Migration Policies, Selection Pressure, and Parallel Evolutionary Algorithms

Erick Cantú-Paz
Department of Computer Science and Illinois Genetic Algorithms Laboratory
University of Illinois at Urbana-Champaign
`cantupaz@illigal.ge.uiuc.edu`

Abstract

This paper investigates how the policy used to select migrants and replacements affects the selection pressure in parallel evolutionary algorithms (EAs) with multiple populations. The four possible combinations of random and fitness-based emigration and replacement of existing individuals are considered. For each combination, the paper quantifies the increase in the selection intensity, which is a domain-independent adimensional quantity that can be used to compare the selection pressure of common selection methods with the pressure caused by migration. The work presented here confirms previous qualitative observations based on calculations of takeover times. The results may help to avoid excessively high (or low) selection pressures that may cause the search to fail, and offer a plausible explanation to the frequent claims of superlinear speedups in parallel EAs.

1 INTRODUCTION

Parallel evolutionary algorithms (EAs) have received considerable attention because of their potential to reduce the execution time in complex applications. One common method to parallelize EAs is to use multiple demes (populations) that occasionally exchange some individuals in a process called migration. There are two alternatives to select the individuals that emigrate from a deme. They can be chosen randomly or by selecting the best individuals in a deme. Likewise, there are two choices to replace existing individuals in the receiving deme with the incoming migrants: choose randomly or replace the worst.

This paper shows that the migration policy affects considerably the selection pressure, and that it influences significantly the speed of convergence of the algorithm. Although it has been recognized before that fitness-based selection of emigrants and replacements increases the selection pressure (Whitley, 1993; Whitley, Rana, & Heckendorn, 1999; Cantú-Paz, 1999), its impact on the algorithm's convergence is not well understood. The objective of this paper is to quantify accurately the additional selection pressure caused by migration. This leads to accurate predictions of the number of generations until convergence. In addition, understanding the effect of migration on the selection pressure is important because excessively slow or fast convergence rates may cause the search to fail (Goldberg, Deb, & Thierens, 1993; Thierens & Goldberg, 1993). If selection is too weak the population may drift aimlessly for a long time, and the quality of the solutions found is not likely to be good. On the other hand, rapid convergence is desirable, but an excessively fast convergence may cause the EA to converge prematurely to a suboptimal solution.

This investigation also sheds some light on the frequent claims of superlinear speedups in parallel EAs (see for example (Tanese, 1989; Belding, 1995; Punch, 1998)). Superlinear speedups imply that the *total* amount of work in the parallel EA is less than the work that a serial EA requires to reach the same solution (Punch, 1998). So we are left with a perplexing question: What caused the reduction in the work? The answer offered by this paper is that the additional selection pressure caused by migration reduces significantly the number of generations until convergence.

This paper assumes that migration occurs every generation, which is an upper bound on the migration frequency. Less frequent migrations are expected to have a lower impact on the convergence of the algorithm. In addition, the calculations assume that the migrants are copies of the individuals selected to migrate. This differs from natural populations, but is the most frequently-used form of migration in EAs.

The next section reviews common selection methods and explains intuitively the concept of selection pressure. The calculation of the selection intensity caused by the different migration policies is in section 3, and the results of experiments that validate the accuracy of the models are in section 4. Section 5 has a brief discussion of superlinear speedups and illustrates how some of them can be explained by an increase in selection pressure. Finally, section 6 summarizes the findings of this study and discusses issues for future research.

2 SELECTION PRESSURE

There are many different mechanisms used in evolutionary algorithms to select the parents of the next generation. They can be classified roughly into two groups: fitness-proportionate and rank-based selection. Fitness-proportionate methods select individuals probabilistically depending on the ratio of their fitness and the average fitness of the population. These methods are amongst the earliest methods used in EAs. Some examples are roulette-wheel selection (or also called proportionate selection) (Holland, 1975), stochastic remainder selection (Booker, 1982), and stochastic universal selection (Baker, 1987). Roulette-wheel selection uses a simulated biased roulette wheel with slots that are sized according to the fitness of each individual. The roulette is spinned once for each individual to be selected. The other schemes were proposed to reduce the stochastic error associated with spinning the roulette wheel numerous times.

