Journal of Southeast University (English Edition) Dec. 2002 Vol.18 No.4 ISSN 1003—7985

Variable Precision Rough Set and a Fuzzy Measure of Knowledge Based on Variable Precision Rough Set

Jian Lirong Da Qingli Chen Weida

(College of Economics and Management, Southeast University, Nanjing 210096, China)

Abstract: Variable precision rough set (VPRS) is an extension of rough set theory (RST). By setting threshold value β, VPRS looses the strict definition of approximate boundary in RST. Confident threshold value for β is discussed and the method for deriving decision-making rules from an information system is given by an example. An approach to fuzzy measures of knowledge is proposed by applying VPRS to fuzzy sets. Some properties of this measure are studied and a pair of lower and upper approximation operators in fuzzy sets are described. Research results reveal that, based on VPRS, fuzzy membership functions can be explicitly interpreted and semantics of membership values can be explicitly stated.

Key words: variable precision rough set, fuzzy set, information system, fuzzy measures

Rough set theory (RST) introduced by Z. Pawlak in 1982 has been described as a new mathematical tool to deal with vagueness and uncertainty. This approach has been successfully applied to machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, pattern recognition, expert systems and decision support systems. However, an object classified using initial RST, is assumed that there is complete certainty that it is a correct classification by an equivalence relation. Namely, an object belongs to or not to a classification. An object cannot be classified in a level of confidence in its correct classification. In its formalism it cannot recognize the presence or absence of non-deterministic relationships, i.e. the ones which can lead to predictive rules with probabilities less than one. In real world decision making, data in an information system acquired by means of random or statistical method are often ambiguous, incomplete, and noisy. So the patterns of classes often overlap, which is not sufficient to produce deterministic rules but may be quite possible to identify strong non-deterministic rules with estimates of decision probabilities. To overcome these problems, an extended variable precision rough set (VPRS) model was proposed in Ref. [1]. By setting threshold value β, VPRS relaxes the strict definition of approximate boundary in initial RST to allow for probabilistic classification. In contrast to RST, when an object is classified in VPRS there is a level of confidence in its correct classification, which perfects the concepts of approximation space, and helps to discover related knowledge from non-related data. Although threshold value β has been investigated by some authors, there is still a lack of systematic study. In this paper, threshold value β is investigated, rules based on VPRS are induced through an example, a fuzzy measure of knowledge is proposed by applying VPRS model to fuzzy sets, some properties are discussed, a sound semantic interpretation of fuzzy membership functions is proposed, and semantics of membership values are explicitly stated.

1 Variable Precision Rough Set

1.1 Basic concepts

When an object is classified in VPRS, a confident threshold value for β needs to be defined. W. Ziarko considered β as a classification error, defined to be in the domain $[0.0, 0.5)$[1]. However, A. An et al. used symbol β to denote the proportion of correct classification, in which case the appropriate range is $(0.5, 1.0]$. They referred to the technique as enhanced RST[2]. VPRS model was extended to incorporate asymmetric bounds on certain classification probabilities in Ref. [3]. Without loss of generality, this paper is restricted to initial VPRS version.

Definition 1 Suppose $S = (U, A, V, f)$ is an information system, where $U = \{U_1, U_2, U_3, \cdots, U_{|u|}\}$ is a finite set of objects(universe), $A = \{a_1, a_2, a_3, \cdots, a_{|A|}\}$ is a finite set of attributes. If the

Received 2002-06-17.

* The project supported by the Project of Education Bureau Foundation of China (01JA630048).

** Born in 1968, female, graduate.

attributes of set A can be partitioned into a set of condition attributes $C \neq \varnothing$ and a set of decision attributes $D \neq \varnothing$, $A = C \cup D$ and $C \cap D = \varnothing$, such a table is also called a decision table. V_a is the domain of the attribute a, $V = \cup V_a$ and $f: U \times A \to V$ is a total function such that $f(x, a) \in V_a$ for $\forall a \in A$ and $\forall x \in U$, called information function.

