Labeling Consequents of Fuzzy Rules Constructed by Using Heuristic Algorithms of Possibilistic Clustering

Dmitri A. VIATTCHENIN[1,2], Stanislau SHYRAI[1], Aliaksandr DAMARATSKI[2]
[1]Belarusian State University of Informatics and Radio-Electronics, Minsk, Belarus
[2]United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus
viattchenin@mail.ru, ashaman410@gmail.com, a.damaratski@gmail.com

The paper deals with the problem of automatic labeling output variables in Mamdani-type fuzzy rules generated by using heuristic algorithms of possibilistic clustering. The labeling problem in fuzzy clustering and basic concepts the heuristic approach to possibilistic clustering are considered in brief. Labeling consequents procedure is proposed. Experimental results are presented shortly and some preliminary conclusions are made.
Keywords: *data mining, clustering, fuzzy rule, consequent, labeling*

1 Introduction

Fuzzy classifiers play an important role in different data mining approaches. Thus, the problem of generation of fuzzy rules is one of more than important problems in the development of fuzzy classifiers.

There are a number of approaches to learning fuzzy rules from data based on techniques of evolutionary or neural computation, mostly aiming at optimizing parameters of fuzzy rules. From other hand, fuzzy or possibilistic clustering seems to be a very appealing method for learning fuzzy rules since there is a close and canonical connection between fuzzy clusters and fuzzy rules. The fact was shown in [1].

Let us consider in brief some basic concepts. We assume that the training set contains n data pairs. Each pair is made of a m_1-dimensional input-vector and a c-dimensional output-vector. We assume that the number of rules in the fuzzy inference system rule base is c. So, Mamdani and Assilian's [2] fuzzy rule l within the fuzzy inference system is written as follows:

$$\begin{aligned}&\textit{If } \hat{x}^1 \textit{ is } B_l^1 \textit{ and} \ldots \textit{and } \hat{x}^{m_1} \textit{ is } B_l^{m_1} \\ &\textit{then } y_1 \textit{ is } C_1^l \textit{ and} \ldots \textit{and } y_c \textit{ is } C_c^l\end{aligned}, \qquad (1)$$

where input variables $\hat{x}^{t_1}$, $t_1 = 1,\ldots,m_1$ are antecedents and output variables y_l, $l = 1,\ldots,c$ are consequents of fuzzy rules, $B_l^{t_1}$, $t_1 \in \{1,\ldots,m_1\}$ and C_l^l, $l \in \{1,\ldots,c\}$ are fuzzy sets that define an input and output space partitioning. A fuzzy classifier which is described by a set of fuzzy classification rules with the form (1) is the multiple inputs, multiple outputs system.

The principal idea of extracting fuzzy classification rules based on fuzzy clustering was outlined in [1] and the idea is the following. Each fuzzy cluster is assumed to be assigned to one class for classification and the membership grades of the data to the clusters determine the degree to which they can be classified as a member of the corresponding class. So, with a fuzzy cluster that is assigned to the some class we can associate a linguistic rule. The fuzzy cluster is projected into each single dimension leading to a fuzzy set on the real numbers. An approximation of the fuzzy set by projecting only the data set and computing the convex hull of this projected fuzzy set or approximating it by a trapezoidal or triangular membership function is used for the fuzzy rules obtaining.

The idea of extracting fuzzy classification rules based on possibilistic clustering [3] is similar to the idea of deriving fuzzy rules based on fuzzy clustering.

On the other hand, a heuristic approach to possibilistic clustering was outlined in [4] and the approach was developed in other publications. Moreover, a method of the rapid extracting fuzzy rules based on results of the heuristic possibilistic clustering of the training data set was also proposed in [4] and the method is very effective in comparison with the method based on fuzzy clustering results. The idea of deriving fuzzy classification rules from the training data can be formulated as follows: the training data set is divided into homogeneous group and a fuzzy rule is associated to each group.

However, names should be assigned to each output variable y_l, $l = 1,\ldots,c$. The process of assigning names to output variables is connected with the problem of interpretation of classification results and a labeling procedure in clustering.

The main goal of this paper is a consideration of an approach to automatic labeling consequents of fuzzy rules generated by heuristic algorithms of possibilistic clustering. The contents of this paper is as follows: in the second section a labeling problem in fuzzy clustering is described, in the third section basic concepts of the heuristic approach to possibilistic clustering are considered, in the fourth section a labeling procedure for fuzzy rules consequents is described, in the fifth a numerical example of application of the proposed procedure to fuzzy rules generated from the Anderson's Iris data set are given, and some final remarks are stated in the sixth section.

2. Related works

The most widespread approach in fuzzy clustering is the optimization approach. Most optimization fuzzy clustering algorithms aim at minimizing an objective function that evaluates the partition of the data into a given number of fuzzy clusters.

