

陆金桂 李谦
王浩 曹一家

编著

遗传算法原理及其工程应用

YICHUAN SUANFA YUANLI JIQI GONGCHENG YINGYONG



遗传算法原理及其工程应用

陆金桂 李 谦 编著
王 浩 曹一家

中国矿业大学出版社

遗传算法原理及其工程应用

陆金柱 李 谦 王 浩 曹一家 编著

责任编辑 姜志方

中国矿业大学出版社出版发行

新华书店经销 中国矿业大学印刷厂印刷

开本 850×1168 1/32 印张 5.875 字数 146 千字

1997 年 12 月第一版 1997 年 12 月第一次印刷

印数 1~1 000 册

ISBN 7 - 81040 - 600 - 0

TP · 40

定价:20.00 元

前 言

自 60 年代美国密执根大学 Holland 教授提出遗传算法 (Genetic Algorithm) 的概念体系以来,遗传算法得到了各领域学者的高度关注。目前国际学术界已定期举办了多届遗传算法方面专门的学术会议,并成立了专门的遗传算法学术组织。

与其它优化方法相比,遗传算法具有许多特点。遗传算法只需利用适应度函数值来进行优化计算,而不需要函数导数等其它辅助信息,因而能解决各种优化问题;遗传算法还具有全局寻优能力和隐含的并行性。因此,遗传算法在许多领域都得到了极为广泛的应用。

自 80 年代后期以来,我国一些学者也先后开展了遗传算法的研究,目前正逐渐形成研究热潮。为进一步推动遗传算法的研究及应用工作,我们编写了这本书,希望本书的出版有助于更多的不同学科的研究人员了解遗传算法、应用遗传算法。

与遗传算法在原理和实现策略上有许多相似之处的进化算法还有进化规划 (Evolutionary Programming) 和进化策略 (Evolution Strategy)。进化规划是由美国学者 Fogel 在 60 年代提出的,而进化策略是由德国柏林技术大学 Rechenberg 和 Schwefel 在 1963 年提出的。由于遗传算法研究和应用在进化算法中占有主导地位,因此我们只将遗传算法原理及其工程应用作为本书的内容。

本书包括遗传算法原理、遗传算法工程应用两部分。第一部分主要介绍遗传算法基本概念、遗传算法的基本原理等;第二部分介绍遗传算法在函数优化、结构优化、机械工程、电力系统等方面的

应用。这样安排希望读者能对遗传算法的原理与应用有全面的了解。由于作者学识有限、编著时间仓促,再加上遗传算法的研究与应用工作在不断深入,因此试图较全面地介绍遗传算法原理及其工程应用的本书肯定会有许多不足,希望广大读者能予以谅解并给予批评指正。

在本书的编写过程中,我们参考并引用了一些学者发表的论文中的研究成果;我们的编写工作还得到了南京化工大学机械工程系几位同志的支持和帮助。在此谨向他们表示衷心的感谢。

本书由南京化工大学陆金桂博士、广东电力局试验研究所李谦博士、广东五邑大学王浩博士、英国利物浦大学曹一家博士合作编写,陆金桂任主编。

本书的部分研究工作得到了国家自然科学基金的资助。

编著者

1997年8月

目 录

第一章 绪 论	(1)
第一节 概 述	(1)
第二节 遗传算法的基本概念	(4)
第三节 遗传算法简介	(11)
第四节 遗传算法的特点	(21)
第二章 遗传算法的基本原理	(23)
第一节 遗传算法的基本内容	(23)
第二节 适应度的比例变换	(29)
第三节 个体编码	(33)
第四节 遗传算法的其它繁殖算子	(35)
第五节 实 例	(45)
第三章 遗传算法若干理论问题	(51)
第一节 遗传算法的基因模式定理	(51)
第二节 遗传算法的隐含并行性分析	(56)
第三节 遗传算法的收敛性分析	(58)
第四章 遗传算法在函数优化中的应用	(66)
第一节 概 述	(66)
第二节 遗传算法在连续变量优化中的应用	(67)
第三节 遗传算法在混合离散变量优化中的应用	(72)
第四节 遗传算法在组合优化中的应用	(80)
第五节 并行遗传算法	(84)

