Fuzzy Sets and Systems 51 (1992) 147–153
North-Holland

Fuzzy rough sets: Application to feature selection

Ludmila Ilieva Kuncheva
Central Laboratory of Bioinstrumentation and Automation, Bulgarian Academy of Sciences, Sofia, Bulgaria

Received June 1991
Revised January 1992

Abstract: Fuzzy rough sets are a notion introduced as a further extension of the idea of rough sets. Its purpose is fuzzy pattern recognition, where classes are described by fuzzy sets on the feature space. A measure of classification accuracy, which can serve also as a criterion for feature selection, is developed in order to be used in fuzzy recognition problems. It is shown that the fuzzy concepts of Positive, Negative and Boundary region of a set, proposed in the paper, coincide with the original definitions if appropriately applied to ordinary rough classification. Feature selection via fuzzy rough sets is illustrated using real data from medical practice. The problem of evaluating the hypoxic resistance of a patient on the basis of the values of his blood pressure during a barocamera examination was considered. The measurements were evaluated by the fuzzy rough sets criterion. The results obtained are closely related with the expert opinion on the problem.

Keywords: Foundations of fuzzy sets; rough sets; fuzzy pattern recognition; feature selection.

1. Introduction

The concept of rough sets have been introduced and developed by Pawlak and co-workers [14, 15, 17] in deterministic and probabilistic sense. This is an attractive tool to assess the highest classification capacity and to select a minimal set of significant features in pattern recognition problems, where data are described by qualitative features.

An appealing point behind the rough sets idea is the application opportunity. Several meaningful medical applications of rough sets are reported in [6, 16, 20, 21]. In [5, 8, 10] applications of rough classification to knowledge engineering and inference models are described.

A number of theoretical investigations have been carried out, which build a parallel between rough sets and fuzzy sets [5, 13, 22, 24, 25]. In these investigations the rough set classification is being 'fuzzified' in different ways, including fuzziness into some of its stages and elements. Interrelations and subordination between fuzzy sets and rough sets are being searched for.

In spite of the significant achievements of fuzzy pattern recognition both in theoretical and application plans [18], the feature selection criterion is still an open problem. Among the great amount of publications devoted to fuzzy pattern recognition only few discuss this issue. In this paper an attempt is made to integrate the approaches of rough sets and fuzzy sets in order to build a criterion oriented to fuzzy pattern recognition. A measure of classification accuracy, which can serve also as a criterion for feature selection, is developed.

Feature selection via fuzzy rough sets is illustrated using real data from medical practice.

In Section 2 the theoretical basis of fuzzy rough sets is presented. The idea of application of fuzzy rough sets to feature selection in fuzzy pattern recognition is described in Section 3. The experimental illustration is given in Section 4.

2. Theoretical ground

2.1. Rough sets

A brief recall of some rough sets definitions is given, in order to enable the easy acceptance of their fuzzy variants.

Let U be a universum and

$$A = \{A_1, \ldots, A_M\}$$

be a partition on U, defined by an equivalence

Correspondence to: Dr. L.I. Kuncheva, Central Laboratory of Bioinstrumentation and Automation, Bulgarian Academy of Sciences, Acad. G. Bonchev str., Bl. 105, 1113 Sofia, Bulgaria.

relation (i.e. $A_i \subseteq U$, $A_i \cap A_j = \emptyset$, $\bigcup A_i = U$). For any subset $X \subseteq U$ the following notions are defined:

– *lower approximation of* X in terms of the partition A (also called *positive region of* X):

$$\mathrm{POS}_A(X) = A^{\mathrm{L}}(X) = \bigcup_{A_i \subseteq X} A_i;$$

– *upper approximation of* X in terms of the partition A:

$$A^{\mathrm{U}}(X) = \bigcup_{A_i \cap X \neq \emptyset} A_i;$$

– *Negative region of* X:

$$\mathrm{NEG}_A(X) = U - A^{\mathrm{U}}(X);$$

– *Boundary region of* X:

$$\mathrm{BND}_A(X) = A^{\mathrm{U}}(X) - A^{\mathrm{L}}(X);$$

– *quality of approximation of* X *by* A:

$$\mu_A(X) = \frac{\mathrm{card}(A^{\mathrm{L}}(X))}{\mathrm{card}(A^{\mathrm{U}}(X))}.$$