All objective function-based fuzzy clustering algorithms can in general be divided into two types: object versus relational.

The object data clustering methods can be applied if the objects are represented as points in some multidimensional space $I^{m_1}(X)$. In other words, the data which is composed of n objects and m_1 attributes is denoted as $\hat{X}_{n\times m_1} = [\hat{x}_i^{t_1}]$, $i = 1,\ldots,n$, $t_1 = 1,\ldots,m_1$ and the data are called sometimes the two-way data [5]. Let $X = \{x_1,\ldots,x_n\}$ is the set of objects. So, the two-way data matrix can be represented as follows:

$$\hat{X}_{n\times m_1} = \begin{pmatrix} \hat{x}_1^1 & \hat{x}_1^2 & \ldots & \hat{x}_1^{m_1} \\ \hat{x}_2^2 & \hat{x}_2^2 & \ldots & \hat{x}_2^{m_1} \\ \ldots & \ldots & \ldots & \ldots \\ \hat{x}_n^1 & \hat{x}_n^2 & \ldots & \hat{x}_n^{m_1} \end{pmatrix}. \quad (2)$$

So, the two-way data matrix can be represented as $\hat{X} = (\hat{x}^1,\ldots,\hat{x}^{m_1})$ using n-dimensional column vectors $\hat{x}^{t_1}$, $t_1 = 1,\ldots,m_1$, composed of the elements of the t_1-th column of $\hat{X}$.

The traditional optimization methods of fuzzy clustering are based on the concept of fuzzy c-partition [1]. The initial set $X = \{x_1,\ldots,x_n\}$ of n objects represented by the matrix of similarity coefficients, the matrix of dissimilarity coefficients or the matrix of object attributes, should be divided into c fuzzy clusters. Namely, the grade u_{li}, $1 \le l \le c$, $1 \le i \le n$ to which an object x_i belongs to the fuzzy cluster A^l should be determined. For each object x_i, $i = 1,\ldots,n$ the grades of membership should satisfy the conditions of a fuzzy c-partition:

$$\sum_{l=1}^{c} u_{li} = 1, \quad 1 \le i \le n, \quad 0 \le u_{li} \le 1, \quad 1 \le l \le c. \quad (3)$$

In other words, the family of fuzzy sets $P(X) = \{A^l \mid l = \overline{1,c}, c \le n\}$ is the fuzzy c-partition of the initial set of objects

$X = \{x_1,...,x_n\}$ if condition (3) is met. Fuzzy c-partition $P(X)$ may be described with the aid of a partition matrix $P_{c\times n} = [u_{li}]$, $l = 1,\ldots,c$, $i = 1,\ldots,n$. The set of all fuzzy c-partitions will be denoted by Π. So, the fuzzy problem formulation in cluster analysis can be defined as the optimization task $Q \to \underset{P(X)\in\Pi}{extr}$ under the constraints (3), where Q is a fuzzy objective function.

The best known optimization approach to fuzzy clustering is the method of fuzzy c-means [6]. The FCM-algorithm is based on an iterative optimization of the fuzzy objective function, which takes the form:

$$Q_{FCM}(P,\mathrm{T}) = \sum_{l=1}^{c}\sum_{i=1}^{n} u_{li}^{\gamma}\left\| x_i - \bar{\tau}^{l} \right\|^2, \quad (4)$$

where u_{li}, $l = 1,\ldots,c$, $i = 1,\ldots,n$ is the membership degree, x_i, $i \in \{1,\ldots,n\}$ is the data point, $\mathrm{T} = \{\bar{\tau}^1,\ldots,\bar{\tau}^c\}$ is the set fuzzy clusters prototypes, and $\gamma > 1$ is the weighting exponent.

The purpose of the classification task is to obtain the solutions $P(X)$ and $\bar{\tau}^1,\ldots,\bar{\tau}^c$ which minimize equation (4). Some other similar objective function-based fuzzy clustering algorithms are considered in [1], [5] and [6] in detail.

However, the condition of fuzzy c-partition is very difficult from essential positions. So, a possibilistic approach to clustering was proposed by Krishnapuram and Keller in [3] and developed by other researchers. Major algorithms of possibilistic clustering are objective function-based procedures.

A concept of possibilistic partition is a basis of possibilistic clustering methods and membership values μ_{li}, $l = 1,\ldots,c$, $i = 1,\ldots,n$ can be interpreted as the values of typicality degree. For each object x_i, $i = 1,\ldots,n$ the grades of membership should satisfy the conditions of a possibilistic partition:

$$\sum_{l=1}^{c}\mu_{li} > 0,\ 0 \le \mu_{li} \le 1. \quad (5)$$

So, the family of fuzzy sets $\mathrm{Y}(X) = \{A^l \mid l = \overline{1,c}, c \le n\}$ is the possibilistic partition of the initial set of objects $X = \{x_1,...,x_n\}$ if condition (5) is met. Obviously that the conditions of the possibilistic partition (5) are more flexible than the conditions of the fuzzy c-partition (3).

