evropský sociální fond v ČR

EVROPSKÁ UNIE

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

FUZZY CLASSIFICATION AND ITS APPLICATION IN HR MANAGEMENT

Pavel Holeček, Jan Stoklasa, Jana Talašová

In both science and practice, it is common to work with classes of objects, defined by verbally specified values of the objects' characteristics. These linguistic expressions can be interpreted as values of linguistic variables representing the characteristics of interest. The classes may be denoted by numbers, usually integers are used. A fuzzy classification system can then be described by means of a rule base in the following way: On the left-hand side of each rule, there is a combination of values of linguistic variables that corresponds to a particular class. A value of an integer-valued variable, which identifies the same class, is on the right-hand side. For any object described by crisp or fuzzy values of its characteristics, its assignment to a class is determined by correspondence between these values and the left-hand sides of the rules. The final output of the fuzzy classification system depends on whether we are solving a problem of objects identification or whether we are classifying the objects for the purpose of their evaluations. In both of these cases, a verbal description of the result is desirable. The solution to this problem will be illustrated on examples from the area of academic staff management.

Tento materiál vznikl v rámci IPN „Zajišťování a hodnocení kvality v systému terciárního vzdělávání"

FUZZY CLASSIFICATION AND ITS APPLICATION IN HR MANAGEMENT

Pavel Holeček, Jan Stoklasa, Jana Talašová

In both science and practice, it is common to work with classes of objects, defined by verbally specified values of the objects characteristics. These linguistic expressions can be interpreted as values of linguistic variables representing the characteristics of interest. The classes may be denoted by numbers, usually integers are used. A fuzzy classification system can then be described by means of a rule base in the following way: On the left-hand side of each rule, there is a combination of values of linguistic variables that corresponds to a particular class. A value of an integer-valued variable, which identifies the same class, is on the right-hand side. For any object described by crisp or fuzzy values of its characteristics, its assignment to a class is determined by correspondence between these values and the left-hand sides of the rules. The final output of the fuzzy classification system depends on whether we are solving a problem of objects identification or whether we are classifying the objects for the purpose of their evaluations. In both of these cases, a verbal description of the result is desirable. The solution to this problem will be illustrated on examples from the area of academic staff management.

Keywords: fuzzy classification, HR management, academic staff evaluation

Classification: 90B90, 03E75

1. INTRODUCTION

Classification problems are very common in the real world. Because the practical classification problems contain elements of uncertainty, it is natural to study the classification methods that make use of the fuzzy sets theory.

Many papers were written on fuzzy classification. The book [5] gives a broad overview of the topic. Various methods of fuzzy classification and their performance are compared e.g. in [3] and in [4].

The vast majority of authors focus mainly on deriving fuzzy rules for the fuzzy classification from the given data (e.g. in [6], [7], [8], and [17]). Nevertheless, this is just the first step in solving the problem. If the fuzzy rule base is already given (either derived from the data or defined expertly) it is necessary to study different ways how the practical classification of the objects by means of the fuzzy rule base can be done. This will be the topic of the paper. The choice of the applied classification method depends on the type of the problem to be solved. In this paper, the type of a problem will be distinguished by a potential existence and type of relationships (ordering, distance) among the classes. The relationships among the classes can be transferred to the set of the numeric identifiers of these classes. Successively, we will describe the mathematical models suitable for the cases when the numeric class identifiers form a nominal, an ordinal, or a cardinal scale. The theory will be accompanied by real-world examples.

First, we should explain the term of classifier. Let us start with the definition of a crisp classifier. Let $\Re^n$ be the space of n-dimensional real vectors. These vectors describe the objects that should be classified. Let $\Omega = \{\omega_1, ..., \omega_k\}$ be a set of class labels. Then, a crisp classifier is described by a mapping $D : \Re^n \to \Omega$ [5]. Every object is assigned to exactly one class and this classification is unambiguous.

A possibilistic classifier is defined as a mapping

$$D^p : \Re^n \to \Re^k - \{\mathbf{0}\}. \tag{1}$$

For any object x described by its features values, its membership degrees $D_1^p(x), ..., D_k^p(x)$ to each of the classes $\omega_1, ..., \omega_k$ are assigned. The case that all of these membership degrees would be zero, i.e. that the object would not belong to any of the classes, is excluded by this definition.

Although the definitions of the crisp and possibilistic classifiers require all objects to be classifiable, in this paper, we will study also the case of objects that cannot be classified good enough. A method for this case will be proposed. The classifiers will be able to refuse the classification if its result could be misleading.

Concerning fuzzy classifiers, one can encounter two different interpretation of this term. In a broad sense, a fuzzy classifier is any classifier that uses fuzzy sets either during its initial training (i.e. during deriving the fuzzy rule base from the data) or during the object classification itself [5]. In the paper, we will consider the fuzzy classifiers in the broad sense. Their common feature will be that a linguistically defined fuzzy rule bases will be used for the description of the classes. However, the difference will be whether crisp or fuzzy values of the features are used for the description of classified objects and whether the classification of the objects is unambiguous or ambiguous.

For the sake of completeness, let us mention that a fuzzy classifier in the strict sense is defined as a possibilistic classifier such that $\sum_{i=1}^{k} D_i^p(x) = 1$ for any $x \in \Re^n$, i.e. the object can be assigned to more than one class but the sum of its membership degrees equals one [5].

