



普通高等教育“十一五”国家级规划教材

· 计算机新技术应用系列丛书 ·

人工神经网络与 微粒群优化

刘希玉 刘 弘 编著



北京邮电大学出版社
www.buptpress.com



普通高等教育“十一五”国家级规划教材

· 计算机新技术应用系列丛书 ·

人工神经网络与微粒群优化

刘希玉 刘 弘 编著

本书主要介绍人工神经网络和微粒群优化的基本原理、算法及应用。全书共分8章。第1章介绍神经网络的基本概念、分类、特点及应用；第2章介绍神经网络的数学模型、学习算法及应用；第3章介绍神经网络的优化算法及应用；第4章介绍微粒群优化的基本原理、算法及应用；第5章介绍微粒群优化的改进算法及应用；第6章介绍微粒群优化的应用；第7章介绍微粒群优化的应用；第8章介绍微粒群优化的应用。

本书可作为高等院校计算机专业及相关专业的教材，也可供从事计算机工作的工程技术人员参考。

北京邮电大学出版社

定价：38.00元

· 北京 ·

ISBN 978-7-5632-1388-0

内 容 简 介

本书介绍了几类常用的和基本的人工神经网络模型,即感知机、前馈型、反馈型、随机神经网络,还介绍了比较新并有发展前途的支持向量机、非线性泛函网络、人工免疫系统,独立成章地阐述了微粒群算法的基本原理及人工神经网络的微粒群优化技术,每种模型都围绕结构、基本原理、学习算法的次序来阐述,同时介绍了在数据挖掘、创新概念设计中的应用。书中部分内容是几年来作者及其所指导的研究生们研究成果的总结。

本书选材精、内容新、阐述系统,力争深入浅出和突出应用,可作为相应学科的研究生和高年级本科生的课程教材,也可作为从事该领域研究的科学技术人员的参考书。

图书在版编目(CIP)数据

人工神经网络与微粒群优化/刘希玉,刘弘编著. —北京:北京邮电大学出版社,2008.3

ISBN 978-7-5635-1368-0

I. 人… II. ①刘…②刘… III. ①人工神经网络—研究②智能控制—算法—研究 IV. TP183 TP273

中国版本图书馆 CIP 数据核字(2008)第 022430 号

书 名: 人工神经网络与微粒群优化

编 著: 刘希玉 刘 弘

责任编辑: 李欣一

出版发行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路 10 号(邮编:100876)

发 行 部: 电话:010-62282185 传真:010-62283578

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 北京源海印刷有限责任公司

开 本: 787 mm×1 092 mm 1/16

印 张: 19

字 数: 473 千字

印 数: 1—3 000 册

版 次: 2008 年 3 月第 1 版 2008 年 3 月第 1 次印刷

ISBN 978-7-5635-1368-0

定价: 38.00 元

• 如有印装质量问题,请与北京邮电大学出版社发行部联系 •

前言

浩瀚宇宙,奇妙无穷,星移斗转,万物丛生。在这个蔚蓝色的地球上,人为万物之灵。人类用智慧认识客观,改造世界,用机械减轻体力劳动,用电气改善生活环境,科学技术正在日新月异地向发展着。然而,正像“不识庐山真面目,只缘身在此山中”所描述的那样,人类对周围客观世界认识越来越清楚的同时,对自己天天都使用的大脑产生智慧的机理,却仍然一知半解。我们希望智能起源问题(科学技术中三大起源难题——宇宙起源、生命起源和智能起源——之一)能在 21 世纪解决。

在世纪之交的时刻,对以仿效大脑神经系统为目的的人工神经网络的研究历程进行一些分析和反思,探讨下一步可能采取的方法步骤,对今后智能信息科学的进一步发展将是有益的。

如果把 1985 年霍普菲尔德(J. J. Hopfield)成功地用全互连式人工神经网络求解推销员问题(Traveling Salesman Problem, TSP)作为再次掀起人工神经网络(Artificial Neural Networks, ANN)研究高潮起点的话,那么世界范围内的 ANN 热潮已经历了 15 年。对国内来说,自从 1990 年在北京召开首届中国神经网络学术大会(C2N2-90)以后,全国 13 个学会按照“携手探智能,联盟攻大关”的共同目标,又举行了八次学术年会(其中四次年会与国际神经网络会议联合召开)。九次年会论文集共收入论文两千多篇,内容涉及 ANN 的模型原理、学习算法、实现方法和专业应用等方面。作为一种新型的信息处理手段,ANN 已在众多的专业领域,如模式识别、自动控制和组合优化等方面得到越来越广泛的普及应用。“神经网络”已成为在近期科技刊物上出现频率甚高的关键词。