To every subset $\varnothing \neq P \subseteq C$, the equivalence relation is denoted by

$I = \{(x, y) \in U \times U: f(x, p) = f(y, p) \ \forall p \in P\}$

The corresponding equivalence class is denoted by $I(P)$. In terms of the basic ideas of VPRS, the definitions are given as follows.

Definition 2 For a given information system $S = (U, A, V, f), A = C \cup D, X \subseteq U, P \subseteq C, 0.5 < \beta \leqslant 1$. β-lower approximation and β-upper approximation of X in S are defined, respectively, by

$$\underline{\mathrm{apr}}_P^{\beta}(X) = \cup \left\{ I(P): \frac{|I(P) \cap X|}{|I(P)|} \geqslant \beta \right\}$$

$$\overline{\mathrm{apr}}_P^{\beta}(X) = \cup \left\{ I(P): \frac{|I(P) \cap X|}{|I(P)|} > 1 - \beta \right\}$$

Definition 3 β-negative region and β-boundary region of X are defined in S, respectively, by

$$\mathrm{neg}_P^{\beta}(X) = \cup \left\{ I(P): \frac{|I(P) \cap X|}{|I(P)|} \leqslant 1 - \beta \right\}$$

$$\mathrm{bnd}_P^{\beta}(X) = \cup \left\{ I(P): 1 - \beta < \frac{|I(P) \cap X|}{|I(P)|} < \beta \right\}$$

where $|\cdot|$ denotes the cardinality of a set.

Definition 4 The measure of quality of classification is defined by

$$\gamma^{\beta}(P, D) = \left| \cup \left\{ I(P): \frac{|X \cap I(P)|}{|I(P)|} \geqslant \beta \right\} \right| \Big/ |U|$$

The value $\gamma^{\beta}(P, D)$ measures the proportion of objects in the universe for which classification is possible at the specified value of β.

Definition 5 An approximate reduct $\mathrm{red}^{\beta}(C, D)$ denoted by β-reduct is defined as the minimal subset of C which keeps the quality of classification unchanged at the specified value β.

$\mathrm{red}^{\beta}(C, D)$ has the twin properties that:

1) $\gamma^{\beta}(C, D) = \gamma^{\beta}(\mathrm{red}^{\beta}(C, D), D)$;

2) No proper subset of $\mathrm{red}^{\beta}(C, D)$ can give the same quality of classification at the same β value.

Reducts are important to final objective of constructing a series of rules to classify a number of objects in the model, the related references have discussed two particular aspects of reducts. On one hand there is the problem of finding the reducts for a given system. This problem is NP-complete[4]. On the other hand there is the problem of finding local reducts for each object, such as set theoretical approach. Further research has been conducted into dynamic reducts and tolerance reducts.

1.2 Research into value β

If $\beta = 1$, $\underline{\mathrm{apr}}_P^{\beta}(x)$ and $\overline{\mathrm{apr}}_P^{\beta}(x)$ coincide with the lower and upper approximation sets in RST. VPRS model comes back to the original RST. For the inconsistent rules of RST, if the inconsistency degree is weak according to the setting threshold value β, an indeterministic rule can be considered as a deterministic one originally but becomes a little inconsistent because of some noises mixed in the given data. This rule or the main of it can be viewed as deterministic one. However, if the inconsistency degree is strong, the corresponding rule is real indeterministic and should be treated as a random rule.

β value is inversely related to the quality of classification. There are two different directions that can be taken. In one direction, when β value increases, the quality of classification is decreasing. Positive region and negative region of set X will become narrower. In other words, the boundary region of set X will become wider. A small number of objects are classified. In the other direction, more objects which are classified can be classified incorrectly.

