



普通高等教育“十一五”国家级规划教材

# 智能计算

吴 微 周春光 梁艳春 编著

3-43



高等教育出版社

普通高等教育“十一五”国家级规划教材

# 智能计算

吴微 周春光 梁艳春 编著

近年来，神经网络、遗传算法、模糊推理、进化计算等智能计算方法，在几乎所有领域都取得了广泛的应用。智能计算已成为当前最热门、最活跃的、最富有挑战性的课题。有关智能计算的专著很多，有的从原理入手，或者材料全面深入，或者专精于某一方面应用，但是适合于大学工科专业教材的并不多见。本书从工科专业教材的角度出发，本书主要着眼于智能计算的原理、设计思路、实施方案以及应用中的注意事项。而对于智能计算的某些具体应用，本书仅在绪论中集中加以阐述，其余各章则偶尔提及。因此，本书命名为“智能计算”，而不是“智能计算方法”。本书力求做到深入浅出、循序渐进、由浅入深、由易到难、由点到面、由理论到应用。本书可作为高等院校工科专业及相关专业的教材，也可供从事智能计算工作的工程技术人员参考。

本书可以分为三大部分：  
第一部分：智能计算的原理、设计思路、实施方案以及应用中的注意事项。  
第二部分：智能计算的某些具体应用。  
第三部分：智能计算的某些具体应用。  
本书可作为高等院校工科专业及相关专业的教材，也可供从事智能计算工作的工程技术人员参考。

本书可作为高等院校工科专业及相关专业的教材，也可供从事智能计算工作的工程技术人员参考。

TP183-43  
W883

高等教育出版社

2009年9月  
ISBN 7-309-07123-0

## 内容简介

本书着眼于实用,从计算方法的角度简要介绍近年来得到广泛应用的神经网络计算、模糊计算和进化计算,着重讨论这些智能计算方法的设计思路、性能特点、实施方案以及应用中的注意事项。书中包含了作者的一些相关研究工作成果。本书可作为理工科高年级本科生和研究生的选修课教材,也可供各相关领域研究者 and 应用工作者参考。

## 图书在版编目(CIP)数据

智能计算 / 吴微, 周春光, 梁艳春编著. —北京: 高等教育出版社, 2009. 12

ISBN 978 - 7 - 04 - 027953 - 5

I. 智… II. ①吴…②周…③梁… III. 人工智能 - 计算 - 高等学校 - 教材 IV. TP183

中国版本图书馆 CIP 数据核字(2009)第 192436 号

策划编辑 杨波 责任编辑 蒋青 封面设计 张楠 责任绘图 黄建英  
版式设计 王莹 责任校对 胡晓琪 责任印制 毛斯璐

---

出版发行 高等教育出版社  
社 址 北京市西城区德外大街 4 号  
邮政编码 100120  
总 机 010 - 58581000

经 销 蓝色畅想图书发行有限公司  
印 刷 国防工业出版社印刷厂

开 本 787 × 960 1/16  
印 张 12.5  
字 数 220 000

---

购书热线 010 - 58581118  
咨询电话 400 - 810 - 0598  
网 址 <http://www.hep.edu.cn>  
<http://www.hep.com.cn>  
网上订购 <http://www.landaco.com>  
<http://www.landaco.com.cn>  
畅想教育 <http://www.widedu.com>

版 次 2009 年 12 月第 1 版  
印 次 2009 年 12 月第 1 次印刷  
定 价 16.50 元

---

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 27953 - 00

# 前 言

近年来,以人工神经网络、模糊系统和进化计算三者为代表的所谓计算智能,在几乎所有应用领域都得到了广泛应用,在理论上也提出了许多重要而有趣的课题。有关计算智能的专著很多,并且各有特点,或者材料全面深入,或者专精于某一方面应用,但是适合于大学理工科教材的似乎不多。本书试图在这一方面做初步的努力和尝试。

计算智能的着眼点在于它们都是用某种数值计算方法来模仿或实现人的智能或自然规律。反过来说,也可以说是模仿人的智能或自然规律,给出解决实际问题的一套计算方法,称之为智能计算方法。应该将这些智能计算方法与传统计算方法一样,视为现代计算方法中有益的甚至是不可缺少的组成部分。从这一角度出发,本书主要着眼于这些智能计算方法的性能、设计思路、实施方案以及应用中的注意事项。而对于这些方法如何模仿人的智能或自然规律这一方面,我们仅在绪论中集中加以简略的讨论,在其余各章则偶尔提及。因此,本书命名为“智能计算”,而不是“计算智能”。这样的处理方式必然要忽略掉许多重大、深入而且有趣的课题,但是也有利于在有限的篇幅内,把智能计算方法的特性与用法讲得更清楚。这一点对于将智能计算作为一种方法来解决问题,而不是作为一种理论来深入研究的大多数理工科学生和科研工作者来说,也许更为实用和迫切。书中引用了作者的一些相关研究工作成果和论著。

