

Diffusion model analysis of OneMax problem

QingLian Ma¹, Kiminobu Koga², Kunihiro Yamamori³, Makoto Sakamoto³, Hiroshi Furutani³, and Yu-an Zhang⁴

¹ Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, Japan

² Graduate School of Engineering, University of Miyazaki, Japan

³ Faculty of Engineering, University of Miyazaki, Japan

⁴ Department of Computer Science and Technology, Qinghai University, China

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

Abstract: A mathematical study is carried out for Genetic Algorithm (GA) on OneMax function within the framework of diffusion model. By using a partial differential equation, we obtain a distribution of the first order schema frequency. We consider the probability that a population includes the optimum solution by applying Markov chain model. We call this probability as the success probability of GA. Effects of mutation on the success probability were studied analytically and experimentally.

Keywords: genetic algorithms, schema theory, OneMax problem, Wright-Fisher model, diffusion model

1 INTRODUCTION

Genetic algorithms are stochastic optimization techniques that simulate the biological evolution. We have studied the effects of stochastic fluctuation in the process of GA evolution. In the analysis, we applied the Wright-Fisher model and the diffusion model. In the diffusion model, we derived an analytical expression of the distribution of the first order schema frequency.

In the previous work, we showed that GA on the OneMax problem is equivalent to the asymmetric mutation, and proposed the method to predict the convergence time of OneMax GA [1, 2]. This paper treats the method for calculating the probability that a population includes the optimum solution, which we call the success probability.

In general, mutation plays an important role in computation of GAs. This paper reports the comparison of numerical experiments and the theoretical calculations. We found that mutation is a very important factor determining the work of stochastic fluctuation.

2 MATHEMATICAL 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 $n = 2^\ell$ genotypes,

$$i = \langle i(\ell), \dots, i(1) \rangle, \quad i(k) \in \{0, 1\}.$$

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 process of fitness proportionate selection is

$$x_i(t+1) = f_i x_i(t) / \bar{f}(t), \quad (1)$$

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

$$\bar{f}(t) = \sum_{i=0}^{n-1} f_i x_i(t). \quad (2)$$

The OneMax fitness function f_i is defined as

$$f_i = \sum_{k=1}^{\ell} i(k). \quad (3)$$

Thus the string of all ones $\langle 1, 1, \dots, 1 \rangle$ is the optimum solution of this function.

2.1 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 [3]. Crossover and mutation have the effect of making the population in linkage equilibrium.

The distribution of a population in linkage equilibrium is given by

$$x_i(t) = \prod_{k=1}^{\ell} h_{i(k)}(t), \quad (4)$$

where $h_{i(k)}(t)$ is a frequency of the first order schema at position k , and $i = \langle i(\ell), \dots, i(1) \rangle$. 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.

2.2 Evolution Equation of the First Order Schema

In the deterministic schema theory, the evolution of the first order schema in linkage equilibrium is given by [4, 5]. The relative frequency of the first order schema at position k is determined by

$$h_1^{(k)}(t+1) = ah_1^{(k)}(t) + b, \quad (5)$$

where constants a and b are

$$a = \left(1 - \frac{1}{\ell}\right)(1 - 2\mu), \quad b = \frac{1}{\ell}(1 - 2\mu) + \mu.$$

The solution is given in terms of a ,

$$b_0 = 1 - \frac{\mu}{2\mu + (1 - 2\mu)/\ell},$$

and the initial value $h_1(0)$

$$h_1(t) = a^t \{h_1(0) - b_0\} + b_0. \quad (6)$$

Since $a < 1$, this solution converges to

$$t \rightarrow \infty \quad h_1(t) \rightarrow b_0.$$

3 MARKOV MODEL

3.1 Wright-Fisher Model

The stochastic model like Markov model explicitly treats the number of schemata. We consider the frequencies of two alleles at some locus. Two alleles are denoted by A and a , and the number of individuals having allele A and a are N_0 and N_1 , respectively. Since $N = N_0 + N_1$ is constant, we consider N_1 in this analysis.

The probability $Q(j|i)$ that $N_1 = i$ at generation t becomes $N_1 = j$ at generation $t + 1$ is given by the binomial distribution.

