In David W. Pearson, Nigel C. Steele, Rudolf Albrecht (eds.), Proceedings of the Sixth International Conference on Neural Networks and Genetic Algorithms (ICANNGA'03), pp. 175-181, Roanne, France, 23-25 April, Springer, 2003.

Improving the Genetic Algorithm's Performance when Using Transformation

Anabela Simões[1,2], Ernesto Costa[2]
[1] Dept. of Informatics and Systems Engineering, Coimbra Polytechnic, Quinta da Nora, 3030 Coimbra, Portugal
[2] Centre for Informatics and Systems of the Univ. of Coimbra, Pinhal de Marrocos, 3030 Coimbra, Portugal
E-mail: abs@isec.pt; ernesto@dei.uc.pt

Abstract. Transformation is a biologically inspired genetic operator that, when incorporated in the standard Genetic Algorithm can promote diversity in the population. Previous work using this genetic operator in the domain of function optimization and combinatorial optimization showed that the premature convergence of the population is avoided. Furthermore, the solutions obtained were, in general, superior to the solutions achieved by the GA with standard 1-point, 2-point and uniform crossover. In this paper we present an extensive empirical study carried to determine the best parameter setting to use with transformation in order to enhance the GA's performance. These parameters include the gene segment length, the replacement rate (percentage of individuals of the previous population used to update the gene segment pool), and the mutation and transformation rates.

1 Introduction

One of the key aspects related with the performance of Evolutionary Algorithms (EA) is the diversity of the evolved population. Loss of diversity may imply the premature convergence of the EA to local optima. The solutions converge to a region of the search space and the EA cannot continue exploring other promising areas. It is also true that a homogeneous population is less adapted to dynamic changes in the environment.

In EA, the diversity of a population is mostly achieved by the application of the genetic operators. The traditional genetic operators used in EA are crossover and mutation. Crossover is used to explore different promising areas of the search space, allowing the convergence to the optima (exploration). Mutation is used to “shake” the GA avoiding eventual traps in local optima (exploitation) [1]. Nevertheless, the premature convergence isn't always avoided with the application of these mechanisms. For instance, when faced with problems involving a dynamic environment, the traditional EAs aren't suitable to solve them. The population converges to the desired optimum of the problem, but when this optimum changes by some reason, the EA has troubles in readapting to the new solution.

In order to promote diversity in the individuals of the population other mechanisms have been used. Recently, Simões and Costa proposed a biologically inspired genetic operator called transformation that promotes diversity and avoids premature convergence [5]. The mechanism mimics the biological process and consists in the capacity of the individuals to absorb fragments of DNA from the environment. These gene segments are then reintegrated in the individuals' genome. The proposed mechanism was tested with several classes of problems, namely function optimization, the static and dynamic 0/1 knapsack problem ([5], [6]). The results showed that this new genetic operator preserves the diversity in the population during all the generations and the EA with transformation achieved, in general, better solutions than the standard GA (SGA). Nevertheless, the implementation of transformation used a parameter setting chosen based on some preliminary empirical studies. For instance, the choice of the gene segment length was always random. In this paper we report the results of an extensive empirical study that allows the user to choose the proper parameter setting in order to enhance the GA's performance when using transformation. The studied parameters are the gene segment length, the percentage of gene segments that are updated using the genetic information of the individuals of previous population (replacement rate), and the mutation and transformation rates.

The remaining paper is organized in the following manner. In section 2, we briefly explain the biological and computational functioning of the transformation mechanism. Section 3, briefly refers to previous work about transformation. In section 4, we detail the characteristics of the experimental study, including the test problems, the parameters and the experimental plan. Section 5 shows the results and states the main conclusions regarding the correct choice of the parameters. Finally, we present the relevant conclusions of the work.

2 Transformation

In nature, some bacteria can absorb fragments of DNA from the surrounding environment and reintegrate this DNA in its genetic material. These bacteria are called competent and the reintegration of this DNA fragments may confer some evolutionary advantages to them. The DNA fragments or gene segments proceed mostly from dead bacteria [3].

Simões and Costa proposed a computational implementation of the transformation mechanism and used it as the main genetic operator in the EA, instead of crossover [5].

