

# Performance Comparisons of Genotype to Phenotype Mapping Schemes for Diploid Representations in Changing Environments

A. Sima UYAR and A. Emre HARMANCI

Istanbul Technical University

Computer Engineering Department

Maslak TR80626 Istanbul - Turkey

e-mail: {etaner, harmanci}@cs.itu.edu.tr

**Abstract:** When working in changing environments, it is important for a solution approach to adapt to the change and follow the new optima. For genetic algorithms, using diploid representations is one of the methods for addressing performance and diversity issues in such environments. When using a diploid representation for individuals, the choice of a good genotype to phenotype mapping mechanism is a very important factor. The main aim of this study is to give performance comparisons of some of the existing genotype to phenotype mapping approaches, show their weaknesses and propose a new, adaptive approach that will perform better in the cases where these fail. The proposed approach is compared against the Ng-Wong and the additive diploidy approaches with a dominance change mechanism, using a controlled 0/1 Knapsack problem. All three approaches tested for this study, show similar performances for oscillating change. In the discussed random change case, the proposed approach shows better performance in automatically adapting to the change.

**Keywords:** genetic algorithms, diploid representations, genotype to phenotype mapping, adaptive dominance values, penetrance, changing environments, diversity.

## 1. Introduction

Most studies in application of genetic algorithms deal with static environments. However there's a class of problems where the environment changes in time. From a genetic algorithm point of view, the change may be in the fitness function, in the constraints or the problem instance itself. In such environments, it becomes important for the solution approach to adapt to this change and follow the new optima. Classical genetic algorithms do not have the necessary mechanisms to address the issues encountered when working in such environments. Thus it becomes necessary either to make modifications to existing algorithms or to incorporate other features. The main issues when dealing with changing environments is preserving diversity in the gene pool of the population and being able to converge to a solution on the phenotype level. The modifications or the new features have to take this issue into account and incorporate a balance between preserving diversity and converging to a solution.

There're several different approaches in literature which deal with variations of genetic algorithms suitable for working in changing environments. Since the choice of a suitable approach depends on analyzing the nature of the change in the environment, it is worthwhile to categorize the change based on specific properties. A good set of criteria for this purpose is given in [1] as follows:

- frequency of change,
- severity of change,
- predictability of change,
- cycle length / cycle accuracy.

Using diploid representations is one of the approaches for addressing performance and diversity issues in changing environments. When using a diploid representation for individuals, the choice of a good genotype to phenotype mapping mechanism is a very important factor in performance. There has been

some research done in this area. Results and discussions can be found in literature, some of which are given in [1,2,3,5,6].

The main aim of this study is to give performance comparisons of some of the existing genotype to phenotype mapping approaches, show their weaknesses and propose a new, adaptive approach that will perform better in the cases where these fail.

## 2. The Proposed Approach

The adaptive dominance approach proposed in this study is loosely based on a natural phenomenon known as the *penetrance* of a gene [12]. The main factor for this phenomenon in nature is environmental change. The level of penetrance for a gene may be calculated as the proportion of individuals with a given genotype which exhibit a particular phenotype. Under incomplete penetrance, a dominant trait may not be expressed in the phenotype for a given genotype. This implies that the level of penetrance of a gene may change when environmental conditions are altered.

For the proposed approach in this study, a binary representation of the alleles for each gene location is used. Each individual has three binary strings, i.e. two chromosome strings making up the genotype and one phenotype string. The alleles are not taken to be either dominant or recessive but a level of penetrance for each is defined. This will be referred to as the dominance factor of an allele for a specific location. These levels are determined at each generation based on the effects of the environment. A global domination map is used for genotype to phenotype mapping. The length of the map is the same as the length of the chromosomes. Each location on the map may take on a real value in [0.0,1.0]. This value shows the domination factor of the allele 1 over the allele 0 for the corresponding location on the chromosomes. For example if the value for the 5<sup>th</sup> location on the domination map is 0.7 and the corresponding 5<sup>th</sup> locations on the chromosomes of an individual are different, the individual will show the allele 1 in its phenotype for the 5<sup>th</sup> location with probability 0.7 and the allele 0 with probability 0.3.

