

Using Genetic Algorithms with Sexual or Asexual Transposition: a Comparative Study

Anabela Simões

Centre for Informatics and Systems of the University of
Coimbra, Polo II - Pinhal de Marrocos
3030 Coimbra - Portugal
abs@eden.dei.uc.pt

Ernesto Costa

Centre for Informatics and Systems of the University of
Coimbra, Polo II - Pinhal de Marrocos
3030 Coimbra - Portugal
ernesto@dei.uc.pt

Abstract - This paper presents the results obtained with a modified GA which uses a biologically inspired mechanism called transposition as the main genetic operator. Previous work has already focused the comparative analysis between the sexual and asexual forms of transposition and the standard crossover operators. The present work completes the comparative study, presenting the results obtained by the GA when using the different transposition mechanisms. The comparative study will analyze the similarities and differences of sexual and asexual transposition concerning the GA efficiency and the choice of one important parameter: the flanking sequence length.

1 Introduction

Genetic diversity is essential for the evolutionary process. When using genetic algorithms, a population evolves through the application of two main genetic operators: mutation and crossover. These operators allow changes in the individuals, creating evolutionary advantages in some of them. The fittest individuals are more likely to be selected allowing the evolution of the population to the best solution (Goldberg 1989).

In nature genetic diversity is caused and maintained by several mechanisms besides crossover and mutation. Some of those mechanisms are inversion, transduction, transformation, conjugation, transposition and translocation (Gould et al. 1996).

Some researchers highlighted the importance of these latest discoveries of molecular biology. Mitchell (1996), Mitchell et al. (1994) and Banzhaf et al. (1998) stress that it would be important to analyze if some of the mechanisms of rearranging genetic material present in the biological systems, when implemented and used with a GA, improve its performance.

Several authors have already used some biologically inspired mechanisms besides crossover and mutation in genetic approaches. For instance, inversion (Holland 1992), conjugation (Harvey 1996; Smith 1996a; Smith 1996b), transduction (Furuhashi et al. 1994; Nawa et al. 1997; Nawa et al. 1998; Nawa et al. 1999), translocation (Oates et al. 1999; Voss et al. 1999) and transposition (Simões et al.

1999a; Simões et al. 1999b, Simões et al. 1999c, Simões 1999, Simões et al. 2000) were already used as the main genetic operators in the GA.

Transposition consists in the presence of genetic mobile units called transposons, capable of relocating themselves, or transposing, onto the chromosome and subsequently jumping into new zones of the same or a different chromosome.

Simões et al. (1999a; 1999b; 1999c; 2000), following the biological inspiration, proposed different ways of using the transposition mechanism. In the sexual form of transposition, the process always occurs between two selected individuals (Simões et al. 1999a; 1999c). The authors introduce two distinct manners of using the sexual mechanism: simple and tournament-based transposition.

Asexual transposition proposed by Simões et al. 2000 works in the same individual.

Previous work used the transposition mechanism as the main genetic operator, by replacing the standard crossover operators. The obtained results showed that the different forms of the transposition allow the GA to reach higher solutions than the crossover operators, even when using smaller populations.

In this paper we will complete the comparative analysis about the transposition mechanism. We will compare the achieved solutions using asexual transposition and both forms of sexual transposition (simple and tournament-based). Moreover, we will try to understand when one of the forms can be more advantageous than the other and the relation between the studied heuristics (Simões et al. 1999b; Simões et al. 2000) to compute one important parameter: the flanking sequence length.

The GA performance will be compared in the domain of function optimization applying a standard test suite with either sexual or asexual transposition. In the studied domain the GA used binary representation. Nevertheless, transposition mechanism can be used with other kind of representations, such as vectors of real numbers or tree-based data structures and in other domains as well.

This paper is organized in the following manner. First, in section 2, we summarize our previous work related to the transposition mechanism. In section 3, we describe how transposition works in nature and how the previous versions were implemented. Section 4 describes the characteristics of

the experimental environment. In section 5, we make an exhaustive comparison of the results obtained with asexual and sexual transposition. Finally we present the relevant conclusions of the work.

