Selected Algorithms of Computational Intelligence in Gastric Cancer Decision Making

Elisabeth Rakus-Andersson
Blekinge Institute of Technology
Sweden

1. Introduction

Due to the latest research (Engelbrecht, 2007; Rutkowski, 2008) the subject of Computational Intelligence has been divided into five main regions, namely, neural networks, evolutionary algorithms, swarm intelligence, immunological systems and fuzzy systems.

Our attention has been attracted by the possibilities of medical applications provided by immunological computation algorithms. Immunological computation systems are based on immune reactions of the living organisms in order to defend the bodies from pathological substances. Especially, the mechanisms of the T-cell reactions to detect strangers have been converted into artificial numerical algorithms.

Immunological systems have been developed in scientific books and reports appearing during the two last decades (de Castro & Timmis, 2002; Dasgupta & Nino, 2008; Engelbrecht, 2007; Forrest et al., 1997). The basic negative selection algorithm NS was invented by Stefanie Forrest (Forrest et al., 1997) to give rise to some technical applications. We can note such applications of NS as computer virus detection (Antunes & Correia, 2011; Harmer et al., 2002; Zhang & Zhao, 2010), reduction of noise effect (Igawa & Ohashi, 2010), communication of autonomous agents (Ishida, 2004) or identification of time varying systems (Wakizono et al., 2006). Even a trial of connection between a computer and biological systems has been proved by means of immunological computation (Cohen, 2006).

Hybrids made between different fields can provide researchers with richer results; therefore associations between immunological systems and neural networks (Gao et al., 2008) have been developed as well.

In the current chapter we propose another hybrid between the NS algorithm and chosen solutions coming from fuzzy systems (Rakus-Andersson, 2007, 2009, 2010a, 2010b, 2011; Rakus-Andersson & Jain, 2009). This hybrid constitutes the own model of adapting the NS algorithm to the operation decisions "operate" contra "do not operate" in gastric cancer surgery. The choice between two possibilities to treat patients is identified with the partition of a decision region in self and non-self, which is similar to the action of the NS algorithm. The partition is accomplished on the basis of patient data strings/vectors that contain codes of states concerning some essential biological markers. To be able to identify the strings that characterize the "operate" decision we add the own method of computing the patients' characteristics as real values. The evaluation of the patients' characteristics is supported by

inserting importance weights assigned to powerful biological indices taking place in the operation decision process. To compute the weights of importance the Saaty algorithm (Saaty, 1978) is adopted.

We introduce the medical task to solve in Section 2. In order to establish the code systems for clinical data the fuzzification of biological markers is discussed in Section 3. In Section 4 we analyze the way of determining the patient characteristics, which should connect the mix of different codes in one value. The adaptation of the NS algorithm to surgery assumptions is made in Section 5. Finally, in Section 6 we test clinical data to prove the action of the model introduced in the paper as an applicable novelty.

2. The description of the medical objective in gastric cancer surgery

Gastric cancer patients are mostly cured by operating on them. Different types of surgery are taken into account. Two of them, namely, the partial resection surgery contra the radical surgery are considered by surgeons when evaluating biological markers in the context of their deviations from normal values (Do-Kyong Kim et al., 2009; de Mello et al., 1983).

Nevertheless, a surgeon often must decide if any operation on a patient is possible. The choice between the status "operate" and "do not operate" will constitute the main problem to solve by engaging different algorithms with their origins in Computational Intelligence. The selection will be made on the basis of three biological markers listed as $X = age$, $Y = CRP\text{-}value$ (C reactive proteins), and $Z = body\ weight$ (Do-Kyong Kim et al., 2009; de Mello et al., 1983). These are considered as the most important indices in gastric cancer surgery decision making.

As a leading method, which should provide us with decisions "operate" against "do not operate", we adapt the NS (Negative Selection) algorithm of immunological computation. To comprehend better some associations between the body immunological system and artificially invented algorithms based on the body protection system let us recall the most essential definitions of immunity.

Immunity refers to the condition, in which the organism can resist diseases. A broader definition of immunity is a reaction to foreign substances (pathogens). The biological immune system (BIS) has the ability to detect foreign substances and to respond them. One of the main capabilities of the immune system is to distinguish own body cells from foreign substances, which is called self/non-self discrimination (Dasgupta & Nino, 2008; Engelbrecht, 2007; Forrest et al., 1997).

This particular ability is assigned to a special kind of lymphocytes called T-cells produced in the bone marrow. The T-cells can differentiate own body cells from pathogenic cells; therefore they play the role of detectors. Both own cells belonging to the self region and foreign pathogen cells forming non-self domain have their special characteristics given in the form of vectors of coded or measured properties.

Let us adapt the meaning of distribution into self and non-self in the medical application sketched as follows. To make a decision concerning an individual patient we assign the immunological region of self to "operate", whereas the non-self field will be identified with "do not operate" (Rakus-Andersson, 2011).

To be able to use the self/non-self discrimination, accomplished by the NS algorithm, we need to create vectors of coded patient data. The own fuzzy technique will be involved to

divide reference sets of *X*, *Y* and *Z* into subintervals assisting growth levels of these biological indices. To the subintervals, in turn, the codes are added. We will arrange a code vector assisting the patient's features after examining his/her values of *X*, *Y* and *Z*.

When implementing the NS algorithm we assume that vectors characteristic of self region are available as input data. We do not intend to test too many casual vectors to decide their similarity with self vectors since we want to use the NS algorithm in the effective way. We thus try to generate the strongest population of strings being representatives of the self region.

In order to select this population we provide another own algorithm that converts the code vector to a real value. This value will be recognized as "characteristics of the patient".

3. Fuzzification of *X*, *Y* and *Z* in the creation of code vectors

Before studying the technique of making the self/non-self discrimination to state if the patient can be operated or not we should first be able to compare different strings $v = (x = age, y = CRP, z = body\ weight)$, $x \in X$, $y \in Y$, $z \in Z$, to decide their grades of affinity (coverage). We thus should design sets of codes for each biological parameter.

The markers *age*, *CRP*, and *body weight* are measurable features. Hence, we intend to determine the collections of codes assisting intervals, which correspond to the markers' levels. We want to accomplish a process of fuzzification of the measurable markers in order not to decide lengths of the level intervals intuitively (Rakus-Andersson, 2009, 2010a, 2010b, Rakus-Andersson & Jain, 2009).

A fuzzy set, say, *A* in the universe *X* is a collection of elements followed by the membership degrees that are computed by means of the membership function $\mu_A : X \rightarrow [0,1]$. Therefore *A* is denoted as $A = \{(x, \mu_A(x)), x \in X\}$. *A* is called normal if at least one element in the set *A* is assigned to the membership degree equal to 1. The support of *A* is a non-fuzzy set that consists of elements accompanied by membership degrees greater than 0.