In order to be applying the found cluster prototype as classifiers, they need to be given reasonable names. One can then use these names as column titles of the membership matrix when using the recall function of the FCM-algorithm. This helps in the interpretation of the results.

The process of assigning class names to cluster prototypes is called labeling. A labeling method for the fuzzy c-means method is to inspect the cluster prototypes and their respective membership values of the various attributes, and to assign a label manually.

However, usually, it is already known when training a classifier which objects belong to which classes. This information can be taken into account to use so as automatic the fuzzy c-means labeling process. The corresponding labeling procedure is described in [7] in detail. The principal idea of the procedure is to present a sample of objects to the FCM-classifier whose class membership are known in the hard form of 0 or 1 values and have also been calculated by the procedure. By means of the given cluster membership values for each fuzzy cluster prototype, the fuzzy cluster prototypes can be associated with their respective classes.

On the other hand, all objective function-based fuzzy clustering algorithms are iterative procedures and the initial fuzzy c-partition $P(X)$ is initialized randomly. So, coordinates of fuzzy clusters prototypes and values of membership functions will be different in each experiment for the same data set, because the result of classification is sensitive to initialization. Moreover, major objective function-based fuzzy

clustering algorithms are need for using some validity measures [1] for determining the most "plausible" number c of fuzzy clusters in the sought fuzzy c-partition $P(X)$. So, a problem of rapid automatic labeling is arises.
The effective labeling procedure for heuristic algorithms of possibilistic clustering was proposed in [8] and the procedure is the basis of the labelling procedure for consequents of derived fuzzy rules. However, basic definitions of the heuristic approach to possibilistic clustering should be considered in the first place.

3. Basic concepts of the heuristic approach to possibilistic clustering

Let us remind the basic concepts of the heuristic method of possibilistic clustering [4]. Let $X = \{x_1,...,x_n\}$ be the initial set of elements and $T: X \times X \to [0,1]$ some fuzzy tolerance on X with $\mu_T(x_i, x_j) \in [0,1]$, $\forall x_i, x_j \in X$ being its membership function. Let α be the α-level value of the fuzzy tolerance T, $\alpha \in (0,1]$. Columns or rows of the fuzzy tolerance matrix are fuzzy sets $\{A^1,...,A^n\}$ on the universal set X. Let A^l, $l \in \{1,\ldots,n\}$ be a fuzzy set on X with $\mu_{A^l}(x_i) \in [0,1]$, $\forall x_i \in X$ being its membership function. The α-level fuzzy set $A^l_{(\alpha)} = \{(x_i, \mu_{A^l}(x_i)) \mid \mu_{A^l}(x_i) \geq \alpha, x_i \in X\}$ is fuzzy α-cluster. So, $A^l_{(\alpha)} \subseteq A^l$, $\alpha \in (0,1]$, $A^l \in \{A^1,\ldots,A^n\}$ and $\mu_{A^l}(x_i)$ is the membership degree of the element $x_i \in X$ for some fuzzy α-cluster $A^l_{(\alpha)}$, $\alpha \in (0,1]$, $l \in \{1,\ldots,n\}$. The membership degree will be denoted μ_{li} in further considerations. The membership degree of the element $x_i \in X$ for some fuzzy α-cluster $A^l_{(\alpha)}$, $\alpha \in (0,1]$, $l \in \{1,\ldots,n\}$ can be defined as a

$$\mu_{li} = \begin{cases} \mu_{A^l}(x_i), & x_i \in A^l_\alpha \\ 0, & otherwise \end{cases}, \tag{6}$$

where $A^l_\alpha = \{x_i \in X \mid \mu_{A^l}(x_i) \geq \alpha\}$, $\alpha \in (0,1]$ is the α-level of a fuzzy set A^l and the α-level is the support of the fuzzy α-cluster $A^l_{(\alpha)}$, $A^l_\alpha = Supp(A^l_{(\alpha)})$. The value of α is the tolerance threshold of fuzzy α-cluster elements.
Let $\{A^1_{(\alpha)},...,A^n_{(\alpha)}\}$ be the family of fuzzy α-clusters for some α. The point $\tau^l_e \in A^l_\alpha$, for which

$$\tau^l_e = \arg\max_{x_i} \mu_{li}, \ \forall x_i \in A^l_\alpha, \tag{7}$$

is called a typical point of the fuzzy α-cluster $A^l_{(\alpha)}$, $\alpha \in (0,1]$, $l \in [1,n]$. A set $K(A^l_{(\alpha)}) = \{\tau^l_1,\ldots,\tau^l_{|l|}\}$ of typical points of the fuzzy cluster $A^l_{(\alpha)}$ is a kernel of the fuzzy cluster and $card\big(K(A^l_{(\alpha)})\big) = |l|$ is a cardinality of the kernel. If the fuzzy cluster have an unique typical point, then $|l| = 1$.
Let $R^\alpha_z(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, 2 \leq c \leq n\}$ be a family of fuzzy α-clusters for some value of tolerance threshold α, which are generated by a fuzzy tolerance T on the initial set of elements $X = \{x_1,...,x_n\}$. If condition

