FUZZY REASONING APPLIED TO MULTISTAGE DIAGNOSIS OF ACUTE RENAL FAILURE IN CHILDREN

MAREK W. KURZYŃSKI

Chair of Systems and Computer Networks,
Faculty of Electronics, Wroclaw University of Technology,
Wyb. Wyspiańskiego 27, 50-370 Wroclaw, Poland
kumar@zssk.pwr.wroc.pl

(Received 12 September 2003; revised manuscript received 30 January 2004)

Abstract: The paper deals with fuzzy inference systems for multistage recognition based on a decision tree scheme. Two conceptually different fuzzy methods are presented and discussed for the given learning set. The first method is developed according to the multistage approach known as the Mamdani inference engine, with rules generated from the learning set. In the second approach, we first construct a fuzzy relation between the decision set and the feature space, which is then used for decision making. Both methods were practically applied to computer-aided medical diagnosis of acute renal failure. Results of comparative experimental analysis are given.

Keywords: multistage recognition, fuzzy systems, medical application

1. Introduction

In many practical pattern recognition problems, a method of classification which is traditional as far as the adopted model and procedure are concerned turns out not to be effective enough or to be insufficient. Hence the swift development of compound methods of recognition, in which a decision on the class of an object is not a single activity but is the result of a more or less complex decision process. Multistage recognition, which is the subject of this work, is one of such methods.

The paper is a sequel to the author's earlier publications [1–4] and it presents new results concerning the application of fuzzy inference systems to decision making at particular stages of multistage recognition procedures.

The contents of the work are as follows. Section 2 describes the multistage classification technique. In Section 3 we introduce the necessary background and notations. In Sections 4 and 5 two different fuzzy inference procedures are presented and discussed. The first method uses a fuzzy rule system generated from the learning set. In the second approach, a fuzzy relation between the set of decisions and the feature space is determined as a solution of an appropriate optimization problem. In the presented example, the genetic algorithm was applied to find an optimal solution.

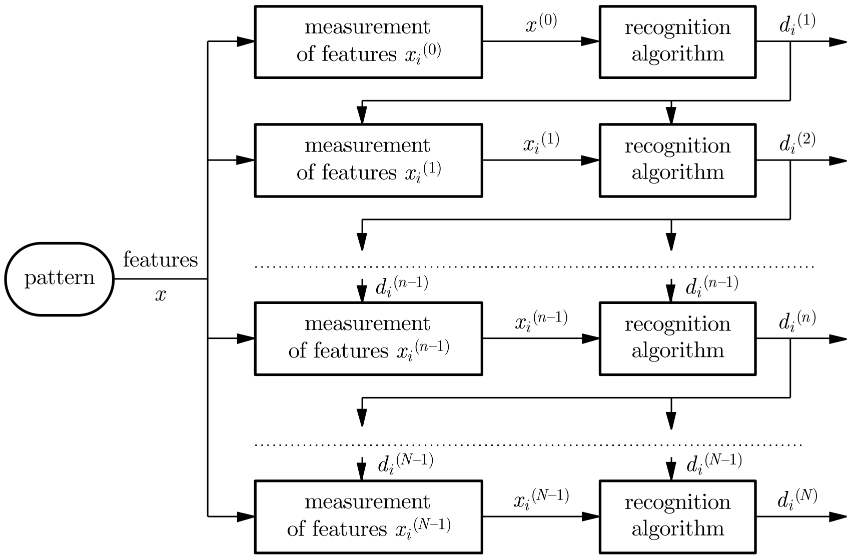

Figure 1. The procedure of multistage recognition

In Section 6 we discuss the results of application of the proposed fuzzy decision systems to computer-aided two-stage diagnosis of acute renal failure.

