第9卷第3期 软件学报 Vol. 9, No. 3
1998年3月 JOURNAL OF SOFTWARE Mar. 1998

解决一类遗传算法早熟收敛的混合法及其推广*

徐川育

(杭州商学院 杭州 310035)

摘要 本文分析了变型标准遗传算法VCGA(variants of canonical genetic algorithms)有时会产生早熟收敛的机理,提出了混合法HVCSDA(hybrid VCGA combined with steepest descent approach),并进行了推广.该方法可使最优保存的超级个体时间序列离开早熟收敛状态而继续接近全局最优解.仿真实例表明了本文算法的有效性.在30城市TSP(traveling salesman problem)的基准测试问题中,本文得到了路径为6.82的结果,它好于用新的现代的启发式搜索方法——TABU搜索法得到的6.99的结果.

关键词 VCGA,早熟收敛,微调能力,HVCSDA及其推广.

中图法分类号 TP18

文献[1]提出的VCGA(variants of canonical genetic algorithms)是CGA(canonical genetic algorithms)[2]、超级个体和提升机制三合一的一种遗传算法.本文称VCGA为一类遗传算法.VCGA有时会发生早熟收敛的问题.关于VCGA早熟收敛的研究结果尚未见报道.但是,关于VCGA基础之一的CGA的早熟收敛问题已经得到研究.[3,4]De Jong和Goldberg等研究了适合度函数比例变换(Fitness Scaling)方法.Kreinovich等又作了进一步的改进.但是比例变换函数(Scaling Functions)解决问题的效果有好有差.[4]如何克服遗传算法早熟收敛仍是个难题,文献[5]总结了有待进一步探索的5个主要课题.在前期研究中,内容大多是怎样阻止遗传算法产生早熟收敛.

本文的动机是:一旦在VCGA发生早熟收敛时,如何使最优保存的超级个体时间序列能离开早熟收敛状态而继续接近全局最优解?混合法(Hybrid Approach)为解决上述问题提供了可能性.混合法的概念是在1989年由Goldberg提出来的.[6]它是遗传算法与特定的问题搜索技术相结合而形成的.它既利用了全局性,又利用了特定的问题搜索技术的收敛性.在最优解附近采用各种登山策略的混合法是当前遗传算法集中研究的问题之一.[7]最近问世的有IENS.[8]本文在分析VCGA早熟收敛机理和参考IENS的基础上,提出了VCGA和最陡下降法相结合的混合法HVCSDA(hybrid VCGA combined with steepest descent approach)及其推广.HVCSDA及其推广与VCGA的不同之处是,前者比后者多了一个"最陡下降法搜索过程".HVCSDA及其推广与IENS也有所不同:在方法上,前者是混合的VCGA及其推广,后者是混合的倒位遗传算法.VCGA和倒位遗传算法也是不同的,VCGA基于Markov链,倒位遗传算法基于Holland基本理论.[8]

1 VCGA的早熟收敛问题

1.1 VCGA的概念[1]

VCGA被用于解决静态优化问题:
$$\max\{f(b)\mid b\in IB^l\} \tag{1}$$
其中$b\in IB^l=\{0,1\}^l$,$0<f(b)<\infty$,$f(b)\neq$常数.

设个体基因数为l,个体为二进制链码,群体大小为n,群体状态为i(其中$i=1,\ldots,2^{nl}$),群体的第k个个体记作$\pi_k(i)$(其中$k=1,\ldots,n$),第k个个体的适合度值为$f(\pi_k(i))$.

VCGA是在CGA基础上增加了超级个体(Super-individual)和提升(Upgrade)机制.

超级个体是指往大小为n的群体中添加的一个不参与3种进化操作(选择、交叉和变异)的个体.它被放在群体的最前边,于是新的群体大小为$n+1$.为不失一般性,仍记新群体的状态为i($i=1,\ldots,2^{(n+1)l}$),群体的第k个个体记作$\pi_k(i)$(其中$k=0,\ldots,n$),$\pi_0(i)$为超级个体.