Proposition 1[2] If condition class X is not given a classification with $0.5 < \beta \leqslant 1$, then X is also indiscernible at any level $\beta < \beta_1 \leqslant 1$. In contrast, if condition class X is given a classification with $0.5 < \beta \leqslant 1$, then X is also discernible at any level $0.5 < \beta_2 < \beta$.

If a condition class is not given a classification for every β, a condition class X is called absolutely indiscernible or absolutely rough. In other words, if and only if $\mathrm{bnd}_P^{\beta}(X) \neq \varnothing$, a condition class X is absolutely rough. In contrast, those only given a classification for a range of β is called relatively rough or weak discernible. For each relatively rough set X, there is a threshold value on the β value on which set X is discernible. Associated with each conditional class is an upper bound on the β value. If any β value chosen which is equal or below the threshold means the set X is discernable. Otherwise, there is no opportunity for majority inclusion, hence the set X is indiscernible. The highest of these upper bounds on the β values is defined as $\beta_{\max}$.

Proposition 2 For any $0.5 < \beta \leqslant 1$, the following holds:

① $\overline{\mathrm{apr}}_P^\beta(X \cup Y) \supseteq \overline{\mathrm{apr}}_P^\beta(X) \cup \overline{\mathrm{apr}}_P^\beta(Y)$;

② $\underline{\mathrm{apr}}_P^\beta(X \cap Y) \subseteq \underline{\mathrm{apr}}_P^\beta(X) \cap \underline{\mathrm{apr}}_P^\beta(Y)$;

③ $\underline{\mathrm{apr}}_P^\beta(X \cup Y) \supseteq \underline{\mathrm{apr}}_P^\beta(\mathrm{X}) \cup \underline{\mathrm{apr}}_P^\beta(Y)$;

④ $\overline{\mathrm{apr}}_P^\beta(X \cap Y) \subseteq \overline{\mathrm{apr}}_P^\beta(X) \cap \overline{\mathrm{apr}}_P^\beta(Y)$.

Proof of ① For any two sets $X \subseteq U, Y \subseteq U$, given β value,

$$\frac{|I(P) \cap (X \cup Y)|}{|I(P)|} \geqslant \frac{|I(P) \cap X|}{|I(P)|}$$

and

$$\frac{|I(P) \cap (X \cup Y)|}{|I(P)|} \geqslant \frac{|I(P) \cap Y|}{|I(P)|}$$

Therefore, $\overline{\mathrm{apr}}_P^\beta(X \cup Y) \supseteq \overline{\mathrm{apr}}_P^\beta(X) \cup \overline{\mathrm{apr}}_P^\beta(Y)$.

Proof of ② For any two sets $X \subseteq U, Y \subseteq U$, given β value,

$$\frac{|I(P) \cap (X \cap Y)|}{|I(P)|} \leqslant \frac{|I(P) \cap X|}{|I(P)|}$$

and

$$\frac{|I(P) \cap (X \cap Y)|}{|I(P)|} \leqslant \frac{|I(P) \cap Y|}{|I(P)|}$$

Therefore, $\underline{\mathrm{apr}}_P^\beta(X \cap Y) \subseteq \underline{\mathrm{apr}}_P^\beta(X) \cap \underline{\mathrm{apr}}_P^\beta(Y)$.

③ and ④ can be proved in a similar way.

1.3 An example

An information system is given in Tab.1, $S = (U, A, V, f)$, where $U = \{n_1, n_2, n_3, n_4, n_5, n_6, n_7\}$, the set of condition attributes $C = \{a_1, a_2, a_3, a_4, a_5, a_6\}$, the set of decision attributes $D = \{d\}$.