本书可以分为三大部分,分别处理神经网络(1—4章)、模糊计算(5—8章)和遗传算法(9—11章),各部分相对独立。

本书编辑杨波、蒋青以及匿名审稿人提出许多详尽而又中肯的修改意见,尤其是责任编辑蒋青耐心细致的工作态度令人叹服。中国石油大学(华东)王健老师仔细校对了全书。本书作者对他们的辛勤劳动与卓越见解表示深深的感激。另外,本书作者感谢国家自然科学基金的资助(资助号 10871220, 60673099, 60873146)。

编者

2009年9月

第三章 径向基函数网络 .....	42
3.1 径向基函数 .....	42
3.2 径向基函数参数的选取 .....	44
3.3 Gauss 函数 .....	47
习题3 .....	49

# 目 录

12	.....9.3 梯度定理与极值问题.....	133
21	.....9.4 遗传算法收敛性分析.....	140
22	.....习题9.....	146
27	.....4.3 ART神经网络.....	148
101	.....第15章 遗传算法的改进与取用.....	148
101	.....10.1 遗传算法的改进.....	148
63	<b>绪论</b> .....	1
63	.....0.1 什么是计算智能.....	1
75	.....0.2 计算智能与智能计算.....	3
77	.....0.3 人工神经网络.....	3
77	.....0.4 模糊系统.....	5
79	.....0.5 进化计算.....	5
79	.....0.6 人工神经网络、模糊系统和进化计算的相互融合.....	7
81	.....1.1 线性感知器.....	8
87	.....1.2 BP网络.....	14
90	.....1.3 BP网络的应用.....	20
92	.....习题1.....	22
92	<b>第一章 前传网络</b> .....	8
92	.....2.1 简单线性联想网络(LAM).....	25
99	.....2.2 Hopfield 联想记忆模型.....	26
101	.....2.3 利用外积和的双极性 Hopfield 网络.....	27
103	.....2.4 Hopfield 网络的存储容量.....	29
103	.....2.5 Hopfield 网络的收敛性.....	31
106	.....2.6 二次优化问题的 Hopfield 网络解法.....	35
106	.....2.7 双向联想记忆(BAM)网络.....	37
106	.....习题2.....	41
111	<b>第二章 Hopfield 联想记忆神经网络</b> .....	25
111	.....3.1 径向基函数.....	42
119	.....3.2 径向基函数参数的选取.....	44
119	.....3.3 Gauss 条函数.....	47
122	.....习题3.....	49
122	<b>第三章 径向基函数网络</b> .....	42

<b>第四章 自组织竞争网络</b> .....	51
4.1 SOFM 网络 .....	51
4.2 SOFM 网络的应用 .....	55
4.3 ART 神经网络 .....	57
习题 4 .....	61
<b>第五章 模糊集合与模糊逻辑</b> .....	63
5.1 模糊集合与隶属度函数 .....	63
5.2 模糊逻辑与模糊推理 .....	72
习题 5 .....	77
<b>第六章 模糊神经网络</b> .....	79
6.1 神经网络与模糊系统的结合 .....	79
6.2 模糊神经元 .....	81
6.3 模糊前馈网络 .....	84
6.4 模糊联想记忆网络 .....	87
6.5 模糊加权推理网络 .....	90
习题 6 .....	95
<b>第七章 模糊聚类</b> .....	96
7.1 引言 .....	96
7.2 模糊 $C$ -均值聚类法 .....	99
7.3 传递闭包法 .....	101
7.4 最大树法 .....	103
习题 7 .....	105
<b>第八章 模糊控制理论与应用</b> .....	106
8.1 模糊控制原理 .....	106
8.2 蒸汽发动机模糊控制系统 .....	108
8.3 还原炉温度的模糊控制系统 .....	114
习题 8 .....	117
<b>第九章 遗传算法基础</b> .....	119
9.1 遗传算法概述 .....	119
9.2 遗传算法的实现技术 .....	122

9.3 模式定理与积木块假设 .....	133
9.4 遗传算法收敛性分析 .....	140
习题 9 .....	146
<b>第十章 遗传算法的改进与应用 .....</b>	<b>148</b>
10.1 遗传算法的改进 .....	148
10.2 用于求解函数优化的遗传算法 .....	157
10.3 遗传算法求解旅行商问题 .....	160
习题 10 .....	164
<b>第十一章 其它进化算法与群智能优化算法 .....</b>	<b>165</b>
11.1 遗传规划 .....	165
11.2 进化策略 .....	173
11.3 进化规划 .....	174
11.4 遗传算法与进化策略和进化规划的比较 .....	175
11.5 蚁群优化算法 .....	176
11.6 粒子群优化算法 .....	178
习题 11 .....	178
<b>参考文献 .....</b>	<b>180</b>