第五章 遗传算法在结构优化设计中的应用	(87)
第一节 概 述	(87)
第二节 基于相对遗传力理论的遗传算法	(88)
第三节 基于遗传算法和神经网络的结构优化策略 ..	(95)
第四节 动态种源的遗传算法	(105)
第五节 遗传算法在结构多目标优化中的应用	(107)
第六章 遗传算法在机械工程中的应用	(114)
第一节 概 述	(114)
第二节 遗传算法在机械设计中的应用	(115)
第三节 遗传算法在机床切削参数优选中的应用	(120)
第四节 遗传算法在设备故障诊断中的应用	(123)
第七章 遗传算法在电工领域中的应用	(130)
第一节 概 述	(130)
第二节 遗传算法在电力系统中的应用	(132)
第三节 遗传算法在电机优化设计中的应用	(144)
第四节 遗传算法在电力变压器优化设计 中的应用	(148)
第八章 遗传算法在神经网络等领域的应用	(156)
第一节 概 述	(156)
第二节 遗传算法在多层神经网络中的应用	(157)
第三节 遗传算法在机器学习中的应用	(164)
第四节 遗传算法在系统辨识中的应用	(174)

第一章 绪 论

第一节 概 述

自然界的生物群体是现实世界中最复杂的系统。在自然界中,复杂的生物群体在自身繁衍的过程中能不断地进化,是因为生物群体的繁殖过程蕴含着自然优化的机制。早在 60 年代,人们就已关注自然界生物群体进化中蕴含着的内在的、朴素的优化思想,并开始将生物进化思想引入工程领域。

1962 年,美国密执根大学 (Michigan University) 的 John Holland 教授认识到生物群体的遗传、进化和人工系统自适应间的相似性,因而借鉴生物遗传的基本理论来研究人工自适应系统,并与 Bagley 等人一起提出了遗传算法 (Genetic Algorithm) 的概念。虽然在 50 年代初期,已有一些生物学家开始利用计算机技术来模拟生物的遗传和进化过程,但研究的主要目的是为了更深入地了解生物遗传进化的机理。因此, Holland 教授等人提出的遗传算法思想和概念具有创新性,他们的研究作为工程领域应用生物遗传进化思想奠定了重要的基础。70 年代初, Holland 教授提出了基因模式理论 (Schema Theorem)。该基因模式理论以二进制位串为基础,探讨了模拟生物染色体的人工染色体的表示、人工染色体的繁殖等,揭示了遗传算法的内在机制。基因模式理论为遗传算法奠定了坚实的理论基础。1975 年, Holland 教授出版了《自然和人工系统的自适应》(Adaption in Natural and Artificial System) 著作,

该著作系统地介绍了遗传算法的理论、原理和方法。《自然和人工系统的自适应》专著的出版进一步推动了遗传算法的研究与应用,遗传算法开始被用来解决各种优化问题。1989年,Goldberg出版了《搜索、优化和机器学习中的遗传算法》(Genetic Algorithm in Search, Optimization and Machine Learning)著作,该著作对遗传算法的理论和应用进行了全面的阐述,为遗传算法的发展奠定了重要的基础。

80年代中期,国际上已经举办了遗传算法方面专门的学术会议。1985年,第一届遗传算法的国际学术会议在美国召开。在遗传算法国际学术会议召开期间,国际遗传算法学会(International Society of Genetic Algorithms)宣告成立。自此,来自不同学科和工程应用领域的各国学者在遗传算法方面有了交流、探讨的国际论坛。此后国际遗传算法学术会议每两年定期召开一次。80年代后期、90年代初期,遗传算法在算法的复杂性、收敛性、算法的混合形式等理论方面都取得了重要的研究成果;在工程实践方面遗传算法也得到了最为广泛的应用。遗传算法已在机械工程的优化设计、切削加工、制造过程规划、设备故障诊断,自动控制的自适应控制、系统辨识、模糊控制、分类系统,人工智能的机器学习、专家系统、神经网络,结构工程的结构设计,电工学科的电机、变压器和电磁设备设计、电网规划、优化调度、潮流计算、电力系统控制,计算机学科的并行计算、图像处理、模式识别、文档处理,电子学科的超大规模集成电路设计,生物学科的分子生物学,计算数学的非线性规划、整数规划、组合规划,社会科学的人口学、交通系统规划等领域都得到了初步成功的应用。几乎所有工程领域的研究人员都曾尝试过利用遗传算法来解决各自专业领域的工程问题。1993年,人们以遗传算法等进化算法为基础提出了进化计算概念(Evolutionary Computation)。《进化计算》国际杂志也随之在美国问世。1994年,国际电工、电机和电子学会(IEEE)神经网络委员

会召开了第一届进化计算国际学术会议。进化计算国际学术会议为世界计算智能(Computational Intelligence)学术大会的重要组成部分。

目前,国际上掀起了以遗传算法为主要内容的进化算法的研究热潮。人们纷纷利用国际信息交互网络(Internet)进行遗传算法方面的信息交流。美国海军后勤研究中心早在1985年就建立了全球性的有关遗传算法的信息交流节点,并不定期地编辑出版遗传算法文摘(GA Digest)供交流遗传算法研究与应用成果。

