

Analysis and Design
of Feedforward Neural Networks

前馈神经网络分析与设计

乔俊飞 韩红桂◎著



科学出版社

前馈神经网络分析与设计

Analysis and Design of Feedforward Neural Networks

乔俊飞 韩红桂 著

科学出版社

北京

内 容 简 介

本书系统地论述了前馈神经网络的主要理论、设计基础及应用实例,旨在使读者了解神经网络的发展背景和研究对象,理解和熟悉它的基本原理和主要应用,掌握它的结构模型和设计应用方法,特别是前馈神经网络的参数学习算法和结构设计方法,为深入研究和应用开发打下基础。为了便于读者理解,书中尽量避免烦琐的数学推导,加强了应用举例,并在内容的选择和编排上注意到读者初次接触新概念的易接受性和思维的逻辑性。作为扩充知识,书中还介绍了前馈神经网络的基本概念、体系结构、控制特性及信息模式。

本书适合高校控制与信息类专业研究生、智能科学技术专业本科生以及各类科技人员阅读。

图书在版编目(CIP)数据

前馈神经网络分析与设计=Analysis and Design of Feedforward Neural Networks/乔俊飞,韩红桂著. —北京:科学出版社,2013

ISBN 978-7-03-033593-7

I. ①前… II. ①乔… ②韩… III. ①前馈-人工神经网络-研究
IV. ①TP183

中国版本图书馆 CIP 数据核字 (2012) 第 025536 号

责任编辑:钱 俊 / 责任校对:包志虹
责任印制:钱玉芬 / 封面设计:东方人华

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2013年1月第 一 版 开本: B5 (720×1000)

2013年1月第一次印刷 印张: 18 1/2

字数: 373 000

定价: 78.00 元

(如有印装质量问题,我社负责调换)

总 序

“211工程”是中华人民共和国成立以来教育领域唯一的国家重点建设工程，面向21世纪重点建设一百所高水平大学，使其成为我国培养高层次人才，解决经济建设、社会发展和科技进步重大问题的基地，形成我国高等学校重点学科的整体优势，增强和完善国家科技创新体系，跟上和占领世界高层次人才培养和科技发展的制高点。

中国高等教育发展迅猛，尤其是1400所地方高校已经占全国高校总数的90%，成为我国高等教育实现大众化的重要力量，成为区域经济和社会发展服务的重要生力军。

在北京市委、市政府的高度重视和大力支持下，1996年12月北京工业大学通过了“211工程”部门预审，成为北京市属高校唯一进入国家“211工程”重点建设的百所大学之一。北京工业大学紧紧抓住“211工程”建设和举办2008年奥运的重要机遇，实现了两个历史性的转变：一是实现了从单科性大学向以工科为主，理、工、经、管、文、法相结合的多科性大学的转变；二是实现了从教学型大学向教学研究型大学的转变。“211工程”建设对于北京工业大学实现跨越式发展、增强服务北京的能力起到了重大的推动作用，学校在学科建设、人才培养、科学研究、服务北京等方面均取得了显著的成绩，综合实力和办学水平得到了大幅度的提升。

至2010年底，北京工业大学的学科门类已经覆盖了8个：工学、理学、经济学、管理学、文学、法学、哲学和教育学。现拥有8个一级学科博士学位授权点、37个二级学科博士学位授权点、15个博士后科研流动站、15个一级学科硕士学位授权点和81个二级学科硕士学位授权点；拥有6种类型硕士研究生专业学位授权资格，工程硕士培养领域19个，拥有3个国家重点学科、16个北京市重点学科和18个北京市重点建设学科。

目前，学校有专职教师1536人、全职两院院士5名、博士研究生导师220人、有正高职称294人和副高职称580人，专任教师中具有博士学位教师的比例达到54.6%。有教育部“长江学者”特聘教授4人、国家杰出青年基金获得者6人、入选中组部“千人计划”1人、北京市“海聚工程”3人、教育部新(跨)世纪优秀人才支持计划15人。

