

文章编号: 1001-0920(2002)06-908-04

一种新的优胜劣汰遗传算法

巩敦卫, 孙晓燕, 郭西进

(中国矿业大学 信息与电气工程学院, 江苏 徐州 221008)

摘要: 为防止进化种群早熟收敛, 并考虑进化种群多样性与进化代数的关系, 提出一种新的近亲交叉回避策略, 该策略中实施近亲交叉回避操作的下限随进化代数和种群平均海明距离变化。在此基础上提出一种新的优胜劣汰遗传算法, 该算法能有效地避免近亲繁殖且体现了优胜劣汰思想。业已证明, 提出的算法可以保证收敛到全局最优解。仿真结果表明, 与简单遗传算法相比该算法是有效的。

关键词: 遗传算法; 优胜劣汰; 近亲交叉回避

中图分类号: TP 18 **文献标识码:** A

Novel survival of the fittest genetic algorithm

GON G Dun-w ei, SUN X iao-yan, GUO X i-jin

(College of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou 221008, China)

Abstract: In order to prevent premature convergence of evolution population, a novel closed crossing avoidance strategy is presented by considering the relationship between diversity of evolution population and evolution times. The lower limit of closed crossing avoidance varies with evolution times and average Hamming distance of evolution population. A novel survival of the fittest genetic algorithm is presented. The algorithm can avoid close breeding effectively and externalize the thought of survival of the fittest. It has been proved that the algorithm can converge to globally optimal solution. Simulation result shows that the algorithm presented is efficient contrast with simple genetic algorithm.

Key words: genetic algorithm; survival of the fittest; closed crossing avoidance

1 引言

遗传算法(GA)是根据生物进化理论和遗传变异理论提出的一种基于种群搜索的优化算法,其思想是通过选择复制和遗传算子的共同作用使种群不断进化,最终收敛到优化解^[1]。遗传算法是一种全局概率搜索优化算法,它针对染色体进行遗传操作,而不直接针对决策变量本身。更重要的是,遗传算法对优化函数本身既不要求连续也不要求可微。因此,在

函数优化^[2]、参数辨识^[3]、机器人控制^[4]、神经网络训练^[5]、模糊逻辑系统^[6]等方面得到了广泛的应用。

最早研究的遗传算法是简单遗传算法(SGA)。简单遗传算法思路直观,操作简单,但其收敛速度慢且受参数选择的影响。更重要的是,简单遗传算法不能保证收敛到全局最优解^[7]。为此,文献[8]提出了一种改进的遗传算法,该算法中的交叉操作采用近亲交叉回避策略,并采用优选变异,提高了遗传操作的效率,加快了收敛速度。但近亲交叉回避策略的海

收稿日期: 2001-08-27 修回日期: 2001-12-03

基金项目: 中国矿业大学科研基金项目(01A 04)

作者简介: 巩敦卫(1970—),男,江苏铜山人,副教授,博士,从事智能控制、遗传算法和CM理论与技术的研究;孙晓燕(1978—),女,江苏丰县人,助教,硕士,从事智能控制、遗传算法的研究。

明距离下限由于不随进化代数和本代平均海明距离变化, 因而不利于产生多样性种群。此外, 由于替换串采用随机方法, 因而替换后的新串不一定是父代串, 这样子代个体的产生失去了依据。文献[9]提出采用混沌变异的进化算法, 该算法的染色体编码采用实数编码, 变异算子采用混沌变异, 混沌序列为一维 Logistic 映射, 考虑到收敛精度与进化代数的关系, 混沌变异结合了尺度收缩思想。因此, 从一定程度上提高了收敛速度。但由于采用线性交叉算子, 在进行交叉之前不进行近亲判断和回避, 因此交叉操作的效率不高。

本文提出一种新的近亲交叉回避策略, 该策略与进化代数和本代种群平均海明距离有关。考虑到种群进化的效率, 提出一种新的优胜劣汰遗传算法, 该算法可以保证收敛到全局最优解。

2 一种新的优胜劣汰遗传算法

为便于描述, 引入以下定义^[7]:

定义 1(个体和个体空间) 所谓 l - 个体, 是指长度为 l 的 0 和 1 字符串, 简称个体。 l - 个体的全体称为个体空间, 记作 $S = \{0, 1\}^l$ 。