Let us specify the problem that will be solved in this paper. In the scientific research and the real life it is common to classify objects into classes which are defined rather vaguely by verbally specified values of the objects characteristics. Our task will be to assign an object, described either by crisp or by vaguely given values of its characteristics, to some of these classes; or in more general way, to determine its location in the frame of these classes. Because the verbal expressions used for description of the classes can be interpreted as vectors of linguistic variables values, it is natural to describe the classes by means of a fuzzy rule base. On the left-hand side of each rule, there is a combination of values of linguistic variables that defines a particular class. The numeric value of a usually integer-valued real variable, which identifies the same class, is on the right-hand side. Labels of the classes can be set also verbally. However, in this case, they are usually interpreted mathematically by real numbers, too. The output of the fuzzy classification system depends on whether we are solving a problem of object identification or whether we are classifying the objects for the purpose of their evaluations. In the first case, the result is a single class for the object or information that the object cannot be classified good enough; the set of the class identifiers can be viewed as a nominal scale. In the latter case, where the classification is used as a certain way of evaluation of the objects, the class identifiers can form an ordinal or a cardinal scale and that affects the form of the classification results. In case of the ordinal scale, several neighboring classes (together with the membership degrees) can be the fuzzy output of the classification. In case of the cardinal scale, the location of the object in the frame of the classes can be calculated. Since the definition of the classes and potentially of the object itself involves uncertainty, the idea of an uncertain classification of the object into classes is meaningful.

Three real-world applications of fuzzy classification will be shown, all of them stemming from the HR management in the academic area. In the first one, academic staff members will be classified according to whether they achieve significantly good results in the research, teaching, or whether their results in both of these areas does not differ substantially. Three classes will be used: *Researcher*, *Teacher* and *Nonspecific*. The result of this classification can be used in HR management. The superordinates can give a possibility to the academic staff members to engage in that area for which they have the best aptitude. In the second application, it will be calculated for academic staff members on various positions (*Assistant professor*, *Associate professor*, or *Professor*) to which of the working positions they were closest according to their performance in the particular time period. The information can be used in the HR management to find promising academic staff members who are aspiring to the higher academic rank. Contrary to the previous example, the classes do not form a nominal but an ordinal evaluation scale and therefore a different model of classification will be chosen. In the last example, the academic staff members will be divided into classes according to their overall performance [13]. The overall performance is calculated from their performance in the areas of pedagogical activities and R&D (research and development). For solving all of the mentioned tasks, the FuzzME software [12, 2] was used. This software product was developed at the Palacky University and have already been applied for solving several practical problems of fuzzy evaluation and fuzzy classification (e.g. [16]). The academic staff performance evaluation system in the frame of which the problems mentioned in this paper are solved is currently being applied also at the Palacky University in Olomouc, Czech Republic.

2. PRELIMINARIES

Fundamentals of the fuzzy set theory (introduced in [14]) are described in detail, e.g., in [1]. A fuzzy set A on a universal set X is characterized by its membership function $A : X \to [0, 1]$. $Ker\ A$ denotes a kernel of A, $Ker\ A = \{x \in X \mid A(x) = 1\}$. For any $\alpha \in [0, 1]$, A_α denotes an α-cut of A, $A_\alpha = \{x \in X \mid A(x) \geq \alpha\}$. A support of A is defined as $Supp\ A = \{x \in X \mid A(x) > 0\}$. The symbol $hgt\ A$ denotes a height of the fuzzy set A, $hgt\ A = \sup\{A(x) \mid x \in X\}$.

A fuzzy number is a fuzzy set C on the set of all real numbers $\Re$ which satisfies the following conditions: a) the kernel of C, $Ker\ C$, is not empty, b) the α-cuts of C, C_α, are closed intervals for all $\alpha \in (0,1]$, c) the support of C, $Supp\ C$, is bounded. A fuzzy number C is called to be defined on $[a,b]$, if $Supp\ C \subseteq [a,b]$.

A center of gravity of a fuzzy number C that is not a real number, is defined as follows

$$t_C = \frac{\int_0^1 C(x).x\ dx}{\int_0^1 C(x)\ dx}. \tag{2}$$

If $C = \{1/c\}$ where $c \in \Re$, then $t_C = c$.

An ordering of fuzzy numbers is defined as follows: a fuzzy number C is greater than or equal to a fuzzy number D, if for the closed intervals C_α and D_α it holds that $C_\alpha \geq D_\alpha$ for all $\alpha \in (0,1]$.

A fuzzy scale on $[a,b]$ is defined as a set of fuzzy numbers $T_1, T_2, ..., T_s$ defined on this interval that form a fuzzy partition on the interval, i.e. for all $x \in [a,b]$ it holds that $\sum_{i=1}^{s} T_i(x) = 1$, and the T's are indexed according to their ordering.

A linguistic fuzzy modeling will be used in this paper frequently. One of its main notions is a linguistic variable. A linguistic variable (see [15]) is defined as a quintuple $(\mathcal{X}, \mathcal{T}(\mathcal{X}), U, G, M)$, where $\mathcal{X}$ is a name of the variable, $\mathcal{T}(\mathcal{X})$ is a set of its linguistic values (linguistic terms), U is a universal set on which the mathematical meanings of the linguistic terms are modeled, G is a syntactic rule for generating linguistic terms from $\mathcal{T}(\mathcal{X})$, and M is a semantic rule which to every linguistic term $\mathcal{A} \in \mathcal{T}(\mathcal{X})$ assigns its mathematical meaning, $A = M(\mathcal{A})$, which is a fuzzy set on U. In this chapter, the linguistic term $\mathcal{A}$ will be distinguished from its mathematical meaning A, which is a fuzzy set, by a different font. In real-life applications, the universe U is usually a closed interval of real numbers, i.e. $U = [a,b]$, and the meanings of the linguistic terms are fuzzy numbers on U.

