171029）（2002-2004）（主要参加人）
- 2 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：“量子神经网络模型研究”（2000-2002）（主要参加人）
- 3 973 国家重点基础研究发展规划项目：“信息技术的基础理论与高性能软件研究”（子课题）（G1998030413）（1999-2003）（参加人）
- 4 教育部博士点基金项目：“面向 Internet 的智能信息检索算法研究”（1999035808）（2000-2002）（参加人）

## 在读期间已发表论文汇总

### 期刊论文

- 1 解光军, 庄镇泉, “量子神经网络”, *计算机科学*, 2001,28(7):1-6
- 2 解光军, 庄镇泉, “一种量子竞争学习算法”, *量子电子学报*, 已录用
- 3 解光军, 李斌, 庄镇泉, “量子系统实现神经计算的理论分析”, *电子与信息学报*, 已录用
- 4 解光军, 庄镇泉, “期权定价中的分步式预测模型”, *系统工程*, 2000,18(4):28-31
- 5 解光军, 庄镇泉, “差分 RBF 神经网络的预测算法及其应用”, *信息与控制*, 2000,29(5):421-424
- 6 解光军, 庄镇泉, 李斌, “基于进化 RBF 神经网络的非平稳时间序列分析”, *电路与系统学报*, 2001,6(2):1-4
- 7 解光军, 庄镇泉, 李海鹰, “一种进化 RBF 神经网络的模型及其训练算法”, *小型微型计算机系统*, 2001,22(11):1315-1317
- 8 李海鹰, 周辉, 解光军, 李动恒, 庄镇泉, “基于 EC-MAS 的半结构化信息处理模式研究”, *计算机工程*, 2001,27(12):15-17
- 9 李斌, 谭立湘, 解光军, 李海鹰, 庄镇泉, “非同步多时间序列中频繁模式的发现算法”, *软件学报*, 2002,13(3):410-416
- 10 高清维, 李斌, 解光军, 庄镇泉, “基于平稳小波变换的信号消噪分析”, *计算机研究与发展*, 已录用

### 会议论文

- 1 Xie Guang-jun, Zhuang Zhen-quan, "The Artificial Neural Network Model Based on Quantum Gates Cell and Its Applications", Proceedings of the 2002 International Conference on Control and Automation, Xiamen University, Xiamen, 2002, (IEEE published, ISBN:0-7803-7413-4)
- 2 Xie Guang-jun, Ye Xue-yi, Gao Qing-wei, Zhuang Zhen-quan, "A Quantum Neural Network Model Based-on Qubit-like Neuron", Proceedings of the International Symposium on Quantum Information, USTC, Huangshan, 2001,51-52
- 3 Li Bin, Zhuang Zhen-quan, Xie Guang-jun, "Quantum Genetic Algorithm and Its Applications", Proceedings of the International Symposium on Quantum Information, USTC, Huangshan, 2001,35
- 4 解光军, 庄镇泉, “量子理论与人工神经网络结合的研究进展”, 中国电子学会电路与系统第 16 届学术年会论文集, 宁波大学, 宁波, 2001,220-224
- 5 解光军, 庄镇泉, “计算智能方法在金融市场分析中的应用”, 经济物理学高级研讨班及金融复杂性国际学术讨论会论文集, 中国科学技术大学, 合肥, 2000,89-97