Obviously $\mu_A(X) \in [0, 1]$ and $\mu_A(X) = 1$ iff X is definable by A (i.e. $A^{\mathrm{U}}(X) = A^{\mathrm{L}}(X)$).

Let $B = \{B_1, \ldots, B_N\}$ be a second partition on U, induced by some other equivalence relation. Then the following definitions hold:

$$\mathrm{POS}_A(B) = \bigcup_{B_i \in B} A^{\mathrm{L}}(B_i),$$

$$\mathrm{BND}_A(B) = \bigcup_{B_i \in B} (A^{\mathrm{U}}(B_i) - A^{\mathrm{L}}(B_i)),$$

$$\mathrm{NEG}_A(B) = U - \bigcup_{B_i \in B} A^{\mathrm{U}}(B_i).$$

Then a plausible measure of dependency of B on A [15] is given by

$$\gamma_A(B) = \frac{\mathrm{card}(\mathrm{POS}_A(B))}{\mathrm{card}(U)}.$$

2.2. Fuzzy rough sets

Let $A = \{A_1, \ldots, A_M\}$ be a weak fuzzy partition on U, i.e.

$$\mu_{A_i}(x) \in [0, 1], \quad x \in U, \quad A_i \in A,$$

$$\bigcup_i \mathrm{supp}\, A_i = U.$$

The weak fuzzy partition is a convenient model of fuzzy classification problems, where A_i, $i = 1, \ldots, M$, correspond to the classes and $\mu_{A_i}(x)$ expresses a certain degree of membership (e.g. severity of illness) rather than a kind of uncertainty. That is why the normalization restriction of ordinary fuzzy partition,

$$\sum_{i=1}^{M} \mu_{A_i}(x) = 1, \quad \forall x \in U,$$

is redundant for these problems.

Definition 1. The *positive region* of a fuzzy set X in terms of the fuzzy partition A on the same universe U is expressed by

$$\mathrm{POS}_A^{\lambda_1}(X) = \bigcup_{I(A_i, X) \geq \lambda_1} A_i,$$

where $\cup$ stands for the union of fuzzy sets, $I(A_i, X)$ is a measure of inclusion of A_i in X (see for example [4]), and $\lambda_1 \in [0, 1]$ is a threshold, which expresses how 'strictly' the positive region of X is to be determined.

Definition 2. The *negative region* of a fuzzy set X in terms of the fuzzy partition A on the same universe U is expressed by

$$\mathrm{NEG}_A^{\lambda_2}(X) = \bigcup_{I(A_i, X) \leq \lambda_2} A_i,$$

where λ_2 is the respective threshold.

Definition 3. The *boundary region* of a fuzzy set X in terms of the fuzzy partition A on the same universe U is expressed by

$$\mathrm{BND}_A^{\lambda_2\lambda_1}(X) = \bigcup_{I(A_i, X) \in (\lambda_2, \lambda_1)} A_i.$$

It should be noted that when A and X are crisp, the definitions introduced above coincide with Pawlak's original definitions. The following simple considerations confirm this statement.

