A Roughness Measure for Fuzzy Sets

Van-Nam Huynh *, Yoshiteru Nakamori,

School of Knowledge Science
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Tatsunokuchi, Ishikawa 923-1292, Japan

Abstract

Recently, an attempt of integration between the theories of fuzzy sets and rough sets has resulted in providing a roughness measure for fuzzy sets (Banerjee and Pal, 1996). Essentially, Banerjee and Pal's roughness measure depends on parameters that are designed as thresholds of definiteness and possibility in membership of the objects to a fuzzy set. In this paper we first remark that this measure of roughness has several undesirable properties, and then propose a parameter-free roughness measure for fuzzy sets based on the notion of the mass assignment of a fuzzy set. Several interesting properties of this new measure are examined. Furthermore, we also discuss how the proposed approach is used to describe the rough approximation quality of a fuzzy classification.

Key words: Rough sets, fuzzy sets, roughness measure, rough fuzzy sets

1 Introduction

The notion of fuzzy sets aimed at mathematically modeling vague concepts was first introduced by Zadeh in [27], in connection with the representation and manipulation of human knowledge automatically. Formally, the theory of fuzzy sets is a generalization of classical set theory, making use of the notion of partial degrees of membership. Practically, the theory of fuzzy sets provides a systematic framework for dealing with complex phenomena in describing the behavior of systems which do not lend themselves to analysis by classical methods based on probability theory and bivalent logic. Since its inception, the mathematical foundation as well as extensive application of the theory to many different areas have already been well established [11].

* Corresponding author.
Email address: huynh@jaist.ac.jp (Van-Nam Huynh).

To appear in Information Sciences

After nearly twenty years since the introduction of fuzzy sets, Pawlak [17] introduced the notion of a rough set as a new mathematical tool to deal with the approximation of a concept in the context of incomplete information. Basically, while a fuzzy set models the ill-definition of the boundary of a concept often described linguistically, a rough set characterizes a concept by its lower and upper approximations due to indiscernibility between objects arose because of incompleteness of available knowledge. The rough set theory is also proven to be of substantial importance in many areas of application (cf. [16,18,20]).

Since the introduction of rough set theory, many attempts to establish the relationships between the two theories, to compare each to the other, and to simultaneously hybridize them have been made (e.g. [7,14,16,19,22,21,24,25]). As an attempt in the line of integration between the two theories, Banerjee and Pal [4] have recently proposed a roughness measure for fuzzy sets, making use of the concept of a rough fuzzy set [7]. Although this measure of roughness satisfies some interesting properties and has potential applications in the field of pattern recognition as mentioned in [4]. At the same time, it also has some undesirable properties as remarked in the following. In this paper we propose an alternative roughness measure for fuzzy sets based on the notions of the mass assignment of a fuzzy set and its α-cuts. Several properties of this measure are examined. We moreover discuss how to extend the notion of *rough approximation quality* γ within the context considered in this paper.

The structure of the rest of this paper is as follows. Section 2 briefly introduces necessary notions of rough sets and fuzzy sets, the mass assignment of a fuzzy set as well as its relation to the probabilistic based semantics of a membership function. In Section 3, after recalling the notion of a rough fuzzy set [7], Benerjee and Pal's approach to roughness of a fuzzy set [4] is reviewed. In Section 4, a new roughness measure of a fuzzy set is proposed based on the mass assignment of the fuzzy set, and several of its properties are proved. Section 5 discusses some possible extensions of the approximation quality, which may be used to describe the rough approximation quality of a fuzzy classification. Finally, some concluding remarks are presented in Section 6.

2 Preliminaries

In this section we recall basic notions in the theories of rough sets and fuzzy sets. For the purpose of this paper we confine ourselves the consideration on only the finite version of universes of discourse.

Formally, the theory begins with the notion of an approximation space, which is a pair $\langle U, R\rangle$, where U be a non-empty set (the universe of discourse) and R an equivalence relation on U, i.e., R is reflexive, symmetric, and transitive. The relation R decomposes the set U into disjoint classes in such a way that two elements x, y are in the same class iff $(x, y) \in R$. Let denote by U/R the quotient set of U by the relation R, and

$$U/R = \{X_1, X_2, \ldots, X_m\}$$

where X_i is an equivalence class of $R, i = 1, 2, \ldots, m$.

If two elements x, y in U belong to the same equivalence class $X_i \in U/R$, we say that x and y are indistinguishable. The equivalence classes of R and the empty set $\emptyset$ are the elementary sets in the approximation space $\langle U, R\rangle$.

Note that in the context of rough-set-based data analysis, the equivalence relation in an approximation space is often interpreted via the notion of information systems [1]. An information system with U as universe would be a triple $\langle U, \mathcal{A}, \{V_a\}_{a\in\mathcal{A}}\rangle$, where U is a set of objects, $\mathcal{A}$ is a set of attributes and V_a the set of attribute values for attribute a undestood as a mapping $a : U \to V_a$. It is easily seen that each subset A of the attribute set $\mathcal{A}$ induces an equivalence relation $\mathrm{ind}(A)$ called *indiscernability relation* as follows

$$\mathrm{ind}(A) = \{(x, y) \in U \times U | a(x) = a(y), \text{ for all } a \in A\}$$

and $\mathrm{ind}(A) = \bigcap\limits_{a\in A} \mathrm{ind}(a)$, where $\mathrm{ind}(a)$ means $\mathrm{ind}(\{a\})$. If $(x, y) \in \mathrm{ind}(A)$ we then say that the objects x and y are indiscernible with respect to attributes from A. In other words, we cannot distinguish x from y, and vice versa, in terms of attributes in A.

Let us return to an arbitrary approximation space $\langle U, R\rangle$. Given an arbitrary set $X \in 2^U$, in general it may not be possible to describe X precisely in $\langle U, R\rangle$. One may characterize X by a pair of lower and upper approximations defined as follows [17]:

$$\underline{R}(X) = \bigcup_{X_i \subseteq X} X_i; \qquad \overline{R}(X) = \bigcup_{X_i \cap X \neq \emptyset} X_i$$

That is, the lower approximation $\underline{R}(X)$ is the union of all the elementary sets which are subsets of X, and the upper approximation $\overline{R}(X)$ is the union of all the elementary sets which have a non-empty intersection with X. The interval

[1] also called data tables, databases, attribute-value systems, etc.

$[\underline{R}(X), \overline{R}(X)]$ is the representation of an ordinary set X in the approximation space $\langle U, R\rangle$ or simply called the rough set of X.

