

Derivation of Fuzzy Rules from Interval-Valued Data

Dmitri A. Viattchenin

Senior Researcher of the Laboratory of Images Recognition and Processing
United Institute of Informatics Problems of the National Academy of Sciences of Belarus
Surganov St. 6, 220012 Minsk, BELARUS

ABSTRACT

Fuzzy inference systems are widely used for classification and control. They can be designed from the training data. This paper describes a technique for deriving fuzzy classification rules from the interval-valued data. The technique based on a heuristic method of possibilistic clustering and a special method of the interval-valued data preprocessing. Basic concepts of the heuristic method of possibilistic clustering based on the allotment concept are described and the method of the interval-valued data preprocessing is also given. The method of constructing of fuzzy rules based on clustering results is presented. An illustrative example of the method's application to the Sato and Jain's interval-valued data is carried out. Preliminary conclusions are formulated.

General Terms

Soft Computing, Data Mining, Pattern Recognition

Keywords

Possibilistic clustering, fuzzy cluster, typical point, tolerance threshold, fuzzy classification rule, interval-valued data

1. INTRODUCTION

Some remarks on methods of extracting of fuzzy rules based on fuzzy clustering are considered in the first subsection. The second subsection includes a brief review of uncertain data clustering methods and the aims of the paper.

1.1 Fuzzy Rules and Fuzzy Clustering

Fuzzy inference systems are very effective tool for identification and control problems solving. Fuzzy rule base is the key element of any fuzzy inference system. So, the problem of generation of fuzzy classification rules is one of the more important problems in the development of fuzzy systems models.

In general, Mamdani's [4] rule l within the fuzzy inference system is written as follows:

$$\begin{aligned} & \text{If } \hat{x}^1 \text{ is } B_t^1 \text{ and } \dots \text{ and } \hat{x}^m \text{ is } B_t^m \\ & \text{then } y_1 \text{ is } C_l^1 \text{ and } \dots \text{ and } y_c \text{ is } C_l^c \end{aligned} \quad (1)$$

where B_t^l , $t \in \{1, \dots, m\}$ and C_l^l , $l \in \{1, \dots, c\}$ are fuzzy sets that define an input and output space partitioning. A fuzzy inference system which is described by a set of fuzzy rules with the form (1) is the multiple inputs, multiple outputs system.

Fuzzy classification rules can be designed either from the expert knowledge or from the data. However, in many cases, expert knowledge either does not exist, or this knowledge is incomplete.

In such situations, the fuzzy inference system must be created from the training data.

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 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.

Fuzzy classification rules can be obtained directly from fuzzy clustering results or possibilistic clustering results. The principal idea of extracting fuzzy classification rules based on fuzzy clustering is the following [1]. 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 rules obtaining [1].

Objective function-based fuzzy clustering algorithms are the most widespread methods in fuzzy clustering [1]. Objective function-based fuzzy clustering algorithms are sensitive to initial partition selection and fuzzy rules depend on the selection of the fuzzy clustering method. All algorithms of possibilistic clustering are also objective functions-based algorithms.

Heuristic algorithms of clustering display low level of a complexity. An outline for a heuristic method of possibilistic clustering was presented in [8], where a basic version of direct possibilistic clustering algorithm was described and the version of the algorithm is called the D-AFC(c)-algorithm [10].

Fuzzy rules can be obtained from heuristic possibilistic clustering results immediately. The corresponding method of the rapid prototyping fuzzy inference systems was proposed in [18]. The method is based on deriving fuzzy classification rules from the data on a basis of clustering results obtained from the D-AFC(c)-algorithm.

1.2 Notes on the Uncertain Data Clustering

Most fuzzy clustering methods are designed for treating crisp data. However, we often have to deal with objects that cannot be described by the quantitative, large or binary signs. In other words, there exists a sign of the object that may assume several values at the same time or, for a given sign; there exists uncertainty in representing the values of this sign. Traditional fuzzy clustering methods cannot be applied directly to such types

of objects. So, a problem of fuzzy clustering of uncertain data arises. Such a need occurs mostly in medicine, biology, chemistry, economy, sociology and some other domains.

Several kinds of uncertainty exist and a number of approaches to the uncertain data fuzzy clustering problem solving were proposed by different researchers. Firstly, Žak [20] uses fuzzy sets to describe the uncertainty of the data and introduces the concepts of fuzzy objects and fuzzy dissimilarity. Using these notions, he proposed hierarchical and non-hierarchical clustering methods of fuzzy objects. Secondly, Yang and Ko [19] proposed a class of fuzzy c-number clustering procedures for fuzzy data clustering. Notable that an extended methodology of fuzzy data clustering based on the D-AFC(c)-algorithm was also proposed in [13]. Thirdly, Sato and Sato [5] consider a problem of fuzzy clustering for three-way data. The purpose of clustering for three-way data is to reveal the latent structure through all the time situations by constructing clusters which take into account not only the similarity between the pair of objects at individual time instants but also the similarity between the patterns of change of observation in time. An alternative approach to clustering of the three-way data was also outlined in [14].

From other hand, Kreinovich and Kosheleva [3] note that interval uncertainty is one of the basic types of uncertainty in clustering problems. In these situations, the only information that we have about the actual value \hat{x}_i^t of some attribute \hat{x}^t , $t \in \{1, \dots, m\}$ of the object x_i , $i \in \{1, \dots, n\}$ is that the value belongs to some interval, $\hat{x}_i^t \in [\hat{x}_i^{t \min}, \hat{x}_i^{t \max}]$. Different clustering methods for the interval-valued data considered in [2], [6], [16] and [17].