The three quantitative markers *X*, *Y* and *Z* will be then differentiated into levels expressed by lists of terms. The terms from the lists are represented by fuzzy sets (Rakus-Andesson, 2007, 2010b), restricted by the membership functions lying over the domains $[x_{min}, x_{max}]$, $[y_{min}, y_{max}]$ and $[z_{min}, z_{max}]$ respectively.

In conformity with the physician's suggestions we introduce five levels of *X*, *Y* and *Z* as the collections

$$X = \text{"age"} = \{ X_1 = \text{"very young"}, X_2 = \text{"young"}, X_3 = \text{"middle-aged"}, X_4 = \text{"old"}, X_5 = \text{"very old"}\}$$

$$Y = \text{"CRP-value"} = \{ Y_1 = \text{"very low"}, Y_2 = \text{"low"}, Y_3 = \text{"medium"}, Y_4 = \text{"high"}, Y_5 = \text{"very high"}\}$$

and

$$Z = \text{"body weight"} = \{ Z_1 = \text{"very underweighted"}, Z_2 = \text{"underweighted"}, Z_3 = \text{"normal"}, Z_4 = \text{"over weighted"}, Z_5 = \text{"very over weighted"}\}.$$

To accomplish a formal mathematical design of level restrictions let us study the special own technique of their implementations (Rakus-Andersson, 2007, 2010b). In general, we suggest that the linguistic list of terms is converted to a sampling of fuzzy sets $L_1,…,L_m$, where m is an odd positive integer. Each term is represented by the corresponding fuzzy set, whose restriction is supposed to be created as the common formula depending on the lth value, where $l = 1,…,m$. We assume that supports of restrictions $\mu_{L_l}(w)$, $l = 1,…,m$, will cover parts of the reference set $L = [\min(L_1),\max(L_m)]$, $w \in L$. We introduce $E = |L|$ as the length of L.

We divide all expressions L_l in three groups, namely, a family of "*leftmost*" sets $L_1,…, L_{\frac{m-1}{2}}$, the set $L_{\frac{m+1}{2}}$ "*in the middle*" and a collection of "*rightmost*" sets $L_{\frac{m+3}{2}},…,L_m$. To design the membership functions of L_l the s-class function

$$s(w,\alpha,\beta,\gamma) = \begin{cases} 0 & \text{for} \quad w \le \alpha, \\ 2\left(\frac{w-\alpha}{\gamma-\alpha}\right)^2 & \text{for} \quad \alpha \le w \le \beta, \\ 1-2\left(\frac{w-\gamma}{\gamma-\alpha}\right)^2 & \text{for} \quad \beta \le w \le \gamma, \\ 1 & \text{for} \quad w \ge \gamma, \end{cases} \tag{1}$$

will be adopted. The point $(a, 0)$ starts the graph of the s-function, whereas the point $(\gamma, 1)$ terminates this graph. The parameter β is found as the arithmetic mean of a and γ. In $w = \beta$ the s-function reaches the value of 0.5.

When designing parameters of each class function we want to consider the possibility to obtain the equal lengths of these parts of L_l's supports, which assist membership values greater than or equal to 0.5. The parts are regarded as the important representatives of fuzzy sets as they possess the largest index of the relationship to the set. We thus determine the breadth of each L_l to be $\frac{E}{m}$ on the membership level equal to 0.5.

Let us first design the parameters of the membership function "*in the middle*". The function of $L_{\frac{m+1}{2}}$ is constructed as a π-function

$$\pi(w) = \begin{cases} s(w,\alpha_1,\beta_1,\gamma_1 = \alpha_2) & \text{for} \quad w \le \gamma_1 = \alpha_2, \\ 1-s(w,\alpha_2 = \gamma_1,\beta_2,\gamma_2) & \text{for} \quad w \ge \gamma_1 = \alpha_2. \end{cases} \tag{2}$$

We suppose that $L_{\frac{m+1}{2}}$ will be a normal fuzzy set in $\gamma_1 = \alpha_2 = \frac{E}{2}$.

In order to guarantee the breadth $\frac{E}{m}$ on the membership level 0.5, function $L_{\frac{m+1}{2}}$ should take $\beta_1 = \frac{E}{2} - \frac{E}{2m} = \frac{E(m-1)}{2m}$ and $\beta_2 = \frac{E}{2} + \frac{E}{2m} = \frac{E(m+1)}{2m}$. Since $L_{\frac{m+1}{2}}$ is expected to preserve the uniform and symmetric shape then $\alpha_1 = \frac{E}{2} - 2\frac{E}{2m} = \frac{E(m-2)}{2m}$ and $\alpha_2 = \frac{E}{2} + 2\frac{E}{2m} = \frac{E(m+2)}{2m}$. We state $L_{\frac{m+1}{2}}$'s formula as

$$\mu_{L_{\frac{m+1}{2}}}(w) = \begin{cases} 0 & \text{for} \quad w \leq \frac{E(m-2)}{2m}, \\ 2\left(\frac{w-\frac{E(m-2)}{2m}}{\frac{E}{m}}\right)^2 & \text{for} \quad \frac{E(m-2)}{2m} \leq w \leq \frac{E(m-1)}{2m}, \\ 1-2\left(\frac{w-\frac{E}{2}}{\frac{E}{m}}\right)^2 & \text{for} \quad \frac{E(m-1)}{2m} \leq w \leq \frac{E}{2}, \\ 1-2\left(\frac{w-\frac{E}{2}}{\frac{E}{m}}\right)^2 & \text{for} \quad \frac{E}{2} \leq w \leq \frac{E(m+1)}{2m}, \\ 2\left(\frac{w-\frac{E(m+2)}{2m}}{\frac{E}{m}}\right)^2 & \text{for} \quad \frac{E(m+1)}{2m} \leq w \leq \frac{E(m+2)}{2m}, \\ 0 & \text{for} \quad w \geq \frac{E(m+2)}{2m}. \end{cases} \tag{3}$$

For the "*leftmost*" family $L_1,...,L_{\frac{m-1}{2}}$ we make suggestions that the top segments of functions lying on the membership level 1 will have the same lengths. Moreover, the last "*left*" function $L_{\frac{m-1}{2}}$ should have the intersection point with "*in the middle*" function on the membership level 0.5. Each upper segment of L_t, $t = 1,...,\frac{m-1}{2}$, will be thus equal to $\frac{\frac{E}{2}}{\frac{m-1}{2}+1} = \frac{E}{m+1}$. Particularly, $\alpha_{L_{\frac{m-1}{2}}} = \frac{E(m-1)}{2(m+1)}$ after multiplying the length of the distinct upper segment by the number of the last left function. We have already found $\beta_1 = \frac{E(m-1)}{2m}$ of the "*in the middle*" function $L_{\frac{m+1}{2}}$. Due to the previously made assumption the functions $L_{\frac{m-1}{2}}$ and $L_{\frac{m+1}{2}}$ should intersect each other in point $\left(\frac{E(m-1)}{2m}, 0.5\right)$; therefore $\beta_{L_{\frac{m-1}{2}}} = \frac{E(m-1)}{2m}$. The difference between $\alpha_{L_{\frac{m-1}{2}}}$ and $\beta_{L_{\frac{m-1}{2}}}$ is evaluated to be $\frac{E(m-1)}{2m(m+1)}$. To find a uniform slope of $L_{\frac{m-1}{2}}$ we determine $\gamma_{L_{\frac{m-1}{2}}} = \beta_{L_{\frac{m-1}{2}}} + \frac{E(m-1)}{2m(m+1)} = \frac{E(m-1)(m+2)}{2m(m+1)}$.

