

“好”基因对自现及其在遗传算法 协同演化中的应用研究

Emergence of “Good” Gene Pair and Its Application to Co-evolution in Genetic Algorithm

陈秋莲 李陶深
Chen Qiulian Li Taoshen

(广西大学计算机与信息工程学院 南宁 530004)
(College of Computer and Information Engineering, Guangxi University, Nanning, 530004)

摘要 简述协同演化及自现的概念, 提出以计算高适应类染色体基因对作用力, 区别“好”基因对, 寻找演化基因, 实现自现的方法, 并给出自现在协同演化中实现的应用模型, 为自现在协同演化中的实现提供理论基础, 对染色体较长的复杂工程应用问题有重要指导意义。

关键词 协同演化 自现 演化基因 遗传算法

中图法分类号 TP 301.6

Abstract The concepts of co-evolution and emergence of genes in genetic algorithm are explained. The “good” gene pair are selected from chromosomes of high adaptability individuals according to effort of gene pair and used to be evolved genes in a process. A module for emergence of “good” gene pair in co-evolution is developed.

Key words co-evolution, emergence, evolved genes, genetic algorithm

自 70 年代初, 密彻根大学的教授 John. Holland 在用计算机模拟生物系统的基础上完成以达尔文的物竞天择, 适者生存理论为基础的求解优化问题的仿生学类算法——遗传算法(称简单遗传算法, SGA)以来, 进入 90 年代, 遗传算法广泛受到重视, 而其应用的广泛, 也暴露了许多问题: 首先是在复杂工程应用问题上, 由于参数增多, 染色体加长, 会出现收敛速度慢、搜索性能过分依赖于初始种群选择等情况; 其次简单遗传算法不符合这样一种解决问题的规律, 即在不断深入的寻解过程中, 随着解的演化, 问题也在不断地变化发展, 典型例子的如版本升级等。为克服第 2 种缺陷, Mary Lou Maher 等^[2]提出协同演化(简称 co-evolution)的思想。

在考虑克服第一种缺陷时, 受遗传生物学的启发: 发现并利用物种中具有优良特性的基因进行重组, 可以迅速获得优良品种, 我们将同样的原理用于协同演化: 找出个体适应值高的内在原因——“好”基因对, 并把它作为演化基因, 对相关染色体进行基因重构, 则可减少染色体长度, 加快算法收敛。这种自现与协同演化结合的方法, 既消除了第 1 种缺陷, 又克

服了第2种缺陷。

1 协同演化及自现的有关概念

1.1 协同演化

如图1所示,协同演化将整个演化系统分成2个部分:问题空间和解空间,它们都随时间演化。协同演化的基础是简单遗传算法(SGA),不同的是在协同演化中要特别考虑适应函数的描述,使问题空间与解空间具有协同性。实际上协同演化是2个SGA进程的相互作用,其中一个SGA进程作用于问题空间,另一个SGA进程作用于解空间,都如同SGA一样演化,只是在协同演化中,它们还彼此发生联系:一个空间的适应函数取决于另一空间的当前种群。

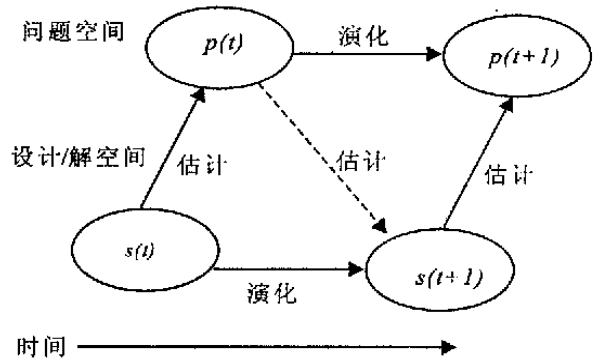


图1 协同演化的问题空间和解空间

另一个SGA进程作用于解空间,都如同SGA一样演化,只是在协同演化中,它们还彼此发生联系:一个空间的适应函数取决于另一空间的当前种群。

1.2 自现

若在分析协同演化的2个空间的个体编码表示时,发现适应值较大的个体其染色体间存在某些共同特征,且经检验证实那些染色体中含有这些共同特征的个体其适应值都比较高,这种使个体适应高的共同特征出现的现象称为自现。本质上说自现是染色体内部各相邻基因(基因对)间的局部相互作用对全局适应性产生的有用结果,它是面向目标的,有什么样的设计目标才有怎样的自现行为,不能在适应函数外定义自现。