## 0.1 什么是计算智能

计算智能(Computational Intelligence, 简称CI)并不是一个新的术语,早在1988年,加拿大的一种刊物便以CI为名。1993年, Bob Marks写了一篇关于计算智能和人工智能之区别的文章,并在文中给出了对CI的理解。1994年的国际计算智能会议(WCCI)的命名就部分地源于Bob的文章。这次IEEE会议将

## 绪 论

计算机的诞生和发展是 20 世纪科学技术最伟大的成就之一。半个多世纪以来,计算机本身几经更新换代,其性能日益优越,越来越广泛地应用于社会生活的各个领域,对推动科学、技术和社会的发展起到了难以估量的作用。

计算机是按 von Neumann 的串行体系结构来实现算术和逻辑运算的,现在的运算速度已达每秒几千万亿次,其结果的精确和可靠程度更是人工无法比拟的。但是,计算机的形象思维能力却与人脑相差甚远。在人的知觉、记忆、语言、思想与获取的心理过程中,尤其是在实时处理中,人脑这一慢速和充满噪声干扰的硬件是在进行着极其复杂的宏并行处理。人们对于十分复杂的事物可以不假思索、一目了然地予以识别。但是,即使很简单的物体让先进的计算机来识别却常常相当困难。因此,要使计算机能在交互式的自然过程中,提取简单信号,表达知识及其结构,并使其与有关知识结合起来产生智能,具有较强的形象思维能力,就必须突破 von Neumann 机的结构体系,另辟蹊径。

制造具有智能的计算机一直是人类的梦想。1956 年人工智能技术的出现,标志着人们为此做出的巨大努力。近年来,随着人工智能应用领域的不断拓广,传统的基于符号处理机制的人工智能方法在表示知识、处理模式信息及解决组合爆炸等方面遇到的障碍已变得越来越突出,这些困难甚至使得某些学者对人工智能提出了强烈的批判,对人工智能的可能性提出了质疑。

基于上述原因,人们自然地转而寻求具有自组织、自适应、自学习等智能特征的大规模并行算法。近年来,出现了一些新的研究方向如神经网络、模糊控制和进化计算等,由于它们都是从模拟某一自然现象或生命过程而发展起来的,并且具有高度并行化与智能化等特征,因而引起了人们的极大兴趣。这些新方法通过一些相对简单的“拟物”与“仿生”策略,为解决许多复杂问题提供了新的契机、新的思路。由此,计算智能应运而生。

### 0.1 什么是计算智能

计算智能(Computational Intelligence,简称 CI)并不是一个新的术语,早在 1988 年,加拿大的一种刊物便以 CI 为名。1993 年, Bob Marks 写了一篇关于计算智能和人工智能之区别的文章,并在文中给出了对 CI 的理解。1994 年的国际计算智能会议(WCCI)的命名就部分地源于 Bob 的文章。这次 IEEE 会议将



国际神经网络学会(NNC)发起的神经网络(ICNN)、模糊系统(FUZZ)和进化计算(ICEC)三个年度性会议合为一体,并出版了名为《计算智能》的论文集。此后,CI这个术语就开始被频繁地使用,同时也出现了许多关于CI的解释。

1992年,James C. Bezdek提出,CI是依靠生产者提供的数字、数据材料进行加工处理,而不是依赖于知识,而AI则必须用知识进行处理。1994年,Bezdek在Florida, Orlando, 94 IEEE WCCI会议上再次阐述他的观点,即智能有三个层次:

- ① 生物智能(Biological Intelligence, 简称BI),是由人脑的物理化学过程反映出来的,人脑是有机物,它是智能的基础;
- ② 人工智能(Artificial Intelligence, 简称AI),是非生物的,人造的,常用符号来表示,AI的来源是人类知识的萃取;
- ③ 计算智能(Computational Intelligence, 简称CI),是由数学方法和计算机实现的,CI的来源是数值计算的传感器。

这三个层次从复杂程度来看, $BI > AI > CI$ ,CI与AI的差距要比AI与BI的差距小得多;而从所属关系来看,CI是AI的一个子集,虽然AI不是BI的子集,但是BI经常用来指导AI模型,同样也指导了CI。Bezdek还认为,模糊系统特别适合从CI到AI的平滑过渡。

Bob Marks于1993年提出,将神经网络、遗传算法、模糊系统、进化规划和人工生命统称为计算智能。这一定义得到许多学者的认同,包括IEEE的领导们,其中NNC的主席Walter Karplus在1996年的ADCOM会议上重述了这种观点,认为CI着重研究的是系统的工作方式,而不是高强度计算。