Since the beginning of $L_{\frac{m-1}{2}}$ is planned to be placed in $(\min(L_1), 1)$ then $\mu_{L_{\frac{m-1}{2}}}(w) = 1 - s\left(w, \frac{E(m-1)}{2(m+1)}, \frac{E(m-1)}{2m}, \frac{E(m-1)(m+2)}{2m(m+1)}\right)$. The membership function of $L_{\frac{m-1}{2}}$ is thus expanded as

$$\mu_{L_{\frac{m-1}{2}}}(w) = \begin{cases} 1 & \text{for} \quad w \leq \frac{E(m-1)}{2(m+1)}, \\ 1-2\left(\frac{w-\frac{E(m-1)}{2(m+1)}}{\frac{E(m-1)}{m(m+1)}}\right)^2 & \text{for} \quad \frac{E(m-1)}{2(m+1)} \leq w \leq \frac{E(m-1)}{2m}, \\ 2\left(\frac{w-\frac{E(m-1)(m+2)}{2m(m+1)}}{\frac{E(m-1)}{m(m+1)}}\right)^2 & \text{for} \quad \frac{E(m-1)}{2m} \leq w \leq \frac{E(m-1)(m+2)}{2m(m+1)}, \\ 0 & \text{for} \quad w \geq \frac{E(m-1)(m+2)}{2m(m+1)}. \end{cases} \tag{4}$$

All constraints characteristic of the "*leftmost*" family of fuzzy sets will be given after inserting parameter $\delta(t) = \frac{2}{m-1} \cdot t$, $t = 1,\ldots,\frac{m-1}{2}$, in (4) to form it as (Rakus-Andersson, 2007)

$$\mu_{L_t}(w) = \begin{cases} 1 & \text{for} \quad w \le \frac{E(m-1)}{2(m+1)}\delta(t), \\ 1 - 2\left(\dfrac{w - \frac{E(m-1)}{2(m+1)}\delta(t)}{\frac{E(m-1)}{m(m+1)}\delta(t)}\right)^2 & \text{for} \quad \frac{E(m-1)}{2(m+1)}\delta(t) \le w \le \frac{E(m-1)}{2m}\delta(t), \\ 2\left(\dfrac{w - \frac{E(m-1)(m+2)}{2m(m+1)}\delta(t)}{\frac{E(m-1)}{m(m+1)}\delta(t)}\right)^2 & \text{for} \quad \frac{E(m-1)}{2m}\delta(t) \le w \le \frac{E(m-1)(m+2)}{2m(m+1)}\delta(t), \\ 0 & \text{for} \quad w \ge \frac{E(m-1)(m+2)}{2m(m+1)}\delta(t). \end{cases} \tag{5}$$

Parameter $\delta(t)$ takes the value of 1 for $t = \frac{m-1}{2}$, which means that $\delta(\frac{m-1}{2})$ in (5) has no influence on the shape of the last left function. However, the introduction of $\delta(t)$ in (5) induces the narrowing effects in the supports of the other left function shapes. To preserve the same lengths of upper segments corresponding to membership 1 and middle segments attached to membership 0.5 we adjust $\delta(t)$, assisting the left function L_t, to be equal to $\frac{1}{\frac{m-1}{2}}$ multiplied by the function number t.

In order to start the implementation of the "*rightmost*" family functions let us note that the first right function $\mu_{L_{\frac{m+3}{2}}}(w)$ should possess $\mu_{L_{\frac{m-1}{2}}}(w)$'s inverted shape. We generate the membership function of $L_{\frac{m+3}{2}}$ by

$$\mu_{L_{\frac{m+3}{2}}}(w) = \begin{cases} 0 & \text{for} \quad w \le E - \frac{E(m-1)(m+2)}{2m(m+1)}, \\ 2\left(\dfrac{w - \left(E - \frac{E(m-1)(m+2)}{2m(m+1)}\right)}{\frac{E(m-1)}{m(m+1)}}\right)^2 & \text{for} \quad E - \frac{E(m-1)(m+2)}{2m(m+1)} \le w \le E - \frac{E(m-1)}{2m}, \\ 1 - 2\left(\dfrac{w - \left(E - \frac{E(m-1)}{2(m+1)}\right)}{\frac{E(m-1)}{m(m+1)}}\right)^2 & \text{for} \quad E - \frac{E(m-1)}{2m} \le w \le E - \frac{E(m-1)}{2(m+1)}, \\ 1 & \text{for} \quad w \ge E - \frac{E(m-1)}{2(m+1)}. \end{cases} \tag{6}$$

The function of $L_{\frac{m+3}{2}}$ is symmetrically inverted to the function of $L_{\frac{m-1}{2}}$ over interval $[\min(L_1), \max(L_m)]$.

Hence, the membership function $\mu_{L_{\frac{m-1}{2}}}(w) = 1 - s\left(w, \frac{E(m-1)}{2(m+1)}, \frac{E(m-1)}{2m}, \frac{E(m-1)(m+2)}{2m(m+1)}\right)$ will be changed into $\mu_{L_{\frac{m+3}{2}}}(w) = s\left(w, E - \frac{E(m-1)(m+2)}{2m(m+1)}, E - \frac{E(m-1)}{2m}, E - \frac{E(m-1)}{2(m+1)}\right)$.

