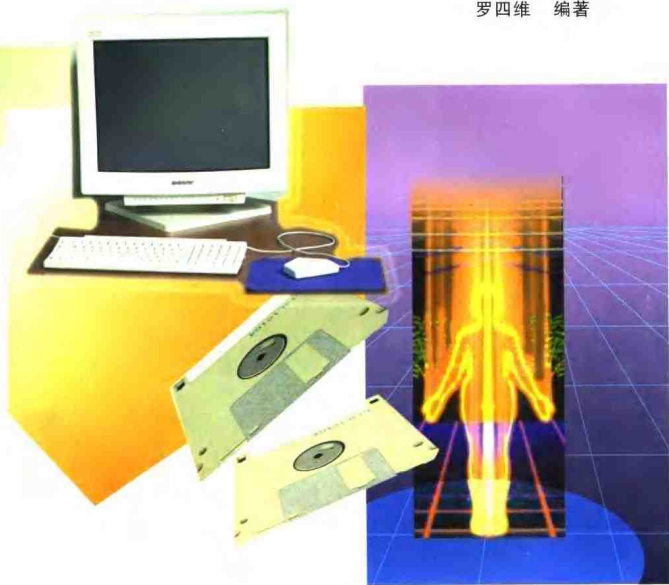


# 人工神经网络建造

罗四维 编著



中国铁道出版社

铁路科技图书出版基金资助出版

# 人工神经网络建造

罗四维 编著



中国铁道出版社

1998年·北京

(京)新登字 063 号

内 容 简 介

人工神经网络是近年来再度兴起并得到迅速发展的前沿交叉学科。它涉及到多学科的基础理论、方法、计算应用等众多问题。围绕这些问题,国内外的许多学者为发展人工神经网络编写了很多高水平的著作。

本书论述了大规模人工神经网络建造所关联的有关理论、方法,以及适合大规模并行处理的人工神经网络结构,主要内容包括人工神经网络的分布系统理论(热力学方法),适合并行分布处理的人工神经网络典型模型和大规模人工神经网络的实现技术。

本书可作为计算机信息处理、人工智能、系统工程等专业高年级本科生、研究生教材,也可作为有关科技人员的参考书。

图书在版编目(CIP)数据

人工神经网络建造/罗四维编著. —北京:中国铁道出版社,1998.4

ISBN 7-113-02925-6

I. 人… II. 罗… III. 人工神经网络计算机 N. TP387

中国版本图书馆CIP数据核字(98)第04531号

书 名:人工神经网络建造

著作责任者:罗四维

出版·发行:中国铁道出版社(100054,北京市宣武区右安门西街8号)

策划编辑:崔洪九

责任编辑:傅立涛

封面设计:马利

印 刷:北京市燕山联营印刷厂印刷

开 本:787×1092 1/16 印张:13 字数:310千

版 本:1998年5月第1版 1998年5月第1次印刷

印 数:1—1500册

书 号:ISBN7-113-02925-6/TP·291

定 价:26.80元

版权所有 盗印必究

凡购买铁道版的图书,如有缺页、倒页、脱页者,请与本社发行部调换。

# 前 言

人工神经网络是一门新兴的信息处理科学,它以人的大脑工作模式为基础,研究自适应的、非程序的信息处理本质。这种工作机制的主要优点是,不需要预先编写软件用先验的知识来处理问题,而是通过网络中神经元群体的相互作用来体现它自身的处理功能。因此,如何建造人工神经网络成为人工神经网络研究、实际应用的关键。

人工神经网络是介于传统处理方法和人脑处理方法之间的一种过渡。当把人的大脑看为一个由大量神经细胞相互作用组成的系统时,可以借用适用范围广泛的分布系统的观点来分析人工神经网络。而当把人工神经网络作为具有某种处理能力的大量单元相互作用组成的系统时,我们又可以使网络结构分层化,简化处理方法。

本书讨论了建造人工神经网络的基础和实际模型,以建造人工神经网络为主线,给出了热力学的分布系统分析方法,并以此作为大规模人工神经网络的理论基础;同时讨论了适合建造大规模人工神经网络的多种模型和大规模人工神经网络的实现方法。

本书既不是人工神经网络原始资料的储存,也不是一些论文的选取。它系统地提供了一种建造人工神经网络的方法论,体现了作者的哲理、信念和经验。从内容上讲,本书强调了分布系统论的分析方法论,以及实用的观点。

作者自1987年在国外就已经开始从事人工神经网络的研究,回国后为本科生、研究生编写了以人工神经网络为核心的《新一代计算机》和《计算机最新系统结构与技术》两本讲义;主持并完成了国家自然科学基金项目“人工神经网络同步并行计算机”的研究;在此基础上,吸收国内外,包括我们的研究成果写成此书。全书共分四章。第一章“绪论”简述了神经科学及人工神经网络发展的历史和有关的基础知识。第二章“人工神经网络的分布系统模型”给出了分布系统论所需要的基本数学工具、分布系统论方法,同时在本章讨论了基于分布系统论的典型模型,如 Ising 模型、Hopfield 网络和波尔兹曼机器。在本章的最后讨论了网络优化问题。第三章“人工神经网络的其它模型”讲述了人工神经网络的一般框架、感知器算法、误差反传递算法、竞争学习算法、遗传算法的神经网络构造方法和基于自适应共振理论的自组织网络。第四章“大规模人工神经网络的实现”讨论了计算机并行机制、Systolic 阵列结构神经网络处理和人工神经网络的专用硬件。

