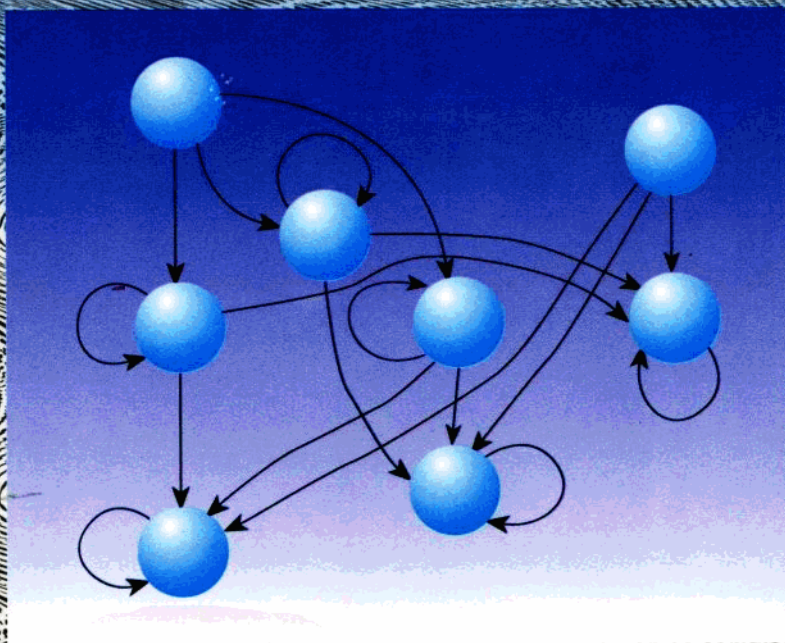


神经网络 应用基础

朱双东 编著



NEUPRESS
东北大学出版社

神经网络应用基础

朱双东 编著

东北大学出版社

图书在版编目(CIP)数据

神经网络应用基础/朱双东编著. —沈阳: 东北大学出版社, 2000. 3
ISBN 7-81054-527-2

I. 神… II. 朱… III. 神经网络 IV. TP183

中国版本图书馆 CIP 数据核字(2000)第 13260 号

内 容 提 要

本书是一本典型的面向应用的教材, 以 δ 学习规则和目前应用最广的 BP 网络为主线编写。书中各章均配有习题, 书后还给出了实验指导书。本书包括 8 章内容: 第 1 章, 绪论; 第 2 章, 神经网络基础; 第 3 章, 感知机; 第 4 章, BP 网络; 第 5 章, Hopfield 网络; 第 6 章, 随机神经网络; 第 7 章, 竞争型神经网络; 第 8 章, 神经网络应用。其中, 前 5 章和第 8 章是本书的重点内容。实验部分包括 3 个实验: M-P 神经元模型的建立, 感知机和 BP 网络。

本书是高等学校神经网络应用技术专业的研究生和高年级本科生的教学用书, 也可以作为相关专业学生的学习参考书。同时, 本书的内容还适合于有关科技人员自学。

©东北大学出版社出版

(沈阳市和平区文化路 3 号巷 11 号 邮政编码 110006)

电话: (024) 23890881

传真: (024) 23892538

沈阳农业大学印刷厂印刷

东北大学出版社发行

开本: 787mm×1092mm 1/16 字数: 275 千字 印张: 11

印数: 1~1000 册

2000 年 3 月第 1 版

2000 年 3 月第 1 次印刷

责任编辑: 孟 颖

责任校对: 冯 伟

封面设计: 唐敏智

责任出版: 秦 力

定价: 19.80 元

前 言

自从 1946 年世界上第一台电子计算机诞生的那一刻起,人们就一直在憧憬着能有那么一天,让机器来代替人的智力,就像 200 多年前产业革命的结果使机器代替了人类的体力、使人类从繁重的体力劳动中解放出来那样。于是,1956 年人工智能学科创立,并在其后的 20 多年时间里得到了蓬勃的发展和长足的进步。然而,人们逐渐发现,以专家系统为核心的人工智能远未能反映出人类智能的本质,其发展逐渐受到了限制。

实际上,早在人工智能学科创立之前,有关智能机理及其实现方法即神经网络的研究就已经开始了。人工智能学科创立之后,神经网络作为人工智能学科的一个分支继续发展。20 世纪 80 年代中后期,神经网络的研究取得了突破性的进展,开始走上了实际应用的发展阶段。目前,神经网络与模糊系统、进化计算等学科比翼齐飞,已经初步形成了智能科学领域的一个新的分支——计算智能,并稳步地发展着。在计算智能领域中,神经网络科学处于中心地位。10 余年来,神经网络的应用技术得到了良好的发展和推广,可以说,普及是当前神经网络研究中的一项非常重要的工作。

1992~1995 年,作者曾在日本东京大学进修,主要从事图像测量与处理以及神经网络应用方面的研究。当时,由于研究工作的需要,作者学习了神经网络的有关理论和知识,并向福岛邦彦(FUKUSHIMA)、北村新三(KITAMURA)等著名学者请教过相关问题。在理论上,神经网络的确是一门比较高深的学问,它不仅对数理基础的要求很高,而且涉及的学科也非常广泛,所以要学好神经网络这门学科,存在很大的困难。然而,神经网络的普及工作势在必行,面对的困难必须克服。因此,让神经网络这门新学科、新技术从神坛上走下来,是教育工作者义不容辞的责任。