At the beginning of the EA a population of individuals and a gene segment pool are created randomly. Then, using a roulette-wheel selection method some individuals are chosen to be transformed. Those individuals form the “mating” pool. The transformation was applied with a fixed probability of 70% and works as follows: select one individual from the “mating” pool and a gene segment from the gene segment pool. It is also chosen randomly a transformation point in the selected individual. The genes of the gene segment replace the genes in the selected individual, after the transformation point. After that, the gene pool is updated using the genetic information of the individuals of the previous population. These individuals are used to create 70% of the new segments, being the remaining created at random. These percentages were fixed to all the executed experiments. Furthermore, the length of each segment was also randomly chosen. The GA modified with this new genetic operator will be referred as Transformed-based GA (TGA)
Figure 1 describes the mechanism. For more details see [5]

3 Previous work

The genetic operator was tested in two domains: function optimization and combinatorial optimization. The chosen functions were four well-known functions: Ackley, Griewangk, Rastrigin and Schwefel and the 0/1 Knapsack problem (0/1 KP) for combinatorial optimization. The goals were to minimize the four functions (which minimum value is zero) and to maximize the value of the objects selected to a knapsack of limited capacity. Several instances of the 0/1 KP were implemented, namely with different number of items: 50, 100, 250 and 500 ([2]).
We compared the quality of the solutions found with transformation against the traditional crossover operators: 1-point, 2-point and uniform crossover. The obtained results are summarized in the tables 1 and 2 ([5]).
In Table 1 we show the results of the function minimization problems. As we can see transformation allowed the GA to achieve much better results. The principal conclusion to support these results was the great diversity that transformation introduces in the population avoiding its convergence, and allowing the evolution during a long period.
Nevertheless, either the GA or the TGA were able to achieve the optimum solution.

```
Transformed-Based GA (TGA)
Begin
1. P(t)=Generate initial population
2. S(t)=Generate initial gene segment pool
3. While (NOT stop_condition)
   a. Evaluate P(t)
   b. Select individuals from P(t)
   c. P’(t)=Transform individuals
   d. P’’(t)=Mutate new individuals
   e. Update S(t) using segments from individuals from
      P(t) and new random segments
   f. Replace P(t) with P’’(t)

End
```

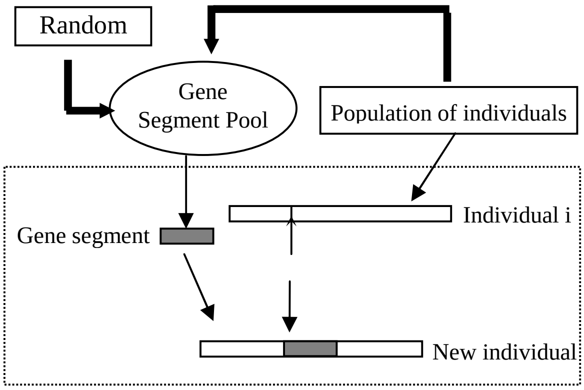

Fig. 1. Computational Transformation

In Table 2 we show the results for the 0/1 KP. In this case the Uniform Crossover was the genetic operator that achieved higher performance.
In order to enhance the transformation performance and achieve better solutions, in particular in the 0/1 KP problem, we performed an extensive parametric study to observe the influence that parameters, such as the gene segment’s length, the transformation rate, mutation rate or replacement rate, could have in the obtained results.
The parametric study will be detailed in the next section.

Table 1. Results obtained by a SGA (using 1-point, 2-point and Uniform Crossover) and the GA using Transformation in the function optimization domain

		Genetic Operator											
		One-point Crossover			*Two-point Crossover*			*Uniform Crossover*			*Transformation*		
	Nº evals	*50000*	*100000*	*200000*	*50000*	*100000*	*200000*	*50000*	*100000*	*200000*	*50000*	*100000*	*200000*
Function	Rastrigin	23.464	16.502	11.599	23.164	15.879	10.739	31.079	22.791	16.676	38.401	18.828	**6.540**
	Griewangk	0.006	0.001	0.001	0.004	0.002	0.001	0.007	0.004	0.003	0.010	0.003	**0.001**
	Schwefel	23.611	2.565	0.409	36.140	13.178	5.484	53.705	11.166	0.521	36.212	0.475	**0.077**
	Ackley	2.921	1.034	0.278	2.699	0.922	0.119	2.946	1.416	0.429	3.128	0.300	**0.002**

