A Novel Approach to Generating Fuzzy Rules Based on Dynamic Fuzzy Rough Sets

Yi Cheng
Department of Computer Science Technology
Tongji University, Shanghai, 201804, P.R.China
chengyimail@sohu.com

Duoqian Miao, Qinrong Feng
Department of Computer Science Technology
Tongji University, Shanghai, 201804, P.R.China
miaoduoqian@163.com

Abstract

We propose a novel approach to generating fuzzy rules which, different from most known fuzzy rules induction, is not based on attributes reduction (AR) but granulation order and variational universe. Most rule induction algorithms based on fuzzy rough sets (FRS) usually include two steps: AR and fuzzy rules induction. It's helpful to shorten the time of rule mining to some extent by AR.. However, AR may make against the induction of fine rules due to its limitation. Avoiding AR in fuzzy rules induction permits to improve the adaptability of fuzzy rules and reduce computational complexity. In this paper, the dynamic FRS is presented in two different ways. Then, an algorithm based on dynamic FRS is put forward for decision rule mining. At last, an initial experimentation is conducted, comparing the new method with a conventional FRS-based rule mining.

1. Introduction

Rough sets theory was originally based on the notions of classical sets theory. Dubois and Prade [1] were among the first who showed that the basic idea of rough sets can be extended in order to describe concept in fuzzy approximation space. The idea of fuzzy rough sets (FRS) was pursued and investigated in many papers [2-7].

Based on FRS, many rule induction algorithms are presented, which usually include two steps: attributes reduction (AR) based on FRS and fuzzy rules induction based on conventional rules mining algorithms. AR usually acts as a preprocessor of fuzzy rules induction. However, for most AR algorithms, three facts should be pointed out. The first is that usually only one reduction can be obtained by the algorithm. The second is that the algorithm may not be convergent on many real data sets or the selected attributes are unreliable. The last is that the computational complexity of the algorithm often increases exponentially with the number of input variables and the size of data patterns. Therefore, avoiding the process of AR permits to improve the adaptability of fuzzy rules and reduce computational complexity.

This paper, based on the most recent work as reported in Refs.[8], puts forward a new approach to generating fuzzy rules which, different from most known fuzzy rules induction, is not based on AR but granulation order and variational universe. From granular computing point of view, a concept is characterized by upper and lower approximations under static granulation in FRS theory defined by Dubois and Prade. Provided that the granulation is unchangeable, whether the granulation is too small or not may be unacceptable. The former will increase the time and cost, the latter may not satisfy the requirements. Therefore, a method of describing a concept in fuzzy approximation space by variational upper and lower approximations under dynamic granulation is proposed which nominated as dynamic fuzzy rough sets (DFRS). This means that a proper granulation family can be selected for a target concept approximation according to the practical requirements.

The rest of this paper is organized as follows. In section 2, a review of rough sets approximation based on dynamic granulation is given. In section 3, FRS under dynamic granulation is proposed by two means: one is based on cut sets, another is a direct generalization of dynamic rough sets. Some important properties are obtained consequently. Then a fuzzy rules induction algorithm based on DFRS is designed. An initial experimental investigation is conducted. Section 4 shows the algorithm is effective, supported by comparisons to the application of fuzzy rules induction based on attributes reduction in Refs.[4]. Section 5 concludes the paper.

0-7695-3032-X/07 $25.00 © 2007 IEEE
DOI 10.1109/GrC.2007.60

2. Rough sets approximation based on dynamic granulation

In rough sets theory, a concept is always characterized via the upper and lower approximations under static granulation.

Let $S=(U,A)$ be an information system, where U is a set called universe and A is a non-empty finite set of attributes. X is a subset of U and $P\subseteq A$ is an attribute set. X is characterized by $\overline{P}(X)$ and $\underline{P}(X)$, where

$$\overline{P}(X)=\cup\{Y\in U/P \mid Y\cap X\neq\phi\},$$

$$\underline{P}(X)=\cup\{Y\in U/P \mid Y\subseteq X\}.$$

However, the fixed granulation usually limits the application of rough sets theory. In Refs.[10], rough sets based on dynamic granulation was proposed. Some basic concepts are given below.

Let $S=(U,A)$ be an information system, $P,Q\in 2^A$ are two attribute subsets. Define a partial relation $\prec$ on 2^A as follows: $P\prec Q$ if and only if, for every $P_i\in U/P$, there exists $Q_j\in U/Q$ such that $P_i\subseteq Q_j$, where $U/P=\{P_1,P_2,\cdots,P_m\}$ and $U/Q=\{Q_1,Q_2,\cdots,Q_n\}$ are partitions induced by P and Q.

In an information system, a partition U/R induced by the equivalence relation $R,R\in 2^A$, provides a granulation space for describing a concept X. So a sequence of attribute sets $R_i\in 2^A\,(i=1,2,\cdots,n)$ with $R_1\succ R_2\succ\cdots\succ R_n$ can determine a sequence of granulation spaces, from the biggest to the smallest one. The upper and lower approximations of a concept under a granulation order are defined as follows.