Assume that given P, Q, ordinary subsets of U, the inclusion grade of P in Q is

$$I(P, Q) = \frac{\mathrm{card}(P \cap Q)}{\mathrm{card}(P)}.$$

Then stating $\lambda_1 = 1.0$ and $\lambda_2 = 0.0$ the following relation (i.e. $A_i \subseteq U$, $A_i \cap A_j = \emptyset$, $\bigcup A_i = U$). For any subset $X \subseteq U$ the following notions are defined:

– *lower approximation of* X in terms of the partition A (also called *positive region of* X):

$$\mathrm{POS}_A(X) = A^{\mathrm{L}}(X) = \bigcup_{A_i \subseteq X} A_i;$$

– *upper approximation of* X in terms of the partition A:

$$A^{\mathrm{U}}(X) = \bigcup_{A_i \cap X \neq \emptyset} A_i;$$

– *Negative region of* X:

$$\mathrm{NEG}_A(X) = U - A^{\mathrm{U}}(X);$$

– *Boundary region of* X:

$$\mathrm{BND}_A(X) = A^{\mathrm{U}}(X) - A^{\mathrm{L}}(X);$$

– *quality of approximation of* X *by* A:

$$\mu_A(X) = \frac{\mathrm{card}(A^{\mathrm{L}}(X))}{\mathrm{card}(A^{\mathrm{U}}(X))}.$$

Obviously $\mu_A(X) \in [0, 1]$ and $\mu_A(X) = 1$ iff X is definable by A (i.e. $A^{\mathrm{U}}(X) = A^{\mathrm{L}}(X)$).

Let $B = \{B_1, \ldots, B_N\}$ be a second partition on U, induced by some other equivalence relation. Then the following definitions hold:

$$\mathrm{POS}_A(B) = \bigcup_{B_i \in B} A^{\mathrm{L}}(B_i),$$

$$\mathrm{BND}_A(B) = \bigcup_{B_i \in B} (A^{\mathrm{U}}(B_i) - A^{\mathrm{L}}(B_i)),$$

$$\mathrm{NEG}_A(B) = U - \bigcup_{B_i \in B} A^{\mathrm{U}}(B_i).$$

Then a plausible measure of dependency of B on A [15] is given by

$$\gamma_A(B) = \frac{\mathrm{card}(\mathrm{POS}_A(B))}{\mathrm{card}(U)}.$$

2.2. *Fuzzy rough sets*

Let $A = \{A_1, \ldots, A_M\}$ be a weak fuzzy partition on U, i.e.

$$\mu_{A_i}(x) \in [0, 1], \quad x \in U, \quad A_i \in A,$$

$$\bigcup_i \mathrm{supp}\, A_i = U.$$

The weak fuzzy partition is a convenient model of fuzzy classification problems, where A_i, $i = 1, \ldots, M$, correspond to the classes and $\mu_{A_i}(x)$ expresses a certain degree of membership (e.g. severity of illness) rather than a kind of uncertainty. That is why the normalization restriction of ordinary fuzzy partition,

$$\sum_{i=1}^{M} \mu_{A_i}(x) = 1, \quad \forall x \in U,$$

is redundant for these problems.

Definition 1. The *positive region* of a fuzzy set X in terms of the fuzzy partition A on the same universe U is expressed by

$$\mathrm{POS}_A^{\lambda_1}(X) = \bigcup_{I(A_i,\, X) \geq \lambda_1} A_i,$$

where $\cup$ stands for the union of fuzzy sets, $I(A_i, X)$ is a measure of inclusion of A_i in X (see for example [4]), and $\lambda_1 \in [0, 1]$ is a threshold, which expresses how 'strictly' the positive region of X is to be determined.

Definition 2. The *negative region* of a fuzzy set X in terms of the fuzzy partition A on the same universe U is expressed by

$$\mathrm{NEG}_A^{\lambda_2}(X) = \bigcup_{I(A_i,\, X) \leq \lambda_2} A_i,$$

where λ_2 is the respective threshold.

Definition 3. The *boundary region* of a fuzzy set X in terms of the fuzzy partition A on the same universe U is expressed by

$$\mathrm{BND}_A^{\lambda_2\lambda_1}(X) = \bigcup_{I(A_i,\, X) \in (\lambda_2,\, \lambda_1)} A_i.$$

It should be noted that when A and X are crisp, the definitions introduced above coincide with Pawlak's original definitions. The following simple considerations confirm this statement.