Table 2. Results obtained by a SGA (using 1-point, 2-point and Uniform Crossover) and the GA using Transformation solving the 0/1 Knapsack Problem

		Genetic Operator												
		One-point Crossover			*Two-point Crossover*			*Uniform Crossover*			*Transformation*			
	P. Size→	*50*	*100*	*200*	*50*	*100*	*200*	*50*	*100*	*200*	*20*	*50*	*100*	*200*
Nº of Items	50	207.17	195.20	185.40	206.87	199.13	187.77	**207.20**	199.80	189.13	197.30	197.80	197.23	207.17
	100	447.93	410.03	373.67	449.63	415.93	381.73	**453.07**	402.90	377.67	413.00	408.40	409.70	447.93
	250	976.40	844.33	780.93	980.93	862.43	785.27	**981.40**	824.33	774.93	838.50	834.87	835.17	976.40
	500	1914.43	1671.53	1535.50	1936.87	1696.20	1537.03	**1955.30**	1622.27	1533.30	1666.20	1669.07	1661.03	1914.43

4 Experimental setting

The GA was implemented with roulette-wheel selection with elite of size 2 and population size of 200 individuals in the minimization of the functions and with population of size 50 and 100 in the maximization of the 0/1 KP. Instead of the classical crossover operator we used transformation. The remaining parameters were change to study their influence in the GA's performance when using this mechanism.

We started analyzing the effect of the gene segment length in both problem domains. The length of the chromosome differs from problem to problem, and so the variation of the segment length was dependent on that.

In previous work the length of each segment was always chosen in a random way. So, we could have in the segment pool segments with several lengths. In this study, we fixed all the remaining parameters and changed the gene segment size from 5 to a maximum value depending on the chromosome length, but the segment was fixed for each experiment. The variation of this value was made in jumps of five units, i.e, 5, 10, 15, … until the maximum allowed value. All the experiments were repeated over 30 runs.

Table 3 shows the chromosome length for each problem and the interval of variation for the gene segment's length.

As referred before, the replacement rate is the percentage of individuals of the previous generation that contribute for the modification of the gene segment pool at each generation.

Table 3. Interval of variation for the gene segment length

Problem	Chromosome Length	Interval of variation for the gene segment length
Ackley's Function	480	[5, 125]
Griewangk's Func.	160	[5, 125]
Rastrigin's Func.	280	[5, 125]
Schwefel's Func.	160	[5, 125]
0/1 KP – 50 Items	50	[5, 50]
0/1 KP – 100 Items	100	[5, 100]
0/1 KP – 250 Items	250	[5,125]
0/1 KP –500 Items	500	[5,125]

In our original work we used 70% of the individuals of the previous population to update the gene segment pool at each generation. In order to determine the influence of this parameter, we changed the replacement rate from 0% to 100%, in jumps of 5 units.

The third parameter to be analyzed was the mutation rate. We changed this rate among six values: 0%, 1%, 5%, 10%, 15% and 20%. Our previous work used 0.1% of mutation rate.

Finally, we altered the transformation rate from 0% to 100% in jumps of 10 units.

5 Results

This section will present some of the results obtained. We will show representative graphics obtained in the function optimization domain and in the 0/1 KP. To illustrate the results obtained in the function optimization domain we choose the Ackley's test function. For the 0/1 KP we show the results achieved with 100 items.

5.1 Segment length

The variation of this parameter had great influence in the obtained results. Smaller sizes allowed, always, better results. Figure 2 shows the variation of the best solutions found minimizing the Ackley's test function. The gene segment's length varied from 5 to 125. As we can see, segments of length equal to 5, 10 and 15 were the best choice for this parameter.

Figure 3, shows the results obtained in the maximization of the 0/1 KP, with 100 items. Once more, smaller segments allowed best results. In this case, segments of length 5 or 10 were the appropriated choice.