The problem of clustering for interval-valued data is very important in medicine, sociology, economics, technology, and military applications. Moreover, the training data for constructing fuzzy inference system can be given as the interval-valued data.

The main goal of the paper is developing method of the rapid prototyping fuzzy inference systems in the case of the interval-valued training data set. The contents of this paper is as follows: in the second section basic concepts of the possibilistic clustering method based on the concept of allotment among fuzzy clusters are outlined and a method of the interval-valued data preprocessing is outlined, in the third section a method of constructing of fuzzy rules is proposed, in the fourth section an illustrative example of deriving fuzzy rules from the Sato and Jain's interval-valued data are given, in the fifth section preliminary conclusions are formulated and some perspectives are discussed.

2. AN APPROACH TO CLUSTERING OF THE INTERVAL-VALUED DATA

The basic concepts of the heuristic method of possibilistic clustering are considered in the first subsection. The method of the interval-valued data preprocessing is given in the second subsection of the section.

2.1 Basic Concepts of the D-AFC(c)-algorithm

Let us remind the basic concepts of the D-AFC(c)-algorithm. The concept of fuzzy tolerance is the basis for the concept of fuzzy α -cluster. That is why definition of fuzzy tolerance must be considered in the first place.

Let $X = \{x_1, \dots, x_n\}$ be the initial set of elements and $T: X \times X \rightarrow [0,1]$ some binary fuzzy relation on X with $\mu_T(x_i, x_j) \in [0,1]$, $\forall x_i, x_j \in X$ being its membership function. Fuzzy tolerance is the fuzzy binary intransitive relation which possesses the symmetricity property

$$\mu_T(x_i, x_j) = \mu_T(x_j, x_i), \quad \forall x_i, x_j \in X, \quad (2)$$

and the reflexivity property

$$\mu_T(x_i, x_i) = 1, \quad \forall x_i \in X. \quad (3)$$

The notions of powerful fuzzy tolerance, feeble fuzzy tolerance and strict feeble fuzzy tolerance were proposed in [7] and [8], as well. In this context the classical fuzzy tolerance in the sense of (2) and (3) was called usual fuzzy tolerance and this kind of fuzzy tolerance was denoted by T_2 . The feeble fuzzy tolerance is the fuzzy binary intransitive relation which possesses the symmetricity property (2) and the feeble reflexivity property

$$\mu_T(x_i, x_j) \leq \mu_T(x_i, x_i), \quad \forall x_i, x_j \in X. \quad (4)$$

This kind of fuzzy tolerance is denoted by T_1 . The strict feeble fuzzy tolerance is the feeble fuzzy tolerance with strict inequality in (4):

$$\mu_T(x_i, x_j) < \mu_T(x_i, x_i), \quad \forall x_i, x_j \in X. \quad (5)$$

This kind of fuzzy tolerance is denoted by T_0 .

Fuzzy tolerances T_1 and T_0 are subnormal fuzzy relations if the condition $\mu_T(x_i, x_i) < 1, \forall x_i \in X$ is met. The fact was demonstrated in [14] and it is important for representation of the structure of the uncertain data. However, the essence of the method here considered does not depend on the kind of fuzzy tolerance. That is why the method herein is described for any fuzzy tolerance T .

Let us consider the general definition of fuzzy cluster, the concept of the fuzzy cluster's typical point and the concept of the fuzzy allotment of objects. The number c of fuzzy clusters can be equal to the number of objects, n . This is taken into account in further considerations.

Let $X = \{x_1, \dots, x_n\}$ be the initial set of objects. Let T be a fuzzy tolerance on X and α be α -level value of T , $\alpha \in (0,1]$. Columns or lines of the fuzzy tolerance matrix are fuzzy sets $\{A^1, \dots, A^n\}$. Let $\{A^1, \dots, A^n\}$ be fuzzy sets on X , which are generated by a fuzzy tolerance T . The α -level fuzzy set $A_{(\alpha)}^l = \{(x_i, \mu_{A^l}(x_i)) \mid \mu_{A^l}(x_i) \geq \alpha, l \in [1, n]\}$ is fuzzy α -

cluster or, simply, fuzzy cluster. So $A_{(\alpha)}^l \subseteq A^l$, $\alpha \in (0,1]$, $A^l \in \{A^1, \dots, A^n\}$ and μ_{li} is the membership degree of the element $x_i \in X$ for some fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$. Value of α is the tolerance threshold of fuzzy clusters elements.

The membership degree of the element $x_i \in X$ for some fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$ can be defined as a

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

where an α -level $A_{(\alpha)}^l = \{x_i \in X \mid \mu_{A^l}(x_i) \geq \alpha\}$, $\alpha \in (0,1]$ of a fuzzy set A^l is the support of the fuzzy cluster $A_{(\alpha)}^l$. So, condition $A_{(\alpha)}^l = \text{Supp}(A_{(\alpha)}^l)$ is met for each fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$. Membership degree can be interpreted as a degree of typicality of an element to a fuzzy cluster. The value of a membership function of each element of the fuzzy cluster in the sense of (6) is the degree of similarity of the object to some typical object of fuzzy cluster. Membership degree defines a possibility distribution function for some fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$. The fact was demonstrated in [11] and the possibility distribution function is denoted by $\pi_l(x_i)$.

