On Designing of Flexible Neuro-Fuzzy Systems for Classification

KRZYSZTOF CPAŁKA, LESZEK RUTKOWSKI
Department of Computer Engineering
Częstochowa University of Technology
POLAND

Abstract:- In the paper we propose a new method for designing flexible neuro-fuzzy systems applied to classification. The systems are characterized by weighted triangular norms describing the importance of linguistic variables and rules. An algorithm for complexity reduction of such systems is developed and simulation results are presented.

Key-Words:- neuro-fuzzy systems, triangular norms, algorithm for reduction of neuro-fuzzy systems

1. Introduction

In literature various neuro-fuzzy systems have been developed. Some of them are known in literature under short names such as ANFIS [8], ANNBFIS [5], DENFIS [9], FALCON [11], GARIC [2], NEFCLASS [12], NEFPROX [12], [13], SANFIS [18] and others. The original concept of flexible neuro-fuzzy systems have been proposed and studied in [3], [4], [14], [15], [16]. It is well known that neuro-fuzzy systems combine the natural language description of fuzzy systems and the learning properties of neural networks. In the paper we propose a new method for designing flexible neuro-fuzzy systems applied to classification. The systems are characterized by weighted triangular norms describing the importance of linguistic variables and rules. An algorithm for complexity reduction of such systems will be developed. In subsequent stages of the algorithm we reduce number of discretization points, number of inputs, number of rules and number of antecedents. The algorithm will be tested using well known benchmarks.

2. Mamdani and Logical Type Neuro-Fuzzy Systems

In this paper we consider multi-input, single-output neuro-fuzzy system mapping $\mathbf{X} \to \mathbf{Y}$, where $\mathbf{X} \subset \mathbf{R}^n$ and $\mathbf{Y} \subset \mathbf{R}$. The fuzzifier performs a mapping from the observed crisp input space $\mathbf{X} \subset \mathbf{R}^n$ to the fuzzy sets defined in $\mathbf{X}$. The most commonly used fuzzifier is the singleton fuzzifier which maps $\bar{\mathbf{x}} = [\bar{x}_1, \ldots, \bar{x}_n] \in \mathbf{X}$ into a fuzzy set $A' \subseteq \mathbf{X}$ characterized by the membership function

$$\mu_{A'}(\mathbf{x}) = \begin{cases} 1 & \text{if} \quad \mathbf{x} = \bar{\mathbf{x}} \\ 0 & \text{if} \quad \mathbf{x} \neq \bar{\mathbf{x}} \end{cases} \tag{1}$$

The fuzzy rule base consists of a collection of N fuzzy IF-THEN rules in the form

$$R^{(k)} : \text{IF}\, \mathbf{x}\, \text{is}\, \mathbf{A}^k \text{ THEN } y \,\text{is}\, B^k \tag{2}$$

where $\mathbf{x} = [x_1, \ldots, x_n] \in \mathbf{X}$, $y \in \mathbf{Y}$, $A_1^k, A_2^k, \ldots, A_n^k$ are fuzzy sets characterized by membership functions $\mu_{A_i^k}(x_i)$, whereas B^k are fuzzy sets characterized by membership functions $\mu_{B^k}(y)$, respectively, $k = 1, \ldots, N$.

The fuzzy inference determines a mapping from the fuzzy sets in the input space $\mathbf{X}$ to the fuzzy sets in the output space $\mathbf{Y}$. Each of N rules (2) determines a fuzzy set $\bar{B}^k \subset \mathbf{Y}$ given by the compositional rule of inference

$$\bar{B}^k = A' \circ \left(\mathbf{A}^k \to B^k\right) \tag{3}$$

where $\mathbf{A}^k = A_1^k \times A_2^k \times \ldots \times A_n^k$. Fuzzy sets $\bar{B}^k$, according to the formula (3), are characterized by membership functions expressed by the *sup-star* composition:

$$\mu_{\bar{B}^k}(y) = \sup_{\mathbf{x} \in \mathbf{X}} \left\{ T\left\{ \mu_{A'}(\mathbf{x}), \mu_{A_1^k \times \ldots \times A_n^k \to B^k}(\mathbf{x}, y) \right\} \right\} \tag{4}$$

where T can be any operator in the class of t-norms. It is easily seen that for a crisp input $\bar{\mathbf{x}} \in \mathbf{X}$, i.e. a singleton fuzzifier (1), formula (4) becomes

$$\begin{aligned} \mu_{\bar{B}^k}(y) &= \mu_{A_1^k \times \ldots \times A_n^k \to B^k}(\bar{\mathbf{x}}, y) \\ &= \mu_{\mathbf{A}^k \to B^k}(\bar{\mathbf{x}}, y) \\ &= I\left(\mu_{\mathbf{A}^k}(\bar{\mathbf{x}}), \mu_{B^k}(y)\right) \end{aligned} \tag{5}$$

where $I(\cdot)$ is an "engineering implication" (Mamdani approach) or fuzzy implication [6]. The aggregation operator, applied in order to obtain the fuzzy set B' based on fuzzy sets $\overline{B}^k$, is the t-norm or t-konorm operator, depending on the type of fuzzy implication.

