The schema deceptiveness and deceptive problems of genetic algorithms

LI Minqiang (李敏强) & KOU Jisong (寇纪淞)

Institute of Systems Engineering, Tianjin University, Tianjin 300072, China
Correspondence should be addressed to Li Minqiang (email: mqli@tju.edu.cn)

Received November 27, 2000

Abstract Genetic algorithms (GA) are a new type of global optimization methodology based on nature selection and heredity, and its power comes from the evolution process of the population of feasible solutions by using simple genetic operators. The past two decades saw a lot of successful industrial cases of GA application, and also revealed the urgency of practical theoretic guidance. This paper sets focus on the evolution dynamics of GA based on schema theorem and building block hypothesis (Schema Theory), which we thought would form the basis of profound theory of GA. The deceptiveness of GA in solving multi-modal optimization problems encoded on $\{0,1\}$ was probed in detail. First, a series of new concepts are defined mathematically as the schemata containment, schemata competence. Then, we defined the schema deceptiveness and GA deceptive problems based on primary schemata competence, including fully deceptive problem, consistently deceptive problem, chronically deceptive problem, and fundamentally deceptive problem. Meanwhile, some novel propositions are formed on the basis of primary schemata competence. Finally, we use the trap function, a kind of bit unitation function, and a NiH function (needle-in-a-haystack) newly designed by the authors, to display the affections of schema deceptiveness on the searching behavior of GA.

Keywords: genetic algorithms, schema competition, schema deceptiveness, GA deceptive problems.

Genetic algorithms (GA) have become a general randomized searching method for solving optimization problems[1—3], and it is characterized by the power of global searching, implicit parallelism, robustness, and scalability.

The schema theorem and building block hypothesis, proposed by Holland, are the sufficient conditions for GA finding the global optima of optimization problems[1,2], and they also constitute the basic theory (schema theory) for the evolutionary mechanism of GA based on population with binary chromosomes or strings encoded on $\{0,1\}$. It is often the case that high order schemata or building blocks cannot be recombined by low order ones in population-based searching, or more seriously the recombination of low order building blocks guides GA searching away from the global optima. Those phenomena that appeared in the searching processes mean that it will be very difficult for GA to locate the global optima, and these optimization problems are GA-hard ones resulting from schema deceptiveness, which are called generally GA deceptive problems[4,5].

As to most of the GA deceptive problems, GA can usually get the global optima with high probability. But when schema deceptiveness is too grave, the probability will become lower, and searching efficiency will fall down quickly. Goldberg, Deb, Whitley, Das, et al.[4—7] started the initial studies about schema deception with a series of new concepts and typical optimization problems specially designed. On the basis of these early research results, this paper will try to give the mathematical definitions and descriptions of schema deceptiveness, deduce some theorems on schema deception and GA, and present the numerical experiments of two typical optimization functions.

1 Schema searching in encoded space

In the binary encoded space $\{0,1\}^L$ of the solution domain with string length of L, the schemata represent hyperplanes of 1 to L dimensions. The population searching of GA starts with low order schemata, or low dimension of hyperplanes in the general subspace. In the evolution of population, low order schemata compete with each other, and form gradually the higher ones. Through the recombination of low order schemata or the exchanges of low dimensional hyperplanes by genetic operators, GA searching tries to find high order schemata with high fitness, or move to the high dimensional hyperplanes in more specified subspace, and finally reach the global optima.

Liepins and Vose[8] present a simple translation function (bitwise addition modulo 2) on binary encoded space, which can map the global optimum to any point in the space. So, the string that represents the global optimum in the binary encoded space can be translated into the chromosome with all of its effective genes being "1", and all of the theorems and conclusions about "111···111" as the global optimum are fit for all other cases[4]. In the following discussions, the string "111···111" is taken as the global optimal string.

