Proceedings of the 2007 IEEE Symposium on
Foundations of Computational Intelligence (FOCI 2007)

Foundations of Immunocomputing

Alexander Tarakanov
St. Petersburg Institute for Informatics and Automation
Russian Academy of Sciences
14th line 39, St. Petersburg, 199178, Russia
Tel: +7 812 328 47 45, Fax: +7 812 328 44 50
E-mail: tar@iias.spb.su

Giuseppe Nicosia
Department of Mathematics and Computer Science
University of Catania
Viale A. Doria 6, 95125 Catania, Italy
Tel: + 39 095 738 30 30, Fax: + 39 095 33 00 94
E-mail: nicosia@dmi.unict.it

Abstract— This paper presents the mathematical basis of the *immunocomputing* using feature extraction and pattern recognition. The key notions of the approach are the *formal immune network* (FIN) and the coding theory for *machine learning*. The training of FIN includes *apoptosis* (programmed cell death) and *autoimmunization* both controlled by cytokines (messenger proteins), whereas parameters of FIN can be optimized by *Kullback entropy*. Recent results suggest that the approach outperforms (by training time and accuracy) state-of-art approaches of computational intelligence.

I. Introduction

Two types of biological systems, the neural system and the vertebrate immune system, possess the capabilities of "computational intelligence", which include memory, the ability to learn, to recognize, and to make decisions with respect to unknown situations, and dynamical and noisy environments. In practice, an immune system protects the organism against pathogen threats and autoimmune diseases. From a computational point of view, the natural immune system constitutes a robust integrated defense system.

The potential of the natural neural system as a biological prototype of a computing scheme has already been well established in the field of Artificial Neural Networks (ANN), or *neural computing* (Cloete & Zurada, 2000). However, the computing capabilities of the natural immune system (Jerne, 1974; Farmer, Packard & Perelson, 1986; De Boer, Segel, & Perelson, 1992) have only recently been appreciated as a field of Artificial Immune Systems (AIS) (Dasgupta, 1999; Cutello & Nicosia 2002; De Castro & Timmis, 2002; Nicosia 2004). The mathematical formalization of these capabilities (Tarakanov & Dasgupta, 2000) forms the basis of ImmunoComputing (IC) as a new computing approach that replicates the principles of information processing by proteins and immune networks (Tarakanov, Skormin, & Sokolova, 2003).

This IC approach looks rather constructive as a basis for a new kind of computational intelligence. A number of successful applications of IC to real world tasks have been reported, including detection of dangerous ballistic situations in near Earth space (Tarakanov & Dasgupta, 2002), computing of ecological map and optical response of laser diode (Tarakanov A. & Tarakanov Y., 2004, 2005), and reconstruction of hydroacoustic fields (Tarakanov, Prokaev, & Varnavskikh, 2006). It is also worth noting that a connection between IC and cellular automata has led to encouraging results in three-dimensional computer graphics (Tarakanov & Adamatzky, 2002). As for biological applications of IC, a concept of "biomolecular immunocomputer" as a computer controlled fragment of the natural immune system has recently been proposed (Goncharova, Melnikov, & Tarakanov, 2003; Goncharova, Jacques, Martin-Vide, Tarakanov, & Timmis, 2005). A connection of IC with brain research has also shed light upon principles of organization of receptor mosaics and molecular networks (Agnati, Tarakanov, Ferre, Fuxe, & Guidolin, 2005). In such background, this paper presents the foundations of the IC approach to intelligent signal processing.

II. The Immunocomputing Approach

In general, the IC approach to signal processing consists of two usual parts: feature extraction and pattern recognition.

Feature extraction is inspired by a mode of biomolecular "computing" where immune cells chop unknown antigen to its local singularities and expose them to the immune system. Analogously, the IC approach represents unknown signal as a tree of data, and chops the branches of the tree to detect local singularities of the signal. The approach is based on the rigorous mathematical methods of Discrete Tree Transform (DTT) (Atreas, Karanikas, & Tarakanov, 2003; Karanikas & Proios, 2003) and Singular Value Decomposition (SVD) (Horn & Johnson, 1986).

