Asexual Versus Sexual Reproduction in Genetic Algorithms[1]

Wendy Ann Deslauriers (wendyd@alumni.princeton.edu)
Institute of Cognitive Science,Room 2201, Dunton Tower
Carleton University, 1125 Colonel By Drive
Ottawa, Ontario, K1S 5B6

Abstract

This study examined the effectiveness of standard genetic algorithms in resolving a 20 city travelling salesperson scenario. The results of standard algorithms were compared to that of competitive/co-operative sexual selection for genetic algorithms developed by Sanchez-Velazco and Bullinaria (2003). A range of parameters was used with both the standard and genetic algorithms to determine the validity of their conclusion that sexual selection is a superior evolutionary approach. This study suggests that gendered reproduction may be useful in some, but not all, contexts.

Introduction

Genetic algorithms can be a successful form of computational modelling. The standard genetic algorithm consists of a population of genetic individuals designed to resolve a problem. Every generation, genes are evaluated for fitness in order to determine quality. Selection methods are used to determine which genes will reproduce. Reproduction involves a pairing up of selected individuals to swap information through a crossover operation. Between successive generations, each individual's genes are subject to probabilistic mutations.

Sanchez-Velazco and Bullinaria (2003) used a roulette selection method to identify individuals for reproduction. Other possibilities investigated in this study include tournament selection and biased tournament selection, in which a roulette method is used to select individuals for tournament evaluation. Variations on selection can be created by altering the space assigned to each individual on the roulette wheel. Sanchez-Velazco and Bullinaria (2003) used a roulette wheel that was divided according to individual fitnesses in linear proportion. Turning on ranking, changes this relationship to one in which individuals are given space according to their order within the population. In order to extend Sanchez-Velazco and Bullinaria's ideas this study investigated the same problem with a wider range of selection methods (tournament, biased tournament and rank selection). In some instances, elite individuals may be carried over from generation to generation.

Method

The Problem

As indicated by Sanchez-Velazco and Bullinaria (2003), 20 cities were arranged on the circumference of a circle. The task for the algorithm to solve was that of the travelling salesperson who visits each city once. In this study the cities were located equidistantly around the circle so that an ideal complete circuit (i.e. the shortest distance) would have a distance, d , of:

$$d = \frac{20r\sin 18^\circ}{\sin 81^\circ}$$

where r represents the radius of the circle. The fitness of each individual in the genetic algorithm was determined by dividing the shortest distance by the distance travelled by that individual's complete circuit. Thus, for each individual, its fitness, f , will have a value as follows:

$$0 < f \leq 1$$

The shorter the travelled route, the higher the fitness. An agent who has completely solved the problem will follow the edge of a 20-sided polygon and have a fitness of 1.

Original Standard Genetic Algorithm

In an attempt to replicate the findings of Sanchez-Velazco and Bullinaria (2003), an initial simulation was run using a standard genetic algorithm with the parameters indicated in their paper. These parameters included roulette selection, a population of 200, uniform crossover with a probability of 0.75, a mutation probability of 0.0505 and an elite count of 14.

Preliminary simulations suggested much greater success than that achieved by Sanchez-Velazco and Bullinaria (2003). Attempting to improve the replication an age-limit (of 5) was added to the standard genetic algorithm in line with the production capabilities of the sexual selection process. The age-limit brought simulation results closer to those of Sanchez-Velazco and Bullinaria (2003) and was therefore run for 2400 generations averaged over 100 experimental runs and used for all subsequent trials.

Since the results of this study were still achieving more successful results than the earlier study an attempt was made to alter mutation to occur between neighbouring cities. However, since this adaptation did not match Sanchez-Velazco and Bullinaria's (2003) results, and since their paper indicated that mutations consisted of swaps of two cities in the list, it was abandoned after the initial attempt and not included in future trials.

