

快速遗传算法研究*

吴斌** 吴坚

(西南工学院信息与控制工程系 绵阳 621002)

涂序彦

(北京科技大学信息工程学院 北京 100083)

【摘要】 提出了一种称为广义自适应遗传算法的快速遗传算法,它首先产生均匀分布的初始种群,其次根据种群模式的状态决定是否引入“高品质”移民,最后自适应地进行交换和变异运算。其搜索性和全局收敛性比现有的许多遗传算法都有明显的改善,并通过仿真说明了该改进遗传算法的有效性。

关键词 广义自适应; 遗传算法; 初始种群; 移民; 适应度函数

中图分类号 TP301.6

在遗传算法中,算法的收敛性主要通过复制实现,而搜索性主要通过交换和变异实现。由于复制和交换的作用,适应度高于种群平均适应度的优良个体得到保留,并且其数量随着进化而不断增加,最后成为种群中的超级个体,产生封闭竞争,即出现“近亲繁殖”,减慢甚至导致进化地停滞,过早地收敛于局部极值解;而变异操作可以增加新的搜索空间,扩大搜索范围,以利于找到全局最优解,但增加变异会影响收敛的速度。现有的各种改进遗传算法有助于解决算法搜索性与收敛性间的矛盾,但许多算法只有在迭代步数很多时才有可能搜索到全局极值点附近,给算法的实际应用造成了困难。为此,本文提出了一种快速遗传算法——广义自适应遗传算法(GSAGA)。

1 广义自适应遗传算法的原理

1.1 初始种群的产生

在许多改进的遗传算法中,初始种群的产生依然采用完全的随机方式,而没有解决初始种群中各个体在解空间中的分布情况。这有可能让许多个体都集中在某一局部区域内,不利于扩大搜索空间和收敛到全局最优解。广义自适应遗传算法首先在初始种群的产生上要求各个体之间保持一定的距离,使各个体尽可能均匀地分布在整个解空间上。

定义 相同长度的以 a 为基的两个字符串中对应位不相同的数量称为两者间的广义海明距离,记为 GH 。

设个体是以 a 为基的字符串,个体的长度为 k ,种群的规模大小为 N ,则要求入选种群的所有个体之间的广义海明距离 GH 必须满足

$$GH_{ij} \geq (k-b) \quad i \neq j \quad (1)$$

式中 i, j 为两个个体, $i, j=1, 2, \dots, N$; b 为一常数,视不同的编码形式而定。一般若为二进制编码,则 $b = \text{int}(k/2)$; 若为十进制编码,则可取 $b \geq 2$ 。

种群的大小 N 影响着算法的有效性,当 N 太小时,算法会很差或找不出问题的解;而 N 太大时,则会使收敛时间增长,因此设定 $N=3K(K>4)$ 。

对于一个以 a 为基,长度为 k 的字符串,共有 a^k 个编码串。在这 a^k 个编码串中,相互间广义海明距离大于或等于 $(k-b)$ 的字符串编码总共有 $a^{k-(k-b)+1}$ 个。所以在广义自适应遗传算法中,种群的大小 $N=3k$ 远小于 $a^{k-(k-b)+1}$,即要求初始种群中所有个体间的广义海明距离大于或等于 $(k-b)$ 是完全能满足要求的。

1998年10月7日收稿,1998年12月3日修改定稿

* 四川省应用基础研究基金资助项目

** 男 33岁 博士生 讲师

设两个长度为 l 的字符串个体间的广义海明距离是 GH ，则这两个字符串中所包含的相同模式的数量为 2^{l-GH} ，而长度为 l 的字符串个体所含的模式数为 2^l ，故这两个字符串中所包含的不相同模式的数量就是 $(2^l - 2^{l-GH})$ 。广义海明距离 GH 越大，字符串间所包含的不相同模式的数量就越多，种群中的模式也越多。

初始种群采取这种方式产生就能保证随机产生的各个体间有较明显的差别，使它们能比较均匀地分布在解空间上，保证初始种群含有较丰富的模式，从而增加搜索收敛于全局最优解的可能。

1.2 适应度函数 f

在遗传算法中，适应度函数用来评价各个解的优劣。适应度计算的简单或复杂，取决于问题本身。对有些问题，只需要一个数学解析公式计算出来；有些问题本生不存在这样的解析式，需通过一系列基于规则的步骤才能求得；还有些问题的计算，是上述两种方法的结合。