Tab.1 An information system

U	Condition attributes C						Decision attribute D
	a_1	a_2	a_3	a_4	a_5	a_6	d
n_1	1	2	1	1	2	1	N
n_2	1	2	2	2	2	2	N
n_3	2	2	1	2	2	2	N
n_4	1	2	1	1	2	1	P
n_5	2	2	2	2	1	1	P
n_6	1	2	1	1	2	1	P
n_7	2	2	2	2	1	2	P

The universe U is partitioned into the following equivalence classes:

$$U/C = \{X_1, X_2, X_3, X_4, X_5\}$$

where $X_1 = \{n_1, n_4, n_6\}, X_2 = \{n_2\}, X_3 = \{n_3\}, X_4 = \{n_5\}, X_5 = \{n_7\}$.

$$U/D = \{Y_N, Y_P\}$$

where $Y_N = \{n_1, n_2, n_3\}, Y_P = \{n_4, n_5, n_6, n_7\}$.

β-reduct and quality of classification computed using the same method as depicted in 1.1 are shown in Tab.2.

As an example, Tab.3 provides the minimal rules associated with β-reduct $\{a_1, a_3\}$.

Tab.2 β-reduct and quality of classification

β-reduct	Quality of classification	$\beta_{\max}$
$\{a_2\}$	1.00	0.57
$\{a_1, a_3\}$	1.00	0.67
$\{a_1, a_4\}$	1.00	0.67
$\{a_4, a_5\}$	0.57	1.00

Tab.3 Decision rules for the β-reduct $\{a_1, a_3\}$

Rules	Support	Degree of confidence
$a_1 = 1 \wedge a_3 = 2 \xrightarrow{100\%} N$	1	1.00
$a_1 = 2 \wedge a_3 = 1 \xrightarrow{100\%} N$	1	1.00
$a_1 = 1 \wedge a_3 = 1 \xrightarrow{67\%} P$	3	0.67
$a_1 = 2 \wedge a_3 = 2 \xrightarrow{100\%} P$	2	1.00

2 Fuzzy Measure of Knowledge Based on VPRS

Definition 6 Let $S = (U, A, V, f)$ be an information system. Inspired by Ref. [5], a fuzzy set FX is defined as

$$\mathrm{FX} = \{(p, \mu_{\mathrm{FX}}(p)) : p \in U, \mu_{\mathrm{FX}}(p) = \frac{|I(P) \cap X|}{|I(P)|}\}$$

where $\mu_{\mathrm{FX}}(p)$ is fuzzy membership function.

$\mu_{\mathrm{FX}}(p)$ can be interpreted as the conditional probability that an arbitrary element of $I(P)$ belongs to a given class X. By definition, the membership values are all rational numbers. If elements are in the same equivalence class they must have the same degree of membership. That is indiscernible elements should have the same membership value. Obviously, $0 \leqslant \mu_{\mathrm{FX}}(p) \leqslant 1$. For each element $x \in U$, if $\mu_{\mathrm{FX}}(p) = 0$ then p does not belong to FX, and if $\mu_{\mathrm{FX}}(p) = 1$ then p surely belongs to FX. If $0 < \mu_{\mathrm{FX}}(p) < 1$ then p possibly belongs to FX, in this case, there is a transitive state between $p \in \mathrm{FX}$ and $p \notin \mathrm{FX}$.

Definition 7[6] Suppose $S = (U, A, V, f)$ is an information system, and FX is a fuzzy set of U. Then lower approximation and upper approximation of FX in S are defined, respectively, by

$$\mu_{\underline{\mathrm{FX}}}(p) = \inf\{\mu_{\mathrm{FX}}(p) \mid y \in I(P)\} \quad \forall p \in U$$

$$\mu_{\overline{\mathrm{FX}}}(p) = \sup\{\mu_{\mathrm{FX}}(p) \mid y \in I(P)\} \quad \forall p \in U$$

Proposition 3 For any two sets X and Y in an information system $S = (U, A, V, f)$, if $X \subseteq Y$ then $\mathrm{FX}_x \subseteq \mathrm{FX}_Y$.