Standard Algorithm Parameter Search

Sanchez-Velazco and Bullinaria (2003) stated that the parameters they used were selected because they had proven to produce good results. However, it was unclear how this proof was obtained. An exploratory process was used to determine good parameters for resolving this 20 city travelling salesperson problem. The range of parameters included three selection methods (roulette, tournament and biased tournament), five populations (2, 10, 50, 100 and 200), five crossover rates (0.1, 0.3, 0.5, 0.7, 0.9), six elite

[1] Carleton University Cognitive Science Technical Report 2006-09. http://www.carleton.ca/ics/TechReports

counts (0, 2, 6, 10, 20, 50), five mutation rates (0.0001, 0.001, 0.01, 0.1, 0.5), and two rank options (on and off), for a total of 4500 initial simulations. When the elite count was greater than or equal to the population the populations would have been unable to evolve and those trials were eliminated. These initial simulations were averaged over 10 experimental runs for 50 generations each.

The results of the first round were organized to determine the parameters with the top 20 best individual fitnesses and the top 20 average fitnesses. After duplicates had been eliminated a list of 29 parameter settings remained. Since Sanchez-Velazco and Bullinaria (2003) had used roulette selection and only one roulette option remained, the four next best sets of parameters involving roulette selection were added to the list. These 34 parameter settings were run for 500 generations, also averaged over 10 experimental runs.

The results of the second round were organized to determine the top 5 best individual fitnesses and top 5 average fitnesses. It should be noted, that more than 5 of the 34 settings had achieved a best fitness of 1. These parameter settings were sorted secondarily by the average fitness of the same settings. An eleventh parameter setting was added to ensure that there were a minimum of two settings using each of the selection procedures. In order to explore the empty spaces in the six-dimensional parameter field, each of the selected 11 settings was used as an origin point for expansion along the axes of four parameters (population, crossover rate, elite count and mutation rate). Expansion along the axes involved using values halfway to the nearest value already used (i.e. parameters with a crossover rate of 0.7 were expanded to test crossover rates of 0.6 and 0.8.). Note, since the first round included the full spectrum of options regarding selection and rank they were excluded from further expansion. The results of 10 experimental runs were averaged for 500 generations using the resultant parameter settings.

The majority of this third round of trials resulted in a best fitness value of 1. Therefore, results were sorted primarily by best fitness value and secondarily by average fitness value. The best three parameter settings, plus a fourth to ensure inclusion of a roulette option, were selected for use in the fourth round. The parameter space was further explored by expanding one-quarter of the distance to the nearest original value along the axes of crossover rate, elite count and mutation rate. Population was eliminated from further exploration because a population of 200 was by far the most successful in all rounds to this point. The resultant parameter settings were evaluated for 500 generations, averaged over 10 experimental runs.

For the final round, four parameter settings were chosen. For tournament and roulette selection, the parameter settings with a best fitness of 1 and the best average fitness were chosen (tournament selection, population of 200, crossover probability of 0.8, elite count of 20, mutation probability of 0.1 and rank on; roulette selection, population of 200, crossover probability of 0.5, elite count of 28, mutation probability of 0.3 and rank on). However, for biased tournament selection the parameters that achieved a best value of 1, did not have the highest average fitness. Therefore, two sets of parameters were used for biased tournament selection (biased tournament selection 1, population of 200, crossover probability of 0.9, elite count of 8, mutation probability of 0.2 and rank off; biased tournament selection 2, population of 200, crossover probability of 0.8, elite count of 8, mutation probability of 0.1 and rank off). These four parameters settings were used for 100 experimental runs evolved through 2400 generations.

Original Gendered Algorithm

The gendered algorithm was an attempt to directly replicate the work of Sanchez-Velazco and Bullinaria (2003). The initial population of 200 was randomly separated as half male and half female and the gender ratio was held constant during successive generations. Other parameters used for this simulation were a crossover probability of 0.75, elite count of 14, female mutation probability of 0.001 and a male mutation probability of 0.1.