Definition 1[8]. Let $S=(U,A)$ be an information system, X be a subset of U and $P=\{R_1,R_2,\cdots,R_n\}$ be a family of attribute sets with $R_1\succ R_2\succ\cdots\succ R_n$ $(R_i\in 2^A,\ i=1,2\cdots,n)$, P-upper approximation $\overline{apr_P}(X)$ and P-lower approximation $\underline{apr_P}(X)$ of X are defined as follows:

$$\overline{apr}_P(X)=\overline{R_n}X,$$

$$\underline{apr}_P(X)=\bigcup_{i=1}^{n}\underline{R_i}X_i,$$

Where $X_1=X$ and $X_i=X-\bigcup_{k=1}^{i-1}\underline{R_k}X_k$, $i=2,3,\cdots,n$.

3. Dynamic fuzzy rough sets (DFRS)

This section proposes the FRS under dynamic granulation by two means: one is based on cut sets, another is a direct generalization of dynamic rough sets.

3.1. Cut sets of a fuzzy set [7]

Let U be a set called universe. A fuzzy set A on U is defined by a membership function $\mu_A:U\to[0,1]$. Given a number $\alpha\in[0,1]$, an $\alpha-$cut of a fuzzy set is defined by: $A_\alpha=\{x\in U \mid \mu_A(x)\geq\alpha\}$.

A fuzzy set A can be reconstructed from its $\alpha-$cut sets as follows: $\mu_A(x)=\sup\{\alpha \mid x\in A_\alpha\}$.

Theorem 1. Let $(A_\alpha)_\alpha,\alpha\in[0,1]$, be a family of subset of U. The necessary and sufficient conditions for the existence of a fuzzy set F such that $F_\alpha=A_\alpha$, $\alpha\in[0,1]$, are:

(1) $\alpha_1\leq\alpha_2\Rightarrow A_{\alpha_1}\supseteq A_{\alpha_2}$,

(2) $\alpha_1\leq\alpha_2\leq\cdots$, and $\alpha_n\to\alpha\Rightarrow\bigcap_{n=1}^{\infty}A_{\alpha_n}=A_\alpha$.

Theorem 2. Let $\psi:[0,1]\to[0,1]$ be a given function, $(A_\alpha)_\alpha,\alpha\in[0,1]$ be a family of subset of U. The necessary and sufficient conditions for the existence of a fuzzy set F such that $F_{\psi(\alpha)}=A_\alpha$, $\alpha\in[0,1]$, are:

(1') $\psi(\alpha_1)\leq\psi(\alpha_2)\Rightarrow A_{\alpha_1}\supseteq A_{\alpha_2}$,

(2') $\psi(\alpha_1)\leq\psi(\alpha_2)\leq\cdots$, and $\psi(\alpha_n)\to\psi(\alpha)\Rightarrow\bigcap_{n=1}^{\infty}A_{\alpha_n}=A_\alpha$.

3.2. Dynamic fuzzy rough sets based on cut sets

Approximations of crisp sets in fuzzy approximation spaces are called FRS[1]. Consider a fuzzy approximation space $apr_R=(U,R)$, Where U is the universe, R is a fuzzy similarity relation. Each of R's $\beta-$cut sets is an equivalence relation. One can represent R by a family of equivalence relations: $R=(R_\beta)_\beta$, $\beta\in[0,1]$. Based on R's β-cut sets, we now research how to construct dynamic fuzzy rough sets (DFRS).

First we extend partial relation $\prec$ on 2^A to fuzzy case: Let $S=(U,P)$ be a fuzzy information system, P be a fuzzy attribute set and $I,Q\in P$. Define a partial relation $\prec$ as follows: $I\prec Q\,(Q\succ I)$ if and only if, for every $I_k \in U/I$, there exists $Q_j \in U/Q$ such that $\forall x \in U, \mu_{I_k}(x) \le \mu_{Q_j}(x)$, where $U/I = \{I_1, I_2, \cdots, I_m\}$ and $U/Q = \{Q_1, Q_2, \cdots, Q_n\}$ are fuzzy equivalence classes induced by I and Q.

Consider the approximation of a crisp set A in fuzzy approximation space $apr_P = (U, P)$, where $P = \{R_1, R_2, \cdots; R_n\}$ be a family of fuzzy similarity relation with $R_1 \succ R_2 \succ \cdots \succ R_n$. Then for $\beta \in [0,1]$, we have $R_{1\beta} \succ R_{2\beta} \succ \cdots \succ R_{n\beta}$. Based on dynamic granulation, a rough set approximation is obtained. P-upper approximation:

$\overline{apr}_{P_\beta}(A) = \overline{R_{n\beta}} A = \{x \mid [x]_{R_{n\beta}} \cap A \ne \phi\}$,

P-lower approximation:

$\underline{apr}_{P_\beta}(A) = \bigcup_{i=1}^{n} \underline{R_{i\beta}} A_i = \bigcup_{i=1}^{n} \{x \mid [x]_{R_{i\beta}} \subseteq A_i\}$,

Where $A_1 = A$ and $A_i = A_{i-1} - \underline{R_{i-1\beta}} A_{i-1}$, $i = 2,3,\cdots,n$.

