

An Immuno-Genetic Hybrid Algorithm

E. Nabil, A. Badr, I. Farag

Emad Nabil

Misr University for Science and Technology, Information Technology Faculty, Computer Science Department.
E-mail: emadnabilcs@gmail.com

Amr Badr, Ibrahim Farag

Cairo University, Computers and Information Faculty, Computer Science Department.
5 Dr. Ahmed Zewail Street, Postal Code: 12613, Orman, Giza, Egypt
E-mail: ruaab@rusys.eg.net, i.farag@fci-cu.edu.eg

Abstract: The construction of artificial systems by drawing inspiration from natural systems is not a new idea. The Artificial Neural Network (ANN) and Genetic Algorithms (GAs) are good examples of successful applications of the biological metaphor to the solution of computational problems. The study of artificial immune systems is a relatively new field that tries to exploit the mechanisms of the natural immune system (NIS) in order to develop problem-solving techniques. In this research, we have combined the artificial immune system with the genetic algorithms in one hybrid algorithm. We proposed a modification to the clonal selection algorithm, which is inspired from the clonal selection principle and affinity maturation of the human immune responses, by hybridizing it with the crossover operator, which is imported from GAs to increase the exploration of the search space. We also introduced the adaptability of the mutation rates by applying a degrading function so that the mutation rates decrease with time where the affinity of the population increases, the hybrid algorithm used for evolving a fuzzy rule system to solve the well-known Wisconsin Breast Cancer Diagnosis problem (WBCD). Our evolved system exhibits two important characteristics; first, it attains high classification performance, with the possibility of attributing a confidence measure to the output diagnosis; second, the system has a simple fuzzy rule system; therefore, it is human interpretable. The hybrid algorithm overcomes both the GAs and the AIS, so that it reached the classification ratio 97.36, by only one rule, in the earlier generations than the two other algorithms. The learning and memory acquisition of our algorithm was verified through its application to a binary character recognition problem. The hybrid algorithm overcomes also GAs and AIS and reached the convergence point before them.

Keywords: genetic algorithms, artificial immune system, fuzzy logic, breast cancer diagnosis, memory acquisition.

1 Introduction

Computing and engineering have been enriched by the introduction of the biological ideas to help developing solutions to various problems. This can be exemplified by the Artificial Neural Networks (ANN), Evolutionary Algorithms (EA) [11], Artificial Life (ALife), and cellular automata (CA) [13]. There exist three different approaches; the first is: biologically motivated computing, under this umbrella the EA, ANN and artificial immune system (AIS) [21]; the second is computationally motivated biology, where computing provides models and inspiration for biology (i.e. ALife and CA). The third approach is computing with biological mechanisms, which involves the use of information processing capabilities of biological systems to replace or supplement the current silicon-based computers (e.g. Membrane

computing, Quantum computing and DNA computing) [8], [9] [14], [18]. Our research point will be under the umbrella of the first approach.

In this paper, we combine two methodologies which are Genetic Algorithms and Artificial Immune System (AIS), so as to automatically produce a fuzzy system for breast cancer diagnosis. The major advantage of fuzzy systems is that they favor interpretability [3], [4] and provide what is called confidence measure which means, in our case, the degree of benignity or malignancy. Finding good fuzzy systems is quite a hard task. So, this is where GA and AIS algorithms work, enabling the automatic production of fuzzy systems, based on a database of training cases.

In this paper we also ensure the ability of memory acquisition and learning of the algorithm by applying it to a binary pattern recognition problem.

The paper is organized as follows: in the next two sections we provide an overview of the clonal selection algorithm and the genetic algorithm. In section [4] we present our proposed hybrid algorithm between GA and AIS that will be tested on the Wisconsin Breast Cancer Diagnosis (WBCD) problem described in section [5]. Evolving fuzzy system of the WBCD, parameters setup and testing also included in section [5]. Section [6] speaks about the learning and memory acquisition of the hybrid algorithm. The algorithm testing is delineated also in section [6], followed by concluding remarks in Section [7].

