

机械优化设计

遗传算法

 陈伦军 等编著

 机械工业出版社
CHINA MACHINE PRESS



机械优化设计遗传算法

陈伦军 罗延科 陈海虹 张莉 编著



机械工业出版社

机械优化设计遗传算法是基于自然进化论和计算机科学相互渗透, 生命科学与机械工程相交所产生的一种新计算方法。

本书介绍遗传算法的基本原理和数学知识, 根据遗传算法的概念和方法来解决机械优化设计问题, 可以避免传统的优化方法出现的局部最优现象, 从而达到整体最优解。为了结合实际应用, 书中给出了机械优化设计遗传算法实例供读者参考。

本书可供机械工程技术人員、机械专业高年级大学生及研究生学习使用。

图书在版编目 (CIP) 数据

机械优化设计遗传算法/陈伦军等编著. —北京: 机械工业出版社, 2005.3

ISBN 7-111-15995-0

I. 机... II. 陈... III. 遗传-算法-应用-最佳化-机械设计 IV. TH122

中国版本图书馆 CIP 数据核字 (2005) 第 000806 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

责任编辑: 曾 红 版式设计: 霍永明 责任校对: 李秋荣

封面设计: 陈 沛 责任印制: 杨 曦

北京机工印刷厂印刷·新华书店北京发行所发行

2005 年 3 月第 1 版第 1 次印刷

1000mm × 1400mm B5·6.375 印张·246 千字

0 001—3 000 册

定价: 25.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

本社购书热线电话 (010) 68326294

封面无防伪标均为盗版

前 言

在科学技术迅速发展的时代，多学科相互交叉、相互渗透和相互促进已成为当今科学发展的显著特征。通过科学家们的不懈努力，将生物自然选择和遗传进化与计算技术相结合，创立了生命科学和工程学科相互渗透的优化算法——遗传算法。

遗传算法根据生物进化理论，应用群体的搜索技术，通过选择、交叉及变异等遗传操作，使群体一代一代进化，逐步接近搜索空间中越来越好的区域，直到取得最优解。由于遗传算法对非线性不连续多峰函数和无解析表达式的优化问题有很强的通用性，对目标函数具有全局优化性和稳健性，它的搜索是多途径进行，所以其运算并行性好、直观简单、可操作性强。

鉴于比传统优化算法更为突出的优越性，所以广大学者对遗传算法的研究仍然是方兴未艾。当今遗传算法视为智能的新颖优化算法之一，在优化技术领域有着极其广阔的应用前景。

本书采用循序渐进的方法，在第1章和第2章中以简单的例子，描述遗传算法的过程，使读者初步了解遗传算法的基本理论；第3章简述遗传基本算法，从可操作性观点出发详述了编码、适应度函数及遗传算子等各种方法和参数的选择；第4章和第5章介绍组合遗传算法及并行算法的实用技术；第6章基于神经网络的遗传算法和第7章模糊优化的遗传算法编入了扩展遗传算法智能计算应用的内容；第8章结合机械优化实际例子汇集了遗传算法在机械优化中的应用。

本书紧密结合机械设计应用实例，通俗易懂。

书中第1章、第2章、第8章由陈伦军编写；第6章、第7章由罗延科编写；第4章、第5章由陈海虹编写；第3章由张莉编写。

本书得到原贵州工业大学领导的关心和经费上的支持，在此表示衷心的感谢。

由于我们水平有限，书中不足之处，恳请读者批评指正。

目 录

前言

第 1 章 绪论	1
1.1 生物进化的基本知识	1
1.2 遗传算法简述	3
1.3 遗传算法的特点	7
1.4 遗传算法的发展概况	8
1.5 遗传算法在机械工程中的应用前景	10
第 2 章 遗传算法的数学基础理论	12
2.1 模式定理	12
2.1.1 模式的定义	12
2.1.2 模式定理的引证	14
2.2 隐含并行性	19
2.3 基因块假设	20
2.4 欺骗性问题	22
2.5 遗传算法的收敛性	27
2.5.1 基本定义	27
2.5.2 守恒交叉算子	29
2.5.3 完全变异算子	30
2.5.4 遗传算法的马尔可夫链分析	31
第 3 章 遗传算法的实现方法	34
3.1 编码方法	34
3.1.1 编码原则	34
3.1.2 二进制编码方法	35
3.1.3 格雷码编码方法	37
3.1.4 实数编码	38
3.1.5 多参数映射编码	39
3.1.6 可变长编码方法	39
3.2 群体设定	40
3.3 适应度函数	41