无论是从国内神经网络盛会的议题,还是从相关国际学术刊物论文的内容分析,都显示出如下特点:在各专业领域的应用多,用硬件方式实现的成果少;将 ANN 当做一般的并行分布处理手段的报道多,在显现智能方面的突破进展少。这就是说,迄今为止,按生物神经网络(Biological Neural Networks, BNN)巨量并行分布方式构造的各种人工神经网络,虽然已经在信息处理中扮演着越来越重要的角色,但是并没有显示出人们所期望的聪明和智慧来。

本书主要介绍了几种常用的、基本的、有发展前途的神经网络的设计和应用,并介绍了用于神经网络优化的微粒群技术,形象生动,图文并茂,深入浅出,脉络清晰,可读性强。前两章介绍了人工神经网络的基本知识;第 3~9 章介绍了感知机、前馈型、反馈型、随机神经网络、支持向量机、非线性泛函网络及其应用;第 10、11 章分别介绍了人工免疫系统和微粒群算法及其应用。相信广大读者通过认真学习本书,可以深入了解各种神经网络,建立牢固

的知识基础,真正做到事半功倍。

本书由刘希玉、刘弘编写,王慧、牟荣参与编写第1章,李菲菲参与编写第2章,张宏参与编写第3章,周国亮参与编写第4章,许文杰参与编写第5章,王慧、唐培霞参与编写第6章和第11章,李玲玲参与编写第7章,何旭堂、王少飞参与编写第8章,刘栋参与编写第9章,许珂参与编写第10章,李丽丽、王宗利、牛雪丽、林令娟、崔海青、吴晓燕、张丽、魏欣、戴芬、岳凤、王晓敏、王敏参与了部分编写工作,刘希玉、刘弘、王慧负责全书的统稿与审校工作。另外,还有许多同志在本书的编校过程中付出了大量的劳动,在此一并表示衷心的感谢。

由于时间仓促加之作者本身水平有限,书中错误之处在所难免。在此,敬请各领域专家和广大读者批评指正。

编著者

目 录

第 1 章 概论

1.1 人工神经网络的基本概念	1
1.1.1 人工神经网络的基本概念和特征	1
1.1.2 神经元数学模型	4
1.2 人工神经网络研究的发展简史	5
1.2.1 人工神经网络研究的发展	5
1.2.2 我国人工神经元网络研究的情况	10
1.2.3 存在的问题	11
1.2.4 几种著名的神经网络	11
1.3 并行及分布处理理论	14
1.3.1 并行分布处理理论	14
1.3.2 PDP 的基本概念	14
1.3.3 联结性和并行分布性	15
1.3.4 非符号性和连续性	15
1.4 研究人工神经网络的意义	15
1.4.1 神经网络与传统计算的比较	15
1.4.2 人工神经元网络与人工智能系统的比较	16
1.4.3 研究人工神经网络的意义	17
1.4.4 人工神经元网络在控制领域中的应用	17

第 2 章 人工神经网络的基本原理

2.1 神经元	20
2.1.1 神经元的结构	20
2.1.2 大脑的信息处理原理	22
2.2 人工神经网络模型	23
2.2.1 形式神经元模型	24
2.2.2 神经网络的基本结构	26
2.2.3 神经元的学习法则	27
2.2.4 神经元的计算原理	30

2.3 十种常用的学习规则	31
2.3.1 Hebb 学习规则	32
2.3.2 感知器的学习规则	35
2.3.3 δ 学习规则	38
2.3.4 Widrow-Hoff 学习规则	40
2.3.5 相关学习规则	40
2.3.6 “胜者为王”学习规则	41
2.3.7 内星和外星学习规则	41
2.3.8 梯度下降算法	42
2.3.9 随机训练算法	42
2.3.10 模拟退火算法	43
2.3.11 学习规则的一览表	44
2.4 学习任务	45
2.4.1 模式联想	45
2.4.2 模式识别	45
2.4.3 函数逼近	46
2.4.4 控制	48
2.4.5 滤波	49
2.4.6 波束形成	50
2.5 记忆	51
2.5.1 相关矩阵记忆	54
2.5.2 回忆	55
2.6 自适应	56
2.7 学习过程的统计性质	57
2.8 统计学习理论	60
2.8.1 一些基本定义	61
2.8.2 经验风险最小化原则	62
2.9 VC 维	63
2.9.1 VC 维的估计	65
2.9.2 可能近似正确的学习模型	66
2.9.3 样本复杂性	67
2.9.4 计算复杂性	67