---

* 本文研究得到国家自然科学基金资助.作者徐川育,1945年生,博士,教授,主要研究领域为系统辨识,计算机过程控制和进化算法.

本文通讯联系人:徐川育,杭州310035,杭州商学院数学教研组

本文1996-11-28收到原稿,1997-04-28收到修改稿

提升机制可通过拷贝算子(Copy Operator)的提升矩阵 U 来表示.

$$U=\begin{bmatrix} U_{11} & & & \\ U_{12} & U_{22} & & \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ U_{2^{nl},1} & U_{2^{nl},2} & \cdots\cdots & U_{2^{nl},2^{nl}} \end{bmatrix} \tag{2}$$

U 为 $2^{(n+1)l}\times 2^{(n+1)l}$ 方阵,其中子块矩阵 U_{ab} 是 $2^{nl}\times 2^{nl}$ 方阵. 设 i 为当代群体状态,定义 $b\equiv\arg\max\{f(\pi_k(i))|k=1,\ldots,n\}\in IB^l$. U 的任一元素 u_{ij} 这样被确定:若 $f(\pi_0(i))<f(b)$,则重新定义下代群体状态 $j=(b,\pi_1(i),\ldots,\pi_n(i))$,并且令元素 $u_{ij}=1$;否则令 $u_{ii}=1$.

VCGA 可用马尔可夫链模型来表示,其概率转移矩阵为:

$$P^+=\begin{bmatrix} CMS & & & \\ & CMS & & \\ & & \cdots & \\ & & & CMS \end{bmatrix}\begin{bmatrix} U_{11} & & & \\ U_{12} & U_{22} & & \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \\ U_{2^{nl},1} & U_{2^{nl},2} & \cdots\cdots & U_{2^{nl},2^{nl}} \end{bmatrix} \tag{3}$$

或

$$P^+=\begin{bmatrix} CM & & & \\ & CM & & \\ & & \cdots & \\ & & & CM \end{bmatrix}\begin{bmatrix} U_{11} & & & \\ U_{12} & U_{22} & & \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \\ U_{2^{nl},1} & U_{2^{nl},2} & \cdots\cdots & U_{2^{nl},2^{nl}} \end{bmatrix}\begin{bmatrix} S & & & \\ & S & & \\ & & \cdots & \\ & & & S \end{bmatrix} \tag{4}$$

其中 S,C 和 M 分别为 CGA 的选择、交叉和变异算子的概率转移矩阵,它们的大小均为 $2^{nl}\times 2^{nl}$. 在(3)(4)两式中的3个准对角矩阵中,对角线上分别有 2^l 个大小为 $2^{nl}\times 2^{nl}$ 的方阵 CMS,CM 和 S.

(3)式是先运行选择、交叉和变异算子,然后运行拷贝算子,(4)式是先运行交叉、变异算子,然后运行拷贝、选择算子,这样,VCGA 就出现了两种运行方式,但是,它们的运行结果却是相同的. 其原因解释如下:(4)式的等号右边是3个矩阵的乘积,其中第3个矩阵是对角元素为子块阵 S 的准对角矩阵,它和第2个矩阵 U 的乘积是可交换的. 于是这两种不同运行方式的概率转移矩阵 P^+ 却是相同的.

定义. 设在时刻 t、状态为 i 的群体中最好的个体适合度值 $z_t=\max\{f(\pi_k(i))\}$,则遗传算法收敛于全局最优解当且仅当

$$\lim_{t\to\infty}P\{z_t=f^*\}=1 \tag{5}$$

其中 $f^*=\max\{f(b)|b\in IB^l\}$ 是(1)式的全局最优解.

通过对 CGA 的上述改进,时间序列 $\{z_t\}$ 的意义发生了变化. 原先 CGA 的 $\{z_t\}$ 是每代群体的最好个体适合度值组成的时间序列;而 VCGA 的 $\{z_t\}$ 是最优保存的超级个体适合度值组成的时间序列.

顺便说明,VCGA 与杰出策略遗传算法不同,前者的超级个体不参与进化,后者的杰出个体参与进化.

