

一种考虑环境作用的协同免疫遗传算法

夏虎, 庄健, 王立忠, 于德弘

(西安交通大学机械工程学院, 710049, 西安)

摘要: 在综合考虑了环境对生物进化的影响、免疫算法的结构以及遗传算法部分算子的基础上, 提出一种考虑环境作用的协同免疫遗传算法(ESIGA), 以实现提高算法搜索速度和全局搜索能力的目标. 在该算法中, 设计了克隆环境演化算子和自适应探索算子, 并构造了 3 个子种群协同进化以发挥克隆环境演化算子的影响, 从而提高算法的全局搜索能力. 引入的自适应探索算子和克隆环境演化算子, 使算法具备了一定的学习能力, 可加速搜索和防止早熟. 构建的主种群和协同种群相互影响, 使得算法对环境具有改良能力, 加强了克隆环境演化算子的性能, 而精英种群则加强了算法在优质个体邻域的搜索能力. 采用 13 个常用无约束优化问题测试函数对算法做了检验, 测试数据表明: ESIGA 算法与正交遗传算法相比, 其搜索速度要快于正交遗传算法 1~2 倍, 并能够处理 1 000 维的高维优化问题.

关键词: 环境; 免疫; 遗传算法; 协同

中图分类号: TP391. 41 **文献标志码:** A **文章编号:** 0253-987X(2009)11-0080-05

Environment-Based Synergic Immune Genetic Algorithm

XIA Hu, ZHUANG Jian, WANG Lizhong, YU Dehong

(School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: The function of environment effecting creature evolution, the process of immune clone algorithm and some operators of genetic algorithm are considered synthetically. An environment-based synergic immune genetic algorithm (ESIGA) is presented to improve the searching rate and global searching ability of optimal algorithm. The clone-environment operator and self adaptive search operator are designed to make the algorithm possess learning capability to accelerate the searching rate and avoid prematurity. And three groups of sub-population are used for improving the global searching ability, where the interaction between main sub-population and synergic sub-population improves the environment and exerts clone-environment operator effect, and the elitist sub-population enhances the neighborhood searching. Then 13 classical benchmarks of non-constraint optimal problem are executed by the proposed algorithm. The test results demonstrate that ESIGA searching rate is twice as QGA/Q. ESIGA is suit to deal with 1 000-dimension non-constraint optimal problems.

Keywords: environment; immune; genetic algorithm; synergy

人工免疫系统和进化计算是近十几年来人工智能众多研究领域中具有代表性的 2 个方向, 已经被广泛应用于多个领域^[1-2], 但传统的进化计算仍然存在着收敛速度慢、数值优化精度不高及早熟^[3-4]等缺

点, 而人工免疫系统却对遗传算法性能的提高具有重要的启迪作用. 目前, 免疫遗传算法虽然对遗传算法的性能有一定的提高^[5-7], 但是这 2 种算法的组合还不够深入, 性能提高有限且缺乏对高维复杂优化

问题的实验验证.

在进化过程中,一个物种在其所处环境中得以长期生存,其免疫系统必须有足够的适应性,对环境中的各种不利因素(如各种病毒)具有足够的抵抗力,而不适应环境的个体却无法抵抗某些病毒的攻击,会逐渐被淘汰.因此,生存下来的个体可以有效地抵御病毒的攻击,适应性较强.研究表明^[8-10],不同种族人群或生活在不同区域的同种族人群对一些疾病(如疟疾)的免疫能力有显著的差别,这是因为影响免疫系统的基因在不同人群中的组织结构或基因频率不同,这也是不同环境给进化以一定选择压力的结果.鉴于这种环境作用和现有免疫遗传算法的不足,本文不仅设计了新算子,还构造了多个种群协同进化,从而得到一种考虑环境作用的协同免疫遗传算法(ESIGA),它体现了环境作用,提高了抗体的寻优能力.