Let T is a fuzzy tolerance on X , where X is the set of elements, and $\{A_{(\alpha)}^1, \dots, A_{(\alpha)}^n\}$ is the family of fuzzy clusters for some $\alpha \in (0,1]$. The point $\tau_e^l \in A_{(\alpha)}^l$, for which

$$\tau_e^l = \arg \max_x \mu_{li}, \quad \forall x_i \in A_{(\alpha)}^l \quad (7)$$

is called a typical point of the fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$. A fuzzy cluster can have several typical points. That is why symbol e is the index of the typical point. A set $K(A_{(\alpha)}^l) = \{\tau_1^l, \dots, \tau_{|l|}^l\}$ of typical points of the fuzzy cluster $A_{(\alpha)}^l$ is a kernel of the fuzzy cluster and $\text{card } \mathfrak{K}(A_{(\alpha)}^l) = |l|$ is a cardinality of the kernel. Obviously, if the fuzzy cluster have an unique typical point, then $|l|=1$. The kind of a fuzzy cluster depend on the kind of fuzzy tolerance and cardinality of the kernel of the fuzzy cluster. The fact demonstrated in [12].

Let $R_z^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}, 2 \leq c \leq n, \alpha \in (0,1]\}$ be a family of fuzzy clusters for some value of tolerance threshold α , $\alpha \in (0,1]$, which are generated by some fuzzy tolerance T on the initial set of elements $X = \{x_1, \dots, x_n\}$. If condition

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

is met for all fuzzy clusters $A_{(\alpha)}^l \in R_z^\alpha(X)$, $l = \overline{1, c}$, $c \leq n$, then the family is the allotment of elements of the set $X = \{x_1, \dots, x_n\}$ among fuzzy clusters $\{A_{(\alpha)}^l, l = \overline{1, c}, 2 \leq c \leq n\}$ for some value of the tolerance threshold α .

It should be noted that several allotments $R_z^\alpha(X)$ can exist for some tolerance threshold α . That is why symbol z is the index of an allotment. The condition (8) requires that every object x_i , $i = 1, \dots, n$ must be assigned to at least one fuzzy cluster $A_{(\alpha)}^l$, $l = \overline{1, c}$, $c \leq n$ with the membership degree higher than zero. The condition $2 \leq c \leq n$ requires that the number of fuzzy clusters in each allotment $R_z^\alpha(X)$ must be more than two. The allotment among fuzzy clusters can be considered as the special case of the possibilistic partition and fuzzy clusters in the sense of (6) are elements of the possibilistic partition.

If condition

$$\sum_{l=1}^c \text{card}(A_{(\alpha)}^l) \geq \text{card}(X), \quad \forall A_{(\alpha)}^l \in R_z^\alpha(X), \quad (9)$$

$$\alpha \in (0,1], \text{card}(R_z^\alpha(X)) = c$$

and condition

$$\text{card}(A_{(\alpha)}^l \cap A_{(\alpha)}^m) \leq w, \quad \forall A_{(\alpha)}^l, A_{(\alpha)}^m, l \neq m, \alpha \in (0,1], \quad (10)$$

are met for all fuzzy clusters $A_{(\alpha)}^l, l = \overline{1, c}$ of some allotment $R_z^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}, c \leq n\}$ then the allotment is the allotment among particularly separate fuzzy clusters and $0 \leq w \leq n$ is the maximum number of elements in the intersection area of different fuzzy clusters. Obviously, if $w=0$ in conditions (9) and (10) then the intersection area of any pair of different fuzzy cluster is an empty set and fuzzy clusters are fully separate fuzzy clusters.

Detection of fixed c number of fuzzy clusters can be considered as the aim of classification. So, the allotment $R^*(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}\}$ among the given number c of fuzzy clusters and the corresponding value of tolerance threshold α are the results of classification. A plan of the D-AFC(c)-algorithm is presented, for example, in [8] and [14]. Notable that the D-AFC(c)-algorithm is the basic algorithm in the heuristic approach to possibilistic clustering and some other algorithms are proposed in [9], [10], and [15].

2.2 A Method of the Interval-Valued Data Preprocessing

The problem of clustering of interval-valued data can be formulated as follows. Let $X = \{x_1, \dots, x_n\}$ is set of objects, where objects are indexed i , $i = 1, \dots, n$; each object x_i is described by m attributes, indexed t , $t = 1, \dots, m$, so that an object x_i can be represented by vector

$x_i = (\hat{x}_i^1, \dots, \hat{x}_i^t, \dots, \hat{x}_i^m)$; every attribute \hat{x}^t , $t=1, \dots, m$ can be characterized by an interval of values, so that $\hat{x}_i^{t_{\min}(t_{\max})} = [\hat{x}_i^{t_{\min}}, \hat{x}_i^{t_{\max}}]$. Thus, the interval-valued data can be presented by a poly-matrix as follows:

$$\hat{X}_{n \times m} = [\hat{x}_i^{t_{\min}(t_{\max})}], \quad i=1, \dots, n. \quad (11)$$

The purpose of the clustering is to classify the set $X = \{x_1, \dots, x_n\}$ into c fuzzy clusters. So, an allotment $R^*(X)$ among c fuzzy clusters $A_{(\alpha)}^1, \dots, A_{(\alpha)}^n$, $\alpha \in (0,1]$ must be detected.

The D-AFC(c)-algorithm can be applied directly to the data given as a matrix of tolerance coefficients. This means that it can be used with the objects by attributes data by choosing a suitable metric to measure similarity. So, the main problem is the problem of the interval-valued data preprocessing.

