

391327

全国高技术重点图书·通信技术领域

遗传算法及其应用

陈国良 王煦法 编著
庄镇泉 王东生



人民邮电出版社

内 容 提 要

本书系统全面地介绍了遗传算法的基本原理、设计方法及其并行实现,以及它在组合优化、机器学习、图像处理、过程控制、进化神经网络、模糊模式识别和人工生命等方面的应用。

本书可作为高等院校计算机、无线电电子学、自动控制、生物医学工程等有关专业高年级学生或研究生的教材和参考书,也可供从事人工智能、信息处理研究和应用的科技人员学习参考。

全国高技术重点图书·通信技术领域

遗传算法及其应用

陈国良 王煦法 编著

庄镇泉 王东生

责任编辑 徐修存 王晓丹

*

人民邮电出版社出版发行

北京朝阳门内南竹杆胡同 111 号

北京顺义向阳胶印厂

新华书店总店北京发行所经销

*

开本:850×1168 1/32 1996年6月 第 1 版

印张:14.125 1996年6月 北京第1次印刷

字数:367千字 印数:1—2 000 册

ISBN7-115-05964-0/TN·1013

定价:26.00 元

《全国高技术重点图书》 出版指导委员会

主 任： 朱丽兰

副主任： 刘 杲 卢鸣谷

委 员： (以姓氏笔划为序)

王大中	王为珍	王守武	牛田佳	卢鸣谷
叶培大	刘 仁	刘 杲	朱丽兰	孙宝寅
师昌绪	任新民	杨牧之	杨嘉墀	陈芳允
陈能宽	张钰珍	张效详	罗见龙	周炳琨
欧阳莲	赵忠贤	顾孝诚	谈德颜	龚 刚
梁祥丰				

总干事： 罗见龙 梁祥丰

《全国高技术重点图书·通信技术领域》 编审委员会

主 任： 叶培大

委 员： 陈俊亮 徐大雄 姚 彦
程时昕 陈芳烈 李树岭

前 言

近代科学技术发展的显著特点之一是生命科学与工程科学的相互交叉、相互渗透和相互促进。遗传算法(GA——Genetic Algorithms)的蓬勃发展正体现了学科发展的这一特征和趋势。

遗传算法的内涵哲理乃是启迪于自然界生物从低级、简单,到高级、复杂,乃至人类这样一个漫长而绝妙的进化过程,借鉴于达尔文的物竞天演、优胜劣汰、适者生存的自然选择和自然遗传的机理,其本质是一种求解问题的高效并行全局搜索方法。它能在搜索过程中自动获取和积累有关搜索空间的知识,并自适应地控制搜索过程以求得最优解。

使用遗传算法求解科学研究和工程技术中各种组合搜索和优化计算问题的这一基本思想,早在 60 年代初期就由美国 Michigan 大学的 J. Holland 教授提出,其数学框架也于 60 年代中期形成,并且在 1975 年 J. Holland 教授编写的名著《自然与人工系统中的自适应》中予以系统的介绍。80 年代中期,遗传算法的研究蓬勃发展,吸引了大批的科学研究者和工程技术人员从事该领域的研究和开发应用工作。由于遗传算法的整体搜索策略和优化计算时不依赖于梯度信息,所以它的应用范围非常广泛,尤其适合于处理传统搜索方法难以解决的高度复杂的非线性问题。它在自适应控制、组合优化、模式识别、机器学习、规划策略、信息处理和人工生命等领域的应用中展示出优越性和魅力,从而也确定了它在 21 世纪的智能计算技术中的关键地位。

遗传算法的两大主要特点是群体搜索策略和群体中个体之间的信息相互交换,它实际上是模拟由个体组成的群体的整体学习过程,其中每个个体表示给定问题搜索空间中的一个解点。遗传算法从任一初始化的群体出发,通过随机选择(使群体中优秀的个体有更多的

机会传给下一代)、交叉(体现了自然界中群体内个体之间的信息交换)和变异(在群体中引入新的变种确保群体中信息的多样性)等遗传操作,使群体一代一代地进化到搜索空间中越来越好的区域,直至抵达最优解点。

遗传算法和其它的搜索方法相比,其优越性主要表现在以下几个方面:首先,遗传算法在搜索过程中不易陷入局部最优,即使在所定义的适应度函数非连续、不规则和伴有噪声的情况下也能以极大的概率找到全局最优解;其次,由于遗传算法固有的并行性,使得它非常适合于大规模并行分布处理;此外,遗传算法易于和别的技术(如神经网络、模糊推理、混沌行为和人工生命等)相结合,形成性能更优的问题求解方法。

