Variable Precision Fuzzy Rough Set Model with Linguistic Labels

1st Alicja Mieszkowicz-Rolka
The Faculty of Mechanical Engineering and Aeronautics
Rzeszów University of Technology
Rzeszów, Poland
alicjamr@prz.edu.pl

2nd Leszek Rolka
The Faculty of Mechanical Engineering and Aeronautics
Rzeszów University of Technology
Rzeszów, Poland
leszekr@prz.edu.pl

***Abstract*—This paper presents an approach which combines unified variable precision fuzzy rough set model together with the concept of fuzzy linguistic labels. A real world application of the standard fuzzy rough sets can be problematic, especially in the case of large universes and noisy data. Due to relaxation of strict inclusion requirement in determining approximations of sets, a more tolerant variable precision fuzzy rough set model is better suited to be useful in analysis of this kind of data. Furthermore, a crucial issue at the initial stage of the fuzzy rough set approach consists in generating a fuzzy partition of a universe, with respect to condition and decision attributes. It requires comparing of elements by using a suitable fuzzy similarity relation. We simplify this process by applying the concept of fuzzy linguistic labels for determining the family of fuzzy similarity classes. This is done by performing a comparison of elements of the universe to a subset of representative elements which are described with the help of dominating linguistic values of attributes. The notions of the variable precision fuzzy rough set model, which is expressed in a unified parameterized form, can be used to determine the quality of the considered information system by evaluating its consistency, and to obtain a system of fuzzy decision rules.**

***Index Terms*—information systems, linguistic labels, fuzzy rough sets, fuzzy decision rules**

I. Introduction

Important areas of application of the rough set theory include modeling and analysis of control actions of human operators (experts) and diagnostics of industrial processes which is based on recorded signals.

In the face of a high popularity of the fuzzy set theory, a natural step was an attempt to compare the ideas of rough sets and fuzzy sets. Both paradigms focus on different aspects of uncertainty, and their combined use in the form of a hybrid approach has begun the next important stage in the development of these theories. Dubois and Prade [1], as well as Nakamura [2], were the first researchers who proposed generalization of the Pawlak's concepts of rough set approximation, in order to adapt them to approximation of fuzzy sets. The idea of fuzzy rough sets, in the sense given by Dubois and Prade, gained a major interest of specialists working on the borderline of the two theories, e.g., [3]–[8].

Unfortunately, the inconsistency of noisy data obtained from real decision processes can be a severe obstacle for a successful application of standard crisp or fuzzy rough sets models. Strict adherence to requirements of the rough set theory, with respect to set inclusion, can lead to losing substantial parts of the decision model represented by the analyzed information system. In other words, the obtained lower approximations of sets become too small because of rejection of large approximating indiscernibility classes, due to the lack of inclusion of even individual elements of such classes in the approximated set. On the other hand, the determined upper approximations of sets may be too large, because even single common elements of large approximating indiscernibility classes and the approximated set are sufficient to include such classes in the upper approximation. Therefore, extensions of the basic rough set concept have been proposed, by relaxing strong inclusion requirement and admitting of tolerance. A significant approach to overcome the problems appearing in practical applications of the Pawlak's rough set theory is the variable precision rough-set model (VPRS) introduced by Ziarko [9], [10]. The VPRS model allows certain level of misclassification when determining the approximation of sets. This is particularly important when generating large indiscernibility classes of elements of a universe. The idea of relaxation that was proposed by Ziarko for crisp rough sets has been then applied in the framework of fuzzy rough sets and studied by many authors, see, e.g., [11]–[21].

However, another significant issue in application of the standard fuzzy rough set approach is the complexity of generated families of fuzzy similarity classes, in the case of large decision tables with many values of fuzzy attributes. Furthermore, determination of similarity between elements of a universe is not unambiguous as in the case of crisp sets. The calculated degrees of similarity can be values in the interval $[0, 1]$. They also depend on the chosen fuzzy T-norm and implication operators. Having many free parameters, which is a characteristic of the fuzzy set theory, can lead to problems of interpretation of the obtained results, when determining the fuzzy rough approximations.

In order to avoid these drawbacks, we introduced [22] the labeled fuzzy rough set approach, which does not use a standard similarity relation for determining fuzzy similarity between all pairs of elements of a universe. Instead, we try to discover characteristic (representative) elements of the universe that share the same combinations of "active" linguistic values of attributes.

978-1-7281-6932-3/20/$31.00 ©2020 IEEE

In this paper, we go a step further by putting the idea of fuzzy linguistic labels into the variable precision fuzzy rough set model which is expressed in a unified parameterized form. The hybrid approach consists of two stages. Firstly, we determine fuzzy partitions of the universe with respect to condition and decision attributes, respectively. By assuming a required similarity level β, we can obtain the families of positive, boundary, and negative linguistic values for all elements of the universe. We find out which elements of the universe have the same combination of positive linguistic values of attributes. Such elements have the same linguistic label and constitute characteristic elements of a corresponding similarity class. In the second stage, we apply a tolerance level ε to determine positive area of classification and to generate a system of fuzzy decision rules.

We should start by giving the definition of a fuzzy decision system, together with the labeled fuzzy description of information systems. Next, we recall the unified variable precision fuzzy rough set model, and introduce fuzzy linguistic labels into the approximation of fuzzy sets. The approach will be illustrated by computational examples of analysis of fuzzy information systems.

II. FUZZY DECISION SYSTEMS

In a real-world application, a decision model of a human operator is usually expressed in the form of a decision table. For the sake of a formal analysis, we prefer to use a more general notion of information system. Hence, each row of a decision table corresponds to an element of a finite universe. The elements of the universe are characterized by condition and decision attributes. Since we want to use attributes with linguistic values in decision systems, we define a fuzzy decision system FDS which is expressed as a 4-tuple [22]

$$\text{FDS} = \langle U, A, \mathbb{V}, f \rangle, \tag{1}$$

where:

U – is a nonempty set, called the universe,

A – is a sum of two sets of fuzzy attributes: $A = C \bigcup D$,
C denotes condition attributes, D – decision attributes,

$\mathbb{V}$ – is a set of linguistic values of attributes, $\mathbb{V} = \bigcup_{a \in A} \mathbb{V}_a$,
$\mathbb{V}_a$ is the set of linguistic values of an attribute $a \in A$,

f – is an information function, $f : U \times \mathbb{V} \to [0, 1]$,
$f(x, V) \in [0, 1]$, for all $x \in U$, and $V \in \mathbb{V}$.

We connect every fuzzy attribute $a_i \in A$, where $i = 1, 2, \ldots, z$, with a family of its linguistic values: $\mathbb{A}_i = \{A_{i1}, A_{i2}, \ldots, A_{iz_i}\}$. Let us assume that the membership degrees of elements $x \in U$ in all linguistic values of fuzzy attributes will be assigned by an expert.