Some common rank-based selection methods are linear ranking (Baker, 1985), tournament selection (Brindle, 1981), $(\mu \overset{+}{,} \lambda)$ selection (Schwefel, 1981), and truncation selection (Mühlenbein & Schlierkamp-Voosen, 1993). In linear ranking selection, individuals are selected with a probability that is linearly proportional to the rank of the individuals in the population. The desired expected number of copies of the best (n^+) and worst (n^-) individuals are supplied as parameters to the algorithm. In tournament selection a random sample of s individuals is selected (with or without replacement), and the best individual is selected. The process is repeated until the mating pool is filled. In $(\mu + \lambda)$ selection, λ offspring are created from μ parents, and the μ best individuals out of the union of parents and offspring are selected. In (μ, λ) selection $(\lambda \geq \mu)$ the μ best offspring are selected to survive. Truncation selection selects the top $1/\tau$ of the population and creates τ copies of each individual. It is equivalent to (μ, λ) selection with $\mu = \lambda/\tau$.

All of these selection mechanisms have the same purpose of creating more copies of the individuals with higher fitness than of those with low fitness. However, the selection mechanisms differ in the manner in which they allocate copies to the fittest individuals. A selection method has a higher selection pressure than another if it makes more copies of the best individuals, thereby eliminating rapidly the low-fit individuals. A strong selection method reaches equilibrium faster than a weaker method, but it also sacrifices genetic diversity that may be needed to find an adequate solution.

The parameters of selection methods regulate the selection pressure, which in turn determines how fast the algorithms converge. The speed of convergence of different selection schemes was first studied by Goldberg and Deb (1991), who introduced the concept of takeover time. The takeover time is the number of generations that selection alone requires to replicate a single individual of the best class until the population is full. Later, Mühlenbein and Schlierkamp-Voosen (1993) used concepts from population genetics to study convergence properties of a particular EA and introduced the use of the selection intensity to study the convergence of selection schemes. Of course, many others have studied and compared different selection methods used in EAs (for example, see the papers by Bäck (1994) and by Hancock (1997))

3 SELECTION INTENSITY

Cantú-Paz (1999) extended Goldberg and Deb's calculation of takeover times for tournament selection to the case with multiple demes. Those simple calculations suggested that the time until convergence is affected greatly by the migration policy, and therefore it must be taken into account when designing parallel EAs. Those results also suggested that the choice of migrants is a greater factor in the convergence speed than the choice of replacements.

This section follows a different approach to estimate the convergence times of parallel GAs. The assumptions used here are more realistic than in the previous study, and the results are very accurate predictions of the convergence time. First, the concept of selection intensity is defined and closed-form expressions for the intensity of common selection methods is provided as a reference. Then, we derive the additional selection intensity caused by migration. We show how these values can be used to predict the convergence times, and the next section verifies empirically the predictions.

Selection	**Parameters**	**I**
Tournament	s	$\mu_{s:s}$
(μ, λ)	μ, λ	$\frac{1}{\mu}\sum_{i=\lambda-\mu+1}^{\lambda} \mu_{i:\lambda}$
Linear Ranking	n^+	$(n^+ - 1)\frac{1}{\sqrt{\pi}}$
Proportional	σ_t, μ_t	σ_t / μ_t

Table 1: Selection intensity for common selection schemes.

The first observation needed to calculate the intensity of a selection method is that the average fitness of the individuals that are selected to survive is greater than the average fitness of the population. The increase depends on the selection method being used, and it can be quantified as the *selection differential*

$$s^t = \bar{f}_s^t - \bar{f}^t, \tag{1}$$

which is the difference between the mean fitness of the selected individuals and the mean fitness of the population (Mühlenbein & Schlierkamp-Voosen, 1993). The superscript t denotes the generation number. Assuming that the fitness values have a normal distribution, the selection differential can be calculated as

$$s^t = I\sigma_t, \tag{2}$$

where σ_t is the standard deviation of the population at time t, and the factor I is the *selection intensity.* The selection intensity of some common selection schemes has been calculated analytically. Bäck (1995) and Miller and Goldberg (1995) independently derived the selection intensity for tournament selection, and Bäck (1995) also derived I for (μ, λ) selection. Blickle and Thiele (1997) calculated the intensity of linear ranking, and Mühlenbein and Schlierkamp-Voosen (1993) calculated I for proportional selection. Table 1 contains the known selection intensities (adapted from (Miller & Goldberg, 1997)). Note that I is independent of the distribution of the current population, except for proportional selection.

