

文章编号: 1001-0920(2005)09-0971-10

遗传算法的机理与收敛性研究

于志刚, 宋申民, 段广仁

(哈尔滨工业大学 航天学院, 哈尔滨 150001)

摘要: 采用一种新的基于解空间分解的定量分析方法, 对遗传算法的种群进化过程进行分析, 阐明了选择、交叉和变异操作的寻优机理, 给出了子代种群在解空间上的概率分布情况; 理论上, 证明了遗传算法具备寻找全局最优解的能力, 并给出了具备寻找全局最优解能力的充分必要条件, 即证明了积木块假设的结论是成立的。同时, 建立了二进制编码有限群体的Markov链模型, 计算出在用于静态优化问题的交叉和变异操作下, 种群在解空间上概率分布情况以及收敛到最优解的概率, 并讨论了产生早熟现象和GA-欺骗问题的原因。

关键词: 遗传算法; 解空间的分区; 积木块假设; Markov链; GA-难和GA-易问题; 早熟收敛现象
中图分类号: TP18 **文献标识码:** A

On the Mechanism and Convergence of Genetic Algorithm

YU Zhi-gang, SONG Shen-min, DUAN Guang-ren

(Aerospace College, Harbin Institute of Technology, Harbin 150001, China Correspondent: YU Zhi-gang, E-mail: yzgly@163.com)

Abstract A new approach based on decomposition of solution space according to difference of genotype is presented to analyze quantitatively the evolutionary course of populations of GA. The analysis yields new insight into the properties of the three phases, mutation, crossover, and fitness selection of a genetic algorithm by representing them as acting on the solution space. The probability distribute of populations over solution space can be calculated by this approach. Theoretically, the capability of finding the global optimum is proved, and a necessary and sufficient condition is obtained namely, the conclusion of building block hypothesis is proved. Meanwhile, under crossover or mutation operator applied to static optimization problems, the probability distribution of populations over solution space can be calculated by means of a finite Markov chain model, and the probability of GA converging to the global optimum also can be estimated. Finally, the reasons about GA deceptive problem and premature convergence are discussed.

Key words: Genetic algorithm; Decomposition of solution space; Building block hypothesis; Markov chain; GA-hard and GA-easy problem; Premature convergence

1 引言

遗传算法自诞生以来, 在各个领域中已得到了广泛应用, 但相关的基础理论研究却远落后于算法发展。由于遗传算法(GA)的理论基础不完善^[1], 导致它在应用中暴露出许多问题, 其主要表现是对遗传算法随机搜索机理的研究不够深入, 原因在于缺少有效的分析方法和工具^[2-6]。在搜索机理的研究

方面, Holland 提出了模式定理^[7], 以及积木块假设和隐并行性的分析。模式定理和积木块假设分别说明遗传算法具有寻优的可能性和寻找到最优样本的能力, 但该结论并未得到证明。在解决建筑积木块假设成立与否的问题上, 目前主要集中于对所谓遗传算法的欺骗函数研究上。应用性研究仅限于特定的实现^[8], 对其在大量应用中所表现出的全局优化能

收稿日期: 2004-10-25; 修回日期: 2005-02-28

作者简介: 于志刚(1972—), 男, 黑龙江哈尔滨人, 博士生, 从事智能控制和软计算的研究; 宋申民(1968—), 男, 山东滕州人, 副教授, 博士, 从事鲁棒控制、智能优化与智能控制的研究。

力未做出理论证明, 并尚未找到一个恰当的度量和论证方法精确刻画 GA 在不同实现下的寻优过程

针对这些问题, 提出了一种新的分析遗传算法的进化过程和收敛性的方法, 可精确地刻画种群的进化行为, 理解 GA 是如何解决问题的. 该方法首先对解空间进行分区, 将整个解空间分成不同的子空间, 这样避免了因解空间过大给分析带来的复杂性. 在子空间上, 容易阐明选择、交叉和变异操作的寻优本质, 并严格地证明了遗传算法具有寻找最优解的能力, 给出了遗传算法寻找最优解的充分必要条件, 即证明了积木块假设的结论成立. 进而分析了产生 GA 难于求解的(所谓 GA-难问题)和易于求解(所谓 GA-易问题)的原因

Rudolph 在文献[9]中证明了标准遗传算法在概率意义下不收敛, 采用精英保留策略可保证算法在概率意义下的收敛. Fogel 证明了, 当不使用变异算子时, 所生成的 Markov 链将存在吸收态^[10]. Goldberg 在文献[11]中采用 Markov 链分析方法对一种只有变异和复制的遗传算法的收敛性进行了分析. 大致地说, 遗传算法过早收敛现象发生在算法种群演化到一种非全局最优状态, 它使得算法的进一步迭代已不能产生更佳可行解. 还有大量文献研究了遗传算法的收敛性^[12~14], 并提出一些克服方法. 但本文提出了一种新的收敛性定量分析方法, 它根据遗传算法的寻优机理, 在解空间分区和所建立的二进制编码有限群体 Markov 链模型基础上进行分析的. 与文献[9, 11, 15, 16]中方法的最大区别是当用于静态优化问题的交叉和变异操作时, 能够计算出种群在解空间上的概率分布情况和种群收敛到某个吸收态的概率等, 并得到过早收敛现象的起因、表现特征及预防措施, 因此更加深刻地揭示了遗传算法的寻优本质及其特点. 该方法具有一般性, 可用于进一步分析和证明已有文献中的一些关于 GA 收敛性的结论和产生早熟收敛现象的本质原因

2 遗传算法机理的分析

遗传算法是由 Holland 教授于 1969 年提出, 并经 De Jong 和 Goldberg 等归纳总结所形成的一类模拟进化算法, 其运行过程为一个典型的迭代过程. 算法的准备工作包括:

- 1) 选择编码策略, 将参数集合 X 映射到个体空间 $H = \{0, 1\}^L$;
- 2) 定义适应度函数 $f(x_i)$;
- 3) 确定遗传策略, 选择操作、交叉操作和变异操作的方法

遗传算法的基本流程如下:

- 1) 初始化群体大小 N , 交叉概率 p_c , 变异概率

p_m 等参数, 随机生成初始种群 $X(0)$;

- 2) 计算种群中个体的适应度;

3) 按照遗传策略, 对第 t 代种群 $X(t)$ 进行选择操作、交叉操作和变异操作, 形成下一代的种群 $X(t+1)$;

4) 判断算法是否满足停止准则, 如果不满足, 则返回到步骤 2);

5) 如果满足, 则输出种群中的最大适应度值的个体作为最优解 x^* , 终止计算

以下分析中, 在遗传算法中采用一点交叉和均匀变异操作. 其他改进算法也可得到与之类似的结论

2.1 遗传操作的机理分析

设 $H = \{0, 1\}^L$ 为个体空间, H^N 为 N 个个体种群空间. 若设个体 $x_i = (x_{i0}, x_{i1}, \dots, x_{iL-1}) \in H$, 则 $X(t) = (x_0(t), x_1(t), \dots, x_N(t))$ 或简记为 $X(t) = (x_0, x_1, \dots, x_N)$, $t = 0$ 表示第 t 代种群

定义 1 个体 $x_i = (x_{i0}, x_{i1}, \dots, x_{iL-1})$ 中占有固定位置的基本遗传单位称为基因, 染色体中其余的部分称为该基因的互补基因, 两个染色体中位置相同的基因为等位基因, 基因的类型称为基因型