2010年学校的到校科研经费为6.2亿元。“十一五”期间，学校承担了国家科技重大专项28项、“973计划”项目16项、“863计划”项目74项、国家杰出青年基金2项、国家自然科学基金重点项目8项、科学仪器专项2项、重大国际合作项目

1 项、面上和青年基金项目 347 项、北京市自然科学基金项目 180 项、获国家级奖励 14 项。现有 1 个共建国家工程研究中心、7 个部级或省部共建科研基地、11 个北京市重点实验室和 3 个行业重点实验室。

为了总结和交流北京工业大学“211 工程”建设的科研成果,学校设立了“211 工程”专项资金,用于资助出版系列学术专著。这些专著从一个侧面代表了北京工业大学教授、学者的学科方向、研究领域、学术成果和教学经验。

展望未来,我们任重而道远。我坚信,只要珍惜“211 工程”建设的重要机遇,构建高层次学科体系,营造优美的大学校园,北京工业大学在建设国际知名、有特色、高水平大学的进程中就一定能够为国家,特别是为北京市的经济建设和社会发展作出更大的贡献。

中国工程院院士
北京工业大学原校长 左铁镛

2011 年 6 月

前 言

在人类几千年的文明发展史中,人们始终在探索人类自身高级智能的奥秘。人们从认知科学、生物学、生物物理和生物化学、医学、数学、信息与计算科学等领域进行广泛的探索和研究。在这个过程中逐步形成了一门具有广泛学科交叉特点的学科——“人工神经网络”(artificial neural network, ANN)。它力图构建“人造”的生物神经细胞(即神经元)和神经网络,在不同程度和不同层次上实现人脑神经系统在信息处理、学习、记忆、知识的存储和检索方面的功能。随着生产力发展水平的提高和实验手段的进步,人们在这个技术领域的各个方面都取得了巨大的进步。但是由于人脑结构和运行机理无比的复杂性,应该说到目前为止,人们对人脑活动的深层次机理和规律的认识还是相当粗浅的。

人工神经网络包括神经网络模型结构与神经网络学习算法,是在细胞的水平上模拟脑结构和脑功能的科学。人工神经网络模型结构与人工神经网络学习算法两者相互联系,人工神经网络模型结构是人工神经网络学习算法的前提,而人工神经网络学习算法是人工神经网络模型结构中神经信息运动或演化的过程。人工神经网络的中心目标是,在神经细胞的水平上,模拟生物神经系统的结构特征和生物神经信息的演化规律,构造人工神经网络模型结构,并建立那些在人工神经网络模型结构中有效的人工神经网络学习算法。

从20世纪40年代M-P神经元模型的提出开始,人工神经网络的发展过程可谓是一波三折。1965年Minsky和Papert的《感知机》使得人工神经网络的研究停滞了10余年,直到80年代初误差反向传播算法等的提出,人工神经网络的研究才步入恢复期。时至今日,人工神经网络系统研究的重要意义已经得到广泛承认,涉及电子科学与技术、信息与通信工程、计算机科学与技术、电气工程、控制科学与技术等诸多学科,其应用领域包括:建模、时间序列分析、模式识别和控制等,并在不断拓展。可以说人工神经网络作为目前非线性科学和计算智能研究的主要内容之一,已经成为一种解决许多实际问题的必要的技术手段。

本书系统地论述了前馈神经网络的主要理论、设计基础及应用实例,旨在使读者了解前馈神经网络的发展背景和研究对象,理解和熟悉它的基本原理和主要应用,掌握它的结构模型和设计应用方法,特别是前馈神经网络的参数学习算法和结构设计方法。为了便于读者理解,书中尽量避免烦琐的数学推导,加强了应用举例,并在内容的选择和编排上注意到读者初次接触新概念的易接受性和思维的逻辑性。在编写过程中,有许多内容取材于最近的国内外文献。各章后都附有较多

