Diversity Adaptation in Genetic Algorithms with Preference Mating

Chaiwat JASSADAPAKORN
Department of Computer Engineering, Chulalongkorn University
Bangkok 10330, Thailand

and

Prabhas CHONGSTITVATANA
Department of Computer Engineering, Chulalongkorn University
Bangkok 10330, Thailand

ABSTRACT

The diversity of the population affects the convergence rate in genetic algorithms. The determination of the proper diversity is still a trial and error process. The objective of this work is to study a method to adapt suitable population diversity for a given problem. The proposed method is based on a modified restricted mating called "*preference mating*". Three well-known test problems, which have different requirement for diversity, are used to evaluate the proposed method.

Keywords: Genetic Algorithms, Population Diversity, Diversity Control, Restricted Mating, Adaptive System.

1. INTRODUCTION

An important issue in applying genetic algorithms (GAs) to solve problems is a phenomenon called "*premature convergence*" [1]. The most common cause of premature convergence is the lack of diversity coupled with ineffectiveness of the crossover operator to search for a new solution. The traditional GAs do not directly employ a method to maintain diversity during the evolutionary process. Without adequate diversity a few fitter individuals dominate the population in a short period of time. When the population diversity is lost, the evolutionary process cannot progress. It is because some necessary genetic materials, which may be the part of solution, are lost. To improve the performance of GAs, many works proposed enhanced strategies by embedding the diversity maintenance feature in different forms. The most well known strategy is the sharing approach [2].

The sharing method is the most frequently used technique which is inspired by natural ecosystem. Each individual is forced to share its fitness value to its neighbors. The survival probability of an individual depends on its fitness value and its difference from others in the neighborhood. This approach encourages the exploration of the new region in a solution space.

The ranked space method [3] is another strategy. This approach embeds the diversity maintaining mechanism explicitly by the use of two ranks in the selection process. The first rank orders individuals by their fitness values, called the quality rank. The second rank called the diversity rank, orders individuals by the difference between each individual and the previously selected individual. The combined rank of these two ranks is used to influence the selection probability. With this approach, the fitter individual is selected and at the same time the population diversity is maintained.

Another strategy is the approach called restricted mating. The restricted mating applies conditions such as restriction or encouragement, to select an individual and its mate partner. For example, the difference between pairs measured by the number of different bits (Hamming distance) is used in [4]. In that work, an individual and its partner are selected with traditional method. However, if the number of different bits of the pair does not pass the threshold, the second partner will be reselected until the condition is satisfied. The threshold is high for the early period and gradually decreases when no pairs of individual can be found to satisfy the condition.

In the work [5], a new selection scheme inspired by the animal mating behavior is proposed. This inspiration causes a new idea to apply dissimilar measurement to the pair of individual. The first mate is selected with the traditional scheme -- the higher fitness value, the more chance to be selected. The second mate is selected by consideration of another feature which can be depended on the first partner (this process is called seduction). Subsequently, in the work [6], the chance to be selected as the second partner is affected from the combination of the fitness value and the difference from the first partner with predefined weights of these two values. Since the mating procedure depends on the difference in each pair of individual, it can maintain the population diversity in an indirect way.

Many enhance strategies are proposed, unfortunately, all these works require the knowledge to setup parameters that affect the degree of population diversity in the evolutionary process such as the radius of neighborhood in niche method, conditions and threshold in restricted mating. Setting parameters incorrectly leads to unsuitable population diversity for the problem and causes poor performance.

The objective of this work is to study a method to adapt suitable population diversity for a problem without using the knowledge of problem's structure. A new mating strategy which extends the restricted mating is proposed, called "*preference mating*". This new strategy enables the system to adapt the suitable diversity for problems automatically. The experiment is carried out on three test problems: one-max, multimodal function, and deceptive function to evaluate the proposed method.

The organization of this paper is as follows. In the next section, the preference mating is explained. The diversity control is present in Section 3. The experiment is described in Section 4. In Section 5, the result and discussion are presented. The conclusion is given in Section 6.