David Fogel在1995年发表评论认为,智能就是对环境的适应能力。但他认为CI高于AI,CI包含AI。

Eberhart, Dobbins和Simpson关于CI的理解是,将智能系统置于一个环境中,智能行为的标准是改变或作用于环境的能力。而CI只是智能系统的一个内部节点,适应性只是CI的衡量指标。CI是一种方法论,是通过计算实现适应和处理新形势的能力,具有推理的属性,能得到预测或决定的结果。

另外,也有些人认为CI和AI仅有部分重合。他们认为,AI是符号主义,基于知识、规则和推理,相当于人的左脑;而CI属于连接主义,基于数据、学习和记忆,相当于人的右脑。

计算智能主要包括以下三方面:

- ① 人工神经网络(Artificial Neural Network, 简称ANN),用来模仿人脑的由大量简单神经元组成的网状结构和并行信息处理过程,着重于计算智能的拓扑结构;
- ② 模糊系统(Fuzzy System, 简称FS),试图描述和处理人的语言和思维中

存在的模糊性概念,着重于计算智能的表述语言;

③ 进化计算(Evolution Computing, 简称 EC),是一种模仿生物进化过程的优化方法,着重于计算智能的演化机制。

## 0.2 计算智能与智能计算

将人工神经网络、模糊系统和进化计算统称为计算智能,着眼点在于它们都是用某种数值计算方法来模仿和实现人的智能或自然规律。反过来也可以说,这三者是模仿人的智能或自然规律,给出解决实际问题的一套计算方法,称之为智能计算方法。应该将这些智能计算方法与传统计算方法一样,视为现代计算方法中有益的甚至是不可缺少的组成部分。从这一角度出发,本书主要着眼于这些智能计算方法的性能、设计思路、实施方案以及应用中的注意事项。而对于这些方法如何模仿人的智能或自然规律这一方面,我们仅在本章集中加以简略的讨论,在其余各章则偶尔提及。因此,本书命名为“智能计算”,而不是“计算智能”。这样的处理方式必然要忽略掉许多重大、深入而且有趣的课题。但是我们期望这样处理有利于在有限的篇幅内,把智能计算方法的特性与用法讲得更清楚,这一点对于将智能计算作为一种方法来解决问题,而不是作为一种理论来深入研究的大多数人来说,也许更为实用和迫切。

## 0.3 人工神经网络

人工神经网络是指模拟人脑神经系统的结构和功能,运用大量的处理部件,由人工方式构造的网络系统。神经网络是一个非线性动力学系统,它突破了传统的、串行处理的数字电子计算机的局限,以分布式存储和并行协同处理为特色,虽然单个神经元的结构和功能相当简单有限,但是大量的神经元构成的网络系统所实现的行为却是极其丰富多彩的。神经网络的研究已有五十多年的历史,但其发展是不平衡的,它的兴衰还与“人工智能走什么路”这一争议问题有关。起始阶段进展不快,并一度陷入低谷。但仍有不少有识之士在极其艰难的条件下坚持研究,并在模型建立等理论方面取得了突破性的成果。时至今日,人工神经网络成了信息领域的热门研究课题。

以下是人工神经网络发展的三个阶段。

(1) 第一阶段:初始发展期(20世纪40年代—20世纪60年代)

早在1943年,美国心理学家 McCulloch 和数学家 Pitts 联合提出了形式神经元的数学模型,即 M-P 模型,从此开创了神经科学理论研究的新纪元。M-P 模型能够完成有限的逻辑运算。1944年,心理学家 Hebb 提出了改变神经元间连

接强度的 Hebb 规则,它们至今仍在各种神经网络模型中起着重要的作用。1957年,计算机科学家 Rosenblatt 用硬件完成了最早的神经网络模型,即感知器(Perceptron),并用来模拟生物的感知和学习能力。1962年,电机工程师 Windrow 和 Hoff 提出了自适应线性元件 Adaline,它是一个连续取值的线性网络,在信号处理系统中用于抵消通信中的回波和噪声,应用十分广泛。

### (2) 第二阶段:低谷期(20世纪60年代末—20世纪70年代末)

1969年,人工智能之父 Minsky 和 Papert 发表的《感知器》一书指出,感知器无科学价值而言,连 XOR 逻辑分类都做不到,只能作线性划分。由于 Minsky 在学术界的地位和影响,故其后若干年内,这一研究方向处于低潮。另一方面,传统的 von Neumann 电子数字计算机正处在发展的全盛时期,整个学术界都陶醉在成功的喜悦之中,从而掩盖了新型计算机发展的必然。

尽管如此,在此期间仍然有不少有识之士不断努力,在极端艰难的条件下致力于神经网络研究,为这一研究的发展奠定了理论基础。Boston 大学的 Grossberg 和 Carpenter 提出了自适应共振理论 ART 网络。芬兰 Heisinki 大学的 Kohonen 提出了自组织映射网络。日本大阪大学的 Fukushima 提出了神经认知机神经网络模型。日本东京大学的 Amari 对神经网络进行了数学理论的研究。