The defuzzifier performs a mapping from a fuzzy set B' to a crisp point $\bar{y}$ in $\mathbf{Y} \subset \mathbf{R}$. The COA (centre of area) method is defined by the following formula

$$\bar{y} = \frac{\int_{\mathbf{Y}} y \mu_{B'}(y) dy}{\int_{\mathbf{Y}} \mu_{B'}(y) dy} \tag{6}$$

or by

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \mu_{B'}\left(\bar{y}^r\right)}{\sum_{r=1}^{N} \mu_{B'}\left(\bar{y}^r\right)} \tag{7}$$

in the discrete form, where $\bar{y}^r$ denotes centres of the membership functions $\mu_{B^r}(y)$, i.e. for $r = 1, \ldots, N$

$$\mu_{B^r}\left(\bar{y}^r\right) = \max_{y \in \mathbf{Y}} \left\{\mu_{B^r}(y)\right\} \tag{8}$$

In [14-16] we proposed a general architecture of neuro-fuzzy structures. It includes both the Mamdani and logical type of inference

$$\bar{y} = f(\bar{\mathbf{x}}) = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \operatorname{agr}_r\left(\bar{\mathbf{x}}, \bar{y}^r\right)}{\sum_{r=1}^{N} \operatorname{agr}_r\left(\bar{\mathbf{x}}, \bar{y}^r\right)} \tag{9}$$

where

$$\operatorname{agr}_r\left(\bar{\mathbf{x}}, \bar{y}^r\right) = \begin{cases} \overset{N}{\underset{k=1}{S}} \left\{I_{k,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right)\right\}, \\ \text{for Mamdani approach} \\ \overset{N}{\underset{k=1}{T}} \left\{I_{k,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right)\right\}, \\ \text{for logical approach} \end{cases} \tag{10}$$

and

$$I_{k,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right) = \begin{cases} T\left\{\tau_k(\bar{\mathbf{x}}), \mu_{B^k}\left(\bar{y}^r\right)\right\}, \\ \text{for Mamdani approach} \\ I_{fuzzy}\left(\tau_k(\bar{\mathbf{x}}), \mu_{B^k}\left(\bar{y}^r\right)\right), \\ \text{for logical approach} \end{cases} \tag{11}$$

The firing strength of rules is given by

$$\tau_k(\bar{\mathbf{x}}) = \overset{n}{\underset{i=1}{T}} \left\{\mu_{A_i^k}(\bar{x}_i)\right\} \tag{12}$$

In this paper, starting with a description (9)-(12), we develop a new method for designing and reduction of neuro-fuzzy systems. The method is based on the concept of the weighted triangular norms [16]. In subsequent stages we reduce number of discretization points, number of inputs, number of rules and number of antecedents. The method is tested using well known benchmarks. It should be emphasized that in this paper we do not assume that the number of terms in formula (7) is equal to the number of rules N. We allow to discretize the integrals in formula (6) using R points. In the simulations we investigate various neuro-fuzzy systems for different values of N and R. To our best knowledge such problems have not been studied yet in the literature.

3. Flexibility Parameters in Neuro-Fuzzy Systems

3.1. Weighted triangular norms

In [16] we introduced a new concept of the weighted t-norm defined by

$$T^*\left\{a_1, \ldots, a_n; w_1^\tau, \ldots, w_n^\tau\right\} = \overset{n}{\underset{i=1}{T}} \left\{1 - w_i^\tau \left(1 - a_i\right)\right\} \tag{13}$$

to connect the antecedents in each rule, $k = 1, \ldots, N$, and the weighted t-norm and t-conorm

$$T^*\left\{a_1, \ldots, a_N; w_1^{\text{agr}}, \ldots, w_N^{\text{agr}}\right\} = \overset{N}{\underset{k=1}{T}} \left\{1 - w_k^{\text{agr}} \left(1 - a_k\right)\right\} \tag{14}$$

and

$$S^*\left\{a_1, \ldots, a_N; w_1^{\text{agr}}, \ldots, w_N^{\text{agr}}\right\} = \overset{N}{\underset{k=1}{S}} \left\{w_k^{\text{agr}} a_k\right\} \tag{15}$$

to aggregate the individual rules in the logical and Mamdani models, respectively. It is easily seen that formula (13) can be applied to the evaluation of an importance of input linguistic values, and the weighted t-norm (14) or t-conorm (15) to a selection of important rules. The results can be depicted in the form of diagrams. In Fig. 1 we show an example of a diagram for a fuzzy system having four rules ($N = 4$) and two inputs ($n = 2$) described by

$$
\begin{aligned}
R^1 &: \begin{bmatrix} \text{IF } x_1 \text{ is } A_1^1\left(w_{1,1}^\tau\right) \text{AND } x_2 \text{ is } A_2^1\left(w_{2,1}^\tau\right) \\ \text{THEN } y \text{ is } B^1 \end{bmatrix} w_1^{\text{agr}} \\
R^2 &: \begin{bmatrix} \text{IF } x_1 \text{ is } A_1^2\left(w_{1,2}^\tau\right) \text{AND } x_2 \text{ is } A_2^2\left(w_{2,2}^\tau\right) \\ \text{THEN } y \text{ is } B^2 \end{bmatrix} w_2^{\text{agr}} \\
R^3 &: \begin{bmatrix} \text{IF } x_1 \text{ is } A_1^3\left(w_{1,3}^\tau\right) \text{AND } x_2 \text{ is } A_2^3\left(w_{2,3}^\tau\right) \\ \text{THEN } y \text{ is } B^3 \end{bmatrix} w_3^{\text{agr}} \\
R^4 &: \begin{bmatrix} \text{IF } x_1 \text{ is } A_1^4\left(w_{1,4}^\tau\right) \text{AND } x_2 \text{ is } A_2^4\left(w_{2,4}^\tau\right) \\ \text{THEN } y \text{ is } B^4 \end{bmatrix} w_4^{\text{agr}}
\end{aligned} \tag{16}
$$

Observe that the third rule is "weaker" than the others and the linguistic value A_2^1 corresponds to a low value of $w_{2,1}^\tau$.