3.3.1	目标函数映射成适应度函数	41
3.3.2	适应度标度变换	42
3.3.3	适应度函数的设计对遗传算法的影响	44
3.4	遗传算子	45
3.4.1	选择算子	45
3.4.2	交叉算子	47
3.4.3	变异算子	49
3.5	遗传算法运行参数的选择	52
3.6	约束条件的处理问题	54
3.6.1	区域约束的处理方法	55
3.6.2	处理一般约束的方法	55
3.6.3	罚函数方法	56
第4章	组合优化的遗传算法	58
4.1	混合遗传算法	58
4.1.1	概述	58
4.1.2	混合遗传算法的基本构成原则	58
4.2	模拟退火算法	59
4.2.1	概述	59
4.2.2	模拟退火算法 (SA)	60
4.2.3	冷却进度表	62
4.2.4	模拟退火算法的改进和变异	69
4.3	遗传模拟退火算法	73
4.4	二倍体与显性操作算法	74
4.4.1	概述	74
4.4.2	遗传算法中的二倍体结构与显性技术	76
4.5	小生境技术	77
4.5.1	概述	77
4.5.2	遗传算法中小生境技术的实现方法	78
4.6	背包问题 (Knapsack problem)	81
4.6.1	问题描述	81
4.6.2	遗传编码	82
4.6.3	适应度函数	82
4.6.4	混合遗传算法在求解背包问题中的应用	82
4.7	装箱问题	84
4.7.1	问题描述	84
4.7.2	装箱问题的启发式算法	85
4.7.3	遗传编码	85

4.7.4 遗传算子	87
4.7.5 适应度函数	88
4.7.6 初始群体	89
4.8 遗传模拟退火算法解平面连杆变幅机构轨迹优化问题	89
4.8.1 求解连杆曲线非线性方程组的同伦算法	89
4.8.2 平面连杆变幅机构轨迹优化模型	91
4.8.3 遗传退火算法	93
4.8.4 实际工程算例	95
第5章 并行遗传算法	97
5.1 概述	97
5.1.1 遗传算法源于自然的并行性	97
5.1.2 遗传算法理论中的并行性	97
5.1.3 遗传算法并行化的目的和存在的问题	98
5.2 遗传算法的并行性分析和实现方法	99
5.2.1 遗传算法的并行性分析	99
5.2.2 并行遗传算法的实现方法	100
5.2.3 并行遗传算法的硬件支持环境及性能评价	101
5.3 单群体并行遗传算法	102
5.3.1 主从式并行方法 (Master-slave)	102
5.3.2 异步同时式并行遗传算法	103
5.3.3 单群体并行遗传算法的效能特点	104
5.3.4 单群体并行遗传算法的研究现状	104
5.4 多群体并行遗传算法	105
5.4.1 多群体并行方法的基本原理	105
5.4.2 多群体并行遗传算法模型描述	107
5.4.3 踏脚石群体模型 (Stepping-stone Model)	109
5.4.4 粗粒度孤岛模型	111
5.4.5 细粒度邻域模型	113
5.5 并行遗传算法的层次模型	115
5.6 伪并行遗传算法简介	117
5.6.1 伪并行遗传算法的基本思想	117
5.6.2 伪并行遗传算法描述	117
5.7 并行遗传算法在天线结构优化设计中的应用	118
5.7.1 目标函数	118
5.7.2 约束条件	119
5.7.3 多群体并行遗传算法	119
5.7.4 算法实例	120