例如, x_{i1} 为一个基因, 个体 $x_i = (x_{i0}, x_{i1}, \dots, x_{iL-1})$ 处于高 k 位的基因 $(x_{i0}, x_{i1}, \dots, x_{i,k-1})$ 称为高位基因, 其余部分为与此高位基因对应的低 $L-k$ 位基因 $(x_{ik}, x_{i,k+2}, \dots, x_{i,L-1})$, 即高位基因的互补基因

高位基因在个体编码中的权值较大, 它决定个体的在解空间全局位置, 是个体全局分布性的基因型; 低位基因在个体编码中的权值较小, 它决定了个体在高位基因附近的局部位置. 交叉操作实际上是等位基因互相交换位置. 交叉和变异操作可分别表示为

$$x_a = C(x_m, x_n), x_a \in X(t+1), \\ x_m, x_n \in X(t); \quad (1)$$

$$x_b = M(x_m), x_b \in X(t+1), \\ x_m \in X(t). \quad (2)$$

定义 2 设编码长度为 L 的解空间(个体空间) $H = \{x_0, x_1, \dots, x_{2^L-1}\}$, 令 $V_0^L = x_0, V_1^L = x_1, \dots, V_{2^L-1}^L = x_{2^L-1}$. 将解空间 H 按个体的高位基因长度 j ($0 \leq j \leq L-1$) 分解成 2^j 个不同的子区间, $V_i^j, i = \{0, 1, \dots, 2^j-1\}$, 且 V_i^j 满足

$$V_i^j = \{V_k^{j+1} : k = 2i, 2i+1\}, \quad (3)$$

该分区方法称为 j 级分区, 每个子区间称为一个分区

在 j 级子空间 $V_0^j, V_1^j, \dots, V_{2^j-1}^j$ 中, 与其中的 V_i^j

具有相同的低 $s (s > L)$ 位基因的分区集合记为 $U_{i,s}^j$, 且满足 $U_{i,s}^j = \{V_{i,r}^j \mid k_i\}$, 其中

$$k_i = \{n = 2^{L-s}m + i, 0 \leq n < 2^j - 1, 0 \leq i < 2^{L-s}\}, \quad (4)$$

即

$$\begin{aligned} k_0 &= \{n = 2^{s-(L-j)}m\}, \\ k_1 &= \{n = 2^{s-(L-j)}m + 1\}, \\ &\vdots \\ k_{2^{L-s}-1} &= \{n = 2^{s-(L-j)}m + 2^{L-s} - 1\}. \end{aligned} \quad (5)$$

以编码长度为 $L = 4$ 的解空间为例, 其分区情况如表 1 所示

表 1 编码长度为 $L = 4$ 的解空间 H 的分区

$V_0^0 = H$															
V_0^1					V_1^1										
V_0^2		V_1^2		V_2^2		V_3^2		V_4^2		V_5^2		V_6^2		V_7^2	
V_0^3	V_1^3	V_2^3	V_3^3	V_4^3	V_5^3	V_6^3	V_7^3	V_8^3	V_9^3	V_{10}^3	V_{11}^3	V_{12}^3	V_{13}^3	V_{14}^3	V_{15}^3
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

其中: $V_0^3 = V_0^4, V_1^3 = V_2^4, V_2^3 = V_3^4, \dots, V_7^3 = V_{14}^4, V_{15}^4$, 与 V_0^3 具有相同低 s 位基因分区的集合分别为 $U_{0,2}^3$ 和 $U_{0,3}^3$. 因为 $L = 4, j = 3$, 所以当 $s = 2$ 时, $k_0 = \{n = 2^{s-(L-j)}m\} = \{0, 2, 4, 6\}, U_{0,2}^3 = \{V_0^3, V_2^3, V_4^3, V_6^3\}$; 当 $s = 3$ 时, $k_0 = \{n = 2^{s-(L-j)}m\} = \{0, 4\}, U_{0,3}^3 = \{V_0^3, V_4^3\}$.

对解空间分区进行如下说明:

1) 通过将解空间分成不同的子空间, 使种群中的个体按遗传基因的性质进行分类. 不同的分区有不同的遗传特性, 对产生后代的作用也不同.

2) 假设优化问题的最优解附近某个邻域 δ 称为最优解区间. 最优解区间 δ 内的每一个解均满足优化问题的精度或性能要求, 因此, 利用遗传算法时, 关键是要得到最优解附近某个邻域 δ 内的至少一个解. 邻域 δ 的大小根据具体问题和求解的精度要求确定.

3) 当 δ 充分小时, 即 δ 内仅包含一个个体, $\delta = \{x_i(t) \mid X(t)\}$.

遗传算法本质上是一个具有导向的随机搜索技术. 其导向机制趋向于在适应度高的父代个体“家族”的邻域内进行搜索. 交叉操作实现了遗传的功能, 即将二进制编码的个体看作一个染色体, 单点交叉时, 保留一部分的原来的基因的性状, 基因的性状可描述为: 高位基因在个体编码中的权值大, 限定了

个体的全局位置; 低位基因的权值小, 仅表示个体在对应高位基因分区中的局部位置. 定理 1 对此进行了定量描述.

定理 1 设在种群 $X(t) = (x_0, x_1, \dots, x_N)$ 中任意选中两个个体 x_i 和 x_j 作为交叉操作的父代个体, p_c 为交叉概率, L 为个体编码长度, 进行交叉操作后其子代个体可能保留在 x_i 或 x_j 所在分区内的概率为 $p_{\text{cross}} = 1 - \left[1 - \frac{j}{L-1}\right] p_c$.

根据一点交叉操作的结果很容易证明定理 1 的结论. 如设 $L = 10, j = 3, p_c = 0.6$ 时, $p_{\text{cross}} = 1 - \left[1 - \frac{3}{10-1}\right] \times 0.6 = 0.6$. 这说明单纯交叉算子作用是使子代出现在父代附近的概率较大, 也就是说交叉的局部搜索作用较强, 因此可将单纯的交叉操作看作局部搜索算子. 但事实上, 遗传算法总是存在选择操作, 它导致交叉操作后, 子代个体易集中出现于有较大选择优势的基因所组合的分区内. 由定理 2 可知, 选择操作在交叉操作时, 对决定种群的进化方向有重要作用.

定理 2 (基因重组定理) 设任意种群 $X(t) = (x_0, x_1, \dots, x_N)$, 通过一次选择、交叉操作, 种群搜索到解空间的任一分区 V_i^j 的概率如下:

$$p_{\text{scross}}\{V_i^j\} = \frac{1}{L-1} \sum_{k=1}^j (p_s\{V_{[i/2^{k-j}]}^k\} \times p_s\{U_{i,L-k}^j\}) p_c + p_s\{V_i^j\} \left[1 - \frac{j}{L-1} p_c\right]. \quad (6)$$

其中: p_c 为交叉概率, $p_s\{V_{[i/2^{k-j}]}^k\}$ 表示集合 $V_{[i/2^{k-j}]}^k$ 中个体被选择的概率, $[i/2^{k-j}]$ 表示取 $i/2^{k-j}$ 的整数部分, $p_s\{U_{i,L-k}^j\}$ 表示集合 $U_{i,L-k}^j$ 中个体被选择的概率.

