A Novel Genetic Algorithm Based on Immunity

Licheng Jiao, *Senior Member, IEEE,* and Lei Wang

***Abstract*—A novel algorithm, the immune genetic algorithm (IGA), is proposed based on the theory of immunity in biology which mainly constructs an immune operator accomplished by two steps: 1) a vaccination and 2) an immune selection. IGA proves theoretically convergent with probability 1. Strategies and methods of selecting vaccines and constructing an immune operator are also given. IGA is illustrated to be able to restrain the degenerate phenomenon effectively during the evolutionary process with examples of TSP, and can improve the searching ability and adaptability, greatly increase the converging speed.**

***Index Terms*—Convergence, genetic algorithm, immune genetic algorithm, immunity, TSP.**

I. Introduction

DURING the last three decades, there has been a growing interest in algorithms which rely on analogies to natural phenomena such as evolution, heredity, and immunity. The emergence of massively parallel computers made these algorithms of practical interest. The genetic algorithm (GA) belongs to one category of these best known algorithms, whose beginnings can be traced back to the early 1950s when several biologists used computers for simulations of biological systems [1]. However, the work done in the late 1960s and the early 1970s at the University of Michigan under the direction of J. Holland led to GA as it is known today [2]. With the characteristics of easier application, greater robustness, and better parallel processing than most classical methods of optimization, GA has been widely used for combinatorial optimization [3], [4], structural designing [5], machine learning rule-based classifier systems [6], [7], and other engineering problems [8]–[10].

It is well known that GA pertains to searching algorithms with an iteration of generation-and-test. Two operators—crossover and mutation—give each individual the chance of optimization and ensure the evolutionary tendency with the selection mechanism of survival of the fittest. GA also proves to be convergent under the condition of maintaining the best individual found over time after selection [11]. Because the two genetic operators make individuals change randomly and indirectly during the whole process, they not only give the individuals the evolutionary chance but also cause certain degeneracy. In some cases, these degenerative phenomena are very obvious. On the other hand, there are many basic and obvious characteristics or knowledge in a pending problem. However the crossover and mutation operators in GA lack the capability of meeting an actual situation, so that some torpidity appears when solving problems, which is conducive to the universality of the algorithm but neglects the assistant function of the characteristics or knowledge. The loss due to the negligence is sometimes considerable in dealing with some complex problems. It is also realized from practice that only using GA or other evolutionary algorithms is far deficient for simulating the ability of human beings to deal with problems. Therefore, the demand for deep mining and exploiting the intelligence resource of the humanity is very urgent. In view of the above, it is a perpetual theme in the research field of evolutionary algorithms and even intelligent computation to further learn, develop and utilize the human intelligence.

Based on the consideration above, this paper aims at the concept of immunity into the canonical GA. On condition of preserving GA's advantages, it utilizes some characteristics and knowledge in the pending problems for restraining the degenerative phenomena during evolution, so as to improve the algorithmic efficiency. This novel algorithm is presently called the immune genetic algorithm (IGA). This paper also presents its detailed steps, proves the global convergence and shows the strategies of selecting immune vaccines and the methods of constructing an immune operator. Examples of TSP show that IGA is not only feasible but also effective.

II. The Immune Genetic Algorithm and its Convergence

A. The Immune Genetic Algorithm

The aim of leading immune concepts and methods into GA is theoretically to utilize the locally characteristic information for seeking the ways and means of finding the optimal solution when dealing with difficult problems. To be exact, it utilizes the local information to intervene in the globally parallel process and restrain or avoid repetitive and useless work during the course, so as to overcome the blindness in action of the crossover and mutation. During the actual operation, IGA refrains the degenerative phenomena arising from the evolutionary process, thus making the fitness of population increase steadily. Because this course is very similar to that of immune phenomenon in nature [12], the algorithm based on the above idea is named the IGA for the purpose of simplicity and direct-perception.

To be more exact, the idea of immunity is mainly realized through two steps based on reasonably selecting vaccines, i.e., a vaccination and an immune selection, of which the former is used for raising fitness and the latter is for preventing the deterioration. Now they are explained as follows.

1) The Vaccination: Given an individual x, a vaccination means modifying the genes on some bits in accordance with priori knowledge so as to gain higher fitness with greater probability. This operation must satisfy the following two conditions.

Manuscript received November 5, 1999; revised July 5, 2000. This work was supported by the National Science Foundation of China under Grant 69772029 and the Project "863." This paper was recommended by Associate Editor W. Pedrycz.

The authors are with the Key Lab for Radar Signal Processing, Xidian University, Xi'an 710071, China.

Publisher Item Identifier S 1083-4427(00)07992-3.

Firstly, if the information on each gene bit of an individual y is wrong, i.e., each bit of it is different from that of the optimal one, then the probability of transforming from x to y is 0. Secondly, if the information on each gene bit of x is right, i.e., x is the optimal one, then the probability of transforming from x to x is 1. Suppose a population is $c = (x_1, x_2, \cdots, x_{n_0})$, and then the vaccination on c means the operation carried out on $n_\alpha = \alpha n$ individuals which are selected from c in proportion as α. A vaccine is abstracted from the prior knowledge of the pending problem, whose information amount and validity play an important role in the performance of the algorithm.

2) The Immune Selection: This operation is accomplished by the following two steps. The first one is the immune test, i.e., testing the antibodies. If the fitness is smaller than that of the parent, which means serious degeneration must have happened in the process of crossover or mutation, then instead of the individual the parent will participate in the next competition; the second one is the annealing selection [13], i.e., selecting an individual x_i in the present offspring $E_k = (x_1 \cdots x_{n_0})$ to join in the new parents with the probability below

$$P(x_i) = \frac{e^{f(x_i)/T_k}}{\sum_{i=1}^{n_0} e^{f(x_i)/T_k}} \tag{1}$$

where $f(x_i)$ is the fitness of the individual x_i and $\{T_K\}$ is the temperature-controlled series approaching 0.

Algorithm 1: The immune genetic algorithm

1) Create initial random population A_1.
2) Abstract vaccines according to the prior knowledge.
3) If the current population contains the optimal individual, then the course halts; or else, continues.
4) Perform crossover on the kth parent A_k and obtain the results B_k.
5) Perform mutation on B_k and obtain C_k.
6) Perform vaccination on C_k and obtain D_k.
7) Perform immune selection on D_k and obtain the next parent A_{k+1}, and then go to 3).

We can refer to the flowchart shown in Fig. 1 about the course of executing the above algorithm, and the specific contents about the immune operation will be discussed in detail in Section III.

B. The Convergence of IGA

Given that the size of a population is n_0, which is the same as the size of the initial population, the encoding of an individual results in the l-bit q-scale code. The crossover is carried out by choosing either one point or multipoints. Mutation is performed on each gene bit independently with probability P_M, after which the probability of being in any other state is $1/q-1$.

