J. Edu. & Sci., Vol. 28, No. 2 , 2019 : 300-315

مقارنه بين الخوارزمية الجينيه وخوارزمية الانتقاء النسيلي لتمييز انماط الارقام اللاتينيه

مها عبدالاله محمد

جامعة الموصل

كلية علوم الحاسوب والرياضيات

قسم هندسة البرمجيات

mahamaster2000@yahoo.com

تاريخ الاستلام **تاريخ القبول**

2012/11/13 **2013/09/04**

الخلاصه

تم في هذا البحث استخدام بعض من خوارزميات التقنيات الذكائية الاصطناعية وهي الخوارزمية الجينية وخوارزمية النظام المناعي الاصطناعي – خوارزمية الانتقاء النسيلي حيث تم اجراء مقارنة بين النمط الأصلي للرقم اللاتيني والأنماط المشوهة المدخلة للرقم نفسه باستخدام الخوارزميتين أعلاه.

من خلال نتائج الأنماط التي تم الحصول عليها بعد تنفيذ الخوارزميتين نلاحظ ان خوارزمية النظام المناعي الاصطناعي(خوارزمية الانتقاء النسيلي) جيدة وبنسبة 85% في الحصول على أنماط مقاربة للنمط الاصلي للرقم اللاتيني المدخل علماً ان الخوارزميتين قائمتين على مبدأ الامثلية في الحصول على نتائج.

الكلمات المفتاحيه: تمييز الانماط، الخوارزمية الجينيه، النظام المناعي الاصطناعي، خوارزمية الانتقاء النسيلي.

Compare Between Genetic Algorithm and Clonal Selection Algorithm To Pattern Recognition Latin's Numbers

Maha Abd_alelah Mohammed
University of Mousel
Computer sciences and math.
Software Engineering
1234@uomosul.edu.iq

Received **Accepted**
13/11/2012 **04/09/2013**

Abstract

This work involves the use some of Artificial intelligence techniques algorithms which are genetic algorithm and artificial immune system algorithm- clonal selection algorithm.

A comparison was done between the original mode of the Latin number and the distorted modes applied for the same number and with the use of above algorithms.

Mode results obtained from implementing both algorithms showed that the artificial immune system algorithm(Clonal selection algorithm) was good and 85% in getting closer modes to that of the applied Latin number original mode.

Both above algorithms are based on optimization principle in getting the results.

Keywords: Patterns Recognition, Genetic Algorithm, Artificial Immune System (AIS), Clonal Selection Algorithm.

1. Introduction

The world has witnessed, in recent times, a clear development in all aspects of life computers and programming are parts of that Industrial control and medical of diagnosis were the prime aspects of development as well as other aspects that required recognition between pictures, patterns and fingerprints in various fields[1].

The increase in complex technical difficulties and the inability of classic programming solutions to deal with these difficulties pushed researches to deepen the search for efficient algorithms. That is in order to help in finding ideal and suitable solutions for complex problems as well as speeding up the process of reaching, storing and feeding back those solutions.

Then to arrive at assembling a solid structure for smart programming building blick in the field of artificial smart techniques which are Artificial Immune System, Genetic Algorithm, Artificial Neural Network, Swarm Intelligent ….etc.

In the 1990s and at the beginning of the 21st century, artificial intelligence progressed rapidly and that is due to the success in many aspects which are: today's high-speed computers, the increase in focus on solving specified secondary problems and the creation of new relationships between artificial intelligence and other fields and above all, researches started to implement strong mathematical curriculum with strict scientific standards [2].

2. Aim of Research

This work involves the use of Artificial intelligence techniques algorithms which are genetic algorithm and clonal selection algorithm that represents one of the artificial immune system algorithms.

A comparison was done between the original mode of the Latin number and the distorted modes applied for the same number and with the use of above algorithms.

3. Patterns Recognition

It is not the science of pattern recognition or the technique of distinguishing patterns from the sciences and techniques commonly mentioned and related to public culture, but has become a science used in many applications and related technologies, including techniques for distinguishing faces in control devices or techniques of drawing drawings and geometric shapes and chromosomes and all the processes in which the computer understands the patterns In the same way that human beings understand and at some high efficiency.

The term "pattern recognition" can be defined as a science used to describe, distinguish and classify scales. Pattern discrimination can be classified into three types: the distinction of statistical patterns, the distinction of syntactic patterns, and the differentiation of neural patterns [3,4].

The pattern is defined as everything to be classified, so discrimination is the ability to classify. Discrimination is used in a number of areas, including diagnosis of diseases, handwriting discrimination, identification of eye prints… etc[5].

Distinguish patterns it is part of the artificial intelligence field. Techniques and research in this field aim at finding or developing new ways to identify specific modes or structures in digital signals.

The signal may represent an image that contains a handwritten letter or musical section that representation a word or even a computer text. It is possible that the mode which is required to be identified is the letter that the image contains or the instrument used to the musical section or the uttered word in the verbal section. To describe the idea, a system is needed to be built that keeps of models P1, P2, … , Pn and when the used applied a signal P , then , the system will have to decide the model Pi that the signal represent [3,6].

There are many algorithms in use for mode recognition. A template matching algorithm was used in Latin numbers modes recognition with in genetic and clonal selection algorithms which are the main aspect of this work.

4. Genetic Algorithm

Genetic algorithms are a method of optimization and research. This method can be classified as one of the methods of evolutionary algorithms that rely on the tradition of nature's work. The genetic algorithm uses a search technique to find optimal solutions. Genetic algorithms are categorized as global search heuristics. It is also a particular category of evolutionary algorithms known as the evolutionary computation that uses technology inspired by evolutionary biology such as inheritance, mutations, and crossovers [4,1]. The items in the genetic algorithm are defined here:

Selection is the process of selecting individual chromosomes which solutions will participate in the reproduction process, which will later be the process of mixing or changing parts with other parts of solutions. **Individual or chromosome** solutions are available and are processed.

Gene is the smallest part of the individual carrying the information where the function variables are usually encoded to be in binary form (zero or one), the bit is called a gene. **Population** is the sum of available solutions [5,6].

The fitness function is the function that results in the probability of an individual entering the selection and inheriting its characteristics. As the optimal solution gives greater luck to enter the process of reproduction and succession of characteristics or change.