定义 2(种群和种群空间) 所谓 N - 种群, 是指 N 个个体组成的集合(个体允许重复), 简称种群。 N - 种群的全体称为种群空间, 记作

$$S^N = \{X = (X_1, X_2, \dots, X_N), \\ X_i \in S, i = 1, 2, \dots, N\}$$

若 x_{ij} 表示个体 X_i 的第 j 个基因, $j = 1, 2, \dots, l$, 则 X_i 可表示为

$$X_i = (x_{i1}, x_{i2}, \dots, x_{il})$$

定义 3(齐次种群) 考虑种群 X , 若任意 $i, j = 1, 2, \dots, N$, 有 $X_i = X_j$, 则称 X 是齐次种群。齐次种群的全体用 β^N 表示。

2.1 近亲交叉回避

若 $X \in \beta^N$, 则任意 $i, j = 1, 2, \dots, N$, X_i, X_j 间的交叉操作都不会产生新的子代个体, 若 X 接近于一个齐次种群, 其个体间的交叉操作也很难产生新的子代个体, 这种母体通过交叉操作产生新的子代个体的进化方式称为近亲繁殖。近亲繁殖不利于产生多样性种群。

当进化代数接近终止代数时, 种群接近于一个齐次种群。为了提高进化效率, 使每代通过交叉操作产生多样性种群, 同时考虑种群多样性与进化代数

的关系, 本文在文献[8]的基础上, 提出一种新的近亲交叉回避策略。其思想是: 若两个个体的海明距离较大, 则选作待交叉母体, 否则在该种群中选择另一个体。海明距离的下限随进化代数而变化。

记 $X(t)$ 为第 t 代进化种群, 其第 i 和第 j 个体分别为 $X_i(t)$ 和 $X_j(t)$, 它们的海明距离为

$$H(X_i(t), X_j(t)) = \sum_{k=1}^l |x_{ik}(t) - x_{jk}(t)| \quad (1)$$

$X(t)$ 的平均海明距离为

$$H(t) = \frac{N-1}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^N H(X_i(t), X_j(t)) / \sum_{i=1}^{N-1} (N-i) \quad (2)$$

记 $\rho(t)$ 为反映进化代数的作用因子, 其表达式为

$$\rho(t) = \exp(-t^2/(2\sigma^2)), \quad \sigma = T/3 \quad (3)$$

这里 T 为进化终止代数。若 $H(X_i(t), X_j(t)) > \rho(t)H(t)$, 选择 $X_i(t), X_j(t)$ 作为待施行交叉操作的母体; 否则选择 $X_{j_0}(t) \in X(t)$, 使得

$$j_0 = \arg \max_j \{H(X_i(t), X_j(t))\} \quad (4)$$

$X_i(t), X_{j_0}(t)$ 作为待施行交叉操作的母体。

在对非近亲的两个个体实施交叉操作时, 若选择的交叉点使得两个交换段的海明距离为 0, 则通过交叉操作不会产生新的子代个体, 这虽不属于近亲繁殖, 但却是失败繁殖, 此时可通过重新选择交叉点来避免。

记 $X_i(t), X_j(t)$ 为待施行交叉操作的母体, c 为交叉点, $\text{Seg}(X_i^c(t))$ 和 $\text{Seg}(X_j^c(t))$ 分别为 $X_i(t), X_j(t)$ 自交叉点之后的基因段, 其海明距离为

$$H(\text{Seg}(X_i^c(t)), \text{Seg}(X_j^c(t))) = \sum_{k=c+1}^l |x_{ik}(t) - x_{jk}(t)| \quad (5)$$

若 $H(\text{Seg}(X_i^c(t)), \text{Seg}(X_j^c(t))) = 0$, 则选择

$$c_0 = \min \{c; H(\text{Seg}(X_i^c(t)), \text{Seg}(X_j^c(t))) > 0, \\ c \in \{1, 2, \dots, l-1\}\} \quad (6)$$

作为交叉点。

2.2 算法描述

1) 确定种群规模, 决策变量采用二进制编码, 选择、交叉和变异的方法分别采用按比例选择、单点近亲回避交叉和单点变异, 交叉、变异概率待定, 算法终止条件是进化代数等于进化终止代数。

2) 取 $t = 0$, 以一定的方式产生初始种群 $X(0)$ 。

3) 计算种群 $X(t)$ 中个体 $X_i(t)$ 的适应度 $f(X_i(t))$, $i = 1, 2, \dots, N$, 再根据个体的适应度按比例选择 N 个母体。