回国后,作者自编了讲义,先后为本科生和研究生开设了“神经网络应用基础”(即“神经网络及其应用”)课。在编写讲义时,作者尽量避免涉及更多的学科和高深的数理基础,学生只要学过“计算机基础”、“线性代数”和“概率论”即可。“神经网络应用基础”是一门典型的面向应用的介绍神经网络的基本理论和应用技术的课程。不过,目前国内有关神经网络的书籍虽然不少,但大部分都属于学术著作,适用于多种学科的面向应用的教材却是凤毛麟角。刚开始上这门课时,最令人感到头疼的问题之一就是没有适用的教材。可以说,当前神经网络应用技术的普及非常需要一本面向应用的教材。这本教材应该是深入浅出、主线清晰、重点突出、重于实践的,它无需面面俱到,也无需将理论介绍得十分透彻。它的读者不是从事智能理论与技术研究的专业人员,而是各行各业希望能够采用智能技术来解决自己面临的各种问题的技术人员以及高等院校的教师和学生。

本书就是在作者编写的讲义的基础上进一步充实修改而成的。本书以 δ 学习规则和目前应用最广的 BP 网络为主线编写,各章均配有习题,书后给出了实验指导书,其中包括 3 个实验。

第 1 章“绪论”,主要介绍智能科学和神经网络科学的概况以及一些入门知识,以期使读者能对神经网络及本书在总体上有一个初步的了解。

第 2 章“神经网络基础”,采用将生物神经元与人工神经元对比、将生物神经系统(神经

网络)和人工神经网络对比的方法,由生物神经元引出人工神经元模型,由生物神经网络引出人工神经网络模型,来介绍相关的基础知识。在这一章中, δ 学习规则和 M-P 模型神经元是重点内容。此外,还介绍了网络结构、学习与记忆等一些基本概念。

第 3 章“感知机”,介绍的是第一个在实践和理论上都被证明了可行性的一种前向型神经网络——感知机,并采用 δ 学习规则来介绍它的算法。本章还对感知机的收敛特性等问题进行了较多的讨论。构成感知机这种神经网络的人工神经元就是第 2 章介绍的 M-P 模型神经元。

第 4 章“BP 网络”,介绍的是目前应用最为广泛的一种神经网络,它通常简称为 BP 网络。BP 网络采用的学习规则是在 δ 学习规则的基础上发展起来的,通常称为广义 δ 学习规则,也称为 BP 算法。BP 网络是在感知机的基础上发展起来的一种神经网络。本章对 BP 网络进行了较详细的介绍和讨论,其中包括 BP 算法的原理、BP 网络的实现方法、BP 网络的优劣以及 BP 算法的改进等。

第 5 章“Hopfield 网络”,介绍的是一种最为出名的反馈型神经网络——Hopfield 网络。本章的主要内容有 Hopfield 网络的模型与分类、反馈型神经网络的稳定性判据及其能量函数、Hopfield 网络的学习与记忆以及 Hopfield 网络的设计方法等。

第 6 章“随机神经网络”,介绍一类基于随机学习方法的神经网络,它也是一种反馈网络。主要内容有随机神经网络的学习算法——模拟退火算法和这类神经网络的典型代表波尔兹曼机。实际上, Hopfield 网络就是波尔兹曼机的一个特例。

第 7 章“竞争型神经网络”,介绍了另一种典型的前向网络——竞争型神经网络。本章的主要内容有竞争学习机理及特点,并介绍了几种典型的竞争网络:基本竞争网络、自组织特征映射网络和对向传播网络。

第 8 章“神经网络应用”,以 BP 网络的应用为主线,介绍了应用 BP 网络解决实际问题时应注意的一些基本问题,并列举了几个有代表性的例子。

实验部分的内容是在作者编写的实验指导书的基础上进一步充实而写成的。其中包括 3 个实验:M-P 神经元模型的建立、感知机和 BP 网络。这 3 个实验是一脉相承的,前一个实验是后一个实验的基础。另外,实验均未给出参考程序,这是为了让读者有更充分的发挥余地。从教学实践的效果来看,这样的安排是合理的,每年都有一些本科学生在实验中编制出质量很高的程序,为毕业设计打下坚实的基础。而研究生都发表了相关的科研论文。从教学实践来看,这门课中的实验有助于教学效果的提高。

本书反映了作者近几年的研究成果,主要集中在第 4 章和第 8 章,其中包括辽宁省科学技术基金资助研究项目。同时,本书的出版得到了抚顺石油学院研究生教材建设基金的资助。