In a fuzzy decision system, every element x of the universe U is described by fuzzy condition attributes $C = \{c_1, c_2, \ldots, c_n\}$, and fuzzy decision attributes $D = \{d_1, d_2, \ldots, d_m\}$. We denote by $\mathbb{C}_i = \{C_{i1}, C_{i2}, \ldots, C_{in_i}\}$ the family of linguistic values of the i-th condition attribute c_i, and by $\mathbb{D}_j = \{D_{j1}, D_{j2}, \ldots, D_{jm_j}\}$ the family of linguistic values of the j-th decision attribute d_j, where $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots m$, respectively. Any element $x \in U$ has a membership degree in every linguistic value of all fuzzy attributes. Membership degrees are values in the interval $[0, 1]$.

In our considerations, we impose the following requirements on all linguistic values:

$$\exists A_{ik} \;\; (A_{ik} \in \mathbb{A}_i, \;\mu_{A_{ik}}(x) \geq 0.5)\,, \tag{2}$$

$$\text{power}\,(\mathbb{A}_i(x)) = \sum_{k=1}^{n_i} \mu_{A_{ik}}(x) = 1. \tag{3}$$

Our motivation for introducing the requirements (2), and (3) comes from generalization of the properties of crisp decision systems, in which every element $x \in U$ has a unique value of each attribute. In fuzzy decision systems, an element $x \in U$ can possess more than one value, but we assume that there is always a dominating linguistic value.

In analysis of a crisp decision model, the basic operation consists in comparing elements of the universe with the help of an indiscernibility relation which is reflexive, symmetric, and transitive. In the case of a fuzzy decision model, similarity between elements of the universe is evaluated by using a fuzzy similarity relation. However, the obtained degree of similarity, in the interval $[0, 1]$, depends on the form of selected fuzzy connectives. Therefore, there is no unique way of determining the fuzzy rough approximations. Moreover, for a large universe, one gets a vast number of fuzzy similarity classes generated with respect to condition, and decision attributes, respectively. As the interpretation of results becomes hardly in such a case, an alternative approach to analysis should be considered, which is based on a simpler method of classifying the elements of a fuzzy information system. In contrast to the standard method, a human operator neither uses a fuzzy similarity relation nor performs a detailed comparison of elements to each other. Instead, he or she rather tries to find out characteristic elements, that correspond to labels that are tuples of dominating (active) linguistic values of attributes. We can imagine that every label is represented by an ideal element (prototype) with membership degree equal to 1 in selected linguistic values of considered fuzzy attributes. The human operator assesses the similarity of elements $x \in U$ to a limited subset of such ideal elements, in order to discover characteristic elements that have a certain level of similarity to the labeled prototypes.

We will obtain the characteristic elements from the decision table, basing on the membership degree of particular elements $x \in U$ in the linguistic values of attributes. In order to distinguish the linguistic values, we apply a similarity threshold β which satisfies the inequality

$$0.5 < \beta \leq 1\,. \tag{4}$$

A selected value of the parameter β will be used as a similarity threshold for classifying particular linguistic values of attributes.

Given a fuzzy information system FDS, we define [23] for any element $x \in U$, and any fuzzy attribute $a \in A$:
the set $\widehat{\mathbb{V}}_a(x) \subseteq \mathbb{V}_a$ of positive linguistic values

$$\widehat{\mathbb{V}}_a(x) = \{V \in \mathbb{V}_a : f(x, V) \geq \beta\}, \tag{5}$$

the set $\overline{\overline{\mathbb{V}}}_a(x) \subseteq \mathbb{V}_a$ of boundary linguistic values

$$\overline{\overline{\mathbb{V}}}_a(x) = \{V \in \mathbb{V}_a : 0.5 \leq f(x, V) < \beta\}, \tag{6}$$

and the set $\check{\mathbb{V}}_a(x) \subseteq \mathbb{V}_a$ of negative linguistic values

$$\check{\mathbb{V}}_a(x) = \{V \in \mathbb{V}_a : 0 \leq f(x, V) < 0.5\}. \tag{7}$$

Due to the constraints (2) and (3), the sets $\widehat{\mathbb{V}}_a(x)$, $\overline{\overline{\mathbb{V}}}_a(x)$, and $\check{\mathbb{V}}_a(x)$ have the following properties [23]:

$$\text{(P1)} \qquad \operatorname{card}\left(\widehat{\mathbb{V}}_a(x)\right) \leq 1,$$
$$\text{(P2)} \qquad \operatorname{card}\left(\overline{\overline{\mathbb{V}}}_a(x)\right) \leq 2,$$
$$\text{(P3)} \qquad \operatorname{card}\left(\check{\mathbb{V}}_a(x)\right) < |\mathbb{V}_a|.$$

Now, every element $x \in U$ can be described with the help of a combination of those linguistic values that are positive for that particular element. In this way, we determine the linguistic labels for all elements of the universe. In the following, we consider linguistic labels generated with respect to a subset of fuzzy attributes $P \subseteq A$.

Formally, the set of linguistic labels $\widehat{\mathbb{L}}^P(x)$ is equal to the Cartesian product of the sets of positive linguistic values $\widehat{\mathbb{V}}_p(x)$, for all $p \in P$:

$$\widehat{\mathbb{L}}^P(x) = \prod_{p \in P} \widehat{\mathbb{V}}_p(x). \tag{8}$$

It should be noted, due to property (P1), that each element $x \in U$ can possess at most one linguistic label: $\operatorname{card}\left(\widehat{\mathbb{L}}^P(x)\right) \leq 1$. By inspecting the decision table, the family $\mathbb{L}^P$ of linguistic labels for the entire universe U will be generated. It is obvious that the number of linguistic labels in the family $\mathbb{L}^P$ can be only decreasing when the similarity threshold β is increased.
Furthermore, we denote by:

$L^P(x)$ — the linguistic label for an element $x \in U$,
$\mathbb{L}^P$ — the set of linguistic labels for the universe U,
L^P — a particular linguistic label from the set $\mathbb{L}^P$.

When inspecting the decision table, we can also discover elements $x \in U$ which have a common linguistic label $L^P(x)$.

By X_{L^P}, we denote the subset of the elements $x \in U$ that correspond to a linguistic label $L^P \in \mathbb{L}^P$, for selected fuzzy attributes $P \subseteq A$:

$$X_{L^P} = \left\{x \in U : L^P(x) = L^P\right\}. \tag{9}$$

The subset X_{L^P} is called the set of characteristic elements of the linguistic label L^P.

Observe that there may be elements $x \in U$ which do not have a positive linguistic value for some attributes, especially for a higher value of the similarity threshold β. It may be necessary to discard such "weak" elements, to assure that the obtained labeled similarity classes retain the property of covering sufficiently the universe U. By removing any elements $x \in U$ that do not have a linguistic label, we get a restricted universe $U' \subseteq U$. In other words, it should be required that every element of the universe is a characteristic element of some linguistic label.