Pattern recognition is used because a feature vector (pattern) has to be extracted from a raw signal by the above part. The IC approach to pattern recognition is based on a notion of cytokine FIN (Formal Immune Network) proposed in (Tarakanov [et al.] 2005). Cytokines (messenger proteins) are a group of biologically active mediator molecules that provide the intercellular interactions within the immune system. They are the central regulators of leukocyte growth and differentiation, being produced by a wide variety of cell

types, targeting various cell subsets and exhibiting numerous biological activities. Up to now more than *100* different human cytokines are identified. Apoptosis is a natural mechanism by which cells *commit suicide* when they have outlived their purpose, become defective, or have aged. A mathematical model of apoptosis and autoimmunization (to correct mistakes of apoptosis) has also been proposed in the frame of FIN (Tarakanov [et al.] 2005).

However, the IC approach contains a set of parameters to be adjusted depending upon the properties of the signal type (e.g., audio, video, radar, sonar, etc.). This theoretical gap can be filled using the Kullback entropy (Kullback, 1959), as has been proposed for AIS and evolutionary algorithms (Cutello, Nicosia, & Pavone, 2003, 2007; Nicosia 2004). In the IC approach, the entropic functions measure the quantity of information that FIN discovers during the training and testing processes.

III. Formal Immune Network

Henceforth, vector/matrix transposing will be designated by upper stroke $[]'$. For example, if X is a column vector then X' is a row vector.

Definition 1. A *Cell* is a pair $V = (f, P)$, where f is real value $f \in R$, whereas $P = (p_1,..., p_q)$ is a point of a q-dimensional space: $P \in R^q$, and P lies within the unit cube: $\max\{|p_1|,...,|p_q|\} \le 1$.

Let the distance, *affinity measure,* $d_{ij} = d(V_i, V_j)$ between cells V_i and V_j be defined by a norm:

$$d_{ij} = \|P_i - P_j\|.$$

For example, it can be Euclidian $\|P\|_E$, Manhattan $\|P\|_M$, Tchebyshev $\|P\|_T$, or any appropriate norm.

Fix a some finite non-empty set of cells, *innate immunity*, $W_0 = (V_1,...,V_m)$ with non-zero distance between cells: $d_{ij} \neq 0$, $\forall i, j : i \neq j$.

Definition 2. FIN is a set of cells: $W \subseteq W_0$.

Definition 3. Cell V_i recognizes cell V_k if the following conditions are satisfied:

$|f_i - f_k| < -$, $d_{ik} < h$, $d_{ik} < d_{ij}$, $\forall V_j \, \forall W$, $j \neq i,k$,

where $\rho \, \rho \, 0$ and $h = 0$ are non-negative real values (the *recognition threshold* and the *affinity threshold*).

Let the behavior of FIN be defined by the following two rules.

Rule 1 (apoptosis). If a given cell $V_i \subseteq W$ recognizes cell $V_k \in W$ then remove V_i from FIN.

Rule 2 (autoimmunization). If cell $V_k \in W$ is nearest to cell $V_i \in W_0 \setminus W$ among all cells of FIN: $d_{ik} < d_{ij}$, $\forall V_j \, \forall W$, whereas $|f_i - f_k| - -$, then add V_i to FIN.

Let W_A be FIN as a consequence of the application of apoptosis to all cells of W_0. Let W_I be FIN as a consequence of the autoimmunization of all cells of W_A by all cells of W_0.

$$W_A = Apoptosis(W_0);$$
$$W_I = Autoimmunization(W_A, W_0);$$

Note that the resulting sets W_A and W_I depend on the ordering of the cells in W_0. In the rest of the paper it will be assumed that the ordering is given.

Considering the general mathematical properties of FIN, it is obvious that neither the result of apoptosis W_A nor the result of autoimmunization W_I can overcome W_0 for any innate immunity and thresholds:

$$W_A \subseteq W_0, \quad W_I \subseteq W_0, \quad \subseteq W_0, h, \subseteq.$$

The following Propositions give more important and less evident properties of FIN. Their proofs can be found in (Tarakanov [et al.] 2005).