2 Previous Related Work

Simões et al. presented a new biologically inspired genetic mechanism, called transposition, as an alternative to crossover (Simões et al. 1999a). In a preliminary work, using a GA as a function optimizer, with a single test function, very promising results were obtained. Such work compared the GA performance when using 1-point, 2-point, uniform crossover or a simple form of transposition. Transposition allowed the GA to reach better results than crossover, even with smaller populations. Later, this preliminary work was enlarged to a test bed containing eighteen test functions and an extensive comparative study showed that, if the right parameters were chosen, transposition always performed better than crossover. Moreover, the authors introduced a new form of transposition, inspired in Harvey's work (Harvey 1996), and called tournament-based transposition, which also proved to be a good substitute to crossover (Simões et al. 1999b). For a detailed description of this work see Simões (1999).

Both mechanisms used sexual reproduction, i.e., two individuals were selected for mating and the transposition mechanism occurs between these individuals. In simple transposition the exchange of the genetic mobile unit is made in a random manner between the two strings. In tournament-based transposition the best individual will give the genetic material to the worst one.

Later, the authors introduced an asexual version of the transposition mechanism (Simões et al. 2000) where all the process of transposing a set of genes occurred in the same individual. This new way of transposition was integrated in the GA, replacing again the standard crossover operators and the achieved results demonstrate once more that the GA performance was also superior even when using smaller populations. The GA with 50 individuals and transposition could achieve better performances than when using 200 individuals and crossover.

3 Transposition

3.1 Biological Transposition

As referred before, transposition is characterized by the presence of mobile genetic units inside the genome, moving themselves to new locations or duplicating and inserting themselves elsewhere. These mobile units are called transposons (Gould et al. 1996; Russell 1998).

One or several genes or just a control unit can form transposons (also known as jumping genes). The movement can take place in the same chromosome or to a different one.

Barbara McClintock first discovered transposition in the 50's (when the DNA structure was not yet completely understood). She proved that certain phenomena present in living beings exposed to UV radiation could not be the result of the normal recombination and mutation processes. She found that certain genetic elements in corn occasionally move producing kernels with unusual colors that could not have resulted from crossover or mutation. Transposons were for a long time considered as some sort of abnormality, but in 1983, when she was awarded the Nobel Prize, many such transposons had been discovered and their possible role in evolution was beginning to be recognized. For instance, the genetic alterations caused by transposons are responsible for the growth of cancer in humans and for the resistance to antibiotics in bacteria (Gould et al. 1996; Russell 1998). In order for a transposable element to transpose as a discrete entity it is necessary for its ends to be recognized. Therefore, transposons within a chromosome are flanked by identical or inverse repeated sequences, some of which are actually part of the transposon. See Figure 1 below.

When the transposon moves to another zone of the genome one of the flanking sequences goes with it.

The insertion point can be chosen at random, but there are transposons that show a regional preference when inserting into the same gene. Other method can be a correspondence in the new position with the flanking sequence.

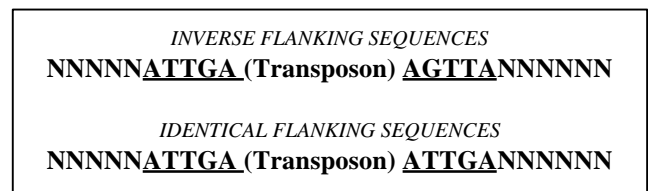


Figure 1: Inverse and Equal Flanking Sequences

The point into which the transposon is inserted requires no homology with the point where the transposon was excised. This is in evident contrast to classical recombination, where relatively long sequences of DNA must share homology to permit a recombination event to occur (same cut point(s)). Consequently, transposition is sometimes referred to as illegitimate recombination.

3.2 Simple Transposition

The first form of computational transposition proposed by Simões et al. (1999a) was directly inspired in biology. After the selection of two parents for mating, the transposon is formed in one of them. The insertion point is found in the second parent. According to this point, the same amount of genetic material is exchanged between the two chromosomes. The transposon is recognized by the presence of equal or inverse flanking sequences with a fixed length. The insertion point is searched in the second chromosome

and is chosen when a sequence of genes equal or inverse to the flanking sequence is found. The insertion point will be the first gene after that sequence. After that, the movement of the transposon occurs. Since it was used a fixed chromosome length, the same amount of genetic material is exchanged between the two selected parents. Figure 2 shows how simple transposition works. The detailed functioning of transposition is described in (Simões et al. 1999a).

