

Neural Network Design

# 神经网络设计


HZ BOOKS  
华章教育

国外经典教材

Classical Texts From Top Universities

Martin T. Hagan  
(美) Howard B. Demuth 著  
Mark H. Beale

戴葵 等译  
李伯民 审校

 机械工业出版社  
China Machine Press





北京华章图文信息技术有限公司

国外经典教材



Classical Texts From Top Universities

# 神经网络设计

## Neural Network Design

Martin T. Hagan

(美) Howard B. Demuth 著

Mark H. Beale

戴葵 等译

李伯民 审校



机械工业出版社  
China Machine Press

本书主要讲述神经网络的基本概念,介绍实用的网络模型、学习规则和训练方法。全书分 19 章,内容涵盖神经元模型和网络结构、感知机学习规则、有监督的 Hebb 学习、Widrow - Hoff 学习算法、反向传播算法及其变形、联想学习、竞争网络、Grossberg 网络、自适应谐振理论和 Hopfield 网络。书中注重对数学分析方法和性能优化的讨论,强调神经网络在模式识别、信号处理以及控制系统等实际工程问题中的应用。同时本书包含大量例题、习题,并配有基于 MATLAB 软件包的“神经网络设计演示”程序。本书可以作为大学高年级本科生或一年级研究生的神经网络课程教材,也可供从事相关研究工作的科技人员参考。

Martin T. Hagan, Howard B. Demuth, Mark Beale: Neural Network Design (ISBN 0-534-94332-2).

Original edition copyright © 1996 by PWS Publishing Company. All rights reserved.

First published by PWS Publishing Company, an imprint of Thomson Learning, United States of America. Simplified Chinese edition published by Thomson Learning Asia and China Machine Press under the authorization of Thomson Learning. No part of this book may be reproduced in any form without the express written permission of Thomson Learning Asia and China Machine Press.

本书中文简体字版由汤姆森学习出版社与机械工业出版社合作出版。未经出版者书面许可,不得以任何方式复制或抄袭本书内容。

版权所有,侵权必究。

本书版权登记号:图字:01-1999-2361

### 图书在版编目(CIP)数据

神经网络设计/(美)哈根(Hagan, M. T.)等著;戴葵等译.-北京:机械工业出版社, 2002.9

(国外经典教材)

书名原文:Neural Network Design

ISBN 7-111-07585-4

I. 神… II. ①哈…②戴… III. 人工神经网络-系统设计 IV. TP183

中国版本图书馆 CIP 数据核字(2002)第 051841 号

机械工业出版社(北京市西城区百万庄大街 22 号 邮政编码 100037)

责任编辑:李伯民

北京第二外国语学院印刷厂印刷·新华书店北京发行所发行

2002 年 9 月第 1 版第 1 次印刷

787mm × 1092mm 1/16 · 30 印张

印数:0 001-4000 册

定价:49.00 元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换

# 出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭橥了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短、从业人员较少的现状下，美国等发达国家在其计算机科学发展的几十年间积淀的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章图文信息有限公司较早意识到“出版要为教育服务”。自1998年始，华章公司就将工作重点放在了遴选、移译国外优秀教材上。经过几年的不懈努力，我们与Prentice Hall, Addison-Wesley, McGraw-Hill, Morgan Kaufmann等世界著名出版公司建立了良好的合作关系，从它们现有的数百种教材中甄选出Tanenbaum, Stroustrup, Kernighan, Jim Gray等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及收藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专诚为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍，为进一步推广与发展打下了坚实的基础。

随着学科建设的初步完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都步入一个新的阶段。为此，华章公司将加大引进教材的力度，在“华章教育”的总规划之下出版三个系列的计算机教材：针对本科生的核心课程，剔抉外版菁华而成“国外经典教材”系列；对影印版的教材，则单独开辟出“经典原版书库”；定位在高级教程和专业参考的“计算机科学丛书”还将保持原来的风格，继续出版新的品种。为了保证这三套丛书的权威性，同时也为了更好地为学校和老师服务，华章公司聘请了中国科学院、北京大学、清华大学、国防科技大学、复旦大学、上海交通大学、南京大学、浙江大学、中国科技大学、哈尔滨工业大学、西安交通大学、中国人民大学、北京航空航天大学、北京邮电大学、中山大学、解放军理工大学、郑州大学、湖北工学院、中国国家信息安全测评认证中心等国内重点大学和科研机构在计算机的各个领域的著名学者组成“专家指导委员会”，为我们提供选题意见和出版监督。