Any linguistic label $L^P(x) \in \mathbb{L}^P$ can be given in the form of an ordered tuple of positive linguistic values, for $p \in P$:

$$L^P(x) = \left(\hat{V}_{p_1}, \hat{V}_{p_2}, \ldots, \hat{V}_{p_{|P|}}\right). \tag{10}$$

By aggregating the membership degrees of positive linguistic values for all attributes $p \in P$, the resulting membership degree of x in the linguistic label $L^P(x)$ can be determined:

$$\mu_{L^P(x)}(x) = \operatorname{aggr}\left(\mu_{\hat{V}_{p_1}}(x), \mu_{\hat{V}_{p_2}}(x), \ldots, \mu_{\hat{V}_{p_{|P|}}}(x)\right). \tag{11}$$

A suitable aggregation operator `aggr` can be based on a distance measure between the element $x \in U$ and an ideal element (prototype), which corresponds to the linguistic label $L^P(x)$. Depending on the metric selected for expressing the distance, we can apply different operators, such as `min` or `ave` (arithmetic mean).

By calculating the membership degree in a linguistic label $L^P(x)$ for all $x \in U$, we get a fuzzy similarity class $\widetilde{L^P}(x)$

$$\widetilde{L^P}(x) = \{\mu_{L^P(x)}(x_1)/x_1, \ \ldots, \mu_{L^P(x)}(x_N)/x_N\}. \tag{12}$$

In analysis of an information system given in the form of a decision table, two families of linguistic labels ought to be generated: $\mathbb{L}^C$ for the condition attributes C, and $\mathbb{L}^D$, for the decision attributes D, respectively.

Linguistic labels can be also used to distinguish selected paths in a flow graph which represents a fuzzy information system [23], [24].

Although, the full example will be presented and discussed in Section IV, we want to demonstrate in advance the use of selected concepts in calculations.

A. Illustrative Example

We consider a fuzzy decision system given in Table I. Positive linguistic values of all attributes, found for the similarity threshold β equal to 0.55, were bolded.

The element x_1 corresponds to a linguistic label $L^C(x_1) = (C_{12}, C_{23}, C_{33})$, obtained with respect to the condition attributes: c_1, c_2, and c_3. For x_5, and x_{10}, we get the same linguistic labels: $L^C(x_1) = L^C(x_5) = L^C(x_{10})$. So, the first linguistic label $L_1^C = (C_{12}, C_{23}, C_{33})$ is connected with the set $X_{L_1^C} = \{x_1, x_5, x_{10}\}$ of its characteristic elements.

Using the aggregation operator `min` in (11), the membership degree in L_1^C can be calculated for every $x \in U$, e.g.:

$$\begin{aligned}
\mu_{L_1^C}(x_1) &= \min(0.80,\ 0.75,\ 0.90) = 0.75,\\
\mu_{L_1^C}(x_2) &= \min(0.10,\ 0.35,\ 0.15) = 0.10,\\
\mu_{L_1^C}(x_5) &= \min(0.75,\ 0.80,\ 1.00) = 0.75,\\
\mu_{L_1^C}(x_{10}) &= \min(0.90,\ 0.75,\ 0.70) = 0.70.
\end{aligned}$$

Finally, we get the fuzzy similarity class $\widetilde{L_1^C}$

$$\widetilde{L_1^C} = \{\ 0.75/x_1,\ 0.10/x_2,\ 0.00/x_3,\ 0.00/x_4,\ 0.75/x_5,\ 0.00/x_6,\ 0.00/x_7,\ 0.00/x_8,\ 0.00/x_9,\ 0.70/x_{10}\}.$$

III. Parameterized Approximation of Crisp and Fuzzy Sets

The basic procedure of the rough set theory consists in finding classes of elements of a finite universe which cannot be discerned with respect to a subset of attributes. In the standard (crisp) rough set approach proposed by Pawlak [25], the attributes of elements can only have crisp values. By checking relationship between the classes of indiscernible elements, we can discover inconsistency in the decision table and determine redundant attributes.

Formally, the classes of indiscernible elements are produced by an indiscernibility (equivalence) relation $R \subseteq U \times U$. Any crisp subset of an universe U can be described with the help of the obtained indiscernibility classes.

Having granules of information in the form of indiscernibility classes, we are able to classify the elements of any subset A of the universe U. This is done by distinguishing between the full accordance and partial accordance of indiscernibility classes with the considered subset A. In consequence, we have two variants of approximation of the subset A.

Lower approximation $\underline{R}(A)$, and upper approximation $\overline{R}(A)$ of a set $A \in U$ are defined [25] as:

$$\underline{R}(A) = \{x \in U : [x]_R \subseteq A\}\,, \tag{13}$$

$$\overline{R}(A) = \{x \in U : [x]_R \cap A \neq \emptyset\}\,, \tag{14}$$

where $[x]_R$ denotes an indiscernibility class that contains the element $x \in U$.

In definitions (13), and (14) two different operations: set inclusion and intersection are applied. However, we proposed [26] alternative definitions of crisp rough approximations which turned out to be better suited for introducing a generalized unified form of approximations in the framework of fuzzy rough sets.

For an indiscernibility relation R, lower approximation $\underline{R}(A)$, and upper approximation $\overline{R}(A)$ of a crisp set A are defined [26] as follows

$$\underline{R}(A) = \{x \in U : \forall\, S \subseteq [x]_R \wedge S \neq \emptyset,\ \ S \subseteq A\}\,, \tag{15}$$

$$\overline{R}(A) = \{x \in U : \exists\, S \subseteq [x]_R \wedge S \neq \emptyset,\ \ S \subseteq A\}\,. \tag{16}$$

In a similar fashion, we can also base the definitions on the notion of membership in a set. Given an indiscernibility relation R, the lower approximation $\underline{R}(A)$ and upper approximation $\overline{R}(A)$ of a crisp set A are defined [26] as follows

$$\underline{R}(A) = \{x \in U : \forall\, y \in [x]_R\,,\ \ y \in A\}\,, \tag{17}$$

$$\overline{R}(A) = \{x \in U : \exists\, y \in [x]_R\,,\ \ y \in A\}\,. \tag{18}$$

We either require membership of *all* elements, or accept membership of even *a single* element, in the case of the lower (17), and the upper (18) approximations, respectively.

Hence, the formulae (15), and (16), as well (17), and (18), differ only in the used quantifier, which highlights two ideal cases of approximation, generated by the indiscernibility relation R.

The definitions of crisp set approximations in the form given above are important for developing a consistent fuzzy rough set model. This is because there are usually several ways of performing an operation on fuzzy sets. Therefore, applying a single fuzzy connective instead of two different is crucial for creating a uniform variable precision fuzzy rough set model.

