Abstract:
Through mechanism analysis of simple genetic algorithm (SGA), every genetic operator can be considered as a linear function. So some disadvantages of SGA may be solved if genetic operators are modified to nonlinear function. According to the above method, nonlinear genetic algorithm (NGA) is introduced. By the analysis of the concept of schema, new schema theorem is proved. The above work establishes a completely new mathematical analytic method of genetic algorithms schema theory and contributes much to perfection and development of the theory of genetic algorithm.

Keywords:
Schema Theorem, Nonlinear Genetic Algorithm, Distance Space

1. Introduction.

Genetic algorithms (GAs) are search and optimization techniques based on the mechanics of natural selection and natural genetics, which effectively and parallelly exploit historical information to speculate on new search points with expected improved performance.

During the iterative process of genetic algorithms, short-defining-length, low-order, and above-average schema receive at least exponentially increasing numbers of trials in subsequent generations. These building blocks are sampled, recombined, and resembled to form strings of potentially higher fitness by the genetic operations of reproduction, crossover, mutation and so on.

Though mechanism analysis of GA operators, we found that every genetic operator is equal to a linear function. So, if we want to reduce the emergence of premature convergence and other disadvantages, it is a useful method to modify the structure of genetic operators as nonlinear function. So nonlinear genetic algorithm (NGA) is introduced, and cursory analysis of NGA is done. For the further understanding NGA, it is important to give a new schema theorem to explain the mechanism of NGA.

In Section 2, a simple framework of nonlinear genetic algorithm, including the form of nonlinear crossover and mutation operators, the structure of the nonlinear genetic algorithm is given; in Section 3, some traditional theoretical results are introduced; in Section 4, some new theoretical results about the schema theorem are given.

2. The Real-Coded Nonlinear Genetic Algorithm (RCNGA).

Without loss of generality, we will only deal with global minimization problems, which can be formalized as

$$\min F(x), x \in D \subseteq \mathbb{R}^n$$

(1)

where D is a bounded and continuous subset, $F : D \rightarrow \mathbb{R}^n$ is a bounded real-valued function. The objective function F does not need to be continuous.
In briefly, we’ll give some concept of RCNGA, more details can be found in [4].

*Nonlinear crossover operator: let $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ are parent vectors, then the offspring $w = (w_1, w_2, ..., w_n), v = (v_1, v_2, ..., v_n)$ are computed by

\[ w_j = f^{-1}(\alpha f(x_j) + (1 - \alpha)f(y_j)) \]
\[ v_j = f^{-1}(\alpha f(y_j) + (1 - \alpha)f(x_j)) \]

where $f$ is normalized function, $f : D_j \rightarrow [0,1], D = D_1 \times D_2 \times ... \times D_n, \alpha \in [0,1]$.

*Nonlinear mutation operator:

let $x = (x_1, x_2, ..., x_n)$ is parent vector, then the offspring $w = (w_1, w_2, ..., w_n)$ is computed by:

\[ w_j = f^1(g(f(x_j))) \]

$f, f^1$ are the same function as above, and $g : [0,1] \rightarrow [0,1]$ is a mutation function.

The algorithm can be sketched as follows:

**PROGRAM RCNGA**

**BEGIN**

Initialize parameters: population size $N$, crossover probability $P_c$, and mutation probability $P_m, f, g, f^1, k := 0$;

Choose an initial population $P(0)$;

**REPEAT**

Perform nonlinear crossover;
Perform nonlinear mutation;
Determine the fitness of each individual;
Perform selection;
$K := K + 1$;
**UNTIL** some stopping criterion applies;
Output the best answer;

**END.**

3. **Traditional Analysis.**

Holland[1] provided the initial formal analysis of the behavior of GAs by characterizing how they biased the makeup of new offspring in response to feedback on the fitness of previously generated individuals. By focusing on hyperplane subspaces of $L$-dimensional spaces (i.e., subspaces characterized by hyperplanes of the form $"-d_1-\ldots-d_L\ldots\"$, where the $d_i$ are the defining positions of the hyperplane), Holland showed that the expected number of samples(along individuals) allocated to a particular $k$-th order hyperplane $H$ at time $t+1$ is given by:

\[ m(H_k, t+1) > m(H_k, t) \frac{f(H_k)}{\bar{f}} \times (1 - kp_m - p_c p_d(H_k)) \]

In this expression, $f(H_k)$ is the average fitness of the current samples allocated to $H_k$, $\bar{f}$ is the average fitness of the current population, $p_m$ is the probability of using the mutation operator, $p_c$ is the probability of using the crossover operator, and $p_d(H_k)$ is the probability that the crossover operator will be “disruptive” in the sense that the children produced will not be members of the same subspace as their parents.

4. **New Schema Theorem.**

When we use SGA to solve question (1), the first step that is typically made is that the variable representing parameters can be represented by bit string, suppose $\Omega$ is the set with all bit strings, then the continuous space $D$ will be replaced by discrete space $\Omega$, and SGA will search within $\Omega$. Of course, a schema can be think as a finite subset of $\Omega$, for example, schema $1*1*$ is the set $\{1010,1110,1011,1111\}$. Because of the element of
discrete space $\Omega$ is finite, and element of schema is finite, so number of schema is finite, this means in the worst status GAs can search every schema if time is large enough.