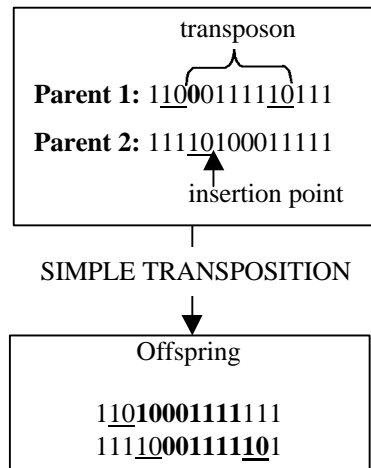


Figure 2: Simple Transposition

3.3 Tournament-based Transposition

In order to come closer to the biological mechanism, the authors proposed a new form of transposition: *tournament-based transposition*.

The two selected parents become competitors in a tournament of size two. The transposon will be searched in the winner chromosome and the insertion point will be located in the loser parent. Only this individual will be altered, by the insertion of the transposon, which replaces the same number of bits after the insertion point. Figure 3 shows this mechanism.

3.4 Asexual Transposition

In nature the transposition mechanism can also occur in the same chromosome. Therefore, Simões et al. 2000 implemented an asexual form of transposition, which kept the main functioning of the biological process. In this case, the entire search for the flanking sequence, the transposon and the insertion point occurs in the same individual.

After selecting one individual for reproduction, the asexual transposition will be applied. The beginning of the transposon will be randomly chosen (gene **T**) in the selected individual. According to the flanking sequence length, the **FSL** bits before the gene **T** make the first flanking sequence. The search of the second flanking sequence begins after gene **T** and stops when an equal or inverse sequence is found. The genes enclosed by gene **T** and the

last gene of the second flanking sequence constitute the transposon.

The insertion point is searched in the same chromosome and this process starts in the first bit after the second flanking sequence. The insertion point is defined when an equal or inverse sequence of bits is found in the chromosome. Notice that the chromosome is viewed as having a circular form. Therefore, after reaching the end of the chromosome the search continues in its first bit. When the insertion point is found, the transposon excises from its original position and will integrate in the insertion point.

Figure 4 synthesizes the complete process of asexual transposition.

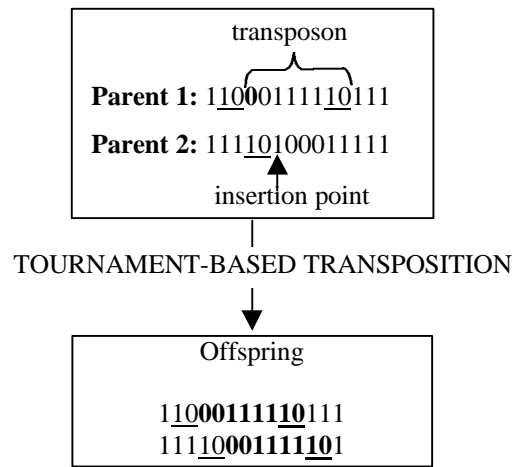


Figure 3: Tournament-based Transposition

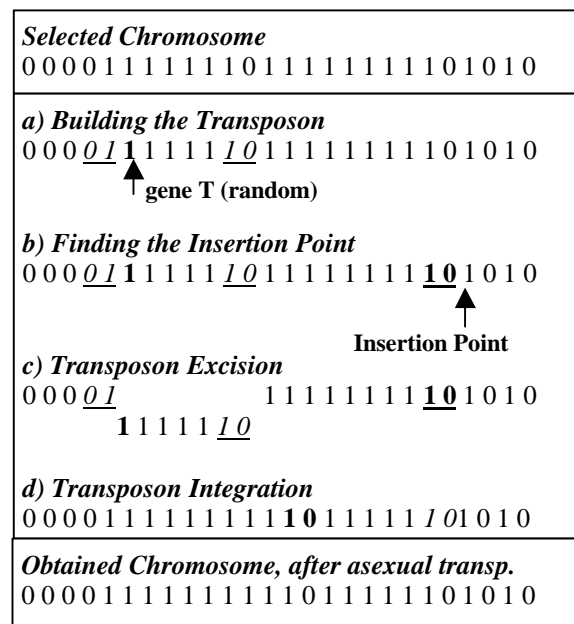


Figure 4: Asexual Transposition

4 The Experimental Environment

4.1 Case study

To compare the results obtained with sexual and asexual transposition we used the GA as function optimizer. We chose a test suite containing seven functions (see appendix), characterized as continuous/discontinuous, unimodal/multimodal, high/low/scalable, dimensional, stochastic/deterministic, quadratic/non-quadratic and convex/non-convex. These functions are a well-known benchmark for genetic approaches (De Jong 1975; Michalewicz 1999; Whitley et al. 1995) and were selected in order to cover such a large set of characteristics.