Calculating the standard deviation is more complicated, and it depends on the fitness function. For the case of a l-bit OneMax function $f = \sum_{i=0}^{l} x_i$ (x_i is the i-th bit of the string x), uniform crossover creates an approximately binomial fitness distribution with probability P_t, where P_t is the proportion of bits set to one in generation t. Therefore, the standard deviation may be calculated as (Mühlenbein & Schlierkamp-Voosen, 1993; Thierens & Goldberg, 1994):

$$\sigma_t = \sqrt{l \; P_t(1 - P_t)}. \tag{3}$$

The selection differential $\bar{f}^{t+1} - \bar{f}^t = \sigma_t \cdot I$ can be written as

$$P_{t+1} - P_t = \frac{I}{\sqrt{l}}\sqrt{P_t(1 - P_t)}.$$

This can be approximated as a differential equation that can be solved to obtain the proportion of correct bits at time t as

$$P_t = 0.5\left(1 + \sin\left(\frac{I}{\sqrt{l}}t + \arcsin(2P_0 - 1)\right)\right), \tag{4}$$

where P_0 is the initial proportion of bits correct. In the case of the OneMax function this is usually 0.5. The number of generations until convergence can be calculated making $P_t = 1$ and solving for t to obtain

$$G = \left(\frac{\pi}{2} - \arcsin(2P_0 - 1)\right)\frac{\sqrt{l}}{I}, \tag{5}$$

which in the usual case when $P_0 = 0.5$ simplifies to $G = \frac{\pi}{2}\frac{\sqrt{l}}{I}$.

Our interest in this paper is to use the equation above to calculate the number of generations until convergence. For this, we must calculate the selection intensity caused by migration.

3.1 MIGRATION AND SELECTION INTENSITY

The remainder of the section follows closely the notation and method used by Bäck (1995) in his study of (μ, λ) selection. This should not imply that fitness-based migration is equivalent to (μ, λ) selection. In fact it is not. Bäck's paper provides a good framework and some useful approximations that we adapt here for our own purposes.

To calculate the selection intensity, we must calculate the selection differential between the mean fitness of a deme after migration $\bar{f}^{t'}$ and before migration $\bar{f}^t$:

$$s^t = \bar{f}^{t'} - \bar{f}^t. \tag{6}$$

The average fitness before migration is simply $\bar{f}^t = \frac{1}{n}\sum_{i=1}^{n} f_i^t$, where f_i^t is the fitness of the i-th individual of the population at generation t. The average fitness after migration can be written as a weighted sum of the average fitness of the migrants and the average fitness of the individuals that survive migration (i.e., are not replaced by migrants). Let δ denote the degree of the topology (the number of neighbors of a deme), and $m = \rho n$ the number of migrants from *one* deme. We can write

$$\bar{f}^{t'} = \frac{1}{n}\left(\delta m \bar{f}^t_{mig} + (n - \delta m)\bar{f}^t_{sur}\right), \tag{7}$$

where $\bar{f}^t_{mig}$ is the mean fitness of migrants from one deme, and $\bar{f}^t_{sur}$ is the mean fitness of the $(n - \delta m)$ survivors. Similarly we can write

$$\bar{f}^t = \frac{1}{n}\left(\delta m \bar{f}^t + (n - \delta m)\bar{f}^t\right). \tag{8}$$

Grouping similar terms we can decompose the selection differential into two parts. One that corresponds to the selection of emigrants s_e^t and another that corresponds to the selection of replacements s_r^t:

$$\begin{aligned} s^t &= s_e^t + s_r^t \\ &= \frac{1}{n}\delta m(\bar{f}^t_{mig} - \bar{f}^t) + \frac{1}{n}(n - \delta m)(\bar{f}^t_{sur} - \bar{f}^t). \end{aligned} \tag{9}$$

By writing the selection differential in this way, we can separate the calculation of the selection intensity into two independent steps. First, we consider the intensity caused by selecting the best individuals to emigrate. We shall see that selecting emigrants randomly has no effect on the selection intensity. Later, similar calculations are performed to calculate the intensity caused by replacing individuals.