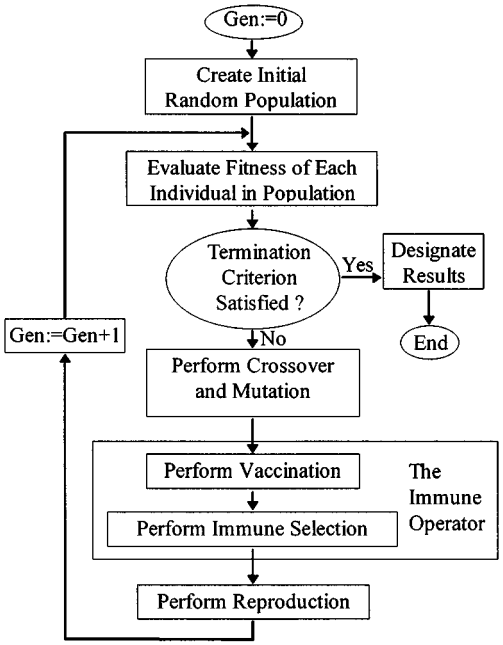

Fig. 1. Flowchart of IGA.

The transformation of the states in the algorithm is shown at the bottom of the page, where the transformation of the states from A_k to D_k constitutes a Markov chain, while the state A_{k+1} is related to each of the former states. However, the random process $\{A_k | k = 1, 2, \cdots\}$ is still a Markov process. Suppose that X is a searching space, namely the space in which all the individuals are included, and we consider the population with the size of n_0 as a point in the state space $S = X^{n_0}$, in which each coordinate is an individual in X. Suppose $|S|$ indicates the number of the states in S; $s_i \in S$ $(i = 1, 2, \cdots, |S|)$ expresses S_i is a certain state in S; $s_i \subseteq s_j$ shows an inclusive relationship when s_i and s_j are both the subsets in X; V_k^i suggests that a random variable V is just in the state s_i of the kth generation, and f is the fitness function on X. Let

$$S^* = \{x \in X | f(x) = \max_{x_i \in X} f(x_i)\} \tag{2}$$

and then the convergence of the algorithm can be defined as follows:

Definition 1: For any initial distribution, if the following equation holds:

$$\lim_{k \to \infty} \sum_{s_i \cup S^* \neq \emptyset} P\{A_k^i\} = 1 \tag{3}$$

then the algorithm is convergent.

This definition of the convergence means that if the algorithm is implemented for an enough iteration, then the probability that the population contains the optimal individual will verge on 1.

$$A_k \overset{Crossover}{\longrightarrow} B_k \overset{Mutation}{\longrightarrow} C_k \overset{Vaccination}{\longrightarrow} D_k \overset{Immune\ Selection}{\longrightarrow} A_{k+1}$$

So the definition shown above is usually called the convergence with probability 1.

Theorem 1: The immune genetic algorithm is convergent with probability 1.

The proof appears in Appendix A. It is necessary to point out that if the immune operator is cut off from this algorithm, then it can be proved that IGA does not converge to the global optimal individual [11], or it is strongly nonconvergent [13].

III. An Immune Operator

A. Mechanism of an Immune Operator

In IGA, an immune operator is composed of two operations, a vaccination and an immune selection, and it utilizes vaccines to intervene aptly in the variation of genes in individual chromosome. During the actual operation, a detailed analysis is firstly carried out on the pending problem, and at the same time, as many basic characteristics of the problem as possible ought to be found. Then, the characteristics are abstracted to be a schema. Finally, the schema is made the basis for the immune operator to generate new individuals. Here it is necessary to note that there is usually not only one characteristic in a certain problem, that is to say, there may not be only one vaccine. Therefore, during the course of vaccination, the injection can be carried out by either selecting any vaccine randomly or getting them together according to a certain logic relationship. On the other hand, a vaccine can be regarded as an estimation on some genes of the optimal individual $x_{\max}$, and the accuracy of this estimation depends on the further test in the subsequent immune selection. Thus it can be seen that the correctness of selecting a vaccine plays an important role in the operational efficiency. However, the accuracy and the quality of vaccines selected only affect the functions of the vaccination, and would not influence the convergence of the algorithm that is ensured by the immune selection in the final analysis.

We can find out the influence of correct selection of a vaccine on the functions of the vaccination with a straightforward example. Given that the encoding of a pending problem is binary with n bits, it can form a space with 2^n pending solutions. If we could define the gene of a certain bit by analyzing the problem, then the population with this gene will centralize in the half space in which the optimal individual is forecast, and therefore the searching efficiency will be improved greatly. On the contrary, if our estimation is wrong, then the vaccination will hold back the searching actions, and even exert a negative influence. Although the influence can be offset by continuing immune selection, the searching efficiency of the algorithm will be affected, however, and the algorithm now is also convergent. Then we can have a look at the actions taken by the immune selection during running the algorithm in the following.

Theorem 2: Under the immune selection, if the vaccination makes the fitness of an individual vaccinated higher than the average fitness of the current population, then the schema with which the vaccine is corresponding will be diffused at an index level in the population. Or else, it will be restrained or be attenuated at the index level.

The proof is shown in Appendix B. It can be seen from this theorem that the immune selection can improve the advantages of vaccination and shows the allowance and robustness in eliminating its negative influence.

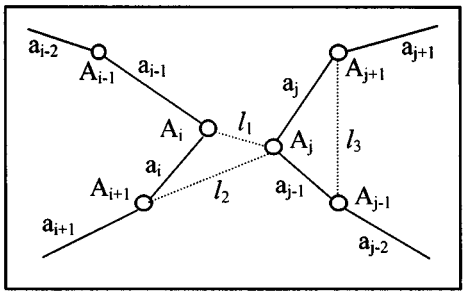

Fig. 2. Sketch map of the mechanism of a vaccine.

B. Example of Selecting a Vaccine

Now in what follows we discuss a general strategy for selecting vaccines with an example of TSP.

1) Analyzing the pending problem and collecting the characteristic information: In view of the general knowledge of humans, when a person is going to travel from one city to another, the first target to be selected is the city to which the distance from the present one is the shortest among all the candidate cities. If the current target is just the city visited last time, then the target changes to the city to which the distance from the present one is the shortest city among all the candidate cities but the visited one, and so on. This idea cannot act as the answer to the global problem, and however, it may act as a strategy for a small problem such as one only relating to three or four cities. With regard to the global problem, the latter belongs to a localized one.