Proposition 1. For any innate immunity W_0 and *recognition threshold* ρ there exists an *affinity threshold* h_0 such that apoptosis does not change W_0 for any h less than h_0: $W_A = W_0, \forall h < h_0$.

Proposition 2. For any innate immunity W_0 and *recognition threshold* ρ there exists *affinity threshold* h_1 such that the consequence of apoptosis and autoimmunization $W_1 = W_I(h_1)$ provides the minimal number of cells $|W_1|$ for given W_0 and ρ and any h: $|W_1| = |W_I(h)|$, $\forall h$, $\forall W_I \, \forall W_0$.

Actually, Proposition 2 states that the minimal number of cells after apoptosis and autoimmunization is a kind of *inner invariant* of any FIN, which depends on the innate immunity and the recognition threshold but does not depend on the affinity threshold. Practically, it means that such invariant can be found for any FIN by apoptosis and autoimmunization without considering any affinity threshold (in Definition 3) at all.

It is worth noting that Rule 1 and Rule 2 should be applied *consequently*, but not simultaneously. In their own turn, Rule 1 and Rule 2 should be applied consequently to all members of FIN. These are defined by the software implementation of FIN in standard (serial) PCs. However, such sequential FIN could be revisited for the implementation in digital signal processors which allow any multi-sequencing (Tarakanov, Kvachev & Sukhorukov 2005).

IV. Coding and Information Theory for Learning

In coding theory, one models the information being passed from a sender to a receiver through a *channel*. The channel may introduce noise, distorting the value of some of the bits (data, in general) during the transmission. The channel can be

a wireless or wired connection, a storage system or a biological system. In this section we describe the immune system as a *noisy channel*. So, the *signal S* is the population of lymphocytes B, the *channel* is the global IS, the *noise source N* is the antigen (Ag), and the *received signal E* is the antibody (Ab). Hence, in this simple model of the natural immune system we consider three immunological entities, B cells, Ag and Ab, defined as binary strings. We consider the case when the signal is perturbed by noise during transmission, and does not always undergo the same change in transmission. In this case, one may assume that *E=f(S,N)*. In a noisy channel there are two parallel, conflicting and sometime non-commensurable processes at work: the source and the noise, in this article, the population of lymphocytes and the pathogen organisms. Coding theory for machine learning studies the trade-offs (e.g., the observed Pareto Fronts) between three conflicting objectives: the *source-signal*, the *noise-problem*, and the *expected-received-signal*. Actually, in the natural immune system there are several parallel processes at different temporal and spatial scales; in this paper, we will consider only two parallel processes: the ones that generate the B cell and the antibody populations.

If the input of the channel, X^t, at generation t is the B cell population then the entropy, $H(X^t)$, and the entropy of the received signal, $H(Y^t)$, can be defined as

$$H(X^t) = \sum_{m=0}^{\ell} f_m^t \log f_m^t$$

$$H(Y^t) = \sum_{m=0}^{\ell} g_m^t \log g_m^t$$

where the B cells distribution function, f_m^t, is defined as the number of B cells at time *t* with distance m from the antigen, B_m^t, over the total number of B cells:

$$f_m^t = \frac{B_m^t}{\sum_{m=0}^{\ell} B_m^t}$$

and

$$g_m^t = \frac{Ab_m^t}{\sum_{m=0}^{\ell} Ab_m^t}$$

where Ab_m^t is the number of antibodies that at time *t* have fitness function value *m*. We note that in the *noiseless case*

$$H(X^{\tilde{}}) \approx H(X^{\tilde{}})$$

where τ is the temporal window with a fixed number of generations. Moreover, there are two conditional entropies $H(X|Y)$ and $H(Y|X)$, the entropy of the received signal when the input known and conversely. Finally, among these entropies we have the relation:

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$$

where $H(X,Y)$ is the joint entropy, while the *mutual information* (or the rate of *actual transmission R*) is defined as follows:

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) =$$
$$= H(X) + H(Y) - H(X,Y)$$