本书可以作为研究生和本科生的《神经网络应用基础》这类应用型课程的教材,也可以作为相关课程的教学参考书,同时,也适用于有关的技术人员自学。

由于作者水平所限、经验不足,书中定会存在疏漏之处,敬请各位读者和专家学者批评指正。

编 者

2000 年 1 月于抚顺石油学院

目 录

前 言

1 绪 论	1
1.1 从“更深的蓝”谈起	1
1.2 人工智能	2
1.3 计算智能	4
1.4 神经网络研究的历史沿革	6
1.5 人脑的信息处理	9
1.6 Von Neumann 机及其局限性	12
1.7 神经计算机	15
习题一	16
2 神经网络基础	17
2.1 神经元模型	17
2.2 神经网络结构	23
2.3 学习与记忆的基本原理	26
2.4 神经网络的学习与记忆	29
2.5 几个概念	37
2.6 ANN 与 BNN 之异同	40
习题二	41
3 感知机	43
3.1 感知机模型	43
3.2 感知机的学习	45
3.3 感知机的收敛性	48
3.4 感知机的局限性	52
习题三	53
4 BP 网络	55
4.1 概 述	55
4.2 BP 算法的原理	57
4.3 BP 算法的实现	62
4.4 隐含层的作用	65
4.5 训练与测试	67

4.6 BP 网络的性能与局限	70
4.7 改进方案	71
习题四	75
5 Hopfield 网络	77
5.1 概 述	77
5.2 DHNN	78
5.3 能量函数	81
5.4 DHNN 的学习	85
5.5 联想记忆	87
5.6 CHNN	90
习题五	91
6 随机神经网络	93
6.1 概 述	93
6.2 模拟退火算法	94
6.3 波尔兹曼机	96
6.4 波尔兹曼机的学习	100
习题六	104
7 竞争型神经网络	105
7.1 概 述	105
7.2 竞争学习网络的结构	106
7.3 基本竞争网络的学习	109
7.4 SOM 网络的学习	112
7.5 CP 网络的学习	114
7.6 模式分类	117
7.7 竞争学习的特点	120
习题七	122
8 神经网络应用	123
8.1 概 述	123
8.2 模式对的规格化	126
8.3 胆甾液晶色/温特性的标定	130
8.4 经济规模评估	133
8.5 图像数据压缩	136
8.6 图像边缘检测	139
习题八	140

实 验	142
实验一 M-P 模型的建立	142
实验二 感知机	147
实验三 BP 网络	154
参考文献	165

1 绪 论

1.1 从“更深的蓝”谈起

1997年夏,一条震惊世界的消息在一夜之间迅速传遍全球,中国中央电视台也在黄金新闻时段作了特别报导。这条消息其实很简单,即:国际象棋世界冠军卡斯帕罗夫终于败在IBM公司的超级计算机“更深的蓝”的手下。这表明,机器终于抓住一次机会战胜了人类。

据说,除了专家系统之外,“更深的蓝”还采用了并行处理等新的智能技术,还有可能采用了神经网络技术。

毋庸置疑,此事对全世界的智能领域的研究者们来说是一个极大的鼓舞。人工智能科学,从1956年创立至今已有40余年,这期间,围绕着机器能不能代替人类的智能这类问题,在科学界一直存在着很多争论。智能科学的发展几上几下,经历了风风雨雨,如今,终于有了一个可靠的例证。这怎么能不令人欢欣鼓舞呢?

实际上,自从半个世纪前世界上第一台数字电子计算机诞生的那一天起,人们就渴望着能有这么一天,像200多年前机器代替了人的体力那样,让机器(计算机)也能代替人的智力。为此,人们进行了不懈的艰苦努力。然而,虽然半个世纪以来计算机技术取得了令人瞠目结舌的飞速发展和进步,但是,经过了大量的科学研究和实践之后,人们还是得出了一个十分令人遗憾的结论:现行的数字电子计算机尽管功能强大,但仍然无法代替人的智力。也就是说,这种计算机无法重现人类丰富的智能活动。

例如,一个婴儿能迅速地在一群人中辨认出自己的母亲来,但是,如果要求一台计算机辨认一个人,那就麻烦多了,即使采用现代的高新技术也是如此。仅以面部的辨识为例,计算机系统首先需要事先将辨识对象——婴儿的母亲的各种面部特征尽可能详尽地录入并存储起来(这代表计算机系统事先已认识婴儿的母亲),其中包括各种不同的角度、不同的表情等。然后将包括婴儿母亲在内的一群人用一台摄像机依次录入,并依次与已经存储在计算机内的婴儿母亲的所有正确模板进行比较,最后才能以概率的形式给出判断结果。这一过程是相当麻烦和费时的,而且面对复杂情况的适应能力很差,亦即很容易出错。如果不是根据面部特征而是根据形体特征,那对计算机来说可就太勉为其难了。

在40余年的发展历程中,智能科学的理论与技术逐步形成了两大分支领域,人工智能与计算智能。这两个领域的发展既是相互对立的,又是相辅相成的,在某种意义上可以说,它们代表了智能科学领域的两大派别,功能主义与结构主义。

所谓功能主义的基本观点是,认为任何复杂系统都可以看做是一种结构不明或结构不肯定的“黑匣子”或“灰匣子”,只要掌握了它的输入-输出关系,就可以根据这种关系设计出与该系统具有相同功能的人工系统。换言之,人工系统的结构可以与原型系统的结构完全不同或没有关系。基于上述思想,功能主义者主要致力于采用软件方法来构筑智能系统。基于现行计算机的人工智能就是功能主义的代表。

