

遗传算法的自适应代沟的替代策略研究

李 兵 谢剑英

(上海交通大学自动化系·上海 200030)

摘要: 遗传算法的替代策略对遗传算法的性能具有重要意义. 提出自适应代沟的替代策略. 仿真表明, 与现有的其他替代策略相比, 该方法在不损失解的质量的情况下, 使基于实数编码的遗传算法求解效率得到明显的改善.

关键词: 遗传算法; 自适应代沟; 竞争选择

文献标识码: A

The Replacement Strategy with Adaptive Generation Gap for Genetic Algorithm

LI Bing and XIE Jianying

(Department of Automation, Shanghai Jiaotong University, Shanghai 200030, P.R. China)

Abstract: The replacement strategy is of great importance for the performance of the genetic algorithms. We introduce a new replacement strategy with adaptive generation gap. Simulation shows, compared with other strategies, the new method can improve the efficiency of the genetic algorithms based on the real-valued coding without loss of quality.

Key words: genetic algorithm; adaptive generation gap; tournament selection

1 遗传算法简介(Introduction of genetic algorithms)

遗传算法是一种非数值并行优化算法, 首先是由 Holland 提出. 它是模拟自然界生物物种进化过程, 利用编码将问题的解用字符串表示, 然后应用选择、交叉、变异、替代等对个体进行操作, 实现在解空间中的搜索. 再根据“适者生存”原则, 保留计算过程中出现的最优解或“潜在最优解”, 并在此解基础上, 进行不断迭代, 直到找到最优解或满意解为止. 标准遗传算法的优化过程如图 1 所示. 由于遗传算法对解决特别困难问题和搜索全局最优解具有极强的能力, 而且, 对所要解决的问题要求了解的知识很少, 这些都是其他如梯度法、迭代搜索等搜索算法所没有的优点. 所以, 遗传算法在计算机网络^[1]、机器学习^[2]、结构优化设计^[3]、大规模信息处理、电子电路计算机辅助设计、经济预测以及各个学术领域的优化计算等方面得到广泛应用.

目前遗传算法的研究主要分为工作原理^[4]、计算模型^[5]、编码方式^[6]、操作符^[7]、早熟收敛问题、问题结构类型等领域. 根据编码方式的不同, 遗传算法主要分为二进制编码遗传算法(binary genetic algorithms)和实数编码遗传算法(real-valued genetic

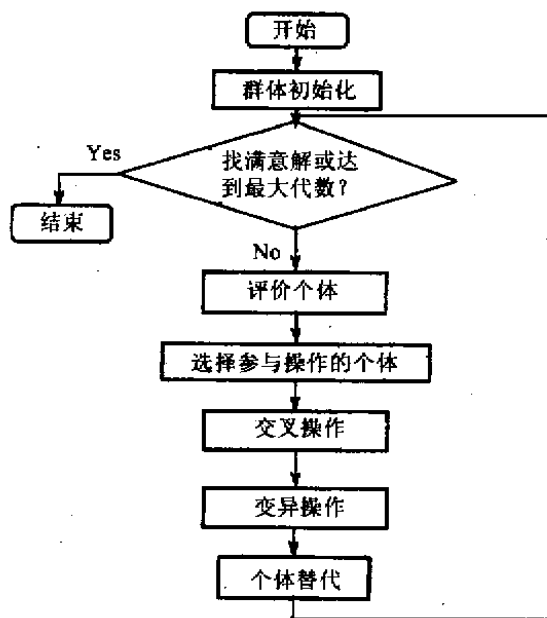


图 1 标准遗传算法流程图

Fig. 1 Standard genetic algorithm flow chart algorithms)两种形式, 并分别有自己的优势: 二进制编码遗传算法适合用在组合优化方面, 而实数编码主要适合于方程优化. 遗传算法在实现方式上分为串行方式和并行方式两种. 随着问题复杂程度的提高及计算规模的增大, 近来并行方式正在越来越

受到高度的重视.但无论是串行计算还是并行计算方式,选择模式、交叉、变异以及个体的替代方式作为遗传算法的重要环节,对于遗传算法的性能都具有十分重要的意义.

2 自适应代沟替代策略的分析研究(Analysis of replacement strategy with adaptive generation gap)

遗传算法研究的主要目标是既要使算法解的质量高,又要求算法具有高的计算效率.由于遗传算法是一种随机搜索,因此解的好的质量通常需要大的计算量;另一方面,收敛速度高的算法往往导致早熟,使解的质量降低.而所谓早熟收敛,是指算法收敛到非全局最优解的情况,因此解的质量必定不如全局最优解.因此好的算法是计算效率和解的质量之间的平衡.从实现手段来讲,就是要在计算过程中,采取合适的个体多样性.我们可以从遗传算法的选择操作来理解收敛速度与解的质量间的关系.从个体选择方面来讲,高的收敛速度往往通过采用高选择压力,即给予当前代中的具有较高适合度的个体更高的再生概率来实现.但另一方面,由于群体大小不变,使得当前代中的具有较低适合度的个体降低了生存概率,从而降低了群体中个体的信息含量,使算法探索新的解的区域的能力降低,即降低了算法搜索全局最优解的能力.因此,容易导致早熟收敛.解决收敛速度和解的质量之间的矛盾,可以从许多方面着手,这里我们考虑改进替代策略.