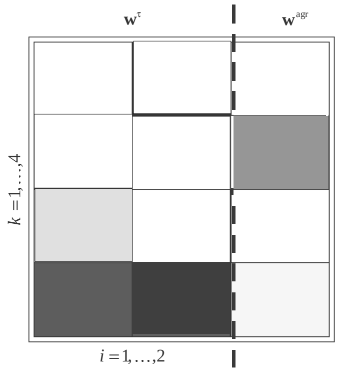

Fig. 1. Exemplary weights representation in a fuzzy system with four rules and two inputs (dark areas correspond to low values of weights and vice versa)

3.2. Soft triangular norms

In this section we recall a concept of soft fuzzy norms proposed by Yager and Filev [19]. Let $a_1,\ldots,a_n$ be numbers in the unit interval that are to be aggregated. The soft version of triangular norms suggested by Yager and Filev is defined by

$$
\tilde{T}\{\mathbf{a};\alpha\} = (1-\alpha)\frac{1}{n}\sum_{i=1}^{n} a_i + \alpha \mathop{T}_{i=1}^{n}\{a_i\} \tag{17}
$$

and

$$
\tilde{S}\{\mathbf{a};\alpha\} = (1-\alpha)\frac{1}{n}\sum_{i=1}^{n} a_i + \alpha \mathop{S}_{i=1}^{n}\{a_i\} \tag{18}
$$

where $\alpha \in [0,1]$. They allow to balance between the arithmetic average aggregator and the triangular norm aggregator depending on parameter α.

4. New Flexible Fuzzy Systems

Neuro-fuzzy architectures developed so far in the literature are based on the discretization of formula (6) with the assumption that number of terms in a corresponding formula (7) is equal to the number of rules N. In this paper we relax that assumption and replace formula (7) by

$$
\bar{y} = \frac{\sum_{r=1}^{R} \bar{y}^r \cdot \mu_{B'}\left(\bar{y}^r\right)}{\sum_{r=1}^{R} \mu_{B'}\left(\bar{y}^r\right)} \tag{19}
$$

where $R \geq 1$. For further investigations we choose neuro-fuzzy systems of a logical type with an S-implication used in formula (5). Moreover, we incorporate flexibility parameters, presented in Section 3, into construction new neuro-fuzzy systems. These parameters have the following interpretation:

1) weights in antecedents of the rules $w_{i,k}^\tau \in [0,1]$, $i = 1,\ldots,n$, $k = 1,\ldots,N$,
2) weights in aggregation of the rules $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots,N$,
3) soft strength of firing controlled by parameter α_k^τ, $k = 1,\ldots,N$,
4) soft implication controlled by parameter α_k^I, $k = 1,\ldots,N$,
5) soft aggregation of rules controlled by parameter α^{agr}.

In view of above assumptions, we derive a flexible neuro-fuzzy system given by

$$
\bar{y} = \frac{\sum_{r=1}^{R} \bar{y}^r \cdot \text{agr}_r\left(\bar{\mathbf{x}}, \bar{y}^r\right)}{\sum_{r=1}^{R} \text{agr}_r\left(\bar{\mathbf{x}}, \bar{y}^r\right)} \tag{20}
$$

where

$$
\tau_k(\bar{\mathbf{x}}) = \begin{pmatrix} \left(1-\alpha_k^\tau\right)\text{avg}\left(\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\right) + \\ +\alpha_k^\tau T^* \begin{Bmatrix} \mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ; \\ w_{1,k}^\tau,\ldots,w_{n,k}^\tau \end{Bmatrix} \end{pmatrix} \tag{21}
$$

$$
I_{k,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right) = \begin{pmatrix} \left(1-\alpha_k^I\right)\text{avg}\left(1-\tau_k(\bar{\mathbf{x}}), \mu_{B^k}\left(\bar{y}^r\right)\right) + \\ +\alpha_k^I S\left\{1-\tau_k(\bar{\mathbf{x}}), \mu_{B^k}\left(\bar{y}^r\right)\right\} \end{pmatrix} \tag{22}
$$

$$
\text{agr}_r\left(\bar{\mathbf{x}}, \bar{y}^r\right) = \begin{pmatrix} \left(1-\alpha^{\text{agr}}\right)\text{avg}\left(I_{1,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right)\right) + \\ +\alpha^{\text{agr}} T^* \begin{Bmatrix} I_{1,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right)\ ; \\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}} \end{Bmatrix} \end{pmatrix} \tag{23}
$$

The general architecture of the above system is depicted in Fig. 2.

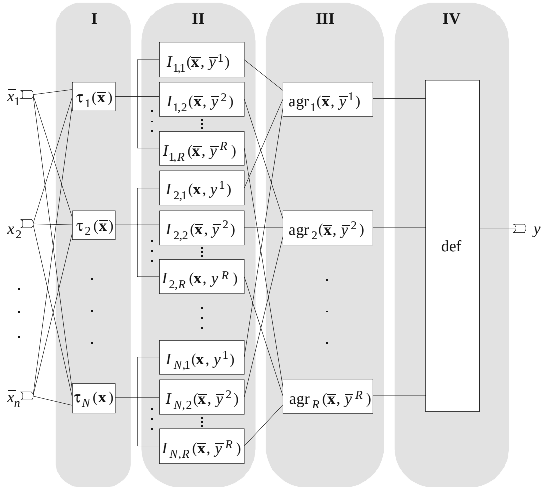

Fig. 2. The scheme of neuro-fuzzy system

It is easily seen that system (20)-(23) contains $N(3n+5)+R+1$ or $3p+5N+R+1$ parameters to be determined in the process of learning, where p is a number of antecedents.

