

An Analysis of Multi-Chromosome GAs on Deceptive Problems

Menglin Li
Information Technology
NUI Galway
Galway, Ireland
M.Li1@nuigalway.ie

Colm O’Riordan
Information Technology
NUI Galway
Galway, Ireland
colm.oriordan@nuigalway.ie

Seamus Hill
Information Technology
NUI Galway
Galway, Ireland
seamus.hill@nuigalway.ie

ABSTRACT

This paper discusses a new approach to using GAs to solve deceptive fitness landscapes by incorporating mechanisms to control the convergence direction instead of simply increasing the population diversity. In order to overcome some of the difficulties that GAs face when searching deceptive landscapes, we introduce two new multi-chromosome genetic algorithms. These multi-chromosome genetic algorithms have been designed to accelerate the GA’s search speed in more complicated deceptive problems by looking for a balance between diversity and convergence. Five different problems are used in testing to illustrate the usefulness of our proposed approaches. The results show that the lack of diversity is not the only reason that normal GAs have difficulty in solving deceptive problems but that convergence direction is also important.

Categories and Subject Descriptors

I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search-Heuristic methods; G.1.6 [Numerical Analysis]: Optimization-Global optimization

General Terms

Algorithms

Keywords

Genetic Algorithms, empirical analysis, multi-chromosome representations, deceptive problems, diversity

1. INTRODUCTION

Despite the success of genetic algorithms [7] in many domains across a wide variety of classes of problems, genetic algorithms have difficulty with both deceptive problems [12] and in changing environments. These kinds of problems have been called GA-Hard problems [4]. These are often characterised by the presence of many deceptive attractors (local

optima that lead the GA away from the global optimum). [12].

Much research has shown that population diversity plays a significant role in preventing GAs from becoming trapped in local optima in deceptive problems [5] and in helping GAs track changes in dynamic environments [1]. Previous research on how to solve deceptive or dynamic environment problems has focused on maintaining diversity [6] [4] [8] [2] [10] [11] [3]. One possible approach is to maintain a many-to-one relationship between genotype and phenotype using multi-layered genotype-phenotype models [6] [4] [8]. There are also several other explicit diversity maintenance methods such as using multi-chromosomes with dominance [2] [10], using multiple populations [11], and utilising reserve selection in the algorithm [3].

In his paper [7], Holland first introduced the “building block” (or schema) concept to explain how GAs work. As the genetic algorithm converges to certain “building blocks” during evolution, particular schemata become more and more represented. Therefore, the average fitness of schema will converge, but the population diversity will decrease, which may reduce the GA’s probability of reaching the global optimum. The ideal GA is a GA that has a fast convergence speed but does not become trapped in deceptive attractors. Moreover, in some domains, it is advantageous to maintain sufficient diversity to deal with any potential environmental change. Maintaining a suitable balance between sufficient diversity and suitably efficient convergence is an ongoing important problem in GA research.

This paper introduces a new approach to solve deceptive and dynamic environment problems based on multi-chromosome representations. The dominant chromosome is used to converge to optima in the space while the recessive chromosome is used to maintain and promote diversity. In our second proposed approach, a third chromosome is added; this is used to search for local minima near the current optima. This is to check whether the current fittest individual is a local optimum in order to reduce the time spent searching near this local optimum.

The paper’s structure is as follows: some operators that used in the later section are introduced in Section 2. Section 3 discusses the role of diversity and convergence and presents an analysis on how GAs can be improved in these kinds of problems. Section 4 discusses our two new GAs, one with two chromosomes (diploid) and the other with three chromosomes (triploid). Experimental results are presented and analysed in Section 5. Section 6 presents conclusions and outlines possible further work.

Permission to make digital or hard copies of all or part 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. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO’11, July 12–16, 2011, Dublin, Ireland.

Copyright 2011 ACM 978-1-4503-0557-0/11/07 ...\$10.00.

2. BACKGROUND CONCEPTS

2.1 Diversity

Much research has focused on solving deceptive problems by maintaining population diversity [6] [4] [8] [2] [10] [11] [3]. There is no doubt that increasing diversity is one of the key approaches to solving deceptive problems or problems in changing environments. However, there are several factors that have a direct impact on GAs' performance in deceptive problems. In addition to the population diversity, convergence is another important property of genetic algorithms. We use the term convergence direction to describe the path to the building block to which the GA is currently being attracted. If the convergence direction repeatedly changes, the GAs will patrol between several solutions and will not fully converge to any of them. In general, increasing diversity will slow down the GA's convergence speed, and if the GA converges to one direction too fast, the GA will lose its diversity very quickly.

