BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 63, No. 3, 2015
DOI: 10.1515/bpasts-2015-0090

Handling fuzzy systems' accuracy-interpretability trade-off by means of multi-objective evolutionary optimization methods – selected problems

M.B. GORZAŁCZANY* and F. RUDZIŃSKI

Department of Electrical and Computer Engineering, Kielce University of Technology, 7 1000-lecia P.P. Ave., 25-314 Kielce, Poland

Abstract. The paper addresses several open problems regarding the automatic design of fuzzy rule-based systems (FRBSs) from data using multi-objective evolutionary optimization algorithms (MOEOAs). In particular, we propose: a) new complexity-related interpretability measure, b) efficient strong-fuzzy-partition implementation for improving semantics-related interpretability, c) special-coding-free implementation of rule base and original genetic operators for its processing, and d) implementation of our ideas in the context of well-known MOEOAs such as SPEA2 and NSGA-II. The experiments demonstrate that our approach is an effective tool for handling FRBSs' accuracy-interpretability trade-off, i.e, designing FRBSs characterized by various levels of such a trade-off (in particular, for designing highly interpretability-oriented systems of still competitive accuracy).

Key words: accuracy and interpretability of fuzzy rule-based systems, multi-objective evolutionary optimization, genetic computations, fuzzy systems.

1. Introduction

Discovering, in an automatic way, knowledge in data sets is nowadays one of central issues in designing intelligent systems in various application areas, cf. [1–3]. As far as the representation of such a knowledge is concerned, among the most commonly used structures are conditional rules – in particular, linguistic fuzzy conditional rules, cf. e.g. [4] – due to their high readability and modularity. In recent years, in an automatic design of fuzzy rule-based systems (FRBSs) from data, more and more attention has been paid not only to aspects of their accuracy (i.e., the ability to adequately represent the modelled system, decision making process, etc.) but also to issues of their interpretability (i.e., the ability to present the functioning of the modelled systems in an understandable way). Whereas there are well-defined and widely accepted measures of FRBSs' accuracy, the definition of a standard measure of much more subjective property such as their interpretability is still an open problem [5]. Usually, two main aspects of FRBSs' interpretability are considered [5]: the complexity of their rule bases and the semantics associated with membership functions representing linguistic terms that describe particular attributes.

The automatic design of FRBSs from data can be presented as a structure and parameter optimization or search problem in large search spaces. For this reason, genetic algorithms have been successfully applied to solve such problems [6–11]. Since accuracy and interpretability are somehow complementary/contradictory objectives, formulating a single-objective optimization task is fully justified (using a fitness function defined as a combination of measures of both objectives, e.g. [8]). As shown in [8], such a solution gives very good accuracy-interpretability trade-offs in many application areas. Nevertheless, obviously, it explores only a reduced part of the whole search space and some interesting solutions may never be discovered. Multi-objective evolutionary optimization algorithms (MOEOAs) are better fitted for this task. They are able to obtain, in a single run of the algorithm, a set of solutions (approximating Pareto-optimal solutions) characterized by various levels of accuracy-interpretability trade-off. However, there are still many open problems regarding the use of MOEOAs in the automatic design of FRBSs from data, see [12] for review. To the list of [12], we can add the problem of computationally efficient FRBSs' representation for genetic computations using MOEOAs.

In this paper, we address some of the above-mentioned open problems. Firstly, we propose a new complexity-related interpretability measure. Secondly, we present a computationally efficient implementation of the so-called strong fuzzy partition (SFP) condition [13] that perfectly meets the semantics-related interpretability constraints [5]. Thirdly, we propose direct and special-coding-free representation of the fuzzy rule base structure and original genetic crossover and mutation operators for its processing. Fourthly, we present our ideas in the context of well-known MOEOAs such as SPEA2 [14] and NSGA-II [15] and compare our approach with several alternative ones [16] using some well-known benchmark data sets.

2. Main components of FRBSs designed from data

Consider a fuzzy rule-based classifier (FRBC) with n inputs $x_1, x_2, \ldots, x_n$ and an output, which has the form of a fuzzy

*e-mail: m.b.gorzalczany@tu.kielce.pl

set over the set $Y = \{y_1, y_2, \ldots, y_c\}$ of c class labels. Each input attribute x_i ($x_i \in X_i$), $i = 1, 2, \ldots, n$ is described by numerical values. The "values" of symbolic attributes are encoded using integer numbers.

2.1. Learning data. The classifier is designed from K input-output learning samples:

$$\boldsymbol{L}_1 = \left\{ \boldsymbol{x}_k^{(lrn)}, y_k^{(lrn)} \right\}_{k=1}^{K}, \tag{1}$$

where $\boldsymbol{x}_k^{(lrn)} = \left(x_{1k}^{(lrn)}, x_{2k}^{(lrn)}, \ldots, x_{nk}^{(lrn)} \right) \in \boldsymbol{X} = X_1 \times X_2 \times \cdots \times X_n$ ($\times$ stands for Cartesian product of ordinary sets) is the set of input numerical attributes and $y_k^{(lrn)}$ is the corresponding class label ($y_k^{(lrn)} \in Y$) for the k-th data sample. Expression (1) can be rewritten in an equivalent way as follows:

$$\boldsymbol{L} = \left\{ \boldsymbol{x}_k^{(lrn)}, B_{(singl.)k}^{(lrn)} \right\}_{k=1}^{K}, \tag{2}$$

where $\boldsymbol{x}_k$ is as in (1) and $B_{(singl.)k}^{(lrn)}$ is the fuzzy singleton for the class label $y_k^{(lrn)}$, i.e., $\mu_{B_{(singl.)k}^{(lrn)}}(y) = 1$ for $y = y_k^{(lrn)}$ and 0 elsewhere ($\mu_{B_{(singl.)k}^{(lrn)}}(y)$ denotes the membership function of the fuzzy singleton $B_{(singl.)k}^{(lrn)}$).