The method the three-way data preprocessing was proposed in [14]. The method can be also extended for the interval-valued data [17]. For example, the interval-valued data can be normalized as follows

$$x_i^{t_{\min}(t_{\max})} = \frac{\hat{x}_i^{t_{\min}(t_{\max})}}{\max_{i, t_{\max}} \hat{x}_i^{t_{\min}(t_{\max})}}. \quad (12)$$

So, each object x_i , $i=1, \dots, n$ from the initial set $X = \{x_1, \dots, x_n\}$ can be considered as an interval-valued fuzzy set and $\mu_{x_i}(x^t) = [\mu_{x_i}(x^{t_{\min}}), \mu_{x_i}(x^{t_{\max}})]$, $i=1, \dots, n$, $t=1, \dots, m$ is its membership function, where $\mu_{x_i}(x^{t_{\min}}) \in [0,1]$, $\mu_{x_i}(x^{t_{\max}}) \in [0,1]$. Obviously, that if $\mu_{x_i}(x^{t_{\min}}) = \mu_{x_i}(x^{t_{\max}})$ for all $i=1, \dots, n$ and $t=1, \dots, m$, then x_i is ordinary fuzzy set. So, the matrix (11) will be ordinary matrix of the two-way object data.

The matrix of coefficients of pair wise dissimilarity between objects $I = [\mu_t(x_i, x_j)]$, $i, j=1, \dots, n$ can be obtained after application of some generalization of a distance between fuzzy sets to the matrix of normalized data. The generalizations for type-two fuzzy sets are functions of dissimilarities and the functions of dissimilarities were proposed in [14]. Moreover, these functions of dissimilarities can be extended for interval-valued fuzzy sets. In particular, a generalization of the normalized Euclidean distance for interval-valued fuzzy sets can be described by the expression

$$e_{G_t}(x_i, x_j) = \sqrt{\frac{1}{m} \sum_{t=1}^m \left(\frac{1}{2^2} \sum_{r \in \{t_{\min}, t_{\max}\}} \mu_{x_i}^r(x^t) - \mu_{x_j}^r(x^t) \right)^2}. \quad (13)$$

Obviously, for the ordinary two-way data a usual normalized Euclidean distance for fuzzy sets will be obtained.

The matrix of similarity coefficients $T = [\mu_t(x_i, x_j)]$, $i, j=1, \dots, n$ can be obtained after application of complement operation

$$\mu_T(x_i, x_j) = 1 - \mu_t(x_i, x_j), \quad \forall i, j=1, \dots, n \quad (14)$$

to the matrix of dissimilarity coefficients $I = [\mu_t(x_i, x_j)]$, $i, j=1, \dots, n$ obtained from previous operations.

3. A METHOD OF EXTRACTING FUZZY RULES

A technique of fuzzy rules antecedents learning is presented in the first subsection. A method of consequents learning is given in the second subsection of the section.

3.1 Antecedents Learning

Let us consider a fuzzy rule (1) where B_l^t , $t=1, \dots, m$, $l \in \{1, \dots, c\}$ is a fuzzy set associate with the feature variable \hat{x}^t . Let B_l^t be characterized by the trapezoidal membership function $\gamma_{B_l^t}(\hat{x}^t)$ which is presented in Figure 1.

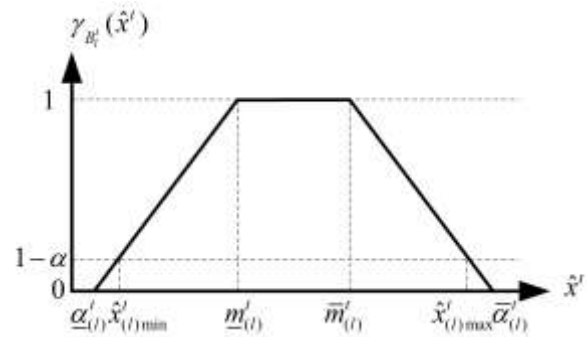


Figure 1. A trapezoidal membership function for an antecedent fuzzy set

So, the fuzzy set B_l^t can be defined by four parameters, $B_l^t = (\underline{a}_{(l)}^t, \underline{m}_{(l)}^t, \bar{m}_{(l)}^t, \bar{a}_{(l)}^t)$. A triangular fuzzy set $B_l^t = (\underline{a}_{(l)}^t, m_{(l)}^t, \bar{a}_{(l)}^t)$ can be considered as a particular case of the trapezoidal fuzzy set where $\underline{m}_{(l)}^t = \bar{m}_{(l)}^t$.

The idea of deriving fuzzy rules from fuzzy clusters is the following [18]. We apply the D-AFC(c)-algorithm to the given data and then obtain for each fuzzy cluster $A_{(\alpha)}^l$, $l \in \{1, \dots, c\}$ a kernel $K(A_{(\alpha)}^l)$ and a support $A_{(\alpha)}^l$. The value of tolerance threshold $\alpha \in (0,1]$, which corresponds to the allotment $R^*(X) = \{A_{(\alpha)}^1, \dots, A_{(\alpha)}^c\}$, is the additional result of classification. We calculate the interval $[\hat{x}_{(l)\min}^t, \hat{x}_{(l)\max}^t]$ of

values of every attribute \hat{x}^t , $t \in \{1, \dots, m\}$ for the support A_{α}^l . The value $\hat{x}_{(l)\min}^t$ can be obtained as follows

$$\hat{x}_{(l)\min}^t = \min_{x_i \in A_{\alpha}^l} \hat{x}^t, \quad \forall t \in \{1, \dots, m\}, \quad \forall l \in \{1, \dots, c\}, \quad (15)$$

and the value $\hat{x}_{(l)\max}^t$, $t \in \{1, \dots, m\}$ can be calculated using a formula

$$\hat{x}_{(l)\max}^t = \max_{x_i \in A_{\alpha}^l} \hat{x}^t, \quad \forall t \in \{1, \dots, m\}, \quad \forall l \in \{1, \dots, c\}. \quad (16)$$