The lower (resp., upper) approximation $\underline{R}(X)$ (resp., $\overline{R}(X)$) is interpreted as the collection of those elements of U that definitely (resp., possibly) belong to X. Further, we also define:

- a set $X \in 2^U$ is said to be *definable* (or *exact*) in $\langle U, R\rangle$ iff $\underline{R}(X) = \overline{R}(X)$.
- for any $X, Y \in 2^U$, X is said to be *roughly included* in Y, denoted by $X \tilde{\subset} Y$, iff $\underline{R}(X) \subseteq \underline{R}(Y)$ and $\overline{R}(X) \subseteq \overline{R}(Y)$.
- X and Y is said to be *roughly equal*, denoted by $X \approx_R Y$, in $\langle U, R\rangle$ iff $\underline{R}(X) = \underline{R}(Y)$ and $\overline{R}(X) = \overline{R}(Y)$.

In [18], Pawlak discusses two numerical chacracterizations of imprecision of a subset X in the approximation space $\langle U, R\rangle$: *accuracy* and *roughness*. Accuracy of X, denoted by $\alpha_R(X)$, is simply the ratio of the number of objects in its lower approximation to that in its upper approximation; namely

$$\alpha_R(X) = \frac{|\underline{R}(X)|}{|\overline{R}(X)|}$$

where $|\cdot|$ denotes the cardinality of a set. Then the roughness of X, denoted by $\rho_R(X)$, is defined by subtracting the accurcy from 1:

$$\rho_R(X) = 1 - \alpha_R(X) = 1 - \frac{|\underline{R}(X)|}{|\overline{R}(X)|}$$

Note that the lower the roughness of a subset, the better is its approximation. Further, the following observations are easily obtained:

(1) As $\underline{R}(X) \subseteq X \subseteq \overline{R}(X)$, $0 \leq \rho_R(X) \leq 1$.
(2) By convention, when $X = \emptyset$, $\underline{R}(X) = \overline{R}(X) = \emptyset$ and $\rho_R(X) = 0$.
(3) $\rho_R(X) = 0$ if and only if X is definable in $\langle U, R\rangle$.

In [23], Yao has interpreted Pawlak's accuracy measure in terms of a classic distance measure based on sets, called *Marczewski-Steinhaus metric* [13], which is defined by

$$D_{MZ}(X, Y) = \frac{|X \cup Y| - |X \cap Y|}{|X \cup Y|} = 1 - \frac{|X \cap Y|}{|X \cup Y|}$$

Using the Marczewski-Steinhaus metric, the roughness measure of a set X in $\langle U, R\rangle$ is the distance between its lower and upper approximations. Moreover, it would be worth noting that this metric has been also used to re-interpret the so-called rough approximation quality γ and ascertain its statistical significance in [8]. Related to the context of this paper, a recent survey on probabilistic approaches to rough sets could be found in [26].

2.2 Fuzzy Sets

Let U be a finite and non-empty set called universe. A fuzzy set F of U is a mapping from U into the unit interval $[0, 1]$:

$$\mu_F : U \longrightarrow [0, 1]$$

where for each $x \in U$ we call $\mu_F(x)$ the membership degree of x in F. Practically, we may consider U as a set of objects of concern, and a crisp subset of U represent a "non-vague" concept imposed on objects in U. Then a fuzzy set F of U is thought of as a mathematical representation of a "vague" concept described linguistically.

Usually in the setting of fuzzy sets, the set theoretic operations union, intersection and complement can be realized with the aid of the so-called De Morgan triples [11], among which the system $(\max, \min, 1-)$ is the most widely-used.

Given a number $\alpha \in (0, 1]$, the α-cut, or α-level set, of F is defined as follows

$$F_\alpha = \{x \in U | \mu_F(x) \geq \alpha\}$$

which is a subset of U. A strong α-cut is defined by

$$F_{\alpha^+} = \{x \in U | \mu_F(x) > \alpha\}$$

An implication of the $(\max, \min, 1-)$ system is that the set theoretic operations on fuzzy sets can be defined by those on α-cuts. However, this may be not the case when another pair of t-norm and t-conorm is used.

In [9], a formal connection between fuzzy sets and random sets was established. Furthermore, viewing the body of evidence as a random set [15], a fuzzy set F is a consonant random set; the family of its α-cuts forms a nested family of focal elements [6,7]. Note that in this case the normalization assumption of F is imposed due to the body of evidence does not contain the empty set. Interestingly, this view of fuzzy sets has been used by Baldwin in [1,2] to introduce the so-called mass assignment of a fuzzy set with relaxing the assumption, and to provide a probability based semantics for a fuzzy concept defined as a family of possible definitions of the concept. These varying definitions may be provided by a population of voters accompanied with a probability distribution where each voter is asked to give a crisp definition of the concept. The mass assignment of a fuzzy set is defined as follows.

Let F be a fuzzy subset of a finite universe U such that the range of the membership function μ_F, denoted by $\text{rng}(\mu_F)$, is $\text{rng}(\mu_F) = \{\alpha_1, \ldots, \alpha_n\}$, where $\alpha_i > \alpha_{i+1} > 0$, for $i = 1, \ldots, n-1$. Then the mass assignment of F, denoted by m_F, is a probability distribution on 2^U satisfying

$$m_F(\emptyset) = 1 - \alpha_1, m_F(F_i) = \alpha_i - \alpha_{i+1}, \text{ for } i = 1, \ldots, n$$

with $\alpha_{n+1} = 0$ by convention, where

$$F_i = \{x \in U | \mu_F(x) \geq \alpha_i\}$$

for $i = 1, \ldots, n$. $\{F_i\}_{i=1}^n$ (or $\{F_i\}_{i=1}^n \cup \{\emptyset\}$ if F is a subnormal fuzzy set, i.e. $\max_{x\in U}\{\mu_F(x)\} < 1$) are referred to as the focal elements of m_F. The mass assignment of a fuzzy concept is then considered as providing a probability based semantics for membership function of the fuzzy concept. A formal connection between this semantics of membership functions and the modal logic based interpretation of fuzzy concepts has been established in [10]. The mass assignment theory has been applied in some fields such as induction of decision trees [3], computing with words [12], among others.

In the following we utilize this notion of the mass assignment of a fuzzy set to introduce a new roughness measure for fuzzy sets within a given approximation space.

3 Roughness Measure of a Fuzzy Set: Banerjee and Pal's Approach

First, we briefly recall the notion of a rough fuzzy set and allied notions [7] that forms the basis of a roughness measure of a fuzzy set proposed in [4].

3.1 Rough Fuzzy Sets

Given a finite approximation space $\langle U, R\rangle$. Let F be a fuzzy set in U with the membership function μ_F. The upper and lower approximations $\overline{R}(F)$ and $\underline{R}(F)$ of F by R are fuzzy sets in the quotient set $U/R = \{X_1, X_2, \ldots, X_m\}$ with membership functions defined [7] by

$$\mu_{\overline{R}(F)}(X_i) = \max_{x\in X_i}\{\mu_F(x)\}$$

$$\mu_{\underline{R}(F)}(X_i) = \min_{x\in X_i}\{\mu_F(x)\}$$

for $i = 1, \ldots, m$. $(\underline{R}(F), \overline{R}(F))$ is called a rough fuzzy set.

