

一种可寻得全局最优解的改进变异算子*

周激流 郭晶

(四川大学电子信息学院·成都, 610064)

摘要: 文献[1]分析了典型遗传算法不成熟收敛的起因, 据此提出了一种改进的遗传变异算子, 计算机仿真实验证明改进后的遗传算法能有效地克服不成熟收敛, 进而搜索到全局最优解。

关键词: 遗传算法; 人工智能; 变异

文献标识码: A

A Study on New Mutation Operator That Can Find Global Optimal

ZHOU Jiliu and GUO Jing

(College of Electronic and Information, Sichuan University, Chengdu, 610064, P. R. China)

Abstract: The reference [1] analyses the reason of premature convergence in canonical genetic algorithms. An improved mutation operator based on the reference [1] is proposed in this paper. Computer simulation results show that the improved genetic algorithm can efficiently find global optimal beyond premature convergence.

Key words: genetic algorithm; artificial intelligence; mutation

1 引言 (Introduction)

遗传算法(GA: genetic algorithm)是一种以达尔文自然进化论和孟德尔遗传变异理论为基础的全局随机搜索优化计算技术, 它同时考察多个候选解, 模拟生物遗传进化过程, 淘汰劣质解, 鼓励发展优质解, 逐步提高解群体的质量, 直至收敛, 以此获得(准)最优解。同时, 由于 GA 仅需知道如何根据解来求得其相应的适应度, 而不要求适应度函数满足连续可微等条件, 因而具有广泛的应用。

但是, 典型遗传算法(CGA: canonical genetic algorithm)在不断搜索更优解的过程中将不可避免地使种群的多样性渐趋于零, 常使种群进化收敛于局部极值, 造成不成熟收敛 (premature convergence) 现象, 这是长期困扰 GA, 进而影响了 GA 进一步应用的主要原因。

本文根据文献[1]分析的造成 CGA 不成熟收敛的原因, 提出了一种以改进变异算子为核心的遗传算法, 该算法在发生不成熟收敛时能从当前局部解中跳出来并向其附近的另一个具有更高适应度的局部/全局解收敛, 因而具有能迅速地搜索到全局最优解的能力。

2 不成熟收敛的原因 (The reason of premature convergence)

典型遗传操作由选择、交叉、变异三算子构成。由于通常变异概率较小, 故变异使种群发生显著改变的概率极小, 所以, 遗传操作中选择和交叉算子起主要作用。

文献[1]证明了: 若当前种群各个体具有的共同模式(schema)记作 H , 则交叉算子有能力搜索且仅能搜索 H 包含的所有个体。当种群多样性越大时, H 的阶就越小, 从而它包含更多的潜在解, 交叉算子的搜索能力就越强。但是 H 以概率 1 单调递减到单点型早熟集(从而种群多样性以概率 1 递减到零), 因此, 选择和交叉算子在起搜索作用的同时, 不可避免地使种群多样性渐趋于零, 从而不断缩小自己的搜索范围, 最终导致不成熟收敛。

由此, 我们得知: 选择和交叉算子具有二重性。既可以强迫算法收敛, 使之达到问题的全局最优, 而又可能在搜索过程中发生不成熟收敛而使算法终止于非(全局或局部)最优状态, 并且它们的成熟化效应是造成遗传算法不成熟收敛的起因。

* 基金项目: 国家自然科学基金(69572024)资助项目。

收稿日期: 1999-07-20; 收修改稿日期: 2000-07-13。

3 对典型遗传算法(CGA)的改进(Improvement on CGA)

从理论上讲,只要有变异存在,进化必然是不收敛的,从而也不会发生不成熟收敛.但在进化过程中一直使用较大的变异概率,则迭代过程趋于纯随机搜索,使算法丧失模拟进化的仿生实质^[2].

因此,我们的策略为:首先不考虑不成熟收敛的发生,而是采取各种办法加速收敛过程,以缩短总的搜索时间;当不成熟收敛发生时,加大变异概率,使用特殊的变异算子,使种群从局部极值中跳跃到新的搜索空间;特殊的变异算子将保证变异产生的个体有相当大的概率落入全局最优解附近,从而能向该全局最优解收敛.