第 6 章 基于神经网络的遗传算法	124
6.1 基本神经元模型	124
6.2 BP 网络	126
6.3 神经网络的遗传算法	129
6.3.1 神经网络连接权的进化方法	129
6.3.2 神经网络结构的进化方法	132
6.3.3 神经网络学习规则的进化	137
6.4 基于神经网络和遗传算法的斜齿圆柱齿轮传动优化设计	138
6.4.1 数学模型	139
6.4.2 斜齿圆柱齿轮传动优化模型的遗传算法实现	140
6.4.3 神经网络模型	141
6.4.4 优化结果	143
第 7 章 模糊优化的遗传算法	144
7.1 引言	144
7.2 模糊优化的概述	145
7.2.1 对称模糊优化的基本思想	145
7.2.2 非对称模糊优化的最优水平截集法	146
7.3 多目标遗传算法的模糊优化	148
7.3.1 多目标优化问题	148
7.3.2 遗传多目标优化	148
7.4 行星齿轮多目标模糊优化的遗传算法	151
7.4.1 建立多目标模糊优化数学模型	151
7.4.2 基于遗传算法的多目标优化模型求解	153
7.4.3 工程算例	154
第 8 章 机械优化设计的遗传算法	155
8.1 遗传算法在机构优化设计中的应用	155
8.1.1 平面六杆机构优化设计的遗传算法	155
8.1.2 平面四杆机构优化设计的遗传算法	159
8.1.3 带液压缸平衡机构优化设计的遗传算法	162
8.2 遗传算法在螺栓弹簧优化设计中的应用	166
8.2.1 紧固螺栓优化设计的遗传算法	166
8.2.2 弹簧优化设计的遗传算法	168
8.3 遗传算法在机械传动方案中的应用	170
8.4 遗传算法在齿轮传动优化设计中的应用	173
8.4.1 锥齿轮传动模糊优化的遗传算法	173
8.4.2 少齿差行星齿轮传动参数优化设计的遗传算法	176

8.4.3 圆柱齿轮减速机优化设计的遗传算法	180
8.5 新型节能电磁换向阀优化设计的遗传算法	183
8.6 机械零件可靠性设计的遗传算法	185
8.6.1 自适应遗传算法用于机械零件的可靠性设计	185
8.6.2 牙嵌式离合器可靠性优化设计的遗传算法	189
参考文献	194

第 1 章

绪 论

遗传算法的术语来源于自然遗传学。1975 年由美国 J. Holland 教授提出的遗传算法 (Genetic Algorithm, 简称 GA) 是基于自然选择原理、自然遗传机制和自适应搜索 (寻优) 的算法。GA 启迪于生物学的新达尔文主义 (达尔文的进化论、魏茨曼的物种选择学说和孟德尔的基因学说), 模仿物竞天演、优胜劣汰、适者生存的生物遗传和进化的规律性。

1.1 生物进化的基本知识

生物界的遗传是指父代与子代之间继承特性或性状。有遗传才能保证各种不同生物的稳定延续, 才有种的属性, 即种瓜得瓜、种豆得豆。生物在环境中生存, 通过生存斗争, 逐渐适应于生活环境, 使其有利变异的个体容易存活下来, 并且有更多的机会有利变异传给后代; 具有不利变异的个体容易被淘汰, 产生后代的机会少得多。有变异才有生物的进化发展, 这就是达尔文的自然选择学说。

生物体的生命活动主要是以细胞为基本单位进行的。根据生物学家的研究, 生物体的细胞由化学元素的各种化合物所组成, 具有不同形式的水、糖、无机盐、脂质、蛋白质及核酸等, 其中核酸是遗传信息的载体, 存在每一个细胞之中, 它是一切生物的遗传物质, 对于生物体的遗传变异和蛋白质的生物合成有极其重要的作用。