The rough fuzzy set $(\underline{R}(F), \overline{R}(F))$ then induces two fuzzy sets F^* and F_* in U with membership functions are defined respectively as follows

$$\mu_{F^*}(x) = \mu_{\overline{R}(F)}(X_i) \text{ and } \mu_{F_*}(x) = \mu_{\underline{R}(F)}(X_i)$$

if $x \in X_i$, for $i = 1, \ldots, m$. That is, F^* and F_* are fuzzy sets with constant membership degree on the equivalence classes of U by R, and for any $x \in U$, $\mu_{F^*}(x)$ (respectively, $\mu_{F_*}(x)$) can be viewed as the degree to which x possibly (respectively, definitely) belongs to the fuzzy set F [4].

Under such a view, we now define the notion of a *definable fuzzy set* in $\langle U, R\rangle$. A fuzzy set F is called a *definable* if $\underline{R}(F) = \overline{R}(F)$, i.e. there exists a fuzzy set $\mathcal{F}$ in U/R such that $\mu_F(x) = \mu_{\mathcal{F}}(X_i)$ if $x \in X_i$, $i = 1 \ldots, m$. Further, as defined in [4], fuzzy sets F and G in U are said to be *roughly equal*, denoted by $F \approx_R G$, if and only if

$$\underline{R}(F) = \underline{R}(G) \text{ and } \overline{R}(F) = \overline{R}(G)$$

The following proposition states fundamental properties on rough fuzzy sets taken from [4].

Proposition 1 *Let F and G be fuzzy sets in U. Then*

(1) $F_* \subseteq F \subseteq F^*$,
(2) $\overline{R}(F \cup G) = \overline{R}(F) \cup \overline{R}(G)$,
(3) $\underline{R}(F \cap G) = \underline{R}(F) \cap \underline{R}(G)$,
(4) $\underline{R}(F) \cup \underline{R}(G) \subseteq \underline{R}(F \cup G)$,
(5) $\overline{R}(F \cap G) \subseteq \overline{R}(F) \cap \overline{R}(G)$,
(6) $\overline{R}(F^c) = \underline{R}(F)^c$, $\overline{R}(F)^c = \underline{R}(F^c)$.

where fuzzy set operations union, intersection and complement are defined via max*,* min *and* $1-$*, respectively.*

3.2 Banerjee and Pal's Roughness Measure

In [4], Banerjee and Pal have proposed a roughness measure for fuzzy sets in a given approximation space. Essentially, this measure of roughness of a fuzzy set depends on parameters that are designed as thresholds of definiteness and possibility in membership of the objects in U to the fuzzy set.

More explicitly, let us be given a finite approximation space $\langle U, R\rangle$ and a fuzzy set F in U. We now consider parameters α, β such that $0 < \beta \leq \alpha \leq 1$. The α-cut $(F_*)_\alpha$ and β-cut $(F^*)_\beta$ of fuzzy sets F_* and F^*, respectively, are called to be the *α-lower approximation*, the *β-upper approximation* of F in $\langle U, R\rangle$, respectively. Then a roughness measure of the fuzzy set F with respect to parameters α, β with $0 < \beta \leq \alpha \leq 1$, and the approximation space $\langle U, R\rangle$ is defined by

$$\rho_R^{\alpha,\beta}(F) = 1 - \frac{|(F_*)_\alpha|}{|(F^*)_\beta|}$$

It is obvious that this definition of roughness measure $\rho_R^{\alpha,\beta}(\cdot)$ strongly depends on parameters α and β.

Remark 2 *By the assumption made on parameters, we have*

(1) $0 \leq \rho_R^{\alpha,\beta}(F) \leq 1$.
(2) If F is such that there is a member x in each equivalence class X_i ($i = 1, \ldots, m$) with $\mu_F(x) < \alpha$, then $\rho_R^{\alpha,\beta}(F) = 1$.
(3) If F is a definable fuzzy set, i.e. μ_F is a constant function on each equivalence class X_i ($i = 1 \ldots, m$) and $\alpha = \beta$, then $\rho_R^{\alpha,\beta}(F) = 0$.

Note that while the third statement seems interesting as it says that the measure $\rho_R^{\alpha,\beta}(\cdot)$ inherits a property of Pawlak's roughness measure, that is the roughness of a definable subset is equal to 0, the second one may not be well-justified. Furthermore, the following property of $\rho_R^{\alpha,\beta}(\cdot)$ proved in [4] may be also undesired, unless the support of a constant fuzzy set, i.e. its strong 0-cut, is definable in the approximation space.

Proposition 3 *If F is a constant fuzzy set, say $\mu_F(x) = \delta$, for all $x \in U$, then $\rho_R^{\alpha,\beta}(F) = 0$, with the exception when $\beta < \delta < \alpha$, in which case $\rho_R^{\alpha,\beta}(F) = 1$.*

Several properties of the measure $\rho_R^{\alpha,\beta}(\cdot)$ and its potential applications in the field of pattern recognition have been reported and mentioned in [4], and more recently in [28].

4 A New Roughness Measure for Fuzzy Sets

In this section we introduce a parameter-free measure of roughness of a fuzzy set that in fact is a generalization of Pawlak's notion of roughness measure and avoids the undesirable properties held by Banerjee and Pal's roughness measure as mentioned above.

Given a finite approximation space $\langle U, R\rangle$ and a normal fuzzy set F in U. Assume that the range of the membership function μ_F is $\mathrm{rng}(\mu_F) = \{\alpha_1, \ldots, \alpha_n\}$, where $\alpha_i > \alpha_{i+1} > 0$, for $i = 1, \ldots, n-1$, and $\alpha_1 = 1$. Let us denote m_F the mass assignment of F defined as in the preceding section. Let

$$F_i = \{x \in U | \mu_F(x) \geq \alpha_i\}, \text{ for } i = 1, \ldots, n$$

It is of interest that if $\{F_i\}_{=1}^n$ could be interpreted as a family of possible definitions of the concept F, then $m_F(F_i)$ is the probability of the event "the concept is F_i", for each i.

Under such an interpretation, we now define the roughness measure of F with respect to the approximation space $\langle U, R\rangle$ as follows

$$\hat{\rho}_R(F) = \sum_{i=1}^{n} m_F(F_i)(1 - \frac{|\underline{R}(F_i)|}{|\overline{R}(F_i)|}) \equiv \sum_{i=1}^{n} m_F(F_i)\rho_R(F_i)$$

That is, the roughness of a fuzzy set F is the weighted sum of the roughness measures of nested focal subsets which are considered as its possible definitions.

Remark 4 • *Clearly,* $0 \leq \hat{\rho}_R(F) \leq 1$.