2 The clonal selection algorithm

The standard clonal selection algorithm CLONALG [5], [6], [15], [16], [17] can be summarized as follows.

```

Begin
  t=0;
  Initialize the initial population p(t) randomly;
  Identify antigen S;
  Evaluate affinity p(t) versus S;
  While (not finished) do
    Begin
      t= t+1;
      Select C(t) from p(t-1);
      Proportional cloning of C(t) forming C'(t);
      Mutation C'(t) forming c"(t);
      Select P(t) from c"(t) and P(t-1);
      Select memory cell from P(t);
      Metadymanics;
    End.
End.

```

3 The genetic algorithm (GA)

The standard genetic algorithm [7], [23] can be summarized as follows.

```

Begin
  t=0;
  Initialize the initial population p(t) randomly;
  Evaluate structures in p (t);
  While (not finished) do
    Begin
      t= t+1;
      Select parents C(t) from p(t-1);
      Crossover and mutate structures in C (t) forming C' (t);
    End.
  End.

```

```

    Replace C' (t) by P (t-1);
End.
End

```

4 The proposed hybrid algorithm

$$D = \sum_{i=1}^L \delta \quad \text{where } \begin{cases} 1, & ab_i \neq ag_i \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

The affinities of individuals are measured using the hamming distance depicted in equation 1.

The proposed algorithm modifies clonal selection algorithm mutation method. The mutation in nature occurs at small percentage value = 0.002 and this is rational from the computational point of view to ensure that the good solutions are not distorted too much. However, researches have shown that an initial large mutation rate that decreases exponentially as a function of the generation number improves the convergence speed and accuracy [1].

The initial large mutation rate ensures that a large space is covered, while the mutation rate becomes smaller when the individuals start to converge to the optimum. This is accepted solution for the trade off between the exploration and exploitation We used the time-decaying formula in equation (2) [18], [22], [24] where τ is a positive constant, $m(0)$ is the initial large mutation rate and t is the generation number. The equation is depicted in Figure 1. We have imported the crossover operator from the genetic algorithms in order to increase the exploration of the landscape and to add a recombination operator in the clonal selection algorithm.

$$m(t) = m(0)e^{-t/\tau^2} \quad (2)$$

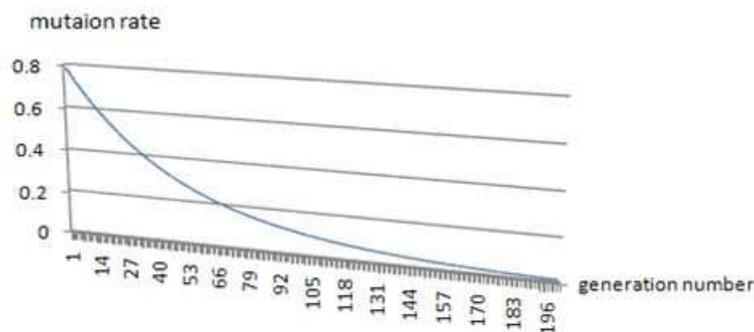


Figure 1: The effect of the degraded function on mutation value

The proposed algorithm can be summarized as follows.

```

Begin
  t=0;
  Initialize the initial population p(t) randomly;
  Identify antigen S;
  Evaluate fitness p (t) versus S;
  While (not finished) do
    Begin
      1. t= t+1;
      2. Select C(t) from p(t-1);
      3. Proportional cloning of C(t) forming C'(t);
    End
  End

```

Table 1: The WBCD data representation

case	V_1	V_2	...	V_9	Diagnosis
1	1	2	...	8	Benign
2	2	4	...	3	Benign
...
683	4	8	...	1	Malignant

4. Degraded Proportional Mutation $C'(t)$ forming $c''(t)$;
5. Crossover $c''(t)$ forming $C^*(t)$;
6. Select $P(t)$ from $c^*(t)$ and $P(t-1)$;
7. Select memory cell from $P(t)$;
8. Metadymanics;

End.

End.

