

# A Homologous Gene Replacement based Genetic Algorithm

Sumaiya Iqbal

University of New Orleans  
2000 Lakeshore Dr, New Orleans, LA 70148  
+1-(504)-261-0184  
siqbal1@uno.edu

Md Tamjidul Hoque

University of New Orleans  
2000 Lakeshore Dr, New Orleans, LA 70148  
+1-(504)-280-2406  
thoque@uno.edu

## ABSTRACT

This paper introduces a new genetic operator, called homologous gene replacement (hGR) applied to the chromosome of genetic algorithm (GA). The new genetic algorithm is referred as hGRGA. This operator aims to extend the ground idea behind the biological evolutionary process based classical genetic algorithm that relies on localizing and utilizing good local schema present in the genes of a chromosome. The operator refurbishes the chromosomes in gene level to boost their overall functionality. The proposed hGRGA is evaluated by widely-used benchmark functions. The simulation results was promising in terms of convergence speed and preciseness in finding optima.

## Keywords

Genetic algorithm; homologous gene replacement; optimization; benchmark test functions.

## 1. INTRODUCTION

Evolutionary computing has become a popular tool for solving real world combinatorial and global optimization problems with rapidly increasing size and complexity in present days. The stochastic search based evolutionary algorithms are well applicable to handle wide variety of scientific research and engineering applications. A rich literature is available on various evolutionary heuristics and swarm intelligence based search algorithms with numerous applications [1-7].

The idea of GA, first described by John Holland [8], was inspired by the Darwinian principles of biological evolution and adaptation in nature. The classical GA, also referred as simple GA (SGA), incorporates two operators, *crossover* and *mutation*, to address two primary challenges involved in the search process, intensify the existing knowledge (*exploitation*) and diversify to discover new knowledge (*exploration*). The underlying principle of GA in [9] states that there exists good ‘*schema*’ in chromosomes that contribute to their higher fitness. In nature, chromosome consists of genes that are connected together within it. Each of these genes is responsible for a function of that organism. Inspired by this idea, we introduce a genetic operator, *homologous gene replacement* (hGR). This operator exploits a chromosome by locating the best gene template in a chromosome and replacing relatively less fit genes of that particular chromosome with the better gene schemata. In addition to the

local exploitation operator, hGR, we include *twin removal* (TR) [10, 11] to have balanced exploration in the proposed algorithm, referred as *hGRGA*. We tested hGRGA on 15 benchmark functions for numerical optimization.

## 2. HOMOLOGOUS GENE REPLACEMENT based hGRGA

Even with a good selection procedure, better solution schema or, chromosomes can be lost due to the disruptive effects of crossover and mutation. We enhance the benefit provided by elitism that can reduce this loss, one step ahead by boosting up the fitness of the elites. In hGR, we replace relatively worse genes present in the elite-set by the best gene schema that is a homolog of the weak genes from same elite chromosome. We apply hGR on the elites to prioritize the schema fragments of elites only and to avoid the saturation of population with the schema of the best gene that may occur by applying hGR to all the individuals. This operator implicitly propagates the good local schema present in the gene of an elite contributing higher fitness to it, from local gene to the full chromosome, improving the overall fitness or, functional capacity of that chromosome. This phenomenon is reasonable since the combination of fully functional local genes can form the fitter chromosome of an organism in nature.

To apply hGR operator, we evaluate the relative fitness of the individual genes of the elite chromosomes and sort them. To compute the fitness contribution of a particular gene (one variable in the solution), we deactivated the effect of the rest of the genes using zero as the value of the corresponding variable. Then, we insert the best gene with healthier schema in place of the unhealthy genes from the sorted sequence. We define the rate of gene replacement ( $r$ ) that determines the number of unhealthy genes to be replaced,  $n = r \times d$  in each elite, where  $d$  is the number of genes of a chromosome. We keep on applying hGR with increased rate of gene replacement till the overall fitness of the elite keeps on increasing with respect to the original one or, there are no more genes to be replaced ( $n > d$ ). The proposed GA variation, hGRGA, combines hGR with crossover, mutation and TR. We apply hGR on the elites before crossover to guide the selection procedure towards better chromosomes for next generation. Therefore, the hGR operator can effectively enhance the exploitation capacity of GA. In addition, we utilize the TR operator to have balanced diversification.