Since the GA was used in the function optimization domain, we chose roulette wheel with elitism as the selection method, in order to keep track of the best solution found (De Jong 1993).

To show the achieved results we will use two functions representative of the global test suite. These functions are Michalewicz's and Griewangk's functions.

Michalewicz's function is a continuous, non-convex, multimodal, non-quadratic and bidimensional test function. Figure 5 shows its graphical representation.

Griewangk's function is a continuous, unimodal, scalable, convex and quadratic test function. In our experiments we used a 10 variable dimension. Figure 6 illustrates the bidimensional form of this function.

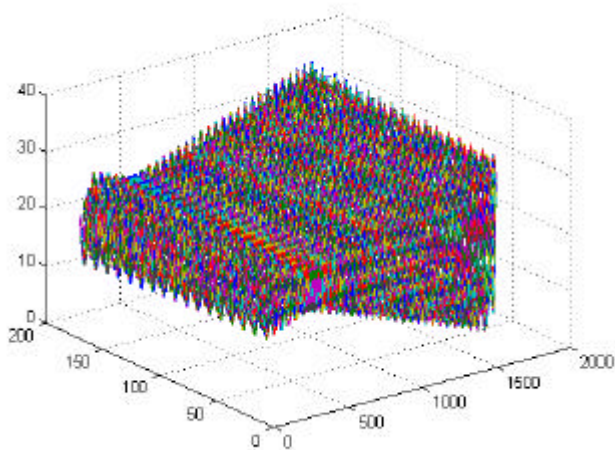


Figure 5: Michalewicz's Test Function

4.2 The Genetic Algorithm's Parameters

The GA was first implemented with sexual transposition (simple and tournament-based) and then with asexual transposition. We executed experiments to study the effect of the population size in the GA efficiency. Therefore, the population size varied between 50, 100 and 200 individuals, either for asexual or sexual transposition. The elite size was 20% of the complete population. The mutation and transposition rate used was 0.01 and 0.7, respectively.

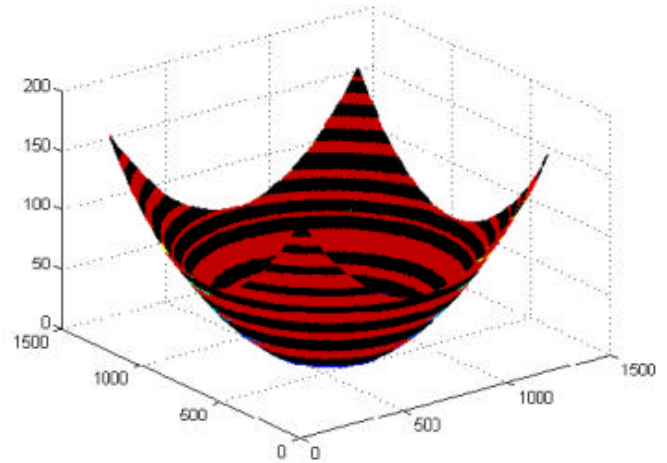


Figure 6: Griewangk's Test Function

An important parameter to the achieved results, when applying the transposition mechanisms, is the flanking sequence length. In our experiments we used flanking sequences size from 1 to 10, 15 or 20, depending on the chromosome length. In order to understand the role of this parameter, in each set of experiments the flanking sequence size was kept constant. We will refer, in the next section, to the conclusions about the role of the flanking sequence length in next section.

Table 1 summarizes all the executed tests for each function of the test suite.

Mechanism	Population	Flanking Sequence
<u>Sexual Transposition:</u> Simple Transposition	50	1 to Maximum (10, 15 or 20)
	100	1 to Maximum (10, 15 or 20)
	200	1 to Maximum (10, 15 or 20)
<u>Sexual Transposition:</u> Tournament-based Transposition	50	1 to Maximum (10, 15 or 20)
	100	1 to Maximum (10, 15 or 20)
	200	1 to Maximum (10, 15 or 20)
<u>Asexual Transposition</u>	50	1 to Maximum (10, 15 or 20)
	100	1 to Maximum (10, 15 or 20)
	200	1 to Maximum (10, 15 or 20)

Table 1: Summary of all the experiments

All the tests were run over 500 or 1000 generations, depending on the test function. The analyzed results are the average of the ten runs we executed.