Goldberg introduced the concept of "schema deceptiveness" based on the earlier work of Bethke[4], in order to better describe what kinds of situations are likely to create difficulty for GA as function optimizer, and he designed a special GA deceptive function, the so-called minimum deception problem. Whitley and other experts analyzed the searching behavior on encoded space of GA population by showing that GA changes the sampling rate on $\{0,1\}^L$ of schemata in $\{0,1,*\}^L$ (called schema space) in line with schema theorem[5,6], and pointed out that it is the schema deceptiveness that affects directly the searching trajectory and performance. Meanwhile, schema deceptiveness and GA deceptive problems are defined conceptually, and explained with the computation of GA in solving some typical deceptive functions.

According to the analysis of Goldberg and Whitley, low schemata, representing the general hyperplanes, are recombined to form higher schemata, in the process of which the population searching of GA is gradually concentrated on the smaller subspace that consists of the high dimensional hyperplanes. When some lower hyperplanes guide GA towards some noncompetitive higher hyperplanes or schemata of high order (called deceptive attractors), we thought that there appears the GA deceptiveness in solving optimization problems.

For instance, suppose "111" is the optimal string on $\{0,1\}^3$. If the relationship holds for the low order schemata fitness:

$$
\begin{aligned}
&f(0**) > f(1**),\ f(00*) > f(11*),\ f(01*),\ f(10*),\\
&f(*0*) > f(*1*),\ f(0*0) > f(1*1),\ f(0*1),\ f(1*0),\\
&f(**0) > f(**1),\ f(*00) > f(*11),\ f(*01),\ f(*10),
\end{aligned}
$$

which means that the competitive low order schemata containing effective gene "0" will get exponential sampling in population, then the GA searching guided by 1,2 order schemata diverges away from the global optima, which forms the 3 order schema deception.

2 Schema containment, competition and relevance

Whitley first used the concepts of schema containment, competition and relevance to show

how schema deceptions take shape[5,6]. Here, mathematics is used to give the exact definitions.

Assume that the binary encoded space and schema space are $\{0,1\}^L$, $\{0,1,*\}^L$, respectively, and schema $H=(a_1,a_2,\cdots,a_L)\in\{0,1,*\}^L$ has determinate bit locus set $L(H)=\{l_1,l_2,\cdots,l_{O(H)}\}$ (the alleles on these loci are 0 or 1), where L is the length of binary string, symbol $*$ represents the indeterminate bits, $O(H)$ is the order of schema H.

Definition 1. For schemata H_1, $H_2\in\{0,1,*\}^L$, their determinate bit locus sets are $L(H_1)=\{l_{11},l_{12},\cdots,l_{1O(H_1)}\}$, $L(H_2)=\{l_{21},l_{22},\cdots,l_{2O(H_2)}\}$. If $O(H_1)\leqslant O(H_2)$ and $L(H_1)\subseteq L(H_2)$, and they have equal alleles on the same locus in $L(H_1)\cap L(H_2)$, then we say H_1 contains H_2, or H_2 is contained in H_1, which is denoted by $H_1\supseteq H_2$.

Definition 2. For schemata H_1, $H_2\in\{0,1,*\}^L$, their determinate bit locus sets are $L(H_1)=\{l_{11},l_{12},\cdots,l_{1O(H_1)}\}$, $L(H_2)=\{l_{21},l_{22},\cdots,l_{2O(H_2)}\}$. If $L(H_1)\cap L(H_2)\neq\phi$ or $|L(H_1)\cap L(H_2)|>0$, and they have equal alleles on the same locus in $L(H_1)\cap L(H_2)$, then we say H_1 intersects with H_2.

For instance, schema "0 * *" contains "01 *" and "01 *" intersects with "* 10". Now, we can use the containment and intersection between schemata to describe the competitive relations.

Definition 3. For schemata H_1, $H_2\in\{0,1,*\}^L$, their determinate bit locus sets are $L(H_1)=\{l_{11},l_{12},\cdots,l_{1O(H_1)}\}$, $L(H_2)=\{l_{21},l_{22},\cdots,l_{2O(H_2)}\}$. If $O(H_1)=O(H_2)=k$ and $L(H_1)=L(H_2)$, or in other words they have the same locus set but different alleles, then we say there exists k-order primary schemata competition between H_1 and H_2.