第3章 感知机

3.1 最简单的感知机	69
3.2 感知机的学习算法	70
3.3 收敛定理	74
3.4 简单感知机的局限性	76
3.5 凸集分离定理	79

3.5.1	凸集与凸性	79
3.5.2	凸集分离定理	80
第4章 前馈型神经网络		
4.1	多层感知器	82
4.2	多层感知器的学习算法	84
4.3	激活函数	88
4.3.1	S型函数	88
4.3.2	其他激活函数	9
4.3.3	非线性误差函数	92
4.4	BP算法的改进	92
4.5	网络设计及数据预处理	96
4.5.1	输入层和输出层的设计	96
4.5.2	隐含层数和层内节点数的选择	98
4.6	网络映射及容量分析	107
4.6.1	映射作用	107
4.6.2	前向网络的容量分析	111
第5章 BP网络的应用		
5.1	XOR问题	115
5.2	天气预报问题	120
5.3	地下水位的预测	123
5.4	室内温度的预测	126
5.5	产品成本的预测	127
第6章 反馈型神经网络		
6.1	Hopfield网络	130
6.2	联想记忆与神经计算	138
6.2.1	联想记忆	138
6.2.2	神经计算	139
6.3	连续时间Hopfield神经网络模型	140
6.3.1	非线性连续时间Hopfield神经网络	140
6.3.2	Hopfield型神经网络的设计与稳定性	144
6.4	Hopfield神经网络论文分析	148
6.5	Hopfield网络在动态联盟伙伴中的应用	153
6.5.1	模型的建立	154
6.5.2	模型的简化	154
6.5.3	对多目标的动态联盟问题求解	155
6.5.4	基于Hopfield网络的多目标动态联盟求解	156

第7章 随机神经网络

7.1 模拟退火算法	160
7.1.1 模拟退火算法	160
7.1.2 改进的模拟退火算法	163
7.1.3 SA 算法的收敛法	164
7.2 玻耳兹曼机	165
7.2.1 玻耳兹曼机模型	166
7.2.2 能量函数	166
7.2.3 学习算法	168
7.2.4 玻耳兹曼机学习算法推导	169
7.3 NN 的概率统计法	171
7.4 并行分布柯西机	174
7.5 神经网络的熵理论	176
7.5.1 NN 计算能量与熵	177
7.5.2 同步并行计算	178
7.5.3 异步串行计算	178
7.6 动力系统的分维学	181
7.6.1 Hausdorff 维数	181
7.6.2 分维的量度	182
7.7 分维神经网络	184
7.7.1 分维 NN 结构	184
7.7.2 信息的存储	185

第8章 非线性泛函网络

8.1 非线性可分性	188
8.1.1 ϕ 可分性	188
8.1.2 RBF 网络的结构及工作原理	189
8.1.3 函数逼近与内插	190
8.2 Cover 定理	191
8.3 正规化理论	195
8.4 RBF 网络的学习算法	200
8.5 泛函连接网络	203
8.6 小波网络	204
8.6.1 小波理论	204
8.6.2 小波网络	209
8.6.3 小波网络的性能分析	211
8.6.4 小波网络在股市预测中的应用	212
8.6.5 小波理论在创新概念设计评价中的应用	215

第9章 支持向量机

9.1 简介	218
9.2 线性可分模式的最优超平面	219
9.2.1 用于寻找最优超平面的二次最优化	220
9.2.2 最优超平面的统计特性	222
9.3 不可分模式的最优超平面	223
9.4 怎样建立用于模式识别的支持向量机	226
9.4.1 内积核	226
9.4.2 Mercer 定理	227
9.4.3 支持向量机的最优设计	228
9.4.4 支持向量机的例子	229
9.5 例子: XOR 问题	230
9.6 ϵ -不敏感损失函数	232
9.7 用于非线性回归的支持向量机	233
9.8 小结和讨论	235