$$\sum_{l=1}^{c} \mu_{li} > 0, \ \forall x_i \in X, \tag{8}$$

is met for all $A^l_{(\alpha)}$, $l = \overline{1,c}$, $c \leq n$, then the family is the allotment of elements of the set $X = \{x_1,...,x_n\}$ among fuzzy α-clusters $\{A^l_{(\alpha)}, l = \overline{1,c}, 2 \leq c \leq n\}$ for some value of the tolerance threshold α. It should be noted that several allotments $R^\alpha_z(X)$ can exist for some tolerance threshold α. That is why symbol z is the index of an allotment.
Obviously, the definition of the allotment among fuzzy clusters (8) is similar to the definition of the possibilistic partition (5). So, the allotment among fuzzy clusters can be considered as the possibilistic partition

and fuzzy clusters in the sense of (6) are elements of the possibilistic partition.
Allotment $R_I^{\alpha}(X)=\{A_{(\alpha)}^{l}\,|\,l=\overline{1,n},\alpha\in(0,1]\}$ of the set of objects among n fuzzy clusters for some tolerance threshold $\alpha\in(0,1]$ is the initial allotment of the set $X=\{x_1,...,x_n\}$. In other words, if initial data are represented by a matrix of some fuzzy T then lines or columns of the matrix are fuzzy sets $A^l\subseteq X$, $l=\overline{1,n}$ and α-level fuzzy sets $A_{(\alpha)}^{l}$, $l=\overline{1,c}$, $\alpha\in(0,1]$ are fuzzy clusters. These fuzzy clusters constitute an initial allotment for some tolerance threshold α and they can be considered as clustering components. If some allotment $R_{c(z)}^{\alpha}(X)=\{A_{(\alpha)}^{l}\,|\,l=\overline{1,c},c\le n\}$ corresponds to the formulation of a concrete problem, then this allotment is an adequate allotment. In particular, if a condition

$$\bigcup_{l=1}^{c}A_{\alpha}^{l}=X, \quad (9)$$

and a condition

$$card(A_{\alpha}^{l}\cap A_{\alpha}^{m})=0,\ \forall A_{(\alpha)}^{l},A_{(\alpha)}^{m},\ l\ne m,\alpha\in(0,1] \quad (10)$$

are met for all fuzzy clusters $A_{(\alpha)}^{l}$, $l=\overline{1,c}$ of some allotment $R_{c(z)}^{\alpha}(X)=\{A_{(\alpha)}^{l}\,|\,l=\overline{1,c},c\le n\}$ for a value $\alpha\in(0,1]$, then the allotment is the allotment among fully separate fuzzy clusters.

Fuzzy clusters in the sense of definition (6) can have an intersection area. If the intersection area of any pair of different fuzzy clusters is an empty set, then conditions (9) and (10) are met and fuzzy clusters are called fully separate fuzzy clusters. Otherwise, fuzzy clusters are called particularly separate fuzzy clusters and $w\in\{0,\ldots,n\}$ is the maximum number of elements in the intersection area of different fuzzy clusters. For $w=0$ fuzzy clusters are fully separate fuzzy clusters. Thus, the conditions (9) and (10) can be generalized for a case of particularly separate fuzzy clusters. So, a condition

$$\sum_{l=1}^{c}card(A_{\alpha}^{l})\ge card(X),\ \forall A_{(\alpha)}^{l}\in R_{c(z)}^{\alpha}(X),\ \alpha\in(0,1],card(R_{c(z)}^{\alpha}(X))=c \quad (11)$$

and a condition

$$card(A_{\alpha}^{l}\cap A_{\alpha}^{m})\le w,\ \forall A_{(\alpha)}^{l},A_{(\alpha)}^{m},\ l\ne m,\alpha\in(0,1] \quad (12)$$

are generalizations of conditions (9) and (10). Obviously, if $w=0$ in conditions (11) and (12) then conditions (9) and (10) are met. The adequate allotment $R_{c(z)}^{\alpha}(X)$ for some value of tolerance threshold $\alpha\in(0,1]$ is a family of fuzzy clusters which are elements of the initial allotment $R_I^{\alpha}(X)$ for the value of α and the family of fuzzy clusters should satisfy the conditions (11) and (12). So, the construction of adequate allotments $R_{c(z)}^{\alpha}(X)=\{A_{(\alpha)}^{l}\,|\,l=\overline{1,c},c\le n\}$ for every α is a trivial problem of combinatorics.