The calculation of the domination map involves a statistical scan of the fitnesses and the phenotypes of the individuals in each generation. In determining the dominance factor of an allele, a weighted proportion based on relative fitnesses of individuals having that allele is calculated. The introduction of the proportions being weighted according to relative fitnesses, artificially incorporates the effects of environmental factors into the calculation of the dominance values. The domination map is re-evaluated at the end of each generation using Equation 1.

$$dom(i) = \frac{\sum_j (p_{ij} * f_j)}{\sum_j f_j} \quad i = 1, \dots, ChrmLen \text{ and } j = 1, \dots, PopSize \quad (1)$$

where  $p_{ij}$  shows the  $i$ th phenotypic value of the  $j$ th individual and  $f$  shows the fitness of the  $j$ th individual.  $dom(i)$  will be higher if the individuals with allele 1 for the  $i$ th location on their chromosomes have higher fitnesses than those having the allele 0. Since the domination map is one of the driving forces of the population, the dominance values corresponding to locations on the phenotype that are optimally 1 are expected to approach 1.0 and 0.0 for the case where the optimal is 0. All locations on the domination map are initialized to 0.5, meaning that allele 1 and allele 0 are equally dominant in the beginning. As a result of re-calculating the domination map in each generation, the adaptation is automatic and it is not required to detect the change to make the necessary modifications.

## 3. Diploid Approaches Chosen for Comparisons

There're several diploid approaches in literature and references for these may be found in [1,2,3,4,5,6,8,9].

It has been shown in [7] and [10] that in the case of changing environments, some type of a dominance change mechanism is needed. The approaches which do not incorporate this feature do not perform better than a simple, haploid genetic algorithm. Based on these results, the approaches with fixed dominance mechanisms will not be used for comparisons in this study.

In [7], the domination approach proposed by Ng-Wong [8] and the additive diploidy approach proposed by Ryan [9] are modified to incorporate a dominance change mechanism as is explained in the following subsections. The results reported in [7] conclude that both approaches are good for environments where the change oscillates between two solutions with fixed intervals but fail to give the same level of performance for the case where the change instances occur randomly.

### 3.1 Modified Ng-Wong Approach

In the basic Ng-Wong approach [8], there are four alleles, two of which are dominant given by **0** and **1** and the others recessive given by *i* and *o*. In the phenotype the dominant allele is expressed. The dominance map to obtain the phenotype from the genotype is given in Table 1 where 0/1 means that 0 and 1 are equally likely).

Table 1 Dominance map for Ng-Wong

	<b>0</b>	<i>o</i>	<b>1</b>	<i>i</i>
<b>0</b>	0	0	0/1	0
<i>o</i>	0	0	1	0/1
<b>1</b>	0/1	1	1	1
<i>i</i>	0	0/1	1	1

As given in [7], Ng-Wong allows a dominance change to occur when the fitness of a population member drops by a particular percentage between successive evaluation cycles. Dominance change is achieved by inverting the dominance values of all allele pairs, i.e. 11 becomes *ii*, 00 becomes *oo*, 1*o* becomes *i0* and 0*i* becomes *o1*.

### 3.2 Modified Additive Diploidy

In the basic additive diploidy approach [9], alleles are represented by ordered values that are combined using a pseudo-arithmetic to obtain the phenotype. In [9], each value is associated with a number (A:2, B:3, C:7, D:9) and addition is performed. If the resulting sum is above a threshold (chosen as 10), the phenotype becomes 1 and if it is below, it becomes 0. The resulting map is given in Table 2.