核酸——基本组成单位是核苷酸, 1 个核苷酸是由一分子含氮的碱基, 一分子五碳糖和一分子磷酸组成。根据核酸中所含五碳糖的种类不同, 可以将核酸分为脱氧核糖核酸 (简称 DNA) 和核糖核酸 (在低等生物中为 RNA), 许多核苷酸通过磷酸二酯键前后相接成链状。两条主干链在外, 碱基朝内, 围绕主轴顺时针方向缠绕成双链螺旋结构。

脱氧核糖核酸 (DNA) ——1 个 DNA 分子由两条平行的多核苷酸组成, 链中的碱基是双环的嘌呤 (A.G), 通过对应的 2 个或 3 个氢链连接单环的嘧啶 (T.C), 以 A = T 或 G = C 配对, 且按照互补原则彼此成为互补链。双链 DNA 分子是个生物大分子, 在 A、T、G、C 四种碱基的排列顺序中蕴藏着生物界无穷无尽的遗传信息, DNA 中遗传信息在一条长链上按一定模式排列, 亦即进行遗传编码。

染色质和染色体——在细胞核中的 DNA 和若干组蛋白质分子结合成染色质, 由染色质的小纤维在组蛋白 H_1 的作用下进一步螺旋, 构成 30nm 的染色质纤维, 称为螺旋管。细胞在分裂时, 染色质纤维将进一步折叠, 构建成染色体, 分配给子细胞。构建染色体时, 螺旋管折叠成襻状, 若干个大小不同的襻状按一定方式组成微带, 若干个微带排成染色体。

基因和基因组——1909 年丹麦植物和遗传学家 Johannsen 把遗传因子改称为“基因”。1926 年 Morgan 通过果蝇的研究, 阐明了基因是直线排列在染色体上的遗传颗粒, 它是具有生物遗传信息的结构单位, 又是控制生物体性状的功能单位, 通过生物学和遗传学家进一步研究揭示了基因的分子基础是核酸, 其中蕴藏着指导蛋白质生物合成的遗传信息。一个或多个基因决定了组成蛋白质生物合成的遗传信息, 以及组成蛋白质 20 种氨基酸的组成比例及排列顺序。遗传基因在染色体中所占据的位置称为基因座, 同一基因座可能有的全体基因称为等位基因。某种生物所特有的基因及其构成形式称为该生物的基因型, 而该生物在环境中呈现的相应性状称为该生物的表现型。一个细胞核中所带有的全部遗传信息称为基因组, 通常是指核内单倍数染色体包含的所有遗传信息, 它含有一种生物完整的一套基因。

遗传物质 DNA 的复制发生在体细胞分裂周期, 复制时一条染色体上某起始处的 DNA 双螺旋解开成倒“Y”形, 复制同时向两侧相反方向进行。对于有性生殖生物, 两个同源染色体通过交叉复制而重组, 亦即在两个染色体的某一相同位置处 DNA 被切断, 从而形成两个新的染色体, 这些新染色体表现出新的性状。

细胞在完成了 DNA 复制后, 通过细胞分裂将复制的具有遗传物质的染色体平均地分配给两个子细胞, 因此每个子细胞都含有与原来母细胞相同的染色体数, 结果子细胞都保持与母细胞相同的全套遗传物质, 从而保证了机体所有细胞间染色体数目恒定和遗传性状的稳定。

生物体通过基因的复制和交叉使其性状的遗传得到选择和控制。在重组和变异后的新基因依据与环境的适应程度决定其增殖能力, 有利于生存环境的基因增多, 而不利于生存环境的基因减少。通过这种自然选择所形成的物种, 推动了生物的进化和发展。

1.2 遗传算法简述

以模拟自然界生物遗传和进化过程形式的遗传算法,是依据生物进化以集团的形式即群体共同进化的。组成群体的单个生物称为个体基本特性的遗传继承,由个体性质的染色体所决定。具有遗传基因染色体的个体对环境有不同的适应性,通过基因杂交和变异产生适应强者,在遗传进化中“适者生存”的自然选择的统制作用下,使得更适应环境的个体被保留下来。遗传算法正是基于自然选择和自然遗传这种生物进化机制的搜索算法,将优化问题开创成新的全局优化搜索算法。