   Based on the consideration above and as to the characteristics of TSP, the final answer, i.e., the optimal entire path, must includes (actually to a great extent) the sub-paths which are the shortest between every two cities. This characteristic is not only one of the properties of TSP, but also can be used as the information or knowledge for dealing with the problem. So it may act as an approach to abstracting vaccines.
2) Producing vaccines according to the characteristic information: Suppose A_j is the nearest city to the city A_i among all its candidate cities, and the two cities are not joined together but in two sections of a certain path such as A_{i-1}—A_i—A_{i+1} and A_{j-1}—A_j—A_{j+1}, which is shown as the solid lines in Fig. 2. Then the current path is

$$\pi=\{A_0,\cdots,A_{i-1},A_i,A_{i+1},\cdots,A_{j-1},A_j,A_{j+1},\cdots,A_N\},$$

whose corresponding length is

$$D_\pi=\sum_{k=1}^{i-1}a_k+a_i+\sum_{k=i+1}^{j-2}a_k+a_{j-1}+a_j+\sum_{k=j+1}^{N}a_k. \quad (4)$$

If the condition of immune probability P_i is satisfied, as to the city A_i, the immune operator regards its neighbor A_j as the

next target city, then the original path will be changed as follows and shown as the dotted lines in Fig. 2:

$$\pi_c = \{A_0, \cdots, A_{i-1}, A_i, A_j, A_{i+1}, \cdots, A_{j-1}, A_{j+1}, \cdots, A_N\}$$

and the corresponding length is

$$D_{\pi_c} = \sum_{k=1}^{i-1} a_k + l_1 + l_2 + \sum_{k=i+1}^{j-2} a_k + l_3 + \sum_{k=j+1}^{N} a_k. \quad (5)$$

Because A_j is the city among all the candidate cities to which the distance from the city A_i is the shortest, and in the triangle constituted by A_i—A_j—A_{i+1}, l_1 must be either the shortest or the sub-shortest side (in the latter case l_2 must be the shortest one, and if $a_i < l_1$, then the nearest city from which A_i is away is A_{i+1} instead of A_j). The relationship among A_i, A_j and A_{i+1} is not certain to be suitable for A_{j-1}, A_j and A_{j+1}. Therefore in most cases, the quantity of reduction from $a_{j-1} + a_j$ to l_3 is greater than that of increase from a_i to $l_1 + l_2$. Then the following relation holds:

$$P(D_{\pi_c} < D_\pi) \gg P(D_{\pi_c} > D_\pi) \quad (6)$$

where $P(A)$ is the probability that a fact A occurs.

The above adjustment is just the process by which the immune operator takes effect on TSP based on a certain vaccine. The statistics of all the connections between every city and its neighbor constructs a series of vaccines: $H = \{h_j | j = 1, 2, \cdots, m\}$ (m is the number of candidate cities).

By the way, in the knapsack problem, we can regard the individuals with a high value per weight ratio as a kind of vaccines.

C. An IGA with Self-Adaptive Selecting Vaccines

In most cases of dealing with some problems, it is difficult to abstract the characteristic information of them because we know little about the priori knowledge. On the other hand, the work of searching the local scheme used for the global solution makes the workload increase greatly and the efficiency decrease, so that the value of this work is lost. At this time, we can abstract information from genes of the present optimal individual to make vaccines during the evolutionary process. The whole algorithm with self-adaptive abstracting vaccines is expressed as follows.

For the purpose of convenient expression, we first give some specific symbols. Suppose $a^i_{H,k}$ is a middle individual produced by vaccinating on the ith individual of the kth generation. P_V is the probability of vaccination. *Vaccine* (a^i_k, h_j) suggests the operation of altering some genes of an individual a^i_k according to the schema h_j. n and m are the size of the population and the number of genes, respectively. *Random*(m) means generating randomly an arbitrary positive integer between 1 and m. Then the process of selecting self-adaptive vaccines and immune evolution can be shown as follows:

Algorithm 2: IGA with self-adaptive abstracting vaccines

Begin:
$k = 0$;
While (Conditions = True)
  $a^{optimal}_k = Statistics(a^i_k | i = 1, \cdots, n)$;
  $H = \{h_j = a^{optimal}_{k,j} | j = 1, 2, \cdots, m\}$;
  Crossover: $a^i_k = Crossover(a^i_k), i = 1, \cdots, n$;
  Mutation: $a^i_k = Mutation(a^i_k), i = 1, \cdots, n$;
  For $i = 1$ to n
    If $\{P_V\}$ = True
    $J = Random(m)$;
    Vaccination: $a^i_{H,k} = Vaccine(a^i_k, h_J)$;
    Immune test: If $a^i_{H,k} < a^i_{k-1}$, then $a^i_k = a^i_{k-1}$;
      Else $a^i_k = a^i_{H,k}$;
  Annealing selection: $A_{k+1} = S(A_k)$;
  $k = k + 1$;
End

The halting condition of the algorithm above can be either the maximal iterative number or the maximal number of the statistics of the quasioptimal fitness that remains steady. It is necessary to point out that the probability of immunization should not be too large to affect the genetic diversity, or else, once vaccines are not selected correctly, the immunization will delay the evolution process.

IV. Simulations

A. The Solution to TSP with IGA

TSP was documented as early as 1759 by Euler whose interest was in solving the knights' tour problem. The term *traveling salesman* was first used in a German book written by a veteran traveling salesman in 1932 [14]. TSP was later proved to be NP-hard and it is very difficult to gain a solution with normal algorithms. In recent years, there have been many results about TSP based on evolutionary algorithms [15], [16]. To a certain extent, it is considered as an indirect criterion among the intelligent algorithms. TSP is described as searching a permutation of integers $\pi = \{p_1, p_2, \cdots p_n\}$ which satisfies the following equation:

$$D_\pi = \sum_{i=1}^{n-1} d(p_i, p_{i+1}) + d(p_n, p_1) = \lim \left(\sum_{\substack{i,j=1 \\ i \neq j}}^{n} d(p_i, p_j) \right) \quad (7)$$

where $d(A, B)$ means the distance between the cities with the names of A and B, respectively, and p_i is the name of the ith city in the entire path.

1) The coding and the fitness function: In order to make the coding easy and clear, we take the permutation of the order of visiting the cities for the coding of TSP and regard the following equation as the fitness function

$$f(\pi_i) = \frac{(76.5 \times L \times \sqrt{N})}{D_{\pi_i}} \quad (8)$$

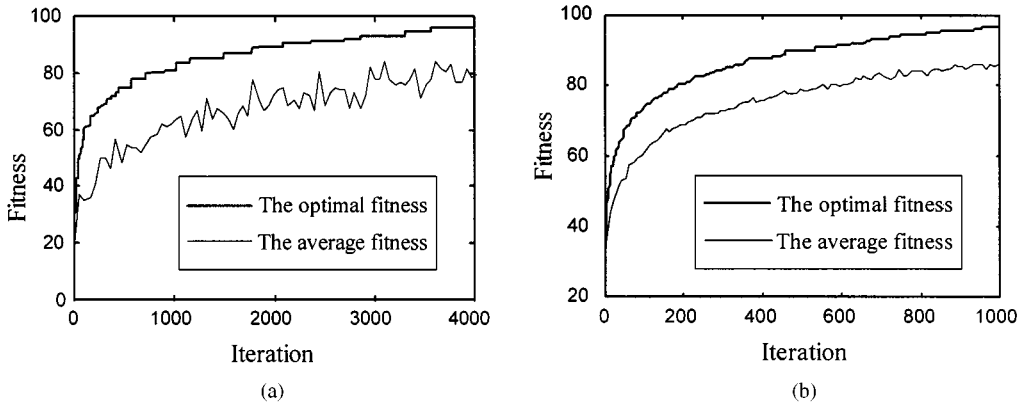

Fig. 3. Comparison between GA and IGA about the variation of the fitness with iteration (a) with canonical GA and (b) with IGA.

where L means the side of the smallest square which can contain all the cities, N is the number of cities and D_{π_i} is the length of the path in the current permutation.