在本书编写过程中,始终得到师生和朋友的鼓励、帮助,对本书的编写作出了重要贡献,在此表示衷心感谢。编写过程中还得到了出版社的大力支持,他们对工作认真、负责、高标准、严要求的精神给我留下深刻印象,对他们的辛勤劳动表示诚挚感谢。

限于作者水平有限,错误和不妥之处敬请读者不吝赐教。

本书由铁路科技出版基金资助出版。

作者

1996年8月于北京

# 目 录

<b>第一章 绪 论</b> .....	1
1.1 历史回顾 .....	1
1.2 生物神经 .....	2
1.2.1 神 经 元 .....	2
1.2.2 信息传递 .....	3
1.3 神经组织 .....	4
1.4 视觉神经 .....	5
1.5 脑记忆的生理机制 .....	6
1.6 分布系统的特点 .....	7
1.7 分布系统的研究方法 .....	9
1.8 人工神经网络与分布系统.....	10
1.9 人工神经网络信息处理原理.....	10
1.9.1 神 经 元.....	11
1.9.2 人工神经网络的分类.....	13
参考文献 .....	14
<b>第二章 人工神经网络的分布系统模型</b> .....	15
2.1 基本数学工具.....	15
2.1.1 概率过程.....	15
2.1.2 连续时间的马尔科夫过程.....	16
2.1.3 离散状态系统与连续状态系统.....	18
2.1.4 概率微分方程.....	21
2.2 势条件和吉布斯分布.....	23
2.2.1 强势条件.....	23
2.2.2 弱勢条件.....	24
2.2.3 细致平衡条件.....	25
2.2.4 正则系统和正则-散逸系统.....	26
2.2.5 Ito,Stratonovich 概率微分方程及它们的福克-普朗克方程 .....	27
2.3 系 统 熵.....	31
2.3.1 最大熵原理.....	31
2.3.2 最小相对信息原理.....	33
2.3.3 最小平均“能量”原理.....	35
2.3.4 有序与无序平衡原理.....	35

2.3.5	系统平衡态的熵	35
2.3.6	平衡状态的平均能量	36
2.3.7	最大熵分布	37
2.4	概率网络	38
2.4.1	网络	38
2.4.2	Ising 模型	39
2.4.3	利用平均场近似	43
2.4.4	马尔科夫概率场和概率网络	45
2.5	Hopfield 网络	49
2.5.1	Hopfield 权值公式证明	50
2.5.2	连续 Hopfield 网	51
2.5.3	Hopfield 网络优化应用	53
2.6	波尔兹曼机器	54
	附录 1	60
	附录 2	61
	附录 3	61
	参考文献	62
2.7	网络优化	62
2.7.1	目标函数	63
2.7.2	最优化问题的概率模型	64
2.7.3	分布最优网络	65
2.7.4	模拟退火法	65
2.7.5	网络结构变换	66
<b>第三章</b>	<b>人工神经网络的其它模型</b>	<b>69</b>
3.1	人工神经网络的一般框架	69
3.1.1	一般框架	69
3.1.2	PDP 模型分类	72
3.1.3	PDP 模型分层机构	73
3.1.4	一般并行活动模型范例	75
	参考文献	80
3.2	感知器算法	80
3.2.1	感知器基本性质	80
3.2.2	感知器梯度算法	82
3.2.3	线性阈值元件感知器	88
3.3.4	最小二乘分类算法	88
3.3	误差反传递算法	89
3.3.1	两层网的缺点	89
3.3.2	扩展误差( $\Delta$ )规则	90

3.3.3	模拟结果	93
3.3.4	进一步扩展	99
3.3.5	改良 BP 算法	100
3.3.6	模拟程序	105
3.4	竞争学习算法	113
3.4.1	竞争学习机构	113
3.4.2	竞争学习	114
3.4.3	形式分析	115
3.4.4	实验结果	117
3.4.5	模拟程序	125
	参考文献	133
3.5	遗传算法的神经网络构造方法	133
3.5.1	遗传算法概述	134
3.5.2	遗传算法应用实例	136
3.5.3	遗传算法的形式描述	139
3.5.4	遗传算法神经网络	144
	参考文献	146
3.6	基于自适应共振理论的自组织网络	147
3.6.1	ART 的提出	147
3.6.2	ART 模型结构	147
3.6.3	竞争学习模型	149
3.6.4	任意输入环境中的自稳定学习	150
3.6.5	交替学习模型	151
3.6.6	2/3 规则	152
3.6.7	假定测试的自动控制	153
3.6.8	ART 模型的数学方法	155
3.6.9	ART 模型的学习算法	159
	参考文献	160
<b>第四章 大规模人工神经网络的实现</b>		<b>161</b>
4.1	计算机并行机制	161
4.1.1	并行性等级	161
4.1.2	并行结构	162
4.1.3	处理机阵列	163
4.1.4	开关网络	167
4.2	Systolic 阵列结构神经网络处理	175
4.2.1	Systolic 阵列结构原理	176
4.2.2	波前阵列	177
4.2.3	人工神经网络的 Systolic 阵列实现	178

