

遗传算法的一种非单调适应值标度变换方法*

李敏强 寇纪淞

天津大学系统工程研究所, 天津 300072

摘要 讨论了遗传算法(GA)中选择算子对于改善群体整体品质、实现群体进化和导致群体早熟与模式欺骗的双重作用. 以 One-Max 函数为例分析了选择压力对群体进化能力的影响. 针对复杂优化问题, 提出了一种非单调适应值标度变换方法, 给出了均值规范变换和可调规范变换两种形式, 并结合典型函数进行了实验计算.

关键词 遗传算法 选择算子 非单调适应值变换

遗传算法(Genetic Algorithms, GA)是在模拟 Darwin 的进化论和 Mendel 的遗传学理论的基础上, 产生和发展起来的一种优化问题求解的随机化搜索方法^[1-3]. 算法的编码空间、群体、个体位串相当于生物种群和染色体基因, 以选择来表示“自然选择”的生物进化机理, 交叉和变异相当于生物的有性繁殖和突变, 适应值的评价相当于自然界对进化中的生物群体所提供的反馈信息以及开放的生物进化的过程, 遗传算法获得了全局搜索能力、信息处理的隐并行性、鲁棒性和可规模化等优良性能, 是一种具有强大的复杂优化问题求解能力和良好普适性的方法.

在 GA 的各种遗传算子中, 选择算子是其中的一个基本算子. 选择过程体现了生物进化过程中“适者生存, 优胜劣汰”的思想, 并且保证优良基因遗传给下一代个体^[3-5]. 选择算子对 GA 搜索性能具有决定性的影响, 选择算子及其参数设计不当将使得 GA 丧失全局收敛性, 搜索效率大幅度下降.

本文具体研究了选择算子与群体早熟的关系, 给出了一种面向克服群体早熟和模式欺骗性的非单调适应值标度变换方法, 以陷阱函数和大海捞针类问题为例进行实验计算.

1 选择算子与 GA 性能

遗传算法应用于优化问题求解, 可以视为一种随机化搜索过程. 在该搜索过程中, GA 不仅需要探索解空间上的全局最优解, 而且应当充分利用已获得的解空间信息逼近当前局部最优解^[3-5], 我们分别称之为 GA 的求泛和求精的能力. 通过扬弃性的探索, 克服局部极值陷阱和模式欺骗, 实现广域搜索, 提高全局寻优能力. 通过继承性的开发, 维持群体的可进化性, 消除早熟问题, 实现邻域搜索, 提高逼近搜索能力.

由于选择算子引导着群体的搜索方向, 所以选择算子在很大程度上影响着 GA 的求精和

2000-06-15 收稿, 2000-10-08 收修改稿

* 国家自然科学基金基金项目资助(批准号: 69974026)

求泛能力. 对于一维优化函数 $\max f(x)$ ($x \in [a, b]$), 从搜索区域来分析, 在任何阶段, 随机搜索的整个区域均为 $[a, b]$, GA 搜索则是以当前最优解 x_0 为中心的 $[a, b]$ 上的一个概率分布 (对称或非对称), 传统基于可行解的单点搜索是以 x_0 为中心的半径为 σ ($\sigma \ll b - a$) 邻域上的一个概率分布.

在进化的不同阶段, GA 搜索区域不一样. 在早期阶段, 群体具有较好的多样性, GA 搜索区域为整个编码空间 (不考虑采样误差). 在选择算子的作用下, 群体的多样性不断减少, 群体中的个体趋于相似, GA 搜索区域仅为编码空间中的某个子空间, GA 搜索行为也就从探索进入开发阶段. 当选择算子和参数设计不合理, GA 过早结束探索阶段, 难以实现全局寻优; 过迟进入开发阶段, 难以完成邻域搜索和提高当前最优解的精度. 因此, 以提高 GA 求解问题全局最优解的能力为目标, 选择算子一直是一个重点研究方向, 也是微观遗传策略设计的一个重要内容^[5,6].