第10章 人工免疫系统及克隆选择算法

10.1 概述	238
10.2 生物免疫系统的基本原理	239
10.2.1 免疫系统的功能	239
10.2.2 固有性免疫响应和适应性免疫响应	240
10.2.3 免疫系统的结构	240
10.3 人工免疫系统	242
10.3.1 人工免疫系统进展	242
10.3.2 人工免疫系统的研究领域	244
10.3.3 人工免疫系统与其他方法的比较	250
10.4 克隆选择学说与克隆选择算子	251
10.4.1 克隆选择	251
10.4.2 克隆算子	253
10.5 简单克隆选择算法及其性能分析	255
10.5.1 简单克隆选择算法	255
10.5.2 简单克隆选择算法的收敛性	256
10.5.3 多克隆算子与单克隆算子的比较	257
10.5.4 克隆选择算法与进化算法	259
10.5.5 克隆选择算法的优缺点	261
10.6 小结	261
10.6.1 人工免疫系统存在的问题	261
10.6.2 基于人工免疫系统的综合集成	261

10.6.3 人工免疫系统应用研究	262
10.6.4 人工免疫系统进一步研究的方向	263

第 11 章 微粒群算法

11.1 基本的微粒群算法	264
11.1.1 引言	264
11.1.2 基本微粒群算法	265
11.1.3 基本微粒群算法的社会行为分析	266
11.2 改进的微粒群算法	268
11.2.1 对基本微粒群算法进化方程的改进	268
11.2.2 利用小生境思想所做的改进	272
11.2.3 离散变量的微粒群算法	277
11.3 微粒群算法的应用	283
11.3.1 进化计算用于神经网络的优化	283
11.3.2 用 PSO 算法优化神经网络	284
11.3.3 协同 PSO 算法优化神经网络	286
11.4 微粒群算法在建筑设计上的应用探索	291
11.4.1 群体智能算法总体模式	291
11.4.2 群体智能算法之 PSO 算法及其在建筑设计上的应用探索	292

第1章 概 论

【内容提要】

人工神经网络的基本概念;人工神经网络研究的发展简史;并行及处理理论;研究人工神经网络的意义;人工神经网络的研究概况

以冯·诺依曼型计算机为中心的信息处理技术的高速发展,使得计算机在当今的信息化社会中起着十分重要的作用,但是在用它来解决某些人工智能问题时却遇到了很大的困难。

例如,一个人可以很容易地识别他人的脸孔,但计算机则很难做到这一点。这是因为脸孔的识别不能用一个精确的数学模型加以描述,而计算机工作则必须有对模型进行各种运算的指令,如果得不到模型,程序也就无法编制。而大脑是由生物神经元构成的巨型网络,它在本质上不同于计算机,是一种大规模的并行处理系统,它具有学习、联想记忆、综合等能力,并有巧妙的信息处理方法。人工神经网络(简称神经网络)也是由大量的、功能比较简单的形式神经元互相连接而构成的复杂网络系统,用它可以模拟大脑的许多基本功能和简单的思维方式。尽管它还不是大脑的完美无缺的模型,但它可以通过学习来获取外部的知识并存储在网络内,可以解决计算机不易处理的难题,特别是语音和图像的识别、理解,知识的处理,组合优化计算和智能控制等一系列本质上非计算的问题。因此,神经网络技术已成为当前人工智能领域中最令人感兴趣和最富有魅力的研究课题。

1.1 人工神经网络的基本概念

1.1.1 人工神经网络的基本概念和特征

人的大脑是自然界所造就的最高级产物。人的思维是由人脑来完成的,思维是人类智能的集中体现。人的思维主要可概括为逻辑思维(包括联想)和形象思维两种。过去以规则为基础的知识系统可被认为是致力于模拟人的逻辑思维,而人工神经网络则可被认为是探索人的形象思维,前者由左脑主管,后者则是由右脑主管。人的智能是多种多样的和分层的。人的行为有时是单种智能,有时是多种智能综合应用的结果。根据19世纪末创建的神

元学说,人们认识到大脑是由大约 10^{11} 数量级的神经元和 $10^{14} \sim 10^{15}$ 个突触组成的巨型系统。

人工神经网络是由生理学上的真实大脑神经网络的结构和功能,以及若干基本特性的某种理论抽象、简化和模拟而构成的一种信息处理系统。从系统观点看,人工神经网络是由大量神经元通过极其丰富和完善的连接而构成的自适应非线性动态系统。由于神经元之间有着不同的连接方式,所以组成不同结构形态的神经网络系统是可能的。