的参考文献,以便读者查阅。

本书第1章为绪论,主要讲述了前馈神经网络结构和学习算法的发展历史和背景,以及编写本书的动机,可作为后续章节的阅读基础。本书包括以下3个部分。

(1) 第2章~第4章构成了本书的第一部分,主要介绍前馈神经网络的一些经典方法。具体介绍如下:

第2章描述感知器神经网络的结构以及学习算法。首先介绍单神经元的结构和表述,其次介绍了感知器神经网络的结构,分别介绍单隐含层和多隐含层感知器神经网络,最后分析感知器神经网络的学习算法:BP算法和BP算法的改进算法。重点介绍感知器神经网络的结构特点和学习算法。

第3章详细讲述了RBF神经网络。首先讨论了RBF神经网络的原理,给出了RBF神经网络的结构。其次,基于RBF神经网络结构介绍其学习算法,讨论了采用插值理论的RBF神经网络隐含层神经元中心学习算法。最后,讨论了神经网络隐含层和输出层连接权值的学习算法。

第4章详细介绍了模糊神经网络。首先,详细讨论了模糊推理系统,主要对模糊集合、隶属函数和模糊运算进行阐述。其次,基于模糊系统和人工神经网络的理论基础引入模糊神经网络,并对一种典型性的模糊神经网络结构进行了剖析。然后,基于模糊神经网络结构介绍其学习算法,讨论了标准型模糊神经网络通常采用BP算法(FBP)。

这三章通过对不同概念的介绍揭示了其共同特点:它们都是前馈神经网络。更为重要的是,它们从各自的角度深入、细致地讨论了学习过程的深层知识——这一特征将在后续章节中进一步探讨。

(2) 第二部分包括第5章和第6章,讨论了前馈神经网络的两种有效的学习算法——快速下降算法和改进型递归最小二乘算法。

第5章介绍一种用于结构动态设计的前馈神经网络学习算法——快速下降算法。快速下降算法主要与参数修改项、学习率、隐含层输出以及当前神经网络输出误差有关,避免了求解导数的过程,减少运算量,提高了人工神经网络的训练速度。

第6章基于递归最小二乘算法提出了一种改进型递归最小二乘算法,并且对该算法进行了理论和实验分析。最后将基于递归最小二乘算法的RBF神经网络应用于非线性函数逼近、双螺旋模式分类以及污泥膨胀预测。

这两章描述了两种前馈神经网络的两种有效的学习算法——快速下降算法和改进型递归最小二乘算法,这两种学习算法不但适合固定结构的前馈神经网络,而且满足人工神经网络结构动态调整的需要。

(3) 第7章~第10章构成本书的第三部分,讨论自组织前馈神经网络。从第7章开始介绍了四种自组织神经网络。具体介绍如下:

第7章描述了一种基于显著性分析的快速修剪型感知器神经网络。以规模过大的感知器神经网络为研究对象,构造误差曲面的模型,分析网络连接权值的扰动对网络输出误差所造成的影响,进行隐含层神经元显著性分析,直接剔除冗余的隐含层神经元实现神经网络结构自组织设计。基于显著性分析的神经网络快速修剪方法适用于规模过大的神经网络结构设计。实验结果表明,快速修剪算法与常规的最优脑外科算法等修剪算法相比,具有更简单的网络结构和更快的学习速度。

第8章描述了一种增长-修剪型多层感知器神经网络。介绍了一种神经网络输出敏感度分析方法,以感知器神经网络为研究对象,进而获得一种增长-修剪型感知器神经网络。增长-修剪型感知器神经网络通过分析隐含层神经元输出对神经网络输出的贡献,对贡献太小的神经元予以删除,对贡献值较大的神经元利用最邻近法在其附近插入新的神经元,从而调整神经网络结构,并利用快速下降算法修改神经网络连接权值,实现了感知器神经网络的结构和参数自校正。通过对非线性函数逼近、数据聚类,以及污水处理过程关键参数预测证明了该方法的有效性。