5. Algorithm of Reduction of Neuro-Fuzzy Systems

In this section we develop an algorithm of reduction of neuro-fuzzy systems. The algorithm is based on analysis of weights in antecedents of the rules $w_{i,k}^{\tau} \in [0,1]$, $i=1,\ldots,n$, $k=1,\ldots,N$, and weights in aggregation of the rules $w_k^{\mathrm{agr}} \in [0,1]$, $k=1,\ldots,N$. The flowchart of the algorithm is depicted in Fig. 3.

The flowchart in Fig. 3 comprises 4-parts. First, we determine performance of the initial system (before the reduction process); for example, in a case of the classification we determine a percentage of mistakes of the system. The weights $w_i^x \in [0,1]$, $i=1,\ldots,n$ are calculated using

$$w_i^x = \frac{1}{N}\sum_{k=1}^{N} w_{i,k}^{\tau} \tag{24}$$

In subsequent stages we reduce number of discretization points, number of inputs, number of rules and number of antecedents.

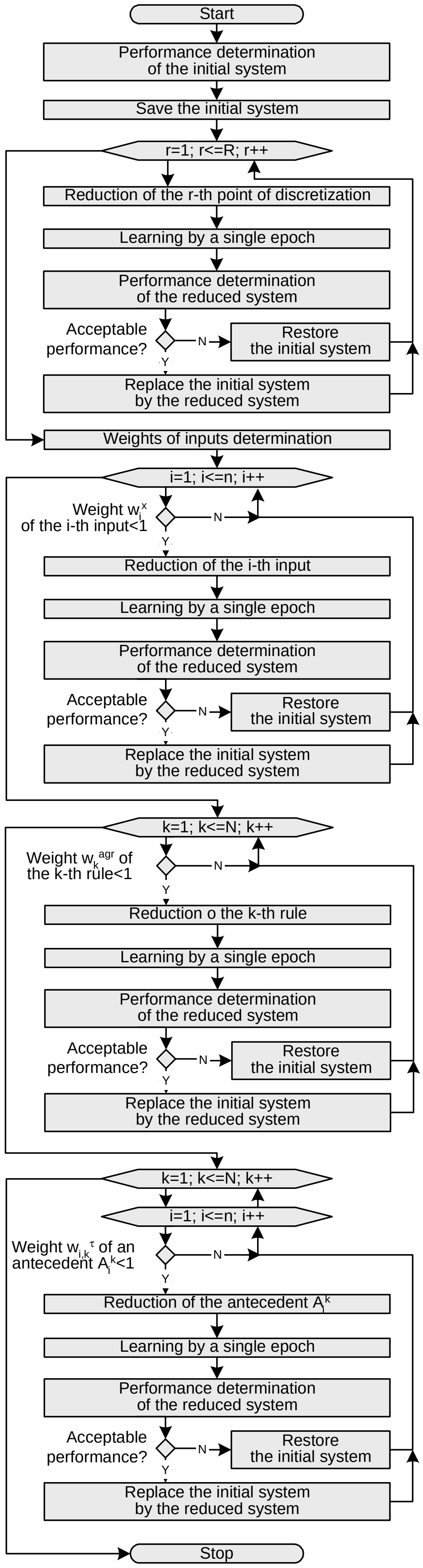

Fig. 3. The algorithm of reduction of neuro-fuzzy systems

6. Simulation Results

The neuro-fuzzy system is simulated on Glass Identification problem and Wisconsin Breast Cancer problem [17].

6.1. Glass Identification Problem

The Glass Identification problem contains 214 instances and each instance is described by nine attributes (RI: refractive index, Na: sodium, Mg: magnesium, Al: aluminium, Si: silicon, K: potassium, Ca: calcium, Ba: barium, Fe: iron). All attributes are continuous. There are two classes: the window glass and the non-window glass. In our experiments, all sets are divided into a learning sequence (171 sets) and a testing sequence (43 sets). The study of the classification of the types of glass was motivated by criminological investigation. At the scene of the crime, the glass left can be used as evidence if it is correctly identified.

The experimental results for the Glass Identification problem are depicted in tables 1, 2, 3, 4, 5 and figures 4, 5, 6, 7, 8. In Table 1 we show the percentage of mistakes in the learning and testing sequences before and after reduction, e.g. for $N = 3$ and $R = 2$ we have 2.92%/2.34% for the learning sequence before and after reduction and 0.00%/0.00% for the testing sequence before and after reduction. In Table 2 we present number of inputs, rules, points of discretization, number of antecedents and number of parameters before and after reduction. In Table 3 we show degree of learning time reduction [%] and degree of learning time reduction per a single parameter [%] for a reduced system. In Table 4 we present reduced inputs and antecedents. In Table 5 we depict percentage of neuro-fuzzy systems having a particular input (attribute) after the reduction process and percentage of inputs (attributes) corresponding to a particular neuro-fuzzy system after the reduction process.