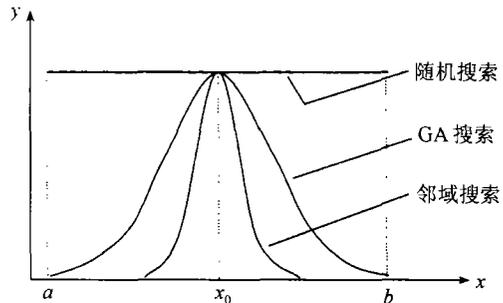


图1 3种搜索方式与能力比较

2 选择算子与群体早熟

在实际优化问题计算中, 存在着遗传算法并不总是收敛到全局最优解或满意解的现象. 多模态函数 GA 求解则更为复杂, 往往会产生严重的模式欺骗问题. 当前群体的最佳位串距最优解位串仍然存在着比较显著的海明距离, 但是当前最优解在很长的一个搜索阶段得不到改进. 我们统称为群体早熟或者成熟前收敛^[7,8].

GA 早熟的本质特征指群体中的各个个体非常相似, 群体的多样性急剧减少, 当前群体缺乏有效等位基因 (最优解位串上的等位基因), 在遗传算子作用下不能生成高阶竞争模式^[7,8]. GA 早熟除与群体规模、模式欺骗、适应函数性质等有关外, 选择算子是一个重要的影响因素. 当变异概率比较小时, 个体选择压力太大, 将导致群体的多样性迅速降低. 个体选择压力太小, 模式竞争减弱, 遗传算子重组生成高阶模式的能力减小, 也会出现进化停滞现象. 我们主要采用实验方式分析 GA 早熟与选择压力的关系.

以 One-Max 函数的 GA 求解为例进行实验计算^[5]. One-Max 函数是一类特殊的析取位函数, 该函数是位串中包含的基因值为“1”的基因位数量的线性函数, 显然是 GA-容易问题. 函数形式为:

$$f(\mathbf{a}) = \sum_{i=1}^L a_i, \quad (1)$$

其中 $\mathbf{a} = (a_1, a_2, \dots, a_L)$, 最优解为 $\mathbf{a}^* = (1, 1, \dots, 1)$, $f(\mathbf{a}^*) = L$.

考虑不同选择压力下的 One-Max 优化问题的 GA 求解情况. 参数设定: 位串长度 $L = 120$, 群体规模 $n = 120$; 采用排序选择方式和精英保留策略, 当前群体最佳个体在下一代生存的期望数量分别为 $\eta^+ = \{1.1, 1.6, 2.0\}$; 采用概率 $p_c = 0.6$ 的两点交叉方式; 变异概率 $p_m = 1/(L \times 50) = 0.000167$. 进化代数 300 的结果如图 2 所示.

可见, 选择压力太大或太小, 均不利于提高 GA 群体进化过程的搜索能力, 在遗传算子参

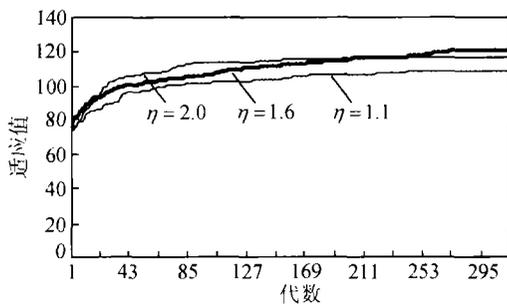


图 2 One-Max 函数在不同选择压力下 GA 进化情况

数组组合不合适的情况下,往往出现 GA 早熟现象.大量实例计算表明,针对不同优化问题设计合适的选择压力非常困难.

在单纯考虑克服群体早熟的情况下,通常采用大变异概率或者适应性变异^[7],但是其效果非常有限. Whitley 等建议从适应值标度变换的创新入手,可以解决因复杂优化问题存在的模式欺骗性而产生的早熟问题^[8].