All of the schemata that have k-order primary schemata competition relations between any pairs of them are called k-order primary schemata competition set, denoted by $\mathrm{H}^k=\{H_1,H_2,\cdots,H_k\}$, $|\mathrm{H}^k|=2^k$. It can be seen that we have C_L^k of H^k, all of which form the k-order primary schemata competition set group. In the following discussion, H^k means any set in the group with k-order primary schemata competition. When $k=L$, we have the only H^L and $|\mathrm{H}^L|=2^L$.

For instance, schemata "* 0 * * 0", "* 1 * * 0", "* 1 * * 1", "* 0 * * 1" have the relation of 2-order primary schemata competition, and $\mathrm{H}^2=\{*0**0,\ *1**0,\ *1**1,\ *0**1\}$ is the 2-order primary schemata competition set. For any $H\in\mathrm{H}^2$, $L(H)=\{2,5\}$.

Similarly, the secondary and indirect competitions between schemata can be defined[6]. Since this paper focuses on primary schemata competition and associated theorems, those definitions will not be presented here.

Definition 4. $\forall H_i\in\mathrm{H}^k$, $\exists H_0\in\mathrm{H}^k$, and $f(H_0)\geqslant f(H_i)$ holds, H_0 is called the global winner on H^k, denoted by $H_0^k=\max(\mathrm{H}^k)$.

Definition 5. For H^{k_1}, H^{k_2}, $(k_1<k_2)$, $\forall H_i\in\mathrm{H}^{k_1}$, $\exists H_j\in\mathrm{H}^{k_2}$, such that $H_i\supset H_j$. And $\forall H_j\in\mathrm{H}^{k_2}$, $\exists H_i\in\mathrm{H}^{k_1}$, such that $H_j\subset H_i$. Then H^{k_1}, H^{k_2} are called to be relevant.

For instance, $\mathrm{H}^2=\{11***, 10***, 01***, 00***\}$ and $\mathrm{H}^3=\{111**, 110**, 101**, 100**, 011**, 010**, 001**, 000**\}$ are the 2-order and 3-order primary schemata competition sets respectively, then H^2, H^3 are relevant by Definition 5. The relevance of H^{k_1}, $\mathrm{H}^{k_2}(k_1<k_2)$ means that the direct competition among schemata in H^{k_1} contains that in H^{k_2}, or the direct competition among schemata in H^{k_2} is the development and

intensification of that in H^{k_1}.

3 Schema deceptiveness and GA deceptive problems

Definition 6. For relevant primary schemata competition set H^{k_1}, $\mathrm{H}^{k_2}(k_1 < k_2)$, $H_0^{k_1} = \max(\mathrm{H}^{k_1})$, $H_0^{k_2} = \max(\mathrm{H}^{k_2})$, if $H_0^{k_2} \not\subset H_0^{k_1}$, then there exists the schema deceptiveness from H^{k_1} to H^{k_2}, and this kind of optimization problem is called the GA deceptive problem.

For instance, if $H_0^2 = \max(\mathrm{H}^2) =$ "11 * * *", $H_0^3 = \max(\mathrm{H}^3) =$ "000 * *" for an optimization problem when solved by GA, then $H_0^3 \not\subset H_0^2$, so that there exists schema deceptiveness from H^2 to H^3, and this optimization is a GA deceptive problem.

With GA deceptive problems, the global winner on a low order primary schemata competition set has different schema characteristics from that on a relevant high order one, which means that the guiding direction of the low order schemata is not consistent with that of high order schemata. So the building block hypothesis does not hold any more, and the corresponding optimization by population searching constitutes the GA hardness.

As to non-monotone or multi-modal functions, there exist schema deceptiveness from H^{k_1} to H^{k_2}, and they will make up some degree of hardness and difficulty for GA. Schema deceptiveness may appear between different orders of schemata, and it will be intensified with the increasing of $k_2 - k_1$.

For continuous parameters optimization problem opt $f(X)(X \subseteq \mathbb{R}^m)$, suppose that its global optima is X^*, the binary encoded space and schema space are $\{0,1\}^L$, $\{0,1,*\}^L$ separately, the associated optimal string or schema is H^* of L order.