Crossover is a process in which the value of variables between individuals or chromosomes selected to enter the process is exchanged. **Mutation** is a change in the gene or one of its variants [4,7].

the sequence execution of genetic algorithm is[1,2]:

START
Generate the initial population
Compute fitness
REPEAT
  Selection
  Crossover
  Mutation
  Compute fitness
UNTIL population has converged
STOP

In this research input-distorted patterns were selected randomly, the number of patterns depend upon the system user. For instance, in this work four distorted modes were applied as for the crossover

method which is of "the crossover with two point " type, the distorted modes two-point crossover position changes at each stage.

As for the mutation, the bit inverse mutation method was considered in which the bit position is Selected randomly at each stage and for each distorted mode. Regarding the fitness function, it is calculated by the maching ratio between the correct position for each bit in the original mode and the correct position for each bit in the distorted mode .

This ratio is calculated in percentage rate. As for the stop criterion, by getting the highest matching rate between the distorted and original modes, then the algorithm stops.

5. Artificial Immune System(AIS)

The artificial immune system is a new method of computational intelligence, and a definition of the artificial immune system can be developed as an adaptive system inspired by the theory of immunology, models, principles, and immune mechanisms used to solve real world problems [8].

Most of the research of the artificial immune system relates to the field of pattern recognition, due to the close association of the artificial immune system with the process of pattern recognition[2]. In the past years, the artificial immune system has been widely used in many areas including control, signal processing, disease diagnosis, digital circuit design, and the development of many algorithms in the field of classification (pattern recognition)[8,9].

For the artificial immune system, the design starts with representation[8,10]. A simple model is made for the immune organs, cells and molecules, and then a set of functions that define the affinity to describe the quantitative interaction of the elements and then put a number of general purpose algorithms to give the dynamic of the immune system. The design structure of the artificial immune system can be represented by the layers shown in (figure 1)[11,12]:

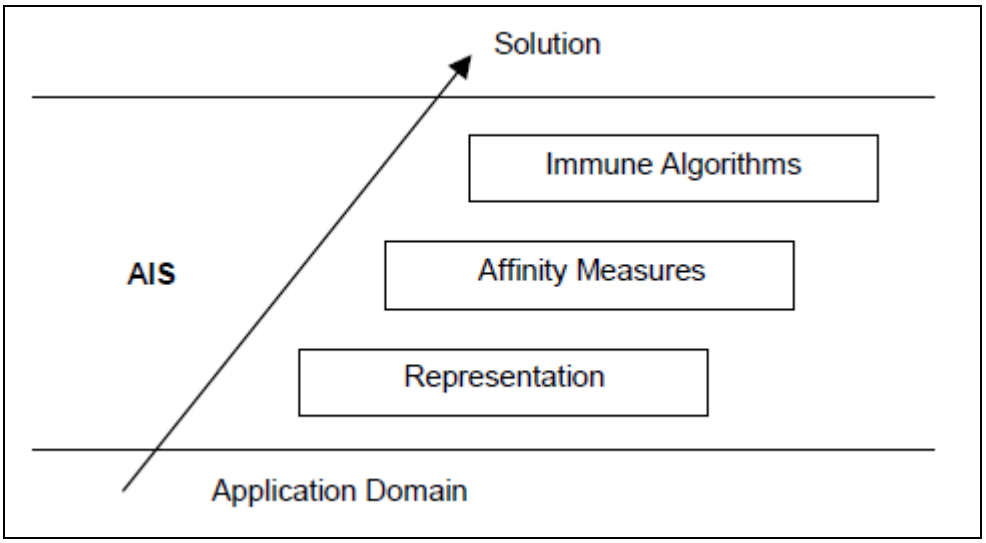


Fig.1 a structure for artificial immune system [8]

To move from the field of application to the field of computer algorithm, the data of the question must be represented in the optimal form so that we can deal with it through the immune algorithms and then we need the mechanism to measure the affinity between the antibody and the antigen and then the artificial immune algorithm is applied and these algorithm are negative and positive selection , clonal selection, and immune network algorithms in order to solve these questions [13,8].

6. Clonal Selection Algorithm

The principle of clonal selection used by the immune system is that the cells that distinguish the antigen are selected for breeding, but those that are not marked are ignored[9,14].
During the process of reproduction cells are subjected to the process of affinity maturation, a process that increases the ability to distinguish [9,15].

The first computer-based algorithm was developed in 2000 by Leandro N. de Castro [13,6].
Algorithm relied on two basic processes in the biological immune system for the purpose of finding solutions: the principle of clonal selection and the process of affinity maturation. The clonal selection algorithm was used to distinguish patterns and optimization. The steps of the clonal selection algorithm are as follows [11,9,16]:

1. Generate a random group of individuals.
2. Calculate the affinity values of each individual.
3. Sort the individual starting from the highest affinity.
4. Clone the better matching individual more with some predefined ratio.
5. Mutate the individual with some predefined ratio.
6. Calculate the new affinity values of each individual.
7. Repeat Steps 3 through 6 while the minimum error criterion is not met.