One approach to maintaining diversity is to build a many-to-one mapping between genotype and phenotype. The ability to solve deceptive problems with this approach has been shown [6] [4] [8]. The inefficiency of this approach due to the lack of convergence has also been shown. In order to find a GA which can solve deceptive problems more accurately and efficiently, the balance between convergence and diversity must be found.

2.2 Diversity Rate

In order to help explain our approach and results, some concepts are first introduced. In order to measure the level of diversity present in a population, a diversity measurement has been devised.

Wineberg and Oppacher [13] defined gene count $c_k(a)$ as the number of the genes with the value 'a' in position k. The diversity rate can be defined as follows:

$$(\text{diversityrate}) = \frac{\sum_{i=1}^L \delta_i}{\lceil \frac{P}{2} \rceil \lfloor \frac{P}{2} \rfloor L}$$

$$\delta_i = c_i(0)c_i(1)$$

$$(L = \text{ChromosomeLength}, P = \text{PopulationSize})$$

A diversity rate value of zero indicates that the population has completely converged. The greater the diversity rate, the greater the diversity in the population. The diversity rate is a universal measurement and can be used to compare different GAs in different problems regardless of the population size and the chromosome length. We use this measure of diversity in our preliminary analysis outlined in the next section.

2.3 GA-HARD PROBLEMS

There are many kinds of problem that GAs find difficult to solve. Deceptive problems [12] are probably the best known GA-Hard Problem. There are deceptive problems defined in both discrete spaces (e.g. Order 3 [6] and Order 4 [4]) and continuous spaces (e.g. Rastrigin and Schwefel's function [4]). They have a common trait that their solution space is not monotonic. Moreover, a deceptive problem may include many potential deceptive attractors. Examples include the one dimensional Rastrigin's function (illustrated in Fig 1).

In this plot, A ($x = 0$) is the global minimum, while there are several local minima such as B and C, to which the genetic algorithm may converge and become trapped.

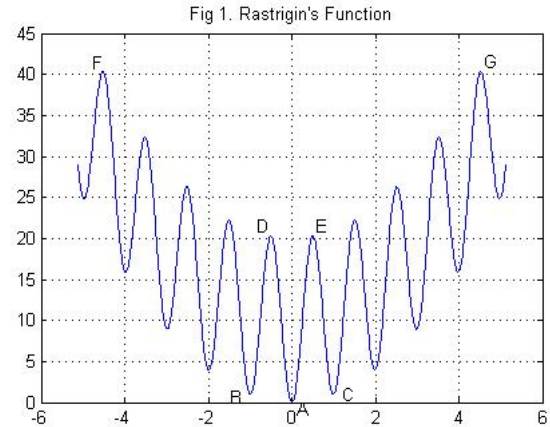


Figure 1: 1 dimensional Rastrigin's function

Another GA-Hard problem is that of dynamic environments [1] [6]. Normally, once a GA has found an optimum, the population will quickly converge to the optimum. Consequently, the population loses its diversity. Therefore, if the environment changes after the GA has converged, the global optimum may move far from the current global optimum [6], and then the GA will find it difficult to find the new optimum.

A problem can be represented as a pattern; this may be repeated several times to create a new problem. These kinds of problems are called multi dimensional problems. Increasing the dimensionality can increase the difficulty of a deceptive problem. However, a high dimensional problem, in which each pattern is not deceptive, may more difficult than a single dimensional deceptive problem.

3. DIVERSITY AND CONVERGENCE ANALYSIS

In this section we detail and discuss a number of deceptive problems that we use in our later experiments. We also provide some preliminary analysis regarding the relationship between convergence and diversity. Some preliminary results that motivate our algorithms are also included.

3.1 Single dimensional discrete problems

The order-3 and order-4 problems are common discrete deceptive problems. Both of the solution spaces are based on a small deceptive pattern which is repeated several times. For the individual patterns, there are no easy hills to climb to the pattern optimum. All points in the space, other than the global optimum, point to the deceptive attractor. For each pattern, the probability of randomly finding the pattern optimum is $1/8$ (order-3) and $1/16$ (order-4). Increasing diversity allows the GAs to have more attempts to "guess" the global optimum. For a one dimensional order-N problem, the chance of reaching the global optimum is $(1/2^N)$. Merely increasing diversity does not help GAs to solve the problem.

