

# 基于生态竞争模型的遗传强化学习\*

曹先彬<sup>1</sup> 高隽<sup>2</sup> 王煦法<sup>1</sup>

<sup>1</sup>(中国科学技术大学计算机科学与技术系 合肥 230027)

<sup>2</sup>(合肥工业大学计算机与信息系 合肥 230009)

**摘要** 未成熟收敛和收敛速度慢是目前遗传算法的明显缺点,借鉴生物在环境生态系统中的生长模式,文章提出一种生态竞争模型。该模型认为,竞争行为在生物的成长中占有十分重要的地位,在子群内实现了个体层次的先天遗传进化和后天竞争学习,在种群层次实现进一步的竞争强化学习。实验结果显示了该模型在解决收敛性问题时的有效性。

**关键词** 遗传算法,收敛性,生态竞争,强化学习,函数优化。

中图法分类号 TP18

目前,以遗传算法(GA)为代表的进化算法发展很快,但它的理论基础还很不完善。在应用中体现出的最明显的缺点是它的收敛性问题,包括收敛速度慢且存在未成熟收敛<sup>[1,2]</sup>。其根本原因就在于,进化算法是从群体的全局散布性出发,而遗传进化是采用定向性操作算子,这两者是相互冲突的。前者保证GA有较大的搜索范围,使群体不易落入局部优化区域,后者又强制将群体引入某一特定子空间。正因如此,目前有很多相关的改进工作,主要集中在3个方面<sup>[3~5]</sup>:①调整算法中的操作参数,如遗传操作概率、集团大小、适应度函数定标等;②进一步维持算法进行中的个体多样性,如增加种群规模、增大配对个体距离,但这些都是经验方法,既增加了计算量,又降低了算法的隐并行性;③引入子群(sub-population)概念,每个子群独立进行遗传进化。这种方法可使未成熟收敛局部化,但仍然不易获得全局收敛,所以仍未解决这一问题。要想更好地加以解决,有必要引进新的思想。

本文在分析环境生态系统的基础上提出一种生态竞争模型,认为生物在不同层次上都存在着相互协作与竞争。该模型在个体层次上实现了生物的先天遗传进化和后天竞争学习,同时又在种群层次上实现了子群之间的进一步竞争强化学习。基于本模型的算法能在很大程度上加速收敛,解决未成熟收敛的问题。

## 1 生态竞争模型

作为一种搜索性算法,GA是通过模拟生物自然选择机制而提出的,但它实际上只考虑了生物进化机制的一个侧面。我们从环境生态系统中得到启示,认为生物在自然界中的成长过程至少应包括以下几点:

- (1) 生物进化不仅要考虑个体的进化行为,而且要考虑生物集团的进化行为;
- (2) 生物在生态环境中进化时,其活动范围是局部的和可变的;
- (3) 竞争在环境生态系统中占有很重要的地位,这种竞争既存在于个体之间,也存在于群体之间,竞争的强度是动态的;并且这种竞争绝对不仅是GA的自然淘汰策略,而是一种相互协作与竞争的关系;
- (4) 生物进化是一个在环境生态系统中结合先天遗传进化和后天竞争学习的复杂过程。

对应于上面几点,GA主要模拟了个体进化行为和生物的先天遗传进化,竞争主要体现在取代

\* 本文研究得到国家自然科学基金和安徽省“九五”重点攻关项目基金资助。作者曹先彬,1969年生,博士,讲师,主要研究领域为进化计算、神经网络、专家系统。高隽,1963年生,副教授,主要研究领域为神经网络。王煦法,1948年生,教授,博士生导师,主要研究领域为计算智能,人工生命。

本文通讯联系人:曹先彬,合肥230027,中国科学技术大学计算机科学与技术系

本文1998-04-20收到原稿,1998-06-29收到修改稿

(replacement)机制上.引入子群的 GA 又模拟出生物活动范围的局部性,但由于未考虑竞争和后天学习的重要性,目前的 GA 搜索能力弱,而且全局收敛性差.