在广义自适应遗传算法中，我们采用解析公式计算适应度。设实际问题的目标函数为 $J(x)$ ，适应度函数为 $f(x)$ ，则有

$$f(x_i) = J(x_i) - J_{\min} + \frac{J_{\max} - J_{\min}}{N} \quad (2)$$

式中 N 为种群大小； $J_{\min} = \min_{i=1}^N J(x_i)$ ， $J_{\max} = \max_{i=1}^N J(x_i)$ 。

1.3 复制算子

为了保证搜索到的最优个体不会因为选择、变异、交换算子的操作而被破坏，我们将父代种群中适应度最大的 $0.1N$ 个（10%）优良个体直接传递到子代种群中，成为子代种群中的个体。对父代种群中剩下的 $0.9N$ 个（90%）个体，按各自的适应度进行从小到大排序。设各个体相应的排序序号为 $r_i (i=1, 2, \dots, 0.9N)$ ，则每个个体按如下的数量复制到匹配池中

$$\text{int} \left(\frac{r_i}{\sum_{i=1}^{0.9N} r_i} \times 0.9 \times N + 0.5 \right) \quad (3)$$

1.4 “高品质”移民

在广义自适应遗传算法中，为了避免出现超级个体，防止发生封闭竞争，将根据匹配池中各个体间的差异来决定是否引入移民和移民的数量，并且要求被引入的移民具有较高的“品质”，即移民个体的适应度应该大于或等于当前匹配池中个体的平均适应度。当匹配池中各个体间的广义海明距离小于一特定值 λ 时，就不断地引入“高品质”移民随机地替代匹配池中的某些个体，直到匹配池中个体间的广义海明距离大于或等于 λ 为止。其中 λ 为一常数，一般取为 0.01。

1.5 自适应交换算子

交换是对被选择到匹配池中的 $0.9N$ 个个体进行的。对于随机选择的交换匹配对的两个个体来说，当两者之间的广义海明距离很小时（即两者非常相似），交换的作用变得不明显，因此应减少交换操作，甚至不进行交换操作，亦即交换的概率 p_c 应较小或为零；而当两者间的距离较大时，交换后产生新个体的机会也较大，有助于提高搜索效率，因而应有较大的交换概率 p_c 。

在广义自适应遗传算法中，交换概率 p_c 具有如下形式

$$p_{jc} = \begin{cases} 0 & \alpha + \frac{(CH_{ij} - \overline{GH})}{(0.001 + \overline{GH})} < 0 \\ \alpha + \frac{(CH_{ij} - \overline{GH})}{(0.001 + \overline{GH})} & 0 \leq \alpha + \frac{(CH_{ij} - \overline{GH})}{(0.001 + \overline{GH})} \leq 1 \\ 1 & \alpha + \frac{(CH_{ij} - \overline{GH})}{(0.001 + \overline{GH})} > 1 \end{cases} \quad (4)$$

式中 i, j 分别为匹配对中的两个个体; GH_{ij} 为它们间的广义海明距离; \overline{GH} 为匹配池中所有个体间的平均广义海明距离; α 为一个常数, 取值范围为 $0.2 \sim 0.8$ 。

这样确定的交换概率 p_c 将依据匹配对个体间的情况而变化, 距离大的交换可能性大; 反之, 交换的可能性就小。

1.6 自适应变异算子

变异算子保证算法能搜索到问题解空间的每一点, 使算法具有全局收敛性。当种群的平均广义海明距离 \overline{GH} 很小时, 说明种群中的各个体基本上趋于一致, 种群的基因模式单一性增强, 因而可能导致进化停滞, 过早地收敛于局部的极值解, 为此必须通过变异操作来改变这种情况。

在交换操作前, 用广义海明距离来评价要交换的双亲个体的差异, 根据这个差异, 以及当前匹配池中个体间的平均广义海明距离和适应度的情况来确定交换后的后代变异概率 p_m 。

在广义自适应遗传算法中, 经交换后的各个体按如下变异概率 p_m 进行变异操作

$$p_m = \frac{\beta}{(f_{\max} - \bar{f})\overline{GH} + (GH - \overline{GH}) + 0.001} \quad (5)$$

式中 f_{\max} 、 \bar{f} 、 \overline{GH} 、 GH 分别为匹配池中交换前所有个体的最大适应度、平均适应度和平均广义海明距离和配对个体间的广义海明距离; β 为常数, 一般取为 0.005。 $f_{\max} - \bar{f}$ 和 \overline{GH} 体现了群体的收敛程度, 当它们较小时, 说明群体已趋向收敛, p_m 应加大。

