

Article ID: 1000-8152(2001)03-0443-04

The Unified Method Analyzing Convergence of Genetic Algorithms *

GUO Guanqi

(College of Information Science and Engineering, Central South University · Changsha, 410083, P. R. China;
Department of Electromechanical Engineering, Yueyang Teachers' College · Hunan Yueyang, 414000, P. R. China)

YU Shouyi

(Department of Electromechanical Engineering, Yueyang Teachers' College · Hunan Yueyang, 414000, P. R. China)

Abstract: This paper models the homogeneous finite Markov chain of the best individuals in the populations. Based on the limit distribution of transition probability that any state on the chain transfers into the state subset containing all global optima, a precise definition of the global convergence of genetic algorithms is presented regardless of the topologic structures of search landscapes. Two unified criterion theorems judging the global convergence are proposed and proved strictly, which are independent of encoding schemes and selection mechanisms. The results of analyzing the convergence of different genetic algorithms illustrate that the unified criterion theorems are generally practical and convenient.

Key words: genetic algorithms; convergence; unified criterions

Document code: A

遗传算法收敛性分析的统一方法

郭观七

(中南大学信息科学与工程学院 · 长沙, 410083; 岳阳师范学院机电系 · 湖南岳阳, 414000)

喻寿益

(岳阳师范学院机电系 · 湖南岳阳, 414000)

摘要: 建立了种群中最佳个体的马尔可夫链模型, 定义了仅包含所有全局最优解的状态子集; 根据从任意状态转移至该子集的概率的极限分布, 给出了独立于搜索曲面拓扑结构的遗传算法全局收敛性的精确定义; 提出并严格地证明了与编码方式和选择策略无关的、统一的全局收敛性判据定理. 对几种不同的遗传算法进行全局收敛性分析的结果表明, 统一的判断方法具有普遍的适用性.

关键词: 遗传算法; 收敛性; 统一判据

1 Introduction

Genetic algorithms (GAs) are a class of probabilistic search and optimization algorithms based on the model of organic evolution. The convergence Analysis of GAs is still one of the most interesting theoretic research topics, some results under additional assumptions were reported in literature [1 ~ 3]. In this paper, we analyzed the global convergence by means of the Markov chain of the best individuals in the population. After a brief description of Markov chain, two fundamental theorems about the limit distribution are given in Section 2. Section 3 is devoted to derive two basic criterion theorems

for judging the global convergence of GAs representation and selection-independently. Finally, we give some examples that aim at illustrating how to apply the criterion theorems to judge the global convergence of GAs in Section 4.

2 Homogeneous finite Markov chain

A finite Markov chain is a realization of random variables $X(t)$ defined over a finite discrete state space S of cardinality $|S| = N$ and discrete time $t \in \mathbb{N}$, which describes a probabilistic transition track of $X(t)$ over the state space S and exhibits Markov dependence, i. e., for any $t, k \in \mathbb{N}$, s. t.,

* Foundation item: supported by the National Natural Science Foundation of China (59835170).

Received date: 1999-11-01; Revised date: 2000-01-15.

$$\begin{aligned}
 P(X(t+k) = i_{t+k} | X(0) = i_0, \\
 X(1) = i_1, \dots, X(t-1) = i_{t-1}, \\
 X(t) = i_t) = P(X(t+k) = i_{t+k} | X(t) = i_t),
 \end{aligned}$$

where $i_0, i_1, \dots, i_{t-1}, i_t \in S$. Assuming that the states are numbered from 1 to N , the transition probability from state $i \in S$ to $j \in S$ at step t can be represented by $p_{ij}(t)$. If $p_{ij}(t)$ is independent of t , i.e., for all $i, j \in S$ and $t_1, t_2 \in \mathbb{N}$, s.t., $p_{ij}(t_1) = p_{ij}(t_2)$, the Markov chain is said to be homogeneous.