2) The crossover and the mutation operators: In simulations, the two-point crossover is used in principle, whose positions are selected randomly (so one-point crossover may occur during the actual implementation). A novel method of only altering the partial path is adopted in mutation, based on evaluating the inheritance to the characteristic of genotypes in the genetic individuals and the diversification of characteristics necessary for further evolution. One part of the entire path is selected every time, whose beginning and end are defined in accordance with the results of evaluation. The mode of exchanging by n times is adopted in actual operation, with n calculated as follows:

$$n = [N/M + \exp(-\alpha K)] \tag{9}$$

where N is the number of the cities, M means the number of sub-paths, K suggests the number of generations and α is a constant denoting the variation of n with K.

3) The immune operator: We first adopt the idea shown in Section III-B to select vaccines. Then, after every genetic operation, we select some individuals for vaccination in accordance with the immune probability, after which the immune test is to be continued to test the individual vaccinated. If the fitness rises, then go to the next operation; or else make the parent take part in the competition of selection instead of the offspring. During the selection, the probability of the individual in offspring to be selected is calculated according to (1).

In the actual test, the 75-city TSP is firstly adopted, where the annealing temperature T in the immune selection is calculated as follows:

$$T_k = \ln\left(\frac{T_0}{k} + 1\right), \qquad T_0 = 100 \tag{10}$$

where k is the evolutionary generations. With the basic parameters fixed, TSP is solved with GA which maintains the best individual found over time after selection [11] and IGA, respectively, in which the crossover and mutation used in canonical GA are the same as those in IGA and the selection is the proportional selection based on fitness. A comparison between the canonical GA and IGA about the variation of the fitness with generations is shown in Fig. 3. IGA finds the final best solution which is the same as that by Fogel [16] with the length 549.180 and the fitness 96.5 after 960 generations in 215.892 s(the average for 20 times and with Pentium 266) while GA finds it after 3550 generations in 541.182 s. All the vaccines used in IGA are shown in Fig. 4. Here, it is necessary to point out that the vaccination and immunization operators create an overhead in computing time, so the duration for a generation in IGA is a litter longer than that in GA. However, the whole efficiency of IGA is still superior to that of GA.

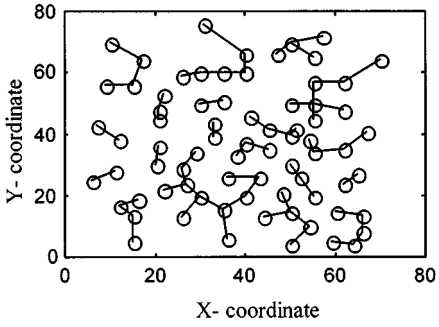

Fig. 4. Vaccines used for the 75-city TSP.

Then *Grötschel's* 442-city TSP is also solved with canonical GA and IGA, respectively, by the method similar to that described above. This time, we use evolutionary programming for abstracting vaccines based on the idea as in Section III-B. With iteration by 70 times, we can gain the vaccines as in Fig. 6(a). A comparison between GA and IGA about the variation of the optimal and the average fitness among individuals in offspring with generations is shown in Fig. 5 (the length of the optimal path is 5154.1 and the corresponding fitness is 124.8187, and it can be seen from this figure that GA finds the optimal solution at about the 28 000th generation while IS at the 3900th), in which the final optimal path is as shown in Fig. 6. From Fig. 5, it can be found that IGA is conducive to raising the searching efficiency and restraining degradation in the later stage of the evolutionary process with GA, thus increasing the convergent speed to some extent.

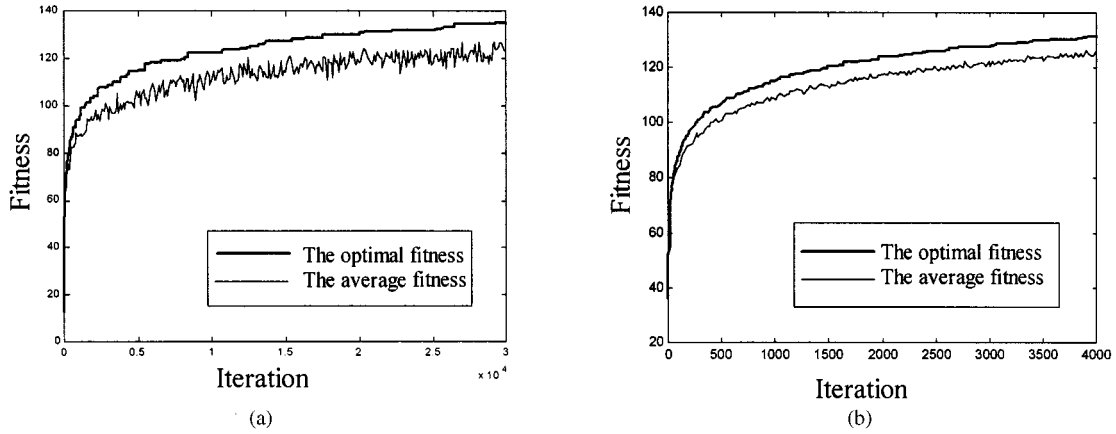

Fig. 5. Comparison between canonical GA and IGA about the variation of the optimal and the average fitness among individuals in offspring with generations (a) GA calculating curve and (b) IGA calculating curve.

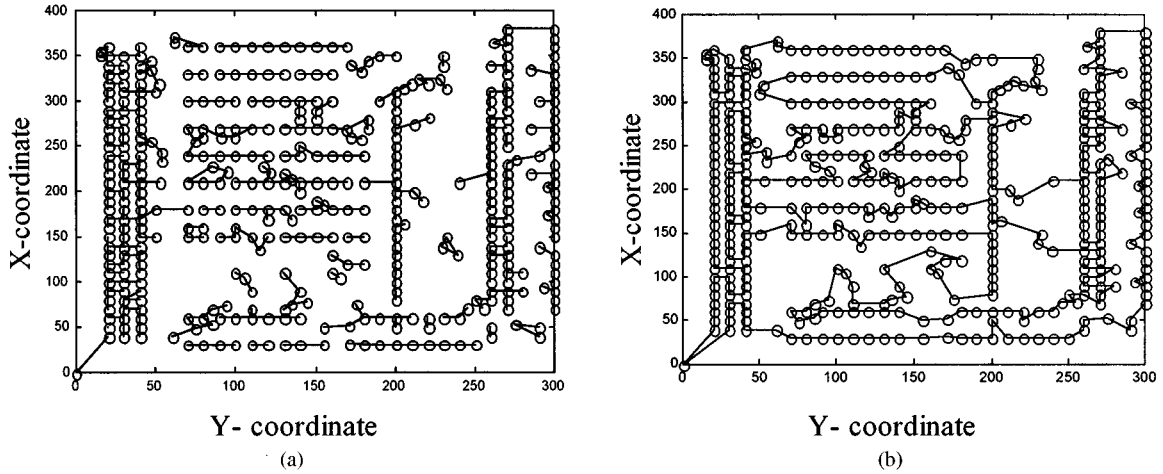

Fig. 6. Solution to *Grötschel's* 442-city TSP (a) vaccines and (b) the optimal path.