在模拟环境生态系统方面,文献[6]提出的 ECO(ecological)模型是相对较好的工作,ECO 模型完全模拟了上面的(1)、(2)和(4)的先天遗传进化,但它同样未意识到竞争的复杂性和后天学习的必要性.

我们提出如下的生态竞争模型 ECM(ecological competition model),如图 1 所示. ECM 考虑到生态环境实际上有两个层次:一个是局部环境之间,另一个是局部环境内的个体之间.同样引入子群概念,将完整种群分为若干子群,子群的初始数目和规模可随机设定.在每一子群内首先执行下面的生态竞争算法 ECA(ecological competition algorithm),使子群能很快收敛,并且子群规模是自适应调整的.在每一子群都达到收敛后,由于完整种群的全局收敛点依存于全部子群的局部收敛点,所以,ECM 又采用一种能反映所有局部收敛点间相互竞争关系的强化学习,使全局收敛点体现出来. ECM 的算法框架如图 2 所示,下面分别说明.

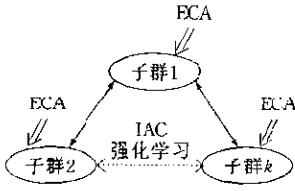


图1 生态竞争模型ECM

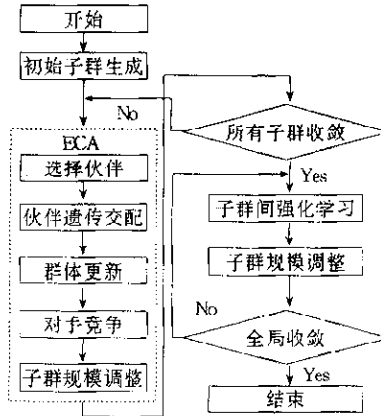


图2 ECM的算法框架

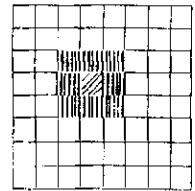


图3 初始子群的组成

### 1.1 初始子群生成

设种群规模为  $M$ ,随机生成  $k$  个子群,每个子群由一个代表个体和其周围的若干个体组成.若以 2 维  $8 \times 8$  栅格描述,则首先在栅格中随机选择一格作为子群的代表个体,而代表个体边缘的 8 个个体是它的局部环境,这 9 个个体组成一个子群,如图 3 所示.图中短斜线的栅格为代表个体,短竖线的栅格为其局部环境.

### 1.2 生态竞争算法

ECA 根据配对个体的行为将个体间的关系分为伙伴和对手两种.伙伴个体通过遗传交配生成新个体,这是 ECM 中个体的先天遗传进化.对手个体之间同时存在相互激活与竞争行为,结果是对手个体的表现型被动态地自适应调整,这是个体的后天竞争学习. ECA 对每一子群进行以下操作:

Step 1. 选择伙伴个体.以代表个体为父本,再按照门当户对原则,选出一个最接近于父本的个体作为母本,父本和母本成为一对伙伴.

Step 2. 伙伴个体间进行交叉、变异等遗传操作,生成新个体.

Step 3. 运用取代机制进行群体更新.

Step 4. 对手个体竞争.

在一个子群内,个体之间是相互联系、相互影响的,它们既可能相互协作,使各自的表现型都得到提升,也可能相互竞争,一方抑制另一方的成长.为体现对手个体间的相互协作与竞争关系,自适应修正对手个体的动作值,我们将个体表现型看作个体的动作值,并引进 IAC 神经网络中的 Grossberg 公式来完整描述对手个体间竞争的进展<sup>[7]</sup>.Grossberg 公式描述的是神经元之间相互兴奋与抑制关系,在将神经元与个体对应后,我们提出以下公式:

$$\Delta f = (\max - f)e - (f - \min)c - \text{decay}(f - \text{rest}). \tag{1}$$

其中  $f$  是个体的动作值,可直接对应适应度函数.  $\max, \min$  分别为动作的最大、最小值.  $e$  为其他个体对本个体的兴奋性输入,它使本个体的动作值朝着最大值的方向增加.  $c$  为其他个体对本个体的抑制性输入,它使本个体的动作值朝着最小值的方向减少. 规定第  $i$  个个体接收的兴奋性输入为  $e_i = \sum_j f_j$ , 其中的  $f_j \geq (\max - \min)/2$ .