2.2. Fuzzy knowledge base. Linguistic fuzzy classification rules that are to be discovered in the learning data $\boldsymbol{L}$ (2) by the presented later in the paper multi-objective genetic optimization approaches have the following form (we consider the r-th rule, $r = 1, 2, \ldots, R$; R changes during the learning process) [17]:

$$\begin{aligned} &\textbf{IF } [x_1 \text{ is } [\text{not}]_{(sw_1^{(r)}<0)} \ A_{1,|sw_1^{(r)}|}]_{(sw_1^{(r)} \neq 0)} \ \textbf{AND...AND} \\ &\quad [x_n \text{ is } [\text{not}]_{(sw_n^{(r)}<0)} \ A_{n,|sw_n^{(r)}|}]_{(sw_n^{(r)} \neq 0)} \\ &\quad \textbf{THEN } y \text{ is } B_{(singl.)j^{(r)}}. \end{aligned} \tag{3}$$

The components $[expression]_{(condition)}$ in (3) mean conditional inclusion (i.e., $expression$ is included into the rule if and only if $condition$ is fulfilled). $|\cdot|$ returns the absolute value. $sw_i^{(r)} \in \{0, \pm 1, \pm 2, \ldots, \pm a_i\}$ (a_i denotes the number of fuzzy sets/linguistic terms defined for the i-th attribute), $i = 1, 2, \ldots, n$ is a switch which controls the i-th input attribute of the r-th fuzzy rule. For $sw_i^{(r)} = 0$, the i-th attribute is excluded from (not active in) that rule, whereas for $sw_i^{(r)} > 0$ the component x_i is A_{ik_i} ($k_i = |sw_i^{(r)}|$) is included (active) and for $sw_i^{(r)} < 0$ the component x_i is not A_{ik_i} is used in that rule (not $A_{ik_i} = \overline{A}_{ik_i}$ and $\mu_{\overline{A}_{ik_i}}(x_i) = 1 - \mu_{A_{ik_i}}(x_i)$). $B_{(singl.)j^{(r)}}$ is the fuzzy singleton representing the class label $y_{j^{(r)}}$, $j^{(r)} \in \{1, 2, ..., c\}$. At least one input attribute is active in a given rule. In different rules, different subsets of attributes are active. In experiments presented later in the paper, the rules (3) will be presented without switches $sw_i^{(r)}$; simply, only the active input attributes will be shown.

$A_{ik_i} \in F(X_i)$, $k_i = 1, 2, \ldots, a_i$, $i = 1, 2, \ldots, n$ ($F(X_i)$ denotes the family of all fuzzy sets defined in the universe X_i) are the S-, M-, or L-type fuzzy sets (see below). Therefore, $FP(X_i) = \{A_{i1}, A_{i2}, \ldots, A_{ia_i}\}$ is a fuzzy partition of X_i, in which A_{i1} is S-type, A_{ia_i} – L-type, and $A_{i2}, A_{i3}, \ldots, A_{i,a_i-1}$ – M-type fuzzy sets. For simplicity, A_{ik_i} denote also the corresponding linguistic terms: "*Small*" (S-type fuzzy set), "*Medium*" (M-type set), and "*Large*" (L-type set). Trapezoidal membership functions of S-, M-, and L-type fuzzy sets are presented in Fig. 1.

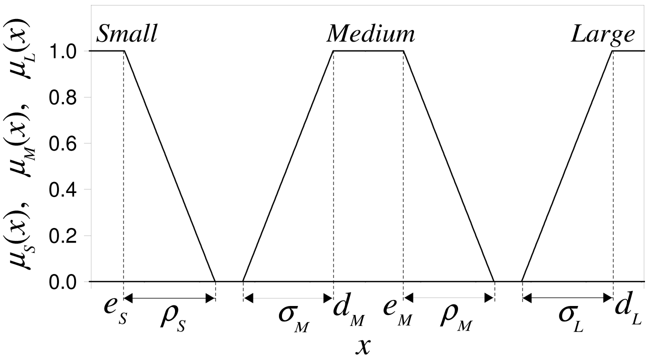

Fig. 1. Trapezoidal membership functions of S-type, M-type, and L-type fuzzy sets and their parameters

In fuzzy knowledge-based systems one can distinguish two components: linguistic rule bases (RBs) and data bases (DBs). In our approach, the RB is represented by the following set of parameters:

$$RB = \left\{ sw_1^{(r)}, sw_2^{(r)}, \ldots, sw_n^{(r)}, j^{(r)} \right\}_{r=1}^{R}. \tag{4}$$

Therefore, we propose direct, simple, special-coding-free and thus computationally efficient RB's representations. Dedicated original crossover and mutation operators for its processing are presented later in the paper.

In turn, the DB contains parameters: e_{i1}, ρ_{i1} (for S-type fuzzy set), d_{i2}, e_{i2}, σ_{i2}, ρ_{i2} (for the first M-type fuzzy set), $\ldots$, d_{i,a_i-1}, e_{i,a_i-1}, σ_{i,a_i-1}, ρ_{i,a_i-1} (for the last M-type fuzzy set), and d_{ia_i}, σ_{ia_i} (for the L-type fuzzy set) of antecedents membership functions, $i = 1, 2, \ldots, n$, for numerical attributes. The DB contains also information on domains of symbolic attributes $X_i = (x_{i1}, x_{i2}, \ldots, X_{ia_i})$ and set of class labels $Y = \{y_1, y_2, \ldots, y_c\}$; these parameters, obviously, are not being tuned.