2. PREFERENCE MATING

The preference mating is named by its property which is different from the traditional GAs. That is, the individuals have a preference for its partner depends on "*preference type*". The higher value of preference type indicates the preference to recombine with the individual that is different from it.

Given the first selected individual x_1 which is selected by a traditional selection method (tournament selection in the experiment), the preference type assigned to x_1 is used to calculate the chance of another individual to be selected as its partner. Let *d* represents the difference between the first selected individual and a candidate, τ represents the preference type, and *D* represents a function of d and τ, called the "*difference function*". The candidate who has a higher *D* value has more chance to be selected as the second partner. The selection criterion depends on the difference function and the fitness value (Eq. (1)).

$$x_2 = \underset{i \in s_t}{\operatorname{argmax}}[f(y_i) \cdot D(\tau, d_i)] \qquad \textbf{Eq. (1)}$$

where x_2 represents the selected partner, y_i is the i^{th} candidate which are randomly selected from the population, f is the fitness function, and s_t is the tournament size. In the experiment, a basic difference function is used (Eq. (2)).

$$D(\tau, d_i) = d_i^{\tau^2} \qquad \textbf{Eq. (2)}$$

where d_i is the difference between the first selected individual and the candidate y_i :

$$d_i = \frac{h(y_i, x_1)}{l} \qquad \textbf{Eq. (3)}$$

where *h* is the Hamming distance of two individuals, and *l* is the length of chromosome. An example of relationship between *D*, *d* and is shown in Fig. (1).

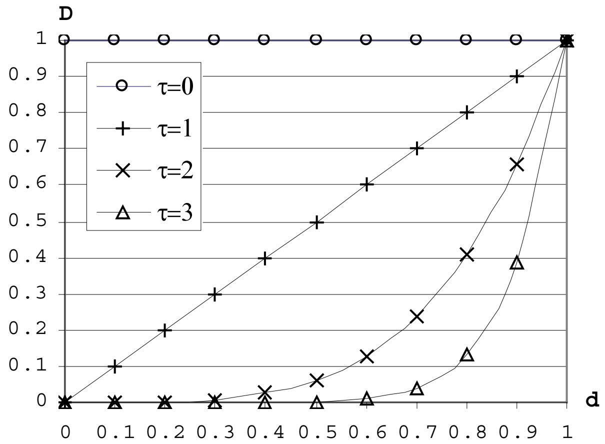

Fig. (1). An example of difference function (Eq. (2)).

Please note that when τ is 0, the probability of selection does not depend on *d*, which is equivalent to a traditional selection method, the chance to be selected depends only on the fitness value. The higher value of τ gives more weight to the difference between individuals, which influences the population towards more diversity.

3. DIVERSITY CONTROL

The preference mating strategy is used to make a diversity control system which has capability to adapt the suitable degree of diversity for a given problem. Since the degree of diversity in preference mating is controlled by the specification of preference type, the goal of this diversity control system is to search for the suitable preference type for a given problem. The idea is to use multiple preference type for the mating process. Each preference type is used equally first. The effectiveness of each preference type is evaluated, that causes the change of the frequency of use in the next iteration. This leads to the design of measurement called "*contribution*".

Contribution is the measurement which counts how good of each preference type in the term of how often they construct better individuals in the next generation population.

$$Contribution(\tau, t) = \frac{\#SuccCross(\tau, t)}{\#Cross(\tau, t)} \qquad \textbf{Eq. (4)}$$

As shown in Eq. (4), the contribution of each preference type τ at the generation t can be calculated by two terms which are the number of successful crossover (denoted by $\#SuccCross(\tau, t)$) and the number of crossover times (denoted by $\#Cross(\tau, t)$). The successful crossover is the one that produce at least one child which is better than both parent considered by the fitness value. This term is normalized by the total number of crossover of each preference type. The contribution is the measurement of the ratio of the better individual creation. The preference type with higher contribution means the higher effectiveness.