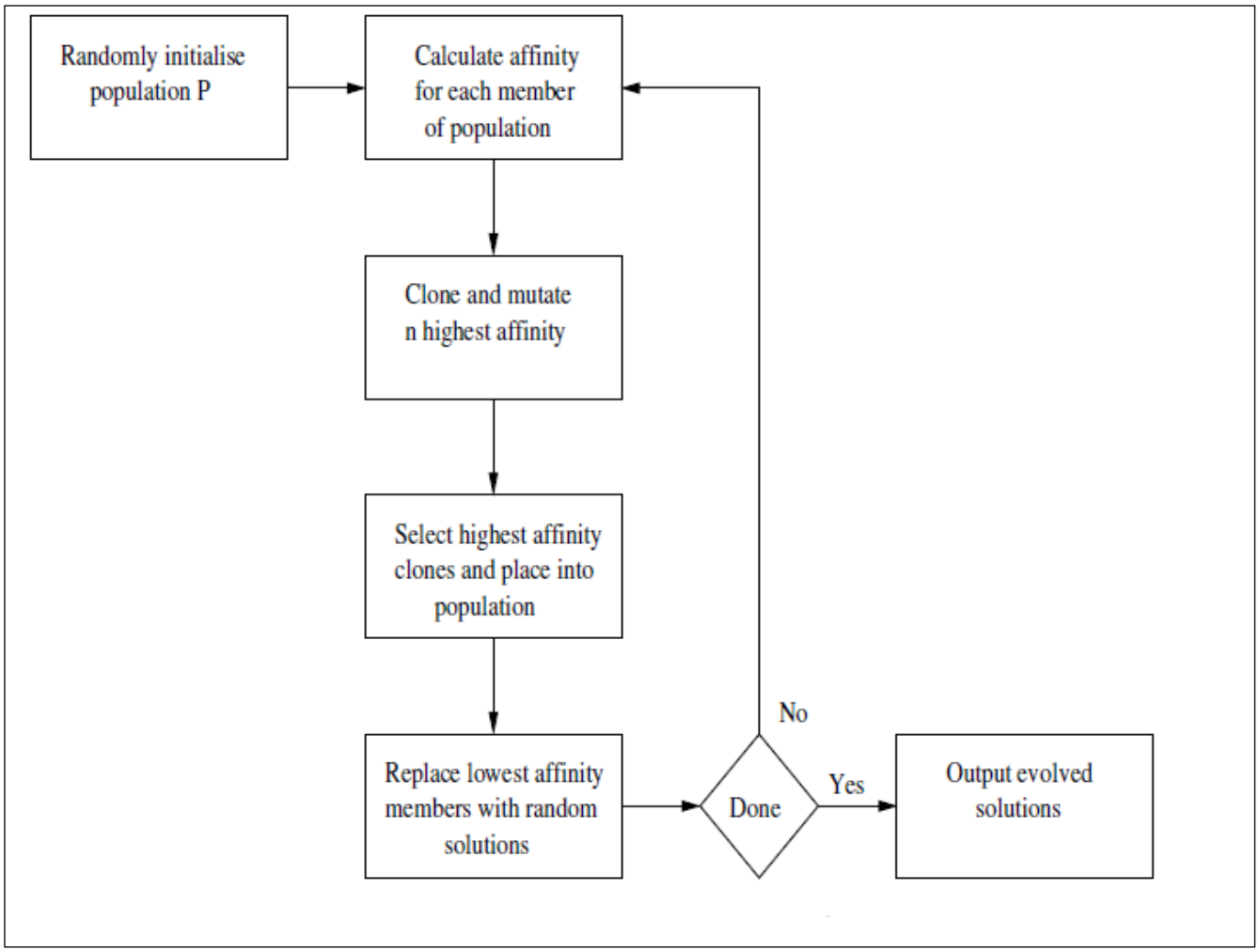


Fig. 2. Flowchart of CLONALG

7. Practical Procedure

In this work , a comparison between artificial intelligence techniques algorithms was achieved. A comparison between genetic and clonal selection algorithms was carried out in order to classify Latin numbers modes.

A Visual Basic.Net language was used for programming. The mode is treated as 8*8 matrix represented by the values (0,1) in which when the bit value is (1), that corresponds to the black color in the image, while the (0) bit corresponds to the white color. At the beginning , the two algorithm are identified with original Latin numbers. Figure(5) illustrates the process of inputing the binary values for the Latin number mode "X" which corresponds to the number " 10" in English.

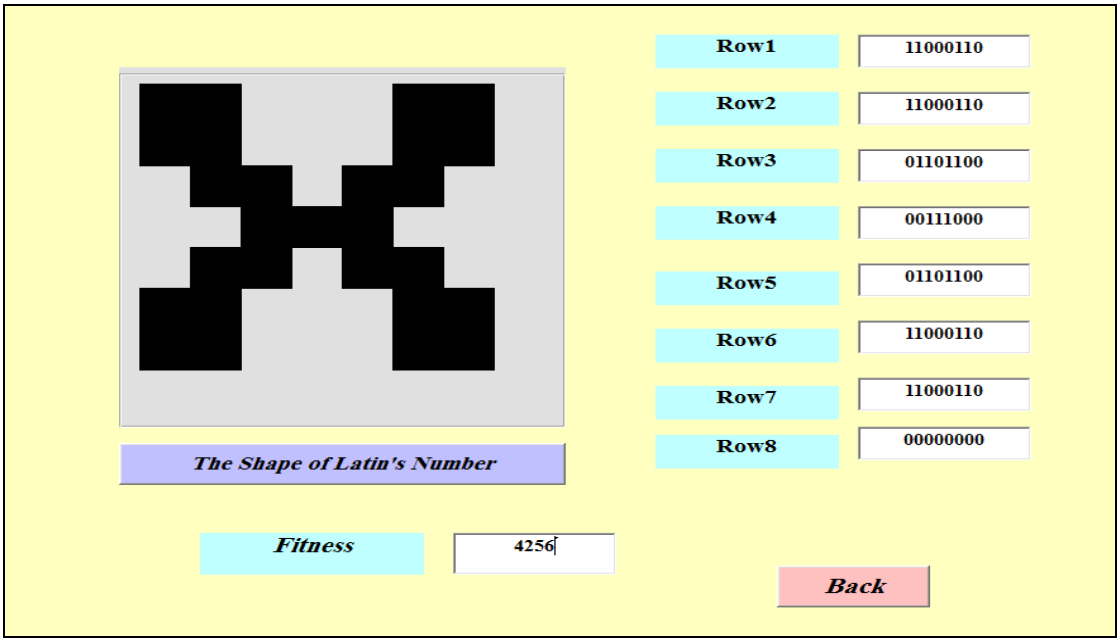
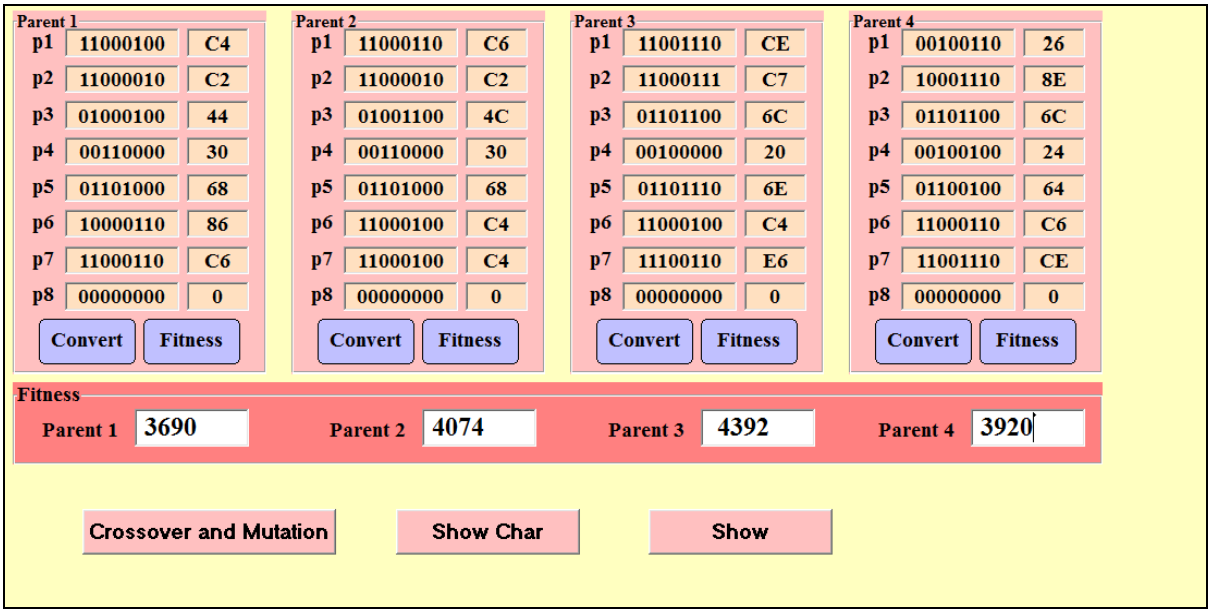