2.3. Fuzzy approximate inference. An evaluation of particular individuals (fuzzy rule bases in our Pittsburgh-type approach) in each generation of genetic computations must be performed. For this reason, a fuzzy-set-theory representation of fuzzy rule base (3) and fuzzy inference mechanism must be employed. Both, compositional rule of inference and similarity-based reasoning with various definitions of fuzzy implications, t-norms, and t-conorms (see, e.g., [18]) can be implemented in our approach. In the case of widely used Mamdani's model (with min-type t-norm, max-type t-conorm and min operator playing the role of fuzzy implication), we obtain – for the input numerical data $\boldsymbol{x}' = (x_1', x_2', \ldots, x_n')$

– a FRBC's fuzzy-set response B' characterized by its membership function $\mu_{B'}(y)$, $y \in Y = \{y_1, y_2, \ldots, y_c\}$:

$$\begin{aligned}\mu_{B'}(y) &= \max_{r=1,2,\ldots,R} \mu_{B'^{(r)}}(y) \\ &= \max_{r=1,2,\ldots,R} \min[\alpha^{(r)}, \mu_{B_{(singl.)j^{(r)}}}(y)],\end{aligned} \tag{5}$$

where

$$\alpha^{(r)} = \min_{\substack{i=1,2,\ldots,n,\\ sw_i^{(r)}\neq 0}} \alpha_i^{(r)}, \quad \text{and}$$

$$\alpha_i^{(r)} = \begin{cases} \mu_{A_{i,sw_i^{(r)}}}(x'_i), & \text{for } sw_i^{(r)} > 0, \\ \mu_{\bar{A}_{i,|sw_i^{(r)}|}}(x'_i), & \text{for } sw_i^{(r)} < 0. \end{cases} \tag{6}$$

If a FRBC's non-fuzzy response y' is required, it is calculated as follows:

$$y' = \arg\max_{y\in Y} \mu_{B'}(y). \tag{7}$$

3. Main components of genetic learning process

3.1. Definition of optimization objectives.

Accuracy. The following measure (a fitness function subject to maximization) of the FRBC's accuracy is used:

$$ff_{ACU} = 1 - Q_{RMSE}^{(lrn)}, \tag{8}$$

where

$$Q_{RMSE}^{(lrn)} = \sqrt{\frac{1}{Kc}\sum_{k=1}^{K}\sum_{j=1}^{c}\left[\mu_{B_{(singl.)k}^{(lrn)}}(y_j) - \mu_{B'_k}(y_j)\right]^2}. \tag{9}$$

$Q_{RMSE}^{(lrn)} \in [0,1]$, B'_k is the system's fuzzy-set response (5) for the learning data sample $\boldsymbol{x}_k^{(lrn)}$, and $B_{(singl.)k}^{(lrn)}$ is the desired fuzzy-singleton response – see (3).

Complexity-related interpretability. The following measure (a fitness function subject to maximization) is used:

$$ff_{INT} = 1 - Q_{CPLX}. \tag{10}$$

$Q_{CPLX} \in [0,1]$ denotes the FRBC's complexity ($Q_{CPLX} = 0$ and 1 represent minimal and maximal complexity, respectively) determined on the basis of three indices that measure average complexity of particular rules (Q_{RINP}) and complexity of the whole system in terms of active inputs (Q_{INP}) and active fuzzy sets (Q_{FS}):

$$Q_{CPLX} = \frac{Q_{RINP} + Q_{INP} + Q_{FS}}{3}, \tag{11}$$

where

$$\begin{gathered} Q_{RINP} = \frac{1}{R}\sum_{r=1}^{R}\frac{n_{INP}^{(r)} - 1}{n-1}, \\ Q_{INP} = \frac{n_{INP} - 1}{n-1}, \\ Q_{FS} = \frac{n_{FS} - 1}{a_{ALL} - 1}, \qquad a_{ALL} = \sum_{i=1}^{n} a_i, \\ \text{and} \quad n > 1. \end{gathered} \tag{12}$$

$n_{INP}^{(r)}$ is the number of active input attributes in the r-th rule, n_{INP} and n_{FS} are the numbers of active inputs and fuzzy sets (linguistic terms), respectively, in the whole system.

Semantics-related interpretability. Fuzzy partitions in which the sum of the values of all membership functions for any domain value is equal to 1 (they are referred to as strong fuzzy partitions (SFPs)), satisfy the desired semantics-related interpretability demands at the highest level. We propose simple and thus computationally efficient implementation of SFP condition for trapezoidal membership functions as follows (see Fig. 2 for three-set SFP of x_i):

$$\begin{gathered} \sigma_{ik_i} = \rho_{i,k_i-1} = d_{ik_i} - e_{i,k_i-1}, \\ k_i = 2, 3, \ldots, a_i \end{gathered} \tag{13}$$

and, obviously,

$$\begin{gathered} e_{i1} \le d_{i2} \le e_{i2} \le \cdots \le d_{i,a_i-1} \le e_{i,a_i-1} \le d_{ia_i}, \\ i = 1, 2, \ldots, n. \end{gathered} \tag{14}$$

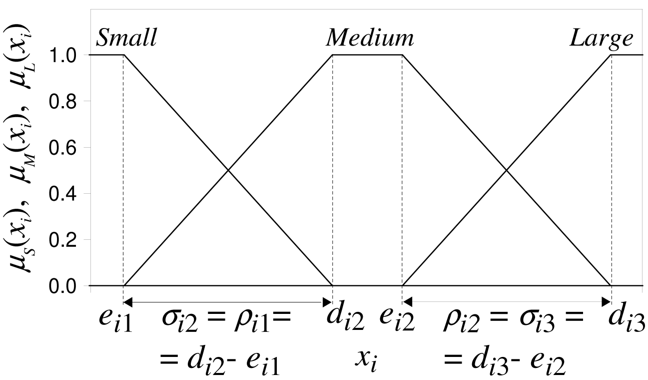

Fig. 2. Implementation of three-fuzzy-set SFP

3.2. Genetic operators. A single individual (representing FRBC) consists of two parts that represent system's RB and DB, respectively. Non-binary genetic operators are defined separately for RB and DB processing (see [17] for details).