To generate the "*rightmost*" family of sets $L_{\frac{m+3}{2}},\ldots,L_m$ we need to create a new parameter $\varepsilon(t) = 1 - \frac{2}{m-1}(t-1)$, $t = 1,\ldots,\frac{m-1}{2}$, which will be inserted in (6). The construction of $\varepsilon(t)$, when

comparing to the creation of $\delta(t)$, is authorized by the fact that $t = 1$ should be followed by $\varepsilon(1) = 1$, whereas $t = \frac{m-1}{2}$ is helped by $\varepsilon(\frac{m-1}{2}) = \frac{2}{m-1}$. Formula (7) constitutes a common base for deriving membership functions

$$
\mu_{L_{\frac{m+3}{2}+t-1}}(w) =
\begin{cases}
0 & \text{for} \quad w \leq E - \frac{E(m-1)(m+2)}{2m(m+1)}\varepsilon(t), \\[2ex]
2\left(\frac{w-\left(E-\frac{E(m-1)(m+2)}{2m(m+1)}\varepsilon(t)\right)}{\frac{E(m-1)}{m(m+1)}\varepsilon(t)}\right)^2 & \text{for} \quad E - \frac{E(m-1)(m+2)}{2m(m+1)}\varepsilon(t) \leq w \leq E - \frac{E(m-1)}{2m}\varepsilon(t), \\[2ex]
1 - 2\left(\frac{w-\left(E-\frac{E(m-1)}{2(m+1)}\varepsilon(t)\right)}{\frac{E(m-1)}{m(m+1)}\varepsilon(t)}\right)^2 & \text{for} \quad E - \frac{E(m-1)}{2m}\varepsilon(t) \leq w \leq E - \frac{E(m-1)}{2(m+1)}\varepsilon(t), \\[2ex]
1 & \text{for} \quad w \geq E - \frac{E(m-1)}{2(m+1)}\varepsilon(t).
\end{cases}
\tag{7}
$$

The functions of fuzzy sets $L_1,...,L_m$ intend to maintain the same distances on the membership level 0.5. This property allows assigning to $L_1,...,L_m$ the relevant parts of their supports possessing the same length. The relevant parts of fuzzy sets consist of the sets' elements that reveal the membership degree values greater than or equal to 0.5. When forming the supports of the same length, in turn, we warrant the partition of $[\min(L_1),\max(L_m)]$ in equal subintervals standing for L_l levels, $l = 1,...,m$. Apart from that, the "*leftmost*" and "*rightmost*" functions also keep the same distances on the membership level 1. This feature provides us with a harmonious arrangement of function shapes.

All steps of the discussed algorithm, which initiates three sets of membership functions corresponding to a list of terms, can be sampled in the block scheme. We need to follow the steps of the scheme together with formulas (3), (5) and (7) to write the excerpt of a computer program. We emphasize that the only data, used in the algorithm, are the length of the reference set and the number of functions. We do not need to specify the sets' borders in the process of the program initialization, as most of programmers do, since the borders are computed automatically by formulas (3), (5) and (7). The steps of the algorithm flow chart are sampled in Fig. 1.

The procedure discussed above has started introduction of membership functions typical of levels of X, Y and Z which, in turn, represent *age*, *CRP* and *body weight*.

For five levels of $X = [0, 100]$, $L = X$, $w = x$, $m = 5$, $E = 100$, the leftmost family is revealed by

$$
\mu_{X_t}(x) =
\begin{cases}
1 & \text{for} \quad x \leq 33.33(0.5t), \\[1ex]
1 - 2\left(\frac{x-33.33(0.5t)}{13.33(0.5t)}\right)^2 & \text{for} \quad 33.33(0.5t) \leq x \leq 40(0.5t), \\[1ex]
2\left(\frac{x-46.66(0.5t)}{13.33(0.5t)}\right)^2 & \text{for} \quad 40(0.5t) \leq x \leq 46.66(0.5t), \\[1ex]
0 & \text{for} \quad x \geq 46.66(0.5t),
\end{cases}
\tag{8}
$$
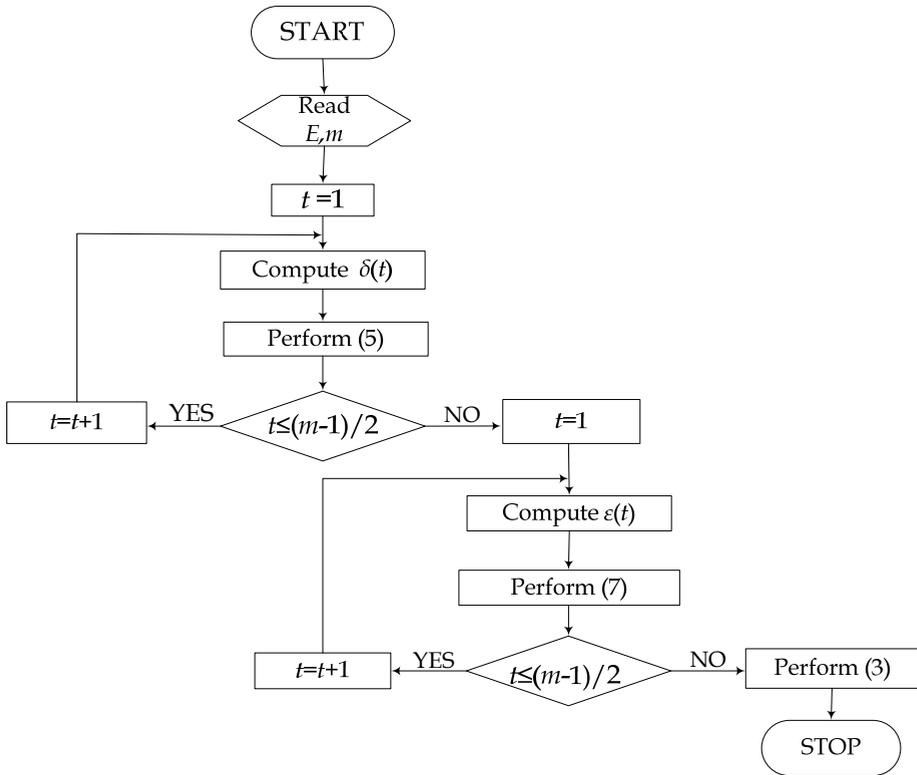
for $t = 1, 2$ due to (5).

Fig. 1. The flow chart of the $L_1,\ldots,L_m$ implementation

The rightmost family of X-levels, composed with conformity with (7), is stated as

$$\mu_{X_{4+t-1}}(x) =$$
$$\begin{cases} 0 \text{ for } x \le 100 - 46.66(1 - 0.5(t-1)), \\ 2\left(\frac{x - (100 - 46.66(1 - 0.5(t-1)))}{13.33(1 - 0.5(t-1))}\right)^2 \\ \text{for } 100 - 46.66(1 - 0.5(t-1)) \le x \le 100 - 40(1 - 0.5(t-1)), \\ 1 - 2\left(\frac{x - (100 - 33.33(1 - 0.5(t-1)))}{13.33(1 - 0.5(t-1))}\right)^2 \\ \text{for } 100 - 40(1 - 0.5(t-1)) \le x \le 100 - 33.33(1 - 0.5(t-1)), \\ 1 \text{ for } x \ge 100 - 33.33(1 - 0.5(t-1)), \end{cases}$$
(9)

for $t = 1, 2$.