尽管目前对遗传算法的研究还存在着一些有争议的问题,某些截然不同的甚至形成鲜明对照的学术观点和设计原则一时尚难统一,而且整个遗传算法的理论基础还显得较为薄弱,但是近 30 年来的不断研究和大量的实例应用充分表明,模拟自然进化的搜索过程往往可以产生非常简单、通用和鲁棒性很强的计算算法。我们认为,虽然在遗传算法研究的过程中,还存在着某些疑虑,但是人们不得不正视大量的研究成果为此研究领域所展示出的巨大潜力。我们相信,利用广泛的数学方法和强大的计算机模拟工具,必将使遗传算法的研究取得长足的进展。

本书根据我们近期有关遗传算法的研究和教学实践,并在综合大量有关文献资料的基础上,进行归纳、总结编著而成。全书共分 8 章,分别介绍了遗传算法的一般概念,遗传算法的基本原理和方法,遗传算法与组合优化,遗传算法与机器学习,遗传算法与并行处理,神经网络、模糊集理论和进化算法,遗传算法与人工生命以及遗传算法的应用实例。在附录中给出了 SGA(标准遗传算法)、TSP(货郎担问题)和 CLS(概念学习系统)3 个典型的遗传算法程序范例。

本书系集体编写,从开始拟定章节目录到分工撰写,大家相互讨论,并几经修改,最后由陈国良教授统一定稿。参与此书有关章节编

写工作的还有黄俊斌(2.1~2.4和2.10)、陈先富(第三章)、陈恩红(第四章)、张群(第五章)、曹先斌(6.1和6.3)、曹磊(7.3)和孙启斌(8.1和8.2)等同志,对于他们的辛勤劳动和贡献,作者表示诚挚的谢意。

由于作者水平有限,加之时间仓促,书中不妥之处在所难免,恳请读者批评指正。

本书的出版得到了国家攀登计划认识科学(神经网络)重大关键项目和国家自然科学基金的资助。

作者

1995年4月于中国科技大学

目 录

第一章 绪论

1.1 引言	1
1.2 生物进化	2
1.3 遗传算法	3
1.3.1 基础用语	3
1.3.2 标准遗传算法	5
1.4 遗传算法的特点	10
1.4.1 遗传算法和其它传统搜索方法的对比	10
1.4.2 遗传算法和若干搜索方法的亲近关系	13
1.4.3 遗传算法和自律分布系统的亲近关系	14
1.5 遗传算法的研究历史和现状	14
1.5.1 遗传算法的研究概况	14
1.5.2 遗传算法研究的新焦点	21
1.6 遗传算法今后研究的主要课题	23
参考文献	26

第二章 遗传算法的基本原理和方法

2.1 模式定理(schemate theorem)	28
2.1.1 模式	29
2.1.2 模式定理	30
2.2 积木块假设	36
2.3 骗问题	41
2.4 隐并行性	48
2.5 性能评估	49

2.6 编码	51
2.6.1 编码问题	51
2.6.2 编码(译码)评估规范和编码原理	52
2.6.3 编码技术	55
2.7 群体设定	67
2.7.1 初始群体设定	67
2.7.2 群体多样性	68
2.8 适应度函数	69
2.8.1 目标函数映射成适应度函数	70
2.8.2 适应度函数定标(scaling)	71
2.8.3 适应度函数的设计对遗传算法的影响	74
2.9 遗传操作	75
2.9.1 选择算子	76
2.9.2 交叉算子(crossover operator)	81
2.9.3 变异算子(mutation operator)	85
2.10 收敛性	88
2.10.1 未成熟收敛	88
2.10.2 有限马尔柯夫链	90
2.10.3 标准遗传算法的收敛性	92
参考文献	98

第三章 遗传算法与组合优化

3.1 基于遗传算法的组合优化方法	101
3.1.1 遗传算法的关键参数确定	102
3.1.2 几种流行的选择机制	103
3.1.3 适应度函数的定标	106
3.1.4 二倍体(diploidy)与显性(dominance)技术	107
3.1.5 物种形成(speciation)与小生境(niche)技术	111
3.2 函数优化(function optimization)	117