Crossover operator for RB transformation (C-RB). It processes two individuals (two RBs) of R_1 and R_2 rules, respectively, by performing one of five randomly selected sub-operations C-RB1, C-RB2, ..., and C-RB5:

C-RB1 (exchange of many rules): In the first stage, for the r-th rule in both RBs, $r = 1, 2, \ldots, min(R_1, R_2)$, the so-called *random-switch condition* (equivalent to the random selection of 1 from the set $\{0, 1\}$ is checked. If this condition is fulfilled, the r-th tules from both RBs are exchanged. In the second stage, each of the remaining rules of the larger RB, assuming that the *random-switch condition* is fulfilled, is moved to the smaller RB.

<u>C-RB2</u> (exchange of a single rule): Analogous as C-RB1 but the activities of C-RB1 are performed unconditionally only once for randomly selected r-th rule in the larger RB.

<u>C-RB3</u> (exchange of many fuzzy sets in many fuzzy rules): If the same condition as in the first stage of C-RB1 is fulfilled then – for the i-th input attribute, $i = 1, 2, \ldots, n$ and for the output class label – the *random-switch condition* is run independently again. If it is fulfilled, the fuzzy sets describing a given input attribute or class label in both RBs are exchanged.

<u>C-RB4</u> (exchange of many fuzzy sets in a single rule): Analogous as C-RB3 but the activities of C-RB3 are performed unconditionally only once for randomly selected r-th rule in both RBs.

<u>C-RB5</u> (exchange of a single fuzzy set): Analogous as C-RB4 but the activities of C-RB4 are performed unconditionally only once for randomly selected i-th input attribute or output class label.

Crossover operator for DB transformation (C-DB). It randomly selects two fuzzy sets, each from one DB. New values of d- and e-parameters are calculated as linear combinations of their old values from both sets; they also must fulfil condition (14) [17]. New values of σ- and ρ-parameters are calculated from (13) using new values of parameters d and e.

Mutation operator for RB transformation (M-RB). It processes a single RB by performing one of four randomly selected sub-operations M-RB1, M-RB2, M-RB3, and M-RB4:

<u>M-RB1</u> (rule insertion): It inserts into RB a new rule (3) with randomly selected values of switches sw_i and class label j ($sw_i \in \{0, \pm1, \ldots, \pm a_i\}$, $i = 1, 2, \ldots, n$, $j \in \{1, 2, \ldots, c\}$).

<u>M-RB2</u> (rule deletion): It removes a randomly selected rule from the RB.

<u>M-RB3</u> (change a single fuzzy set): It randomly selects one rule from the RB and its one i-th input attribute or output class label j. Next, it randomly selects a new value of switch sw_i or class label j.

<u>M-RB4</u> (change of an input in a fuzzy rule): It randomly selects: one rule in the RB, its one active (i.e., with $sw_{i_1} \neq 0$) and one non-active (i.e., with $sw_{i_2} = 0$) input attributes. Then, the first attribute is off ($sw_{i_1} = 0$) and the second – is on ($sw_{i_2} \neq 0$) in that rule.

Mutation operator for DB transformation (M-DB). It randomly selects one fuzzy set from the DB and one of its two parameters d and e (say, d is selected). Its new value $d_{new} = d + rand(-0.2, 0.2)[x_{i,\max} - x_{i,\min}]$, where $rand(\cdot)$ returns a random number from the assumed interval and $[x_{i,\min}, x_{i,\max}]$ is a range of the domain of the selected set [17]. New values of σ and ρ are calculated from (13).

After performing crossover and mutation operations on RBs and DBs, empty rules (without antecedents) and rule duplicates are removed from the RBs.

4. Application to selected classification problems

Our ideas will now be presented in the context of two, presently most advanced, MOEOAs, i.e., SPEA2 [14] and NSGA-II [15], and compared with several alternative approaches using two benchmark data sets such as *Breast Cancer Wisconsin* (*Diagnostic*) (*BCWD*, for short) and *Wine* [19]. *BCWD* data set has 569 records, 30 numerical attributes, and 2 classes, whereas *Wine* data set – 178 records, 13 numerical attributes, and 3 classes. 10-fold cross-validation is performed for both data sets. The experiments for a single learning/test data split are presented in detail.

Figures 3a and 3b present the evolution of the Pareto-front approximations during the learning process using SPEA2 and NSGA-II for the *BCWD* data. Pareto front is the set of Pareto-optimal (i.e., non-dominated by any other) solutions; they are characterized by various levels of accuracy-interpretability trade-off. Figure 3c presents the best SPEA2/NSGA-II-based 10-point Pareto-front approximation, whereas Table 1 – the measures describing in detail the interpretability and accuracy of particular solutions of Fig. 3c. We can see that the more interpretability-oriented systems are generated by SPEA2 and the more accuracy-oriented ones – by NSGA-II. Additionally, Table 2 presents the fuzzy rule base of the most interpretability-oriented FRBC, i.e., the system no. 1 of Table 1 and Fig. 3c. Figure 4b presents the SFP of input attribute x_{25} (*Worst Area*) that occurs in the fuzzy rule base of Table 2 (obviously, $\mu_{\text{not } Small} = 1 - \mu_{Small}$). It is worth stressing that the solutions of Fig. 3c are characterized by complete and consistent fuzzy rule bases, cf. e.g. [20].