The proposed algorithm will be tested on the famous Wisconsin breast cancer diagnosis problem (Section 5) and simple binary pattern recognition problem (Section 6) to ensure the memory acquisition ability of the algorithm.

5 The Wisconsin breast cancer diagnosis problem

In this section, we present the Wisconsin breast cancer diagnosis problem [3] which is the test case of our proposed algorithm. Breast cancer is the most common cancer among women, excluding skin cancer. The presence of a breast mass is an alert sign of a cancer, but it does not always indicate a malignant one. Fine Needle Aspiration (FNA) is an outpatient procedure that involves using a small-gauge needle to extract fluid directly from a breast mass. FNA procedure over breast masses is a cost-effective, non-traumatic, and mostly non-invasive diagnostic test that obtains information needed to evaluate malignancy. The Wisconsin Breast Cancer Diagnosis (WBCD) database [4] is the result of the efforts made at the university of Wisconsin Hospital for accurately diagnosing breast masses based solely on an FNA test. Nine visually assessed characteristics of an FNA sample considered relevant for diagnosis were identified, and assigned an integer value between 1 and 10. The measured variables are as follows:

1. Clump thickness (V_1);
2. Uniformity of cell size (V_2);
3. Uniformity of cell shape (V_3);
4. Marginal adhesion (V_4);
5. Single epithelial cell size (V_5);
6. Bare nuclei (V_6);
7. Bland chromatin (V_7);
8. Normal nucleoli (V_8);
9. Mitosis (V_9).

The database itself consists of 683 cases. The general form of the database is described in Table 1. There exist some previous systems that achieved high classification ratio, but these systems look like black boxes and with no explanation or interpretation about how the decision was taken. Further, the degree of benignity or malignancy is not provided. These two points are covered in this study besides high performance classification ratio.

5.1 Evolutionary fuzzy modeling

Evolutionary algorithms are used to search large, and often complex, search spaces. They have proven worthwhile on numerous diverse problems and are able to find near-optimal solutions with an adequate performance measure. Fuzzy modeling can be seen as an optimization problem where part or all of the parameters of a fuzzy system constitute the search space.

5.2 Applying evolution to fuzzy modeling

Three of the four types of fuzzy parameters can be used to define targets for evolutionary fuzzy modeling: structural parameters, connective parameters, and operational parameters. Logical parameters are usually predefined by the designer based on experience.

The evolutionary algorithm is used to tune the knowledge contained in the fuzzy system by finding membership function values (μ , d values) and the relevant variables. Evolutionary structure learning is carried out by encoding within the genome an entire fuzzy system. This is known as the Pittsburgh approach.

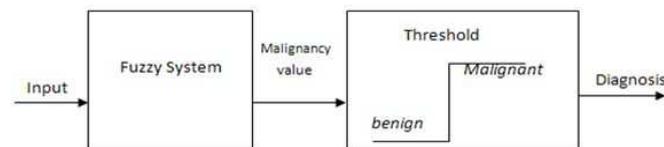


Figure 2: The proposed diagnosis system. Note that the fuzzy subsystem displayed to the left is the fuzzy inference system in Figure 3.

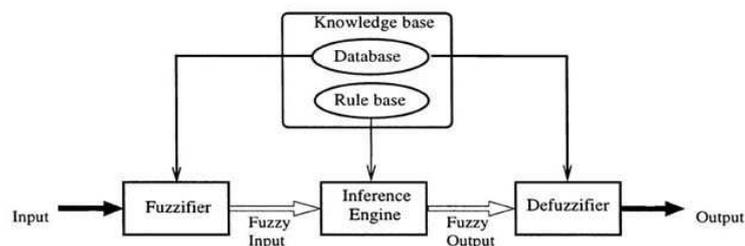


Figure 3: Basic structure of a fuzzy inference system

5.3 Evolving fuzzy systems for the WBCD problem

The solution scheme we propose for the WBCD problem is depicted in Figure 2, Note that the fuzzy subsystem displayed to the left of figure 2 is the fuzzy inference system of Figure 3 [10], [20]. Figure 2 consists of a fuzzy system and a threshold unit. The fuzzy system computes a malignancy value of the

malignancy of a case, based on the input values, the threshold unit then outputs a benign or malignant diagnostic according to the fuzzy system's output. If the malignancy value is less than or equals 3, it is considered a benign case. Other than that, it is diagnosed as a malignant one.