Reproduction occurred between individuals of opposite gender and involved the creation of two offspring, one daughter and one son. Selection of males for crossover was through a standard roulette wheel. However, selection of the female partner involved a complicated combination of the female's individual fitness, $f(x)$, the improvement in fitness of son over father from the preceding generation, $\Delta f(y)$, and an age related fertility factor, $g(Age(x))$. The equation used for selection was:

$$x_{sel} = Sel\ \frac{0.75f(x)+0.55\Delta f(y)+0.18g(Age(x))}{0.75+0.55+0.18}$$

The relation between age and fertility can be seen in Table 1 and a more detailed description of the selection process can be found in Sanchez-Velazco and Bullinaria (2003).

Table 1. Age and Fertility

$Age(x)$	0	1	2	3	4	5	6+
$g(Age(x))$	.5	.75	1.0	.75	.5	.25	0

Gendered crossover was accomplished by randomly choosing a single split point in each of the parents (s_x and s_y). The daughter is created from her mother's genetic material. She inherits her mother's trailing genes (those from the split point to the end), followed by the genes from the first part of the chromosome in reverse order. The son, inherits his mother's genes following the father's split point and then completes his chromosome using genes from his father in the order they appear. If his father's gene would create an illegal individual then the gene is discarded in favour of the next, until a complete, legal son is formed. This process can be seen in Figure 1. Results of the gendered simulation were averaged over 100 experimental runs through 2400 generations.

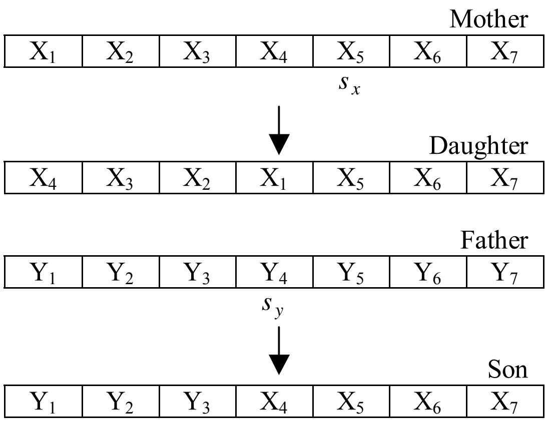

Figure 1.

Gendered Algorithm with Determined Parameters

In order to create comparative data an additional gendered algorithm was run for 2400 generations and averaged over 100 trials. The parameters for this algorithm were selected to match the parameters from the second biased tournament simulation in the standard genetic algorithm parameter search. This set of parameters was selected because its average fitnesses and best fitnesses showed little variance and the best fitness reached a level above 0.995. The ratio of the male:female mutation rates was set at 100:1, as in Sanchez-Velazco and Bullinaria (2003) for the original gendered algorithm, to average to 0.1 as a total mutation probability for the entire population.

Natural Parameters

Since Sanchez-Velazco and Bullinaria (2003) referred to natural evolution as a rationale for their genetic algorithm, both a standard and a gendered genetic algorithm were run to compare their success rates. As before, both were averaged over 100 experimental runs through 2400 generations. Population was set to 200, since that had been the most successful population value throughout. Crossover probability was set to 1.0 since natural gendered reproduction does not permit non-gendered reproduction. The elite count was set to 8 since this was the smallest elite number to produce good results in earlier trials and natural systems provide some incentive to retain elite individuals but are usually less successful than computer simulations. The combined mutation probability was set to a 0.000 004 with a male:female ratio, for the gendered version, of 50:3 as per information from the human genome project. Selection was set to biased tournament since it was the most widely successful method during the parameter search.