匹配池中的 $0.9N$ 个个体经移民、交换、变异操作后进入新一代种群。

1.7 停止条件

在广义自适应遗传算法中, 算法停止条件为 M 代内最优适应度值无显著提高。 M 太大则收敛时间太长, 而太小则所求得的最优结果与实际最优解相差太大。我们根据种群规模 N 来确定, 取 $M = \frac{N}{3}$ 。

2 广义自适应遗传算法仿真

文献[3]比较了几种遗传算法, 我们也用其中的函数来检验广义自适应遗传算法, 并与它的结果进行了比较。

2.1 实际问题

待优化函数为

$$J = 100 - \sum_{i=1}^3 x_i^2 \quad (6)$$

式中 $x_i (i=1, 2, 3)$ 的取值范围均为: $x_i \in [-5.12, 5.12]$, 求函数的最大值。函数的最大值在 $(0,0,0)$ 处取得, 为 100。

2.2 算法实现

每个变量用 8 位十进制数表示, 把三个变量的十进制码串接构成一个个体, 于是编码长度 $K=24$,

种群大小 $N = 3K = 72$, 结束迭代次数 $M = \frac{N}{3} = 24$ 。

适应度函数定义为

$$f_i = J_i - J_{\min} + \frac{J_{\max} - J_{\min}}{N} \quad (7)$$

式中 i 代表种群中第 i 个个体; J_{\min} 和 J_{\max} 分别为种群中最小和最大的目标函数值。

算法流程图如图 1 所示, 程序用 VB 编写。

2.3 结果分析

为消除随机性带来的干扰, 算法重复执行了 30 次, 将仿真实验数据取平均值记录于表 1 中。从表 1 可以看出, 当指定的目标函数值小于 99.99 时, 广义自适应遗传算法能迅速地收敛到指定极值点, 并且在实验中发现如果没有移民, 则对于搜索大于或等于 99.99 的目标函数值将要进化很多代才能搜索到,

