
A Novel Mating Approach for Genetic Algorithms

Severino F. Galán*

seve@dia.uned.es

Department of Artificial Intelligence, UNED, Madrid, 28040, Spain

Ole J. Mengshoel

ole.mengshoel@sv.cmu.edu

CMU, NASA Research Park, Moffett Field, California 94035, USA

Rafael Pinter

pinter@dia.uned.es

Department of Artificial Intelligence, UNED, Madrid, 28040, Spain

Abstract

Genetic algorithms typically use crossover, which relies on mating a set of selected parents. As part of crossover, random mating is often carried out. A novel approach to parent mating is presented in this work. Our novel approach can be applied in combination with a traditional similarity-based criterion to measure distance between individuals or with a fitness-based criterion. We introduce a parameter called the mating index that allows different mating strategies to be developed within a uniform framework: an exploitative strategy called best-first, an explorative strategy called best-last, and an adaptive strategy called self-adaptive. Self-adaptive mating is defined in the context of the novel algorithm, and aims to achieve a balance between exploitation and exploration in a domain-independent manner. The present work formally defines the novel mating approach, analyzes its behavior, and conducts an extensive experimental study to quantitatively determine its benefits. In the domain of real function optimization, the experiments show that, as the degree of multimodality of the function at hand grows, increasing the mating index improves performance. In the case of the self-adaptive mating strategy, the experiments give strong results for several case studies.

Keywords

Genetic algorithms, premature convergence, mating strategies, mating index, self-adaptive mating.

1 Introduction

Genetic algorithms (GAs; Holland, 1975; Goldberg, 1989) use stochastic search methods based on natural evolution in order to solve problems in fields like optimization, design, learning, or scheduling, among others. A GA creates a set of candidate solutions each generation. The quality of a solution, its fitness, determines its chance to survive and reproduce. Two processes form the basis of genetic algorithms: variation (recombination and mutation) and selection. While the former facilitates diversity and novelty, the latter favors quality. Ideally, at the end of running a GA, a solution with optimal or near-optimal fitness is found.

Premature convergence to local optima is one of the most frequent difficulties that arise when applying GAs to complex problems. It occurs when genetic operators can no

*Corresponding Author.

longer generate offspring that are fitter than their suboptimal parents. Premature convergence is associated with the loss of diversity in the population. However, too much population diversity can lead to a dramatic deterioration of GA efficiency. Therefore, an important issue in the design and application of GAs is the trade-off between exploitation of the best individuals and exploration of alternative regions of the search space.

By focusing on the mating phase of GAs, the present work deals with achieving a proper balance between exploitation and exploration. Traditionally, mating takes place after parent selection and prior to recombination. Normally, parents are mated in pairs so that each pair can subsequently be recombined. A key question is how mating should be carried out in order to give strong GA performance. The traditional mating approach consists of selecting a parent's mate uniformly at random from the set of remaining parents. In addition to the traditional random mating approach, other approaches exist that perform mating restriction based on similarity relations between parents (Deb and Goldberg, 1989; Eshelman and Schaffer, 1991; Smith and Bonacina, 2003). Although these methods have been shown to benefit GA performance, they are costly in computational terms. This disadvantage is due to the fact that similarity comparisons between two parents' chromosomes are performed gene by gene. Furthermore, these methods were designed for rather specific contexts such as fitness sharing (Deb and Goldberg, 1989) and incest prevention (Eshelman and Schaffer, 1991) and, therefore, their impact has been quite limited.

The goal of this work is to develop, analyze, and evaluate a novel and general approach to mating in GAs. The novel approach uses a parameter called the mating index, which allows the degree of exploration to be controlled depending on the hardness of the problem to be solved. In this way, we hope that our approach can easily be applied to a wide variety of problems of different complexity. In addition, by using fitness-based comparisons between parents, rather than only similarity-based comparisons, the computational complexity of the applied mating algorithm can be reduced. Furthermore, the novel approach lends itself to a self-adaptive algorithm which gives rise to a useful mating strategy.

The domain of real function optimization is used in this work to experimentally study the benefits of our new approach. ANOVA tests have been performed in order to determine the statistical significance of the experimental results obtained. The main results of the study are the following:

- The main parameters of the new approach, mating size and mating index, have a strong influence on GA performance.
- The degree of multimodality (number of local optima) for the function at hand determines which mating strategy produces the best performance. In general, the higher the degree of multimodality, the higher the mating index should be.
- Experiments with the self-adaptive mating strategy give good results for unimodal functions. For multimodal functions, the results under self-adaptive mating are highly dependent on the mating size. (The self-adaptive strategy implemented in our work keeps a constant mating size and self-adapts the mating indices contained in the chromosomes.)
- From a qualitative point of view, fitness-based mating produces analogous results to those of similarity-based mating for an important set of the studied cases. However, fitness-based mating needs less computation time.

The rest of this paper is structured as follows. Section 2 reviews previous work on restricted mating in GAs. Section 3 introduces our novel approach to mating, along with the different mating strategies derived from it. Section 4 analyzes the mating approach. Section 5 includes an extensive empirical evaluation of the novel mating strategies. We discuss our experimental results for mating in GAs in Section 6. Finally, Section 7 contains the main conclusions and discusses future work.

2 Restricted Mating in Genetic Algorithms

The traditional way of mating parents in GAs consists of taking a parent from the mating pool and selecting its mate by choosing uniformly at random one of the remaining parents. The mated parents are then removed from the mating pool, and the same process is repeated until all the individuals have been mated. Restricted mating techniques, which do not select a mate uniformly at random, have been successfully developed for specific contexts such as fitness sharing (Deb and Goldberg, 1989) and incest prevention (Eshelman and Schaffer, 1991). Other approaches that incorporate mating preferences into evolutionary systems are: assortative mating GAs (Fernandes et al., 2001; Huang, 2001; Ochoa et al., 2005), correlative tournament selection (Matsui, 1999), seduction (Ronald, 1995), TABU GA (Ting et al., 2003), and evolving agents (Smith et al., 2000; Smith and Bonacina, 2003; Unemi and Nagayoshi, 1997).

Fitness sharing (Deb and Goldberg, 1989) is a method that forces the population to maintain different niches. In multimodal optimization problems, where a number of high-fitness individuals corresponding to various local optima are identified, niches are search space regions around local optima, and high-fitness niches are of particular interest. Fitness sharing adjusts the fitness of individuals prior to parent selection, so that individuals are allocated to niches in proportion to the niche fitness. In order to improve the efficiency of fitness sharing, Deb and Goldberg used a restricted mating approach whose goal was to avoid the creation of lethal (low fitness) individuals. Once niches are formed in the population, the recombination of two parents from different niches is likely to form lethal offspring. Therefore, restricted mating among individuals of the same niche is promoted. This is achieved by following the same scheme as random mating but, given a parent, a candidate mate is accepted only if the phenotype/genotype distance between them is smaller than a given threshold. Otherwise, another candidate is sought. If no candidate is found, one is chosen uniformly at random as in random mating. In the case of fitness sharing for real functions optimization, the phenotype space corresponds to the real values of the variables, while the genotype space uses a binary representation for them. If similarity is measured within the phenotypic space, Euclidean distance is used. Hamming distance is employed when similarity between individuals is measured within the genotypic space. Other GA approaches applying restricted mating in the specific context of multimodal optimization problems are island models (Cohon et al., 1987; Martin et al., 1997), diffusion models (Manderick and Spiessens, 1989; White and Pettey, 1997), and automatic speciation models (Booker, 1982; Spears, 1994).

In contrast to fitness sharing, incest prevention (Eshelman and Schaffer, 1991) was defined in the context of global optimization rather than niching. Incest prevention promotes restricted mating between dissimilar enough individuals. In general, when two similar individuals are mated, their offspring may not introduce significant new information about the search space, which provokes a reduction in the performance of the GA. Incest prevention follows a dual scheme to that used in fitness sharing. A candidate mate is accepted only if its phenotype/genotype distance to the current

parent is greater than a given threshold. Usually, this threshold is reduced when a better offspring is not obtained during the search process.

In comparison to random mating, similarity-based restricted mating has been shown to produce a more effective exploration of the search space, both in fitness sharing (Deb and Goldberg, 1989) and in incest prevention (Eshelman and Schaffer, 1991). However, these similarity-based approaches are in some sense polar opposites, and in this work we develop a uniform framework for the similarity-based approach, thus improving the understanding of it. At the same time, the time cost associated with measuring the distances between individuals is a disadvantage of the similarity-based approach. This work explores fitness-based mating as an alternative in order to establish mating preferences with a lower computational cost. Although fitness-based restricted mating was addressed earlier (De et al., 1998; Chakraborty and Chakraborty, 1999), this technique has not been sufficiently investigated in the past, due to the widespread use of similarity in the definition of mating approaches. One of the goals of this work is to thoroughly compare fitness-based mating strategies with their similarity-based counterparts.

The present work aims at formalizing a general mating approach which allows a wide range of mating strategies to be defined and effectively applied to the task of global optimization in GAs. Finally, a self-adaptive mating method is developed.

3 The New Mating Approach

This section presents a novel approach to mating in GAs. Our novel approach has three main characteristics. Firstly, a parameter named the mating index allows the degree of exploration to be controlled in a simple way, which makes our approach flexible and general. Secondly, it allows mating preferences to be defined either in terms of similarity between individuals or in terms of fitness of individuals, in contrast to most of the mating strategies reviewed in Section 2, which are typically based on similarity between individuals. Thirdly, the novel approach lends itself to a self-adaptive implementation, in which each individual in the population has its own mating preference; in this way, different mating strategies can be applied depending on the hardness of the fitness function and the current state of the search process.

The novel approach is defined by Algorithm 1, which constitutes a GA's mating phase, taking place between parent selection and parent recombination. In this algorithm, γ (mating size), different parents are randomly chosen for the next round of mating, and the fittest of them is mated with another individual as determined by cr (mating criterion) and α (mating index). Note that P_s in Algorithm 1 could contain multiple copies of individuals after parent selection, depending on the fitness value.

Similarity in the phenotype space is the traditional criterion used to establish mating preferences in GAs (Deb and Goldberg, 1989; Eshelman and Schaffer, 1991). This is why similarity has been included in the domain of cr in our algorithm. Due to the computational complexity of similarity comparisons, a new fitness-based criterion for establishing mating preferences is also introduced. While determining the similarity of two individuals requires examining their chromosomes gene by gene, comparing their fitnesses involves examining only two numbers.

Note that the novel mating approach defined by Algorithm 1 becomes the traditional mating approach when $\gamma = 2$, since parents are then mated uniformly at random. Therefore, the novel mating approach is a generalization of the traditional approach. The values of mating index α induce different mating strategies, corresponding to different degrees of exploitation versus exploration. A specific α value defines a mating

Algorithm 1 Mating step of the novel mating approach for GAs**Input:**

- P_s Population of selected parents
- γ Mating Size: number of eligible parents for next mating
- cr Criterion used for defining mating preferences
($cr \in \{\text{similarity, fitness}\}$)
- α Mating Index: integer used for defining a mating preference
($2 \leq \alpha \leq \gamma$)

Output:

- P_m Population of mated parents
 $\{\langle p_1(1), p_2(1) \rangle, \dots, \langle p_1(i), p_2(i) \rangle, \dots, \langle p_1(|P_s|/2), p_2(|P_s|/2) \rangle, p\}$

begin

$P_m \leftarrow \emptyset$

$i \leftarrow 0$

while ($|P_s| > 1$)

$i \leftarrow i + 1$

$Ch \leftarrow$ Set of γ parents chosen uniformly at random from P_s without replacement

$p_1(i) \leftarrow$ Parent in Ch with highest fitness

Remove $p_1(i)$ from P_s and Ch , and include it in P_m . (Set Ch now includes the candidate mates for $p_1(i)$.)

From the set of $\gamma - 1$ candidates in Ch , obtain the ordered subset of $\alpha - 1$ best candidates under criterion cr . When $cr = \text{similarity}$, candidates are ranked according to decreasing phenotype similarity with p_1 . When $cr = \text{fitness}$, candidates with higher fitness are ranked first. Let

$$B_{\alpha-1}(p_1(i)) = \{b_2, b_3, \dots, b_{\alpha-1}, b_\alpha\}$$

be the resulting ordered subset. (Note that b_i is the i th order statistic of the original sample set Ch when $cr = \text{fitness}$ and a minimization problem is considered, or when $cr = \text{similarity}$ and phenotype distance is used to measure similarity.)