Definition 7. Given the k-order primary schemata competition set H^k, for any $j < k (j = 1,2,\cdots,k-1)$ and relevant H^j, if there exists the schema deceptiveness from H^j to H^k, and $H_0^j \supset H_1^k$, $H_1^k \in \mathrm{H}^k$, $H_1^k \neq H_0^k$, $f(X)$ is called the k-order fully deceptive problem. And more, if $k = L$, $H^* = H_0^L = \max(\mathrm{H}^L)$, $f(X)$ is called the fully deceptive problem.

In the above definition, H_1^k is called the k-order deceptive attractor.

Definition 8. Given the k-order primary schemata competition set H^k, for any $j < k (j = 1,2,\cdots,k-1)$ and relevant H^j, if there exists the schema deceptiveness from H^j to H^k, $f(X)$ is called the k-order consistently deceptive problem. And more, if $k = L$, $H^* = H_0^L = \max(\mathrm{H}^L)$, $f(X)$ is called the consistently deceptive problem.

The k-order consistently deceptive problem dose not require that there exists the same k-order deceptive attractor H_1^k for any $j < k$ $(j = 1,2,\cdots,k-1)$. The k-order deceptive attractor H_1^k for relevant sets H^j and H^k may be any schema in H^k, or other k-order schema in $\{0,1,*\}^L$. Obviously, fully deceptive problems are certainly the consistently deceptive problems, and their propositional definitions are described as follows:

(i) Fully deceptive problem:

$$\forall (H_{i_1}^j, H_{i_2}^j)\{f(H_{i_1}^j) > f(H_{i_2}^j) \mid H_{i_1}^j, H_{i_2}^j \in \mathrm{H}^j;\ H_1^k \subset H_{i_1}^j,\ H_1^k \not\subset H_{i_2}^j;\ i_1 \neq i_2\}. \quad (1)$$

(ii) Consistently deceptive problem:

$$\forall (H_{i_1}^j, H_{i_2}^j)\{f(H_{i_1}^j) < f(H_{i_2}^j) \mid H_{i_1}^j, H_{i_2}^j \in \mathrm{H}^j;\ H_0^k \subset H_{i_1}^j,\ H_0^k \not\subset H_{i_2}^j;\ i_1 \neq i_2\}. \quad (2)$$

Non-consistently deceptive problems are called the partially deceptive, and have two types as below.

Definition 9. Given the k-order primary schemata competition set H^k, for most $j < k(j = 1,2,\cdots,k-1)$ and relevant H^j, if there exists the schema deceptiveness from H^j to H^k, $f(X)$ is called the k-order chronically deceptive problem. And more, if $k = L$, $H^* = H_0^L = \max(\mathrm{H}^L)$, $f(X)$ is called the chronically deceptive problem.

In other words, with the chronically deceptive problem, there exists at least one $\mathrm{H}^j(j < k(j = 1,2,\cdots,k-1))$ which forms schema deceptiveness to H^k. Experiential computations indicate that the chronically deceptive problem is also hard to some extent for GA[6].

Definition 10. Given the k-order primary schemata competition set H^k, for $j = 1$ and relevant H^1, if there exists the schema deceptiveness from H^1 to H^k, $f(X)$ is called the k-order fundamentally deceptive problem. And more, if $k = L$, $H^* = H_0^L = \max(\mathrm{H}^L)$, $f(X)$ is called the fundamentally deceptive problem.

4 Some theorems about schema deception

Theorem 1. Suppose that $f(X)$ is a k-order fully deceptive problem, the global winner on H^k is $H_0^k = (a_1, a_2, \cdots, a_L)(a_l \in \{0,1,*\};\ l = 1,2,\cdots.L)$, then its k-order deceptive attractor can be got by the complementary operation:

$$H_1^k = O(\mathrm{com},\ H_0^k) = \begin{cases} 1 - a_l, & \text{if } a_l = 0,1, \\ a_l, & \text{else.} \end{cases}$$

For instance, if $H_0^k =$ "$**1*1*1**$", then $H_1^k = **0*0*0**$. For $H^* =$ "$111\cdots111$", its fully deceptive attractor is $H_1^L =$ "$000\cdots000$".

