Analysis of Speciation and Niching in the
Multi-Niche Crowding GA

Walter Cedeño
HEWLETT-PACKARD
Research and Development
2850 Centerville Road MS-2C10
Wilmington, DE 19808-1610
(wcedeno@lf.hp.com)

V. Rao Vemuri
Department of Applied Science
University of California, Davis and
Lawrence Livermore National Laboratory
Livermore, CA 94550
(vemuri1@llnl.gov)

Abstract

The *multi-niche crowding* genetic algorithm (MNC GA) has demonstrated its ability to maintain population diversity and stable subpopulations while allowing different species to evolve naturally in different niches of the fitness landscape. These properties are a consequence, in part, to the effect of *crowding selection* and *worst among most similar replacement* genetic operators. In this paper we take a closer look at these genetic operators and present mathematical results that show their effect on the population when used in the MNC GA. We also present some guidelines about the parameter values to use in these genetic operators to achieve the desired niching pressure during a run. We conclude with a list of unexplored avenues that might be helpful in a future analysis of the behavior of the MNC GA.

Keywords: Genetic algorithms, multimodal functions, niching, speciation.

1. Introduction

In this paper we analyse the influence of the selection and replacement operators used by the multi-niche crowding (MNC) genetic algorithm (Cedeño, 1995; Cedeño and Vemuri, 1992, 1996; Cedeño et al, 1996). The purpose of this analysis is to gain a deeper understanding of the niching capabilities of MNC GA as influenced by the parameters of these operators. Specifically, we show how the "crowding selection" operator promotes mating among individuals from the same niche. The analysis provides guidance on choosing the appropriate "crowding selection size" parameter in order to achieve the desired "mating pressure" during the selection operation. Similarly, we show how the "WAMS (worst among most similar) replacement" operator

promotes the replacement of individuals by offspring from the same niche, thus applying "fitness pressure" on members from a given niche. The performance of the WAMS replacement operator is moderated by the "crowding group size" and "crowding factor" parameters. Selection of the appropriate values for these parameters is crucial in promoting niching and therefore the survival of the fittest.

Section 2 begins with a brief overview of the MNC GA along with an introduction to the selection and replacement operators. Section 2 concludes with a brief summary of the historical context surrounding the development of this algorithm. Section 3 introduces the notation and in Section 4 we analyse the crowding selection operator and develop a mathematical expression for the probability that any two individuals in the population are selected for mating. The WAMS replacement operator is analysed in Section 5 and the probability of an individual being replaced by an offspring is calculated. In Section 6, the crowding selection operation is treated in depth and mathematical expressions for the expected value and variance of the similarity rank are derived. Section 7 develops an analogous analysis for WAMS replacement. In Section 8 an attempt is made to compare these analytical results with some empirical results, and Section 9 summarises the conclusions.

2. Overview of the MNC GA

There are many versions of genetic algorithms. Here, we only concern ourselves with the analysis of a steady state genetic algorithm (Whitley, 1988; Syswerda, 1989), namely the multi-niche crowding genetic algorithm (MNC GA). In a nutshell, all steady state genetic algorithms have three basic steps: selection, recombination (or, reproduction) and replacement. During the *selection step*, a decision is made as to who, in the population, is allowed to produce offspring. During the *recombination step,* offspring are produced via the operations of crossover and mutation. During the *replacement step* another decision is made as to which of the members in the current population are forced to perish (or vacate a slot) in order to make room for an offspring to compete (or, occupy a slot) in the next iteration. These steps are applied until a suitable condition is satisfied, say, the number of function evaluations. Various versions of steady state GA's differ from each other in the details of how these steps are implemented.

In the MNC GA both the selection and replacement steps are modified by some type of crowding (De Jong, 1975). The idea is to ameliorate the selection pressure caused by fitness proportionate reproduction (FPR) (Holland, 1975) and allow the population to maintain diversity throughout the search. This objective is achieved in part by encouraging mating and replacement within members of the same niche while allowing some competition for the population slots among the niches. The result is an algorithm that:

1. maintains stable subpopulations within different niches,
2. maintains diversity throughout the search, and
3. converges to multiple local optima.

No prior knowledge of the search space is needed and no restrictions are imposed during selection and replacement thus allowing exploration of other areas of the search space while converging to the best individuals in different niches.

In MNC, the FPR selection is replaced by what we call *crowding selection*. In crowding selection each individual in the population has the same chance for mating in every generation. Application of this selection rule is done in two steps. First, a parent I_i is selected for mating. This selection can be either sequential or random. Second, its mate I_j is selected, not from the entire population, but from a small group of individuals of size *s* (crowding selection group size). The individuals in the crowding selection group are picked uniformly at random (with replacement) from the population. The mate I_j thus chosen must be the one who is the most "similar" to I_i. The similarity metric used here is not a genotypic metric such as the Hamming distance, but a suitably defined phenotypic distance metric.

Crowding selection promotes mating between members of the same niche while still allowing individuals from different niches to mate. Unlike mating restriction (Deb and Goldberg, 1989) that only allows individuals from the same niche to mate, crowding selection allows some amount of exploration to occur while at the same time exploiting the similarity between individuals of a niche.

During the replacement step the MNC GA uses a replacement policy called *worst among most similar* (WAMS). The goal of this step is to pick an individual from the population for replacement by offspring. Implementation of this policy follows these steps. First, *f* "crowding

groups" are created by randomly picking *g* (crowding group size) individuals per group from the population. Second, one individual from each group that is most similar to the offspring is identified and placed in the "crowding factor group". This gives *f* individuals that are candidates for replacement by virtue of their similarity to the offspring. The offspring will replace one of them. From the crowding factor group of most similar candidates, we pick the one with the lowest fitness to die and be replaced by the offspring. Figure 1 shows a pictorial view of this replacement policy.
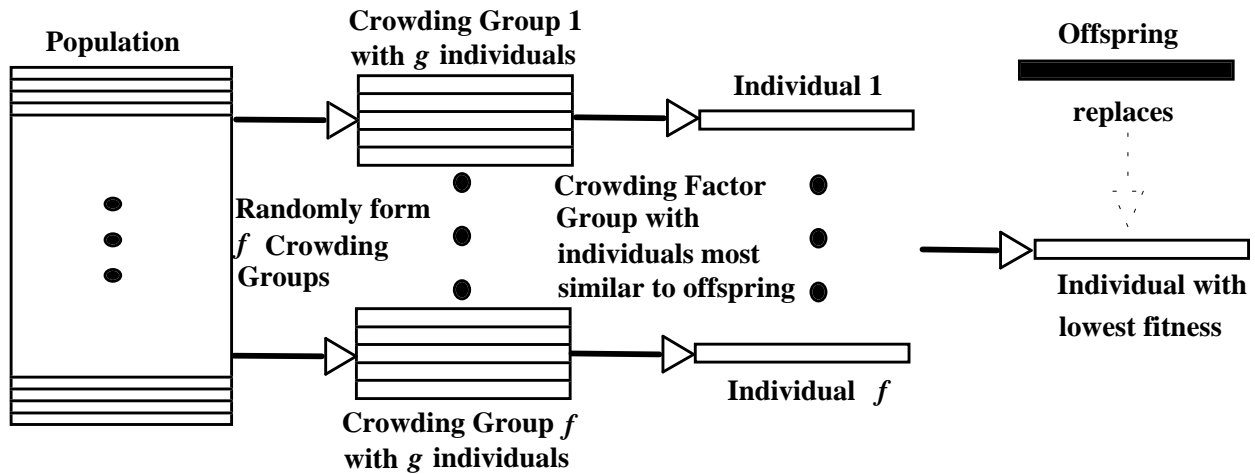