若  $f_j < (\max + \min)/2$ , 则兴奋性输入转化为抑制性输入,这时  $c_i = \sum_j f_j$ . 另外,  $rest$  为个体在无输入时的基本值,称之为静息值,一般规定  $\min \leq rest \leq 0$ .  $decay$  是衰减率,它使个体的动作值趋于静息值水平,体现个体生长期间的时间衰减关系,  $0 \leq decay \leq 1$ .  $rest$  和  $decay$  在算法中是两个可调参数,如取  $rest = \min/2, decay = 0.4$ .

可以看出,在公式(1)中,一个输入的作用不是确定性的,当个体的动作值较低时,它可能是兴奋的,反之,则可能是抑制的. 其动力学系统体现出了相互协作与竞争的特性:考虑竞争中的两个个体  $a, b$ , 假设它们的兴奋性输入分别为  $e_a, e_b$ , 但在强度上有差别,  $e_a > e_b$ , 并设其他个体对  $a, b$  的抑制性输入都为  $r$ , 个体  $a$  对  $b$  的抑制性输入为  $c_b = f_a$ , 个体  $b$  对  $a$  的抑制性输入为  $c_a = f_b$ , 则

$$\begin{aligned} \Delta f_a &= (\max - f_a)e_a - (f_a - \min)(r + c_a) - decay(f_a - rest) \\ &= (\max - f_a)e_a - (f_a - \min)(r + f_b) - decay(f_a - rest), \\ \Delta f_b &= (\max - f_b)e_b - (f_b - \min)(r + c_b) - decay(f_b - rest) \\ &= (\max - f_b)e_b - (f_b - \min)(r + f_a) - decay(f_b - rest). \end{aligned}$$

若初始  $f_a = f_b$ , 则第 1 次竞争时,  $\Delta f_b < \Delta f_a$ , 使个体  $b$  在竞争中处于劣势,而个体  $a$  由于在竞争初始具有优势,所以在竞争过程中将逐步强化这种优势. 但随着  $f_a, f_b$  的不断变化,有可能  $\Delta f_a < \Delta f_b$ , 这时又使个体  $b$  具有优势.

在公式(1)中,令  $\Delta f = 0$ , 即

$$(\max - f)e - (f - \min)c - decay(f - rest) = 0,$$

则 
$$f = (\max \cdot e + \min \cdot c + decay \cdot rest) / (e + c + decay),$$

即当某一个体的动作值等于这个比例时,它就达到平衡;当然,这并不代表所有个体都达到了平衡.

Step 5. 子群规模调整.

由代表个体在其局部环境中的地位重新决定子群的规模. 以 2 维栅格为例,调整策略为:

- (1) 在子群内选择一个最佳个体作为代表个体;
- (2) 计算代表个体适应度  $f_i$  与子群平均适应度  $\bar{f}$  的比例  $f_i/\bar{f}$ ;
- (3) 若  $f_i/\bar{f} < 1$ , 则依次以概率  $1 - f_i/\bar{f}$  随机选择子群内边缘的个体,并将其删除,这样使子群规模缩小;
- (4) 若  $f_i/\bar{f} = 1$ , 则保持子群规模不变;
- (5) 若  $f_i/\bar{f} > 1$ , 设  $f_i/\bar{f} = m, n$ , 则依次以概率  $0, n$  随机选择子群外边缘的  $m$  个个体,并将它们加入子群,像这样使子群规模扩大.