Proof. Suppose that H_0^k is the global winner on H^k, and H_1^k, H_0^k have an identical allele g on a locus, the complementary bit value of which is $c = O(\mathrm{com},\ g)$. According to the definition of schemata relevance and deceptiveness, for any two k' ($k' < k$) order schemata "$*\cdots*g*\cdots*$", "$*\cdots*c*\cdots*$"in $\mathrm{H}^{k'}$, $f(*\cdots*g*\cdots*) < f(*\cdots*c*\cdots*)$ holds. Then the allele of H_1^k at this locus must be c but not g. So we have $H_1^k = O(\mathrm{com},\ H_0^k)$.

Theorem 2. For k-order consistently deceptive problem, if the fitness of its k-order deceptive attractor is equal to or lower than that of the schema which is two determinant bits different from it, then this k-order consistent deceptiveness will not hold any more.

Proof. Here we only need to consider the consistent deceptiveness between H^{k-1} and H^k.

Suppose the global winner on H^k is $H_0^k =$ "$*\cdots*1111*\cdots*$". Then its deceptive attractor is $H_1^k =$ "$*\cdots*0000*\cdots*$". Give any schema that is two bits different from H_1^k: $H_z^k =$ "$*\cdots*1001*\cdots*$", and two of one bit different schemata $H_2^k =$ "$*\cdots*$"$1000*\cdots*$, $H_3^k =$ "$*\cdots*0001*\cdots*$". It might as well be assumed that $f(H_2^k) > f(H_3^k)$.

Suppose $H_2^{k-1} =$ "$*\cdots*000\ **\cdots*$", $H_3^{k-1} =$ "$*\cdots*100\ **\cdots*$. Since there exists schema deceptiveness between H^{k-1} and H^k, $f(H_2^{k-1}) > f(H_3^{k-1})$ holds. According to the calculation of schema, we have

$$f(H_2^{k-1}) = (f(H_1^k) + f(H_3^k))/2,\ f(H_3^{k-1}) = (f(H_z^k) + f(H_2^k))/2.$$

Hence

$$f(H_1^k) + f(H_3^k) > f(H_z^k) + f(H_2^k), \text{ or } f(H_1^k) > f(H_z^k) + (f(H_2^k) - f(H_3^k)).$$

Theorem 2 shows that for a k-order consistently deceptive problem with H^k, its deceptive attractor $H_1^k = O(\mathrm{com},\ H_0^k)$ is a local optimal hyperplane in the Hamming space of $\{0,1\}^L$. Or if it is not the local optimal hyperplane, then it can only be one bit different from, or located 1 of Hamming distance away from the local optimal hyperplane.

Theorem 3. For $f(X)$ with binary encoded space $\{0,1\}^L$, if

(i) there exists no schema deceptiveness between any order relevant

$$\mathrm{H}^j(j < k;\ j = 1,2,\cdots,\ k-1) \text{ and } \mathrm{H}^L;$$

(ii) and there exists only one $H^* = H_0^L = \max(\mathrm{H}^L)$;

then $H^* = H_0^L = \bigcap_{l=1}^{L} H_0^1(l)$, or H^* can be calculated by the intersection of the global winners on the 1-order primary schemata competition sets H^1 of L.

Proof. Here, schemata H^1 and H^k are considered first.

Suppose the global winner on H^L is H_0^L = "$1\cdots111\cdots1$". If there does not exist schema deceptiveness from H^1 to H^L, for any $H_{l,1}^1 H_{l,2}^1 \in \mathrm{H}_l^1$($l$ is the determinate bit), it might be as well that we take $H_{l,1}^1$ = "$* \cdots * 1 * \cdots *$", $H_{l,2}^1$ = "$* \cdots * 0 * \cdots *$". Then according to the definition of schema deceptiveness, $f(H_{l,1}^1) > f(H_{l,2}^1)$ holds. So, $H_{l,1}^1$ is the global winner on H_l^1, $H_0^1(l) = \max(\mathrm{H}_l^1) = H_{l,1}^1$.