### (3) 第三阶段:兴盛期(20世纪80年代以后)

20世纪70年代末期,试图模拟视听觉的人工智能研究遇到挫折,人们习以为常的知识难以教给计算机。计算机的设计者和制造商也发现前面有不可逾越的线路微型化的物理极限,人们开始思考 von Neumann 机到底还能走多远。同时,VLSI、脑科学、生物学、光学的迅速发展也为人工神经网络的发展打下了物质基础。

1982年,加州大学的物理学家 Hopfield 提出了 Hopfield 网络模型,并用电路实现。1985年,Rumelhart 提出了 BP 算法,把学习的结果反馈到神经网络的隐层,来改变权系数矩阵,这是迄今为止应用最普遍的网络。Hinton 等人提出了 Boltzman 机模型。1988年,蔡少堂提出了细胞神经网络模型。

近年来,神经网络理论引起了美国、欧洲与日本等国家和地区的科学家和企业家的巨大热情。新的研究小组、实验室、风险公司等与日俱增,世界各国也正在组织和实施与此有关的重大研究项目。如美国 DARPA 计划、日本 HFSP 计划、法国尤里卡计划、德国欧洲防御计划和前苏联高技术发展计划等。1986年4月,美国物理学会在 Snowbirds 召开了国际神经网络学术会议。1987年6月,IEEE 在 San Diego 召开了国际神经网络学术会议,并成立了国际神经网络学会。1988年起,IEEE 和国际神经网络学会每年召开一次国际会议。1990年3月,IEEE 神经网络会刊问世。

神经网络理论的应用已经渗透到各个领域,在智能控制、模式识别、计算机

视觉、自适应滤波、信号处理、非线性优化、语音识别、知识处理、传感技术与机器人等方面取得了令人鼓舞的进展。神经网络代表一种新的主义,即连接主义,用于建立诸如知识表达、推理学习、联想记忆乃至复杂的社会现象等复杂系统的统一模型,它预示着一个新的工业的到来。

## 0.4 模糊系统

模糊概念其实是我们日常生活中司空见惯的,如“天气很热”“味道很怪”等。人的大脑能在信息不完整不精确的情况下作出判断与决策,即进行模糊信息处理。模糊数学是用数学方法研究和处理客观存在的模糊现象的一门新兴学科。它所研究的事物的概念本身是模糊的,即对于一个对象是否符合这个概念难以确定,这种由于概念外延的模糊而造成的不确定性称为模糊性。

模糊理论诞生于1965年。美国控制论专家、加利福尼亚大学教授 L. A. Zadeh 首先提出模糊集合的概念,发表了开创性论文《模糊集合论(Fuzzy Sets)》。他提出,模糊数学的核心思想就是运用数学手段,仿效人脑思维,对复杂事物进行模糊处理。模糊数学在基础理论和实际应用等方面引起了各国学者的极大兴趣,并产生了许多有价值的应用和惊人的成果。1973年,Zadeh 教授又提出模糊逻辑(Fuzzy Logic)理论,并积极倡导将模糊理论向人工智能方向发展。模糊逻辑的研究虽然时间还不很长,但在智能模拟和智能控制等领域却已有了飞快的发展。1974年,印度裔英国学者 E. H. Mamdani 首先将模糊理论用于锅炉和蒸汽机的控制,并实验成功,开创了模糊控制的新领域。20世纪80年代后期以来,在日本采用模糊控制技术的家电产品大量上市,模糊技术在图像识别、自动控制、市场预测、人工智能等领域普遍应用,掀起了一股模糊热。日本、美国和我国都成功地研制出了智能化的模糊推理机,这表现了模糊理论的强大生命力。另一方面,模糊理论在学术界也得到了普遍的认同和重视。1992年,IEEE 召开了第一届关于模糊系统的国际会议(FUZZ-IEEE),并决定以后每年举行一次。1993年 IEEE 创办了专刊“IEEE Transactions on Fuzzy Systems”。当前,模糊理论和应用正向深度和广度进一步发展,研究成果大量涌现,已经成为世界各国高科技竞争的重要领域之一。

## 0.5 进化计算

进化计算(Evolution Computing)采用简单的编码技术来表示各种复杂的结构,并通过简单的遗传操作和优胜劣汰的自然选择来指导学习和确定搜索方向。由于它采用种群(即一组可能解)的方式组织搜索,这使得它可以同时搜索解空