所谓结构主义,即结构决定论,主张有什么样的结构就有什么样的功能。基于这种观点,为了研究智能,首先就需要弄清脑的结构,然后再设法按照脑的有关结构来构筑相应的人工系统,从而达到模拟或复制人类智能的目的。因此,结构主义者主要致力于采用硬件方法来构筑智能系统。以人工神经网络为中心的計算智能就是结构主义的代表。

下面,对人工智能和計算智能作一个简单的介绍。

1.2 人工智能

研究人的智能的重要目的之一,就是使其能为人类的进步服务,为科学技术、生产生活服务,这就是说,要实现“人工智能”。

1.2.1 人工智能的诞生

众所周知,产业革命使机械化和自动化得到了迅速的发展,其结果是使得机器代替了人类繁重的体力劳动。人们只要以轻微的体力操纵机器,即可搬动异常沉重的物体,这也可以说是人的体力被机器“放大”了。20世纪40年代中期,数字电子计算机诞生了,随即迅速发展并获得了日益广泛的应用。于是,一个令人憧憬的设想便顺理成章地出现了:能否用机器(即计算机)来代替人类的智力劳动呢?也就是说,机器是否也能像“放大”人类的体力那样,来进一步“放大”人类的智力呢?

计算机能不能有智能,能不能思维,至今尚无完整的论证,在学术界是有争议的。这涉及到对“智能”认识的发展以及对脑、体劳动界限划定的不断推移。不过,对于目前计算机完成的数字计算功能,一般地已经不认为是高级的智能行为了。

当前的计算机虽然速度快得惊人,但它只是个计算熟练工,而现实世界需要处理的大多数实际问题并不是简单的数值计算,像语音、语声、图像、图形、判断决策、医疗诊断……等等,均非数值计算范畴,往往需要专门特有的经验知识才能得出正确的结论。这种由“数据世界”到“知识世界”的转化,由数据处理范畴扩展到符号知识处理范畴的转变,是人工智能诞生的重要因素。

试探性搜索,启发式的不精确的模糊的甚至是允许出错的推理方法,实际上更符合人类的思维过程,这一观点指导着人工智能的求解方法。启发式方法“既天才洋溢,又疏忽大意,不受拘束,巧妙而富有人性”,能把复杂问题大为简化,可在浩瀚的搜索空间中迅速找到解答。这是一种来自专门领域的经验知识,自然是限于特定场合,经常会取得成功但又不能保证必然成功,常能求得有关问题的满意解,但又不是数学上的最优解。上述这些观点促成了人工智能学科的诞生。

1956年夏天,美国的几位著名科学家(包括心理学、数学、计算机科学和信息论等)在美国 Dartmouth 大学召开了一次为期两个月的学术会议,讨论如何用计算机来模拟和实现人脑的某些智能。在这次会议上,McCarthy 建议列出一个计算机学科的新的分支,叫做人工智能(Artificial Intelligence,即 AI),并得到了与会者的赞同。于是,人工智能这一学科就诞生了。目前,学术界已经公认人工智能学科就是在此时建立的,至今已有40余年。

1.2.2 人工智能的发展

40年来,人工智能学科的发展是非常曲折的,几上几下,可以说是呈大起大落之势。其

发展历程既漫长而曲折,同时也十分有趣、发人深省。

首先,远在人工智能学科创立之前,人们就在结构主义思想的指导下,开始了用硬件方法来实现人工神经网络的探索,并取得了一些非常有价值的成果。这一过程一直延续到人工智能学科创立之后。到了20世纪60年代,研究进展逐渐变得缓慢起来,研究工作面临着越来越大的困难。于是,人们(当然不是所有的人)便来了一个彻底的大转变,在功能主义思想的指导下开始寻求利用软件方法来实现智能系统的探索,其典型的代表就是专家系统。后来,到了70年代末和80年代初,专家系统的缺陷开始越来越明显地暴露出来,于是,结构主义又以新的面貌东山再起……。总之,结构主义和功能主义在这一发展过程中此起彼伏、此落彼涨。从表面上看,结构主义和功能主义好像是对立的,但实际上它们是相辅相成的,是一个事物的两个方面。

尽管人工智能科学发展的道路是曲折的,但仍然取得了有效的进展。实际上,也只有建立起有说服力的智能程序和广泛应用人工智能,才能证明人工智能学科存在的真实意义。

20世纪50年代以游戏、博弈为对象开始了人工智能的研究。这期间,以电子线路模拟神经元及人脑均未获成功。1956年Logic Theorist程序获得了成功的应用。

20世纪60年代初期以搜索法、一般问题的求解研究为主。1960年,人工智能程序设计语言Lisp语言诞生。60年代后期,机器定理的证明取得了重大的进展,对规划问题也进行了研究。归结原理和语义网络的知识表示方法诞生,并出现了早期的专家系统。

20世纪70年代前期,人工智能受到责难,但人工智能的研究仍取得了有效的进展。这期间的研究以自然语言理解、知识表示为主。特别是认识到仅靠推理是不够的,知识对于实现智能是最为重要的。1974年,Prolog设计语言诞生。70年代后期,以知识工程、认知科学的研究为主,专家系统开始获得了较为广泛的应用,人工智能产业日益兴起。