对于求函数最大值的优化问题:

$$\begin{aligned} & \max f(X) \\ \text{s.t. } & g(X) \leq 0 \\ & X \in R^n \end{aligned} \quad (1-1)$$

式中, $X = [x_1, x_2, \dots, x_n]^T$ 为决策变量, R^n 为 n 维欧氏空间, $g(X) \leq 0$ 为约束条件, $f(X)$ 为目标函数。

遗传算法中,求函数的极大值可以用适应度代替目标函数。适应度是生物个体对环境适应程度,用来评估生物群体中每个个体适应环境所表现出的不同生命力,从而决定其遗传机会的大小。

优化问题中的 n 维矢量 $X = [x_1, x_2, \dots, x_n]^n$, 遗传算法是用 n 个记号 $X_i (i = 1, 2, \dots, n)$ 所组成的符号串 X 来表示:

$$X = X_1 X_2 \dots X_n$$

这样, X 就可以看成由 n 个遗传基因 X_i 所组成的染色体。每一个 X_i 就是一个遗传基因,所有可能的取值即为等位基因,这里等位基因是一组整数,也可以是某一范围内的实数值或者是纯粹的一个记号。最简单的等位基因是由 0 和 1 这两个整数组成的,相应的染色体就可表示为一个二进制符号串。染色体 X 也称为个体 X ,对于每一个个体 X ,要按一定规则确定出其适应度。个体适应度与其相应的个体表现型 X 的目标函数相关, X 越接近于目标函数的最优点,其适应度越大,反之亦然。

遗传算法是以决策变量 X 组成优化问题的解空间。对问题最优解的搜索是通过对染色体 X 的搜索过程来进行的,从而由所有的染色体 X 就组成了问题的搜索空间。与生物一代一代的自然进化过程相类似,遗传算法的运算过程也是一个反复迭代过程。这个群体不断地经过遗传和进化操作,并且每次都按照优胜劣汰的规则,将适应度较高的个体更多地遗传到下一代,这样最终在群体中将会得到一个优良个体 X ,它所对应的表现型 X 将达到或接近于问题的最优解

X^* 。标准的遗传算法的主要步骤可描述如下 (图 1-1):

- 1) 随机产生一组初始个体构成初始群体。
- 2) 计算适应度, 判断算法收敛准则是否满足, 若满足就输出搜索结果, 否则, 进行下一步骤。
- 3) 选择运算, 按优胜劣汰原则执行复制操作。
- 4) 交叉运算, 按一定方式进行交叉操作。
- 5) 变异运算, 按一定规则执行变异操作。
- 6) 返回步骤 2)

这里, 仅以一个简单的二维函数极值求解为例, 说明遗传算法的概念及处理过程。

例 求下述函数的最大值。

$$\max f(X) = x_1^2 + x_2^2 + 1$$

$$X \in [x_1, x_2]^T$$

$$x_1 \in \{0, 1, 2, \dots, 7\}$$

s. t.

$$x_2 \in \{0, 1, 2, \dots, 7\}$$

下面我们对计算的主要过程简述如下:

(1) 个体编码

遗传算法不能直接处理解空间的数据, 必须通过编码将它们表示成遗传空间的基因型个体符号串。根据本例的变量 x_1 、 x_2 编码为二进制整数符号串, 则 x_1 、 x_2 取 0~7 之间的整数, 可取 3 位无符号二进制整数来表示, 两个变量 x_1 、 x_2 组成一组 6 位二进制整数形成个体的基因型, 表示一个可行解。例如, 基因型 $X = 011010$ 所对应的表现型是: $X = [3, 2]^T$ 。

(2) 初始群体的生成

因遗传算法是对群体进行操作, 所以需要准备一些起始搜索的初始群体。为了简化说明, 本例中群体的规模大小取为 4, 也就是群体由 4 个个体组成。每个个体是通过随机方法产生, 初始群体作为进化第一代如表 1-1 第②栏所示。