2. Multistage pattern recognition

The procedure of multistage pattern recognition, presented in Figure 1, consists of the following sequences of activities [2, 3]. At the initial stage, some specified features x_0 are measured, chosen from among all the accessible features, x, describing the pattern being classified. These data constitute the basis for making decision i_1. This decision, being the result of recognition at the initial stage, defines a certain subset in the set of all classes and simultaneously indicates features x_{i_1} (from among x) which should be measured in order to make a decision at the next stage. Now, at the second stage, features x_{i_1} are measured, which – together with i_1 – are the basis for making the next decision, i_2. This decision – like i_1 – indicates features x_{i_2} necessary to make the next decision (at the third stage) and – again, as at the previous stage – defines a certain subset of classes, not in the set of all classes, however, but in the subset indicated by decision i_1, and so on. The whole procedure ends at the last (N^{th}) stage, where the decision made (i_N) indicates a single class, which is the final result of multistage recognition. Thus, multistage recognition means a successive narrowing of the set of potential classes from stage to stage, down to a single class, simultaneously indicating at every stage features which should be measured to make the next decision more precise.

The action of a multistage classifier can be conveniently described by means of a decision-tree (see Figure 4).

The synthesis of a multistage classifier is a complex problem. It involves specification of the following components [3]:

- the decision logic, *i.e.* a hierarchical ordering of classes,

- features used at each decision stage,
- the decision rules (strategy) for performing the classification.

The present paper is devoted to the last problem only. This means that we shall restrict ourselves to a presentation of decision algorithms, assuming that both the tree skeleton and the features used at each non-terminal node have been specified. Moreover, our considerations deal with the case when a fuzzy inference system is applied as a diagnostic algorithm. Fuzzy systems have successful aplications in a wide variety of fields, for example in automatic control, pattern recognition, signal and image processing, to name just a few [5–7]. The differences between these systems lie in the consequences of if-then rules. In accordance with practical requirements and the character of the considered pattern recognition problem, we assume systems with crisp inputs and discrete consequences of rules.

In the following section we introduce the neccessary notations and present multistage recognition algorithms in accordance with the probabilistic model. These algorithms will constitute the conceptual basis for appropriate recognition procedures with a fuzzy inference engine.

3. Preliminaries and problem statement

Let us consider a pattern recognition problem in which $x \in \mathbf{X} = \mathbb{R}^d$ denotes a feature vector describing the pattern to be recognized and $j \in \mathbf{M} = \{1, 2, \ldots, M\}$ is its class number. Let us additionally assume that classes are organized in a N-level decision tree with terminal nodes labelled by class numbers M and that features used at each non-terminal node are given.

Throughout this paper we shall use the following symbols referring to the tree structure and the action of a multistage classifier:

- $\mathbf{M}^{(i_{n-1})} = \{i_{n-1}^{(1)}, i_{n-1}^{(2)}, \ldots, i_{n-1}^{(n')}\}$ – the set of decision numbers at the n^{th} stage determined by decision i_{n-1} made at the previous stage (i_0 denotes the root node);
- $\mathbf{M}_{i_n}$ – the set of class numbers accessible after decision i_n is made at the n^{th} stage (classes connected with terminal nodes of the sub-tree with node i_n as the root node);
- $x^{(i_{n-1})} = (x_1^{(i_{n-1})}, x_2^{(i_{n-1})}, \ldots, x_{d_n}^{(i_{n-1})}) \in \mathbf{X}^{(i_{n-1})}$ – vector of features (numerical variables) used at the n^{th} stage determined by decision i_{n-1} made at the previous stage;
- $\bar{x}^{(i_{n-1})} \in \bar{\mathbf{X}}^{(i_{n-1})}$ – as previously, but now the features are linguistic variables.

If we adopt a probabilistic model of the recognition task and assume that there exist *a priori* probabilities of classes, p_j, and conditional probability density functions of features, $f_j(x)$ ($x \in \mathbf{X}, j \in \mathbf{M}$), then recognition algorithms appropriate at the particular stages of the classification procedure can be obtained by solving a certain optimization problem. Now, however, in contrast to one-stage recognition, the optimality criterion can be formulated in different ways, and various manners of action of the algorithms can be assumed, which in effect give different optimal decision rules for the particular stages of classification [3]. Let us consider two cases.