证明

$$\begin{aligned} p_{\text{scross}}\{V_i^j\} &= p\{x_a \in V_i^j; x_a = C(x_m, x_n), (x_a \in X(t+1), x_m, x_n \in X(t))\} \\ &= p_s\{V_i^j\} (1 - p_c) + \frac{L-j-1}{L-1} p_s\{V_i^j\} p_c + \frac{1}{L-1} \sum_{k=1}^j (p_s\{V_{[i/2^{k-j}]}^k\} \times p_s\{U_{i,L-k}^j\}) p_c \\ &= \frac{1}{L-1} \sum_{k=1}^j (p_s\{V_{[i/2^{k-j}]}^k\} \times p_s\{U_{i,L-k}^j\}) p_c + p_s\{V_i^j\} \left[1 - \frac{j}{L-1} p_c\right]. \end{aligned}$$

由定理 2 可知, 某代种群搜索到最优解所在 V_i^j 区间内概率 $p_{\text{scross}}\{V_i^j\}$ 的大小与下面 3 个因素有关:

- 1) 与 V_i^j 中个体具有相同高 k 位基因的所有个体组成的集合被选择的概率;
- 2) 与 V_i^j 中个体具有相同低 $L - k$ 位基因的所

有个体组成的集合被选择的概率;

3) 交叉概率

定义 3 设在 j 级分区下, 种群 X(t) 中有个体出现的分区集合 A(t) = {V_i^j: x_m(t) ∈ V_i^j}, 子代种群 X(t+1) 中有个体出现的分区集合 A(t+1) = {V_i^j: x_m(t+1) ∈ V_i^j}, 则称

λ_cross = p_cross{V_i^j} / p_cross{V_i^j} (7)

为交叉搜索率

1) 当 λ_cross > 1 时, 种群的交叉操作表现为全局搜索性, λ_cross 越大全局搜索性越强;

2) 当 λ_cross = 1 时, 种群的交叉操作表现为局部搜索性, λ_cross 越小局部搜索性越强;

3) 当 λ_cross = 0 时, 种群的交叉操作完全收敛

选择和交叉操作有时表现出全局搜索性, 有时表现出局部搜索性, 它可用 λ_cross 的大小衡量. 如果某时刻的种群在解空间上接近均匀分布, 种群的基因类型较多, 且选择作用较小时, 经过交叉操作后, 由定理 2 可知种群的子代个体在解空间的各个分区内出现的概率几乎相等. 此时, 相当于在解空间上进行全局搜索, 这种情况一般发生在种群的进化初期. 但当种群过分集中于某个区域内, 种群中的基因类型比较单一或因选择较强作用时, 都可能使得种群在交叉操作后, 子代个体集中于解空间的某个局部区域, 此时交叉操作体现出很强的局部搜索性, 这种情况一般发生在种群进化的后期.

推论 1 设在 j 级分区下, 对于种群 X(t) = (x_0, x_1, ..., x_N) 有:

1) 若 ∀ x_i 满足 x_i ∈ V_i^j, 则 p_cross{V_i^j} = 1;

2) 若 {x_i: x_i ∈ V_i^j} = ∅, 则 p_cross{V_i^j: k = i} = 0

证明 1) 由已知得, p_s{V_i^j} = 1; 又当 1 ≤ k ≤ j 时, p_s{V_{[i/2^k - j]}^j} = p_s{U_{i,L-k}^j} = 1, 于是, 根据定理 2, 有

p_cross{V_i^j} = 1 / (L - 1) * sum_{k=1}^j p_c + [1 - 1 / (L - 1) * p_c] = 1

2) 当 1 ≤ k ≤ j 时, 由于 {x_i: x_i ∈ V_i^j} = ∅, 则 p_s{U_{i,L-k}^j} = 0, 当 j + 1 ≤ k ≤ L 时, P_s(V_{[1/2^k - j]}^j) = P_s{∅} = 0, 又因为 p_s{V_i^j} = P_s{∅} = 0, 所以 p_cross{V_i^j: k = i} = 0

推论 1 说明了, 如果种群全部集中于某个分区 V_i^j 内, 那么经交叉操作, 后代个体全部集中在该分区内, 也就是说不可能在该分区以外的同级别的分区内有个体出现. 交叉操作与种群基因多样性的关系密切, 交叉操作在解空间搜索范围只限于当代种群所包含基因的可能组合区域.

定理 3 (遗传定理) 通过选择和交叉操作的作用, 遗传算法能够寻找到全局最优解的充分必要条件是种群中至少包含一对全局最优解的高位基因及其互补基因

证明 1) 充分性: 设全局最优解 x* ∈ V_i^j, 种群中包含全局最优解基因及其互补基因, 那么, 包含全局最优解的高 k 位基因个体被选择概率 p_s{V_{[i/2^k - j]}^j} > 0, 同时, 包含低 L - k 位基因个体被选择的概率 p_s{U_{i,L-k}^j} > 0, 又因为 p_s{V_i^j} > 0, 根据定理 2, 遗传算法通过选择和交叉操作能够寻找到最优解 x* 的概率 p_cross{V_i^j} > 0. 所以, 欲使算法通过选择和交叉能够寻找到最优解, 必须满足: 种群中至少包含一对全局最优解基因及其互补基因

2) 必要性: 将充分性的证明进行反向推导, 可知在父代种群中必然存在最优解高位及其对应的低位的基因

遗传定理说明了如果种群中包含全局最优解的高位基因及其互补基因, 那么, 通过选择、交叉操作的作用, 遗传算法能够找到最优解. 证明了积木块假设是成立的, 同时给出了积木块假设成立的充分条件是种群中至少包含一个全局最优解的基因及其互补基因. 如果初始种群中不包含最优解, 那么在进化过程中, 在最优解的基因被全部淘汰掉之前, 通过交叉操作必然可能产生最优解, 产生最优解的概率可通过定理 2 求得; 否则, 交叉操作将失效, 产生最优解或最优解的基因只能依靠变异操作完成. 如果最优解的基因被选择的概率较小, 则最优解的基因将在进化过程中很容易被淘汰; 如果变异的作用较小, 则此时的算法可能发生未成熟收敛现象.

定理 4 (变异的作用) 设任意种群 X(t) = (x_0, x_1, ..., x_N), 父代个体 x_m ∈ V_i^j, 通过变异作用, 子代个体 x_b = M(x_m), x_b ∈ X(t+1), x_m ∈ X(t), 则 x_b 出现在解空间的任意一个分区 V_i^j 内的概率为

p_mut{V_i^j} = sum_{i=j}^L [1/2 + (delta_{m,i}^{x_{bi}} - 1/2) * (1 - 2p_m)] (8)

其中

delta_{m,i}^{x_{bi}} = { 0, x_{mi} = x_{bi}; 1, x_{mi} != x_{bi}

证明 设 x_m 个体的每个分量 x_{mi} 的变异是互相独立的, 则有

p{x_{bi} = M(x_{mi})} = { 1 - p_m, x_{bi} = x_{mi}; p_m, x_{bi} != x_{mi}

令

delta_{m,i}^{x_{bi}} = { 0, x_{mi} = x_{bi}; 1, x_{mi} != x_{bi}

于是

$$p\{x_{bi} = M(x_{mi})\} = p_m + \delta_{m^i, x_{bi}}(1 - 2p_m) = \frac{1}{2} + (\delta_{m^i, x_{bi}} - \frac{1}{2})(1 - 2p_m),$$

所以

$$p_{mut}\{V_i^j\} = p\{x_b = M(x_m), x_b \in X(t+1), x_m \in X(t)\} = \prod_{i=j}^L [\frac{1}{2} + (\delta_{m^i, x_{bi}} - \frac{1}{2})(1 - 2p_m)]$$

推论 2 种群 $X(t)$ 的个体 $x_m \in V_i^j$ 的子代个体 x_b 出现在其父代个体 x_m 所在分区 V_i^j 内的概率为 $p_{mu} = (1 - p_m)^{L-j}$.