4.3 Evaluation Measure

We used the De Jong's off-line measure to compare GA efficiency (De Jong 1975). This measure is defined by:

$$X_e^*(g) = \frac{1}{T} * \sum_{t=1}^T f_e^*(t)$$

Where $f_e^* = \text{best} \{f_e(1), f_e(2), \dots, f_e(n)\}$ and T is the number of runs. This means that off-line measure is the average of the best individuals in each generation. Due to the total of ten trials, the average of the tens runs was evaluated.

5 The Results

The results obtained with the seven studied functions demonstrate that asexual transposition is, in general, the mechanism that allows better solutions. Nevertheless, asexual and tournament-based transposition allows the GA to achieve similar results. When using simple transposition, in most cases (except F5), the results were slightly worse. Both mechanisms, asexual and tournament-based transposition, are less disruptive than simple transposition. This seems to be the reason for the similarity of the results and for the best performance of the GA when using one of these mechanisms.

We will show the comparative analysis for asexual and sexual transposition (simple and tournament-based) focusing the following issues. First, we will analyze the importance of the choice for the flanking sequence length and the relation between the achieved heuristics for each case, presented by Simões et al. (1999b) and Simões et al. (2000). Then, we will show for both functions, the different solutions obtained when using sexual or asexual transposition with different population's size.

5.1 The Importance of the Flanking Sequence Length

When using one of the transposition mechanisms as the main genetic operator for the GA, the choice of the flanking sequence length must be done through one of the heuristics studied in previous works (Simões et al. 1999b, Simões et al. 2000). Such heuristics compute the flanking sequence length, for each form of transposition, which will lead the GA to the best results.

Observing the values calculated for the heuristics in each situation, we could see that, when the chromosome length has smaller size, all the heuristics compute similar values to the flanking sequence length. When using larger chromosomes to codify the function (F7, F5 and F3), once

more, asexual and tournament-based transposition exhibits similar behavior. As table 2 shows, the heuristics for these two mechanisms compute alike values.

Test Function	Chrom. Length	Best Sequence Length		
		Asexual Transp.	Simple Transp. (sexual)	Tournament-Based Transp. (sexual)
F1	24	2, 3	3, 4	4, 5, 6
F4	33	3, 4	4, 5	4, 5, 8
F2	50	5, 6	2, 3, 4	3, 4
F7	200	10, 11	4, 5, 6	10, 11
F5	210	11, 12	6, 7, 8	10, 12
F3	240	12, 13	2, 7, 8	20, 12, 13
F6	280	14, 15	9, 11	11, 12, 13, 14

Table 2: The Best Flanking Sequence Length for Asexual and Sexual Transposition

5.2 Comparing Asexual Transposition with Sexual Transposition: Michalewicz's Test Function

In this section we will show the results obtained with the GA using asexual and sexual transposition for the Michalewicz's function.

For this test function, using 50, 100 or 200 individual the results obtained were very close. Figures 7, 8 and 9 show that results. We can see that asexual transposition allowed the GA to reach better solutions than the two variants of sexual transposition. Nevertheless, this superiority is not very expressive.

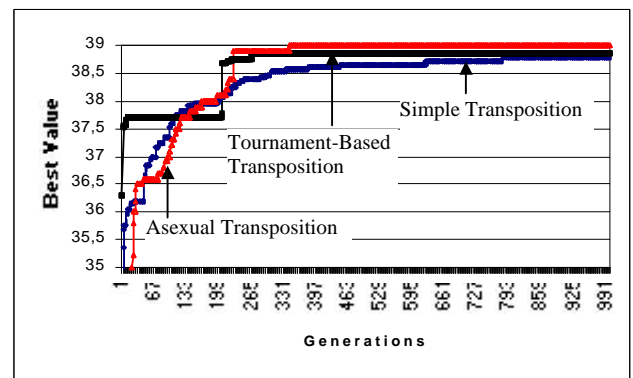


Figure 7: Sexual vs. Asexual Transposition for Michalewicz's Function: Population size = 50 Individuals