参考文献 .....	184
4.3 专用硬件 .....	185
4.3.1 数字 VLSI 神经元处理器 .....	185
4.3.2 模拟电路 VLSI 神经元处理器 .....	188
4.3.3 系统实例 .....	191
4.3.4 光技术机器 .....	193



# 第一章 绪 论

人工神经网络是一门崭新的信息处理科学,近年来,由于神经科学,数理科学,信息科学,计算机科学的快速发展,使得研究以大脑工作模式,非程序的信息处理的人工神经网络成为可能。本章主要介绍人工神经网络的一些基本特点、概念和人工神经网络的发展历史。

## 1.1 历史回顾

第一台计算机于1946年问世以来,至今已走过半个多世纪。在这期间几乎所有的信息处理都是采用程序方式,即在人们对问题明确认识的基础上设计适合计算机处理的规划和算法,并以程序的方式实现信息的处理。

电子计算机是以冯·诺伊曼原理,用逻辑规划进行计算。计算机本身已多次更新换代,性能不断完善和提高,为本世纪科学技术,人类生活作出了巨大贡献,并推动社会、科学技术的迅猛前进。尽管如此,电子计算机的形象思维能力与人相去甚远。一台巨型计算机在识别能力上甚至无法与一个三岁玩童相比。人和计算机之间如此之大的反差,其原因是什么呢?这一问题已经对多方面的科学家提出了挑战,这是一个必须回答的问题。现在的一些研究结果对这一巨大的差别有了一些解释。应该说这是可喜的,但还是很肤浅的,表面的,有的甚至是假想的。就现在的科学水平而言,回答这个问题还为时过早。

人是一个复杂的系统,以大脑神经为控制中枢,人对自身的研究应该说是历史悠久,但对模拟人的人工神经网络研究也仅有30年的历史,迟后于电子计算机的发明。企图制作人脑功能的电脑这样梦想的大胆开拓者虽然为数不多,却代表了人类对科学的大胆幻想和期望。

人工神经网络可以说是从常规的信息处理到实现电子人脑的一种中间过渡,一方面人工神经网络的基本机理是模拟人脑的部分认知原理,另一方面又参与了对人脑的许多假设,并采用数学的,物理的方法达到信息处理的目的。这样在现阶段,人工神经网络已成为一种信息处理方法,一种为实现人脑功能的过渡技术方法。我们可以作这样一种比喻,人与计算机交流信息最简便,最自然的途径是使计算机具有视觉、听觉和说话的能力。但是图像和语音的识别远没有达到自由实用阶段,那么,多媒体技术可以看成是实现图像和语音的识别的过渡技术。

早在1943年麦可洛奇(McCulloch)和皮兹(Pitts)合作提出了神经元的M-P数学模型,从此开创了人工神经网络研究的新时代。在此以后,经过长期的努力,人工神经网络取得了不小的进展,形成了当今的信息处理学科。

从M-P模型开始,人们就用逻辑的数学工具研究神经网络对客观世界的表述,而且人工神经网络具有学习功能。1944年Hebb提出了改变神经元连接强度达到学习目的Hebb学习规则,至今在一些人工神经网络的模型中依然发挥重要作用。50年代中期,人工神经网络的研究有了较大的进展,主要表现在1958年后,Rosenblatt在他的“The perceptron: a probabilistic model for information storage and organization in the brain”一文中首先提出了感知器(per-

ceptron)的概念,用以模拟动物或人脑的感知和学习能力。感知器的学习过程是改变神经元之间的连接强度,适用于模式识别、联想记忆等人们感兴趣的实用技术。感知器模型包含了现代神经计算机的基本原理,在结构上也大体符合神经生理学知识。因此,掀起了人工神经网络研究的第一次高潮。好像人工神经网络无所不能,第一次进入了感知器的经典时期。此后,就转入了低谷,这一突然的降温源于 M. Minsky 和 S. Papert 发表的“Perceptrons”一书,这是一本辉煌的著作,指出感知器只能解决一阶谓词逻辑,对高阶谓词逻辑无能为力。这一结论导致了人们对人工神经网络研究的降温。另一方面,始终与人工神经网络相比较的冯·诺伊曼式数字计算机得到了飞快的发展,这又动摇了人们对神经网络研究的信心。到了 70 年代,一些新兴的科学不断出现,微电子技术也在飞速发展。这些都在不间断地刺激人工神经网络的发展。到了 80 年代,终于迎来了人工神经网络的第二次高潮。1982 年美国加州工学院物理学家 Hopfield 进行了开创性的工作,在他的“Neural Networks and Physical Systems With Emergent Collective Computational Abilities”<sup>[1]</sup>论文中使用了计算能量函数解决了离散人工神经网络的建造问题。1984 年 Hopfield 又在他的论文“Neurons With Graded Response Have Collective Computational Properties Like of Two-state Neurons”<sup>[2]</sup>中提出了连续状态的人工神经网络模型,该模型可用电路实现。Hopfield 的这些研究成果奠定了他在人工神经网络研究中的地位,同时也标志着人工神经网络研究的再度兴起。1985 年 D. H. Ackley, G. E. Hinton 和 T. J. Sejnowski 在“A Learning Algorithm for Boltzmann Machines”<sup>[3]</sup>一文中借用了统计物理的方法论提出了波尔兹曼机器学习算法。可以认为波尔兹曼机器学习是在 Hopfield 人工神经网络的基础上引入了随机变量。另外, Hopfield 人工神经网络又与研究磁体迁移现象的 Ising 模型相类似。此时可以认为研究复杂分布系统的热力学模型已经有效地应用于人工神经网络的研究。1986 年, D. E. Rumelhart, G. E. Hinton 和 R. J. Williams 在“Learning Internal Representations by Error Propagation”一文中提出了多层网络的误差反传播算法。这一算法解决了多年来人们未曾解决的多层网络学习算法,表明人工神经网络的计算能力具有很宽的应用范围。