(3) 适应度计算

遗传算法用适应度的大小来评估个体 (或解) 的优劣, 从而决定其遗传机会的多少。因本例的目标函数总取非负值, 并且以求函数最大值为优化目标, 所以直接利用目标函数值作为个体适应度。为了计算目标函数, 需先对个体基

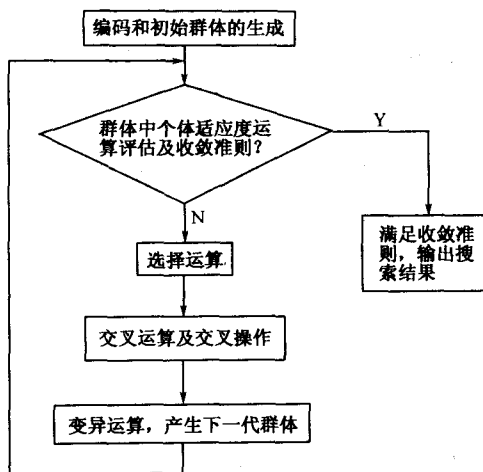


图 1-1 遗传算法的运算基本流程

因型 X 中 x_1 、 x_2 进行解码（如表 1-1 中第③、④所示）。在表 1-1 中⑤栏所示为各个个体所对应的目标函数 $f_i(x_1, x_2)$ 的值，同时也给出了群中的最大值 $f_{\max}$ 和平均值 $\bar{f}$ 。

表 1-1 遗传算法求 $f(X) = x_1^2 + x_2^2 + 1$ 极值计算过程

① 个体 编号	② 初始群体	③ x_1	④ x_2	⑤ 适应度 $f_i(x_1, x_2)$	⑥ 选择概率 $f_i/\sum f_i$	⑦ 适应度期望值 $f_i/\bar{f}$	⑧ 选择 次数
1	011101	3	5	35	0.239	0.96	1
2	101011	5	3	35	0.239	0.96	1
3	011100	3	4	26	0.178	0.71	0
4	111001	7	1	51	0.344	1.4	2
$\sum f_i = 147 \quad f_{\max} = 51 \quad \bar{f} = 36.25$							
⑨ 选择结果	⑩ 配对情况	⑪ 交叉点位置	⑫ 交叉结果	⑬ 变异点	⑭ 变异结果	⑮ 子代群体	
01 : 1101	1—3	1—3:2	011001	4	011101	011101	
1010 : 11			101001	2	111001	111001	
11 : 1001			111101	5	111111	111111	
1110 : 01			111011	6	111010	111010	
⑯ x_1	⑰ x_2	⑱ $f_i(x_1, x_2)$ 适应度			⑲ 选择概率 $f_i/\sum f_i$		
3	5	35			0.146		
7	1	51			0.213		
7	7	99			0.415		
7	2	54			0.226		
$\sum f_i = 239 \quad f_{\max} = 99 \quad \bar{f} = 59.75$							

(4) 选择运算的复制

选择的目的是在当前群体中选出优良个体，使它们有机会作为父代产生下一代个体。判断个体优良与否的准则就是各自的适应度，这一操作是借用进化原则，即个体适应度越高，其被选择的机会就越多。在本例中采用和适应度值成比例的概率来进行选择，其具体操作是：首先计算群体中所有个体适应度总和 $\sum f_i$ ；其次计算每个个体的适应度所占的比例 $f_i/\sum f_i$ ，如表 1-1 中第⑥栏所示，并以此作为每个个体遗传到下一代群体中的概率，每个概率值为一个区域，全部概率之和为 1；最后再产生一个 0 到 1 之间的随机数，由随机当选出现在哪一个概率区来确定各个个体中被选中的次数。表 1-1 中⑧、⑨栏所示选择结果，个体 1 和 2 各复制一次，个体 3 最差不复制被淘汰，因个体 4 最优秀，获得最多的生存繁殖机会，所以复制两次。