1 算法定义

无约束优化问题可以定义为如下求最小值问题的形式

$$\min\{f(\mathbf{X}) \mid \mathbf{X} \in \mathbf{R}^m\} \quad (1)$$

式中: $f(\mathbf{X})$ 是目标函数, $\mathbf{X}=(x_1, x_2, \dots, x_m)^T$ 为在 $\mathbf{R}^m$ 空间的列向量. $\mathbf{X}$ 也称为基因,即为遗传算法所述个体和免疫算法所述抗体, $x_i=c_{i1}, c_{i2}, \dots, c_{iN_i}$ 称为子基因, c_{ij} 称为基因位.本文采用十进制编码,所有变量归一化在 $[0, 1]$ 区间,所以 $c_{ij} \in \{0, 1, 3, 4, 5, 6, 7, 8, 9\}$, x_i 可译码为

$$x_i = c_{i1}10^{-1} + c_{i2}10^{-2} + \dots + c_{iN_i}10^{-N_i} \quad (2)$$

抗体种群空间向量 $\mathbf{A}_i=(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{N_i})^T$,其中 N_i 为第 i 个种群的规模.对于无约束优化问题,本文采用目标函数值作为代价函数,即

$$C(\mathbf{X}) = f(\mathbf{X}) \quad (3)$$

排序策略为目标函数的值小,即亲和度大的抗体放在前面.本文所提算法包含主种群、协同种群和精英种群,3个种群共有的操作定义如下.

(1)免疫疫苗选择 T_c^s :将每代亲和度最高的抗体认为是疫苗,因此该操作可以描述为如下的形式(其中 t 为进化代数)

$$T_c^s(\mathbf{A}_t) = \max_{j \in [1, N_i]} C(\mathbf{X}_{jt}) \quad (4)$$

(2)将 Θ 表示为克隆操作,具体为 $\Theta(\mathbf{A}_t)=[\Theta(\mathbf{X}_1), \dots, \Theta(\mathbf{X}_{N_i})]^T$,其中 $\Theta(\mathbf{X}_{N_i})=\mathbf{I}_k \times \mathbf{X}_i$, $\mathbf{I}_k$ 是 k 维单位行矩阵, k 为克隆规模.克隆后 $\mathbf{A}_t$ 种群包括:①克隆产生的集合,记做 $\mathbf{A}_t^\Theta$;②未被克隆的原

始集合 $\mathbf{A}'_t$.

(3)更新 T_c^s :如果进化后群体中最大亲合度的个体优于母体的亲合度,则由其来替换母体.

1.1 主种群 $\mathbf{A}_1$ 的算子

主种群的算子主要有以下几种

(1)克隆交叉算子 T_c^C 只作用于克隆群体 $\mathbf{A}_1^\Theta$ 上,首先根据 $\mathbf{A}'_1$ 中抗体亲和度的大小降序排列,得到 $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{N_1}$.选取排在前面的70%的抗体对应的克隆个体进行双点交叉操作,而对排在后面的30%的抗体与主种群的免疫疫苗进行交叉操作,其定义为

$$T_c^C(\mathbf{X}_i^\Theta, \mathbf{X}_j^\Theta) = \begin{cases} c_n^\Theta \Leftrightarrow c_n^\Theta & i, j \in [1, 0.7N_1] \\ n = p, p+1, \dots, q; p, q \in [1, m] \\ c_n^\Theta + (T_c^s(\mathbf{A}_{1t}) - c_n^\Theta)\omega & i, j \in [0.7N_1, N_1] \\ n = 1, 2, \dots, m; \omega \in [0, 1] \end{cases} \quad (5)$$

克隆交叉算子中随机2个个体的交叉操作提高了种群的多样性,而与免疫疫苗进行的交叉操作使得抗体按梯度信息进行搜索,增加了当前最优解附近的抗体分布密度.

(2)克隆环境演化算子 T_c^a 对 $\mathbf{A}_1^\Theta$ 中的抗体做基于环境作用的变异操作,即

$$\left. \begin{aligned} T_c^a(\mathbf{X}_t^\Theta) &= \mathbf{X}_t^\Theta + R_1 E^t & \mathbf{X}_t^\Theta \subseteq \mathbf{A}_{1t}^\Theta \\ E^t &= R_2 E^{t-1} + R_3 (T_c^s(\mathbf{A}_{1t}) - \mathbf{X}_t^\Theta) \end{aligned} \right\} \quad (6)$$

式中: E 表示环境变量; R_1 表示学习系数; R_2 表示遗忘系数; R_3 表示修改系数.