5.4 Fuzzy system parameters

According to information obtained from previous work [3], we have deduced the following points.

- Small number of rules: Systems with no more than four rules have been shown to obtain high performance [2], [19].
- Small number of variables: Rules with no more than 4 antecedents have proven to be adequate[2].
- Nature of the input variables: higher-valued variables are associated with malignancy. Some fuzzy models forgo interpretability in the interest of improved performance. Where medical diagnosis is concerned, interpretability, also called linguistic integrity, is the major advantage of fuzzy systems. This motivated us to take into account the following semantic criteria, defining constraints on the fuzzy parameters [12]:
- Distinguishability: To what extent the system is understood and has interpretability.
- Justifiable number of elements: The number of membership functions of a variable. This number should not exceed the limit of 7 ± 2 distinct terms. The same criterion is applied to the number of variables in the rule antecedent; this is to be familiar for humans.
- Orthogonality. For each element of the universe of discourse, the sum of all its membership values should be equal to one.

5.5 The fuzzy system setup

Logical parameters

- Reasoning mechanism: singleton-type fuzzy system, i.e. Output membership functions are real values, rather than fuzzy ones.
- Fuzzy operators: min.
- Input membership function type: orthogonal, trapezoidal.
- Defuzzification method: weighted average.

Structural parameters

- Relevant variables: there is insufficient a priori knowledge to define them; therefore, this will be one of the algorithm's objectives.
- Number of input membership functions: two membership functions denoted Low and High.
- Number of output membership functions: two singletons are used, corresponding to the benign and malignant diagnostics.
- Number of rules: in our approach, this is a user-configurable parameter. Will there be only one rule? The rule itself is to be found by the genetic algorithm.

Connective parameters

- Antecedents of rules: to be found by the algorithm.
- Consequent of rules: the algorithm finds rules for the benign diagnostic; the malignant diagnostic is an else condition.
- Rule weights: active rules have a weight of value 1 and the else condition has a weight of 0.25.

Operational parameters

- Input membership function values: to be found by the evolutionary algorithm.
- Output membership function values: following the WBCD database, we used a value of 2 for benign and 4 for malignant.

5.6 The evolutionary algorithm setup

We apply Pittsburgh-style structure learning, using our algorithm to search for three parameters. The relevant variables, the input membership function values, and the antecedents of rules. They are constructed as follows:

- Membership function parameters. There are nine variables (V_1-V_9), each with two parameters P and d, defining the start point and the length of the membership function edges, respectively.
- Antecedents. The i^{th} rule has the form: **if** (V_1 is A_1^i) **and...and** (V_9 is A_9^i) **then** (output is benign) where A_j^i represents the membership function applicable to variable V_j . A_j^i can take the values: 1 (Low), 2 (High), or 0 or 3 (Other).
- Relevant variables are searched for implicitly by letting the algorithm choose non-existent membership functions as valid antecedents; in such a case, the respective variable is considered irrelevant.

Table 2: Parameters encoding of a genome, total genome length is $54+18=72$

Parameter	Values	Bits	Quantity	Total bits
P	1-8	3	97	27
d	1-8	37	9	27
A	0-3	2	9	18

The parameters encoding are described in Table 2, which form a single individual's genome. Table 3 shows a sample genome. We used a genetic algorithm with a fixed population size of 200 individuals to evolve the fuzzy inference system, and fitness-proportionate selection. The algorithm terminates when the maximum number of generations is reached.

An example of a genome for a rule system depicted in table 3, The first 18 positions encode the parameters P and d for the nine variables V_1-V_9 . The rest encode the membership function applicable for the nine antecedents of the rule; table 4 is an interpretation of the database and the rule base of the rule system encoded in table 3.