B. Application in Function Optimization

In what follows, we may also examine the searching ability of IGA with an example of function optimization:

$$f(x) = 10 + \frac{\sin\left(\frac{1}{x}\right)}{(x-0.16)^2 + 0.1} \tag{11}$$

where the variation of the function with the independent variable x in (0,1) is shown in Fig. 7. Our purpose is to search x_{max} in (0,1) which satisfies the following inequality:

$$f(x_{\max}) \geq f(x), \quad \forall\, x \in (0,1). \tag{12}$$

For the convenience of comparison between IGA and GA, we adopt a unified coding for the problem as follows:

0.	x_k^1	x_k^2	x_k^3	x_k^4

where x_k^j $(j = 1, 2, 3, 4)$ is an integer in [0,9] which expresses the jth gene of the kth individual.

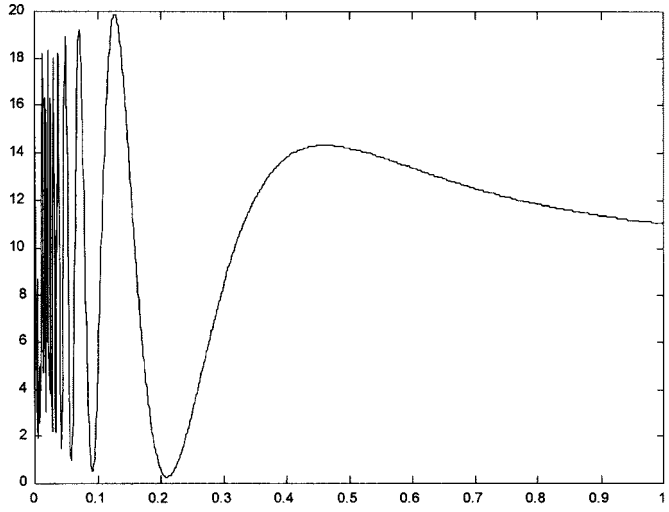

Fig. 7. Function curve on (0,1].

In the actual test, we let the size of population be five for convenient observation, the fitness function be just the original function and the halt condition be the maximum iteration (i.e.,

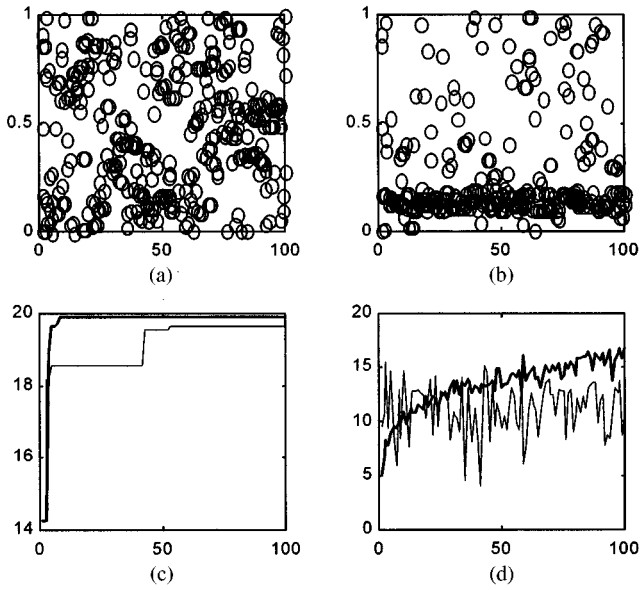

Fig. 8. Process of accepting a self-adaptive vaccine. Curves of the evolutionary process based on GA and IGA (a) distribution of individuals during the evolutionary process of GA, (b) distribution of individuals during the evolutionary process of IGA, (c) comparison in the optimal fitness between GA and IGA, and (d) comparison in the average fitness between GA and IGA. —— The curve of IGA; —— The curve of GA.

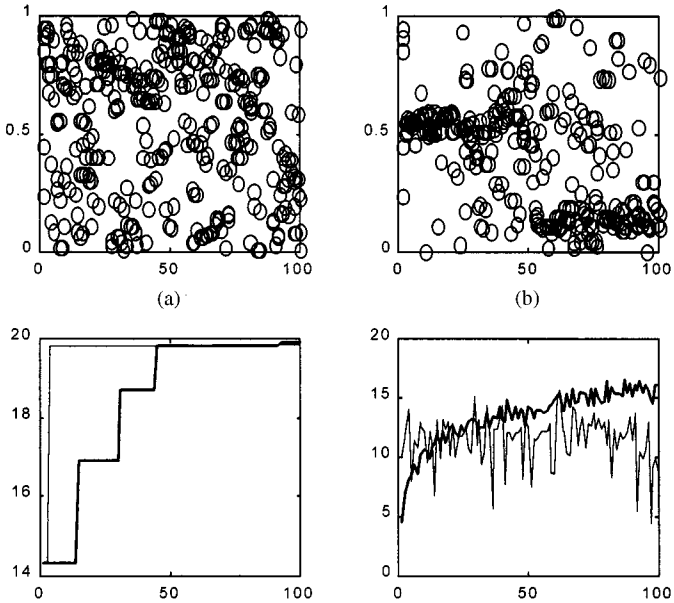

Fig. 9. Process of accepting an error vaccine $x_k^1 = 5$. Curves of the evolutionary process based on GA and IGA (a) distribution of individuals during the evolutionary process of GA, (b) distribution of individuals during the evolutionary process of IGA, (c) comparison in the optimal fitness between GA and IGA, and (d) comparison in the average fitness between GA and IGA. —— The curve of IGA; —— The curve of GA.

100). In the first group, we generate randomly 20 initial populations, produce vaccines with the algorithm of self-adaptive selecting vaccines as shown in Section III-C, and observe in the following four aspects:

1) The distribution of the individuals during evolution with no vaccination (i.e., the process of GA).
2) The distribution of the individuals during evolution with vaccination (i.e., the process of IGA and the probability of vaccination is 0.5).
3) The variation of the optimal fitness with the reproduction of offspring.
4) The variation of the average fitness with the reproduction of offspring.

The results by 20 times vary in detail in the above ways, but the whole tendency is identical. One of the results is shown in Fig. 8, where IGA finds the globe optimum $(f(x)_{\max} = 19.8949; x_{optimal} = 0.1275)$ after 12 iterations, while GA only finds the local optimum $(f(x) = 19.8903; x = 0.1273)$ after 53 iterations.