许多科学家长期以来从事认识人类自身特别是大脑的工作,有了不小的进展,但还远远没有认识清楚,尽管如此,人们还是借鉴了对大脑的结构和功能的一些了解,去设计和建造具有一定智能的机器,使之具有像人那样的信息处理能力,尽可能多地去做人所能做的各种事情。随着大脑工作机制奥秘的不断揭开,将会帮助人们尽早制造出具有更高智能的机器。

现在已从符号和连接两大方面来研究人类的智能,正像明斯基在 1990 年的文章中指出的“这两个方面都是同一个事业的一部分”。尽管 1990 年又出现了从行为方面来研究人类的智能,但我们认为毫不例外地也是同一个事业的一部分,因为人工智能既应该包括基于思维的研究内容,也应该包括基于行为的研究内容,此外,还应该研究思维与行为的交互关系。目前,对人类智能的研究,需要多层次、多途径而且应该将现有方法结合在一起,取长补短,这样才能更好和更高程度地模拟人类的智能。正像明斯基 1990 年文章的编者按中所说的:“明斯基不打算去寻找一条笔直的路,他相信用多种多样的部件来建造系统的时候已经来到,这些部件中有的是联接主义的,有的是符号的,每个部件都有它存在的不同理由。”

人的智能来自于大脑,大脑是由大量的神经细胞或神经元组成的。每个神经元可以看作一个小的处理单元,这些神经元按照某种方式互相连接起来,构成了大脑内部的生理神经网络,它们中各神经元之间连接的强弱,按照外部的激励信号做自适应变化,而每个神经元又随着接收到的多个激励信号的综合大小呈现兴奋或抑制状态。据现在的了解,大脑的学习过程就是神经元之间连接强度随外部激励信息做自适应变化的过程,而大脑处理信息的结果则由神经元的状态表现出来。显然,神经元是信息处理系统的最小单元。虽然神经元的类型有很多种,但其基本结构相似,如图 1-1 所示。

神经元(即神经细胞)由细胞体、树突、轴突和突触 4 部分组成。每个细胞体都有一个细胞核在进行着呼吸和新陈代谢等许多生化过程。整个细胞的外部叫做细胞膜。从细胞体向外伸出许多树突和一条长的轴突,树突和轴突分别负责传入和传出兴奋或抑制信息到细胞体。神经元的树突较短,分支很多,是信息的输入端。轴突较长,是信息的输出端。突触是一个神经元与另一个神经元相联系的特殊结构部位,突触包括突触前(成分)、突触间隙和突触后(成分)3 个部分。突触前(成分)是第一个神经元的轴突末梢部分,突触后(成分)是指第二个神经元的受体表面。突触前(成分)通过化学接触或者电接触,将信息传往突触后受体表面,实现神经元的信息传输。树突和轴突一一对接,从而靠突触把众多的神经元连成一个神经网络。神经元群或者神经网络系统对外界有兴奋和抑制两种反应,兴奋指的是相对静止变为相对活动,抑制是指由相对活动变为相对静止。神经元之间信息的传递形式有正、负两种连接。正连接呈相互激发;负连接呈相互抑制。各神经元之间的连接强度和极性