遗传算法的替代策略同样必须考虑的重要因素是保持个体多样性.遗传算法的个体替代操作方式体现了自然界父代与子代之间的继承与相互竞争的关系.这种普遍的生物现象,在物种进化过程中产生了大量具有广泛适应性的鲁棒性结构.并且由遗传算法流程图可以看出,个体替代策略是算法所必需的一个重要环节.遗传算法的性能在很大程度上依赖于个体替代策略,也就是依赖于如何在每一代中选择被替换的个体来提高算法的效率和求解质量.其实质是怎样既要保留进化过程中父代个体携带的有效模式,又要用子代中产生的个体所携带的更有效的新模式来替换父代中已不再有生命力个体中的旧有模式.在遗传算法中,用代沟的概念来表示每代中个体被替代的比例.目前,替代策略分为两种^[8]: 1) Grefenstette将代沟取为 1,即每一代的个体由下一代全部替换.这种策略是基于这样一种假设:所有子代个体比父代个体都好. 2) L. Davis 等只将每代中

少量固定数量的个体替换掉,比较典型的是将被替换个体数取为 2.这种操作策略被称作“稳态替代”.但是 D. E. Goldberg 等证明稳态替代策略的性能,可以通过完全替代策略,采用提高竞争选择的选择压力(selection pressure)来达到,并不比策略 1) 更有优势.综观 1) 2) 两种替代策略,有一个共同点,就是人为的决定每代所被替代的个体数量,这种人为地确定被替代个体的数量,实际上忽略了遗传算法优化过程的动态性.也就是说,遗传算法每一代对解的改善程度是不确定的.与上一代相比,具有更有效模式的个体数量也是动态变化的.人为确定被替代个体的数量,导致在拥有较多有效模式的群体中,对有些个体所拥有的有效模式的信息被浪费掉,使得算法计算效率和解的质量受到损失.

“自适应代沟替代策略”将父代与子代个体分级,父代中所有适合度低于子代同级个体适合度的个体全部替换.具体过程分为两步:①将父代与子代和子代分别根据适合度值大小排序;②比较父子两代中的排序相同的个体,保留两者中的高适合度个体.由于保留的个体是从父、子两代中选择,所以其实质也是一种滚动优化的思想体现.与完全替代策略相比,自适应代沟替代策略的特点是代沟不是固定常数,而是对应于算法的动态过程而动态变化的.在平均适合度提高较多的子代,有更多的个体来代替父代的个体,而平均适合度提高较少的子代,其父代就会有更多的个体被保留.这种替代策略与前面所述的 1) 2) 两种策略相比,更类似于生物进化过程中的自然选择和优胜劣汰的优化过程.

3 仿真结果与分析(Simulation results and analysis)

为比较自适应代沟替代策略与完全替代策略的性能,我们将两者在轮盘赌选择与竞争选择两种情况下进行仿真比较.

我们将“自适应代沟的替代策略”与代沟为 1 的完全替代策略进行性能比较.因为根据 D. E. Goldberg 的研究, L. Davis 的稳态替代策略的性能可以通过完全替代策略,利用提高竞争选择的规模来达到.

根据多样性损失公式(公式 1),当 Tour 取到 5 时,个体多样性损失就达到大约 50%.为尽量避免遗传算法的早熟收敛,需要保持个体多样性.因此在仿真中,其中 Tour 取 30,是为比较自适应代沟替代与完全替代在极限情况的高选择压力下的性能.

$$\text{LossDiv}(\text{Tour}) = \text{Tour}^i(-1/(\text{Tour}-1)) - \text{Tour}^i(-\text{Tour}/(\text{Tour}-1)).$$

(1)

LossDiv——个体多样性损失, Tour——选择规模, 且 $\text{IndNum} \geq \text{Tour} \geq 2$, IndNum——群体中的个体数, 这里我们取 $\text{IndNum} = 30$.

这里我们采用典型的测试方程 $F6^9$ (公式 6) 来进行仿真比较. 图 2 为 $F6$ 的三维图形. 由于 $F6$ 为多模式方程, 我们可以利用它检验遗传算法解决困难问题的能力. 图 3, 图 4, 图 5, 图 6, 图 7 (Tour 分别为 2, 3, 4, 5, 30) 为采用竞争选择方式时两种不同替代策略运行 30 次的平均性能曲线. 图 8 为轮盘赌选择方式下两种替代策略运行 30 次的平均性能曲线.