3.1. Globally optimal strategy (GOS)

A minimization of the mean probability of misclassification of the whole multistage decision process leads to an optimal decision strategy, whose recognition algorithm at the n^{th} stage is as follows:

$$\begin{gathered}\Psi^*_{i_{n-1}}(x_{i_{n-1}})=i^{(k)}_{n-1}\\ \text{if}\quad Pc\left(i^{(k)}_{n-1}\right)\sum_{j\in\mathbf{M}_{i^{(k)}_{n-1}}}p_j(x_{i_{n-1}})=\max_l Pc\left(i^{(l)}_{n-1}\right)\sum_{j\in\mathbf{M}_{i^{(l)}_{n-1}}}p_j(x_{i_{n-1}}),\end{gathered}\tag{1}$$

where $Pc(i)$ is the probability of correct classification at the next stages if at the n^{th} stage decision i is made, and $p_j(x)$ denotes *a posteriori* probability of classes which can be calculated from the given data using the Bayes rule.

The manner of operation of the above decision rule is interesting. Its decision indicates the node for which *a posteriori* probability of a set of classes attainable from it, multiplied by the respective probability of correct classification at the next stages of the recognition procedure, is the greatest. In other words, the decision at any interior node of a tree depends on the future to which this decision leads.

3.2. Locally optimal strategy (LOS)

Formally, a locally optimal strategy can be derived by minimizing the local criteria which denote probabilities of misclassification for particular nodes of a tree. Its recognition algorithm at the n^{th} stage is as follows:

$$\begin{gathered}\bar{\Psi}_{i_{n-1}}(x_{i_{n-1}})=i^{(k)}_{n-1}\\ \text{if}\quad \sum_{j\in\mathbf{M}_{i^{(k)}_{n-1}}}p_j(x_{i_{n-1}})=\max_l \sum_{j\in\mathbf{M}_{i^{(l)}_{n-1}}}p_j(x_{i_{n-1}}).\end{gathered}\tag{2}$$

The LOS strategy disregards the context and its decision rules are mutually independent.

Let us now assume that the learning set is known:

$$S=\{(x_1,j_1),(x_2,j_2),\ldots,(x_L,j_L)\},\quad x_i\in\mathbf{X},\ j_i\in\mathbf{M},\tag{3}$$

where x_i denotes the feature vector of the i^{th} learning pattern and j_i is its correct classification. For example, in a medical diagnosis task these may be the results of experienced physicians which we consider correct or the results of diagnoses using additional examinations which together with x constitute conclusive examinations, *i.e.* the results taken together unequivocally determine the patient's pathological state (disease). Additionally, let $S^{(i_{n-1})}$ denote a subset of learning patterns from classes belonging to the $\mathbf{M}_{(i_{n-1})}$ set.

Now, our purpose – assuming that the learning set (3) is given – is to derive fuzzy inference engine procedures for multistage recognition which will be related to LOS and GOS strategies and will in a sense imitate their ideas of local and global optimization. In the following sections we present two different methods, which in effect lead to four algorithms corresponding to LOS and GOS strategies.

4. Fuzzy Method 1 (FM1)

4.1. Classification algorithm

We shall now consider decision algorithms for a multistage diagnosis task using an inference engine that makes inferences on a fuzzy rule system. We assume that the form of the k^{th} $(k=1,2,\ldots,K)$ fuzzy if-then rule at the n^{th} stage $(n=1,2,\ldots,N)$ of the recognition procedure is as follows:

$$\textbf{IF}\; x_1^{(i_{n-1})}\;\textbf{IS}\; A_{1,k}\;\textbf{AND}\; x_2^{(i_{n-1})}\;\textbf{IS}\; A_{2,k}\;\textbf{AND}\ldots\textbf{AND}\; x_{d_n}^{(i_{n-1})}\;\textbf{IS}\; A_{d_n,k}\;\textbf{THEN}\; B_k^{(i_{n-1})}. \quad (4)$$

$A_{i,k}$ denote fuzzy sets (whose membership functions are designated by $\mu_{A_i,k}$) that correspond to the nature of particular observations (for simplicity we assume the sets to be triangular fuzzy numbers), whereas B is a discrete fuzzy set defined on the set of decision numbers $\mathbf{M}^{(i_{n-1})}$ determined by decision i_{n-1} made at the previous stage, with the μ_B membership function.