80 年代是人工神经网络研究光辉的年代,引起了世界各国科学家、企业家的巨大热情,特别是科技发达国家纷纷成立研究小组、实验室,组织实施重大科研项目,我国自然科学基金项目委员会近年来一直把人工神经网络的研究作为资助课题。为了交流人工神经网络研究的结果,推动向深层次发展,1986 年 4 月美国物理学会在 Snowbirds 召开了国际神经网络会议,1987 年 6 月 IEEE 在 San Diego 召开了神经网络国际会议,成立了国际神经网络学会,1988 年 1 月“神经网络”杂志问世。1990 年 3 月 IEEE 神经网络会刊问世。自 1988 年起除了国际神经网络学会和 IEEE 一年召开一次的国际学术年会外,多种学术讨论会上都设有神经网络的论坛专题。

## 1.2 生物神经

人工神经网络的研究源于人脑神经网络。了解人脑神经网络的组成和原理有助于对人工神经网络的理解。

### 1.2.1 神经元

人脑是认识客观世界的器官。研究表明,人的意识、思维、行为等脑的高级功能都与客观世

界密切相关。神经科学起始于上个世纪末,1875年意大利解剖学家C. Golgi用染色体法首先识别出单个神经细胞,1889年Cajal创立了神经元学说,指出了神经系统是由结构上相对独立的神经细胞构成,在最近几十年来的研究结果认为人脑的神经元数量为 $10^{12}$ 。每个神经元包含了以下几个结构特性:

1. 细胞体(Cell Body),其大小在5至100微米的直径不等。细胞体由细胞核,细胞质和细胞膜组成。

2. 轴突(Axon),是细胞体向其它细胞伸出的最长一条分支,即神经纤维,相当于细胞的输出,每个神经元只有一个。

3. 树突(也称枝晶,Dendrites),是细胞体向外伸出的许多较短的树状分支,相当于细胞的输入。

4. 突触(Synaptic),是神经元之间连接的接口。整个脑内突触的数目大约在 $10^{14}\sim 10^{15}$ 之间,通过突触互连,连接方式不同,其生理作用也不同。突触的信息传递特性可变,因此细胞之间的连接强度可变,这是一种柔性连接,也称为神经元结构的可塑性。

另外,研究表明,神经元细胞膜内外之间存在电位差,称为膜电位,膜外为正,膜内为负。其大小约为几十微伏。膜电压接受神经其它神经元的输入后,电位上升或下降,当转入冲动的时空整合结果,使膜电位上升,而且当超过叫做动作电位的阈值时,细胞进入兴奋状态,产生神经冲动,由轴突输出,这个过程称为兴奋。动作阈值电位约为40微伏,当传入的冲动时空整合结果使膜电压下降并低于动作电压的阈值时,细胞进入抑制状态,无神经冲动输出。

### 1.2.2 信息传递

突触是神经细胞间传递信息的结构,突触由三部分构成,即突触前成分,突触间隙和突触后成分。突触所传递的信息采用电传递和化学传递两种方式。突触前成分是神经末梢上一个特化了的部分,突触末梢形成许多球形的小体,小体上直接进入突触连接部分的质膜叫做突触前膜,小体原浆中含有大量的突触小泡,小泡的直径约为200~800埃,内含神经递质。突触前膜外面是突触间隙,是突触前后之间的一个区域,其宽度为100~500埃。突触间隙的液体与细胞外液体是连通的,因此具有相同的离子组成。突触后细胞的一边是突触下膜,它是突触后细胞质膜特化的区域,含有特殊的分子受体。突触的结构示于图1.1.1。

突触信息传递的过程如图1.1.2所示。

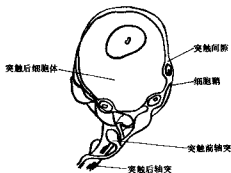


图 1.1.1 突触结构

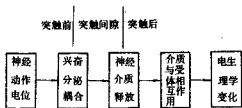


图 1.1.2 突触信息传递过程

在高等动物神经系统中,突触前的电活动不直接引起突触后成分的活动,不存在电学耦连。突触传递一般通过特殊的化学物质中介,这种物质叫做神经介质或递质,突触的信息传递只能由突触前到突触后,不存在反向活动的机制。因此突触传递是单方向的。兴奋—分泌耦连,神经介质释放和介质在间隙的扩散直到突触后膜的极化约需0.5~1毫秒,这就是突触延迟。用微电极技术研究脊髓前角细胞的膜电位,根据突触后电位的反应,将突触分为两种:兴奋性突触和抑制性突触。神经末梢释放介质使突触后膜产生极化反应,即兴奋性突触后电位,这是兴奋性突触。如果释放介质使突触后膜产生超极化反应,即抑制性突触后电位,则是抑制性突触。在许多可兴奋的细胞之间发现电学耦连。现已证明,缝隙连接在形态上代表这种耦连。电学突触能提供更大的传递速率,并在神经细胞间产生整合作用。