如当 $p_m = 0.001, L = 10, j = 3$ 时, $p_{mu} = (1 - 0.001)^{10-3} = 0.993$. 定理 4 及其推论 2 说明了变异的作用在于搜索范围是整个解空间, 变异搜索作用与变异的概率关系密切. 当变异概率较大时, 全局搜索性能较强; 当变异概率较小时, 局部搜索性能较强, 所以可认为变异的搜索既有全局搜索性, 又有局部搜索性.

综上所述, 证明交叉操作是一种有条件的全局搜索算子, 但经常表现出局部搜索性; 变异算子完全是一种全局搜索算子, 其作用更为重要. 遗传算法是一种有导向的随机搜索方法, 其导向作用是指交叉操作实际上是一种启发式的随机搜索方法, 还有变异操作的准随机搜索, 因为根据定理 4 可知, 它不是完全在解空间上的随机搜索, 高位基因和低位基因变异的结果是不同的.

2.2 分析在选择、交叉和变异的联合作用下的结果

定理 5 设任意种群 $X(t) = (x_0, x_1, \dots, x_N)$, 通过 r 次选择、交叉和 s 次变异操作, 种群搜索到解空间的任一分区 V_i^j 的概率为

$$p_{gen}\{V_i^j\} = 1 - (1 - p_{cross}\{V_i^j\})^r (1 - p_{mut}\{V_i^j\})^s \quad (9)$$

证明 根据定理 2, 一次交叉后 V_i^j 分区内有个体出现的概率为 $p_{cross}\{V_i^j\}$, 则经过 r 次的交叉后的出现概率为 $p_1 = 1 - (1 - p_{cross}\{V_i^j\})^r$; 再根据定理 4, 经过一次变异操作后在解空间的 V_i^j 分区范围内有个体出现的概率为 $p_{mut}\{V_i^j\}$, 则经过 s 次的变异操作后, 个体出现的概率为 $p_2 = 1 - (1 - p_{mut}\{V_i^j\})^s$, 于是

$$p_{gen}\{V_i^j\} = p_1 + p_2 - p_1 p_2 = 1 - (1 - p_{cross}\{V_i^j\})^r + 1 - (1 - p_{mut}\{V_i^j\})^s - [1 - (1 - p_{cross}\{V_i^j\})^r][1 - (1 - p_{mut}\{V_i^j\})^s] = 1 - (1 - p_{cross}\{V_i^j\})^r (1 - p_{mut}\{V_i^j\})^s$$

当 $j = L$ 时, $p_{gen}\{V_i^L\}$ 描述了在种群中的每个个体在整个解空间上的进化情况.

定义 4 设在 j 级分区下, 种群 $X(t)$ 中个体出现的分区集合 $A(t) = \{V_i^j: x_m(t) \in V_i^j\}$, 子代 $X(t+1)$ 中个体出现的分区集合 $A(t+1) = \{V_i^j: x_m(t+1) \in V_i^j\}$, 则称

$$\lambda_{gen} = \frac{p_{gen}\{V_i^j\}}{p_{gen}\{V_i^j\}} \quad (10)$$

为交叉变异搜索率. 并且有:

- 1) 当 $\lambda_{gen} > 1$ 时, 种群表现为全局搜索性, λ_{gen} 越大全局搜索性越强;
- 2) 当 $\lambda_{gen} = 1$ 时, 种群表现为局部搜索性, λ_{gen} 越小局部搜索性越强;
- 3) 当 $\lambda_{gen} = 0$ 时, 种群完全收敛.

推论 3 种群在选择、交叉和变异操作的联合作用下, 遗传算法必然能够寻找到最优解.

证明 设全局最优解所在的分区为 V_i^j , 第 t 代种群中包含全局最优解的基因及其互补基因, 其基因及其互补基因被选择的概率分别为 $p_s\{V_i^j\} > 0$ 和 $p_s\{U_{i,L-k}^j\} > 0$, 根据定理 2, 有 $p_{cross}\{V_i^j\} > 0$; 又因为 $p_{mut}\{V_i^j\} > 0$, 所以根据定理 4, 有 $0 < p_{gen}\{V_i^j\} < 1$. 设种群进化 T 代后, 产生最优解的概率为 p_T^* , 则

$$p_T^* = 1 - \prod_{i=1}^T (1 - p_{gen}\{V_i^j\}),$$

所以

$$\lim_{T \rightarrow \infty} p_T^* = 1 - \lim_{T \rightarrow \infty} \prod_{i=1}^T (1 - p_{gen}\{V_i^j\}) = 1$$

即当进化的世代数趋于无穷大时, 遗传算法必然收敛到全局最优解.

2.3 产生 GA-难问题和 GA-易问题的原因

定理 2 和定理 4 分别对交叉和变异操作的机理进行了定量分析, 定理 5 及其推论定量分析并且证明了遗传算法具有寻求最优解的能力. 可利用这些结论讨论产生 GA-难问题和 GA-易问题的原因.

2.3.1 GA-难问题的原因

各种研究结果表明, GA-难问题一般是在最优解周围一些适应度较差的解. 根据定理 2 可知, 这种情况很难发挥交叉操作的局部搜索作用 (即利用这些较差的解产生最优解的概率很小). 如果种群中组成最优解所在的 V_i^j 区间内的高位基因或其互补基因缺失, 根据定理 2 得 $p_{cross}\{V_i^j\} = 0$, 即不可能由此代种群通过交叉操作在其子代中得到最优解. 而且, 根据推论 1 可知, 在以后的进化过程中也不可能通过交叉产生最优解. 最优解基因的缺失可能是由于在种群初始化时该基因就不存在, 也可能是在进化

过程中丢失了. 丢失原因是含有该基因的个体被选择的概率太小, 在未形成包含该基因的高适应度个体前即被淘汰了. 此时, 种群重新出现该基因只能依靠变异的作用, 但根据定理 4 可知变异作用很有限. 另外导致 GA - 难问题现象发生的两种情况: 一是种群中高适应度个体不具有组成最优解的基因, 或具有组成最优解基因的被选择概率较小, 则此代种群中的个体通过交叉产生最优解的概率也较小; 二是种群通过交叉操作虽然没有产生最优解, 但以较大概率产生了具有最优解基因的个体, 它有利于在以后的进化中通过交叉产生最优解, 但如果此时具有最优解的基因个体被选择的概率较小, 则不利于下一代产生最优解, 因为仍可能在进化过程中发生该最优解基因的丢失. 以上任意一种情况发生都将产生 GA - 难问题的现象. 对于大多数的具有 GA 欺骗性的问题, 在足够大的种群规模下 GA 仍能获得全局最优解, 但是严重的欺骗问题, 发现全局最优解的概率将大幅度减小, 同时搜索效率也大大降低.

2.3.2 GA - 易问题产生的原因

优化过程中除了发生 GA - 难问题的情况外, 使用 GA 进行优化均可取得较好的效果, 只是程度不同. 根据定理 2 和定理 3, 可界定 GA - 易问题的上界, 最容易得到最优解的情况是: 种群进化是沿着产生最优解基因的方向进行的, 而且包含最优解基因的个体具有较高的被选择概率. 这与种群的初始化和选择的作用密切相关.

对于二进制编码的单调函数或单峰函数的优化问题, GA 具有良好的寻找全局最优解能力. 完全可利用定理 2 和定理 4 解释. 必须指出, 此种情况并不是 GA 最容易求解的问题. 大多数实际优化问题的目标函数常为多峰函数或非单调函数, 因此搜索过程中可能存在搜索方向偏离全局最优解存在位置. 另外, De Jong K A^[17] 等人通过大量实验发现, 对于具有平滑包络线类型的多峰函数 GA 也具有良好的求解能力, 仍可利用定理 2 和定理 4 解释. 这是模式定理难以说明的.