$p_2(i) \leftarrow b_\alpha$

Remove $p_2(i)$ from P_s and include it in P_m . Note that $p_1(i)$ and its mate $p_2(i)$ are placed in contiguous positions in P_m . (If $|B_{\alpha-1}(p_1(i))| < \alpha - 1$, the last element in $B_{\alpha-1}(p_1(i))$ is chosen as mate for $p_1(i)$.)

endwhile

if ($|P_s| = 1$) **then** include in P_m the remaining individual p of P_s

return P_m

end

strategy that we name best- $(\alpha - 1)$ th, since the best parent is mated with the $(\alpha - 1)$ -th mating candidate. Several variants of the general scheme in Algorithm 1 are possible. This paper focuses on three of them:

1. When $\alpha = 2$, the best parent is mated with the first mating candidate under criterion cr . Thus, the resulting scheme is called best-first mating.

2. When $\alpha = \gamma$, the best parent is mated with the last mating candidate under criterion cr . This strategy is called best-last mating.
3. When the parameter α is made local to each individual, encoded into the chromosome, and subjected to recombination and mutation, a self-adaptive mating strategy results.

In the rest of this section, these three mating strategies are discussed in more detail.

3.1 Best-First Mating

Exploitation of the best solutions in the current population can be achieved by setting $\alpha = 2$ in Algorithm 1. In this way, the fittest of the chosen parents, p_1 , is mated with the first of the candidates under criterion cr . If a fitness-based criterion is used, p_1 's mating preference is clearly an exploitative strategy, since fitter candidates are preferred over the rest. If a similarity-based criterion is used, p_1 's mating preference is exploitative as well, since it is implicitly assumed that fitter candidates are more similar to p_1 than the rest. Since P_s in Algorithm 1 could contain multiple copies of individuals, it is possible that some clones are mated by best-first.

In the Algorithm 1, an ordering of the first $\alpha - 1$ candidates in Ch under criterion cr is in fact not necessary in best-first mating. Only the first of the candidates under criterion cr is sought. Thus, only a variable storing the current first candidate is needed to implement best-first.

Best-first mating with a similarity-based criterion is inspired by the mating strategy used by Deb and Goldberg (1989) in the context of fitness sharing. Whereas Deb and Goldberg used a similarity threshold to guide the mating process within niches, best-first mating employs a mating size parameter in order to obtain a certain degree of exploitation. Similarity-based best-first mating is also similar to positive assortative mating (Fernandes et al., 2001; Huang, 2001; Ochoa et al., 2005), which chooses the most similar candidate as the mate of an individual. At the same time, best-first mating with a fitness-based criterion has common characteristics with some of the mating methods developed previously (De et al., 1998; G. Chakraborty and B. Chakraborty, 1999).

3.2 Best-Last Mating

Exploration of alternative solutions to the best ones in the current population can be performed by setting $\alpha = \gamma$ in Algorithm 1. By doing that, the fittest of the chosen parents, p_1 , is mated with the last of the candidates under criterion cr . If a fitness-based criterion is used, p_1 's mating preference is clearly an explorative strategy, since the fittest parent prefers less fit candidates over the rest. If a similarity-based criterion is used, p_1 's mating preference is explorative as well, since the most distant candidate in the phenotype space is chosen for mating.

In Algorithm 1, an ordering of the first $\alpha - 1$ candidates in Ch under criterion cr is not necessary in best-last mating. Only the last of the candidates under such a criterion is sought. Therefore, a unique variable storing the currently last candidate is needed to implement best-last.

Best-last mating with a similarity-based criterion is inspired by the mating strategy used by Eshelman and Schaffer (1991) for incest prevention. While Eshelman and Schaffer used a similarity threshold to prevent incest, best-last mating achieves a particular degree of exploration by setting the mating size value. Similarity-based best-last mating is also similar to negative assortative mating (Fernandes et al., 2001; Huang, 2001;

Algorithm 2 Self-adaptive mate election

Input:

- p_1 Individual to be mated, with genotype $\langle x_{p_1,1}, \dots, x_{p_1,n}, \alpha_{p_1} \rangle$
 Ch Candidate mates for p_1
 cr Criterion used for defining mating preferences
 ($cr \in \{\text{similarity with respect to } p_1, \text{ fitness}\}$)
 P_s Population of selected parents
 P_m Population of mated parents

Output:

- p_2 Mate for p_1

begin

- $p_2 \leftarrow (\alpha_{p_1} - 1)$ -th candidate in Ch under criterion cr
 Remove p_2 from P_s
 Include p_2 in P_m

end

Ochoa et al., 2005), which chooses the most dissimilar candidate as the mate for the current individual. On the other hand, to the best of our knowledge, best-last mating with a fitness-based criterion has not yet been investigated in the literature.

3.3 Self-Adaptive Mating

A GA's parameters can either be manually tuned in advance or automatically controlled during execution. Compared to manual parameter tuning, the advantages of automatic parameter control are that (i) it is less taxing on the GA's user and (ii) parameters can be adapted to the state of the search process. A classification of parameter setting techniques for evolutionary algorithms can be found elsewhere (Eiben et al., 1999; Eiben and Smith, 2003, Chapter 8). This section discusses self-adaptive control of mating parameters. Self-adaptive parameter control consists of encoding the parameters into the chromosomes and performing recombination and mutation on them. In this way, the values of the parameters leading to better individuals will have a greater chance to survive.

If an individual j is represented as $\langle x_{j,1}, \dots, x_{j,n} \rangle$, its extended representation under self-adaptive mating would be $\langle x_{j,1}, \dots, x_{j,n}, x_{j,n+1} \rangle$, where $x_{j,n+1} = \alpha_j$ is the mating index for individual j . In other words, the mating index is now a local parameter, and each individual has an independent mating preference. The algorithm performing self-adaptive mating can easily be obtained from Algorithm 1 by removing α from the input and substituting α with $x_{p_1,n+1}$ in the body of the algorithm. Algorithm 2 shows the algorithm for the mate election step under self-adaptive mating.

It remains to consider how mating indices are initialized, recombined, and mutated. As far as initialization is concerned, each mating index is assigned an integer generated uniformly at random from the range $[2, \gamma]$. Recombination of the mating indices of two parents can be carried out in several ways: by assigning to the two children the mean of the parents' mating indices, or by letting the two children inherit the parents' mating indices, among other possibilities. This work uses the latter method, since we have found it to produce better experimental results. Mutation of mating indices

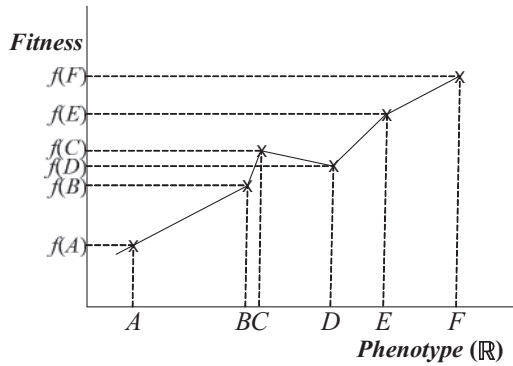


Figure 1: Population of six parents $\{A, B, C, D, E, F\}$ to be mated. These parents belong to a real interval $I \subset \mathbb{R}$.

is implemented by setting a probability $p_{=}$ that the mating index is unchanged, a probability p_{+} that the mating index is incremented by one, a probability p_{-} that the mating index is decremented by one, and a probability $1 - p_{=} - p_{+} - p_{-}$ that the mating index is changed uniformly at random. Values $p_{=} = 0.5$ and $p_{+} = p_{-} \lesssim 0.25$ were employed, since they gave better performance in the experiments.

3.4 An Example of the Novel Mating Strategies

We now discuss an example illustrating best-first mating, best-last mating, and self-adaptive mating. Consider in Algorithm 1 a population of six selected parents, $P_s = \{A, B, C, D, E, F\}$, resulting after parent selection in a GA. The six parents have to be mated before recombination. Figure 1 depicts the parents according to their phenotype (x axis) and their fitness (y axis), where it is assumed that there is a bijection between phenotypes of individuals and a certain interval of real numbers.

The random mating strategy mates parents by choosing a mate uniformly at random among the remaining parents. A possible mating resulting from this strategy is $\{CF, EA, BD\}$. It is important to note that fitness or similarity information is not used at any step of random mating.

If a mating size $\gamma = 6$ is assumed for simplicity, the best-first, the best-last, and the self-adaptive mating strategies create $Ch = \{A, B, C, D, E, F\}$. The first mate, p_1 , is the parent with highest fitness in Ch ; $p_1 = F$ in this case. In best-first, the second mate, p_2 , is the first of the candidates in $\{A, B, C, D, E\}$ under criterion cr . First, we discuss the similarity-based criterion, under which $p_2 = E$. As a result, F and E are mated, and the same process continues until all of the parents have been mated. A new set Ch would be formed prior to each pairing between p_1 and p_2 . Ultimately, the mating resulting from this strategy is $\{FE, CB, DA\}$. Second, we discuss the fitness-based criterion, under which $p_2 = E$ for $p_1 = F$ as in the similarity-based criterion. Next, C is paired with D , instead of with B as in the similarity case. The final pairing is $\{FE, CD, BA\}$.

The best-last mating strategy works analogously to best-first, but now the last of the candidates under criterion cr is assigned to p_2 at each iteration of the mating algorithm. In this way, under similarity-based mating, F and A are first mated. In the end, the mating resulting from this strategy is $\{FA, EB, CD\}$. For fitness-based mating, the final pairing is $\{FA, EB, CD\}$ as well.

Table 1: Pairings obtained for the example in Figure 3.

Mating strategy	Criterion	Pairing
Random		{CF, EA, BD}
Best-first	Similarity	{FE, CB, DA}
	Fitness	{FE, CD, BA}
Best-last	Similarity	{FA, EB, CD}
	Fitness	{FA, EB, CD}
Self-adaptive	Similarity	{FE, CD, BA}
	Fitness	{FE, CB, DA}

The self-adaptive mating strategy is now considered. The following mating indices will be assumed: $\{\alpha_A = 5, \alpha_B = 2, \alpha_C = 3, \alpha_D = 3, \alpha_E = 4, \alpha_F = 2\}$. The first mate, p_1 , is again the parent with the highest fitness in Ch ; $p_1 = F$ in this case. The second mate, p_2 , is the $(\alpha_F - 1)$ -th candidate in $\{A, B, C, D, E\}$ under criterion cr . First, we discuss the similarity-based criterion, under which $p_2 = E$. As a result, F and E are mated, and the same process continues until all of the parents have been mated. A new set Ch is formed prior to each pairing between p_1 and p_2 . When the number of candidates for p_1 in Ch is smaller than α_{p_1} , the last element in Ch under criterion cr is selected as mate for p_1 . The mating resulting from this strategy is $\{FE, CD, BA\}$. Second, we discuss the fitness-based criterion, under which $p_2 = E$ as in similarity-based criterion. Then, C is paired with B , instead of with D as in similarity-based mating. The final pairing is $\{FE, CB, DA\}$.

The pairings obtained for the different mating strategies considered in this section are summarized in Table 1. It should be noted that best-first produces the best potential mating for the simple fitness function in Figure 1, since mating parents with high fitness will favor the creation of children with high fitness with higher probability. However, in more realistic and interesting fitness functions, it is clear that best-first is not always optimal, as Section 5 shows.

4 Analysis

The present section analyzes the novel mating approach introduced in this work. The analysis is developed for the two types of domains that can be encountered in function optimization: on the one hand, unimodal fitness functions with a unique optimum and, on the other hand, multimodal fitness functions with many local optima. In both cases, the influence of mating on the effectiveness to reach the global optimum is studied. For simplicity, this section only discusses similarity-based mating. As in Section 3.4, it will be assumed that there is a bijection between phenotypes of individuals and a certain interval of real numbers.

4.1 Analysis for Unimodal Fitness Functions

We divide the analysis for unimodal fitness functions into two cases: functions with no peaks and functions with one peak. In either case, and for simplicity, linear functions will be used.