We note that if *X* and *Y* are independent then *I(X;Y)=0*, while if *X* and *Y* are not perfectly independent then:

$$I(X;Y) = H(X) - H(Y)$$

Hence, H(X,Y) measures the dependency between the two variables. Since, X^t and Y^t represent the population of B cells and antibodies at generation t respectively, we assume they are dependent. The termination condition of an Artificial Immune System modeled as a noisy channel is straightforward: if for a certain number of generations, τ, one have that

$$H(X^{\tilde{}}) \approx H(Y^{\tilde{}})$$

then the received signal is the same of the input signal hence the *noise-antigen* has been *corrected-recognized*. We fail in the case of transmission in a *noiseless channel*. We note that the noise-antigen is corrected-recognized by a B cell (or antibody) if the B cell receptor represents the optimal (or suboptimal) solution for the given Ag-problem-instance.

In this framework, the Shannon Entropy is used to measure the flatness of the *information distribution* provided by a set of solution points, hence it is desirable to maximize it. We do not consider the entropy *H* as a measure of the uncertainty in a stochastic event with a given probability distribution. Now we have a metric for the goodness of the spread of points in a *k*-dimensional objective space.

The Kullback entropy is perhaps the most frequently used information-theoretic *distance* measure from a theoretical point of view. The Kullback entropy *K* is a functional of two PDFs p(x) and q(x),

$$K = \int dx p(x) \log[p(x)/q(x)]$$

The discrete form of the previous equation is

$$K = \sum_m p(x_m) \log[p(x_m)/q(x_m)]$$

To analyze the learning processes, we use the notion of *information gain*, an entropic function associated to the quantity of information the algorithm discovers during the learning phase. It follows that the Information Gain can be defined as:

$$I_G(t,t_0) = \sum_{m=0}^{\ell} f_m^t \log[f_m^t / f_m^{t_0}]$$

$K(t,t_0)$ is the quantity of information the algorithm discovers during the convergence process. The gain is the amount of information the system has already learned from the given optimization problem, with respect to the initial (time step t_0) distribution function. Once the learning process starts, the information gain (I_G) increases monotonically until it reaches a final steady state. Moreover, in general the information gain is a kind of entropy function useful both *on-line* and at *run-time* to understand the algorithms' behavior and to set the parameters of the procedures. In fact, in some application domains it is important to compute the total amount of information between two *adjacent time steps*:

$$I_G(t,t\sum\sum) = \sum\sum_{m=0}^{\ell} f_m^t \log[f_m^t / f_m^{t\Sigma\Sigma}]$$

V. General Model

Consider the mathematical model of signal processing by the IC approach. Step 1 corresponds to *feature extraction*. Steps 2-5 and Step 9 correspond to *training*. Steps 6-8 correspond to *recognition*.

1. Let $T = \{t_1,...,t_u\}$ be a fragment of the real-valued signal, where $u = 2^{N_T}$ and N_T is some number exponent so that u is a power of 2. Compute DTT of this fragment as the values of feature vector $X = [x_1,...,x_n]'$, where $n = 2^{N_X}$, $N_X \le N_T$ so that N_X is a level of the discrete tree.

2. Form a training matrix $A = [X_1...X_m]'$ of dimension $m \times n$, where $X_1,...,X_m$ are training vectors (obtained by DTT of signal parts $T_1,...,T_m$) with known values $f_1,...,f_m$ of some training function $f(X)$.

3. Compute the first q singular values $s_1,...,s_q$ and the corresponding left and right singular vectors $L_1,...,L_q$ and $R_1,...,R_q$ by SVD of the training matrix, where $q \le r$ and r is rank of the matrix. Such SVD can be computed by a rather simple and robust iterative scheme (Tarakanov [et al.] 2003).

4. For any training vector X_i, compute its mapping $Y(X_i) = [y_{i1}...y_{iq}]'$ into a q-dimensional space of FIN:

$$y_{i1} = \frac{1}{s_1} X_i' R_1, \ldots, \; y_{iq} = \frac{1}{s_q} X_i' R_q .$$

5. Using procedures of apoptosis and autoimmunization, reduce m training points of FIN to $k - m$ points, where the points number k is *self-defined* by the inner invariant of FIN (see Proposition 2 in previous section).