间的多个区域,特别适合大规模并行计算。进化计算具有自组织、自适应、自学习的特点,并且不受其搜索空间限制性条件(如可微、单峰等)的约束,不需要其它辅助信息(如导数)。这使得进化计算不仅能获得较高的效率,而且操作简单、通用性强。进化计算在 20 世纪六、七十年代并未受到普遍的重视。其主要原因一是这些方法本身还不够成熟;二是由于这些方法需要较大的计算量,而当时的计算机还不够普及且速度较慢,这样便限制了它们的应用;三是当时基于符号处理的人工智能方法正处于其顶峰时期,使得人们难以认识到其它方法的有效性及其适应性。到了 20 世纪 80 年代,人工智能方法的局限性越来越突出,并且随着计算机速度的提高和并行计算机的普及,进化计算对机器速度的要求不再是制约其发展的因素。由于进化计算在机器学习、过程控制、经济预测、工程优化等领域取得的成功,引起了各领域科学家们的极大兴趣,自 20 世纪 80 年代中期以来,世界上许多国家都掀起了进化计算的研究热潮。目前,有数种以进化计算为主题的国际会议在世界各地定期召开,并已出版了多种有关进化计算的杂志。可以预料,随着进化计算理论研究的不断深入和应用领域的不断拓广,进化计算必将取得更大的成功。

进化计算包括三大分支:遗传算法(Genetic Algorithm,简称 GA)、进化策略(Evolution Strategies,简称 ES)和进化规划(Evolution Programming,简称 EP)。这三个分支在算法实现方面具有一些细微的差别,但它们具有一个共同的特点,即都是借助生物进化的思想和原理来解决实际问题。

遗传算法是一类模拟生物界自然选择和自然遗传机制的随机化搜索算法,由美国 J. Holland 教授于 1975 年首次提出。它是利用某种编码技术作用于称为染色体的二进制数串,其基本思想是模拟由这些串组成的种群的进化过程,通过随机的、然而是有组织的信息交换来重新组合那些适应性好的串。遗传算法对求解问题的本身几乎一无所知,它所需要的仅是对算法所产生的每个染色体进行评价,并根据适应性来选择染色体,使适应性好的染色体比适应性差的染色体有更多的繁殖机会。遗传算法尤其适用于处理传统搜索方法难于解决的复杂的非线性问题,可广泛用于组合优化、机器学习、自适应控制、规划设计和人工生命等领域。

进化策略由德国柏林工业大学的 I. Rechenberg 等人在 1964 年提出。他们在求解流体动力学柔性弯曲管的形状优化问题时,发现用传统的方法很难优化设计中描述物体形状的参数,于是利用生物变异的思想来随机地改变参数值并获得了较好的结果。随后,他们便对这种方法进行了深入的研究和发展,形成了进化计算的另一个分支——进化策略。进化策略与遗传算法的一个不同之处在于,进化策略直接在解空间上进行操作,而遗传算法是将原问题的解空间映射到位串空间之中,然后再施行遗传操作。进化策略主要用于求解数值优化问题。

进化规划的方法最初是由美国 L. J. Fogel 等人在 20 世纪 60 年代提出的。他们在人工智能的研究中,期望智能行为具有能预测其所处环境的状态,并按照给定目标作出适当响应的能力。在进化规划研究中,他们将模拟环境描述成由有限字符集中的符号组成的序列。

## 0.6 神经网络、模糊系统和进化计算的相互融合

神经网络能够通过学习和训练获得用数据表达的知识,除了可以记忆已知的信息之外,还具有较强的概括能力和联想记忆能力。但神经网络的推理知识表示体现在网络连接权值上,难以直接、直观地理解,这是它的一个缺点。

模糊系统以扎德创立的模糊集合论为基础,抓住了人类思维中的模糊性特点,以模糊推理来处理常规方法难以解决的问题,能对复杂事物进行模糊识别、模糊度量。模糊系统的显著特点是能够直接表示逻辑,适于直接的或高级的知识表达,具有较强的逻辑功能。但是对于模糊系统来说,模糊推理虽然是一种善于表现知识的推理方法,但它本身没有获取知识的能力,模糊规则的确定也比较困难,通常需要专家知识的指导。如何构造可自动处理模糊信息的模糊系统,即实现模糊规则的自动提取和模糊变量隶属度函数的自动生成及优化,一直是困扰模糊信息处理技术进一步推广的难题。

解决上述两个难题的途径之一是利用神经网络的自学习自适应功能,将符号逻辑推理方法与连接机制方法相结合,将数值方法和模糊逻辑方法相结合,从而导致模糊神经网络的产生。

神经网络的网络结构的设计和权值的训练是一个十分重要而困难的问题,传统的方法多是凭经验来设计网络,用梯度法来确定其中的权值,常常需要进行反复试验,而且很难找到最优的网络结构和权值。而进化计算模拟生物进化的过程,以全局并行搜索方式来搜索优化种群中的最优个体,其主要优点是简单、鲁棒性强、搜索空间大。进化计算为神经网络的自动设计和训练提供了一种新的途径。

