

文章编号: 1005-3026(2002)07-0628-04

基于可进化性的快速遗传算法

何大阔, 王福利

(东北大学 信息科学与工程学院, 辽宁 沈阳 110004)

摘 要: 给出三类可进化性较强的个体的定义, 提出充分重视个体的可进化性并对可进化性强的个体加以引导, 来提高遗传算法搜索效率的思想。同时考虑算法全局收敛性能而采用一种并行群体进化结构, 设计了旨在提高遗传算法寻优效率的基于可进化性的快速遗传算法, 仿真结果验证了这种快速遗传算法的良好性能。

关 键 词: 快速遗传算法(EGA); 个体; 可进化性; 加速群体; 并行群体进化结构

中图分类号: TP 13 **文献标识码:** A

遗传算法是目前应用较为广泛的优化方法^[1~5]。它以其所需领域知识较少, 对目标函数无过多要求及在解决多维问题时良好的性能, 在解决一些复杂优化问题时较传统数学优化方法更加有效^[6]。但遗传算法也存在许多问题, 其中搜索速度较慢, 寻优效率较低就是目前研究的重要课题之一。所以, 了解寻优效率问题的原因, 对遗传算法进行改良以提高其寻优速度, 开发快速遗传算法具有重要的理论和现实意义。

传统遗传算法的基本思想是源于自然界生物遗传进化规律, 自然界生物遗传进化是一个漫长的历史过程, 它的最终目的是通过遗传进化使生物群体整体趋于优良同时对时间并无限制。但是, 作为一种优化方法目的是在尽可能少的时间内找到最优解或满意解, 即便是多解问题也只需一部分群体达到优良以找出所有最优解即可, 不要求群体整体趋于优良。优化方法的这种速度性与自然界生物遗传进化的整体性和无时间性较为矛盾。而传统遗传算法的寻优过程是完全模拟生物遗传进化过程, 于是就不可避免地继承了自然界生物遗传进化强调整体、速度较慢的特点。所以, 要提高遗传算法的搜索效率, 必须在遵循物竞天择、适者生存的自然进化法则的同时, 改进具体的进化操作方式, 降低盲目性加强对进化速度贡献较大的因素, 以便达到缩短进化过程的目的。本文分析了遗传算法寻优效率与个体可进化性的关

系, 提出一种基于可进化性的快速遗传算法(EGA)。

1 遗传算法搜索效率与个体可进化性

自然界的生物群体能否逐步趋于优良是源于群体的可进化能力, 而个体的可进化能力是群体可进化的原动力。所谓可进化能力是个体可以改善和优化自己的能力, 也就是个体具有的可进化的性质即可进化性。可进化性在生物学中是指一个生物群体产生比群体中所有现存个体更好的子个体的能力^[7], 正是由于这些具有可进化性的个体产生出更优的个体, 才使整个种群得以延续和进化, 逐渐趋于优良。可见, 个体的可进化性是物种进化的关键, 个体的可进化性越强, 群体进化的质量越高和速度越快。既然生物进化的速度取决于群体中个体的可进化性, 那么如何充分发挥这些可进化性强的个体的作用就是加快进化速度的重点。显然, 作为以生物进化为基础的遗传算法而言, 个体的可进化性及如何充分发挥其作用当然也是影响遗传算法搜索速度的重要因素。所以, 只要充分重视个体的可进化性并对可进化性强的个体加以引导, 加强可进化性强的个体对群体进化的参与, 即可实现提高搜索效率的目的。事实上, 群体中个体的可进化性越强, 遗传算法的寻优速度也就越快。