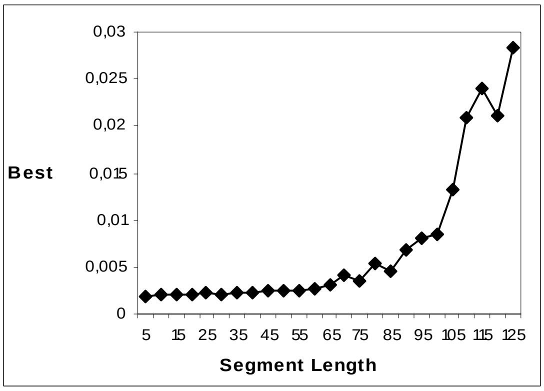

Fig. 2. Influence of the Gene Segment's Length in the Best Solution Found (minimization of Ackley's test function)

The main reason for the degradation of the results using larger segments appears to be the disruption introduced in the population. As we can see, smaller segments can also preserve the diversity of the population, but the levels are not as high as in the case of larger segments.

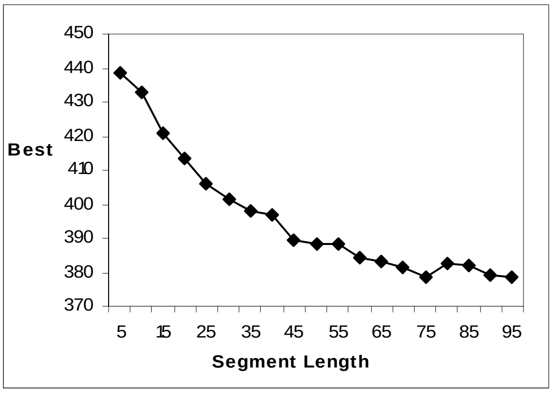

Fig. 3. Influence of the Gene Segment's Length in the Best Solution Found (maximization of the 0/1 KP – 100 items)

In order to explain these results we measure the diversity of the population using a standard measure defined by:

$$Div(Pop) = \frac{1}{LP(P-1)} \sum_{i=1}^{P} \sum_{j=1}^{P} HD(p_i, p_j)$$

where **L** is the length of the chromosome, **P**, the population size, $\mathbf{p_i}$, the i^{th} individual in the population and **HD** the hamming distance.

Figures 4 and 5 illustrate how the diversity of the population evolved during the entire process, using gene segments of length 5 and 125 for the Ackley's function and 5 and 100 for the 0/1 KP, respectively.

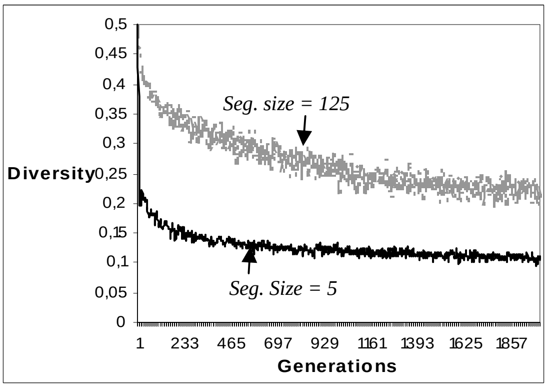

Fig 4. Population's Diversity using Gene segments of size 5 and 125 (Ackley's Function)

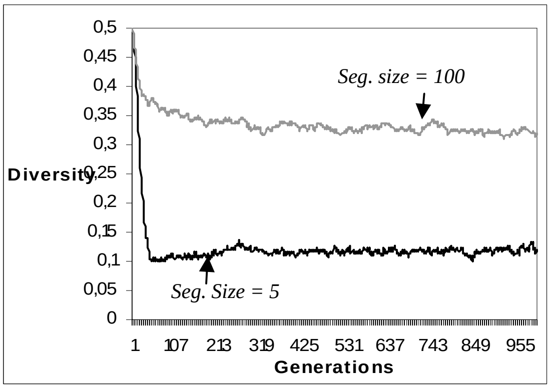

Fig. 5. Population's Diversity using Gene segments of size 5 and 100 (0/1 KP – 100 items)

5.2 Replacement rate

The replacement rate was changed from 0% (all the gene segments are created at random) to 100% (all the segments proceed from individuals of previous generations). The results obtained in the two domains were quite different: when minimizing the test functions, replacement rates superior to 70% allowed the best results. In the 0/1 KP, the best solutions were found when using replacement rates inferior to 60%.
Figures 6 and 7 show the behavior of the GA minimizing the Ackley's test function and maximizing the 0/1 KP, respectively.