进化算法包括遗传算法、进化规划(Evolutionary Programming)和进化策略(Evolution Strategy)等主要内容。进化规划是由美国学者Lawrence J. Fogel在60年代提出的。Fogel在对自然界生物进化现象进行研究的基础上,借鉴生物进化思想提出了进化规划。Fogel认为智能行为必须包括环境预测能力和在一定目标指导下对环境的合理响应能力。Fogel提出的进化规划注重进化过程中父代与子代表现行为的联系,而遗传算法则侧重于父代与子代的遗传操作。1966年,Fogel等出版了《基于模拟进化的人工智能》(Artificial Intelligence Through Simulated Evolution)著作,系统阐述了进化规划思想和内容。由于一些学者对利用进化规划研究人工智能持怀疑态度,因此进化规划思想和方法未能引起学术界的关注。直到1992年,进化规划的第一届学术年会才在美国召开。

进化策略是由德国柏林技术大学Ingo Rechenberg和H. P. Schwefel在1963年提出的一种进化算法。由于当时一些基于梯度策略的优化方法难以解决气流中物体外形最优问题,Rechenberg提出了按照自然突变和自然选择的生物进化策略来对物体的外形参数进行随机调整的思想并进行了尝试,进化策略方法由此诞生。当时,人们对这种随机策略进化方法无法接受,但Rechenberg等人仍坚持研究。1970年,Rechenberg完成了有关进化策略研究的

博士论文。1974年, Schwefel 对有关进化策略研究的成果进行了归纳整理。

由于遗传算法、进化规划和进化策略是在 60 年代由不同领域的研究人员独立提出的, 因而这些进化算法的提出者在相当长的时期里相互之间没有直接进行交流。直到 1990 年, 遗传算法、进化规划和进化策略的提出者才得以进行深入交流。遗传算法、进化规划和进化策略在理论上都是借鉴生物进化思想, 在内容上亦有许多相似之处。目前, 在进化算法中, 遗传算法方面的研究和应用占有主导地位, 因此本书只涉及遗传算法原理及其工程应用。

第二节 遗传算法的基本概念

遗传算法是基于生物遗传进化思想的一种优化方法, 因此遗传算法与数学规划类优化方法在原理、实现手段等方面有着明显的差别。本节给出遗传算法中常用的一些名词和基本概念。

一、个体(Individual String)

定义: 个体是遗传算法中用来模拟生物染色体的一定数目的二进制位串, 该二进制位串用来表示优化问题的设计点。

由于个体是用来模拟生物群体进化的最基本的单位, 因而个体是遗传算法中的基本概念。在遗传算法中, 个体又称为人工染色体(Artificial Chromosome)。

例 1 用六位二进制串来表示的个体。

显然, 六位的二进制串最多可用来表示 2^6 个个体。下面是六位二进制串表示的 64 个个体。

个体 1 1 1 1 1 1 1

个体 2 1 1 1 1 1 0

个体 3 1 1 1 1 0 1

个体 4 1 1 1 0 1 1

个体 5	1 1 0 1 1 1
...	1 0 1 1 1 1
...	0 1 1 1 1 1
...	1 1 1 1 0 0
...	
个体 61	0 0 0 0 1 1
个体 62	0 0 0 0 1 0
个体 63	0 0 0 0 0 1
个体 64	0 0 0 0 0 0

遗传算法中个体与优化设计模型中的设计变量集存在着对应关系,一个个体可以用来描述优化设计问题中的一个解。

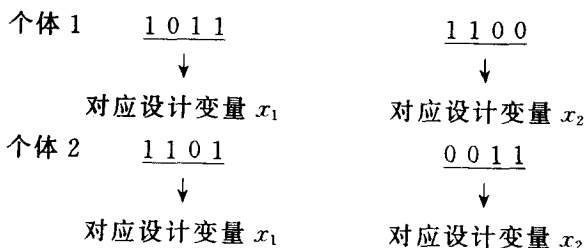
例 2 用个体来表示优化问题的设计点。

优化问题为:

$$\min F(x)$$

$$x = [x_1, x_2]$$

假设每个设计变量都以 4 位二进制码来表示,则两个设计变量需用 8 位二进制码来表示,因此与设计变量解对应的个体将是一个 8 位二进制位串。



在例 2 中,个体对应着优化问题的两个设计变量,个体 1 可表示设计解 $[x_1^{(1)}, x_2^{(1)}]$,个体 2 表示设计解 $[x_1^{(2)}, x_2^{(2)}]$ 。