5.7 Testing

The proposed algorithm has been tested on the WSBC problem. The three algorithms have been implemented and have been tested in Wisconsin database. The three algorithms have reached a valid classification ratio equal to 97.36% i.e. 665 valid diagnosis cases from 683 cases. And the results of the

three algorithms were depicted in Figure 4. It is clear that the hybrid algorithms reached the maximum classification ratio in the earlier generations before the GA and the AIS. Also the AIS reached before the GA.

Table 3: Database

p_1	d_1	p_2	d_2	p_3	d_3	p_4	d_4	p_5	d_5	p_6	d_6	p_7	d_7
3	5	4	1	2	8	5	1	7	7	2	5	5	5
p_8	d_8	p_9	d_9	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	
7	2	4	7	1	1	3	3	3	1	3	1	1	

Table 4: Rule Base

Rule	If ((v_1 is low) and (v_2 is low) and (v_6 is low) and (v_8 is low) and (v_9 is low)) then(output is benign)
Default	Else(output is malign)

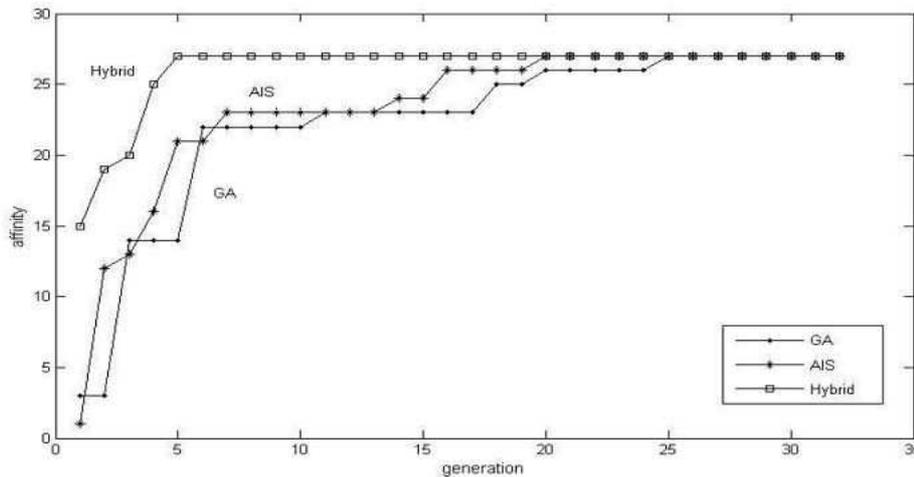


Figure 4: The execution of the three algorithms

6 The Pattern Recognition Problem

The learning and memory acquisition was verified through its application to a binary character recognition problem. In this case, we assumed that the antigen population was represented by a set of ten binary characters ($N = 10$) to be learned. Each character is represented by a bit string of length $L = 121$. The population size =200.

The original characters are depicted in Figure 5. Figure 6 illustrates the initial memory set. Figure 7 illustrates the input patterns. (i.e. the antigens) for which the learning will take place. Figures 8,9 and 10 represent the maturation of the memory set through 200 generations. The affinity here refers the degree of matching of the antigens, i.e. the affinity measure is the hamming distance (discussed in Section 4) between the antigens and antibodies, the. Note that the exact matching is not important for recognition; partial matching is enough. The hybrid algorithm converged at generation 200.

Figure 8, 9 and 10 presents the application of the GA, AIS and the hybrid algorithm to the binary character recognition problem respectively where (a) presents Memory set after 50 cell generations; (b)



Figure 5: The original digits



Figure 6: The initial input patterns



Figure 7: The input patterns (i.e. the antigens) for which the learning will take place



Figure 8: Application of the GA to the binary character recognition problem



Figure 9: Application of the AIS to the binary character recognition problem



Figure 10: Application of the Hybrid Algorithm to the binary character recognition problem.

presents Memory set after 100 cell generations; (c) presents Memory set after 150 cell generations and (d) presents Memory set after 200 cell generations. It is clear that the hybrid algorithm overcomes the other two algorithms. We have used static mutation values, also proportional to affinity, in the binary recognition problem because results showed that static mutation is better than dynamic one. Figure 11 represents the affinity of GA, AIS and the Hybrid algorithms in recognition of the pattern zero, note that the affinity refers the degree of matching with antigens. I.e. the hamming distance with antigens.