4) 对选中的 N 个母体用近亲交叉回避策略进行交叉操作和变异操作后得到种群

$$Y(t) = (Y_1(t), Y_2(t), \dots, Y_N(t)) \quad (7)$$

对于 $X(t)$, 记

$$k_0 = \arg \min_k \{ \max_k \{ f(X_k(t)) \} \} \quad (8)$$

对于 $Y(t)$, 记

$$i_0 = \arg \min_i \{ f(Y_i(t)) \} \quad (9)$$

若 $f(X_{k_0}(t)) > f(Y_{i_0}(t))$, 得到子代种群为

$$X(t+1) = (Y_1(t), Y_2(t), \dots, Y_{i_0-1}(t), X_{k_0}(t), Y_{i_0+1}(t), \dots, Y_N(t)) \quad (10)$$

否则, 得到子代种群为

$$X(t+1) = Y(t) \quad (11)$$

5) 若 $t = T$, 则进化终止, 输出优化结果, 否则, 令 $t = t + 1$, 转到步骤3)。

算法的计算框图如图1所示。

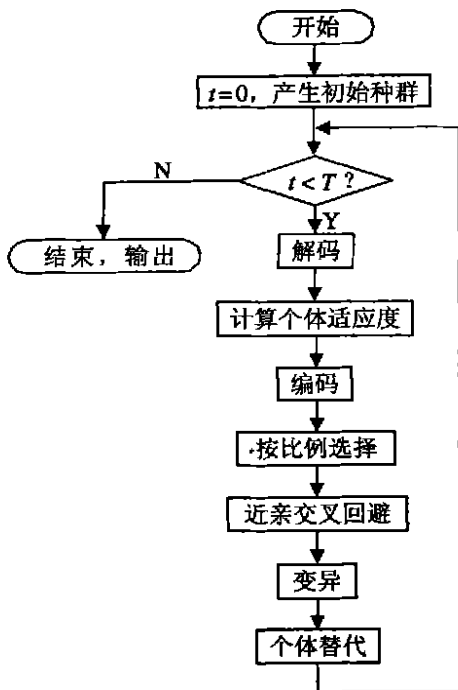


图1 算法的计算框图

2.3 算法分析

在进化过程的前期, 增加种群的多样性是主要目的, 此时应使下限大一些, 而到进化过程的后期,

种群趋于齐次种群, 增加种群的多样性已不是主要目的, 此时应使下限很小。因此, 本算法由于使用近亲交叉回避操作, 增加了种群的多样性, 从而提高了种群进化的效率。实施近亲交叉回避操作的下限是一个变量, 让它随进化代数的增加而减少, 最终趋于0, 这是符合进化规律的。

本算法体现了优胜劣汰思想, 当父代种群最优个体的适应度大于子代最劣个体的适应度时, 以前者取代后者, 从而使子代种群的性能不劣于父代种群。文献[7]中证明, 这种通过优胜劣汰遗传算法得到的种群序列 $\{X(t); t = 0, 1, \dots\}$ 以概率1收敛到满意种群集 $M^* = \{X; \max\{f(X); X \in S\}\}$, 即

$$\lim_{t \rightarrow \infty} P\{X(t) \in M^* / X(0)\} = 1$$

因而, 该算法可以收敛到全局最优解。

3 仿真

为验证本文算法的有效性, 就算法的全局收敛性和收敛速度与简单遗传算法进行比较。取被优化函数

$$\max f(x) = (\sin 2x + \cos 5x)^3, \quad x \in [0, 9]$$

图2为该函数的图形, 它是一个多峰函数, 共有8个局部极大值, 但其中只有一个最大值为7.5029, 对于搜索能力不强的算法来讲, 很难收敛到全局最优解。

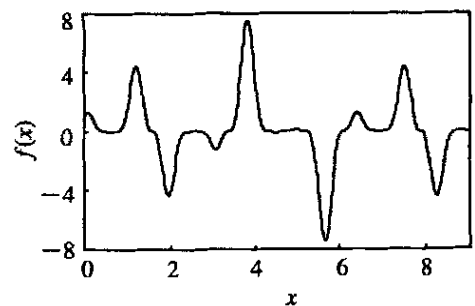


图2 函数的图形

仿真时个体采用二进制编码, 计算精度为小数点后4位, 考虑到变量的取值区间, 码长为 $l = 17$, 种群规模为 $N = 30$, 交叉概率为 $P_c = 0.9$, 变异概率为 $P_m = 0.0026$, 终止代数为 $T = 120$ 。两种算法在计算时初始种群相同, 记录每次最大值, 收敛到最大值的迭代步数和种群的平均适应度, 共计算10次, 然后取10次的平均值, 对比结果如表1所示。在计算过程中, 简单遗传算法有2次到迭代终止步数时