Allotment $R_P^{\alpha}(X)=\{A_{(\alpha)}^{l}\,|\,l=\overline{1,c}\}$ of the set of objects among the minimal number c, $2\le c\le n$ of fully separate fuzzy clusters for some tolerance threshold $\alpha\in(0,1]$ is the principal allotment of the set $X=\{x_1,...,x_n\}$.

Several adequate allotments can exist. Thus, the problem consists in the selection of the unique adequate allotment $R_c^{*}(X)$ from the set B of adequate allotments, $B=\{R_{c(z)}^{\alpha}(X)\}$, which is the class of possible solutions of the concrete classification problem. The selection of the unique adequate allotment $R_c^{*}(X)$ from the set $B=\{R_{c(z)}^{\alpha}(X)\}$ of adequate allotments must be made on the basis of evaluation of allotments. In particular, the criterion

$$F(R_{c(z)}^{\alpha}(X),\alpha)=\sum_{l=1}^{c}\frac{1}{n_l}\sum_{i=1}^{n_l}\mu_{li}-\alpha\cdot c, \quad (13)$$

where c is the number of fuzzy clusters in the allotment $R_{c(z)}^{\alpha}(X)$ and $n_l=card(A_{\alpha}^{l})$,

$A^l_{(\alpha)} \in R^{\alpha}_{c(z)}(X)$ is the number of elements in the support of the fuzzy cluster $A^l_{(\alpha)}$, can be used for evaluation of allotments. Maximum of criterion (13) corresponds to the best allotment of objects among c fuzzy clusters. So, the classification problem can be characterized formally as determination of the solution $R^*_c(X)$ satisfying

$$R^*_c(X) = \arg \max_{R^{\alpha}_{c(z)}(X) \in B} F(R^{\alpha}_{c(z)}(X), \alpha), \qquad (14)$$

The problem of cluster analysis can be defined in general as the problem of discovering the unique allotment $R^*_c(X)$, resulting from the classification process and detection of fixed or unknown number c of fuzzy clusters can be considered as the aim of classification.

Thus, the problem of cluster analysis can be defined as the problem of discovering the unique allotment $R^*_c(X)$, resulting from the classification process and detection of fixed or unknown number c of fuzzy α-clusters can be considered as the aim of classification.

Direct heuristic algorithms of possibilistic clustering can be divided into two types: relational versus prototype-based. A fuzzy tolerance relation matrix is a matrix of the initial data for the direct heuristic relational algorithms of possibilistic clustering and a matrix of attributes (2) is a matrix for the prototype-based algorithms. In particular, the group of direct relational heuristic algorithms of possibilistic clustering includes

- D-AFC(c)-algorithm: using the construction of the allotment among given number c of partially separate fuzzy clusters;
- D-PAFC-algorithm: using the construction of the principal allotment among an unknown minimal number of at least c fully separate fuzzy clusters;
- D-AFC-PS(c)-algorithm: using the partially supervised construction of the allotment among given number c of partially separate fuzzy clusters.

On the other hand, the family of direct prototype-based heuristic algorithms of possibilistic clustering includes

- D-AFC-TC-algorithm: using the construction of the allotment among an unknown number c of fully separate fuzzy clusters;
- D-PAFC-TC-algorithm: using the construction of the principal allotment among an unknown minimal number of at least c fully separate fuzzy clusters;
- D-AFC-TC(α)-algorithm: using the construction of the allotment among an unknown number c of fully separate fuzzy clusters with respect to the minimal value α of the tolerance threshold.

It should be noted that these direct prototype-based heuristic possibilistic clustering algorithms are based on a transitive closure of an initial fuzzy tolerance relation.

On the other hand, a family of direct prototype-based heuristic possibilistic clustering algorithms based on a transitive approximation of a fuzzy tolerance is proposed in [9].

So, the matrix of memberships $R^*_c(X) = [\mu_{li}]$, the value α of the tolerance threshold and the set of kernels $\{K(A^1_{(\alpha)}), \ldots, K(A^c_{(\alpha)})\}$ are results of classification. The results will be constant in each experiment for the same data set, because the sought clustering structure $R^*_c(X)$ of the set of objects X is based directly on the formal definition of fuzzy cluster and the possibilistic memberships are determined directly from the values of the pairwise similarity of objects.

The training data matrix (2) and clustering results are a basis for constructing of Mamdani-type fuzzy rules (1). The corresponding methodology described in [4] in detail.

4. A labeling procedure for fuzzy rules consequents

Fuzzy classifier can be generated directly by some heuristic algorithm of possibilistic clustering [4]. A fuzzy rule is associated to each fuzzy α-cluster of the obtained allotment, $R_c^*(X)$. So, a number of fuzzy rules is equal to a number of fuzzy α-clusters and equal to a number of output variables y_l, $l = 1,\ldots,c$.