Figure 1: Worst among most similar (WAMS) replacement policy

The WAMS replacement operator can also be expressed mathematically as follows. Let M = *RandSet*(p, g, f) be the a function that returns a *f* by *g* matrix M with individuals in the population, p, selected at random with replacement. This function corresponds to the step in WAMS where *f* sets of size *g* are created from the population. Each row of the matrix is one of the crowding groups. Let V = *MaxSimilar*($M, offspring$) be the function that returns the most similar individual, to the offspring, from each row in the matrix M. This function corresponds to the step in WAMS where we select the most similar out of each crowding group to form the crowding factor group with *f* individuals. Let i = *MinFit*(V) be the function that returns the least fit individual in vector V. Combining the functions above, we get

i = *MinFit*(*MaxSimilar*(*RandSet*(p, g, f), *offspring*))

a minmax representation of WAMS replacement operator.

After an offspring becomes part of the population it competes for survival with other individuals when the next offspring is inserted in the population. In WAMS replacement an offspring is likely to replace a low fitness individual from the same niche. It can also happen that it replaces a high fitness individual from the same niche or an individual from another niche. This stratagem allows a more diverse population to exist throughout the search. At the same time it promotes competition between members of the same niche and between members belonging to different niches. A similar technique was used by Goldberg (1989) in classifier systems but he replaced the most similar individual out of a group of lowest fitness candidates. The pseudo code for the MNC GA is shown in Figure 2.

```
Generate initial population of size n at random.
Evaluate initial population.
For evals = 1 to MAX_EVALUATIONS
    Select individual at random and its mate using crowding selection.
    Apply crossover and mutation.
    Insert offspring in population using WAMS replacement
```

Figure 2: Pseudo code for steady state Multi-Niche Crowding Genetic Algorithm

2.1 Summary of Related Work

There have been many attempts to apply GAs to multimodal search problems and a thorough review of the state of the art can be found in Cedeño (1995) along with complete citations. For completeness, this section provides a brisk review of the context for the analysis presented here.

When classical optimization methods are used to locate the global maximum of multimodal functions, they tend to converge to a local peak. In several applications, information about the location and height of local peaks are as useful as the corresponding information about the global peak. In any event, when GA's are used on multimodal functions, they tend to exhibit the classical behavior of converging, at times prematurely, to a local extremum. Many investigators have proposed modifications to the 'standard GA' in order to render them suitable for multimodal search. These methods seek either to reduce the selection pressure or increase the population diversity.

In Cavicchio's *preselection* (1970), only the offspring with higher fitness than its parent can replace a parent. De Jong's *crowding* (1975) is a generalization of preselection, and the phrase *crowding factor* has its origins in this work. Goldberg and Richardson (1987) used the *sharing*

concept, first proposed by Holland (1975), as a way of reducing the selection pressure caused the FPR rule. Mahfound's *deterministic crowding* (1992) attempts to address the selection pressure issue by allowing any individual to mate with any other individual. Beasely et al's *fitness derating* (1993) allows unimodal optimization methods to be applied to multimodal problems by using the knowledge gained in a run to avoid re-searching the same area. Spears' *subpopulation scheme* (1994) strives to maintain diversity by creating subpopulations in a classical GA using tags. Harik's *retsricted tournment selection* (RTS) (1995) modifies the selection and replacement steps of the Steady State GA to ameliorate the effects of selection pressure. This method uses the greedy stratagem of allowing replacement only when the offspring fitness exceeds the fitness of the individual it is replacing.

Another line of attack is to devise methods that seek to increase population diversity. which is expected to permit the GA to discover new peaks while preserving the good solutions found thus far. In one of the approaches part of the population is re-initialized (Eshelman, 1991; Grefenstette, 1992) after it has converged. Maresky et. al. (1995) introduced an operator called *selectively destructive re-start* that improved the previous approach by reinitializing the chromosome in a solution with certain probability. Cobb and Grefenstette (1993) compared a partial re-start of the population with two approaches that manipulated the mutation rate of the GA. Some other approaches used schemes to encode, in the chromosomes, previous history of the individual (Goldberg and Smith, 1987; Ng and Wong, 1995). In a separate study Dasgupta and McGregor (1992) used a tree structure representation of the population.

In MNC GA, analyzed in this paper, niches are formed by promoting mating and replacement among members of similar phenotype. The result is an algorithm that maintains solutions in multiple peaks while at the same time allowing a subset of the individuals in the population to explore other regions of the search space. This balance between inter-niche and intra-niche selective pressure is a result of crowding selection and WAMS replacement. In the sections that follow we examine these genetic operators in detail and provide information about their effect on the population. Refer to Cedeño (1995) for empirical results about the effect of different parameters on the MNC GA and a comparison to other techniques.

3. Notation

In this section we define some notation and terms that will be used throughout this paper. First, let Ω be the set of possible solutions in the search space. The elements of Ω are called *chromosomes*. For example, let Ω be the search space represented by the binary strings of length l. The number of strings in Ω is given by $r = |\Omega| = 2^l$. For $l = 3$ the cardinality of set Ω is $r = 2^3 = 8$ chromosomes. Figure 3 shows the search space for binary strings of length $l = 3$. For easy identification, each element of Ω is associated with a tag in the interval $[0, r-1]$, which in this case is the decimal value of the binary string. The GA manipulates elements of Ω to find highly fit chromosomes. The members of the population, composed of elements of Ω, are called *individuals*.

$$\Omega = \begin{Bmatrix} 000 \\ 001 \\ 010 \\ 011 \\ 100 \\ 101 \\ 110 \\ 111 \end{Bmatrix} \qquad Tags = \begin{Bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{Bmatrix}$$

Figure 3: GA search space, Ω, for binary strings of length $l = 3$ and their corresponding tag values.

One way to denote the population of size n is by the column vector $\boldsymbol{I} = [I_0\ I_1\ ...\ I_{n-1}]^T$, where T stands for vector transpose. A component of this vector, namely I_j, represents the jth individual in the population with a value in Ω. Using elements of Ω in Figure 3, the vector $\boldsymbol{I} = [4\ 1\ 7\ 5\ 4\ 3]^T$ is a valid population with six individuals whose tag values are 4, 1, 7, 5, 4, and 3. Clearly, there are different vectors $\boldsymbol{I}$ representing the same population, like for example $\boldsymbol{I}' = [5\ 4\ 3\ 4\ 1\ 7]^T$ which is a permutation of the previous vector. The size of vector $\boldsymbol{I}$ is n, the number of individuals in the population.