Table 1. Simulation results

GLASS IDENTIFICATION PROBLEM				
R	***N***			
	1	2	3	4
2	6.43%/**5.85%**	2.34%/2.34%	2.92%/**2.34%**	2.34%/2.34%
	9.30%/9.30%	2.33%/2.33%	0.00%/0.00%	0.00%/0.00%
3	5.85%/5.85%	2.92%/2.92%	2.34%/2.34%	2.34%/2.34%
	9.30%/9.30%	2.33%/2.33%	0.00%/0.00%	0.00%/0.00%
4	5.85%/5.85%	2.34%/2.34%	2.34%/2.34%	2.34%/**1.75%**
	9.30%/9.30%	2.33%/2.33%	0.00%/0.00%	0.00%/0.00%

Table 2. Simulation results

GLASS IDENTIFICATION PROBLEM				
R	***N***			
	1	2	3	4
2	9/1/2/9/35 **2**/1/2/**2**/**14**	9/2/2/18/67 **4**/2/2/**6**/**31**	9/3/2/27/99 **5**/3/2/**12**/**54**	9/4/2/36/131 **4**/**3**/2/**11**/**51**
3	9/1/3/9/36 **2**/1/3/**2**/**15**	9/2/3/18/68 **5**/2/3/**9**/**41**	9/3/3/27/100 **5**/3/3/**10**/**49**	9/4/3/36/132 **6**/4/3/**19**/**81**
4	9/1/4/9/37 **2**/1/4/**2**/**16**	9/2/4/18/69 **5**/2/4/**6**/**33**	9/3/4/27/101 **4**/3/4/**10**/**50**	9/4/4/36/133 **7**/4/**3**/**14**/**66**

Table 3. Simulation results

GLASS IDENTIFICATION PROBLEM				
R	***N***			
	1	2	3	4/
2	66%	63%	53%	68%
	28%	27%	19%	24%
3	50%	39%	53%	41%
	-5%	5%	6%	9%
4	45%	47%	47%	55%
	-12%	0%	0%	15%

Table 4. Simulation results

GLASS IDENTIFICATION PROBLEM				
R	***N***			
	1	2	3	4
2	$\overline{x}_1$, $\overline{x}_2$, $\overline{x}_4$, $\overline{x}_5$, $\overline{x}_6$, $\overline{x}_7$, $\overline{x}_9$	$\overline{x}_1$, $\overline{x}_2$, $\overline{x}_4$, $\overline{x}_5$, $\overline{x}_6$, A_7^1, A_8^2	$\overline{x}_1$, $\overline{x}_2$, $\overline{x}_5$, $\overline{x}_6$, A_4^1, A_9^1, A_4^3	$\overline{x}_1$, $\overline{x}_2$, $\overline{x}_5$, $\overline{x}_6$, $\overline{x}_8$, A_4^1, $rule_4$
3	$\overline{x}_1$, $\overline{x}_2$, $\overline{x}_4$, $\overline{x}_5$, $\overline{x}_6$, $\overline{x}_7$, $\overline{x}_9$	$\overline{x}_1$, $\overline{x}_4$, $\overline{x}_5$, $\overline{x}_6$, A_9^1	$\overline{x}_2$, $\overline{x}_5$, $\overline{x}_6$, $\overline{x}_7$, A_1^1, A_9^1, A_3^2, A_8^2, A_9^2, A_4^3	$\overline{x}_2$, $\overline{x}_5$, $\overline{x}_6$, A_1^1, A_4^1, A_1^2, A_3^3, A_4^4
4	$\overline{x}_1$, $\overline{x}_2$, $\overline{x}_4$, $\overline{x}_5$, $\overline{x}_6$, $\overline{x}_7$, $\overline{x}_9$	$\overline{x}_1$, $\overline{x}_5$, $\overline{x}_6$, $\overline{x}_7$, A_3^1, A_4^1, A_9^1, A_2^2	$\overline{x}_1$, $\overline{x}_2$, $\overline{x}_4$, $\overline{x}_5$, $\overline{x}_6$, A_7^1, A_3^2	$\overline{x}_2$, $\overline{x}_6$, A_1^1, A_4^1, A_5^1, A_7^1, A_8^1, A_9^1, A_1^2, A_3^2, A_4^2, A_5^2, A_7^2, A_1^3, A_3^3, A_5^4, $\overline{y}^1$

Table 5. Simulation results

GLASS IDENTIFICATION PROBLEM													
N R	1 2	1 3	1 4	2 2	2 3	2 4	3 2	3 3	3 4	4 2	4 3	4 4	
$\overline{x}_1$	-	-	-	-	-	-	-	v	-	-	v	v	25
$\overline{x}_2$	-	-	-	-	v	v	-	-	-	-	-	-	17
$\overline{x}_3$	v	v	v	v	v	v	v	v	v	v	v	v	100
$\overline{x}_4$	-	-	-	-	-	v	v	v	-	v	v	v	50
$\overline{x}_5$	-	-	-	-	-	-	-	-	-	-	-	v	8
$\overline{x}_6$	-	-	-	-	-	-	-	-	-	-	-	-	0
$\overline{x}_7$	-	-	-	v	v	-	v	-	v	v	v	v	58
$\overline{x}_8$	v	v	v	v	v	v	v	v	v	-	v	v	92
$\overline{x}_9$	-	-	-	v	v	v	v	v	v	v	v	v	75
	22	22	22	44	56	56	56	56	44	44	67	78	[%]

In Fig. 4 we show degree of parameter number reduction [%], in Fig. 5 degree of learning time reduction [%], in Fig. 6 degree of learning time reduction per a single parameter [%], in Fig. 7

percentage of neuro-fuzzy systems having a particular input (attribute) after the reduction process, in Fig. 8 percentage of inputs (attributes) corresponding to a particular neuro-fuzzy system after the reduction process.