1.2 VCGA 早熟收敛的机理

VCGA 的超级个体的来源是 CGA 产生的冒尖个体. 研究表明,CGA 存在的问题是:在进化前期,常常会导致一些个体支配了选择过程并使它变慢;在进化后期,群体包含的大多数个体的适合度值接近最优解,以致实际上缺乏竞争,CGA 使选择过程进一步变慢.[4] 本文指出:当 CGA 的进化过程足够慢时,就产生了早熟收敛. 此时 CGA 无冒尖个体提供给 VCGA,造成 VCGA 的最优保存的超级个体时间序列徘徊不前. 于是 VCGA 的早熟收敛也就相继发生.

2 HVCSDA

本文提出解决 VCGA 早熟收敛问题的思路如下:当 VCGA 接近最优解且进化过程变得足够慢时,应及时改变搜索策略,辅之以有效的局部搜索策略,以增强 VCGA 微调能力. 这样,本文自然地提出了 VCGA 和最陡下降法相结合的混合法 HVCSDA.

HVCSDA 由两部分组成:(1) VCGA;(2) 最陡下降法 SDA(steepest descent approach).

HVCSDA 先用 VCGA 进行全局搜索. 当 VCGA 的收敛速度放慢到一定程度时,就中断 VCGA,并转向执行 SDA,以对当前最优保存的超级个体的邻域进行搜索.

SDA 要解决的实际上是一个无约束条件的非线性 0-1 规划问题.

设一个二进制基因链码形式的个体为 x,它的第 j 个基因为 x_j(其中 $j=1,\ldots,l$),适合度函数为 $f(x_1,\ldots,x_l)$,不失一般性,优化问题为

$$\min f(x_1,\ldots,x_l)\quad x_j=0\text{ 或 }1;\quad j=1,\ldots,l \tag{6}$$

由 Taylor 公式，有

$$f(x_1+\triangle x_1,\ldots,x_l+\triangle x_l)=f(x_1,\ldots,x_l)+\sum\nolimits_{j=1}^{l}\partial f/\partial x_j\cdot\triangle x_j+O(\triangle x^2) \tag{7}$$

其中 f 是二次可微函数. 若忽略 $O(\triangle x^2)$ 项，就有

$$f(x_1+\triangle x_1,\ldots,x_l+\triangle x_l)=f(x_1,\ldots,x_l)+\sum\nolimits_{j=1}^{l}\partial f/\partial x_j\cdot\triangle x_j \tag{8}$$

由于适合度函数未必满足二次可微的条件，加之 X 也是离散的，所以无法运用数值方法的 SDA.

克服上述困难的途径是：HVCSDA 中的 SDA 采用搜索法来解决(6)式的问题. 在 SDA 中，相当于(8)式中梯度的算子为 $x_j\pm1$. 其中 $j=1,\ldots,l$. SDA 对当前超级个体的邻域搜索 $\triangle x_j$，使得 $x_j^{new}=x_j\pm\triangle x_j$ 能改善适合度值. SDA 每次调整一个基因，逐个调整. 随着不断地调整，超级个体的适合度值呈下降趋势，直至超级个体的适合度值达到一个稳定状态 $f^*(x)$，此时，$x^*=\arg f^*(x)$ 为所求个体.

为确保 SDA 不致陷入局部极小状态，在每次调用 SDA 前，都要先由 VCGA 全局搜索多次. 但搜索子代的数目过多，又可能花费较多的计算机时. 一个解决的办法是对子代的数目设置一个阀值. 当 VCGA 搜索的子代的数目小于阀值时，它继续运行；当超过阀值时，VCGA 就被中断. HVCSDA 以当前最优保存的超级个体作为 SDA 新的“登山”起点，调用 SDA.