3.2 Single dimensional continuous problems

Rastrigin's function is a continuous GA-hard problem. The phenotype of a one dimensional Rastrigin's function is plotted in Fig 1. Once the convergence begins, all of the initial individuals will converge to the nearest optima¹. In most deceptive problems, the deceptive attractor's fitness value is very close to the global optimum (otherwise it will be easily passed by the selection function because it is not fit enough to be selected). Hence, those deceptive attractors and the global optimum have a similar chance of being selected by the selection function. This may introduce instability into the GA process, i.e., the GA may fluctuate between different optima. In other words, the GA may converge to different building blocks in different generations. As a result, the GA will patrol between the global optimum and several deceptive attractors but does not know to which it should converge. Even if it discovers the global optimum in some generation, the GA, without elitism, may not remain at the optimum.

3.3 Completely deceptive problems

Completely deceptive problem landscapes are defined as problem spaces where all the low order hyperplanes lead to some hyperplane of order n which is not the global optimum. For example, in discrete problem spaces, an order-N problem is a completely deceptive problem.

Continuous deceptive problems differ from discrete deceptive problems in that there usually exists a small area which leads to the global optimum rather than just the one point (the optimum itself) typical in most discrete deceptive problems. Still, there are completely deceptive problem in continuous solution spaces too, such as:

$$f(x) = \begin{cases} x^2 & \text{if } x \neq 0 \\ 102 & \text{if } x = 0 \end{cases} \quad x \in [-10, 10]$$

3.4 Multi-dimensional problems and patterns

Multi-dimensional problems can be created as a combination of many single problems. It may be the same problem combined to create a new problem or a set of different problems. GAs have difficulty solving these kind of problems as they may be too many deceptive attractors causing the GA to lose its ability to converge. In other words, the GA does not have a definite convergence direction. There may be many solutions in the space that have different building blocks but similar fitness scores, this will make the GA converge to different building blocks over time. The GA's convergence direction changes over time. In that case, the GA may not lose its diversity quickly, but it still cannot find the global optimum.

3.5 Elitism in non-completely deceptive problems

Through the analysis above the following points can be made. If the GA does not have a fixed convergence direction, it will patrol between the optima; hence it will not lose its diversity quickly. This is well-known fact in genetic algorithms research [9]. If the GA still has enough diversity, especially around the global optimum, following a certain

¹Disclaimer: this statement and several of our other intuitions regarding convergence are stated as working hypotheses and have not been formally proven

number of generations, it will have a high chance of reaching the global optimum, but it may lose said optima after several generations due to the instability in the convergence. As long as the fittest individual is always recorded, the global optimum can be found.

The following experiment (over 5000 generations) was undertaken to confirm the above points. We use the CGA (canonical genetic algorithm) with, and without, elitism on a 10-dimensional order-3 deceptive problem (detailed later). We record the chromosomes throughout the evolution. We represent all the patterns that reached the pattern optima (111) with '+'s, all the patterns that are trapped in pattern deceptive attractor (000) with '-'s, and the others with '.'s. So, the more '+'s in an individual, the closer it is to the global optimum, the more '-'s in an individual, the closer it is to the deceptive attractor, and the more '.'s in an individual, the more diversity the chromosome has. Single point crossover has been used in this experiment, which means '+' and '-' can only become '.' during crossover.

The results show that without elitism, the CGA converges very quickly to the local optima (in roughly 100 generations) and remains there for the remaining 4,900 generations. There tend to be many different chromosomes with similar fitnesses involving a few '+' and many '-'. At times the CGA may, via crossover, converge to contain more '+'s but is unable to maintain them. The addition of elitism enables the CGA to maintain '+'s at more locations.

When we reset the problem to be a non-deceptive 10 dimensional order-3 problem, the results are similar. This is due to the large number of small patterns present. The simplest order-N problem is the order-1 problems, when there is only one bit in each pattern. This problem can be viewed as a Hamming distance problem and can be stated as follows:

```
For each bit in the chromosome
  If this bit is 1
    Fitness of this chromosome++;
  End if
End for
```

If the chromosome length is L, the problem corresponds to looking for the greatest Hamming distance from the string of zeros. If we chose random solutions in the solution space, the probability distribution of its fitness score will follow a normal distribution as L tends to infinity. That means 99% of the set is distributed less than 3 standard deviations from the expected value (which is L/2). The probability of reaching the global optimum is very low when the chromosome length is long. Even if the global optimum has been found, because it do not have a huge advantage in terms of fitness score, without elitism, the GA may lose it very quickly. It has been shown in many bodies of work that elitism can be helpful. Furthermore, as we have shown, increasing diversity alone is insufficient for many problems. There is still potential to improve performance by finding the global optimum more frequently and more quickly while avoiding becoming trapped in local optima.

4. DIPLOID, TRIPLOID AND PEAK MAINTAINING

Based on the analysis in Section 3, two genetic algorithms

have been designed to improve the efficiency of solving hard deceptive problems.