A linguistic scale (see [11]) is a special case of a linguistic variable. A linguistic scale offers simplified description of a continuous real variable with values on $[a,b]$ by specifying a finite number of linguistic values that are modeled by fuzzy numbers on $[a,b]$. We say that a linguistic variable $(\mathcal{X}, \mathcal{T}(\mathcal{X}), [a,b], G, M)$ where $\mathcal{T}(\mathcal{X}) = \{\mathcal{T}_1, \mathcal{T}_2, ..., \mathcal{T}_s,\}$ is a linguistic scale on $[a,b]$ if the fuzzy numbers $T_1, T_2, ..., T_s$, representing meanings of its linguistic values, form a fuzzy scale on $[a,b]$.

Another important tool of the linguistic fuzzy modeling is a linguistically defined fuzzy rule base. The fuzzy rule base for the classification problems will be the topic of the next section.

3. FUZZY CLASSIFICATION

In the following text, the classes will be described by fuzzy rules. On the left-hand side of each rule, there are linguistic variables together with their linguistic values specifying the class of interest. On the right-hand side, there is a label of the class, whose mathematical meaning is usually modeled by a numeric value of an integer-valued variable. It is possible to describe one class by multiple rules; different weights of rules can be defined. Under these circumstances, we will be solving the problem of classification of an object described by values of its characteristics. Let C, be a set of numeric identifiers of the classes of interest, usually $C = \{1, ..., k\}$, $k \in N$. A fuzzy classification system can then be described by means of a fuzzy rule base in the following form:

$$
\begin{array}{l}
\text{If } \mathcal{F}_1 \text{ is } \mathcal{A}_{1,1} \text{ and } \ldots \text{ and } \mathcal{F}_m \text{ is } \mathcal{A}_{1,m}, \text{then class } \mathcal{D}_1 \text{ with weight } w_1 \\
\text{If } \mathcal{F}_1 \text{ is } \mathcal{A}_{2,1} \text{ and } \ldots \text{ and } \mathcal{F}_m \text{ is } \mathcal{A}_{2,m}, \text{then class } \mathcal{D}_2 \text{ with weight } w_2 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\text{If } \mathcal{F}_1 \text{ is } \mathcal{A}_{n,1} \text{ and } \ldots \text{ and } \mathcal{F}_m \text{ is } \mathcal{A}_{n,m}, \text{then class } \mathcal{D}_n \text{ with weight } w_n
\end{array}
$$

where for $i = 1, 2, \ldots, n, j = 1, 2, \ldots, m$:

- $(\mathcal{F}_j, \mathcal{T}(\mathcal{F}_j), [p_j, q_j], M_j, G_j)$ are linguistic variables, usually linguistic scales, for the individual features,
- $\mathcal{A}_{i,j} \in \mathcal{T}(\mathcal{F}_j)$ are linguistic values, and $A_{ij} = M_j(\mathcal{A}_{i,j})$ are fuzzy numbers on $[p_j, q_j]$ representing their meanings,
- $\mathcal{D}_i$ are the class labels, and D_i are the corresponding numeric class identifiers, usually $D_i \in \{1, ..., k\}$.
- $w_i \in (0, 1]$ are weights of the rules.

The weights are typically used in cases when the fuzzy rules are derived from the data. In the application described in this paper, we will expect that all rules have their weight equal to 1.

If the goal of the classification is just an identification of the object as a member of one of the classes, and there are no relationships among the classes, or their relationships are not related to the given problem, then the scale formed by the numeric identifiers of the classes is considered to be a nominal one. Although the above-mentioned fuzzy rule base with the numeric classes identifiers on the right-hand sides of the rules reminds of the Sugeno fuzzy rule base [9], it is obvious that the Sugeno fuzzy inference algorithm, whose output is calculated as a weighted average of the numeric values on the right-hand sides of the rules, cannot be used for the fuzzy classification. In the next chapter, fuzzy classification algorithms appropriate for this case will be described. The result of classification will be the class to which the object characteristics match the best, or alternatively, the information that the object cannot be classified unambiguously. This type of fuzzy classification will be illustrated on the example of determining the type of an academic staff member.

If the goal of the classification is a certain type of objects evaluation, it makes sense to assume that the numbers, which identify the classes, form an ordinal, or even a cardinal scale. It is meaningful to consider the object as lying between two neighboring classes. Moreover, in the case of a cardinal evaluation scale, it is also meaningful to calculate the location of the objects in relation to these classes and the Sugeno inference algorithm [9] can be used. If we demand a natural verbal description of the evaluation process and the fuzzy classification result, it is possible to use the Sugeno-Yasukawa model [10]. Concerning the calculation of the results, this model is analogous to the Sugeno approach. However, on the right-hand sides of the rules, there are linguistic values. For the calculation of the fuzzy weighted average in the Sugeno fuzzy inference algorithm, real numbers are used. These numbers are the crisp representatives of the fuzzy numbers that model the mathematical meanings of those linguistic values. The application of a slightly modified Sugeno-Yasukawa approach will be shown in the area of academic staff performance evaluation (for more details see [13]). The linguistic values on the right-hand sides of the rules will be used for a linguistic interpretation of the classification results.

The great advantage of using the tools of linguistic fuzzy modeling in all of the mentioned cases is that the fuzzy classification rules and final results are described in the most natural way for the humans, i.e. verbally.

4. APPLICATIONS IN THE AREA OF HR MANAGEMENT

The various types of fuzzy classification mentioned in the previous section will be illustrated by examples originating from the system of academic staff performance evaluation that is applied at the Palacký University in Olomouc, Czech Republic.