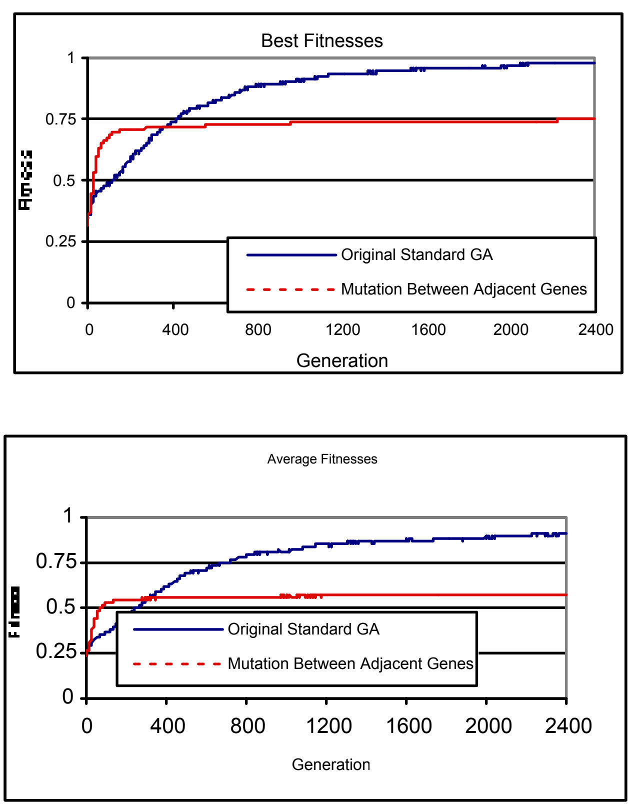

Figure 2. Comparison of initial standard genetic algorithms.

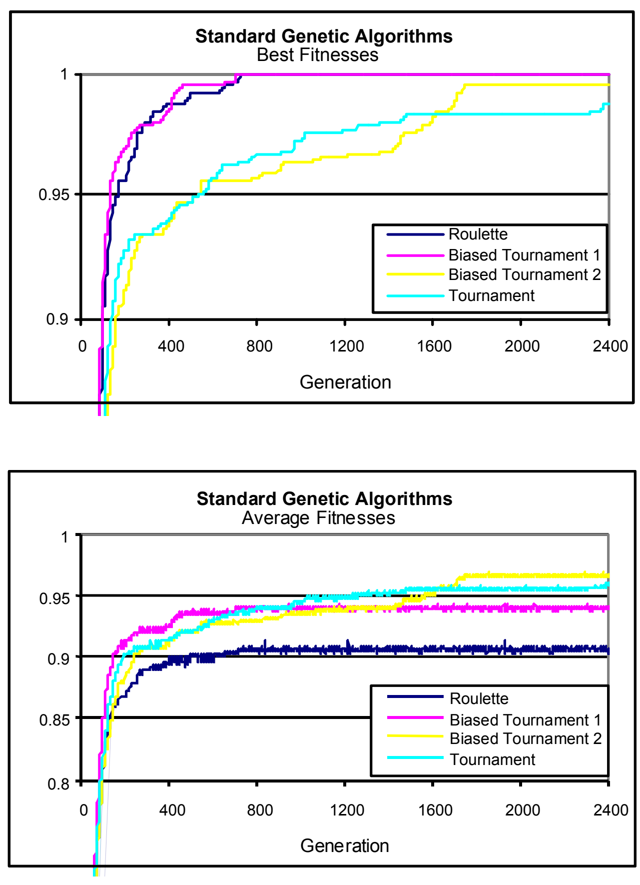

Figure 3. Comparison between selection methods in the standard genetic algorithm parameter search.

Results

Adjusting the mutation process to occur between all adjacent genes, was initially useful but created a flattening of the fitness curve (Figure 2). The adjusted algorithm improved quickly at the beginning but then hit a success limit within 100 generations. As previously indicated, this technique was not used for further experiments.

