Fundamental Research on Genetic Algorithms

Zhou Chang

1.Department of Mathematics, Northwest University, Xi'an,China
2. College of Science, Xi'an University of Posts &Telecommunications , Xi'an,China
maytheday@163.com

Abstract—**Genetic Algorithms uses the population selection technology, newer population is generated from current population by cross, mutation and other genetic operations, and gradually make the population to evolve the optimization result. Compared with traditional optimization methods, Genetic Algorithm has two notable characters, one is the latent parallelism and the other is searching in the whole area. And because of its independence, global optimization and implicit parallelism, GA is developed and applied by more and more people, so has been widely used in many fields, such as adaptive control, combinatorial optimization, pattern recognition, machine learning, artificial life and management decisions,etc.**

Keywords- Genetic Algorithms, Implicit Parallelism, Global Searching, Robustness

I. INTRODUCTION

In recent years, with the expansion of application fields of artificial intelligence, the use of traditional AI methods in knowledge representation, information processing and combinatorial explosion faces more difficult situation, so the object of study is to seek an algorithm with the self-organization, self-adaptation and self-learning which suitable for numerical computing. The nature has contributed to inspire our research, the genetic algorithms were first promulgated by J.Holland in 1975, and he was inspired by the evolution theory [1]. GA is an optimization method with the advantages by high degree of parallel, random and self-adapting, which based on the "Survival of the fittest". It represent the problem solving as the evolution of chromosomes, The optimization result was searched by selection ,cross and mutation and other genetic operations over the generations.GA is a commonly used optimization algorithm, its encoding technique and genetic operations proved reasonably easy, and the optimal result was independent of the restrictive conditions. Implicit parallelism and global searching are two characteristics of GA. Recently, with the development of computer technique,GA was receive more and more attention, and it has been used extensively in many fields such as adaptive control, combinatorial optimization, pattern recognition, machine learning, artificial life, management decisions, neural network and VLSI design,etc.

II. THE CHARACTERISTIC OF GENETIC AND EVOLUTION MECHANISM OF BIOLOGY

1) The genetic information was all included in chromosomes, the biological characters was depended on the chromosomes.

2) Chromosomes are organized by genes and their sequencing, the genetic and evolve are occur in chromosomes.

3) Multiplicative process of biological was completed by the reproduction process of genes.

4) New species with new character was generated by the cross and mutation of the homologous chromosomes [2].

5) Genes or chromosomes with better adaptability have a better chance than that with poor adaptability to pass from one generation to the next.

III. THE FORMAL DEFINITION OF GA

GA can be defined an eight-component tuple:

In GA= （C，E，P_0，M，Φ , Γ , Ψ , T ） ,
C——encoding method of individual
E——fitness evaluation function
P0——initial population
M——population size [3]
Φ ——selection operator
Γ ——crossover operator
Ψ ——mutation operator
T ——the conditions for termination of heredity

IV. THE BASIC MECHANISM OF GA

A. Encoding method of chromosomes

The common method of GA is binary code, represent the individual of population as the binary symbol string with the fixed length[4],whose gene alleles was consists of binary symbol set{0，1}.The genes value of each individual in initial population was generated by even-distributed random number. For example, X=100111001000101101 express an individual, the length of its chromosome is n=18. Apart from the binary code, there are many other coding methods such as Float coding, Gray coding, Symbolic coding and Multiparameter coding, etc.

B. Fitness function

We introduce a fitness function which can measure each chromosome to display the adaptive ability of chromosomes. The algorithm evaluates the fit and unfit quality of the

individual by fitness function, which is "survival of the fittest". In optimization problem, fitness function is objective function which can reflect the difference between each chromosome and the best solution chromosome. If the difference is smaller, then the corresponding difference of fitness function value is smaller, or else is bigger. The value range of fitness function has great relationship with the object of problem.

C. Genetic operation

The operations of GA are the following three：selection, crossover and mutation. The improved genetic algorithm has extended the operations of genetic algorithm for high efficiency.Seclection operation also be called reproduction，the individual which be eliminated or inherited from one generation to the next was depended on the fitness. Usually, the offspring that have bigger fitness are chosen to put in next evolution. The cross operation is that select two individuals as the parent individual, and exchange the code in part. Mutation operation is that change some numerical code in number strings.