克隆环境演化算子的目的有:①在群体中较优个体周围探索前进方向,以提高算法效率;②通过环境变量获取进化过程中的历史经验,使算法在一定程度上具有学习能力.

(3)克隆直接变异算子 T_c^m 针对 $\mathbf{A}_1^\Theta$ 中的抗体进行变异,这相当于在单一抗体周围产生一个变异群体,进行局部试探性搜索.克隆直接变异为

$$T_c^m(c_n^\Theta) = c_n^m; f_p(p_m) = 1 \quad (7)$$

式中: $f_p(p_m)$ 为贝努力试验函数; p_m 为变异概率; c_n^Θ 为相应抗体的第 n 个子基因; c_n^m 为变异后的子基因.

(4)自适应探索算子 T_c^e 对免疫疫苗的克隆个体在某些维上添加可控扰动,引导种群的寻优方向.将该算子描述为

$$T_c^e(c_n^\Theta) = c_n^\Theta \pm r_1 j \quad (8)$$

$$j \in Z^+; n = q, q \pm 1, \dots, q \pm (s_1 - 1)$$

$$q \in [1, m]$$

$$\text{若 } n > m, n = n - m$$

$$\text{若 } n < 1, n = n + m$$

$$\text{若 } r_1 < \frac{1}{5(u_n - d_n)}, \text{ 则 } r_1 = \frac{\omega}{5(u_n - d_n)}$$

$$\text{否则 } r_1 = 0.5\omega/j$$

$$\omega \in [0, 1)$$

式中: s_1 为连续选中的子基因个数; r_1 为扰动步长; j 为扰动调节范围; u_n, d_n 分别为对应维变量的上、下限。

1.2 协同种群 A_2 的算子及操作

协同种群的算子及操作主要有以下几种。

(1) 克隆变异算子 T_c^* , 将每个抗体的克隆群体分为: ①一般变异, 以保持种群的多样性; ②按一定概率将免疫疫苗替换, 以加快搜索速度, 表达式为

$$T_c^*(c_n^o) = \begin{cases} [0, 1), j = 1, 2, \dots, k_1 \\ T_c^*(A_{2t}, c_n), j = k_1 + 1, k_1 + 2, \dots, k \end{cases} \quad (9)$$

式中: k 表示克隆规模, 本文取 $k=5, k_1=3$; $T_c^*(A_{2t}, c_n)$ 表示协同种群中免疫疫苗的第 n 个子基因。

(2) 借鉴文献[11]中的倒位算子, 本文设计了克隆倒位算子 T_c^* , 以加强个体中各个子基因内部的信息交互, 其表达式与文献[11]中的一致, 只是倒位概率 $p_c=0.3$ 。

(3) 遗传交叉算子 T_c^* 与克隆交叉算子类似, 但作用于 A'_2 时可参照式(5)。

(4) 遗传变异算子 T_g^m 作用于 A_2^1 , 目的是为了防

$$\left. \begin{aligned} T_g^m(c_n^a) &= \omega \quad \omega \in [0, 1), f_p(p_g) = 1 \\ p_g &= e^{i/N_2}/m \quad i \in [1, N_2] \end{aligned} \right\} \quad (10)$$

(5) T_c^* 的表达式可参考式(8)。

(6) 不对称信息的交互操作: 将主种群排在前面 10% 的优质个体替换协同种群中亲和度最差的 10% 的个体; 将协同种群排在前面的 5% 的优质个体替换主种群中亲和度最差的 5% 的个体。不对称信息交互是由于主种群是进化的主体, 因此性能优于协同种群。