1.3 子群层次上的强化学习

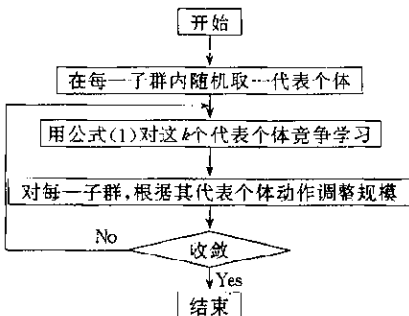


图4 ECM中子群层次上的强化学习

在所有子群都收敛以后,需要在多个局部收敛点中动态选择出全局收敛点,这时,子群内和子群间同样存在着相互刺激与竞争. 一方面子群内个体间的关系是兴奋性的,其作用是使子群中代表个体的动作(表现型)更特殊. 而子群间的联系是抑制性的,这可使低动作的子群代表个体受到抑制,高动作的子群代表个体得到强化,从而使全局收敛点体现出来,这就实现了子群之间的进一步竞争强化. 这一过程如图 4 所示.

1.4 ECM 的收敛能力

在 ECM 中,子群的引入使未成熟收敛局部化,对每一子群而言,由于规模相对较小,其收敛速度较快,这就是一些引入子群的改

进遗传算法的优点.同时,ECM又借鉴 Grossberg 对神经元间相互刺激与抑制关系的描述,对于群内的个体进行后天竞争学习,使个体的表现型被动态地自适应调整,从而大大加快了子群的收敛速度.子群规模调整策略的执行使子群规模自适应于环境,同时也使子群内个体越来越相似,ECO就采用了这一策略.另一方面,除 ECM 以外的其他算法都没有意识到了群之间的复杂竞争关系,而 ECM 通过对子群间竞争行为的分析,在种群层次上实现了子群之间的进一步竞争强化学习,使各个子群的行为在相互协作与竞争中得到动态调整,从而使全局收敛点体现出来.

## 2 实验

我们以函数优化问题为例,来考察本文提出的生态竞争模型及强化学习的收敛能力.将 GA 应用于函数优化的工作很多,其中 De Jong 的工作最有影响,我们选择他在文献[8,9]中提出的多峰函数  $F_5$ :

$$f_5(X_i) = 0.002 \times \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{ij})^6}, \quad -65.536 \leq x_i \leq +65.536.$$

De Jong 发现,SGA(standard genetic algorithm)及一些改进策略在收敛速度和全局收敛方面性能都不高.因此,他提出用“排挤因素模型(crowding factor model)” $R_s$ . $R_s$ 限定采用二进制数编码表示、赌轮选择机制、一点交叉和位点变异.在  $R_s$  中,当一个个体诞生时,便选择一个个体消亡,消亡个体为随机选取的  $n/CF$  个个体中与新生个体最相似的个体,其中  $n$  为群体规模, $CF$  是 De Jong 引入的一个参数.另外,De Jong 还使用了代沟值  $G = 0.1$ . De Jong 的研究表明,当  $CF = 2$  时,基于  $R_s$  的遗传算法对函数  $F_5$  获得了一个逼近全局最优的性能.

我们采用 ECM 来优化  $F_5$ .为评估不同算法的收敛性,在此采用 De Jong 提出的离线性能评估准则  $F_5(T)$ :

$$F_5(T) = \frac{1}{T} \sum_{t=1}^T \text{best} \{f_e(1), f_e(2), \dots, f_e(t)\}.$$

其中  $f_e(t)$  是第  $t$  代中相应于环境  $e$  的平均适应度函数.图 5 给出了基于 ECM 的结果,同时显示出 ECO 的结果和 De Jong 的结果.可以看出,ECM 能在很大程度上加速收敛.在算法执行过程中,我们可以观察到如第 1.4 节所分析的 ECM 的收敛能力,当算法趋于局部极值点时,ECM 通过个体间的竞争使个体的表现型发生变化,从而很快脱离局部极值点.表 1 给出采用 SGA, ECO 和 ECM 在优化  $F_5$  时的全局收敛性,表中列出的是进化 100 代时算法达到全局极值点的次数,表明 ECM 的全局收敛性能最佳.

