



“关于遗传算法的报告”

课 程	数据结构与算法
学 院	环境科学与工程
学 号	1200092702
作 者	崔炫娥
指 导 老 师	陈斌
联 系 方 式	13126523643 1200092702@pku.edu.cn

【摘要】

遗传算法（GA）是基于自然界的进化现象及遗传机制而形成的一种全局性的概率搜索最优解的算法。其基本思想简单，具有全局并行搜索、简单通用、鲁棒性强等优点，其过程使种群像自然进化一样，后代种群比前代更加适应于环境，末代种群中的最优个体经过解码可以作为问题近似最优解。本文为了了解遗传算法，首先介绍遗传算法与生物进化的基本内容，其次阐述了基本框架及运算过程，然后介绍选择、交叉、变异等算子的类型及特点，最后指出遗传算法目前存在的问题及展望。

【关键词】

遗传算法 生物进化 遗传算子 优缺点

【正文】

一. 遗传算法与生物进化

遗传算法是模拟在自然环境中生物的优胜劣汰、适者生存的遗传和进化过程来形成的一种过程搜索最优解的算法。1970年初，John Holland（当时 Michigan大学的教授）提出的一种基于自然选择及自然遗传的随机搜索算法。GA是模拟在生物进化中的繁殖、交叉和基因突变现象，在每次迭代中都保留一组候选解，并按某种指标从解群中选取较优的个体，利用选择，交叉，变异（遗传算子）对这些个体进行组合，产生新一代的候选解群，重复此过程，直到满足某种收敛指标为止。

我们生活的自然界里，生命千姿百态，万物相互依存。数亿年的进化，使各种生物适应在现有的环境而生存下去并繁衍后代。进化的过程随时变化而进化的结果也各不一样，但所有生物的进化都具有一些共同的特点就是：进化的发生地点是信息载体，即染色体，而不是被编码的生物个体上；染色体能被复制，成功适应环境的被编码个体的染色体有更多的机会进行复制，形成有优势的后代；染色体能产生突变，使子代染色体不同于母代。如下表是生物学和遗传算法用法的不同点：

生物学	遗传算法
染色体 Chromosome	字符串 String
遗传基因 Gene	特征 Feature, 性格 character
等位基因 Allele	特征值 Feature value
位点 Locus	字符串位置 String position
基因型 Genotype	结构 Structure
显型 Phenotype	参数集 Parameter set

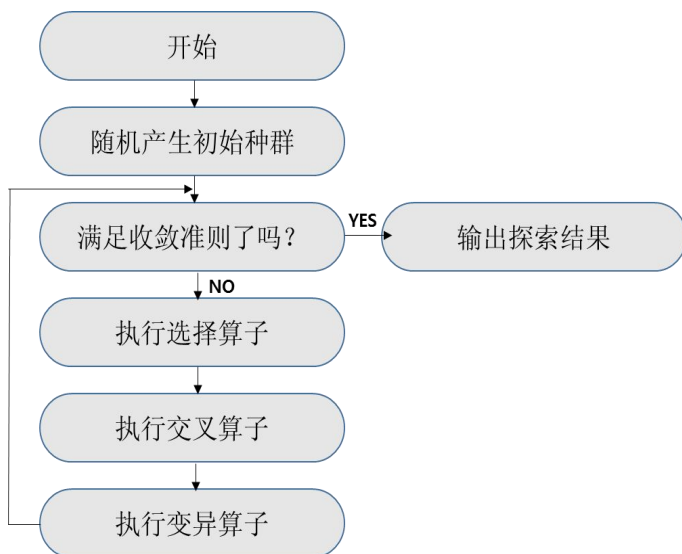
二. 遗传算法的基本框架及运算过程

GA只是一具有固定种群规模及个体用固定长度的基因链的抽象模型。它的操作算子包括选择、交叉和变异三种基本形式，构成了遗传算法强大搜索能力的核心，是模拟自然选择和遗传过程中发生的繁殖、杂交和突变现象的主要载体。

1. Selection（选择）：选择操作体现适者生存的原理，通过适应度选择优质个体而抛弃劣质个体，其主要作用是避免基因缺失，提高全局收敛性和计算效率。

2. Crossover（交叉）：交叉能使两个个体之间的遗传物质进行交换从而产生更好的个体。

3. Mutation（变异）：变异使基因的某个值强制地变化起来，以防止个体在形成最优解过程中过早收敛。



其基本框架是跟左图一样：建立模型；用设计好的算法将表现型映射到个体基因型；遗传算子只对编码后的染色体起作用，由个体表现型计算目标函数值后就可以判断染色体的优劣；确定适应度转换规则，染色体所对应的解空间的值可能相差很大，需要一定的转换使其适合定量评估个体的优劣；设计交叉、变异和选择算子。这基本框架具有完全性，封闭性，紧致性，可扩展性，可重性等特点。下边再次具体地介绍一下上述的三个算子类型及特点。