Figure (3) A template for inputting the undistorted and original Latin number's"

The figure shows the binary values for the Latin number "X" and when pressing on the button " The shape of Latin's number" , the latin number will be drawn that represents the undistorted and orginal.

The next step of the programmer is to input the distorted modes (initial generation) of the X Latin number so that they can be recognized first using clonal selection algorithm steps and then using genetic algorithm steps. Figure(4) shows the inputted distorted modes.



Figure(4) A template for inputting the distorted modes(initial generation) for the Latin number "X"

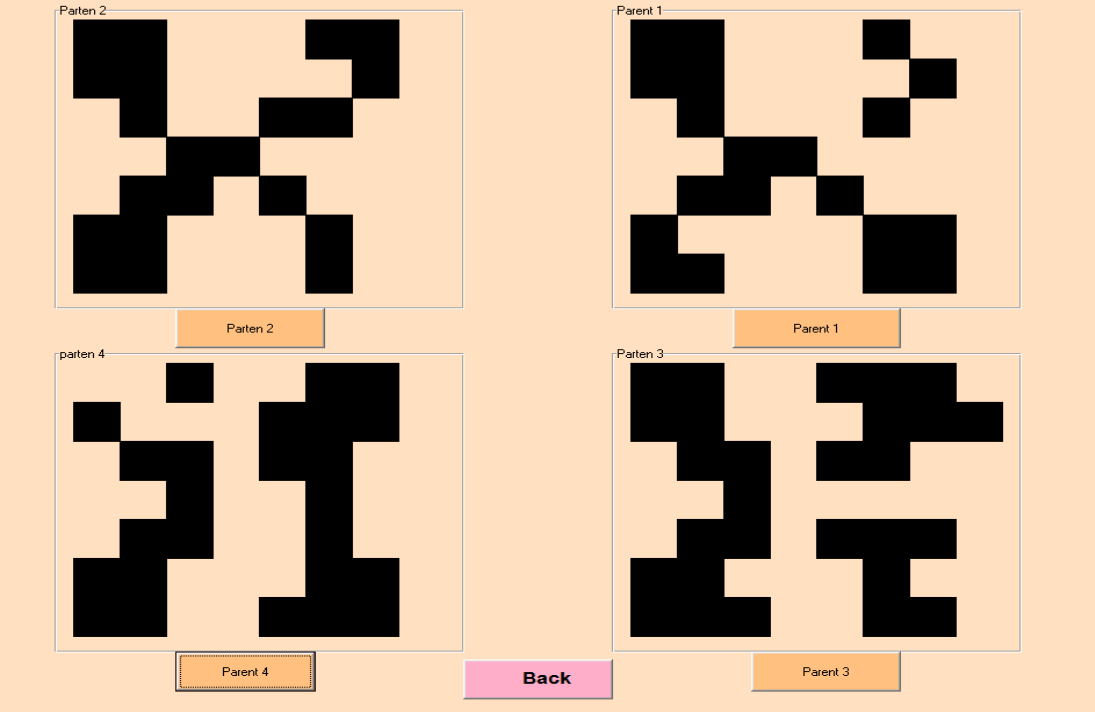


Figure (5) The initial generation pattern for the inputted distorted modes

The applied distorted modes that represent initial generation will be inputted on clonal selection algorithm. It is important to notice that the inputted modes application is done in random selection, while the mutation process is done according to the algorithm. However, the positions of genes and chromosomes are varied in each sequence and according to the clonal selection algorithm steps. After implementing the clonal selection algorithm steps, the infinity value is shown to represent a greater choice (Affinity) for the selected mode, while a pattern represent a sequence of patterns. As for clon_number, it represents the calculated number of copies for patterns from the following equation:

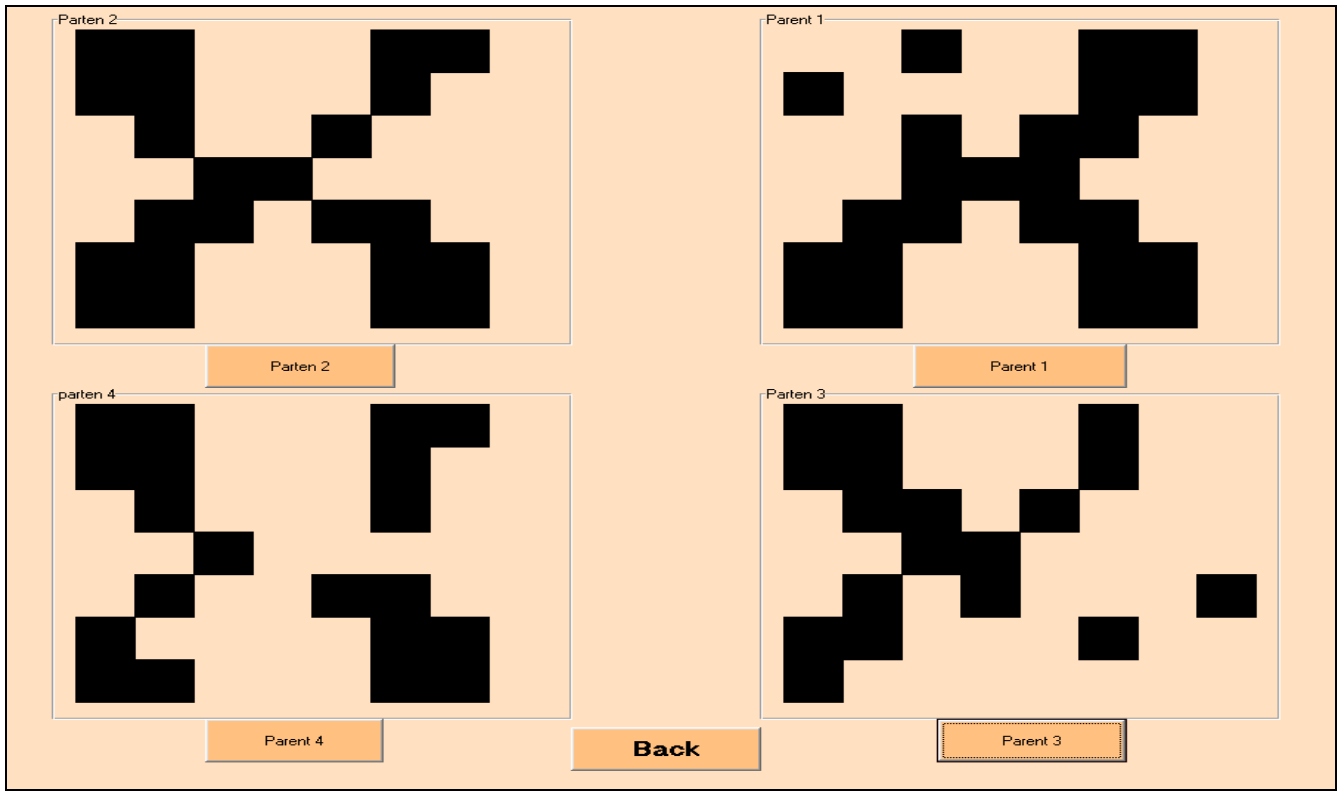
Number of copies =(Affinity value/10) ……. (1)

The mutation_number represents the mutate as a result of the following equation:

Mutate value = (Total bits number- Affinity value) ……. (2)

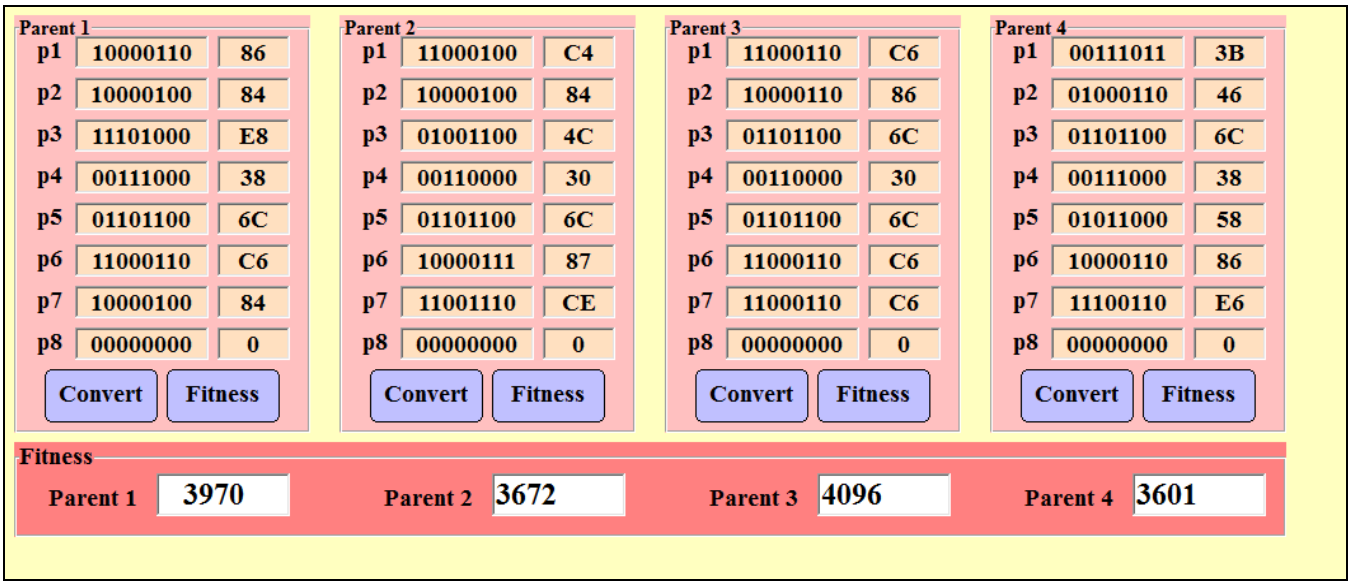
Figure (6) represents the values generated by executing the clonal selection algorithm (first generation). Figure (7) shows the modes pattern generated by executing the clonal selection algorithm (first generation).

Figure (6) The values generated by executing the clonal selection algorithm (first generation)



Figure(7) The pattern of modes generated by the clonal selection algorithm(first generation)

Then the distorted modes are applied which represent initial generation. They are applied on genetic algorithm in which the value of fitness function is observed for each mode. The fitness function is considered as the maximum function. Modes inputting was done in random selection as well as a crossover and mutation process according to algorithm, but positions of genes and chromosomes are varied in each repetition of genetic algorithm steps. Figure (8) represent first generation produced by the genetic algorithm.



Figure(8) The first generation generated after implementing the genetic algorithm.

As for the modes pattern produced by implementing the genetic algorithm, it is shown in figure(9).