The probability $q_i(t)$ of $N_1 = i$ at generation t obeys the evolution equation

$$q_j(t+1) = \sum_{i=0}^N Q(j|i) q_i(t). \quad (7)$$

This equation is represented by using $(N + 1)$ -dimensional vector $\mathbf{q}(t)$

$$\mathbf{q}(t) = (q_0(t), \dots, q_N(t))^T,$$

where T is transposition, and the matrix $Q = [Q(j|i)]$

$$\mathbf{q}(t+1) = Q \mathbf{q}(t), \quad (8)$$

This model is called as the Wright-Fisher model[3].

Next, we derive the transition matrix on the OneMax fitness. Replacing

$$h_1(t) \rightarrow i/N$$

in the evolution equation of the first order schema (5), we have the transition matrix

$$Q(j|i) = \binom{N}{j} u(y)^j \{1 - u(y)\}^{N-j}, \quad (9)$$

$$u(y) = ay + b = a\frac{i}{N} + b,$$

where $y = i/N$.

3.2 Diffusion Model

In genetics, the approximation of Markov model by the diffusion equation is used frequently [3, 6], and we have derived the diffusion approximation for OneMax problem [4].

Using two functions of y

$$M(y) = (a - 1)y + b, \quad V(y) = u(y)\{1 - u(y)\}/N,$$

the diffusion approximation of Markov process in OneMax problem is given by

$$\frac{\partial \phi(y, t)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial y^2} \{V(y)\phi(y, t)\} - \frac{\partial}{\partial y} \{M(y)\phi(y, t)\}. \quad (10)$$

The function $\phi(y, t)$ stands for the probability density function of the relative frequency y at time t .

We derive the solution of Kolmogorov forward equation (10). Approaching $t \rightarrow \infty$, $\phi(y, t)$ converges to the stationary solution $\psi(y)$. At this, $\partial \phi(y, t)/\partial t = 0$, and $\psi(y)$ satisfies

$$\frac{d^2}{dy^2} \{V(y)\psi(y)\} - 2 \frac{d}{dy} \{M(y)\psi(y)\} = 0.$$

The solution is

$$\psi(y) = \frac{C}{V(y)} \exp \left\{ 2 \int^y \frac{M(x)}{V(x)} dx \right\}, \quad (11)$$

where C is the normalization constant.

$$\int_0^1 \psi(y) dy = 1.$$

By integrating this equation, we have

$$\psi(y) = C (ay + b)^{2c_1-1} (1 - ay - b)^{2c_2-1} \quad (12)$$

where C is a normalization constant, and c_1 and c_2 are

$$c_1 = Nb/a^2, \quad c_2 = N(1 - a - b)/a^2.$$

3.3 Success Probability

In condition of mutation existing, the GA becomes an ergodic Markov chain, and it drifts among all of placement states of population. Therefore, if there is enough time, it can get the transition state which contain optimum solution. We discuss the convergence between stationary distribution and stationary state by ergodic Markov chain. In other words, we

repeat the same calculation with different random numbers and examine distribution state of convergence.

Now we consider some partial bit string from l bit to $l - 1$ bite. The string in which all bits are 1, $\langle 1, 1, \dots, 1, 1 \rangle$, has $l - 1$ bit partial optimum solution.

In population, partial string of length is X_ℓ , the probability of having j units of optimum solution is expressed by

$$S_j^{(\ell)} = \Pr \{X_\ell = j\}. \quad (13)$$

We have defined success probability and failure probability before, it can be calculated

$$S = \sum_{j=1}^N S_j^{(L)}, \quad F = S_0^{(L)}. \quad (14)$$

For obtain S and F , it is necessary to solve $S_j^{(\ell)}$.

$S_j^{(\ell)}$ can be calculated by regression, the initial condition is

$$S_j^{(1)} = \pi_j^{(1)}, \quad (0 \leq j \leq N). \quad (15)$$

when the number of $l - 1$ bit partial optimum solution is i , in the string length X_ℓ , the probability of having j units of optimum solution is

$$Q_{i,j}^{(\ell)} = \Pr X_\ell = j | X_{\ell-1} = i. \quad (16)$$

Transition probability $Q_{i,j}^{(\ell)}$ use the distribution of first order schemata in the ℓ th bit,

1. if $j > i$

$$Q_{i,j}^{(\ell)} = 0, \quad (17)$$

2. if $j \leq i$

$$Q_{i,j}^{(\ell)} = \sum_{m=j}^{N-i+j} \binom{N}{m}^{-1} \binom{i}{j} \binom{N-i}{m-j} \pi_m^{(\ell)}. \quad (18)$$

The probability $S_j^{(\ell)}$ of having j units of ℓ - bit partial optimum solution can be calculated by this equation

$$S_j^{(\ell)} = \sum_{i=0}^N S_i^{(\ell-1)} Q_{i,j}^{(\ell)}. \quad (19)$$

4 NUMERICAL EXPERIMENT

In this section, we compare results of the theoretical prediction with GA experiments. We use the fitness proportionate selection and uniform crossover. The length of string is $L = 20$, and population size $N = 100$. We performed 10000 runs for each parameter set, and averaged over them. The success probability S of the GA is calculated by summing the number of run, where the optimum solution appears in the population, and deviding it by 10000.

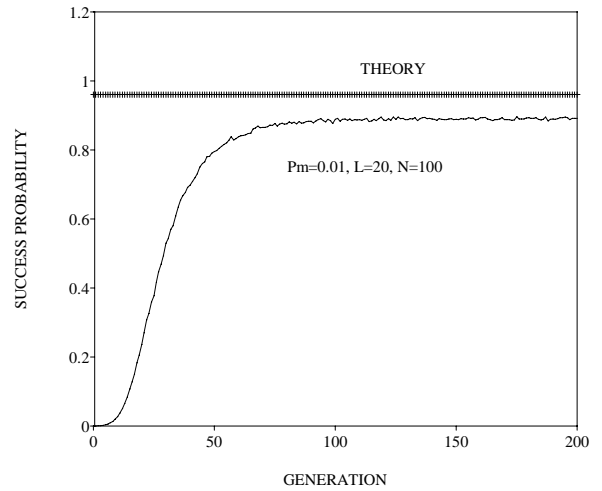


Fig. 1. Generation dependence of the success probability S with $L = 20$, $N = 100$ and $p_m = 0.01$. The thin line is the result of the GA calculation. The thick line is the theoretical prediction.

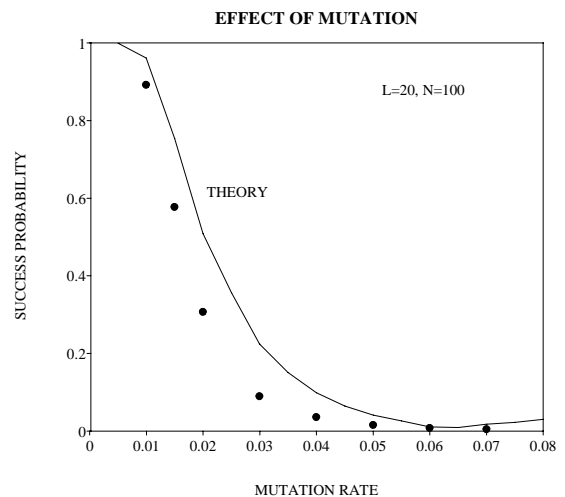


Fig. 2. Dependence of the success probability S on the mutation rate p_m . Calculation with $N = 100$ and $L = 20$. The solid line shows the theoretical prediction, and black circles are GA calculations.

Figure 1 shows the time dependence of the success probability S . The thin line is the result of GA calculation. The thick line is the theoretical prediction, which slightly overestimates the numerical calculation.

Figure 2 is the mutation rate p_m dependence of the success probability S . We observe that the theoretical prediction can reproduce the p_m dependence quite well. It should be noted that the success probability decreases sharply as p_m increases.

5 SUMMARY

In this paper, we demonstrate that the stochastic methods, Wright-Fisher model and its diffusion approximation, can reproduce the success probability S of GA on the OneMax function. Our next aim is to predict the first-appearing time T of the optimum solution in a population by applying the stochastic methods used in this study.

REFERENCES

- [1] Ma, Q., Zhang, Y., Koga, K., Yamamori, K., Sakamoto, M., and Furutani, H. (2012): Analysis of first order schema for OneMax problem, Proceedings of the 17th International Symposium on Artificial Life and Robotics
- [2] W. Y. Tan (2002), *Stochastic Models with Applications to Genetics, Cancers, AIDS and Other Biomedical Systems*, World Scientific, Singapore.
- [3] J. W. J. Ewens (2004), *Mathematical Population Genetics. I. Theoretical Introduction*, Second Edition. Springer-Verlag, New York.
- [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] 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
- [6] J. F. Crow and M. Kimura (1970), *An Introduction to Population Genetics Theory*, Harper and Row, New York.