Now consider the relation of 2-order primary schemata competition set H^2 and H^L. For any $H_{l,1,1}^2, H_{l,2,1}^2 \in \mathrm{H}^2$($\mathrm{H}_l^1$ and H^2 are relevant), we may as well take, $H_{l,1,1}^2$ = "$* \cdots * 11 \cdots *$", $\mathrm{H}_{l,2,1}^2$ = "$* \cdots * 00 \cdots *$". Then $f(H_{l,1,1}^2) > f(H_{l,2,1}^2)$ holds, so that $H_0^2 = \max(\mathrm{H}^2) = H_{l,1,1}^2$.

Hence, there does not exist any schema deceptiveness from H^1 to H^2. Similarly, it can be proved that there exists no schema deceptiveness between relevant H^k and $\mathrm{H}^{k+1}(L-1 \geqslant k > 1)$. Therefore, the feedback information afforded by the competition among one order schemata is consistent with that of the higher order schemata competition.

The intersection of L 1-order global winner on H^1 is $\bigcap_{l=1}^{L} H_0^1(l)$ = "$1\cdots111\cdots1$", so

$$H^* = H_0^L = \bigcap_{l=1}^{L} H_0^1(l). \tag{3}$$

Besides, it can be deduced that if $f(X)$ is a chronically deceptive problem, even though there exists schema deceptiveness between middle order primary competition sets, which will guide the searching of GA away from the global optima, we still can get correct estimation of the optimal string by eq. (3)[9].

With the above definition and theorems, monotone and single-peak functions are not schema deceptive problems of any type and any order, so that such problems are GA-easy. It should be pointed out that we assume $H_0^k = \max(\mathrm{H}^k)$ as the only global winner on H^k in discussion. If there are many global winners on H^k, our conclusions are still correct when taking other global winners as deceptive attractors. But the analyses based on schemata competition will become more complicated, and need further study.

5 Schema deceptiveness and the implicit parallelism, convergence of GA

According to the schema theorem, GA can process $O(n^3)$ of schemata with the population of size n, which is called the implicit parallelism. For optimization problems with no schema deceptiveness, GA would process schemata competition and sample schemata in line with the schema theorem. Meanwhile, the global winner of the different order primary competition sets have consistent structure of alleles. Therefore, the implicit parallelism for GA holds exactly.

As to multi-modal functions, there usually exists schema deceptiveness of different types and varying extents. When the population of size n is created by uniformly random sampling on $\{0,1\}^L$, 1-order schemata have equal survival numbers. Because of the schema competition under deception, the sampling of 2-order and above schemata will not keep with uniform sampling. Under the guiding of deceptive attractor, and because of the inconsistent alleles structure of effective genes, only the winners transfer their genetics to the individuals in next generations and sampling according to schema theorem. Other schemata containing some alleles of the global optimal string would disappear gradually. Hence, implicit parallelism dose not hold any more for GA.

Similarly, under the guiding of deceptive attractor, the population searching of GA gets away from the subspace where global optima is located. It can be seen that schema deceptiveness affects the convergence of GA searching to the global optima. However, compared with traditional randomized searching methods, GA shows typical robustness. By designing proper genetic strategies, GA can be endowed with high capability to overcome schema deceptiveness, which is an important research direction hereafter.

6 Numerical examples

The practical optimization problems in automatic control, machine learning and management of economics and enterprises are all complicated ones, and will form series schema deceptiveness when solved with GA. Now, two examples are designed to show the effect of schema deceptiveness on the population searching of GA.

6.1 Trap function

Trap function is a special bit unitation function[8,9], and takes the form:

$$f(u) = \begin{cases} \frac{a}{z}(z - u), & \text{if } u \leqslant z, \\ \frac{b}{L - z}(u - z), & \text{otherwise,} \end{cases} \tag{4}$$

where $a < b$. The global optimal string is "111⋯111", and its deceptive attractor is "000⋯000", $f(L) = b$, $f(0) = a$. For $u < z$, $f(u)$ is a monotonically decreasing function. When $u \geqslant z$, $f(u)$ is a monotonically rising function.