Another verification of our ideas implemented by means of SPEA2 and NSGA-II has been carried out using *Wine* data set. The results are presented in Fig. 5, Table 3, Table 4, and Fig. 6. General conclusions are the same as in the case of *BCWD* data.

Moreover, Table 5 presents a comparative analysis of our most interpretability-oriented solutions with several alternative techniques reported in [16]. The experiments show that our approach generates FRBCs of significantly improved interpretability, while still characterized by competitive accuracy.

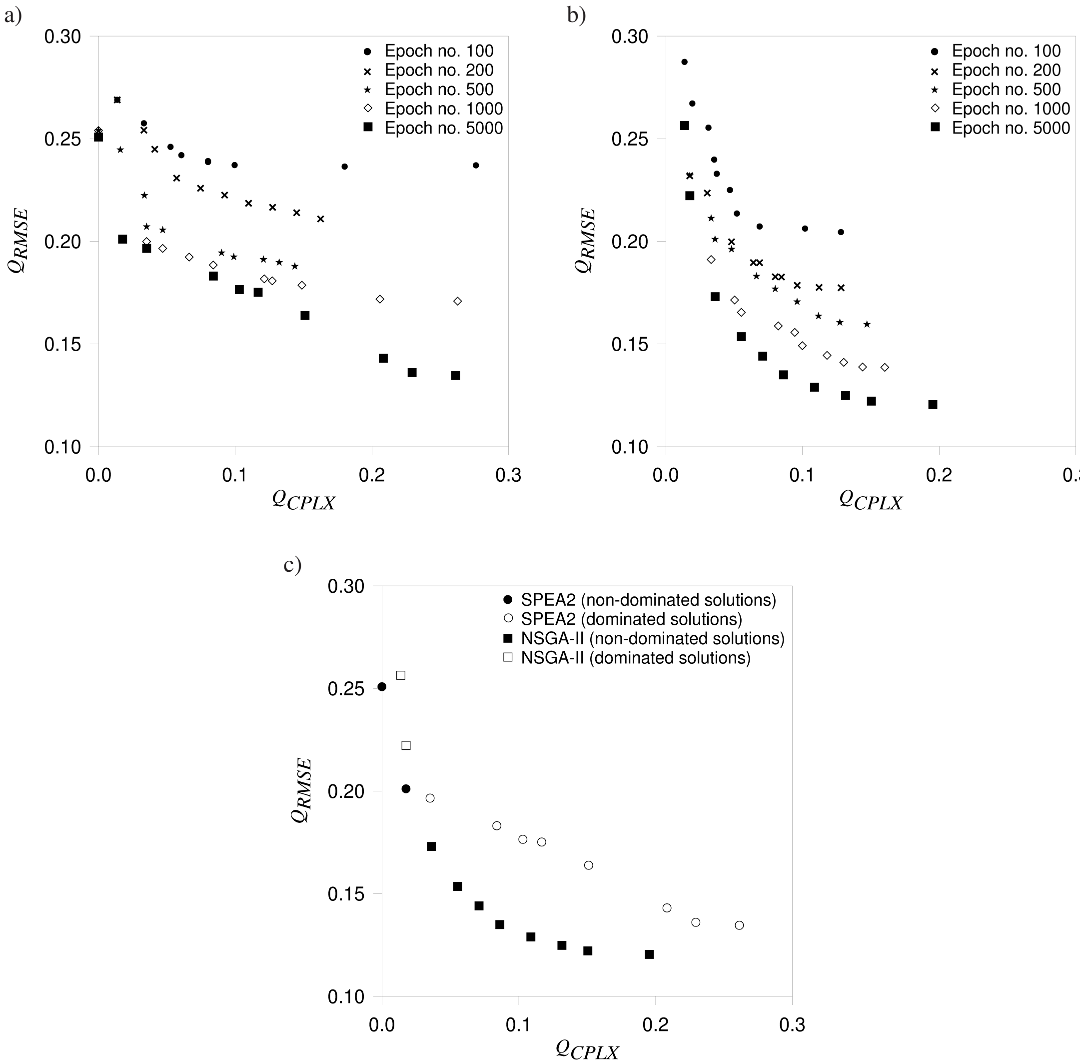

Fig. 3. Evolution of the Pareto-front approximations during the learning process using SPEA2 (a) and NSGA-II (b), as well as the best SPEA2/NSGA-II-based approximation (c) for *BCWD* data set

Table 1
Interpretability and accuracy measures of solutions from Fig. 3c

No.	Q_{CPLX}	R	n_{INP}	n_{FS}	$n_{INP/R}$	$Q_{RMSE}^{(lrn)}$	$CD^{(lrn)}$	$Q_{RMSE}^{(tst)}$	$CD^{(tst)}$
	interpretability measures					accuracy measures			
1	0	2	1	1	1	0.2508	92.1%	0.225	92.3%
2	0.0175	3	2	2	1.3	0.2011	95.0%	0.1492	96.9%
3	0.0360	4	3	3	1.7	0.1746	96.4%	0.1929	95.4%
4	0.0552	6	4	6	1.8	0.1535	97.6%	0.1939	96.9%
5	0.0708	7	5	7	2	0.1441	97.8%	0.1985	96.9%
6	0.0860	8	6	8	2.1	0.1349	97.8%	0.1969	95.4%
7	0.1088	7	7	10	2.7	0.1289	98.4%	0.2187	93.8%
8	0.1315	8	9	11	2.4	0.1249	98.2%	0.2013	95.4%
9	0.1504	9	11	15	2.2	0.1221	98.2%	0.2178	93.8%
10	0.1954	9	13	17	2.9	0.1204	98.4%	0.2176	93.8%