A population can also be represented by a vector of size r, the cardinality of Ω, containing the number of copies of each tag value in the population. This representation allows similar populations, those that differ only by a permutation of their individuals, to be represented

uniquely. In this representation a population is denoted by the column vector $p = [p_0\ p_1\ ...\ p_{r-1}]^T$. The components of this vector, namely p_j, are the number of copies of a chromosome with tag value j in the population. Using the example cited above we have $p = [0\ 1\ 0\ 1\ 2\ 1\ 0\ 1]^T$, see Figure 4, which represents one copy of chromosome 1, one copy of chromosome 3, two copies of chromosome 4, one copy of chromosome 5, and one copy of chromosome 7 in the population. The order of the individuals in the population is lost when using the representation given by vector p. Evidently

$$\sum_{j=0}^{r-1} p_j = n,$$

the size of the population.

$$I = \begin{bmatrix} 4 \\ 1 \\ 7 \\ 5 \\ 4 \\ 3 \end{bmatrix} \Rightarrow p = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Figure 4: Different representation of a population. The vector I denotes the individuals in the population by their tag value. The vector p represents the same population by the number of copies of each chromosome in the population I.

4. Analysis of Crowding Selection

In this section we define the selection probability for any pair of individuals under crowding selection. As indicated previously crowding selection selects the pair of individuals that will undergo crossover. For notational convenience we will identify the first individual selected into the pair as the *parent* and the other the *mate*. The selection step of MNC GA, described in Section 2, can be summarised as follows. For each mating pair the parent is chosen uniformly at random from the population and its mate is chosen using crowding. That is, the mate is chosen as the one that is most similar to the parent from a group of s candidates taken at random from the

population, one at a time, with replacement. It is not hard to see that fitness plays no role in this selection step.

Let $\boldsymbol{u} = [u_0 \ u_1 \ ... \ u_{r-1}]^T$ be the vector defining the parent selection probability for each chromosome in Ω. That is u_j, the jth element, is the probability that chromosome j is selected as the parent for mating. We emphasise that the quantity u_j refers only to the probability of selecting a parent, not its mate. Notice also that in MNC GA, it is always true that

$$\boldsymbol{u} = \tfrac{1}{n}\boldsymbol{p},$$

<div align="right">Equation 1</div>

where $\boldsymbol{u}$ and $\boldsymbol{p}$ are vectors and n is a scalar. From this formulation it is clear that

$$\sum_{j=0}^{r-1} u_j = 1.$$

To facilitate the analysis of crowding selection, we introduce the function $P_s(j,k,s,n,p)$, which defines the probability that chromosome k is selected as the mate, given parent j, from a group of s individuals from the population, represented by $\boldsymbol{p}$, of size n. Recall that in crowding selection the mate is the most similar individual to parent j from a group of s candidates taken at random (with replacement) from the population. To calculate $P_s(j,k,s,n,p)$ we first need to rank the members of the population according to their similarity to parent j. Towards this end we define the function $Sr(j,k,p)$ which returns the similarity ranking of mate k with respect to parent j in the population.

The ranking of chromosome k with respect to chromosome j can be obtained by sorting the members of the population in ascending order of their phenotypic distance to chromosome j (assume for now that there are no ties). Then, a rank value ranging from 0 to n-1 is assigned to the sorted list. The value of 0 is given to the population member closest to j (always itself) and n-1 to the one farthest from j. The smaller the distance, the more similar a chromosome is and the lower its rank value. The value $Sr(j,k,p)$ is problem dependent and is assumed (for the moment) to be a unique value between 0 and n-1. The rank value returned by $Sr(j,k,p)$ indicates the number of individuals in the population with lower rank than chromosome k, i.e., there are $Sr(j,k,p)$ population members closer to j (lower distance) than chromosome k.

To calculate $P_s(j,k,s,n,p)$ it is only necessary to know how many of the possible crowding selection groups, where the order of the chromosomes matters, will have chromosome k as the lowest ranked member of the group. This value can be obtained by counting all possible groups where chromosome k appears among higher ranked members of the population. In such cases chromosome k appears at least once and up to s times in that group. All other positions in the group are filled with an arbitrary combination of the higher ranked chromosomes. We also know that there are n^s possible ways of selecting s chromosomes, one at a time with replacement, from a population with n chromosomes. Given that chromosome k has rank $Sr(j,k,p)$ we have $(n - Sr(j,k,p) - 1)$ chromosomes with higher rank. The number of ways of arranging m copies of k in a group with s positions is given by $\binom{s}{m}$. The number of ways of selecting $(s - m)$ higher rank chromosomes, one at a time with replacement, is given by $(n - Sr(j,k,p) - 1)^{s-m}$. Adding over all possible values of m, the number of times chromosome k is in the crowding selection group we get the following expression;

$$P_s(j,k,s,n,p) = \frac{1}{n^s} \sum_{m=1}^{s} \binom{s}{m}(n - Sr(j,k,p) - 1)^{s-m},$$

which is easily evaluated using the binomial theorem, to yield

$$P_s(j,k,s,n,p) = \frac{(n - Sr(j,k,p))^s - (n - Sr(j,k,p) - 1)^s}{n^s}.$$

Equation 2

Equation 2 does not account for chromosomes having the same distance to parent j as mate k or duplicate copies of k. In these cases we have the chromosomes with the same rank value (ties). If we break ties at random we can calculate $P_s(j,k,s,n,p)$ by averaging over all possible rankings of chromosome k and multiplying the result by the number of copies of chromosome k. Let $Er(j,k,p)$ be the function that returns the number of chromosomes in the population with the same distance to j as k (including all copies of chromosome k). Redefining $Sr(j,k,p)$ to represent the number of population members that are closer to j than to k, we can expand Equation 2 to obtain the average probability under crowding selection and get

$$P_s(j,k,s,n,p) = \frac{p_k}{Er(j,k,p)n^s} \sum_{m=0}^{Er(j,k,p)-1} (n-(Sr(j,k,p)+m))^s - (n-(Sr(j,k,p)+m)-1)^s$$

$$= p_k \frac{(n-Sr(j,k,p))^s - (n-Sr(j,k,p)-Er(j,k,p))^s}{Er(j,k,p)n^s}.$$

<div align="right">Equation 3</div>

In summary, the probability that chromosome j (the parent) and k (its mate) are selected is given by Equation 1 and Equation 3 respectively. From these two equations the probability that chromosomes j and k undergo crossover is

$$\chi \frac{1}{n} p_j p_k \frac{(n-Sr(j,k,p))^s - (n-Sr(j,k,p)-Er(j,k,p))^s}{Er(j,k,p)n^s},$$

where χ denotes the crossover probability.

It is useful to note, in passing, that Nix and Vose (1992) have shown that the number of possible populations of size n is given by

$$\binom{n+r-1}{r-1}.$$

This result was not used here because we account for all the possible ways to form the crowding selection group. Each chromosome in the group is selected at random with replacement from the population. Given that we have s positions in the group and n possible choices for each position, we get a total of n^s possible crowding selection groups.

5. Analysis of Worst Among Most Similar Replacement

In this section we calculate, for all individuals in the population, $P_r(j,k,g,f,n,p)$, the probability that chromosome k is replaced by offspring j under WAMS replacement. Recall that during replacement the MNC GA creates f crowding groups with g individuals each, chosen at random with replacement, from the population. Then the individual most similar to the offspring in each crowding group is chosen to form the crowding factor group of f most similar candidates. From the crowding factor group the least fit individual is replaced by the offspring in the population.