The parameters of GA are set as: string length of $L = 30$, population size of $n = 100$, two-points crossover with $p_c = 0.6$, mutation with $p_m = 1/L$ 0.0333.

Assume $a = 80$, $b = 100$. Let us analyse the relation of the searching capability of GA and $z = \text{int}(L/2) + k$, where proportionate selection and elitist retaining are used, and the maximum generation of evolution is prescribed as 200. Each case is repeated 120 times of calculation. The statistical results are listed in table 1.

For $z > \text{int}(L/2) = 15$, the schema deceptiveness of trap function gets stronger with the increasing z, and the population finally converges to $a = 80$. When $z \leqslant \text{int}(L/2) = 15$, the schema deceptiveness is weakened as z decreases, and $b = 100$ becomes the converged point of GA population. The competitions of the same order of schemata are intensified around $z = \text{int}(L/2) = 15$, and if the schemata containing effective alleles or the alleles of deceptive attractor become dominant, the population would converge quickly to the global optima or to the deceptive attractor.

Table 1 The relation of GA searching capability and parameter z

(1)	$z > \text{int}(L/2) = 15$					
	$z = 15 + 1$	$z = 15 + 3$	$z = 15 + 5$	$z = 15 + 7$	$z = 15 + 9$	$z = 15 + 12$
The averaged generations and probability of converging to a	31.25 33.33%	29.73 91.67%	30.06 100%	32.08 100%	33.48 100%	35.83 100%
The averaged generations and probability of converging to b	27.50 66.67%	34.00 8.33%				
(2)	$z \leqslant \text{int}(L/2) = 15$					
	$z = 15 - 9$	$z = 15 - 7$	$z = 15 - 5$	$z = 15 - 3$	$z = 15 - 1$	$z = 15$
The averaged generations and probability of converging to a						32.50 5.56%
The averaged generations and probability of converging to b	35.35 100%	32.91 100%	31.86 100%	28.08 100%	26.83 100%	24.03 94.44%

6.2 Needle-in-a-haystack problem (NiH)

As to the Needle-in-a-haystack problems (NiH), the global optima are surrounded by the worst solutions, which blocks the way of schema recombination from low order to high order. In these cases, it is very difficult for GA to move from local optima to the neighborhood of the global optima in randomized searching. For analyzing the searching process of GA, a typical NiH optimization problem is designed as below:

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

where $a = 3.0$, $b = 0.05$, $\max f(0,0) \approx 3600$, the four local optimal points are $(-5.12, 5.12)$, $(-5.12, -5.12)$, $(5.12, 5.12)$, $(5.12, -5.12)$, and their function values are 2748.78. With the changing of control parameter $\{a, b\}$, this function will be featured with different degrees of schema deceptiveness, and the above four points are the deceptive attractors.

The parameters of GA are set as: string length of $L = 40$ (x, y are all encoded with 20 of binary bits), population size of $n = \{20, 40, 60, 80, 100, 120\}$, two-points crossover with $p_c = 0.6$, mutation with $p_m = 1/L = 0.025$. Proportionate selection and elitist retaining is used, and maximum generation of population evolution is prescribed as 300. The computation is repeated 120 times of each case, and statistical results are listed in table 2.

Table 2 The performance of GA for NiH with different population sizes

	$n = 20$	$n = 40$	$n = 60$	$n = 80$	$n = 100$	$n = 120$
The averaged generations and probability of converging to global optima	88.60 20.83%	83.43 52.50%	81.21 57.50%	82.59 60.83%	83.37 64.17%	81.43 66.67%

While population size is fixed at 60, the searching trajectories of GA under schema deception and after overcoming schema deception are displayed in fig.1, and the fitness growing of the current best individual in population along with generations of evolution is recorded in fig. 2.