3 适应值非单调标度变换

为了在群体进化过程中改善 GA 的性能,无论是极大或极小优化问题,都需要对适应函数进行调节.关于目标函数值到适应数值映射问题,一般表示为 $T: g(x) \rightarrow f(x)$, $x \in \mathbb{R}$ (不妨设 $f(x) > 0$). 由于个体适应值是选择算子运行的基础,所以合理确定目标函数值到适应值的映射关系具有重要意义.传统的数值映射一般要求保持群体中个体序的不变性,即对于 $x_1, x_2 \in \mathbb{R}$, 若 $g(x_1) > g(x_2)$, 则 $f(x_1) > f(x_2)$. 若 $g(x_1) = g(x_2)$, 则 $f(x_1) = f(x_2)$.

按照新 Darwin 主义的观点^[1,2,9],在自然界生物进化过程中存在着 3 种选择方式:稳定选择、定向选择、破坏性选择,规范的 GA 选择操作就是一种定向选择,即促使群体向着适应值增大的方向进化.稳定选择是一种随机选择,并不改变群体的整体性质和水平,它和个体的适应值不存在直接联系.破坏性选择导致群体中当前最佳个体或染色体被破坏或丢失,群体整体进化水平下降或者发生显著变异.

自然界中的物种进化需要一定的条件,比如一定的群体规模、适宜的生态环境、个体的免疫能力等.群体规模太小难以实现基因重组,难以抵抗自然灾害.恶劣的生态环境使得种群难以在短期内完全适应,导致种群损失太大,直至物种的灭亡和消失.从整个环境来看,物种在整体上可能是进化的,并不排除阶段性或者局部小生态环境下的退化.适者生存法则则是以定向选择为基础的,但是不能排除其他选择的存在性.

采用 GA 求解优化问题,定向选择方式保证了群体的进化收敛性,但也导致了 GA 早熟或者模式欺骗问题.为了克服 GA 进化过程中的早熟和欺骗问题,必须改变传统的选择策略,通过参考有关文献的研究思想和大量的实验分析^[3,9],我们提出了一种非单调适应值标度变换方法.

定义 1 任意优化函数 $\text{opt}g(x)$ ($x \in \mathbb{R}$, $g(x) \geq 0$),适应值函数为 $f(x)$,对于 $x_1, x_2 \in \mathbb{R}$, 存在 $x_0 \in \mathbb{R}$, $x_1 < x_0 < x_2$, 若原适应值 $f(x_1) > f(x_0) > f(x_2)$, 则新的适应值 $f'(x_1) \geq f'(x_0) \leq f'(x_2)$, 称该适应值标度变换为非单调变换.

显然,非单调适应值标度变换是一种典型的破坏性选择,它不再遵循传统上的目标函数值到适应数值映射的群体个体序的不变性,使得某些目标函数值比较差的个体也具有较大的生存概率.这里给出两种基本的变换方式.

(1) 均值规范变换:

设规模为 n 的群体 $P = \{a_1, a_2, \dots, a_n\}$, 个体适应值为 $f(a_1), f(a_2), \dots, f(a_n)$; 该适应值

的非单调标度变换为

$$f'(a_j) = |f(a_j) - \bar{f}|, \quad a_j \in P, \quad (2)$$

其中, $\bar{f} = \frac{\sum_{i=1}^n f_i}{n}$ 为群体的平均适应值.

对于 $a_j \in P$, $f'(a_j)$ 越大生存机会越高, 即适应值远离群体均值的个体被选择的概率高, 均值附近的个体被选择的概率低. 适应值处于中间位置的个体对群体改进贡献较小, 因此在非单调标度变换中的生存机会被大幅度降低.

(2) 可调规范变换:

设规模为 n 的群体 $P = \{a_1, a_2, \dots, a_n\}$, $f_{\min} = \min\{f(a_j), j = 1, 2, \dots, n\}$, $f_{\max} = \max\{f(a_j), j = 1, 2, \dots, n\}$; 适应值的非单调标度变换为

$$f'(a_j) = |f(a_j) - \bar{f}'|, \quad a_j \in P, \quad (3)$$

其中 $\bar{f}' = f_{\min} + \alpha(f_{\max} - f_{\min})$, $\alpha \in [0, 1]$. 当 $\alpha = 0$ 时, 该变换为传统的标度变换(最差个体不参与选择); $\alpha = 1$ 时, 为破坏程度最大的标度变换. 选择合适的 α 值, 可以控制非单调变换的选择效果.