- $\hat{\rho}_R(\cdot)$ *is a natural extension of Pawlak's roughness measure for fuzzy sets, i.e. if* F *is a crisp subset of* U *then* $\hat{\rho}_R(F) = \rho_R(F)$.
- F *is a definable fuzzy set if and only if* $\hat{\rho}_R(F) = 0$.

For the sake of illustration, let us consider a very simple example depicting the introduced notion.

Example 5 *Suppose we are given an approximation space* $\langle U, R\rangle$*, where* $U = \{0, 1, 2, \ldots, 10\}$ *and* R *is such that*

$$U/R = \{\{0, 2, 4\}, \{1, 3, 5\}, \{6, 8, 10\}, \{7, 9\}\}$$

Let us consider a linguistic value small *whose membership function is defined by Table 1:*

Table 1
The membership function of *small*

u	0	1	2	3	4	5	6	7	8	9	10
$\mu_{small}(u)$	1	1	0.75	0.5	0.25	0	0	0	0	0	0

The approximations of the fuzzy set μ_{small} *in* $\langle U, R\rangle$ *are given in Table 2.*

Table 2
The approximations of the fuzzy set μ_{small}

	{0,2,4}	{1,3,5}	{6,8,10}	{7,9}
μ_{small_*}	0.25	0	0	0
μ_{small^*}	1	1	0	0

Then we obtain the mass assignment for the linguistic value small, *and approximations of its focal sets given in Table 3 as follows.*

Table 3
Mass assignment for *small* and approximations of its focal sets

$\mathrm{rng}(\mu_{small})$	1	0.75	0.5	0.25
$small_\alpha$	$\{0,1\}$	$\{0,1,2\}$	$\{0,1,2,3\}$	$\{0,1,2,3,4\}$
$m_{small}(small_\alpha)$	0.25	0.25	0.25	0.25
$\underline{R}(small_\alpha)$	$\emptyset$	$\emptyset$	$\emptyset$	$\{0,2,4\}$
$\overline{R}(small_\alpha)$	$\{0,1,2,3,4,5\}$	$\{0,1,2,3,4,5\}$	$\{0,1,2,3,4,5\}$	$\{0,1,2,3,4,5\}$

Using Banerjee and Pal's notion, we obtain

$$\rho_R^{\alpha,\beta}(small) = \begin{cases} 1 & \text{for } \alpha > 0.25 \\ 0.5 & \text{for } 0.25 \geq \alpha > 0 \end{cases}$$

where the constraint $\alpha \geq \beta > 0$ is always assumed. On the other hand, the new measure of roughness yields

$$\hat{\rho}_R(small) = \sum_{\alpha \in \mathrm{rng}(\mu_{small})} m_{small}(small_\alpha)\left(1 - \frac{|\underline{R}(small_\alpha)|}{|\overline{R}(small_\alpha)|}\right) = 0.875$$

We now need the following [7].

Lemma 6 *For any $\alpha \in (0,1]$, we have*

$$\overline{R}(F)_\alpha = \overline{R}(F_\alpha) \text{ and } \underline{R}(F)_\alpha = \underline{R}(F_\alpha)$$

Let F^* and F_* be fuzzy sets induced from the rough fuzzy set $(\underline{R}(F), \overline{R}(F))$ as above. Denote

$$\mathrm{rng}(\mu_{F_*}) \cup \mathrm{rng}(\mu_{F^*}) = \{\omega_1, \ldots, \omega_p\}$$

such that $\omega_i > \omega_{i+1} > 0$ for $i = 1, \ldots, p-1$. Obviously, $\{\omega_1, \ldots, \omega_p\} \subseteq \mathrm{rng}(\mu_F)$, and $\omega_1 = \alpha_1$ and $\omega_p \geq \alpha_n$. Assume that $\omega_j = \alpha_{i_j}$, for $j = 1, \ldots, p$ and $i_1 = 1$. Then numbers in the set $\{\omega_1, \ldots, \omega_p\}$ makes a partition of $\{\alpha_1, \ldots, \alpha_n\}$, say

$$\{\{\alpha_1, \ldots, \alpha_{i_2-1}\}, \ldots, \{\alpha_{i_j}, \ldots, \alpha_{i_j-1}\}, \ldots, \{\alpha_{i_p}, \ldots, \alpha_n\}\}$$

where the enumeration preserves strict descreasing order. With these notations, we have

Lemma 7 *For any $1 \leq j \leq p$, if there exists $\alpha_i, \alpha_{i'} \in \mathrm{rng}(\mu_F)$ such that $\omega_{j+1} < \alpha_i < \alpha_{i'} \leq \omega_j$ then we have $F_i \approx_R F_{i'}$, and so $\rho_R(F_i) = \rho_R(F_{i'})$.*

PROOF. Indeed, since $\alpha_i < \alpha_{i'}$ we have $F_{i'} \subseteq F_i$ [2], so $\underline{R}(F_{i'}) \subseteq \underline{R}(F_i)$, and $\overline{R}(F_{i'}) \subseteq \overline{R}(F_i)$. The inverse inclusions are proved by the contrary as follows.

Assume that there exists $x \in U$ such that $[x]_R \subseteq \underline{R}(F_i)$ and $[x]_R \nsubseteq \underline{R}(F_{i'})$. By Lemma 6 we have $\underline{R}(F_{i'}) = \underline{R}(F)_{\alpha_{i'}}$. By definition, we get

$$\underline{R}(F_{i'}) = \underline{R}(F)_{\alpha_{i'}} = (F_*)_{\alpha_{i'}}$$

It follows from $[x]_R \nsubseteq \underline{R}(F_{i'})$ that $\min_{y\in[x]_R}\{\mu_F(y)\} < \alpha_{i'}$, which implies $\mu_{\underline{R}(F)}([x]_R) \leq \omega_{j+1}$. On the other hand, since $\alpha_i \notin \text{rng}(\mu_{F_*})$ and $\omega_{j+1} < \alpha_i < \omega_j$ we also have

$$\underline{R}(F_i) = \underline{R}(F)_{\alpha_i} = (F_*)_{\alpha_i} = (F_*)_{\omega_j}$$

which implies $\mu_{\underline{R}(F)}([x]_R) \geq \omega_j$, since $[x]_R \subseteq \underline{R}(F_i)$, a contradiction. An analogous argument is also applied for the remaining inclusion. Consequently, the lemma is proved completely.

Now we can represent the roughness $\hat{\rho}_R(F)$ in terms of level sets of fuzzy sets F_* and F^* in the following lemma.