Q_{CPLX}, R, n_{INP}, and n_{FS} – as in (12); $n_{INP/R} = \sum_{r=1}^{R} n_{INP}^{(r)}/R$ – the number of active input attributes per rule; $Q_{RMSE}^{(lrn)}$ and $Q_{RMSE}^{(tst)}$ – learning and test errors; $CD^{(lrn)}$ and $CD^{(tst)}$ – percentages of correct decisions for learning and test data; solutions 1 and 2 – generated by SPEA2; solutions $3, 4, \ldots, 10$ – generated by NSGA-II.

Table 2
Fuzzy rule base of the most interpretability-oriented FRBC for *BCWD* data set

No.	Fuzzy classification rules	Number (percentage) of correct decisions	
		learning data	test data
1	**IF** x_{25} (*Worst Area*) is $Small$ **THEN** Class 1 (*Benign*)	306 (91.1%)	41 (89.1%)
2	**IF** x_{25} (*Worst Area*) is not $Small$ **THEN** Class 2 (*Malignant*)	158 (94.0%)	19 (100%)
	Overall:	464 (92.1%)	60 (92.3%)
	Q_{RMSE}:	0.251	0.233

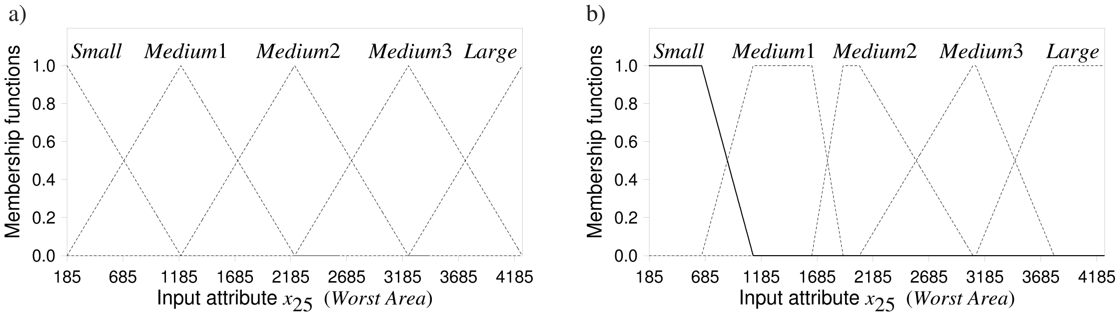

Fig. 4. The initial (a) and final (b) shapes of strong fuzzy partition (SFP) of input attribute x_{25} (*Worst Area*) in *BCWD* data set