Discussion of the uniform parametric approximation of sets should be started with analysis of the case of crisp sets and relations. To this aim, we recall the idea of the variable precision rough set model (VPRS) proposed by Ziarko [10]. We express it [27], basing on the notion of inclusion degree $\mathrm{INCL}(A, B)$ of a nonempty crisp set A in a crisp set B

$$\mathrm{INCL}(A, B) = \frac{\mathrm{card}(A \cap B)}{\mathrm{card}(A)}\,. \tag{19}$$

In the following, we refer to an extended VPRS model [9] with asymmetric bounds l and u for an admissible inclusion error. The lower limit l and the upper limit u satisfy the following requirement:

$$0 \leq l < u \leq 1\,. \tag{20}$$

By applying the limits l and u limits, we express the u-lower and l-upper approximation of any subset A of the universe U by means of the indiscernibility relation R.

The u-lower approximation of A by R is a set defined as follows

$$\underline{R}_u(A) = \{x \in U\colon\ \ \mathrm{INCL}([x]_R, A) \geq u\}\,, \tag{21}$$

and the l-upper approximation of A by R is a set

$$\overline{R}_l(A) = \{x \in U\colon\ \ \mathrm{INCL}([x]_R, A) > l\}\,. \tag{22}$$

Definitions (21) and (22) are expressed using the same notion of inclusion degree. They correspond to definitions (15), and (16), which can be perceived as special cases.

The uniform parametric variable precision crisp and fuzzy rough set models should be based only on the degree of set inclusion as the fundamental notion. In the generalized approach, the notion of rough inclusion function, introduced in [28], proved to be useful. It is defined on the Cartesian product of the powersets $\mathbb{P}(U)$ of the universe U

$$\nu : \mathbb{P}(U) \times \mathbb{P}(U) \to [0, 1]\,. \tag{23}$$

It is assumed that the first parameter represents a nonempty set. Furthermore, the rough inclusion function should be monotonic with respect to the second parameter

$$\nu(X, Y) \leq \nu(X, Z) \quad \text{for any } Y \subseteq Z, \quad \text{where } X, Y, Z \subseteq U.$$

The inclusion degree (19) satisfies the requirements to be a rough inclusion function (23).

Now, we define lower and upper approximations of a crisp set A, by applying the rough inclusion function ν, as follows

$$\underline{R}(A) = \{x \in U : \nu([x]_R, A) = 1\}\,, \tag{24}$$

$$\overline{R}(A) = \{x \in U : \nu([x]_R, A) > 0\}\,. \tag{25}$$

Finally, we abandon the standard interpretation of the classic rough set model by introducing a parameterized unified form of approximation of crisp sets [27].

Given an indiscernibility relation R, the ε-approximation $R_\varepsilon(A)$ of a crisp set A is defined as follows

$$R_\varepsilon(A) = \{x \in U : \nu([x]_R, A) \geq \varepsilon\} , \qquad (26)$$

where $\varepsilon \in (0, 1]$.

The ε-approximation R_ε can be used for expressing every approximation, according to the following properties:

$$\begin{array}{llll} \text{(W1)} & R_\varepsilon(A) = \underline{R}(A) & \text{for } \varepsilon = 1 , \\ \text{(W2)} & R_\varepsilon(A) = \overline{R}(A) & \text{for } \varepsilon = 0+ , \\ \text{(W3)} & R_\varepsilon(A) = \underline{R}_u(A) & \text{for } \varepsilon = u , \\ \text{(W4)} & R_\varepsilon(A) = \overline{R}_l(A) & \text{for } \varepsilon = l+ . \end{array}$$

In a practical applications, several values of the parameter ε can be selected in repeated calculation of ε-approximation. Let us consider a series of n ε-approximations of a crisp set A. Owing to monotonicity of the inclusion function, the following property is satisfied: $R_{\varepsilon_1}(A) \subseteq R_{\varepsilon_2}(A) \subseteq \ldots \subseteq R_{\varepsilon_n}(A)$ for $\varepsilon_1 \geq \varepsilon_2 \geq \ldots \geq \varepsilon_n$.

With the aim of creating a unified approach to parameterized approximation of fuzzy sets, we need to generalize the single notion of crisp ε-approximation. In this way, we can obtain a variable precision fuzzy rough set model in a consistent form.

First, we recall the definition of fuzzy rough set, proposed by Dubois and Prade [1]. For a given fuzzy set A and a fuzzy partition $\Phi = \{F_1, F_2, \ldots, F_n\}$ on the universe U, the membership functions of the lower and upper approximations of A by Φ are defined as follows [5]

$$\mu_{\underline{\Phi}(A)}(F_i) = \inf_{x \in U} \mathrm{I}(\mu_{F_i}(x), \mu_A(x)) , \qquad (27)$$

$$\mu_{\overline{\Phi}(A)}(F_i) = \sup_{x \in U} \mathrm{T}(\mu_{F_i}(x), \mu_A(x)) , \qquad (28)$$

where T and I denote a T-norm operator and an implicator, respectively. The pair $(\underline{\Phi}F, \overline{\Phi}F)$ is called a fuzzy rough set.

Fundamental issue in the conception of variable precision fuzzy rough set model is the way of determining the degree of inclusion of one fuzzy set in another fuzzy set. There are different measures of fuzzy set inclusion (see, e.g., [11], [29]). The crucial point of our approach consists in describing inclusion of sets with the help of a fuzzy set rather than a single number. To this end, we introduced [30] a notion of a fuzzy inclusion set, denoted here by $\mathrm{INCL}(A, B)$, which expresses the inclusion of a fuzzy set A in a fuzzy set B. The set $\mathrm{INCL}(A, B)$ is determined with respect to elements of the set A. Such a notion of fuzzy inclusion allows to express precisely how one fuzzy set is included in another fuzzy set.

Implication-based inclusion set $\mathrm{INCL}(A, B)$ of a nonempty fuzzy set A in a fuzzy set B is defined as follows

$$\mu_{\mathrm{INCL}(A,B)}(x) = \begin{cases} \mathrm{I}(\mu_A(x), \mu_B(x)) & \text{if } \mu_A(x) > 0 , \\ 0 & \text{otherwise.} \end{cases} \qquad (29)$$

Next, we need a counterpart of the function (23) in the form of a fuzzy rough inclusion function defined on the Cartesian product of the families of all fuzzy subsets of $\mathbb{F}(U)$ in the domain of the universe U.

$$\nu_\alpha : \mathbb{F}(U) \times \mathbb{F}(U) \to [0, 1] . \qquad (30)$$

The value $\nu_\alpha(A, B)$ should express how many elements of the nonempty fuzzy set A belong, at least to the degree α, to the fuzzy set B.