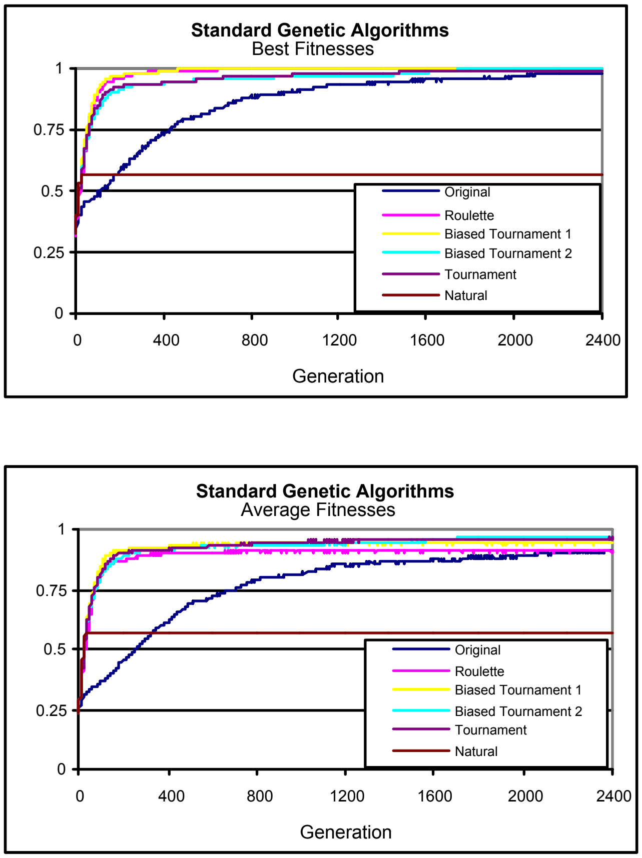

Figure 4. Comparison between all standard genetic algorithms.

The standard genetic algorithms, were all successful in solving the 20 city travelling salesperson problem. The set of parameters labelled Biased Tournament 1 achieved success fastest. The best individual in the population, achieved a perfect solution by 800 generations, however, the average fitness of the population levelled out at about the same point. Roulette selection produced the next most effective solution when only the best individual in the population is considered. However, the average fitness of the population was worst when roulette selection was used. Biased Tournament 2 and Tournament selection produced roughly similar results. The full results from the standard genetic algorithm parameter search can be seen in Figure 3.

When compared to both the original and natural genetic algorithms, all four of the parameter search settings were more successful (Figure 4). The natural algorithm improved quickly at the beginning, but flattened out, with both an average fitness and a best fitness of approximately 0.565, after 20 to 30 generations.

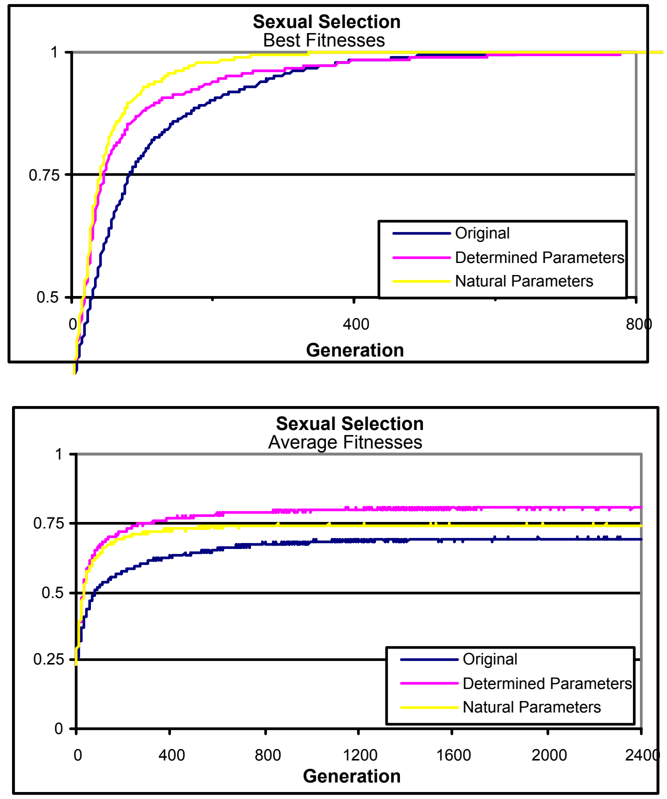

Figure 5. Comparison of gendered algorithms.

When sexual selection was used, the fitness variance between the assorted gendered algorithms was smaller than between the standard algorithms (Figure 5). However, the variance between the best fitnesses and the average fitnesses is greater than with the standard algorithm. When the best individuals in the population are examined, the natural parameters improve most rapidly and all three gendered algorithms reach a fitness of 1 by the 800th generation. The average fitness of the population is highest with the Biased Tournament 2 settings.