6. For any n-dimensional test-vector Z (antigen obtained by DTT of any signal fragment), compute its mapping $Y(Z) = [y_1...y_q]'$ into the q-dimensional space of FIN:

$$y_1 = \frac{1}{s_1} Z' R_1, \ldots, \; y_q = \frac{1}{s_q} Z' R_q .$$

7. Among the reduced training points of FIN $Y_1,\ldots,Y_k$, determine the p nearest points to $Y(Z)$, $Y_1,...,Y_p$, and their distances:

$d_1 = \|Y_1 = Y\|, \ldots, \; d_p = \|Y_p = Y\|$.

8. Interpolate $f(Z)$ by the following sum:

$$f = \sum_{i=1}^{p} a_i f_i ,$$

where $f_i = f(Y_i)$ are the training values in the nearest points of FIN, whereas the coefficients a_i are determined by the distances:

$$a_i = \frac{1}{1 + d_i \not\equiv \sum_{j \ne i}^{p} \frac{1}{d_j}} .$$

9. Use the Information Gain (Nicosia 2004) to adjust the parameters u, n, q, p (dimensions of signal parts, feature vector, space of FIN, and number of nearest points of FIN). Such parameter tuning can be performed using a test set of signal fragments (which FIN has not been trained) but with known values of function f. As future work we plan to use cross-fold evaluation techniques.

Note the following useful features of FIN. It can be shown that $\sum_{i=1}^{p} a_i = 1$. It can be also shown that $f = f_i$ if $d_i = 0$ for any i (then $d_j \ne 0$ for any $j \ne i$).

Consider a special case when the input antigen represents a row of the training matrix and, thus, it is equal to a training vector: $Z = X_i$, $i = 1,\ldots,m$. According to the SVD properties (Steps 2-4), the projection of such antigen into the space of FIN coincides with the corresponding training point of FIN: $Y(Z) = Y(X_i)$. In such a case, Step 5 calculates the following distances of the nearest points of FIN: $d_1 = 0$, $d_2 \ne 0, \ldots, d_p \ne 0$. Then, according to Step 8: $a_1 = 1$, $a_2 = 0, \ldots a_p = 0$, and the output of FIN is equal to the value of the function $f(X_i)$ for corresponding training vector X_i: $f = f_i$.

VI. Comparative Performance of the Approach

According to (Tarakanov A. & Y., 2004, 2005), consider the following example. Let the task be to predict the structure of a laser diode which would provide us with the output signal of maximum optical power. The test data for the laser diode structure with 5 parameters is given in Table 1. This data was obtained using the computational physics method. The experimental results for such structures do not yet exist. However, the computational results for a simpler structure, which has only 1 internal barrier, correspond well to the experimental structures.

The optical power depends on the properties of the internal structure (barriers) in the laser diode. The input data composes these barriers, as well as being the emitters, namely the percentage of aluminium in the ternary solution AlGaAs, which defines the energy offset of these barriers. Therefore, the indicators are defined as follows:

- x_1 and x_5 are aluminium percentages in emitters of electrons and holes, respectively;
- x_2, x_3, and x_4 are aluminium percentages in 1st, 2nd and 3rd internal barriers respectively.

The class (index) number corresponds to the output optical power in a response to the nanosecond current pulse with amplitude of 3.2 Ampere (A) as follows: class 1: 0-2 Watt (Wt); class 2: 2-3 Wt; class 3: 3-4 Wt; class 4: 4-5 Wt; class 5: 5-6 Wt; class 6: 6-7 Wt; class 7: more than 7 Wt.

We used the first 15 structures in Table 1 as the training patterns (the class is marked in bold) and all 19 structures as the test patterns.

Table 1. Optical power of different laser diode structures

$X_{\#}$	x_1	x_2	x_3	x_4	x_5	Class c* (optical power)
1	40	40	30	30	40	**1**
2	40	20	40	30	40	**1**
3	40	30	55	40	40	**1**
4	30	40	40	30	55	**1**
5	40	30	40	30	40	**2**
6	40	40	40	30	55	**3**
7	40	40	30	40	40	**3**
8	40	20	30	40	40	**4**
9	40	30	40	55	40	**4**
10	40	30	40	40	40	**4**
11	55	40	40	30	30	**4**
12	30	40	40	30	30	**5**
13	40	40	40	30	40	**6**
14	55	40	40	30	55	**6**
15	40	40	40	30	30	**7**
16	30	40	40	30	40	1
17	40	40	40	20	40	1
18	40	40	40	40	40	3
19	40	20	40	40	40	4

Comparative performance of the IC approach, ANN trained by error back propagation, and the genetic algorithm (GA) for this example is given in Table 2, where the error is the absolute difference between the computed class and the actual class given in Table 1. These results show that both ANN and GA give inadmissible errors, whereas ANN is too slow in comparison with IC and GA.