The major assumption that we make is that the fitness values of the population at time t have a normal distribution. Under this assumption, the fitness values f_i^t can be interpreted as samples of random variables F_i^t with a common distribution $N(\bar{f}^t, \sigma_t)$. We may arrange the random variables in increasing order as

$$F^t_{1:n} \leq F^t_{2:n} \leq \cdots \leq F^t_{n:n}.$$

These are the order statistics of the F_i^t variables, and we can use them to calculate the average fitness of the emigrants and the survivors. Without loss of generality we assume a maximization problem. The mean fitness of the $m = \rho n$ best individuals selected to migrate from *one* deme is

$$\bar{f}^t_{mig} = \frac{1}{m}\sum_{i=n-m+1}^{n} E(F^t_{i:n}). \tag{10}$$

The random variables can be normalized as

$$Z_{i:n} = \frac{F^t_{i:n} - \bar{f}^t}{\sigma_t} \sim N(0, 1),$$

and the average fitness of the migrants may be rewritten in terms of the normalized variables

$$\begin{aligned} \bar{f}^t_{mig} &= \frac{1}{m}\sum_{i=n-m+1}^{n} (E(Z_{i:n})\sigma_t + \bar{f}^t) \\ &= \sigma_t \cdot \frac{1}{m}\sum_{i=n-m+1}^{n} E(Z_{i:n}) + \bar{f}^t. \end{aligned} \tag{11}$$

Now, we can calculate the selection differential caused by the migrants as

$$s_e^t = \frac{\delta m}{n}(\bar{f}^t_{mig} - \bar{f}^t) = \frac{1}{n}\cdot\sigma_t\cdot\delta\sum_{i=n-m+1}^{n} E(Z_{i:n}). \tag{12}$$

Since the selection differential is also defined as $s^t = I \cdot \sigma_t$ (equation 2), the selection intensity caused by selecting the best individuals to emigrate is

$$I_e = \frac{1}{n}\cdot\delta\cdot\sum_{i=n-m+1}^{n} E(Z_{i:n}). \tag{13}$$

The expected value of the i-th order statistic of a sample of size n is defined as

$$\begin{aligned} \mu_{i:n} &= E(Z_{i:n}) \\ &= n\binom{n-1}{i-1}\int_{-\infty}^{\infty} z\phi(z)\Phi^{i-1}(z)[1 - \Phi(z)]^{n-i}dz, \end{aligned} \tag{14}$$

where $\phi(z) = \exp(-z^2/2)/\sqrt{2\pi}$ and $\Phi(z) = \int_{-\infty}^{z}\phi(x)dx$ are the PDF and CDF respectively of a standard normal distribution with mean 0 and standard deviation of 1. The values of $\mu_{i:n}$ are computationally expensive to calculate, but they are tabulated for $n \leq 400$ (Harter, 1970). Nevertheless, computing the sum in equation 13 can be cumbersome, but fortunately the following approximation exists[1] (Burrows, 1972; Bäck, 1995):

$$\sum_{i=n-m+1}^{n} \mu_{i:n} \approx n\phi(\Phi^{-1}(1-\rho)), \tag{15}$$

and therefore equation 13 can be approximated as

$$I_e \approx \delta\phi(\Phi^{-1}(1-\rho)). \tag{16}$$

It is important to realize that the selection intensity is an adimensional quantity that does not depend on the fitness function or on the generation t. The only assumption made to calculate the intensity is that the fitness values have a normal distribution, but any other distribution may be used as long as $E(F_{i:n})$ may be computed (by substituting the appropriate PDF and CDF in equation 14). The selection intensity depends on the population size (see equation 13), but its approximation depends only on the migration rate.

In the case where individuals are chosen randomly to emigrate, the expected fitness of the migrants is the same as the mean fitness of the population, and therefore the selection differential is $s^t = \bar{f}^t_{mig} - \bar{f}^t = 0$, and there is no additional selection intensity ($I_e = 0$). However, there may be an increase in the overall selection intensity if the migrants replace the worst individuals in the target deme. The replacement of individuals is treated in the following paragraphs.

Replacing the *worst* individuals in a deme with migrants causes an increase in the average fitness of the deme. In a manner similar as above, we can calculate the mean fitness of the individuals that survive (i.e., are not replaced by the δm migrants) as

$$\begin{aligned}\bar{f}^t_{sur} &= \frac{1}{n-\delta m}\sum_{i=\delta m+1}^{n} E(F_{i:n}) \\ &= \sigma_t \cdot \frac{1}{n-\delta m}\sum_{i=\delta m+1}^{n} \mu_{i:n} + \bar{f}^t.\end{aligned} \tag{17}$$