In the frame of the system, the performance of each member of academic staff is evaluated in both pedagogical, and research and development (R&D) areas of activities. Input data are acquired from a form filled in by the staff where particular activities are assigned a score according to their importance and time-consumption. Three areas are taken into consideration for pedagogical performance evaluation: (a) lecturing, (b) supervising students, and (c) work associated with the development of fields of study. The evaluation of research and development activities is based on the R&D methodology of evaluation valid in the Czech Republic (papers in important journals, monographs, and patents are valued very highly) but other important activities (grant project management, editorial board memberships etc.) are also included. Both pedagogical and R&D areas are assigned standard scores different for senior assistant professors, associate professors, and professors. The number representing a partial evaluation of a member of academic staff in a certain area is determined as a multiple of the respective standard for his or her position. For better clarity and easier interpretation, linguistic fuzzy scales are defined on the domains of the partial evaluations (see Figures 1 and 2). If for example the performance of an academic worker in R&D is 1.25 times of the standard, using the scale in the Figure 2, it can be linguistically interpreted that the performance is *75 % standard and 25 % high.* In the pedagogical area (see Figure 1), the evaluation of the activities is based namely on their time consumption; so the double of the standard performance is considered to be an extreme performance. In R&D (see Figure 2), the mentioned methodology is used. In frame of this methodology, the evaluations of journals grows sharply with their importance; so the triple of the standard score is still achievable for the academic staff members.

In the following chapters we are going to apply different types of fuzzy classification to answer the following questions that are important from the point of view of the HR management: (1) Is an academic staff member more teacher or researcher? In which area does he/she perform especially well? Where should an additional space be given to him/her for his/her further development? (2) Which working position corresponds to the behavior of the academic staff member - assistant professor, associate professor or professor? Is he/she a promising academic worker whose behavior corresponds better to the higher academic rank than his/her real one? And finally, we are going to answer the most important question: (3) What is the overall performance, i.e. the one calculated from his/her evaluations in the areas of pedagogical activities and R&D, of the academic staff member? These questions were deliberately chosen so that the different types of fuzzy classification could be shown in an illustrative way.

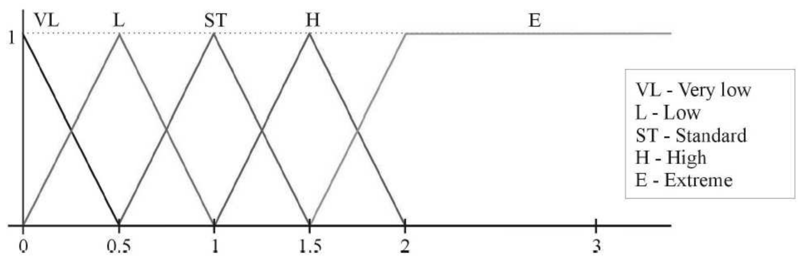

Fig. 1. A linguistic fuzzy scale used for the evaluation of academic staff members in the area of pedagogical activities

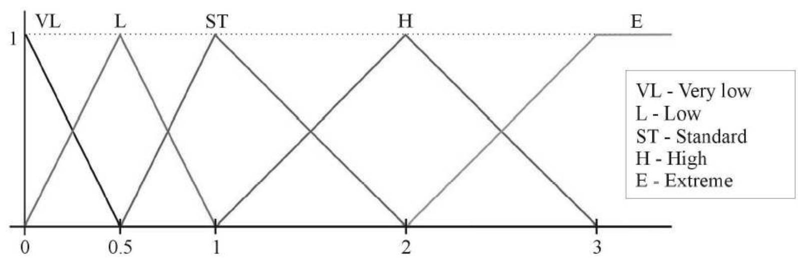

Fig. 2. A linguistic fuzzy scale used for the evaluation of academic staff members in the area of R&D

4.1. Is an academic staff member more teacher or researcher?

In this chapter, we will use a fuzzy classifier to solve an identification problem. The academic staff member will be classified according to the area where they perform better. The possible classes are *Teacher*, *Researcher*, and *Nonspecific*; the third class means that the academic staff member has balanced evaluation in both of these areas. The knowledge of the type of an academic staff member can be used in human resource management at the university. If academic staff members perform significantly well in one area and have not-so-good results in the other, then their supervisor can allow them more space to focus on that area of activities for which they are better suited. Moreover, to calculate their overall evaluation, different rule bases can be used than in the case of the *Nonspecific* academic staff members.

This classification is based on the evaluations of the academic staff members in the areas of pedagogical activities and R&D. The designed fuzzy rule base is shown in Figure 3.

Academic staff member type		Research and Development Performance				
		Very low	Low	Standard	High	Extreme
Pedagogical activities performance	Very low	Nonspecific	Nonspecific	Nonspecific	Researcher	Researcher
	Low	Nonspecific	Nonspecific	Nonspecific	Researcher	Researcher
	Standard	Nonspecific	Nonspecific	Nonspecific	Nonspecific	Researcher
	High	Teacher	Teacher	Nonspecific	Nonspecific	Nonspecific
	Extreme	Teacher	Teacher	Teacher	Nonspecific	Nonspecific

Fig. 3. Fuzzy rule base used for determining the type of academic staff members

It is obvious that the classes (*Teacher*, *Researcher*, and *Nonspecific*) form only a nominal scale. Two classification algorithms suitable for this type of classification were tested - *Single Winner* and *Voting by Multiple Fuzzy Rules* [4].