Schema deception led the searching process towards deceptive attactors, and then GA was trapped in the local optimal points (located randomly at one of the four points). When schema deception was overcome by genetic operators, GA shifted its searching to the neighborhood of global optima, and finally converged to global optimal point. It can be seen that the global opti-

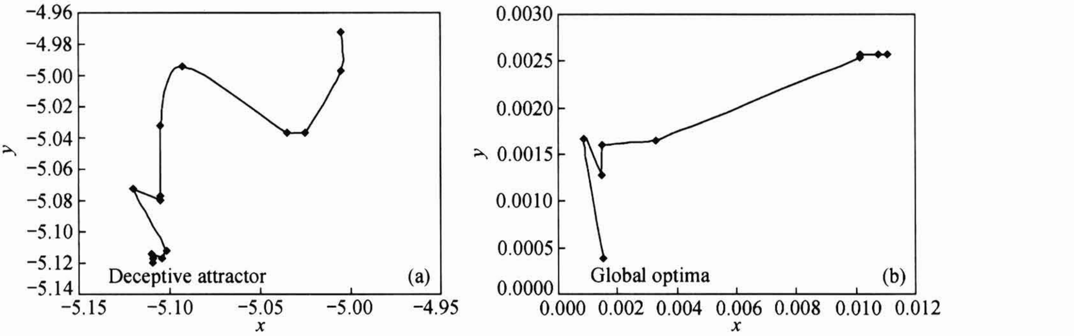

Fig. 1. The searching trajectory of GA. (a) With schema deception; (b) after overcoming schema deception.

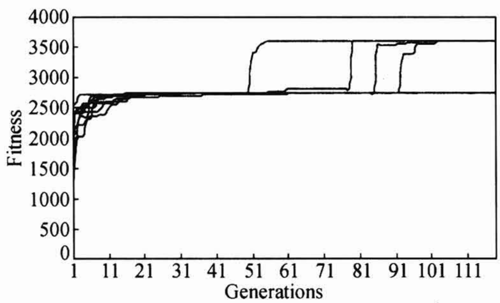

Fig. 2. The performance of GA for NiH.

mal convergence of GA becomes a stochastic event under the circumstance of hard schema deceptiveness.

7 Conclusion

In this paper, some new concepts and theorems are defined and improved to describe schema deceptiveness and GA deceptive problems, which will help to explain the searching behavior of GA population, and to enrich the basic theory of evolution dynamics of GA. Meanwhile, these results can assist the designing of genetic strategies and operators[8,9]. As to the case that there exists partially schema deceptiveness, the population searching of GA needs to be analyzed in detail. Besides, there are other factors, such as population size, population initializing, schema sampling errors, the error in stochastic operations of crossover and mutation, etc., which affect the performance of GA. Different types of genetic operators and parameters, designed properly or not, can produce contradictory effects in overcoming schema deceptiveness. All of these problems need to be investigated theoretically and experimentally.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 69974026).

References

1. Belew, R. K., Vose, M. D., Foundation of Genetic Algorithms, 4, San Francisco: Morgan Kaufmann, 1997.
2. Mitchell, M., An Introduction to Genetic Algorithms, Cambridge: The MIT Press, 1996.
3. Bethke, A. D., Genetic algorithm as function optimizers, Ph.D. Dissertation, University of Michigan, 1980.
4. Goldberg, D. E., Simple genetic algorithm and the minimal deceptive problem, in Genetic Algorithms and Simulatied Annealing (ed: Davis, L.), San Francisco: Morgan Kaufman, 1987, 74—88.
5. Das, R, Whitley, D., The only challenging problems are deceptive: global search by solving order-1 hyperplanes, Proceedings of ICGA (eds. Belew, R., Booker, L.), San Francisco: Morgan Kaufman, 1991, 166—173.
6. Whitley, D., Fundamental principles of deception in genetic search, in Foundations of Genetic Algorithms (ed. Rawlins, G.), San Francisco: Morgan Kaufmann, 1991, 221—241.
7. Deb, K., Goldberg, D. E., Analyzing deception in trap functions, IlliGAL Report No.91009, Urbana: University of Illinois Genetic Algorithms Laboratory, 1991.
8. Liepins, G. E., Vose, M. D., Representational issues in genetic optimization, Journal of Experimental Theory and Instruments, 1990, (2): 4—30.
9. Goldberg, D. E., Korb, B., Deb, K., Messy genetic algorithms: motivation, analysis, and first results, Complex Systems, 1989, (3): 493—530.