The results obtained from heuristic algorithms of possibilistic clustering are stable. The set of kernels $\{K(A^1_{(\alpha)}),\ldots,K(A^c_{(\alpha)})\}$ and the set of labels $\{label\,1,\ldots,label\,c\}$ are inputs for a labeling procedure [8]. We assume that a condition $card\big(K(A^l_{(\alpha)})\big)=1$ is met for each kernel $K(A^l_{(\alpha)})$, $l=\overline{1,c}$. In other words, the set of typical points $\{\tau^1,\ldots,\tau^c\}$ is given.

Each output variable $\{y_1,\ldots,y_c\}$ corresponds to a fuzzy α-cluster of the obtained allotment, $R_c^*(X)$. There is a two-step procedure which can be described as follows:

1. Perform the following operations for each typical point τ^l, $l=\overline{1,c}$ and each label $label\,m$, $m=\overline{1,c}$:
   1.1 Let $l:=1$ and $m:=1$;
   1.2 Check the following condition:
      if τ^l corresponds to $label\,m$
      then the label $label\,m$ is the label for the typical point τ^l and go to step 1.3 **else** $m:=m+1$ and go to step 1.2;
   1.3 Check the following condition:
      if the typical point τ^l is labeled
      then $l:=l+1$ and go to step 1.2
      else go to step 1.4;
   1.4 Check the following condition:
      if all typical points τ^l, $l=\overline{1,c}$ are labeled **then** go to step 2.
2. Perform the following operations for each output variable y_l, $l=\overline{1,c}$ and each typical point τ^l, $l=\overline{1,c}$:
   2.1 Let $l:=1$;
   2.2 A label of typical point τ^l should be assigned to output variable y_l;
   2.3 Check the following condition:
      if a condition $l<c$ is met **then** $l:=l+1$ and go to step 2.2 **else** stop.

That is why the proposed labeling procedure for consequents of fuzzy rules can be considered as an extended version of the procedure for labeling fuzzy α-clusters [8].

5. An illustrative example

The Anderson's Iris database [10] is the most known database to be found in the pattern recognition literature. The data set represents different categories of Iris plants having four attribute values. The four attribute values represent the sepal length, sepal width, petal length and petal width measured for 150 irises. It has three classes Setosa, Versicolor and Virginica, with 50 samples per class. Examples of records in the database are presented in Table 1.

Table 1. Examples of records in the Iris database

Numbers of objects	Attributes				Labels of classes
	Sepal length	Sepal width	Petal length	Petal width	
…	…	…	…	…	…
18	5.1	3.3	1.7	0.5	SETOSA
…	…	…	…	…	…
48	5.5	2.6	4.4	1.2	VERSICOLOR
…	…	…	…	…	…

108	5.6	2.8	4.9	2.0	VIRGINICA
…	…	…	…	…	…

The Anderson's Iris data form the matrix of attributes $\hat{X}_{150\times 4} = [\hat{x}_i^{t_1}]$, $i = 1,\dots,150$, $t_1 = 1,\dots,4$, where the sepal length is denoted by $\hat{x}^1$, sepal width – by $\hat{x}^2$, petal length – by $\hat{x}^3$ and petal width – by $\hat{x}^4$. The data was normalized as follows:

$$x_i^{t_1} = \frac{\hat{x}_i^{t_1}}{\max_i \hat{x}_i^{t_1}}. \tag{15}$$

So, each object can be considered as a fuzzy set x_i, $i = 1,\dots,150$ and $x_i^{t_1} = \mu_{x_i}(x^{t_1}) \in [0,1]$, $i = 1,\dots,150$, $t_1 = 1,\dots,4$, are their membership functions. The matrix of coefficients of pair wise dissimilarity between objects $I = [\mu_I(x_i, x_j)]$, $i, j = 1,\dots,150$ can be obtained after application of some distance to the matrix of normalized data $X_{150\times 4} = [\mu_{x_i}(x^{t_1})]$, $i = 1,\dots,150$, $t_1 = 1,\dots,4$. In particular, the normalized Euclidean distance [11]

$$e(x_i, x_j) = \sqrt{\frac{1}{m_1}\sum_{t_1=1}^{m_1}\left(\mu_{x_i}(x^{t_1}) - \mu_{x_j}(x^{t_1})\right)^2}. \tag{16}$$

was applied to the normalized data.

The matrix of fuzzy tolerance $T = [\mu_T(x_i, x_j)]$ was obtained after application of complement operation

$$\mu_T(x_i, x_j) = 1 - \mu_I(x_i, x_j), \tag{17}$$

to the matrix of fuzzy intolerance $I = [\mu_I(x_i, x_j)]$, $i, j = 1,\dots,150$ obtained from previous operations. The labeled training data is shown in Fig. 1.