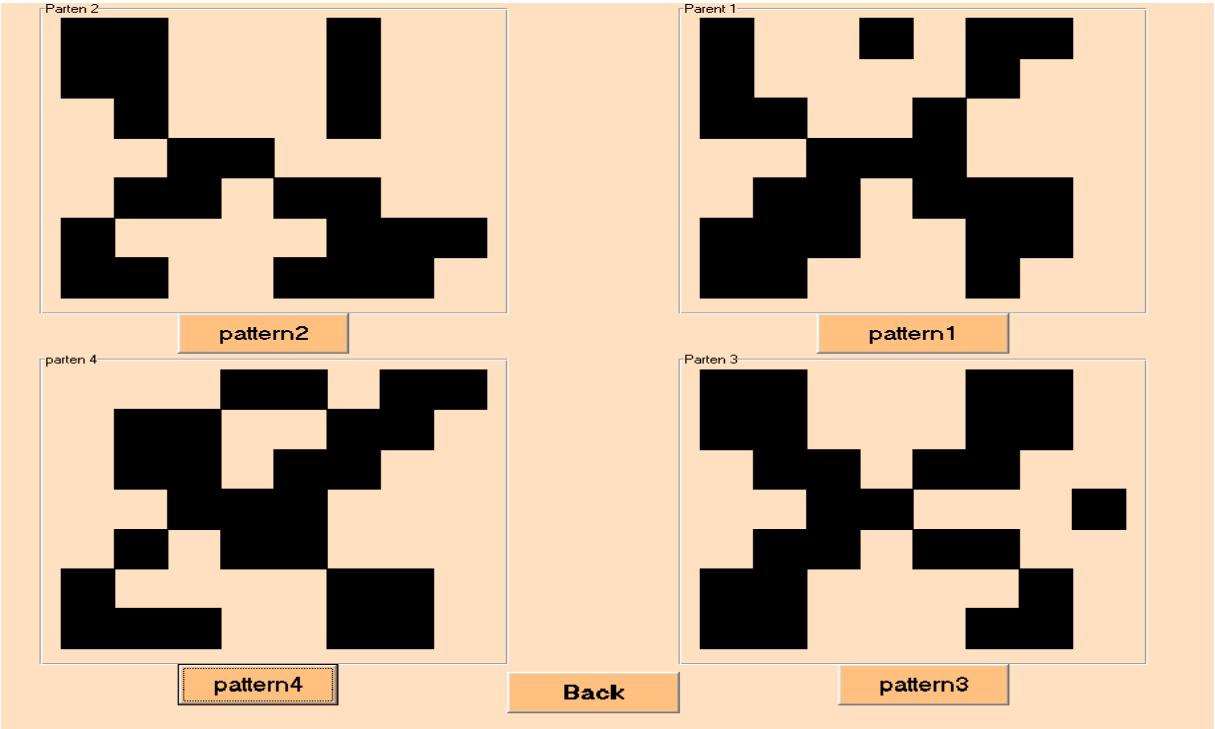


Figure (9) The modes pattern produced by implementing the genetic algorithm (first generation).

As for the second generation , it is shown in figure(10) for which the clonal selection algorithm was used.