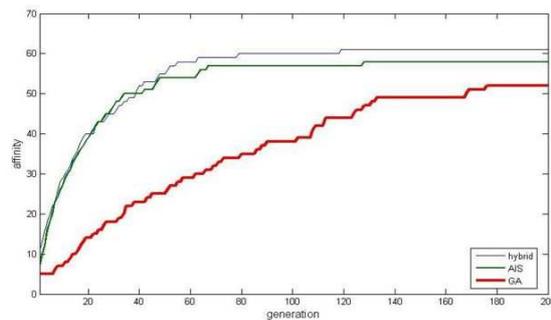


Figure 11: the affinity of GA, AIS and the Hybrid algorithms in recognition of the pattern zero

7 Conclusions

In this research, artificial immune system is combined with genetic algorithms in one hybrid algorithm. A modification is proposed to the clonal selection algorithm which is inspired from the clonal selection principle and affinity maturation of the human immune responses. The adaptability of the mutation rate is introduced by simple degrading function. Also, the crossover is merged into the clonal selection algorithm, two-point crossover applied after the mutation process, to increase the exploration of the landscape. The hybrid algorithm is combined with fuzzy logic and applied to the well-known Wisconsin breast cancer diagnosis problem. We claim that our evolved system exhibits two important characteristics; first, it attains high classification performance, with the possibility of attributing a confidence measure to the output diagnosis; second, the system has a simple fuzzy rule system; therefore, it is interpretable. The hybrid algorithm overcomes both the genetic algorithm and the artificial immune system and reached the highest classification ratio 97.36, by only one rule, in the earlier generations than the two other algorithms.

The proposed system also applied to a binary character recognition problem. The mutation in the hybrid algorithm was adapted using a degraded function so that the mutation decreases with time, but in

the binary character recognition problem, the results showed that it is better to keep the mutation value small and static through all generations. The hybrid algorithm overcomes the other two algorithms.

The hybrid system can solve the WBCD problem with more than one fuzzy rule. This is to increase the classification accuracy. Also, there are many gene representation techniques that can be used instead of the Pittsburgh approach like the Michigan approach, the iterative rule learning approach and hybridization between them. This can be considered as a future work. We claim that our hybrid algorithm is highly effective and better than GAs and AIS, for sure not in all cases, future experiments may prove that GAs or AIS separately is better, but at least in memory acquisition, the WCDB and similar problems we claim that our algorithm is better.

Bibliography

- [1] A.P. Engelbrecht, *Computational Intelligence: an Introduction*, England, John Wiley & Sons; 2003.
- [2] C.A. Pena Reyes, M.A. Sipper, Evolving fuzzy rules for breast cancer diagnosis, *Proc Nonlinear Theory and Applications*, 2, pp 369-372, 1998.
- [3] C.A. Pena Reyes, M.A. Sipper, fuzzy-genetic approach to breast cancer diagnosis, *Artificial Intelligence in Medicine*; vol: 17, num:2, 131-155, 1999.
- [4] C.J. Merz, P.M. Murphy, UCI repository of machine learning database, <http://www.ics.uci.edu/~Mlearn/MLRepository.html>, 1996.
- [5] D. Dasgupta, *Artificial Immune systems and their applications*, Springer-Verlag, inc., 1999.
- [6] D. Dasgupta, N. Attoh-Okine, Immunity-Based Systems, *IEEE International Conference on Systems, Man, and Cybernetics*, Orlando, Florida, pp 363-374, October 12-15, 1997.
- [7] D.A. Coley, An introduction to genetic algorithms for scientists and engineers, *World Scientific Publishing Co., inc.*, 2001.
- [8] E. Gutuleac, Descriptive Timed Membrane Petri Nets for Modelling of Parallel Computing, *International Journal of Computers, Communications & Control*, Vol. I, No. S: Suppl. issue, pp. 256-261, 2006.
- [9] G. Ciobanu, A Programming Perspective of the Membrane Systems, *International Journal of Computers, Communications & Control*, Vol. I, No. S: Suppl. issue, pp.13-22, 2006.
- [10] H. Zhang, D. Liu, *Fuzzy Modeling and Fuzzy Control*, Birkhauser, 2006.
- [11] J. Rennard, Genetic Algorithm Viewer: Demonstration of a Genetic Algorithm, <http://www.rennard.org/alife/english/gavgb.pdf>, 2000.
- [12] J.J. Espinosa, J. Vandewalle, Constructing fuzzy models with linguistic Integrity, *IEEE Transactions on Fuzzy Systems*; vol. 7, no. 4, pp. 377-393, 1999.
- [13] L.N. De Castro, *Fundamentals of natural computing: basic concepts, algorithms, and applications*, CRC Press LLC; 2007.
- [14] L.N. De Castro, F.J. Zuben, Artificial Immune Systems: Part I – Basic Theory and Applications, EEC/Unicamp, Campinas, SP, Tech. Rep. – RT DCA 01/99, p. 95. 1999.
- [15] L.N. De Castro, F.J. Zuben, Learning and optimization using the clonal selection principle, *IEEE transactions on evolutionary computation*, vol.:6, num.:3, pp 239-251, Jun, 2002.