The one step transition probabilities of a homogenous finite Markov chain can be described in a transition matrix $P = (p_{ij})$, where, for all $i, j \in S$, s.t., $p_{ij} \in [0, 1]$ and $\sum_{j=1}^N p_{ij} = 1$. The matrices with the above properties are called stochastic. Given an initial distribution p^0 as a row vector, the distribution of the Markov chain after t steps is $p^t = p^0 P$. Therefore, a homogenous finite Markov chain is completely determined by (p^0, P) , the limit distribution depends on the structure of matrix P .

For homogenous finite Markov chains, the following two theorems^[4] (To ease understanding, the expressions herein are slightly different from the original.) hold:

Theorem 1 Let P be a primitive stochastic matrix of order N , then P^k converges as $k \rightarrow \infty$ to a unique stable stochastic matrix which has nonzero entries regardless of the initial distribution, i.e., for all $i, j \in S$, s.t.,

$$P^\infty = \lim_{k \rightarrow \infty} (p_{ij}^k) = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}_{N \times 1},$$

where $\pi = (p_1, \dots, p_N)$ and for all $1 \leq j \leq N$, s.t., $p_j \neq 0$.

Theorem 2 Let P be a reducible stochastic matrix of order N as the structure $P = \begin{pmatrix} C & 0 \\ R & T \end{pmatrix}$, where C is a primitive stochastic matrix of order m and $R, T \neq 0$, then P^k converges as $k \rightarrow \infty$ to a unique stable stochastic matrix regardless of the initial distribution, i.e., for all $i, j \in S$, s.t.,

$$P^\infty = \lim_{k \rightarrow \infty} (p_{ij}^k) = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}_{N \times 1},$$

where $\pi = (p_1, \dots, p_m, 0, \dots, 0)$, when $1 \leq j \leq m$,

s.t., $p_j \neq 0$.

Theorem 1 and 2 serve as the mathematical foundation of analyzing the convergence of GAs.

3 Judging criteria of convergence

Assuming the feasible region of a global optimization problem is I of cardinality $N = |I|$, which denotes arbitrary finite set; A genetic algorithm with invariable population size n and crossover probability $p_c \in [0, 1]$ and mutation probability $p_m \in (0, 1)$ does not vary the selection mechanism in the process of running. Let us give the formal definition of some terms as follows:

Definition 1 Let

$P(t) = \{a_i(t) | i \in [1, n], t \in \mathbb{N}, a_i(t) \in I\}$ is the population at generation t , where $a_i(t)$ denotes an individual representing a solution of the feasible region; $f: I \rightarrow \mathbb{R}$ denotes a fitness function assigning real values to individuals; $I^* = \{a^* | a^* = \arg \max \{f(a) | a \in I\}\}$ is a subset in the feasible region I , each of which has the global maximum fitness; The best individual in the population at generation t is $a^*(t) = a_j(t) \in P(t)$, where $a_j(t)$ subjects to $f(a_j(t)) \geq f(a_i(t) \in P(t))$.

Under the action of mutation, crossover and selection operations, $a^*(t)$ and its fitness will change randomly over time. But there always exists at least one $a^*(t)$ in each population, therefore, as $t \rightarrow \infty$, whether or not $a^*(t)$ converges to the subset I^* implies the genetic algorithm is or not globally convergent, i.e.,

$$P(\lim_{t \rightarrow \infty} a^*(t) \in I^*) = I \Rightarrow P(a^* \in \lim_{t \rightarrow \infty} P(t)) = 1$$

The evolution process of $a^*(t)$ can be described by a homogeneous finite Markov chain, for notation convenience, such a Markov chain is simply said to be $a^*(t)$ -chain. Clearly, each state in the state space S of $a^*(t)$ -chain uniquely corresponds to a solution in the feasible region I , and $|S| = |I| = N$. Let all the states of S be ordered descendent by $f(a^*(t))$ and numbered from 1 to N . This leads to the definition and criterion theorems of global convergence of GAs as follows:

Definition 2 Let $P = (p_{ij})$ be the transition matrix of an $a^*(t)$ -chain with state space S , and S^* be the state subspace each of which corresponds respectively to a solution in the subset I^* . As $t \rightarrow \infty$, if and only if

$a^*(t)$ transfers from any state $i \in S$ into S^* with probability one, i.e., iff

$$\lim_{t \rightarrow \infty} \sum_{j \in S^*} p_{ij}^t = 1, \text{ where } i \in S, \quad (1)$$

the genetic algorithm converges to one of the global optima.