通过图 3 至图 7 可以看到, 采用竞争选择方式, 无论是在高选择压力还是低选择压力情况下,

$$F6 = 0.5 - (\sin^2(\sqrt{x^2 + y^2}) - 0.5)(1 + 0.001(x^2 + y^2)) \quad (-4 < x < 4, -4 < y < 4)$$

(2)

自适应代沟替代的性能(收敛时间和解的质量)都高于完全替代策略. 而在轮盘赌选择方式下, 自适应代沟替代策略的优势是显而易见的.

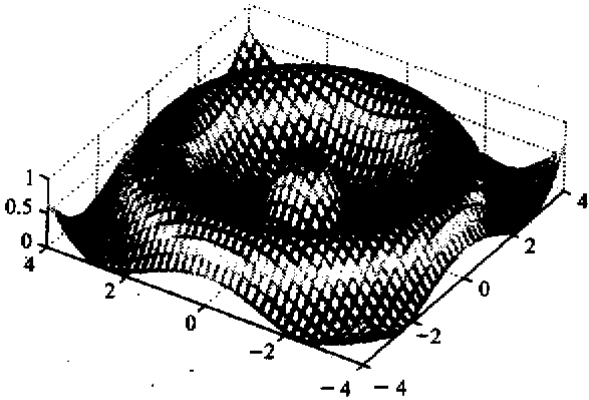


图 2 $F6$ 三维图形 $(-4 < x < 4, -4 < y < 4)$
Fig. 2 Three-dimensioned plot of $F6$
 $(-4 < x < 4, -4 < y < 4)$

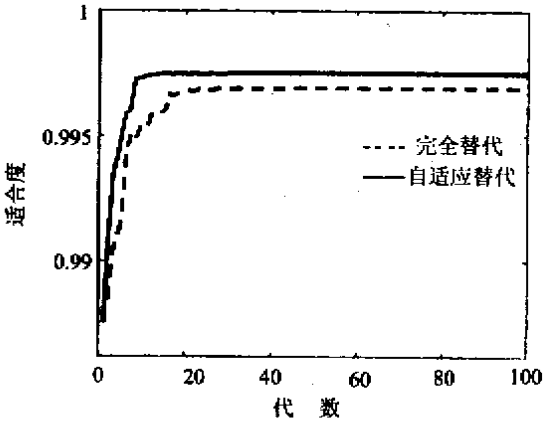


图 3 竞争选择 (Tour=2)
Fig. 3 Tournament selection (Tour=2)

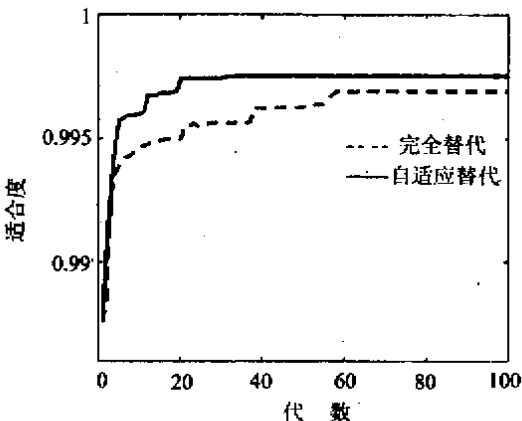


图 4 竞争选择 (Tour=3)
Fig. 4 Tournament selection (Tour=3)

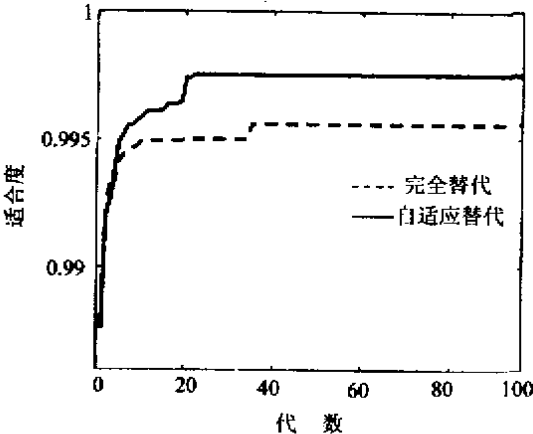


图 5 竞争选择 (Tour=4)
Fig. 5 Tournament selection (Tour=4)