In this case the response to selection is the difference between the mean of the individuals that survive and the mean fitness of the population:

$$\begin{aligned}s^t_r &= \frac{1}{n}(n-\delta m)(\bar{f}^t_{sur} - \bar{f}^t) \\ &= \sigma_t \cdot \frac{1}{n}\sum_{i=\delta m+1}^{n} \mu_{i:n} \\ &= \sigma_t \cdot I.\end{aligned} \tag{18}$$

Therefore, the selection intensity caused by choosing the worst individuals to be replaced is

$$\begin{aligned}I_r &= \tfrac{1}{n}\textstyle\sum_{i=\delta m+1}^{n} \mu_{i:n} && (19)\\ &\approx \phi(\Phi^{-1}(1-\delta\rho)). && (20)\end{aligned}$$

In this case, the maximum of I_r is $\phi(0) = 1/\sqrt{2\pi} = 0.3989$, which is a fairly low value, but it is not negligible. When the migration policy is to replace individuals randomly in the target deme, there is no differential between the average fitness of the individuals that survive and the average fitness of the entire deme. Therefore, in this case $I_r = 0$.

The overall selection intensity caused by migration is simply

$$I_m = I_e + I_r. \tag{21}$$

To predict the number of generations until convergence, we should add I_m to the intensity from the selection method used to select the parents of the next generation in each deme (given in table 1). The total intensity is then used in equation 5.

3.2 COMPARING MIGRATION POLICIES

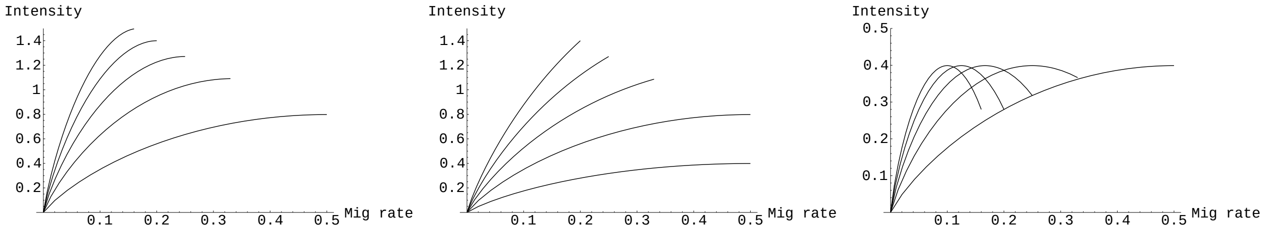
Figure 1 presents plots of I_m for topologies with different degrees and varying the migration rate[2]. From the plots it is easy to see that the migration policy with the highest intensity is when the best individuals migrate and replace the worst, followed closely by the case when the best migrants replace random individuals. The difference between these two policies is not as large as the difference with the policy where migrants are selected randomly, as predicted from the takeover time calculations.

Observe that the magnitudes of I_m are fairly large. For example, in the case where the best migrants replace the worst individuals, a hypercube of degree $\delta = 3$

---

[1] Bäck shows that for $n > 50$ the approximation is indistinguishable from the real values.

[2] $\Phi^{-1}(x)$ in equations 16 and 20 was calculated numerically using Mathematica 3.0 as $\sqrt{2}$`InverseErf[0,2x-1]`



(a) Best replace worst (eq. 21) (b) Best replace random (eq. 16) (c) Random replace worst (eq. 20)

Figure 1: Selection intensity for different migration policies varying the number of neighbors (δ=1,2,..,5 from bottom to top) and the migration rate.

and using only 5% migration produces approximately the same selection intensity as tournament selection with $s = 3$. Migration alone is capable of producing significant selection pressures and can cause the populations to converge fast. Notice also that when the best individuals migrate, the relation between δ and ρ is not linear. The selection intensity produced by two neighbors using $\rho = 0.05$ is greater than the selection intensity caused by one neighbor sending $\rho = 0.10$ of its individuals.

4 EXPERIMENTS

This section presents experimental evidence that verifies the accuracy of the predictions of the previous section. Experiments are performed with the four migration policies, and the results shown are the average of 20 independent runs for each parameter setting.

All experiments use a 500-bit OneMax function, and the populations are initialized randomly (on average $P_0 = 0.5$). Each deme is a generational GA with $n = 100$ individuals, which is sufficient to ensure convergence to the optimum in all cases. The GAs use pairwise tournament selection ($s = 2$, $I = 0.5642$), uniform crossover with probability 1.0, and no mutation. Migration occurs every generation.