3 遗传算法的 Markov 链收敛性分析

3.1 基于二进制编码的有限群体 Markov 链模型

设 $H = \{0, 1\}^L$ 为个体空间, H^N 为 N 个个种群空间; 若个体 $x_i \in H, 0 \leq i \leq N - 1$, 则 $X(t) = (x_0, x_1, \dots, x_{N-1}), (t \geq 0)$ 表示第 t 代种群; $\{X(t), t \geq 0\}$ 为遗传算法种群空间上的种群序列, 将所求问题的最优解记作 x^* .

将个体空间 H 进行 $r(0 < r < L)$ 级分解, 得到其子空间 $\{V_0^r, \{V_1^r, \dots, \{V_{2^r-1}^r\}$. 将子空间进行组合, 并将子空间及其组合方式作为一个新的随机变

量 $Y(t)$ 的状态空间, 记为

$$S^r = \{\{V_0^r\}, \{V_1^r\}, \dots, \{V_{2^r-1}^r\}, \{V_0^r, V_1^r\}, \dots, \{V_0^r, V_1^r, \dots, V_{2^r-1}^r\}\}, \quad (11)$$

其中的状态总数 $N = 2^{2^r} - 1$. 为了简便起见, 记 $S^r = \{1, 2, \dots, 2^{2^r} - 1\}$, 对于状态 $i \in \{1, 2, \dots, 2^{2^r} - 1\}$ 按照从小到大的顺序一一对应于 $\{V_0^r\}, \{V_1^r\}, \dots, \{V_w^r\}, \{V_0^r\}, \dots, \{V_0^r, V_1^r, \dots, V_{2^r-1}^r\}$. 在 S^r 中前 2^r 个状态称为基本状态, 它是由 H 的子空间组成; 其余状态是基本状态的一个组合.

种群 $X(t)$ 与随机变量 $Y(t)$ 的对应关系满足: 当 $Y(t) = i(i \in S^r)$ 时, 对于状态 i 中包含的每个基本状态 $\{V_k^r\}$, 在种群 $X(t)$ 中至少存在一个个体 $x_m(t)$, 使得 $x_m(t) \in V_k^r$. 则 $\{Y(t), t \geq 0\}$ 为定义在概率空间上离散参数的随机过程. 状态空间 S^r 为可列集, 对于任意非负整数 t 及任意状态 $i_0, i_1, \dots, i_{t+1} \in S^r, p\{Y(0) = i_0, Y(1) = i_1, \dots, Y(t) = i_t\} > 0$, 有

$$p\{Y(t+1) = i_{t+1} | Y(0) = i_0, Y(1) = i_1, \dots, Y(t) = i_t\} = p\{Y(t+1) = i_{t+1} | Y(t) = i_t\} \quad (12)$$

成立, 则 $\{Y(t), t \geq 0\}$ 为遗传算法的离散参数 Markov 链.

定义 5 对于 $\forall i, j \in S^r$, 称

$$p_{ij}(t) = p\{Y(t+1) = j | Y(t) = i\} \quad (13)$$

为 t 时刻的一步状态转移概率. 记 $P(t) = (p_{ij}(t))$, 称 $P(t)$ 为马尔可夫链 $\{Y(t), t \geq 0\}$ 的一步状态概率转移矩阵. 显然有 $p_{ij}(t) \geq 0, \sum_{j \in S^r} p_{ij}(t) = 1$, 所以

$P(t)$ 是一个随机矩阵.

记 $\pi_t = p\{Y(t) = i\}$, t 时刻 $Y(t)$ 的概率分布向量表示为 $\pi(t) = \{\pi_0(t), \pi_1(t), \dots, \pi_i(t), \dots\}$, 称 $\pi(0)$ 为 $\{Y(t), t \geq 0\}$ 的初始分布. 记

$$p_{ij}^{(n)}(t) = p\{Y(t+n) = j | Y(t) = i\} \quad (14)$$

为 n 步转移概率; $P^{(n)}(t) = (p_{ij}^{(n)}(t))$ 为 n 步转移概率矩阵.

定义 6 若 $\lim_{t \rightarrow \infty} \pi_t(t) = \pi_t^*$ 存在, 则称

$$\pi^* = \{\pi_0^*, \pi_1^*, \dots, \pi_i^*\} \quad (15)$$

为马尔可夫链的极限分向量.

一个马尔可夫链运动规律的概率特性取决于其状态转移矩阵的特性. 给定一个初始分布 $\pi(0)$, 则马尔可夫链在 t 个时间步后的分布 $\pi(t) =$

$\pi(0) \prod_{m=0}^{t-1} P(m)$. 因此遗传算法的有限马尔可夫链

$\{Y(t), t \geq 0\}$ 完全由 $\pi(0)$ 和 $P(t)$ 决定, $\{Y(t), t \geq 0\}$ 的渐近性质取决于概率转移矩阵 $P(t)$ 的结构.

在遗传算法的优化计算中, 关键问题有两个: 1) 当 $t \rightarrow \infty$ 时, $p\{Y(t) = i\} = \pi_i(t)$ 的极限是否存在? 2) 在什么条件下, 遗传算法的马尔可夫链是一个平

稳序列?对于问题 1), 因 $\pi_i(n) = \pi_i(0)p_{ij}^{(n)}$, 故可转化为研究概率转移矩阵 $p_{ij}^{(n)}$ 的渐近性质, 即 $\lim_n p_{ij}^{(n)}(t)$ 是否存在?若存在, 其极限是否与 i 有关? 问题 2) 是关于一个平稳分布是否存在的问题 $\{Y(t), t \geq 0\}$ 的状态转移矩阵 $P(t)$ 是时变矩阵, 它与第 t 代种群 $X(t)$ 的个体的分布情况有关, 并且状态转移概率可以求得. 因为种群 $X(t)$ 个体分布情况决定了种群 $X(t+1)$ 中个体在各分区内的概率分布.

3.2 种群 $X(t)$ 在选择的作用下, 交叉操作的收敛性分析

定理 6 在选择和交叉作用下的遗传算法的 Markov 链 $\{Y(t), t \geq 0\}$ 的极限分布存在, 并且此极限分布为其平稳分布.

证明 设对解空间进行 r 级分区, 在 t 时刻后的 n 步状态转移矩阵, 记为 $P^{(n)}(t)$, $P^{(n)}(t) = (p_{ij}^{(n)}(t))$ 为随机矩阵, 经初等变换后, 可简化为

$$P^{(n)}(t) = \left[\begin{array}{ccc|c} & & & 0 \\ & & & \vdots \\ & & & 0 \\ \hline p_{2^{2^r-1,1}}^{(n)}(t) & p_{2^{2^r-1,2}}^{(n)}(t) & \dots & p_{2^{2^r-1,2^r}}^{(n)}(t) \end{array} \right] \tilde{p}^{(n)}(t)$$