Lemma 8 $\hat{\rho}_R(F) = \sum\limits_{j=1}^{p} (\omega_j - \omega_{j+1})(1 - \frac{|(F_*)_{\omega_j}|}{|(F^*)_{\omega_j}|})$, *where* $\omega_{p+1} = 0$, *by convention.*

PROOF. Indeed, we have

$$\begin{aligned}
\hat{\rho}_R(F) &= \sum_{i=1}^{n} m_F(F_i)(1 - \tfrac{|\underline{R}(F_i)|}{|\overline{R}(F_i)|}) \\
&= \sum_{i=1}^{n} m_F(F_i)\rho_R(F_i) \\
&= \sum_{j=1}^{p} \sum_{i=i_j}^{i_{j+1}-1} m_F(F_i)\rho_R(F_i) \\
&= \sum_{j=1}^{p} \sum_{i=i_j}^{i_{j+1}-1} m_F(F_i)\rho_R(F_{\omega_j}) \\
&= \sum_{j=1}^{p} (\omega_j - \omega_{j+1})\rho_R(F_{\omega_j}) \\
&= \sum_{j=1}^{p} (\omega_j - \omega_{j+1})(1 - \tfrac{|\underline{R}(F_{\omega_j})|}{|\overline{R}(F_{\omega_j})|}) \\
&= \sum_{j=1}^{p} (\omega_j - \omega_{j+1})(1 - \tfrac{|\underline{R}(F)_{\omega_j}|}{|\overline{R}(F)_{\omega_j}|}) \\
&= \sum_{j=1}^{p} (\omega_j - \omega_{j+1})(1 - \tfrac{|(F_*)_{\omega_j}|}{|(F^*)_{\omega_j}|})
\end{aligned}$$

[2] Note that F_i stands for F_{α_i}

As mentioned in [8], by Lemma 8 it is worth noting that $\hat{\rho}_R(F)$ can also be used in all three steps of modelling–learning, testing and applying a model–because it is properly defined with the knowledge of the fuzzy set F in the learning and testing stage, and without knowing F but its approximations in the application stage.

Lemmas 7 and 8 are illustrated by the following example.

Example 9 *Let us continue with the approximation space $\langle U, R\rangle$ and the fuzzy set* small *given in Example 5. We have*

$$\mathrm{rng}(\mu_{small}) = \{1, 0.75, 0.5, 0.25\}$$

By Table 2, we obtain

$$\mathrm{rng}(\mu_{small_*}) \cup \mathrm{rng}(\mu_{small^*}) = \{1, 0.25\}$$

which makes a partition of $\mathrm{rng}(\mu_{small})$ *as* $\{\{1, 0.75, 0.5\}, \{0.25\}\}$*. It is easily to see that Table 3 illustrates for Lemma 7, and by Lemma 8 and Table 2 we get*

$$\hat{\rho}_R(small) = (1 - 0.25)(1 - \frac{(small_*)_1}{(small^*)_1}) + 0.25(1 - \frac{(small_*)_{0.25}}{(small^*)_{0.25}}) = 0.875$$

which coincides with that given in Example 5.

More interestingly, we have the following proposition.

Proposition 10 *If fuzzy sets F and G in U are roughly equal in $\langle U, R\rangle$, then we have $\hat{\rho}_R(F) = \hat{\rho}_R(G)$.*

PROOF. Assume that the ranges of the membership functions μ_F and μ_G are

$$\mathrm{rng}(\mu_F) = \{\alpha_1, \ldots, \alpha_n\} \text{ and } \mathrm{rng}(\mu_G) = \{\beta_1, \ldots, \beta_k\}$$

respectively. Further, let us denote m_F and m_G the mass assignments of F and G, respectively. By definition, $F \approx_R G$ follows $\underline{R}(F) = \underline{R}(G)$ and $\overline{R}(F) = \overline{R}(G)$. This directly implies $\alpha_1 = \beta_1$. Furthermore, we have $F_* = G_*$ and $F^* = G^*$. Clearly also, the ranges of the membership functions μ_{F_*} and μ_{F^*} are subsets of the set $\mathrm{rng}(\mu_F)$. It follows by the fact $F \approx_R G$ that

$$\mathrm{rng}(\mu_{F_*}) = \mathrm{rng}(\mu_{G_*}) \subseteq \{\alpha_1, \ldots, \alpha_n\} \cap \{\beta_1, \ldots, \beta_k\}$$

$$\mathrm{rng}(\mu_{F^*}) = \mathrm{rng}(\mu_{G^*}) \subseteq \{\alpha_1, \ldots, \alpha_n\} \cap \{\beta_1, \ldots, \beta_k\}$$

These immediately imply

$$\mathrm{rng}(\mu_{F_*}) \cup \mathrm{rng}(\mu_{F^*}) = \mathrm{rng}(\mu_{G_*}) \cup \mathrm{rng}(\mu_{G^*})$$

Then Lemma 7 and Lemma 8 conclude the proof.

5 Rough Approximation Quality of a Fuzzy Classification

5.1 Extended Measures

Recall that roughness of a crisp set is defined as opposed to its accuracy. First, in the following we will see that it is possible to make a similar correspondency between the roughness and accuracy of a fuzzy set.

It should be noticed that if F is a subnormal fuzzy set, we have $m_F(\emptyset) > 0$, and then the empty set may be also considered as a possible definition of F. In this case, we should define the roughness measure of F as

$$\hat{\rho}_R(F) = \sum_{i=1}^{n} m_F(F_i)\rho_R(F_i) + m_F(\emptyset)\rho_R(\emptyset)$$

which trivially turns back to the normal case above as, by convention, $\rho_R(\emptyset) = 0$. However, we should take the case into account when once we want to consider the accuracy measure instead of roughness, with the convention that $\alpha_R(\emptyset) = 1$.

Under such an observation, it is eligible to define the accuracy measure for a fuzzy set F by

$$\hat{\alpha}_R(F) = \sum_{i=1}^{n} m_F(F_i)\alpha_R(F_i)$$

if F is a normal fuzzy set, or

$$\hat{\alpha}_R(F) = \sum_{i=1}^{n} m_F(F_i)\alpha_R(F_i) + m_F(\emptyset)\alpha_R(\emptyset)$$

if F is a subnormal fuzzy set. With this definition we have

$$\hat{\alpha}_R(F) = 1 - \hat{\rho}_R(F)$$

for any fuzzy set F in U.

In the rough set theory, the approximation quality γ is often used to evaluate the classification success of attributes in terms of a numerical evaluation of the dependency properties generated by these attributes.

Let us turn back to an approximation space $\langle U, R\rangle$. Assume now that there is another partition (or, classification) $\mathcal{C}$ of U, say $\mathcal{C} = \{Y_1, \dots, Y_n\}$. Note that, as mentioned above, R (or equivalently, U/R) and $\mathcal{C}$ may be induced respectively by sets of attributes applied to objects in U. In [18] Pawlak defines two measures to describe inexactness of approximation classifications as follows. The first one is called the *approximation accuracy* of $\mathcal{C}$ by R defined by

$$\alpha_R(\mathcal{C}) = \frac{\sum_{i=1}^{n} |\underline{R}(Y_i)|}{\sum_{i=1}^{n} |\overline{R}(Y_i)|} \tag{1}$$

which is easily represented in terms of accuracies of sets as follows

$$\alpha_R(\mathcal{C}) = \sum_{i=1}^{n} \frac{|\overline{R}(Y_i)|}{\sum_{i=1}^{n} |\overline{R}(Y_i)|} \alpha_R(Y_i) \tag{2}$$

That is, the approximation accuracy of a classification can be regarded as the convex sum of accuracies of its classes.