## 3. SIMULATION RESULTS

Fifteen benchmark test functions are used to assess and compare the performance of hGRGA collected from the basic functions used in the latest Competition, CEC 2013 and 2014 [12, 13]. We categorize the fifteen functions into four types based on their properties. The details of the functions with their global minima and the minimum function values are available in [14]. We report

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s).  
GECCO'16 Companion, July 20-24, 2016, Denver, CO, USA  
ACM 978-1-4503-4323-7/16/07.  
<http://dx.doi.org/10.1145/2908961.2909020>

simulation results for reduced number of functions in **Table 1**. We set the number of genes in each chromosome (number of variables or, dimensions),  $d$  equal to 10, and the other GA parameter values can be found in [14]. We ran SGA, TRGA [10] and hGRGA for 20 times with 2000 epochs in each time. We report the best (minimum) and mean function value found along with the average number of epochs required to achieve the minimum function value out of the 20 runs.

**Table 1:** Numerical performance of SGA, TRGA and hGRGA on benchmark test functions.

Function	Method	Function value		Epoch ( <i>mean</i> )
		<i>best</i>	<i>mean</i>	
Type I: unimodal, separable				
$f_2$	hGRGA	0	<b>0</b>	<u>25.85</u>
(Discus)	TRGA	0	<b>0</b>	174.15
	SGA	0	<b>0</b>	172.65
Type II: unimodal, nonseparable				
$f_5$	hGRGA	0	<b>0</b>	<u>25.75</u>
(Schwefel 2.2)	TRGA	0	<b>0</b>	161.65
	SGA	0	<b>0</b>	188.70
Type III: multimodal, separable				
$f_6$	hGRGA	0	<b>0</b>	<u>33.40</u>
(Rastrigin)	TRGA	0.9950	3.5837	632.20
	SGA	0.9952	4.2306	195.90
$f_7$	hGRGA	1.27e-04	<b>0.0601</b>	<u>65.80</u>
(Schwefel 2.6)	TRGA	129.309	309.905	1410.6
	SGA	308.918	780.969	422.70
Type IV: multimodal, nonseparable				
$f_{11}$	hGRGA	0	<b>0</b>	<u>57.45</u>
(Griewank)	TRGA	0.1299	0.4884	364.20
	SGA	0.0426	0.4896	363
$f_{14}$	hGRGA	0	<b>0.009</b>	<u>77.6</u>
(Expanded Schaffer F6)	TRGA	0.3587	1.1864	1307.6
	SGA	0.3757	1.6178	547.75

The best mean function value is highlighted by bold.

The fastest convergence in terms of number of epoch is underlined.

For type I and II functions, hGRGA could locate the unique global minima,  $f_2(x^*) = 0 = f_5(x^*)$ . For  $f_2$ , hGRGA converged about 85% faster than the other two algorithms and for  $f_5$ , it converged to the global minima in 84.07% (and 86.35%) fewer mean epochs than those of TRGA (and SGA). Type III and IV are the most challenging functions that endorse the searching strength of the algorithms to avoid the local traps and to obtain the global minima. Both SGA and TRGA converged prematurely in the local minima for such four functions (**Table 1**). To compare, hGRGA successfully found the global minima ( $f_6(x^*) = 0$ ) for  $f_6$  and converged only close to the global minima ( $f_7(x^*) = 0$ ) for  $f_7$  with faster convergences. This close result is reasonable as The No-Free-Lunch (NFL) theorem [15, 16] states that it is only possible to develop a promising global optimization technique for a class of problems. For  $f_{11}$  and  $f_{14}$ , hGRGA reached the global minima,  $f_{11}(x^*) = 0 = f_{14}(x^*)$  as its best performance. Moreover, the convergence speed of hGRGA exceeds the SGA and TRGA algorithms by no less than 84.17% and 88.83% for  $f_{11}$  and  $f_{14}$  respectively.

## 4. CONCLUSION

In this paper, we introduce a novel homologous gene replacement (hGR) operator that is dynamically adaptable to large number of variables. The proposed hGRGA algorithm appropriately balances the exploitation (or, intensification) and exploration (or, diversification) tasks which is crucial for any optimization algorithm. The resulting hGRGA delivered promising performance both in terms of finding minimum function value and

speed of convergence for benchmark test functions having various complex properties. Therefore, it would be interesting to apply hGRGA to solve challenging modern combinatorial optimization.

