Runtime Analysis of OneMax Problem in Genetic Algorithm

Yifei Du
Graduate School of Engineering, University of Miyazaki, Gakuen Kibanadai-nishi 1-1
Miyazaki City, Miyazaki Prefecture, 889-2192, Japan

QinLian Ma
Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, Gakuen Kibanadai-nishi 1-1
Miyazaki City, Miyazaki Prefecture, 889-2192, Japan

Makoto Sakamoto and Hiroshi Furutani
Faculty of Engineering, University of Miyazaki, Gakuen Kibanadai-nishi 1-1
Miyazaki City, Miyazaki Prefecture, 889-2192, Japan

Yu-an Zhang
Department of Computer Science and Technology, Qinghai University, No. 251, Ning Road
Xining City, Qinghai Province, 810016, China

E-mail: furutani@cs.miyazaki-u.ac.jp

Abstract

Genetic algorithms (GAs) are stochastic optimization techniques, and we have studied the effects of stochastic fluctuation in the process of GA evolution. A mathematical study was carried out for GA on OneMax function within the framework of Markov chain model. We obtained the steady state solution, which represents a distribution of the first order schema frequency. We treated the task of estimating convergence time of the Markov chain for OneMax problem, and studied the effects of mutation probability and string length on the convergence time.

Keywords: genetic algorithms, schema theory, OneMax problem, Markov model, convergence time

1. Introduction

We have studied the probability that a population includes the optimum solution by applying Markov chain model[1]. We call this probability as the success probability of GA. We obtained the analytical form of the transition matrix P of Markov chain[2], which represents the evolution of the population in OneMax problem. We also found that GA for the OneMax problem is equivalent to the asymmetric mutation model[3]. There are rich studies of the asymmetric mutation model, and we can apply these results to the OneMax problem in GA. The relation between the convergence time and success probability was studied analytically and experimentally. In these analyses, we applied the Wright-Fisher model[4].

In this paper, first, we considered the effects of crossover on the evolution speed of OneMax problem.

This task was carried out by using linkage analysis. Next, we studied the convergence time of the Markov chain for OneMax problem. This analysis was performed by using the eigenvalues of transition matrix representing the behavior of population in the GA. From this analysis, we found that the convergence of GA to the stationary state can be represented approximately by mutation rate and string length. We used this theoretical result for the analysis of real GA calculations.

2. Mathematical Model

2.1. *OneMax model*

We treat the evolution process of a population with N individuals. The individuals are represented by binary strings of length l, and there are $n = 2^l$ genotypes,
$$i = <i(l), \dots, i(1)>, i(k) \in \{0,1\}.$$

We use the relative frequency $x_i(t)$ for describing the evolution

$$x_i(t) = N_i(t)/N,$$

where $N_i(t)$ is the number of individuals of genotype i at generation t. The relative frequencies satisfy the normalization condition

$$\sum_{i=0}^{n-1} x_i(t) = 1.$$

The average fitness $\bar{f}(t)$ of the population at generation t is

$$\bar{f}(t) = \sum_{i=0}^{n-1} f_i x_i(t). \qquad (1)$$

The OneMax fitness function f_i is defined as

$$f_i = \sum_{k=1}^{l} i(k). \qquad (2)$$

Thus the string of all ones $< 1,1,...,1 >$ is the optimum solution of this function.

2.2. *Linkage equilibrium*

We derive here the evolution equation for the first order schema frequency. To do this, we introduce the notion of linkage equilibrium. Linkage means the correlation between the different loci in a chromosome, and if there is some correlation we call this state as linkage disequilibrium[5]. Crossover and mutation have the effect of making the population in linkage equilibrium.

The distribution of a population in linkage equilibrium is given by

$$x_i(t) = \prod_{k=1}^{l} h_{i(k)}(t), \qquad (3)$$

where $h_{i(k)}(t)$ is a frequency of the first order schema at position k, and $i = < i(l),...,i(1) >$. We also use the notation of $h_0^{(k)}$ and $h_1^{(k)}$ for the first order schema frequencies of bit 0 and bit 1, respectively.