The second measure called the approximation quality of $\mathcal{C}$ by R is defined by

$$\gamma_R(\mathcal{C}) = \frac{\sum_{i=1}^{n} |\underline{R}(Y_i)|}{|U|} \tag{3}$$

which is also represented in terms of accuracy as follows

$$\gamma_R(\mathcal{C}) = \sum_{i=1}^{n} \frac{|\overline{R}(Y_i)|}{|U|} \alpha_R(Y_i) \tag{4}$$

In this case the measure $\gamma_R(\mathcal{C})$ can be regarded as the weighted mean of the accuracies of approximation of $\mathcal{C}$ by R [8].

Now let us consider a fuzzy classification $\mathcal{FC}$ of U instead of a crisp one $\mathcal{C}$, i.e., $\mathcal{FC} = \{Y_1, \ldots, Y_n\}$ is a fuzzy partition of U. This situation may come up in a natural way when a linguistic classification must be approximated in terms of already existing knowledge R. For example, assume that we have a personnel database given as $\mathcal{D} = \text{PERSONNEL}[ID; Name; Position; Salary]$, and Position attribute induces an approximation space $\langle \mathcal{D}, \text{ind}(Position) \rangle$. Given a linguistic description on the attribute Salary, say '*high*', it defines a fuzzy set on $\mathcal{D}$ denoted by μ_{high}. Then we can measure the accuracy of the fuzzy set μ_{high}, namely $\hat{\alpha}_{\text{ind}(Position)}(\mu_{high})$, which may express the degree of completeness of our knowledge about the statement "Salary is *high*", given the granularity of $\mathcal{D}/\text{ind}(Position)$. Further, a linguistic classification, say $\{low, medium, high\}$, may be imposed on the Salary attribute that induces a fuzzy partition of $\mathcal{D}$. Now one may want to measure a degree of dependency between 'knowledge on Salary attribute expressed linguistically' and 'knowledge on Position attribute'.

In such a situation, quite naturally with the spirit of our proposal in the preceding section, one may define the accuracy and quality of approximation of $\mathcal{FC}$ by R respectively as

$$\hat{\alpha}_R(\mathcal{FC}) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n_i} m_{Y_i}(Y_{i,j})|\underline{R}(Y_{i,j})|}{\sum_{i=1}^{n}\sum_{j=1}^{n_i} m_{Y_i}(Y_{i,j})|\overline{R}(Y_{i,j})|} \tag{5}$$

and

$$\hat{\gamma}_R(\mathcal{FC}) = \frac{1}{|U|}\sum_{i=1}^{n}\sum_{j=1}^{n_i} m_{Y_i}(Y_{i,j})|\underline{R}(Y_{i,j})| \tag{6}$$

where for $i = 1\ldots, n$, m_{Y_i} and $\{Y_{i,j}\}_{j=1}^{n_i}$ stand for the mass assignment of Y_i and the family of its focal elements, respectively.

By Lemma 7 and a simple transformation, we easily infer

$$\sum_{j=1}^{n_i} m_{Y_i}(Y_{i,j})|\underline{R}(Y_{i,j})| = |Y_{i*}|$$
$$\sum_{j=1}^{n_i} m_{Y_i}(Y_{i,j})|\overline{R}(Y_{i,j})| = |Y_i^*|$$

where the cardinality of a fuzzy set Y_i^* is defined as

$$|Y_i^*| = \sum_{x\in U} \mu_{Y_i^*}(x)$$

Therefore, we have

$$\hat{\alpha}_R(\mathcal{FC}) = \frac{\sum_{i=1}^{n} |Y_{i*}|}{\sum_{i=1}^{n} |Y_i^*|} \tag{7}$$

and

$$\hat{\gamma}_R(\mathcal{FC}) = \frac{1}{|U|}\sum_{i=1}^{n} |Y_{i*}| \tag{8}$$

Furthermore, similar as mentioned in [18], the measure of approximation quality (also called the *measure of rough dependency*) $\hat{\gamma}_R$ does not capture how this partial dependency is actually distributed among fuzzy classes of $\mathcal{FC}$. To capture this information we need also the so-called *precision measure* $\hat{\pi}_R(Y_i)$, for $i = 1, \ldots, k$, defined by

$$\hat{\pi}_R(Y_i) = \sum_{j=1}^{n_i} m_{Y_i}(Y_{i,j})\frac{|\underline{R}(Y_{i,j})|}{|Y_{i,j}|} \tag{9}$$

which may be considered as the expected relative number of elements in Y_i approximated by R. Clearly, we have $\hat{\pi}_R(Y_i) \geq \hat{\alpha}_R(Y_i)$, for any $i = 1, \ldots, k$. As such the two measures $\hat{\gamma}_R$ and $\hat{\pi}_R$ give us enough information about the "classification power" of the knowledge R with respect to the linguistic classification $\mathcal{FC}$.

Table 4
Relation in a Relational Database

ID	**Degree**	**Experience** (n)	**Salary**
1	Ph.D.	$6 < n \leq 8$	63K
2	Ph.D.	$0 < n \leq 2$	47K
3	M.S.	$6 < n \leq 8$	53K
4	B.S.	$0 < n \leq 2$	26K
5	B.S.	$2 < n \leq 4$	29K
6	Ph.D.	$0 < n \leq 2$	50K
7	B.S.	$2 < n \leq 4$	35K
8	M.S.	$2 < n \leq 4$	40K
9	M.S.	$2 < n \leq 4$	41K
10	M.S.	$8 < n \leq 10$	68K
11	M.S.	$6 < n \leq 8$	50K
12	B.S.	$0 < n \leq 2$	23K
13	M.S.	$6 < n \leq 8$	55K
14	M.S.	$6 < n \leq 8$	51K
15	Ph.D.	$6 < n \leq 8$	65K
16	M.S.	$8 < n \leq 10$	64K

On the other hand, the representations (2) and (4) may also suggest other possible extensions of α_R and γ_R for a fuzzy classification by

$$\hat{\alpha}_R(\mathcal{FC}) = \sum_{i=1}^{n} \frac{|\overline{R}(Y_i)|}{\sum_{i=1}^{n} |\overline{R}(Y_i)|} \hat{\alpha}_R(Y_i) = \sum_{i=1}^{n} \frac{|Y_i^*|}{\sum_{i=1}^{n} |Y_i^*|} \hat{\alpha}_R(Y_i) \tag{10}$$

and

$$\hat{\gamma}_R(\mathcal{FC}) = \sum_{i=1}^{n} \frac{|\overline{R}(Y_i)|}{|U|} \hat{\alpha}_R(Y_i) = \sum_{i=1}^{n} \frac{|Y_i^*|}{|U|} \hat{\alpha}_R(Y_i) \tag{11}$$

respectively. However, these extensions will not be considered in the rest of this paper.