“国外经典教材”是响应教育部提出的使用外版教材的号召，为国内高校的计算机本科教学度身订造的。在广泛地征求并听取丛书的“专家指导委员会”的意见后，我们最终选定了这20多种篇幅内容适度、讲解鞭辟入里的教材，其中的大部分已经被M.I.T.、Stanford、U.C. Berkley、C.M.U.等世界名牌大学采用。丛书不仅涵盖了程序设计、数据结构、操作系统、计算机体系结构、数据库、编译原理、软件工程、图形学、通信与网络、离散数学等国内大学计算机专业普遍开设的核心课程，而且各具特色——有的出自语言设计者之手、有的历三十年而不衰、有的已被全世界的几百所高校采用。在这些圆熟通博的名师大作的指引之下，读者必将在计算机科学的宫殿中由登堂而入室。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑。这些因素使我们的图书有了质量的保证，但我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。教材的出版只是我们的后续服务的起点。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方式如下：

电子邮件：[hzedu@hzbook.com](mailto:hzedu@hzbook.com)

联系电话：(010) 68995265

联系地址：北京市西城区百万庄南街1号

邮政编码：100037

# 专家指导委员会

(按姓氏笔画顺序)

尤晋元	王 珊	冯博琴	史忠植	史美林
石教英	吕 建	孙玉芳	吴世忠	吴时霖
张立昂	李伟琴	李师贤	李建中	杨冬青
邵维忠	陆丽娜	陆鑫达	陈向群	周伯生
周克定	周傲英	孟小峰	岳丽华	范 明
郑国梁	施伯乐	钟玉琢	唐世渭	袁崇义
高传善	梅 宏	程 旭	程时端	谢希仁
裘宗燕	戴 葵			

# 译者序

本书全面、系统地介绍了神经网络系统的相关知识。主要内容包括：神经元模型和网络结构，感知机学习规则，有监督的 Hebb 学习，Widrow - Hoff 学习算法，反向传播神经网络，反向传播算法的变形，联想学习，竞争网络，Grossberg 网络，自适应谐振理论以及 Hopfield 网络等。重点是对这些神经网络的数学分析、训练方法、性能优化，以及神经网络在模式识别、信号处理以及控制系统等工程实践问题中的应用。

本书在叙述神经网络每个特定的主题之前，都会有一些相关材料的介绍和数学基础知识的讨论，使读者能够由浅入深地学习。书中每一章都由“目的”、“理论和实例”、“小结”、“例题”、“结束语”、“参考文献”和“习题”等部分组成。“理论和实例”部分是各章的主体部分，包含神经网络结构和算法的基本思想，以及相应的说明实例。“小结”部分列出每一章的重要公式和概念，为读者归纳出一章的核心内容。每章的“例题”部分占用大量篇幅，给出详细说明关键概念的应用实例。

本书是为高年级本科生或一年级研究生编写的神经网络课程教材，也可供从事相关研究工作的科技人员参考。本书附有使用软件包 MATLAB 的多种“神经网络设计演示”（*Neural Network Design Demonstration*）程序。另外，和本书配套的有教学投影胶片和习题答案，为实施课程的教学提供便利。

参加本书翻译工作的人员有戴葵、宋辉、谭明峰、陆洪毅、沈立、刘芳、王蓉晖、王蕾、赵学秘、李宗伯等同志。本书的翻译，由于时间仓促，难免还存在不少不足之处，恳请广大读者给予批评指正。

译者

2002 年 3 月

## 注意：

- 本书页边标明英文版页码，便于读者参考原书内容。
- 本书索引中标注英文版出处页码，用于查阅术语。

# 前 言

本书介绍了神经网络的基本结构和学习规则，重点是对这些神经网络的数学分析、训练方法和神经网络在模式识别、信号处理以及控制系统等工程实践问题中的应用。

本书尽力用清晰和一致的方式来组织材料，以易于阅读和使用。对每个讨论的主题，使用大量例题来阐明。

由于这是一本关于神经网络设计的书，因此在选择主题时我们依据了两个原则：首先，尽量采用最实用的神经网络结构、学习规则和训练方法；其次，尽量保证该书的完整性，使读者从一章到下一章的学习感觉流畅。为此，在特定主题前，都会有一些相关的介绍性材料和应用数学基础的章节。总之，在我们选择的主题中，某些部分在神经网络实际应用中极其重要，而另一些部分对解释神经网络如何运算十分有用。

书中省去了很多本来可以收入的主题。比如，我们并没有把书写成有关所有已知神经网络结构和学习规则的分类和纲要，而是集中介绍一些基本概念。其次，我们没有讨论神经网络的实现技术，比如说 VLSI 实现、光学器件实现和并行计算机实现等。另外，我们也没有深入阐述神经网络的生物学和心理学基础。上述内容虽然重要，但本书并不包含这些内容，因为我们希望能集中力量把我们认为在神经网络设计中对读者最重要的主题阐述清楚。