Table 2 Dominance map for additive diploidy

	A	B	C	D
A	0	0	0	1
B	0	0	0	1
C	0	0	1	1
D	1	1	1	1

In [7], the above approach is modified to incorporate a dominance change mechanism where alleles in the genotype are demoted or promoted by one grade (demoting a B makes it an A and promoting it makes it a C). As in the Ng-Wong approach, dominance change is allowed to occur when the fitness of a population member drops by a particular percentage between successive evaluation cycles. To achieve the dominance change, for each locus one of the genotypic alleles is chosen at random and if the phenotypic expression for this locus is 1, then the chosen genotypic allele is demoted (unless it is

an A) or if the phenotypic expression for this locus is 0, then the chosen genotypic allele is promoted (unless it is a D).

## 4. Test Method

In the tests performed for this study, the modified Ng-Wong and additive diploidy approaches are compared against the proposed domination mechanism through a controlled set of changes using a modified 0/1 knapsack problem. In these tests, change is introduced through changing the weight constraint for the knapsack instance. The knapsack is chosen so that for each weight constraint, there's only one solution. Three types of change are used for the comparisons. In the first case, the optima oscillate between two values with fixed intervals, in the second the change is random and is low to moderate in severity and in the third the change is random and is high in severity.

The change instances that are used for the discussions is given in Table 3. The generation no. shows the generation at which the change occurs, the weight constraint shows the new constraint to be applied, the optimum value is the optimum solution for the knapsack for the given constraint and the hamming distance is the hamming distance between the previous and the current optimum solutions.

Table 3 Discussed Change Instances for Random Change Case

Generation no.	Weight Constraint	Optimum Value	Hamming Distance
0	22	112648	0
1500	100	131066	9
2100	40	128512	11

The comparisons are done based on the ability of the algorithms in tracking the change and conclusions are given as to why and how these results are obtained.

## 5. Tests and Results

Results of the tests are given as plots of performance for each approach. In the plots, the x-axis shows the number of generations and the y-axis shows the maximum values obtained. The optimum solution and the best values obtained for each generation are plotted. For each test case the size of the populaton is 250, the mutation probability is 0.001 and the crossover probability is 0.9. For the Ng-Wong and the additive diploidy, domination change occurs when the fitness of an individual drops by 20% between successive evaluations.

For the first set of tests, the weight constraint is changed between  $W=115$  and  $W=40$  every 500 generations. A representative part of the resulting plots for the Ng-Wong, the additive diploidy and the proposed dominance approach are given in Fig. 1, Fig. 2 and Fig. 3 respectively. The results obtained for Ng-Wong and the additive diploidy are in keeping with those reported in [7]. It is seen that the proposed approach also gives similar performance in this case.

For the second set of tests, the weight constraint is changed randomly. It takes on random values at random intervals as given in Table 3. It should be noted from the hamming distance values given, the severity of the change is moderate. For the discussions, only a representative part of the whole generations is considered. The best values obtained and the generations at which they are found for each change instance are given in Table 4. The resulting plots for the Ng-Wong, the additive diploidy and the proposed dominance approach are given in Fig. 4, Fig. 5 and Fig. 6 respectively. It is observed that the proposed approach outperforms the other two in the given test case. It adapts to the change better and is able to follow the new solutions after a change instance while it also reaches acceptable solutions much quicker than the other two. These results are based on the case where the change is random and it is low to moderate in severity. However further change instances where the change is highly severe (given by high hamming distances between subsequent solutions) were also considered in the scope of this study. Plots are not provided, however it is found that even though the proposed approach shows good performance for the low to moderately severe changes, it too fails considerably in the case where the change is high in severity .

## 6. Conclusion

The results obtained for the modified Ng-Wong and additive diploidy approaches are in keeping with the ones found in [7]. The proposed approach outperforms these two methods in the tested random case mainly due to the fact that its domination map adapts to the change in the environment automatically and the dominance values evolve along with the individuals in the population through generations. The re-calculation of the domination map at the end of each generation brings extra computational load, however unlike the other two approaches, it is not required to detect the change in the environment and alter the genotypic values for each individual. This is especially good for cases where the change is not severe and the new solution is close to the old one where it may be missed by approaches that rely on detecting it. The adaptive dominance map approach presented in this paper proves itself to be suitable for use in environments similar to the one presented in this study. However when the change is severe, the proposed approach also fails. This is mainly because the domination map approach introduced is not sufficient to provide the very high levels of genotypic diversity required in such cases. A possible remedy for introducing this required diversity on the genotypic level while preserving convergence on the phenotypic level is proposed in [11] by the same authors.