When the gendered algorithm is compared to the standard algorithm with the same parameter settings the results of this study differ from those of Sanchez-Velazco and Bullinaria (2003). Using both Sanchez-Velazco and Bullinaria's original settings, and the Biased Tournament 2 parameter settings, this study found that sexual selection improved the best individual in the population but slowed the improvement in average fitness. These results can be seen in Figures 6 and 7.

Only when using the natural parameter settings, does the gendered algorithm consistently outperform the standard algorithm. The gendered natural parameter algorithm achieved a best fitness of 1 by the 450th generation and the average fitness of the population levelled out around 0.75 after 800 generations.

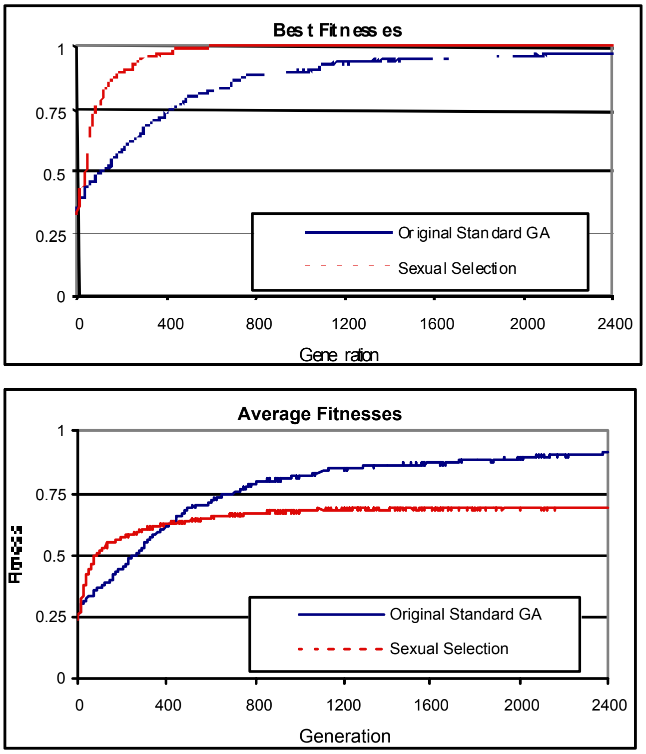

Figure 6. Comparison between the standard and gendered genetic algorithms using the original parameter set.

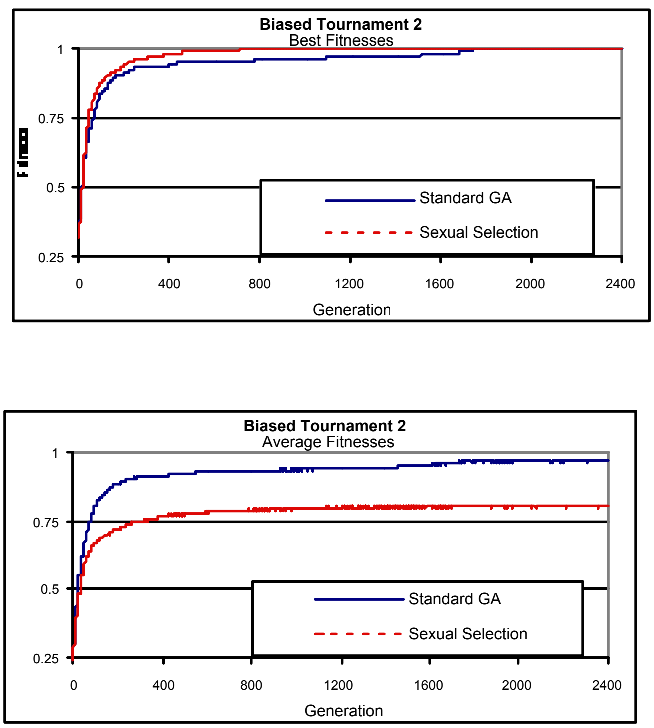

Figure 7. Comparison between the standard and gendered genetic algorithms using the Biased Tournament 2 parameter set.