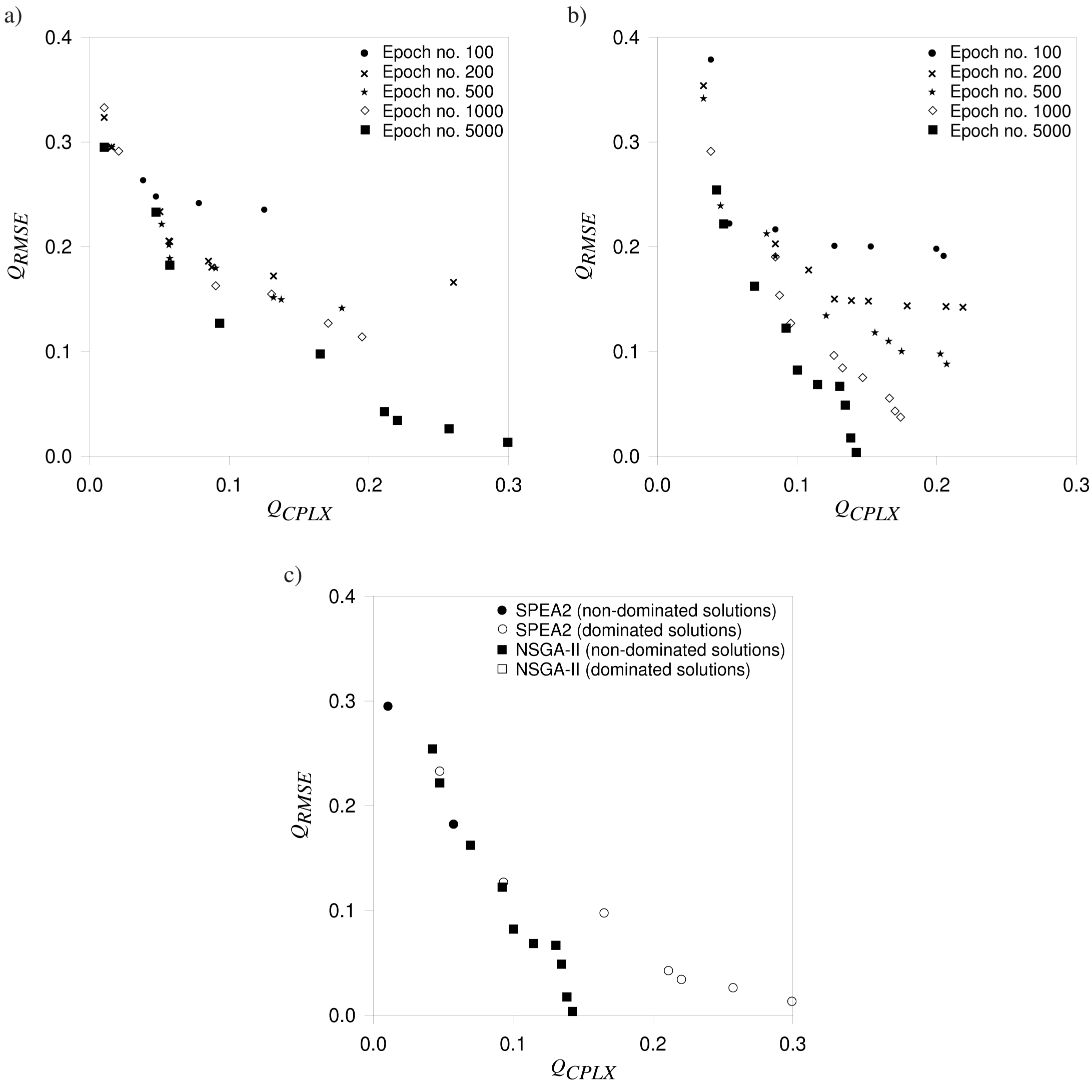

Fig. 5. Evolution of the Pareto-front approximations during the learning process using SPEA2 (a) and NSGA-II (b), as well as the best SPEA2/NSGA-II-based approximation (c) for *Wine* data set

Table 3
Interpretability and accuracy measures of solutions from Fig. 5c

No.	Q_{CPLX}	R	n_{INP}	n_{FS}	$n_{INP/R}$	$Q_{RMSE}^{(lrn)}$	$CD^{(lrn)}$	$Q_{RMSE}^{(tst)}$	$CD^{(tst)}$
	interpretability measures					accuracy measures			
1	0.0104	3	1	3	1	0.2949	80.1%	0.3521	70.6%
2	0.0451	4	2	3	1.2	0.2541	90.7%	0.3083	88.2%
3	0.0474	3	2	3	1.3	0.2217	93.8%	0.2633	94.1%
4	0.0694	4	2	5	1.7	0.1621	95.6%	0.2513	88.2%
5	0.0920	4	3	4	1.7	0.1222	98.7%	0.2663	88.2%
6	0.1001	6	3	6	1.7	0.0822	98.7%	0.1980	94.1%
7	0.1145	6	3	7	2	0.06846	99.4%	0.1980	94.1%
8	0.1344	7	4	7	1.7	0.04871	99.4%	0.1980	94.1%
9	0.1383	7	4	7	1.8	0.01738	100%	0.1980	94.1%
10	0.1426	7	4	7	2	0.003542	100%	0.2008	94.1%

see Table 1 for description of parameters; solution 1 – generated by SPEA2; solutions 2, 3, . . . , 10 – generated by NSGA-II.

Table 4
Fuzzy rule base of the most interpretability-oriented FRBC for *Wine* data set

No.	Fuzzy classification rules	Number (percentage) of correct decisions: learning data	test data
1	**IF** x_7 (*Flavanoids*) is $Medium2$ **THEN** Class 1	44 (80.0%)	5 (62.5%)
2	**IF** x_7 (*Flavanoids*) is $Medium1$ **THEN** Class 2	46 (76.7%)	5 (71.4%)
3	**IF** x_7 (*Flavanoids*) is $Small$ **THEN** Class 3	39 (78.3%)	2 (100%)
	Overall:	129 (80.1%)	12 (70.6%)
	Q_{RMSE}:	0.294	0.352

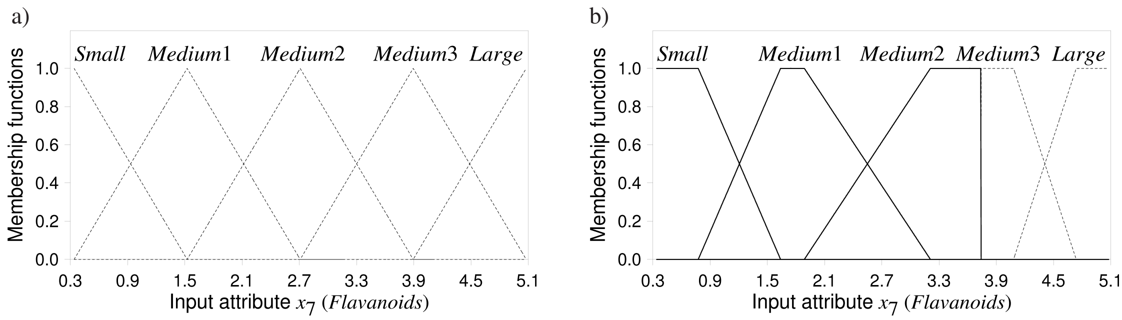

Fig. 6. The initial (a) and final (b) shapes of strong fuzzy partition (SFP) of input attribute x_7 (*Flavanoids*) in *Wine* data set

Table 5
Comparative analysis

Approach	$BCWD$ data set: $\overline{R}$	$\overline{n_{INP/R}}$	$\overline{CD^{(tst)}}$	$Wine$ data set: $\overline{R}$	$\overline{n_{INP/R}}$	$\overline{CD^{(tst)}}$
Our	**2.8**	**1.3**	**93.90%**	**3.0**	**1.4**	**90.56%**
SGERD	3.7	2.0	90.68%	4.2	2.0	91.88%
2SLAVE	5.2	8.1	92.33%	5.5	10.3	92.52%
FH-GBML	7.2	4.9	92.26%	9.2	4.7	92.61%
FARC-HD	10.4	1.7	95.25%	8.7	1.6	94.35%

$\overline{R}$, $\overline{n_{INP/R}}$, and $\overline{CD^{(tst)}}$ are average values of R, $n_{INP/R}$, and $CD^{(tst)}$ obtained in 10-fold cross-validation experiment; 2SLAVE – Structural Learning Algorithm on Vague Environment; FH-GBML – Fuzzy Hybrid Genetic-Based Machine Learning algorithm; SGERD – Steady-state Genetic algorithm for Extracting fuzzy classification Rules from Data; FARC-HD – Fuzzy Associative Rule-based Classification method for High-Dimensional datasets.