第9章描述了一种弹性 RBF 神经网络。针对 RBF 神经网络的结构设计问题,介绍了一种神经元修复准则。基于神经元的活跃度以及神经元修复准则,判断增加或删除 RBF 神经网络隐含层中的神经元,获得一种弹性 RBF 神经网络。利用快速下降算法修改神经网络连接权值,快速下降算法保证了最终 RBF 网络的精度。弹性 RBF 神经网络实现了神经网络的结构和参数自校正,解决了 RBF 神经网络结构过大或过小的问题,并给出了神经网络结构动态变化过程中收敛性证明。通过对非线性函数的逼近、非线性系统的建模以及污水处理过程溶解氧(DO)的预测控制,结果证明了弹性 RBF 神经网络具有良好的自适应能力和逼近能力,尤其是在泛化能力、最终网络结构等方面较之其他自组织,RBF 神经网络有较大的提高。

第10章描述了一种自组织模糊神经网络。针对模糊神经网络的结构设计,介绍了一种模糊神经网络结构自组织设计方法,自组织模糊神经网络利用神经网络输出敏感度方法分析规则化层神经元的敏感度,判断规则层中需要增加或删除的神经元,并优化网络中的模糊规则。同时,自组织模糊神经网络利用改进型递归最小二乘算法修改神经网络连接权值,并给出了其结构动态变化过程中收敛性证明。利用自组织模糊神经网络对非线性系统建模、Mackey-Glass 混沌系统预测、污水处理过程关键水质参数预测以及污水处理 DO 控制,结果证明了自组织模糊神经网络具有良好的自适应能力和逼近能力。

这四章通过对三种典型的前馈神经网络——感知器神经网络、RBF 神经网络、模糊神经网络的结构自组织神经和参数自调整进行讨论,运用不同的自组织方法实现了各自的自组织。这部分内容是本书的一个重点;更为重要的是,这部分内容也是作者近几年的工作积累。

作者在本书写作过程中特别注重基础知识的积累,增加了基础的介绍,使得读者能够较快理解前馈神经网络结构基础、前馈神经网络学习算法基础以及数学基础;介绍两种有效的前馈神经网络学习算法,并重点分析了其设计方法,以便读者通过学习与练习获得独立设计前馈神经网络算法的能力;主要介绍了作者取得较大突破的前馈神经网络结构设计方法,这些设计方法已经取得业界的认可,并且都是近几年的研究成果,以便读者通过学习获得较新的知识;在内容的选择和编排上注意到读者初次接触新概念的易接受性和思维的逻辑性,力求深入浅出,自然流畅。

本书可作为电子科学与技术、信息与通信工程、计算机科学与技术、电气工程、控制科学与技术等专业研究生和高年级本科生的教材,同时对有关专业领域的研究人员和工程技术人员也有重要的参考价值。

本书部分工作得到了国家自然科学基金重点项目(No. 61034008)、国家自然科学基金项目(No. 61203099, No. 60873043)、北京市自然科学基金项目(No. 4122006, No. 4092010)、北京市创新人才建设计划项目(No. PHR201006103)、教育部新世纪优秀人才支持计划项目(No. NCET-08-0616)、教育部博士点基金项目(No. 200800050004)的支持。在此表示衷心的感谢!

本书第1章~第4章由乔俊飞编写,第5章~第10章由韩红桂编写。北京工业大学博士研究生杨刚为本书的编写出版做了大量工作,在此表示衷心的感谢。同时,感谢参与本书前期准备的全体人员,他们是高学金、于建均、李民爱、柴伟、武利等老师,张昭昭、博迎春、陈启丽、李荣、韩广等博士研究生,没有他们的辛勤工作,本书的编写是无法顺利完成的;还要感谢北京工业大学电子信息与控制工程学院的支持。