The first GA uses a diploid chromosome representation. It has been shown that the canonical GAs without elitism cannot find the global optimum in fully deceptive order-3 problem [6]. In these runs, the average diversity rate of the population is approximately 0.6. When the elitism has been included in the GAs, the problem can be solved, but after several generations, the diversity rate decreases to 0.3~0.4. We explored the effect of adding an immigrant in each generation, but this addition had little impact on diversity. In this paper, diploid chromosome representation has been designed maintain diversity.

4.1 Diploid genetic algorithm

The dominance of this diploid representation is decided by the phenotype. Elitism and immigration have also been used in this representation.

The main idea of the diploid GA (DGA) is to use the recessive chromosome to help the dominant chromosome converge in the early stages. Once the dominant chromosomes have converged, the recessive chromosomes help maintain diversity. The algorithm's pseudo code is:

```

Initialize DGA;
Calculate both dominant and recessive chromosome's fitness;

If the recessive chromosome has a better
fitness score than the dominant chromosome
    Swap the dominant and recessive chromosome;
End if
While not reached the max generation
    Select the most fit individual in current generation
    put it into next generation
    Add an immigrant in next generation;
    While next single generation does not reach the-
    population size
        Select 2 individuals from current generation;
        Crossover the individuals to create new-
        individual by external crossover rate;
        Mutate the new individual by mutate rate;
        Crossover the dominant and recessive-
        chromosome by internal crossover rate;
        Put the new individual into next generation;
    End while
End while
End DGA

```

The selection function used here is the normal roulette wheel selection based on the dominant chromosome's fitness, and both dominant chromosome and recessive chromosome are subjected to crossover separately using one-point crossover. When putting the individual selected via elitism into the next generation, the recessive chromosome is set to be the same value as the dominant chromosome.

When the DGA starts to converge, only the dominant chromosomes are converging, because the selection is based only on the dominant chromosome's fitness. For the recessive chromosome, it was selected passively according to the fitness of the dominant chromosome. An individual having a better dominant chromosome does not mean it has a better recessive chromosome. The individual selected by elitism is set to have the same chromosome for the domi-

nant and recessive chromosome. So the recessive chromosome will give the fittest chromosome a chance to crossover with other chromosomes which do not have a outstanding fitness score. Also, when a recessive chromosome have a better fitness score than the dominant chromosome in the same individual, the dominant chromosome will be swapped with the recessive chromosome, the new dominant chromosome may have more chance of being selected, that means the recessive chromosome of the same individual, which used to be the dominant chromosome of this individual (it usually has a good fitness score) will be selected too. This will cause some convergence in the recessive chromosomes.

When the whole population's dominant chromosomes are highly converged, fitness scores will not be widely scattered. The probability of each individual being selected will be similar. For recessive chromosomes, the selection function will be more like a random selection. Because only the selection has significant impact on the diversity, both the crossover and mutation do not have a huge effect on the diversity score, so the recessive chromosomes do not converge due to the random-like selection.

There is internal crossover between each individual's dominant and recessive chromosomes; the internal crossover rate is the key to balancing the convergence of the recessive chromosome and the diversity of the dominant chromosome. When this rate has been set higher, the dominant chromosome's convergence speed will decrease, and the recessive chromosomes may lose its diversity.

4.2 Triploid Representation (three chromosomes)

The new triploid genetic algorithm (TGA) is designed to make search more efficient. It has the same diversity maintaining mechanism (recessive chromosome) as the DGA. An extra chromosome (termed a reverse chromosome) has been added in TGA, which is used to search the nearest local minimum. To maintain the minima, the reverse chromosome should converge to the local minima. Following every N generations (reverse generation), the reverse chromosome should crossover with the dominant chromosome in the same individual by a certain rate (reverse rate). With the reverse chromosomes, the TGA can solve completely deceptive problems efficiently.

For example, the reverse generation is set to 100 generations, and the reverse rate is 75 percent. Following 100 generations, for each individual, the reverse chromosome will crossover with the dominant chromosome with 75 percent probability. This will bring the entire generation's reverse chromosome toward to the solution space that the dominant chromosomes are currently searching, which could make the reverse chromosome search for the local minima, instead of the global minimum. There is an extra swapping mechanism between the reverse and dominant chromosomes in the same individual with a specified rate. The solution near the local optima may have a better chance to converge to another optimum, which may be the global optimum.