三. 选择、交叉、变异算子

1. 选择算子：选择操作是通过适应度来选择优质个体为下阶段抛弃劣质个体。常见的选择操作主要有以下几种，这种操作也能够重复使用：

1.1 排序选择：对个体适应值取正值或负值以及个体适应度之间的数值差异程度无特殊要求，对群体中的所有个体按其适应度大小进行排序，根据排序来分配各个体被选中的概率。

1.2 轮盘赌选择：选择某假设的概率而通过这个假设的适应度与当前群体中其他成员的适应度的比值而得到。这方法可以结合最优保存策略以保证当前适应度最优的个体能够进化到下一代而不被遗传操作的随机性破坏。

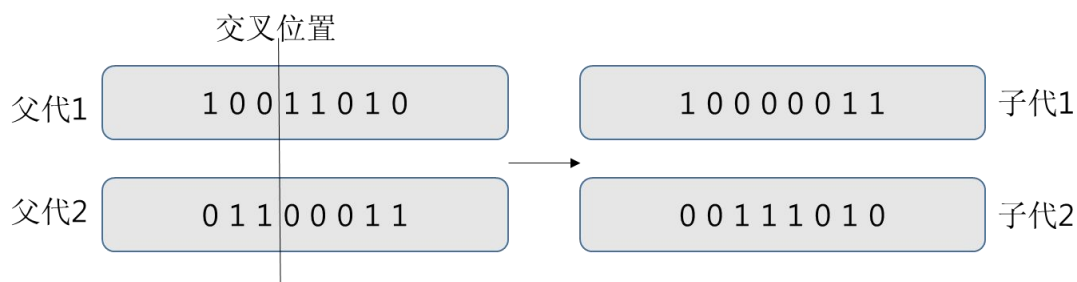
1.3 最优个体保存：父代群体中的最优个体直接进入子代群体中。这方法可以保证在遗传过程中所得到的个体不会被交叉和变异操作所破坏，它是遗传算法收敛性的一个重要保证条件。

1.4 随机联赛选择：每次选取N个个体（取2个的概率是最多的）中适应度最高的

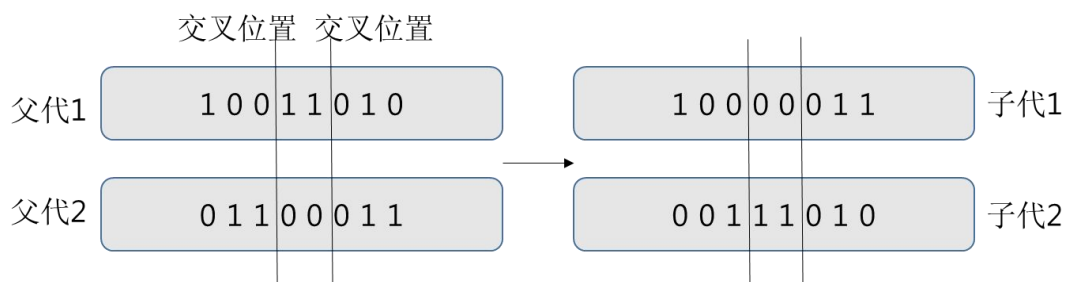
个体遗传到下一代群体中。

2. 交叉算子：交叉是产生新个体的主要方法，它对两个相互交叉的染色体按某种方式相互交换其部分基因，从而形成两个新的个体。常见的交叉操作主要有以下几种：

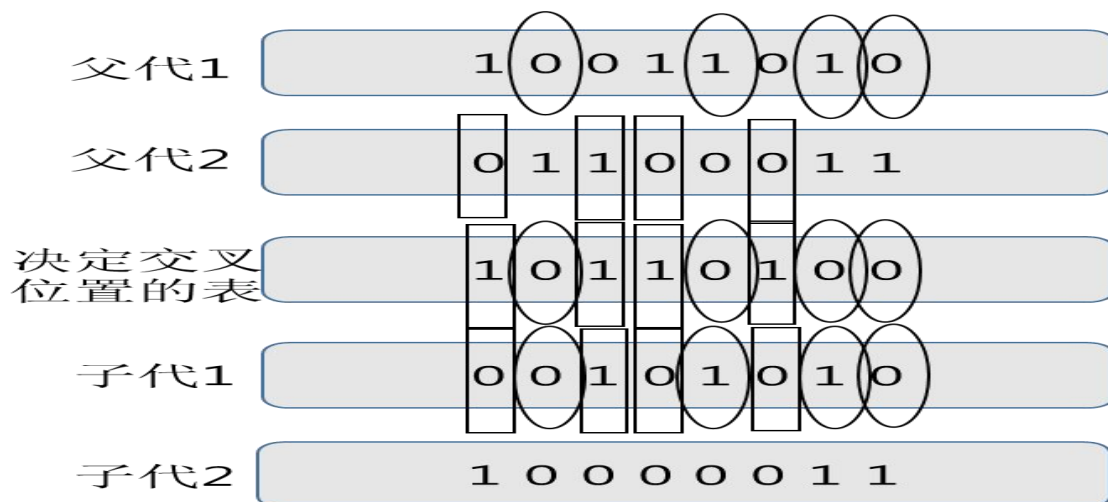
2.1 简单交叉（Simple Crossover）：又称为单点交叉，是指在个体编码串中随机设置一个交叉点后，在该点相互交换两个配对个体的部分染色体。