4.1.1 Functions Without Peaks

Consider the maximization problem of the linear fitness function depicted in Figure 2, $f(x) = m \cdot x$, where x represents individuals' phenotype defined in the range $[x_1, x_2] \subset \mathbb{R}$ (with $x_1 < x_2$) and where m is a positive constant. Let $x' \in (x_1, x_2)$ denote the fittest

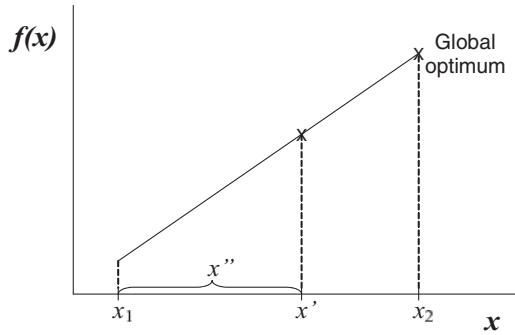


Figure 2: A linear unimodal fitness function with no local optima except for the global optimum. Individual $x' \in [x_1, x_2]$ is mated with individual $x'' \in [x_1, x']$.

individual within set Ch of Algorithm 1; in other words, x' is the best individual resulting from selecting (uniformly at random) γ individuals from the current population of parents. Therefore, as shown in Figure 2, x' is the next individual to be assigned a mate x'' . Candidate mates for x' can only belong to range $[x_1, x']$, since $f(x') \geq f(x'')$ from Algorithm 1.

Assume that the recombination of two parents, $x_{p_1} \in [x_1, x_2]$ and $x_{p_2} \in [x_1, x_2]$ with $x_{p_1} < x_{p_2}$, is performed in the phenotype space and produces children, $x_{\text{child}(x_{p_1}, x_{p_2})}$, uniformly distributed over $[x_{p_1}, x_{p_2}]$. Consequently, the expected fitness of a child of x_{p_1} and x_{p_2} can be calculated as:

$$\bar{f}(x_{\text{child}(x_{p_1}, x_{p_2})}) = \frac{f(x_{p_1}) + f(x_{p_2})}{2}, \quad (1)$$

since the probability that $x_{\text{child}(x_{p_1}, x_{p_2})} \notin [x_{p_1}, x_{p_2}]$ is equal to zero and f is linear with respect to x . Given that $f(x) = m \cdot x$, Equation (1) turns into:

$$\bar{f}(x_{\text{child}(x_{p_1}, x_{p_2})}) = m \cdot \frac{x_{p_1} + x_{p_2}}{2}. \quad (2)$$

From Equation (2), the expected fitness of a child resulting from the recombination of x' and x'' , with $x'' \in [x_1, x']$, is the following:

$$\bar{f}(x_{\text{child}(x', x'')}) = m \cdot \frac{x' + x''}{2}. \quad (3)$$

It is important to note that, for a fixed x' , $\bar{f}(x_{\text{child}(x', x'')})$ in Equation (3) increases as x'' approaches x' and reaches a maximum when $x'' = x'$. This indicates that the closer x'' is to x' , the fitter are their children (in expectation). Furthermore, since we are dealing with a unimodal fitness function, higher values of $\bar{f}(x_{\text{child}(x', x'')})$ indicate children closer to the global optimum. It can be concluded from this result that, in the case of unimodal fitness functions with no peaks, those mating strategies under the novel approach favoring recombination with similar individuals produce a more effective search for the global optimum.

4.1.2 Functions With One Peak

Consider the maximization problem of the fitness function shown in Figure 3, defined as follows:

$$f(x) = \begin{cases} mx & \text{if } x \in [x_1, x_p] \\ -mx + 2mx_p & \text{if } x \in [x_p, x_2] \end{cases},$$

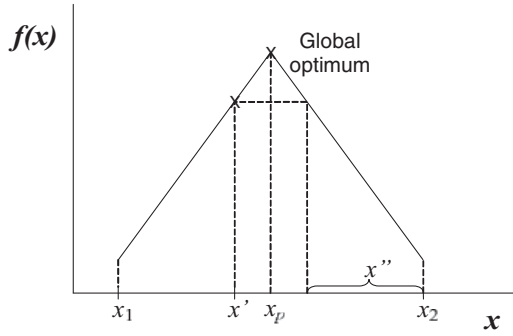


Figure 3: A one-peak fitness function. The case is shown where individual $x' \in (x_1, x_p)$ is mated with individual $x'' \in [2x_p - x', x_2]$.

where m is a positive constant and $x_p \in (x_1, x_2)$ is a peak for the function. If we assume that $x' \in (x_1, x_p)$, then $x'' \in [x_1, x'] \cup [2x_p - x', x_2]$. Since the case where $x'' \in [x_1, x']$ was already studied in the previous section, we will focus on the case where $x'' \in [2x_p - x', x_2]$.

When $x'' \in [2x_p - x', x_2]$, the expected fitness for a child of x' and x'' can be calculated as follows:

$$\begin{aligned} \bar{f}(x_{\text{child}(x', x'')}) &= \frac{\int_{x'}^{x''} f(x) dx}{x'' - x'} = \frac{\int_{x'}^{x_p} mx dx + \int_{x_p}^{x''} (-mx + 2mx_p) dx}{x'' - x'} \\ &= m \cdot \left\{ \frac{x'' + x'}{2} - \frac{(x'' - x_p)^2}{x'' - x'} \right\}, \end{aligned} \quad (4)$$

which takes positive and monotonically decreasing values with x'' in the interval $[2x_p - x', x_2]$ provided that $f(x_2) > 0$. The following important results can be derived from Equation (4):

1. If $x'' = 2x_p - x'$, then $\bar{f}(x_{\text{child}(x', x'')}) = m \cdot \frac{x' + x_p}{2} > m \cdot x'$. Since $m \cdot x'$ is the maximum value that $\bar{f}(x_{\text{child}(x', x'')})$ can take when $x'' \in [x_1, x']$ (see Section 4.1.1), in the case of the type of unimodal fitness functions with one peak considered in this section, symmetric mates with respect to the optimum produce the best expected children. These symmetric mates can be easily identified by adopting the following values for the parameters of Algorithm 1: a mating size γ as high as possible, a mating index α as low as possible, and $cr = \text{fitness}$. The superiority of $cr = \text{fitness}$ with respect to $cr = \text{similarity}$ in unimodal functions is confirmed by Figures 6(a) and 7(a) of Section 5.1.1, where convergence is reached earlier when $cr = \text{fitness}$.
2. For $x'' \in [2x_p - x', x_*]$, mating always produces better expected children than those obtained for $x'' \in [x_1, x']$ and $x'' \lesssim x'$ (see Section 4.1.1). The value of x_* can be calculated by making $\bar{f}(x_{\text{child}(x', x'')}) = m \cdot x'$ in Equation (4), which yields:

$$x_* = 2x_p - x' \pm \sqrt{2}(x' - x_p).$$

Since $x_* > 2x_p - x'$ and $x_p > x'$, the first solution is not valid. The second solution can be written as follows:

$$x_* = (2 + \sqrt{2})x_p - (1 + \sqrt{2})x'.$$

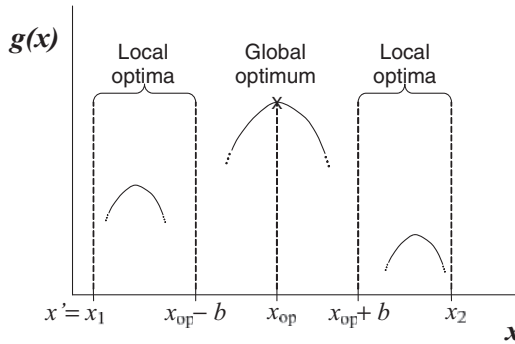


Figure 4: A multimodal fitness function.

It should be noted that the length of interval $[2x_p - x', x_*']$ is equal to $\sqrt{2}(x_p - x')$, which is proportional to $x_p - x'$. Given an individual x' , interval $[2x_p - x', x_*']$ constitutes a promising area for mate election. The resulting children are expected to be fitter than parent x' if $x'' \in [2x_p - x', x_*']$. In terms of the novel mating approach presented in this paper, if mating size γ is high enough, then the lowest mating indices (giving rise to $x'' \lesssim x'$) would produce worse expected children than those obtained with somewhat higher mating indices (giving rise to $x'' \in [2x_p - x', x_*']$).

4.2 Analysis for Multimodal Fitness Functions

The method used in Section 4.1 to establish how beneficial it is for x' to elect x'' as a mate consists of obtaining the expected fitness for the children of x' and x'' , $\bar{f}(x_{\text{child}(x', x'')})$, which in principle can be calculated for any fitness function shape. However, it is important to note that, for multimodal or deceptive fitness functions, higher $\bar{f}(x_{\text{child}(x', x'')})$ values do not necessarily correspond to children that are closer to the global optimum. (For this type of function, higher $\bar{f}(x_{\text{child}(x', x'')})$ values might even correspond to children farther from the global optimum than their parent x' .) As a consequence, in the rest of this section we adopt an alternative analysis method to that used in Section 4.1. The new analysis method is centered on the concept of hitting the basin of attraction for the global optimum.

Consider the maximization problem of a multimodal fitness function, $g(x)$, where x represents individuals' phenotype defined in the range $[x_1, x_2]$ with $x_1 < x_2$, as depicted in Figure 4. Assume that function $g(x)$ has a global optimum at x_{op} , whose basin of attraction lies in the range $[x_{\text{op}} - b, x_{\text{op}} + b]$. The basin of attraction for a maximum is the set of points in the fitness landscape from which a steepest-ascent hill-climbing search would finish reaching that maximum.

Without loss of generality, consider that individual $x' = x_1$ is to be assigned a mate x'' . Assume that the function $g(x)$ has several local optima between x' and $x_{\text{op}} - b$. Contrary to Section 4.1, whose argument can only be applied to individuals in the basin of attraction for the global optimum, the individual x' has now been chosen outside that basin. This is the usual situation for individuals when a multimodal fitness function is optimized.

As in Section 4.1, we assume that the recombination of two parents, $x_{p_1} \in [x_1, x_2]$ and $x_{p_2} \in [x_1, x_2]$ with $x_{p_1} < x_{p_2}$, is performed in the phenotype space and produces

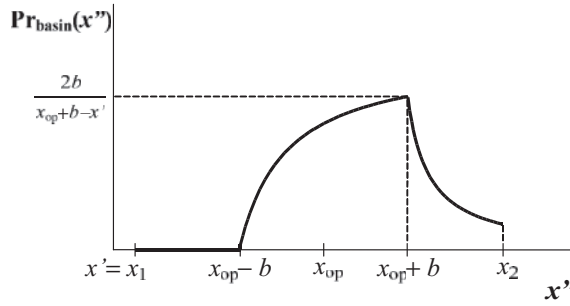


Figure 5: Probability that a child of x' and x'' reaches the basin of attraction for the global optimum $[x_{op} - b, x_{op} + b]$.

children, $x_{\text{child}(x_{p_1}, x_{p_2})}$, uniformly distributed over $[x_{p_1}, x_{p_2}]$. Since the problem of global optimization is to be solved, it is convenient to calculate the probability of reaching the basin of attraction for the global optimum by recombining x' with a candidate mate $x'' \in [x_1, x_2]$. (We are not interested in reaching the regions of the fitness landscape located outside the basin of attraction for the global optimum, since that would lead to premature convergence.) Such a probability is zero for $x'' \in [x_1, x_{op} - b]$, since $[x', x''] \cap [x_{op} - b, x_{op} + b] = \emptyset$. For $x'' \in [x_{op} - b, x_{op} + b]$, the probability that the child reaches $[x_{op} - b, x_{op} + b]$ is equal to $\frac{x'' - (x_{op} - b)}{x'' - x'}$. Finally, for $x'' \in (x_{op} + b, x_2]$, the probability is equal to $\frac{2b}{x'' - x'}$, since $[x', x''] \cap [x_{op} - b, x_{op} + b] = [x_{op} - b, x_{op} + b]$. In summary, as shown in Figure 5,

$$\Pr_{\text{basin}}(x'') = \begin{cases} 0 & \text{if } x'' \in [x_1, x_{op} - b) \\ \frac{x'' - (x_{op} - b)}{x'' - x'} & \text{if } x'' \in [x_{op} - b, x_{op} + b] \\ \frac{2b}{x'' - x'} & \text{if } x'' \in (x_{op} + b, x_2] \end{cases} \quad (5)$$

Figure 5 shows that, while mates in $[x_1, x_{op} - b)$ lead to local optima, there is always a probability greater than zero that the basin of attraction for the global optimum is reached if mates are taken from $[x_{op} - b, x_2]$. Note that interval $[x_1, x_{op} - b)$ corresponds to the more similar individuals to x' , which are chosen as mates by x' when using low mating indices under the novel mating approach. The other interval, $[x_{op} - b, x_2]$, contains distant individuals to x' corresponding to high mating indices. (Note also that the higher the mating size is in Algorithm 1, the more probable it is that a low mating index produces a similar mate for x' and that a high mating index produces a dissimilar mate for x' .) Therefore, mating strategies using high mating indices produce in general a more effective search for the global optimum in the case of multimodal functions.