另外研究也表明了神经网络的复杂多样,不仅在于神经元和突触的数量大,组合方式复杂和联系广泛,还在于突触传递的机制复杂。已知的传递机制有:突触后兴奋,突出后抑制,突触前兴奋,突触前抑制和远程抑制等等。在突触传递机制中,释放神经递质是实现突触传递机制的中心环节,不同的神经递质有着不同的作用性质和特点。神经递质在维持正常的生理功能方面起着重要的作用,通过电信号—化学信号系统传递神经信息,进行复杂的信息加工,实现对机体的调节控制。

### 1.3 神经组织

神经元与神经纤维构成的神经组织具有两种基本特性,即兴奋与传导。当神经元的某一部分受到某种刺激时,在受刺激的部位就产生兴奋。这种兴奋会沿着神经元扩散开来,并在一定的条件下通过突触传达到相连的神经细胞。神经纤维的一部分兴奋起来时产生电位或电流的变化,这就是生物电流,是研究神经兴奋状态的标志。电位的变化非常迅速且短暂,叫做峰形电位。神经兴奋的结果产生神经冲动,神经冲动是能量传递的一种方式。通常,一条神经纤维接受一个适度的刺激时,则会产生一个冲动。但当对一个神经来进行刺激时,则会引起若干纤维同时冲动。刺激越强,则发生冲动的神经纤维越多,反之亦然。实验表明,神经冲动的能量并非来自刺激,而是神经纤维本身产生的,刺激的作用是引发神经纤维产生冲动。

对于人体来说,在外界、内界环境刺激下,通过神经系统,机体对刺激产生规律性回答。外部和内部刺激物作用于感受器,引起神经冲动。神经冲动沿着传入神经传达到神经中枢。通过中枢的神经联系,再经传出神经传达到效应器官,引起反应。这就是一个完整的反射过程。产生反射活动的基本结构组成有感受器,传入神经,神经中枢,传出神经和效应器。感受器一般是神经组织末梢的特殊结构,它把刺激的能量转变为神经的兴奋过程,所以感受器是一种换能装置。

某一特定的反射往往是在刺激特定的感受器后产生的,该特定感受器所在的部位成为这个感受器的感受野。神经中枢是指调节某一特定生理机能的神经细胞群。神经中枢的活动可以通过神经纤维直接影响效应器,在某些情况下,也可以通过体液的道路间接影响效应器,这种体液调节就是指内分泌腺的调节。这时反射是按感受器、传入神经、神经中枢、传出神经、内分泌腺、激素在血液中转运、效应器这样的过程进行。

反射可以分为两种:无条件反射和条件反射。无条件反射是先天性的,一定刺激作用于一定的感受野时,常引起一定的反射。例如,食物入口引起唾液分泌反射,机械刺激角膜产生眨眼

反射等。无条件反射使人们能初步适应环境。条件反射是在机体的生活中形成的,它可以随着机体的外部环境和内部状况的变化而变化。条件反射的建立大大扩展了机体的反应范围,比无条件反射有更大的预见性和灵活性,更适应于复杂变化的生存环境。

在实际活动中,无条件反射和条件反射的划分有相对的意义,肌体的每一活动都具有这两种反射的性质。在肌体内,无条件反射只有在新生时出现,在这以后由于条件反射不断建立,条件反射和无条件反射越来越不可分割地融合在一起。每次无条件反射出现,都有条件反射参与,而条件反射归根到底是在无条件反射的基础上建立的,它的构成已经把某些无条件反射的成分包括了进去。所以,几乎所有的生理机能都是无条件反射和条件反射的有机统一。

## 1.4 视觉神经

眼是人接收来自外部信息的最主要的接收器官,是最为复杂的感官器官。外界物体的光线射入眼中,聚焦后在视网膜上成像,视网膜发出神经冲动达到大脑皮层视区,产生视觉。人眼的简要水平切面如图 1.1.3 所示。角膜与晶状体之间是前房,虹膜与晶状体之间是后房,前房、后房之间都充满液体叫房水。

晶状体的后方直到视网膜充满透明的胶状物质,叫玻璃体。角膜,房水与晶状体等构成折光系统,它是透明的组织,能把物像形成在视网膜上。在所有的感官系统中,视网膜的结构最复杂。视网膜为感光系统,能感受光的刺激,发放神经冲动。它不仅有一级神经元(感光细胞),还有二级神经元(双极细胞)和三级神经元(神经节细胞)。