By applying the notion of α-cut of a fuzzy set, and generalizing the measure of inclusion degree (19), we proposed [27] the following form of the fuzzy rough α-inclusion function $\nu_\alpha(A, B)$ of any nonempty fuzzy set A in a fuzzy set B, which is defined for any $\alpha \in (0, 1]$

$$\nu_\alpha(A, B) = \frac{\mathrm{power}(A \cap \mathrm{INCL}(A, B)_\alpha)}{\mathrm{power}(A)} , \qquad (31)$$

It can be proven [27] that the implication-based fuzzy rough inclusion function ν_α is monotonic with respect to the second parameter, for any $\alpha \in (0, 1]$

$$\begin{array}{l} \nu_\alpha(X, Y) \leq \nu_\alpha(X, Z) \quad \text{for any } Y \subseteq Z, \\ \text{where } X, Y, Z \subseteq \mathbb{F}(U) . \end{array}$$

It is easy to check that the rough inclusion degree used in formulae (21) and (22) is a special case of the fuzzy rough inclusion function (31). For any nonempty crisp set A, any crisp set B, and for $\alpha \in (0, 1]$, the implication-based inclusion function $\nu_\alpha(A, B)$ is equal to the inclusion degree $\mathrm{INCL}(A, B)$.

Relaxation of strong inclusion requirement in the case of fuzzy rough approximations is a generalization of the idea given by Ziarko for crisp rough sets. It consists in ignoring the influence of selected elements of the approximating fuzzy similarity classes, depending on the value of fuzzy rough inclusion function. In consequence, membership degree of the approximating fuzzy similarity class in the approximated fuzzy set will be determined taking into account only part of the elements. Hence, we get a resulting membership degree [27] that can be expressed formally with the help of a function called `res`. It is defined on the Cartesian product $\mathbb{P}(U) \times \mathbb{F}(U)$, where $\mathbb{P}(U)$ denotes the powerset of the universe U, and $\mathbb{F}(U)$ the family of all fuzzy subsets of the universe U, respectively

$$\mathrm{res} : \mathbb{P}(U) \times \mathbb{F}(U) \to [0, 1] . \qquad (32)$$

The function `res` must satisfy the following requirements:

$$\begin{array}{l} \mathrm{res}(\emptyset, Y) = 0 , \\ \mathrm{res}(X, Y) \in \{0, 1\}, \quad \text{if } Y \text{ is a crisp set} , \\ \mathrm{res}(X, Y) \leq \mathrm{res}(X, Z) \quad \text{for any } Y \subseteq Z , \\ \text{where } X \in \mathbb{P}(U), \text{ and } Y, Z \in \mathbb{F}(U) . \end{array}$$

Value of the function $\mathrm{res}(X, Y)$ is calculated for a crisp set X and a fuzzy set Y. It expresses the resulting membership degree in the set Y, taking into account not all elements of the universe U, but only the elements of the set X.

If we assume the limit-based approach of Dubois and Prade, then the function `res` should have the following form:

$$\operatorname{res}(X,Y) = \inf_{x \in X} \mu_Y(x)\,. \tag{33}$$

However, referring only to a single value of membership degree is not always acceptable, e.g., in the case of large information systems. In general, we should admit of different variants of the function `res`, in which many values of membership degree are used in calculation. We follow this assumption in definition of ε-fuzzy rough approximation of a fuzzy set [27].

For $\varepsilon \in (0,1]$, the ε-approximation $\Phi_\varepsilon(A)$ of a fuzzy set A, by a fuzzy partition $\Phi = \{F_1, F_2, \ldots, F_n\}$, is a fuzzy set on the domain Φ with membership function expressed as

$$\mu_{\Phi_\varepsilon(A)}(F_i) = \operatorname{res}(S_\varepsilon(F_i,A), \mathrm{INCL}(F_i,A))\,, \tag{34}$$

where

$$S_\varepsilon(F_i,A) = \operatorname{support}(F_i \cap \mathrm{INCL}(F_i,A)_{\alpha_\varepsilon})\,,$$

$$\alpha_\varepsilon = \sup\{\alpha \in [0,1]:\ \nu_\alpha(F_i,A) \geq \varepsilon\}\,.$$

Membership degree $\mu_{\Phi_\varepsilon(A)}(F_i)$ is calculated by taking into account selected elements of the approximating class F_i which are included in A at least to the degree α_ε. Selection of elements is done with the help of the crisp set $S_\varepsilon(F_i,A)$, being support of the intersection of F_i with the α-cut of $\mathrm{INCL}(F_i,A)$ (*for any fuzzy set* F, $\operatorname{support}(F)$ *is a crisp subset of the universe:* $\operatorname{support}(F) = \{x : \mu_F(x_i) > 0\}$). The resulting membership $\mu_{\Phi_\varepsilon(A)}(F_i)$ is obtained using only the elements from $S_\varepsilon(F_i,A)$ instead of the entire class F_i.

If we assume the limit-based definition (33) of the function `res`, then the the ε-approximation (34) has the following form:

$$\mu_{\Phi_\varepsilon(A)}(F_i) = \sup\{\alpha \in [0,1]:\ \nu_\alpha(F_i,A) \geq \varepsilon\}\,. \tag{35}$$

Due to a single definition of fuzzy rough approximation, we can avoid problem in interpretation of results that can be caused by applying different fuzzy connectives as in the standard fuzzy rough set approach. Moreover, calculation of variable precision fuzzy rough ε-approximation can be simplified, when fuzzy partition $\Phi = \{F_1, F_2, \ldots, F_n\}$ is obtained as a family of similarity classes corresponding to linguistic labels discussed in Section II.

IV. EXAMPLE

A. *Case I*

To illustrate the details of the hybrid labeled variable precision fuzzy rough set approach, we present analysis of a small fuzzy information system given in Table I. The universe U contains ten elements, which are characterized by three fuzzy condition attributes c_1, c_2, c_3, and one decision attribute d_1. All the fuzzy attributes have three linguistic values.

We assume the required similarity level β to be equal to 0.55. Positive linguistic values of the condition and decision attributes are marked in Table I.

Four linguistic labels L_1^C, L_2^C, L_3^C, and L_4^C with respect to the condition attributes C can be found. The corresponding fuzzy similarity classes are determined by aggregating membership degrees of positive linguistic values for all attributes C. We apply `min` as the operator `aggr` in (11).

$L_1^C = (C_{12}, C_{23}, C_{33})$,

$\widetilde{L_1^C} = \{\ 0.75/x_1,\ 0.10/x_2,\ 0.00/x_3,\ 0.00/x_4,\ 0.75/x_5,\ 0.00/x_6,\ 0.00/x_7,\ 0.00/x_8,\ 0.00/x_9,\ 0.70/x_{10}\}$.