可以有所不同,并且都可进行调整,因此大脑才可以有存储信息的功能。

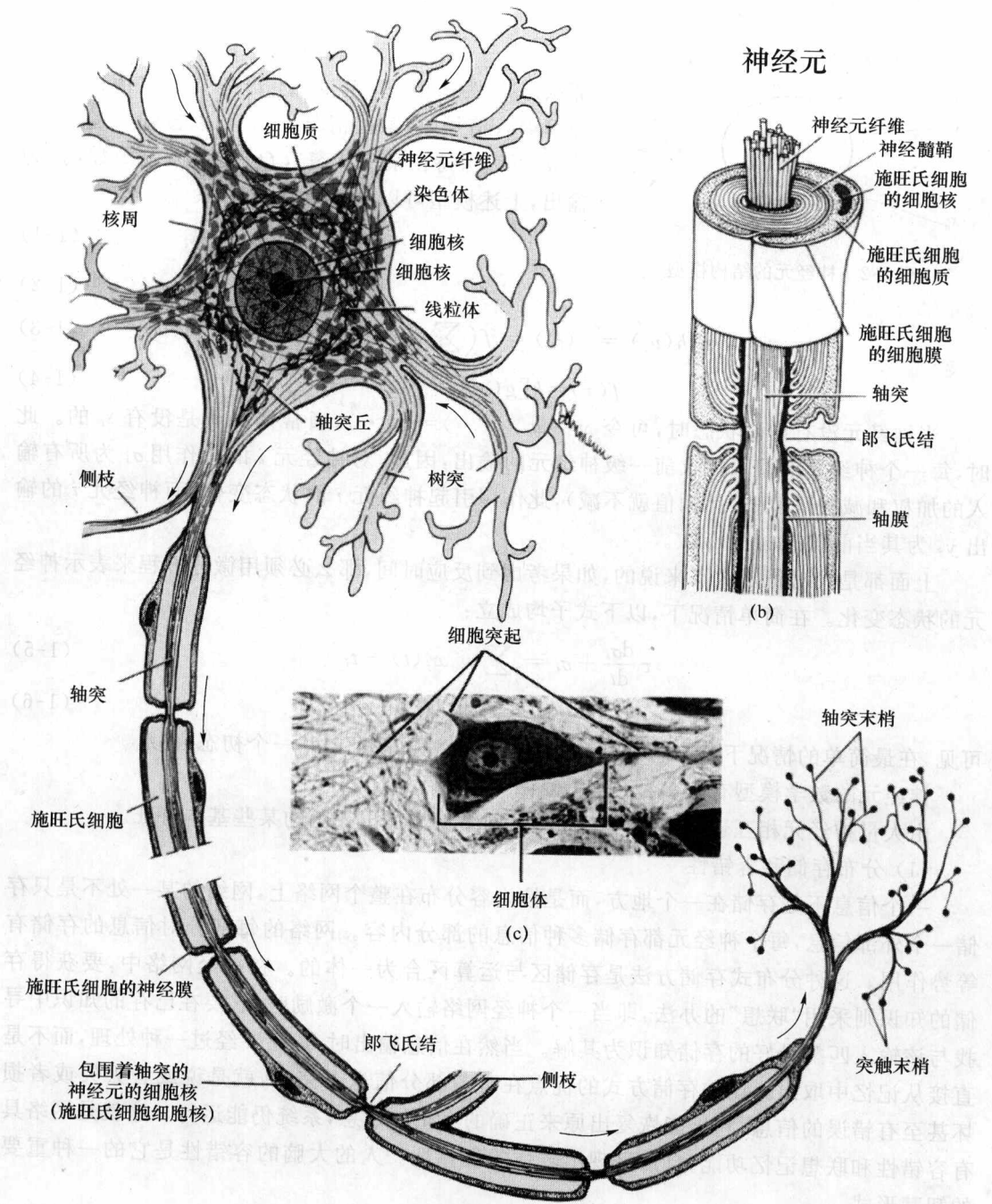


图 1-1 神经元结构示意图

1.1.2 神经元数学模型

简化的神经元数学模型如图 1-2 所示。其中, $x_1, x_2, \dots, x_n$ 为输入信号, u_i 为神经元内部状态, t_i 为阈值, w_{ij} 为 u_i 到 u_j 连接的权值, s_i 表示外部输入信号(在某些情况下,它可以控制神经元 u_i , 使 u_i 可以保持在某一状态), $f(\cdot)$ 为激发函数, y_i 为输出, 上述模型可以描述为

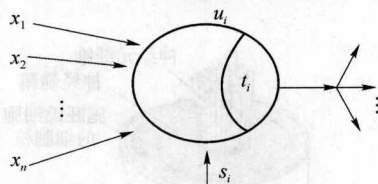


图 1-2 神经元的结构模型

$$\sigma_i = \sum_j w_{ij} x_j + s_i - t_i \quad (1-1)$$

$$u_i = g(\sigma_i) \quad (1-2)$$

$$y_i = h(u_i) = f(\sigma_i) = f\left(\sum_j w_{ij} x_j + s_i - t_i\right) \quad (1-3)$$

$$f(x) = h[g(x)] = f \circ g \quad (1-4)$$

当神经元没有内部状态时,可令 $y_i = u_i$, $f(\cdot) = g(\cdot)$, 通常情况下是没有 s_i 的。此时,每一个神经元的输入接收前一级神经元的输出,因此,对神经元 i 的总作用 σ_i 为所有输入的加权和减去阈值(若无阈值就不减),此作用引起神经元 i 的状态变化,而神经元 i 的输出 y_i 为其当前状态 σ_i 的函数。