It is not hard to see a relation between the replacement step and crowding selection. Each one of the individuals in the crowding factor group is selected using crowding, but here the similarity ranking is based on the offspring. Therefore, the creation of the crowding factor group can be viewed as the application of crowding selection f times with $s = g$ and the offspring being the parent. Given this relationship we can use the results from Section 4 to get the probability that a chromosome in the population is selected from one of the crowding groups. Therefore, Equation 3

$$P_s(j,k,g,n,p) = p_k \frac{(n - Sr(j,k,p))^g - (n - Sr(j,k,p) - Er(j,k,p))^g}{Er(j,k,p)n^g},$$

gives us the probability of selecting chromosome k from the population to a crowding group given the offspring j. There is only one distinction, here the offspring j can be any chromosome of the entire set Ω, while during crowding selection parent j can only be a chromosome in the population.

Now we need to define the probability of selecting a chromosome from the crowding factor group. Let $Fr(k,p)$ be the function that returns the fitness rank of the individual k in the population. The function $Fr(k,p)$ returns the number of individuals in the population with lower fitness than chromosome k. The least fit individual is assigned a rank value of 0 and the most fit individual a rank value of n - 1. Here again we break ties at random. Let $F(k)$ be the function used to calculate the fitness value of chromosome k in Ω. Then we have

$$Fr(k,p) = \sum_{j=0}^{r-1} p_j \theta_1(j,k),$$

where

$$\theta_1(j,k) = \begin{cases} 1 & \text{if } F(j) < F(k) \\ 0 & \text{otherwise} \end{cases}$$

Equation 4

Next, define the function $Ef(k,p)$ that returns the number of individuals in the population with the same fitness value as chromosome k. Using the fitness function $F()$ we have

$$Ef(k, p) = \sum_{j=0}^{r-1} p_j \theta_2(j, k),$$

where

$$\theta_2(j, k) = \begin{cases} 1 & \text{if } F(j) = F(k) \\ 0 & \text{otherwise} \end{cases}$$

<div align="right">Equation 5</div>

To calculate $P_r(j, k, g, f, n, p)$, the probability of replacing chromosome k with offspring j, we need to know all possible permutations of the crowding factor group where chromosome k has the lowest fitness (lowest rank) value. Then we add the probability of each of the crowding factor groups to obtain $P_r(j, k, g, f, n, p)$. Unlike forming the crowding groups, the probability of being selected to the crowding factor group is not random, but based in the similarity to offspring j and given by $P_s(j, k, g, n, p)$. Given the offspring j, the probability of a particular crowding factor group $(i_1, i_2, ..., i_f)$, where i_k is a chromosome in the population, is given by the product

$$\Lambda(j, g, n, p, i_1, i_2, ..., i_f) = \prod_{k=1}^{f} \frac{P_s(j, i_k, g, n, p)}{p_{i_k}}.$$

<div align="right">Equation 6</div>

Enumerating all possible crowding factor groups where chromosome k has the lowest fitness rank and adding the probability of each group will give us the value of $P_r(j, k, g, f, n, p)$. For chromosomes with equal fitness rank, we need to average over all possible fitness rank values to obtain the actual value of $P_r(j, k, g, f, n, p)$. Assume, without loss of generality, that the chromosomes $(i_1, i_2, ..., i_{Fr(k,p)})$ are the individuals in the population with lower rank than chromosome k. Assume also that the chromosomes $(i_{Fr(k,p)+p_k+1}, ..., i_{Fr(k,p)+Ef(k,p)})$ are the individuals with equal rank as chromosome k (not including the copies of chromosome k) and $(i_{Fr(k,p)+Ef(k,p)+1}, ..., i_n)$ are the individuals with higher rank than chromosome k. Using Equation 4, Equation 5, and Equation 6 we can now calculate the probability that offspring j replaces chromosome k in the population using WAMS replacement with the following equation:

$$P_r(j,k,g,f,n,p) = \frac{1}{a} \sum_{l=0}^{a-p_k} \sum_{m=1}^{f} \binom{f}{m} \left(\sum_{\ell_1=h-l}^{n} \sum_{\ell_2=h-l}^{n} \cdots \sum_{\ell_{f-m}=h-l}^{n} \Lambda(j,g,n,p,i_{\ell_1},i_{\ell_2},\ldots,i_{\ell_{f-m}},i_k^m) \right),$$

where

$$h = Fr(k,p) + Ef(k,p) + 1,$$

$$a = Ef(k,p),$$

i_k^m denotes m copies of chromosome i_k.

<div align="right">Equation 7</div>

In summary, Equation 7 specifies the replacement probability for chromosome k in the population, given that offspring j was generated after crossover and mutation. It is not hard to see that fitness plays an important role during WAMS replacement. A high fitness value results in a high fitness rank ($Fr(k,p)$) value in the population. A high fitness rank value results in a lower probability of being selected from the crowding factor group for replacement. Closer similarity to the offspring, on the other hand, increases the probability of being selected into the crowding factor group. The combination of similarity to the offspring and fitness determines the replacement probability of an individual in the population.

6. A Closer Look at Crowding Selection

In this section we examine the properties of crowding selection more closely. Specifically, we calculate, under crowding selection, the bounds for the selection probability ($P_s(j,k,s,n,p)$) of the mate as well as the expected value and variance of the similarity rank. Using Equation 2 we can calculate the lower and upper bounds of $P_s(j,k,s,n,p)$, the probability of selecting chromosome k using crowding selection given that chromosome j is the parent. The lower bound can be obtained from the highest similarity rank value ($Sr(j,k,p) = n-1$) and the upper bound from the lowest similarity rank value ($Sr(j,k,p) = 0$). Using these values we have

$$\frac{1}{n^s} \leq P_s(j,k,s,n,p) \leq \frac{n^s - (n-1)^s}{n^s}.$$

<div align="right">Equation 8</div>

<div align="center">14</div>

Take for example the case $n = 10$ and $s = 2$, we have $0.01 \leq P_s(j,k,s,n,p) \leq 0.19$. Although the probability of selecting the highest ranked individual is small, it is nevertheless non-zero and given enough trials it will be selected.

It is not hard to show that $P_s(j,k,s,n,p)$ summed over all chromosomes k in Ω is equal to unity for any chromosomes in the population, i.e.,

$$\sum_{k=0}^{r-1} P_s(j,k,s,n,p) = 1 \text{ when } p_j \neq 0.$$

<div align="right">Equation 9</div>

We will prove Equation 9 by adding Equation 2 over all possible rank values. Using m to denote all possible rank values for $Sr(j,k,p)$ we can verify that the numerators of Equation 2 sum to the total number permutations under crowding selection (n^s). Since all possible rank values for m are in the range [0, n-1] we have that

$$\sum_{m=0}^{n-1} (n-m)^s - (n-m-1)^s = n^s.$$

<div align="right">Equation 10</div>

Of importance to us is the expected value of the similarity rank of a mate as s (crowding selection group size) varies. Knowing the expected value of the similarity rank, for different values of s, will allow us to select an appropriate value for a given fitness function. Using Equation 2, the expected value of the similarity rank, $E(Sr)$, of the mate for any parent is

$$E(Sr) = \frac{1}{n^s} \sum_{m=0}^{n-1} m[(n-m)^s - (n-m-1)^s] = \frac{1}{n^s} \sum_{m=1}^{n-1} (n-m)^s = \frac{1}{n^s} \sum_{m=1}^{n-1} m^s.$$

<div align="right">Equation 11</div>

In the same manner we can calculate the variance $Var(Sr)$ to get

$$Var(Sr) = E(Sr^2) - E(Sr)^2 = \frac{1}{n^s} \sum_{m=1}^{n-1} (2n-2m-1)m^s - (\frac{1}{n^s} \sum_{m=1}^{n-1} m^s)^2.$$

<div align="right">Equation 12</div>

From the variance we can calculate the standard deviation as $STD(Sr) = \text{SQRT}(Var(Sr))$.