$L_2^C = (C_{13}, C_{22}, C_{32})$,

$\widetilde{L_2^C} = \{\ 0.00/x_1,\ 0.65/x_2,\ 0.00/x_3,\ 0.00/x_4,\ 0.00/x_5,\ 0.80/x_6,\ 0.00/x_7,\ 0.00/x_8,\ 0.00/x_9,\ 0.10/x_{10}\}$.

$L_3^C = (C_{11}, C_{21}, C_{31})$,

$\widetilde{L_3^C} = \{\ 0.00/x_1,\ 0.00/x_2,\ 0.75/x_3,\ 0.65/x_4,\ 0.00/x_5,\ 0.00/x_6,\ 0.90/x_7,\ 0.75/x_8,\ 0.00/x_9,\ 0.00/x_{10}\}$.

$L_4^C = (C_{13}, C_{21}, C_{33})$,

$\widetilde{L_4^C} = \{\ 0.00/x_1,\ 0.00/x_2,\ 0.00/x_3,\ 0.00/x_4,\ 0.00/x_5,\ 0.00/x_6,\ 0.00/x_7,\ 0.00/x_8,\ 0.85/x_9,\ 0.00/x_{10}\}$.

For the decision attribute d_1, three linguistic labels L_1^D, L_2^D, and L_3^D are obtained.

$L_1^D = (D_{11})$,

$\widetilde{L_1^D} = \{\ 0.00/x_1,\ 0.00/x_2,\ 0.90/x_3,\ 0.40/x_4,\ 0.00/x_5,\ 0.00/x_6,\ 1.00/x_7,\ 0.85/x_8,\ 0.00/x_9,\ 0.00/x_{10}\}$.

$L_2^D = (D_{12})$,

$\widetilde{L_2^D} = \{\ 0.15/x_1,\ 0.90/x_2,\ 0.10/x_3,\ 0.60/x_4,\ 0.10/x_5,\ 0.95/x_6,\ 0.00/x_7,\ 0.15/x_8,\ 0.00/x_9,\ 0.45/x_{10}\}$.

$L_3^D = (D_{13})$,

$\widetilde{L_3^D} = \{\ 0.85/x_1,\ 0.10/x_2,\ 0.00/x_3,\ 0.00/x_4,\ 0.90/x_5,\ 0.05/x_6,\ 0.00/x_7,\ 0.00/x_8,\ 1.00/x_9,\ 0.55/x_{10}\}$.

In the next stage, we determine ε-approximations of the fuzzy similarity classes $\widetilde{L_1^D}$, $\widetilde{L_2^D}$, and $\widetilde{L_3^D}$ by the family of fuzzy similarity classes obtained with respect to the condition attributes C. We apply the R-implicator of Łukasiewicz: $I(x,y) = \min(1, 1-x+y)$, and assume the limit-based form (33) of the function `res`.

In order to get a deeper insight into the process of determining ε-approximation, let us inspect the inclusion degree of single similarity classes $\widetilde{L_C}$ in particular similarity classes $\widetilde{L^D}$ for five different values of the parameter ε from the set $\{1.0,\ 0.9,\ 0.8,\ 0.7,\ 0.65\}$.

We take a closer look at calculating the inclusion of the similarity class $\widetilde{L_1^C}$ in the similarity class $\widetilde{L_3^D}$. According to (29), we get the implication-based inclusion set $\mathrm{INCL}(\widetilde{L_1^C}, \widetilde{L_3^D}) = \{1.00/x_1,\ 1.00/x_2,\ 0.00/x_3,\ 0.00/x_4,\ 1.00/x_5, 0.00/x_6,\ 0.00/x_7,\ 0.00/x_8,\ 0.00/x_9,\ 0.85/x_{10}\}$.

For ε equal to 1.0, there is no relaxation of strong inclusion requirement. By seeking for the biggest $\alpha \in [0,1]$,

TABLE I
DECISION TABLE.

	c_1			c_2			c_3			d_1		
	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}	D_{11}	D_{12}	D_{13}
x_1	0.20	**0.8**	0.00	0.00	0.25	**0.75**	0.00	0.10	**0.90**	0.00	0.15	**0.85**
x_2	0.00	0.1	**0.90**	0.00	**0.65**	0.35	0.00	**0.85**	0.15	0.00	**0.90**	0.10
x_3	**0.75**	0.25	0.00	**0.75**	0.25	0.00	**0.80**	0.20	0.00	**0.90**	0.10	0.00
x_4	**0.65**	0.35	0.00	**0.65**	0.35	0.00	**0.65**	0.35	0.00	0.40	**0.60**	0.00
x_5	0.25	**0.75**	0.00	0.00	0.20	**0.80**	0.00	0.00	**1.00**	0.00	0.10	**0.90**
x_6	0.00	0.00	**1.00**	0.00	**0.80**	0.20	0.00	**0.90**	0.10	0.00	**0.95**	0.05
x_7	**1.00**	0.00	0.00	**0.90**	0.10	0.00	**0.90**	0.10	0.00	**1.00**	0.00	0.00
x_8	**0.90**	0.10	0.00	**0.90**	0.10	0.00	**0.75**	0.25	0.00	**0.85**	0.15	0.00
x_9	0.00	0.10	**0.90**	**1.00**	0.00	0.00	0.00	0.15	**0.85**	0.00	0.00	**1.00**
x_{10}	0.00	**0.90**	0.10	0.00	0.25	**0.75**	0.00	0.30	**0.70**	0.00	0.45	**0.55**

for which $\nu_\alpha(\widetilde{L_1^C}, \widetilde{L_3^D}) \geq 1.0$, we obtain $\alpha_\varepsilon = 0.85$, and $\mathrm{INCL}(\widetilde{L_1^C}, \widetilde{L_3^D})_{0.85} = \{x_1, x_2, x_5, x_{10}\}$. When we use the function `res` defined by (33), the membership degree of $\widetilde{L_1^C}$ in $\widetilde{L_3^D}$ is obtained as $\mu_{L_3^D}(L_1^C) = 0.85$.

If ε equal to 0.65 is used, then we get $\alpha_\varepsilon = 1.00$, and $\mathrm{INCL}(\widetilde{L_1^C}, \widetilde{L_3^D})_{1.00} = \{x_1, x_2, x_5\}$.

The fuzzy rough α-inclusion function (31) can be determined as $\nu_{1.00}(\widetilde{L_1^C}, \widetilde{L_3^D}) = \frac{\mathrm{power}(\{0.75/x_1, 0.10/x_2, 0.75/x_5\})}{\mathrm{power}(\{0.75/x_1, 0.10/x_2, 0.75/x_5, 0.70/x_{10}\})} = \frac{1.60}{2.30} = 0.69$.

Summarizing, the similarity class $\widetilde{L_1^C}$ is included in the class $\widetilde{L_3^D}$ with the membership degree equal to 0.85, when ε is equal to 1.0, 0.9, 0.8, and 0.7. For ε equal to 0.65, the membership degree increases to 1, due to the relaxation of inclusion requirement.