In the following we consider a simple example to illustrate the proposed notions.

5.2 An Illustration Example

Let us consider a relation in a relational database as shown in Table 4 (this database is a variant of that found in [5]).

Then by the attributes **Degree** and **Experience** we obtain an approximation space

$$\langle U, \text{ind}(\{\textbf{Degree}, \textbf{Experience}\})\rangle$$

where $U = \{1, \ldots, 16\}$, and the corresponding partition as follows

$$U/\text{ind}(\{\textbf{Degree}, \textbf{Experience}\}) = \{\{1, 15\}, \{2, 6\}, \{3, 11, 13, 14\},\\ \{4, 12\}, \{5, 7\}, \{8, 9\}, \{10, 16\}\}$$

Further, consider now for example a linguistic classification

$$\{Low, Medium, High\}$$

defined on the domain of attribute **Salary**, say $[20K, 70K]$, with membership functions of linguistic classes graphically depicted as in Fig. 1. Then the linguistic classification induces a fuzzy partition on U whose membership functions of fuzzy classes are shown in Table 5.

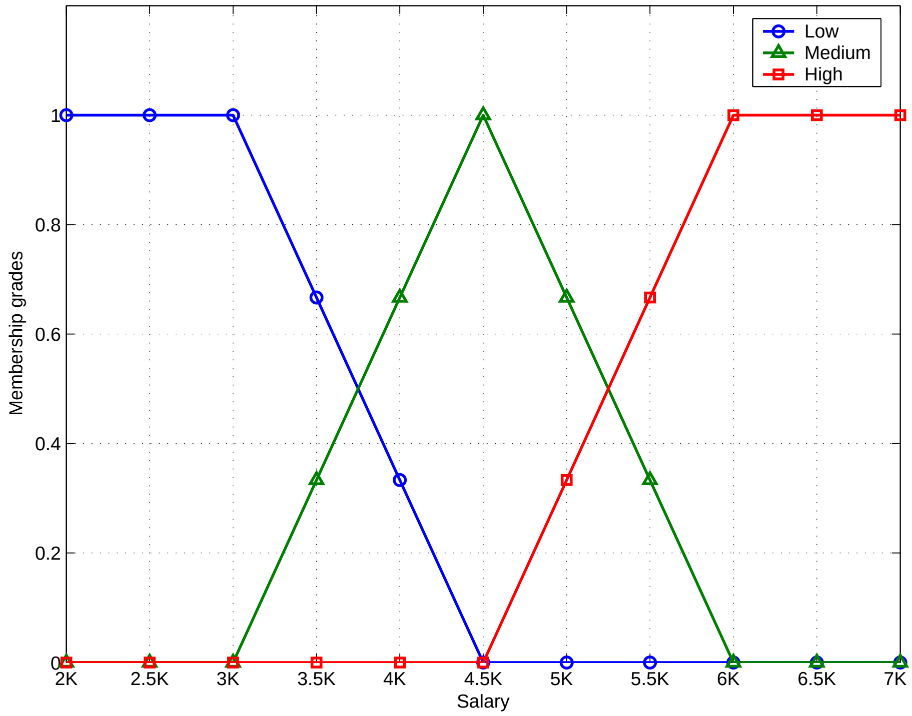

Fig. 1. A Linguistic Partition of **Salary** Attribute

Then approximations of the fuzzy partition induced by **Salary** in the approximation space defined by **Degree** and **Experience** are given in Table 6.

Table 5
Induced Fuzzy Partition of U Based on **Salary**

U	μ_{Low}	μ_{Medium}	μ_{High}
1	0	0	1
2	0	0.87	0.13
3	0	0.47	0.53
4	1	0	0
5	1	0	0
6	0	0.67	0.33
7	0.67	0.33	0
8	0.33	0.67	0
9	0.27	0.73	0
10	0	0	1
11	0	0.67	0.33
12	1	0	0
13	0	0.33	0.67
14	0	0.6	0.4
15	0	0	1
16	0	0	1

Table 6
The approximations of the fuzzy partition based on **Salary** in $\langle U, \mathrm{ind}(\{\mathbf{Degree}, \mathbf{Experience}\})\rangle$

X_i	$\{1,15\}$	$\{2,6\}$	$\{3,11,13,14\}$	$\{4,12\}$	$\{5,7\}$	$\{8,9\}$	$\{10,16\}$
μ_{High_*}	1	0.13	0.33	0	0	0	1
μ_{High^*}	1	0.33	0.67	0	0	0	1
μ_{Medium_*}	0	0.67	0.33	0	0	0.67	0
μ_{Medium^*}	0	0.87	0.67	0	0.33	0.73	0
μ_{Low_*}	0	0	0	1	0.67	0.27	0
μ_{Low^*}	0	0	0	1	1	0.33	0

Using (7) and (8) we easily obtain

$$\hat{\alpha}_{\{\mathbf{Degree},\mathbf{Experience}\}}(\mathbf{Salary}) = \frac{13.46}{18.21} = 0.739$$

and

$$\hat{\gamma}_{\{\mathbf{Degree},\mathbf{Experience}\}}(\mathbf{Salary}) = \frac{13.46}{16} = 0.84$$

respectively.

Table 7
Mass assignment for μ_{Low} and approximations of its focal sets

α	1	0.67	0.33	0.27
Low_α	$\{4,5,12\}$	$\{4,5,12,7\}$	$\{4,5,12,7,8\}$	$\{4,5,12,7,8,9\}$
$m_{Low}(Low_\alpha)$	0.33	0.34	0.06	0.27
$\underline{R}(Low_\alpha)$	$\{4,12\}$	$\{4,5,12,7\}$	$\{4,5,12,7\}$	$\{4,5,12,7,8,9\}$
$\overline{R}(Low_\alpha)$	$\{4,5,12,7\}$	$\{4,5,12,7\}$	$\{4,5,12,7,8,9\}$	$\{4,5,12,7,8,9\}$

That is we have the following partial dependency in the database

$$\{\textbf{Degree}, \textbf{Experience}\} \Rightarrow_{0.84} \textbf{Salary} \tag{12}$$

To calculate the precision measures for the fuzzy classes we need to obtain the mass assignment for each fuzzy class and approximations of its focal sets respectively. For example, the mass assignment of *Low* and approximations of its focal sets are shown in Table 7.