甚至有时在指定的进化代数内搜索不到指定的极值。

与文献[3]的结果相比较，广义自适应遗传算法不仅收敛速度快，而且能够搜索到更优越的解，其搜索性和收敛性比许多遗传算法都有很大的改善。

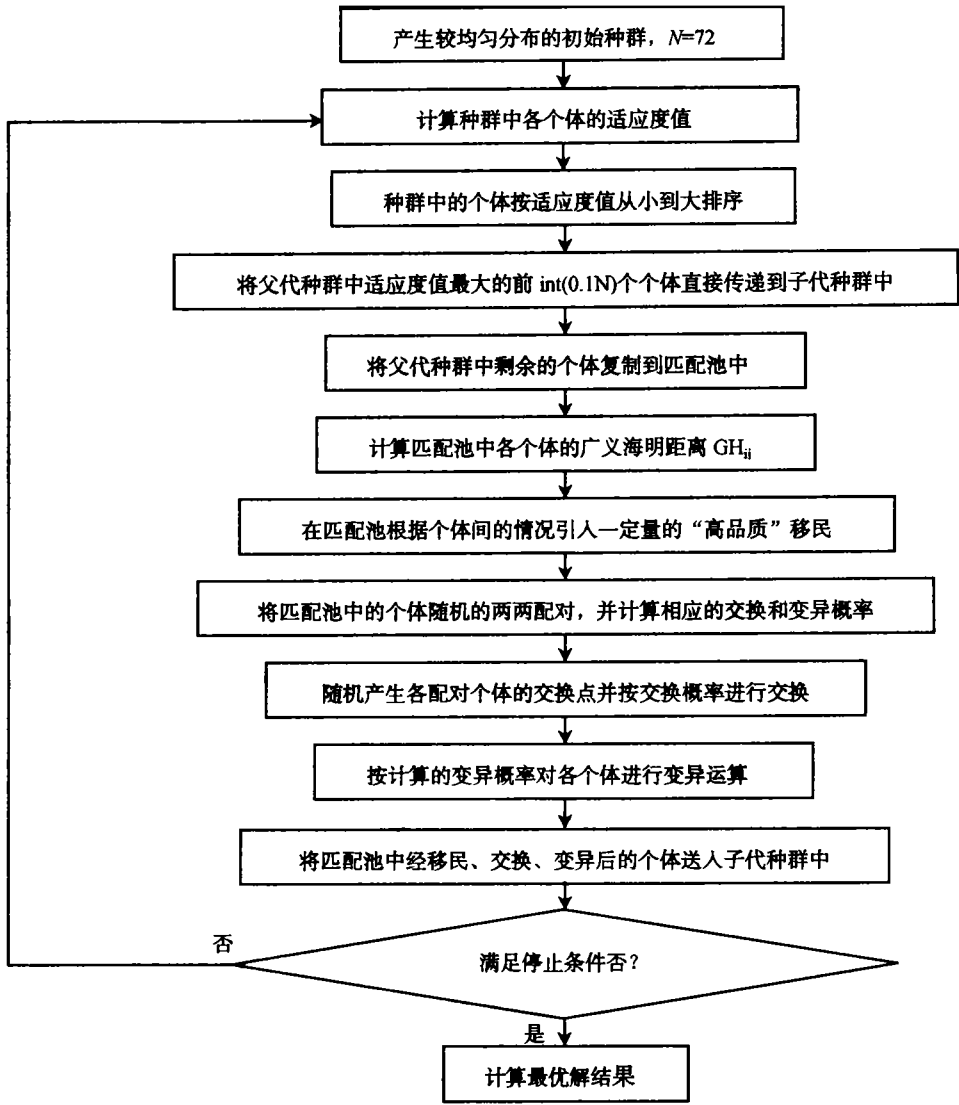


图 1 广义自适应遗传算法实现流程图

表 1 不同遗传算法达到指定函数值的进化代数

目标函灵敏值	99	99.5	99.9	99.95	99.99	99.995	99.999	99.999 5
简单遗传算法	3	8	30 、	45	185	—	—	—
均匀交换遗传算法	3	8	54	90	157	—	—	—
引入普通移民的算法	4	7	40	78	179	—	—	—
文献[3]的遗传算法	3	7	26	58	134	163	—	—
广义自适应遗传算法	2.63	4.47	8.4	11.63	38.83	46.1	51.27	52.87

注：“—”表示在进化 200 代后该算法的最大目标函数值仍不能达到指定值。

3 结 论

从上面的分析和仿真实验结果可以看出, 由于广义自适应遗传算法首先通过产生能适应问题解空间分布状况的初始种群, 保证了初始种群中个体模式的合理分布; 其次是通过保护优秀个体直接进入下一代种群, 避免了交换、变异操作破坏优良个体, 同时, 有条件地引入“高品质”移民, 有效地解决了种群中个体的多样性; 最后, 对匹配池中的个体用自适应的方式确定各个体的交换概率 p_c 和变异概率 p_m , 从而较好地解决了算法收敛性和搜索性之间的协调问题。在广义自适应遗传算法的整个过程中, 无论是初始种群的产生, 还是外来移民的引入, 以及交换和变异概率的确定都是根据种群中各个体的具体情况而定的, 并且自适应地随着进化而不断地改变。所有这一切使得广义自适应遗传算法具有良好的搜索性和收敛性, 其性能明显优于现存的许多遗传算法。

参 考 文 献

- 1 Holland John H. Adaptation in natural and artificial systems. Michigan:The University of Michigan Press,1975
- 2 Goldberg , David E. Genetic algorithms on search, optimization and machine learning. New York:Addisen-Wesley, 1989
- 3 蔡 弘, 李衍达, 一种快速收敛的遗传算法, 智能控制与智能自动化. 第二届全球华人智能控制与智能自动化大会论文集. 西安: 西安交通大学出版社, 1997:1 689~1 692

Research of Fast Genetic Algorithm

Wu Bin Wu Jian

(Dept. of Information and Control Engineering , SWIT Mianyang 621002)

Tu Xuyan

(Information Engineering School, UST Beijing Beijing 100083)

Abstract A fast genetic algorithm——GSAGA(generalized self-adaptive genetic algorithm) is presented in this paper. First, evenly distributed initial population is generated. Then, high quality immigrants are introduced according to the condition of the population schema. Finally, crossover and mutation are operated on self-adaptively. In GSAGA, the searching performance and global convergence are greatly improved compared with many existing genetic algorithms. Through emulation, the validity of this modified genetic algorithm is proved.

Key words generalized self-adaptive; genetic algorithm; initial population; immigration; fitness function