3.2.1	问题描述	117
3.2.2	编码与适应度函数	121
3.2.3	基本遗传算法(SGA)的搜索性能	122
3.2.4	基本遗传算法的若干变体形式的搜索性能	127
3.3	背包问题(knapsack problem)	130
3.3.1	问题描述	130
3.3.2	遗传编码	131
3.3.3	适应度函数	131
3.3.4	基于基本遗传算法求解背包问题	132
3.3.5	基本遗传算法的搜索能力	135
3.3.6	基于“与/或”交叉方法求解背包问题	136
3.4	货郎担问题	137
3.4.1	编码与适应度函数	137
3.4.2	交叉策略	138
3.4.3	变异技术	141
3.4.4	选择机制和群体构成	142
3.4.5	混合 GA 技术	143
3.4.6	基于遗传算法求解 TSP 的算法实现	143
3.5	混合搜索方法	147
3.5.1	概述	147
3.5.2	启发式搜索法简介	148
3.5.3	混合遗传算法(Hybrid GA)	151
3.5.4	实验与讨论	153
3.6	图的划分问题	158
3.6.1	问题描述	158
3.6.2	编码与适应度函数设计	158
3.6.3	遗传操作	159
3.6.4	实验结果	159
	参考文献	161

第四章 遗传算法与机器学习

4.1 概述	164
4.2 分类器系统	166
4.2.1 规则与消息	168
4.2.2 桶队算法	171
4.2.3 遗传算法	174
4.3 学习系统 LS-1	177
4.3.1 LS-1 与 CS-1 的区别	177
4.3.2 LS-1 的工作原理	178
4.4 基于遗传算法的概念学习系统	182
4.4.1 搜索空间的表示	184
4.4.2 遗传操作	186
4.4.3 执行过程	187
4.4.4 非标准操作	188
4.4.5 GABIL 系统的自适应性	189
4.5 小结	191
参考文献	192

第五章 遗传算法与并行处理

5.1 遗传算法固有的并行性及其并行化的困难	195
5.1.1 源于自然的并行性	195
5.1.2 遗传算法理论中的并行性	196
5.1.3 遗传算法在并行实现上的困难	197
5.2 遗传算法的并行化途径	198
5.2.1 主从式(master-slave)并行化方法	198
5.2.2 粗粒度模型	200
5.2.3 细粒度模型	201
5.3 粗粒度的孤岛模型	202

5.3.1	粗粒度模型的生物学依据	203
5.3.2	粗粒度模型的研究现状	204
5.3.3	孤岛模型在 MIMD 机器上的实现	205
5.3.4	扩展的分布式遗传算法	211
5.4	细粒度的邻域模型	214
5.4.1	细粒度模型的理论基础	215
5.4.2	细粒度模型的研究现状	216
5.4.3	MIMD 上的细粒度模型的实现	217
5.4.4	SIMD 上的细粒度模型的实现	220
5.5	粗粒度模型与细粒度模型的性能比较	222
5.6	实现并行遗传算法的一个例子	224
5.6.1	迁入式算法	224
5.6.2	迁出式算法	226
5.6.3	扩散式算法	227
5.7	LCS 的并行实现	229
5.7.1	执行系统	230
5.7.2	信用系统中的分配策略	232
5.7.3	遗传算法在 LCS 中的应用	232
5.7.4	LCS 的一个 MIMD 实现	233
5.7.5	LCS 在 CM 机器上的实现	235
	参考文献	237

第六章 神经网络、模糊集理论和进化算法

6.1	遗传算法与神经网络	241
6.1.1	神经网络的发展	241
6.1.2	神经网络连接权的进化	245
6.1.3	神经网络结构的进化	250
6.1.4	神经网络学习规则的进化	257
6.2	遗传算法与模糊集理论	258

6.2.1	基于遗传算法的模糊推理规则的优化	259
6.2.2	遗传算法在模糊模式识别中的应用	262
6.3	进化算法	271
6.3.1	引言	271
6.3.2	进化算法的总框架	274
6.3.3	遗传算法	275
6.3.4	进化规划	277
6.3.5	进化策略	279
6.3.6	交叉和变异的关系	282
6.3.7	小结	284
	参考文献	285

第七章 遗传算法与人工生命

7.1	人工生命的研究内容和方法	287
7.1.1	人工生命及其特征	287
7.1.2	人工生命研究的内容与方法	289
7.2	遗传算法与人工生命进化模型	292
7.3	L系统与形态生成模型	295
7.3.1	L系统与植物形态	295
7.3.2	植物的形态生成模型	296
7.3.3	讨论	303
7.4	博弈型人工生态系统	307
7.4.1	博弈与策略	308
7.4.2	博弈型生态系统	311
7.4.3	生态动力学与自组织化	314
7.5	人工生命与遗传信息处理	318
7.5.1	人类信息世界	319
7.5.2	监视遗传	321
7.5.3	遗传信息处理模型	322