在早熟收敛的情况下，HVCSDA 一般比 VCGA 较为有效. 事实上，此时 HVCSDA 继续保留 VCGA 的概率搜索机制，又鉴于 VCGA 的微调能力较差，HVCSDA 把在局部搜索空间上的微调问题化为基于搜索法解决 0-1 规划的优化问题. 换言之，HVCSDA 要比 VCGA 多一种局部优化机制，从而微调能力得到加强. 在 HVCSDA 进化过程中，由群体的最好个体适合度值组成的时间序列 $\{Z'_t\}$，已不再是原先 VCGA 的最优保存的超级个体适合度值时间序列 $\{Z_t\}$，而是最优保存的超级个体及其邻域最优解所共同组成的新超级个体适合度值时间序列. 在一般情况下，全局最优解附近的 0-1 规划优化问题有解，于是 HVCSDA $\{Z'_t\}$ 的单调性也就好于 VCGA $\{Z_t\}$ 的单调性.

SDA 的框图如图 1 所示. HVCSDA 的框图如图 2 所示.

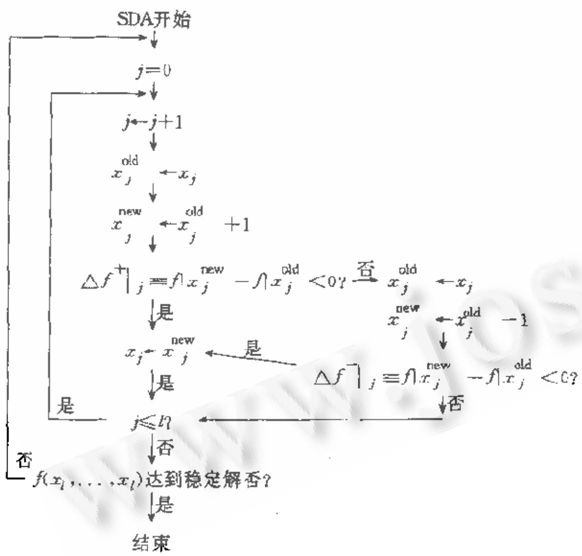

图1　SDA框图

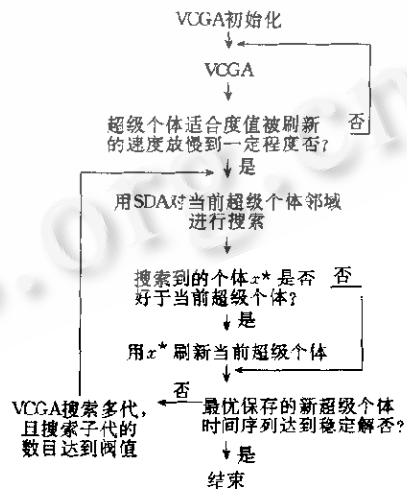

图2　HVCSDA框图

3　HVCSDA 的推广

在多数问题中，问题解的表示对于个体的编码都有限制条件. 如 TSP 的解是包含所有城市在内的、且不允许有城市重复出现的一条闭合路径. 为了防止 VCGA 的遗传算子产生非法解，需要对它们进行修正，以保证产生合法解. 本文将 HVCSDA 推广到这类修正的 VCGA 上. 具体作法是：在 HVCSDA 框图结构不变的前提下，将 HVCSDA 的两个框里的内容进行如下置换.

(1) GA 部分，将 VCGA 置换成修正的 VCGA.

(2) SDA 部分,置换过程详述如下.

在个体编码有约束条件的情况下,求解适合度函数的极值问题是一个有约束条件的非线性整数规划问题.

引入名词"关键值"(Key Value).个体的基因链码被划分成有意义的最短基因片段,$x=(k_1,\ldots,k_m)$,称每个基因片段 k_ν(其中 $\nu=1,\ldots,m$),为一个关键值.如 TSP 中,表示一个城市的一个基因片段就是一个关键值.