二、群体(Population)

定义:群体是由一定数量的个体组成的集合。

生物进化过程是一代生物群体繁殖产生另一代生物群体的过程。因此,基于生物进化思想的遗传算法从本质上看是模拟生物的一代群体向另一代群体变化的历程,因而群体是遗传算法中体现算法特点的一个重要概念。

遗传算法中的群体与优化模型中若干个设计解组成的集合相对应,因而群体与优化模型中设计解集合存在着一一对应的关系。

群体规模是与群体有关的一个概念。下面介绍群体规模的定义。

定义:群体规模是指群体中包含的个体的数量。

在例2中,如果将个体1和个体2作为一个群体,那么群体规模数为2。

初始化群体是一类特殊的群体,下面给出初始化群体的定义。

定义:初始化群体是指遗传算法中的第一代群体。

初始化群体反映着优化问题的初解,因而初始化群体与优化模型的初始解相对应。初始化群体又称为原始群体;而初始化群体的过程,被称为群体的初始化。

三、基因模式(Schemata)

定义:基因模式是指二进制位串表示的个体中,某一或某些位置上具有相似性的个体组成的集合。

例3 基因模式。

下面给出如下六位二进制串中包含的6个基因模式。

$$H_1 = * 1 * * * *$$

$$H_2 = * 1 0 * * *$$

$$H_3 = * 1 0 0 * *$$

$$H_4 = 1 0 0 1 * *$$

$$H_5 = * 1 0 1 1 0$$

$$H_6 = 1 0 0 1 1 0$$

在上述基因模式中,符号*代表在该位置处二进制数值可以

是 0 或 1。

基因模式 H_1 代表在位置 2 二进制值为 1 的个体集合, 基因模式 H_2 代表在位置 2、3 二进制值为 1、0 的个体集合, 基因模式 H_3 代表在位置 2、3、4 二进制值为 1、0、0 的个体集合, 基因模式 H_4 代表在位置 1、2、3、4 二进制值为 1、0、0、1 的个体集合, 基因模式 H_5 代表在位置 2、3、4、5、6 二进制值为 1、0、1、1、0 的个体集合, 基因模式 H_6 代表在位置 1、2、3、4、5、6 二进制值为 1、0、0、1、1、0 的个体集合。

在 6 位二进制位串表示的个体中, 1 个个体包含例 3 中的基因模式 H_6 , 3 个个体包含例 3 中的基因模式 H_5 , 9 个个体包含例 3 中的基因模式 H_4 , 27 个个体包含例 3 中的基因模式 H_3 , 81 个个体包含例 3 中的基因模式 H_2 , 243 个个体包含例 3 中的基因模式 H_1 。

例 4 两个个体:

个体 1 1 1 0 0 0 0

个体 2 0 1 0 0 1 1

在例 4 中, 由于个体 1 在第二位上与基因模式 H_1 相同, 因而个体 1 包含了基因模式 H_1 ; 个体 1 在第 2、3 位上与基因模式 H_2 相同, 因而个体 1 也包含了基因模式 H_2 ; 个体 1 在第 2、3、4 位上与基因模式 H_3 相同, 因而个体 1 也包含了基因模式 H_3 ; 个体 1 在相应的位置上与基因模式 H_4 、 H_5 、 H_6 不相同, 因而个体 1 不包含基因模式 H_4 、 H_5 、 H_6 。同样, 个体 2 在第二位上与基因模式 H_1 相同, 因而个体 2 包含了基因模式 H_1 ; 个体 2 在第 2、3 位上与基因模式 H_2 相同, 因而个体 2 也包含了基因模式 H_2 ; 个体 2 在第 2、3、4 位上与基因模式 H_3 相同, 因而个体 2 也包含了基因模式 H_3 ; 个体 2 在相应的位置上与基因模式 H_4 、 H_5 、 H_6 不相同, 因而个体 1 不包含基因模式 H_4 、 H_5 、 H_6 。

从例 4 可以看出, 个体中相似位越多, 个体所包含的基因模式

数也越多。对于长度为 l 的个体,在某一位置上存在着 0、1、* (* 代表数值可以是 0 或 1) 三种可能的模式,因而长度为 l 的个体共有 3^l 个基因模式。

基因模式是遗传算法中最重要的概念,基因模式影响着群体进化的效率。在遗传算法中,基因模式常用来进行算法收敛性分析等。

四、模式阶次 (Schema Order)