In [2], authors give the schema of real-coded GAs, there is a question: since element of continuous space $D$ is infinite, and element of schema is finite, so the number of schema is infinite, this means in the worst status real-coded GAs can’t search every schema even time is large enough. This means the Holland’s schema has some defaults which limit to use NGA. In the next part we will discuss the question and attempt to give some results.

If $T$ is a hyperplane (subset) of $\Omega$, we have the schema characterized by hyperplanes of the form $\ldots -d_1 - d_2 - \ldots - d_k - \ldots$, where the $d_i$ are the defining positions of the hyperplane, of course schema $\ldots -d_1 - d_2 - \ldots - d_k - \ldots$ (call it $P$) defines a subset of $\Omega$, now it has two situations: (1) $T \subset P$; (2) $T = P$. It means for an arbitrary subset, it is satisfied: belongs to one schema or equals to one schema. It’s called schema conditions. Now, from the concept of subset, definition of schema can be extended: For a discrete space $\Omega$, subset group $\Sigma$ (if and only if $A$ is a schema, schema group) is given, and $\Sigma$ is satisfied: for any subset $M \subseteq \Omega$, $M \notin \Sigma \Rightarrow \exists N \in \Sigma, M \subseteq N \quad \forall 5\|$ then if $A \in \Sigma$, $A$ is defined a abstract schema. So, (5) is schema condition, and Holland’s schema of GAs is a subset of new schema. Now, from another point, the distance is considered. The distance of individual $x$ and $y$ can be defined:

$$d(x, y) = \frac{1}{k} \left| X OR(x, y) \right| \quad (6)$$

$|X|$ is the number of element of set $X$.

Thus a distance space is defined, a subset by schema defined is some region although the space is a discrete space. Then by the introduction of distance, discrete space and continuous space are uniform into distance space, and schema is some region. So, we give a specific definition of schema.

For an arbitrary subset $[\alpha_j, \beta_j]$ of $[0, 1]$, $U = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] \times \ldots \times [\alpha_n, \beta_n]$ is defined a schema. Thus if we suppose the population of $t$th generation is $\{x_1, x_2, \ldots, x_n\}$, and $x_j = (x_j^1, x_j^2, \ldots, x_j^n)$, then the maximum schema of population is $M = [y_1, z_1] \times [y_2, z_2] \times \ldots \times [y_n, z_n]$.

$$y_j = \min \{x_j^k \}_{k=1}^n, z_j = \max \{x_j^k \}_{k=1}^n \quad \text{Possible schema is defined as } N = g(M).$$

In fact the schema as defined above is a hypercube in domain. For an arbitrary subset, because of $[0, 1]$ is Euclidian distance, there must exist a hypercube, the condition (5) is satisfied.

Now we consider RCNGA’s schema theorem. Because of the definition of $M$, the results of nonlinear crossover operator is the subset of $M$, but schema $H$ is subset of domain $D$, so the probability of nonlinear crossover results are not belong to schema $H$ should satisfied:

$$P \leq \frac{m(M \setminus H)}{m(M)} \quad (7)$$

where $M \setminus H$ means the different set of $M$ and $H$, $m(.)$ is Lebesgue measure. So we have:

**Theorem 1.** If only selection and crossover occur, the number of schema $H$ in the next generation is

$$m(H, t + 1) \geq m(H, t) \frac{F(H)}{F} \left[1 - p_c \frac{m(M \setminus H)}{m(M)} \right] \quad (8)$$

By the same reason, the probability of nonlinear mutation result is not belong to schema $H$ should satisfied:

$$P \leq \frac{m(N \setminus H)}{m(N)} \quad (9)$$

So we have:
Theorem 2. If only selection and mutation occur, the number of schema H in next generation is:

\[
m(H, t + 1) \geq m(H, t) \frac{F(H)}{F} \left[1 - p_m \frac{m(N \setminus H)}{m(N)}\right]
\]

(10)

Now, RCNGA’s Schema theorem is:

For a given schema H, let:

1. \(m(H, t)\) be the relative frequency of the schema H in the population of the \(t^{th}\) generation,
2. \(F(H)\) be the mean fitness of the elements of H,
3. \(m(.)\) be the measure,
4. \(M\) be the max schema of the current population,
5. \(N\) be the Possible schema,
6. \(P_c\) be the nonlinear crossover probability,
7. \(P_m\) be the nonlinear mutation probability,
8. \(\bar{F}\) be the mean fitness of the current population.

Then:

\[
m(H, t + 1) \geq m(H, t) \frac{F(H)}{F} \left[1 - p_c \frac{m(M \setminus H)}{m(M)} - p_m \frac{m(N \setminus H)}{m(N)}\right]
\]

(11)

5. Conclusions.

In this paper, an extend of Holland’s schema is given. Further research will include the essential of schema and the implementation of nonlinear genetic algorithm.

References:


