A Robust Solution Searching Scheme in Genetic Search

Shigeyoshi Tsutsui[1], Ashish Ghosh[2] and Yoshiji Fujimoto[3]

[1] Department of Management and Information Science, Hannan University
5-4-33 Amamihigashi, Matsubara, Osaka 580 Japan
tsutsui@hannan-u.ac.jp

[2] Department of Industrial Engineering, College of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Sakai, Osaka 593 Japan

[3] Department of Applied Mathematics and Informatics
Faculty of Science and Technology, Ryukoku University
1-5 Yokoya, Seta Ooe, Ohtsu, Shiga 520-21 Japan

Abstract. Many of the studies on GAs give emphasis on finding the global optimal solution. In this paper, we propose a new method which extend the application of GAs to domains that require detection of *robust solutions*. If a global optimal solution found is on a sharp-pointed location, there may be cases where it is not good to use this solution. In nature, the phenotypic feature of an organism is determined from the genotypic code of genes in the chromosome. During this process, there may be some perturbations. Let $\boldsymbol{X}$ be the phenotypic parameter vector, $f(\boldsymbol{X})$ a fitness function and $\boldsymbol{\Delta}$ a noise vector. As can be easily understood from the analogy of nature, actual fitness function should be of the form $f(\boldsymbol{X}+\boldsymbol{\Delta})$. We use this analogy for the present work. Simulation results confirm the utility of this approach in finding robust solutions.

Keywords: Genetic Search, Robust Solutions, Adding noise

1 Introduction

Over the years, genetic algorithms (GAs) have proved useful in a variety of search and optimization problems [6]. There are many theoretical and empirical studies to improve the performance of GAs to solve difficult problems such as the multimodal and deceptive problems [5, 7, 10, 12, 13]. As a matter of fact, these researches put emphasis on finding the global optimal solution.

There are a number of investigations which emphasizes on finding multiple solutions (peaks) including local optima. These include the crowding method of De Jong [3], the sharing method of Goldberg et al. [8], the deterministic crowding method of Mahold [11], and the sequential niche method of Beasley et al. [2]. These approaches extend the application of GAs to domains that require the location of multiple solutions.

In this paper, we propose a new method which extend the application of GAs to domains that require the location of *robust solutions*. If the global optimal solution detected is a very sharp point, then there may be cases where it is not good to adopt this solution. This is especially true in areas such as setting the control parameters of aerospace control system, nuclear power control system, where a kind of robustness against the perturbation of the environmental features are important.

In nature, the phenotypic feature of an organism is determined from the genotypic code of genes in the chromosome. In this transformation, there may be some perturbations, for example, caused by an abnormal temperature, a nutritionally imbalance condition, existence of injurious matter etc. If in an organism these perturbed phenotypic features have low fitness, then the organism can not survive or produce offspring. Thus a species having a good genotypic material, if its phenotypic features become sensitive to perturbations, it would die out. On the other hand, the species which is robust against this perturbations would survive and evolve.

To emulate this sort of genetic feature, we can develop a genetic search scheme which can produce *robust solutions*. The proposed approach uses the effect of noise in the calculation of fitness values. Approaches on evolutionary computation which give consideration to existence of noise in calculating the fitness function are discussed in [3, 9]. Let $\boldsymbol{X} = (x_1, x_2, ..., x_m)$ be a phenotypic parameter vector, $f(\boldsymbol{X})$ the fitness function and $\boldsymbol{\Delta} = (\delta_1, \delta_2, ...,\delta_m)$ a noise vector. They used the form $f(\boldsymbol{X}) +\delta$, i.e. add noise δ with fitness value. As can be understood from the analogy of nature that noise is added during transformation from genotype to phenotype. Thus we can use an evaluation function of the form $f(\boldsymbol{X}+\boldsymbol{\Delta})$. By this approach, we can search for robust solutions. We call this approach *Robust Solution Searching Scheme* (RS^3) in genetic search.

In the following sections, first a model of the RS^3 in genetic search is outlined. Next, a mathematical model of RS^3 is described. Finally, empirical results are shown which confirm the effectiveness of the approach.

2 Description of the Robust Solution Searching Scheme (RS^3)

In this paper, we are primarily concerned with finding the robust solutions (against perturbations of $\boldsymbol{X}$) which maximize the function $f(\boldsymbol{X})$. Then the problem becomes as follows:

Find X which maximizes $f(X+\boldsymbol{\Delta})$.

Let $\boldsymbol{G}$ be a genotype string (or *chromosome*) which generates parameter $\boldsymbol{X}$. The model of the GA with RS^3 becomes as shown in Fig. 1. Here, it should be noted that adding noise in the form of $f(\boldsymbol{X}+\boldsymbol{\Delta})$ may look like a mutation operation on real valued coding, but it is completely different from mutation, since it does not affect the individual. The perturbations are used only for selection operation purpose.

```
gen = 0;
Pop(gen) = randomly initialized population (G^1, G^2, ..., G^N)
Transform each G^i to X^i;
Y^i = X^i + Δ;
Evaluate fitness of each Y^i in Pop(gen);
while( termination condition == false ){
        gen += 1;
        Select Pop(gen) from Pop(gen - 1) based on the fitness value f(Y^i);
        Apply genetic operators to G^1, G^2, ..., G^N in Pop(gen)
        Transform each G^i to X^i;
        Y^i = X^i + Δ;
        Evaluate fitness of each Y^i in Pop(gen);
}
Evaluate fitness of each X^i in Pop(gen)
```

Fig. 1 Model of the GA with RS^3

3 Mathematical Model

In this section we describe a simple mathematical model of RS^3. Although there are major differences between high- and low-dimensional problems, we consider $\boldsymbol{X}$ to be one dimensional to demonstrate the effect of adding noise more clearly and denote $\boldsymbol{X}$ by x. Extension of it to multi-dimension can be performed in a similar manner.

3.1 Effective Fitness Function in GA with RS^3

Let us consider the Schema Theorem of a GA using a proportional selection scheme and single-point crossover [6]:

$$M(H,t+1) \geq M(H,t) \cdot \frac{f(H,t)}{\overline{f(t)}}\left[1 - p_c \frac{d(H)}{L-1} - o(H)p_m\right], \tag{1}$$

where H represents a schema in population $P(t)$, $f(H,t)$ is the average fitness of representatives of H in $P(t)$, $\overline{f(t)}$ denotes the average fitness of the individuals in $P(t)$, p_c is the crossover rate, p_m is the mutation rate, $d(H)$ is the defining length of H, $o(H)$ is the order of H, L is the string length and $M(H,t)$ is the expected number of representatives of schema H at generation t. $\overline{f(t)}$, the average of the individuals in $P(t)$, is taken as

$$\overline{f(t)} = \sum_{i=1}^{N} f(x^i)/N\,, \tag{2}$$

where N is size of the population. If noise is added to the fitness function $f(x)$, it becomes $f(x+\delta)$; then

$$\overline{f(t)} = \sum_{i=1}^{N} f(x^i + \delta) / N \,. \tag{3}$$

The same logic holds true for calculating $f(H,t)$. If we assume the noise δ to be independent of time, the effective average fitness $\overline{F(t)}$, i. e., the average fitness $\overline{f(t)}$ over the distribution of δ, can be obtained as follows:

$$\begin{aligned}\overline{F(t)} &= \int_{-\infty}^{\infty} \sum_{i=1}^{N} f(x^i + \delta) \cdot q(\delta) d\delta / N \\ &= \sum_{i=1}^{N} \left[\int_{-\infty}^{\infty} f(x^i + \delta) \cdot q(\delta) d\delta \right] / N \,,\end{aligned} \tag{4}$$

where $q(\delta)$ is the density function of noise δ.

From Eqs. 2 and 4, we can conclude that the *effective fitness function* $F(x)$ can be formulated as

$$\begin{aligned}F(x) &= \int_{-\infty}^{\infty} f(x + \delta) \cdot q(\delta) d\delta \\ &= \int_{-\infty}^{\infty} f(y) \cdot q(y - x) dy \,.\end{aligned} \tag{5}$$

Thus, we can calculate the effective fitness function $F(x)$ from fitness function $f(x)$ and noise density function $q(\delta)$.

If we assume $q(\delta)$ to be symmetric, i. e., $q(\delta) = q(-\delta)$, then $F(x)$ can be written as

$$F(x) = \int_{-\infty}^{\infty} q(x - y) \cdot f(y) dy \,. \tag{6}$$

Eq. 6 has the form of convolution integral. Let $\hat{F}(\omega)$, $\hat{f}(\omega)$ & $\hat{q}(\omega)$ be Fourier transform of functions $F(x)$, $f(x)$ & $q(x)$, respectively. Then, we can get $\hat{F}(\omega)$ as

$$\hat{F}(\omega) = \hat{q}(\omega) \cdot \hat{f}(\omega) \tag{7}$$

It is very natural to assume the noise to be a Gaussian noise N(0, σ) as it is the sum of various perturbations. Then, $\hat{q}(\omega)$ is obtained as

$$\hat{q}(\omega) = \frac{1}{\sqrt{2\pi}} e^{\frac{-\sigma^2 \omega^2}{2}} \,. \tag{8}$$

Function $\hat{q}(\omega)$ decreases as ω increases. Thus, we can understand that adding Gaussian noise to phenotypic parameters acts as a low-pass filter as in a signal processing system. Also we can see that the effect of low-pass filtering is strengthened as the value of σ increases.

3.2 Calculation of the Effective Fitness Functions

For mathematical manipulation let us chose (for the sake of simplicity) a *peak* of fitness function by a rectangular function with height h and width $2w$ as follows:

$$f(x) = \begin{cases} h: & -w \leq x \leq w \\ 0: & \text{otherwise}. \end{cases} \tag{9}$$

Then, from Eq. 6 we can calculate effective fitness function of Eq. 9 as follows:

$$\begin{aligned} F(x) &= h\int_{-w}^{w} q(x-y)dy \\ &= h\left[\Phi\left(\frac{x+w}{\sigma}\right) - \Phi\left(\frac{x-w}{\sigma}\right)\right], \end{aligned} \tag{10}$$

where a Gaussian noise N(0, σ) is assumed and $\Phi(y)$ is the distribution function of the standard Gaussian noise defined by

$$\Phi(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} dy. \tag{11}$$

By setting the derivative of the function $F(x)$ to zero, the peak point is obtained at $x = 0$ and the peak value $\max F(x)$ is obtained as

$$\begin{aligned} \max F(x) &= F(0) \\ &= h\left[2\Phi\left(\frac{w}{\sigma}\right) - 1\right]. \end{aligned} \tag{12}$$

Fig. 2 plots the $\max F(x)$ values vs. w/σ. Fig. 3 shows the relationship between the fitness function $f(x)$ and the effective fitness function $F(x)$ for w/σ = 4.0, 2.0, 1.0, 0.5 &

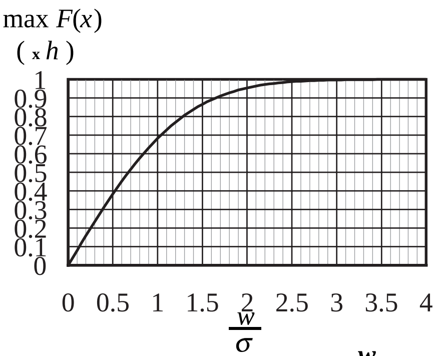


Fig. 2. max $F(x)$ vs. $\frac{w}{\sigma}$

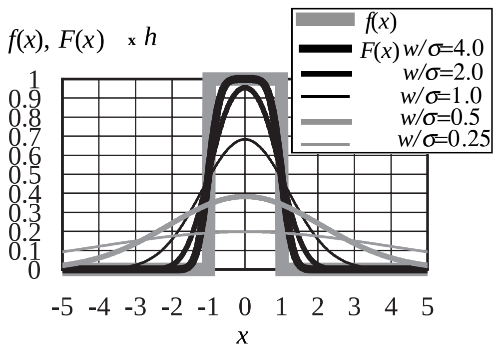


Fig. 3. Relationship between $f(x)$ and $F(x)$

0.25. From Figs. 2 and 3, we can confirm that adding Gaussian noise to phenotypic parameters acts as a low pass filter as discussed in Section 3.1. We can estimate σ value as follows. Let $2w_0$ be the width of a sharp peak with height h. If we take σ in the range $2w_0$ - $4w_0$, then w/σ has values in the range 0.5- 0.25 and the effective peak value would lie between 0.197h and 0.383h. Thus, we can roughly estimate σ value.

3.3 An Absorption Effect

Let us consider a fitness function $f(x)$ with one robust peak with width $2w_1$ and height h_1, one sharp peak with width $2w_2$ and height h_2 and a gap w_3 between them as follows:

$$f(x) = \begin{cases} h_1 : & -w_1 \le x \le w_1 \\ h_2 : & w_1 + w_3 \le x \le w_1 + 2w_2 + w_3 \\ 0 : & \text{otherwise}. \end{cases} \tag{13}$$

Here, we assume $w_1 >> w_2$ and $h_1 < h_2$. The effective fitness function $F(x)$ of $f(x)$ can be obtained as follows:

$$F(x) = h_1\left[\Phi\left(\frac{x+w_1}{\sigma}\right) - \Phi\left(\frac{x-w_1}{\sigma}\right)\right] + h_2\left[\Phi\left(\frac{x-w_1-w_3}{\sigma}\right) - \Phi\left(\frac{x-w_1-2w_2-w_3}{\sigma}\right)\right]. \tag{14}$$

Fig. 4 shows both $f(x)$ and $F(x)$ for w_1 = 1, w_2 = 0.1, w_3 = 0.5, h_1 = 1 and h_2 = 2. Two cases corresponding to σ =1.0 & 0.5, that is, w_1/σ = 1.0 (: w_2/σ = 0.1) and w_1/σ = 2 (: w_2/σ = 0.2) are shown. As seen from the figure, although the highest peak point of $F(x)$ is located in the robust peak area of $f(x)$, the point is a little biased towards the sharp peak of $f(x)$ especially when σ = 1.0. As is easily guessed, this tendency become prominent when the gap between the peaks (w_3) become smaller and σ becomes larger. This *absorption effect* is normally one problem to handle.

To cope with this, we need to limit the size of noise. Thus, we can decrease the mutual influence of peaks on effective fitness function. In this paper, we propose

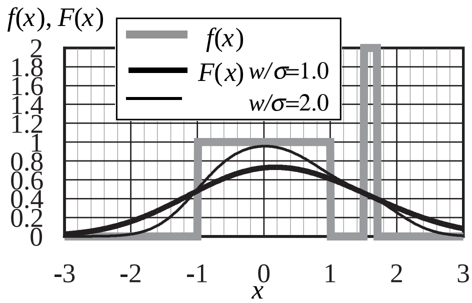


Fig. 4. The Absorption Effect

to use the following *modified Gaussian noise* N'(0, σ):

$$q(y) = \begin{cases} \dfrac{1}{C\sqrt{2\pi}\sigma} e^{\frac{-y^2}{2\sigma^2}} & : \quad -k\sigma \le y \le k\sigma \\ 0 & : \quad \text{otherwise}, \end{cases} \tag{15}$$

where k is a control parameter and C is a normalizing constant calculated as

$$C = 2\Phi(k) - 1. \tag{16}$$

In our experiment, we used k = 1.

Although the form N'(0, σ) is useful, there still remains a point to be considered. Let us consider a fitness function where the gap between a robust peak and a sharp peak (w_3 in this section) is very small. In this case, it may so occur that the peak of the effective fitness function is located in the gape. This problem can also be tackled if we are careful in using the RS^3. Due to the presence of noise, individuals in the converged population are distributed according to the effective fitness function (see Figs. 6 & 7). Thus, each individual is a candidate to be adopted. We should choose a solution from these candidates.

4 Empirical Study

To show the utility of RS^3, we studied two functions. We use a simple GA with stochastic universal sampling [1]. GA parameters were kept constant for all the simulations as p_m = 0.006, p_c = 0.6, population size N = 50, maximum number of trials = 3,000, and string length = 30 (Gray coded).

(1) Function f_a

f_a is a function (shown in Fig. 4) with two peaks in the range -3 ≤ x ≤ 3. We assume σ = 0.4 (w_2/σ = 0.25). Fig. 5 shows the mean value of parameter x in the population with trials. Simple GA converged to the highest sharp peak with mean value 1.6 (the center of the highest peak). GA with RS^3 converged to the lower robust peak. Fig. 6 shows a typical distribution of the individuals in the population after 3,000 trials for both simple GA and the proposed model.

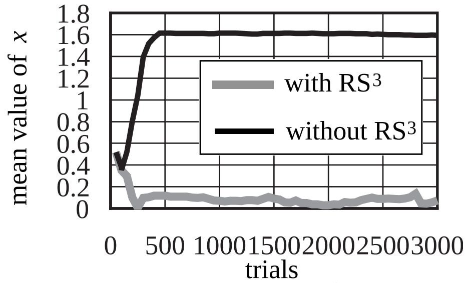


Fig. 5. Convergence Process for Fuction f_a

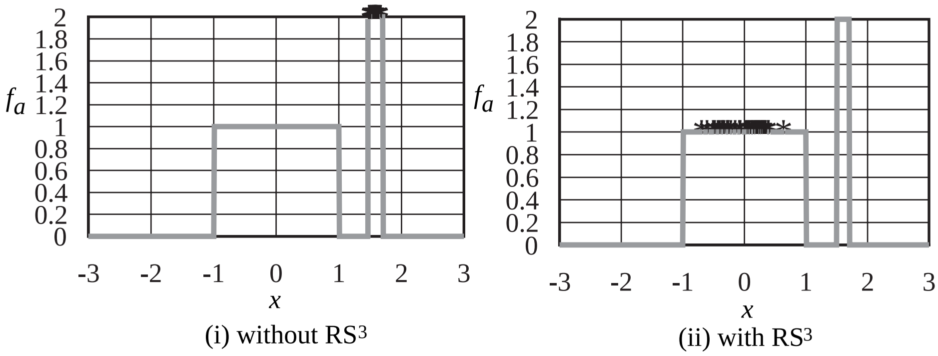


Fig. 6. State of the Population on Function f_a

(2) Function f_b

Function f_b has five unequal peaks in the range $0 \le x \le 1$ and is a variant of the function used in [4] defined as follows:

$$f_b = \begin{cases} e^{-2\ln 2\left(\frac{x-0.1}{0.8}\right)^2} \left|\sin(5\pi x)\right|^{0.5} & : 0.4 < x \le 0.6 \\ e^{-2\ln 2\left(\frac{x-0.1}{0.8}\right)^2} \sin^6(5\pi x) & : \text{otherwise}. \end{cases} \tag{17}$$

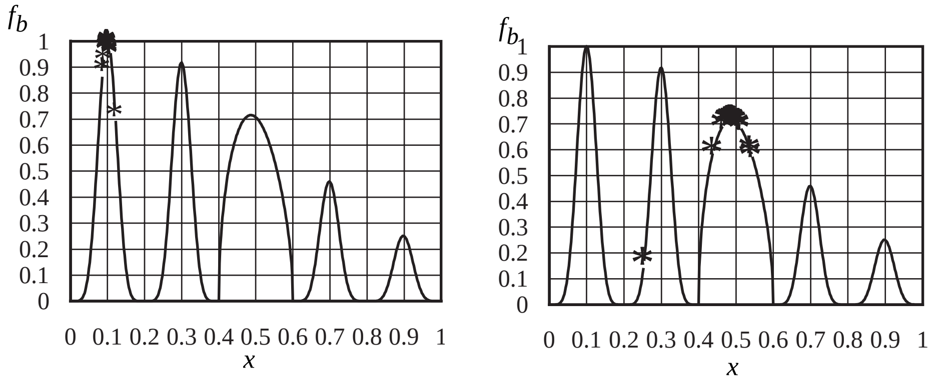
As shown in Fig. 7, global optimum is located at $x = 0.1$ with value 1.0. All peaks except the third one are sharp. The third peak is robust compared to the other four peaks and is located at $x = 0.486$ with peak value 0.715. The effective width $2w$ of the sharp peaks can be estimated as follows:

$$1 \times 2w \approx 1 \times \int_0^{0.2} \sin^6(5\pi x)\,dx = 1/16 \quad \Rightarrow \quad w \approx 1/32\,. \tag{18}$$

Referring to Fig. 2, w/σ can be chosen in (0.25, 0.5] so as to reduce the effective fitness value by 50%. We choose $w/\sigma = 0.5$.Thus, $\sigma = 1/16$ (: $\sigma = w/0.5 = 2 \times 1/32$) is used.

Fig. 7 shows a typical distribution of the individuals in the population after 3,000 trials for both simple GA and the proposed model. From the figures, we can say that in the GA with RS^3 the population converges to the robust peak. Fig. 8 shows a convergence process of the mean values of parameter x in the population. GA without RS^3 converges to the highest peak at 0.1, the center of the highest peak. GA with RS^3 converges to the lower robust peak zone. It can be observed from Fig. 8 that at the first stage of searching, the possible solutions move towards the zone where the highest peak exists. Then they shift towards the robust peak. This phenomenon may be explained as follows. In the early stage of evolution, the effect of noise is low since the diversity of the population is large. As the search proceeds, existence of noise gradually affects and moves the solutions.

Finally, from the studies on both the function f_a and f_b we notice that there is no effect of absorption on the results. This may be due to the fact that we added noise with



(i) without RS^3 (ii) with RS^3

Fig. 7. State of the Population on Function f_b

modified Gaussian distribution and appropriate value of σ.

5 Conclusions

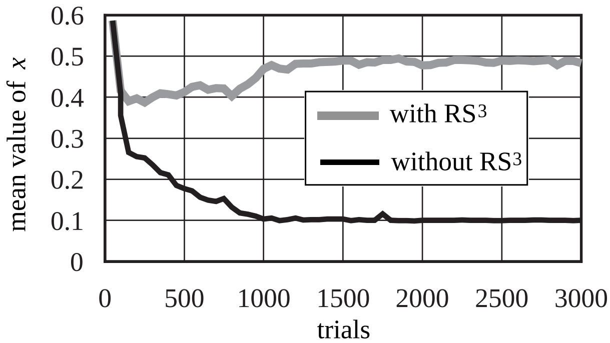


Fig. 8. Convergence Process on Fuction f_b

The *robust solution searching scheme* (RS^3), a new method which extends the application of GAs to domains that require detection of *robust solutions*, is proposed in the present article. In nature, the phenotypic feature of an organism is determined from the genotypic code of genes in the chromosome. In this process, there may be some perturbations. Even with these perturbation, strong organisms can survive. To emulate this sort of genetic feature, we added noise while evaluating the fitness of a function.

We developed a mathematical model of RS^3 and analyzed using simple functions. This analysis provides a guideline to determine the amount of noise to be added. The effectiveness of the RS^3 is demonstrated by maximizing functions having robust and sharp peaks.

Future work will focus on the following items:

(1) Combining the RS^3 with niche-formation methods, such as a sharing method, remains to be an interesting topic to be studied for more fruitful search.

(2) To analyze the influence of RS^3 on the convergence properties of high-dimensional and more complex problems.

(3) To extend the RS^3 for order based representational problems such as a scheduling system.

Acknowledgments

The authors gratefully acknowledge Prof. David E. Goldberg for his useful comments on this work. This research is partially supported by the Ministry of Education, Science, Sports and Culture under Grant-in-Aid for Scientific Research on Priority Areas number 264-08233105.

References

1. Baker, J. E.: Reducing Bias and Inefficiency in the Selection Algorithm, *Proc. of the Fifth ICGA*, pp. 14-21 (1987).
2. Beasley, D., Bull, D. R. and Martin, R. R.: A Sequential Niche Technique for Multimodal Function Optimization, *Evolutionary Computation*, MIT Press, Vol. 1, No. 2, pp. 101-125 (1993).
3. De Jong, K. A.: An Analysis of the Behavior of a Class of Genetic Adaptive Systems (Doctoral Dissertation, University of Michigan). *Dissertation Abstract International* 36(10), 5140B. (University Microfilms No. 76-9381)
4. Deb, K. and Goldberg, D. E.: An Investigation of Niche and Species Formation in Genetic Function Optimization, *Proc. of the Third ICGA*, pp. 42-50 (1989).
5. Eshelman, L. J.: The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination, *Foundations of Genetic Algorithms*, edited by Rawlins, G.J.E., Morgan Kaufmann, pp.265-283 (1991).
6. Goldberg, D. E.: *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Publishing Company, Inc. (1989).
7. Goldberg, D. E., Deb, K. and Korb, B.: Messy Genetic Algorithms Revisited: Studies in Mixed Size and Scale, *Complex Systems*, vol. 4, pp. 415-444 (1990).
8. Goldberg, D. E. and Richardson, J.: Genetic Algorithms with Sharing for Multimodal Function Optimization, *Proc. of the Second ICGA*, pp. 41-49 (1987).
9. Hammel, U. and Back, T.: Evolution Strategies on Noisy Function How to Improve Convergence Properties, *Proc. of Third Conf. on Parallel Problem Solving from Nature*, North-Holland, pp. 159-168 (1994).
10. Mathias, K and Whitley, L. D.: Changing Representations During Search: A Comparative Study of Delta Coding, *Evolutionary Computation*, MIT Press, Vol. 2, No. 3, pp. 249-278 (1994).
11. Mahfoud, S. W.: Crowding and preselection revisited, *Proc. of Second Conf. on Parallel Problem Solving from Nature*, North-Holland, pp. 27-34 (1992).
12. Tsutsui, S. and Fujimoto, Y.: Forking Genetic Algorithm with Blocking and Shrinking Modes, *Proc. of the Fifth ICGA*, pp. 206-213 (1993).
13. Tsutsui, S. and Fujimoto, Y.: Phenotypic Forking Genetic Algorithms (p-fGA), *Proc. of the 1995 IEEE ICEC*, pp. 566-572 (1995).