With the use of contribution measurement, the preference types are in direct competition against each other to be used. The preference type that performs well will be promoted. The preference type that is inferior will be demoted. The promoted and demoted mechanisms cause increasing and decreasing the chance to be used respectively. This scheme leads to the concentration of computational effort to the promising preference type which causes the adaptation of diversity for a given problem.

The process of the diversity control system can be summarized as follows:

1) Randomly generate the population of individual.
2) Evaluate each individual by fitness function.
3) Set the contribution equally for each preference type for the first time.
4) Select an individual and its partner with preference mating procedure. The probability of choosing a preference type is proportional to its contribution.
5) Reproduce two new individuals for the next generation by crossover.
6) Repeat step 4 and 5 for the whole population.
7) Evaluate each new individual by fitness function.
8) Compare the fitness value of the new individuals and their parental individuals. Calculate contribution of each preference type.
9) Repeat step 4-8 until reach the final generation.

4. EXPERIMENT

The proposed method is evaluated using three test functions. They are well-known test problems in GAs: one-max problem, deceptive function, and multimodal function. These functions are range from easy, moderately difficult to very difficult and they require different degree of diversity in the population to solve them efficiently. Each problem will be briefly explained.

One-Max Problem

The one-max problem is an example of an easy problem for GAs. The goal is to evolve an individual that all bits in chromosome are “1”. The fitness evaluation of this problem is straightforward. The number of bit “1” in the chromosome of each individual is assigned to be its fitness value. The 30-bit one-max is used in the experiment.

Deceptive Function

The deceptive function is a hard problem for GAs. The fitness function of this problem does not guide the evolutionary process to the correct direction. An order-3 deceptive function is used in the experiment. The fitness value of each 3-bit binary string is given as Eq. (5). where x denotes the 3-bit binary string and $|x|$ denotes the number of bit “1” in the binary string x. Ten blocks of the 3-bit string are used, the total length of a string is 30. The fitness value of each individual is the summation of the fitness value of each block. The optimum of this problem is all bits set to “1”. The highest fitness value is 10.

$$f(x) = \begin{cases} 0.9 & if \quad |x| = 0 \\ 0.8 & if \quad |x| = 1 \\ 0 & if \quad |x| = 2 \\ 1 & if \quad |x| = 3 \end{cases} \qquad \textbf{Eq. (5)}$$

Multimodal Function

This type of function comprised of many subobtimum peaks. It is difficult for an evolutionary process to find the global optimum point. The source of the multimodal function (Eq. (6)) used in the experiment comes from [7]. This function has 5 peaks.

$$f(x) = e^{-2(ln2)\left(\frac{x-0.08}{0.854}\right)^2} sin^6 [5\pi(x^{3/4} - 0.05)] \qquad \textbf{Eq. (6)}$$

where $0 \le x \le 1$. The plot of this function is shown in Fig. (2).

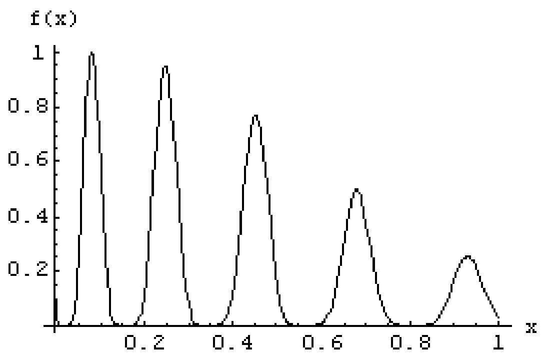

Fig. (2). The plot of the multimodal function.

In the experiment, the real value x between 0 and 1 is encoded with 30-bit chromosome. The goal is to evolve an individual which maximize the function f. The optimum point of this function is approximately x = 0.08.

The objective of the experiment is to investigate the two issues of the adaptive system which are the performance for solving problem and the adaptation behavior. The performance of the adaptive system is compared with the non-adaptive systems. The non-adaptive systems are the GAs using the preference mating with a predefined preference type. As noted before, the non-adaptive system with preference type 0 is equivalent to a traditional GA.