1.3 演化基因

演化过程中自现出与高适应个体统计相关的“好”基因对集,若某一基因对对高适应个体的产生起特殊的重要作用则有演化基因产生。演化基因是“好”基因对的另一种表示方法,是基因对相互间的作用力达到一定程度时的表达方式。在下一轮的演化中用产生的演化基因来代替相关基因,则既减少染色体长度又保持个体的高适应特性。

2 自现的实现

实现自现就是分析高适应个体的染色体,计算基因对作用力,找出“好”基因对,鉴别演化基因,进行基因重构,并以已重构的染色体进入下一轮演化循环。

2.1 查找“好”基因对

2.1.1 “好”基因对的评价标准

根据经验给定一个门槛适应值 F_{\min} ,将种群中适应值 $F > F_{\min}$ 的个体归入高适应类,其余归为低适应类(这里适应值已标准化, $0 \leq F \leq 1$)。利用方程(1)和方程(2)来衡量基因对 i 是否是“好”基因对。若高适应类中有超过 S_{\min} 个个体含有基因对 i 或者基因对 i 的作用力 $C_i \geq C_{\min}$,则称该基因对 i 为“好”基因对。

$$S_{\min} = (1 - F_{\min})S_{\text{pop}}, \quad C_{\min} = S_{\min}. \quad (1)$$

$$C_i \geq C_{\min} \text{ 时基因对 } i \text{ 是“好”基因对。} \quad (2)$$

C_{\min} :“好”基因对的最小作用力; S_{\min} :最小门槛个体数; S_{pop} :种群个体总数。

2.1.2 计算基因对的作用力

方法1 根据方程式(3),由适应值除以基因长度求得单基因作用力;每对基因产生两倍单基因作用力,且在个体中不只出现一次,故由方程式(4)求得个体基因对作用力,最后由方

程式(5)求高适应类中所有个体基因对作用力之和,得基因对的总作用力。

$$C_j^{unit} = F_j/L_j. \quad (3)$$

$$C_{ij} = 2k_{ij}C_j^{unit}. \quad (4)$$

$$C_i = \sum_{j=1}^m C_{ij}. \quad (5)$$

式中 m 为高适应类个体个数; F_j 为第 j 个个体的适应值; L_j 为第 j 个个体染色体长度; C_j^{unit} 为第 j 个个体的单个基因作用力; k_{ij} 为基因对 i 在第 j 个个体中出现的次数; C_{ij} 为基因对 i 在第 j 个个体所受的个体基因对作用力; C_i 为基因对 i 在高适应类中所受的总作用力。

方法2 与方法1类似,只是将方程式(5)改为方程式(6),将基因对接其出现次数标准化其作用力,除去经常出现但适应值却很小的基因对。

$$C_i = \sum_{j=1}^m C_{ij} / \sum_{j=1}^m k_{ij}. \quad (6)$$

方法3 如方程式(7)所示,不问基因对出现的次数,只管基因对存在与否,存在则它的作用力是1,否则是0。

$$C_i = \sum_{j=1}^m N_j, \quad (7)$$

$$\begin{cases} N = 1 & \text{高适应个体 } j \text{ 含有基因对 } i; \\ N = 0 & \text{高适应个体 } j \text{ 中不含某基因对 } i. \end{cases}$$

方法3将注意力集中在多少高适应个体含某基因对上,算法比较简单,所以我们倾向于用方法3。

并非说所有“好”基因对都是演化基因,因此还要再进行验证。

2.2 生成演化基因

验证“好”基因对,确定演化基因,步骤如下:

(i) 随机产生 m 个个体 ($m \geq 2$);

(ii) 给定一个门槛适应值 F_{min} ,将种群分成高适应类与低适应类,计算最小门槛个体数 S_{min} 及“好”基因对的最小作用力标准 C_{min} ;

(iii) 用上述方法之一检查所有高适应类个体的基因对,若其作用力大于 C_{min} ,则假定这些基因对就是使个体适应值高的内在特征,是候选演化基因;

(iv) 随机产生含候选演化基因的 k 个个体,检查这些个体的适应值是否大于 F_{min} ,若是,则候选演化基因就是演化基因,否则回到(1),重新再找演化基因。

(v) 若候选演化基因就是演化基因,则该基因对用演化基因表示,给予一个新的标识,防止它被交叉和变异破坏。

3 协同演化中自现实现

自现问题和自现解与协同演化的基本框架结合构成一个完整整体。在协同演化中实现自现涉及3部分:问题空间,适应函数,解空间。图2给出了协同演化与自现的关系。

初始化时,解空间中的个体仅用初始要求(简称 Rinit)来估计其适应值,将高于门槛适应值 (F_{min}) 的个体归到高适应类中。然后按上述方法之一计算高适应类的每一对基因对对高适应性能的作用力,从中选出一些“好”基因对作为演化基因,加入到基本基因池中,替代原基因对,对解空间的各染色体进行基因重构,重构后的染色体进入下一轮协同演化。与此