1.3 精英种群 A_3 的算子及操作

精英种群包含有邻域交叉算子 T_c^* , 目的是针对一般进化算法在接近全局最优时收敛速度变慢的缺点, 从主种群和协同种群中各取前 10% 的优质个体

组成精英种群, 进行交叉操作, 交叉概率为 0.6, 并定义为

$$T_c^*(c_n^i, c_n^j) = \begin{cases} c_n^i = c_n^j + 2\omega(c_n^i - c_n^j) \\ c_n^j = c_n^i + 2\omega(c_n^j - c_n^i) \end{cases} \quad (11)$$

$$\omega \in [0, 1); i \neq j; i, j \in [1, N_3]; n \in [1, m]$$

该算子实际上是分别以 c_n^i, c_n^j 为中心, 以 $|c_n^i - c_n^j|$ 为半径进行的邻域搜索。

2 算法流程的描述

ESIGA 算法的流程具体描述步骤如下。

步骤 1 初始化主种群和协同种群, 设当前主种群抗体号 $i=1$, 协同种群抗体号 $j=1$, 进化代数 $t=0$ 。

步骤 2 对上述 2 个种群分别排序, 按式(4)规则分别选择各自种群的疫苗 $T_c^*(A_i)$ 。

步骤 3 判断是否满足停止规则, 如果是则得到最终优化结果, 否则: ①对所有个体执行克隆操作, 克隆规模为 1, 并按式(5)执行克隆交叉操作且更新原始抗体; ②判断当前抗体号 i 是否大于主种群规模 N_1 , 如果是跳转到④, 否则跳转到③; ③选择 X_i , 按式(6)、式(7)依次执行克隆环境演化操作和克隆直接变异操作, 并更新原始抗体, $i=i+1$, 返回②; ④对主种群进行排序操作, 然后按式(8)执行自适应探索操作并更新疫苗, 再以 10% 的优质抗体更新协同种群; ⑤判断当前抗体号 j 是否大于协同种群规模 N_2 , 如果是跳转到⑦, 否则跳转到⑥; ⑥选择 X_i , 依次执行克隆变异操作和克隆倒位操作, 并更新原始抗体, $j=j+1$, 返回⑤; ⑦参照式(5)执行遗传交叉操作, 按式(10)执行遗传变异操作, 排序后再执行自适应探索操作并更新疫苗, 然后以 5% 的优质抗体更新主种群, 最后对协同种群中亲和度最低的 15% 的个体用随机生成的新个体替换; ⑧按式(11)执行邻域交叉操作并选择精英种群中的疫苗个体。

步骤 4 从所有种群中选择最优的疫苗作为全体的疫苗, 更新当前代数 $t, t=t+1$, 返回步骤 3。

3 实验结果与分析

为了测试算法的性能, 本文选取 13 个常用的无约束优化测试函数, 函数序号与文献[12]相对应。ESIGA 算法中主种群和协同种群的规模均为 $N_1 = N_2 = 20$, 自适应探索算子的参数为: 当函数维数 $m=30$ 和 100 时, $s_1=1, j=2$; 当 $m=500$ 时, $s_1=1, j=5$; 当 $m=1\ 000$ 时, $s_1=2, j=5$, 其余参数的选取

如文中第 1 节所述. 算法的停止准则为:连续 50 代最优值不变或者误差 ϵ 小于规定值,本文算法对各个测试函数独立运行 50 次(在 500 和 1 000 维时独立运行 20 次),并统计函数的平均调用次数 F_t 和平均最优值 F_m .

表 1 为 ESIGA 算法与正交遗传算法(OGA/Q)的对比结果,其中 f_7 的停止准则为函数的调用次数不超过 3×10^5 . 结果显示,除 f_3 、 f_{13} 以外,其余函数的平均调用次数或平均最优值均明显优于正交遗传算法,可见 ESIGA 算法的改进效果是明显的.