Other examples in (Tarakanov & Prokaev 2007; Tarakanov [et al.] 2007) also show a better accuracy of IC over ANN and the GA. This work also demonstrates the theoretically rigorous feature of training IC with a zero error rate. On the other hand, the training errors of ANN and of the GA are usually too high. In addition, attempts to reduce training errors of ANN may lead to the so called overtraining effect when total error increases drastically. These advantages of IC in training time and accuracy are expected to rise as the dimension of the training data increases. According to the results reported in (Tarakanov, Kvachev & Sukhorukov 2005), the IC algorithm reduces more than 50000 41-dimensional training signals of network traffic to less than 800 points of 3D FIN by 62 sec (AMD 1.5 GHz) without any loss of accuracy of recognition (intrusion detection).

Table 2. Comparative performance of IC, ANN and GA

Algorithm	IC	ANN	GA
Training patterns	15	15	15
Training time (s) for Pentium-4 1.8 GHz	<1	120	<1
Total errors on training set	0	0	0
Maximal error on training set	0	0	0
Test patterns	19	19	19
Errors on non-training patterns	0	9	9
Total errors on test set	0	9	9
Maximal error on test set	0	4	5
Mean error per pattern	0.00	0.47	0.47

VII. CONCLUSION

The results reported in (Nicosia, 2004; Nicosia [et al.] 2002, 2003, 2006, 2007; Tarakanov A. & Y., 2004, 2005; Tarakanov [et al.] 2005, 2006) suggest that the IC approach is competitive and sometime outperforms other robust methods of computational intelligence (in particular, neurocomputing and evolutionary algorithms) in terms of training time and accuracy. Possible ways to reinforce these advantages may be norms other than Euclidian norm together with more delicate and sophisticated methods of interpolation by nearest points of FIN and other entropy functions (e.g., Renyi, 1970).

REFERENCES

[1] Agnati, L.F., Tarakanov, A.O., Ferre, S., Fuxe, K., & Guidolin, D. (2005). Receptor-receptor interactions, receptor mosaics, and basic principles of molecular network organization: possible implication for drug development. *Journal of Molecular Neuroscience*, 26/2-3, 193-208.

[2] Atreas, N.D., Karanikas, C.G., & Tarakanov, A.O. (2003). Signal processing by an immune type tree transform. *LNCS* (Vol. 2787, pp. 111-119). Springer.

[3] Cloete, I., & Zurada, J.M. (Eds.) (2000). *Knowledge-Based Neurocomputing*. Cambridge, MA: MIT Press.

[4] Cutello, V., & Nicosia, G. (2002). An Immunological Approach to Combinatorial Optimization Problems, Proc. of 8th Ibero-American Conf. on Artificial Intelligence (IBERAMIA'02), Springer, LNAI, pp. 361-370.

[5] Cutello, V., Nicosia, G., & Pavone, M. (2003). A Hybrid Immune Algorithm with Information Gain for the Graph Coloring Problem. Proc. of Int. Conf. on Genetic and Evolutionary Computation (GECCO'03), Springer-Verlag, pp. 171-182.

[6] Cutello, V., Nicosia, G., & Pavone, M. (2004). An immune algorithm with hyper-macromutations for the

Dill's 2D hydrophobic-hydrophilic model. *Congress on Evolutionary Computation* (Vol. 1, pp. 1074-1080). IEEE Press.

[7] Cutello, V., Nicosia, G., & Pavone, M. (2007). An immune algorithm with stochastic aging and Kullback entropy for the chromatic number problem. *Journal of combinatorial optimization*, (to appear).