20世纪80年代,以推理技术、知识获取、自然语言理解和机器视觉的研究为主。机器学习或知识获取已成为人工智能界的热门话题,机器翻译系统已成为商品。

实际上,现今对人工智能学科仍有赞成派和反对派。赞成派确信可以建立人工智能,即机器可以实现人工智能。批评者则说他们把天然智能解释得过于简单,从而夸大了技术上的可能性。反对派的代表是哲学家Drayfus,认为人的思维与机器思维并无共同之处,批评乐观主义者的观点是20世纪的炼金术。计算机的奠基人冯·诺依曼(Von Neumann)本人也说过计算机不会有智能。赞成派则批评反对派把思维神秘化了。

当前,人工智能已经开始朝着实用化迈进,不再限于理论讨论。

1.2.3 什么是人工智能

对于人工智能,目前尚难以给出严格的定义。它原本是计算机科学的一个分支,是研究使计算机能完成表现人类智能任务的学科。主要包括计算机实现智能的原理,制造类似人脑的智能计算机,以及使计算机更巧妙地实现高层次的应用。人工智能涉及计算机科学、心理学、哲学和语言学等多门学科,总的目标是增强人的智能。

人工智能是研究如何用计算机模仿人脑的推理、学习、思考、规划等思维活动,以解决那些需人类专家才能处理的复杂问题。从实用的观点来看,人工智能是一门知识工程学,以知识为对象,研究知识的获取、表示方法和使用。人工智能研究的是以符号表示的知识,而不是数值数据,采用的是启发式推理方法而不是常规算法。另外,控制结构与领域知识是分离的,并允许出现不正确的解答。

当前,对人类智能的理解大致有以下两种观点。

①符号主义 认为人类的智能基本单元是符号,认知过程就是符号表示下的符号运算,从而思维就是符号运算。这种观点基于“物理符号系统假设”,如有些成功的专家系统、自然语言理解系统等,都是基于这种观点的。

②连接主义 认为智能的基本单元是神经元,认知过程是由神经网络构成的,是并行分布的,而不是符号的运算。这种观点可避免知识表示问题,但随之而来的新问题是神经元连接的权值的设定比较困难。

这两种观点所构成的智能系统各有长短,目前看来,两者互相取长补短的结合似乎是更有发展前途的。

综上所述可知,神经网络的研究本来就是人工智能学科的一个分支。不过,目前通常将以基于符号主义的专家系统为代表的一类智能理论与技术称为人工智能,而将以基于连接主义的神经网络为代表的一类智能理论与技术称为计算智能。当然,这并非最后的定论,不过,这样来分类应该是有助于智能科学的研究与发展的。

通常,一个程序是否是智能程序,应根据其自然语言理解能力,自动推理能力以及机器学习能力来分辨。由此可见,多数被称为智能化的程序并非真正的智能程序,少数这样的程序可视为智能程序的雏型。

1.3 计算智能

1.3.1 计算智能的构成

计算智能(Computational Intelligence,即 CI)这一概念最早是由 Marks 于 1993 年和 1994 年提出的,随即受到了高度重视。计算智能主要包括神经网络、模糊系统、进化计算,它们都具有结构主义的特点(结构主义的基本观点是:有什么样的结构就有什么样的功能)。也有人认为计算智能还包括混沌理论,它们都是基于生命科学甚至人类智能的交叉学科,是新的智能理论与技术的主要组成部分。IEEE 随即于 1994 年在美国召开了一次规模空前的世界计算智能学术大会。

计算智能这一研究领域形成的积极意义在于,促进基于计算的和基于物理符号相结合的各种智能理论、模型、方法的综合集成,以利于发展思想更先进、功能更强大、解决问题能力更强的智能系统。下面,简单介绍一下计算智能所包含的各个研究领域。

(1) 神经网络

神经网络(Neural Networks,即 NN)的结构主义色彩很浓,它是模仿生物神经网络的结构与机能所构造出来的人工智能系统。在理论上,它本应该采用全硬件技术来实现,但限于目前的技术水平,大量的实现方法和技术还是基于软件技术来实现,因此,通常也称之为神经计算(Neural Computing,即 NC)。神经网络科学的发展和形成与人工智能的发展是密切相关的,它的起源甚至早于人工智能。

(2) 模糊系统

模糊系统(Fuzzy System,即 FS)的数学基础是模糊逻辑(Fuzzy Logic,即 FL)。模糊逻辑虽然也是解决不确定性问题的,但是,它与概率论不同,其主要区别是模糊逻辑强调主观的(语言表达)不确定性,而概率论强调的则是客观的不确定性(事件发生的可能性大小)。

例如,“明天要下雪”的不确定性是由概率来表示的,但是到了明天,这一事件就是确定性的事件了(无论是否下了雪)。而像“年轻”这一概念的不确定性和时间的经过无关,这是由语言造成的模糊概念,是主观的。例如,80岁的人可以说60岁的人还“年轻”,20岁的人则会将60岁的人看做是与自己的祖父同一辈的“老人”,而决不会说60岁的人还“年轻”。同样,20岁的人会认为30岁的人已经很“老成”了,而决不会说30岁的人比自己还“年轻”。