The parameter $\underline{a}_{(l)}^t$ can be obtained as following

$$\gamma_{B_i}(\hat{x}_{(l)\min}^t) = (1 - \alpha), \quad \gamma_{B_i}(\underline{a}_{(l)}^t) = 0, \quad (17)$$

and the parameter $\bar{a}_{(l)}^t$ can be obtained from the conditions

$$\gamma_{B_i}(\hat{x}_{(l)\max}^t) = (1 - \alpha), \quad \gamma_{B_i}(\bar{a}_{(l)}^t) = 0. \quad (18)$$

We calculate the value $\hat{x}_{(l)}^t$ for all typical points $\tau_e^t \in K(A_{(\alpha)}^l)$ of the fuzzy cluster $A_{(\alpha)}^l$, $l \in \{1, \dots, c\}$ as follows:

$$\hat{x}_{(l)}^t = \min_{\tau_e^t \in K(A_{(\alpha)}^l)} \hat{x}^t, \quad \forall e \in \{1, \dots, |l|\}, \quad (19)$$

and the value $\hat{x}_{(l)}^t$ can be obtained from the equation

$$\hat{x}_{(l)}^t = \max_{\tau_e^t \in K(A_{(\alpha)}^l)} \hat{x}^t, \quad \forall e \in \{1, \dots, |l|\}. \quad (20)$$

Thus, the parameter $\underline{m}_{(l)}^t$ can be calculated from the conditions

$$\gamma_{B_i}(\hat{x}_{(l)}^t) = \gamma_{B_i}(\underline{m}_{(l)}^t) = 1, \quad (21)$$

and the parameter $\bar{m}_{(l)}^t$ can be obtained as following

$$\gamma_{B_i}(\hat{x}_{(l)}^t) = \gamma_{B_i}(\bar{m}_{(l)}^t) = 1. \quad (22)$$

So, condition $\hat{x}_{(l)}^t = \underline{m}_{(l)}^t$ and condition $\hat{x}_{(l)}^t = \bar{m}_{(l)}^t$ are met for all input variables \hat{x}^t , $t = 1, \dots, m$. The height $h(A_{(\alpha)}^l) = \sup_{x_i \in A_{\alpha}^l} \mu_{A_{(\alpha)}^l}(x_i)$ of the fuzzy cluster $A_{(\alpha)}^l$, $l \in \{1, \dots, c\}$ must be taken into account because the fuzzy cluster $A_{(\alpha)}^l \in R^*(X)$ can be a subnormal fuzzy set [12], [17]. Obviously, if condition $h(A_{(\alpha)}^l) = 1$ is met for the fuzzy cluster $A_{(\alpha)}^l$ and only one typical point is presented in the fuzzy cluster, then the condition $\hat{x}_{(l)}^t = \hat{x}_{(l)}^t = \underline{m}_{(l)}^t = \bar{m}_{(l)}^t$ is met.

3.2 Consequents Learning

The variables y_l , $l = 1, \dots, c$ are the consequents of fuzzy rules (1), represented by the fuzzy sets C_l^l , $l = 1, \dots, c$ with the membership functions $\gamma_{C_l^l}(y_l)$. Fuzzy sets C_l^l , $l = 1, \dots, c$ can be defined on the interval of memberships $[0, 1]$ and these fuzzy sets can be presented as follows: $C_l^l = (\alpha, \underline{\mu}_l, \bar{\mu}_l, 1)$, where α is the tolerance threshold, $\underline{\mu}_l = \min_{x_i \in A_{\alpha}^l} \mu_{li}$ and $\bar{\mu}_l = \max_{x_i \in A_{\alpha}^l} \mu_{li}$. So, membership functions $\gamma_{C_l^l}(y_l)$ of fuzzy sets C_l^l , $l = 1, \dots, c$ will be trapezoidal membership functions. The situation is shown in Figure 2.

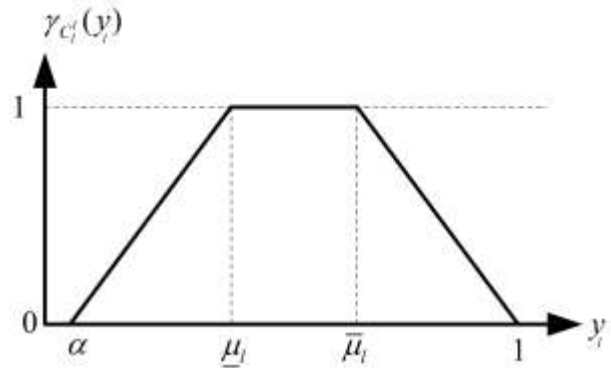


Figure 2. A trapezoidal membership function for a consequent fuzzy set in a general case

Fuzzy clusters can be subnormal fuzzy sets [12]. So, the case which is presented in Figure 2 is the general case.