本书是为高年级本科生或一年级研究生编写的半学期导论性课程教材（也适于作短期教程、自学或参考用书）。希望读者有一定的线性代数、概率论和微分方程的基础知识。

本书每一章都分为以下各节：目的、理论和实例、小结、例题、结束语、参考文献和习题。理论和实例部分是各章的主体部分，包括基本思想的发展和实例。小结部分列出了一些重要的公式和概念，以利于将本书作为实际工作的参考。每章大约三分之一的篇幅是例题部分，这一部分给出了所有关键概念的详细例题。

后面的图说明了各章之间的相互关系。

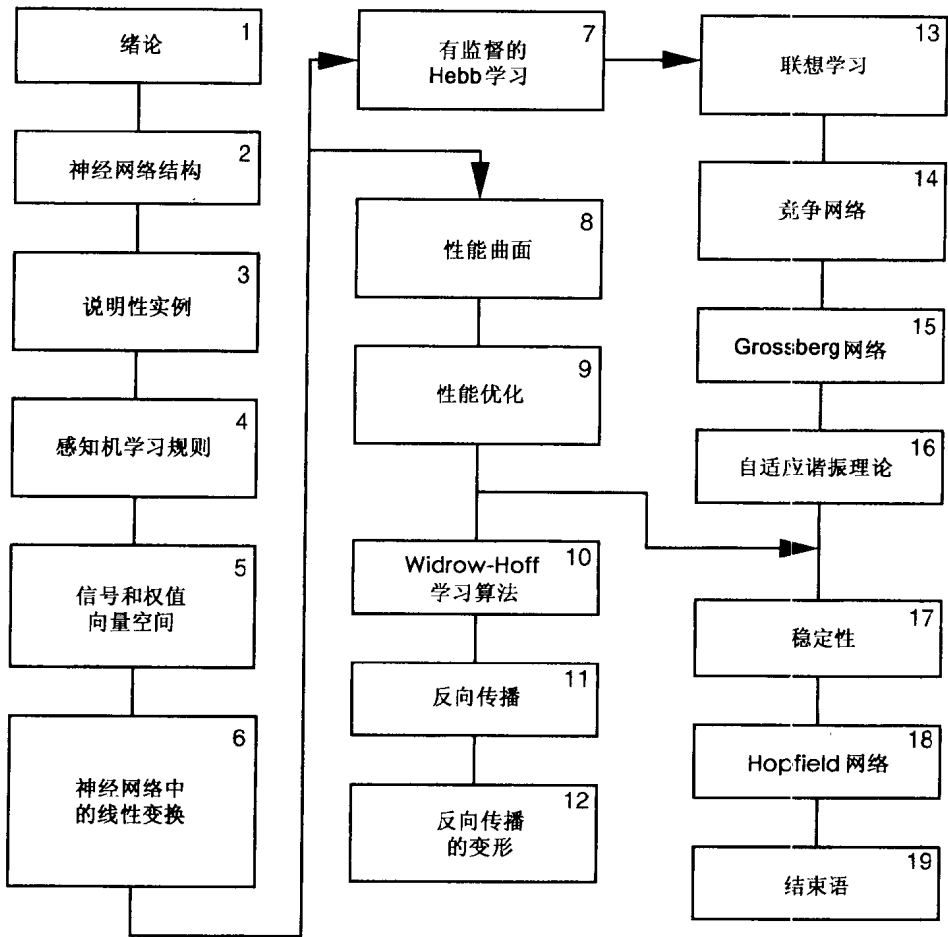
第 1 章到第 6 章覆盖了其余各章所需要的基本概念。第 1 章是绪论部分，简单介绍历史背景和一些基本生物学知识。第 2 章介绍基本的神经网络结构。全书都使用这一章给出的标记方法。第 3 章描述一个简单的模式识别问题，并说明怎样用三种不同类型的神经网络可以求解。这三种网络是本书所述其余神经网络类型的代表。另外，所引入的模式识别问题也为全书提供了一个实验的主线。

本书的许多重点是在使神经网络能完成各种任务的训练方法上。第 4 章介绍学习算法，并给出第一个实用算法（感知机学习规则）。虽然感知机网络存在一些基本的局限性，但它有其自身的历史重要性，并且对于导入某些关键概念也是十分有用的工具，这些概念将要用于其后各章讨论功能更加强大的网络中。

本书的主要目的是解释神经网络的基本工作原理。为此，我们将结合神经网络这一主题给出其他的一些介绍性材料。比如，第 5 和第 6 章要复习线性代数，它是理解神经网络的数学基础的核心。这两章讨论的概念在其余各章被广泛地用到。