The parameters used in the experiment are shown in Table (1). The one-point crossover is used in the experiment. The mutation operation is excluded.

Table (1). The parameters used in the experiment.

Parameter	Value
Population size	400
Chromosome length	30 bits
Number of generation	200
Number of repeated run	500
Crossover probability (P_c)	100%
Tournament size	3
Number of preference type	4 (τ = 0-3)

5. RESULT AND DISCUSSION

To compare the performance for solving problem between adaptive and non-adaptive system, the computational effort [8] is used. It is defined as the average number of individual to be evaluated to obtain the solution.

Let $P(M,i,z)$ be the probability of finding the answer within the generation i, M is the number of individual in the population. P can be observed by repeating the experiment many times. $R(M,i,z)$ be the number of run required to find an answer in the generation i with the confidence z. $R(M,i,z) = \lceil log(1\text{-}z)/ log(1\text{-} P(M,i,z)) \rceil$. The minimum number of individual that must be processed to find an answer with the confidence z is $I(M,i,z) = M \times i \times R(M,i,z)$. The minimum value of $I(M,i,z)$ is defined as the computation effort. The confidence z in this work is 99 %. I^* is the generation that the minimum effort occurs. *#Success* is the number of run that found the answer.

Table (2). shows the computational efforts of the non-adaptive and the adaptive system for the 3 problems. The table shows the good performance of the adaptive system comparing to the non-adaptive system. The computational effort scores are not far from the best non-adaptive system. This shows the success of the adaptive system in the performance issue.

Table (2). The computational efforts.

Problem	I^*	Effort	#Success
One-Max			
τ = 0	14	6,000	500
τ = 1	14	6,000	500
τ = 2	21	8,800	500
τ = 3	25	10,400	500
adaptive	16	6,800	500
Deceptive			
τ = 0	22	64,400	265
τ = 1	32	26,400	475
τ = 2	39	32,000	486
τ = 3	53	21,600	495
adaptive	30	24,800	472
Multimodal			
τ = 0	26	172,800	140
τ = 1	48	39,200	492
τ = 2	63	25,600	497
τ = 3	63	25,600	497
adaptive	37	30,400	484

Fig. (3)-(5) show the adaptive behavior of the adaptive system. They are the plots of the number that a preference type is selected to participate in the crossover. For clarity of the presentation, the data are plotted to the generation 50. They are averaged from 500 runs.

For the one-max problem, the plots show the adaptation toward low diversity. The average generation used of this problem is 13.01 generations. Within this time, the plots show that the lower preference type with are the preference type 0 and 1 are used more frequently. However, once the solution is obtained (from the generation 16 onward), the preference type 0 and 1 (prefer less diversity) cause the loss of diversity and hence they cannot generate any contribution. Their use are declined.

For the deceptive function which is the harder problem. The average generation used of this problem is 23.96. The preference type 0 dominates early on. Once the population hits local minima, more diversity is need, the use of preference type 0 declines and the use of other types increases.

For the multimodal function which is the hardest problem. The plots clearly show the adaptation toward more diversity. The preference type 0 does drop rapidly since the early generation. The average generation used of this problem is 27.66.

The diversity plots (Fig. (6)-(8)) show non-adaptive versus adaptive GAs. They clearly show the effect of adaptation. The adaptive system converges the diversity toward the suitable values in all 3 problems. The calculation of diversity is shown in Eq. (7). The

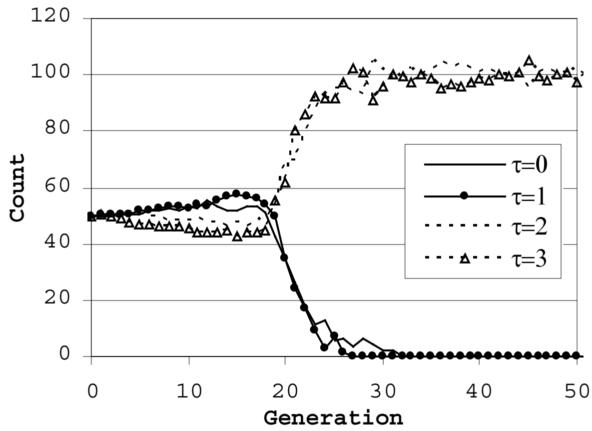

Fig. (3). The number that a preference type is selected of the one-max problem.