2.2 多点交叉（Multi-point Crossover）：又称为两点交叉，是指在相互配对的两个个体编码串中随机设置两个交叉点，然后交换两个交叉点之间的部分基因。



2.3 均匀交叉（Uniform Crossover）：又被称为一一致交叉，是指两个相互配对个体的每一位基因都以相同的概率进行交换，从而形成两个新个体。



2.4 部分映射交叉（Partially Mapped Crossover）：先对父代进行常规的两点交叉，再根据交叉区域内各基因值之间的映射关系来修改交叉区域之外的各个基因座的基因值。举例的话：

$$A = 9 \ 8 \ 4 \mid 5 \ 6 \ 7 \mid 1 \ 3 \ 2 \ 10 \rightarrow A' = 9 \ 8 \ 4 \mid 2 \ 3 \ 1 \ 0 \mid 1 \ 6 \ 5 \ 7$$

$B = 8\ 7\ 1\ |\ 2\ 3\ 10\ |\ 9\ 5\ 4\ 6 \rightarrow B' = 8\ 10\ 1\ |\ 5\ 6\ 7\ |\ 9\ 2\ 4\ 3$

2.5 顺序交叉 (Order Crossover)：其先进行常规的双点交叉，再对个体遍历路径进行有效顺序修改，修改时尽量维持各结点原有的相对遍历顺序。举例的话：

$A = 9\ 8\ 4\ |\ 5\ 6\ 7\ |\ 1\ 3\ 2\ 10$
 $B = 8\ 7\ 1\ |\ 2\ 3\ 10\ |\ 9\ 5\ 4\ 6$
 $\rightarrow B = 8\ H\ 1\ |\ 2\ 3\ 10\ |\ 9\ H\ 4\ H$
 $\rightarrow B = 2\ 3\ 10\ |\ H\ H\ H\ |\ 9\ 4\ 8\ 1$
 $A' = 5\ 6\ 7\ |\ 2\ 3\ 10\ |\ 1\ 9\ 8\ 4$
 $B' = 2\ 3\ 10\ |\ 5\ 6\ 7\ |\ 9\ 4\ 8\ 1$

3. 变异算子：变异的基本内容是对群体的个体串的某些基因座上的基因值变动。基于字符集 $\{0, 1\}$ 的二值码串而言，变异操作就是把某些基因座上的基因值取反，即 $1 \rightarrow 0$ 或 $0 \rightarrow 1$ 。常见的变异操作主要有以下几种：

3.1 基本位变异：对个体编码串以变异概率P随机指定某一位或某几位基因进行变异操作。

3.2 均匀变异：分别用符合某一范围内均匀分布的随机数，以某一较小的概率来替换个体编码串中各个基因座上的原有基因值。

3.3 二元变异：需要两条染色体参与，通过二元变异操作后生成两条新个体中的各个基因分别取原染色体对应基因值的同或及异或。

四. 建议

遗传算法的基本思想简单，不需要复杂的微分等的运算，而且多数的个体之间相互合作来能够探索最优解。但是，还需要进一步改进的地方也有，比如，算法规模小；遗传算法的编码一般二进制编码，但不同的应用应该采用不同的编码方式；遗传算法控制参数的选择问题；遗传算法与优化技术的融合等，如果在这几个方面上有逐步改进的话，遗传算法应该在智能计算领域中展现出更加光明的前景。

【参考文献】

- [1] 周明，孙树栋，遗传算法原理及应用，国防工业出版社，2001
- [2] 蔡良伟，李霞，遗传算法交叉操作的改进，系统工程与电子技术，2006
- [3] 朱灿，梁昔明，周书仁，基于物种选择的遗传算法，小型微型计算机系统，2009
- [4] 任子武，伞冶，实数遗传算法的改进及性能研究，电子学报，2007
- [5] http://www.aistudy.co.kr/biology/genetic/genetic_jeong.htm
- [6] <http://baike.baidu.com/link?url=Y3UfElG5cSLUx-wq8bpNHdutDnTNmCs3jCx4519pYx070zgZ1ZOHEimn-rreRMJtBlxjgnl1WNlgi0Cr34wq>