Equation 11 and Equation 12 allow us to observe the mating pressure imparted during selection by the MNC GA. Figure 5 presents such information graphically for a population size of 100. Clearly, we can see that the expected rank value decreases as the crowding selection size increases. This indicates that during crowding selection a large crowding selection size will more likely generate a mating pair from the same niche. Given a population size, we can calculate the crowding selection size, s, that will give us the expected rank values that will promote mating among localise individuals. Moreover, we can observe that the expected rank value do not change significantly after a crowding selection size of 11. The effect of crowding selection and its benefits are more noticeable at lower s values. This result agrees with the rule of thumb we have been using in our experiments. The rule specified a value of s between 2 and 15% of the population size.
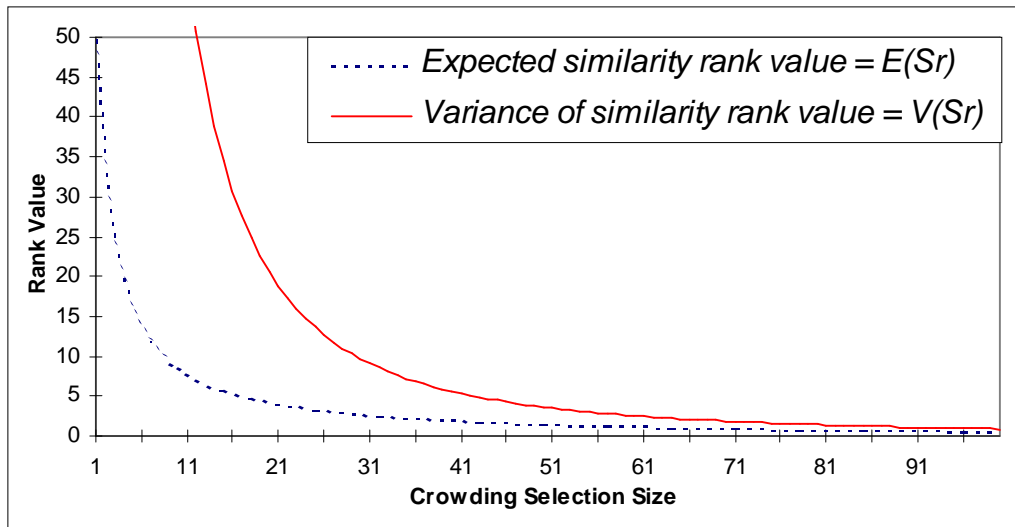

Figure 5: Expected value and variance of the similarity rank of the mate under crowding selection as a function of the crowding selection size (s) using a population size $n = 100$.

Another way to look at the effect of the crowding selection size in crowding selection is by examining the probability distribution for the different rank values. The probability distribution can be used to answer questions about the probability for a specific rank value. For example, one can calculate the probability that chromosome with rank k is selected for mating. Figure 6 shows the probability distribution for a crowding selection size of 1, 6, and 11. Again, we can observe

that a higher crowding selection size will increase the probability that a lower rank individual (which means more similar in our case) is selected as a mate. Similarly, we can use these results to calculate, for a given population size, the crowding selection for a specific rank value and selection probability. These results allow us to select appropriate parameters to control how much localised mating we want for a particular problem. For example, suppose you would like the probability of selecting an individual with rank 10 to be 0.01. Using Equation 2 we can calculate the value of s necessary for crowding selection.
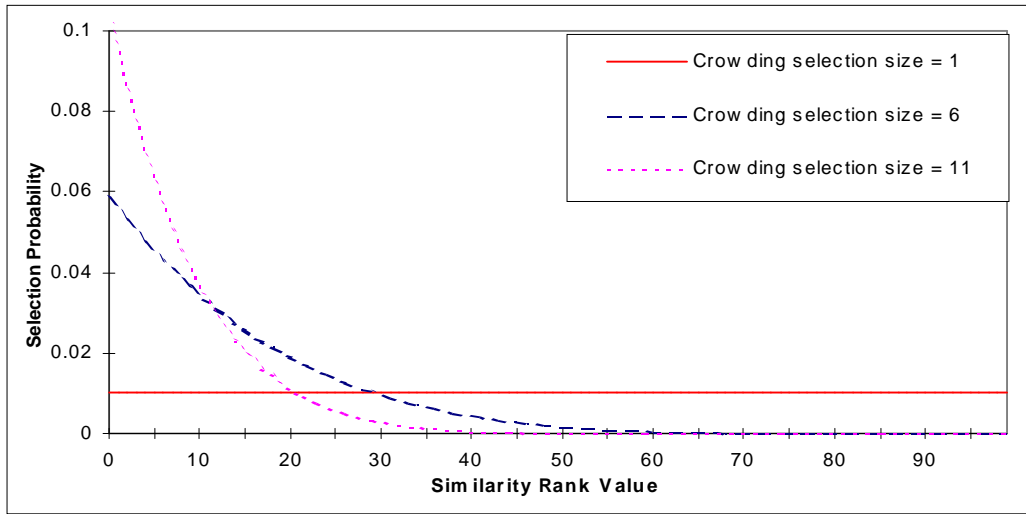

Figure 6: Probability distribution of rank values for the mate under crowding selection using a population of size $n = 100$.

7. A Closer look at Worst Among Most Similar Replacement

In this section we further examine the properties of WAMS replacement. Specifically, we want to look at the effect of the parameters g and f on replacement. Recall that in WAMS replacement f crowding groups, each with g individuals, are formed by choosing individuals at random (with replacement) from the population. Then the most similar individual (to the offspring) in each crowding group is selected to form the crowding factor group with f individuals. The least fit individual in the crowding factor group is replaced by the offspring in the population.

Creating a crowding group is similar to crowding selection. All the results obtained in Section 6 apply directly to the creation and selection of individuals to crowding groups. In summary, increasing the value of the group size, in this case g, decreases the expected similarity rank of the individual, thus increasing the probability of selecting individuals from the same niche. During replacement this means that offspring are more likely to replace members of the same niche when using higher values of g.

Recall that $Sr(j,k,p)$ denotes the similarity rank of chromosome k in the population with respect to chromosome j, the offspring. That is, there are $Sr(j,k,p)$ chromosomes in the population that are more similar to the offspring than chromosome k. Assume also that no two chromosomes in the population have the same similarity rank. Using Equation 2, the probability that chromosome k in the population is selected from a crowding group is

$$P_s(j,k,g,n,p) = \frac{(n - Sr(j,k,p))^g - (n - Sr(j,k,p) - 1)^g}{n^g}.$$

<div align="right">Equation 13</div>

A chromosome will be in the crowding factor group if it is selected from at least one of the crowding groups. Since there are f crowding groups, the probability that chromosome k is selected into the crowding factor group is

$$P_c(j,k,g,f,n,p) = 1 - (1 - P_s(j,k,g,n,p))^f,$$

<div align="right">Equation 14</div>

where

$$(1 - P_s(j,k,g,n,p))^f$$

is the probability that chromosome k is not selected from any of the crowding groups.