定义:模式阶次是指基因模式中包含相似位置的数目。

模式阶次常用 $o(H)$ 来表示。

在例 3 中,基因模式 H_1 的模式阶次为 1,基因模式 H_2 的模式阶次为 2,基因模式 H_3 的模式阶次为 3,基因模式 H_4 的模式阶次为 4,基因模式 H_5 的模式阶次为 5,基因模式 H_6 的模式阶次为 6。

五、模式定义长度 (Defining Length of Schema)

定义:模式定义长度是基因模式中相似位间相距的最大距离。

模式定义长度的单位是位串个数,一般常用符号 $\delta(H)$ 来表示。

在例 3 中,基因模式 H_1 的模式定义长度为 0,基因模式 H_2 的模式定义长度为 1,基因模式 H_3 的模式定义长度为 2,基因模式 H_4 的模式定义长度为 3,基因模式 H_5 的模式定义长度为 4,基因模式 H_6 的模式定义长度为 5。

在遗传算法中,基因模式可用来分析个体间的相似性,而基因模式的阶次、基于模式的定义长度又是用来定量描述基因模式相似性的主要指标。

六、适应度 (Fitness)

定义:适应度是以数值方式来描述个体优劣程度的指标。

在优化模型中,目标函数是对设计解优劣进行比较的指标,在遗传算法中适应度是用来评判个体优劣的重要指标,因此遗传算

法的适应度在物理意义上对应着优化模型中的目标函数。

七、平均适应度(Average Fitness)

定义:平均适应度是若干个个体的适应度值的算术平均值。

平均适应度常用来评判群体中个体的优劣程度。

八、繁殖(Generating Next Population)

定义:繁殖是由一代群体繁衍产生另一代群体的方式总称。

在遗传算法中,由父代个体繁殖产生下代个体是最重要的内容。目前繁殖的方式主要包括选择、杂交、突变等繁殖算子。

九、选择(Reproduction)

定义:选择算子是指在上一代群体中按照某些指标挑选参与繁殖下一代群体的一定数量的个体。

十、杂交(Crossover)

定义:杂交算子是指对优选后的父代个体进行基因模式的重组而产生后代个体的繁殖机制。

在个体繁殖过程中,杂交能引起基因模式的重组,从而有可能产生包含优良性能的基因模式的个体。杂交方式一般有一点杂交、两点杂交、均匀杂交、基于顺序杂交等,其中一点杂交算子实现起来最简单。

十一、一点杂交(One Point Crossover)

定义:一点杂交是指在代表个体的二进制位串中选择一截断位,将截断位前后的二进制位串互相交换产生后代个体的方式。

利用一点杂交算子来产生后代个体如图 1-1 所示。

十二、两点杂交(Two Point Crossover)

定义:两点杂交是指在代表个体的二进制位串中选择两个截断位,将两个截断位间的二进制位串互相交换产生后代个体的方式。

利用两点杂交算子来产生后代个体如图 1-2 所示。

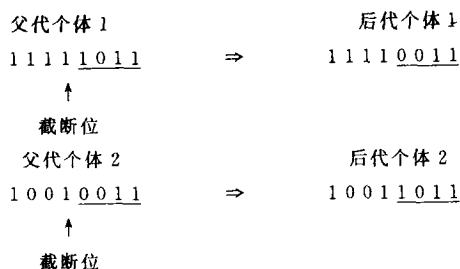


图 1-1 一点杂交

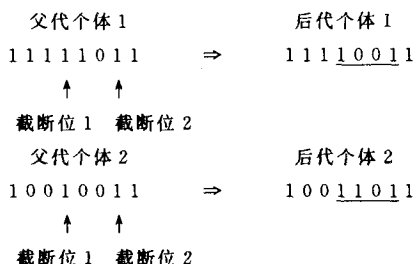


图 1-2 两点杂交

十三、突变(Mutation)

定义:突变算子是指模拟生物在自然的遗传进化环境中由于各种偶然因素引起的基因模式突然改变的个体繁殖方式。

在突变算子中,常以一定的概率在群体中选取个体,随机选择个体的二进制位串中的某一位进行由概率控制的变换(即由 1 改为 0 或由 0 改为 1)从而产生新的个体。

利用突变算子来产生后代个体如图 1-3 所示。在遗传算法中,采用突变算子增加了群体中基因模式的多样性,从而增加了群体进化过程中自然选择的作用,并能避免群体早熟性收敛现象的产生,从而避免群体进化过程过早地陷入局部最优区域。突变产生的优良个体在群体进化过程中将被保留,突变算子产生的非优良个