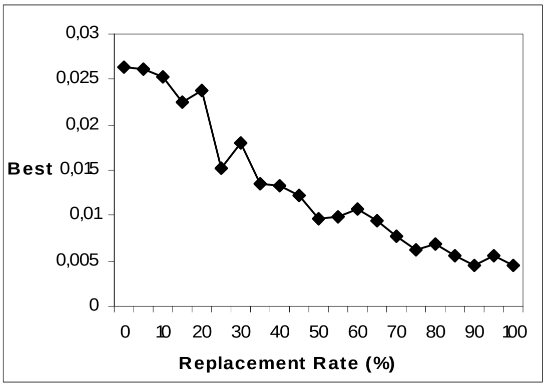

Fig. 6. Influence of the Replacement Rate in the Best Solution Found (minimization of Ackley's test function)

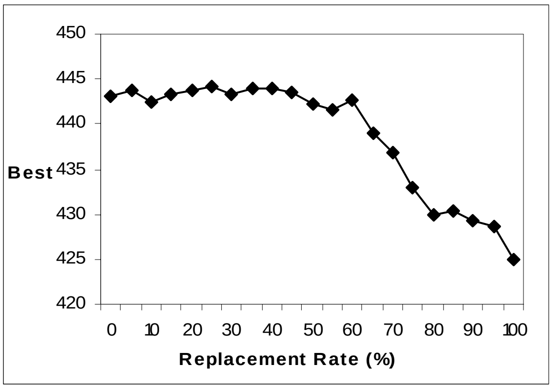

Fig. 7. Influence of the Replacement Rate in the Best Solution Found (maximization of the 0/1 KP – 100 items)

5.3 Mutation rate

The mutation rate was changed from 0% to 20% (0%, 1%, 5%, 10%, 15% and 20%). In both problems mutation rate equal to 0% or 1% allowed the GA to achieve the best performances. As we increase the mutation rate, the results become worst. Figures 8 and 9 show the results.
The observed results can be understand if we analyze the population's diversity when using 0% and 20% of mutation rate. If we observe figures 10 and 11 we can see that with 20% mutation, the diversity in the population is very high and consequently it was introduced a great degree of disruption. Transformation without any mutation was able to maintain the population's diversity within a reasonable value and lead the GA to the best solutions.

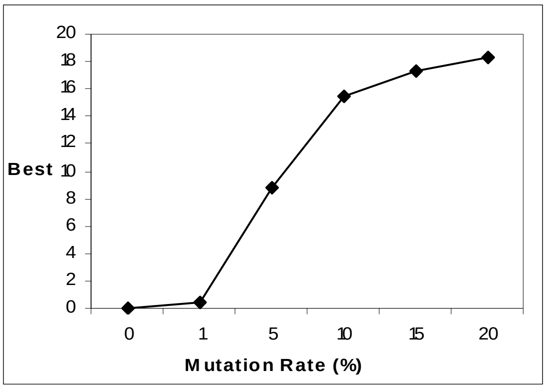

Fig. 8. Influence of the Mutation Rate in the Best Solution Found (minimization of Ackley's test function)

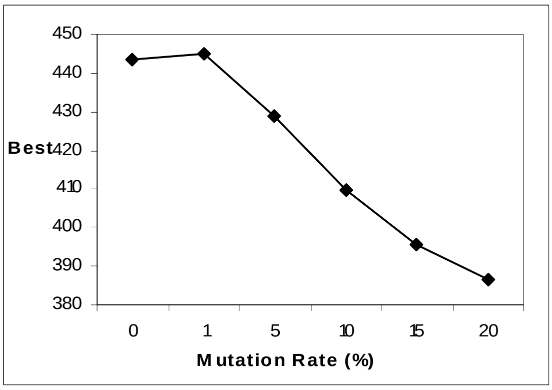

Fig. 9. Influence of the Mutation Rate in the Best Solution Found (maximization of the 0/1 KP – 100 items)