4.1 BEST MIGRANTS REPLACE WORST INDIVIDUALS

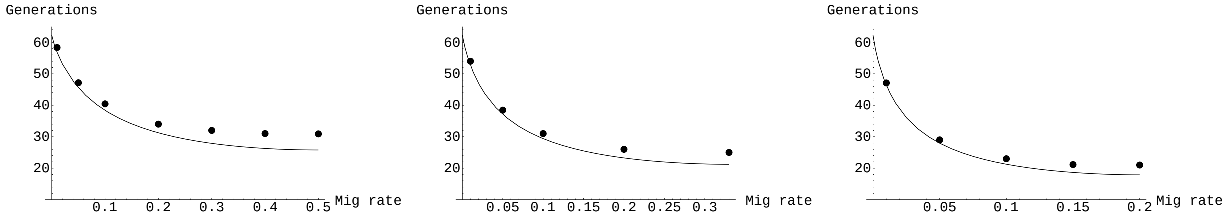
The first set of experiments is designed to verify the prediction of the number of generations until convergence (eq. 5). In this migration policy, both the choices of migrants and replacements increase the selection pressure, and therefore I_m is calculated by adding equations 16 and 20. Then, the selection intensity from pairwise tournament selection is added, and the result is used as I in equation 5. The theoretical predictions and the experimental results are presented in figure 2. Experiments with other topologies of the same degree show no difference. In particular, experiments with 8 demes connected as uni- and bi-directional rings ($\delta = 1$ and 2, respectively) and hypercubes of degree $\delta = 3$ and 4 (16 demes) yield identical results as those with fully-connected topologies in figure 2. For this reason, in the remainder we experiment only with fully connected topologies.

4.2 BEST MIGRANTS REPLACE RANDOM INDIVIDUALS

For this migration policy, I_e is given by equation 16 and $I_r = 0$. The results are presented in figure 3. We can observe that the generations until convergence are only slightly fewer than in the previous case, where the replacements were chosen according to their fitness. This observation supports the hypothesis raised by the analysis of takeover times that indicated that the major component in the increase of selection pressure is the selection of emigrants, not the selection of replacements.

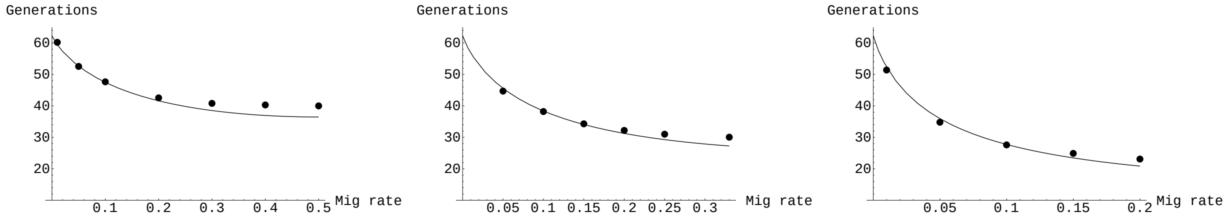
4.3 RANDOM MIGRANTS REPLACE WORST INDIVIDUALS

In this migration policy, there is no selection pressure caused from selecting random emigrants ($I_e = 0$), but replacing the worst individuals in the target deme causes the intensity to increase. I_r is given by equation 20. We can observe in figure 4 that since the additional selection pressure caused by this migration policy is not very strong, the generations until convergence do not decrease as much as in the previous cases. We used migration rates higher than the maximum $\rho^* = 1/(2\delta)$ rate that makes sense in this case to



(a) $\delta = 1$ (b) $\delta = 2$ (c) $\delta = 4$

Figure 2: Generations until convergence when the *best* individuals are selected to migrate and replace the *worst*. The line is eq. 5 and the dots are the experimental results.



(a) $\delta = 1$ (b) $\delta = 2$ (c) $\delta = 4$

Figure 3: Generations until convergence when the *best* individuals are selected to migrate and replace *random* individuals. The line is eq. 5 and the dots are the experimental results.

be consistent with the previous experiments and to illustrate how the selection pressure decreases after this point.