A difficulty in creating the TGA is that the dominant chromosomes converge to the global optimum while the reverse chromosomes converge to the local minimum in the same individual; two different selection functions cannot be used in one population within normal crossover. The problem is circumvented using multi-parent crossover. Two individuals (b1, b2) are selected from the current generation, the better the dominant chromosome's fitness score,

the more opportunity it has of being selected. Another two individuals (w_1 , w_2) are selected using an alternative metrics: the worse the reverse chromosome's fitness score, the more opportunity it has of being selected. During crossover, the children's dominant chromosome is produced by b_1 and b_2 's dominant chromosome, and the children's reverse chromosome is produced by w_1 and w_2 's chromosomes. The children's recessive chromosome is produced by any two of the four recessive chromosomes.

5. EXPERIMENTS AND RESULTS

Five experiments have been undertaken on each of above three algorithms (CGA, DGA, and TGA) for 200 independent runs. For all of the three GAs, the (external) crossover rate has been set to 75%, and the mutation rate to 0.5%. The internal crossover rate of DGA and TGA has been set to 2.5%. The reverse rate and reverse generation of TGA are 75% and 500 respectively. All GAs have only one individual selected by elitism and one chosen using immigration. To allow fair comparison, the population size of CGA, DGA, and TGA were set to 300, 150, and 100, so that each of them will have 300 chromosomes in every generation. The stopping condition for all runs is either when the global optimum has been found or the predefined maximum number of generations has elapsed (25000). The values in all individuals in the initial population are randomly set. Tables 1, 2, 3, 4, and 5 show the number of times that the GAs reaches the global optimum and the average generation when the optimum is reached over 200 independent runs. The plots show the average best fitness score over all generations of 200 independent runs.

The fitness scores are usually in a large range, but the deceptive attractors always have quite similar fitness scores to the global optimum. To make the plot more readable, different plots have different size scales, that means, the "important" area of the plot (which are usually near the global optimum) are "zoomed in" in the output plot. A Wilcoxon rank sum test is undertaken for statistical significance between the TGA and the CGA. The differences are shown to be statistically significant using this test on the following problems (60-bit order-3 problem, the mixed order problem, Rastrigin's function and the multi-level deceptive problem). On the completely deceptive problem the CGA fails all the time and the TGA succeeds all the time so the differences are also significant.

5.1 Order-3 60-bit problem

The order-3 problem is a classic multi-dimensional deceptive problem; each pattern of the order-3 problem is described as:

$$\begin{aligned} f(111) &= 30 \\ f(110) &= 0 & f(101) &= 0 & f(011) &= 0 \\ f(100) &= 14 & f(010) &= 22 & f(001) &= 26 \\ f(000) &= 28 \end{aligned}$$

To convert the problem into a changing environment, the single pattern's landscape changed midway (at generation 12,500) during the experiment to the following:

$$\begin{aligned} f(000) &= 30 \\ f(001) &= 0 & f(010) &= 0 & f(100) &= 0 \\ f(011) &= 14 & f(101) &= 22 & f(110) &= 26 \\ f(111) &= 28 \end{aligned}$$

The pattern has been repeated 20 times to create a tightly-ordered, 20-dimensional, 60 bit-chromosomes order-3 problem. The max generation has been set to 25000, and the environment change occurs at generation 12500.

The results are illustrated in Table 1 and Fig 2 where "B.C" indicates before environmental change and "A.C" refers to after the environmental change.

Table 1: Table (Order-3 60bits)

Functions		Global Opt. Reached	Avg. reach generation
CGA	B.C.	99.5% (199)	3741
	A.C.	98.5% (197)	17981
DGA	B.C.	99.5% (199)	4468
	A.C.	98.5% (197)	18650
TGA	B.C.	100% (200)	1873
	A.C.	100% (200)	14849

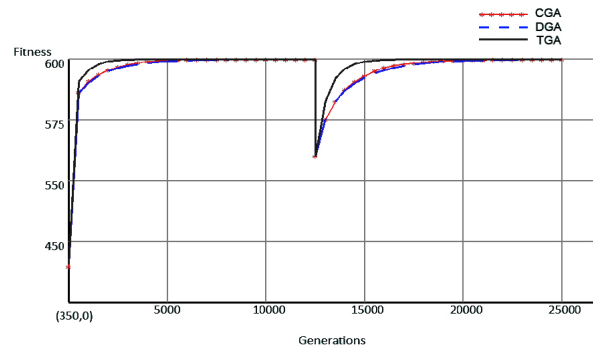


Figure 2: Average best Fitness of Order-3 60bits Problem

The TGA has a much better performance than both CGA and DGA. All of the three GAs spent more generations searching for the global optimum after the environmental change. Unexpectedly, CGA has slightly better performance than DGA.