3.1 收敛速度的改进(Improvement on convergence speed of CGA)

为了加快收敛速度,本文对适应度进行了标准化处理,以突出较优解的优势,具体方法是:

首先,求得种群适应度值的平均值 \bar{f} 和方差 σ :

$$\bar{f} = \frac{\sum_{i=1}^N f_i}{N}, \sigma = \sqrt{\frac{\sum_{i=1}^N (f_i - \bar{f})^2}{N}} \quad (1)$$

然后,再根据下式将各适应度值标准化:

$$f'_i = \frac{f_i - \bar{f}}{\sigma} \quad (2)$$

若标准化后的适应度最小值小于0,还需将所有适应度值都加上适应度最小值的绝对值,以保证适应度函数值非负.即:

设 $f_{\min} = \min\{f'_i\}, i = 1, 2, \dots, N$, 若 $f_{\min} < 0$, 则:

$$f_{\text{new}} = f'_i + |f_{\min}|, i = 1, 2, \dots, N. \quad (3)$$

3.2 新变异算子(New mutation operator)

通常的二进制编码的变异算子是随机选中一个变异位,将该位上的二进制码取反^[3,4].这种做法借鉴于纯随机搜索思想,在实践中对于帮助种群跳出局部解作用不大.

新变异算子基于这样一个事实:当进行函数优化时,对于多峰函数,在某局部解 S_A 的周围、与 S_A 距离为 r (在大多数情况下, r 不能精确确定,这不要紧,只需知道 r 的可能取值范围 $[r_{\min}, r_{\max}]$ 就行了) 处必有一个适应度更高的局部解 S_B , 当陷入局部解 S_A 时,种群中绝大多数个体落在 S_A 的极小邻域内.如果我们使这些个体均发生变异,且变异的范围为一个半径为 $r_{\min} \leq r \leq r_{\max}$ 的圆环,则有相当大的概率使某个个体变异后会落入局部解 S_B 的邻域,从而促使种群向 S_B 收敛.

以两变量问题为例,其具体的变异方法是:将待变异个体 S_i 解码(decoding)得到 (x, y) , 然后随机产生 r 和 θ , 使 $r_{\min} \leq r \leq r_{\max}, 0 \leq \theta \leq 2\pi$, 则, $\Delta x = r\cos\theta, \Delta y = r\sin\theta$, 再将 $(x + \Delta x, y + \Delta y)$ 编码生成 S'_i , 这就是变异结果.

3.3 新变异算子的物理及数学解释(Physical and mathematical explanation of new mutation operator)

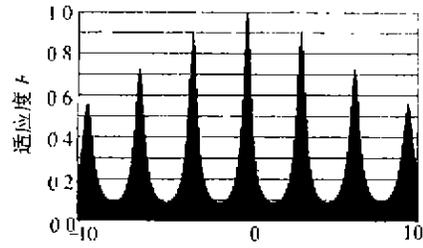
以熟知的 Schaffer 函数极值点求解为例. Schaffer 函数表达式如下:

$$f(x, y) = 0.5 - \frac{\sin^2 \sqrt{x^2 + y^2} - 0.5}{[1 + 0.001(x^2 + y^2)]^2} \quad (4)$$

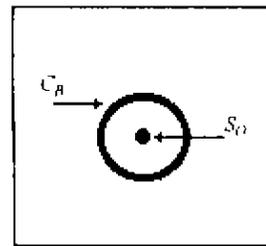
其自变量搜索范围为 $-10 \leq x, y \leq 10$. 该函数在最大值峰(1.0)周围有一圈脊, 取值均为 0.990283. 我们采用如下的适应度变换函数

$$F = \frac{\alpha^*}{\alpha + (1 - f)}, \alpha = 0.1. \quad (5)$$

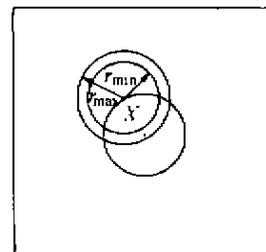
将适应度取值限定在 $[0, 1]$ 范围内, 且最大值(1.00)与次最大值(0.91)之间有较明显的区别. 变换后的适应度函数 F 的图像在 $y = 0$ 处的剖面图如图 1(a)所示, F 的俯视示意图如图 1(b)所示.