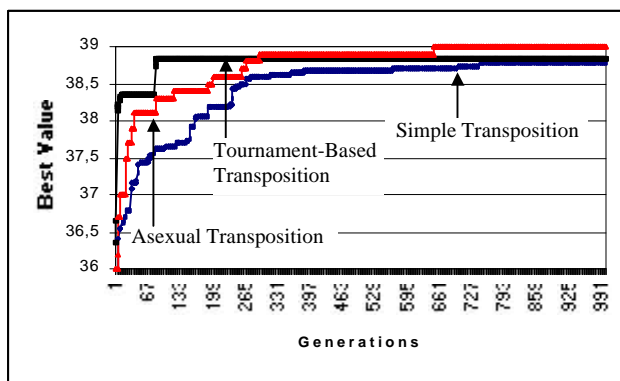


Figure 8: Sexual vs. Asexual Transposition for Michalewicz's Function: Population size = 100 Individuals

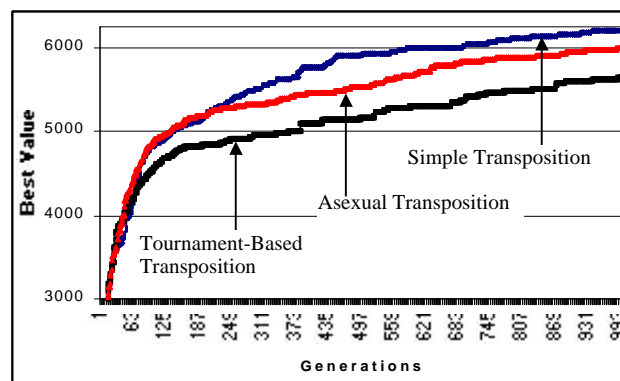


Figure 10: Sexual vs. Asexual Transposition for Griewangk's Function: Population size = 50 Individuals

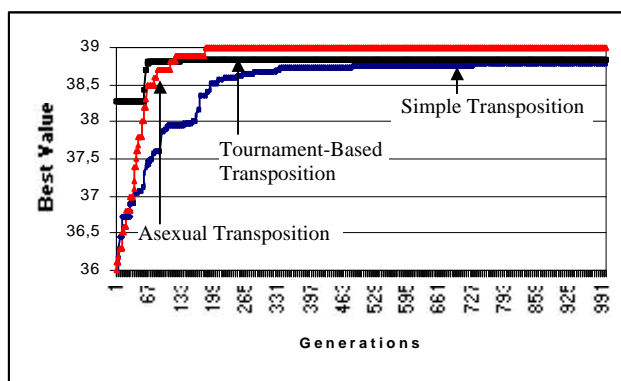


Figure 9: Sexual vs. Asexual Transposition for Michalewicz's Function: Population size = 200 Individuals

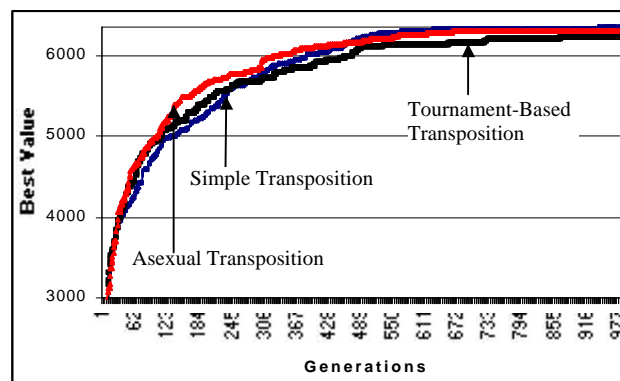


Figure 11: Sexual vs. Asexual Transposition for Griewangk's Function: Population size = 100 Individuals

5.3 Comparing Asexual Transposition with Sexual Transposition: Griewangk's Test Function

In this section we will show the results obtained with the GA using asexual and sexual transposition for the Griewangk's function.

The results are similar to the previous case but, for that function, simple transposition allowed the GA to achieve higher performances than tournament-base or asexual transposition. This advantage is clearer when using only 50 individuals (Figure 10). Using larger populations the results obtained become closer for both sexual and asexual transposition (Figure 11 and 12).