[8] De Boer, R.J., Segel, L.A., & Perelson, A.S. (1992). Pattern formation in one and two-dimensional shape space models of the immune system. *Journal of Theoretical Biology*, 155 295-333.

[9] De Castro, L.N., & Timmis, J. (2002). Artificial Immune Systems: A New Computational Intelligence Approach. London: Springer.

[10] Dasgupta, D. (Ed.) (1999). Artificial Immune Systems and Their Applications. Berlin: Springer.

[11] Farmer, J. D., & Packard, N. H., & Perelson, A. S. (1986) The immune system, adaptation, and machine learning, Physica D, Vol. 22, pp. 187-204.

[12] Goncharova L.B., Jacques Y., Martin-Vide C., Tarakanov A.O., & Timmis J.I. (2005). Biomolecular immune-computer: theoretical basis and experimental simulator. *LNCS* (Vol. 3627, pp. 72-85). Berlin: Springer.

[13] Goncharova, L.B., Melnikov, Y.B., & Tarakanov, A.O. (2003). Biomolecular immunocomputing. *LNCS* (Vol. 2787, pp. 102-110). Berlin: Springer.

[14] Horn, R., & Johnson, Ch. (1986). *Matrix Analysis*. Cambridge University Press.

[15] Jerne, N.K. (1974). Toward a network theory of the immune system. *Annals of Immunology*, 125C, 373-389.

[16] Karanikas, C., & Proios, G. (2003). A non-linear discrete transform for pattern recognition of discrete chaotic systems. *Chaos, Solitons and Fractals*, 17, 195-201.

[17] Kullback, S. (1959). *Statistics and Information Theory*. New York: J. Wiley & Sons.

[18] Nicosia G., Cutello, V., Bentley, P., & J. Timmis (2004). Proceedings of the Third International Conference on Artificial Immune Systems, Catania, Springer-Verlag.

[19] Nicosia, G. (2004). *Immune Algorithms for Optimization and Protein Structure Prediction*, PhD Thesis, Department of Mathematics and Computer Science, University of Catania, Italy, December 2004.

[20] Renyi, A. (1970). *Probability Theory*. North-Holland.

[21] Tarakanov, A., & Adamatzky, A. (2002). Virtual clothing in hybrid cellular automata. *Kybernetes*, 31/7-8, 394-405.

[22] Tarakanov, A., & Dasgupta, D. (2000). A formal model of an artificial immune system. *BioSystems*, 55/1-3, 151-158.

[23] Tarakanov, A., & Dasgupta, D. (2002). An immunochip architecture and its emulation. *NASA/DoD Conference on Evolvable Hardware (EH'02)*, 261-265. Alexandria, Virginia.

[24] Tarakanov, A.O., Goncharova, L.B., & Tarakanov, O.A. (2005). A cytokine formal immune network. *LNAI* (Vol. 3630, pp. 510-519). Berlin: Springer.

[25] Tarakanov, A.O., Kvachev, S.V., & Sukhorukov, A.V. (2005). A formal immune network and its implementation for on-line intrusion detection. *Lecture Notes in Computer Science* (Vol. 3685, pp. 394-405). Berlin: Springer.

[26] Tarakanov, A & Prokaev, A. (2007). Identification of cellular automata by immunocomputing. *Journal of Cellular Automata*, (in press).

[27] Tarakanov, A., Prokaev, A. & Varnavskikh, E. (2007). Immunocomputing of hydroacoustic fields. *International Journal of Unconventional Computing*, (in press).

[28] Tarakanov, A.O., Skormin, V.A., & Sokolova, S.P. (2003). *Immunocomputing: Principles and Applications*. New York: Springer.

[29] Tarakanov, A.O., & Tarakanov, Y.A. (2004). A comparison of immune and neural computing for two real-life tasks of pattern recognition. *Lecture Notes in Computer Science* (Vol. 3239, pp. 236-249). Springer.

[30] Tarakanov, A.O., & Tarakanov, Y.A. (2005). A comparison of immune and genetic algorithms for two real-life tasks of pattern recognition. *International Journal of Unconventional Computing*, 1/4, 357-374.