限于作者的思想境界和学术水平,另外人工神经网络本身也在不断得到丰富和发展,书中不妥之处在所难免,恳请广大读者批评指正。

乔俊飞 韩红桂

2012年2月15日

目 录

总序

前言

| | |
|----------------------------|----|
| 第 1 章 绪论 | 1 |
| 1.1 引言 | 1 |
| 1.2 神经网络及其发展 | 2 |
| 1.2.1 神经网络的定义 | 2 |
| 1.2.2 神经网络的功能 | 3 |
| 1.2.3 神经网络的发展 | 4 |
| 1.2.4 神经网络的应用 | 7 |
| 1.3 人工神经网络的结构设计 | 9 |
| 1.3.1 人工神经网络的结构 | 9 |
| 1.3.2 前馈神经网络结构设计研究现状 | 10 |
| 1.4 本书主要内容..... | 13 |
| 1.4.1 神经网络参数学习算法研究 | 14 |
| 1.4.2 神经网络结构设计方法研究 | 14 |
| 1.4.3 自组织神经网络结构算法研究 | 15 |
| 1.4.4 应用研究..... | 15 |
| 参考文献 | 17 |
| 第 2 章 感知器神经网络 | 23 |
| 2.1 引言..... | 23 |
| 2.2 感知器神经网络分析..... | 23 |
| 2.2.1 单神经元分析 | 24 |
| 2.2.2 单层感知器神经网络 | 25 |
| 2.2.3 多层感知器神经网络 | 28 |
| 2.3 感知器神经网络学习算法..... | 30 |
| 2.3.1 隐含层与输出层之间的权值修正 | 33 |
| 2.3.2 输入层与隐含层之间的权值修正 | 34 |
| 2.3.3 BP 算法的改进 | 36 |
| 2.4 本章小结..... | 43 |
| 附录 A 数学基础 | 44 |

| | |
|---------------------------------|-----------|
| 附录 A.1 泰勒引理 | 44 |
| 附录 A.2 泰勒定理和推论 | 46 |
| 参考文献 | 49 |
| 第3章 RBF神经网络 | 53 |
| 3.1 引言 | 53 |
| 3.2 RBF神经网络原理 | 54 |
| 3.2.1 插值计算 | 54 |
| 3.2.2 模式可分性 | 56 |
| 3.2.3 正规化法则 | 57 |
| 3.2.4 RBF神经网络结构 | 58 |
| 3.3 RBF神经网络学习算法 | 59 |
| 3.3.1 中心值学习策略 | 59 |
| 3.3.2 隐含层和输出层连接权值学习策略 | 63 |
| 3.4 本章小结 | 67 |
| 附录 B 数学运算 | 67 |
| 附录 B.1 域和向量空间 | 67 |
| 附录 B.2 矩阵的表示和运算 | 69 |
| 附录 B.3 矩阵的性质 | 70 |
| 附录 B.4 矩阵范数的运算 | 75 |
| 参考文献 | 76 |
| 第4章 模糊神经网络 | 80 |
| 4.1 引言 | 80 |
| 4.2 模糊推理系统描述 | 81 |
| 4.2.1 模糊集合与隶属函数 | 81 |
| 4.2.2 模糊运算 | 83 |
| 4.3 模糊神经网络结构 | 87 |
| 4.4 模糊神经网络学习算法 | 89 |
| 4.5 本章小结 | 93 |
| 参考文献 | 94 |
| 第5章 前馈神经网络快速下降算法研究 | 98 |
| 5.1 引言 | 98 |
| 5.2 神经网络学习 | 98 |
| 5.2.1 神经网络结构及信息处理 | 98 |
| 5.2.2 神经网络学习算法分析 | 100 |
| 5.3 快速下降算法 | 101 |