4.3 Analysis Discussion

The analysis in this section has covered optimization problems of increasing complexity. Although we have dealt with simplified fitness functions for illustration purposes, we believe that our analysis gives insight into more general fitness functions that one can find in practice.

One way in which the analysis in Section 4.1.2 can be extended is by allowing $f(x)$ to be a nonlinear function. For example, the sphere function (see Section 5.1.1) is a unimodal function that assigns quadratic fitness depending on the distance to the global optimum. Even if a quadratic $f(x)$ is used in Equation (4) instead of a linear $f(x)$, similar results to those explained in Point 1 and Point 2 of Section 4.1.2 can be obtained.

Another way to extend the analysis in Section 4.1.2 is to consider multidimensional functions. In this case, Equation (4), which calculates the expected fitness for the children of x' and x'' , needs to be defined for the multidimensional case. The experiments in Section 5.1.1 confirm that, as in the case of the simplified function studied in Section 4.1.2, low mating indices are more advantageous in the case of multidimensional quadratic unimodal functions. In the case of the multimodal functions analysis of Section 4.2, both the concept of basin of attraction and Equation (5) can be generalized to the multidimensional case. An example of a multimodal function defined over several variables is the Schwefel function, which is experimentally studied in Section 5.1.2. The experiments in this section for the Schwefel function confirm the conclusions derived in Section 4.2 for unidimensional multimodal functions. High mating indices yield the best results in this case.

5 Experiments

The experiments are concerned with optimization of real functions. Given an n -dimensional function, $f: \mathbb{R}^n \rightarrow \mathbb{R}$, global optimization consists of determining $x^* \in \mathbb{R}^n$ such that $f(x^*) \geq f(x)$ for all $x \neq x^*$ with $x \in \mathbb{R}^n$. This definition corresponds to a maximization problem. In the case of minimization, the inequality to be considered is $f(x^*) \leq f(x)$.

In Sections 5.1 and 5.2, a discretized real interval is considered for each dimension of the function domain. Each interval point is encoded as a binary string by using a gray code. The experiments in these two sections were performed by means of a simple GA using tournament parent selection with tournament size equal to two, one-point crossover with crossover probability equal to one, bit-flip mutation with mutation probability equal to the inverse of the chromosome length, generational survivor selection, and elitism for the best individual. Different seeds for the random number generator were used for each run of the simple GA. All of the experiments were carried out on a 2-GHz processor running Windows.

In order to determine whether the results obtained in Sections 5.1 and 5.2 for mating under binary representation extend to real representation of variables, Sections 5.3 and 5.4 report on the same type of experiments for a GA using real-coded variables. In the case of real representation, new crossover and mutation operators need to be introduced.

5.1 Binary Variables: Mating Size Experiments

This section contains a comparative evaluation of the following mating strategies: Random mating, best-first mating, best-last mating, and self-adaptive mating. In general, the best-first mating strategy produces exploitation of the best solutions in the current population, best-last mating produces exploration of alternative solutions to the best ones in the current population, and self-adaptive mating produces a combination of exploration and exploitation that depends on the shape of the fitness function and the state of the search process.

The rest of this section is structured so that the following comparisons are progressively made:

1. **Unimodal versus Multimodal Fitness Functions.** Two different types of functions were tested, namely the sphere function in Section 5.1.1 and the Schwefel function in Section 5.1.2. While sphere is a unimodal function with just one local optimum (the global optimum), Schwefel is a multimodal function that contains a high number of local optima.
2. **Fitness-Based versus Similarity-Based Mating Preferences.** Both cases are explored in Section 5.1.1 and Section 5.1.2.
3. **Traditional (Random) versus Advanced (Best-First, Best-Last, and Self-Adaptive) Mating Strategies.**
4. **Varying Mating Sizes.** The range of explored values is $\gamma \in \{3, 5, 10, 20, 30\}$.

5.1.1 Binary Variables: Mating Size Experiments for the Sphere Function

The sphere function is defined as follows:

$$f_1(x) = \sum_{i=1}^n x_i^2.$$

This function has a minimum at $(x_1 = 0, \dots, x_n = 0)$ whose value is 0. The experiments were designed for $n = 20$, $-10 \leq x_i \leq 10$ for all $i \in \{1, \dots, 20\}$, and 10 bits were used to represent each variable; as a result, chromosomes with 200 genes were employed. The population size selected was 100 individuals.

Figure 6 represents the evolution, generation by generation, of the mean best fitness for the sphere function when fitness-based mating strategies are used. One hundred runs were carried out for each experiment. Whereas the random mating strategy is depicted in the three graphs of Figure 6 for illustrative purposes, the rest of the mating strategies (best-first, best-last, and self-adaptive) are depicted in just one graph. For these three advanced strategies, experiments were performed for different mating size values: $\gamma \in \{3, 5, 10, 20, 30\}$.

The best-first mating strategy performs better than the random strategy, as shown in Figure 6(a). In general, the performance improvement obtained by best-first increases with the mating size. This behavior can also be observed in Figure 6(c) for the self-adaptive strategy, although it takes mating size $\gamma = 10$ to begin to see an improvement over the random strategy. From Figure 6(b), it is clear that the best-last mating strategy performs worse than the traditional random strategy in the case of the sphere function. This behavior gets worse as the mating size increases.

Figure 7 shows the evolution of the mean best fitness for the sphere function when similarity-based mating strategies are utilized. In general, the results for similarity-based mating follow a similar pattern to those in Figure 6 for fitness-based mating. However, in the case of the similarity-based best-first strategy, more generations are needed in order to outperform random mating, as shown in Figure 7(a). This is a disadvantage of similarity-based best-first compared to fitness-based best-first in the case of the sphere function.

5.1.2 Binary Variables: Mating Size Experiments for the Schwefel Function

The Schwefel function (Schwefel, 1981, pp. 292–293) is defined in the following way:

$$f_2(x) = \sum_{i=1}^n x_i \cdot \sin(\sqrt{|x_i|}).$$

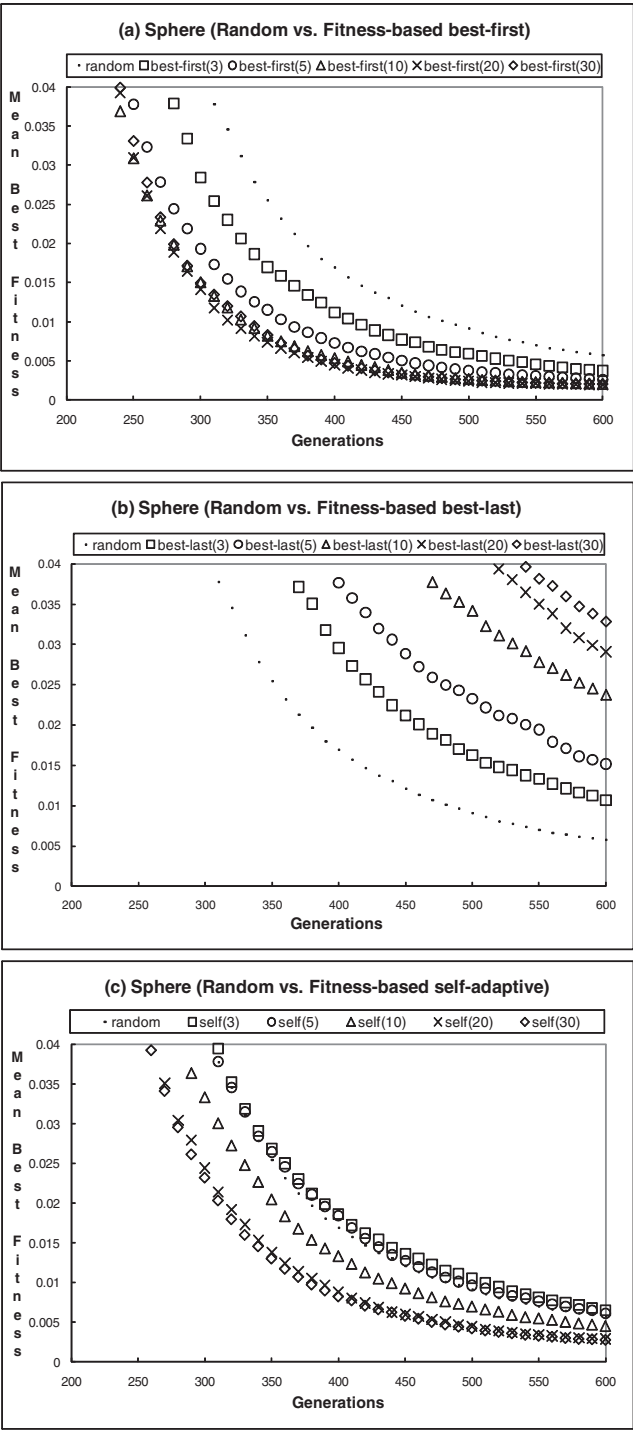


Figure 6: Mean best fitness results for the sphere minimization problem with binary-coded variables under the random, (a) the best-first, (b) the best-last, and (c) the self-adaptive fitness-based mating strategies. For the nonrandom strategies, the mating size varies from 3 to 30.

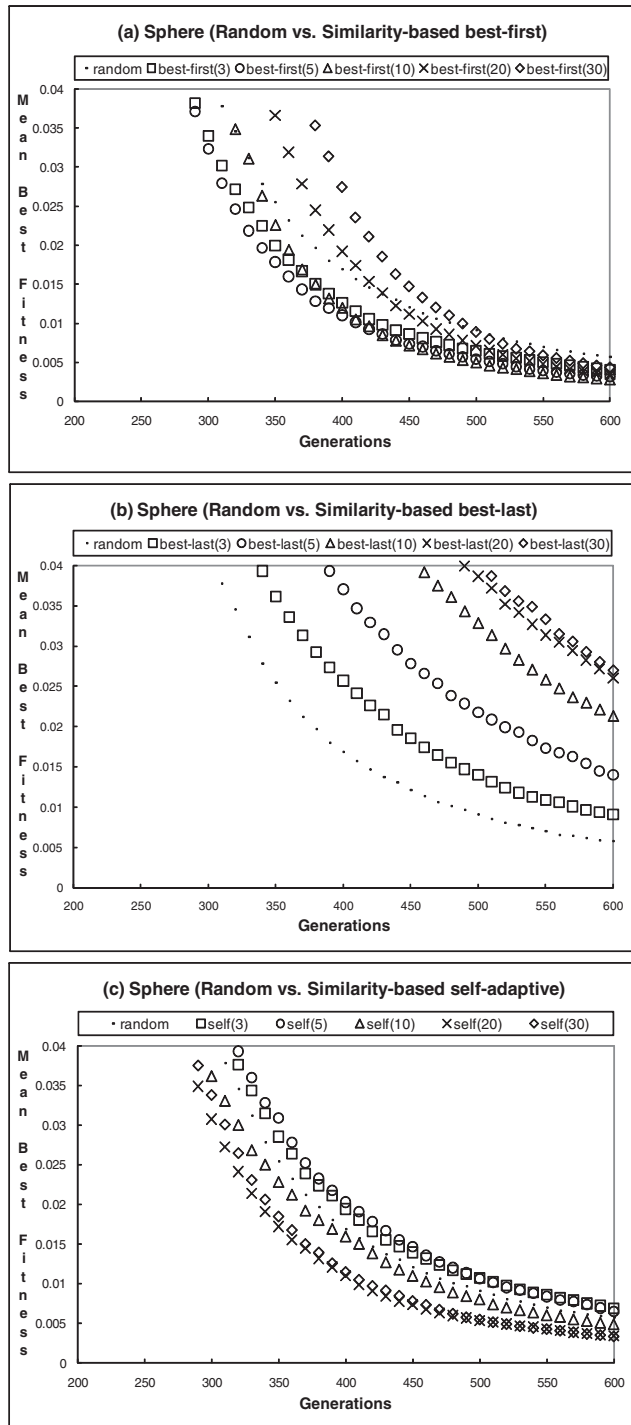


Figure 7: Mean best fitness results for the sphere minimization problem with binary-coded variables under the random, (a) the best-first, (b) the best-last, and (c) the self-adaptive similarity-based mating strategies. For the nonrandom strategies, the mating size varies from 3 to 30.

This function has a maximum at $(x_1 = 420.9687, \dots, x_n = 420.9687)$ whose value is $n \cdot 418.9829$. The experiments in this section were designed for $n = 10$, $-500 \leq x_i \leq 500$ for all $i \in \{1, \dots, 10\}$, and 100 bits were used to represent each variable; consequently, chromosomes with 1000 genes were created. The population size was 100 individuals. Due to the complexity of the Schwefel function, 500 runs were performed for each experiment.