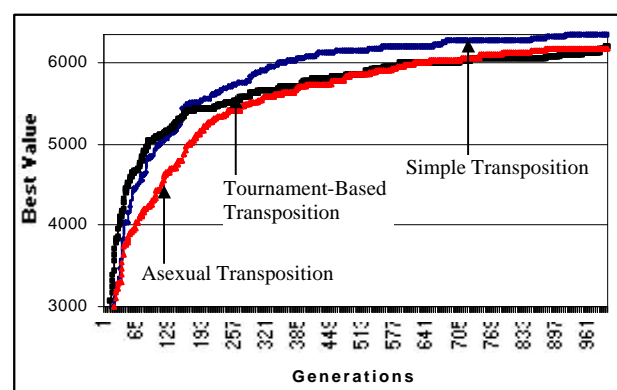


Figure 12: Sexual vs. Asexual Transposition for Griewangk's Function: Population size = 200 Individuals

6 Conclusions

In this paper we compared the GA performance using all the forms of transposition already proposed in previous articles: asexual and sexual transposition (simple and tournament-based). We compared these mechanisms using a test suite containing seven functions. All the transposition operators allowed the GA to achieve similar results. In spite of this resemblance, we observed some similarity in the behavior of the GA when using asexual and tournament-based transposition. In these two situations the GA performance was slightly superior than when applying simple transposition.

Comparing the flanking sequence lengths that allow the best results when applying transposition, we also notice some similarities between asexual and sexual transposition.

In order to see the advantages that these mechanisms can introduce in different practical domains we will apply transposition to different problems. Our studies will focus in the importance that transposition can have in finding "good" solutions knowing that the exchange of genes is controlled by a context defined by the flanking sequences.

Acknowledgements

This work was partially financed by the Portuguese Ministry of Science and Technology under the Program Praxis XXI.

Appendix

F1: De Jong's Test Function F2 (Rosenbrock's Saddle)

$$F_1(x_1, x_2) = 100 * (x_1^2 - x_2^2)^2 + (1 - x_1)^2$$

F2: De Jong's Test Function F3 (Step Function)

$$F_2(x) = \sum_{i=1}^5 \text{integer}(x_i)$$

F3: De Jong's Test Function F4 (Gaussian Quartic)

$$F_3(x) = \sum_{i=1}^{30} ix_i^4 + \text{Gauss}(0,1)$$

F4: Michalewicz's Function

$$F_4(x_1, x_2) = 21.5 + x_1 \cdot \sin(4\pi x_1) + x_2 \cdot \sin(20\pi x_2)$$

F5: Griewangk's Function

$$F_5(x) = 1 + \sum_{i=1}^n \left[\frac{x_i^2}{4000} \right] - \prod_{i=1}^n \left[\cos\left(\frac{x_i}{\sqrt{i}}\right) \right]$$

with $n = 10$

F6: Rastrigin's Function

$$F_6(x) = n * A + \sum_{i=1}^n [x_i^2 - A * \cos(2\pi x_i)]$$

with $n = 20, A = 10$

F7: Schwefel's (Sine Root) Function

$$F_7(x) = V * n + \sum_{i=1}^n [-x_i * \sin(\sqrt{|x_i|})]$$

with $n = 10, V = 418,9829$

Bibliography

- W. Banzhaf, P. Nordin, R. E. Keller, F. D. Francone (1998). *Genetic Programming - An Introduction - On the Automatic Evolution of Computer Programs and its Applications*. Morgan Kaufmann
- K. A. De Jong (1975). *Analysis of the Behavior of a Class of Genetic Adaptive Systems*. Ph.D. Dissertation, Department of Computer and Communication Science, University of Michigan.
- K. A. De Jong (1993). Genetic Algorithms Are NOT Function Optimizers. In L. Darrell Whitley (ed.), *Foundations of Genetic Algorithms 2*, 5-17. CA: Morgan Kaufmann
- T. Furuhashi, K. Nakaoka and Y. Uchikawa (1994). A New Approach to Genetic Based Machine Learning and an Efficient Finding of Fuzzy Rule. In T. Furuhashi (ed.), *Proceedings of the 1994 IEEE/Nagoya-University World Wisepersons Workshop*. Lecture Notes in Artificial Intelligence, vol. 1011, 173-189. Springer-Verlag.
- D. E. Goldberg (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Publishing Company, Inc.
- J. L. Gould and W. T. Keeton (1996). *Biological Science*. W. W. Norton & Company.
- I. Harvey (1996). The Microbial Genetic Algorithm. Submitted as a Letter to *Evolutionary Computation*. MIT Press.
- J. H. Holland (1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. 1st MIT Press edition, MIT Press.
- Z. Michalewicz (1999). *Genetic Algorithms + Data Structures = Evolution Programs*. 3rd Edition Springer-Verlag.
- M. Mitchell and S. Forrest (1994). Genetic Algorithms and Artificial Life. *Artificial Life* 1(3):267-289.