Assume that given P, Q, ordinary subsets of U, the inclusion grade of P in Q is

$$I(P, Q) = \frac{\mathrm{card}(P \cap Q)}{\mathrm{card}(P)}.$$

Then stating $\lambda_1 = 1.0$ and $\lambda_2 = 0.0$ the following equations are obtained:

$$\mathrm{POS}_A^{1,0}(X) = \bigcup_{I(A_i,X)\geq 1.0} A_i = \bigcup_{I(A_i,X)=1.0} A_i = \bigcup_{A_i\subseteq X} A_i,$$

$$\mathrm{BND}_A^{0,1}(X) = \bigcup_{I(A_i,X)\in(0,1)} A_i = \bigcup_{\substack{A_i\cap X\neq\emptyset\\ A_i\setminus X\neq\emptyset}} A_i,$$

$$\mathrm{NEG}_A^{0,0}(X) = \bigcup_{I(A_i,X)\leq 0.0} A_i = \bigcup_{I(A_i,X)=0} A_i = \bigcup_{A_i\cap X=\emptyset} A_i.$$

The above coincidence provides a guarantee that other related notions from rough sets theory can be successfully introduced in the sense of fuzzy sets. In this paper a new measure of approximation of a fuzzy set X in terms of a weak fuzzy partition A is proposed.

Definition 4. The *measure of approximation* of a fuzzy set X in terms of the weak fuzzy partition A on the same universe U is

$$\nu_A^{\lambda_2\lambda_1}(X) = \tfrac{1}{2}(S(\mathrm{POS}_A^{\lambda_1}(X), X) + 1 - S(\mathrm{BND}_A^{\lambda_2\lambda_1}(X), X)), \quad (1)$$

where $S(P, Q)$ is a measure of similarity between two fuzzy sets P and Q on the same universe (see e.g. [4]).

Since $S(P, Q) \in [0, 1]$, it is clear that the measure of approximation defined above varies in the interval $[0, 1]$. Moreover it takes value 1 if X is completely definable by A, and 0 if X is completely undefinable. The extreme values of the measure (from the interval $[0, 1]$) depend on the type of measure of similarity S chosen and on the properties of the weak fuzzy partition.

Let $B = \{B_1, \ldots, B_N\}$ be another weak fuzzy partition on U. The problem of assessment of approximation of B via A can be considered in two ways:

(a) By analogy with Pawlak's concepts [15], the positive, negative and boundary regions of the weak fuzzy partition B can be defined and then the above introduced measure of approximation can be appropriately applied.

(b) The measure $\mu_A^{\lambda_2\lambda_1}(B_i)$ can be determined for every $B_i \in B$ and then an aggregation rule can be applied to obtain the final value of approximation.

In the sense of fuzzy pattern recognition the second approach seems more reasonable. In the first approach all fuzzy sets in A and B are treated as equally significant in determination of the final result, whereas the second one enables the designer to use the achievements of fuzzy decision theory and to treat the sets stating different weights.

3. Application of fuzzy rough sets to feature selection

3.1. Fuzzy pattern recognition

Fuzzy pattern recognition deals with problems where objects and classes are described in fuzzy way and objects may have multiclass membership. The degree of membership stands either for an estimate of certainty, belief, typicalness, strength of confirmation, etc., or for a characteristic which does not stem from uncertainty. In real problems this characteristic may correspond to e.g. degree of illness, severity of the disease, etc.

Let

$$\mathbb{X} = \{X_1, X_2, \ldots, X_n\}$$

be the feature set (which in medical problems is the set of findings, measurements, or signs for the patients),

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_M\}$$

the set of classes (diseases or physiological states) and

$$\mathbb{Z} = \{Z_1, Z_2, \ldots, Z_N\}$$

the sample.