表 1 计算比较结果

| 算 法 | 最大值 | 收敛到最大值的迭代步数 | 种群平均适应度 |
|----------|---------|-------------|----------|
| 简单遗传算法 | 1.861 2 | 63 | 55.934 5 |
| 优胜劣汰遗传算法 | 7.494 6 | 34 | 61.261 8 |

仍未达到最大值, 为便于比较, 此时记收敛到最大值的迭代步数为 120。

从表 1 可以看出, 简单遗传算法收敛不到全局最优解, 本文提出的优胜劣汰遗传算法可以收敛到全局最优解, 且迭代次数少。由此可见, 本文提出的算法是一种较优的搜索算法。

4 结 语

考虑进化种群多样性与进化代数的关系, 提出一种新的近亲交叉回避策略。在此基础上提出一种新的优胜劣汰遗传算法, 该算法不但能有效地避免近亲繁殖, 体现了优胜劣汰思想, 而且可以保证收敛到全局最优解。仿真结果表明本文算法是有效的。在该算法中, 作用因子对下限值的确定有很大影响, 目前只能根据经验确定。因此, 如何合理确定作用因子将是进一步研究的课题。

参考文献(References):

- [1] Holland J H. *A daptation in N atural and A rtificial Systems* [M]. Michigan: The University of Michigan Press, 1975
- [2] 唐加福, 汪定伟, 高振, 等. 面向非线性规划问题的混合式遗传算法[J]. 自动化学报, 2000, 26(3): 401-404
(TANG Jia-fu, WANG Ding-wei, GAO Zhen, et al. Hybrid genetic algorithm for solving nonlinear programming problem [J]. *Acta Automatica Sinica*, 2000, 26(3): 401-404)
- [3] 张永健, 周东华. 基于遗传算法的非线性系统时变时滞的在线估计[J]. 控制与决策, 2000, 15(6): 756-758
(ZHANG Yong-jian, ZHOU Dong-hua. Online estimation approach based on genetic algorithm to time-varying time delay of nonlinear systems[J]. *Control and*

Decision, 2000, 15(6): 756-758)

- [4] 樊晓平, 徐建闽, 毛宗源. 受限柔性机器人基于遗传算法的自适应模糊控制[J]. 自动化学报, 2000, 26(1): 61-67.
(FAN Xiao-ping, XU Jian-min, MAO Zong-yuan. Genetic algorithm based adaptive fuzzy control for constrained flexible-link manipulators[J]. *Acta Automatica Sinica*, 2000, 26(1): 61-67.)
- [5] Dunwei GONG, Shifan XU, Xiaoyan SUN. Research on fast training algorithm for recurrent neural network [A]. *Proc of the 6th IEEE International Symposium on Industrial Electronics*[C]. Piscataway: The Institute of Electrical and Electronics Engineers, Inc, 2001. 446-448
- [6] Chang-wook Han, Jung-il Park. Design of a fuzzy controller using genetic algorithms employing random signal-based learning and simulated annealing[A]. *Proc of the 6th IEEE International Symposium on Industrial Electronics*[C]. Piscataway: The Institute of Electrical and Electronics Engineers, Inc, 2001. 1231-1236
- [7] 张文修, 梁怡. 遗传算法的数学基础[M]. 西安: 西安交通大学出版社, 2000
- [8] 边润强, 陈增强, 袁著祉. 一种改进的遗传算法及其在系统辨识中的应用[J]. 控制与决策, 2000, 15(5): 623-625
(BIAN Run-qiang, CHEN Zeng-qiang, YUAN Zhu-zhi. Improved genetic algorithm and its application in system identification[J]. *Control and Decision*, 2000, 15(5): 623-625)
- [9] 骆晨钟, 邵惠鹤. 采用混沌变异的进化算法[J]. 控制与决策, 2000, 15(5): 557-560
(LUO Chen-zhong, SHAO Hui-he. Evolution algorithm with chaotic mutation[J]. *Control and Decision*, 2000, 15(5): 557-560)