Proof $\forall p \in U$, clearly, $X \subseteq Y$ implies $|I(P) \cap X| \leqslant |I(P) \cap Y|$, the following holds:

$$\frac{|I(P) \cap X|}{|I(P)|} \leqslant \frac{|I(P) \cap Y|}{|I(P)|}$$

That is to say $\mu_{\mathrm{FX}_X} \leqslant \mu_{\mathrm{FX}_Y}$, therefore $\mathrm{FX}_x \subseteq \mathrm{FX}_Y$.

Proposition 4 For any two sets X and Y in an information system $S = (U, A, V, f)$, the following holds:

① $FX_{X\cup Y} \supseteq FX_X \cup FX_Y$;

② $FX_{X\cap Y} \subseteq FX_X \cap FX_Y$.

Proof of ① $\forall p \in U$,

$$\mu_{FX_{X\cup Y}}(p) = \frac{|I(P) \cap (X \cup Y)|}{|I(P)|} = \frac{|(I(P) \cap X) \cup (I(P) \cap Y)|}{|I(P)|} \geqslant \frac{\max\{|I(P) \cap X|, |I(P) \cap Y|\}}{|I(P)|} = \max\left\{\frac{|I(P) \cap X|}{|I(P)|}, \frac{|I(P) \cap Y|}{|I(P)|}\right\} = \max\{\mu_{FX_X}, \mu_{FX_Y}\} = \mu_{FX_X \cup FX_Y}$$

Therefore $FX_{X\cup Y} \supseteq FX_X \cup FX_Y$.

② can be proved in a similar way.

Definition 8 Suppose $S = (U, A, V, f)$ is an information system, FX is a fuzzy set of U, and $0.5 < \beta \leqslant 1$. The lower approximation and upper approximation are defined, respectively, by

$$\mu^{\beta}_{\underline{FX}}(p) = \cup\{I(P) \mid \mu_{\underline{FX}}(p) \geqslant \beta\} \quad \forall p \in U$$

$$\mu^{\beta}_{\overline{FX}}(p) = \cup\{I(P) \mid \mu_{\overline{FX}}(p) > 1 - \beta\} \quad \forall p \in U$$

Proposition 5 Let FX is a classical set, for $0.5 < \beta \leqslant 1$, $\mu^{\beta}_{\underline{FX}}(p)$ and $\mu^{\beta}_{\overline{FX}}(p)$ come back to the lower approximation $\underline{apr}X$ and upper approximation $\overline{apr}X$ of original rough set model.

Proof For $0.5 < \beta \leqslant 1$, FX is a classical set, and $\mu_{FX}(p) \in \{0, 1\}$, hence,

$$\mu^{\beta}_{\underline{FX}}(p) = \{p \in U \mid \mu_{\underline{FX}}(p) = 1\}$$

Then, $\forall y \in I(P)$,

$$\{p \in U \mid \mu_{\underline{FX}}(p) = 1\} = \{p \in U \mid \mu_{\underline{FX}}(y) = 1\} = \{p \in U \mid I(P) \subseteq FX\}$$

Therefore $\mu^{\beta}_{\underline{FX}}(p) = \underline{apr}X$.

The proof of $\mu^{\beta}_{\overline{FX}}(p) = \overline{apr}X$ is similar.

The lower approximation $\mu^{\beta}_{\underline{FX}}(p)$ of fuzzy FX can be interpreted as the union of the equivalence class of the elements whose degree of confidence belonging to FX is not below β, while $\mu^{\beta}_{\overline{FX}}(p)$ can be interpreted as the union of the equivalence class of the elements whose membership degree belonging to FX is above $1 - \beta$.

The theory of fuzzy sets is typically developed as an uninterpreted mathematical theory of abstract membership functions without the above limitations. In contrast, VPRS provides a more specific and more concrete interpretation of fuzzy membership functions. These are illustrated through the given example in 1.3 as follows.

Let fuzzy set $FX = \{FX_N, FX_P\}$, where $FX_N = Y_N$ and $FX_P = Y_P$.