个体的可进化性反映了其对群体进化质量与

收稿日期: 2001-11-09

基金项目: 教育部骨干教师基金资助项目; 辽宁省自然科学基金资助项目(002013)。

作者简介: 何大阔(1975—), 男, 辽宁沈阳人, 东北大学博士研究生; 王福利(1957—), 男, 辽宁辽阳人, 东北大学教授, 博士生导师。

速度的贡献,如何评价个体的可进化性是强调个体可进化性提高效率的重要技术。本文认为在遗传算法中,个体可进化性的含义较生物学概念要广泛。首先,可进化性体现在个体自身的生存能力上,因为遗传算法是通过定向选择适应度较高的个体来启发寻优的。从一定意义上讲自身具有很高的适应度也是对群体进化的贡献,比如采用最优保持策略可使算法具有全局收敛性^[8]。同时,可进化性还在于个体产生较群体中其他个体更优的个体的能力,由于其产生了现存最好的解则可认为他含有一定的较优解的信息,于是利用它可能缩短寻优进程。另外,产生新个体较自身适应度提高很大的个体也是可进化性强的体现。这类个体可能具有快速优化自身的特质。换言之,无论这类个体自身的适应度如何,其被改进的速度很快可能是因为该个体所在的区域解的分布较为陡峭,重视这类个体可以使算法在较短的时间内搜索完可能快速搜索的区域(即陡峭区域)。可进化性的这一定义不仅有利于局部区域的搜索,同时有利于达到全局最优解。因为问题解的分布较为复杂,一些问题就是由于最优解所在区域较为陡峭使得该区域中的个体适应度相差较大。在最优解附近的个体适应度可能很差但却可能较其他个体含有更多关于最优解的信息,而这类个体在算法寻优过程中会因自身适应度差遭淘汰,从而使算法丧失通过这些个体逼近最优解的机会^[9]。这类个体也是加快收敛速度、提高全局收敛性的重要环节。在遗传算法的寻优过程中,只要充分利用和引导上述几类可进化性强的个体,即可提高效率、改善性能。

2 基于可进化性的快速遗传算法

通常在任何先验知识的情况下,无法事先判断最优解的方向或区域,即无法判断当前的搜索方向是否为全局收敛方向。所以,对于一些复杂多峰问题速度过快可能会降低算法的全局收敛性导致早熟的发生。遗传算法作为一种启发式随机优化方法,必须考虑速度性与全局收敛性之间的协调。本文为使算法在强调快速性的同时不致出现早熟问题,采用基于多群体防止早熟理论^[10]的并行群体进化结构来实现以个体可进化性为基础的快速遗传算法。并行群体进化结构是指算法由主群体及加速群体两部分组成:主群体为常规遗传算法的寻优结构并以常规遗传算法的操作来进行寻优;加速群体由可进化性强的个体组成,对这些个体进行寻优操作。

主群体的主体结构为常规遗传算法,其初始种群均匀分布使种群尽可能充分覆盖解空间,提高初始种群的多样性;采用父代与子代共同参与的两代竞争选择法,充分尊重精英个体性状^[11]。另外,为使选择、交叉及变异相互协调又不失相对独立性,本文对各算子采用并行操作方式即选择、交叉、变异都是分别直接对父代进行操作然后将产生的新个体以一定比例直接进入子代,通过各算子产生的新个体进入子代的比例来调节各算子对遗传算法收敛性能的影响。这样既可以充分尊重父代个体的性状^[12],保持交叉寻优的速度,又可以在降低变异算子破坏力的同时保证种群多样性,从而提高遗传算法的寻优性能。主群体的变异算子采用邻域搜索变异算子,使算法不致因个体大幅度变异的随机性而无法正确判断其可进化性。

加速群体的种群规模较主群体要小,由主群体中的可进化性强的个体组成。根据本文对可进化性的定义,加速群体的初始种群由以下个体构成:主群体第一代产生的最优个体、产生最优个体的父代个体(这类个体由参与交叉的个体产生)及产生的子代相对自身适应度提高很大的父代个体(这类个体均由参与邻域变异的个体产生)。这里定义 $J = \left| \frac{f_{old} - f_{new}}{f_{old}} \right|$ 为个体的自我改进系数,其中 f_{old} 为变异的父代个体适应值, f_{new} 为子代个体适应值,子代个体优于父代个体即 $f_{new} \geq f_{old}$, J 越大个体改进自身的能力越强即可进化性越强。加速群体的种群中三类个体的数量与比例可以调节以使算法处于较好的寻优状态。加速群体的个体进行邻域变异,但只容许有改进的子代存活。同时,加速群体引入加速算子即对有改进的父代与子代个体,在父代个体到子代个体的方向上进行几步线性搜索,线性搜索所得到的改进个体作为加速算子产生的个体,从而通过加速算子进一步加快算法的搜索速度。另外,主群体与加速群体之间进行个体交流:每隔几代就将当前主群体中的可进化性最强的一部分个体放入加速群体中直接取代加速群体中较差的个体;将加速群体中的最优个体引入主群体直接取代主群体中的可进化性较差个体;同时,当主群体中出现其 J 值超过某一较大的阈值即对自身改进较大的个体时,该个体直接进入加速群体代替其中较差的个体。主群体起到保持群体的多样性避免早熟的作用,通过与加速群体的交流不断将可进化性强的个体信息提供给加速群体进行加速,同时为加速群体加速后的个体提供进一步进化的土壤,而加速群体是