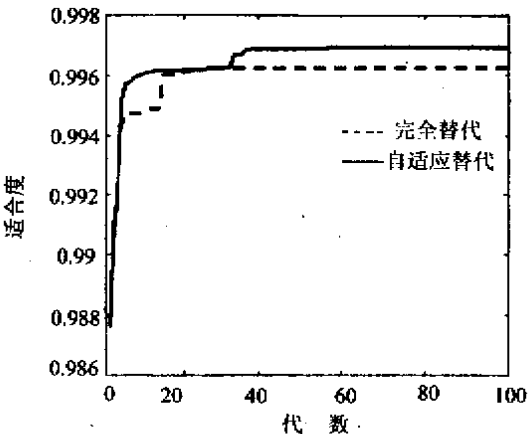


图 6 竞争选择 (Tour=5)
Fig. 6 Tournament selection (Tour=5)

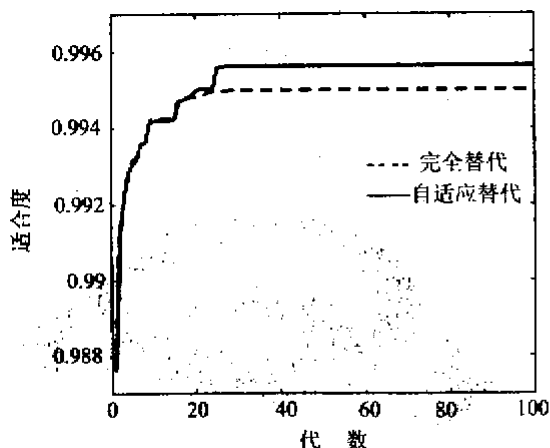


图 7 竞争选择 (Tour=5)

Fig. 7 Tournament selection (Tour=5)

4 结论 (Conclusions)

通过模拟自然界生物进化过程,充分考虑遗传算法优化过程中的动态性,根据遗传算法解的改善程度,综合地比较父、子个体的优劣,我们可以更有效地保留个体中包含的有效信息,从而明显改善遗传算法的性能。由仿真我们可以看出,自适应代沟替代策略与传统的完全替代策略相比,在不损失解的质量的同时,遗传算法具有更高的计算效率。

参考文献 (References)

- [1] Potts J C, Giddens T D and Yadav S B. The development and evaluation of an improved genetic algorithm based on migration and artificial selection [J]. IEEE Trans. on Systems, Man, and Cybernetics, 1994, 24(1): 73-86
- [2] Louis S J. Working from blueprints: evolutionary learning for design [R]. Nevada: Department of Computer Science, University of Nevada, Reno-89557, 1989, 1-13
- [3] Louis S J. Genetic algorithm as a computational tool for design [D]. Indiana: Department of Computer Science, Indiana University, 1993
- [4] Vose M D. Generalizing the notion of schema in genetic algorithms [J]. Artificial Intelligence, 1991, 50(50): 385-395

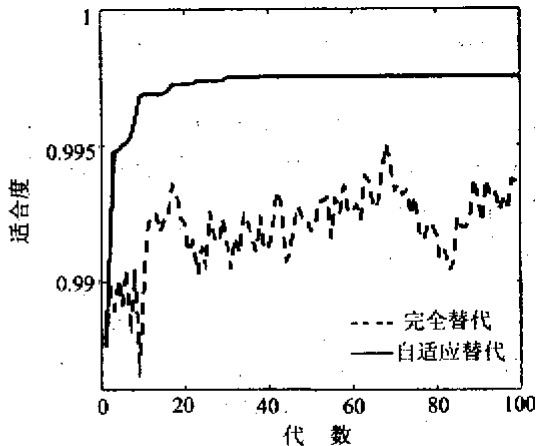


图 8 轮盘赌选择

Fig. 8 Roulette wheel selection

- [5] Potter M A. The design and analysis of a computational model of cooperative coevolution [D]. Virginia: George Mason University, 1997
- [6] Louis S J and Fang Zhao. Domain knowledge for genetic algorithms [R]. Nevada: Department of Computer Science, Mackay School of Mines, University of Nevada, Reno-NV89557, 1989
- [7] Miller B L and Goldberg D E. Genetic algorithms, tournament selection, and the effects of noise [R]. Urbana-Champaign Department of General Engineering, University of Illinois at Urbana-Champaign, IlliGAL Report No. 95006, 1995
- [8] Beasley D, Bull D R and Marton R R. An overview of genetic algorithms, Part 1: fundamentals [J]. University Computing, 1993, 15(2): 58-69
- [9] Liu Yong, Tang Lishan and Chen Yuping. Non-digital Parallel Algorithm-Genetic Algorithm [M]. Beijing: Science Publishing House, 1997, 91 (in Chinese)

本文作者简介

李 兵 1969 年生, 1996 年获哈尔滨工业大学硕士学位, 现为上海交通大学自动化系博士生, 研究方向为进化算法及其在控制工程中的应用, 变结构控制。

谢剑英 1940 年生, 1964 年毕业于上海交通大学自动控制系, 自动控制理论与应用博士生导师, 目前主要从事工业过程优化控制, 和网络技术应用等方面的研究。