$m=500 \sim 1\,000$ 时为高维函数,表 2 为 ESIGA 算法对部分函数取 500 和 1 000 维的测试结果. 在一般情况下,平均每维的函数调用次数应低于 5 000 次,算法才有效. 本文算法在较高精度($\epsilon=10^{-4}$)下,函数的平均调用次数远低于限定值,可见 ESIGA 算法在处理高维无约束优化问题时具有较好的性能. 表 3 说明了克隆环境演化算子和协调种群对算法性能的贡献,其中各个参数的设置和停止条件与表 1 相一致,所测函数的维数也一致. 表 3 说明合理利用进化过程中的历史经验可以指导函数优化的

表 1 2 种算法对部分函数测试结果的对比

函数	ESIGA		OGA/Q		m	$F_{\min}$
	F_t	F_m	F_t	F_m		
f_2	76 118	0	224 710	0	30	0
f_4	52 883	0	134 000	0	30	0
f_{11}	61 102	0	112 559	0	30	0
f_{13}	125 189	0	112 612	0	30	0
f_{14}	81 026	0	112 576	0	30	0
f_3	111 459	0	112 421	4.440 ^①	30	0
f_5	44 896	0	134 556	6.019 ^②	30	0
f_6	60 933	0	134 143	1.869 ^③	30	0
f_8	73 145	22	190 031	4.672 ^④	100	0
f_{12}	27 990	0	112 652	6.301 ^⑤	30	0
f_1	69 497	-12 569.487	302 166	-12 569.454	30	-12 569.487
f_7	300 167	-97.093	302 773	-92.830	100	-99.600
f_9	32 568	-78.331	245 930	-78.300	100	-78.333

①~⑤:表示 F_m 的精度分别为 10^{-16} 、 10^{-6} 、 10^{-4} 、 10^{-7} 和 10^{-3} ; $F_{\min}$:函数理论最小值.

表 2 ESIGA 算法对部分高维函数测试的结果

函数	$m=500$		$m=1\,000$		$F_{\min}$
	F_t	$10^{-5} \times F_m$	F_t	$10^{-5} \times F_m$	
f_2	428 691	8.442	893 709	8.994	0
f_3	675 890	8.493	1 490 256	9.953	0
f_4	575 073	9.833	1 280 071	9.411	0
f_5	651 045	8.867	1 443 192	9.872	0
f_6	1 017 131	9.914	2 435 544	9.964	0
f_{11}	667 081	9.918	1 483 176	9.947	0
f_{12}	416 641	9.058	986 582	9.956	0

进行,而克隆环境演化算子则具有加速函数收敛和防止早熟的性能. 同时,表 3 还说明协同种群具有改良环境的能力,但会增加函数的调用次数. 当主种群所在环境本身较为优良时,协同种群的作用会减弱(如 f_{13} 、 f_{14}),将增加冗余的函数调用次数,但总的来说,协同种群与主种群的交互信息,构成了一个封闭的环境调控系统,使算法能保持较高的收敛精度.

自适应探索算子的目的是在优质个体的某些维上做距离可变的探索,它在长距离搜索时有利于跳出局部极值,而在短距离时有利于加强邻域搜索. s_1 和 j 的取值与维数有关,同时还应兼顾对函数调用次数的影响.

表 3 ESIGA 算法的性能分析

函数	ESIGA ^①		ESIGA ^②	
	F_t	F_m	F_t	F_m
f_2	101 298	0	88 329	0
f_3	208 881	0		
f_6	99 261	0	66 062	0
f_{11}	81 671	0	58 461	0
f_{12}	32 567	0	29 659	0
f_{13}	253 903	4	75 306	0
f_{14}	113 688	0	70 054	0
f_1	94 459	-12 559.783	113 943	-12 569.487
f_4	105 136	0.004	80 470	0.011
f_5	84 235	0.001	78 774	0.046
f_7	300 156	-97.062	300 188	-92.867
f_8	69 441	0.006	343 700	0.001
f_9	32 733	-78.329	81 163	-78.328

注： $f_2, f_3, f_6, f_{11}, f_{12}, f_{13}, f_{14}$ 的最优值精度均为 10^{-8} ；①表示无克隆演化算子；②表示无协同种群。

4 结 论

本文提出了一种考虑环境作用的协同免疫遗传算法,克服了遗传算法的不足,不仅改进了免疫克隆算法与遗传算法中的基本算子,增强了局部搜索能力,还引入了自适应探索算子和克隆环境演化算子,使算法具备了一定的学习能力,可加速搜索和防止早熟. 本文构建的主种群和协同种群相互影响,使得算法对环境具有改良能力,加强了克隆环境演化算子的性能,而精英种群则加强了算法在优质个体邻域的搜索能力. 实验证明,该算法在解决无约束优化问题上具有良好的性能.