Once the individuals are selected from the crowding groups to form the crowding factor group, fitness is used to select the individual being replaced by the offspring. Here the individual selected is the one with the lowest fitness in a group of f individuals. Once selected into the crowding factor group, the chromosome replaced by the offspring must be the one with the lowest fitness in the group. For simplicity, let us assume that each of the chromosomes in the

population can be assigned a unique fitness rank, from 0 to *n-1*. Let us also assume that the fitness rank values are assigned in ascending order of fitness, that is, a value of 0 is assigned to the individual with the lowest fitness, 1 to the second lowest fitness, and so on until the most fit individual gets the fitness rank of *n-1*.

Let $Fr(k,p)$ denote the chromosomes in the population with lower fitness rank (lower fitness value) than chromosome *k*. Let $Ef(k,p)$ denote the chromosomes in the population with equal fitness rank as chromosome *k*. We can use Equation 3 again to calculate the probability, $P_f(k,f,n,p)$, that chromosome *k* with fitness rank $Fr(k,p)$ is selected from a group of *f* chromosomes selected at random with replacement. Similar to crowding selection, but using fitness rank instead of similarity rank. Replacing similarity rank by fitness rank and ignoring chromosome *j* we have,

$$P_f(k,f,n,p) = p_k \frac{(n - Fr(k,p))^f - (n - Fr(k,p) - Ef(k,p))^f}{Ef(k,p)n^f} .$$

<div align="right">Equation 15</div>

A simpler form of Equation 15 exists when the fitness ranks of the chromosomes in the population have different fitness values. In this case $Ef(k,p) = 1$ and $p_k = 1$, and we get

$$P_f(k,f,n,p) = \frac{(n - Fr(k,p))^f - (n - Fr(k,p) - 1)^f}{n^f} .$$

<div align="right">Equation 16</div>

Not accounted for in the above equation is how the fitness rank, of the individuals selected to the crowding factor group, is affected by the similarity rank to the offspring. Intuitively one would expect the fitness rank of the offspring to be relatively close to that of similar individuals. Since the crowding factor group is a group of most similar individuals to the offspring, the fitness rank of the individual selected for replacement is therefore dependent on its similarity rank. Nevertheless, to analyse the effect of WAMS in the population, we will assume that the fitness rank of a chromosome in the population is independent of its similarity rank to the offspring. Then, we can calculate the replacement probability under WAMS replacement by multiplying Equation 14 and Equation 16. The probability, $P_r(j,k,g,f,n,p)$, that a

chromosome k with fitness rank $Fr(k,p)$ and similarity rank $Sr(j,k,p)$ is replaced by a chromosome j, the offspring, under WAMS replacement is given by

$$P_r(j,k,g,f,n,p) = P_c(j,k,g,f,n,p)P_f(k,f,n,p) = (1-(1-P_s(j,k,g,n,p))^f)P_f(k,f,n,p).$$

<div align="right">Equation 17</div>

From Equation 17 we can calculate $E(Sr, Fr)$, the expected similarity rank and fitness rank values of the chromosome replaced by the offspring under WAMS replacement. Given a population size n, crowding group size g, and crowding factor f we have

$$E(Sr,Fr) = \sum_{Sr=0}^{n-1}\sum_{Fr=0}^{n-1} SrFr\left(1-\left(1-\frac{(n-Sr)^g-(n-Sr-1)^g}{n^g}\right)^f\right)\left(\frac{(n-Fr)^f-(n-Fr-1)^f}{n^f}\right).$$

<div align="right">Equation 18</div>

The expected fitness rank value, $E(Fr)$, can also be calculated in the same manner using as the expected similarity rank under crowding selection, Equation 11, with the exception of the group size which is given by the crowding factor, f, in this case. In contrast, the expected similarity rank value, $E(Sr)$, is affected by the number of crowding groups, f, formed.
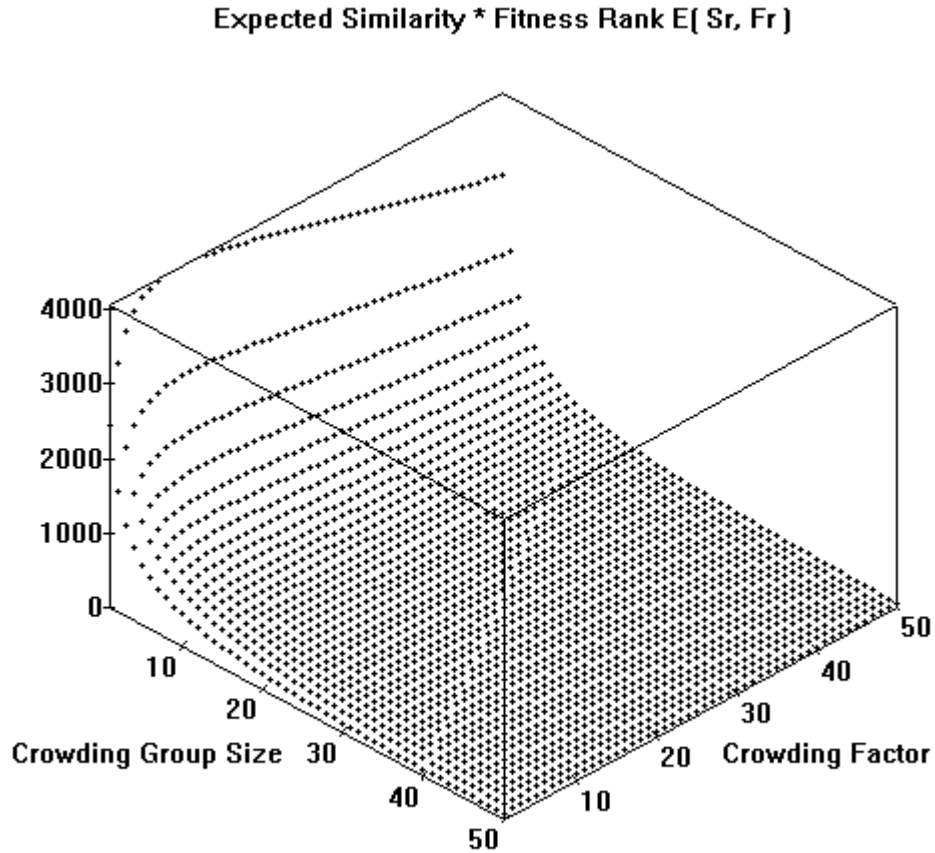
Figure 7: Expected fitness rank values under WAMS replacement as a function of the crowding group size, *g*, and crowding factor, *f* for a population of size 100.