Theorem 3 If the transition matrix $P = (p_{ij})$ of the $a^*(t)$ -chain with state space S and the subspace S^* is a positive stochastic matrix, then the genetic algorithm does not converge to anyone of the global optima.

Proof Since every positive matrix is also a primitive one, it follows by theorem 1 that the limit distribution of P is a unique stable one without nonzero entries. Therefore, for any $i \in S$, s.t.,

$$\lim_{t \rightarrow \infty} \sum_{j \in S^*} p_{ij}^t = 1 - \lim_{t \rightarrow \infty} \sum_{j \in S-S^*} p_{ij}^t = 1 - \sum_{j=|S^*|+1}^{|S|} p_j < 1,$$

so that equation (1) is not fulfilled, the proof is completed.

Theorem 4 If the transition matrix $P = (p_{ij})$ of the $a^*(t)$ -chain with state space S and the subspace S^* is a reducible stochastic matrix as the structure $P = \begin{pmatrix} C & 0 \\ R & T \end{pmatrix}$, where C is a positive stochastic matrix of order $|S^*|$ and $R, T \neq 0$, then the genetic algorithm converges to one of the global optima.

Proof From Theorem 2, we can obtain that for all $i, j \in S$, s.t.,

$$P^\infty = \lim_{t \rightarrow \infty} (p_{ij}^t) = (\pi, \dots, \pi)_{1 \times N}^T,$$

where $\pi = (p_1, \dots, p_{|S^*|}, 0, \dots, 0)$, when $1 \leq j \leq |S^*|$, $p_j \neq 0$, and $\sum_{j=1}^{|S^*|} p_j = 1$. It follows directly that, for any $i \in S$, s.t.

$$\lim_{t \rightarrow \infty} \sum_{j \in S^*} p_{ij}^t = \sum_{j=1}^{|S^*|} p_j = 1.$$

Equation (1) is fulfilled and the proof is completed.

Theorem 3 and 4 can be applied directly to judge the global convergence of GAs if the structure of transition matrix of $a^*(t)$ -chain can be made certain.

4 Application examples

4.1 Proportional reproduction

The simple genetic algorithms^[5] use selection strategy of proportional reproduction with crossover probability $p_c \in [0, 1]$ and mutation probability $p_m \in (0, 1)$. Let S be the state space of $a^*(t)$ -chain with cardinality of N

$= |S| = 2^l$. Moreover, let $C = (c_{ij})$, $M = (m_{ij})$, $R = (r_{ij})$ be the intermediate transition matrices corresponding to crossover, mutation and selection respectively, then the transition matrix P of $a^*(t)$ -chain can be naturally represented as a product of C , M and R , i.e., $P = CMR$. Since crossover maps probabilistically each state of S to another one, C is stochastic. The same holds for the other operators and their transition matrices.

Assuming the length of a string encoded in binary is l bits, each string denotes a state in S . Let $H(i, j)$ represents the Hamming distance between state $i \in S$ and $j \in S$. Then the transition probability from state i to j caused by mutation is

$$m_{ij} = p_m^{H(i,j)} (1 - p_m)^{l-H(i,j)} > 0,$$

so that M is a positive stochastic matrix.

The probability that $a^*(t)$ is selected to be a member of the next generation can be calculated by

$$r_{ii} = \frac{f(a^*(t))}{\sum_{k=1}^N f(a_k(t))} > 0, \text{ for any } i \neq j, r_{ij} = 0.$$

Obviously, R is also stochastic and there does not exist any column with total entries of zero.

Let $A = CM$, $P = AR$, then,

$$A = CM = (a_{ij}) = \sum_{k=1}^N c_{ik} \cdot m_{kj} > 0,$$

$$P = AR = (p_{ij}) = \sum_{k=1}^N c_{ik} \cdot r_{kj} > 0,$$

so that P is a positive stochastic matrix. It follows by theorem 3 that simple genetic algorithms do not converge to anyone of the global optima.