- [16] L.N. De Castro, F.J. Zuben, The Clonal Selection Algorithm with Engineering Applications, Artificial Immune System Workshop, *Genetic and Evolutionary Computation Conference*, A. S. Wu (Ed.), pp. 36-37, 2000.
- [17] L.N. De Castro, J. Timmis, *Artificial Immune Systems (A new computational Approach)*, Springer - Verlag, 2002.
- [18] L.N. De Castro, Natural computing, *Information science and technology*, Idea Group, Inc., 2005.
- [19] R. Setiono, Extracting rules from pruned neural networks for breast cancer diagnosis, *Artificial Intelligence in Medicine*, vol. 8, no. 1, pp. 37-51, Feb. 1996.
- [20] R.R. Yager, L.A. Zadeh, *Fuzzy Sets, Neural Networks, and Soft Computing*, New York, Van Nostrand Reinhold, 1994.
- [21] S. Forrest, S.A. Hofmeyr, A. Somayjit, Architecture for an Artificial Immune System, *Evolutionary Computing*, vol. 8, no. 4, pp 443-473, 2000.
- [22] T. Back, D. Fogel, Z. Michalewicz, Glossary, *Evolutionary Computation 1: Basic Algorithms and Operators*, Institute of Physics Publishing, Bristol and Philadelphia, 2000.
- [23] T. Back, The Interaction of Mutation Rate, Selection & Self-Adaptation within a Genetic Algorithm, In Proc. *2nd Int. Conf. on Parallel Problem Solving From Nature*, North-Holland, Amsterdam, pp. 85-94, 1992.
- [24] W. M. Spears, Adapting Crossover in Genetic Algorithms, Artificial Intelligence Center Internal Report AIC-94-019, Naval Research Laboratory, Washington, DC 20375, 1994.

Emad Nabil Received his BSc and MSc degrees in computer science from computers and information faculty, Cairo University in 2004 and 2008 respectively; His main research interests are in the natural computing area including P systems, GA, AIS and ANN. Now i am working in my Ph.D. in P systems and their applications in optimization.

Amr Badr Currently he is an Associate Professor of Computer Science, computers and information faculty, Cairo university. His research interests include Computational Intelligence, Petri Nets, Bioinformatics and Medical Imaging. He has published about 50 Journal research papers in these areas and supervised nearly 50 MSc and PhD students. He is currently on the Editorial Boards and a reviewer at several Journals.

Ibrahim Farag Professor of Computer Science, computers and information faculty, Cairo university. He is the founder of the Faculty of Computers and Information at Cairo University and one of the pioneers of Computer Science in Egypt. He has supervised more than 200 MSc and PhD students at Cairo University.