Single Winner

In the *Single Winner* method [4], the classification of objects is done as follows. Let us suppose that an object is described by the values of its characteristics, i.e. by real numbers $a_1, ..., a_m$. Moreover, we expect that a fuzzy rule base is given and that it is in the form described in the Section 3. Then, the classification by the *Single Winner* method is done by the following procedure.

First, the degrees of correspondence h_i, $i = 1, ..., n$, between the inputs and the left-hand sides of the rules are calculated

$$h_i = A_{i1}(a_1) \cdot A_{i2}(a_2) \cdot ... \cdot A_{im}(a_m) \cdot w_i, \quad i = 1, ..., n. \tag{3}$$

The coefficients in the formula (3) express the fulfillment of the individual conditions on the left-hand sides of the rules. In the used literature [4], the multiplication was applied for their aggregation. However, it is possible to use another t-norm (e.g. minimum) or even an averaging aggregating operator (e.g. a weighted arithmetic mean) instead. The choice of the suitable operator depends on the nature of the problem to be solved.

Let us note, that the formula (3) can be generalized for the case that the objects that should be classified are described by fuzzy values of their characteristics. If an object is described by fuzzy numbers $A_1^{'}, ..., A_m^{'}$, then the degrees of correspondence h_i, $i = 1, ..., n$, can be calculated by the following formula

$$h_i = hgt(A_1^{'} \cap A_{i1}) \cdot hgt(A_2^{'} \cap A_{i2}) \cdot ... \cdot hgt(A_m^{'} \cap A_{im}) \cdot w_i, \;\; i = 1, ..., n. \qquad (4)$$

In the next step, the so called *number of votes* is calculated for each class as follows:

$$v_T = \max_{\substack{i \in \{1,...,n\}: \\ D_i = T}} h_i, \;\; T = 1, ..., k. \qquad (5)$$

The resulting class T^* for a given object is the one with the maximum value of v_T, i.e. the one for which it holds that

$$v_{T^*} = \max_{T=1,...,k} v_T. \qquad (6)$$

Voting by Multiple Fuzzy Rules

In case of *Voting by Multiple Fuzzy Rules* method [4], the degrees of correspondence h_i are calculated in the same way as with the *Single Winner*:

$$h_i = A_{i1}(a_1) \cdot A_{i2}(a_2) \cdot ... \cdot A_{im}(a_m) \cdot w_i, \;\; i = 1, ..., n. \qquad (7)$$

Next, the number of votes is calculated for each class as the sum of degrees of correspondence pertaining to those rules that voted for this class:

$$v_T = \sum_{\substack{i \in \{1,...,n\}: \\ D_i = T}} h_i, \;\; T = 1, ..., k. \qquad (8)$$

The resulting class T^* is again the one with the maximum value of v_T.

Figure 4 compares the results for a testing example that were obtained by the two mentioned methods. Each dot represents an academic staff member. The dot position is determined by the academic staff member's evaluation in the area of pedagogical activities (x axis) and R&D (y axis). The resulting classes are differentiated by the color of the dots - black for *Teachers*, gray for *Researchers*, and white for the *Nonspecific* academic staff members. It can be seen that the border between the classes is smoother for *Voting by Multiple Fuzzy Rules*.

Objects that cannot be classified

In some cases an object cannot be classified unambiguously, i.e. if it was classified then its membership degree to the chosen class would not be significantly higher than its membership degrees to the other classes. In the above example, academic staff members were classified into three classes. The optimum class for an academic staff member was determined by the largest number of votes; if there were more such classes, it would be possible to select any of them. So far, we have not studied how reliable the assignment of the class to an academic worker was. This will be

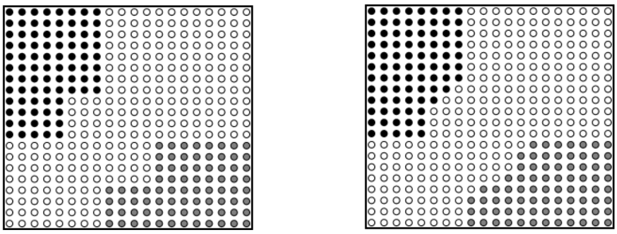

Fig. 4. Results obtained by *Single Winner* (left) and *Voting by Multiple Fuzzy Rules* (right)

discussed in the following text and a method for identifying unclassifiable objects will be proposed.

The boundary between classifiable and unclassifiable objects can be set by choosing a minimal required *distinctiveness of the winner.* The distinctiveness of the winner CF is a real number on $[0, 1]$ defined as follows.

$$CF = 1 - (V'/V), \tag{9}$$

where

$$V = \max_{T=1,\dots,k} v_T = v_{T^*} \tag{10}$$

is the number of votes for the winning class T^* (respecting the definition of the possibilistic classifier (see formula 1), the fuzzy rule base will be defined in such way so that V would never be zero, i.e. the fuzzy rules have to cover the entire input space). The value of $V^{'}$,

$$V^{'} = \max_{\substack{T=1,\dots,k \\ T \neq T^*}} v_T, \tag{11}$$

is the number of votes for the second best fitting class. If the distinctiveness of the winner is lower than the selected one, it means that the classification is ambiguous and the object belongs to more than one class in similar degrees.

Figure 5 shows the results for two different values of the minimal required distinctiveness CF_{min} that were obtained for the case of the *Single Winner* algorithm. The unclassifiable objects are depicted by crosses. For example, let us assume four academic staff members with different evaluations in the area of pedagogical activities and R&D. The tables 1 and 2 compares the results of the fuzzy classification for 4 academic staff members. The proposed class for the first three academic staff members is the same. However, for the last academic staff member, each of the algorithms proposed a different class. Notice that in this case the distinctiveness of the winner is significantly lower.