There are various things that can be pointed out from these results. Increasing the crowding factor group size, *f*, increases the probability that lower fitness individuals are replaced in the population. The size of the crowding group determines the likelihood that the offspring replaces similar individuals. The effect of the crowding factor and crowding group size values can be seen in Figure 7. As the value *f* increase, it is more likely that an individual with low fitness rank is selected. Also shown in Figure 7, as the value of *g* increases it is more likely that an individual with low similarity rank (more similar to the offspring) is selected. Increasing the values of *g* in the WAMS replacement operator increases inter-niche competition, i.e., the probability of selecting a more similar chromosome for replacement. On the other hand, increasing the value of *f* increases intra-niche competition, i.e., the probability that a low fitness individual is selected for replacement.
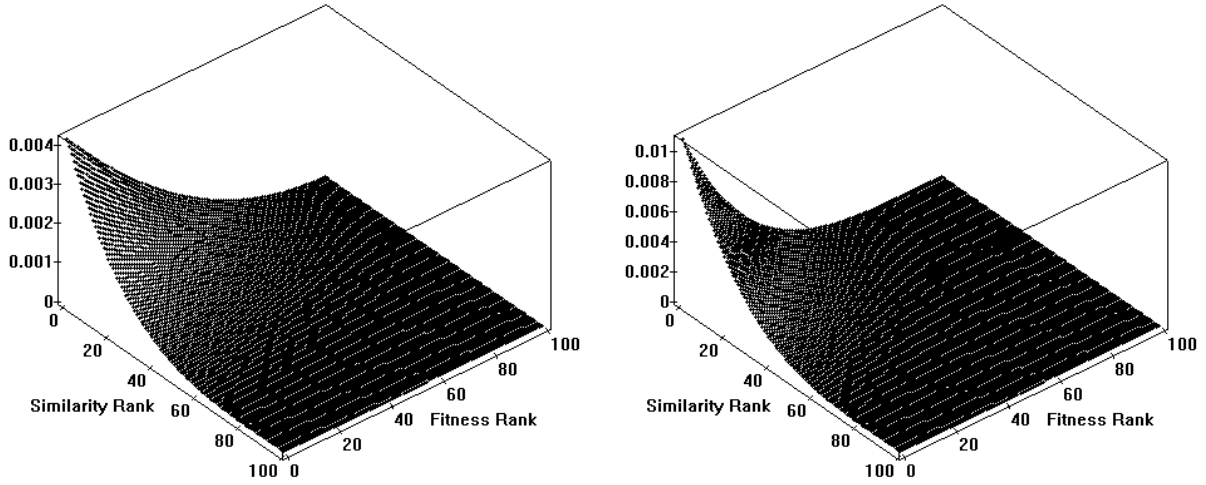
Figure 8: Probability distribution, under WAMS replacement, for the individuals in a population of size 100. Plot A shows the distribution for a crowding group size of 5 and a crowding factor of 3. Plot B shows the distribution for a crowding group size of 5 and a crowding factor of 5.

Figure 8 shows the effect of increasing the value of *f* on the probability distribution of the individuals in the population of size 100. As the value of *f* increases so does the probability of selection for individuals with low fitness rank, i.e., low fitness values. It is the WAMS replacement operator that applies the "survival of the fittest" metaphor to the members of the population. Individuals with higher fitness are more likely to survive from generation to generation. The WAMS replacement operator also increases the likelihood of high fit individuals to reproduce because they are most likely to survive for many more generations.

8. Empirical Results

In this section we apply the MNC GA to a hypothetical function *F(x, y)* and collect empirical data about the ranks of the individuals selected for mating and replacement. We will then compare the empirical results with the results predicted by work presented in the previous sections. Specifically, we will calculate the average similarity rank of the mate during crowding selection and the average fitness rank of the individual selected for replacement using WAMS. The function *F(x, y)* is given by the equation

$$F(x, y) = \sum_{i=1}^{2} H_i / 1 + W_i[(x - X_i)^2 + (y - Y_i)^2],$$

22

where

$$X = [45000\,,\,15000],\ Y = [2000,\,62000],\ H = [100,\,100],\ W = [0.0004,\,0.0004].$$

This function has two peaks of equal height, 100, located at coordinates (45000, 2000) and (15000, 62000). A population size of $n = 100$ was used and the GA was run for 50 generations. A total of 5000 selections were used in each run to calculate the average similarity rank of the mate in crowding selection. Similarly, a total of 5000 replacements were used to calculate the average fitness rank of the individual replaced by the offspring.

Table 1 shows the results for different values of crowding selection size (s), crowding group size (g), and crowding factor (f). The expected values for the similarity rank agreed with the empirical values calculated for the different parameters. The expected values for the fitness rank however did not agree in many places with the empirical values. Only the empirical values for rows 1, 5, and 9 are close to the expected values. Some of the discrepancy can be accounted for by the function used for the experiment. The assumption we made about the independence of similarity and fitness to simplify the results does not apply for $F(x,\,y)$. The fitness of individuals in a neighbourhood is very dependent to their proximity. The larger the value of g during WAMS replacement, the more likely that the individual selected for replacement will belong to the same neighbourhood as the offspring and therefore have similar fitness value. This effect can be observed in rows 5 to 8 in Table 1. The larger the value of g, the large is the discrepancy between the observed fitness rank and expected fitness rank of the selected individual.

Table 1: Comparison of empirical results and expected values for the similarity rank and fitness rank. Column 1 shows the values for crowding selection size (s), crowding group size (g), and crowding factor (f) used in the MNC GA.

(s, g, f)	Empirical Average Similarity Rank in Crowding Selection	Expected Average Similarity Rank in Crowding Selection	Empirical Average Individual Fitness Rank in WAMS	Expected Average Individual Fitness Rank in WAMS
(1, 5, 3)	51.0716	49.5	26.9386	24.50
(5, 5, 3)	16.2428	16.17	34.5012	24.50
(10, 5, 3)	9.5104	8.6	34.87	24.50
(15, 5, 3)	6.5272	5.76	34.722	24.50
(5, 1, 3)	16.9072	16.17	24.3106	24.50
(5, 5, 3)	16.2428	16.17	34.5012	24.50
(5, 10, 3)	16.4124	16.17	39.5894	24.50

(5, 15, 3)	16.5404	16.17	40.1656	24.50
(5, 5, 1)	16.8732	16.17	50.7122	49.5
(5, 5, 3)	16.2428	16.17	34.5012	24.50
(5, 5, 5)	16.442	16.17	26.9186	16.17

To a lesser extent, the convergence properties of the MNC GA can also be attributed to the discrepancy of some of the values. As the individuals in the population start to converge to the peaks in the function, the population becomes homogeneous thus affecting the average fitness rank values. The expected value as calculated in the previous section does not account for duplicate individuals in the population.

Nevertheless the results we obtained show the pattern we presented in Figure 7. The average fitness rank value of the individual selected for replacement decreases as the value of the crowding factor is increased. This means that low fit individuals are more likely to be selected for replacement. In the same manner, when the crowding group size is increased so does the likelihood of selecting and individual from the same neighbourhood as the offspring. It is this balance between inter-niche competition and intra-niche competition that allows the MNC GA the evolve individuals to different niches in the search space.

9. Comments and Conclusions

The results presented here brings us closer to an understanding of the effect of different parameters in the MNC GA. The effect of the crowding selection size, s, is clear. Increasing its value increases the likelihood of selection of a mate from the same niche as the parent. Values between 5% and 15% of the population size are appropriate for selecting mating pairs from the same niche and at the same time allowing mating between pairs of different niches . The higher the value of the crowding selection size, the higher is the mating pressure during selection. The value of s controls the amount of inter-niche and intra-niche breeding in the MNC GA. Using the results from Section 6, the appropriate value of s can be selected to achieve the desire mating pressure during a run.