If the allotment $R^*(X)$ among fully separate fuzzy clusters is obtained and all fuzzy clusters $A_{(\alpha)}^l \in R^*(X)$, $l \in \{1, \dots, c\}$ are normal fuzzy sets then $h(A_{(\alpha)}^l) = 1$, $\forall A_{(\alpha)}^l \in R^*(X)$ and $\bar{\mu}_l = 1$ for each fuzzy cluster $A_{(\alpha)}^l$, $l \in \{1, \dots, c\}$. So, a trapezoidal membership function $\gamma_{C_l^l}(y_l)$ of a fuzzy set C_l^l , $l = 1, \dots, c$ in the case of normal and fully separate fuzzy clusters is presented in Figure 3.

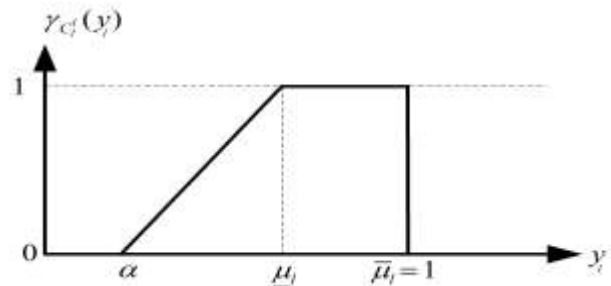


Figure 3. A trapezoidal membership function for a consequent fuzzy set in a case of normal and fully separated fuzzy clusters

Thus, trapezoidal membership functions $\gamma_{C_l^i}(y_l)$ for the fuzzy sets C_l^i , $l=1,\dots,c$ can be constructed on a basis of the clustering results. The empty set $A_\alpha^l = \emptyset$, $l \in \{1,\dots,c\}$ can be correspond to some output variable y_l , $l \in \{1,\dots,c\}$. So, the empty fuzzy set C_l^i will be correspond to the output variable y_l , $l \in \{1,\dots,c\}$ and $\gamma_{C_l^i}(y_l) = 0$ is the membership function of the corresponding fuzzy set C_l^i .

4. AN ILLUSTRATIVE EXAMPLE

The first subsection of the section includes the results of the Sato and Jain's artificial data clustering by the D-AFC(c)-algorithm. The designed fuzzy inference system is presented in the second subsection.

4.1 Results of the Sato and Jain's Data Clustering

The Sato and Jain's interval-valued data will be used for demonstrating the proposed method of fuzzy rules derivation. These data originally appear in [6] and the data shown in Table 1.

Table 1. The Sato and Jain's interval-valued data

Object s	Attributes		
	\hat{x}^1	\hat{x}^2	\hat{x}^3
x_1	[10, 10]	[1, 1]	[2, 2]
x_2	[9, 9]	[3, 3]	[4, 4]
x_3	[13, 13]	[3, 3]	[2, 2]
x_4	[14, 14]	[4, 4]	[5, 5]
x_5	[4, 8]	[11, 11]	[2, 12]
x_6	[6, 10]	[9, 9]	[1, 8]
x_7	[2, 11]	[10, 10]	[1, 11]
x_8	[3, 9]	[8, 8]	[2, 9]

In this table, x_1, \dots, x_8 are eight objects and each value shows the interval-valued data. Notable that $x_1, x_2, x_3,$ and x_4 have the same values for minimum and maximum values.

The cluster structure of the data set is quasi-stable and the number of classes is equal two. The fact was shown in [16]. That is why the number of clusters in the sought allotment $R^*(X)$ assumed to be two.

The data were processed using the formulae (12) – (14). So, the matrix of similarity coefficients was obtained. The matrix is

presented in Figure 4. Conditions (3), (4) and (5) are not met for the matrix of similarity coefficients. The D-AFC(c)-algorithm was applied to the obtained matrix for $c=2$. Let us consider the clustering results which are described in [17].

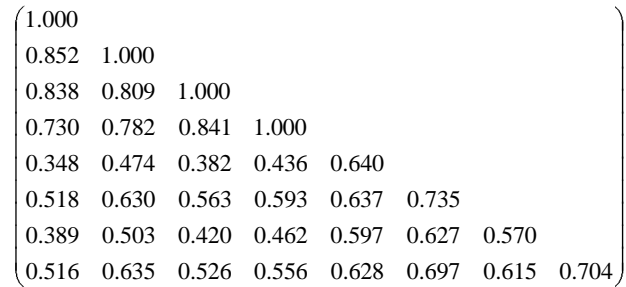


Figure 4. The matrix of similarity coefficients

By executing the D-AFC(c)-algorithm for two classes, we obtain that the first class is formed by 4 elements and the second class is also composed of 4 elements. The allotment among fully separated fuzzy clusters $R^*(X)$, which corresponds to the result, was obtained for the tolerance threshold $\alpha=0.5265$. The value of the membership function of the fuzzy cluster, which corresponds to the first class is maximal for the object x_1 and is equal 1.0000. So, the object x_1 is the typical point of the fuzzy cluster which corresponds to the first class. The membership value of the object x_5 is equal 0.6404 and the value is maximal for the fuzzy cluster which corresponds to the second class. Thus, the object x_5 is the typical point of the fuzzy cluster which corresponds to the second class. So, the second fuzzy cluster is the subnormal fuzzy set. Membership functions of two classes of the allotment are presented in Figure 5 and values which equal zero are not shown in the figure.