通过独立对可进化性强的个体进行操作来加快这些个体的进化速度,然后将这些个体不断反馈给主群体对主群体的搜索进行引导以达到增强快速性的目的。加速群体作为辅助群体以可进化性强的个体为基础对主群体实施作用加速其收敛,两群体各自相对独立地进行寻优计算,通过交流使算法的两个群体联系起来组成整个算法的并行群体进化结构,进而构成基于可进化性的快速遗传算法(EGA)。

3 仿 真

选取如下两个测试问题,在代表收敛速度的平均进化代数和体现全局收敛性的收敛概率中找到全局最优解的次数等方面对本文提出的 EGA 与传统 GA 的寻优性能进行比较。仿真中,EGA 主群规模为 50,加速群体为 20(其中最优个体、产生最优个体的父代个体及 J 值较大的个体比例为 2:1:2),交流个体规模为 5、交流代数为 3,选择、交叉及变异比例为 1:2:2;传统 GA 种群规模为 50、交叉率 0.8、变异率 0.288。两种方法的最大代数为 300 代,定义满意解的偏差数量级为 10^{-4} ,随机运行 20 次,结果如表 1 所示。

表 1 仿真结果
Table 1 simulation results

测试问题	算法	已知最优值	收敛概率 (%)	平均进化代数	最好满意解偏差
F1	GA	0	55	264	1.72×10^{-4}
	EGA	0	100	89	1.14×10^{-15}
F2	GA	1	35	128	0
	EGA	1	100	21	0

$$F1: \min f(x) = 100(x_1^2 - x_2)^2 + (1 - x_1)^2 - 2.048 \leq x_1, x_2 \leq 2.048。$$
$$F2: \max f(x) = 0.5 + \frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{[1.0 + 0.0001(x_1^2 + x_2^2)]^2} - 100 \leq x_1, x_2 \leq 100。$$

其中, $F1$ 为 Rosenbrock 函数,在 (1, 1) 处达到极小值 0,是一个单峰非凸问题; $F2$ 为 Schaffer 函数,在 (0, 0) 处达到极大值 1,是具有大量局部最优点的多峰问题。

上述仿真结果说明,无论是平均进化代数、还是收敛概率,本文的 EGA 均优于传统 GA。从表 1 明显可以看出 EGA 的快速性,同时全局收敛性也很好且精确性较高。

4 结 论

本文以个体可进化性这一影响遗传算法搜索速度的关键因素为基础,通过对常规遗传算法的改进并采用并行群体进化结构提出了旨在提高遗传算法寻优速度的新方法——基于可进化性的快速遗传算法,该方法大大提高了遗传算法的优化效率。

参考文献:

[1] Potts J C, Giddens T D, 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 Cybernetic*, 1994, 24(1): 73—86.

[2] Fumiaki T, Toshiniro N. Banknote recognition by means of optimized masks, neural network, and genetic algorithms[J]. *Engineering Applications of Artificial Intelligence*, 1999, 12(2): 175—184.

[3] Maniezzo V. Genetic evolution of the topology and weight distribution of neural networks[J]. *IEEE Trans on Neural Networks*, 1994, 5(1): 39—53.

[4] Glover F. Genetic algorithm and Tabu search: Hybrid for optimizations[J]. *Computers & Operations Reseach*, 1995, 22(2): 111—134.