考虑到 GA 整个搜索过程不同阶段的要求, 限定采用非单调适应值标度变换的有效代数范围, 以便将探索和开发搜索过程恰当地结合起来. 设最大进化代数为 T , 采用适应值非单调标度变换的有效代数为: $t \leq \beta T$, $0 \leq \beta \leq 1$, 一般取 $\beta = 2/3$. 当 $t \leq \beta T$ 时, GA 处于探索阶段, 最大程度上实现全局搜索. 当 $t > \beta T$ 时, GA 转为正常的定向选择操作, 开始局部求精的搜索过程.

4 实例计算

自动控制、机器学习、经济管理等领域的很多优化问题, 基本上都是多模态类型函数, 一般形式的 GA 均存在着如何避免陷于局部极值点而实现全局最优解搜索的问题. 我们通过对两类问题进行实验计算, 表明适应值非单调变换的意义.

4.1 陷阱函数

陷阱函数(Trap function)是一个分段析取位函数^[2], 其中包括 3 个重要参数 a , b , z . 该函数的基本形式为

$$f(u) = \begin{cases} \frac{a}{z}(z - u), & \text{如果 } u \leq z; \\ \frac{b}{L - z}(u - z), & \text{否则.} \end{cases} \quad (4)$$

其中, $a < b$, 最优解位串为“111...111”, 欺骗吸引子为“000...000”, $f(L) = b$, $f(0) = a$. 当 $u < z$ 时, $f(u)$ 为单调降函数; 当 $u \geq z$ 时, $f(u)$ 为单调升函数.

GA 参数设置为, 染色体串长 $L = 30$, 群体 $n = 100$, 两点交叉概率 $p_c = 0.6$, 变异概率 $p_m = 1/L = 0.0333$.

(1) 设 $a = 80$, $b = 100$, 分析该函数的 GA 收敛性与 $z = \text{int}(L/2) + k$ 变化关系. 其中 GA 采用适应值比例选择和精英个体保留策略, 预定进化代数为 200, 每种条件下的实验计算重复 120 次, 计算结果如表 1 所示.

表 1 陷阱函数 GA 欺骗性与 z 变化关系

(1)	$z > \text{int}(L/2) = 15$					
	$z = 15 + 1$	$z = 15 + 3$	$z = 15 + 5$	$z = 15 + 7$	$z = 15 + 9$	$z = 15 + 12$
收敛到 a 的代数	31.25	29.73	30.06	32.08	33.48	35.83
收敛到 a 的概率/%	33.33	91.67	100	100	100	100
收敛到 b 的代数	27.50	34.00				
收敛到 b 的概率/%	66.67	8.33				
(2)	$z \leq \text{int}(L/2) = 15$					
	$z = 15 - 9$	$z = 15 - 7$	$z = 15 - 5$	$z = 15 - 3$	$z = 15 - 1$	$z = 15$
收敛到 a 的代数						32.50
收敛到 a 的概率/%						5.56
收敛到 b 的代数	35.35	32.91	31.86	28.08	26.83	24.03
收敛到 b 的概率/%	100	100	100	100	100	94.44

当 $z > \text{int}(L/2) = 15$ 时,随着 z 的增大,陷阱函数的模式欺骗性增强,群体最终完全收敛于 $a = 80$. 当 $z \leq \text{int}(L/2) = 15$ 时,随着 z 的减小,该函数的模式欺骗性减弱,群体最终完全收敛于 $b = 100$. 同时,在 $z = \text{int}(L/2) = 15$ 的附近,同阶模式之间的竞争比较激烈,一旦包含最优解基因信息的模式或者包含欺骗吸引子基因信息的模式取得生存优势,将快速收敛到最优解位串或者欺骗吸引子位串.

(2) 设 $z = \text{int}(L/2) + 5 = 20$, $a = 60$ 或者 $a = 80$, $b = 100$,分别采用均值规范、可调规范非单调适应值变换, $\alpha = 0.67$; 预定进化代数为 200,非单调适应值变换范围 $\beta = 2/3$. 其他参数不变,每种条件下的实验重复 120 次,计算结果如表 2 所示.