The Mamdani fuzzy inference system has been applied as a recognition algorithm [5, 6] (see Figure 2). In this system we use the minimum t-norm as the **AND** connection in premises, product operation as conjuctive implication interpretation in rules, the maximum t-conorm as the aggregation operation, and the maximum defuzzification method.

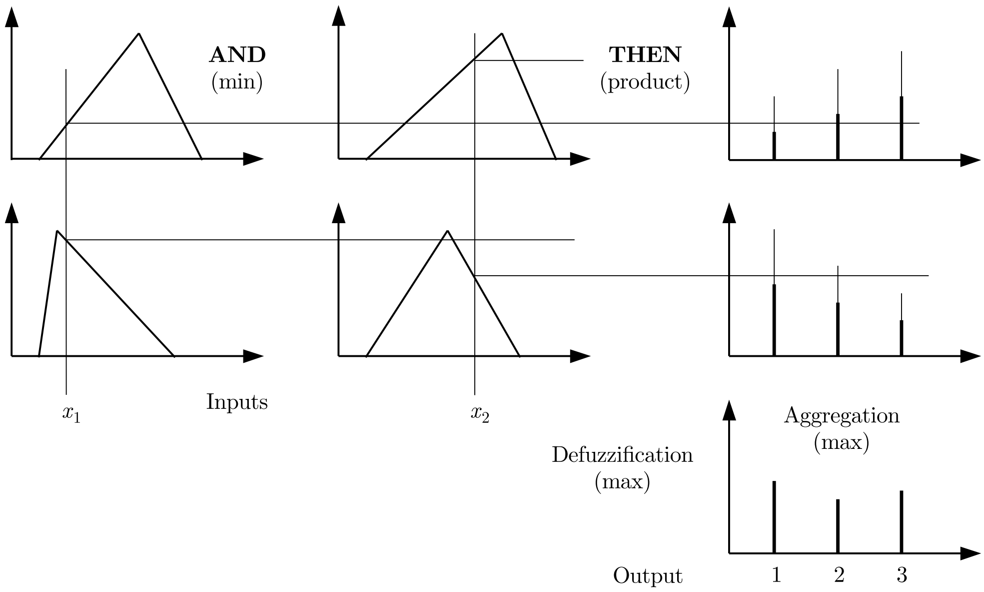

Figure 2. Mamdani inference system with discrete conclusions

4.2. Extraction of rules

Two methods are used to obtain a collection of fuzzy if-then rules (4) in the construction of a fuzzy system:

- from a human expert or based on domain knowledge,
- extraction of rules using numerical input-output data of the desired system.

One of the best known methods of generating rule from the given training patterns (3), is the method proposed by Wang and Mendel [8] (WM method). This method, developed for multistage recognition, leads to the following procedure for the $i_{n-1}{}^{\text{th}}$ node of a decision-tree:

1. divide the spaces of features $x^{(i_{n-1})}$ into fuzzy regions. In the following example we use triangular fuzzy sets with 3 and 5 partitions as depicted in Figure 3. Usually, these fuzzy sets correspond to linguistic "values" of features, which state space $\bar{\mathbf{X}}^{(i_{n-1})}$;

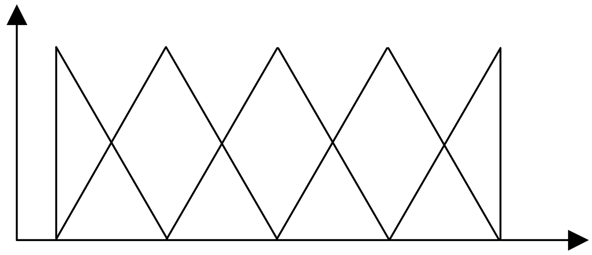

Figure 3. An example of antecedent fuzzy sets with 5 partitions

2. for each example generate a fuzzy rule with premises corresponding to the fuzzy regions with the highest membership grade of the appropriate feature;
3. find rules with the same premises and aggregate them into one rule;
4. determine the fuzzy conclusion of the rule.