## References

- [1] Branke J. (2002), Evolutionary Optimization in Dynamic Environments, Kluwer Academic Publishers, 2002.
- [2] Collingwood E., Corne D., Ross P. (1996), "Useful Diversity via Multiploidy", in Proceedings of AISB Workshop on Evolutionary Computation, 1996.
- [3] Goldberg D. E. (1989), Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, 1989.
- [4] Grefenstette J. J. (1992), "Genetic Algorithms for Changing Environments", in Proceedings of Parallel Problem Solving from Nature 2, North Holland, 1992.
- [5] Hadad B. S., Eick C. F. (1997), "Supporting Polyploidy in Genetic Algorithms Using Dominance Vectors", in Proceedings of the Sixth International Conference on Evolutionary Programming, Springer Verlag, 1997.
- [6] Kim Y., Kim J. K., Lee S., Cho C., Hyung L. (1996), "Winner Take All Strategy for a Diploid Genetic Algorithm", in Proceedings of the First Asia-Pacific Conference on Simulated Evolution and Learning, 1996.
- [7] Lewis J., Hart E., Graeme R. (1998), "A Comparison of Dominance Mechanisms and Simple Mutation on Non-Stationary Problems", in Proceedings of Parallel Problem Solving from Nature, Springer Verlag, 1998.
- [8] Ng K. P., Wong K. C. (1995), "A New Diploid Scheme and Dominance Change Mechanism for Non-Stationary Function Optimization", in Proceedings of the Sixth International Conference on Genetic Algorithms, 1995.
- [9] Ryan C. (1994), "The Degree of Oneness", in Proceedings of the 1994 ECAI Workshop on Genetic Algorithms, Springer Verlag, 1994.
- [10] Smith R. E., Goldberg D. E. (1992), "Diploidy and Dominance in Artificial Genetic Search", Complex Systems, Vol. 6, pp.251-285, 1992.
- [11] Uyar A. S., Harmanci A. E. (1999), "Investigation of New Operators for a Diploid Genetic Algorithm", in Proceedings of SPIE: Application and Science of Neural Networks, Fuzzy Systems and Evolutionary Computation II, 1999.
- [12] Weaver R. F., Hedrick P. W. (1997), Genetics 3rd Ed., Wm. C. Brown Publishers, 1997.

Table 4 Found Solutions by Each Method

Ng-Wong		Additive Diploidy		damGA	
Value Found	Generation	Value Found	Generation	Value Found	Generation
112648	9	112648	25	112648	39
131040	1852	130728	1719	131066	1890
128512	2180	128512	2203	128512	2113

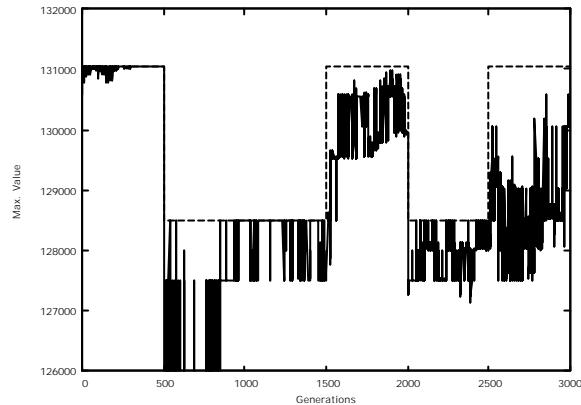


Fig. 1 Plot for Ng-Wong (oscillating change)