表 2 陷阱函数 GA 定向选择与非单调变换的计算结果^{a)}

	定向选择		均值规范变换		可调规范变换	
	$a = 80$	$a = 60$	$a = 80$	$a = 60$	$a = 80$	$a = 60$
收敛到 a 的代数	31.25	46.70	43.74	29.83	120.44	141.33
收敛到 a 的概率/%	100	100	95	76.67	65	5
收敛到 b 的代数			29.0	22.75	140.76	144.87
收敛到 b 的概率/%			5	23.33	35	95

a) 误差范围: ≤ 0.00001

可见,在均值规范非单调适应值变换下,当 $a = 80$ 时,克服模式欺骗性非常困难,仅取得了微小效果. 当 $a = 60$ 时,该变换可以在一定程度上克服模式欺骗性. 在可调规范非单调适应值变换下,当 $a = 80$ 时,克服模式欺骗性已比较显著. 当 $a = 60$ 时,该变换克服模式欺骗性接近 100%,这是传统遗传策略设计所难以达到的效果.

4.2 大海捞针类问题(Needle-in-a-haystack, 简称 NiH 问题)

对于一类大海捞针问题,其全局最优解往往被最差解所包围,隔断了模式的重组过程,使得 GA 搜索长期陷入局部极值点. 为了清晰地分析 GA 搜索过程,我们设计了一个典型的大海捞针类函数(NiH 函数)

$$\max f(x, y) = \left(\frac{a}{b + (x^2 + y^2)} \right)^2 + (x^2 + y^2)^2, \quad x, y \in [-5.12, 5.12]. \quad (5)$$

其中,当 $a = 3.0$, $b = 0.05$, $\max f(0, 0) \approx 3600$; 4 个局部极值点为 $(-5.12, 5.12)$ 、 $(-5.12,$

- 5.12)、(5.12, 5.12)、(5.12, - 5.12), 函数值为 2748.78. 随着参数 $\{a, b\}$ 的变化, 该函数将形成严重程度不同的 GA 欺骗问题, 欺骗吸引子为上述 4 个局部最优解.

GA 参数设置为, 染色体位串长度 $L = 40$ (x, y 的二进制编码长度相等即均为 20), 群体 $n = \{20, 40, 60, 80, 100, 120\}$, 两点交叉概率 $p_c = 0.6$, 变异概率 $p_m = 1/L = 0.025$, 采用适应值比例选择和精英个体保留策略, 预定进化代数 $T = 300$. 每种情况下的实验重复 120 次, 计算结果如表 3 所示.

表 3 NiH 函数不同群体规模下的 GA 性能

	$n = 20$	$n = 40$	$n = 60$	$n = 80$	$n = 100$	$n = 120$
收敛到最优解的代数	88.60	83.43	81.21	82.59	83.37	81.43
收敛到最优解的概率/%	20.83	52.50	57.50	60.83	64.17	66.67

当群体规模为 60, 其他 GA 参数不变, GA 在模式欺骗情况下, 或者克服模式欺骗后, 搜索过程和行为轨迹(群体当前最优解)如图 3 所示, 群体进化中当前最优解随代数变化情况如图 4 所示.