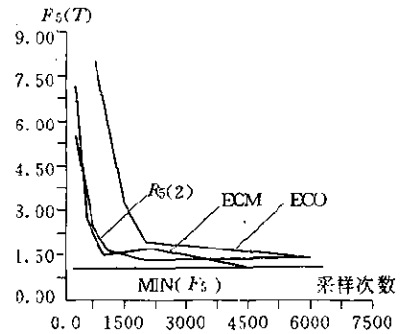


图5 ECM的收敛性能

表 1 3种方法的全局收敛性能比较

实验次数	SGA	ECO	ECM
10	7	10	10
50	37	41	47
100	61	75	92

## 3 结束语

以 GA 为代表的进化算法虽然取得了一些成功,但也存在收敛速度慢和未成熟收敛等特殊缺点.本文在分析环境生态系统的基础上,认为传统进化算法和目前的一些改进策略没有充分考虑生物在环境中的进化模式,从而提出一种生态竞争模型 ECM. ECM 强调竞争行为在生物进化过程中的重要性,这种竞争既存在于局部环境内的个体与个体之间,也存在于子群与子群之间,并且竞争的环境和强度是动态的.我们引进神经元的相互协作与竞争思想,在 ECM 中实现了个体之间的竞争和子群间的竞争,体现出生物的先天气质和后天强化学习思想.实验结果表明,ECM 的收敛速度很快,并更容易趋于全局收敛.

致谢 本文的研究工作得到国家自然科学基金资助, 此项目编号为 69671029。

### 参考文献

- 1 Goldberg D E. Genetic Algorithms in Search, Optimization and Machine Learning. Reading, MA: Addison-Wesley Publishing Company, 1989
- 2 Rudolph G. Convergence analysis of canonical genetic algorithms. IEEE Transactions on Neural Networks, 1994,5(1): 96~101
- 3 Schradoph N N *et al.* Dynamic parameter encoding for genetic algorithms. Machine Learning, 1992,9(1):9~21
- 4 Srinivas M *et al.* Adaptive probabilities of crossover and mutation in genetic algorithms. IEEE Transactions on Systems, Man and Cybernetics, 1994,24(4):656~667
- 5 Cohoon J P *et al.* A multi-population genetic algorithm for solving the k-partition problem on hyper-cubes. In: Belew R K ed. Proceedings of the 4th International Conference on Genetic Algorithms. San Mateo, CA: Morgan Kaufmann Publisher, Inc., 1991. 244~248
- 6 Yural Davidor. An ecological model for evolutionary computing. システム/制御/情報, 1993,31(8):468~474  
(Yural Davidor. An ecological model for evolutionary computing. System/Control/Information, 1993,31(8):468~474)
- 7 Rumelhart D E *et al.* Explorations in Parallel Distributed Processing. Cambridge, MA: MIT Press, 1988
- 8 De Jong. An analysis of the behavior of a class of genetic adaptive system [Ph. D. Thesis]. University of Michigan, 1975
- 9 陈国良, 王煦法等. 遗传算法及其应用. 北京: 人民邮电出版社, 1996  
(Chen Guo-liang, Wang Xu-fa *et al.* Genetic Algorithms and Its Applications. Beijing: People's Post and Telecommunications Publishing House, 1996)

## An Ecological Competition Model for Genetic Reinforcement Learning

CAO Xian-bin<sup>1</sup> GAO Jun<sup>2</sup> WANG Xu-fa<sup>1</sup>

<sup>1</sup>(Department of Computer Science and Technology University of Science and Technology of China Hefei 230027)

<sup>2</sup>(Department of Computer and Information Hefei University of Technology Hefei 230009)

**Abstract** Premature convergence and low converging speed are the distinct weaknesses of the genetic algorithms. Using the living things' growth pattern for reference, a new model called ECM(ecological competition model) is proposed, in which the competition is considered to be in important position. In the ECM model, the congenital genetic evolution and the postnatal competition learning on individuals' level are realized in each sub-population, moreover, the competition reinforcement learning on population level is realized. The experimental results show the ECM model's effectiveness.

**Key words** Genetic algorithms, convergence, ecological competition, reinforcement learning, function optimization.