Figure 8 depicts the evolution of the mean best fitness for the Schwefel function and fitness-based mating strategies. The opposite performance to that of the sphere function is obtained for best-last and best-first with respect to random mating. Firstly, Figure 8(b) shows that the best-last mating strategy performs better than the random strategy in the case of the Schwefel function. In general, the performance improvement obtained by best-last increases with the mating size. Secondly, the best-first mating strategy performs worse than the random strategy, as shown in Figure 8(a). This behavior gets worse as the mating size increases. On the other hand, and in contrast to the sphere function shown in Figure 6(c), fitness-based self-adaptive mating for the Schwefel function behaves worse as mating size grows. As depicted in Figure 8(c), although $\gamma \in \{3, 5\}$ outperforms random mating, that is not the case for $\gamma \in \{10, 20, 30\}$.

Figure 9 contains the evolution of the mean best fitness for the Schwefel function when similarity-based mating strategies are used. In general, the results for similarity-based mating follow a similar pattern to those in Figure 8 for fitness-based mating. However, in the case of similarity-based strategies, the outperformance of both the best-last and the self-adaptive strategies compared to the random strategy is superior to that obtained for their fitness-based counterparts. This represents an advantage of similarity-based strategies over fitness-based strategies in the case of the Schwefel function.

5.2 Binary Variables: Mating Index Experiments

The present section empirically compares random mating and best- α th mating. In general, the greater α is, the higher the degree of exploration is. The following comparisons are progressively made throughout the rest of this section:

- 1 **Unimodal versus Multimodal Fitness Functions.** Besides the sphere and the Schwefel functions, the Rastrigin function is now evaluated. Regarding multimodality, the Rastrigin function lies between the sphere and the Schwefel functions.
- 2 **Fitness-Based versus Similarity-Based Mating.**
- 3 **Traditional (Random) versus Advanced (Best- α th) Mating Strategies.**
- 4 **Varying Mating Indices.** The range of explored values is $\alpha \in \{2, 5, 10, 15, 20\}$ for a constant mating size of value $\gamma = 20$.

The Rastrigin function (Mühlenbein et al., 1991) is defined as follows:

$$f_3(x) = 10 \cdot n + \sum_{i=1}^n (x_i^2 - 10 \cdot \cos(2 \cdot \pi \cdot x_i)).$$

This function has a minimum at $(x_1 = 0, \dots, x_n = 0)$ whose value is 0. The Rastrigin function is based on the sphere function with the addition of a cosine modulation to produce many local optima; therefore, in terms of multimodality, the Rastrigin function

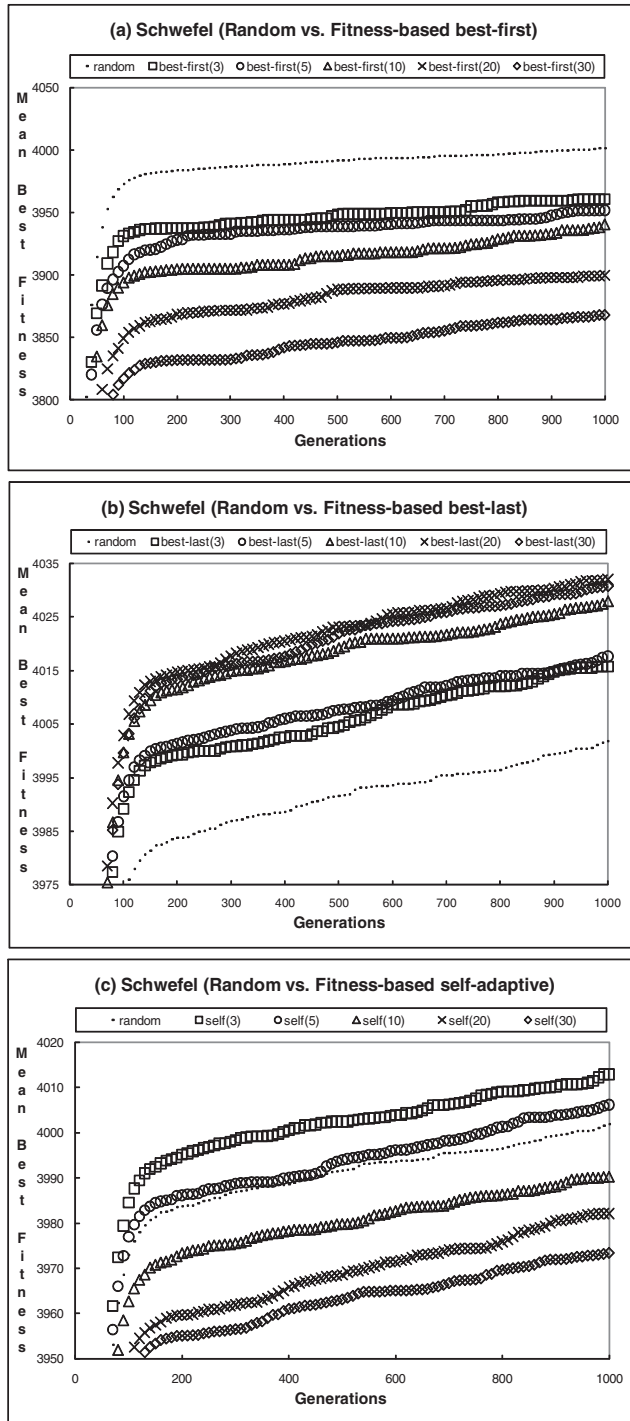


Figure 8: Mean best fitness results for the Schwefel maximization problem with binary-coded variables under the random, (a) the best-first, (b) the best-last, and (c) the self-adaptive fitness-based mating strategies. For the nonrandom strategies, the mating size varies from 3 to 30.

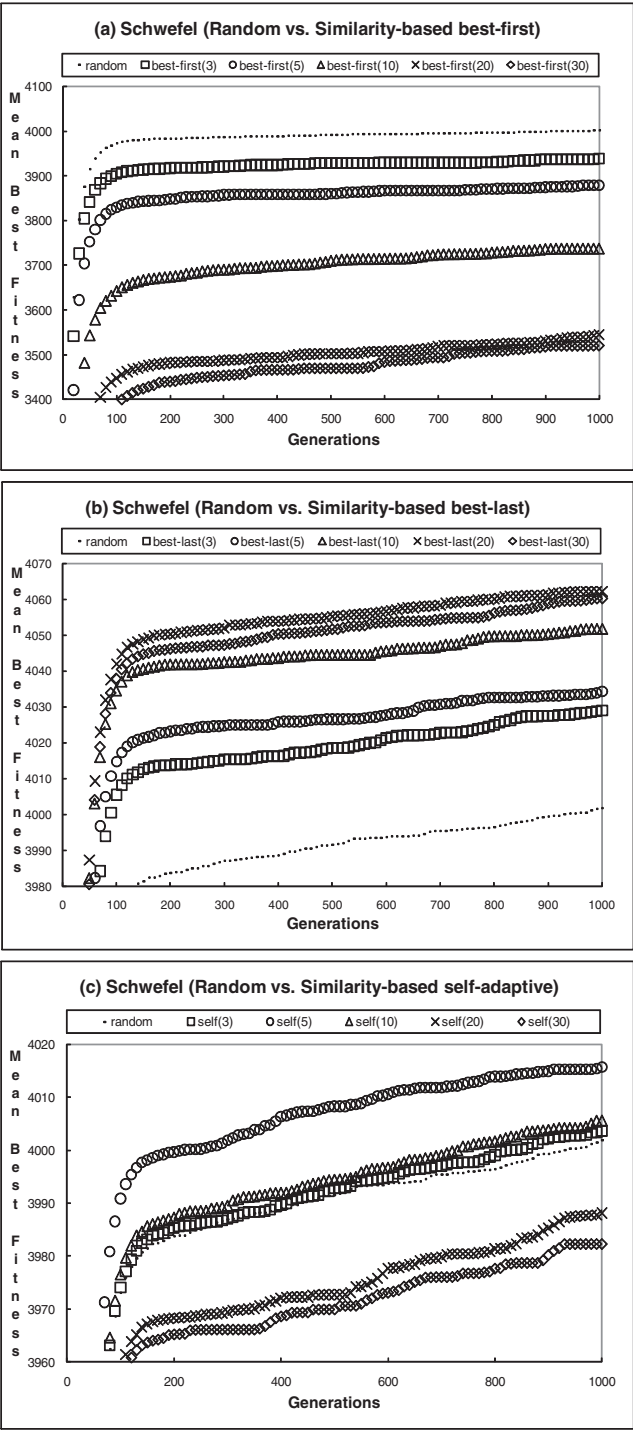


Figure 9: Mean best fitness results for the Schwefel maximization problem with binary-coded variables under the random, (a) the best-first, (b) the best-last, and (c) the self-adaptive similarity-based mating strategies. For the nonrandom strategies, the mating size varies from 3 to 30.

Table 2: Parameters used in the experiments of Section 5.2.

Parameter/function	Sphere	Rastrigin	Schwefel
Dimensions	10	5	5
Bits per dimension	10	10	20
Chromosome length	100	50	100
Runs*	100	250 for similarity criterion 500 for fitness criterion	
Generations	400	400	250
Individuals	50	50	50

*For Rastrigin and Schwefel functions, the number of runs was 250 for the similarity criterion and 500 for the fitness criterion.

is more complex than the sphere function and less complex than the Schwefel function.¹ The experiments were designed for $n = 5$, $-5.12 \leq x_i \leq 5.12$ for all $i \in \{1, \dots, 5\}$, and 10 bits were used to represent each variable; in this way, chromosomes with 50 binary genes were generated. Table 2 summarizes the parameters used in these experiments for the three functions considered: sphere, Rastrigin, and Schwefel.

Figure 10 illustrates the evolution of the mean best fitness for the three functions under both similarity-based and fitness-based criteria. For the best- α th strategies, the mating size is kept constant ($\gamma = 20$), while the mating index is assigned values $\alpha \in \{2, 5, 10, 15, 20\}$.

The following results can be derived from Figure 10:

- 1 As depicted in Figure 10(a), lower mating indices produce better performance in the case of the sphere function, where random mating is outperformed by best- α th when $\alpha \leq 10$. In principle, it could seem paradoxical that $\alpha = 2$ behaves worse than $\alpha = \{5, 10\}$ in Figure 10(a), but an explanation for this effect is offered in Point 2 of Section 4.1.2.
- 2 Figure 10(c) shows that higher mating indices lead to better performance for the Schwefel function, and random mating is outperformed by best- α th when $\alpha > 10$. This is the opposite behavior to that observed for the sphere function.
- 3 Figure 10(b) includes an interesting result. While the middle mating indices ($\alpha \in \{5, 10\}$) outperform random mating, both the lowest ($\alpha \in \{2\}$) and the highest ($\alpha \in \{15, 20\}$) mating indices have poorer performance than random mating. In other words, in the case of the Rastrigin function as defined in this section, neither a pure explorative strategy ($\alpha \sim \gamma$) nor a pure exploitative one ($\alpha \sim 2$) are good options. The mating index parameter provides us with intermediate strategies ($\alpha \sim \frac{\gamma}{2}$) that lead to the best performance. This is a novel result from this work.
- 4 Figures 12(d), 12(e), and 12(f) for fitness-based strategies show qualitatively similar results to those just described for similarity-based strategies in Points 1–3.

¹In the Schwefel function, the global maximum is distant from the next best local maximum, which makes convergence frequently take place in the wrong direction.