By executing the D-AFC(c)-algorithm [4] for $c = 3$ using the normalized Euclidean distance (16), we obtain that the typical point of the first class τ^1 is the object x_{23}, the typical point of the second class τ^2 is the object x_{95}, and the typical point of the third class τ^3 is the object x_{98}. The clustering result is presented in Fig. 2. The set of labels is {SETOSA, VERSICOLOR, VIRGINICA} and these labels were assigned to corresponding output variables. The performance of the generated fuzzy classifier is shown in Fig. 3 in which $l = 1,\dots,3$ is the number rule.

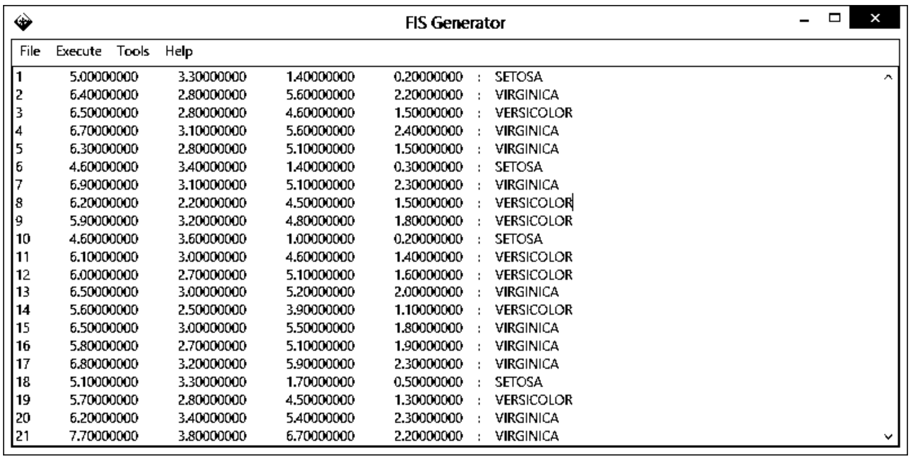

Fig. 1. The training data set

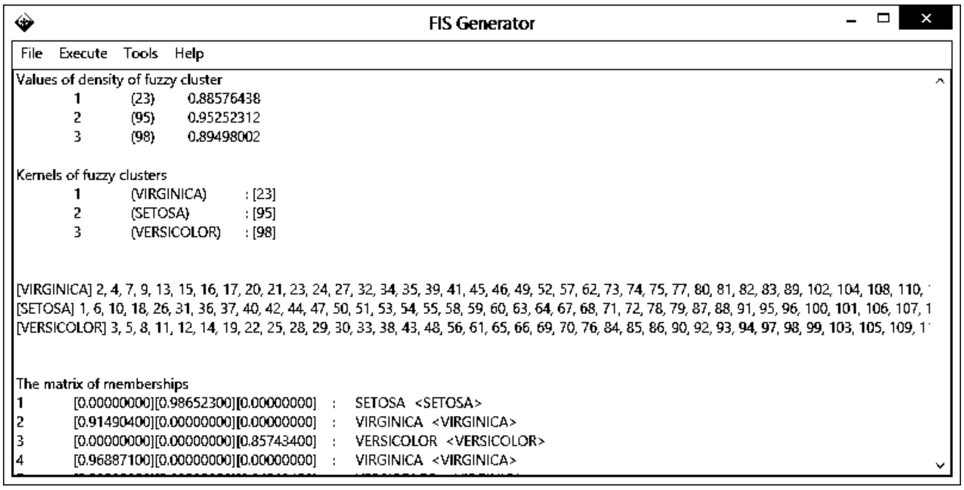

Fig. 2. The clustering result

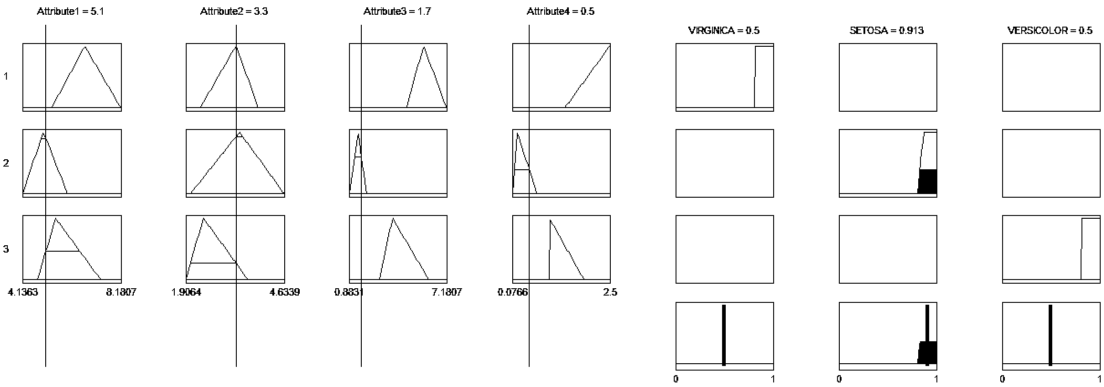

Fig. 3. Performance of the fuzzy classifier which was generated from Anderson's Iris data

So, the result of application the proposed labeling procedure seems to be satisfactory.