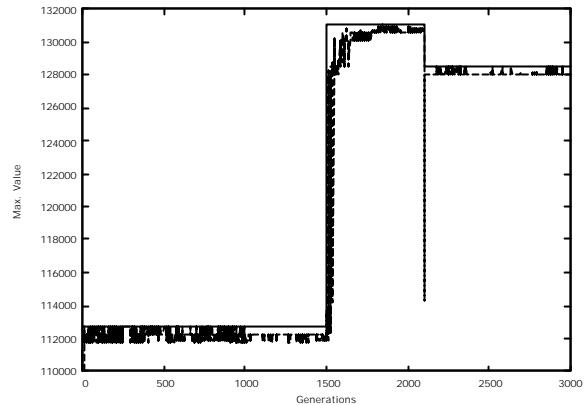


Fig. 4 Plot for Ng-Wong (random change)

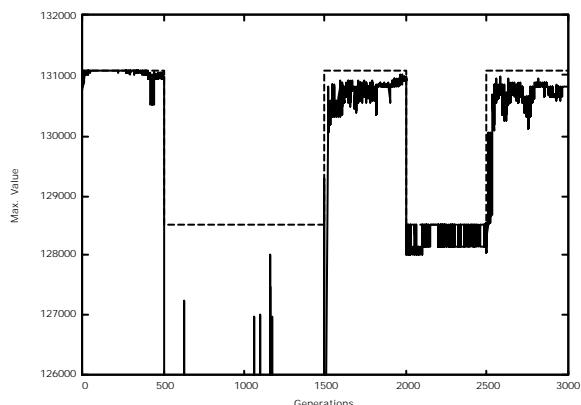


Fig. 2 Plot for additive dipl. (oscillating change)

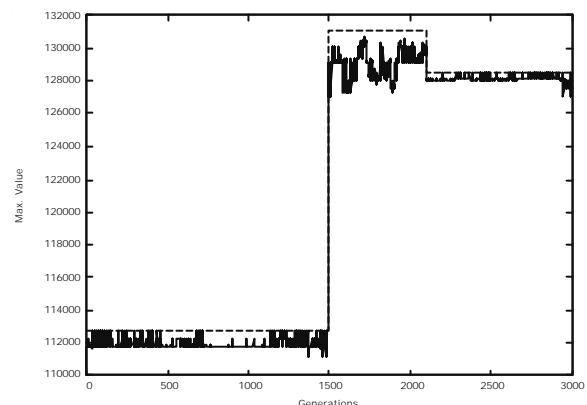


Fig. 5 Plot for additive dipl. (random change)

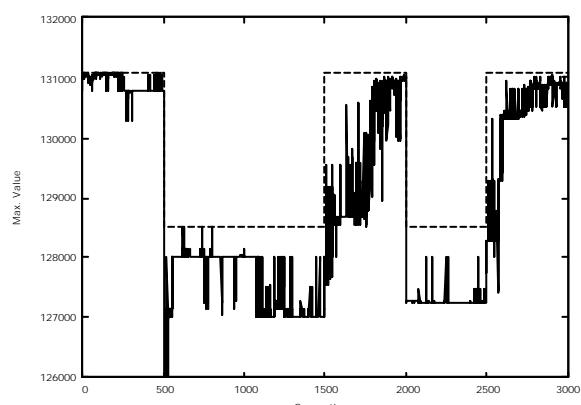


Fig. 3 Plot for proposed approach (oscillating change)

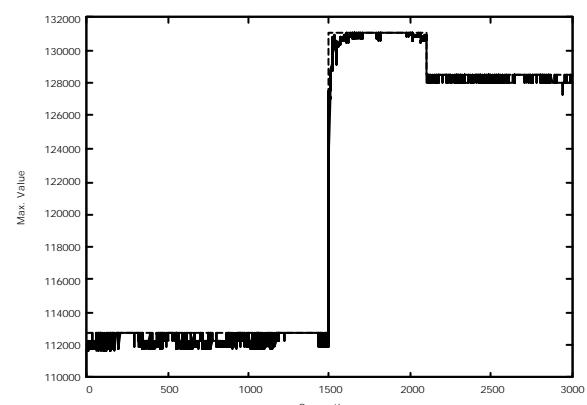


Fig. 6 Plot for proposed approach (random change)