The *prior classification* of objects is the set $\mu = \{\mu_1^a, \mu_2^a, \ldots, \mu_M^a\}$, which consists of M initial fuzzy sets, corresponding to the classes, with membership functions μ_i^a of the form

$$\mu_i^a : Z \to [0, 1].$$

The value $\mu_i^a(Z_j)$ is the degree in which the object Z_j, $j = 1, \ldots, N$, belongs to the class ω_i, $i = 1, \ldots, M$. These values are initially stated by interviewing an expert council or by measuring some real characteristic of the object.

In fact, this is the initial (weak) partition on the set $\mathbb{Z}$. The interpretation of prior classification could be elucidated considering an example from medical practice – a problem of Ischemic Heart Disease (IHD) diagnosis. Each object Z_j from the sample $\mathbb{Z}$ corresponds to a patient, described by a set of features $\mathbb{X}$ (containing anamnesis, laboratory, clinical etc. measurements). Two fuzzy sets ω_1 and ω_2 on $\mathbb{Z}$ for the classes *healthy* and *affected*, respectively, form the weak fuzzy partition. The degree $\mu_{\omega_1}(Z_j)$ denotes the degree of 'health' of the patient, and $\mu_{\omega_2}(Z_j)$ the degree of 'illness' respectively. The weak fuzzy partition is an expression of the following fact. A person can be comparatively healthy, even able to practice some kind of sport ($\mu_{\omega_1}(x) = 0.6$, subjectively assessed) and in the same time suffer from a latent form of IHD ($\mu_{\omega_2}(x) = 0.7$, measured by % of occlusion of main coronary arteries). Obviously, a statistical interpretation could hardly be used in these problems.

A fuzzy classifier is to be designed, which yields degress of membership for a given set of values for the features from $\mathbb{X}$. For the objects from $\mathbb{Z}$ the degrees obtained are to be as close as possible to the initial ones. Thus the posterior classification of objects $\mu^{\mathrm{p}} = \{\mu_1^{\mathrm{p}}, \mu_2^{\mathrm{p}}, \ldots, \mu_M^{\mathrm{p}}\}$ is obtained, which forms a second partition on $\mathbb{Z}$.

A great variety of strategies and methods exists for fuzzy classification which use: a linguistic approach based on fuzzy relational equations and fuzzy logic; fuzzy analogies of Bayes' recognition method; fuzzy k-Nearest Neighbor methods; fuzzy linear classification rules; fuzzy multistage and multi-level classification schemes; etc. Only few papers, however, consider criteria for feature selection (e.g. [3, 12, 18]) and for assessing the accuracy of fuzzy classification [1, 7, 10, 19, 23]. These criteria are predominantly defined to compare two fuzzy sets representing prior and posterior classifications respectively.

Here a feature selection criterion for fuzzy pattern recognition is proposed on the basis of the above defined measure of approximation (1).

3.2. *Feature selection via fuzzy rough sets*

Let $\mu^{\mathrm{a}} = \{\mu_1^{\mathrm{a}}, \mu_2^{\mathrm{a}}, \ldots, \mu_M^{\mathrm{a}}\}$ be a weak fuzzy partition on the sample of objects $\mathbb{Z}$ and let $\mathbb{G} \subseteq \mathbb{X}$ be a subset of the initial feature set $\mathbb{X}$. Using some fuzzy clustering procedure, the weak fuzzy partition

$$\mu^{\mathrm{p}}(\mathbb{G}) = \{\mu_1^{\mathrm{p}}, \mu_2^{\mathrm{p}}, \ldots, \mu_L^{\mathrm{p}}\}$$

on $\mathbb{Z}$ can be obtained. Then a criterion for approximation of μ^{a} in terms of $\mu^{\mathrm{p}}(\mathbb{G})$ is to be designed, as described in the previous section:

$$\mathbb{J}(\mathbb{G}) = \sum_{i=1}^{M} w_i v^{\lambda_2\lambda_1}_{\mu^{\mathrm{p}}(G)}(\mu_i^{\mathrm{a}}),$$

where w_i are weights.