6 Conclusions

The fast procedure for labeling consequents of fuzzy rules generated by using heuristic possibilistic clustering results is proposed in the paper. Stability of results of heuristic possibilistic clustering is a basis of the developed procedure. The proposed procedure can be very useful in process of adding new records to database. For the purpose, all records in a database can be classified by using heuristic possibilistic clustering and fuzzy classifier can be generated on a basis of clustering results. That is why the corresponding label can be assigned to a new record which added to the database. These perspectives for investigations are of great interest both from the theoretical point of view and from the practical one as well.

Acknowledgment

This investigation was done in framework of a project of the scientific program "Monitoring-SG" of the Union State of Russia and Belarus.

References

[1] F. Höppner, F. Klawonn, R. Kruse and T. Runkler, *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*, Chichester, Wiley Intersciences, 1999.

[2] E.H. Mamdani and S. Assilian, "An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller", *International*

Journal of Man-Machine Studies, Vol. 7, No. 1, 1975, pp. 1-13.

[3] R. Krishnapuram and J.M. Keller, "A Possibilistic Approach to Clustering", *IEEE Transactions on Fuzzy Systems*, Vol. 1, No. 2, 1993, pp. 98-110.

[4] D.A. Viattchenin, *A Heuristic Approach to Possibilistic Clustering: Algorithms and Applications*, Heidelberg, Springer, 2013.

[5] M. Sato-Ilic and L.C. Jain, *Innovations in Fuzzy Clustering*, Heidelberg, Springer, 2006.

[6] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, New York, Plenum Press, 1981.

[7] *DataEngine: Tutorials and Theory*, Aachen, MIT GmbH, 1999.

[8] D.A. Viattchenin, A. Damaratski, E. Nikolaenya and S. Shyrai, "An Outline for an Approach to Automatic Labeling for Interpretation of Heuristic Possibilistic Clustering Results," *Proceedings of the 12th Int. Conf. on Pattern Recognition and Information Processing*, pp. 290-294, Minsk, Belarus, 28-30 May 2014.

[9] D.A. Viattchenin and A. Damaratski, "Direct Heuristic Algorithms of Possibilistic Clustering Based on Transitive Approximation of Fuzzy Tolerance", *Informatica Economica Journal*, Vol. 17, No. 3, 2013, pp. 5-15.

[10] E. Anderson, "The Irises of the Gaspe Peninsula", *Bulletin of the American Iris Society*, Vol. 59, No. 1, 1935, pp. 2-5.

[11] A. Kaufmann, *Introduction to the Theory of Fuzzy Subsets*, New York, Academic Press, 1975.

Dmitri A. VIATTCHENIN was born on 13.12.1972 in Moscow, Russia. In 1994 he graduated the Department of Mathematical Modelling and Data Analysis of the Faculty of Applied Mathematics and Informatics in Belarusian State University. He defended his PhD in the field of philosophy of sciences and technology in 1998. He got the title of Associate Professor in the specialty informatics and computer engineering in 2013. He is currently an associate professor in the Department of Software Information Technology of the Belarusian State University of Informatics and Radio-Electronics and a leading researcher in the Laboratory of Information Protection Problems of the United Institute of Informatics Problems of the National Academy of Sciences of Belarus. His research interests include techniques of fuzzy and possibilistic clustering, fuzzy control, intelligent decision making systems, simulation of complex systems, and philosophical aspects of artificial intelligence. He is the author of over 135 publications including papers in international and domestic journals and conference proceedings, 3 monographs and 1 edited volume. He is also a reviewer various journals, for example Fuzzy Sets and Systems, Neurocomputing, Journal of Information and Organizational Sciences.

Stanislau SHYRAI was born on 28.09.1991 in Stavropol, Russia. He graduated from the Faculty of Computer Systems and Networks of the Belarusian State University of Informatics and Radio-Electronics in 2014. At present he is a lecturer in the Department of Software Information Technology of the Belarusian State University of Informatics and Radio-Electronics. His interest domains related to computer science are: software development methodologies, intuitionistic fuzzy sets, fuzzy cluster analysis, fuzzy classifiers, information systems and databases.

Aliaksandr DAMARATSKI was born on 29.11.1983 in Mogilev, Belarus. He has graduated the Faculty of Computer Design of Belarusian State University of Informatics and Radio-Electronics in 2005. He is currently a junior researcher in the Laboratory of Information Protection Problems of the United Institute of Informatics Problems of the National Academy of Sciences of Belarus. His work focuses on possibilistic clustering, image processing, decision making, developing software, and computer networks analysis. He is the author of more than 25 papers in journals and conference proceedings.