[5] 夏靖波, 周铁鹏, 王师. 基于遗传算法神经网络流量测量 [J]. 东北大学学报(自然科学版), 2000, 21(3): 254—256. (Xia J B, Zou T P, Wang S. Neural network flow rate measurement based on genetic algorithms[J]. *Journal of Northeastern University (Natural Science)*, 2000, 21(3): 254—256.)

[6] Fogel D B. An introduction to simulated evolutionary optimization[J]. *IEEE Trans on Neural Networks*, 1994, 5(2): 3—14.

[7] Michalewicz Z. *Genetic algorithms + data structures evolution programs*. 3rd Edition [M]. New York: Springer, 1996.

[8] Rudolph G. Convergence properties of canonical genetic algorithms[J]. *IEEE Trans on Neural Networks*, 1994, 5(1): 96—101.

[9] 恽为民, 席裕庚. 遗传算法的运行机理分析[J]. 控制理论与应用, 1996, 13(3): 289—297. (Yun W M, Xi Y G. The analysis on running mechanism of genetic algorithm [J]. *Control Theory and Applications*, 1996, 13(3): 289—297.)

[10] Potts J C, Giddens T D, Yadav S B. The development and evaluation of an improved genetic algorithm based on migration and artificial selection[J]. *IEEE Trans SMC*, 1994, 24(1): 73—86.

[11] 于海斌, 王浩波, 徐心和. 两代竞争遗传算法及其应用研究 [J]. 信息与控制, 2000, 29(4): 309—313. (Yu H B, Wang H B, Xu X H. A genetic algorithm with competitive selection between adjacent two generations and its applications to tsp[J]. *Information and Control*, 2000, 29(4): 309—313.)

[12] Rudolph G. Convergence properties of canonical genetic algorithms[J]. *IEEE Trans on Neural Networks*, 1994, 5(1): 96—101.

Fast Genetic Algorithm Based on Evolvability

HE Da-kuo, WANG Fu-li
(School of Information & Engineering, Northeastern University, Shenyang 110004, China. Correspondent: HE Da-kuo, E-mail: houy2000@hotmail.com)

Abstract: The searching efficiency of the conventional genetic algorithms was investigated. The searching efficiency depends on the evolvability of individuals. The higher evolvability the individuals have, the higher searching efficiency the genetic algorithm has. three kinds of individuals with high evolvability were defined. the searching efficiency of the genetic algorithm can be improved by attaching importance to the evolvability of the individuals sufficiently and guiding the individuals with high evolvability. When global convergence is taken into account, a paralleled population heuristic structure is adopted, and the fast genetic algorithm based on evolvability is presented to improve the searching efficiency.
Key words: fast genetic algorithm; individuals; evolvability; accelerated population; the paralleled population heuristic structure
(Received November 9, 2001)

待发表文章
摘要预报

液相线半连续铸造 112 合金的二次加热

路贵民, 任栖锋, 崔建忠, 董 杰

用电阻炉恒温加热的方法对液相线半连续铸造法制备的 112 合金进行了二次加热试验。结果表明, 由于铝合金黑度小, 吸热慢, 热导率大, 传热快, 在加热过程中, 试样内部温度分布均匀, 最大温差不超过 3℃, 且加热速度快, 在炉温为 700、740、800℃时, 分别加热 21、18、15 min 后铸锭即可达到半固态状态, 恢复良好的触变性, 得到均匀、细小、球化的组织, 完全适于半固态加工。

轴线直线度误差的理论研究

陈立杰, 张 玉, 张 镭

根据轴线直线度误差的定义, 建立了在直角坐标采样时该项误差的最小二乘评定法数学模型, 该模型的采样数据不需要等角度间隔采样; 并用计算机进行仿真分析。结果表明, 所建立的数学模型具有编程简单; 计算出的轴线直线度误差值与真实情况的误差小于等于 $10^{-6}\mu\text{m}$ 。因此由本模型引入的误差可忽略不计。本文所建立的数学模型为研制形位误差虚拟测量系统提供了理论基础。