第 7 章和第 13 章到第 16 章介绍主要由生物学和心理学的启示所得到的神经网络和学习





规则。它们主要分为两类：相联网络和竞争网络。第 7 章和第 13 章介绍基本概念，第 14 章到第 16 章论述更先进的网络。

第 8 章到第 12 章提出一类叫性能学习 (performance learning) 的学习方法，用它训练网络以优化网络的性能。第 8 章和第 9 章介绍性能学习的基本概念。第 10 章到第 12 章将这些概念用于前馈神经网络中，这将增强网络的能力，但同时也会增加学习的复杂性。

第 17 章和第 18 章讨论递归网络，这些含有反馈连接的网络是一种动态系统。第 17 章研究这些系统的稳定性；第 18 章描述 Hopfield 网络，它是目前最有影响的递归网络之一。

在第 19 章，我们对本书所给出的各种网络进行小结，并讨论它们同本书没有涉及的其他网络之间的关系。同时，我们也要为读者指明进一步研究的一些其他参考资料。如果您想知道“我从此将走向何处？”，请看第 19 章。

## 软件

MATLAB 软件包不是使用本书必需的。上机练习可以用其他任何编程语言实现，而且 *Neural Network Design Demonstration* (神经网络演示) 虽然对读者有帮助，但也是理解本

书涉及材料的关键所在。

然而，我们还是把 MATLAB 软件包作为这本教科书的附件。由于该软件包含矩阵/向量的表示法和图形显示，所以它也为神经网络的实验提供了一种方便的环境。我们以两种不同的方法使用 MATLAB。第一，书中含有大量需要读者用 MATLAB 完成的习题。神经网络许多重要特征只有在解决大规模问题时才能体现出来，这些密集的计算不适于手算。用 MATLAB 能很快实现神经网络算法，并对大量问题方便地进行测试。（若没有 MATLAB，用其他语言也可以完成这些练习。）

使用 MATLAB 的第二种方法是用本书所附磁盘中的 *Neural Network Design Demonstrations*（神经网络设计演示），这些交互式演示阐述了每章的重要概念。使用时，把软件拷到 MATLAB 目录中，在 MATLAB 的提示符下，敲击 nnd 即可激活。所有演示都可以在主菜单下访问。



用左边所示的图标，指明正文中对这些演示的引用。演示需要 MATLAB 4.0 或更高版本，或者 MATLAB 4.0 学生版。另外，一些演示需要用 MathWorks 公司的 *Neural Network Toolbox*（神经网络工具箱）1.0 版本或更高版本。如何使用演示软件请参阅附录 C。

为了帮助使用本书的教师，我们还准备了投影胶片和习题答案。每一章的投影胶片（用微软的 Powerpoint 格式）可以从网址 [www.pws.com/pwsftp.html](http://www.pws.com/pwsftp.html) 获得。也可以获取习题答案。

## 致谢

我们深切地感谢那些阅读了本书全部或部分稿件的人以及对软件的各种不同版本进行测试的人，特别要感谢的是 Canterbury 大学的 John Andreae 教授，AT&T 的 Dan Foresee，Oklahoma 州立大学的 Carl Latino 博士，MCI 的 Jack Hagan，SRI 的 Gerry Andeen 博士，以及 Idaho 大学的 Joan Miller 和 Margie Jenks。我们从 Oklahoma 州立大学的 ECEN 5713 班及 Canterbury 大学的 ENEL 621 班的研究生们那里得到了许多建设性的意见，他们阅读了最早的初稿，并测试了软件，对本书的改进提供了极有帮助的建议。另外，我还要感谢那些虽然没有留下姓名，但对该书提出了许多有益建议的人。

我们竭诚感谢 Peter Gough 博士，感谢他邀请我们加入 Canterbury 大学的电气和电子工程系，感谢 Mike Surety 在计算机方面的帮助，以及全系同仁的大力支持。在 Oklahoma 州立大学的学术假和离开 Idaho 大学的这一年时间里，我们完成了这本书。感谢 TI 公司特别是 Bill Harland 对我们在神经网络研究上的支持。感谢 MathWorks 公司允许我们使用 *Neural Network Toolbox*（神经网络工具箱）中的材料。

我们非常感谢 Joan Pilgram 的鼓励和她商业方面的建议，感谢 Bernice Hewitt 夫人的热情支持。

最后，我们由衷地表达对 PWS 出版公司的同仁们的感激，特别是 Bill Barter，Pam Rockwell，Amy Mayfield，Ken Morton 和 Nathan Wilbur。感谢 Vanessa Pineiro 为本书设计的非常具有艺术品位的封面。