(3) 进化计算

进化计算(Evolutionary Computation, 即 EC)是基于达尔文的进化论,即生物进化的演化过程的一种算法,目前比较活跃的是遗传算法(Genetic Algorithms, 即 GA),它是按照“物竞天择,优胜劣汰”的原则来进行优化计算的。

(4) 混沌

混沌(Chaos)是由确定性非线性系统所产生的一种不确定的现象,它对系统的初始条件具有极端敏感性,初始条件的微小差别,可导致经过长时间后,系统的解完全不同,具有很强的不可测性。通俗地说,混沌是一种介于确定性和随机性之间的“无序”现象。混沌的研究目前也相当活跃,还形成了专门的研究领域。神经网络的研究中也常用到它,用于研究和模拟人脑的信息活动机理。

1.3.2 何谓神经网络

神经网络的基本理论与应用技术是本书要讨论的主要内容,那么,什么是神经网络呢?

神经网络是人工神经网络(Artificial Neural Networks, 即 ANN)的简称,这一概念来源于生物神经网络(Biological Neural Networks, 即 BNN)。在不致引起混淆的情况下,本书中所提到的神经网络一词均指人工神经网络。

在某种意义上可以说,人是世界上最完美的大系统,是仿生学的极限,人脑则是这一系统的核心。现行计算机(即冯·诺依曼机或记为 Von Neumann 机)就是仿人脑功能的一大成果,它导致了信息时代的到来,使信息科学得到了飞速发展,并带动了诸多学科的发展,可以说是20世纪最具代表性的伟大的科学成果之一。然而,在“智能技术”这一意义上来说,其功能则是十分有限的,终使“电脑”徒有虚名。

于是,人们由生物神经网络想到人工神经网络,向生命学习而构造了人工神经网络,它才可以说是人脑。

虽然神经网络与冯·诺依曼机都是向人类智能学习的成果,但是,它们具有相当大的不同之处。简单地说,冯·诺依曼机模拟的是人脑的逻辑思维功能,而神经网络模拟的是人脑的形象思维功能;冯·诺依曼机追求的是“思维(运算)”的精确解,而神经网络追求的是“思维”的满意解。例如,人们在识别一个熟人时,并不需要知道这个人的精确数据,而只要知道这个人的面部主要特征就足够了,这就是满意解的概念。

神经网络的研究领域涉及神经生理科学、认知科学、物理学、数学、信息科学、计算机科学、微电子学和光学等众多学科。

1.3.3 综合智能系统

目前,还有一个研究领域的动向十分值得注意,即综合智能系统(Hybrid Intelligent System, 即 HIS)。其目前的研究主要有以下特点:

①计算智能各分支之间相结合,其典型代表是综合计算智能系统(Hybrid Computa-

tional Intelligent System, 即 HCIS);

- ②计算智能与人工智能相结合;
- ③以 Von Neumann 机为支持系统;
- ④神经网络处于中心地位。

目前,计算智能的应用主要体现在神经网络和模糊系统以及它们的结合上。在计算智能的应用领域中,最活跃的有两个:智能信息处理和智能控制,尽管它们尚处于发展的初期,但都是比较成功的应用领域。

1.4 神经网络研究的历史沿革

神经网络研究是人类探索模仿脑神经系统信息处理智能装置的一个相当重要的领域,在讨论神经网络的有关具体问题之前,有必要先简单地了解一下神经网络科学的发展历程。其发展历程几经兴衰,基本上可分为以下4个阶段。

1.4.1 启蒙期

1890年,美国心理学家 W. James(《心理学》一书的作者)对相关学习、联想记忆的基本原理进行了开创性的研究。他认为,当两个基本脑“突起”(即神经细胞,通常称为神经元)同时或相继被激活时,二者之一将把这种兴奋传递给另一个。他还预言,神经元的活动是其输入的总和。上述假设与当前神经网络中相关学习、联想记忆的概念有着密切的关系。

几乎就在同时,西班牙解剖学家 Cajal 创立了神经元学说,他认为神经元有两极,一个极接受信息的输入,另一个极产生相关的输出。

1943年,美国心理学家 McCulloch 和数学家 W. H. Pitts 首先从信息处理的角度出发,提出了著名的形式神经元的数学模型,开创了神经科学研究的新时代。该模型以他们的姓氏的首写字母命名,简称为 M-P 模型,至今仍是神经科学研究的重要基础之一。

1949年,心理学家 D. O. Hebb(中文译名为“赫布”)根据心理学中的条件反射机理,在其所著的《行为组织(Organization of Behavior)》一书中提出了著名的 Hebb 学习规则:如果两个神经元都兴奋,那么它们之间的连接强度将得到增强。这是第一个神经元学习规则,其正确性在30年后得到了证实,而且至今仍是神经科学研究的重要基础之一。

1958年, F. Rosenblatt 提出了著名的感知机模型,他弄清了当时使人们深感困惑的两个问题:

- ①信息是如何存储或记忆的?
- ②存储的信息是如何影响识别和行为的?