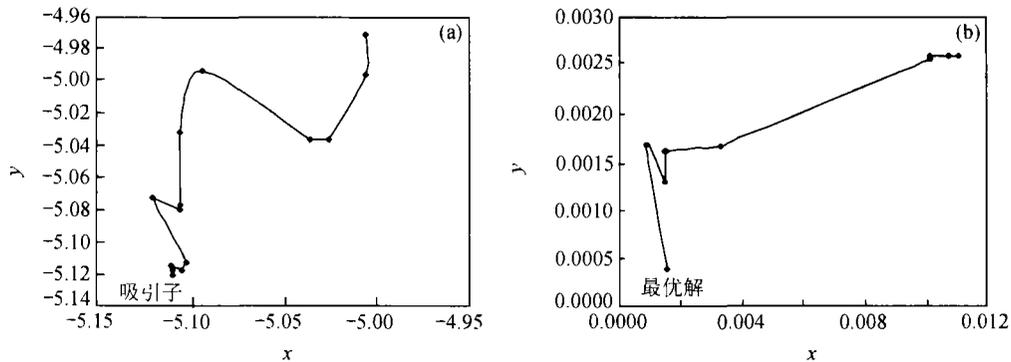


图 3 NiH 函数 GA 搜索行为

当模式欺骗性将搜索过程引向欺骗吸引子, GA 只能在局部极值点邻域内搜索, 最终收敛于局部极值点(4 个局部极值点的随机选择). 当遗传算子克服了模式欺骗之后, 则将群体搜索方向扭转到全局最优解所在的邻域, 最终收敛于全局最优解. 在某些情况下, 遗传算子很难克服模式欺骗性, 所以 GA 能否搜索到全局最优解就成为了一个随机事件.

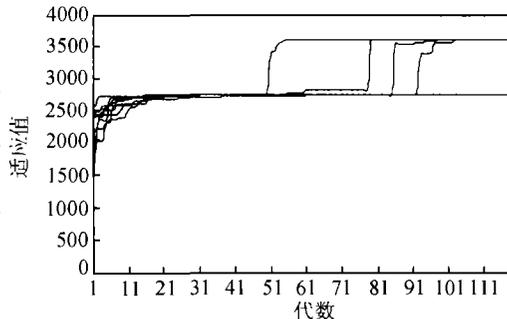


图 4 NiH 函数 GA 进化过程

分别采用均值规范、可调规范非单调适应值变换, $\alpha = 0.33$, 非单调适应值变换范围 $\beta = 2/3$, 其他参数保持不变, 重复实验 120 次, 计算结果如表 4 所示.

对于 NiH 函数, 采用均值规范或者可调规范变换, GA 搜索效果基本上达到了 90% 以上.

表 4 NiH 函数 GA 定向选择与非单调变换的计算结果^{a)}

	定向选择		均值规范变换		可调规范变换	
	$n = 60$	$n = 100$	$n = 60$	$n = 100$	$n = 60$	$n = 100$
收敛到最优解的代数	81.21	83.37	84.15	74.15	81.61	78.22
收敛到最优解的概率/%	57.50	64.17	90.83	98.33	88.89	91.67

a) 误差范围: ≤ 0.1

当 NiH 函数的欺骗性进一步加剧时,定向选择下 GA 搜索到全局最优解的概率迅速下降,而采用适应值非单调变换,GA 可以保持良好的搜索性能.

参 考 文 献

- 1 Richard K B, et al. *Foundation of Genetic Algorithms*. 4. San Francisco, Calif: Morgan Kaufmann Publishers Inc, 1997
- 2 Mitchell M. *An introduction to Genetic Algorithms*. Cambridge: The MIT Press, 1996
- 3 Goldberg D E, et al. Optimizing global-local search hybrids. In: Banzhaf, W. et al. eds. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-99)*. San Fransisco, CA: Morgan Kaufmann, 1999, 26 ~ 41
- 4 恽为民,等. 遗传算法的全局收敛性和计算效率分析. *控制理论与应用*, 1996, 13(4): 455
- 5 Goldberg D E, et al. A comparative analysis of selection schemes used in genetic algorithms. In: Rawlins G J E, ed. *Foundations of Genetic Algorithms*, San Mateo: Morgan Kaufmann Publishers, CA, 1991: 69 ~ 93
- 6 Miller B, et al. Genetic algorithms, selection schemes, and the varying effects of noise. *Evolutionary Computation*, 1996, 4(2): 25
- 7 徐宗本,等. 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): 363
- 8 Whitley D. Fundamental principles of deception in genetic search. In: Rawlins G J E, ed. *Foundations of Genetic algorithms*, Morgan Kaufmann, 1991, 221 ~ 241
- 9 Kuo T, et al. A genetic algorithm with disruptive selection. *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, 1996, 26(2): 65