# 目 录

出版者的话	
专家委员会	
译者序	
前言	
第 1 章 绪论	1
1.1 目的	1
1.2 历史	1
1.3 应用	3
1.4 生物学的启示	5
参考文献	6
第 2 章 神经元模型和网络结构	8
2.1 目的	8
2.1 理论和实例	8
2.2.1 符号	8
2.2.2 神经元模型	8
2.2.3 网络结构	13
2.3 小结	17
2.4 例题	20
2.5 结束语	22
习题	22
第 3 章 一个说明性实例	23
3.1 目的	23
3.2 理论和实例	23
3.2.1 问题描述	23
3.2.2 感知机	24
3.2.3 Hamming 网络	28
3.2.4 Hopfield 网络	30
3.3 结束语	32
习题	32
第 4 章 感知机学习规则	34
4.1 目的	34
4.2 理论和实例	34
4.2.1 学习规则	35
4.2.2 感知机的结构	35
4.2.3 感知机学习规则	39
4.2.4 收敛性证明	44
4.3 小结	47
4.4 例题	47
4.5 结束语	56
参考文献	56
习题	57
第 5 章 信号和权值向量空间	60
5.1 目的	60
5.2 理论和实例	60
5.2.1 线性向量空间	60
5.2.2 线性无关	62
5.2.3 生成空间	62
5.2.4 内积	63
5.2.5 范数	63
5.2.6 正交性	64
5.2.7 向量展开式	65
5.3 小结	68
5.4 例题	70
5.5 结束语	76
参考文献	76
习题	76
第 6 章 神经网络中的线性变换	79
6.1 目的	79
6.2 理论和实例	79
6.2.1 线性变换	79
6.2.2 矩阵表示	80
6.2.3 基变换	82
6.2.4 特征值和特征向量	85
6.3 小结	88
6.4 例题	89
6.5 结束语	96
参考文献	96
习题	97
第 7 章 有监督的 Hebb 学习	99

7.1 目的	99	10.2.4 收敛性分析	173
7.2 理论和实例	99	10.2.5 自适应滤波	175
7.2.1 线性联想器	100	10.3 小结	181
7.2.2 Hebb 规则	100	10.4 例题	182
7.2.3 仿逆规则	103	10.5 结束语	193
7.2.4 应用	104	参考文献	193
7.2.5 Hebb 学习的变形	106	习题	194
7.3 小结	106	第 11 章 反向传播	197
7.4 例题	107	11.1 目的	197
7.5 结束语	116	11.2 理论和实例	197
参考文献	116	11.2.1 多层感知机	197
习题	117	11.2.2 反向传播算法	201
第 8 章 性能曲面和最优点	119	11.2.3 例子	205
8.1 目的	119	11.2.4 反向传播	207
8.2 理论和实例	119	11.3 小结	211
8.2.1 泰勒级数	119	11.4 例题	212
8.2.2 方向导数	121	11.5 结束语	221
8.2.3 极小点	122	参考文献	222
8.2.4 优化的必要条件	124	习题	222
8.2.5 二次函数	126	第 12 章 反向传播算法的变形	227
8.3 小结	131	12.1 目的	227
8.4 例题	132	12.2 理论和实例	227
8.5 结束语	141	12.2.1 BP 算法的缺点	227
参考文献	141	12.2.2 BP 算法的启发式改进	232
习题	141	12.2.3 数值优化技术	235
第 9 章 性能优化	143	12.3 小结	244
9.1 目的	143	12.4 例题	246
9.2 理论和实例	143	12.5 结束语	255
9.2.1 最速下降法	143	参考文献	255
9.2.2 牛顿法	148	习题	257
9.2.3 共轭梯度法	152	第 13 章 联想学习	259
9.3 小结	155	13.1 目的	259
9.4 例题	156	13.2 理论和实例	259
9.5 结束语	165	13.2.1 简单联想网络	260
参考文献	166	13.2.2 无监督的 Hebb 规则	261
习题	166	13.2.3 简单的识别网络	264
第 10 章 Widrow - Hoff 学习算法	168	13.2.4 instar 规则	265
10.1 目的	168	13.2.5 简单回忆网络	268
10.2 理论和实例	168	13.2.6 outstar 规则	268
10.2.1 ADALINE 网络	168	13.3 小结	271
10.2.2 均方误差	170	13.4 例题	272
10.2.3 LMS 算法	171	13.5 结束语	279