In the second group, we generate randomly another 20 initial populations. This time, we adopt an error vaccine, i.e., $x_k^1 = 5$, and also make observations in the above four ways. One of the results is shown in Fig. 9, where IGA finds the globe optimum after 92 iterations, while GA only finds the local optimum $(f(x) = 19.8797; x = 0.1284)$ after 67 iterations. From Fig. 9(b), we can see that if the vaccine is selected incorrectly, then it may be restrained step by step and the population will approach the schema with a high fitness.

In a word, the correct selection of a vaccine plays a most important role in executing the algorithm, but how it is selected would not affect the convergence. On the other hand, the immune selection can restrain the schema with low fitness and strengthen the high one, so it is the basis for ensuring the algorithm convergence.

V. Conclusion

A novel global parallel algorithm, the IGA, combining the immune mechanism and the evolutionary mechanism is proposed, whose global convergence is proved, and the strategies of constructing an immune operator are also given in this paper. The theoretical analysis and simulations of the TSP show that IGA is not only feasible but also effective and is conducive to alleviating the degeneration phenomenon in the original GA, thus greatly increasing the converging speed.

Appendix A
Proof of Theorem 1

Proof of Theorem 1: Given $s_i = (x_1, \cdots, x_{n_0}) \in S$ and $f(s_i) = (f(x_1), \cdots, f(x_{n_0}))$, if the first nonzero part of $f(s_i) = f(s_j)$ or $f(s_i) - f(s_j)$ is positive, then we consider $s_i \geq s_j$. In addition, we suppose $I = \{i | s_i \geq s_j, \forall s_j \in S\}$. From the definition of "$\geq$," if $i \in I$, then $s_i = (x_1, \cdots, x_{n_0})$ satisfies

$$f(x_1) = f(x_2) = \cdots = f(x_{n_0}) = \max_{x \in X} f(x) \tag{A1}$$

so, $s_i \cap S^* \neq \emptyset$ (in fact $s_i \subseteq S^*$). Here we define $f^* = \max_{x \in X} f(x)$ and $\rho_1 = \min\{|f(x_i) - f(x_j)| | x_i, x_j \in X$ and $f(x_i) \neq f(x_j)\}$.

Given that $g_{ij}(k)$ is the probability of transformation for one step in Markov process $\{A_k|k=1,2,\cdots\}$, then

$$g_{ij}(k)=P\{A_{k+1}^j/A_k^i\}=\sum_{d=1}^{|S|} P\{D_k^d/A_k^i\}\cdot P\{A_{k+1}^j/A_k^iD_k^d\}. \tag{A2}$$

Suppose that D_k has transformed into E_k after the immune test, and then

$$\begin{aligned} &P\{A_{k+1}^j/A_k^iD_k^d\} \\ &\quad=\sum_{e=1}^{|S|} P\{A_{k+1}^jE_k^e/A_k^iD_k^d\} \\ &\quad=\sum_{e=1}^{|S|} P\{E_k^e/A_k^iD_k^d\}P\{A_{k+1}^j/A_k^iD_k^dE_k^e\}. \end{aligned} \tag{A3}$$

i) If $i\in I$ and $j\notin I$, then $\forall\ x\in s_i$, and from (A1), $f(x)=f^*$. Since $j\notin I, \exists x_0\in s_j$, such that $f(x_0)<f^*$, then (see (A4) at the bottom of the page) It is clear that $\varepsilon_k\to 0\ (k\to\infty)$. By noticing $\Sigma_{e=1}^{|S|} P\{E_k^e/A_k^iD_k^d\}=1$ and $\Sigma_{d=1}^{|S|} P\{D_k^d/A_k^i\}=1$, from (A2) and (A3), the following holds:

$$\begin{aligned} g_{ij}(k)&\le\sum_{d=1}^{|S|} P\{D_k^d/A_k^i\} \\ &\quad\cdot\left(\varepsilon_k\sum_{e=1}^{|S|} P\{E_k^e/A_k^iD_k^d\}\right)=\varepsilon_k. \end{aligned} \tag{A5}$$

ii) If $i\notin I$ and $j\in I$, given that $s_d=(x_1,x_2,\cdots,x_{n_0})$ and the number of digits in the set $\{m|f(x_m)=f^*\}$ is n_d, and letting $S^1=\{s_d|s_j\subseteq s_d$, and $n_d\ge n_\alpha\}$, then from (A2)

$$g_{ij}(k)\ge\sum_{s_d\in S^1} P\{D_k^d/A_k^i\}P\{A_{k+1}^j/A_k^iD_k^d\}. \tag{A6}$$

By the C-K equation

$$\begin{aligned} P\{D_k^d/A_k^i\}&=\sum_{b=1}^{|S|}\sum_{c=1}^{|S|} P\{B_k^b/A_k^i\}P\{C_k^c/B_k^b\}P\{D_k^d/C_k^c\} \\ &\ge\sum_{b=1}^{|S|} P\{B_k^b/A_k^i\}P\{C_k^d/B_k^b\}P\{D_k^d/C_k^d\} \end{aligned} \tag{A7}$$

where $P\{D_k^d/C_k^d\}$ is the probability of transforming from s_d into s_d through vaccination and ought to be not less than $\binom{n_d}{n_\alpha}/\binom{n_0}{n_\alpha}$. $P\{C_k^d/B_k^b\}$ is the probability from s_b into s_d through mutation and should be equal to $\Pi_{m=1}^{n_0}\ [(P_m/q-1)^{h_i}(1-P_m)^{l-h_i}]$, where h_i is the number of mutating genes in the individual $x_{\in}s_b$. Suppose $\rho_2=[\min(P_M,1-P_M)/(q-1)]^{n_0l}$, and then $P\{C_k^d/B_k^b\}\ge\rho_2.P\{B_k^b/A_k^i\}$ is the probability from s_i into s_b through crossover and should satisfy $\Sigma_{b=1}^{|S|}\ P\{B_k^b/A_k^i\}=1$. Then from (A7)

$$P\{D_k^d/A_k^i\}\ge\rho_2\binom{n_d}{n_\alpha}\binom{n_0}{n_\alpha}. \tag{A8}$$

Since

$$\begin{aligned} P\{A_{k+1}^j/A_k^iD_k^d\}&=\sum_{e=1}^{|S|} P\{E_k^e/A_k^iD_k^d\}P\{A_{k+1}^j/A_k^iD_k^dE_k^e\} \\ &\ge P\{E_k^d/A_k^iD_k^d\}P\{A_{k+1}^j/A_k^iD_k^dE_k^d\}, \end{aligned} \tag{A9}$$

and by (A1)

$$\begin{aligned} &P\{A_{k+1}^j/A_k^iD_k^dE_k^d\} \\ &\quad=\prod_{m=1,x_m\in s_j}^{n_0}\left[\exp(f(x_m)/T_k)\Big/\sum_{m=1,x_m\in s_d}^{n_0}\right. \\ &\qquad\left.\cdot\ \exp(f(x_m)/T_k\right] \\ &\quad\ge\prod_{m=1,x_m\in s_j}^{n_0}\left[\exp(f^*/T_k)\Big/\sum_{m=1,x_m\in s_d}^{n_0}\right. \\ &\qquad\left.\cdot\ \exp(f^*/T_k)\right]=\frac{1}{n_0^{n_0}}. \end{aligned} \tag{A10}$$