The effect of WAMS replacement can also be explained by the values of the crowding group size, g, and crowding factor, f, parameters. Competition between members of same niche can be increased by increasing the value of the crowding group size. On the other hand decreasing the

value of the crowding group size increases competition among members of different niches. The risk here is that niches with lowered average fitness may not be able to maintain any individuals in them. By using appropriate values for the crowding group size and crowding factor we can increase replacement of low-fitness individuals from the same niche allowing the MNC GA to converge to the top of different niches. Values between 5% and 15% of the population size are also acceptable for the crowding group size (Cedeño, 1995).

Increasing the value of the crowding factor, increases the fitness pressure for the individuals in the population. The probability that a low fit individual in the population is replaced increases as a function of the crowding factor. It clearly shows that the WAMS replacement operator applies the "survival of the fittest" metaphor to the members of the population. Moreover, the MNC GA does not use fitness during selection. The likelihood that an individual participates in mating is directly influenced by the WAMS replacement operator. WAMS replacement allows high fit individuals to survive for many more generations allowing them to participate in mating more often. Values between 2 and 10% of the population size are recommended values for the crowding factor.

The value of g controls the amount of inter-niche and intra-niche competition in the MNC GA. The lower the value of g, the higher the competition between members of different niches. The value of f, on the other hand, controls the selective pressure in the MNC GA. The higher the value of f, the higher the chances of eliminating low fit individuals using WAMS replacement. Both values can be combined to achieve the desire selective pressure and niche competition in a run. The values of g and f control the classical tradeoff between exploration and exploitation in GAs.

In order to accurately predict the replacement probability under WAMS it is necessary to determine the dependency between the similarity rank and the fitness rank. This dependency is directly affected by the fitness function and the search space being analyzed by the GA. It seems to be beneficial to incorporate the fitness function as part of the analysis to obtain more accurate results when applying the MNC GA. Knowing how different parameters affect the convergence properties of the algorithm, to different fitness functions, will make its application to other problems easier.

Acknowledgment

The authors are indebted to the two anonymous reviewers whose painstaking and detailed commentary substantially helped in improving the quality of the paper.

References

Beasley, D., Bull, D. R., and Martin, R. R. (1993). A sequential technique for multimodal function optimization, *Evolutionary Computation*, volume 1, number 1, MIT Press, MA.

Cavicchio, D. J. (1970). *Adaptive search using simulated evolution*. Ph.D. thesis, University of Michigan, Ann Arbor, MI.

Cedeño, W. (1995). The multi-niche crowding genetic algorithm: analysis and applications. *UMI Dissertation Services*, 9617947.

Cedeño, W. and Vemuri, V. (1996). Database design with genetic algorithms. D. Dasgupta and Z. Michalewicz (eds), *Evolutionary Algorithms in Engineering Applications*, Springer Verlag, 3/97.

Cedeño, W., Vemuri, V., and Slezak, T. (1995). Multi-Niche crowding in genetic algorithms and its application to the assembly of DNA restriction-fragments. *Evolutionary Computation*, 2:4, 321-345.

Cedeño, W. and Vemuri, V. (1992). Dynamic multimodal function optimization using genetic algorithms. In *Proceedings of the XVIII Latin-American Informatics Conference*, Las Palmas de Gran Canaria, Spain: University of Las Palmas, 292-301.

Cobb, H. J. and Grefenstette, J. J. (1993). Genetic algorithms for tracking changing environments. In S. Forrest (ed.) *Proceedings of the Fifth International Conference on Genetic Algorithms*. Morgan Kaufmann Publishers San Mateo, California, 523-530.

Dasgupta, D. & McGregor, D. R. (1992). Non-stationary function optimization using the structured genetic algorithm. In R. Manner and B. Manderick (eds.), *Parallel Problem Solving from Nature*, 2. Amsterdam: North Holland, 145-154.

De Jong, K. A. (1975). *An analysis of the behaviour of a class of genetic adaptive systems*. Doctoral dissertation, University of Michigan. Dissertation Abstracts International 36(0), 5140B. (University Microfilms No. 76-9381).

Deb, K. and Goldberg, D. E. (1989). An investigation of niche and species formation in genetic function optimization, In J. D. Schaffer (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms*. San Mateo, CA: Morgan Kaufmann, 42-50.

Eshelman, L. J., ed. (1991). The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. In G.J.E. Rawlings (ed.), *Foundations of Genetic Algorithms - 1*. San Mateo, CA:Morgan Kaufmann Publishers.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization & Machine Learning*. Reading MA: Addison-Wesley.

Goldberg, D. E., & Richardson, J. (1987). Genetic algorithms with sharing for multimodal function optimization. In J. J. Grefenstette (Ed.), *Proceedings of the Second International Conference on Genetic Algorithms.* Hillsdale, NJ: Lawrence Erlbaum Associates, 41-49.

Goldberg D. E. & Smith R. E. (1987). Non-stationary function optimization using genetic algorithms with dominance and diploidy. In J. J. Grefenstette (Ed.), *Proceedings of the Second International Conference on Genetic Algorithms.* Hillsdale, NJ: Lawrence Erlbaum Associates

Grefenstette, J.J. (1992). Genetic algorithms for changing environments. In R. Manner and B. Manderick (eds.), *Parallel Problem Solving form Nature*, 2. Amsterdam: North Holland, 137-144.

Harik, G. R. (1995). Finding multimodal solutions using restricted tournament selection. In L. J. Eshelman (ed.), *Proceedings of the Sixth International Conference on Genetic Algorithms*. San Mateo, CA:Morgan Kaufmann Publishers, 24-31.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*, Ann Arbor MI: The University of Michigan Press.

Mahfoud, S. W. (1992). Crowding and preselection revisited. In R. Männer & B. Manderick (Eds.), *Proceedings of Parallel Problem Solving from Nature 2.* New York, NY: Elsevier Science B. V., 27-36.

Maresky, J., Davidor, Y., Gitler, D., Aharoni, G., and Barak, A. (1995). Selectively destructive restart. In L. J. Eshelman (ed.), *Proceedings of the Sixth International Conference on Genetic Algorithms*. San Mateo, CA:Morgan Kaufmann Publishers.

Ng, K. P. & Wong, K. C. (1995). A new diploid scheme and dominance change mechanism for non-stationary function optimization. In L. J. Eshelman (ed.), *Proceedings of the Sixth International Conference on Genetic Algorithms*. San Mateo, CA:Morgan Kaufmann Publishers.

Nix, A. & Vose, M. D. (1992). Modeling genetic algorithms with Markov chains, *Annals of Mathematics and Artificial Intelligence* 5, 79-88.

Spears, W. M. (1994). Simple subpopulation schemes, in *Proceedings of the 94 Evolutionary Programming Conference*, San Diego, CA.

Syswerda, G. (1989). Uniform crossover in genetic algorithms. In J. D. Schaffer (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms.* San Mateo, CA: Morgan Kaufmann, 2-9.

Whitley, D. (1988). GENITOR: a different genetic algorithm. *In Proceedings of the Rocky Mountain Conference on Artificial Intelligence*. Denver Colorado, 118-130.