感光细胞有两种,视杆和视锥细胞。二者都与双极细胞形成突触联系。双极细胞外端与视杆细胞和视锥细胞相连,内端与神经节细胞相接。感光细胞的分布是不均匀的,视锥细胞分布在视网膜的中央凹部分。前视杆细胞则分布在视网膜的比较边缘部分。在视神经进入的地方没有感受细胞,形成盲点。视杆细胞的特点是对弱光有高度的感受性,因而是夜视觉的器官。视锥细胞对光的强度有较弱的感受性,因而是昼视觉的器官。视杆细胞中含有夜视觉所必须的视紫红质。视锥细胞中含有昼视觉所必须的视紫质。视锥细胞可以感受和分辨颜色。

视神经是由成束的神经节细胞的轴突组成,来自两侧的视神经在脑下垂体前方会合成视交叉。在这里组成每一根视神经的神经纤维束在进一步进入脑部之前被重新分组。从视神经交叉再发出的神经束叫作视束。在重新分组时,来自两眼视网膜右侧的纤维合成一束传向脑的右半部,来自两眼视网膜左侧的纤维合成另一束传向脑的左半部(如图 1.1.4 所示)。

这两束经过改组的纤维束继续向脑内行进,大部分终止于丘脑的两个被分成外侧膝状体的神经核。外侧膝体的细胞主要有两类,投射细胞和中间神经元。外侧膝体完成输入信息处理上的第一次分离,然后传送到大脑的第一视区和第二视区。外侧膝体属丘脑,是限到视皮层的中继站。这就是视觉通路。

视网膜上的光感受细胞通过光化学反应和光生物化学反应,产生光感受器电位和神经脉冲,在视网膜上沿垂直和水平两个方向传递信息。这样的信息沿着视觉通路进行传播。

视觉信号按连续步骤进行加工,发生在每一水平的转变和整合,从神经感受野可以作出最完善的分析。中枢神经元的感受野是指能影响某一视神经元反应的视网膜或视野的区域。每个视皮层,外侧膝状体的神经元或视网膜神经节细胞在视网膜上均有其特定的感受野,视网膜上神经细胞的感受野都是同心圆。可分为开中心圆型和闭中心圆型两类。外侧膝状体的

神经元的感受野与神经节细胞相似。皮层神经元的感受野至少可分为三种类型：简单型，复杂型，超复杂型。

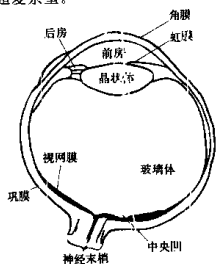


图 1.1.3 人眼结构示意图

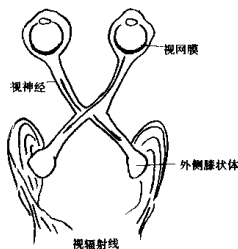


图 1.1.4 视神经

Hubel 和 Wiesel 于 1962 年提出了一个组合的感受野模式<sup>[4]</sup>。具有相同感受野的视皮层神经元在垂直于皮层表面的方向上呈柱状分布，它们是视皮层的基本功能单位，成为超柱。超柱内的神经元对同一感受野中图像和景物的各种特征进行并行的处理和译码，是产生主观感知的重要神经基础。现在大体上有两种超柱模型及译码理论：特征提取理论和空间频率分析器理论。视觉特征提取理论认为，视觉皮层的超柱是由许多不同特征提取功能柱所组成。每种功能柱内的细胞不但感受野相同，其功能也相同，根据功能不同可以分为方位柱，眼势柱和颜色柱等。与上述特征提取的超柱模型不同，视觉空间频率分析器的理论则认为，视皮层的神经元类似于傅立叶分析器，每个神经元敏感的空间频率不同。大量的研究表明，对复杂图形的视知觉即包含特征提取，又包括空间频率分析的译码过程。此外，视皮层以外的皮层联络区也协同活动。

Miller 和 Stryker 总结了眼优势柱的生理学研究成果之后，建立了眼优势柱模型<sup>[5]</sup>。他们认为视皮层活动性差异是形成神经元可塑性的基础，活动性差异的统计学规律是优势柱形成的基础，每只眼传入刺激引起的视皮层神经元单位发放模式的相关性，比两眼同时性传入刺激引起的发放差异，在眼优势柱形成中具有更大的作用；不仅传入刺激引起的突触前神经活动模式是重要的，突触后的皮层神经元反应模式在眼优势柱中也是重要的。这些假设基本符合 Hebb 突触理论观点。

## 1.5 脑记忆的生理机制

神经生理学家一直在从事脑在何处以及如何记忆的研究。研究表明，在脑细胞经受多次某个对象的刺激而保持连续兴奋状态时，只有当这种刺激达到一定强度（阈值）之后，脑细胞里才会留下痕迹。而且当这个对象刺激的频率提高时，阈值降低，记忆更加牢固。

人脑的大量神经细胞都以发射和不发射冲动对行为和精神活动作出贡献。这些细胞之间以多种方式交互作用。动态记忆理论是由桑克<sup>[6]</sup>在 1982 年提的，它描述了记忆是如何组织的。