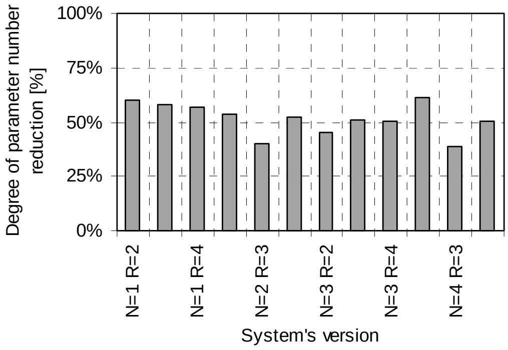

Fig. 4. Degree of parameter number reduction [%] for the Glass Identification problem

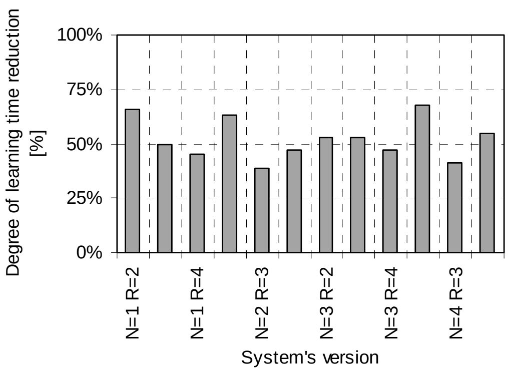

Fig. 5. Degree of learning time reduction [%] for the Glass Identification problem

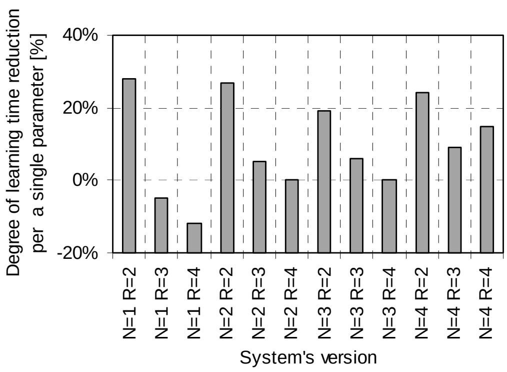

Fig. 6. Degree of learning time reduction per a single parameter [%] for the Glass Identification problem

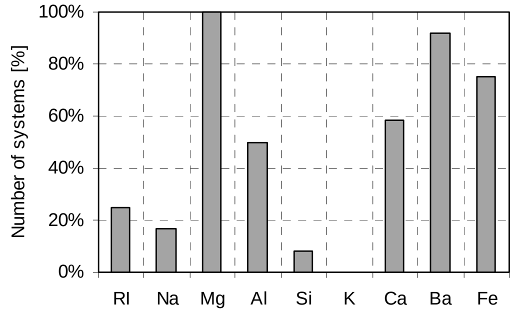

Fig. 7. Percentage of neuro-fuzzy systems having a particular input (attribute) after the reduction process for the Glass Identification problem

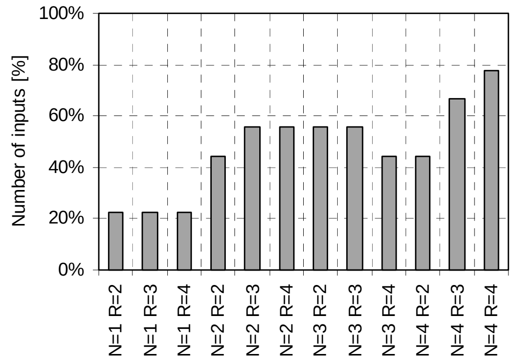

Fig. 8. Percentage of inputs (attributes) corresponding to a particular neuro-fuzzy system after the reduction process for the Glass Identification problem

6.2. Wisconsin Breast Cancer problem

The Wisconsin Breast Cancer data contains 699 instances (of which 16 instances have a single missing attribute) and each instance is described by nine attributes (clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli, mitoses). We removed those 16 instances and used the remaining 683 instances. Out of 683 data samples, 444 cases represent benign breast cancer and 239 cases describe malignant breast cancer. The problem is to classify whether a new case is a benign (class 1) or malignant (class 2) type of cancer. In our experiments, all sets are divided into a learning sequence (478 sets) and a testing sequence (205 sets).

The experimental results for the Wisconsin Breast Cancer problem are depicted in tables 6, 7, 8, 9, 10 and figures 9, 10, 11, 12, 13. The meaning of tables and figures in this example is analogous to that presented in the previous example.

Table 6. Simulation results

WISCONSIN BREAST CANCER PROBLEM				
R	***N***			
	1	2	3	4
2	3.35%/3.35%	2.72%/**2.51**%	2.51%/2.51%	2.51%/2.51%
	0.98%/0.98%	1.46%/1.46%	1.46%/1.46%	1.46%/1.46%
3	2.72%/2.72%	2.72%/2.72%	2.51%/2.51%	3.35%/**3.14**%
	1.46%/**0.98**%	1.46%/**0.98**%	0.98%/0.98%	0.98%/0.98%
4	2.51%/2.51%	2.51%/2.51%	2.72%/**2.51**%	2.51%/2.51%
	1.46%/1.46%	1.46%/**0.98**%	1.46%/1.46%	0.98%/0.98%

Table 7. Simulation results

WISCONSIN BREAST CANCER PROBLEM				
R	***N***			
	1	2	3	4
2	9/1/2/9/35 **5**/1/2/**5**/**23**	9/2/2/18/67 **6**/**1**/2/**6**/**26**	9/3/2/27/99 **7**/**2**/2/**14**/**55**	9/4/2/36/131 **7**/**1**/2/**7**/**29**
3	9/1/3/9/36 **6**/1/3/**6**/**27**	9/2/3/18/68 **6**/2/3/**12**/**50**	9/3/3/27/100 **7**/3/**2**/**18**/**72**	9/4/3/36/132 **5**/**3**/3/**15**/**64**
4	9/1/4/9/37 **6**/1/4/**6**/**28**	9/2/4/18/69 **5**/2/**3**/**10**/**44**	9/3/4/27/101 **6**/**1**/**3**/**6**/**27**	9/4/4/36/133 **8**/**3**/**3**/**23**/**88**