In the deterministic schema theory, the evolution of the first order schema in linkage equilibrium can be obtained[2]. The relative frequency of the first order schema at position k is determined by

$$h_1^{(k)}(t+1) = a h_1^{(k)}(t) + b, \qquad (4)$$

where constants a and b are

$$a = \left(1 - \frac{1}{l}\right)(1 - 2p_m), b = \frac{1}{l}(1 - 2p_m) + p_m.$$

3. Analysis of Evolution Rate

3.1. *Fisher's theorem*

According to Fisher's "Fundamental Theorem of Natural Selection"[5], evolution rate becomes faster if variance of fitness becomes larger. The change $\Delta\bar{f}(t)$ of fitness $\bar{f}(t)$ at generation t is

$$\Delta\bar{f}(t) = \bar{f}(t+1) - \bar{f}(t). \qquad (5)$$

Variance of fitness $VAR(f)$ is

$$VAR(f) = \sum_i f_i^2 x_i(t) - \bar{f}(t)^2. \qquad (6)$$

From eqs. (5) and (6), we have

$$\Delta\bar{f}(t) = \frac{1}{\bar{f}(t)} VAR(f).$$

We call $v(t)$ as evolution rate

$$v(t) \equiv \Delta\bar{f}(t) = \bar{f}(t+1) - \bar{f}(t).$$

If we ignore mutation, evolution rate $v(t)$ is proportional to the variance of fitness. We note that variance of fitness must be increased in order to promote the evolution.

3.2. *Experiments*

We examined the effect of crossover by numerical calculation for OneMax problem. The length of string is $l = 20$, population size $N = 100$ and mutation rate $p_m = 0.02$. We applied roulette selection. We averaged
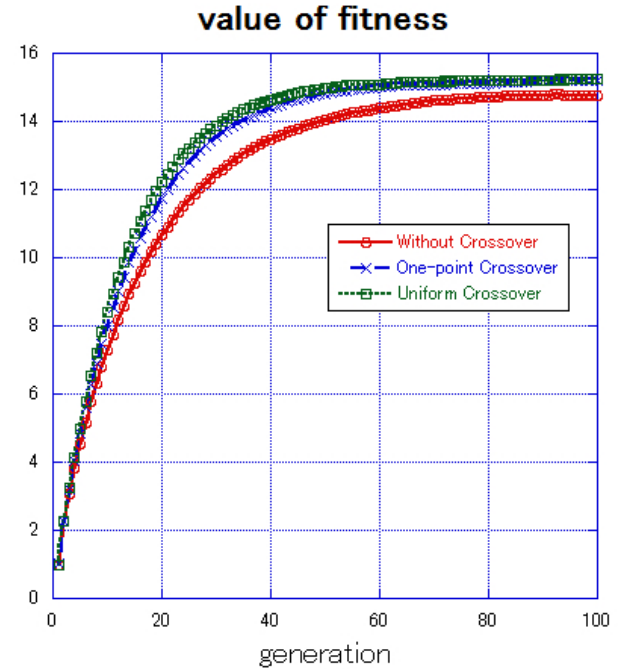

Fig. 1. Average fitness

the results obtained by repeating 1000 calculations. The initial state was randomly generated with $p^{(1)} = 1/l$. We compared three types of calculations: without crossover, one-point crossover and uniform crossover.

Figures 1 and 2 show values of average and variance of fitness, respectively.
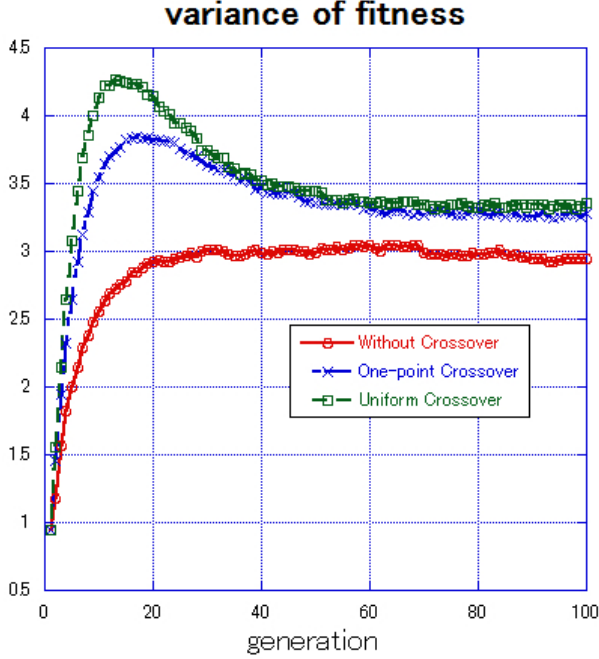

Fig. 2. Variance of fitness

We find that Fisher's theorem explains the results of Figs. 1 and 2.