上面都是针对稳定状态来说的,如果考虑到反应时间,那么必须用微分方程来表示神经元的状态变化。在简单情况下,以下式子均成立:

$$\tau \frac{d\sigma_i}{dt} + \sigma_i = \sum_j w_{ij} x_j(t) - t_i \quad (1-5)$$

$$y_i(t) = f[\sigma_i(t)] \quad (1-6)$$

可见,在最简单的情况下,神经元的状态与输入成比例,而且向某一个初态衰减。

神经元的数学模型有很多种,这里就不一一介绍了。

由大量神经元相互连接组成人工神经网络将显示出人脑的某些基本特征。

(1) 分布存储和容错性

一个信息不是存储在一个地方,而是按内容分布在整个网络上,网络的某一处不是只存储一个外部信息,每个神经元都存储多种信息的部分内容。网络的每部分对信息的存储有等势作用。这种分布式存储方法是存储区与运算区合为一体的。在神经网络中,要获得存储的知识则采用“联想”的办法,即当一个神经网络输入一个激励时,它要在已存的知识中寻找与该输入匹配最好的存储知识为其解。当然在信息输出时,也还要经过一种处理,而不是直接从记忆中取出。这种存储方式的优点在于若部分信息不完全,就是说或者丢失或者损坏甚至有错误的信息,它仍能恢复出原来正确的完整的信息,系统仍能运行。这就是网络具有容错性和联想记忆功能,自然呈现出较强的鲁棒性。人的大脑的容错性是它的一种重要的智慧形式。

(2) 大规模并行处理

人工神经网络在结构上是并行的,而且网络的各个单元可以同时进行类似的处理过程,因此,网络中的信息处理是在大量单元中平行而又有层次地进行的,其运算速度很高,大大超过传统的序列式运算的数字机。虽然每个神经元的信息传递(神经脉冲)速度是以毫秒计算的,比普通序列式计算机要慢很多,但是人通常能在 1 秒内即可作出对外界事物的判断

和决策,这就是能神奇地完成所谓“百步”决策。目前,依靠现有的计算机技术及人工智能技术是做不到这一点的。

(3) 自学习、自组织和自适应性

学习和适应要求在实践中系统内部结构和联系方式有改变,神经网络是一种变结构系统,恰好能完成对环境的适应和对外界事物的学习能力。神经元之间的连接多种多样,各神经元之间连接强度具有一定的可塑性,相当于突触传递信息能力的变化,这样,网络可以通过学习和训练进行自组织以适应不同信息处理的要求。

(4) 非线性动态系统

神经网络是大量神经元的集体行为,并不是各神经元行为的简单的相加,从而表现出一般复杂非线性动态系统的特性,如不可预测性、不可逆性、有各种类型的吸引子(信息正是“存储”在定点吸引子上)和出现混沌现象等。

(5) 处理复杂系统

神经元可以处理一些环境信息十分复杂、知识背景不清楚和推理规则不明确的问题,例如语音识别和手写体识别、医学诊断以及市场估计等,都是具有复杂非线性和不确定性对象的控制。在处理复杂系统时,信源提供的模式丰富多彩,有的相互间存在矛盾,而判定决策原则又无条理可循,通过神经网络学习(按照学习法则)能从典型事例中学会处理具体事例,给出比较满意的解答。

1.2 人工神经网络研究的发展简史

人工神经网络的研究已有近半个世纪的历史,但它的发展并不是一帆风顺的,而是经过两起一落中间呈现马鞍形的过程。下面以时间的顺序和著名人物或者某方面的突出研究成果为线索,简要地介绍其发展史。

1.2.1 人工神经网络研究的发展

现代的神经网络开始于 McCulloch 和 Pitts 的开拓性工作。McCulloch 是一位精神病学家和解剖学家。他用 20 年的时间考虑神经系统中关于事件的表示问题。Pitts 是数学天才,于 1942 年加入 McCulloch 的工作。

在他们的经典论文里,McCulloch 和 Pitts 结合了神经生理学和数理逻辑的研究,描述了一个神经网络的逻辑演算。他们的神经元模型假定遵循一种所谓“有或无”(all-or-none)规则。如果简单的神经元数目足够多和适当设置突触连接并且同步操作,McCulloch 和 Pitts 证明这样构成的网络原则上可以计算任何可计算函数。这是一个有重大意义的结果,它就标志着神经网络和人工智能学科的诞生。