5 SUPERLINEAR SPEEDUPS

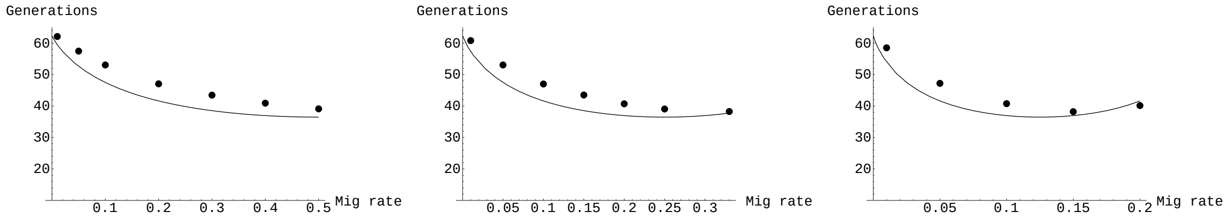
To illustrate how the different convergence times affect the speedups of parallel GAs consider an example with the 500-bit OneMax problem used previously. A simple GA with a single population of $n = 100$ individuals and the same parameters as before reached the global solution in all of 20 independent trials. The execution time of the serial GA can be estimated as $T_s = G_s n T_f$, where G_s is the number of generations until convergence and T_f is the function evaluation time. With $l = 500$, $P_0 = 0.5$, and using pairwise tournaments ($I = 0.5642$), equation 5 gives $G_s = 62$. For the purposes of the example, let $T_f = 1$ unit of time, and therefore $T_s = 6200$.

For the parallel case, consider $r = 4$ demes with $n = 25$ individuals each. The demes are connected by a bidirectional ring ($\delta = 2$) and use a migration rate of $\rho = 0.05$. We validated empirically that this configuration reaches the optimum in all of 20 independent runs. The execution time in the parallel case is $T_p = G_p n T_f + T_c$, where T_c is the time used in communications. It depends on δ and ρ, and for our example we will use a large value $T_c = 100$[3]. For this configuration, $I_m = 0.381$. Substituting $I = 0.5642 + 0.381$ in equation 5 gives $G_p = 37$. With these values $T_p = 1025$, and the ratio $\frac{T_s}{T_p} = 6.05$, which even with the conservative values used here is much higher than the expected ideal of 4.

We should not conclude that all claims of superlinear speedups are caused by an increase of selection pressure due to migration. Other possible causes are implementation particulars (for example, the smaller demes may fit completely in the processors' caches, which was the explanation that Belding (1995) gave for his results) or inappropriate sizing of populations (such that convergence to solutions of the same quality is not guaranteed).

---

[3] Notice that the ratio of T_c/T_f is extremely high. In most practical applications of parallel GAs this ratio is much lower than one. We are being overly conservative here.



(a) $\delta = 1$ (b) $\delta = 2$ (c) $\delta = 4$

Figure 4: Generations until convergence when *random* individuals are selected to migrate and replace the *worst*. The line is eq. 5 and the dots are the experimental results.

6 SUMMARY

The choice of migrants and the replacement of individuals are not often considered important parameters of parallel EAs. However, this paper showed that choosing migrants or replacements according to their fitness increases the selection pressure and may cause the algorithm to converge significantly faster. The migration policy that causes the greatest reduction in work is to choose both the migrants and the replacements according to their fitness, which is also the most common policy. The results of the paper also indicate that the selection pressure increases monotonically with higher migration rates. These results are consistent with previous qualitative observations (Cantú-Paz, 1999). The faster convergence may result in a reduction of the total computational work, and may explain some claims of superlinear speedups in parallel EAs.

ACKNOWLEDGMENTS

I wish to thank Prof. David E. Goldberg for his guidance during the research reported here. I was partially supported by a Fulbright-García Robles Fellowship. The work was sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-97-1-0050. Research funding for this project was also provided by a grant from the US Army Research Laboratory Program, Cooperative Agreement DAAL01-96-2-003. The US Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies and endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the US Government.

References

Bäck, T. (1994). Selective pressure in evolutionary algorithms: A characterization of selection mechanisms. In *Proceedings of the First IEEE Conference on Evolutionary Computation*, Volume 1 (pp. 57–62). Piscataway, NJ: IEEE Service Center.

Bäck, T. (1995). Generalized convergence models for tournament- and (μ, λ)-selection. See Eschelman (1995), pp. 2–8.

Baker, J. E. (1985). Adaptive selection methods for genetic algorithms. In Grefenstette, J. J. (Ed.), *Proceedings of an International Conference on Genetic Algorithms and Their Applications* (pp. 101–111). Hillsdale, NJ: Lawrence Erlbaum Associates.