(a) 适应度图像在 $y=0$ 处的剖面图



(b) 适应度函数俯视示意图



(c) 变异算子变异范围示意图

图 1 Schaffer 函数适应度剖面、俯视示意图及新变异算子变异范围示意图特性

Fig. 1 Sketch map of Schaffer function characteristic

在图 1(b)中,矩形标示出函数的搜索范围(此面积记作 S_T),正中的实心圆点表示全局最大值的适应度值大于 0.91 的邻域(此面积记作 S_O),围绕正中实心圆点的圆表示全局最大值附近的一圈脊(该圆记作 C_B ,其上各点的适应度值均为 0.91),为简洁起见,此图未画出其余的极值点。

在选择算子和最佳个体保存法的作用下,若种群中有一个以上的个体落入 S_O 范围内,则因为其适应度值高于搜索空间的其它区域,种群就可以向着全局最大值逐渐收敛。然而,因为 S_O 与 S_T 相比实在太小,因而单靠初始种群的随机生成及选择、交叉算子作用,很难有个体落入其中,所以,进化较易收敛于 C_B 上,陷入不成熟收敛。又因为常规的变异算子基于纯随机搜索方法, C_B 上某点的变异结果可能为 S_T 中任一点,其落入 S_O 中的概率(S_O/S_T)相当小,结果,落入非 S_O 区域的个体又逐渐收敛于 C_B 上,因而采用常规的变异算子不能使进化有效地跳出不成熟收敛。而我们提出的变异算子实际上把变异范围由全搜索空间 S_T 缩小到一个由内半径 r_{\min} 和外半径 r_{\max} 界定的圆环(此圆环面积记作 S_C),如图 1(c)所示。图中 X 点表示不成熟收敛种群中的被变异的一个个体。该圆环包含 S_O 区域,对位于该圆环圆心的个体 X 采用我们提出的变异算子进行变异,其结果必然落入 S_C 中,因此,其进一步落入 S_O 的概率为 S_O/S_C 。因为 S_C 远小于 S_T ,所以概率 S_O/S_C 远大于概率 S_O/S_T ,所以,在改进变异算子的作用下,从不成熟收敛区域跳跃到全局最优解邻域的概率比采用常规变异算子时的概率大得多。这就是为什么我们提出的改进变异算子比常规变异算子更能有效地跳出不成熟收敛的基于概率的解释。

3.4 新变异算子中动态半径 R 的调整策略(Adjusting policy of dynamic radius R in new mutation operator)

确定 r_{\min} 和 r_{\max} 本身是一个优化问题。因为 r_{\min} 和 r_{\max} 与适应度函数的数学特性有关,在不知其特性的情况下,很难给出具体的推导规则。当然,也可借鉴 Grefenstette 利用 GA 本身来优化 GA 的控制参数^[5]的方法,即利用 GA 来优化 r_{\min} 和 r_{\max} ,但这样的计算量很大。为此,我们提出如下策略,可以有效地动态调整 R ,使改进后的遗传算法能快速跳出不成熟收敛,进而找到全局最优解。

首先,由用户指定搜索范围,以此定出变异半径的最大允许取值 R_{\max} ,而变异半径的最小允许取值 $R_{\min} = 0$ 。然后,取 $r_{\min} = 0, r_{\max} = r_{\min} + r_{\text{step}}$,其中,

r_{step} 是由用户指定(当然,这些参数也可由程序自动优化,不过这样做将使算法更加复杂)的变异圆环的厚度(外半径与内半径之差)值,常取 0.1。变异在此范围内进行。若种群已收敛于某局部解 S_A ,而在其附近的具有更大适应度的另一个局部解 S_B 已包含在变异圆环内,则经过变异后有一个以上个体落入 S_B 邻域的期望迭代次数可近似表示为

$$n = [1 - (1 - \frac{S_O}{\pi r_{\max}^2 - \pi r_{\min}^2})^N] \alpha. \quad (6)$$