McCulloch 和 Pitts 于 1943 年发表的论文从那时直到现在一直被广泛阅读。它影响了冯·诺依曼,使得他在 EDVAC(Electronic Discrete Variable Automatic Computer,电子离散变元自动计算机)中使用从 McCulloch 和 Pitts 的神经元导出的理想化开关延迟元件,这台机器是从 ENIAC(Electronic Numerical Integrator and Computer,电子数值积分器和计算机)发展而来的。ENIAC 是第一台通用电子计算机,1943—1946 年在宾夕法尼亚大学摩

尔电子工程学院建成。1949年,冯·诺依曼在 Illinois 大学中所做的第二个报告的主要内容即为 McCulloch-Pills 的形式化神经网络理论。

1948年,Wiener 的名著《控制论》(*Cybernetics*)出版,为控制、通信和统计信号处理描述了某些重要概念。1961年该书第2版出版发行,添加了关于学习和自组织的新材料。在第1版的第2章中,Wiener 在主题方面抓住了统计力学的物理意义,但是却把统计力学和学习系统联系起来的丰硕成果留给了 Hopfield(在30多年以后)。

神经网络第二个重要发展是在1949年 Hebb 的书《行为组织学》(*The Organization of Behavior*)出版后,他在书中第一次清楚地说明了突触修正的生理学学习规则。特别是,Hebb 提出人脑的连接方式在机体学习不同功能任务时是连续变化的,神经组织就是通过这种变化创建起来的。Hebb 继承了 Ramón Cajál 早期的假设并引入自己现在著名的学习假说,即两个神经元之间的可变突触的作用被突触两端神经元中一个对另一个的重复的激活而加强了。Hebb 的书在心理学家中有巨大的影响,但遗憾的是对工程界影响很少或没有影响。

Hebb 的书是学习系统和自适应系统的计算模型发展的灵感源泉。Rochester、Holland、Haibt 和 Duda 于1956年发表的论文可以说是用计算机模拟测试以 Hebb 学习假说为基础的严格公式化的神经理论的第一次尝试;论文报告的模拟结果表明必须加上抑制理论才能实际工作。同年,Uttley 演示了带有可修改的突触的神经网络,可以学习分类简单的二值模式集。Uttley 引入了所谓泄漏集成和点火神经元(Leaky Integrate and Fire Neuron),后来 Caianiello 对它进行了形式化分析。在较晚的工作中,Uttley 假设了神经系统可变突触的作用依赖于突触两端波动状态的统计关系,因此和香农的信息论联系起来。

1952年,Ashby 的书《脑的设计:自适应行为的起源》(*Design for a Brain: The Origin of Adaptive Behavior*)出版。这本书关注的是基本概念,即自适应行为不是与生俱来而是后天学习的,通过学习动物(系统)的行为变得更好。这本书强调活的机体如同机器的动态方面和有关稳定性的概念。

1954年,Minsky 在普林斯顿大学写了“神经网络”的博士论文,题目是“神经模拟加固系统及其在脑模型问题中的应用”。1961年 Minsky 发表了早期关于人工智能的优秀论文“逐步迈向人工智能”,后面这篇文章包括了有关现在被称为神经网络的一大部分内容。1967年,Minsky 出版了《计算:有限和无限机器》(*Computation: Finite and Infinite Machines*)一书,第一次以书的形式扩展了 McCulloch 和 Pitts 1943年的结果,并把该结果置于自动机理论和计算理论的背景中。

也是在1954年,Gabor 提出了非线性自适应滤波器的思想,他是早期通信理论的先驱者之一和全息照相术的发明者。他接着在合作者的帮助下致力于建立这样的机器,其细节描述在 Gabor 等于1960年发表的文章中有所提及,通过把随机过程样本以及希望机器产生的目标函数一起提供给机器来完成学习。

20世纪50年代,Taylor 开始研究联想记忆,接着,Steinbuch 引入了学习矩阵。这个矩阵由插在成行的“感觉”接收器和“马达”效应器之间的开关平面网络构成。1969年,Willshaw、Buneman 和 Longuet-Higgins 发表了关于非全息照相术的联想记忆的优秀论文。