Baker, J. E. (1987). Reducing bias and ineffiency in the selection algorithm. In Grefenstette, J. J. (Ed.), *Proceedings of the Second International Conference on Genetic Algorithms* (pp. 14–21). Hillsdale, NJ: Lawrence Erlbaum Associates.

Belding, T. C. (1995). The distributed genetic algorithm revisited. See Eschelman (1995), pp. 114–121.

Blickle, T., & Thiele, L. (1997). A comparison of selection schemes used in evolutionary algorithms. *Evolutionary Computation*, *4*(4), 361–394.

Booker, L. B. (1982). *Intelligent behavior as an adaptation to the task environment.* Doctoral dissertation, The University of Michigan. (University Microfilms No. 8214966).

Brindle, A. (1981). *Genetic algorithms for function optimization.* Unpublished doctoral dissertation, University of Alberta, Edmonton, Canada.

Burrows, P. (1972). Expected selection differentials for directional selection. *Biometrics*, *23*, 1091–1100.

Cantú-Paz, E. (1999). Migration policies and takeover times in parallel genetic algorithms. In Banzhaf, W. (Ed.), *Proceedings of the 1999 Genetic and Evolutionary Computation Conference*. San Mateo, CA: Morgan Kaufmann Publishers.

Eschelman, L. (Ed.) (1995). *Proceedings of the Sixth International Conference on Genetic Algorithms*. San Francisco, CA: Morgan Kaufmann.

Goldberg, D. E., & Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms. *Foundations of Genetic Algorithms*, *1*, 69–93. (Also TCGA Report 90007).

Goldberg, D. E., Deb, K., & Thierens, D. (1993). Toward a better understanding of mixing in genetic algorithms. *Journal of the Society of Instrument and Control Engineers*, *32*(1), 10–16.

Hancock, P. J. B. (1997). Selection: a comparison of selection mechanisms. In Bäck, T., Fogel, D. B., & Michalewicz, Z. (Eds.), *Handbook of Evolutionary Computation* (pp. C2.8:1–C2.8:11). Bristol and New York: Institute of Physics Publishing and Oxford University Press.

Harter, H. L. (1970). *Order statistics and their use in testing and estimation*. Washington, D.C.: U.S. Government Printing Office.

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

Miller, B. L., & Goldberg, D. E. (1995). Genetic algorithms, tournament selection, and the effects of noise. *Complex Systems*, *9*(3), 193–212.

Miller, B. L., & Goldberg, D. E. (1997). Genetic algorithms, selection schemes, and the varying effects of noise. *Evolutionary Computation*, *4*(2), 113–131.

Mühlenbein, H., & Schlierkamp-Voosen, D. (1993). Predictive models for the breeder genetic algorithm: I. Continuous parameter optimization. *Evolutionary Computation*, *1*(1), 25–49.

Punch, W. F. (1998). How effective are multiple programs in genetic programming. In Koza, J. R., Banzhaf, W., Chellapilla, K., Deb, K., Dorigo, M., Fogel, D. B., Garzon, M. H., Goldberg, D. E., Iba, H., & Riolo, R. L. (Eds.), *Genetic Programming 98* (pp. 308–313). San Francisco: Morgan Kaufmann Publishers.

Schwefel, H. (1981). *Numerical optimization of computer models*. Chichester: John Wiley and Sons.

Tanese, R. (1989). Distributed genetic algorithms. In Schaffer, J. D. (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms* (pp. 434–439). San Mateo, CA: Morgan Kaufmann.

Thierens, D., & Goldberg, D. E. (1993). Mixing in genetic algorithms. In Forrest, S. (Ed.), *Proceedings of the Fifth International Conference on Genetic Algorithms* (pp. 38–45). San Mateo, CA: Morgan Kaufmann.

Thierens, D., & Goldberg, D. E. (1994). Convergence models of genetic algorithm selection schemes. In Davidor, Y., Schwefel, H.-P., & Männer, R. (Eds.), *Parallel Problem Solving fron Nature, PPSN III* (pp. 119–129). Berlin: Springer-Verlag.

Whitley, D. (1993). An executable model of a simple genetic algorithm. In Whitley, L. D. (Ed.), *Foundations of Genetic Algorithms 2* (pp. 45–62). San Mateo, CA: Morgan Kaufmann.

Whitley, D., Rana, S., & Heckendorn, R. B. (1999). Exploiting separability in search: The island model genetic algorithm. *Journal of Computing and Information Technology*. forthcoming.