7.5.4 基于遗传信息处理模型的人工生命合成	325
7.5.5 人工生命与人工智能	327
参考文献.....	332

第八章 遗传算法应用实例

8.1 遗传算法在图像恢复中的应用	334
8.1.1 引言	334
8.1.2 图像退化模型	335
8.1.3 遗传算法用于图像恢复	335
8.1.4 遗传算法与贝叶斯方法相结合的图像恢复	340
8.2 遗传算法在图像识别中的应用	345
8.2.1 引言	345
8.2.2 数学模型	347
8.2.3 目标函数形成	348
8.2.4 随机全局优化方法	351
8.2.5 实验结果	352
8.3 遗传算法在控制中的应用	354
8.3.1 操作序列的最优化	355
8.3.2 倒立摆控制	356
8.4 调度问题	358
8.4.1 车间作业调度问题	359
8.4.2 两种解法	359
8.4.3 实验	366
8.5 硬件进化	368
8.5.1 硬件进化的特点	369
8.5.2 硬件进化的学习方法	371
8.5.3 实例	375
参考文献.....	383
附录 A SGA 程序	385

附录 B	TSP 程序	397
附录 C	CLS 程序	416

第一章 绪 论

遗传算法是一类借鉴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 教授提出,其主要特点是群体搜索策略和群体中个体之间的信息交换,搜索不依赖于梯度信息。它尤其适用于处理传统搜索方法难于解决的复杂和非线性问题,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,是 21 世纪有关智能计算中的关键技术之一。本章先从生物进化讲起,接着示例介绍简单遗传算法的具体设计方法和步骤,然后归纳出遗传算法的一般特点,最后简要介绍遗传算法的研究历史和现状以及今后将研究的主要有关课题。

1.1 引 言

生命科学与工程科学的相互交叉、相互渗透和相互促进是近代科学技术发展的一个显著特点,而遗传算法的蓬勃发展正体现了科学发展的这一特征和趋势。

遗传算法(Genetic Algorithm——GA),是模拟达尔文的遗传选择和自然淘汰的生物进化过程的计算模型,它是由美国 Michigan 大学的 J. Holland 教授于 1975 年首先提出的^[1]。J. Holland 教授和他的研究小组围绕遗传算法进行研究的宗旨有两个,一是抽取和解释自然系统的自适应过程,二是设计具有自然系统机理的人工系统。毫无疑问,Holland 教授的研究无论对自然系统还是对人工系统都是十分有意义的。

众所周知,在人工智能领域中,有不少问题需要在复杂而庞大的搜索空间中寻找最优解或准最优解。像货郎担问题和规划问题等

组合优化问题就是典型的例子。在求解此类问题时,若不能利用问题的固有知识来缩小搜索空间则会产生搜索的组合爆炸。因此,研究能在搜索过程中自动获取和积累有关搜索空间的知识,并自适应地控制搜索过程,从而得到最优解或准最优解的通用搜索算法一直是令人瞩目的课题。遗传算法就是这种特别有效的算法。它的主要特点是简单、通用,鲁棒性强,适用于并行分布处理,应用范围广。尽管遗传算法本身在理论和应用方法上仍有许多待进一步研究的问题,但它在组合优化问题求解、自适应控制、规划设计、机器学习和人工生命等领域的应用中已展现了其特色和魅力。

1.2 生物进化

生命自从在地球上诞生以来,就开始了漫长的生物进化历程,低级、简单的生物类型逐渐发展为高级、复杂的生物类型。这一过程已经由古生物学、胚胎学和比较解剖学等方面的研究工作所证实。生物进化的原因自古至今有着各种不同的解释,其中被人们广泛接受的是达尔文的自然选择学说。

自然选择学说认为,生物要生存下去,就必须进行生存斗争。生存斗争包括种内斗争、种间斗争以及生物跟无机环境之间的斗争三个方面。在生存斗争中,具有有利变异(mutation)的个体容易存活下来,并且有更多的机会将有利变异传给后代;具有不利变异的个体就容易被淘汰,产生后代的机会也少得多。因此,凡是在生存斗争中获胜的个体都是对环境适应性比较强的。达尔文把这种在生存斗争中适者生存,不适者淘汰的过程叫做自然选择。达尔文的自然选择学说表明,遗传和变异是决定生物进化的内在因素。遗传是指父代与子代之间,在性状上存在的相似现象。变异是指父代与子代之间,以及子代的个体之间,在性状上或多或少地存在的差异现象。在生物体内,遗传和变异的关系十分密切。一个生物体的遗传性状往往会发生变异,而变异的性状有的可以遗传。遗传能使生物的性状不