其中, S_O 表示 S_B 邻域中适应度值大于 S_A 的适应度值的区域的面积, r_{\max} 为变异圆环外半径, r_{\min} 为变异圆环内半径, N 为种群规模, α 为容错系数。

容错系数 α 的设置主要是因为变异的随机性,增大 α 可使迭代在不调整 $[r_{\min}, r_{\max}]$ 的情况下多变异几代,但过大的 α 会降低搜索速度。

得到期望变异迭代次数 n 后,若在 n 代内通过变异都未使种群最佳适应度上升,则说明当前的 $[r_{\min}, r_{\max}]$ 不合适,应调整为 $r'_{\min} = r_{\min} + r_{\text{step}}, r'_{\max} = r_{\max} + r_{\text{step}}$,然后重新计算 n 值并再试。当 r_{\max} 达到 R_{\max} 时,则 $[r_{\min}, r_{\max}]$ 的调整方向应反向,即: $r'_{\min} = r_{\min} - r_{\text{step}}, r'_{\max} = r_{\max} - r_{\text{step}}$,直至 r_{\min} 达到 0 后再反向。

3.5 新算法搜索效率分析(Analysis of efficiency of new algorithm)

公式(6)的物理意义为:某个体在圆环区域内跳入 S_O 的概率为 S_O/S_C , S_C 的面积显然为 $\pi r_{\max}^2 - \pi r_{\min}^2$,则该个体不能跳入 S_O 的概率为 $[1 - S_O/(\pi r_{\max}^2 - \pi r_{\min}^2)]$,因此,如群体规模为 N ,则群体不能跳入 S_O 的概率为 $[1 - S_O/(\pi r_{\max}^2 - \pi r_{\min}^2)]^N$,而群体能跳入 S_O 的概率则为 $\{1 - [1 - S_O/(\pi r_{\max}^2 - \pi r_{\min}^2)]^N\}$ 。因此,只要该概率乘上一个大于 N 的常数,则群体一定能跳入 S_O 。

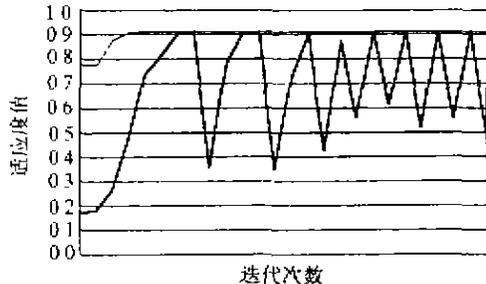
从以上分析,我们可以看出,新算法在每一个圆环区域中的迭代次数是有一定依据且是有限的,而圆环区域的动态调整也是有限次数的,因此,本算法与全局穷尽搜索相比,具有更高的搜索效率。

4 仿真实验结果及讨论(Simulation results and conclusion)

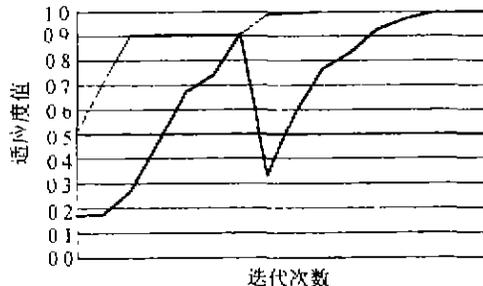
计算机仿真实验中,我们采用二进制编码,种群规模取为 500。

Schaffer 函数采用常规变异算子和新变异算子的迭代过程分别如图 2(a)、2(b)所示。图中粗曲线为种群平均适应度,细曲线为种群当前最佳适应度。

图2(a)表明,采用常规变异算子的CGA算法经过多次变异均不能超越次最优解而向全局最优解搜索.而图2(b)表明,当进化收敛于局部解时,新的变异算子使种群中的部分个体有能力跳跃到全局最优解附近,从而促使种群向全局最优解收敛.