3.3. *Linkage analysis*

The linkage disequilibrium coefficient is

$$D[m,m'] = p^{(11)}_{m,m'} - p^{(1)}_m p^{(1)}_{m'}. \qquad (7)$$

Here, $p^{(1)}_m$ is the probability that the m-th bit is 1, and $p^{(11)}_{m,m'}$ is the probability that the m-th bit and the m'-th bit are both 1. Using Walsh function and Walsh conversion, the variance is given by

$$VAR(f) = V_A + V_I,$$

where

$$V_A = \sum_{m=0}^{L} p^{(1)}_m (1 - p^{(1)}_m),$$

$$V_I = 2 \sum_{m<m'} D[m,m'].$$

Since V_A depends only on $p^{(1)}_m$, which is the function

of one bit, it is not affected directly by crossover. On the other hand, V_I dependents on the correlation of two loci. Therefore, the value varies by crossover. If correlation of two loci is small, the absolute value of V_I is small. In addition, because

$$V_I \leq 0,$$

variance increases with weaker correlation.

3.4. *Analysis of variance*

We used the same conditions as the experiment of section 3.2, and analyzed the variance of Fig.2 by linkage analysis.

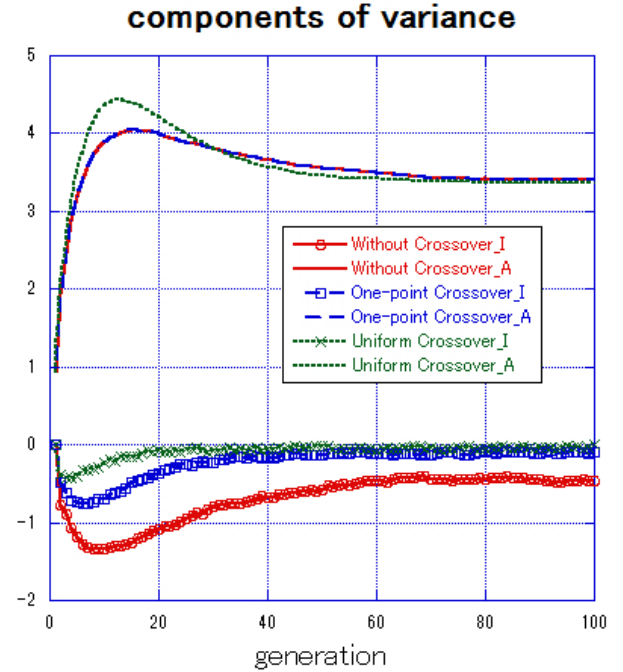Fig. 3 shows the change of V_A and V_I with generation.

Fig. 3. Change of V_A and V_I with generation.

From this calculation, we found that the variance of fitness is influenced by crossover. Crossover breaks the correlation between two loci, and a population will be in linkage equilibrium. In the case of uniform crossover, we found that the variance of fitness is the largest, and evolution rate is the fastest.

4. Analysis of Convergence

4.1. *Markov model*

In this study, we used the concept of Markov chain. Markov chain is a stochastic process, in which the transition only depend on the previous state.

One of the Markov model in population genetics is a

Wright-Fisher model[5,6]. Wright and Fisher put forward their evolution equation by considering the finite individuals. The Wright-Fisher model treats chromosomes having one locus and two alleles, corresponding to the GA of $l = 1$ with genotypes $k \in \{1,0\}^3$. The number of the first genotype 1 takes the values of

$$N_1 = 0,1,\dots,N,$$

and that of the genotype 0 is given by $N_0 = N - N_1$.

We analyze evolution processes by taking into account the effect of random sampling, and consider the fitness proportionate selection. If there are $N_1 = i$ copies of the genotype 1 at the current generation t, the probability $P(j|i)$ of N_1 taking the value of j at the next generation $t + 1$ is given by the binomial distribution

$$P_{i,j} = P(j|i) = \binom{N}{j} p_i^j (1 - p_i)^{N-j} \tag{8}$$

$$p_i = ay + b = a\left(\frac{i}{N}\right) + b.$$

The probability $P(j|i)$ specifies the process of random sampling, and the future behavior of the process only depends on its current frequencies[3]. Thus this process is a Markov chain. Let $\mu_i(t)$ be the probability that the population is in the state of $N_1 = i$ at generation t.