| | | |
|------------|-----------------------------|------------|
| 5.3.1 | 快速下降算法描述 | 101 |
| 5.3.2 | 快速下降算法收敛性分析 | 104 |
| 5.4 | 仿真研究 | 108 |
| 5.4.1 | 感知器神经网络仿真研究 | 109 |
| 5.4.2 | RBF神经网络仿真研究 | 112 |
| 5.5 | 本章小结 | 117 |
| | 参考文献 | 118 |
| 第6章 | 前馈神经网络改进型递归最小二乘算法研究 | 123 |
| 6.1 | 引言 | 123 |
| 6.2 | 递归最小二乘算法 | 124 |
| 6.2.1 | 递归最小二乘算法描述 | 124 |
| 6.2.2 | 递归最小二乘算法分析 | 128 |
| 6.3 | 改进型递归最小二乘算法 | 129 |
| 6.3.1 | 改进型递归最小二乘算法描述 | 130 |
| 6.3.2 | 改进型递归最小二乘算法收敛性分析 | 131 |
| 6.4 | 改进型递归最小二乘算法的应用 | 132 |
| 6.4.1 | 非线性函数逼近 | 133 |
| 6.4.2 | 双螺旋模式分类 | 135 |
| 6.4.3 | 污泥膨胀预测 | 137 |
| 6.5 | 本章小结 | 142 |
| | 参考文献 | 143 |
| 第7章 | 基于显著性分析的快速修剪型感知器神经网络 | 148 |
| 7.1 | 引言 | 148 |
| 7.1.1 | 增长型神经网络 | 148 |
| 7.1.2 | 修剪型神经网络 | 152 |
| 7.2 | 显著性分析 | 153 |
| 7.2.1 | 误差曲面分析 | 153 |
| 7.2.2 | 显著性分析算法 | 155 |
| 7.3 | 基于显著性分析的快速修剪算法 | 156 |
| 7.3.1 | 多层感知器神经网络 | 157 |
| 7.3.2 | 多层感知器神经网络快速修剪算法 | 158 |
| 7.3.3 | 仿真研究 | 160 |
| 7.4 | 本章小结 | 167 |
| | 参考文献 | 168 |

| | |
|------------------------------------|-----|
| 第 8 章 增长-修剪型多层感知器神经网络 | 174 |
| 8.1 引言 | 174 |
| 8.2 敏感度计算 | 175 |
| 8.2.1 敏感度分析方法的分类 | 175 |
| 8.2.2 敏感度分析方法 | 178 |
| 8.2.3 敏感度计算 | 181 |
| 8.3 神经网络输出敏感度分析 | 182 |
| 8.3.1 敏感度分析的频域研究 | 182 |
| 8.3.2 神经网络输出敏感度分析 | 187 |
| 8.4 增长-修剪型多层感知器神经网络分析 | 189 |
| 8.4.1 隐含层神经元的敏感度 | 189 |
| 8.4.2 神经元增长和修剪 | 192 |
| 8.4.3 增长-修剪型感知器神经网络 | 193 |
| 8.4.4 收敛性分析 | 195 |
| 8.5 增长-修剪型多层感知器神经网络应用 | 197 |
| 8.5.1 非线性函数逼近 | 197 |
| 8.5.2 数据分类 | 199 |
| 8.5.3 生化需氧量软测量 | 202 |
| 8.6 本章小结 | 208 |
| 参考文献 | 210 |
| 第 9 章 弹性 RBF 神经网络 | 215 |
| 9.1 引言 | 215 |
| 9.2 RBF 神经网络描述 | 216 |
| 9.3 弹性 RBF 神经网络 | 217 |
| 9.3.1 神经元修复准则 | 217 |
| 9.3.2 神经网络结构优化设计 | 219 |
| 9.3.3 弹性 RBF 神经网络 | 221 |
| 9.3.4 收敛性分析 | 221 |
| 9.4 弹性 RBF 神经网络应用 | 224 |
| 9.4.1 非线性函数逼近 | 224 |
| 9.4.2 非线性系统建模 | 227 |
| 9.4.3 溶解氧模型预测控制 | 231 |
| 9.5 本章小结 | 238 |
| 附录 C 熵 | 239 |
| 附录 C.1 熵的概念 | 239 |