Therefore, by (9) we have

$$\hat{\pi}_{\{\textbf{Degree},\textbf{Experience}\}}(Low) = 0.878$$

Similarly, we also obtain

$$\begin{aligned} \hat{\pi}_{\{\textbf{Degree},\textbf{Experience}\}}(Medium) &= 0.646 \\ \hat{\pi}_{\{\textbf{Degree},\textbf{Experience}\}}(High) &= 0.876 \end{aligned}$$

Now in order to show how the influence of, for example, attribute **Experience** on the quality of approximation, let us consider the partition induced by the attribute **Degree** as follows

$$U/\text{ind}(\{\textbf{Degree}\}) = \{\{1,2,6,15\},\{3,8,9,10,11,13,14,16\},\{4,5,7,12\}\}$$

Then we obtain approximations of the fuzzy partition induced by **Salary** in the approximation space defined by **Degree** given in Table 8.

Thus we have

$$\hat{\gamma}_{\{\textbf{Degree}\}}(\textbf{Salary}) = \frac{3.2}{16} = 0.2$$

Similarly, we also easily obtain

$$\hat{\gamma}_{\{\textbf{Experience}\}}(\textbf{Salary}) = \frac{5.06}{16} = 0.316$$

Table 8
The approximations of the fuzzy partition based on **Salary** in $\langle U, \text{ind}(\{\textbf{Degree}\})\rangle$

X_i	$\{1, 2, 6, 15\}$	$\{3, 8, 9, 10, 11, 13, 14, 16\}$	$\{4, 5, 7, 12\}$
μ_{High_*}	0.13	0	0
μ_{High^*}	1	1	0
μ_{Medium_*}	0	0	0
μ_{Medium^*}	0.87	0.73	0.33
μ_{Low_*}	0	0	0.67
μ_{Low^*}	0	0.33	1

As we can see, both attributes **Degree** and **Experience** are highly significant as without each of them the measure of approximation quality changes considerably.

6 Conclusions

This paper has proposed a new roughness measure of a fuzzy set based on the notion of the mass assignment of a fuzzy set and its α-cuts. It is shown that the proposed measure inherits interesting properties of Pawlak's measure of roughness of a crisp set and, moreover, avoids the undesired properties of the measure proposed by Banerjee and Pal. As roughness measure is defined as opposed to its accuracy in rough set theory, it has been shown that there is also a similar correspondence between the roughness and accuracy of a fuzzy set. Furthermore, some possible extensions of the so-called approximation quality, which may be used to describe the rough approximation quality of a fuzzy classification have been suggested. It would be of interest that such measures may support us as numerical characteristics to realize partial dependency between attributes which is intuitively infered based on background knowledge and often expressed linguistically, while such a dependency could not be expressed in terms of traditional data dependencies. However, further work should be done to explore the theoretical features as well as practical implications of the introduced measures.

Acknowledgements

The authors would like to thank the referees for their constructive comments as well as helpful suggestions from the Editor-in-Chief, which have helped improve this paper significantly.

References

[1] J.F. Baldwin, The management of fuzzy and probabilistic uncertainties for knowledge based systems, in: S.A. Shapiro (Ed.), *The Encyclopaedia of AI*, Wiley, New York, 1992, pp. 528–537.

[2] J.F. Baldwin, J Lawry, T.P. Martin, A mass assignment theory of the probability of fuzzy events, *Fuzzy Sets and Systems* **83** (1996) 353–367.

[3] J.F. Baldwin, J Lawry, T.P. Martin, Mass assignment based induction of decision trees on words, *Proceedings of IPMU'98*, 1996, pp. 524–531.

[4] M. Banerjee, S.K. Pal, Roughness of a fuzzy set, *Information Sciences* **93** (1996) 235–246.

[5] S. M. Chen, C. M. Huang, Generating weighted fuzzy rules from relational database systems for estimating null values using genetic algorithms, *IEEE Transactions on Fuzzy Systems* **11** (2003) 495–506.

[6] D. Dubois, H. Prade, Properties of measures of information in evidence and possibility theories, *Fuzzy Sets and Systems* **24** (1987) 161–182.

[7] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, *International Journal of General Systems* **17** (1990) 191–209.

[8] G. Gediga, I. Düntsch, Rough approximation quality revisited, *Artificial Intelligence* **132** (2001) 219–234.

[9] I.R. Goodman, Fuzzy sets as equivalence classes of random sets, in R. Yager (Ed.), *Fuzzy Set and Possibility Theory* (Pergamon Press, Oxford, 1982) 327–342.

[10] V. N. Huynh, Y. Nakamori, T. B. Ho, G. Resconi, A context model for fuzzy concept analysis based upon modal logic, *Information Sciences* **160** (1–4) (2004) 111–129.

[11] R. Klir, B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice-Hall PTR, Upper Saddle River, NJ, 1995.

[12] J. Lawry, A methodology for computing with words, *International Journal of Approximate Reasoning* **28** (2001) 51–89.

[13] E. Marczewski, Steinhaus, On a certain distance of sets and the corresponding distance of functions, *Colloquium Mathematicum* **6** (1958) 319–327.

[14] T. Murai, H. Kanemitsu, M. Shimbo, Fuzzy sets and binary-proximity-based rough sets, *Information Sciences* **104** (1–2) (1998) 49–80.

[15] H.T. Nguyen, On random sets and belief functions, *Journal of Mathematical Analysis and Applications* **65** (1978) 531–542.

[16] S. K. Pal, A. Skowron (Eds.), *Rough Fuzzy Hybridization: New Trends in Decision Making*, Springer Verlag, Singapore, 1999.

[17] Z. Pawlak, Rough sets, *International Journal of Computer and Information Sciences* **11** (1982) 341–356.

[18] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Boston, MA, 1991.

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

[20] L. Polkowski, *Rough Sets: Mathematical Foundations*, Physica-Verlag, Heidelberg-New York, 2002.

[21] W. Z. Wu, W. X. Zhang, Constructive and axiomatic approaches of fuzzy approximation operators, *Information Sciences* **159** (3-4) (2004) 233–254.

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

[23] Y. Y. Yao, Information granulation and rough set approximation, *International Journal of Intelligent Systems* **16** (2001) 87–104.

[24] Y. Y. Yao, Combination of rough and fuzzy sets based on alpha-level sets, in T. Y. Lin & N. Cercone (Eds.), *Rough Sets and Data Mining: Analysis of Imprecise Data* (Kluwer Academic Publishers, Boston/London/Dordrecht, 1997) 301–321.

[25] Y. Y. Yao, A comparative study of fuzzy sets and rough sets, *Information Sciences* **109** (1998) 227–242.

[26] Y. Y. Yao, Probabilistic approaches to rough sets, *Expert Systems* **20** (5) (2003) 287–297.

[27] L.A. Zadeh, Fuzzy sets, *Information and Control* **8** (1965) 338–353.

[28] H. Zhang, H. Liang, D. Liu, Two new operators in rough set theory with applications to fuzzy sets, *Information Sciences*, in press.