5.2 Mixed order-3/4/5 60bit problem

A mixed order problem is a transmutation of normal order-N problem, it is defined as a random combination of five order-3, five order-4, and five order-5 patterns placed in random order. For each pattern, the pattern optimum, which has the fitness score 30, is randomly generated. The deceptive attractor which has the fitness score 28, is set to be the opposite of the global optimum, i.e. the smaller the Hamming distance to the global optimum, the smaller the fitness score is. In other words, the fitness score of the individuals except the global optimum are inversely proportional to the Hamming distance between the global optimum and itself.

At the midway point in the evolution, the order of the 15 patterns will be randomly arranged again, and for each pattern, the global optimum will also be randomly set. The

maximum generation has been set to 25000 generations, the environmental change happens on generation 12500. The experimental results are described as Table 2 and Fig 3.

Functions		Global Opt. Reached	Avg. reach generation
CGA	B.C.	27.5% (55)	8332
	A.C.	29% (58)	20948
DGA	B.C.	37.5% (75)	9134
	A.C.	36.5% (73)	21036
TGA	B.C.	100% (200)	4077
	A.C.	100% (200)	16563

Table 2: Table (Mix-Order)

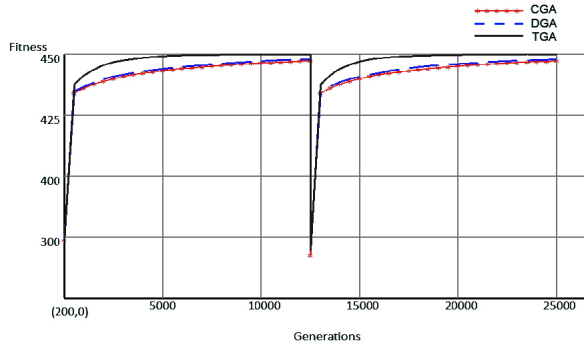


Figure 3: Average best Fitness of Mix-Order Problem

The results shows that the DGA is more accurate than the CGA in the mixed order problem; this is probably because the DGA has better diversity than the CGA. However, the performance of both the CGA and the DGA are still far behind that of the TGA.

5.3 Rastrigin’s function

Rastrigin’s function [4] is defined as:

$$R(\vec{x}) = \sum_{i=1}^N [x_i^2 - 10\cos(2\pi x_i) + 10], x_i \in [-5.12, 5.12]$$

The one dimensional Rastrigin’s function is plotted in Fig 1. In this paper, the experiments involve a 20-dimensional Rastrigin’s function. The maximum generation has been set to 25000 generations and no environmental change is included in these experiments. The chromosomes are represented with binary gray codes. Each number is encoded with 10 bits and transformed to real values in the appropriate range. To limit the size of the chromosomes, the precision used is 0.01. This is sufficient to include all the deceptive attractors in the continuous functions we tested. Increasing the precision does not introduce any new deceptive attractors.

The results are illustrated in Table 3 and Fig 4:

Using the Rastrigin’s function involves looking for the minimal value of the function, so the global optimum is at the bottom of the figure ($f(x) = 0$). TGA still performs much better than other GAs. Again, CGA performs better than the DGA.

	Global Opt. Reached	Avg. reach generation
CGA	65.5% (131)	16382
DGA	60.5% (121)	17747
TGA	90% (180)	13817

Table 3: Table (Rastrigin’s function)

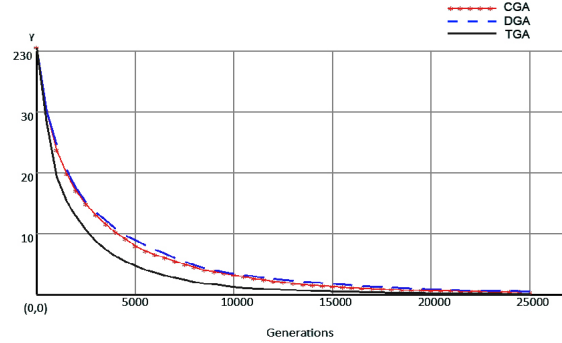


Figure 4: Average best Fitness of Rastrigin’s function

5.4 Multi-level deceptive function

Fig 1 shows the Rastrigin’s function’s plot, the deceptive attractors in Rastrigin’s function converge to the global optimum. We use a new function that has “deceptive deceptive attractor”, which we call a multi-level deceptive function. The chromosomes are also coded using binary gray code with 0.01 precision. The experiments are using 8 dimensional multi-level deceptive functions.

$$f(\vec{x}) = \sum_{i=1}^N [15 + 5\cos(1.5x + 2) + 10\sin(2\pi x)], x \in [0, 10.24]$$