The "*in the middle*" *X*-level "*middle-aged*" has, in accord with (3), the constraint

$$\mu_{X_3}(x) = \begin{cases} 0 & \text{for} \quad x \le 30, \\ 2\left(\frac{x-30}{20}\right)^2 & \text{for} \quad 30 \le x \le 40, \\ 1 - 2\left(\frac{x-50}{20}\right)^2 & \text{for} \quad 40 \le x \le 50, \\ 1 - 2\left(\frac{x-50}{20}\right)^2 & \text{for} \quad 50 \le x \le 60, \\ 2\left(\frac{x-70}{20}\right)^2 & \text{for} \quad 60 \le x \le 70, \\ 0 & \text{for} \quad x \ge 70. \end{cases} \tag{10}$$

All levels of *X* are sketched in Fig. 2.

The parts of X_1–X_5 supports should be consisted of elements, which have the strongest connections with the X_1–X_5 fuzzy sets. Therefore we only select the elements having the membership degrees greater than or equal to 0.5.

To make the partition of *X* in subintervals representing levels X_1–X_5 we return to formulas (8), (9) and (10). Due to (8), to find the subinterval of *X* assisting X_1 when $t = 1$, we concatenate the intervals $x \le 33.33(0.5 \cdot 1)$ and $33.33(0.5 \cdot 1) \le x \le 40(0.5 \cdot 1)$, leading to [0, 20]. We have chosen these intervals, which contain elements of X_1 furnished with membership degrees greater than or equal to 0.5. For $t = 2$, set in two first intervals of (8), we aggregate $x \le 33.33(0.5 \cdot 2)$ and $33.33(0.5 \cdot 2) \le x \le 40(0.5 \cdot 2)$ in [0, 40]. This generates the interval [0, 40]–[0, 20] = [20, 40] typical of X_2. By (10) we find [40, 60] = [40, 50] + [50, 60] as an essential part of X_3. The insertion of $t = 2$ in (9) produces a joint of
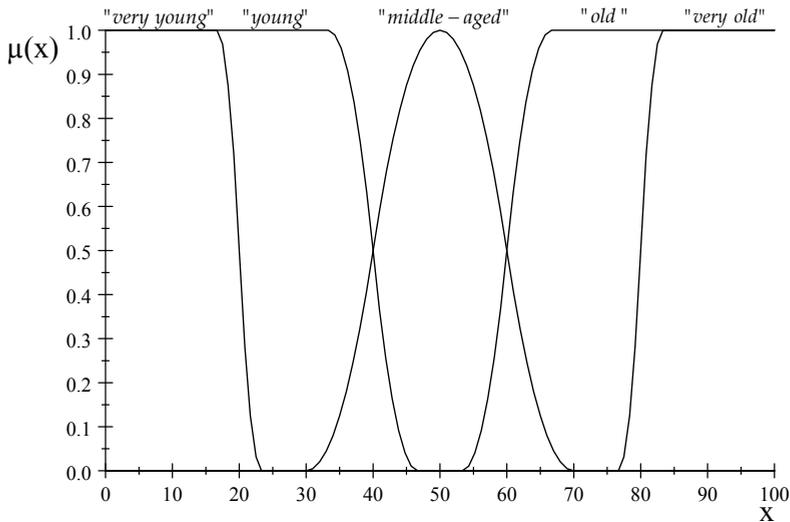

Fig. 2. The fuzzy sets X_1–X_5

$100 - 40(1 - 0.5(2 - 1)) \leq x \leq 100 - 33.33(1 - 0.5(2 - 1))$ and $x \geq 100 - 33.33(1 - 0.5(2 - 1))$ to be a common interval [80, 100] regarded as the domain of X_5. By setting $t = 1$ in the last intervals of (9) we get the field [60, 100]. It means that X_4 will be given by [60, 80] = [60, 100]–[80, 100]. We are furnished with the same intervals after accomplishing the close analysis of Fig. 2 on the membership level 0.5.

Let us now initiate the associations among the terms of X, characteristic intervals of these terms and assigned to them codes due to the scheme

name of X-level	representative interval	code
X_1	0–20	0
X_2	20–40	1
X_3	40–60	2
X_4	60–80	3
X_5	80–100	4

We emphasize the role of an elegant mathematical design of X's membership functions, which allows making the partition of the X-domain in equal intervals. Definitely, we obtain the same results when dividing the length of X by the number of levels to get a length of one part but the effects computed by means of membership functions only confirm this intuitive calculation. Moreover we can modify the arbitrary lengths of X-subintervals by making changes in the formulas of $\delta(t)$ and $\varepsilon(t)$.

By applying the same technique to $Y = [0, 60]$, $L = Y$, $w = y$, $m = 5$, $E = 60$ we generate the code pattern

name of Y-level	representative interval	code
Y_1	0–12	0
Y_2	12–24	1
Y_3	24–36	2
Y_4	36–48	3
Y_5	48–60	4

Lastly, if $Z = [40, 120]$ for men, $L = Z$, $w = z$, $m = 5$, $E = 80$ then

name of Y-level	representative interval	code
Z_1	40–56	0
Z_2	56–72	1
Z_3	72–88	2
Z_4	88–104	3
Z_5	104–120	4

If we collect clinical data, concerning a patient examined then we will be now capable to create code vectors taking place in the discrimination NS algorithm.

Example 1

An eighty one-year-old man, whose *CRP* is 17 and weight is 91, will be given by the vector $v = (4, 1, 3)$.

In order to measure the affinity (coverage) of two code vectors v_1 and v_2 of the same length over the same alphabet we are furnished with the r-contiguous bit matching rule, which provides us with a true match(v_1, v_2) if v_1 and v_2 agree in r contiguous locations.

Example 2

For $v_1 = (3, 1, 3)$ and $v_2 = (3, 1, 2)$, when $r = 2$, match(v_1, v_2) is true.

4. The selection of the most representative data vectors for the decision "operate"

We have already mentioned that we need the "operate" types of patient data vectors as the entries of the NS discrimination algorithm. We thus want to prepare typical data strings for the decision "operate" in advance.

Let us first treat the vector $v = (x, y, z)$ as the string of integers $v = (x\ y\ z)$, where x, y and z can take the code values 0, 1, 2, 3, 4. We form the function $f(x\ y\ z) = x + y + z$ to measure the common code value of the data vector. To make the selection of "operate" type vectors even more accurate let us assign the weights of power-importance to the biological indices considered in the operation decision. In the gastric cancer operation decision we first concentrate our attention on the changes of *CRP*- values, which points out *CRP* as the most decisive factor. The analysis of *CRP* is followed by the judgment of age and, finally, we check the values of body weights. Hence, we state the ranking of the symptom importance as $CRP \succ age \succ body\ weight$, provided that $\succ$ means "more important than".