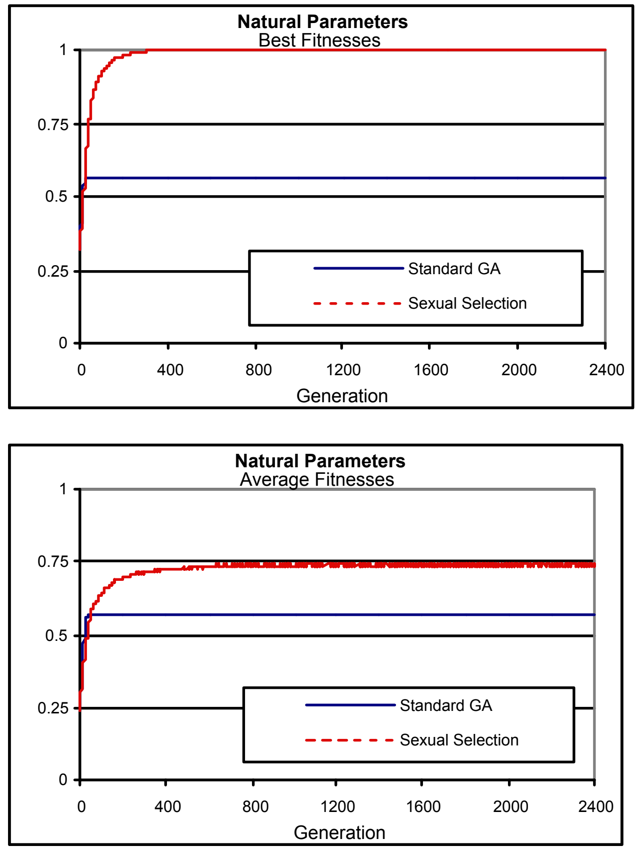

Figure 8. Comparison between the standard and gendered genetic algorithms using the natural parameter set.

Discussion

Although Sanchez-Velazco and Bullinaria (2003) suggested that gender and gendered reproduction are essential components of natural evolution, the results of this study do not support gender as a consistently preferable option. Sexual selection appears to improve the fitness of the best individual in the population but also spreads the variance within the population at the same time. Some biologists have hypothesized that human evolution has essentially stalled and others have suggested that mutation rates may have slowed during the evolutionary process. Further investigation into these options and their connection to genetic algorithms may elicit more information. Another important extension might involve comparing a progression from standard to gendered or from gendered to standard reproduction.

To determine whether a specific type of natural evolution is preferable, a survey of assorted mating patterns (i.e. mate for life, one individual mates with many, etc.) and mate selection (i.e. benefit to offspring, as in Sanchez-Velazco and Bullinaria (2003), similar fitness levels between parents, etc.) should be investigated. Also, natural evolution is normally accompanied by varied life spans, population growth and different environmental stimuli. Significant work would be required before this method could be said to accurately model natural processes.

In evaluation of gendered reproduction this study has indicated both advantages and disadvantages. The elite individuals are improved through sexual selection but the population as a whole was improved in only one of three simulations. Further study is necessary before determining whether gendered algorithms improve the evolutionary process.

Acknowledgements

Thanks to Terry C. Stewart for his assistance with coding.

References

Central Intelligence Agency. (2005, November). The world factbook (ISSN 1553-8133). Retrieved November 30, 2005, from http://www.cia.gov/cia/publications/factbook

Sanchez-Velazco, J., and Bullinaria, J. A. (2003) Sexual selection with competitive/co-operative operators for genetic algorithms. *Proceedings of the IASTED International Conference on Neural Networks and Computational Intelligence (NCI 2003)*, 389-048.

U.S. Department of Energy Office of Science, Office of Biological and Environmental Research, Human Genome Program. (2005, November) The science behind the human genome project: Basic genetics, genome draft sequence, and post-genome science. Retrieved November 30, 2005, from www.ornl.gov/hgmis