记忆是怎样从过去经验中获得智能而自动改变和增长的。当证实过去的经验失败就存储新的经验,即由经验中学习,用联想及从联想中归纳的方法改变结构以适应预测的失败。动态记忆依赖于持续运行的神经冲动,或者依赖于可能由重复到达的冲动所强化的神经元内某些活跃的代谢变化和电位变化的保持。通过联系神经元的封闭环的神经冲动可能是这种动态记忆机制,每一记忆项目依赖于特定的神经元环或网的活动。事实上人们的记忆容量比神经元多,因此,不同的记忆可能共用部分通路。如果冲动实际上没有留下长时持续的痕迹,那么活动一旦停止,记忆便完全地,不可改变地消失掉。

另外,实验证明记忆不是纯动态的。但是这并不排除最初对神经元的依赖性。冲动的运行必然包含在那种留下记忆痕迹的最初经验中。重复造成更好的记忆这一事实告诉我们,冲动为了以后留下持久的物质变化必须在它选定的道路上循环流动。事实上巩固记忆痕迹的确需要时间作为代价。学习机制表明,突触结构的变化,或者神经蛋白质的变化,是这种持久的静态痕迹。那么神经元原来不起作用,由于活动而可能变得起作用并持续起作用。既然是记忆就存在一个记忆位置的问题。

在神经元生理学中,记忆研究最常用的方法是对人脑进行局部破坏,观察归纳障碍的情况。实验表明记忆与海马有密切的关系。另外与记忆有关的是杏仁体。杏仁体把感觉输入信号会聚成同样一些部分,又把神经纤维深入地送进大脑中的丘脑下部。杏仁体的多种联系构成为认为它能为记忆服务的多种作用的基础。从皮质感觉系统最终一个神经站来的神经纤维到达杏仁体。感觉印象在那里启动记忆系统的一条环路,它依靠杏仁体和丘脑之间的联系。杏仁体和丘脑之间的联系,杏仁体和丘脑下部之间的纽带似乎允许把一种体验与情感通路接合起来。那些纽带通过激活从杏仁体到感觉通路之间的反复联系,也可以使感情影响学习。从杏仁体返回到感觉区的联系的存在,有可能解释一个单一的刺激能引出多种多样的记忆。如当嗅到一种熟悉的气味时能唤起对它的外观、质地和味道记忆。

总的来说,两个主要的记忆回路分别起源于海马和杏仁体,它们负责许多种认知学习。除此之外,还存在用于学习的第二系统。复制刺激反应是该系统中的关键部分,我们称这种学习类型为习惯。习惯是刺激和反应的无意识接合。行为主义心理学家早就证明这种接合是所有学习的基础。行为主义观点中排除了通常意义中的“精神”、“认知”、以及“记忆”这样一些术语。学习可能依靠两个完全不同的系统,其中之一源自认知性习惯,另一个则是认知记忆的基础。这样就可调和行为主义和认知主义学派,行为就可能对刺激的无意识反应和由认知和期望所指导的行为的接合体。

## 1.6 分布系统的特点

计算机的问世及发展为人类的科学进步提供了强有力的研究手段。从第一台计算机诞生以来的半个世纪,计算机已经成为从事科学研究、商业活动、工业生产、以及日常生活不可缺少的工具和助手。而且随着科学技术、半导体和微电子技术的迅速发展,计算机体系结构也得到了快速的发展,无疑这是提高计算机处理能力的基础环境。

近年来,由于VLSI技术的进步,由多个具有智能处理器构成的系统已经出现,而且这种趋势在不断加强。切不可低估了这一动向,它将会带来一种质的变化。以前我们接触的多为机械式的被动要素或由这些被动要素以固定连接方式构成的系统。新的系统应该由自律性要素

按照柔性连接组成的系统。神经网络就是这类系统的一个模型。

分布系统是在人类周围存在的最为普遍的系统。分布系统与非分布系统存在着较大的差异。非分布系统往往受一个统一命令的指挥,有一个集中发布命令的机构。与之相反,分布系统有多个发布命令的机构,故而称之为分布的。显然分布系统应该包含异步这一层含义,既然没有统一的指挥机构自然表现为非同步。另外,对于分布系统不能期待整个系统为完全有序。即从宏观上看是有序的,而在微观动作上存在着随机性,以实现与其它系统或环境的协调。此外,分布系统中的每个要素相互竞争、协调,在各种制约下作为整体保持一定秩序,适应外界的变动。社会学中的很多现象,如一个人类集团内部,民族之间、国家之间,以至人与人,人类与自然之间等等普遍存在这种现象。

为了建立具有分布系统协调能力的自组织系统应该采取怎样的规则呢?为了说明这个问题,人们已经研究了一些有效的模型。具有代表性的模型有脑的神经模型、社会现象模型、生物系统模型、Ising 模型以及 Ising 扩展模型等等。很明显,作为分布系统的每一个具体模型,根据描述对象不同肯定存在着千差万别。但是,由于它们都是分布系统,故而应该保留一些共同点,这些共同点应该反映出分布系统的特点。