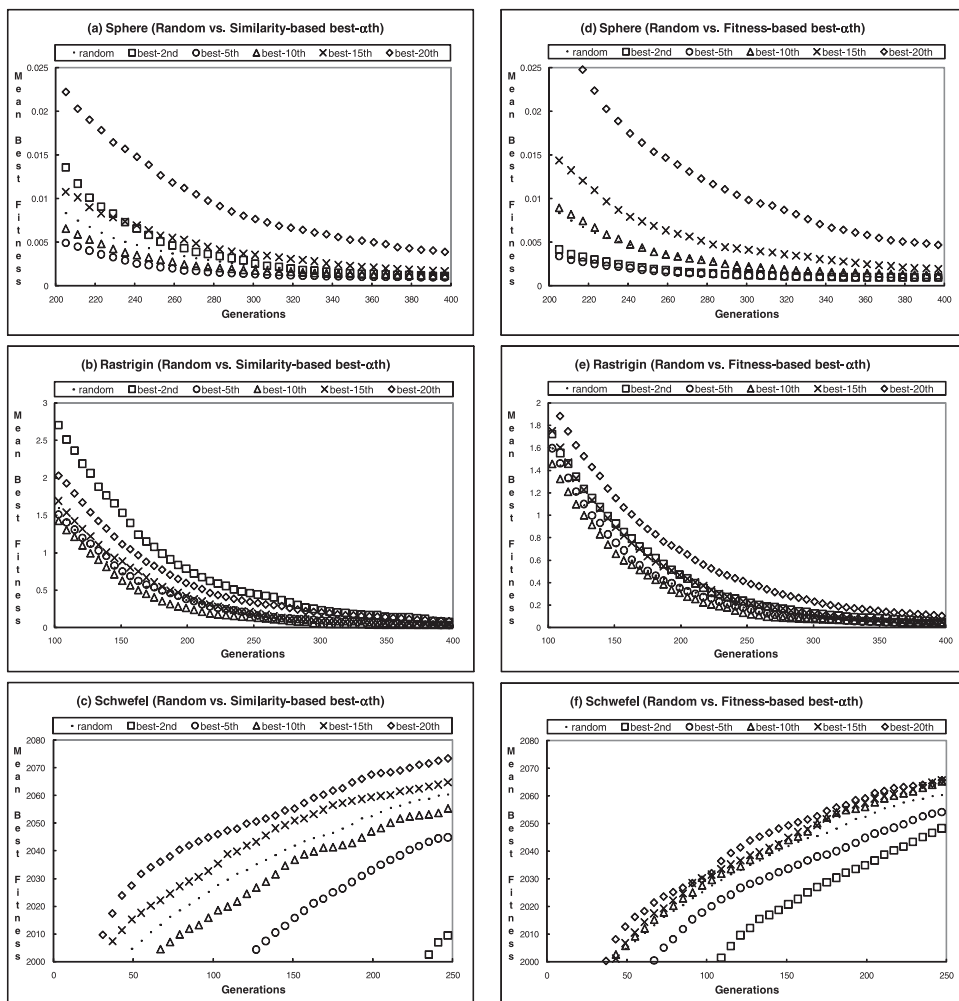


Figure 10: Mean best fitness results for the sphere, the Rastrigin, and the Schwefel optimization problems with binary-coded variables under the random and the (similarity-based and fitness-based) best- α th mating strategies. For the nonrandom strategies, the mating size is equal to 20 and the mating index varies from 2 to 20.

5.3 Real Variables: Mating Size Experiments

The aim of this section and Section 5.4 is to determine whether the qualitative experimental results obtained for mating in Sections 5.1 and 5.2 are also valid when real-coded variables are used within chromosomes. Since the representation of individuals is now based on n real variables, where n is the dimension of the function being optimized, new crossover and mutation operators need to be considered. We implemented crossover as discrete recombination: If an offspring z is created from parents x and y , the allele value (real number) for gene i , z_i , is taken from $\{x_i, y_i\}$ with equal probability. We used a mutation operator that selects uniformly at random one of the n genes (or variables) and adds to its real value an amount drawn randomly from a Gaussian distribution with mean zero and given standard deviation (0.5 in our case).

Figures 11 through 14 depict, for real-coded variables, the same type of experiments for the sphere and Schwefel functions that are shown in Figures 6 through 9, respectively, for binary-coded variables. As we will see below, the use of real variables produces similar qualitative results for how the different mating strategies behave.

5.4 Real Variables: Mating Index Experiments

Like Section 5.3, this section presents experimental results for different mating strategies when real-coded variables are used in the chromosome. Specifically, Figure 15 shows for read-coded variables and the sphere, the Rastrigin, and the Schwefel functions the same type of experiments appearing in Figure 10 for binary-coded variables. Under real representation, the behavior of the different mating strategies evaluated in Figure 15 is qualitatively comparable to that observed in Figure 10 under binary representation.

6 Discussion of the Experiments

This section discusses the influence of mating on the performance of a simple GA applied to real-function optimization. The results obtained in Sections 5.1 through 5.4 demonstrate that, as long as mating preferences are defined in the phenotype space, the genotype has no significant influence on the qualitative behavior of mating. For the real-function optimization problem addressed in this work, we always consider the Euclidean distance as the similarity measure between two individuals, regardless of whether binary or real representation is being used. Therefore, for the sake of simplicity, in the rest of this section we will refer to the experiments using binary-coded variables.

Before discussing the experimental results obtained in Section 5, the statistical significance of these results needs to be established. In other words, for each figure in Section 5 (from Figure 6[a]) to Figure 15[f]), it has to be determined whether or not the six graphs in the figure come from the same distribution and, therefore, the differences are due to random effects (null hypothesis). ANOVA tests were performed on the best-fitness data for the last generation depicted in each of the six graphs contained in Figures 6(a) through 15(f). For each graph, 100 samples were taken. Table 3 includes the results for ANOVA tests on groups of samples constituted by mean best fitness values under the {random, $\gamma = 3$, $\gamma = 5$, $\gamma = 10$, $\gamma = 20$, $\gamma = 30$ } mating strategies (Figures 6[a] through 9[c] in Section 5.1 and Figures 11[a] through 14[c] in Section 5.3) and on groups of samples constituted by mean best fitness values under the {random, $\alpha = 2$, $\alpha = 5$, $\alpha = 10$, $\alpha = 15$, $\alpha = 20$ } mating strategies (Figures 10[a] through 10[f] in Section 5.2 and Figures 15[a] through 15[f] in Section 5.4). In addition to F value, the probability p of the corresponding result assuming the null hypothesis is shown. The values for p in Table 3 can be considered as significant enough to discard the null hypothesis for the data obtained, suggesting that mating has an important influence on GA performance.

6.1 Discussion of the Mating Size Experiments

The experimental results obtained in Section 5.1 for the best-first, the best-last, and the self-adaptive mating strategies suggest that best-first mating is the best option for unimodal problems, while best-last mating is the best option for highly multimodal problems. When the degree of multimodality is unknown, the self-adaptive mating approach behaves differently depending on the problem at hand. In unimodal problems, self-adaptive mating clearly outperforms random mating as γ increases. In multimodal problems, self-adaptive mating produces better results than random mating for middle and low γ values. For mating size values in the range $\gamma \in [3, 10]$, self-adaptive mating

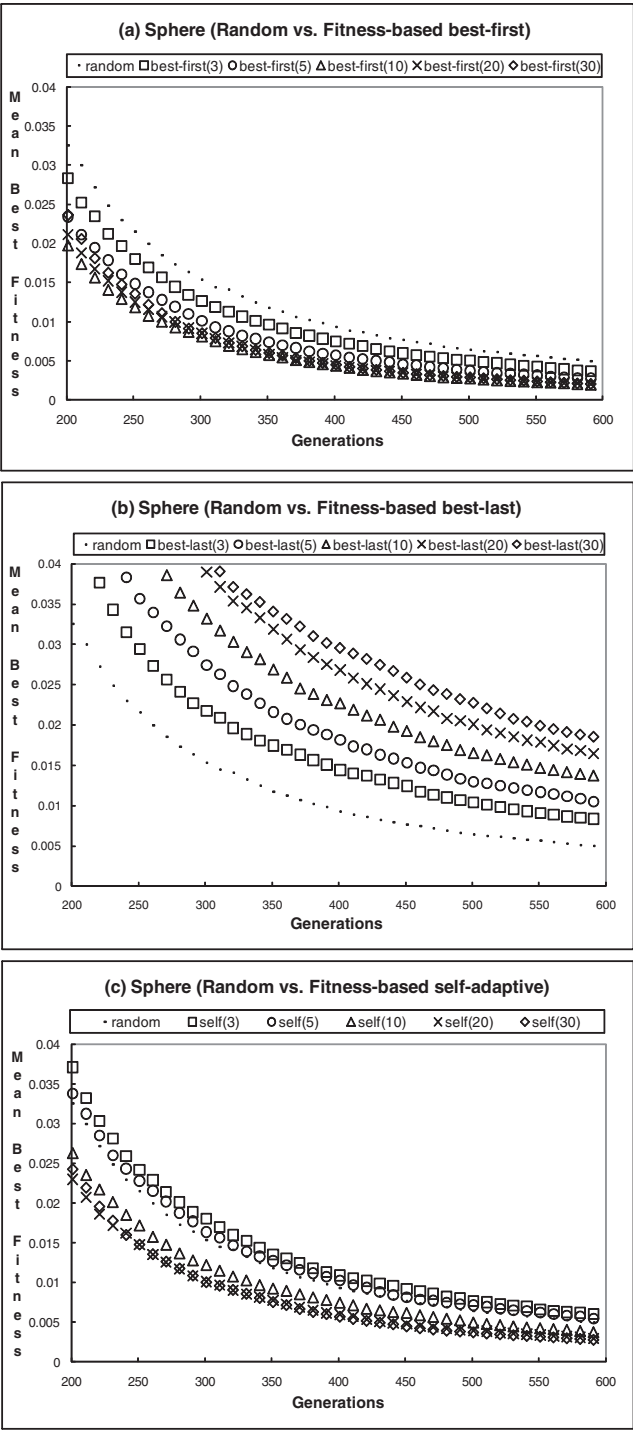


Figure 11: Mean best fitness results for the sphere minimization problem with real-coded variables under the random, (a) the best-first, (b) the best-last, and (c) the self-adaptive fitness-based mating strategies. For the nonrandom strategies, the mating size varies from 3 to 30.

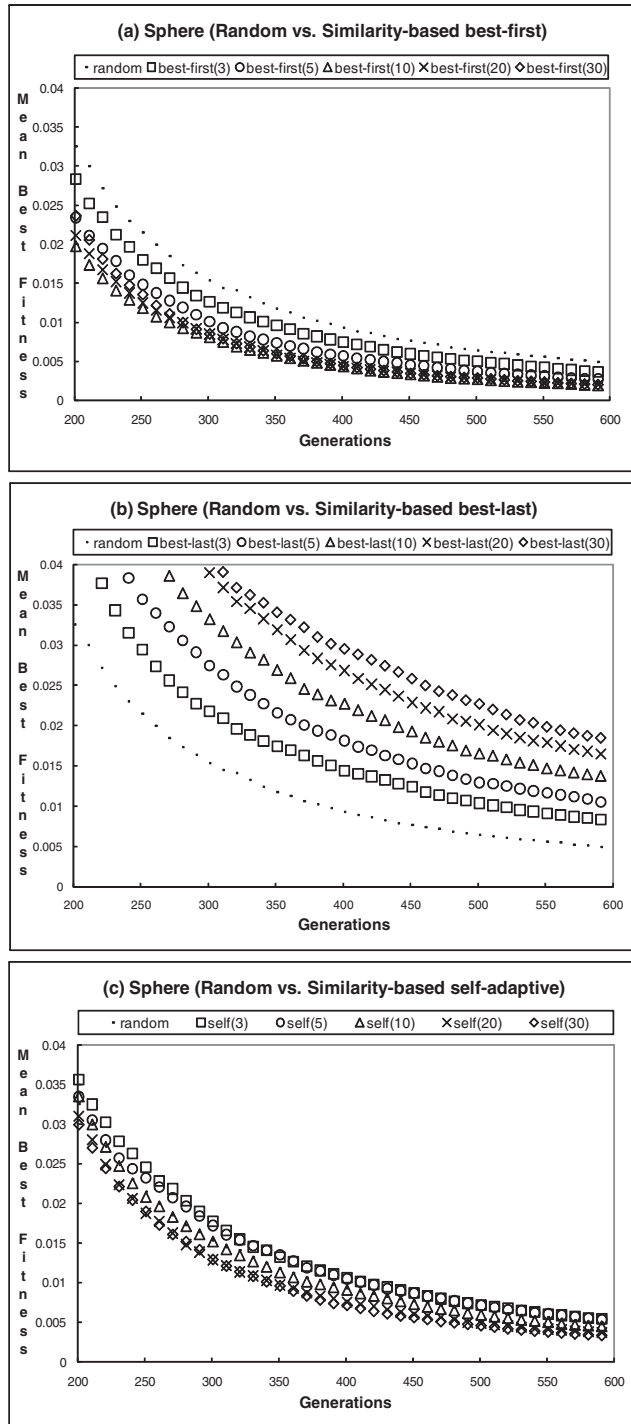


Figure 12: Mean best fitness results for the sphere minimization problem with real-coded variables under the random, (a) the best-first, (b) the best-last, and (c) the self-adaptive similarity-based mating strategies. For the nonrandom strategies, the mating size varies from 3 to 30.