参考文献:

[1] 何珍梅,徐雪松. 人工免疫系统研究综述 [J]. 华东交通大学学报,2007,24(4):79-83.
HE Zhenmei, XU Xuesong. The overview of artificial immune system research[J]. Journal of East China Jiaotong University,2007,24(4):79-83.

[2] 葛继科,邱玉辉,吴春明,等. 遗传算法研究综述 [J]. 计算机应用研究,2008,25(10):2911-2916.
GE Jike, QIU Yuhui, WU Chunming, et al. Summary of genetic algorithms research[J]. Application Research of Computers,2008,25(10):2911-2916.

[3] GOLDBERG D E. Genetic algorithms in search optimization and machine learning [M]. New York, USA: Addison-Wesley, 1989: 1-83.

[4] 赵金帅,鲁瑞华. 一种用于防止早熟收敛的改进遗传算法 [J]. 西南大学学报(自然科学版),2008,30(1): 156-159.

ZHAO Jinshuai, LU Ruihua. An improved genetic algorithm to prevent premature convergence [J]. Journal of Southwest University (Natural Science Edition),2008,30(1):156-159.

[5] 谢刚,武斌,谢克明. 基于免疫遗传算法的 TSP 优化问题求解 [J]. 太原理工大学学报,2007,38(3):199-201.
XIE Gang, WU Bin, XIE Keming. Traveling salesman problem's solution based on immune genetic algorithm [J]. Journal of Taiyuan University of Technology, 2007,38(3):199-201.

[6] 高岩,位耀光,付冬梅,等. 免疫遗传算法的研究及其在函数优化中的应用 [J]. 微计算机信息,2007,23(2):183-184.
GAO Yan, WEI Yaoguang, FU Dongmei, et al. Research of artificial genetic algorithm and its application in function optimization1 [J]. Microcomputer Information, 2007,23(2):183-184.

[7] 马西庚,李媛媛,戴永寿. 一种用于优化计算的自适应免疫遗传算法 [J]. 计算机工程与应用,2006(12):41-43.
MA Xigeng, LI Yuanyuan, DAI Yongshou. An adaptive immune-genetic algorithm applied to optimization computation [J]. Computer Engineering and Applications,2006(12):41-43.

[8] MARCANO T J, MORGADO A, TOSTA C E, t al. Cross-sectional study defines difference in malaria morbidity in two Yanomami communities on Amazonian boundary between Brazil and Venezu [J]. Mem Inst Oswaldo Cruz,2004,99(4):369-76.

[9] TAYLOR A W R. A model of development of acquired immunity to malaria in humans living under endemic conditions [J]. Med Hypotheses,2002,58(2):56-148.

[10] PAGANOTTI G M, BABIKER H A, MODIANO D, et al. Genetic complexity of plasmodium falciparum in two ethnic groups of Burkina Faso with marked differences in susceptibility to malaria [J]. Am J Trop Med Hyg, 2004,71(2):8-173.

[11] 杜海峰,公茂果,焦李成,等. 用于高维函数优化的免疫记忆克隆规划算法 [J]. 自然科学进展,2005,14(8):925-933.
DU Haifeng, GONG Maoguo, JIAO Licheng, et al. Immune memory clonal programming algorithm for high-dimensional function numerical optimization[J]. Progress in Natural Science, 2005,14(8):925-933.

[12] LEUNG Yiuwing, WANG Yuping. An orthogonal genetic algorithm with quantization for global numerical [J]. IEEE Transactions on Evolutionary Computation, 2001,5(1):41-53.