This is an evaluation of the capacity of the feature set $\mathbb{G}$, which is to be taken into account in classifier design. The optimal feature set can be found as

$$\mathbb{G}^* = \arg\max_{\mathbb{G} \subseteq \mathbb{X}} \mathbb{J}(\mathbb{G}).$$

Note that the number of classes for the second partition is not definitely M. This is one of the main advantages of the proposed approach. The criterion evaluates the subset $\mathbb{G}$ taking into account the partition which this subset induces in spite of the prior classification. This advantage will be illustrated by the following example.

Let us consider the hypothetical case of one feature and one class presented in Figure 1. The prior membership function μ^{a} (denoted as 'Initial') is to be approached as close as possible by the posterior function. Suppose that two posterior functions are available, $\mu^{\mathrm{p}1}$ and $\mu^{\mathrm{p}2}$, which induce partition 1 and partition 2, respectively. In order to estimate the closeness between each posterior membership function and the prior one, five measures of similarity

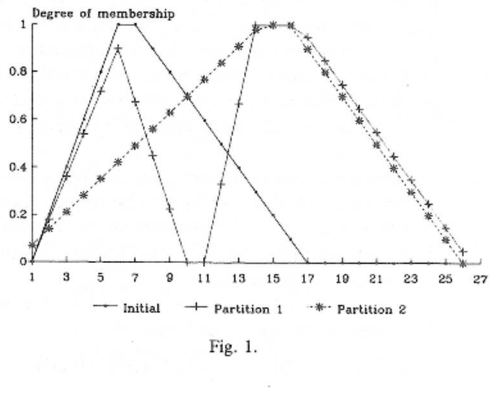

Fig. 1.

Table 1

Similarity measure	$S(\mu^a, \mu^{p1})$	$S(\mu^a, \mu^{p2})$	$S(\mu^a, \mu^{p1})$	
			Cluster 1	Cluster 2
S_1	0.330	0.357	0.471	0·082
S_2	0.583	0.593	0.827	0.429
S_3	0.502	0.518	0.772	0.404
S_4	0.050	0.100	0.300	0.000
S_5	0.050	0.100	0.225	0.000

between two fuzzy sets P and Q are used [4]:

$$S_1(P, Q) = \|P \cap Q\| / \|P \cup Q\|, \tag{2}$$

$$S_2(P, Q) = 1 - \|P \nabla Q\|, \tag{3}$$

$$S_3(P, Q) = 1 - \|P \triangle Q\|,$$

$$S_4(P, Q) = 1 - \sup_x \mu_{P\nabla Q}(x),$$

$$S_5(P, Q) = 1 - \sup_x \mu_{P\triangle Q}(x),$$

where $\cap$ denotes the intersection of fuzzy sets, $\cup$ the union, $\|\cdot\|$ the relative cardinality of a fuzzy set, ∇ stands for Hamming distance between fuzzy sets and $\triangle$ for a symmetrical difference of the form

$$\mu_{P\triangle Q}(x) = \max(\min[\mu_P(x), 1 - \mu_Q(x)], \min[1 - \mu_P(x), \mu_Q(x)]).$$

The values of these measures are presented in Table 1. According to them there is no reason to prefer one posterior membership function to the other.

It could be seen from Figure 1 that the membership function μ^{p1} implies a fuzzy partition consisting of two clusters. If we could detect the objects falling in the left cluster the prior membership function will be approximated in a higher degree. In confirmation of this the above five measures are separately calculated for the left and for the right clusters of partition 1. Results are presented in Table 1. In the multidimensional case, however, the information of this type is hidden and could be detected only if fuzzy clustering is performed to obtain a posterior partition which is the underlying idea of the proposed criterion.