在 SDA 中相应于(8)式梯度的算子是 $k_\nu\pm\triangle k_\nu$,$(\nu=1,\ldots,m)$,$\triangle k_\nu$ 的取值受 k_ν 的邻域和约束条件的制约.设符号"$\pm$"为从"+"和"−"中择二为一.当个体 x 从 k_ν 变到 $k_\nu\pm\triangle k_\nu$ 时,对应的适合度值从 $f|_\nu$ 变到 $f|_{\nu'}$.两者的差值

$$\triangle f^{\pm}=f|_\nu-f|_{\nu'} \tag{9}$$

若 $\triangle f^{\pm}>0$,则确认个体变化到 $k_\nu\pm\triangle k_\nu$;否则个体仍保持不变.SDA 采用搜索法在当前超级个体 x 的邻域内寻找满足约束条件的 $\triangle k_\nu$,使 $k_\nu^{new}=k_\nu\pm\triangle k_\nu$(其中 $\nu=1,\ldots,m$)最大程度地改善适合度值.SDA 每次调整一个关键值,逐个调整.随着不断地调整,超级个体的适合度值呈下降趋势,当适合度值达到稳定状态 $f^*(x)$ 时,所求个体为 $x^*=\arg f^*(x)$.

4 仿真实例

(1) 实例 1[9]

$$y=(x_1^2+x_2^2)/2+\cos(20\times\pi\times x_1)\cos(20\times\pi\times x_2)+2 \tag{10}$$

在 $x_1,x_2\in[-10,10]$ 的区域内,函数有 40 000 个局部极小点.当 $x_1=x_2=0$ 时,数值 1 是全局最优解.

把 x_1 和 x_2 编码成一个二进制基因链码个体.取 $P_c=1.0$,$P_m=0.01$,群体大小为 200.在本例中,VCGA 在第 1 200代的时候,出现了早熟收敛.此时,HVCSDA 改变搜索策略,仅用了 3 个 SDA 循环就找到了全局最优解.仿真结果见表 1.

表1 HVCSDA对于实例1的仿真结果

		子代的数目	超级个体适合度值刷新次数	最优保存的超级个体的适合度值	x_1,x_2
HVCSDA运行结果	VCGA部分运行结果	1	1	1.921 225	
		2	2	1.508 018	
		4	3	1.046 785	
		6	4	1.013 085	
		65	5	1.012 840	
		114	6	1.006 430	
		124	7	1.004 830	
		151	8	1.003 752	
		154	9	1.002 673	
		478	10	1.001 880	
		1 200		1.001 880	$x_1=0.000\ 000$ $x_2=-0.000\ 977$
	SDA部分运行结果	SDA循环轮数	获得的稳定适合度值		x_1,x_2
		3	1.000 000		$x_1=0.000\ 000$ $x_2=0.000\ 000$

(2) 实例 2: 30 城市 TSP(traveling salesman problem).[8]

本文用 30 城市 TSP 作为基准测试问题(Benchmark),对修正的 VCGA 和推广的 HVCSDA 进行对照.修正的 VCGA 采用的遗传算子是常见的:选择算子同 VCGA 的一样;交叉算子为 order-crossover[10],取 $P_c=1.0$;变异算子为对换算子(Swap Operator)[11],取 $P_m=0.01$.推广的 HVCSDA 的 SDA 为局部相邻对换算子(Local Adjacent Swap Operator).[11]在推广的 HVCSDA 中,取群体大小为 200.仿真实例 2 的运行情况是:修正的 VCGA 在第 900 代时出现了早熟收敛.此时,推广的 HVCSDA 转向 SDA,用了 128 个循环,达到了 30 城市路径为 6.822 的结果.结果表明:推广的 HVCSDA 要比一类修正的 VCGA 结果来的好,并好于用 TABU[3]得到的 6.99 的结果.计算机仿真结果见表 2.

5 结束语

一类遗传算法 VCGA 有可能发生早熟收敛的问题.HVCSDA 能使最优保存的超级个体时间序列离开早熟收敛状态而继续接近全局最优解.并且 HVCSDA 并可推广到修正的 VCGA 上.HVCSDA 及其推广的实现也比较容易,只需在(修正的)VCGA 之后添上 SDA.总之,HVCSDA 及其推广是解决一类遗传算法(修正的)VCGA 早熟收敛问题的一种有效的方法.