We get the membership degree equal to 1 for all values of the ε parameter, when we check the inclusion of $\widetilde{L_4^C}$ in $\widetilde{L_3^D}$, and $\widetilde{L_2^C}$ in $\widetilde{L_2^D}$.

The similarity class $\widetilde{L_3^C}$ is included in the class $\widetilde{L_1^D}$ with the membership degree equal to 0.75, for ε equal to 1.0, 0.9, and 0.8. The membership degree increases to 1.0 for ε equal to 0.7, and 0.65.

As we see, due to a tolerant fuzzy rough ε-approximation, additional two certain fuzzy decision rules can be obtained: $\widetilde{L_1^C} \to \widetilde{L_3^D}$, and $\widetilde{L_3^C} \to \widetilde{L_1^D}$.

Superiority of our variable precision fuzzy rough model over the standard approach would become even more visible in the case of large information systems. Furthermore, with the help of fuzzy linguistic labels, relative small fuzzy partitions of the universe can be generated. If we applied a standard fuzzy similarity relation, we would obtain 10 fuzzy similarity classes with respect to condition, and 10 fuzzy similarity classes with respect to decision attributes. Both determination of approximations and interpretation of results would become problematic in such a case, as we have to determine 100 fuzzy inclusion sets. In contrast to this, there are only 12 fuzzy inclusion sets to be determined, when the labeled fuzzy rough set approach is applied.

B. Case II

In order to make the influence of relaxation of inclusion more evident, we also provide a more complex example of approximation in the case of larger similarity classes.

Assume that the similarity class $\widetilde{L^C}$ was generated with respect to condition attributes:

$$\begin{aligned}\widetilde{L^C} = \{\ & 0.7/x_1, && 0.8/x_2, && 0.1/x_3, && 0.0/x_4, && 0.0/x_5,\\ & 0.8/x_6, && 0.0/x_7, && 0.2/x_8, && 0.4/x_9, && 0.75/x_{10},\\ & 0.7/x_{11}, && 0.9/x_{12}, && 0.8/x_{13}, && 0.3/x_{14}, && 0.0/x_{15},\\ & 0.3/x_{16}, && 0.9/x_{17}, && 0.0/x_{18}, && 0.2/x_{19}, && 0.9/x_{20}\},\end{aligned}$$

and the similarity class $\widetilde{L^D}$ with respect to decision attributes:

$$\begin{aligned}\widetilde{L^D} = \{\ & 0.8/x_1, && 0.9/x_2, && 0.2/x_3, && 0.0/x_4, && 1.0/x_5,\\ & 0.9/x_6, && 0.1/x_7, && 0.1/x_8, && 0.6/x_9, && 0.9/x_{10},\\ & 0.9/x_{11}, && 0.1/x_{12}, && 0.9/x_{13}, && 0.1/x_{14}, && 0.1/x_{15},\\ & 0.5/x_{16}, && 1.0/x_{17}, && 0.0/x_{18}, && 0.2/x_{19}, && 1.0/x_{20}\}.\end{aligned}$$

We get the fuzzy inclusion set $\mathrm{INCL}(\widetilde{L^C}, \widetilde{L^D})$:

$$\begin{aligned}&\mathrm{INCL}(\widetilde{L^C}, \widetilde{L^D}) =\\ &\quad\{\ 1.0/x_1, && 1.0/x_2, && 1.0/x_3, && 0.0/x_4, && 0.0/x_5,\\ &\quad\ \ 1.0/x_6, && 0.0/x_7, && 0.9/x_8, && 1.0/x_9, && 1.0/x_{10},\\ &\quad\ \ 1.0/x_{11}, && 0.2/x_{12}, && 1.0/x_{13}, && 0.8/x_{14}, && 0.0/x_{15},\\ &\quad\ \ 1.0/x_{16}, && 1.0/x_{17}, && 0.0/x_{18}, && 1.0/x_{19}, && 1.0/x_{20}\}.\end{aligned}$$

By repeating the same steps as in Example IV-A, we obtain the following results:

$$\begin{aligned}\mu_{L^D}(L^C) &= 0.2 \quad \text{for} \quad \varepsilon = 1.00,\\ \mu_{L^D}(L^C) &= 0.8 \quad \text{for} \quad \varepsilon = 0.89,\\ \mu_{L^D}(L^C) &= 0.9 \quad \text{for} \quad \varepsilon = 0.86,\\ \mu_{L^D}(L^C) &= 1.0 \quad \text{for} \quad \varepsilon = 0.84.\end{aligned}$$

Observe that the decision rule $\widetilde{L^C} \to \widetilde{L^D}$ cannot be accepted, if we require full inclusion by setting ε to 1.00. A slight relaxation of inclusion requirement, by decreasing ε to 0.84, makes $\widetilde{L^C} \to \widetilde{L^D}$ a certain decision rule. This can be easy explained, because even a few "bad" elements of the approximating similarity class can deteriorate the resulting inclusion degree. For a larger universe, this phenomenon has a more negative impact on results. That is why the presented variable precision rough set approach is so important.

V. Conclusions

The concept of fuzzy linguistic labels helps to reduce computational complexity of the standard fuzzy rough set algorithms. By applying a similarity threshold β, families of fuzzy similarity classes can be obtained in a straightforward manner. Every similarity class is connected with a subset of its characteristic elements, which have the same fuzzy linguistic label. A hybrid approach that combines the idea of fuzzy linguistic labels with the concept of unified variable precision approximation of sets turns out to be an effective tool in practical applications, especially in the case of large fuzzy information systems. Its most important advantage over an ordinary noise tolerant fuzzy rough set model is the ability to easily discover a system of fuzzy decision rules. Moreover, it is possible to get more insight into the decision process, by taking into account its most significant features only.

References

[1] D. Dubois and H. Prade, "Putting rough sets and fuzzy sets together," in *Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory*, R. Słowiński, Ed. Boston Dordrecht London: Kluwer Academic Publishers, 1992, pp. 203–232.

[2] A. Nakamura, "Application of fuzzy-rough classifications to logics," in *Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory*, R. Słowiński, Ed. Boston Dordrecht London: Kluwer Academic Publishers, 1992, pp. 233–250.

[3] S. Greco, B. Matarazzo, and R. Słowiński, "Rough set processing of vague information using fuzzy similarity relations," in *Finite Versus Infinite — Contributions to an Eternal Dilemma*, C. S. Calude and G. Paun, Eds. Berlin Heidelberg: Springer-Verlag, 2000, pp. 149–173.

[4] M. Inuiguchi, "Generalizations of rough sets: From crisp to fuzzy cases," in *Rough Sets and Current Trends in Computing*, ser. Lecture Notes in Artificial Intelligence, S. Tsumoto *et al.*, Eds., vol. 3066. Berlin Heidelberg: Springer-Verlag, 2004, pp. 26–37.