神经网络、模糊系统和进化计算三者的研究同步发展、相互渗透,界限日益模糊,由三者按各种方式交叉融合而组成的新系统往往具有更强的功能。这方面研究是当今一大热点。

# 第一章 前传网络

## 1.1 线性感知器

### 1.1.1 概述

图 1.1 给出了一个简单的单层前传网络(神经元)的示意图。它也是许多更复杂的神经网络的基本构件之一。神经元对外界传入的  $N$  个信号经权值向量  $\mathbf{w}$  和阈值  $b$  处理后,用线性求和器得到“综合印象”,再由活化函数  $g(\cdot)$  对此综合印象做出非线性反应。这种反应机制是对真正的生物神经元反应机制的一种简单而又常常有效的模拟。将大量简单神经元按某种方式连接起来,并通过某种学习过程确定单元之间的连接强度(权值  $\mathbf{w}$  和阈值  $b$ ),就得到各种人工神经网络,用来完成逼近、分类、预测、控制和模拟等各种任务。

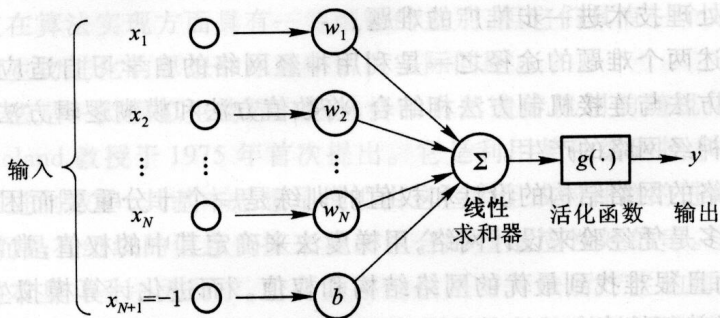


图 1.1 神经元模型

一旦确定了非线性函数  $g(t): \mathbb{R}^1 \rightarrow \mathbb{R}^1$  以及权值向量  $\mathbf{w} = (w_1, \dots, w_N)^T \in \mathbb{R}^N$  和阈值  $b \in \mathbb{R}^1$  之后,对任一输入向量  $\mathbf{x} = (x_1, \dots, x_N)^T \in \mathbb{R}^N$ ,神经元的响应输出为

$$y = g(\mathbf{w} \cdot \mathbf{x} - b) = g\left(\sum_{n=1}^N w_n x_n - b\right) \quad (1.1.1)$$

为了确定权值  $\mathbf{w}$  和阈值  $b$ ,需要给定  $J$  个输入样本  $\{\mathbf{x}^j\}_{j=1}^J$  以及相应的理想输出  $\{o^j\}_{j=1}^J \subset \mathbb{R}^1$ ,其中  $\mathbf{x}^j = (x_1^j, \dots, x_N^j)^T \in \mathbb{R}^N$ 。神经元的學習过程就是利用这些样本,通过某种学习算法来选择权值向量  $\mathbf{w}$  和阈值  $b$ ,使得

$$o^j = y^j \equiv g(\mathbf{w} \cdot \mathbf{x}^j - b) = g\left(\sum_{n=1}^N w_n x_n^j - b\right), j = 1, \dots, J \quad (1.1.2)$$

其中  $y^j$  为网络的实际输出。 $\mathbf{w}$  和  $b$  确定之后,神经网络就可以开始工作过程,即向网络输入  $\mathbb{R}^N$  中其它模式向量  $\mathbf{x}$ ,由式(1.1.1)得到相应输出。

式(1.1.1)中的函数  $g(\cdot)$  称为活化函数,常见的有符号函数、S型函数(Sigmoid函数)、径向基函数、随机值函数等等。网络输出  $y$  可以只取有限个离散值(例如双极值  $\pm 1$  或二进制  $0, 1$ ),这时网络相当于一个分类器;也可以取连续值,这时网络相当于输入  $\mathbf{x}$  与输出  $y$  之间函数关系的一种数值逼近器。当存在  $\mathbf{w}$  和  $b$  使式(1.1.2)成立时,我们说该问题是可解的,或样本集  $\{\mathbf{x}^j\}_{j=1}^J$  是可分的。否则,称为不可解的,或不可分的,这时只能选取  $\mathbf{w}$  和  $b$  使得误差  $o^j - y^j$  ( $j=1, \dots, J$ ) 尽可能地小。

### 1.1.2 线性感知器

下面,我们取  $g(t)$  为如下符号函数(如图 1.2 所示):

$$\text{sgn}(t) = \begin{cases} 1, & t \geq 0 \\ -1, & t < 0 \end{cases} \quad (1.1.3)$$