The values of the measure v for the two partitions using measure S_2 (3) are:

$$v^{0.2,0.8}_{\mu^{p1}}(\mu^a) = 0.914, \qquad v^{0.2,0.8}_{\mu^{p2}}(\mu^a) = 0.204.$$

It seems clear that these two posterior membership functions are no more indiscernible and the partition generated by μ^{p1} is highly preferable.

It should be pointed out that this paper discusses only the criterion for feature selection. The ways of searching through the set of subsets of $\mathbb{X}$ can be referred to e.g. [2].

4. Experimental illustration

The proposed feature selection criterion is illustrated with a set of real data from aviation medicine [9]. The sample consists of 7 measurements of systolic blood pressure of 200 persons, examined in a hypobaric chamber (barocamera) in order to estimate their hypoxic resistance. The 7 values have been measured during the experiment in the following moments:
- before the examination;
- at the 1st, 5th, 10th, 15th and 30th minute in the barocamera;
- after the examination.

Two classes have been formed: Class 1, *high hypoxic resistance* and Class 2, *low hypoxic resistance*. The respective membership grades have been stated by expert opinion.

The results obtained are shown in Table 2. The values of the respective measures of similarity S_1 to S_5 are also provided. Only single features were evaluated using fuzzy rough sets. The clustering was done using frequency tables and frequency polygons of the respective features.

The following formula for inclusion grade [4] is used to calculate the POS(Ω), BND(Ω):

$$I(P, Q) = \frac{\|P \cap Q\|}{\|P\|},$$

where P and Q are fuzzy sets on the universe U. The similarity measure for calculating the respective measure of approximation is S_1 (2).

The results coincide with the expert opinion, in the respect that the systolic blood pressure is more informative in the barocamera than after the examination. Moreover the measurement in the first minute of examination has low significance, because the first reaction of the body to the extreme condition is very different

Table 2

Feature	Similarity measures					Number of clusters	Value of the proposed criterion
	S_1	S_2	S_3	S_4	S_5		
Before examination	0.56	0.68	0.53	0.00	0.00	3	0.58
1st minute	0.50	0.60	0.48	0.05	0.05	4	0.47
5th minute	0.50	0.62	0.48	0.10	0.07	3	0.61
10th minute	0.43	0.57	0.44	0.02	0.02	2	0.60
15th minute	0.40	0.57	0.43	0.00	0.00	2	0.57
30th minute	0.45	0.61	0.46	0.00	0.00	3	0.56
After examination	0.50	0.61	0.49	0.02	0.02	2	0.32

and not very important for the decision. The values of measure of approximation are nearly indistinguishable in the barocamera, but there are considerable differences between values for the classes. This can be taken into account not only in the further feature selection process, but also in the final choice of optimal feature set. It should be mentioned that this medical interpretation does not hold for the values of similarity measures.

Taking into account the presumption that for each method there exists a data set for which this method yields perfect results, the positive experimental results could not be considered as unambiguous evidence in favor of the proposed criterion. On the other hand they allow a sensible interpretation which illustrates the application possibilities of the proposed idea.

5. Conclusions

The present paper dealt with the application of the rough sets idea to fuzzy sets. A measure of classification accuracy, which can serve also as a criterion for feature selection, was developed in order to be used in fuzzy pattern recognition. It was shown that the fuzzy concepts of positive, negative and boundary region of a set proposed in the paper coincide with the original ones if appropriately applied to ordinary rough classification.

Feature selection via fuzzy rough sets was illustrated using real data from medical practice. The results are closely related with the expert opinion on the problem. This gives some hope for further application of fuzzy rough sets in real problems, described by fuzzy recognition model.

References

[1] L.I. Borodkin and O.Ye. Stadnik, Algorithms for construction of decision rule in pattern recognition by using fuzzy sets, *Avtomatika i Telemehanika* (11) (1985) 108–116 (in Russian).