| | |
|------------------------------------|------------|
| 附录 C.2 互信息 | 240 |
| 参考文献 | 242 |
| 第 10 章 自组织模糊神经网络 | 248 |
| 10.1 引言 | 248 |
| 10.2 模糊神经网络 | 249 |
| 10.3 自组织模糊神经网络分析 | 253 |
| 10.3.1 模糊神经网络结构优化 | 253 |
| 10.3.2 模糊神经网络自组织设计算法 | 256 |
| 10.3.3 收敛性分析 | 258 |
| 10.4 自组织模糊神经网络应用 | 260 |
| 10.4.1 非线性系统建模 | 260 |
| 10.4.2 Mackey-Glass 时间序列系统预测 | 263 |
| 10.4.3 污水处理关键水质参数预测 | 266 |
| 10.4.4 污水处理过程溶解氧控制 | 269 |
| 10.5 本章小结 | 272 |
| 参考文献 | 273 |
| 索引 | 278 |

第 1 章 绪 论

1.1 引 言

脑和神经系统是人体结构、功能中最复杂的系统,随着分子生物学、细胞生物学的发展,大规模开展脑研究成为可能^[1]。众所周知,数字计算机具有很强的计算和信息处理能力,但是它对于模式识别、感知以及在复杂环境中作决策等问题的处理能力却远不如人。神经生理学研究结果表明,人的智能主要取决于大脑皮层,而大脑皮层是一个大规模互连的生物神经网络^[2]。探求大脑的组织结构和运行机制,从模仿人脑智能的角度出发,寻求新的信息处理方法是当前人工智能领域的研究热点问题。国家中长期科学和技术发展规划纲要中也明确指出:脑功能的细胞和分子机理、脑学习记忆和思维等高级认知功能的过程及其神经基础、脑信息表达等研究方向属于当前的科学前沿问题^[3]。

高性能、低成本、普适计算和智能化等是当前信息科学发展的主要方向,寻求新的计算与处理方式和物理实现是未来信息技术领域面临的重大挑战^[3]。人工神经网络的研究正是在与传统计算方法挑战的过程中得以发展、壮大,目前已经成为人工智能领域中最活跃的研究方向之一。一般来说,神经网络主要研究 ABC² 等问题,即人工神经网络(artificial neural network, ANN)、生物神经网络(biological neural network, BNN)、认知科学(cognitive science)和混沌(chaos)^[4]。人工神经网络的研究主要集中在神经网络结构和学习算法的研究两方面^[5]。近年来随着脑科学、神经生物学的发展,人工神经网络结构的研究开始全面向生物神经系统靠拢。

人工神经网络是智能科学的重要组成部分,已经成为脑科学、认知科学、计算机科学、数学和物理学等学科共同关注的焦点。其应用研究已经渗透到工业、农业、国防、航空等领域,并且在信号处理、智能控制、模式识别、图像处理、非线性优化、知识处理等方面取得了令人鼓舞的进展^[6~11]。

从 20 世纪 80 年代初复苏以来,人工神经网络在计算能力、对任意连续映射的逼近能力、学习理论以及动态网络稳定性分析等方面都取得了丰硕的成果^[12]。在结构设计方面,除了经典的神经网络外却鲜有突破性进展,但这方面的工作已引起了学者们的广泛关注^[13~15]。神经网络实际应用的需求驱动了其理论研究的发展,其实每一个成功的应用都需要对神经网络进行精心设计。可见,人工神经网络结

构优化设计是神经网络成功应用的核心技术,对其展开研究也是神经网络推广应用的客观需要。

人工神经网络就是借鉴生物神经网络结构和生物神经元工作机理,在一定程度上模拟人脑功能的信息处理系统。人工神经网络已经在模式识别、组合优化、函数逼近、智能控制、过程建模等方面得到成功应用,应用前景不容置疑。但目前的神经网络多数是通过足够的设计经验和充足的数据确定其结构,且神经网络结构一旦确定之后将不再调整。对一些工况变化不大、动态特性比较平稳的任务,理论和实践都已经证明通过调整神经网络参数可以满足实际需要。但是对于工况变化异常剧烈、动态特性呈现出很强的非线性的任务,其效果往往不佳。传统的固定结构人工神经网络性能仅仅由参数学习算法提供,在工作过程中,只是通过改变神经网络的参数以适应任务的变化。而神经计算领域的研究结果显示生物神经网络之所以有如此强大的信息处理能力,与生物神经网络的结构有很大的关系,生物神经网络的信息传输和信息处理能够根据信息量和复杂度进行自组调整神经网络结构连接方式。为了进一步推进人工神经网络对人脑功能的模拟,解决人工神经网络结构动态优化设计的问题,本书详细介绍了几种典型前馈神经网络;并描述了几种前馈神经网络学习算法;最后分析了神经网络性能与神经网络结构之间的联系,介绍了前馈神经网络结构生长和削减的演化机制,基于此获得几种自组织前馈神经网络。