并且假设理想输出  $o^j$  取值亦为  $\pm 1$ 。(也可以令符号函数的取值为  $0, 1$ 。一般来说,取值为  $0, 1$  时电路实现方便,而取  $\pm 1$  时数学处理比较简单。)对输入模式  $\mathbf{x}$ ,网络实际输出为

$$y = \text{sgn}(\mathbf{w} \cdot \mathbf{x} - b) \quad (1.1.4)$$

式(1.1.4)所表示的神经网络称为线性感知器。

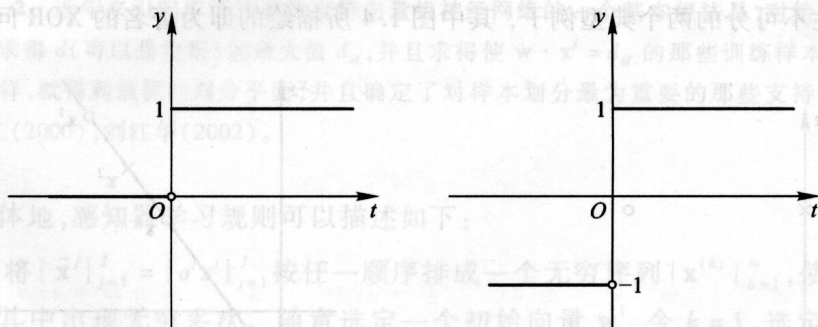


图 1.2 符号函数(函数值为  $\{0, 1\}$  或  $\{1, -1\}$ )

**注 1.1** 式(1.1.4)中主要的运算为便于并行处理的向量乘法。另外,除了在原点附近,符号函数(以及后面将要用到的符号函数的各种逼近)对于自变量的变化并不敏感,即容错



性好。事实上,各种神经网络用到的主要运算就是向量乘法,并且广泛采用符号函数及其各种逼近作为活化函数。典型的神经网络都可以用电路、光路等硬件来实现(参见 Murray (1997),戴葵(1998)),这时不论  $N$  多大,式(1.1.4)中的向量乘法所需的时间基本不变(参看图 1.1),使得便于并行处理的特点更加突出。并行、容错、可以硬件实现以及后面将要讨论的自我学习特性,是神经网络的几个基本优点,也是神经网络计算方法与传统计算方法的重要区别。当然,神经网络(尤其是只需进行一次的学习过程)也可以用计算机来模拟实现,这时并行的优点就不那么突出了。

以  $N=2$  为例。线性感知器的目标就是求法向量  $\mathbf{w}$  和阈值  $b$ ,使得与  $\mathbf{w}$  垂直的直线(一般是  $N-1$  维超平面)  $\mathbf{w} \cdot \mathbf{x} = b$  将样本模式  $\{\mathbf{x}^j\}_{j=1}^J$  分成  $\mathbf{w} \cdot \mathbf{x} > b$  和  $\mathbf{w} \cdot \mathbf{x} < b$  (即  $y^j = 1$  和  $y^j = -1$ ) 两类,分别位于  $\mathbf{w} \cdot \mathbf{x} = b$  的两侧(见图 1.3)。

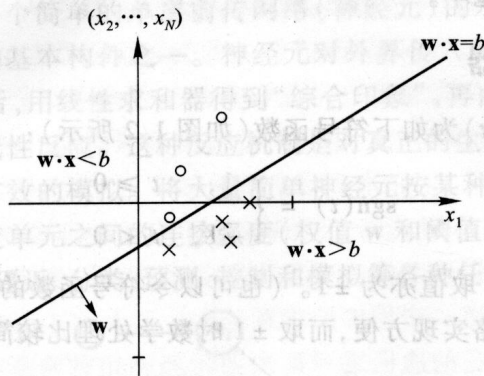


图 1.3 用线性感知器分类

容易证明,若  $\{\mathbf{x}^j\}_{j=1}^J$  线性无关,则一定是线性可分的。在图 1.4 和图 1.5 中给出线性不可分的两个典型例子,其中图 1.4 所描绘的即为著名的 XOR 问题。

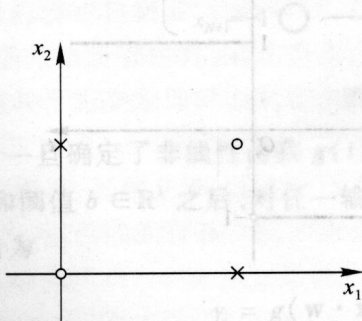


图 1.4 XOR 问题

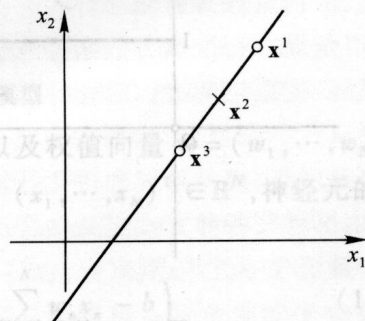


图 1.5  $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3$  线性不可分

权值向量  $\mathbf{w}$  和阈值  $b$  是通过学习得到的。下面讨论所谓感知器学习规则。为简便起见,在本章其余地方,我们总假设  $b=0$ 。这样做并不失一般性。这是