## 5. ACKNOWLEDGMENTS

Both the authors gratefully acknowledge the Louisiana Board of Regents through the Board of Regents Support Fund, LEQSF (2013-16)-RD-A-19. Special thanks to Mr. Denson Smith for critically reviewing the paper.

## 6. REFERENCE

- [1] M. Gong, Q. Cai, X. Chen, and L. Ma, "Complex network clustering by multiobjective discrete particle swarm optimization based on decomposition," *Evolutionary Computation, IEEE Transactions on*, vol. 18, no. 1, pp. 82-97, 2014.
- [2] S. Iqbal, M. Kaykobad, and M. S. Rahman, "Solving the multi-objective Vehicle Routing Problem with Soft Time Windows with the help of bees," *Swarm and Evolutionary Computation*, vol. 24, pp. 50-64, 2015.
- [3] Z. Zhang, N. Zhang, and Z. Feng, "Multi-satellite control resource scheduling based on ant colony optimization," *Expert Systems with Applications*, vol. 41, no. 6, pp. 2816-2823, 2014.
- [4] B. Maher, A. A. Albrecht, M. Loomes, X.-S. Yang, and K. Steinhöfel, "A firefly-inspired method for protein structure prediction in lattice models," *Biomolecules*, vol. 4, no. 1, pp. 56-75, 2014.
- [5] M. A. Rashid, F. Khatib, M. T. Hoque, and A. Sattar, "An Enhanced Genetic Algorithm for Ab initio Protein Structure Prediction," *IEEE Transactions on Evolutionary Computation*, no. 99, pp. 1, 2015.
- [6] Z. Bao, Y. Zhou, L. Li, and M. Ma, "A Hybrid Global Optimization Algorithm Based on Wind Driven Optimization and Differential Evolution," *Mathematical Problems in Engineering*, vol. 2015, 2015.
- [7] A. Mishra, S. Iqbal, and M. T. Hoque, "Discriminate protein decoys from native by using a scoring function based on ubiquitous Phi and Psi angles computed for all atom," *Journal of Theoretical Biology*, vol. 398, pp. 112 - 121, 2016.
- [8] J. Holland, "Adaptation in natural and artificial systems," Univ. of Michigan Press, 1975/1992.
- [9] M. Mitchell, "Genetic Algorithm: An Overview," *Complexity*, vol. 1, no. 1, pp. 31 - 39, 1995.
- [10] M. T. Hoque, M. Chetty, and L. S. Dooley, "Generalized schemata theorem incorporating twin removal for protein structure prediction," *Pattern Recognition in Bioinformatics*, pp. 84-97: Springer, 2007.
- [11] M. T. Hoque, M. Chetty, A. Lewis, and A. Sattar, "Twin removal in genetic algorithms for protein structure prediction using low-resolution model," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 1, pp. 234-245, 2011.
- [12] J. Liang, B. Qu, and P. Suganthan, "Problem definitions and evaluation criteria for the CEC 2014 special session and competition on single objective real-parameter numerical optimization," *Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou China and Technical Report, Nanyang Technological University, Singapore*, 2013.
- [13] J. Liang, B. Qu, P. Suganthan, and A. G. Hernández-Díaz, "Problem definitions and evaluation criteria for the CEC 2013 special session on real-parameter optimization," *Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou, China and Nanyang Technological University, Singapore, Technical Report*, vol. 201212, 2013.
- [14] S. Iqbal, and M. T. Hoque, "A Scalable Gene Replacement Operator for Genetic Algorithm," 2016, [http://cs.uno.edu/~tamjid/TechReport/hGRGA\\_TR20161.pdf](http://cs.uno.edu/~tamjid/TechReport/hGRGA_TR20161.pdf), TR-2016/1
- [15] D. H. Wolpert, and W. G. Macready, "No free lunch theorems for optimization," *Evolutionary Computation, IEEE Transactions on*, vol. 1, no. 1, pp. 67-82, 1997.
- [16] Y.-C. Ho, and D. L. Pepyne, "Simple explanation of the no-free-lunch theorem and its implications," *Journal of optimization theory and applications*, vol. 115, no. 3, pp. 549-570, 2002.