A procedure for obtaining a ratio scale of importance for a group of m elements (in the considered case – biological markers) was developed by Saaty (Saaty, 1978). Assume that we have m objects (symptoms) and we want to construct a scale, rating these objects as to their importance with respect to the decision. We ask a decision-maker to compare the objects in paired comparison. If we compare object j with object k, $j, k = 1,...,m$, then we will assign the values b_{jk} and b_{kj} as follows

1. $b_{kj} = \dfrac{1}{b_{jk}}$.

2. If objective j is more important than objective k then b_{jk} gets assigned a number according to the following scheme:

Intensity of importance *expressed by the value of b_{jk}*	*Definition*
1	Equal importance of x_j and x_k
3	Weak importance of x_j over x_k
5	Strong importance of x_j over x_k
7	Demonstrated importance of x_j over x_k
9	Absolute importance of x_j over x_k
2, 4, 6, 8	Intermediate values

If object k is more important than object j, we assign the value of b_{kj}.

Having obtained the above judgments an $m \times m$ importance matrix $B = \left(b_{jk}\right)_{j,k=1}^{m}$ is constructed. The importance weights are decided as components of this eigenvector that corresponds to the largest in magnitude eigenvalue of the matrix B.

Example 3

For priorities $Y = CRP \succ X = age \succ Z = body\ weight$ we determine the contents of B as

$$B = \begin{array}{c} \\ X \\ Y \\ Z \end{array} \begin{array}{ccc} X & Y & Z \\ \begin{bmatrix} 1 & \frac{1}{3} & 3 \\ 3 & 1 & 5 \\ \frac{1}{3} & \frac{1}{5} & 1 \end{bmatrix} \end{array}$$

The largest eigenvalue ($\lambda = 3.033$) of B has the associated eigenvector $V = (0.37, 0.92, 0.15)$. V is composed of coordinates that are interpreted as the importance weights w_1, w_2, w_3 sought for X, Y, Z.

Let us rearrange the form of function f by adding the weights of importance to the vector code values. The new pattern of f is designed as $f(x\ y\ z) = w_1x + w_2y + w_3z$. The function value yields the patient's characteristics given by a combination of codes stated for different symptoms.

Example 4

The patient vector v = (3, 1, 2) has the characteristics $f(3\,1\,2) = 0.37 \cdot 3 + 0.92 \cdot 1 + 0.15 \cdot 2 = 2.33$.

Due to the physician's expertise we assume that we can operate patients who are characterized by codes of age equal to 1, 2 and 3, codes of CRP recognized as 0, 1 and 2 and codes of body weight determined as 1, 2 and 3. The minimal patient characteristics to be operated is thus $f(1\,0\,1) = 0.37 \cdot 1 + 0.92 \cdot 0 + 0.15 \cdot 1 = 0.52$, whereas the maximal data characteristics, classifying the patient for the operation is given by $f(3\,2\,3) = 0.37 \cdot 3 + 0.92 \cdot 2 + 0.15 \cdot 3 = 3.4$. We conclude that the patient who is capable to be operated should have the characteristics $f(x\ y\ z)$ included in the interval [0.52, 3.4]. It is worth emphasizing that the decisions are made with respect to the decisive power of biological markers *age*, *CRP* and *body weight*.

Example 5

We test v = (4, 2, 4). As $f(4\,2\,4) = 0.37 \cdot 4 + 0.92 \cdot 2 + 0.15 \cdot 4 = 3.92$, which lies beyond the boundaries of the "operate" interval, then the patient with data v = (4, 2, 4) should not be operated. For vectors v_1 = (3, 2, 2), v_2 = (2, 0, 2), v_3 = (3, 1, 3), v_4 = (3, 1, 2) the decision will be made as "operate".

The flow chart, sketched in Fig. 3, will show the selection of vectors typical of the decision "operate".

The vectors v_1, v_2, v_3 and v_4 will be included in the experimental population of representative data strings for the positive decision of the operation. We intend to use them in the next part of the chapter, when discussing the action of the NS algorithm adapted to the operation model.

5. The negative selection algorithm

After coding the patient data and selecting the initial data, which are given into account for the decision "operate", we can make a choice between two alternatives concerning the cure
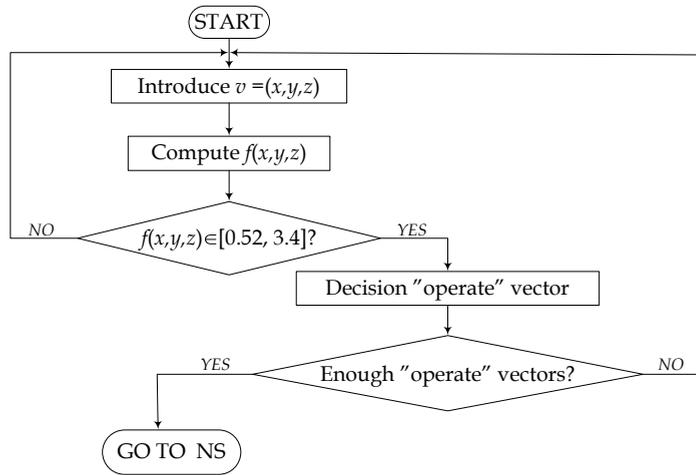
START

Introduce $v = (x,y,z)$

Compute $f(x,y,z)$

$f(x,y,z) \in [0.52, 3.4]$?

NO   *YES*

Decision "operate" vector

YES   Enough "operate" vectors?   *NO*

GO TO NS

Fig. 3. The flow chart of the selection of "operate" type vectors

of gastric cancer patients. We intend to adapt the technique of an immunological algorithm based on the T cell behaviour. We use the negative selection algorithm NS proposed by Forrest (Forrest et al., 1997).

The goal of NS is to cover the non-self space with a set of detectors. For the sake of the surgery aim, already outlined in Section 2, the algorithm should lead to discrimination of the statements "operate" and "do not operate" provided that vectors characteristic of type "operate" are available. This assumption is motivated by the surgeon's intention to cure the patient from his/her cancer disease by making surgery if the patient's state allows accomplishing it. The patient data reports, which register his/her parameters in the case of operating, are clearly interpretable. However, the physician can have some doubts when he denies an operation for the patient. Therefore we have used the strings confirming the "operate" decision as the more convincing vectors in the entrance of NS.

We distinguish two steps in the surgery NS algorithm prepared on the basis of the general NS (Dasgupta & Nino, 2008; Engelbrecht, 2007; Forrest et al., 1997):

1. Generation of detectors, which should possess the property vectors corresponding to the decision "do not operate" on a patient. These strings are not recognized as obviously as the strings of "operate"; that is why we get some help from the algorithm in generating their patterns.
2. Selection of the surgery settlements "operate" or "do not operate" for any patient data vector due to the matching criterion concerning detectors.