感知机与当时的计算机及人工智能完全不同。感知机理论的提出导致神经网络的研究出现了一个高潮,实际上,目前许多成功的神经网络都是在感知机的基础上发展起来的。

1960年, B. Widrow 和 M. Hoff 发表了《自适应开关电路》一文,提出了自适应线性神经元模型 adaline 及著名的 δ 学习规则。同时,他们还制作了一个实际装置,能学会识别字母等简单图形。这是第一位对神经网络的硬件系统作出杰出贡献的学者。可以说,20世纪60年代掀起了神经网络研究的第一次热潮。

1.4.2 低潮期

事物的发展总是会有曲折的,神经网络的研究也不例外。当时,神经网络的研究还是人

工智能科学的一个组成部分,或者说是人工智能科学领域中的一个学派。随着研究工作的进一步深入,人们遇到了来自认识、应用和实现方面的各种困难和迷惑,一时难以解决。对于神经网络的学习能力问题,在学术界引起了很大的争议。于是,神经网络的研究开始进入了缓慢发展的低潮期,有人甚至称之为“冰河期”。神经网络研究形成低潮期的主要原因大致有以下3点:

①当时还没有找到一种行之有效的算法来支持神经网络的进一步研究。

②微电子技术当时取得了飞速的进步,已经到了超大规模集成电路(VLSI)的实用阶段,冯·诺依曼机处于全盛时期,计算机的处理速度和存储容量急速提高,软件日益丰富。此间,基于逻辑符号处理方法的人工智能的研究得到了迅速发展并取得了显著的成就。因此,很多学者都把实现人工智能的发展方向,寄托在不断改进和提高冯·诺依曼机的性能这一途径上。

以上两点是形成低潮期的客观原因,是事物发展的必然结果。辩证唯物主义认为:存在决定意识。因此,低潮期的形成还有一个重要的主观原因。

③1969年,人工智能学科的主要创始人之一、美国麻省理工学院的 M. Minsky 教授经过数年的潜心研究之后,与 S. Papert 共同出版了影响极大的《感知机》一书。Minsky 认为,感知机这种网络功能非常有限,甚至连“逻辑异或(XOR)”这样简单的问题都不能解决。

在书中,进行了大量分析论证之后,他们说:“感知机具有某些局限性,而正是由于这些局限性,使其具有研究的价值。这些局限性也正是感知机一部分魅力之所在:线性、有趣的学习定理、作为一类并行计算例证的简单性等。没有理由认为,任何一个这种特性会导致转向多层形式。这就是说,作者直观认为,将感知机引入扩展到多层装置没有意义。这种看法的正确与否乃是一个大的研究课题。或许在将来,有一些更强有力的定理会被发现,或者能明白无法在多层装置中获得有意义学习定理的道理。”

由于 Minsky 在人工智能领域的巨大威望和他的悲观结论,给当时神经网络沿感知机方向的发展研究当头泼了一盆凉水。此后,美国联邦基金有15年之久未对神经网络的研究投资,前苏联也取消了几项很有前途的研究计划。

值得一提的是,在上述这种“山穷水尽疑无路”之际,少数具有远见卓识的学者并没有放弃努力,仍然在极为艰难的条件下不懈地坚持进行神经网络的研究,他们的研究为以后神经网络研究的复兴奠定了基础。

1969年,美国学者 S. Grossberg 与 G. A. Carpenter 提出了自适应共振理论(ART)模型,并在此后的若干年中发展了这一理论,先后共提出了3个版本的模型,即 ART1, ART2 和 ART3。这是迄今提出的最复杂的神经网络之一。

1972年,芬兰学者 T. Kohonen 提出了自组织映射(SOM)理论,并称其神经网络结构为联想存储器。与此同时,神经心理学家 J. Anderson 也提出了类似的神经网络结构,并命名为交互存储器。SOM 网络是一类很重要的无导师学习网络,主要用于模式识别、语音识别及分类等应用场合。

日本学者福岛邦彦(K. Fukushima, 大坂大学教授)在1972~1984年提出过数种神经网络及其算法,其中最著名的是1979年提出的新认知机理论。该理论在发展过程中,由无导师学习改进为有导师学习,有很强的工程应用观点。

日本学者甘利俊一(S. Amari, 东京大学教授)一直致力于与神经网络有关的数学理论

的研究,谋求以数学的观点来解释神经网络。

另一方面,人们逐渐发现,计算机虽然在复杂的科学计算方面显示了巨大的威力,但是,计算机却很难学会一些很普通的知识和经验,在这一点上,超级计算机甚至都不如一个二三岁的儿童。这一切迫使计算机科学工作者不得不慎重思考:智能问题是否可以完全由人工智能中的逻辑推理规则来描述?人脑的智能是否可以在计算机中重现?通俗地说,就是:人工智能在冯·诺依曼机上能实现吗?