(5) 交叉运算及操作

遗传算法的交叉是产生新个体的主要操作过程，某两个个体之间的部分染色体是以某一概率相交换。本例采用简单的交叉（即单点交叉），其过程是：首先对群体随机配对，如表 1-1 中第⑩栏所示为一种随机配对情况；其次随机设置交叉点位置，如表 1-1 中第⑪栏和⑨栏所示，表中的数值字为交叉点设置在该基因座之后；最后再相互交换配对染色体之间的部分基因。表 1-1 中第⑫栏所示为交叉运算结果。

例如，若第 2 号和第 4 号个体基因座之后进行交叉运算，即可得到两个新个体：

第 2 号个体：1010 : 11 交叉操作 1010 : 01
 第 4 号个体：1110 : 01 1110 : 11

由上看出，所产生的新个体“111011”的适应度较原来两个个体的适应度还高，即 $X = [7, 3], f(7, 3) = 59$ 。

(6) 变异运算

变异的目的是为了防止丢失一些有用的遗传基因，尤其是当群体中的个体经遗传运算可能使某些串位失去多样性，从而可能失去检验有用遗传基因的机会，同时克服有可能限于局限解的弊病，为的是挖掘群体中个体的多样性。变异操作是按位进行，以很小的概率随机地改变某一位的值。对二进制编码的个体来说，若某位原为 0，则通过变异操作变成 1，反之亦然。在表 1-1 中第⑬栏所示为随机产生的变异点位置，⑭栏中为变异运算结果。下面举例说明变异操作。

例如，若第 1 号个体的第 4 位基因座需要进行变异运算，通过变异操作产生一个新个体：

第 1 号个体：0 1 1 0 0 1 第 4 位变异 0 1 1 1 0 1

以上对群体进行了一轮选择交叉、变异运算操作之后，得到新一代（子代）群体，见表 1-1 中第⑮栏，表中第⑯ ~ ⑲ 栏还分别表示出了新群体的解码值、适应度和相对适应度，并给出了适应度的最大值和平均值。由表 1-1 中可以看出，群体经过一代进化之后，其适应度的最大值、平均值得到了明显的改善，从而找到了最佳个体是第 3 号个体“111111”，最大值 $f_{\max} = f[7, 7] = 99$ 。

上述简单例子，为便于直接明了地看出运算操作的结果，其随机数的产生特意选择了一些较好的数值，以便能够得到较好的效果。实际运算过程中可能需要经过多少代的进化才能达到最佳结果。对于例中还有某种规律或指导性的原则，本书后面将深入讨论。

1.3 遗传算法的特点

传统的优化方法的主要问题:

1) 解析性 通常是通过求解目标函数梯度为零的一组非线性方程来进行搜索, 要求目标函数连续可微。当处理变量多、方程较复杂的优化问题时, 它就显得无能为力了。对于多峰问题容易陷入局部最优解。

2) 爬山法 对于单峰性质的空间, 且在更好的解位于当前解附近的前提下, 爬山法才能继续进行行之有效的搜索。它也是属于寻找局部最优解的方法。

3) 穷举法 在一个连续有限或离散的无限搜索空间中, 计算空间中每个点的目标函数值, 并进行逐点比较。对应于搜索空间很大时, 此法的效率很低。

4) 随机搜索法 主要指直接解法中的随机试验法和随机方向搜索法, 它们的计算精度和效率仍然不高, 计算量大, 通常用于小型优化问题。

遗传算法是寻找求解优化问题的效率和稳定性之间的有机协调, 计算方法新颖独特, 与传统优化计算方法相比有以下特点:

1) 对优化问题, 遗传算法不是直接处理决策变量本身的实际值, 而是对它进行编码为运算对象。此编码处理方式, 使优化计算过程可以借鉴生物学中染色体和基因等概念, 通过模拟自然界中生物的遗传和进化等机理, 可以方便地应用遗传操作算子, 特别是对一些无数值概念或很难有数值概念而只有代码概念的优化问题, 这种编码处理方式更显示出了独特的优越性, 使得遗传算法具有广泛的应用领域。