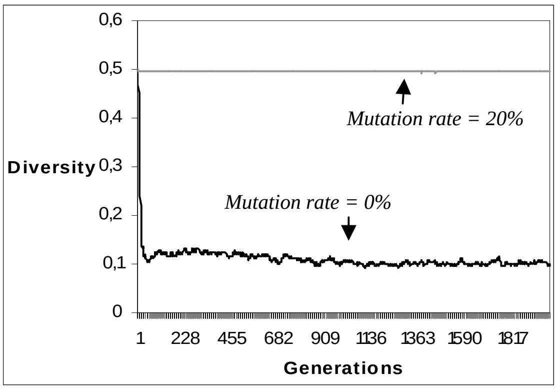

Fig. 10. Population's Diversity using Mutation rate = 0% and 20% (Ackley's Function)

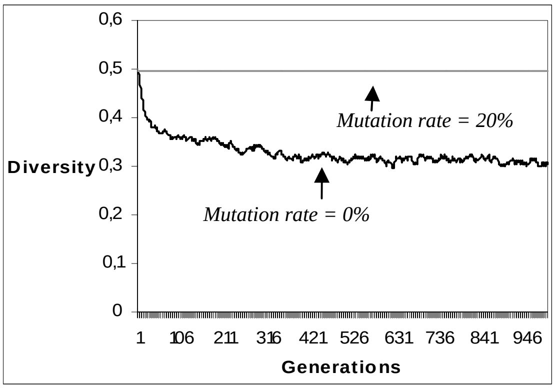

Fig. 11. Population's Diversity using Mutation rate = 0% and 20% (0/1 KP – 100 items)

5.4 Transformation rate

Transformation rate was changed from 0% to 100%. It is obvious that a rate equal to 0% didn't allow the evolution towards the global optimum, because no mutation was used. The best choice for this parameter, for both problems was a value superior to 60%.

Figures 12 and 13 show the influence of the transformation rate in the minimization of Ackley's function and in the maximization of the 0/1 KP with 100 items.

The transformation rate that leads the GA to the minimum values of the functions was 90%. In the case of the 0/1 KP, for all the instances, values of 90% or 100% of transformation rate allowed the GA to achieve the best solutions.

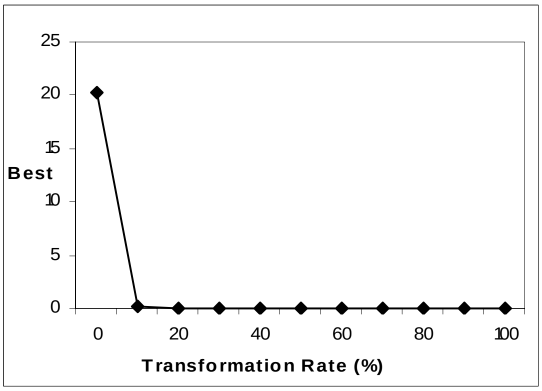

Fig. 12. Influence of the Transformation Rate in the Best Solution Found (Ackley's test function)

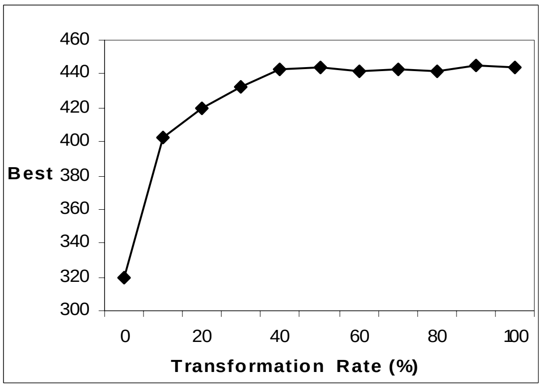

Fig. 13. Influence of the Transformation Rate in the Best Solution Found (maximization of the 0/1 KP – 100 items)

5.5 Global influence of the parameters

After the study we can conclude that the choice of parameters is an important aspect when using transformation in the GA. All the parameters influenced the results in an expressive manner.

We run the GA with a new set of parameters and we compared the obtained results with the solutions found by the GA with the initial parameter setting. Table 4 reports the results for the function optimization domain. As we can see, choosing the appropriate values for the parameters the quality of the results was significantly increased.

Table 5 shows the results obtained maximizing the 0/1 KP.