参考文献	280	习题	375
习题	280	第 17 章 稳定性	378
第 14 章 竞争网络	285	17.1 目的	378
14.1 目的	285	17.2 理论和实例	378
14.2 理论和实例	285	17.2.1 递归网络	378
14.2.1 Hamming 网络	286	17.2.2 稳定性概念	379
14.2.2 竞争层	287	17.2.3 Lyapunov 稳定性定理	381
14.2.3 生物学意义上的竞争层	291	17.2.4 单摆例子	381
14.2.4 自组织特征图	292	17.2.5 LaSalle 不变性定理	385
14.2.5 学习向量量化	295	17.3 小结	390
14.3 小结	299	17.4 例题	391
14.4 例题	301	17.5 结束语	396
14.5 结束语	310	参考文献	396
参考文献	310	习题	397
习题	311	第 18 章 Hopfield 网络	399
第 15 章 Grossberg 网络	315	18.1 目的	399
15.1 目的	315	18.2 理论和实例	399
15.2 理论和实例	315	18.2.1 Hopfield 模型	400
15.2.1 生物学的启发: 视觉	316	18.2.2 Lyapunov 函数	401
15.2.2 基本非线性模型	321	18.2.3 增益效应	406
15.2.3 两层竞争网络	323	18.2.4 Hopfield 网络设计	409
15.2.4 与 Kohonen 规则的关系	331	18.3 小结	413
15.3 小结	332	18.4 例题	415
15.4 例题	335	18.5 结束语	421
15.5 结束语	342	参考文献	422
参考文献	343	习题	423
习题	344	第 19 章 结束语	426
第 16 章 自适应谐振理论	346	19.1 目的	426
16.1 目的	346	19.2 理论和实例	426
16.2 理论和实例	346	19.2.1 前馈和联想网络	426
16.2.1 自适应谐振概述	346	19.2.2 竞争网络	429
16.2.2 第一层	347	19.2.3 动态联想存储器网络	429
16.2.3 第二层	351	19.2.4 神经网络的经典基础	430
16.2.4 调整子系统	354	19.2.5 参考书目和杂志	431
16.2.5 学习规则: L1 - L2	356	19.3 结束语	432
16.2.6 学习规则: L2 - L1	358	参考文献	432
16.2.7 ART1 算法小结	359	附录 A 文献目录	439
16.2.8 其他 ART 体系结构	360	附录 B 符号	447
16.3 小结	361	附录 C 软件	452
16.4 例题	364	索引	456
16.5 结束语	374		
参考文献	375		

# 第1章 绪 论

## 1.1 目的

当你现在看这本书的时候，就正在使用一个复杂的生物神经网络。你有一个约为  $10^{11}$  个神经元的高度互连的集合帮助你完成阅读、呼吸、运动和思考。你的每一个生物神经元都是生物组织和化学物质的有机结合。若不考虑其速度的话，可以说每个神经元都是一个复杂的微处理器。你的某些神经结构是与生俱来的，而其他一些则是在实践中形成的。

科学家们才刚刚开始对生物神经网络工作机理有所认识。一般认为，包括记忆在内的所有生物神经功能，都存储在神经元和及其之间的连接上。学习被看作是在神经元之间建立新的连接或对已有的连接进行修改的过程。这便将引出下面一个问题：既然我们已经对生物神经网络有一个基本的认识，那么能否利用一些简单的人工“神经元”构造一个小系统，然后对其进行训练，从而使它们具有一定有用功能呢？回答是肯定的。本书正是要讨论有关人工神经网络工作机理的一些问题。

我们在这里考虑的神经元不是生物神经元。它们是对生物神经元极其简单的抽象，可以用程序或硅电路实现。虽然由这些神经元组成的网络的能力远远不及人脑的那么强大，但是可对其进行训练，以实现一些有用的功能。本书所要介绍的正是有关于这样的神经元，以及包含这些神经元的网络及其训练方法。

1-1

## 1.2 历史

在人工神经网络的发展历程中，涌现了许多在不同领域中富有创造性的传奇人物，他们艰苦奋斗几十年，提出了许多至今仍然让我们受益的概念。许多作者都记载了这一历史。一本特别有趣的书是由 John Anderson 和 Edward Rosenfeld 撰写的《神经计算：研究的基础》(*Neurocomputing: Foundations of Research*)。在该书中，他们收集并编辑了一组由 43 篇具有特别历史意义的论文，每一篇前面都有一段历史观点的导言。

本书各章开始包括了一些主要神经网络研究人员的历史，所以这里不必赘述。但是，还是有必要简单地回顾一下神经网络的主要发展历史。

对技术进步而言，有两点是必需的：概念与实现。首先，必须有一个思考问题的概念，根据这些概念明确所面临的问题。这就要求概念包含一种简单的思想，或者更具特色，并且引入数学描述。为了解理解这一点，让我们看看心脏的研究历史。在不同时期，心脏被看成灵魂的中心或身体的热源。17 世纪的医生们认识到心脏是一个血泵，于是科学家们开始设计实验，研究泵的行为。这些实验最终开创了循环系统理论。可以说，没有泵的概念，就不会有人们对心脏的深入认识。