2) 许多传统的优化方法是单点搜索法, 遗传算法在搜索空间中同时处理群体中多个个体的方法, 即同时对搜索空间多个解进行评估, 在这之中包括一群搜索点进行寻优, 从而提高了搜索的效率, 有效地防止了搜索过程限于局部最优解, 减少了陷于局部最优解的风险, 而且具有较大的可能求得全局最优解。

3) 遗传算法对目标函数不要求连续, 更不要求可微, 既可以是数学解析所表示的显函数, 也可以是其他方式(映射矩阵或神经网络)的隐函数, 可以说对目标函数几乎没有限制, 仅用适应度来评估个体。适应度的唯一要求是输入可计算出加以比较下的输出, 利用适应度来指导搜索向不断改进的方向前进。

4) 很多传统的优化算法通常使用的是确定性搜索方法, 从一个搜索点到另一个搜索点有确定的转移方法和转移的关系, 这种确定性往往也有可能使得搜索难以达到最优点。然而遗传算法是属于一种自适应概率搜索技术, 采用概率变迁规则而非确定性规则来指导其搜索空间。虽然这种概率特性也会使群体中产生一些适应度不高的个体, 但随着进化过程的进行, 新的群体中总会更多地产生出许多优良个体, 继续沿着最优解方向前进。

5) 遗传算法具有隐含并行性, 不但使优化计算提高搜索效率, 而且易于采用并行机和并行高速计算, 因此适合大规模复杂问题的优化。

1.4 遗传算法的发展概况

早在 20 世纪 40 年代, 生物模拟就成为计算科学的一个组成部分。进化计算的研究起于 20 世纪 50 年代, 当时几个计算机领域的科学家独立地开始研究进化系统, 使之将自然界中的进化过程引入工程研究领域, 以解决工程优化问题。在用进化思想解决优化问题时, 使用了进化过程中的遗传、选择等概念, 并且把它们作为算子参与优化。

在 20 世纪 60 年代, Rechenberg 提出进化策略方法, 在此以后, 这一工作被 Schwefel 继续下去, 该法可以用于优化实值函数。同一时代 Fogel、Owens 和 Walsh 提出了进化规则的方法, 该方法把给定的问题描述成有限状态机制, 通过施加进化算子达到优化的目的。

Holland 在 20 世纪 60 年代运用生物遗传和进化的思想来研究自然和人工自适应系统的生成以及它们与环境的关系, 提出在研究和设计人工自适应系统时, 可以借鉴生物遗传的机制, 以群体的方法进行自适应搜索, 并且充分认识到交叉、变异等运算策略在自适应系统中的重要性。同时 Holland 不但发现了基于适应度的人工遗传选择的基本作用, 而且还对群体操作等进行了认真的研究, 1965 年他首次提出了人工遗传操作的重要性, 并把这些应用于自然系统和人工系统中。1968 年 Holland 教授提出了遗传算法的基本定理——模式定理, 从而奠定了遗传算法的理论基础。模式定理提出了群体中的优良个体的样本数将以指数级规律增长, 因而从理论上保证了遗传算法是一个可以用来寻求最优可行解的优化过程。

1967 年, Bagley 的论文中首次提出遗传算法这一名称, 他构造的遗传算法用来搜索下棋游戏评价函数中的参数集, 这与我们现在应用的遗传算法很相似, 其中利用了类似于前面描述过的复制、杂交和变异等的遗传算子。他还敏锐地意识到在运算开始和结束阶段需要适当的选择率, 为此还引入了适应值比例机制, 在算法的起始阶段减小选择的强制性, 在后期阶段增加选择的强制性, 因而, 在接近群体收敛时, 在类似的高适应值各串之间保持了适当的竞争。

1971 年, Hollstien 写的关于“遗传算法在纯数学优化应用”方面的第一篇学术论文, 主要研究了五种不同的选择方法和八种交配策略。他采用 16 位二进制串, 其中两个 8 位参数是用无符号二进制整数或 Gray 码整数来编码的, 群体规模为 16 个串, 通过计算机实验结果, Hollstien 指出由于群体规模太小 ($N = 16$) 所引起的问题。