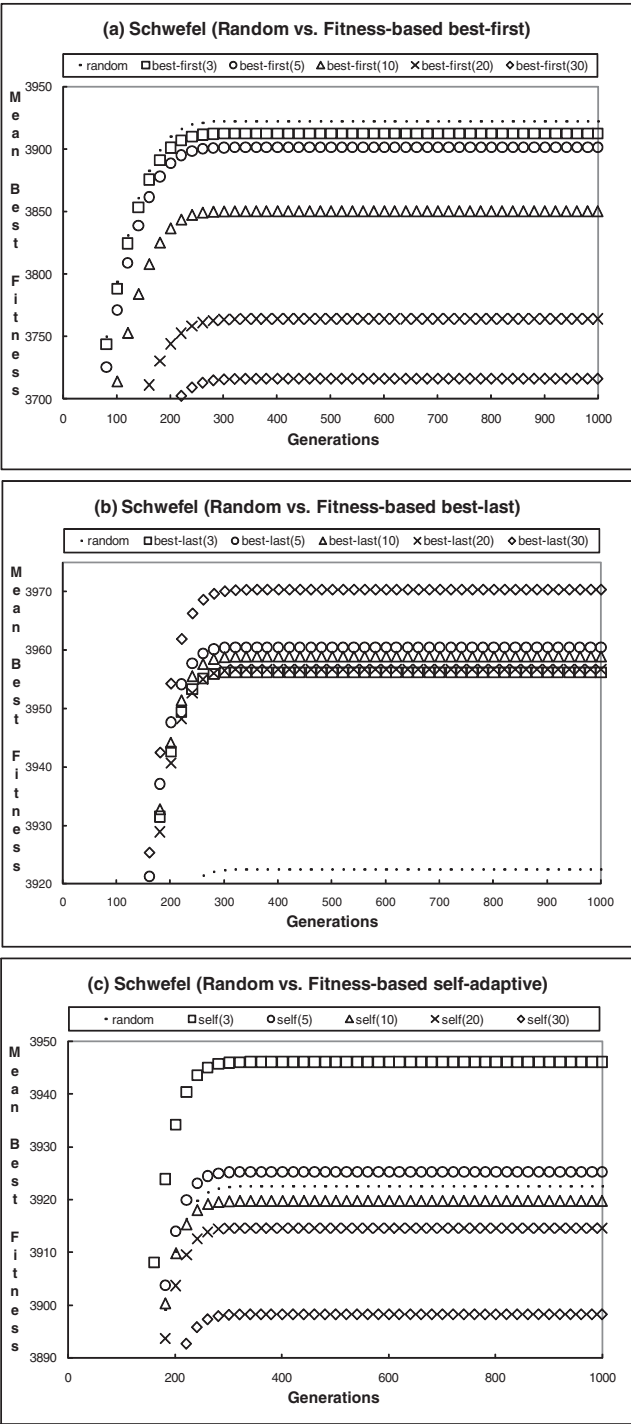


Figure 13: Mean best fitness results for the Schwefel maximization problem with real-coded variables under the random, (a) the best-first, (b) the best-last, and (c) the self-adaptive fitness-based mating strategies. For the nonrandom strategies, the mating size varies from 3 to 30.

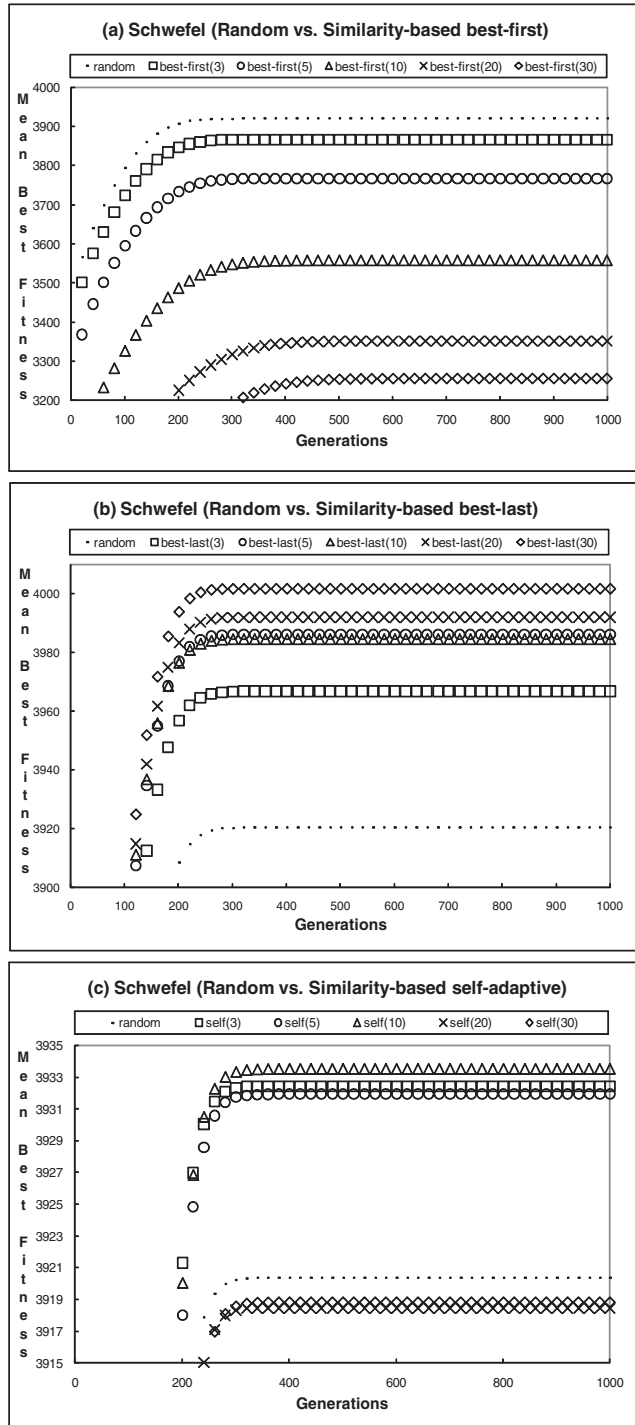


Figure 14: Mean best fitness results for the Schwefel maximization problem with real-coded variables under the random, (a) the best-first, (b) the best-last, and (c) the self-adaptive similarity-based mating strategies. For the nonrandom strategies, the mating size varies from 3 to 30.

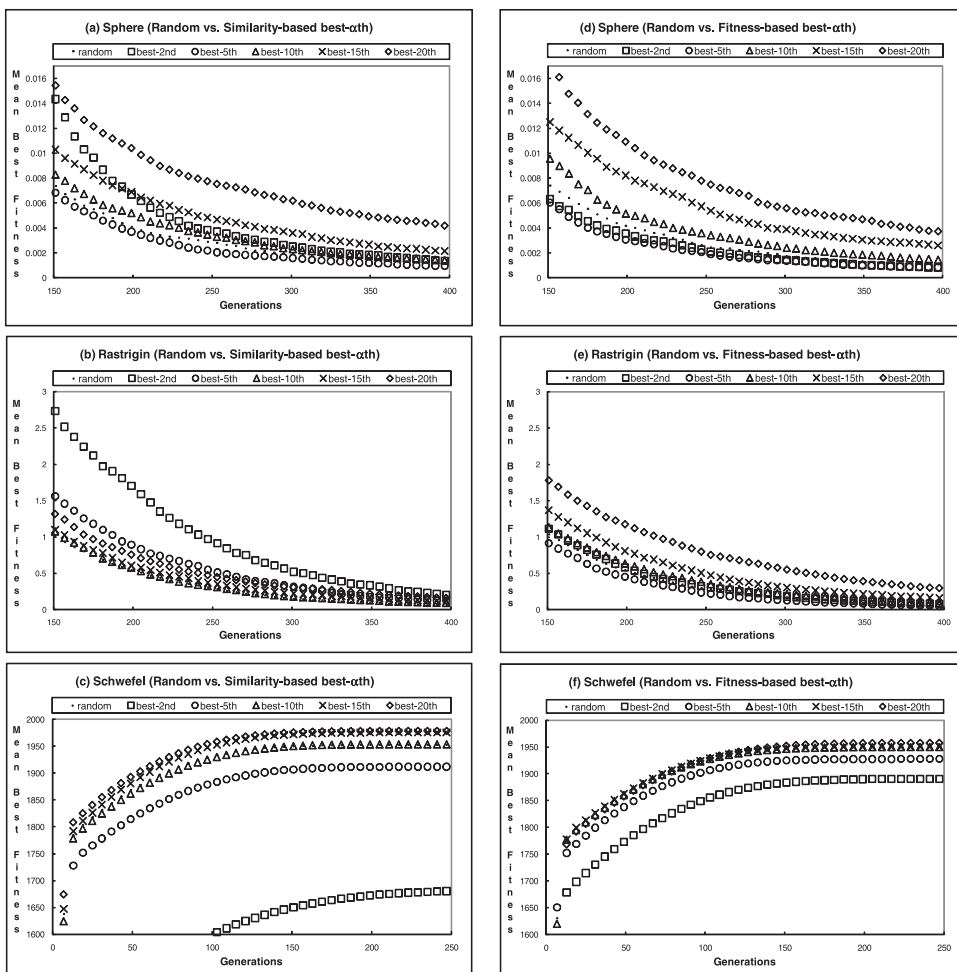


Figure 15: Mean best fitness results for the sphere, the Rastrigin, and the Schwefel optimization problems with real-coded variables under the random and the (similarity-based and fitness-based) best- α th mating strategies. For the nonrandom strategies, the mating size is equal to 20 and the mating index varies from 2 to 20.

performs better than (or at least comparably to) random mating over the experiments in Section 5.1. However, for multimodal problems, self-adaptive mating does not offer a good robust behavior for all the mating sizes.

An example of how self-adaptive mating works is shown in Figure 16. This figure depicts the mean α value in the population, generation by generation, for the two cases in which the self-adaptive strategy had the best behavior in Section 5.1: fitness-based mating for the sphere function, and similarity-based mating for the Schwefel function. The results are averaged over 100 runs for the sphere function and over 500 runs for the Schwefel function. In both cases, the mating size was assigned value $\gamma = 20$. Only the results for the first 100 generations are depicted, since this is the range in which significant changes are obtained for the mean mating index. From Figure 16, the population mean mating index value in the initial population (approximately $\alpha = 11$) decreases rapidly

Table 3: ANOVA test results for the experiments in Section 5. When $F \geq F_{\text{critical}}$, the result is significant at the level of probability $p = .05$.

Figure	Function	F	F_{critical}	p
8(a)	Sphere	148.367	2.229	4.704E-102
8(b)	Sphere	95.771	2.229	6.4498E-74
8(c)	Sphere	88.373	2.229	2.019E-69
9(a)	Sphere	10.373	2.229	1.4855E-9
9(b)	Sphere	40.309	2.229	1.0403E-35
9(c)	Sphere	5.112	2.229	1.3378E-4
10(a)	Schwefel	10.399	2.229	1.4042E-9
10(b)	Schwefel	2.367	2.229	0.03838
10(c)	Schwefel	3.128	2.229	0.00849
11(a)	Schwefel	80.781	2.229	1.2221E-64
11(b)	Schwefel	2.929	2.229	0.01269
11(c)	Schwefel	2.505	2.229	0.02935
12(a)	Sphere	81.266	2.229	5.9728E-65
12(b)	Rastrigin	2.273	2.229	0.04596
12(c)	Schwefel	12.158	2.229	3.1114E-11
12(d)	Sphere	90.772	2.229	6.7666E-71
12(e)	Rastrigin	2.784	2.229	0.01695
12(f)	Schwefel	3.933	2.229	0.00162
13(a)	Sphere	132.875	2.229	2.2126E-94
13(b)	Sphere	113.289	2.229	5.5539E-84
13(c)	Sphere	71.565	2.229	1.3731E-58
14(a)	Sphere	49.813	2.229	4.2947E-43
14(b)	Sphere	175.26	2.229	2.226E-114
14(c)	Sphere	51.582	2.229	2.0078E-44
15(a)	Schwefel	7.404	2.229	9.4783E-7
15(b)	Schwefel	2.468	2.229	0.0315
15(c)	Schwefel	3.089	2.229	0.0092
16(a)	Schwefel	118.977	2.229	4.3373E-87
16(b)	Schwefel	2.295	2.229	0.0441
16(c)	Schwefel	2.326	2.229	0.0415
17(a)	Sphere	100.58	2.229	9.2883E-77
17(b)	Rastrigin	119.007	2.229	4.1807E-87
17(c)	Schwefel	38.55	2.229	2.6903E-34
17(d)	Sphere	112.609	2.229	1.3215E-83
17(e)	Rastrigin	49.542	2.229	6.8876E-43
17(f)	Schwefel	5.728	2.229	3.5761E-5

for the sphere function. This is in accordance with the experimental results obtained in Section 5.1.1, which showed that best-first (where α is small) produces the best results for the sphere function. On the other hand, the mean mating index for the Schwefel function is greater than that for the sphere function throughout the generations. This is what should be expected taking into account that, from Section 5.1.2, best-last (where α is large) produces the best results for the Schwefel function. However, due to the fact that individuals with a high mating index may produce lethal individuals after recombination in the case of the Schwefel function, the graph for this function is not as close to $\alpha = 20$ as the graph for the sphere function is to $\alpha = 2$. (In principle, the lower graph in Figure 16 should be close to $\alpha = 2$; however, note that $\alpha = 2$ and $\alpha = 5$ produce similar results in Figure 12(d). Therefore, a convergence value of 7.5 in Figure 16 is quite close to the optimal behavior.) This explains the results obtained in Section 5.1, in which self-adaptive mating applied to the sphere function produced similar results to