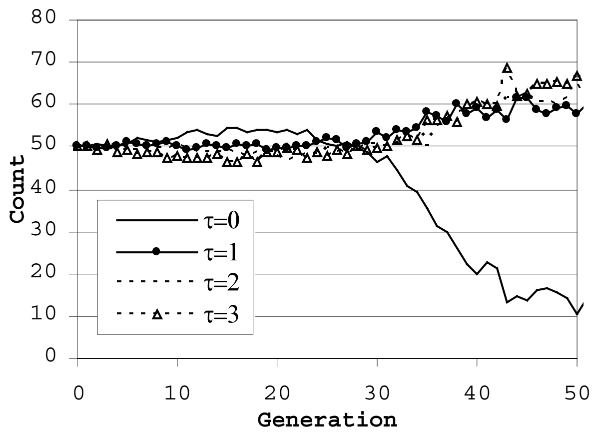

Fig. (4). The number that a preference type is selected of the deceptive function.

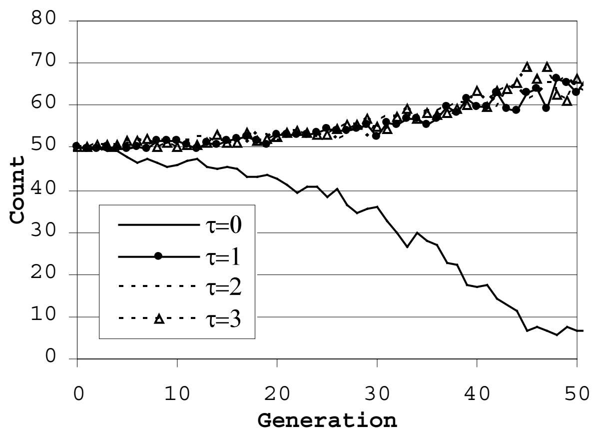

Fig. (5). The number that a preference type is selected of the multimodal function.

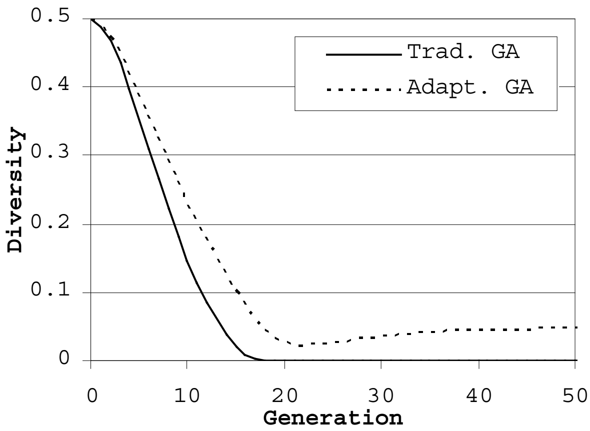

Fig. (6). Diversity comparison between the traditional GA and the adaptive GA of the one-max problem.

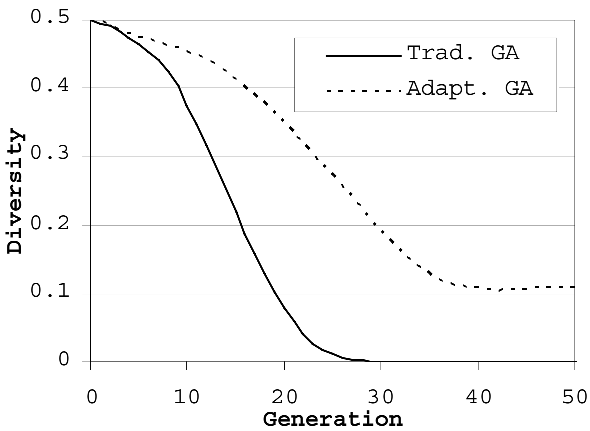

Fig. (7). Diversity comparison between the traditional GA and the adaptive GA of the deceptive function.