In multistage recognition, however, we can determine the fuzzy conclusion of the rules in a different manner. Let:

$$B_k^{(i_{n-1})} = \{i_{n-1}^{(1)}/\mu_k(i_{n-1}^{(1)}), \ldots, i_{n-1}^{(n')}/\mu_k(i_{n-1}^{(n')})\}, \tag{5}$$

where $i_{n-1}^{(j)} \in \mathbf{M}^{(i_{n-1})}$, be the discrete fuzzy set which denotes the conclusion of the k^{th} rule of the system. We propose two different algorithms for determining its membership function, μ. It leads to two different fuzzy reasoning systems, which correspond to the LOS (2) and GOS (1) strategies.

Algorithm 1

$$\mu_k^{(1)}(i_{n-1}^{(j)}) = \frac{K(i_{n-1}^{(j)})}{\sum\limits_j K(i_{n-1}^{(j)})}, \tag{6}$$

where $K(i_{n-1}^{(j)})$ denotes the number of learning patterns fulfiling the k^{th} rule for which the class number belongs to the $\mathbf{M}_{i_{n-1}^{(j)}}$ set.

Algorithm 2

To determine the membership function of (5), the following formula is proposed:

$$\mu_k^{(2)}(i_{n-1}^{(j)}) = \mu_k^{(1)}(i_{n-1}^{(j)}) Pc(i_{n-1}^{(j)}), \tag{7}$$

where $Pc(i_{n-1}^{(j)})$ is the frequency of correct classification (empirical probability) at the next stages if at the n^{th} stage decision $i_{n-1}^{(j)}$ is made (for the sub-tree with the $i_{n-1}^{(j)}$ node as the root node).

Algorithm 1 – similarly as LOS – disregards the context of the decision procedure in a multistage process. Algorithm 2 corresponds to the GOS strategy in the probabilistic approach.

It can easily be observed that Algorithm 2 (and the GOS strategy) can be explicitly determined for particular nodes of a tree starting from the terminal level of the tree, by alternately determining the appropriate frequencies of correct classification.

5. Fuzzy Method 2 (FM2)

In the second method, on the basis of learning set (3), first we find for each nonterminal node a fuzzy relation between the fuzzified feature space and the class numbers set, as a solution of the appropriate optimization problem. Then this relation, expressed as the so-called expert matrix, can be used to make a decision in a multistage recognition procedure – similarly as previously – in twofold manner.

More precisely, this method leads to the following steps for the $i_{n-1}{}^{\text{th}}$ node of a decision-tree:

- as step 1 of FM1;
- calculate observation matrix $O^{(i_{n-1})}$ (a fuzzy relation between feature space $\bar{\mathbf{X}}^{(i_{n-1})}$ and learning subset $S^{(i_{n-1})}$) – the i^{th} matrix row contains grades of membership of features $x^{(i_{n-1})}$ of the i^{th} learning pattern from $S^{(i_{n-1})}$ to the fuzzy sets created at the previous step;
- determine decision matrix $D^{(i_{n-1})}$ (a relation between learning subset $S^{(i_{n-1})}$ and the set of decision numbers $\mathbf{M}^{(i_{n-1})}$) – its i^{th} row contains 0's and the figure one at the position corresponding to decision number $i_{n-1}^{(j)}$, for which the class number of the i^{th} learning pattern from $S^{(i_{n-1})}$ belongs to the $\mathbf{M}_{i_{n-1}^{(j)}}$ set;
- find matrix $E^{(i_{n-1})}$, so as to minimize the following criterion:

$$\rho(O^{(i_{n-1})} \circ E^{(i_{n-1})}, D^{(i_{n-1})}). \tag{8}$$

Operator $\circ$ denotes max$-t$-norm composition of relations, *i.e.* multiplication of matrices O and E with $\cdot$, $+$ operators replaced by t-norm and max [7]. Criterion $\rho(A,B)$ evaluates the difference between matrices A and B, *i.e.* $\rho(A,B) \geq 0$ and $\rho(A,B) = 0$ iff $A = B$. In the following example, we adopt:

$$\rho(A,B) = \sum_{i,j} (a_{ij} - b_{ij})^2 \tag{9}$$

and apply the genetic algorithm as the method of minimization (7).