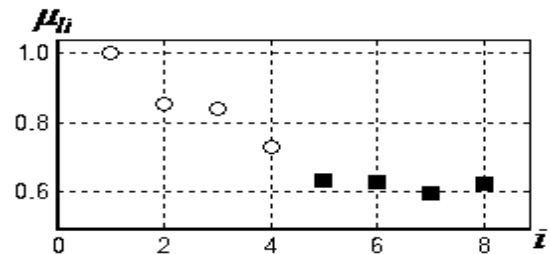


Figure 5. Membership functions of two classes obtained from the D-AFC(c)-algorithm

Membership values of the first class are represented by \circ and membership values of the second class are represented by \blacksquare . The first fuzzy cluster is a power fuzzy cluster with a center and the second fuzzy cluster is a feeble fuzzy cluster with a center [12].

4.2 A Fuzzy Inference System

Let us consider the created fuzzy inference system. Membership functions $\gamma_{B_i}(\hat{x}^t)$ and $\gamma_{C_l^i}(y_l)$ for corresponding fuzzy sets

B_i^l and C_i^l , $i=1, \dots, 3$, $l=1, 2$, can be constructed immediately. The rule base induced by the D-AFC(c)-algorithm clustering result can be seen in Figure 6 where $l=1, 2$ is the

number of rule. Elements of the allotment $R^*(X)$ are fully separated fuzzy clusters. That is why fuzzy sets C_2^1 and C_1^2 are empty fuzzy sets.

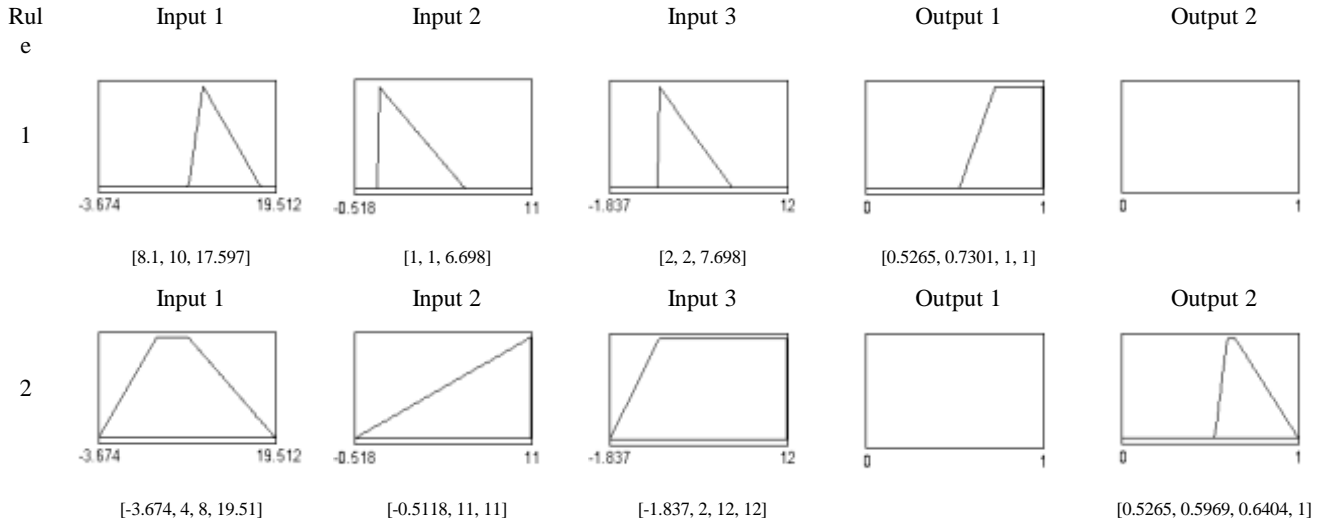


Figure 6. The rule base induced by the clustering result

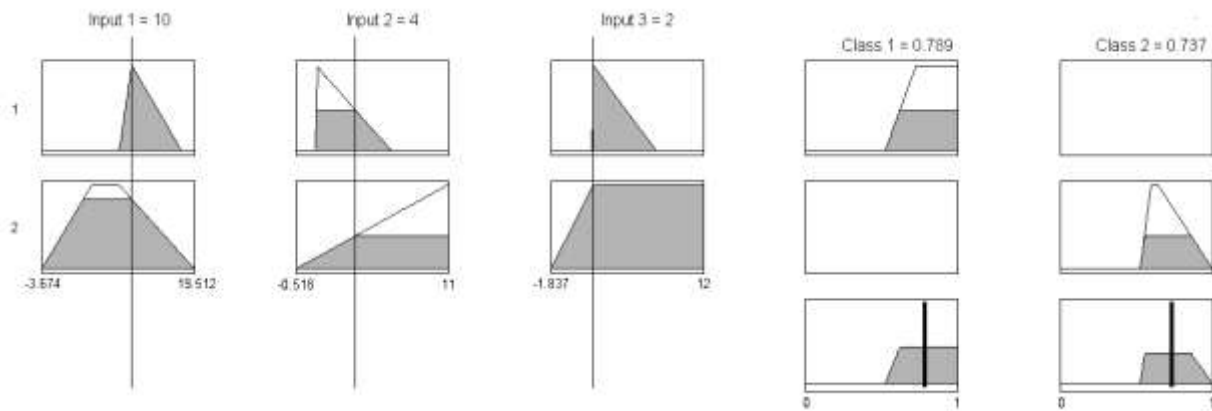


Figure 7. The performance of the fuzzy inference system

The example of classification of some testing object $x_{test} = (\hat{x}_{test}^1 = 10, \hat{x}_{test}^2 = 4, \hat{x}_{test}^3 = 2)$ is shown in Figure 7. Note that membership values of the object are different for both classes.