In the first step a set of detectors is generated. To accomplish this task we use as an input a collection of vectors found by the method of preparing "operate" strings, which have been discussed in Section 4. Candidate detectors that match any of the "operate" type vector

samples are eliminated whereas unmatched ones are kept. We adopt the *r*-contiguous bit matching rule for the patient data vectors as a measure of "the distance" between the "operate" type and the "do not operate" decision.

In the second step of NS the stored detectors, generated in the first stage, are used to check whether new incoming samples of patient data vectors correspond to the "operate" type or to the "do not operate" type. If an input sample, characterizing a patient, matches any detector then the patient should not be operated. When we cannot find a match between detectors and the incoming patient data vector it will mean that the decision about the surgery should be made. Figure 4 collects all steps of the surgery NS algorithm in the flow chart.
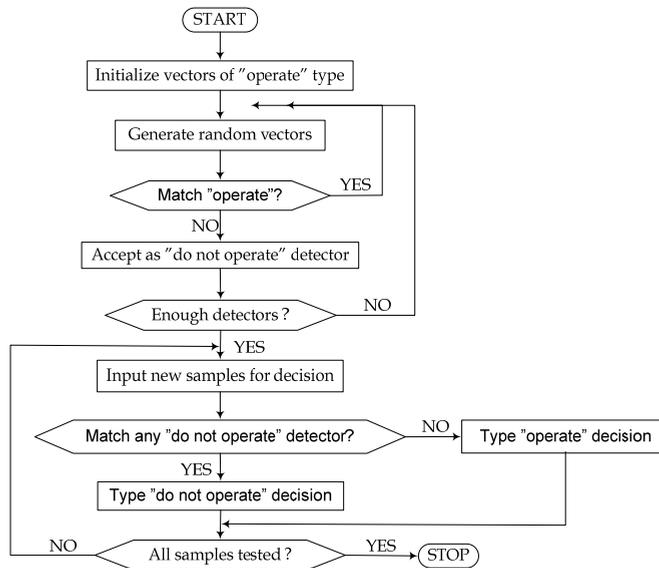

Fig. 4. The flow chart of the surgery NS algorithm

6. The surgery decision based on the NS algorithm

We wish now to follow the steps of the surgery NS algorithm to study its action in practical decision cases concerning the operation decision.

Let us thus go through the following example.

Step 1. *Initialization*

As the input data we introduce the set $V = \{v_1, v_2, v_3, v_4\}$, which consists of four patient data vectors characteristic of the "operate" type. The length of each vector is decided to be three in conformity with previously made suggestions. In Section 4 we have already initialized

$v_1 = (3, 2, 2),$
$v_2 = (2, 0, 2),$

$v_3 = (3, 1, 3)$,
$v_4 = (3, 1, 2)$.

The vectors emerge the clinical data concerning elderly patients whose the *CRP-* values are not very high. The patients' weights are not radically deviated from normal standards either. Hence, they have been operated in conformity with the surgeon's determination.

We now wish to generate the set D of four detectors d_1, d_2, d_3, d_4 that should not match any of v_j, $j =1,\dots,4$. At the beginning of the procedure D is an empty set.

To measure the match grade between v_j and candidates to be detectors we state, e.g., $r = 2$ in the *r*-contiguous bit matching rule.

Step 2. *Introduction of random candidates to act as detectors*

We present $d = (3, 1, 1)$ and check matches between d and each v_j, $j = 1,\dots,4$, as

match((3, 2, 2), (3, 1, 1)) is false,
match((2, 0, 2), (3, 1, 1)) is false,
match((3, 1, 3), (3, 1, 1)) is true,
match((3, 1, 2), (3, 1, 1)) is true.

Since d matches v_3 and v_4 then it cannot be classified as a detector.

We prove the next candidate $d = (4, 3, 1)$ to make matches between d and each v_j, $j = 1,\dots,4$, in the form of

match((3, 2, 2), (4, 3, 1)) is false,
match((2, 0, 2), (4, 3, 1)) is false,
match((3, 1, 3), (4, 3, 1)) is false,
match((3, 1, 2), (4, 3, 1)) is false.

All matches are false, which means that $d_1 = d$ is the first detector placed in D. The set of detectors now contains one element $d_1 = (4, 3, 1)$. We repeat the procedure until we determine four detectors in set D. D is finally formed as

$D = \{(4, 3, 1), (2, 3, 4), (4, 4, 1), (3, 4, 0)\}$.

Step 3. *Operation decision making*

In the second phase of the algorithm we test data strings to organize them in either the "operate" type or in the "do not operate" type decisions. If the data vector matches any detector from D then the decision is made as "do not operate" (the non-self region). Otherwise, for all false matches between the data vector and d_k, $k = 1,\dots,4$, we accept the operation (the self region).

We introduce $v = (3, 2, 3)$. The matches to detectors are determined as

match((4, 3, 1), (3, 2, 3)) is false,
match((2, 3, 4), (3, 2, 3)) is false,
match((4, 4, 1), (3, 2, 3)) is false,
match((3, 4, 0), (3, 2, 3)) is false.

As all matches to detectors are false we conclude the performance of surgery (decision "operate").

Another test vector $v = (3, 4, 1)$ is inserted into the checking system. The match results are shown as

match$((4, 3, 1), (3, 4, 1))$ is false,
match$((2, 3, 4), (3, 4, 1))$ is false,
match$((4, 4, 1), (3, 4, 1))$ is true,
match$((3, 4, 0), (3, 4, 1))$ is true.

Vector v converges to two detectors, which means the decision to be referred to "do not operate".

By setting $r = 2$ in the contiguous bit matching rule we have preserved a margin of imprecision in decision making, since we do not demand all contiguous vector codes to be equal. This gives a certain chance of operating for the patients whose mix of biological indices cannot be precisely judged. For $r = 3$ the decision will be quite strict.

The method of making medical decisions by means of immunological systems is an applicable novelty. The example has a more didactic and experimental meaning than a real medical investigation. If we really want to use the method for making decisions in the surgery discipline we should, at first, extend the length of data strings by introducing more biological markers. A very dense set of initial vectors from "self" ("operate") ought to be chosen by the algorithm belonging to Section 4. Nevertheless, the proposal of combining fuzzy systems and weighted characteristics of vectors with the NS algorithm to create the hybrid can start a new applied domain in medicine.

7. Conclusion

In the process of creation of a new medical application model we have inserted some elements of fuzzy systems into the negative selection immunological algorithm. This hybrid, attached to two disciplines of Computational Intelligence, has found a practical application in surgery decision making. As self and non-self constitute two regions of the NS partition of objects then we could identify these regions with decisions "operate" against "do not operate" in the case of curing gastric cancer patients. The action of the modified NS could help us to determine the surgery or its lack for individual patients with respect to their clinical data entry vectors.