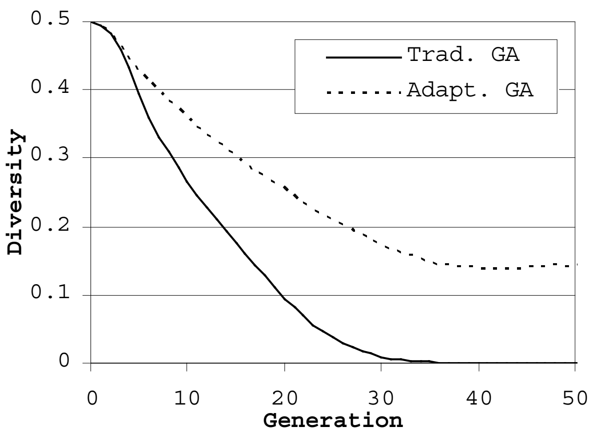

Fig. (8). Diversity comparison between the traditional GA and the adaptive GA of the multimodal function.

maximum value of the diversity is 0.5 and the minimum is 0.

$$Diversity = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} h(I_i, I_j)}{n^2 \cdot l} \qquad \textbf{Eq. (7)}$$

where I_i and I_j are the i^{th} and j^{th} individual in the population respectively, h is the Hamming distance of two individuals, n is the population size, and l is the length of chromosome.

The results demonstrate clearly the ability to adapt the diversity in the population. The proposed method is able to adapt the diversity for a given problem using the preference type and results in the efficient use of resource as can be seen from the computational effort.

6. CONCLUSION

The adaptive system based on the proposed mating procedure, preference mating, works successfully for the standard test problems of GAs. It has capability to adapt the suitable diversity with the good performance for solving problems. Our future work will concentrate on applying this adaptive system to solve complex real-world problems.

ACKNOWLEDGEMENT

The first author would like to acknowledge the support of the Royal Golden Jubilee Ph.D. Graduates program by Thailand Research Fund organization.

REFERENCES

[1] Y. Leung, Y. Gao, Z. Ben Xu, "Degree of Population Diversity - A Perspective on Premature Convergence in Genetic Algorithms and Its Markov Chain Analysis", **IEEE Trans. Neural Networks**, Vol. 8, 1997, pp. 1165-1176.

[2] D.E. Goldberg, J. Richardson, "Genetic Algorithms with Sharing for Multimodal Function Optimization", **Proc. 2nd Int. Conf. Genetic Algorithms**, 1987, pp. 41-49.

[3] P.H Winston, **Artificial Intelligence**, Addison-Wesley, 1992.

[4] L.J. Eshelman, J.D. Schaffer, "Preventing Premature Convergence in Genetic Algorithms by Preventing Incest", **Proc. 4th Int. Conf. Genetic Algorithms**, 1991, pp. 115-122.

[5] E. Ronald, "When Selection Meets Seduction", **Proc. 6th Int. Conf. Genetic Algorithms**, 1995, pp. 167-173.

[6] K. Matsui, "New Selection Method to Improve the Population Diversity in Genetic Algorithms", **Proc. IEEE Int. Conf. Systems, Man, and Cybernetics**, 1999, pp. 625-630.

[7] B. Sareni, L. Krähenbühl, "Fitness Sharing and Niching Methods Revisited", **IEEE Trans. Evolutionary Computation**. Vol. 3, 1998, pp. 97-105.

[8] J.R. Koza, **Genetic Programming II: Automatic Discovery of Reuseable Programs**, MIT Press, 1994.