In the following, we use the vector notation

$$\mu^T = (\mu_0, \mu_1, \dots, \mu_N),$$

with the normalization condition

$$\sum_{i=0}^{N} \mu_i(t) = 1. \tag{9}$$

Then the evolution process is described by

$$\mu_j(t + 1) = \sum_{i=0}^{N} \mu_i(t) P_{i,j}. \tag{10}$$

The evolution equation is given in the vector form

$$\mu(t + 1)^T = \mu(t)^T P, \tag{11}$$

where P is a matrix of the size $(N + 1) \times (N + 1)$. We know the eigenvalues of this matrix[5,7],

$$\lambda_0 = 1, \lambda_1 = a, \lambda_2 = a^2(1 - 1/N), \dots \tag{12}$$

It should be noted that the second largest eigenvalue λ_1 does not depend on the population size N. We denote the left and right eigenvectors

$$\mathbf{u}_i^T P = \lambda_i \mathbf{u}_i^T, \quad P\mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad (0 \le i \le N).$$

These eigenvectors satisfy the orthogonally condition

$$\mathbf{u}_i^T \cdot \mathbf{v}_j = 0 \quad (i \neq j). \tag{13}$$

The explicit form of the right eigenvector $\mathbf{v}_0$ is given by

$$\mathbf{v}_0 = (1,1,\dots,1)^T, \tag{14}$$

and all elements of the left eigenvector $\mathbf{u}_0$ are positive. We adopt the normalization conditions

$$\sum_{i=0}^{N} |u_i| = 1, \quad \mathbf{u}_i^T \cdot \mathbf{v}_i = 1.$$

We consider the GA under positive mutation rate $p_m > 0$. In this case, all elements of the transition matrix P are positive, and the Markov chain of schema evolution is irreducible and aperiodic. The Markov chain theory states that an irreducible and aperiodic Markov chain converges to the stationary distribution π

$$\lim_{t \to \infty} \mu(t) = \pi, \tag{15}$$

and all elements π_i are positive[5].

The initial distribution of the first order schema can be expanded in terms of the left eigenvectors

$$\mu(t = 0)^T = \sum_{i=0}^{N} C_i \mathbf{u}_i^T. \tag{16}$$

Since $\mathbf{u}_i$ is eigenvector, and from eq. (11), we have

$$\mu(t)^T = \sum_{i=0}^{N} \lambda_i^t C_i \mathbf{u}_i^T. \tag{17}$$

Multiplying the right eigenvector $\mathbf{v}_0$ from right side

$$\mu(t)^T \cdot \mathbf{v}_0 = \sum_{i=0}^{N} \lambda_i^t C_i \mathbf{u}_i^T \cdot \mathbf{v}_0 = \lambda_0^t C_0 = C_0,$$

and from eq. (9) and (14), we have $C_0 = 1$. At large t, $\mu(t)$ is approximately given by

$$\mu(t) \approx \mathbf{u}_0 + a^t C_1 \mathbf{u}_1, \tag{18}$$

where $\mathbf{u}_0$ is the stationary distribution π.

4.2. *Total variation distance*

An important goal of the Markov chain theory is to estimate the rate of convergence to the stationary state. To this aim, the total variation distance has been used in many researches[8]. The total variation distance between the stationary distribution and the first order schema at generation t is defined as

$$TV(t) = \frac{1}{2} \sum_{i=0}^{N} |\mu_i(t) - \pi_i|. \tag{19}$$

Substituting eq. (18) into this equation, we have

$$TV(t) = C\, a^t, \qquad (20)$$

at large t, and C is a constant. It should be noted that the convergence behavior is determined by only one parameter a.

4.3. *Experiments*

We compared results of the theoretical prediction with GA experiments. Crossover is uniform crossover with crossover rate $= 1$. Mutation rate is p_m for each bit. Selection is roulette selection. We averaged the results obtained by repeating 10000 calculations. The initial state was randomly generated with $p^{(1)} = 1/2$.
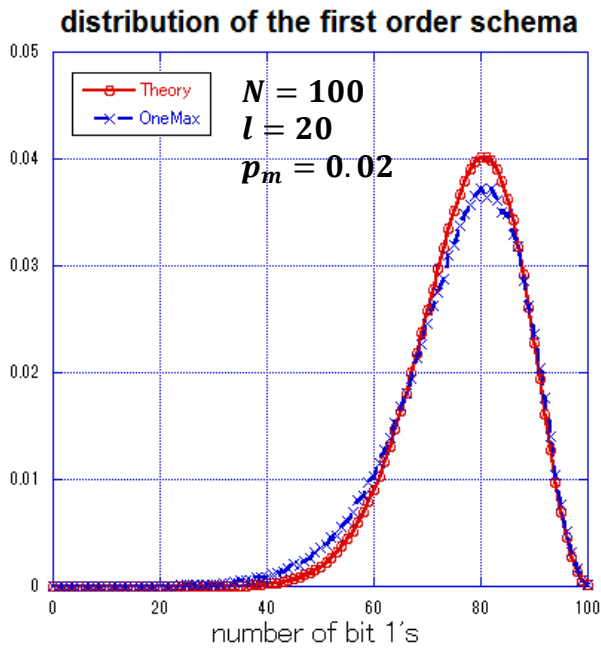