式中: I_{2^r} 为状态: $\{V_0^0\}, \{V_1^1\}, \dots, \{V_{2^r-1}^{2^r-1}\}$ 对应的状态转移矩阵 (单位阵); $p_{2^{2^r-1,1}}^{(n)}(t), p_{2^{2^r-1,2}}^{(n)}(t), \dots, p_{2^{2^r-1,2^r}}^{(n)}(t)$ 为状态 2^{2^r-1} 转移到吸收状态 $i (i = \{1, \dots, 2^r\})$ 的状态转移概率; $\tilde{p}^{(n)}(t) = (p_{2^{2^r-1,k}}^{(n)}(t))_{k=2r+1}^{2^{2^r}-1}$ 为状态集合 $U = \{2^r+1, 2r+2, \dots, 2^{2^r}-1\}$ 中的所有状态的 n 转移概率的代数和. 根据 Kolmogorov-Chapman 方程, 有 Markov 链 $\{Y(t)\}$ 的 n 步状态转移概率矩阵

$$p^{(n)}(t) = p^{(n-1)}(t) \cdot p(t) = \sum_{k=0}^{n-1} p(t+k),$$
 令 $\forall i \in \{1, 2, \dots, 2^r\}$, 有 $0 < p_{2^{2^r-1,i}}^{(n)}(t) < 1, 0 < \tilde{p}^{(n)}(t) < 1$, 又因为
$$p_{2^{2^r-1,i}}^{(n)}(t) = p_{2^{2^r-1,i}}^{(n-1)}(t) (1 + p_{2^{2^r-1,i}}^{(n-1)}(t) + \dots + p_{2^{2^r-1,i}}^{(n-1)}(t)^{n-1}) > p_{2^{2^r-1,i}}^{(n-1)}(t),$$
 随着 n 增大, $p_{2^{2^r-1,i}}^{(n)}(t)$ 是单调增加的有界函数, 所以 $p_{2^{2^r-1,i}}^{(n)}(t)$ 的极限存在, 记为 $p_i = \lim_n p_{2^{2^r-1,i}}^{(n)}(t)$. 又因为 $p^{(n)}(t)$ 为随机矩阵, 则有

$$\tilde{p}^{(n)}(t) = 1 - \sum_{i \in \{1, 2, \dots, 2^r\}} p_{2^{2^r-1,i}}^{(n)}(t),$$

由于 $p_{2^{2^r-1,i}}^{(n)}(t)$ 是单调递增函数, 则 $\tilde{p}^{(n)}(t)$ 是单调递减且有界函数, 所以 $\lim_n \tilde{p}^{(n)}(t) = 0$, 于是

$\lim_n p_{2^{2^r-1,k}}^{(n)}(t) = 0$, 且

$$\lim_n p^{(n)}(t) = \lim_n \sum_{k=0}^{n-1} p(t+k) = \left[\begin{array}{ccc|c} & & & 0 \\ & & & \vdots \\ & & & 0 \\ \hline p_1 & p_2 & \dots & p_{2^r} \end{array} \right] \tilde{\pi}(0) \quad (16)$$

设在 $t = 0$ 后的 n 个时间步时的概率分布 $\pi_i(n) = \pi_i(0)p_{ij}^{(n)}$, 又由极限分布的定义有

$$\pi_i^* = \lim_n p(y(n) = j) = \lim_n \pi_i(n) = \lim_n \pi_i(0)p_{ij}^{(n)}(0) = \pi_i(0) \lim_n p_{ij}^{(n)}(0),$$

$$\pi^* = \{\pi_0^*, \pi_1^*, \dots, \pi_{2^r}^*, \tilde{\pi}^*\} = \pi(0) \lim_n p^{(n)}(0) = (\pi_0(0) \quad \pi_1(0) \quad \dots \quad \pi_{2^r}(0) \mid \tilde{\pi}(0)) \times \left[\begin{array}{ccc|c} & & & 0 \\ & & & \vdots \\ & & & 0 \\ \hline p_1 & p_2 & \dots & p_{2^r} \end{array} \right]$$

$$= (\pi_0(0) (1 + \tilde{\pi}(0)p_1) \quad \pi_1(0) (1 + \tilde{\pi}(0)p_2) \quad \dots \quad \pi_{2^r}(0) (1 + \tilde{\pi}(0)p_{2^r}) \mid 0).$$

因此, 在选择和交叉作用下的遗传算法的 Markov 链 $\{Y(t)\}$ 的极限分布存在, 且当 $j \in \{1, \dots, 2^r\}$ 时, $\pi_j^* = \pi_j(0) (1 + \tilde{\pi}(0)p_j)$; $j = U$ 时, $\pi_j^* = 0$.

对于概率分布向量 $\pi^* = \{\pi_0^*, \pi_1^*, \dots, \pi_{2^r}^*, 0, \dots, 0\}$, 有 $\pi^* P(t) = \pi^*$, 则 π^* 为其平稳分布.

推论 4 仅在选择、交叉的作用下, 种群必然收敛于解空间的某一个分区.

证明 设 $\tilde{\pi}(t) = \sum_{k=2r+1}^{2^{2^r}-1} \pi_k(t)$ 为状态集合 $U = \{2r+1, 2r+2, \dots, 2^{2^r}-1\}$ 中的状态的概率分布之和, 记 $\tilde{\pi}^* = \lim_t \tilde{\pi}(t)$. 对于初始分布 $\pi(0) = (\pi_0(0) \quad \pi_1(0) \quad \dots \quad \pi_{2^r}(0) \mid \tilde{\pi}(0))$, 根据定理 6, 存在极限分布 $\pi^* = \{\pi_0^*, \pi_1^*, \dots, \pi_{2^r}^*, 0, \dots, 0\}$. 由此可见, $\tilde{\pi}^* = 0$, 所以种群必然收敛于某一个基本状态. 并且, 收敛于基本状态 j 内的概率为 $\pi_j(0) (1 + \tilde{\pi}(0)p_j)$.

由推论 4 可知, 选择与交叉操作可导致算法的未成熟收敛现象的发生. 可根据定理 6, 从 $\{Y(t)\}$ 的极限分布得到收敛到最优解的概率, 以及得到局部最优解的概率. 关于未成熟收敛现象在第 3 节中作具体分析.

定义 7 种群 $X(t)$ 在遗传操作下能够搜索到的分区数, 称为该种群多样性, 记为 $d(t)$.



推论 5 设任意初始种群 $X(0)$, 在选择、交叉的作用下, 则有 $\lim_j d(t) = 1$.

证明 设 $j \in U$, 由定理 1 可知 $\{Y(t)\}$ 的极限分布存在, 并且 $\lim_j p\{Y(t) = j\} = 0$, 又由推论 4 可知, $\{Y(t)\}$ 必然收敛于某一个分区 V^i , 于是有 $\lim_j d(t) = 1$.

例 1 当对解空间进行一级分区时, 即 $r = 1$, 则

$$S^1 = \{\{V_0^1\}, \{V_1^1\}, \{V_0^1, V_1^1\}\},$$

$$P(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ p_{31}(t) & p_{32}(t) & p_{33}(t) \end{bmatrix},$$

根据 (16) 有

$$\lim_n P^{(n)}(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ p_{31} & p_{32} & 0 \end{bmatrix},$$

于是

$$\lim_n \pi_1(n) = \lim_n \sum_{i \in S} \pi_i(0) p_{i1}^{(n)}(0) = p_{31},$$

$$\lim_n \pi_2(n) = \lim_n \sum_{i \in S} \pi_i(0) p_{i2}^{(n)}(0) = p_{32},$$

$$\lim_n \pi_3(n) = \lim_n \sum_{i \in S} \pi_i(0) p_{i3}^{(n)}(0) = 0$$

对于初始分布 $\pi(0) = (0 \ 0 \ 1)$ 和 $\pi(0) = (0 \ 1 \ 0)$, 遗传算法的马尔可夫链 $\{Y(t), t \geq 0\}$ 的极限分布分别为

$$\pi^* = \pi(0)P = (0 \ 1 \ 0) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ p_{31} & p_{32} & 0 \end{bmatrix} = (p_{31} \ p_{32} \ 0),$$

$$\pi^* = \pi(0)P = (0 \ 1 \ 0) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ p_{31} & p_{32} & 0 \end{bmatrix} = (0 \ 1 \ 0).$$

上述结论说明, 只有两个基本分区时, 交叉操作必然使种群以概率 p_{31} 或 p_{32} 收敛于分区 $\{V_0^1\}$ 或 $\{V_1^1\}$. 如果最优解在 $\{V_0^1\}$ 分区内, 那么得到最优解的概率等于 p_{31} ; 在交叉操作下, 对于状态 $i \in \{1, 2\}$, 有 $p_{ii} = 1$, 所以 i 为吸收态. 如果种群 $X(t)$ 进化到某一个吸收态, 那么, 种群 $X(t)$ 的子代个体必然全部收敛于该吸收态. 当 $r = L$ 时, $\{Y(t)\}$ 的状态空间 S^L 的每一个基本状态对应于个体空间 H 中的一个个体, $\{Y(t)\}$ 完全可以等价地定量描述 $\{X(t)\}$ 的进化过程. 至此, 可用马尔可夫链 $\{Y(t)\}$ 精确刻画有限群体下遗传算法的交叉操作的动态行为, 及其收敛性.