5. Conclusions

In this paper, we address several open problems regarding the automatic design of FRBSs from data using MOEOAs. In particular, we propose: a) new complexity-related interpretability measure, b) efficient SFP implementation for improving semantics-related interpretability, c) special-coding-free RB's implementation and original genetic operators for its processing, and d) implementation of our ideas in the context of well-known MOEOAs such as SPEA2 and NSGA-II.

The experiments demonstrate that our approach is an effective tool for handling FRBSs' accuracy-interpretability trade-off, i.e, designing FRBSs characterized by various levels of such a trade-off (in particular, for designing highly interpretability-oriented systems of still competitive accuracy). Our approach can also be applied to other types of FRBSs, e.g. those considered in [21, 22].

Acknowledgements. The numerical experiments reported in this paper have been performed using computational equipment purchased in the framework of the EU Operational Programme Innovative Economy (POIG.02.02.00-26-023/09-00) and the EU Operational Programme Development of Eastern Poland (POPW.01.03.00-26-016/09-00).

REFERENCES

[1] K.J. Cios, *Medical Data Mining and Knowledge Discovery*, Physica-Verlag, Springer, New York, 2001.

[2] O. Maimon and L. Rokach, *Data Mining and Knowledge Discovery Handbook*, Springer, New York, 2005.

[3] J. Ponce and A. Karahoca, *Data Mining and Knowledge Discovery in Real Life Applications*, IN-TECH, Vienna, 2009.

[4] D. Dubois and H. Prade, "What are fuzzy rules and how to use them", *Fuzzy Sets and Systems* 84 (2), 169–185 (1996).

[5] M.J. Gacto, R. Alcala, and F. Herrera, "Interpretability of linguistic fuzzy rule-based systems: an overview of interpretability measures", *Information Sciences* 181 (20), 4340–4360 (2011).

[6] F. Herrera, "Genetic fuzzy systems: taxonomy, current research trends and prospects", *Evolutionary Intelligence* 1 (1), 27–46 (2008).

[7] M.B. Gorzałczany and F. Rudziński, "A modified Pittsburg approach to design a genetic fuzzy rule-based classifier from data", *Lecture Notes in Computer Science* 6113, 88–96 (2010).

[8] M.B. Gorzałczany and F. Rudziński, "Accuracy vs. interpretability of fuzzy rule-based classifiers: an evolutionary approach", *Lecture Notes in Computer Science* 7269, 222–230 (2012).

[9] M.B. Gorzałczany and F. Rudziński, "Genetic fuzzy rule-based modelling of dynamic systems using time series", *Lecture Notes in Computer Science* 7269, 231–239 (2012).

[10] M.B. Gorzałczany and F. Rudziński, "Measurement data in genetic fuzzy modelling of dynamic systems", *Measurements, Automatics, Control* 12, 1420–1423 (2010), (in Polish).

[11] F. Rudziński and J. Piekoszewski, "The maintenance costs estimation of electrical lines with the use of interpretability-oriented genetic fuzzy rule-based systems", *Przegląd Elektrotechniczny* 8, 43–47 (2013).

[12] M. Fazzolari, R. Alcala, Y. Nojima, H. Ishibuchi, and F. Herrera, "A review of the application of multiobjective evolutionary fuzzy systems: current status and further directions", *IEEE Trans. on Fuzzy Systems* 21 (1), 45–65 (2013).

[13] E.H. Ruspini, "A new approach to clustering", *Information and Control* 15 (1), 22–32 (1969).

[14] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength pareto evolutionary algorithm for multi-objetive optimization", *Proc. Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems* 1, 95–100 (2001).

[15] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II", *IEEE Trans. on Evolutionary Computation* 6 (2), 182–197 (2002).

[16] J. Alcalá-Fdez, R. Alcalá, and F. Herrera, "A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning", *IEEE Trans. on Fuzzy Systems* 19 (5), 857–872 (2011).

[17] F. Rudziński, "A multi-objective genetic optimization of interpretability-oriented fuzzy rule-based classifiers", *Applied Soft Computing*, (to be published).

[18] M. Baczyński and B. Jayaram, *Fuzzy Implications*, Studies in Fuzziness and Soft Computing, Springer, Berlin, 2008.

[19] *Machine Learning Database Repository*, University of California at Irvine, (ftp.ics.uci.edu).

[20] L.-X. Wang, *A Course in Fuzzy Systems and Control*, Prentice-Hall, New York, 1998.

[21] S. Osowski, K. Brudzewski, and L. Tran Hoai, "Modified neuro-fuzzy TSK network and its application in electronic nose", *Bull. Pol. Ac.: Tech.* 61 (3), 675–680 (2013).

[22] J. Smoczek, "P1-TS fuzzy scheduling control system design using local pole placement and interval analysis", *Bull. Pol. Ac.: Tech.* 62 (3), 455–464 (2014).