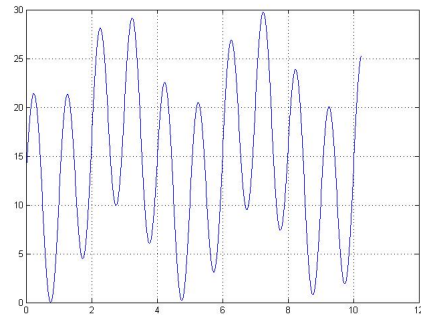


Figure 5: Multi-level deceptive function’s plot

In this deceptive function, the global optimum is $f(x) \approx 0.00072$, at point $x \approx 0.75$. There are 10 local optima including the global optimum in this deceptive function, and the deceptive attractors’ space is also deceptive. The most attractive deceptive attractor is at $x \approx 4.76$, the local optimum at $x = 4.76$ is $f(x) \approx 2.1274$. There are 3 less attractive deceptive attractors between the global optimum and the most

attractive deceptive attractor. Fig 5 demonstrates the plot of a one-dimensional multi-level deceptive function.

The experiments of the 8-dimensional multi-level deceptive function's results are show in Table 4.

	Global Opt. Reached	Avg. reach generation
CGA	37.5% (75)	12702
DGA	66.5% (133)	12067
TGA	88% (176)	10682

Table 4: Table (Multi-level deceptive function)

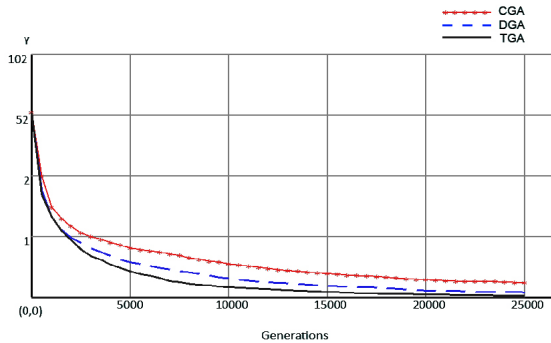


Figure 6: Average best fitness of Multi-level deceptive function

This function was originally designed to check if the TGA will have a high probability of becoming trapped at the “deceptive deceptive attractor”. The results shown in Table 4 and Fig 6 are far from what we had anticipated. TGA does not have a higher tendency to be trapped in any attractor in comparison to CGA and DGA. Also, DGA performs much better than CGA, it reaches the global optimum twice as frequently as CGA.

5.5 Completely deceptive problem

The CGA and DGA approaches have difficulty in solving completely deceptive problems. However, by maintaining the ability the converge to the local minima, the TGA should be able to solve them quickly. To demonstrate this, we created a 60-bit discrete completely deceptive problem.

The results of the completely deceptive problem are described by Table 5.

	Global Opt. Reached	Avg. reach generation
CGA	0%	NA
DGA	0%	NA
TGA	100% (200)	102

Table 5: Table (Completely Deceptive Problem)

The results, as expected, show that the TGA can solve completely deceptive problems quickly while both CGA and DGA fail in our experiments. The data and plot of the results are illustrated in Table 5 and Fig 7 respectively.

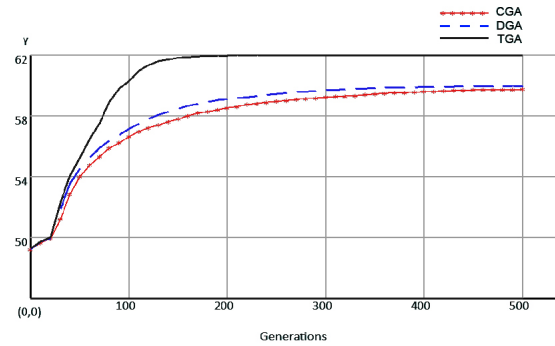


Figure 7: Average best Fitness of Completely Deceptive Problem

5.6 Exploring variants of the TGA

For completeness, we created another two-chromosome genetic algorithm by removing the recessive chromosome from the TGA and relying solely on the dominant chromosome and the reverse chromosome. It has no recessive chromosome to maintain diversity. For the mixed-order problems, this new representation does well (outperforms the CGA and the DGA). It performs very badly, as expected, on the multi-level problem as it converges and remains at local minima.

5.7 Summary

The experiments and results described in this section show that canonical GAs with elitism can solve deceptive problems, if given sufficient generations. The DGA can be considered similar to a CGA with diversity maintaining mechanisms; it does not have any new mechanism to improve the convergence. In most situations, DGA has similar performance to CGA (in the plot, the dashed line and star line are almost superposed). In some problems CGA performs even better than DGA, such as the order-3 problem and Rastrigin’s function. These results confirm our previous assumption that diversity is not the only factor that can affect GAs’ performance in deceptive problems.