By noticing $P\{E_k^d/A_k^iD_k^d\}\ge\binom{n_d}{n_\alpha}/\binom{n_0}{n_\alpha}$, and from (A9), we can obtain

$$P\{A_{k+1}^j/A_k^iD_k^d\}\ge\frac{1}{n_0^{n_0}}\binom{n_d}{n_\alpha}\Big/\binom{n_0}{n_\alpha}. \tag{A11}$$

By substituting (A8) and (A11) into (A6)

$$\begin{aligned} g_{ij}(k)&\ge\sum_{s_d\in S^1}\left(\frac{\rho_2}{n_0^{n_0}}\right)\left[\binom{n_d}{n_\alpha}\Big/\binom{n_0}{n_\alpha}\right]^2 \\ &\ge\rho_2\Big/\left[n_0^{n_0}\binom{n_0}{n_\alpha}^2\right]\ \stackrel{d}{=}\rho_0. \end{aligned} \tag{A12}$$

$$\begin{aligned} P\{A_{k+1}^j/A_k^iD_k^dE_k^e\}&=\begin{cases}\displaystyle\sum_{e=1,x_m\in s_j}^{n_0}\left[\exp\left(f(x_m)/T_k\right)\sum_{e=1,x_m\in s_e}^{n_0}\exp(f(x_m)/T_k)\right] & s_j\subseteq s_e \\ 0 & \text{the other}\end{cases} \\ &\le\frac{\exp(f(x_0)/T_k)}{n_\alpha\exp(f^*/T_k)}=\frac{1}{n_\alpha}\exp(-(f^*-f(x_0))/T_k)\le\frac{1}{n_\alpha}\exp(-\rho_1/T_k)\stackrel{d}{=}\varepsilon_k. \end{aligned} \tag{A4}$$

iii) By expressing $P\{A_k^i\}$ as $P_k(i), P_k \stackrel{d}{=} \Sigma_{i\notin I} P_k(i)$, from the property of Markov chain, we can get

$$\begin{aligned} P_{k+1} &= \sum_{i=1}^{|S|} \sum_{j\notin I} P_k(i) g_{ij}(k) \\ &= \sum_{i\in I} \sum_{j\notin I} P_k(i) g_{ij}(k) + \sum_{i\notin I} \sum_{j\notin I} P_k(i) g_{ij}(k). \end{aligned} \tag{A13}$$

Since

$$\begin{aligned} &\sum_{i\notin I} \sum_{j\notin I} P_k(i) g_{ij}(k) + \sum_{i\notin I} \sum_{j\in I} P_k(i) g_{ij}(k) \\ &\quad = \sum_{i\notin I} P_k(i) = P_k \end{aligned} \tag{A14}$$

$$\sum_{i\notin I} \sum_{j\notin I} P_k(i) g_{ij}(k) = P_k - \sum_{i\notin I} \sum_{j\in I} P_k(i) g_{ij}(k) \tag{A15}$$

by drawing (A15) into (A13), and from (A5) and (A12), $P_{k+1} \le P_k - \rho_0 P_k + |S|\varepsilon_k$. Let $1 - \rho_0 = \beta$ and $|S|\varepsilon_k = \delta_k$. Then $0 < \beta < 1$, $\delta_k \to 0$ $(k \to \infty)$, and $P_{k+1} \le \beta P_k + \delta_k$. Furthermore, $\forall\, k_0$,

$$\begin{aligned} P_{k_0+k+1} &\le \beta P_{k_0+k} + \delta_{k_0+k} \\ &\le \beta^2 P_{k_0+k-1} + \beta\delta_{k_0+k-1} + \delta_{k_0+k} \\ &\le \beta^3 P_{k_0+k-2} + \beta^2\delta_{k_0+k-2} + \beta\delta_{k_0+k-1} + \delta_{k_0+k} \\ &\le \cdots \le \beta^k P_{k_0+1} + \beta^{k-1}\delta_{k_0+1} + \cdots \\ &\quad + \beta\delta_{k_0+k-1} + \delta_{k_0+k}. \end{aligned} \tag{A16}$$

Since $\delta_k \to 0, \forall\, \varepsilon > 0, \exists N_1$, such that if $k_0 > N_1$ and $\delta_{k_0} < \varepsilon$, from (A16)

$$P_{N_1+k+1} \le \beta^k P_{N_1+1} + \varepsilon\frac{1-\beta^k}{1-\beta} \le \beta^k P_{N_1+1} + \frac{\varepsilon}{1-\beta}. \tag{A17}$$

Since $0 < \beta < 1, \exists N_2$, such that if $k > N_2, \beta^k P_{N_1+1} < \varepsilon$. Let $N = N_1 + N_2 + 1$, and if $n > N$, then

$$P_n = P_{N_1+N_2+1} \le \varepsilon + \frac{\varepsilon}{1-\beta} = \varepsilon\left(1 + \frac{1}{1-\beta}\right), \tag{A18}$$

Therefore, $\lim_{n\to\infty} P_n = 0$, and then

$$\begin{aligned} 1 &\ge \lim_{k\to\infty} \sum_{s_i\cap S^*\neq\emptyset} P_k(i) \\ &\ge \lim_{k\to\infty} \sum_{i\in I} P_k(i) = 1 - \lim_{k\to\infty} P_k = 1. \end{aligned} \tag{A19}$$

So $\lim_{k\to\infty} \Sigma_{s_i\cap S^*\neq\emptyset} P_k(i) = 1$.

APPENDIX B
PROOF OF THEOREM 2

Proof: Suppose $A(k) = \{a_k^i | i = 1, 2, \cdots, n\}$ is the kth population, in which a_k^i is the ith individual and n is the size of it. The selecting probability $P(a_k^i)$ of the individual a_k^i is calculated with the (1), i.e.,

$$\begin{aligned} P(a_k^i) &= \frac{e^{f(a_k^i)/T_k}}{\sum_{i=1}^{n} e^{f(a_k^i)/T_k}} = \frac{e^{\overline{f}_k + \Delta f_k^i/T_k}}{\sum_{i=1}^{n} e^{f(a_k^i)/T)k}} \\ &= \frac{e^{\overline{f}_k/T_k}}{\sum_{i=1}^{n} e^{f(a_k^i)/T_k}} \cdot e^{\Delta f_k^i/T_k} \end{aligned} \tag{A20}$$

where $\overline{f}_k$ suggests the average fitness of the kth population $A(k)$, and Δf_k^i is the difference between the fitness $f(a_k^i)$ of the individual a_k^i and $\overline{f}_k$. Because the probability of selecting an individual is calculated only according to its fitness, the average fitness is usually regarded to be corresponding with the average of the probabilities. Suppose the average probability of the current population is $\overline{p}_k$, and then the equation above can be simplified as follows:

$$P(a_k^i) \cong \overline{p}_k \cdot e^{\Delta f_k^i/T_k}. \tag{A21}$$

Given that the schema with which a vaccine is corresponding is H, the number of the individuals vaccinated $\Sigma a_{H,k}^i$ is $m(H,k)$. For the convenience of discussion, suppose the difference, Δf_k, between the fitness of an individual vaccinated and the average remains constant starting from the kth generation, then the number of individuals vaccinated of the next generation is

$$m(H, k+1) = \sum_{i=1}^{m(H,k)} P(a_{H,k}^i)/\overline{p}_k = m(H,k) \cdot e^{\Delta f_k/T_k}. \tag{A22}$$