3.3 在变异操作下的状态转移矩阵

在 r 级分区下, $S^r = \{1, 2, \dots, 2^r - 1\}$. 种群 $X(t)$ 在变异操作下, $\{Y(t), t \geq 0\}$ 的状态转移矩阵 $Q(t) = (q_{ij}(t))$ 为随机矩阵, 且可以求得 $0 < q_{ij}(t) < 1$.

当对解空间进行 1 级分区时, 则 $S^1 = \{\{V_0^1\}, \{V_1^1\}, \{V_0^1, V_1^1\}\}$, 变异操作的状态转移矩阵

$$Q(t) = \begin{bmatrix} q_{11}(t) & q_{12}(t) & q_{13}(t) \\ q_{21}(t) & q_{22}(t) & q_{23}(t) \\ q_{31}(t) & q_{32}(t) & q_{33}(t) \end{bmatrix}.$$

因为 $q_{ij}(t) > 0$, 所以任意两个状态是相通的, 此时的马尔可夫链 $\{Y(t), t \geq 0\}$ 是不可约的. 当 $r = L$ 时, $\{Y(t), t \geq 0\}$ 的状态转移矩阵将描述的是遗传算法马尔可夫链 $\{X(t), t \geq 0\}$ 的变化, 即种群 $X(t)$ 在变异操作下的动态行为.

3.4 遗传算法的选择、交叉和变异操作联合作用下的收敛性分析

设马尔可夫链 $\{Y(t), t \geq 0\}$ 的状态转移矩阵为 $R(t) = (r_{ij}(t))$, 令 $P(t)$ 和 $Q(t)$ 分别表示交叉和变异操作下 $\{Y(t)\}$ 的状态转移矩阵, 选择、交叉和变异联合作用时, 遗传算法的马尔可夫链 $\{Y(t), t \geq 0\}$ 的一步状态转移矩阵 $R(t) = P(t)Q(t)$. 可以证明 $\{Y(t), t \geq 0\}$ 是有限、非齐次、不可约、非周期马尔可夫链.

例 2 当对解空间进行一级分区时, $S^1 = \{\{V_0^1\}, \{V_1^1\}, \{V_0^1, V_1^1\}\}$, 则有

$$R(t) = \begin{bmatrix} r_{11}(t) & r_{12}(t) & r_{13}(t) \\ r_{21}(t) & r_{22}(t) & r_{23}(t) \\ r_{31}(t) & r_{32}(t) & r_{33}(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ p_{31}(t) & p_{32}(t) & p_{33}(t) \end{bmatrix} \times \begin{bmatrix} q_{11}(t) & q_{12}(t) & q_{13}(t) \\ q_{21}(t) & q_{22}(t) & q_{23}(t) \\ q_{31}(t) & q_{32}(t) & q_{33}(t) \end{bmatrix} = \begin{bmatrix} q_{11}(t) & q_{12}(t) & q_{13}(t) \\ q_{21}(t) & q_{22}(t) & q_{23}(t) \\ q_{31}(t) & r_{32}(t) & r_{33}(t) \end{bmatrix}. \quad (17)$$

其中: $r_{3k}(t) = p_{31}q_{1k} + p_{32}q_{2k} + p_{33}q_{3k}, k = 1, 2, 3$. $r_{1k}(t)$ 和 $r_{2k}(t)$ 的大小由变异概率决定, $r_{3k}(t)$ 的大小由交叉和变异操作的状态转移概率的大小决定. 一般变异的概率较小 (0.001) 时, 可求得 $q_{31}(t) = 0, q_{32}(t) = 0$, 且 $q_{33}(t) = 1$, 于是 $r_{3k}(t) = p_{3k}$, 即交叉操作在遗传算法中起决定的作用.

根据式 (2) 可容易总结出一级分区下的选择、

交叉和变异在种群的进化中的作用, 并可做定量和定性分析。如交叉操作下, 对于状态 $i \in \{1, 2\}$ 为吸收态, 那么, 种群在进化过程中, 跳出吸收状态必须依靠变异的作用, 跳出吸收态的能力取决于变异作用的大小。这对设计防止早熟的算法很重要。当 $r \ll L$ 时, $\{Y(t), t=0\} = \{X(t), t=0\}, \{Y(t), t=0\}$ 的状态转移矩阵描述的是遗传算法马尔可夫链 $\{X(t), t=0\}$ 的变化, 即种群 $X(t)$ 在选择、交叉和变异操作下的动态行为。根据状态转移矩阵可容易证明, 标准遗传算法马尔可夫链 $\{X(t), t=0\}$ 为有限、齐次、不可约、非周期的马尔可夫链。

3.5 关于遗传算法的早熟收敛现象分析

遗传算法的未成熟收敛是遗传算法的主要弊病之一, 它也是遗传算法应用的最大障碍。分析遗传算法的本质属性, 找到发生早熟现象的原因有重要的理论意义和实践价值。未成熟收敛具体表现如下:

1) 基于适应度选择作用下, 进化过程中具有最优解基因的使低适应度的个体被逐渐淘汰, 而且使种群中个体趋于一致;

2) 基因型相同个体进行交叉不能生成新个体;

3) 变异产生适应度高的个体数量少, 变异的作用不明显。

下面分析发生早熟现象的原因和时间:

设包含最优解的状态为 j , 当种群 $X(t_1)$ 处于某个状态 $i (i \neq j)$ 时, $p_{ij}(t_1) = \alpha$ (α 为一个较小的正数), 即此时种群 $X(t_1)$ 收敛于状态 j 的概率较小, 如果 $t_2 > t_1$ 时, 均有 $p_{ij}(t_2) = \alpha$ 。那么, t_1 时刻的种群 $X(t_1)$ 将发生准早熟收敛现象, 在 $t_2 > t_1$ 时, 种群 $X(t_2)$ 发生了早熟现象。种群收敛现象的外在表现是种群中的个体所包含的基因型减少, 使种群中个体趋于一致。如果种群不是朝着最优解的方向收敛, 那么种群将收敛到局部最优解, 根据已经发生了早熟收敛的种群不能通过交叉操作生成新个体, 如果依靠变异的作用产生新的基因型的能力有限, 这时就发生了早熟现象。

早熟现象发生原因是由于解空间中具有多个吸收态, 变异操作作用较小时, 种群主要依靠选择和交叉操作, 根据推论 1 可知, 选择和交叉操作将引导种群向着某一个吸收态收敛, 如果包含收敛于最优解的吸收态的概率较小, 则种群发生早熟现象的概率较大。当且仅当种群中具有最优解的基因, 通过交叉操作才有可能搜索到最优解。如果种群收敛到某个吸收态时, 通过交叉产生包含最优解状态的概率近似等于 0。如果此时变异的概率很小, 则得到最优解的概率仍旧很小。