表 2　推广的 HVCSDA 对于 30 城市 TSP 的仿真结果

推广的 HVCSDA 的运行结果					
修正的 VCGA 部分的运行结果			SDA 部分的运行结果		
子代的数目	超级个体适合度值被刷新的次数	最优保存的超级个体的适合度值	SDA 循环轮数	超级个体适合度值被刷新的次数	SDA 搜索到的最好适合度值
1	1	11.092 319	0	0	8.509 102
3	2	10.793 600	0	0	8.450 747
5	3	10.555 657	2	2	8.045 442
8	4	10.058 983	14	3	7.965 511
9	5	9.941 136	15	4	7.886 013
11	6	9.465 212	19	5	7.804 464
31	7	9.451 612	23	6	7.609 268
58	8	9.246 127	30	7	7.505 808
59	9	8.665 101	48	8	7.331 295
152	10	8.509 102	72	9	7.250 476
900		8.509 102	100	11	6.833 110
			128	12	6.821 870
			150		6.821 870

参考文献

1 Rudoulph G C. Convergence analysis of canonical genetic algorithms. IEEE Transactions on Neural Networks, 1994,5(1):96～101

2 Holland J H. Adaptation in natural and artificial systems. Ann Arbor: The University of Michigan Press, 1975

3 徐宗本,高勇.遗传算法过早收敛现象的特征分析及其预防.中国科学(E 辑),1996,26(4):364～375
(Xu Zong-ben, Gao Yong. Characteristic analysis and prevention on premature convergence in genetic algorithms. Science in China (Series E), 1996,26(4):364～375)

4 Kreinovich V, Quintana C, Fuentes O. Genetic algorithms: what fitness scaling is optimal. Cybernetics and Systems, 1993,24(1):9～36

5 孟庆春,贾培发.关于 Genetic 方法的研究及应用现状.清华大学学报(自然科学版),1995,35(5):44～47
(Meng Qing chun, Jia Pei-fa. Actual research status and recent applications of genetic algorithms. Journal of Tsinghua University(Science and Technology), 1995,35(5):44～47)

6 Goldberg D E. Genetic algorithms in search, optimization and machine learning. Reading, MA: Addison Wesley, 1989

7 Editorial. Soft computing: elements of learning systems. International Journal of Systems Science. 1996,27(2):143～144.

8 Lin W, Delgado-Frias J G. Hybrid newton-raphson genetic algorithm for traveling salesman problem. Cybernetics and Systems, 1995,26(5):387～412

9 Percy P C Yip, Yoh-Han Pao. Combinatorial optimization with use of guided evolutionary simulated annealing. IEEE Transactions on Neural Networks, 1995,6(2):290～295

10 Parsons R J, Forrest S, Burrs C. Genetic algorithms, operators, and DNA fragment assembly. Machine Learning, 1995,21:11～33

11 Feng Tse Lin, Cheng Yan-Kao, Ching Chi-Hsu. Applying the genetic approach to simulated annealing in solving some NP-hard problems. IEEE Transactions on System, Man and Cybernetics, 1993,23(6):1752～1767

Hybrid Approach and its Generalization for Solving Premature Convergence of a Class of Genetic Algorithms

XU Chuan-yu

(*Hangzhou Institute of Commerce*　*Hangzhou*　310035)

Abstract　In this paper, the author analyses the mechanism that VCGA (variants of canonical genetic algorithms) may sometimes produce premature convergence, suggests a hybrid approach called HVCSDA (hybrid VCGA combined with steepest descent approach), and generalizes HVCSDA in order to broad its application. The approach can make the time series of super individual best maintained leave the state of premature convergence near to the global optimal solution. Two simulation examples show the efficience of HVCSDA and its generalization. In the benchmark problem of the 30 cities TSP (traveling salesman problem), the lenghth of routing is 6.82 by the HVCSDA's generlization. It is better than one that is 6.99 by new, modern heuristic search method—TABU search.

Key words　VCGA (variants of canonical genetic algorithms), premature convergence, capability of fine tuning, HVCSDA (hybrid VCGA combined with steepest descent approach) and its generalization.

Class number　TP18