概念及其相应的数学描述还不足以使新技术走向成熟，除非能通过某种方式实现这种系统。比如，虽然多年前就从数学上知道根据计算机辅助层析成像(CAT)扫描可以重构图像，但是直到有了高速计算机和有效的算法才使其走向实用，并最终实现了有用的 CAT 系统。

神经网络的发展史同时包含了概念创新和实现开发的进步。但是这些成果的取得并不是一帆风顺的。

神经网络领域研究的背景工作始于 19 世纪末和 20 世纪初。它源于物理学、心理学和神经生理学的跨学科研究，主要代表人物有 Herman Von Helmholtz, Ernst Mach 和 Ivan Pavlov。这些早期研究主要还是着重于有关学习、视觉和条件反射等一般理论，并没有包含有关神经元工作的数学模型。

1-2

现代对神经网络的研究可以追溯到 20 世纪 40 年代 Warren McCulloch 和 Walter Pitts 的工作 [McPi43]。他们从原理上证明了人工神经网络可以计算任何算术和逻辑函数。通常认为他们的工作是神经网络领域研究工作的开始。

在 McCulloch 和 Pitts 之后，Donald Hebb [Hebb49] 指出，经典的条件反射(由 Pavlov 发现)是由单个神经元的性质引起的。他提出了生物神经元的一种学习机制(参见第 7 章)。

人工神经网络第一个实际应用出现在 20 世纪 50 年代后期，Frank Rosenblatt [Rose58] 提出了感知机网络和联想学习规则。Rosenblatt 和他的同事构造了一个感知机网络，并公开演示了它进行模式识别的能力。这次早期的成功引起了许多人对神经网络研究的兴趣。不幸的是，后来研究表明基本的感知机网络只能解决有限的几类问题。(有关 Rosenblatt 和感知机学习规则，请参见第 4 章。)

同时，Bernard Widrow 和 Ted Hoff [WiHo60] 引入了一个新的学习算法用于训练自适应线性神经网络。它在结构和功能上类似于 Rosenblatt 的感知机。Widrow-Hoff 学习规则至今仍然还在使用。(关于 Widrow-Hoff 学习请参见第 10 章。)

但是，Rosenblatt 和 Widrow 的网络都有同样的固有局限性。这些局限性在 Marvin Minsky 和 Seymour Papert 的书 [MiPa69] 中有广泛的论述。Rosenblatt 和 Widrow 也十分清楚这些局限性，并提出了一些新的网络来克服这些局限性。但是他们没能成功找到训练更加复杂网络的学习算法。

许多人受到 Minsky 和 Papert 的影响，相信神经网络的研究已走入了死胡同。同时由于当时没有功能强大的数字计算机来支持各种实验，从而导致许多研究者纷纷离开这一研究领域。神经网络的研究就这样停滞了十多年。

即使如此，在 20 世纪 70 年代，科学家们仍然在该领域开展了许多重要的工作。1972 年 Teuvo Kohonen [Koho72] 和 James Anderson [Ande72] 分别独立提出了能够完成记忆的新型神经网络。(有关 Kohonen 网络更加详细的内容请参见第 13 章和第 14 章。)这一时期，Stephen Grossberg [Gros76] 在自组织网络方面的研究也十分活跃。(参见第 15 章和第 16 章。)

1-3

前面我们说过，在 60 年代，由于缺乏新思想和用于实验的高性能计算机，曾一度动摇了人们对神经网络的研究兴趣。到了 80 年代，随着个人计算机和工作站计算能力的急剧增强和广泛应用，以及不断引入新的概念，克服了摆在神经网络研究面前的障碍，人们对神经网络的研究热情空前高涨。

有两个新概念对神经网络的复兴具有极其重大的意义。其一是：用统计机理解释某些类型的递归网络的操作，这类网络可作为联想存储器。物理学家 John Hopfield 的研究论文 [Hopf82] 论述了这些思想。(第 17 章和第 18 章讨论 Hopfield 网络。)

其二是：在 20 世纪 80 年代，几个不同的研究者分别开发出了用于训练多层感知机的反

传算法。其中最具影响力的反传算法是 David Rumelhart 和 James McClelland [RuMc86]提出的。该算法有力地回答了 60 年代 Minsky 和 Papert 对神经网络的责难。(有关反传算法详细内容请参见第 11 章和第 12 章。)

这些新进展对神经网络研究领域重新注入了活力。在过去的 10 年中,人们发表了成千上万的神经网络研究论文,神经网络也有了应用。许多理论和实践工作蜂拥而至,以致于我们至今还不十分清楚这将会把我们带向何方。