- M. Mitchell (1996). *An Introduction to Genetic Algorithms*. MIT Press.
- N. Nawa, T. Hashiyama, T. Furuhashi and Y. Uchikawa (1997). A Study on Fuzzy Rules Using Pseudo-Bacterial Genetic Algorithm with Adaptive Operator. *Proceedings of the IEEE International Conference on Evolutionary Computation*, 589-593. IEEE Press.
- N. Nawa and T. Furuhashi (1998). Bacterial Evolutionary for Fuzzy System Design. *Proceedings of the 1998 IEEE International Conference on Systems, Man and Cybernetics*. IEEE Press.
- N. Nawa, T. Furuhashi, T. Hashiyama and Y. Uchikawa (1999). A Study of the Discovery of Relevant Fuzzy Rules Using Pseudo-Bacterial Genetic Algorithm. *IEEE Transactions on Industrial Electronics*.
- M. Oates, D. Corne, R. Loader (1999). Skewed Crossover and the Dynamic Distributed Database Problem, In the *Proceedings of the Fourth Int. Conference on Neural Networks and Genetic Algorithms*, pp. 280-287. Springer-Verlag.
- P. J. Russell (1998). *Genetics*. 5th edition, Addison-Wesley.
- A. Simões (1999). *Transposition: A Biologically Inspired Mechanism and its Integration in the Classical Genetic Algorithm*. M.Sc. Dissertation Thesis (in Portuguese), University of Coimbra.
- A. Simões and E. Costa (1999a). Transposition: A Biologically Inspired Mechanism to Use with Genetic Algorithms. In the *Proceedings of the Fourth International Conference on Neural Networks and Genetic Algorithms (ICANN'99)*, pp. 612-619. Springer-Verlag
- A. Simões and E. Costa (1999b). Transposition versus Crossover: An Empirical Study. Banzhaf, W., Daida, J., Eiben, A. E., Garzon, M. H., Honavar, V., Jakiela, M., and Smith, R. E. (eds.), *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'99)*, pp. 612-619, Orlando, Florida USA, CA: Morgan Kaufmann.
- A. Simões and E. Costa (1999c). Enhancing Transposition Performance. *Proceedings of the 1999 Congress on Evolutionary Computation (CEC'99)*, pp. 1434-1441. Piscataway, NJ: IEEE Press.
- A. Simões and E. Costa (2000). Using Genetic Algorithms with Asexual Transposition. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2000)*, Las Vegas, 8-12 July.
- P. Smith (1996a). Finding Hard Satisfiability Problems Using Bacterial Conjugation. Presented at the *AISB96 Workshop on Evolutionary Computing*. University of Sussex.
- P. Smith (1996b). Conjugation - A Bacterially Inspired Form of Genetic Recombination. In *Late Breaking Papers of the First International Conference on Genetic Programming*. Stanford University, California.
- M. Voss and C. Foley (1999). Evolutionary Algorithm For Structural Optimization, Banzhaf, W., Daida, J., Eiben, A. E., Garzon, M. H., Honavar, V., Jakiela, M., and Smith, R. E. (eds.), *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'99)*, pp. 678-685, Orlando, Florida USA, CA: Morgan Kaufmann.
- D. Whitley, K. Mathias, S. Rana and J. Dzubera (1995). Building Better Test Functions. In L. J. Eshelman (ed.), *Proceedings of the Sixth International Conference on Genetic Algorithms*, 239-246. CA: Morgan Kaufmann.
- T. Yoshikawa, T. Furuhashi and Y. Uchikawa (1997). Knowledge Acquisition of Fuzzy Control Rules for Mobile Robots Using DNA Coding Method and Pseudo-Bacterial Genetic Algorithm. *Proceedings of the First Asia-Pacific Conference on Simulated Evolution and Learning*. Lecture Notes in Artificial Intelligence, vol. 1285, 126-135. Springer-Verlag.