Table 8. Simulation results

WISCONSIN BREAST CANCER PROBLEM				
R	***N***			
	1	2	3	4/
2	48%	68%	51%	81%
	25%	22%	14%	20%
3	33%	33%	44%	57%
	14%	12%	24%	15%
4	27%	48%	80%	44%
	6%	22%	30%	17%

Table 9. Simulation results

WISCONSIN BREAST CANCER PROBLEM				
R	***N***			
	1	2	3	4
2	$\bar{x}_2$, $\bar{x}_4$, $\bar{x}_5$	$\bar{x}_1$, $\bar{x}_2$, $\bar{x}_4$, $rule_2$	$\bar{x}_1$, $\bar{x}_4$, $rule_2$	$\bar{x}_1$, $\bar{x}_4$, $rule_2$, $rule_3$, $rule_4$
3	$\bar{x}_2$, $\bar{x}_4$, $\bar{x}_5$	$\bar{x}_2$, $\bar{x}_4$, $\bar{x}_5$	$\bar{x}_4$, $\bar{x}_7$, A_5^1, A_2^2, A_6^2, $\bar{y}^2$	$\bar{x}_2$, $\bar{x}_3$, $\bar{x}_4$, $\bar{x}_5$, $rule_2$
4	$\bar{x}_1$, $\bar{x}_2$, $\bar{x}_4$	$\bar{x}_2$, $\bar{x}_4$, $\bar{x}_5$, $\bar{x}_9$, $\bar{y}^3$	$\bar{x}_1$, $\bar{x}_2$, $\bar{x}_4$, $rule_2$, $rule_3$, $\bar{y}^3$	$\bar{x}_2$, $rule_3$, $\bar{y}^4$

Table 10. Simulation results

WISCONSIN BREAST CANCER PROBLEM													
N *R*	1 2	1 3	1 4	2 2	2 3	2 4	3 2	3 3	3 4	4 2	4 3	4 4	
$\bar{x}_1$	v	v	-	-	v	v	-	v	-	-	v	v	58
$\bar{x}_2$	-	-	-	-	-	-	v	v	-	v	-	-	25
$\bar{x}_3$	v	v	v	v	v	v	v	v	v	v	-	v	92
$\bar{x}_4$	-	-	-	-	-	-	-	-	-	-	-	v	8
$\bar{x}_5$	-	-	v	v	-	-	v	v	v	v	-	v	58
$\bar{x}_6$	v	v	v	v	v	v	v	v	v	v	v	v	100
$\bar{x}_7$	v	v	v	v	v	v	v	-	v	v	v	v	92
$\bar{x}_8$	v	v	v	v	v	v	v	v	v	v	v	v	100
$\bar{x}_9$	-	v	v	v	v	-	v	v	v	v	v	v	83
	56	67	67	67	67	56	78	78	67	78	56	89	[%]

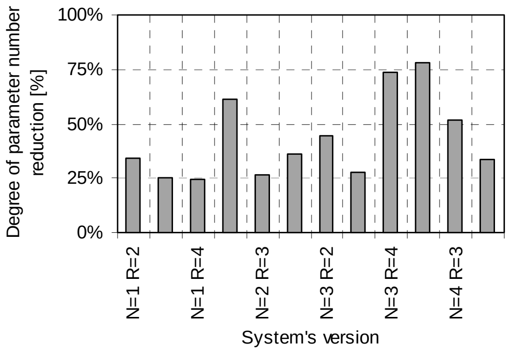

Fig. 9. Degree of parameter number reduction [%] for the Wisconsin Breast Cancer problem

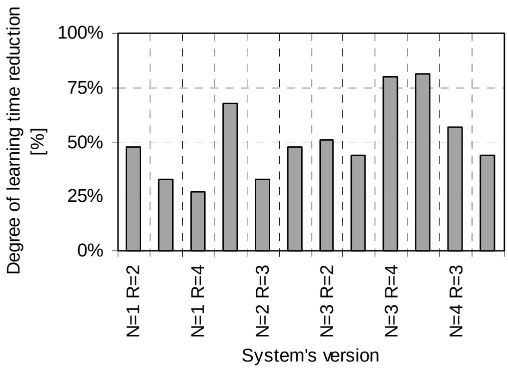

Fig. 10. Degree of learning time reduction [%] for the Wisconsin Breast Cancer problem

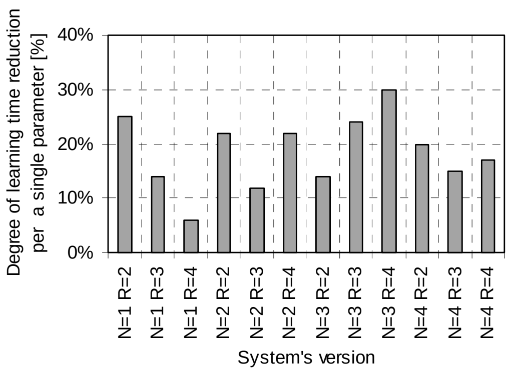

Fig. 11. Degree of learning time reduction per a single parameter [%] for the Wisconsin Breast Cancer problem

7. Final remarks

In the paper a new method for designing flexible neuro-fuzzy systems has been presented. Our algorithm allows to find a compromise between accuracy and complexity of neuro-fuzzy systems. This leads to transparent fuzzy rules and improves interpretability of such systems.