Matrix $E^{(i_{n-1})}$, which will be called an expert matrix, is a fuzzy relation between decision set $\mathbf{M}^{(i_{n-1})}$ and fuzzified feature space $\bar{\mathbf{X}}^{(i_{n-1})}$. Its elements (grades of membership) represent the intensity of fuzzy features (*e.g.* symptoms in medical diagnosis) for each class (*e.g.* disease). The manner of using the expert matrix to make a decision at node i_{n-1} is obvious. Namely, from the given feature vector, $x^{(i_{n-1})}$, of the pattern to be recognized, we must first determine the observation row-matrix, $O_0^{(i_{n-1})}$ (like in step 2 of the method), and then calculate the decision row-matrix, $D_0^{(i_{n-1})}$, as a max$-t$-norm composition of relations (see step 3), *viz.*:

$$O_0^{(i_{n-1})} \circ E^{(i_{n-1})} = D_0^{(i_{n-1})}. \tag{10}$$

As a decision, we choose the number from the $\mathbf{M}^{(i_{n-1})}$ set which corresponds to the maximum value of elements of row-matrix $D_0^{(i_{n-1})}$ (Algorithm 3) or row-matrix

$D_0^{(i_{n-1})} Pc(i_{n-1})$ (Algorithm 4), where $Pc(i_{n-1})$ denotes a diagonal matrix of empirical probabilities defined in (7). Obviously, Algorithm 3 corresponds to LOS of the probabilistic approach and Algorithm 1 of FM1, whereas Algorithm 4 is related to the GOS strategy and Algorithm 2.

In the following section we present results of comparative analysis of the proposed algorithms using a sufficiently rich set of real-life data concerning multistage diagnosis of acute renal failure.

6. An example – diagnosis of acute renal failure (ARF)

6.1. Material and methods

ARF is a syndrome of clinical symptoms caused by an adverse action of factors of the urinary tracts. A quick and proper diagnosis of ARF is essential for an appropriate therapy and prognosis. Unfortunately, the cause of ARF, particularly in the initial phase of the disease, is very often difficult to establish. Therefore, a need for computer-aided diagnosis process is evident.

The diagnosis of ARF as a pattern recognition task includes the following 11 classes (etiologic types of ARF) [9]: 1 – toxicosis, 2 – the nephrotic syndrome, 3 – sepsis, 4 – circulatory failure, 5 – others (prerenal), 6 – acute glomerulonephritis, 7 – the uremic-haemolytic syndrome, 8 – renal vain thrombosis, 9 – the andrenogenital syndrome, 10 – others (intrarenal), 11 – postrenal failure.

The classes have been organized by a team of physician into a two-stage classifier depicted in Figure 4. Its decision logic (decision tree) is a deliberate one, since from the clinical point of view the most important step is to include the cause of the disease into one of the three main categories of ARF, as a partial diagnosis suggests the choice of an appropriate therapy.

At the Department of Pediatric Nephrology of the Wroclaw Medical Academy, a set of case records of children suffering from ARF was collected, which constitutes the learning set (3). Each case record contains administrative data, values of the 34 clinical features and a firm diagnosis. Most of the diagnoses were made during hospitalization according to generally accepted criteria.

Table 1 presents the clinical features collected in case records. We have chosen to enter into the computer the data which the registrar obtained when he first saw the case. This is important since clinical data change and a case which may be puzzling when first seen, may become obvious the following morning.