1.4.3 复兴期

1982年,美国加州工学院的著名物理学家 J. J. Hopfield(霍普菲尔德)提出了一种新的全互联神经网络模型,这是一种反馈型神经网络模型,后来被以他的名字命名为 Hopfield Neural Networks(英文缩写为 HNN,即“霍普菲尔德网络”之意)模型。

在这种新的神经网络模型中,他的突出贡献是引入了“能量函数”概念,并给出了网络的稳定性判据。更可喜的是, Hopfield 用简单的模拟电路实现了这种模型,成功地求解了 TSP(旅行商、巡回推销商)问题,并实现了 4 位 A/D 转换器,取得了令人满意的解。Hopfield 的研究成果为神经计算机(Neurocomputer)的研制奠定了基础,同时,开创了神经网络用于联想记忆和优化计算的新途径。这标志着新的曙光出现了,人们终于迎来了“柳暗花明又一村”的新局面。

1984 年之后的 3 年内,美国 AT&T 公司贝尔实验室利用 HNN 理论实现了基于硅芯片的硬件神经网络。G. E. Hinton 和 T. J. Sejnowski 提出了一种随机型神经网络模型——波尔兹曼机,学习过程中采用了模拟退火技术,有效地克服了局部极小的影响。

1986 年,由 Rumelhart 和 McClelland 领导的 PDP(即 Parallel Distributed Processing,意为“并行分布处理”)小组提出了 PDP 理论,并出版了《并行分布处理》一书共 3 卷。PDP 理论着重于认知微观结构的探索,指出了神经网络模型具有 3 个属性:结构、神经节点传递函数和学习算法;并进而提出了著名的 BP 算法。他们最突出的贡献之一是发展了多层网络学习的 BP 算法,解决了多层网络学习的问题。BP 算法迄今仍是应用最多最广泛的神经网络学习算法。

1.4.4 高潮期

1987 年 6 月,第一届神经网络国际会议在美国加州圣地亚哥召开,约 200 位专家学者发表了研究论文,与会的学者达一千多人,规模之宏大、讨论之热烈实属罕见。当大会执行主席 S. Grossberg 教授向大会通告与会者中有著名的物理学家 J. Hopfield 等人时,全场顿时爆发了经久不息的热烈掌声。

在这次大会上,国际神经网络学会(INNS)正式成立。这标志着神经网络的研究发展进入了一个新纪元,神经计算机时代到来了。

1988 年 1 月,神经网络杂志《Neural Networks》创刊。1990 年 3 月,IEEE 神经网络会刊《IEEE Transaction on Neural Networks》问世。

我国于 1989 年 10 月在北京召开了一个非正式的神经网络学术会议,即神经网络及其应用研讨会,并出版了会议论文集。1990 年 12 月在北京召开了中国首届神经网络学术大会,这次会议确定:今后每年召开一次中国神经网络学术大会。1991 年 12 月在南京召开的中国神经网络 1991 学术大会期间,正式成立了中国神经网络学会。

这期间,各国政府和军方都对神经网络和神经计算机的研究与开发应用给予了高度的重视与支持。现仅举数例如下。

美国政府制定了“神经、信息、行为科学(NIBS)”计划,投资5.5亿美元用于开发有关新一代计算机的基础研究。美国国防部于1987年发表了一个DARPA(高级研究计划局)报告,宣布从1988年11月开始执行一项为期8年的神经计算机研究计划,总投资为4亿美元。美国国防部的DARPA认为,神经网络“看来是解决机器智能的惟一希望”,“这是一项比原子弹工程更重要的技术”。

日本称1988年为“神经计算机年”,当年,日本通产省提出了所谓“人类尖端科学计划(Human Frontier Science Program)”,期望在多学科的广泛的国际合作的基础上,研制出能增强和模拟人类智能行为的计算机系统。

此外,还有前苏联的“高技术发展”计划、法国的“尤里卡”计划、德国的“欧洲防御”计划等,不一而足。

我国的“863高技术研究”计划,于1990年批准了关于人工神经网络的3项课题,自然科学基金和国防科技预研基金也都将神经网络的有关研究列入了选题指南,对中选的课题提供研究上的资助。我国国家攀登计划重大项目“认知科学中神经网络理论与应用基础研究”于1992年10月启动,目前已经取得了可观的研究成果。

在这股热潮中,企业界当然也不甘落后。实际上,无须专注大企业的动向,只要看看下面的一些数据,就已经很能说明问题了。

1989年,全世界经营神经网络产品的公司有45家,1990年就达到了65家,而1991年又猛增至135家。1990年神经网络产品的总销售额为1.5亿美元,利润约为5000万美元,神经网络工业收入约为8500万美元,并以年增长率61%的速度递增。

总之,1987年神经网络的研究开始进入了一个新的时期,由理论研究进入实践应用的水平,软、硬件都有了具体的实现,神经网络的应用日渐活跃。

20世纪90年代以来,神经网络的研究已逐渐进入一个平稳发展的时期,神经网络已经不仅仅是一门科学理论,在技术化和应用发展的道路上开始迈出了坚实的步伐。智能计算机的出现也许还需要一定的时日,但已经不是可望而不可及的事情了。

神经网络的研究之所以能够出现上述新热潮,主要有以下几个方面的原因:

- ①神经科学研究的突破和进展;
- ②计算机科学与人工智能发展的迫切需要;
- ③新技术的迅速发展和进步,为神经网络的硬件实现提供了物质基础和技术上的保障,如VLSI,光学,超导,生物技术等等;
- ④非线性科学的迅速发展。

下面,将Von Neumann机与神经网络作一个简单的比较。为此,需要首先了解一下人类智能信息处理的主要特点。

1.5 人脑的信息处理

1.5.1 人脑的能力

众所周知,现行的Von Neumann机还无法代替人脑的大量的形象思维的功能。从形