By definition 6, the membership functions of FX_n and FX_P are given, respectively, by

$FX_N = \{0.33/n_1, 1/n_2, 1/n_3, 0.33/n_4, 0/n_5, 0.33/n_6, 0/n_7\}$

$FX_P = \{0.67/n_1, 0/n_2, 0/n_3, 0.67/n_4, 1/n_5, 0.67/n_6, 1/n_7\}$

Let $0.5 < \beta \leqslant 0.67$, the lower approximation and the upper approximation of fuzzy set FX are as follows:

$$\mu^{\beta}_{\underline{FX}}(p) = \{n_2, n_2, n_3, n_4, n_5, n_6, n_7\}$$

$$\mu^{\beta}_{\overline{FX}}(p) = \{n_1, n_2, n_3, n_4, n_5, n_6, n_7\}$$

Let $1 \geqslant \beta > 0.67$, the lower approximation and the upper approximation of fuzzy set FX are as follows:

$$\mu^{\beta}_{\underline{FX}}(p) = \{n_2, n_3, n_5\ n_7\}$$

$$\mu^{\beta}_{\overline{FX}}(p) = \{n_1, n_2, n_3, n_4, n_5, n_6, n_7\}$$

3 Conclusion

The notions of fuzzy set theory and rough set theory are not rival ones but two different mathematical tools and aim to two different purposes. Typically, a fuzzy set is defined as $U \to [0,1]$. Although such a fuzzy system provides a sound and consistent mathematical model, there is a lack of semantic interpretation of fuzzy membership values and fuzzy set theoretic operations. This might lead to some difficulties in the applications of the theory. We apply VPRS to fuzzy set. The source of the fuzziness in describing a concept is viewed as the indiscernibility of elements. Thus a more restrictive, but more concrete view of fuzzy set and a more specific interpretation of fuzzy membership functions are provided. Such more concrete views of fuzzy sets, with explicitly stated semantics of membership values, can be more useful for the applications of the theory of fuzzy sets.

References

[1] Ziarko W. Variable precision rough set model[J]. *Journal of Computer and System Sciences*, 1993, **46**(1): 39 - 59.

[2] An A, Shan N, Chan C, et al. Discovering rules for water demand prediction: An enhanced rough-set approach[J]. *Engineering Application and Artificial Intelligence*, 1996, **9**(6): 645 - 653.

[3] Dimitras A I, Slowinski R, Susmaga R, et al. Business failure using rough sets[J]. *European Journal of Operational Research*, 1999, **114**(2): 263 - 280.

[4] Skowron A, Rauszer C. *Intelligent decision support-handbook of applications and advances of the rough sets theory*[M]. Dor-

drecht: Luwer Academic Publishers, 1992.331－362.

[5] Dubois D, Prade H. Rough-fuzzy sets and fuzzy-rough sets [J]. *International Journal of General Systems*, 1990, **17**(2): 191－209.

[6] Zhang Wenxiu, Wu Weizhi, Liang jiye. *Rough set theory and methodology*[M]. Bejing: Science Press, 2001.128－158 (in Chinese)

变精度粗糙集与基于变精度粗糙集的知识模糊度量

管利荣　达庆利　陈伟达

(东南大学经济管理学院,南京 210096)

摘　要　变精度粗糙集是对标准粗糙集理论的一种扩展.它通过设置阈值参数 β,放松了标准粗糙集理论对近似边界的严格定义.文中讨论了变精度粗糙集的置信阈值 β,通过算例给出了信息系统中基于变精度粗糙集的规则提取方法;将变精度粗糙集模型应用于模糊集,提出了在变精度粗糙集中知识的一种模糊度量方法,对这种方法的一些性质进行了研究,并用该模糊度量方法描述了近似算子.研究表明,该方法可合理解释模糊隶属函数,清晰说明了隶属度的含义.

关键词　变精度粗糙集,模糊集,信息系统,模糊度量

中图分类号　TP311