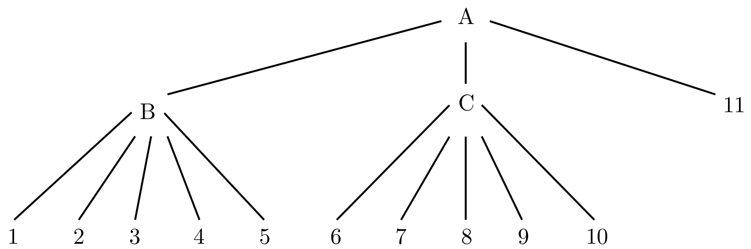

Figure 4. Decision tree for the medical problem in question (A – acute renal failure, B – acute prerenal failure, C – acute intrarenal failure)

Table 1. Clinical features consiered

GENERAL
Age (1), Weight (2)
PHYSICAL EXAMINATIONS
Blood pressure – systolic (3), – diastolic (4), Pulse (5), Body temperature (6),
Urine in bladder (7)
LABORATORY EXAMINATIONS
Sedimentation rate – after 1 hour (8), – after 2 hours (9)
GASOMETRIC EXAMINATIONS OF THE BLOOD
p O_2 (10), p CO_2 (11), pH (12), Stand. HCO_3 (13),
Actual HCO_3 (14), BE (15)
MORPHOLOGY OF THE BLOOD
Leucocytes (16), Reticulocytes (17), Trombocytes (18), Erythrocytes (19), Hemoglobin (20)
SERUM
Urine level (21), Creatinine level (22), Uric acid level (23), Total protein level (24)
SERUM IONOGRAM
Na^+ (25), K^+ (26), Ca^+ (27)
URINE
24-hours amount (28), Specific weight (29), Protein (30), Leucocytes (31), Erythrocytes (32),
Cylinders (33)

First, we have selected the best feature subset for each non-terminal node (34 features were avaliable for selection) using the Kolmogorov criterion [10]. Results are listed in Table 2. Each row presents a list of 7 features, selected and ordered according to their Kolmogorov criterion value.

Table 2. The results of feature selection

Node	List of the ranged feature numbers						
A	8,	14,	17,	30,	10,	20,	28
B	29,	6,	24,	5,	15,	22,	32
C	30,	23,	17,	11,	16,	24,	29

At each non-terminal node, the features listed in Table 2 were used successively, from the best single one to the set of all 7 features. For fuzzy methods, we also change the number of partitions of feature spaces. In order to find expert matrices for non-terminal nodes in FM2 the genetic algorithm was applied, which is a popular method in optimization capable of improving the search procedure.

The genetic algorithm proceeded as follows [11, 12]:

- *the coding method*: the values of elements of matrix E were directly coded to the chromosome;
- *the fitness function*: it was defined as follows:

$$\mathrm{Fit} = Q - \rho(A,B), \tag{11}$$

  where ρ is as in (9) and Q is a suitably selected positive constant;
- *initialization*: the initial population of chromosomes with which the search begins was generated randomly. The size of population – after trials – was set at 40;
- *reproduction*: a roulette wheel with elitism;

- *crossover and mutation*: a two-point crossover was used, while the probability of mutation was 0.05;
- *the stop procedure*: the evolution process was terminated after 1000 generations. In fact, the fitness value usually converged within this value. Figure 5 shows the fitness change against the generation number in one example.

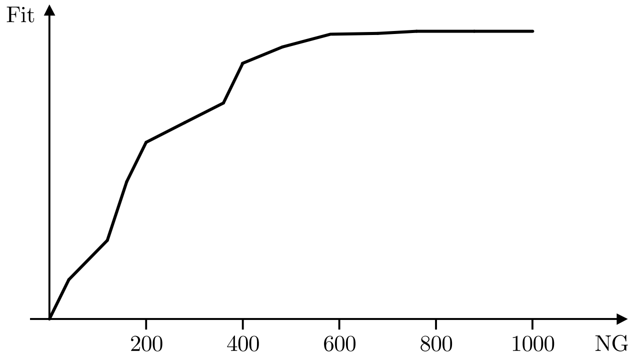

Figure 5. Fitness change versus number of generation (NG)

6.2. Results

In order to study the performance of the proposed recognition concepts and evaluate their usefulness for computer-aided diagnosis of ARF, computer experiments were made using the leave-one-out method [10]. This method does not require dividing the data set into learning and testing sets. The leave-one-out method leaves one pattern out of the learning dataset and uses it as a test pattern every time. The procedure continues until each pattern is tested.