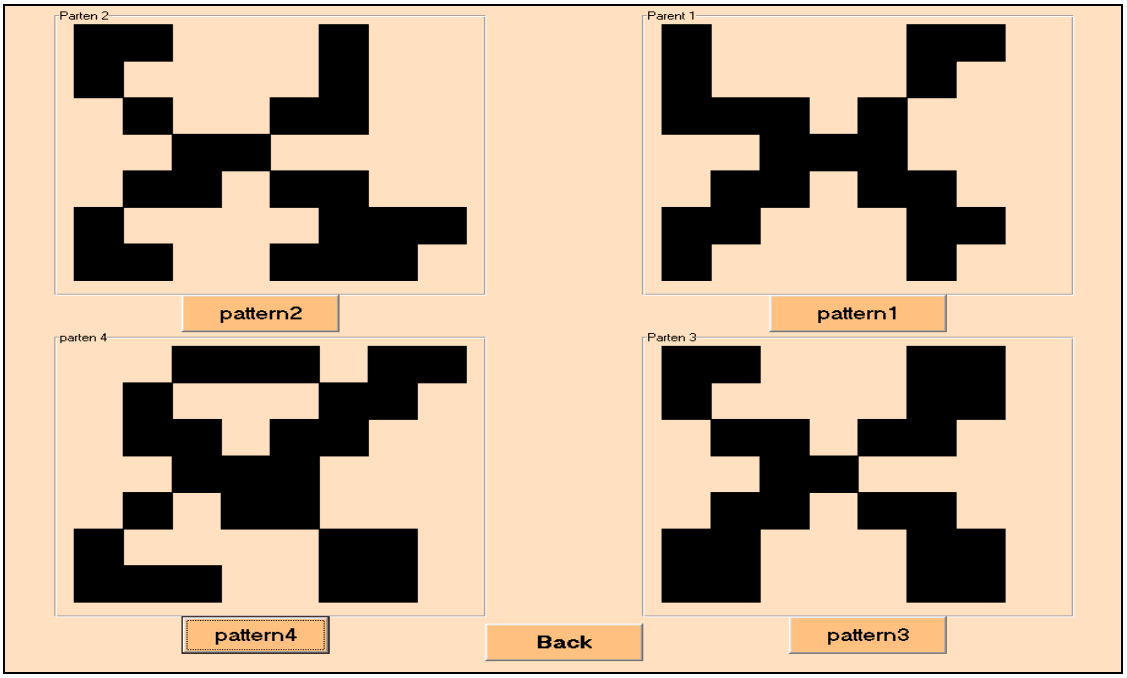
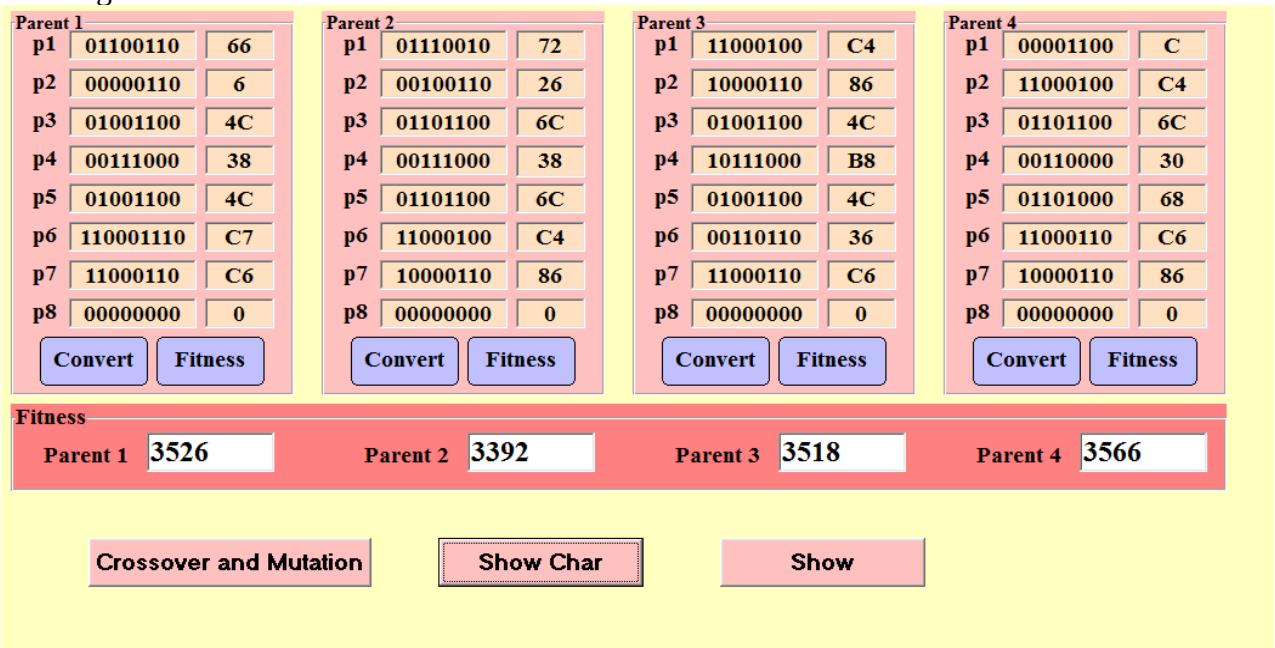
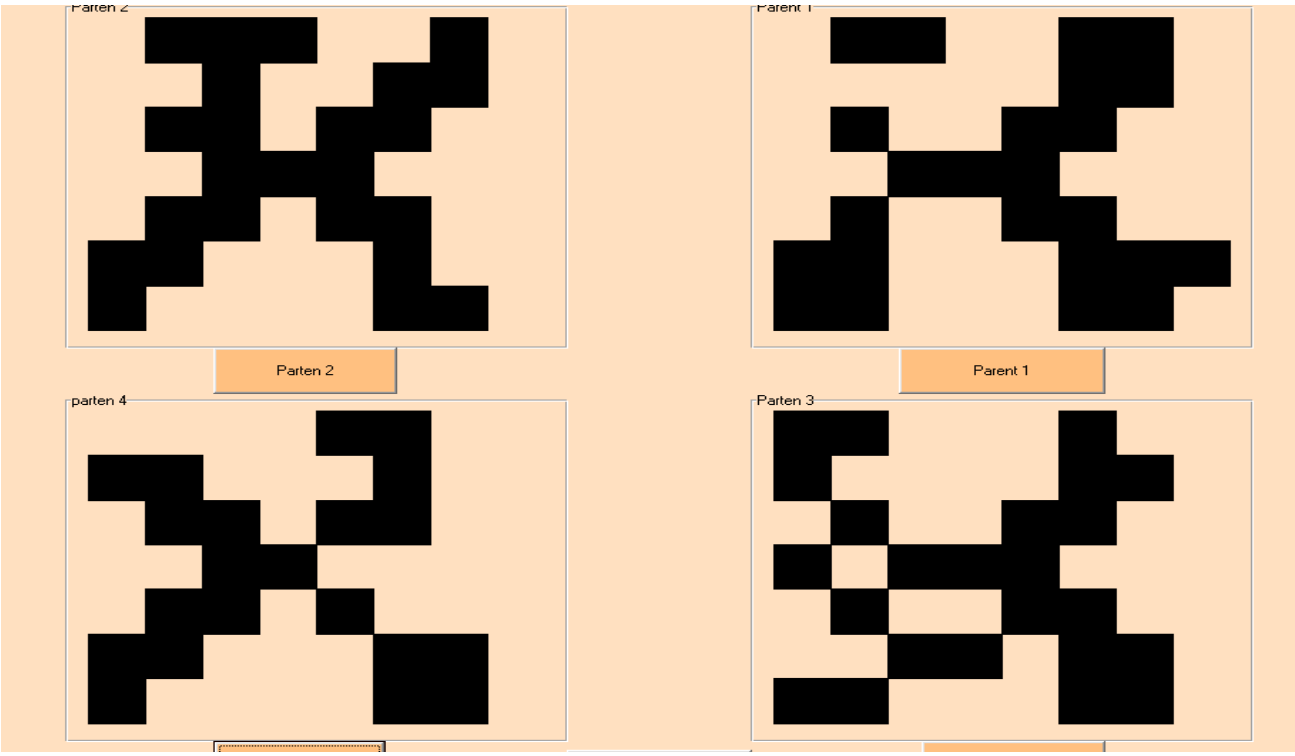


Figure (10) The modes pattern produced by implementing the clonal selection algorithm (second generation).

The following represents the fourth generation results after repeating the implementation of the genetic algorithm.