以上简略的历史回顾并没有列出所有对神经网络作出重要贡献的人,但它能使读者知道神经网络是如何发展而来的。读者或许会注意到,这个发展趋势并不总是“缓慢而坚定”的,而是曾经有急剧发展的时期,也有相对停滞的时期。

许多神经网络研究进展都与新概念的提出有关,如革新的神经网络结构和训练规则。同样十分重要的是,高性能计算机的出现使新概念能够得到检验。

好了,对神经网络的历史就说这么多。真正的问题是:“以后的 10 到 20 年会怎样?”神经网络将演变为一个永久的数学/工程工具,还是像许多曾大有希望的技术那样退出历史舞台?目前来看,似乎神经网络不仅有兴旺的时日,而且能取得一个永久的地位,即使它不能解决所有问题,但在某些适当的场合还是非常有用的工具。另外,要记住我们现在对人脑的认识仍很肤浅,相信将来某一天神经网络将会取得最重要的进展。

尽管很难预料神经网络今后能否成功,但这种新技术的大量而广泛应用还是令人鼓舞的。下面一节将介绍一些神经网络应用。

1-4

### 1.3 应用

最近报纸报道 Aston 大学用神经网络来进行文献研究。这篇报道说“神经网络可以用来识别个人的写作风格,研究人员用它比较了莎士比亚和他同时代人的著作”。一个大众科学电视节目最近报道了某意大利的研究结构用神经网络测试橄榄油的纯度。这些例子从一个侧面说明神经网络有极其广泛的应用领域。正是因为它适合于解决实际问题,所以其应用领域在不断扩大,它不仅可以广泛应用于工程、科学和数学领域,也可广泛应用于医学、商业、金融和文学等领域。神经网络在许多领域的广泛应用,使其极具吸引力。同时,基于高速计算机和快速算法,也可以用神经网络解决过去许多计算量很大的复杂工业问题。

以下神经网络的应用说明来源于 MATLAB 用到的 *Neural Network Toolbox* (神经网络工具箱),已经得到了 MathWorks 公司的允许。

1988 年,在 DARPA 的“神经网络研究报告”(Neural Network Study)[DARP88]中列举了各种神经网络的应用。其中第一个应用就是大约在 1984 年的自适应频道均衡器。这个设备在商业上取得了极大的成功。它用一个单神经元网络来稳定电话系统中长距离传输的声音信号。DARPA 报告还列举了其他一些神经网络在商业领域中的应用,包括一个小规模的单词识别器、过程监测器、声纳分类器和一个风险分析系统。

自 DARPA 报告问世以来,神经网络已被用于许多领域。在文献中所列举的一些应用如下:

#### 航空

高性能飞行器自动驾驶仪,飞行路径模拟,飞机控制系统,自动驾驶优化器,飞行部件模拟,飞行器部件故障检测器



## 汽车

汽车自动导航系统, 担保行为分析器

## 银行

1-5

支票和其他公文阅读器, 信贷申请的评估器

## 国防

武器操纵, 目标跟踪, 目标辨识, 面部识别、新型的传感器, 声纳、雷达和图像信号处理(包括数据压缩、特征提取、噪声抑制、信号/图像的识别)

## 电子

代码序列预测, 集成电路芯片布局, 过程控制, 芯片故障分析, 机器视觉, 语音综合, 非线性建模

## 娱乐

动画, 特技, 市场预测

## 金融

不动产评估, 借贷咨询, 抵押审查, 公司证券分级, 投资交易程序, 公司财务分析, 通货价格预测

## 保险

政策应用评估, 产品优化

## 制造

生产流程控制, 产品设计和分析, 过程和机器诊断, 实时微粒识别, 可视质量监督系统, 啤酒检测, 焊接质量分析, 纸张质量预测, 计算机芯片质量分析, 磨床运转分析, 化工产品设计分析, 机器性能分析, 项目投标, 计划和管理, 化工流程系统动态建模

## 医疗

乳房癌细胞分析, EEG 和 ECG 分析, 修复设计, 移植次数优化, 医院费用节流, 医院质量改进, 急诊室检查建议

## 石油和天然气

1-6

探查

## 机器人

轨道控制, 铲车机器人, 操作手控制器, 视觉系统

## 语音

语音识别, 语音压缩, 元音识别, 文本到语音的综合

## 有价证券

市场分析, 自动证券分级, 股票交易咨询系统

## 电信

图像和数据压缩, 自动信息服务, 实时语言翻译, 客户支付处理系统

## 交通

卡车制动器诊断系统, 车辆调度, 运送系统

## 结论

神经网络应用的数量、投入到神经网络软硬件上的资金和公众对这些设计的兴趣都在快