Table 4. Comparing Previous Results with the Results Obtained with the Appropriate Parameter Setting (Function Optimization)

Parametric Study				Random Choice of Parameters			
Segment length=5 Replacement Rate=90% Transformation Rate= 70% Mutation Rate=0.0%				Segment length=random Replacement Rate=90% Transformation Rate= 70% Mutation Rate=0.1%			
Nº evals->	*50000*	*100000*	*200000*	*Nº evals->*	*50000*	*100000*	*200000*
Ackley	2.678	0.044	**0.002**	Ackley	3.128	0.300	0.002
Griewangk	0.001	0.000	**0.000**	Griewangk	0.010	0.003	0.001
Rastrigin	8.290	0.821	**0.001**	Rastrigin	38.401	18.828	6.540
Schwefel	0.147	0.031	**0.008**	Schwefel	36.212	0.475	0.077

As we can see, choosing the GA parameters with some criteria, the results obtained in the minimization of the four test functions were quite better than the results achieved in our initial work, where the parameters were chosen based only in a small set of experiments.

Table 5. Comparing Previous Results with the Results Obtained with the Appropriate Parameter Setting (0/1 KP)

Parametric Study			Random Choice of Parameters		
Segment length=5 Replacement Rate=50% Transf. Rate= 90% Mutation Rate=0.0%			Segment length=random Replacement Rate=90% Transf. Rate= 70% Mutation Rate=0.1%		
Pop size->	*50*	*100*	*Pop size->*	*50*	*100*
50 items	204.60	**204.90**	50 items	197.30	197.80
100 items	442.50	**444.47**	100 items	413.00	408.40
250 items	**955.20**	954.60	250 items	838.50	834.87
500 items	**1926.87**	1910.00	500 items	1666.20	1669.07

In the 0/1 KP problem, the results were also improved. If we compare the new results with the ones obtained by the TGA, we can see that choosing the correct values for the parameters, the GA was able to reach higher values.

6 Conclusions

We used a new biologically inspired genetic operator called transformation in the GA in two different problem domains: function optimization and combinatorial optimization.

Previous work using this genetic operator showed that it is capable of preserve the population diversity during the entire evolutionary process. Nevertheless, the choice of the parameters to run the GA with transformation was always made without any strongly supported criteria.

In this paper we performed a parametric study to enhance the GA's performance when using transformation. In this study we varied four parameters: the gene segment length, the replacement rate, the mutation rate and the transformation rate. The results showed that the choice of these parameters influenced the results. In fact, using an appropriated parameter setting the GA achieved much better solutions than the ones obtained with the first set of parameters.

All the parameters had great influence in the obtained results.

As we increase the size of the segments the results become worst. Studies involving the population's diversity indicate that larger segments introduce more disruption in the individuals of the population.

Concerning the mutation rate, values of 0% allowed the ETGA to achieve the best solutions. In fact, the algorithm is able of continue evolving only with the application of transformation, and no mutation is necessary.

The replacement rate, i.e., the percentage of individuals that contribute in the generation of the gene segment pool of the next generation, must be chose in the appropriated interval: in the minimization of the test functions, replacement rates superior to 80% were the correct choice: in the maximization of the 0/1 KP, the appropriate choice was a value between 50% to 100%.

The correct choice for the transformation rate was a value in the interval 60% to 100%.

Combining all those parameters and choosing the correct values we obtained better results than the ones achieved previously by the TGA.

Acknowledgments

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

References

1. D. E. Goldberg (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Publishing Company, Inc, 1989.

2. Z. Michalewicz (1999). *Genetic Algorithms + Data Structures = Evolution Programs*. 3rd Edition Springer-Verlag, 1999.

3. P. J. Russell (1998). *Genetics*. 5th edition, Addison-Wesley, 1998.

4. A. Simões, E. Costa (2001a). *Using Biological Inspiration to Deal with Dynamic Environments*. Proceedings of the Seventh International Conference on Soft Computing (MENDEL'01), pp. 7-12, Brno, Czech Republic, 6-8 June, Brno University of Technology, 2001.

5. A. Simões, E. Costa (2001b). *On Biologically Inspired Genetic Operators: Using Transformation in the Standard Genetic Algorithm*. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2001), pp. 584-591, San Francisco, USA, 7-11 July, Morgan Kaufmann Publishers, 2001.