[5] A. M. Radzikowska and E. E. Kerre, "A comparative study of fuzzy rough sets," *Fuzzy Sets and Systems*, vol. 126, pp. 137–155, 2002.

[6] J. T. Starczewski, *Advanced Concepts in Fuzzy Logic and Systems with Membership Uncertainty*, ser. Studies in Fuzziness and Soft Computing. Berlin Heidelberg: Springer-Verlag, 2013, vol. 284.

[7] L. D'eer and C. Cornelis, "A comprehensive study of fuzzy covering-based rough set models: Definitions, properties and interrelationships," *Fuzzy Sets and Systems*, vol. 336, pp. 1–26, 2018.

[8] K. Zhang, J. Zhan, and W.-Z. Wu, "Novel fuzzy rough set models and corresponding applications to multi-criteria decision-making," *Fuzzy Sets and Sytems*, vol. 383, pp. 92–126, 2020.

[9] J. D. Katzberg and W. Ziarko, "Variable precision extension of rough sets," *Fundamenta Informaticae*, vol. 27, pp. 155–168, 1996.

[10] W. Ziarko, "Variable precision rough sets model," *Journal of Computer and System Sciences*, vol. 46, no. 1, pp. 39–59, 1993.

[11] J. M. Fernández Salido and S. Murakami, "Rough set analysis of a general type of fuzzy data using transitive aggregations of fuzzy similarity relations," *Fuzzy Sets and Systems*, vol. 139, pp. 635–660, 2003.

[12] W. N. Liu, J. Yao, and Y. Yao, "Rough approximations under level fuzzy sets," in *Rough Sets and Current Trends in Computing*, ser. Lecture Notes in Artificial Intelligence, S. Tsumoto *et al.*, Eds., vol. 3066. Berlin Heidelberg: Springer-Verlag, 2004, pp. 78–83.

[13] C. Cornelis, M. D. Cock, and A. M. Radzikowska, "Vaguely quantified rough sets," in *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, ser. Lecture Notes in Artificial Intelligence, A. An *et al.*, Eds., vol. 4482. Berlin Heidelberg: Springer-Verlag, 2007, pp. 87–94.

[14] C. Cornelis, N. Verbiest, and R. Jensen, "Ordered weighted average based fuzzy rough sets," in *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, ser. Lecture Notes in Artificial Intelligence, A. An *et al.*, Eds., vol. 6401. Berlin Heidelberg: Springer-Verlag, 2010, pp. 78–85.

[15] Q. Hu, L. Zhang, S. An, and D. Zhang, "On robust fuzzy rough set models," *IEEE Transactions on Fuzzy Systems*, vol. 20, pp. 636–651, 2012.

[16] Y. Yao, J. Mi, and Z. Li, "A novel variable precision (θ, σ)-fuzzy rough set model based on fuzzy granules," *Fuzzy Sets and Systems*, vol. 236, pp. 58–72, 2014.

[17] S. Zhao, E. C. C. Tsang, and D. Chen, "The model of fuzzy variable precision rough sets," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 2, pp. 451–467, 2009.

[18] L. D'eer, N. Verbiest, C. Cornelis, and L. Godo, "A comprehensive study of implicator-conjunctor-based and noise-tolerant fuzzy rough sets: Definitions, properties and robustness analysis," *Fuzzy Sets and Systems*, vol. 275, pp. 1–38, 2015.

[19] T. Feng, H.-T. Fan, and J.-S. Mi, "Uncertainty and reduction of variable precision multigranulation fuzzy rough sets based on three-way decisions," *International Journal of Approximate Reasoning*, vol. 85, pp. 36–58, 2017.

[20] B. W. Fang and B. Q. Hu, "Granular fuzzy rough sets based on fuzzy implicators and coimplicators," *Fuzzy Sets and Systems*, vol. 359, pp. 112–139, 2019.

[21] C. Y. Wang, "A comparative study of variable precision fuzzy rough sets based on residuated lattices," *Fuzzy Sets and Systems*, vol. 373, pp. 94–115, 2019.

[22] A. Mieszkowicz-Rolka and L. Rolka, "A novel approach to fuzzy rough set-based analysis of information systems," in *Information Systems Architecture and Technology. Knowledge Based Approach to the Design, Control and Decision Support*, ser. Advances in Intelligent Systems and Computing, Z. Wilimowska *et al.*, Eds., vol. 432. Switzerland: Springer International Publishing, 2016, pp. 173–183.

[23] ——, "Labeled fuzzy rough sets versus fuzzy flow graphs," in *Proceedings of the 8th International Joint Conference on Computational Intelligence – Volume 2: FCTA*, J. J. Merelo *et al.*, Eds. SCITEPRESS Digital Library – Science and Technology Publications, Lda, 2016, pp. 115–120, www.scitepress.org/DigitalLibrary.

[24] ——, "Flow graphs and decision tables with fuzzy attributes," in *Artificial Intelligence and Soft Computing — ICAISC 2006*, ser. Lecture Notes in Artificial Intelligence, L. Rutkowski *et al.*, Eds., vol. 4029. Berlin Heidelberg: Springer-Verlag, 2006, pp. 268–277.

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

[26] A. Mieszkowicz-Rolka and L. Rolka, "Fuzzy rough approximations of process data," *International Journal of Approximate Reasoning*, vol. 49, pp. 301–315, 2008.

[27] ——, "On representation and analysis of crisp and fuzzy information systems," in *Transactions on Rough Sets VI*, ser. Lecture Notes in Computer Science (Journal Subline), J. F. Peters *et al.*, Eds., vol. 4374. Berlin Heidelberg: Springer-Verlag, 2007, pp. 191–210.

[28] A. Skowron and J. Stepaniuk, "Tolerance approximation spaces," *Fundamenta Informaticae*, vol. 27, pp. 245–253, 1996.

[29] C. Cornelis, C. Van der Donck, and E. Kerre, "Sinha-dougherty approach to the fuzzification of set inclusion revisited," *Fuzzy Sets and Systems*, vol. 134, pp. 283–295, 2003.

[30] A. Mieszkowicz-Rolka and L. Rolka, "Variable precision fuzzy rough sets," in *Transactions on Rough Sets I*, ser. Lecture Notes in Computer Science (Journal Subline), J. F. Peters *et al.*, Eds., vol. 3100. Berlin Heidelberg: Springer-Verlag, 2004, pp. 144–160.

[31] R. Słowiński, Ed., *Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory*. Boston Dordrecht London: Kluwer Academic Publishers, 1992.

[32] S. Tsumoto *et al.*, Eds., *Rough Sets and Current Trends in Computing*, ser. Lecture Notes in Artificial Intelligence, vol. 3066. Berlin Heidelberg: Springer-Verlag, 2004.