Fig. 4. Frequencies of the first order schema

Figure 4 is frequencies of the first order schema calculated by the theoretical prediction and GA experiments. The line of "Theory" is the results of theoretical calculation obtained by using eqs. (10) and (15). The line of "OneMax" is the stationary distribution of the first order schema by GA experiments.

Figures 5, 6, 7 and 8 are the total variation distance TV between the stationary distribution and the first order schema. Experimental results were obtained by the moving average with the window size of 21. We fitted the experimental data by matching the theoretical TV with approximation curve of eq.(20) at arbitrary matching time tm.
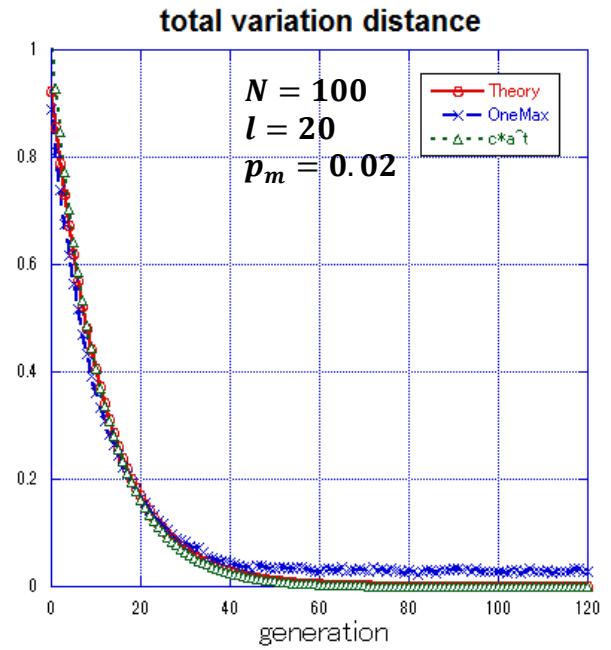

Fig. 5. Convergence of the first order schema

Figure 5 shows the results with N=100, l=20 and $p_m = 0.02$. We used the matching time tm=10. The dotted line is the approximate curve obtained by eq.(20). We found that the convergence behavior of the GA calculation is well reproduced by eq. (20).
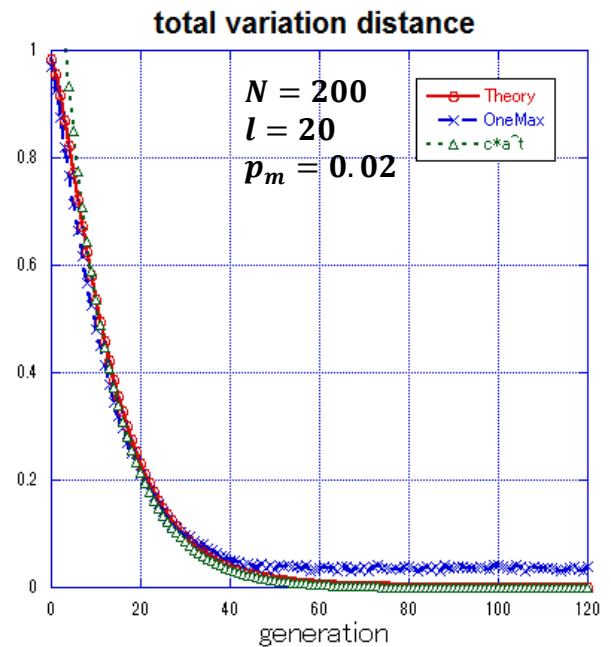