4.2 $(n+n, n)$ selection

We designed a modified genetic algorithm which introduces the niche and migration implemented by crossover of individuals with similar fitness and $(n+n, n)$ selection. Before crossover, n individuals in the current population are ordered in the sequence of their fitness from highest to lowest, and they are orderly conjugated to construct $n/2$ pair of parents. The genes of each pair of parental individuals are recombined randomly to generate two descendants. Two individuals that have higher fitness are determinately chosen into the temporary population by $(2+2, 2)$ selection. Mutation operation is performed over the temporary population proba-

bilistically with $p_m \in (0, 1)$, but only the mutation over $a^*(t)$ in the temporary population applies $(1+1, 1)$ selection to choose the better one from $a^*(t)$ and the mutated $a^*(t)$ into the next population, the other mutated individuals are directly placed into the next population. For the sake of simplicity, assuming that only one $a^*(t)$ exists in the feasible region. Clearly, the intermediate stochastic matrices C , M corresponding to crossover and mutation and the transition matrix P of $a^*(t)$ -chain with cardinality $N = |S|$ are respectively

$$C = \begin{pmatrix} 1 & & & \\ c_{21} & c_{22} & & 0 \\ & \vdots & \ddots & \\ c_{N1} & c_{N2} & \cdots & c_{NN} \end{pmatrix},$$

$$M = \begin{pmatrix} 1 & & & \\ m_{21} & m_{22} & & 0 \\ & \vdots & \ddots & \\ m_{N1} & m_{N2} & \cdots & m_{NN} \end{pmatrix},$$

$$P = CM = \left(\sum_{k=1}^N c_{ik} \cdot m_{kj} \right) =$$

$$\begin{pmatrix} 1 & & & \\ p_{21} & p_{22} & & 0 \\ & \vdots & \ddots & \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{pmatrix},$$

where, when $1 < i \leq N$ and $j \leq i$, s. t., $0 < c_{ij}, m_{ij}, p_{ij} < 1$ and $\sum_{j=1}^i c_{ij} = \sum_{j=1}^i m_{ij} = \sum_{j=1}^i p_{ij} = 1$. The matrix P is the same as the structure of the reducible matrix in Theorem 4. It follows by Theorem 4 that the modified genetic algorithm with the $(n+n, n)$ selection is glob-

ally convergent.

5 Conclusions

The above analyses reveal that, in respect of the qualitative analysis of convergence the Markov chain model in this paper proved to be a practical facility. The derived criterion theorems based on $a^*(t)$ -chain are independent of encoding scheme and selection mechanism as well as the topological structure of search space. They can be conveniently applied to judge the global convergence of GAs. But when time complexity of GAs is concerned, it should be modified properly and is our future research topic.

References

- [1] Qi X and Palmieri F. Adaptive mutation in the genetic algorithms [A]. Fogel D B, Atmar W, Eds., Proc. of the Sec. Ann. Conf. on Evolutionary Programming [C], La Jolla, CA: Evolutionary Programming Society, 1993, 192 - 196
- [2] Rudolph G. Convergence analysis of canonical genetic algorithms [J]. IEEE Trans. on Neural Networks, 1994, 5(1):96 - 101
- [3] David Reynolds and Jagannathan Gomatam. Stochastic modelling of genetic algorithms[J]. Artificial Intelligence, 1996, 82:303 - 330
- [4] Iosifescu M. Finite Markov Processes and Their Applications [M]. Chichester: Wiley, 1980, 123 - 126
- [5] Holland J H. Adaptation in Natural and Artificial Systems[M]. Ann Aber: The University of Michigan Press, 1975

本文作者简介

郭现七 1963年生,副教授,博士研究生.主要研究方向:自适应控制,人工智能,遗传算法的理论与应用.

喻寿益 1940年生,教授,博士生导师.主要研究方向:复杂机电系统的耦合与解耦理论,自适应控制,智能控制,进化算法的理论与应用.