理论上, 只要有变异存在, GA 必然是不收敛

的, 也不会收敛到最优解。但观察到在 GA 的执行实践中, 交叉概率通常取 $0.6 \sim 0.95$, 变异概率取 $0.001 \sim 0.01$, 故变异使种群发生显著改变的概率极小, 此时交叉算子在 GA 中起相对核心的作用; 如果种群发生准早熟收敛现象时, 使种群从某个吸收态中跳出来的变异操作起核心作用。

标准遗传算法的马氏链是时齐、非周期和不可约的, 所以其种群序列 $\{X(t), t=0\}$ 从一个状态出发, 将以概率 1 在有限步内达到 H 中的任一状态 (因为有变异操作存在)。但是, 从理论上讲, 必须在 t

时, 采用精英保留策略的遗传算法依概率收敛到全局最优解, 即 $\lim_{t \rightarrow \infty} p(X^* = X(t)) = 1$ 。遗传搜索效率和时间的复杂性与遗传操作的搜索性能有关。在实际应用中, 关心的是在有限步的迭代内找到全局最优解。也就是说, 必须在种群发生早熟现象之前, 通过选择与交叉操作的全局搜索能力发现最优解。

4 结 论

目前, 大量实践证据表明了遗传算法寻优能力在许多领域获得了成功, 它通过迭代搜索来逐步找到问题的最优解。在这个过程中, 哪个个体更易生存或被淘汰以及种群的进化方向等问题, 定理 2、定理 4 和定理 5 以概率的形式做出了定量分析, 并证明了遗传算法具有寻求最优解的能力。基于这些结论解释了为什么会产生 GA-难问题和 GA-易问题, 在什么条件下产生这些问题。从刻画遗传算法的种群进化过程出发, 分析了选择、交叉、变异操作中基因重组和突变的具体过程, 阐明了遗传算法寻优本质。从理论上更加完善地对遗传算法进行定量的分析。遗传算法的定向搜索机制是在适应度高的父代“家族”的邻域内进行搜索。在解空间的分区基础上建立了二进制编码的有限群体的 Markov 链模型, 可恰当地度量与精确刻画 GA 在不同实现下的收敛过程, 并给出了一种新的有关遗传算法的收敛性理论。同时, 阐明了遗传算法产生早熟收敛现象的本质原因。一旦理解了遗传算法寻优和收敛的本质, 将为遗传算法更广泛地应用与推广奠定扎实的基础。

参考文献 (References)

- [1] Fogel D B. An Introduction to Simulated Evolutionary Optimization [J]. *IEEE Trans on Neural Networks*, 1994, 5(1): 3-14
- [2] Cheong F, Lai R. Constraining the Optimization of a Fuzzy Logic Controller Using an Enhanced Genetic Algorithm [J]. *IEEE Trans on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2000, 30(1): 31-46
- [3] Pan Z J, Kang L S, Nie S X. Evolving Both the Topology and Weights of Neural Network [J]. *Parallel*

- A lgorithms and Applications*, 1996, 9(4): 299-307.
- [4] Pham D T, Jin G. Genetic Algorithm Using Gradient-like Reproduction Operator [J]. *Electronics Letters*, 1995, 31(18): 1558-1559.
- [5] Fogel D B. An Introduction to Simulated Evolutionary Optimization [J]. *IEEE Trans on Neural Networks*, 1994, 5(1): 3-14.
- [6] 张丽萍, 柴跃廷. 遗传算法的现状与发展动向[J]. *信息与控制*, 2001, 30(6): 532-536
(Zhang L P, Chai Y T. Actuality and Developmental Trend for Genetic Algorithms [J]. *Information and Control*, 2001, 30(6): 532-536.)
- [7] Holland J H. *Adaptation in Natural and Artificial Systems* [M]. Ann Arbor: Michigan University Press, 1975.
- [8] Maniezzo V. Genetic Evolution of the Topology and Weight Distribution of Neural Networks [J]. *IEEE Trans Neural Networks*, 1994, 5(1): 54-65.
- [9] Rudolph G. Convergence Analysis of Canonical Genetic Algorithms [J]. *IEEE Trans on Neural Networks*, 1994, 5(1): 96-101.
- [10] Fogel D B. Asymptotic Convergence Properties of Genetic Algorithms and Evolutionary Programming [J]. *An Analysis and Experiments, Cybern, and System*, 1994, 25(2): 389-407.
- [11] Goldberg D E, Segrest P. Finite Markov Chain Analysis of Genetic Algorithms [A]. *Proc of the 2nd International Conf on Genetic Algorithms* [C]. San Mateo, 1987: 1-8.
- [12] Schraudolph N N, Belew R K. Dynamic Parameter Encoding of Genetic Algorithms [J]. *Machine Learning*, 1992, 9(1): 9-21.
- [13] 徐宗本, 高勇. 遗传算法过早收敛现象的特征分析及其预防[J]. *中国科学(E)*, 1996, 26(4): 364-375
(Xu Z B, Gao Y. Analysis and Prevention of Genetic Algorithms Premature Convergence [J]. *Science in China (Series E)*, 1996, 26(4): 364-375.)
- [14] Xu Z B, Nie Z K, Zhang W X. A Most Sure Convergence of Genetic Algorithms: A Martingale Approach [J]. *Chinese J of Computers*, 2001, 25(8): 785-793.
- [15] 徐宗本, 陈志平. 遗传算法基础理论研究的新近发展[J]. *数学进展*, 2000, 29(2): 97-114
(Xu Z B, Chen Z P. Theoretical Development on Genetic Algorithms: A Review [J]. *Advances in Mathematics*, 2000, 29(2): 97-114.)
- [16] Zhao X Y, Nie Z K. The Markov Chain Analysis of Premature Convergence of Genetic Algorithms [J]. *Chinese Quarterly Journal of Mathematics*, 2003, 18(4): 364-368.
- [17] De Jong K A. *The Analysis of the Behavior of a Class of Genetic Adaptive Systems* [D]. Dearborn: University Michigan, 1975: 76-9381.

(上接第970页)

- [6] Fogel D B. An Introduction to Simulated Evolutionary Optimization [J]. *IEEE Trans on Neural Networks*, 1994, 5(1): 3-14.
- [7] 潘伟, 井元伟. 基于遗传算法的降阶 H_2 控制器[J]. *东北大学学报(自然科学版)*, 2004, 25(6): 519-522
(Pan W, Jing Y W. Reduced Order H_2 Controllers Based on Genetic Algorithm [J]. *J of Northeast University (Natural Science)*, 2004, 25(6): 519-522.)
- [8] 王进华, 史忠科, 曹力, 等. 混合 H_2/H_∞ 鲁棒控制在飞行控制中的应用[J]. *飞行力学*, 2000, 18(4): 68-72
(Wang J H, Shi Z K, Cao L, et al. Application of Mixed H_2/H_∞ Robust Control to Aircraft [J]. *Flight Dynamics*, 2000, 18(4): 68-72.)
- [9] 潘伟, 王学勇, 井元伟. 基于遗传算法的混合降阶 H_2/H_∞ 状态反馈控制器[J]. *控制与决策*, 2005, 20(2): 132-136
(Pan W, Wang X Y, Jing Y W. Mixed State Feedback Controllers Based on Genetic Algorithm [J]. *Control and Decision*, 2005, 20(2): 132-136.)