GA used the following three kind of genetic operators:

- Selection operation uses the selection operator;
- Cross operation uses the crossover operator;
- Mutation operation uses the mutation operator.

GA has four parameters will need to be setting.

M：Population size, that is number of individuals in population, values range from 20 to 100.

T：The termination evolutionary generation, values range from 100 to 500.

p_c：Cross probability, values range from 0.4 to 0.99.

p_m：Mutation probability, values range from 0.0001 to 0.1.

V. THE TECHNOLOGICAL PROCESS OF GA

1) To generate the initial population with a group of initial individuals randomly and evaluate each individual's fitness value;

2) To determine whether such convergence rule are satisfied. If satisfied, output the result of searching; if not, the following steps are done;

3) Carry out the copying operation according with the fitness value;

4) Carry out the cross operation according with the cross probability p_c;

5) Carry out the mutation operation according with the mutation probability p_m;

6) Return to step (2)

VI. STEP OF GA

Step1 Select a coding method for problem, give an initial population POP(1),t：=1 with N chromosomes [5].

Step2 Computing the fitness function $f_i = fitness(pop_i(t))$ for each chromosome in population POP (t).

Step3 If the termination condition was satisfied, then the end; if not, computing the probability

$$p_i = \frac{f_i}{\sum_{j=1}^{N} f_i}, \quad i = 1,2,\cdots,N \tag{1}$$

And select some chromosomes randomly with the probability distribution (1) from POP (t) to compose a population New POP (t+1) = { $pop_j(t)$, $J = 1,2,\cdots,N$ }.

Note: Maybe one element in POP (t) is repeated in the set new POP (t+1).

Step4 We can get a Cross POP (t+1) with N chromosomes by mating.

Step5 Mutating one gene of chromosomes with a small probability p ,and formed Mut POP(t+1)；t：= t+1，the new population POP(t)= Mut POP(t)；Return to Step2.

VII. EXAMPLE

Finding the maximum value of the function

$$f(x) = x^2 \quad (x \text{ is an integer, and } x \in [0,31])$$

with the genetic algorithm.

Solution as follows [6]:

(1) Encoding

Since x is an integer which belongs to $[0,31]$, it can be expressed by five binary numbers. So we can use the binary code and the length of coded string is five. For example, $x = 0$ is expressed by 00000, and $x = 31$ is expressed by 11111. 0 and 1 are the genetic value.

(2) Establishing the initial population

If we assume that the population size is N =4, then the initial population is the four generated randomly binary coded string whose length is five. And assume that the generated randomly initial population is

$$S_{01} = 0\ 1\ 1\ 0\ 1$$
$$S_{02} = 1\ 1\ 0\ 0\ 1$$
$$S_{03} = 0\ 1\ 0\ 0\ 0$$
$$S_{04} = 1\ 0\ 0\ 1\ 0$$

(3) Fitness function

We can define the fitness function be $f(x)$, that is

$$f(S) = f(x),$$

where S is a binary string corresponding to the value of x. So the initial populations can be seen the table below:

TABLE I. INITIAL POPULATIONS

number	chromosome	x	fitness	percent	accumulative percent	times of selecting
S_{01}	0 1 1 0 1	13	169	14.3	14.3	1
S_{02}	1 1 0 0 1	25	625	52.88	67.18	2
S_{03}	0 1 0 0 0	8	64	5.41	72.59	0
S_{04}	1 0 0 1 0	18	324	27.41	100	1

The fitness of S_{02} is the largest in the four individuals, so it is the best individual.

(4) Selection operation

Assume that we select the individual using the roulette wheel, and the four random numbers which generated in turn are 0.85, 0.32, 0.12, 0.46, the new population after the selection is

$$S'_{01} = 1\,0\,0\,1\,0$$
$$S'_{02} = 1\,1\,0\,0\,1$$
$$S'_{03} = 0\,1\,1\,0\,1$$
$$S'_{04} = 1\,1\,0\,0\,1.$$

(5) Crossover

Assume that the crossover probability P_c is 50%, and then only half of chromosomes are involved in the crossover operation. If further assume that chromosomes are cross in pairs, and S'_{01} crosses S'_{02}, S'_{03} doesn't cross S'_{04}. See the table below:

TABLE II. CROSSOVER OF INITIAL POPULATIONS

number	chromosome	cross-object	crossover point	generation	fitness
S'_{01}	1 0 0 1 0	S'_{02}	3	1 0 0 0 1	289
S'_{02}	1 1 0 0 1	S'_{01}	3	1 1 0 1 0	676
S'_{03}	0 1 1 0 1	S'_{04}	N	0 1 1 0 1	169
S'_{04}	1 1 0 0 1	S'_{03}	N	1 1 0 0 1	625

Then the new population after the crossover operation is

$$S''_{01} = 1\,0\,0\,0\,1$$
$$S''_{02} = 1\,1\,0\,1\,0$$
$$S''_{03} = 0\,1\,1\,0\,1$$
$$S''_{04} = 1\,1\,0\,0\,1.$$

(6) Mutation

The mutation probability P_m is generally small. If there are no mutations in this cycle, then the population before its mutation is the first generation after its evolution. That is

$$S_{11} = 1\,0\,0\,0\,1$$
$$S_{12} = 1\,1\,0\,1\,0$$

$$S_{13} = 0\,1\,1\,0\,1$$
$$S_{14} = 1\,1\,0\,0\,1.$$

Then the operation of the 0 generation is finished, and repeat the process (4) ~ (6) for the first generation. We can get the second generation. Then the operation of the first generation is finished, and repeat the process (4) ~ (6) for the second generation. As for the second generation, after its selection operation, the maximum value can be found in the operation of crossover. See the table below:

TABLE III. CROSSOVER OF SECOND GENERATION

number	chromosome	cross-object	crossover point	generation	fitness
S'_{21}	1 1 0 0 1	S'_{22}	3	1 1 0 1 0	676
S'_{22}	1 0 0 1 0	S'_{21}	3	1 0 0 0 1	289
S'_{23}	1 1 0 0 1	S'_{24}	4	1 1 0 0 0	576
S'_{24}	1 1 1 1 0	S'_{23}	4	1 1 1 1 1	961

The corresponding chromosome is 1 1 1 1 1, so the optimized solution is given at $x = 31$.

VIII. THE ADVANTAGES OF GA

1) The manipulate objects of GA is a set of feasible solutions, not the single feasible solutions; there are many search tracks not single track, so GA has the excellent parallelism [7];

2) GA used the value information of objects without gradient information, so it can be used in optimization of large-scale, non-linear and discontinuous multimodal function and in optimization of object function without analytic solution.

3) The choosing-the-best system of GA is a "soft" selection and its excellent parallelism, such that GA has the excellent global optimization and robustness.

4) The objective function is explainable as the adaptive value of individual code, so it can lead to greater practicality, feasibility and operability.

IX. CONCLUSIONS

We use the GA to solve optimization problem without strong technique and a good knowledge of this subject, such as scheduling, routing, and layout and so on. The algorithm process of GA is actually simpler by coding the decision variable of problems, and gets the satisfactory solution quickly. Compared with traditional optimization methods, genetic algorithm has two notable characters, one is the latent parallelism and the other is searching in the whole area. And because of its independence, global optimization and implicit parallelism, GA is developed and applied by more and more people, so has been widely used in many fields, such as adaptive control, combinatorial optimization, pattern recognition, machine learning, artificial life and management decisions,etc. Although the genetic algorithm is a rising subject, corresponding theory and methods are not mature, it

can provides a hope to solve many other complicated problems.

REFERENCES

[1] WANG Ling, Intelligent Optimization Algorithm and its Application. Tsinghua University Press. 2001,36

[2] SHI Guangyan, QIAN Weiyi, Pang Liping, Optimization Methods. Higher Education Press.2010,136

[3] ZHOU Ming, SUN Shudong, Genetic Algorithms: Theory and Applications. National Defense Industry Press.2005,18-19

[4] CAIZixing, Xu Guangyou, Artificial Intelligence: Principles&Applications. Tsinghua University Press. 2010, 161

[5] Xing Wenxun,Xie Jinxing, Modern Optimization Algorithms. Tsinghua university Press. 2005,115

[6] WANG Wansen, Artificial Intelligence.Posts&Telecom Press. 2011, 129-132.

[7] ZHANG Wenxiu,LIANG Yi,Mathematical Foundation of Genetic Algorithms. Xi'an Jiao Tong University Press.2003,2