The rest may be reduced by analogy, and the number of individuals after l generations will be

$$\begin{aligned} m(H, k+l) &= m(H, k+l-1) \cdot e^{\Delta f_k/T_k} \\ &= \cdots = m(H,k)[e^{\Delta f_k/T_k}]^l. \end{aligned} \tag{A23}$$

If the vaccination makes the fitness of an individual vaccinated higher than the average fitness of the current population, then $\Delta f_k > 0$, and $e^{\Delta f_k/T_k} > 1$. From the equation above, it can be seen that the number of the individuals vaccinated in the population will increase at an index level, which also means that the schema with which the vaccine is corresponding will be diffused at the same level. On the other hand, there would be two possibilities. If the fitness of the individual vaccinated is not increased, and even is not higher than that of its parent, then the individual vaccinated will be substituted by its parent through the immune test in the immune selection, and therefore the schema with which the vaccine is corresponding will be restrained. If the fitness of the individual vaccinated is higher than that of its parent, but lower than the average fitness of the current population, which means the positive role of the vaccine is not evident, then $\Delta f_k < 0$, and $e^{\Delta f_k/T_k} < 1$, so that the number of the schema will be attenuated at an index level. □

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful suggestions, and also thank Prof. D. Qin, who helped in editing the original manuscripts, and who has given us many useful suggestions on writing.

REFERENCES

[1] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.
[2] J. H. Holland and J. S. Reitman, "Cognitive systems based on adaptive algorithms," in *Pattern-Directed Inference Systems*, D. A. Waterman and F. Hayes-roth, Eds. New York: Academic, 1978.
[3] H. Muhlenbein, "Parallel genetic algorithms in combinatorial optimization," in *Computer Science and Operation Research—New Developments*, O. Balci, R. Sharda, and S. A. Zenios, Eds. New York: Pergamon, 1992, pp. 441–453.
[4] J. J. Grefenstette, R. Gopal, B. Rosmaita, and D. Van-Gucht, "Genetic algorithms for the traveling salesman problem," in *Proc. 1st Int. Conf.Genetic Algorithms and Their Applications*, J. J. Grefenstette, Ed., Piscataway, NJ, 1985, pp. 160–168.
[5] G. Miller, P. Todd, and S. Hedge, "Designing neural networks using genetic algorithm," in *Proc. 3rd Int. Conf. Genetic Algorithms and Their Applications*, D. J. Schaffer, Ed.. San Mateo, CA, 1989, pp. 360–369.
[6] R. L. Riolo, "Modeling simple human category learning with a classifier system," in *Proc. 4th Int. Conf. Genetic Algorithms and Their Applications*, R. K. Belew and L. B. Booker, Eds. San Mateo, CA, 1991, pp. 324–333.
[7] S. Tokinaga and A. B. Whinston, "Applying adaptive credit assignment algorithms for the learning classifier system based upon the genetic algorithm," *IEICE Trans. Fund. Elec. Comm. and Sci.*, vol. E75A, no. 5, pp. 568–577, 1992.
[8] J. R. Koza, *Genetic Programming*. Cambridge, MA: MIT Press, 1992.
[9] K. Kristinsson and G. A. Dumont, "System identification and control using genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, no. 5, pp. 1033–1046, 1992.
[10] K. Krishnakumar and D. E. Goldberg, "Control system optimization using genetic algorithms," *J. Guid. Contr. Dynam.*, vol. 15, no. 3, pp. 735–740, 1992.
[11] G. Rudolph, "Convergence analysis of canonical genetic algorithms," *IEEE Trans. Neural Networks*, vol. 5, no. 1, pp. 96–101, 1994.
[12] R. Chen, *The Basees of Immunology*. Peking: Health Press, 1980.
[13] J. S. Zhang, Z. B. Xu, and Y. Liang, "The whole annealing genetic algorithms and their sufficient and necessary conditions of convergence," *Science in China*, vol. 27, no. 2, pp. 154–164, 1997.
[14] E. L. Lawer, J. K. Lenstra, A. H. R. Kan, and D. B. Shmoys, *The Traveling Salesman Problem*. New York: Wiley, 1985.
[15] P. Jog, J. Y. Suh, and D. V. Gucht, "The effects of population size, heuristic crossover and local improveement on a genetic algorithm for the traveling salesman," in *Proc. 3rd Int. Conf. Genetic Algorithms*, J. D. Shaffer, Ed.. San Mateo, CA, 1989.
[16] D. B. Fogel, "Applying evolutionary programming to selected traveling salesman problems," *Cybern. Syst.*, vol. 24, pp. 27–36, 1993.
[17] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, pp. 721–742, 1984.
[18] C. Y. Percy and Y. H. Pao, "Combinatorial optimization with use of guided evolutionary simulated annealing," *IEEE Trans. Neural Networks*, vol. 6, no. 2, pp. 290–295, 1995.

Licheng Jiao (M'87–SM'91) was born in Shaanxi, China, on October 15, 1959. He received the B.S. degree from Shanghai Jiaotong University, China, in 1982, and the M.S. and Ph.D. degrees from Xi'an Jiaotong University, Xi'an, China, in 1984 and 1990, respectively.

From 1984 to 1986, he was an Assistant Professor in Civil Aviation Institute of China, Tianjing, China. During 1990 and 1991, he was a Postdoctoral Fellow in the Key Lab for Radar Signal Processing, Xidian University. Since 1992, he has been with the Key Lab for Radar Signal Processing, where he became a Full Professor and the Director of Research Center of Neural networks. His current research interests include signal and image processing, nonlinear circuits and systems theory, learning theory and algorithms, optimization problems, wavelet theory, and data mining. He is the author of four books: *Theory of Neural Network Systems* (1990), *Theory and Application on Nonlinear Transformation Functions* (1992), *Neural Computing* (1993), and *Applications and Implementations of Neural Networks* (1996), all from Xidian University Press. He is the author or coauthor of more than 150 scientific papers.

Dr. Jiao is a member of several international scientific committees.

Lei Wang was born in Xi'an, China, on May 22, 1972. He received the B.S. and M.S. degrees in automation engineering from Xi'an University of Technology, Xi'an, in 1994 and 1997, respectively. He is currently pursuing the Ph.D. degree in the Key Laboratory for Radar Signal Processing, Xidian University. His research interests include evolutionary algorithms, neural networks, and data mining.