5. FINAL REMARKS

Some preliminary conclusions are formulated in the first subsection. The second subsection deals with the perspectives on future investigations.

5.1 Discussions

Many techniques to design fuzzy inference systems from data are available; they all take advantage of the property of fuzzy inference systems to be universal approximators. This paper presents an automatic method to design fuzzy inference system from the interval-valued training data set via heuristic possibilistic clustering. The proposed method is simple in comparison with other well-known approaches.

The results obtained with the proposed modeling approach for the Sato and Jain's data set case illustrate the effectiveness of the

proposed method for derivation of fuzzy classification rules from the interval-valued data.

5.2 Perspectives

Constructing a rule base from fuzzy clusters gives a first approximation for the data which can be used as a basis for further improvements. So, a technique of fuzzy rules tuning can be developed.

From other hand, the D-AFC(c)-algorithm can be applied for classification the three-way data [14] and the fuzzy data [13]. So, the proposed method of designing fuzzy inference systems can be generalized for corresponding cases of the training data set. Moreover, the Mamdani's fuzzy inference system can be extended for a case of interval-valued testing data set.

These perspectives for investigations are of great interest both from the theoretical point of view and from the practical one as well.

ACKNOWLEDGMENTS

I am grateful to Prof. Janusz Kacprzyk, Prof. Jan W. Owsinski, and Prof. Frank Klawonn for their interest in the investigations and support. I also thank Mr. Aliaksandr Damaratski for elaborating experimental software. I would like to thank the anonymous referees for their valuable comments.

REFERENCES

- [1] Höppner, F., Klawonn, F., Kruse, R. and Runkler, T. 1999. *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*. John Wiley & Sons.
- [2] Kreinovich, V. 2009. Interval computations and interval-related statistical techniques: tools for estimating uncertainty of the results of data processing and indirect measurements. In: *Data Modeling for Metrology and Testing in Measurement Science*, F. Pavese and A. B. Forbes (Eds.)
- [3] Kreinovich, V. and Kosheleva, O. 2009. Towards dynamical systems approach to fuzzy clustering. In: *Developments in Fuzzy Clustering*, D.A. Viattchenin (Ed.)
- [4] Mamdani, E.H. and Assilian, S. 1975. An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, 7(1), 1-13.
- [5] Sato, M., Sato, Y. 1994. On a multicriteria fuzzy clustering method for 3-way data. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2(2), 127-142.
- [6] Sato-Ilic, M. and Jain, L.C. 2006. *Innovations in Fuzzy Clustering: Theory and Applications*. Springer-Verlag.
- [7] Viattchenin, D.A. 1999. Application of feeble similarity relation to data representation for pattern classification problems. In: *Proceedings of the 6th International Conference ACS'99*.
- [8] Viattchenin, D.A. 2004. A new heuristic algorithm of fuzzy clustering. *Control & Cybernetics*, 33(2), P. 323-340.
- [9] Viattchenin, D.A. 2007. Direct algorithms of fuzzy clustering based on the transitive closure operation and their application to outliers detection. *Artificial Intelligence*, 3, 205-216. (in Russian)
- [10] Viattchenin, D.A. 2007. A direct algorithm of possibilistic clustering with partial supervision. *Journal of Automation, Mobile Robotics and Intelligent Systems*, 1(3), 29-38.
- [11] Viattchenin, D.A. 2008. On possibilistic interpretation of membership values in fuzzy clustering method based on the allotment concept. *Proceedings of the Institute of Modern Knowledge*, 36(3), 85-90. (in Russian)
- [12] Viattchenin, D.A. 2008. Kinds of fuzzy α -clusters. *Proceedings of the Institute of Modern Knowledge*, 37(4), 95-101. (in Russian)
- [13] Viattchenin, D.A. 2008. A heuristic approach to possibilistic clustering for fuzzy data. *Journal of Information and Organizational Sciences*, 32(2), 149-163.
- [14] Viattchenin, D.A. 2009. An outline for a heuristic approach to possibilistic clustering of the three-way data. *Journal of Uncertain Systems*, 3(1), 64-80.
- [15] Viattchenin, D.A. 2009. An algorithm for detecting the principal allotment among fuzzy clusters and its application as a technique of reduction of analyzed features space dimensionality. *Journal of Information and Organizational Sciences*, 33(1), 205-217.
- [16] Viattchenin, D.A. 2009. Analysis of the cluster structure robustness in nonstationary clustering problems. *Doklady BGUIR*, 44(6), 91-98. (in Russian)
- [17] Viattchenin, D.A. and Damaratski, A. 2010. Constructing of allotment among fuzzy clusters in case of quasi-robust cluster structure of set of objects. *Doklady BGUIR*, 47(1), 46-52. (in Russian)
- [18] Viattchenin, D.A. 2010. Automatic generation of fuzzy inference systems using heuristic possibilistic clustering. *Journal of Automation, Mobile Robotics and Intelligent Systems*, 4(3), 36-44.
- [19] Yang, M.-S. and Ko, C.-H. 1996. On a class of fuzzy c-numbers clustering procedures for fuzzy data. *Fuzzy Sets and Systems*, 84(1), 49-60.
- [20] Žák, L. 2002. Clustering of vaguely defined objects. *Archivum Mathematicum*, 38(1), 37-50.