Through all the experiments, TGA has been shown to be better on both accuracy and efficiency, especially in complex problems. In the mixed order problems, CGA and DGA perform very poorly, probably because the order-5 problem, which has 32 solution space requires too many “guesses” in the CGA and DGA, while TGA more easily solves these kinds of problems due to its maintainence of the local minima. The multi-level deceptive function is designed to be hard for the TGA to solve, because the “peak points” also have many “trough”/“crest” (deceptive attractors). Interestingly, TGA still has much better performance than the other two GAs.

6. CONCLUSIONS AND FURTHER WORK

The analysis of deceptive problems shows that the diversity is not the only parameter that may affect the GAs’ performance in solving these kinds of problems. In fact, in completely non-deceptive problems, the GA does not have correct convergence to find the global optimum because it cannot maintain a definite convergence direction. Increase diversity could help to solve the problems but it is not the

only way to do it. Given enough generations, the canonical GA with elitism could solve deceptive problems. Therefore, a new approach to solving deceptive problems in changing environments is by controlling the convergence direction has been proposed. Two multi-chromosome GAs have been designed to solve five different problems in both discrete and continuous spaces. The results show that increasing the diversity can increase the probability that GAs solve deceptive problems, and that the ability to maintain convergence directions affects the efficiency. Maintaining diversity while controlling the convergence direction is much more efficient than only maintaining the diversity.

Our results used fixed settings for mutation, reverse rate etc. However, we have explored a range of values and obtained similar results. However, a more complete exploration of the parameter space is left to future work. Other planned work will include the detailed comparison of this approach against other genetic algorithms other than the canonical genetic algorithm.

7. REFERENCES

- [1] L. T. Bui, J. Branke, and H. A. Abbass. Diversity as a selection pressure in dynamic environments. In *Proceedings of the 2005 conference on Genetic and evolutionary computation*, GECCO '05, pages 1557–1558, New York, NY, USA, 2005. ACM.
- [2] R. Cavill, S. Smith, and A. Tyrrell. Multi-chromosomal genetic programming. In *Proceedings of the 2005 conference on Genetic and evolutionary computation*, GECCO '05, pages 1753–1759, New York, NY, USA, 2005. ACM.
- [3] Y. Chen, J. Hu, K. Hirasawa, and S. Yu. Gars: an improved genetic algorithm with reserve selection for global optimization. In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, GECCO '07, pages 1173–1178, New York, NY, USA, 2007. ACM.
- [4] R. Chow. Evolving genotype to phenotype mappings with a multiple-chromosome genetic algorithm. In *Genetic and Evolutionary Computation - GECCO 2004*, volume 3102 of *Lecture Notes in Computer Science*, pages 1006–1017. Springer Berlin / Heidelberg, 2004. 10.1007/978-3-540-24854-5_100.
- [5] T. Friedrich, P. S. Oliveto, D. Sudholt, and C. Witt. Theoretical analysis of diversity mechanisms for global exploration. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, GECCO '08, pages 945–952, New York, NY, USA, 2008. ACM.
- [6] S. Hill and C. O’Riordan. Solving fully deceptive problems in changing environments. In *21st conference on Artificial Intelligence and Cognitive Science (AICS)*, July 2010.
- [7] J. H. Holland. Genetic algorithms. *Scientific American*, pages 66–72, July 1992.
- [8] J. L. Risco-Martín, J. I. Hidalgo, J. Lanchares, and O. Garnica. Solving discrete deceptive problems with emmrs. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, GECCO '08, pages 1139–1140, New York, NY, USA, 2008. ACM.
- [9] G. Rudolph. Convergence analysis of canonical genetic algorithms. *IEEE Transactions on Neural Networks*, 5:96–101, 1994.
- [10] R. E. Smith and D. E. Goldberg. Diploidy and dominance in artificial genetic search. *Complex Systems*, 6:251–285, 1992.
- [11] P. T and R. K. R. A dual-population genetic algorithm for adaptive diversity control. *Evolutionary Computation, IEEE Transactions on*, Issue:99:1–1, June 2010.
- [12] L. D. Whitley. Fundamental principles of deception in genetic search. In *Foundations of Genetic Algorithms*, pages 221–241. Morgan Kaufmann, 1991.
- [13] M. Wineberg and F. Oppacher. The underlying similarity of diversity measures used in evolutionary computation. In *Proceedings of the 2003 international conference on Genetic and evolutionary computation: Part II*, GECCO'03, pages 1493–1504, Berlin, Heidelberg, 2003. Springer-Verlag.