Results for Algorithm 1 (A1) and Algorithm 2 (A2) of FM1, as well as for Algorithm 3 (A3) and Algorithm 4 (A4) of FM2 are presented in Table 3. Additionally, results for GOS and LOS probabilistic algorithms are also given.

Table 3. The results of classification accuracy in percent

Algorithm	The number of features per node						
	1	2	3	4	5	6	7
LOS	48.9	56.7	62.2	66.7	74.4	80.2	78.3
GOS	47.5	56.1	64.5	69.3	77.9	87.8	84.2
A1 (3 partitions)	45.4	54.5	61.9	70.4	78.3	75.5	70.2
A1 (5 partitions)	47.1	55.8	62.9	73.8	80.2	77.3	72.1
A2 (3 partitions)	47.1	57.2	63.5	71.1	81.3	79.1	70.8
A2 (5 partitions)	48.0	58.2	65.3	74.9	84.1	80.5	73.9
A3 (3 partitions)	49.3	55.4	60.2	68.4	76.4	77.3	72.1
A3 (5 partitions)	49.8	57.8	63.5	70.1	79.2	80.3	79.5
A4 (3 partitions)	48.6	57.2	62.5	71.3	78.8	80.6	80.1
A4 (5 partitions)	50.3	59.8	65.7	73.3	82.0	81.8	80.7

These results imply the following conclusions:

1. There occurs a common effect within each algorithm group: algorithms that disregard the context of the decision procedure in a multistage process (LOS, A1 and A3) are always worse than those that treat the multistage procedure as a compound decision process (GOS, A2 and A4). This confirms the effectiveness and usefulness of the concept and algorithm construction principles presented above for the purposes of multistage diagnosis.
2. Fuzzy algorithms with 5 partitions of feature spaces are better than algorithms with 3 partitions.
3. The difference between probabilistic algorithms and fuzzy methods is insignificant.

7. Final remarks

In this paper we have focused our attention on the fuzzy approach to multistage pattern recognition and the application of the elaborated methods to the diagnosis of acute renal failure. In order to study the performance of the proposed recognition concepts and evaluate their usefulness for computer-aided diagnosis, computer experiments were made using real data. The objective of our experiements was to measure the quality of the tested algorithms, defined by the frequency of correct decisions.

The comparative analysis presented above for multistage diagnosis is also of experimental nature. The algorithm-ranking outcome cannot be treated as having the ultimate character a law; it has been obtained for specific data within a specific diagnostic task. However, although the outcome may be different for other tasks, the presented research may nevertheless suggest some perspectives for practical applications and – as it seems – has proven the proposed concepts to be correct.

References

[1] Kurzyński M 2002 *Proc. 1st Int. Conf. on Fuzzy Systems and Knowledge Discovery*, Singapore (CD ROM)
[2] Kurzyński M 2002 *Proc. IEEE EMBS Conf.*, Vienna, pp. 762–766
[3] Kurzyński M 1988 *Pattern Recognition* **21** 355
[4] Kurzyński M 2003 *Medical Data Analysis*, Springer Verlag, pp. 99–108
[5] Pedrycz W 1993 *Fuzzy Control and Fuzzy Systems*, John Willey & Sons, New York
[6] Wang L-X 1998 *A Course in Fuzzy Systems and Control*, Prentice-Hall, New York
[7] Czogala E and Leski J 2000 *Fuzzy and Neuro-fuzzy Intelligent Systems*, Physica-Verlag, New York
[8] Wang L-X and Mendel J M 1992 *IEEE Trans. SMC* **22** 1414
[9] James J A 1996 *Renal Disease in Childhood*, The C.V. Mosby Comp., London
[10] Devroye L, Gyorfi P and Lugossi G 1996 *A Probabilistic Theory of Pattern Recognition*, Springer Verlag, New York
[11] Mitchell T M 1996 *An Introduction to Genetic Algorithms*, MIT Cambridge, MA
[12] Goldberg D E 1989 *Genetic Algorithms in Seach, Optimization and Machine Learning*, Addison-Wesley, New York