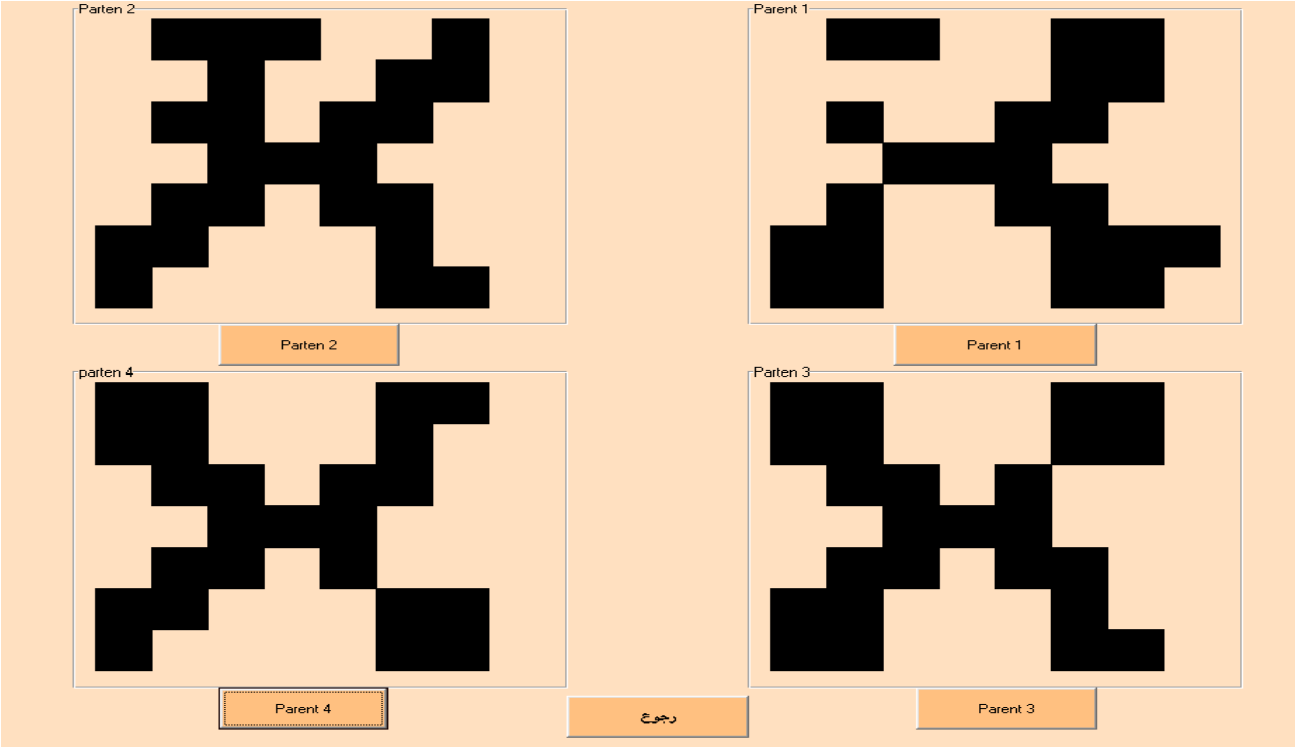


Figure(11) The fourth generation produced after implementing the genetic algorithm.



Figure(12) The modes pattern produced by implementing the genetic algorithm(fourth generation)

We notice some of the modes samples patterns (third generation) after repeating the clonal selection algorithm implementation as follows:



Figure(13) The modes pattern produced by implementing the clonal selection algorithm(third generation)

8. Conclusions

Having the results of the modes pattern produced by implementing the genetic and clonal selection algorithm, the following is concluded:

1. In clonal selection algorithm, there is no crossover for the produced modes that are closer to the original mode, while in genetic algorithm, there is a crossover for the produced modes that are farther form the original mode.

2. In clonal selection algorithm, the third generation is the closest to original mode, while in genetic algorithm; the fourth generation is the farthest from original mode.
3. In the clonal selection algorithm, the stop criterion is defined by obtaining the highest affinity value rate for the distorted modes when compared with the original modes, with the emanation of few number of generations. As for the genetic algorithm, the stop criterion is considered for which the highest fitness value rate can be obtained with few number of generations.
4. The clonal selection algorithm has an objective function, while the genetic algorithm has a fitness function.

5. In clonal selection algorithm, the clone_number represents the calculated patterns number of copies that increases the number of modes close to the original mode, while in genetic algorithm, the crossover and mutation makes the possibilities of distorted modes too many and farther from the original mode.

6. Throughout the results, the clonal selection algorithm achieved 85% rate as the highest matching rate for the input-distorted modes, while the genetic algorithm achieved 52% rate at the highest maching rate for the input-distorted mode regarding the Latin number that is despite of the fact that both algorithms are based on optimization and probability principle.

References

1. David , Beasley, R. Bull, and Ralph R. Martin. University Computing, volume 15(2), pages 58-69, 1993.

2. Carter, J. H. (2000), Journal of the American Medical Informatics Association, 7(3), pp. 28-41.

3. Forrest, S., Javornik, B., Smith, R. E. & Perelson, A. S. (1993), Evolutionary Computation, 1(3), pp. 191-211.

4. D. E. Goldberg (1989), Addison-Wesley Publishing Company, Inc.

5. A. Hadad, C. Eick, (1997), In P. Angeline et al. (eds.) Proc. of the 6th International Conf. on Ev. Programming, vol. 1213 of LNCS. Springer.

6. Christian Tschudin, Thomas Meyer and Lidia Yamamoto , 2009, November 24.

7. J. J. Grefenstette, (1992), In R. Maenner, B. Manderick (eds.), Parallel. Problem Solving from Nature 2, pp. 137-144. North Holland.

8. L. N. de Castro, F. J. Von Zuben. aiNet: "An Artificial Immune Network for Data Analysis". In H. A. Abbass, R. A. Sarker, C. S. Newton (eds) Data Mining: A Heuristic Approach, Chapter XII. Idea Group Publishing, 2001 .

9. L. N de Castro and F. J. Von Zuben, IEEE Transactions on Evolutionary Computation, 6:239–251, June 2002.

10. S. M. Garrett, Evol. Comput., vol. 13, no. 2, pp. 145–177, 2005.

11. Stephanie Forrest, Brenda Javornik, Robert E. Smith, and Alan S. Perelson, 1993, July 28.

12. Steven A. Hofmeyr and Stephanie Forrest, 2000, Dept. of Computer Science, University of New Mexico, Albuquerque, NM 87131-1386.

13. Liangpei Zhang, Yanfei Zhong, Pingxiang LI, 2002.

14. De Castro, L. N. & Von Zuben, F. J. (2000),GECCO'00 – Workshop Proceedings, pp. 36-37.

15. G. G. Rajput, Rajeswari Horakeri and Sidramappa Chandrakant, 2010, International Journal on Computer Science and Engineering Vol. 02, No. 05, 1622-1626.

16. B. Gatos, I. Pratikakis, A.L. Kesidis and S.J. Perantonis,2006, Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, National Center for Scientific Research "Demokritos", GR-153 10 Agia Paraskevi, Athens, Greece.