To make the action of the NS algorithm more efficient we have complemented the method by preparing the population of the most representative vectors standing for the "operate" type. The vectors have been converted to real values giving the common characteristics of a patient. In that characteristics the weights of importance, assigned to biological markers, will play the essential role in the final judgment of the vectors' influence on the decision "operate".

We wish to add that the excerpts from fuzzy systems, involved in NS, come from own research, which has been concentrated on the creation of compact parametric formulas. These formulas concern the generation of a family of membership functions without predetermining their borders in advance.

All parts of the methodology have been prepared in the form of numerical algorithms given by flow charts. This allows composing a common computer program to test large samples of vectors in a real clinical application.

We emphasize that the proposal is a novel contribution in medical applications and should be still tested on larger samples of data. We can expect that, in future investigations, an introduction of the neural artificial perceptron model instead of the NS algorithm will provide us with similar results concerning surgery decisions. As an extension of the model we also wish to adapt the real-value negative selection algorithm in order to insert measured values of biological markers in data vectors instead of codes. This procedure should improve the reliability of a decision. Having results from more models we can select the most efficient one to work on its further development.

8. Acknowledgment

9. References

Antunes, M. & Correia, M.E. (2011). Tunable Immune Detectors for Behavior-Based Network Intrusion Detection, *Artificial Immune Systems: 10th International Conference, ICARIS 2011*, Cambridge, UK, pp. 334–347, LNCS, vol. 6825, Springer, ISBN 978-3-642-22370-9

De Castro, L.N. & Timmis, J. (2002). *Artificial Immune Systems,* Springer, ISBN 978-1-85233-594-6, Berlin Heidelberg

Cohen, I.R. (2006). Immune System Computation and the Immunological Homunculus, LNCS, vol. 4199, pp. 49–52, ISBN 978-3-540-45772-5

Dasgupta, D. & Nino, F. (2008). *Immunological Computation: Theory and Applications*, Auerbach Publishers Inc., ISBN 978-1-420-06545-9, London

Do-Kyong Kim, Sung Yong Oh, Hyuk-Chan Kwon, Suee Lee, Kyung A Kwon, Byung Geun Kim, Seong-Geun Kim, Sung-Hyun Kim, Jin Seok Jang, Min Chan Kim, Kyeong Hee Kim, Jin-Yeong Han & Hyo-Jin Kim (2009). Clinical Significances of Preoperative Serum Interleukin-6 and C-reactive Protein Level in Operable Gastric Cancer, *BMC Cancer* 9, pp. 155–156, ISSN 1471-2407

Engelbrecht, A.P. (2007). *Computational Intelligence*, Wiley & Sons Ltd., ISBN 978-0-470-03561-0, Chichester

Forrest, S., Hofmeyr, A.S. & Somayaji, A. (1997). Computer Immunology, *Communications of the ACM,* vol. 40, nr 10, pp. 88–96, ISSN 1017-4656

Gao, X.Z., Ovaska, S.I., Wang, X. & Chow, M.Y. (2008). A Neural Networks-based Negative Selection Algorithm in Fault Diagnosis, *Neural Computing & Applications,* vol. 17, nr 1, pp. 91–98, ISSN 1433-3058

Harmer, P.K., Williams, P.D., Gunsch, G.H. & Lamont, G.B. (2002). An Artificial Immune System Architecture for Computer Security Applications, *IEEE Transactions on Evolutionary Computation,* vol. 6, nr 3, pp. 252–280, ISSN 1089-778X

Igawa, K. & Ohashi, H. (2009). A Negative Selection Algorithm for Classification and Reduction of the Noise Effect, *Journal of Applied Soft Computing*, vol. 9, nr 1, pp. 431–438, ISSN 1568-4946

Ishida, Y. (2004). *Immunity-Based Systems: A Design Perspective,* Springer, ISBN 3-540-00896-9, Berlin Heidelberg New York

de Mello, J., Struthers, L., Turner, R., Cooper, E.H. & Giles, G.R. (1983). Multivariate Analyses as Aids to Diagnosis and Assessment of Prognosis in Gastrointestinal Cancer, *Br. J. Cancer,* nr 48, pp. 341–348, ISSN 0007-0920

Rakus-Andersson, E. (2007). *Fuzzy and Rough Techniques in Medical Diagnosis and Medication*, Springer, ISBN 978-3-540-49707-3, Berlin Heidelberg

Rakus-Andersson, E. & Jain, L. (2009). Computational Intelligence in Medical Decisions Making, *Recent Advances in Decision Making*, Springer, Berlin Heidelberg, pp. 145–159, ISBN 978-3-642-02186-2

Rakus-Andersson, E. (2009). Approximate Reasoning in Surgical Decisions, *Proceedings of the International Fuzzy Systems Association World Congress – IFSA 2009*, Instituto Superior Technico, pp. 225–230, ISBN 978-989-95079-6-8

Rakus-Andersson, E. (2010a). One-dimensional Model of Approximate Reasoning in Surgical Considerations, *Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics*, vol. II, System Research Institute of Polish Academy of Sciences, pp. 233–246, ISBN 139788389475305

Rakus-Andersson, E. (2010b). Adjusted s-parametric Functions in the Creation of Symmetric Constraints, *Proceedings of the 10th International Conference on Intelligent Systems Design and Applications - ISDA 2010*, Cairo, Egypt, pp. 451–456, ISBN 978-1-4244-8135-4

Rakus-Andersson, E. (2011). Hybridization of Immunological Computation and Fuzzy Systems in Surgery Decision Making, *Proceedings of the 15th Conference on Knowledge Expert Systems – KES 2011*, Kaiserslautern, Germany, LNCS/LNAI, vol. 6884, Springer, Berlin Heidelberg, pp. 399–408, ISBN 978-3-642-23865-9

Rutkowski, L. (2008). *Computational Intelligence: Methods and Techniques*, ISBN 978-3-540-76287-4, Springer, Berlin Heidelberg

Saaty, T.L. (1978). Exploring the Interface between Hierarchies, Multiplied Objectives and Fuzzy Sets, *Fuzzy Sets and Systems* 1, pp. 57–58, ISSN 0165-0114

Wakizono, M., Hatanaka, T. & Uosaki, K. (2006). Time Varying System Identification with Immune Based Evolutionary Computation, *Proceedings of SICE-ICASE, 2006*, Busan, Korea, pp. 5608–5613, ISBN 89-950038-4-7

Zhang, Q. & Zhao, H. (2010). The Research of Generation Algorithm of Detectors in a New Negative Selection Algorithm, *Proceedings of the International Conference on Technology Management and Innovation 2010*, paper 78, ISBN 13-978-0-791-85961-2