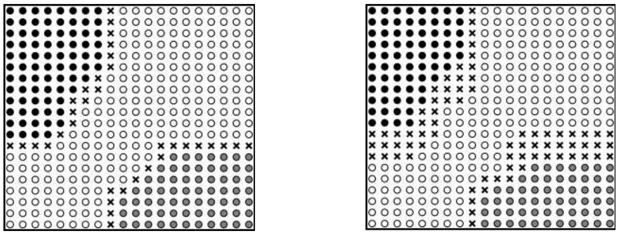

Fig. 5. Results for $CF_{min} = 0.4$ (left) and $CF_{min} = 0.7$ (right).

Pedagogical activities	**R&D**	**Proposed class**	**Distinctiveness**
0.9 (Standard)	2.8 (Extreme)	Researcher	0.75
2.1 (Extreme)	0.5 (Low)	Teacher	1
1.3 (High)	2.2 (High)	Nonspecific	0.83
0.8 (Standard)	2.3 (High)	Nonspecific	0.33

Table 1. Sample results of the academic staff members fuzzy classification by the *Single Winner* method.

Pedagogical activities	**R&D**	**Proposed class**	**Distinctiveness**
0.9 (Standard)	2.8 (Extreme)	Researcher	0.81
2.1 (Extreme)	0.5 (Low)	Teacher	1
1.3 (High)	2.2 (High)	Nonspecific	0.92
0.8 (Standard)	2.3 (High)	Researcher	0.28

Table 2. Sample results of the academic staff members fuzzy classification by the *Voting by Multiple Fuzzy Rules* method.

4.2. Which working position corresponds to the behavior of the academic staff member - assistant professor, associate professor or professor?

In this example, another classification of the academic staff members will be done. The result of the classification will be a working role (*assistant professor*, *associate professor* or *professor*) which corresponds the best with the academic staff member's performance during the last evaluated period. This information can be valuable for the HR management, especially if the performance of an academic staff member corresponds to a higher academic rank than his/her real one. The results can be used to find promising academic staff members who are aspiring to the higher rank. For example, if an assistant professor has the performance typical for the associate professor, the head of department should give him/her time to prepare for the higher academic rank (e.g. in form of a sabbatical).

For the classification, a fuzzy classifier based on a fuzzy rule base is again used. The classifier is applied only to such academic staff members whose performance in both evaluating areas (pedagogy and R&D) is at least acceptable. For each of the working roles (*assistant professor*, *associate professor*, or *professor*), one rule is present in the fuzzy rule base. The rules reflect the typical behavior of the representatives of these working positions. With a higher academic rank, a significant increase of the performance in R&D is expected. In the pedagogical area, with an increase of the academic rank, the focus is moved from lecturing to students supervising (diploma students, doctoral students) or to the work associated with the development of fields of study.

The classification of academic staff members into the above-mentioned classes represents a certain type of evaluation. It is obvious that if an academic worker is assigned to the class *professor*, it represents better results for him/her than if he/she would be assigned to the class *associate professor*. On the other hand, on the basis of the two characteristics used for the classification, it is not possible to quantify the distances between the neighboring classes for the evaluation. The classes can be only ordered, but their distances cannot be measured. This case represents the evaluation only on an ordinal scale. In the academic staff performance evaluation model implemented at the Palacky University in Olomouc, which is discussed in this paper, the standard score in R&D for professors was set as a double of the standard for associate professors, which is again the double of the standard score for assistant professors. Concerning the evaluation of the pedagogical area, the standard scores for all three positions are the same, but they are acquired from different type of activities. The academic rank affects the ratio between supervising diploma and doctoral students and work associated with the development of fields of study on one side, and the direct teaching on the other side.

Two input variables were used - R&D outcomes (shortly *R&D*) and prevailing pedagogical activities (shortly *pedagogics*). The first real-valued input variable is defined as the ratio between the achieved score in R&D and the standard score for the lowest of the three mentioned academic ranks. The second real-valued input variable is defined as the ratio between scores acquired by the particular academic staff member for supervising students and the work associated with the development of fields of study, and the lecturing. The linguistic fuzzy scales for both variables are depicted in the Figures 6 and 7. The final fuzzy rule base then looks as follows:

If (*R&D* is *low*) and (*pedagogics* is *teaching*) then class is *assistant professor*.
If (*R&D* is *medium*) and (*pedagogics* is *balanced*) then class is *associate professor*.
If (*R&D* is *high*) and (*pedagogics* is *students supervising*) then class is *professor*.

It is obvious that the evaluating function defined by the fuzzy rule base is non-decreasing in both of the variables. Contrary to the previous classification problem, the numeric identifiers of the classes (1 – *assistant professor*, 2 – *associate professor*, and 3 – *professor*) form not only nominal scale but an ordinal one. Therefore, the resulting information that we can obtain will be different. If the fuzzy rule base