1.2 神经网络及其发展

1.2.1 神经网络的定义

人工神经网络,简称神经网络,是由大量简单的处理单元——人工神经元(artificial neuron)互相连接而组成的一个高度非线性、并行的自适应的信息处理系统。神经网络是一个非线性的动力学系统,打破了传统的串行处理计算机的局限,以并行分布式存储和处理信息,尽管单个神经元的结构和功能都比较简单,但大量的神经元组合起来却具有强大的处理问题的能力。神经网络旨在模仿人脑或生物的信息处理系统,是对人脑功能的一种模仿与简化,具有学习、记忆、联想、类比、计算以及智能处理的能力,是现代神经科学研究与工程技术相结合的产物。神经网络理论的开创与发展,对智能科学和信息技术的发展产生了重大的影响和积极的推动作用。

1943年美国神经生理学家 McCulloch 和 Pitts 提出的第一个神经网络模型 M-P 模型^[16],开创了微观人工智能的研究工作,奠定了人工神经网络发展的基础。人工神经网络经过几十年的发展,无论是在理论研究还是在工程应用方面都取得了较为丰富的科研成果。

1.2.2 神经网络的功能

神经网络是通过对人类大脑结构和功能的模拟建立起来的一个非线性、自适应的高级信息处理系统。它是现代神经科学研究与工程技术应用相结合的产物，通过对大脑的模拟进行信息处理。神经网络具有强大的计算能力，是由其本身大规模的并行分布式结构和较好的学习能力以及由此延伸而来的泛化能力决定。神经网络具有非线性、分布/处理、容错性、自适应性等显著特点，如图 1-1 所示。

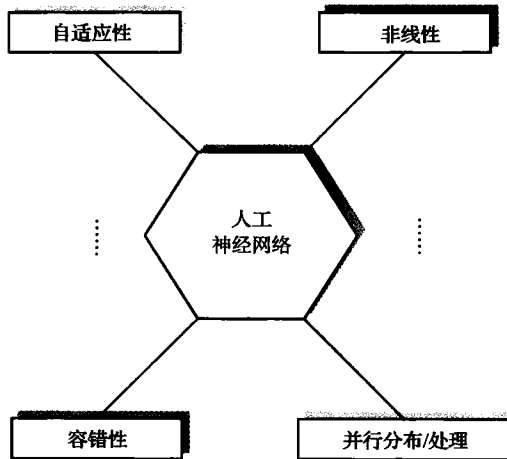


图 1-1 神经网络的特点

1. 非线性

神经网络的单个处理单元——人工神经元可以是线性或非线性的，但是由此互相连接而成的神经网络本身却是非线性的。此外，非线性是一种分布于整个网络的特殊性质。因此，神经网络具有非线性映射能力，且理论研究已经表明一个三层的神经网络能够以任意精度逼近非线性系统。

2. 并行分布/处理

神经网络是为模拟大脑的结构和功能而建立的一种数学模型，大量的人工神经元相互连接成一个高度并行的非线性动力学系统。尽管单个人工神经元的功能都十分简单，但大量神经元的并行活动使得整个网络呈现出较强大的处理能力。神经网络中信息的存储体现在神经元之间互相连接的并行分布结构上，进而使得信息的处理必然采用大规模的并行分布方式进行，即神经网络中信息的存储和处理是在整个网络中同时进行的，信息不是存储在神经网络中的某个局部，而是分布