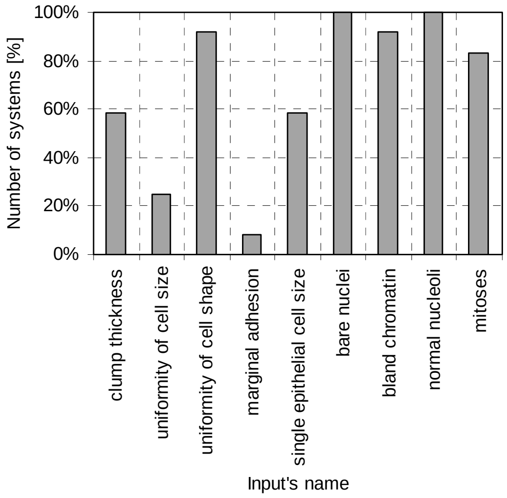

Fig. 12. Percentage of neuro-fuzzy systems having a particular input (attribute) after the reduction process for the Wisconsin Breast Cancer problem

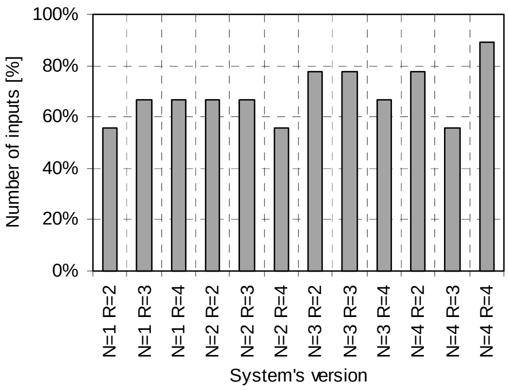

Fig. 13. Percentage of inputs (attributes) corresponding to a particular neuro-fuzzy system after the reduction process for the Wisconsin Breast Cancer problem

Acknowledgement

This work is supported by the Foundation for Polish Science (Professorial Grant 2005-2008), the Polish State Committee for Scientific Research (Grant Nr T11C 04827) and Polish Ministry of Science and Higher Education (Special Research Project 2006-2009).

References:

[1] R. A. Aliev and R. R. Aliev. *Soft Computing and its Applications. World Scientific Publishing,* Singapore-New Jersey-London-Hong Kong, 2001.

[2] H. R. Berenji and P. Khedkar. Learning and tuning fuzzy logic controllers through reinforcements, *IEEE Trans. Neural Networks,* vol. 3. pp. 724-740, October 1992.

[3] K. Cpałka, L. Rutkowski, *A New Method for Designing and Reduction of Neuro-Fuzzy Systems,* 2006 IEEE International Conference on Fuzzy Systems, IEEE World Congress on Computational Intelligence, Vancouver, BC, Canada, 2006.

[4] K. Cpałka, L. Rutkowski, *A New Method for Complexity Reduction of Neuro-Fuzzy Systems*, WSEAS Transactions on Systems, Issue 11, vol. 5, Nov. 2006, pp. 2514-2521.

[5] E. Czogała, J. Łęski, *Fuzzy and Neuro-Fuzzy Intelligent Systems,* Physica-Verlag, A Springer-Verlag Company, Heidelberg, New York, 2000.

[6] C. Fodor, On fuzzy implication, *Fuzzy Sets and Systems,* vol. 42, pp. 293-300, 1991.

[7] M. Gorzałczany, *Computational Intelligence Systems and Applications: Neuro-Fuzzy and Fuzzy Neural Synergisms.* Springer-Verlag 2002.

[8] J. S. R. Jang. ANFIS: Adaptive network based fuzzy inference system, IEEE Trans. Syst., Man, Cybern., vol. 23, pp. 665 685, June 1993.

[9] N. Kasabov. DENFIS: dynamic evolving neural fuzzy inference system and its application for time series prediction, IEEE Trans. on Fuzzy Systems, vol. 10, pp. 144 154, April 2002.

[10] E. P. Klement, R. Mesiar, E. Pap, *Triangular Norms,* Kluwer Academic Publishers, Netherlands 2000.

[11] C. T. Lin and C. S. G. Lee. Neural network based fuzzy logic control and decision system, IEEE Trans. Comput., vol. 40, pp. 1320 1336, December 1991.

[12] D. Nauck, F. Klawon and R. Kruse. *Foundations of Neuro-Fuzzy Systems,* Chichester, U.K.: Wiley, 1997.

[13] D. Nauck and R. Kruse. Neuro-fuzzy systems for function approximation, *Fuzzy Sets and Systems,* vol. 101, pp. 261-271, 1999.

[14] L. Rutkowski, *Flexible Neuro-Fuzzy Systems*. Kluwer Academic Publishers, 2004.

[15] L. Rutkowski and K. Cpałka, Designing and learning of adjustable quasi triangular norms with applications to neuro fuzzy systems, *IEEE Trans. on Fuzzy Systems,* vol. 13, Feb 2005, pp. 140 151.

[16] L. Rutkowski, K. Cpałka, Flexible neuro-fuzzy systems, *IEEE Transactions on Neural Networks,* vol. 14, May 2003, pp. 554 574.

[17] UCI respository of machine learning databases, Available online: *http://ftp.ics.uci.edu/pub/machine-learning-databases/.*

[18] J. S. Wang and C. S. G. Lee. Self-adaptive neuro-fuzzy inference systems for classification applications, *IEEE Trans. on Fuzzy Systems,* vol. 10, pp. 790-802, December 2002.

[19] R. R. Yager, D. P. Filev, *Essentials of fuzzy modelling and control*. John Wiley & Sons, 1994.