(a) Schaffer函数采用常规变异算子的迭代图像



(b) Schaffer函数采用改进变异算子的迭代图像

图2 Schaffer函数采用两种变异算子迭代过程比较示意图像

Fig 2 Comparative sketch map of iterative process of old and new mutation operator

变异是遗传算法的重要操作之一,它既可以产生种群中原本没有的较优基因,也可以恢复在先前

迭代中被破坏的较优基因,所以它对于加快搜索速度、确保种群不至于陷入局部解方面具有重要意义.从以上实验可见,本文提出的新变异算子及其变异半径的动态调整,至少在多峰函数的优化过程中,能有效地克服不成熟收敛进而搜索到全局最优解.

参考文献(References)

- [1] Xu Gaoben and Gao Yong. Precaution and analysis on characteristic of premature convergence of genetic algorithm [J]. Science in China (Series E), 1996, 26(4): 364-375 (in Chinese)
- [2] Bhandari D, Pal N R and Pal S K. Directed mutation in genetic algorithms [J]. Information Sciences, 1994, 79(3): 251-270
- [3] Whitley D, Starkeather T and Bogart C. Genetic algorithms and neural networks: Optimizing connections and connectivity [J]. Parallel Computing, 1990, 14(3): 347-361
- [4] Whitley D and Hanson T. Optimizing neural networks using faster, more accurate genetic search [A]. Proceedings of the Third International Conference on Genetic Algorithms [C], Beijing, 1989, 391-396
- [5] Grefenstette J J. Optimization of control parameters for genetic algorithms [J]. IEEE Trans. Systems, Man, and Cybernetics, 1986, 16(1): 122-128

本文作者简介

周激流 1963年生,博士,四川大学电子信息学院教授,博士生导师.已在国内外学术刊物上发表论文40多篇,出版专著3部.主要研究方向有:图像图形处理,模式识别与人工智能等.

郭晶 1975年生,硕士研究生.感兴趣的研究方向有:模式识别与人工智能等.

(上接第754页)

1993, 70:289-301

- [11] Fang Chongzhi and Xiao Deyun. Procession Control [M]. Beijing: Tsinghua University Press, 1998 (in Chinese)
- [12] Hwang J N, Kung S Y, Mitrnanjan A and Principe J C. The past, present and future of neural networks for signal processing [J]. IEEE Signal Processing Magazine, 1997, 14(11): 28-48
- [13] Zhang Jiashu and Xiao Xianci. Predicting chaotic time series using recurrent neural networks [J]. Chinese Physics Letters, 2000, 17(2): 88-90
- [14] Zhang Jiashu and Xiao Xianci. Predicting low-dimensional chaotic time series using Volterra adaptive filters [J]. Acta Physic Sinica, 2000, 49(3): 403-406 (in Chinese)
- [15] Zhang Jiashu and Xiao Xianci. Predicting high-dimensional chaotic time series using nonlinear Volterra adaptive filters [J]. Acta Physic Sinica, 2000, 49(7): 1220-1226 (in Chinese)

本文作者简介

伍为根 1963年生,副教授.1984年7月毕业于重庆大学,1989年7月毕业于山西矿业学院,现为电子科技大学博士生.主要从事工业控制技术和混沌控制研究.已发表论文10余篇.

张家树 1965年生,教授.1987年7月本科,1990年1月硕士毕业于电子科技大学电子工程,生物医学工程与仪器专业,1990年1月~1998年9月在西南师范大学物理系从事信息工程专业的教学与科研工作,1993年获总后勤部科技进步二等奖一项,1995年破格晋升为副教授.现在西南交通大学通信与计算机学院工作.主要从事图像分析与处理,非线性自适应信号处理,混沌信号分析设计与处理,计算智能等在通信,雷达中的应用研究.已发表学术论文50余篇.

古天祥 1942年生,教授,博士生导师.主要从事电子测控理论及应用研究.