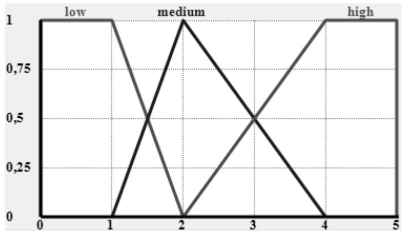

Fig. 6. Linguistic fuzzy scale for the evaluation of R&D outcomes

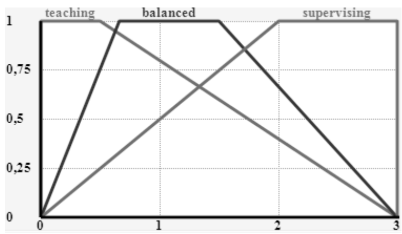

Fig. 7. Linguistic fuzzy scale for the prevailing pedagogical activities

expresses an evaluation on an ordinal scale, then it is meaningful to expect that this linguistic evaluating function is non-decreasing in all its variables (similarly as in our example). Then it is obvious that an object can be (partially) assigned either to one class or to a sequence of mutually neighboring classes. The membership degrees to these classes will be given by the formulas 5 or 8. The assignment of objects into such a sequence of several neighboring classes is meaningful especially in case of objects described by fuzzy values of their characteristics. Concerning the linguistic description of the classification results for an object with non-zero membership degrees to the classes $i, i+1, ..., j-1, j$, we say that this object belongs to the classes i to j. Specifically in our example, the result will be either one class with the membership degree less than or equal to one, or two neighboring classes with the corresponding membership degrees.

In the following chapter, we will study the use of fuzzy classification for an evaluation on a cardinal scale. We will show that in this case it is possible to apply the Sugeno or Sugeno-Yasukawa inference algorithms in the way that is described in the chapter 3.

4.3. What is the overall performance of the academic staff member?

This application is the main part of the academic staff performance evaluation model which was developed at the Palacky University in Olomouc, Czech Republic [13]. In this application, several performance classes are defined for academic staff members. The classification is based on their evaluations in the areas of pedagogical activities and R&D, whose calculation is described in the introduction of the chapter 4. Since the evaluating scales used for these two areas differ in their character, the aggregation of these two partial evaluations is difficult. That is why a model that uses a fuzzy rule base (see Figure 8) was designed. Specifically, the fuzzy classification based on the Sugeno-Yasukawa approach [10] was applied. Apart from the mentioned basic fuzzy rule base, which gives the same importance to both the areas of the academic staff members' activities, it is also planned to create special fuzzy rule bases for the academic staff members of the type *teacher* and *researcher* (the identification of these types was described in the chapter 4.1).

Overall performence of an academic staff member		Research and Development Performance				
		Very low	Low	Standard	High	Extreme
Pedagogical activities performance	Very low	Unsatisfactory	Unsatisfactory	Substandard	Standard	Very Good
	Low	Unsatisfactory	Unsatisfactory	Substandard	Very Good	Excellent
	Standard	Substandard	Substandard	Standard	Very Good	Excellent
	High	Standard	Very Good	Very Good	Excellent	Excellent
	Extreme	Very Good	Excellent	Excellent	Excellent	Excellent

Fig. 8. Fuzzy rule base used for classification according to the overall performance of academic staff members

In this application, the indicators of classes were defined as significat values on a continuous cardinal evaluating scale. In this case for an easier interpretation, the class identifiers are not integers; numeric values 0, 0.5, 1, 1.5, and 2 were used. By fuzzification of these values, elements of the fuzzy scale were obtained. These elements were subsequently described by the linguistic terms *unsatisfactory*, *substandard*, *standard*, *very good*, and *excellent*. The original numeric values lie in the kernels of the triangular fuzzy numbers that form the fuzzy scale. The fuzzy scale is depicted in Figure 9.

The overall evaluation of academic staff members will be calculated as follows. First, the rule base described in the Figure 8 was modified so that instead of the linguistic terms stated in the table, there would be only the significant values of the classes, i.e. the above-mentioned real numbers 0, 0.5, 1, 1.5, and 2, on the right-hand sides of the rules. Then, the Sugeno inference algorithm [9] will be applied to the crisp evaluations of a given academic staff member in the area of pedagogical activities (pa) and R&D (rd). In this way, a crisp value of the overall evaluation ($eval(pa, rd)$) will be calculated. This procedure can be expressed by the following formula:

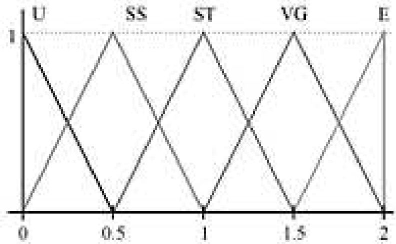

Fig. 9. The linguistic fuzzy scale used for performance classes.

$$eval(pa, rd) = \frac{\sum_{i=1}^{n} A_{i1}(pa) \cdot A_{i2}(rd) \cdot ev_i}{\sum_{i=1}^{n} A_{i1}(pa) \cdot A_{i2}(rd)} = \sum_{i=1}^{n} A_{i1}(pa) \cdot A_{i2}(rd) \cdot ev_i \qquad (12)$$

where for $i = 1, ...n$: A_{i1} is the fuzzy number representing the meaning of the linguistic term describing evaluation in the pedagogical area in rule i, and A_{i2} is the fuzzy number representing the meaning of the linguistic term describing evaluation in the area of R&D in rule i, and ev_i is the real number representing the most typical value of the linguistic term describing the resulting class $\mathcal{D}_i$ in rule i; ev_i lies in the kernel of the respective triangular fuzzy number.

From the numeric evaluation $eval(pa, rd)$ we will proceed to its linguistic description, which, in the context of the HR management, is more suitable. For that purpose, we will make use of the linguistic scale in Figure 9. If $eval(pa, rd) = ev_i$, for some $i = 1, .., n$, then the academic staff member fully belongs to the class with the characteristic value ev_i and the linguistic interpretation of the result is clearly given by the corresponding term. Otherwise, it belongs to two neighboring classes which are the closest, i.e. where the value $eval(pa, rd)$ belongs with a non-zero membership degree. Membership degrees of $eval(pa, rd)$ to these two classes are used for the linguistic description of the resulting evaluation. For example, a possible result can be that the overall performance of a given academic staff member is *70 % standard and 30 % very good*. Both the linguistically defined evaluation function and the real evaluating function given by the formula 12 are non-decreasing in both variables. Because of this property of the fuzzy rule base, only a single class or several neighboring classes can have a non-zero weight in the Sugeno inference algorithm.