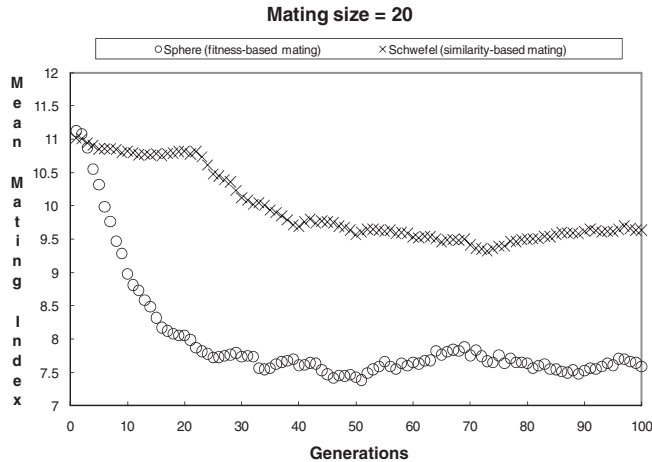


Figure 16: Mean α value in the population for the sphere and the Schwefel functions under the self-adaptive strategy with $\gamma = 20$.

those obtained through the best-first strategy, while self-adaptive mating applied to the Schwefel function could not reach the good results produced by the best-last strategy.

While fitness-based strategies produce better results than similarity-based strategies for unimodal problems like the sphere function optimization, similarity-based strategies outperform fitness-based strategies in the case of multimodal problems like the Schwefel function optimization. However, an advantage of fitness-based strategies is that they lead to computation time savings as shown in Table 4. The differences in computation times between fitness-based and similarity-based mating are due to the fact that similarity comparisons operate on long binary strings. These differences are smaller in the case of real-coded variables.

6.2 Discussion of the Mating Index Experiments

The mating index experiments in Section 5.2 show an important result derived from the present work. As a consequence of parameterizing the degree of exploration applied in the mating phase, new mating strategies have arisen apart from the traditional random one and apart from the pure exploratory/exploitative strategies also found in the literature. The introduction of the mating index α enables us to define different degrees of exploration in a GA, and it turns out that some fitness landscapes can be better optimized by adopting intermediate exploratory strategies. To the best of our knowledge, such intermediate strategies have never before been discussed in the literature. Given an arbitrary fitness landscape, it remains to investigate how to calculate the most appropriate α value for the optimization problem.

7 Conclusion

Most of the existing approaches to mating in GAs apply restrictions based on similarity between individuals. The novel mating approach introduced in this work also considers an alternative fitness-based criterion for defining mating strategies, which is compared to the widespread similarity-based criterion. The fitness-based criterion offers

Table 4: Mean computation times (s) for each run of the experiments in Section 5.1, with t_f and t_s referring to fitness-based and similarity-based mating, respectively. The number of function evaluations is the same for all cases in this table, number of generations \times population size, due to the fact that establishing the fittest parent in set Ch of Algorithm 1 requires all the individuals in the current population to be evaluated previously to mating.

		Sphere			Schwefel		
		t_f	t_s	t_s/t_f	t_f	t_s	t_s/t_f
Best-first	$\gamma = 3$	7.37	22.05	2.99	37.94	111.32	2.93
	$\gamma = 5$	7.23	36.85	5.1	38.39	198.82	5.18
	$\gamma = 10$	7.46	67.15	9	38.79	369.9	9.54
	$\gamma = 20$	7.62	115.94	15.21	38.74	680.56	17.57
	$\gamma = 30$	7.45	154.87	20.79	38.29	944.99	24.68
Best-last	$\gamma = 3$	7.5	21	2.8	38.24	123.79	3.24
	$\gamma = 5$	7.42	36.86	4.97	37.98	203.7	5.36
	$\gamma = 10$	7.57	66.16	8.74	37.86	389.39	10.28
	$\gamma = 20$	7.63	124.88	16.37	35.88	687.93	19.17
	$\gamma = 30$	7.87	166.38	21.14	36.33	956.08	26.32
Self-adaptive	$\gamma = 3$	8.81	26.86	3.05	45.32	129.42	2.86
	$\gamma = 5$	8.2	57.82	7.05	43.8	298.66	6.82
	$\gamma = 10$	8.89	203.54	22.89	43.05	1,056.43	24.54
	$\gamma = 20$	9.06	675.39	74.55	43.91	3,636.85	82.82
	$\gamma = 30$	9.38	1,304.7	139.09	44.31	7,235.56	163.29

computation time savings and, in cases like unimodal function optimization, greater efficiency to approach the optimum in fewer generations.

An important group of mating methods for GAs, for instance, assortative mating (Fernandes et al., 2001; Huang, 2001; Ochoa et al., 2005), use mating strategies that select just the most similar or the most dissimilar individual from a set of candidates. In our novel approach, a parameter called the mating index (see Section 3) allows any of the candidates to be chosen. In this way, if a similarity-based criterion is considered, a candidate with an arbitrary degree of similarity can be obtained or, if a fitness-based criterion is considered, a candidate with an arbitrary fitness can be selected. Therefore, a wide spectrum of mating strategies can be investigated by varying the mating index. Intermediate mating indices, which are supported by our approach but not by previous approaches, happen to be the best option in cases where either a pure exploratory strategy or a pure exploitative strategy yield poor performance.

Our novel mating approach also facilitates the definition of a self-adaptive mating strategy in which each individual has its own mating preference (or mating index). In this way, the fittest individuals carry with them the most successful mating strategies from generation to generation. While self-adaptive mating does not perform well on all problems we investigated, it is a good strategy in the case of unimodal problems and also in the case of multimodal problems when low or middle mating sizes are used. Strategies such as best-first and best-last greatly outperform random mating in different types of problems. While unimodal problems greatly benefit from using the best-first strategy, the same applies to multimodal problems in the case of the best-last strategy.

A future research topic is the exploration of a mating strategy that more deterministically controls the mating index parameter throughout the GA generations. In this way, although the mating index would be the same for every individual in the population,

it could change from generation to generation. A possible scheme consists of assigning $\alpha = \gamma$ at GA initialization and letting α be a monotonically nonincreasing function of the number of generations. (Both linear and nonlinear reduction schemes could be possible.) This deterministic scheme applies more exploration in the initial generations of the GA, when promising search areas are sought, and applies more exploitation in the final generations, when population diversity has decreased.

Another future research topic is the inclusion of the mating size parameter γ as a local parameter in the chromosome of each individual. The performance of the new self-adaptive strategy, resulting from including mating size γ along with mating index α as local parameters, should be compared to the strategies defined in this work.

Finally, we suggest as another important future research direction the application of the novel mating approach to other domains, in addition to those considered in this paper, using other evolutionary algorithm techniques.

Acknowledgments

Severino F. Galán was supported by the Spanish Ministry of Science and Innovation through grant JC2007-00110 from its José Castillejo Program.

References

- Booker, L. B. (1982). *Intelligent behaviour as an adaptation to the task environment*. PhD thesis, Department of Computer and Communication Science, University of Michigan, Ann Arbor.
- Chakraborty, G., and Chakraborty, B. (1999). Ideal marriage for fine tuning in GA. In *Proceedings of the 1999 IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC'99)*, pp. 631–636.
- Cohon, H. P., Hedge, S. U., Martin, W. N., and Richards, D. (1987). Punctuated equilibria: A parallel genetic algorithm. In *Proceedings of the 2nd International Conference on Genetic Algorithms (ICGA'87)*, pp. 148–154.
- De, S., Pal, S. K., and Ghosh, A. (1998). Genotypic and phenotypic assortative mating in genetic algorithm. *Information Sciences: An International Journal*, 105(1–4):209–226.
- Deb, K., and Goldberg, D. E. (1989). An investigation of niche and species formation in genetic function optimization. In *Proceedings of the 3rd International Conference on Genetic Algorithms (ICGA'89)*, pp. 42–50.
- Eiben, A. E., Hinterding, R., and Michalewicz, Z. (1999). Parameter control in evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 3(2):124–141.
- Eiben, A. E., and Smith, J. E. (2003). *Introduction to evolutionary computing*. Berlin: Springer.
- Eshelman, L. J., and Schaffer, J. D. (1991). Preventing premature convergence in genetic algorithms by preventing incest. In *Proceedings of the 4th International Conference on Genetic Algorithms (ICGA'91)*, pp. 115–122.
- Fernandes, C., Tavares, R., Munteanu, C., and Rosa, A. (2001). Using assortative mating in genetic algorithms for vector quantization problems. In *Proceedings of the 2001 ACM Symposium on Applied Computing (ACM SAC'2001)*, pp 361–365.
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Reading, MA: Addison-Wesley.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems*. 2nd ed. Ann Arbor, MI: University of Michigan Press.

- Huang, C. F. (2001). An analysis of mate selection in genetic algorithms. Technical Report CSCS-2001-002, Center for the Study of Complex Systems, University of Michigan.
- Manderick, B., and Spiessens, P. (1989). Fine-grained parallel genetic algorithms. In *Proceedings of the 3rd International Conference on Genetic Algorithms (ICGA'89)*, pp. 428–433.
- Martin, W. N., Lienig, J., and Cohoon, J. P. (1997). Island (migration) models: Evolutionary algorithms based on punctuated equilibria. In T. Bäck, D. B. Fogel, and Z. Michalewicz (Eds.), *The handbook of evolutionary computation* (pp. C6.3:1–16). London: Institute of Physics Publishing and Oxford University Press.
- Matsui, K. (1999). New selection method to improve the population diversity in genetic algorithms. In *Proceedings of the 1999 IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC'99)*, pp. 625–630.
- Mühlenbein, H., Schomisch, D., and Born, J. (1991). The parallel genetic algorithm as function optimizer. *Parallel Computing*, 17:619–632.
- Ochoa, G., Mädler-Kron, C., Rodriguez, R., and Jaffe, K. (2005). Assortative mating in genetic algorithms for dynamic problems. In F. Rothlauf, J. Branke, S. Cagnoni, D. W. Corne, R. Drechsler, Y. Jin, P. Machado, E. Marchiori, J. Romero, G. D. Smith, and G. Squillero (Eds.), *Applications of Evolutionary Computing*, Vol. 3449 (pp. 617–622). Berlin: Springer.
- Ronald, E. (1995). When selection meets seduction. In *Proceedings of the 6th International Conference on Genetic Algorithms (ICGA'95)*, pp. 167–173.
- Schwefel, H. P. (1981). *Numerical optimization of computer models*, Vol. 26. New York: Wiley.
- Smith, R. E., and Bonacina, C. (2003). Mating restriction and niching pressure: Results from agents and implications for general evolutionary computation. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2003)*, pp. 1382–1393.
- Smith, R. E., Bonacina, C., Kearney, P., and Merlat, W. (2000). Embodiment of evolutionary computation in general agents. *Evolutionary Computation*, 8(4):475–493.
- Spears, W. M. (1994). Simple subpopulation schemes. In *Proceedings of the 3rd Annual Evolutionary Programming Conference*, pp. 296–307.
- Ting, C. K., Li, S. T., and Lee, C. (2003). On the harmonious mating strategy through TABU search. *Information Sciences: An International Journal*, 156(3–4):189–214.
- Unemi, T., and Nagayoshi, M. (1997). Evolution of learning robot team via local mating strategy. Poster presented at the *Fourth European Conference on Artificial Life*, Brighton, UK.
- White, P. M., and Pettey, C. C. (1997). Double selection vs. single selection in diffusion model GAs. In *Proceedings of the 7th International Conference on Genetic Algorithms (ICGA'97)*, pp. 174–180.