同时, 选出解空间的当代最佳个体 (简称 CBS), 与 Rinit 结合作为问题空间的适应函数。

类似的, 在同代的问题空间演化中, 用上面得出的 CBS 和 Rinit 结合来估计问题空间的个体适应值, 按适应值分类, 计算高适应类中每对基因对的作用力, 找出演化基因, 重构问题空间的染色体。同时亦选出问题空间的当代最佳个体 (简称 CBR), 与 Rinit 结合作为下一代协同演化中解空间的适应函数, 进行新一轮的协同演化。

问题自现和解自现综合到协同演化中, 随着演化的进一步深入, 基因对中含有演化基因, 从而新演化基因含有旧演化基因, 形成基因的自现分层结构, 染色体长度将越来越短, 迅速提高了算法的收敛速度。

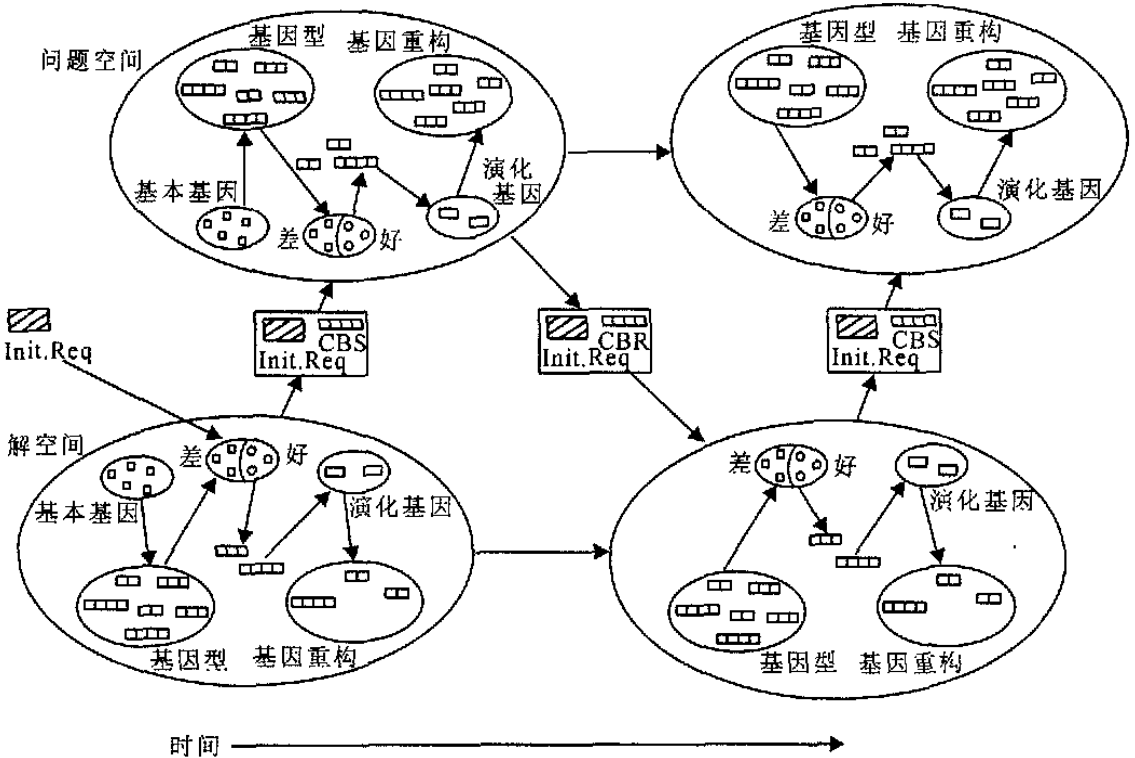


图 2 协同演化与自现

4 结语

本文对协同演化及自现作了探讨, 提出了以计算高适应类染色体基因对作用力, 区别“好”基因对, 寻找演化基因, 实现自现的方法, 并给出了自现在协同演化中实现的应用模型, 为自现在协同演化中的实现提供了理论基础。对染色体较长的复杂工程应用问题有重要指导意义。

参考文献

- 1 潘正君, 康立山, 陈毓屏. 演化计算. 北京: 清华大学出版社, 1994.
- 2 Maher Mary Lou, Poon Josiah. Modelling design exploration as co-evolution. Submitted to the Special Issues of Microcomputers in Civil Engineering on Evolutionary Systems in Design, 1994.
- 3 Poon J, Maher M L. Co-evolution and emergence in design. Preprint Submitted to Elsevier Science, 1996.

(责任编辑: 邓大玉)