The procedure used to calculate the overall evaluations, including the linguistic description of the result, is very similar to the Sugeno-Yasukawa approach described in [10]. This approach uses a rule base with linguistic variables (preferably linguistic scales) on both sides of the rules. To calculate the crisp outputs, which correspond to the given crisp inputs and the rule base, only the representatives of the fuzzy numbers that model the meaning of the right-hand sides of the rules are used. In the Sugeno-Yasukawa approach, the representatives are the centers of gravity of the given fuzzy numbers contrary to our approach where the elements from the kernels of these triangular fuzzy numbers are used instead. The linguistic values on the right-hand sides of the rules are used subsequently for the linguistic description of the calculated crisp results in the same way as described in our case.

5. CONCLUSION

On a real example from the human resource management in the academic area, three fuzzy classification problems were described in this paper. The first case represents an identification problem when it is necessary to decide to which of the classes that are described verbally by the fuzzy values of their characteristics does a given object belongs (or alternatively, to decide that the object cannot be classified satisfactorily). No relationships among the classes are considered; their identifiers form a nominal scale. In the second and third case, fuzzy classification is applied for solving of evaluation problems. An ordinal evaluating scale is used in the second case, whereas a cardinal evaluating scale is applied in the third one. Suitable mathematical models were described for all of them.

The identification and evaluation represent the two typical problems, where the fuzzy classification is applied in the practice. In this paper we have studied the classification problems without and with existing relationships among the classes (orderings or metrics). We have shown the algorithms for solving such problems on a real-world example. In the further research, the issues of the structures existing on the set of the classes and the corresponding classification algorithms will be solved systematically.

(Received ????)

REFERENCES

[1] D. Dubois and H. Prade, editors. *Fundamentals of Fuzzy Sets. The Handbook of Fuzzy Sets Series.* Kluwer Academic Publishers, Boston-London-Dordrecht, 2000.

[2] P. Holeček and J. Talašová. Designing fuzzy models of multiple-criteria evaluation in FuzzME software. In *28th International Conference on Mathematical Methods in Economics 2010*, pages 250–256, 2010.

[3] H. Ishibuchi, T. Murata, and M. Gen. Voting in fuzzy rule-based systems for pattern classification problems. *Fuzzy Sets and Systems*, 35(3-4):575–578, 1998.

[4] H. Ishibuchi, Nakashima T., and Morisawa T. Performance evaluation of fuzzy rule-based classification systems obtained by multi-objective genetic algorithms. *Computers & Industrial Engineering*, 103(2):223–238, 1999.

[5] L. I. Kuncheva, editor. *Fuzzy classifier design.* Springer-Verlag, Heidelberg, 2000.

[6] Z. Lei and L. Ren-hou. Designing of classifiers based on immune principles and fuzzy rules. *Information Sciences*, 178(7):1836–1847, 2008.

[7] T.-H. Li, Guo N. R., and Kuo C.-L. Design of adaptive fuzzy model for classification problem. *Engineering Applications of Artificial Intelligence*, 18(3):297–306, 2005.

[8] J. A. Roubos, Setnes M., and Abonyi J. Learning fuzzy classification rules from labeled data. *Information Sciences*, 150(1-2):77–93, 2003.

[9] M. Sugeno. An introductory survey on fuzzy control. *Information Sciences*, 36:59–83, 1985.

[10] M. Sugeno and T. Yasukawa. A fuzzy-logic-based approach to qualitative modeling. *IEEE Transactions on fuzzy systems*, 1(1):7–31, 1993.

[11] J. Talašová. *Fuzzy metody vícekriteriálního hodnocení a rozhodování (in Czech)*. Publishing House of Palacký University, ISBN 80-244-0614-4, Olomouc, 2003.

[12] J. Talašová and P. Holeček. Multiple-criteria fuzzy evaluation: The FuzzME software package. In *Joint International Fuzzy Systems Association World Congress/European Society for Fuzzy Logic and Technology*, pages 681–686, 2009.

[13] J. Talašová and J. Stoklasa. Fuzzy approach to academic staff performance evaluation. In *Proceedings of the 28th International Conference on Mathematical Methods in Economics 2010*, pages 621–626, 2010.

[14] L.A Zadeh. Fuzzy sets. *Inform. Control*, 8:338–353, 1965.

[15] L.A Zadeh. The concept of linguistic variable and its application to approximate reasoning. *Information sciences*, 8:199–249, 1975.

[16] B. Zemková and J. Talašová. Application of fuzzy sets in human resources management. In *28th International Conference on Mathematical Methods in Economics 2010*, pages 682–687, 2010.

[17] E. Zhou and A. Khotanzad. Fuzzy classifier design using genetic algorithms. *Pattern Recognition*, 40(12):3401–3414, 2007.

Pavel Holeček, Palacký University Olomouc, Faculty of Science, Dept. of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/121, 77146 Olomouc, Czech Republic

e-mail: pavel.holecek@upol.cz

Jan Stoklasa, Palacký University Olomouc, Faculty of Science, Dept. of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/121, 77146 Olomouc, Czech Republic

e-mail: jan.stoklasa@upol.cz

Jana Talašová, Palacký University Olomouc, Faculty of Science, Dept. of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/121, 77146 Olomouc, Czech Republic

e-mail: jana.talasova@upol.cz