[2] P.A. Devijver and J. Kittler, *Pattern Recognition. A Statistical Approach* (Prentice-Hall, London–Englewood Cliffs, 1982).

[3] V. Di Gesu and M.C. Maccarone, Feature selection and possibility theory, *Pattern Recognition* **19** (1986) 63–72.

[4] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications* (Academic Press, New York 1980).

[5] D. Dubois and H. Prade, Twofold fuzzy sets and rough sets – Some issues in knowledge representation, *Fuzzy Sets and Systems* **23** (1987) 3–18.

[6] J. Fibak, Z. Pawlak, K. Słowiński and R. Słowiński, Rough sets based decision algorithm for treatment of duodenal ulcer by HSV, *Bulletin of PAS* **34** (1986) 227–248.

[7] A. Jóźwik, A learning scheme for a fuzzy *k*-NN rule, *Pattern Recognition Lett.* **1** (1983) 287–289.

[8] M. Kubat, Floating approximation in time-varying knowledge bases, *Pattern Recognition Lett.* **10** (1989) 223–227.

[9] K.H. Kunchev, Investigations on the resistance of the flying staff in respect to a moderate degree of hypoxia, Thesis, Sofia (1981) (in Bulgarian).

[10] L.I. Kuncheva, Evaluation of computerized medical diagnostic decisions via fuzzy sets, *Internat. J. Biomed. Comput.* **28** (1991) 91–100.

[11] A. Mrozek, Rough sets and dependency analysis among attributes in computer implementations of expert's inference models, *Internat. J. Man–Machine Stud.* **30** (1989) 457–473.

[12] S.K. Pal and B. Chakraborty, Intraclass and interclass ambiguities (fuzziness) in feature evaluation, *Pattern Recognition Lett.* **2** (1984), 275–279.

[13] Z. Pawlak, Rough sets and fuzzy sets, *Fuzzy Sets and Systems* **17** (1985) 99–102.

[14] Z. Pawlak, Rough sets, *Internat. J. Comput. and Inform. Systems* **11** (1982) 341–356.

[15] Z. Pawlak, Rough classification, *Internat. J. Man–Machine Stud.* **20** (1984) 469–483.

[16] Z. Pawlak, K. Słowiński and R. Słowiński, Rough classification of patients after highly selective vagotomy for duodenal ulcer, *Internat. J. Man–Machine Stud.* **24** (1986) 413–433.

[17] Z. Pawlak, S.K.M. Wong and W. Ziarko, Rough sets: Probabilistic versus deterministic approach, *Internat. J. Man–Machine Stud.* **29** (1988) 81–95.

[18] W. Pedrycz, Fuzzy sets in pattern recognition: Methodology and methods, *Pattern Recognition* **23** (1990) 121–146.

[19] W. Pedrycz, Classification in a fuzzy environment, *Pattern Recognition Lett.* **3** (1985) 303–308.

[20] K. Słowiński, R. Słowiński and J. Stefanowski, Rough sets approach to analysis of data from peritoneal lavage in acute pancreatitis, *Med. Inform.* **13** (1988) 143–159.

[21] K. Słowiński and R. Słowiński, Sensitivity analysis of rough classification, *Internat. J. Man–Machine Stud.* **32** (1990) 693–705.

[22] R. Słowiński and J. Stefanowski, Rough classification in incomplete information systems, *Math. and Comput. Modeling J.* (1989).

[23] O.Ye. Stadnik, Fuzzy recognition rules in medical diagnosis and prediction, *Avtomatika i Telemehanika* (3) (1986) 129–135 (in Russian).

[24] S.K.M. Wong and W. Ziarko, Comparison of the probabilistic approximate classification and the fuzzy set model, *Fuzzy Sets and Systems* **21** (1987) 357–362.

[25] M. Wygralak, Rough sets and fuzzy sets – Some remarks on interrelations, *Fuzzy Sets and Systems* **29** (1989) 241–243.