分布系统的主要特点大致可归纳为以下几个方面。

#### (1) 分布性

分布性体现在,系统由多个独立智能要素构成。独立的智能要素可以是简单要素,也可以是子系统,它们有某种自主的决定权利。

#### (2) 蠕动性

就分布系统本身而言,有其不确定性,即系统并非固定不变。在此,系统存在某种程度的流动性、模糊性、异步性,各要素之间通过相互作用、相互通信达到协调工作。

#### (3) 自组织

分布系统能适应环境的变化进行自组织,按一种规律或秩序来达到某一目的。自组织与自适应概念相似。自适应是通过自组织(或重构)实现的。系统在自组织过程中不断地从外界环境中吸取信息,不断地重新组织本身。

#### (4) 集中与分布共存

“集中”的含义是指用一个统一的命令体系来控制系统动作。在此我们对这一解释稍加扩展。如果在一个系统中有多个体、自律要素或子系统,它们执行一个动作,这种情况在客观效果上与执行统一命令相同,我们也使用“集中”这个术语。“分布”与“集中”恰好相对,如果所有的要素接受不同的、独立的信息,进行不同的动作,我们称之为“完全”的分布系统。对于一般的分布系统而言,它往往表现在“集中”与“完全”分布系统之间,即在分布系统中某些要素接受不同的、独立的信息,进行不同的动作,而另外一些要素可能在一个或多个不同的统一的命令体系控制下进行动作。恰当选择分布系统的组成形式有助于使系统的功能可以达到某一最终目的或实现某一秩序。

#### (5) 宏观性

我们在研究分布系统时并不注重每个要素的微观动作,如同在热力学研究中不关心每个气体分子的状态一样。气体分子在微观层次上完全无序,而在宏观层次上看来是均匀的。在分布系统研究中着眼于系统整体的宏观动作或秩序。当然,微观和宏观是相对的概念。例如,一个分子同组成它的原子相比又是宏观的。总之,我们从宏观层次出发,这样可以实现系统的聚



合化或宏观模型化。

## 1.7 分布系统的研究方法

分布系统所研究的对象是由大量自律要素组成的系统。科学的各个分支都给我们提供了大量的事例。在中学物理课程中我们已经知道,气体由大量的分子组成,比如说每立方厘米就有  $10^{23}$  个分子。气分子以完全无规则的形式飞来飞去,因而不可避免地造成分子之间频繁的碰撞。这样的系统具有巨大数目的动力学自由度。在通常情况下我们感兴趣的是气体的几个物理量,如压强、温度、密度等。利用这些物理量就可以描述系统的态,这就是热力学研究的方法。相比之下,晶体中的原子或分子排列得井然有序,它们仅轻轻地振动,我们感兴趣的是晶体的压缩率。细胞由复杂的细胞膜、细胞核和细胞质组成,它们之中的每一个都含有许多低一层的成分。在细胞中,在同一时间有条不紊地进行着几十到几千个新陈代谢过程。器官由以非常规则的方式合作着的大量细胞组成,各个器官又进一步服务于各种特定的目的,并在一个动物体内合作着,这些动物本身又组成了动物世界。世界上最复杂的系统大概是人的大脑,  $10^{12}$  个甚至更多的神经细胞组成了奇妙的大脑世界,通过这些神经细胞的合作才使人们能够说话,识别文字、图像,或者进行其它各种思维活动。

在工程学中仍然存在着这样的系统。它可能是机器,比如一台机车的内燃机,或整个工厂。再例如经济活动中的参与者,商品和货币的流通、贸易,商品的生产、消耗和储备。同样,包含有政治、宗教、职业或文化习惯的社会各种活动,也属于这类系统。

我们根据系统产生的方式,可将其分为两大类。一类是人类设计制造的人工系统,另一类是自然界自身形成的系统,我们称之为自组织系统。

在这里我们所关心的不是某一个具体的物理、化学或者生物系统等。我们强调的是对分布系统的研究方法。为此,我们应该首先考虑热力学方法,因为热力学是一门以普适的观点处理任意复杂系统的科学。另一方面,依据动力学的观点,可以通过尽可能精确地指定系统的所有动力学量来定义系统的态,这就是统计力学研究方法。

热力学和统计力学这两种方法有着明显的不同。热力学是唯象理论,统计力学则从物质的原子或分子结构以及原子世界的基本动力学原理出发,并结合概率论的逻辑推理。热力学为一种宏观的可以处理任意系统的普适理论。在热力学中,一个重要的概念是熵。熵表示了系统的特征状态。熵的变化由下式决定:

$$ds = (dq/T)_{re}$$

其中  $T$  为绝对温度,  $q$  为系统热量。在热力学中有两个重要的定律,那就是热力学第一定律和热力学第二定律。热力学第一定律指出孤立系统的能量是守恒的;热力学第二定律表明封闭系统熵增原理。利用能量守恒定律和最大熵原理就可以从宏观数据推导出系统的某些微观性质。统计力学则是通过微观理论去推导唯象热力学宏观定律,然而在这里的核心概念仍然是熵。

对于分布系统若从微观角度来考虑,它是由大量自律要素或子系统构成的系统;若从宏观上讲,它表现了整体的一致性,可实现聚合化。因此可以利用物理学中的两种研究方法:宏观的经典热力学方法和微观的统计力学方法。

由大量自律要素构成的分布系统,根据概率统计的观点,从宏观上看,各自律要素所处的状态可以看作是随机变量。正如统计热力学中把大量气体分子看作一个系统的每个分子状态