Fig. 6. Convergence of the first order schema

Figure 6 is the results when the population size N is changed to 200 from the experiment of Figure 5. The matching time is 10.
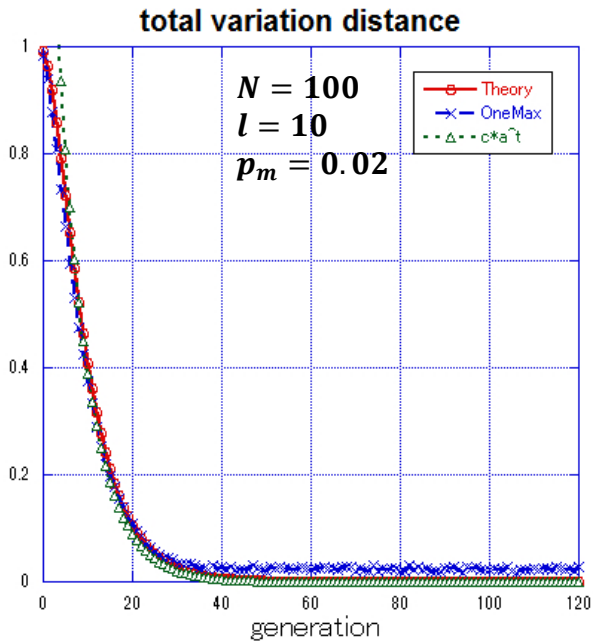
Fig. 7. Convergence of the first order schema

Figure 7 shows the results when the length of string l is changed to 10 from the experiment of Figure 5, and *tm*=8.
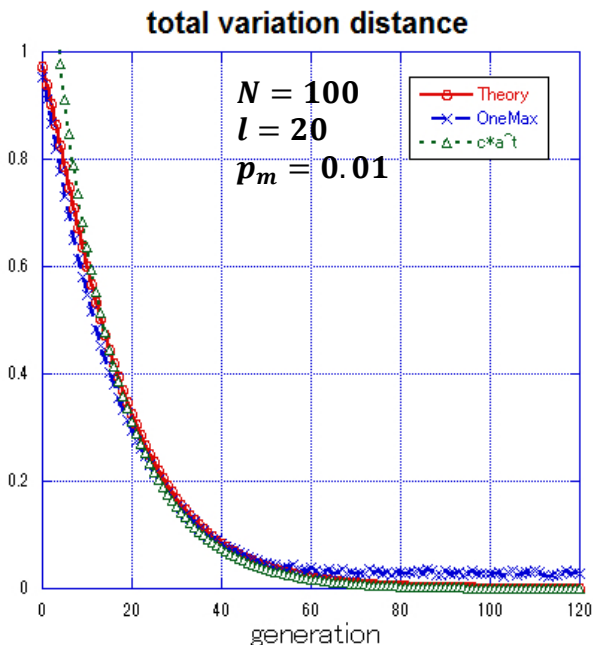

Fig. 8. Convergence of the first order schema

Figure 8 is the results when the mutation rate p_m is changed to 0.01 from the experiment of Figure 5. The mating time *tm*=15.

5. Summary

We studied the speed of GA calculation using the fitness function of OneMax. First, we showed that the speed of increase in average fitness function can be explained by separating the variance of fitness into two parts. Second, we studied the GA process to converge to stationary state. We obtained the approximate expression for predicting the convergence properties of Markov chain in terms of one parameter *a*.

References

1. H. Furutani, Y. Zhang and M. Sakamoto (2009), Study of the Distribution of Optimum Solution in Genetic Algorithm by Markov Chains, IPSJ, Transactions on Mathematical Modeling and its Applications, 2:54-63.

2. H. Furutani (2003), Schema Analysis of OneMax Problem–Evolution Equation for First Order Schemata. in Foundations of Genetic Algorithms 7, Morgan Kaufmann, San Francisco, 9–26.

3. Q. Ma, Y. Zhang, K. Koga, K. Yamamori, M. Sakamoto and H. Furutani (2013), Stochastic analysis of OneMax problem by using Markov chain, Artificial Life and Robotics 17: 395-399.

4. H. Furutani, S. Katayama, M. Sakamoto, M. Ito (2007), Stochastic Analysis of Schema Distribution in a Multiplicative Landscape, Artificial Life and Robotics 11:101–104.

5. J. W. J. Ewens (2004), *Mathematical Population Genetics. I. Theoretical Introduction*, Second Edition. Springer-Verlag, New York.

6. J. F. Crow and M. Kimura (1970), *An Introduction to Population Genetics Theory*, Harper and Row, New York.

7. W. Y. Tan (2002), *Stochastic Models with Applications to Genetics, Cancers, AIDS and Other Biomedical Systems*, World Scientific, Singapore.

8. D. A. Levin, Y. Peres and E. L. Wilmer (2008), *Markov Chains and Mixing Times*, American Mathematical Society, Providence.