That is, $(\underline{apr}_{P_\beta}(A), \overline{apr}_{P_\beta}(A))_\beta$ is a family of rough sets with reference to set A. Whether they are the families of $\beta-$cut sets of two fuzzy sets, we have following theorem:

Theorem 3. Let $(\underline{apr}_{P_\beta}(A))_\beta$ and $(\overline{apr}_{P_\beta}(A))_\beta, \beta \in [0,1]$ be a family of P-lower and P-upper approximations with respect to A respectively, then exists a pair of fuzzy sets $\underline{apr}_P(A)$ and $\overline{apr}_P(A)$ such that:

$(\underline{apr}_P(A))_{1-\beta} = \underline{apr}_{P_\beta}(A)$, $(\overline{apr}_P(A))_\beta = \overline{apr}_{P_\beta}(A)$.

Proof. Consider the family of lower approximations $(\underline{apr}_{P_\beta}(A))_\beta, \beta \in [0,1]$.

(1)Recall that $\underline{apr}_{P_\beta}(A)$ equals to $\underline{R_{n\beta}}(A)$ in essence and $R_{n\beta}$ is derived from a fuzzy similarity relation R_n, then $\beta_2 \le \beta_1 \Rightarrow R_{n\beta_2} \supseteq R_{n\beta_1}$. Let $\psi(\beta) = 1-\beta$, we have $\psi(\beta_1) \le \psi(\beta_2) \Rightarrow \underline{apr}_{P_{\beta_1}}(A) \supseteq \underline{apr}_{P_{\beta_2}}(A)$.

(2)Since $R_{n\beta}$ are derived from a fuzzy similarity relation R_n, they satisfy property (2) in Theorem 1, that is: for $\beta_1 \le \beta_2 \le \cdots$, and $\beta_m \to \beta$, we have $\bigcap_{m=1}^{\infty} R_{n\beta_m} = R_{n\beta}$.

Then we obtain: $\psi(\beta_1) \le \psi(\beta_2) \le \cdots$, and $\psi(\beta_m) \to \psi(\beta) \Rightarrow \bigcap_{m=1}^{\infty} \underline{apr}_{P_{\beta_m}}(A) = \underline{apr}_{P_\beta}(A)$. By Theorem 2, there exists a fuzzy set $\underline{apr}_P(A)$ such that $(\underline{apr}_P(A))_{\psi(\beta)} = (\underline{apr}_P(A))_{1-\beta} = \underline{apr}_{P_\beta}(A)$.

Similarly, we can show the existence of a fuzzy set $\overline{apr}_P(A)$ for the family of upper approximations $(\overline{apr}_{P_\beta}(A))_\beta$ such that $(\overline{apr}_P(A))_\beta = \overline{apr}_{P_\beta}(A)$. $\underline{apr}_P(A)$ and $\overline{apr}_P(A)$ are defined as:

$$\mu_{\underline{apr}_P(A)}(x) = \sup\{\psi(B) \mid x \in (\underline{apr}_P(A))_{\psi(\beta)}\}$$
$$= \sup\{1-\beta \mid x \in \underline{apr}_{P_\beta}(A)\},$$
$$\mu_{\overline{apr}_P(A)}(x) = \sup\{\beta \mid x \in (\overline{apr}_P(A))_\beta\}$$
$$= \sup\{\beta \mid x \in \overline{apr}_{P_\beta}(A)\}.$$

Definition 2: Let $S = (U, P)$ be a fuzzy information system, A be a crisp set of U and $P = \{R_1, R_2, \cdots, R_n\}$ be a family of fuzzy attribute sets with $R_1 \succ R_2 \succ \cdots \succ R_n$, where $R_k (k = 1,2\cdots,n)$ is a fuzzy similarity relation, P-upper approximation $\overline{apr}_P(A)$ and P-lower approximation $\underline{apr}_P(A)$ of A are defined by the following membership functions:

$\mu_{\overline{apr}_P(A)}(x) = \sup\{\beta \mid x \in \overline{apr}_{P_\beta}(A)\}$,

$\mu_{\underline{apr}_P(A)}(x) = \sup\{1-\beta \mid x \in \underline{apr}_{P_\beta}(A)\}$,

Where $\overline{apr}_{P_\beta}(A) = \overline{R_{n\beta}} A = \{x \mid [x]_{R_{n\beta}} \cap A \ne \phi\}$,

$$\underline{apr}_{P_\beta}(A) = \bigcup_{i=1}^{n} \underline{R_{i\beta}} A_i = \bigcup_{i=1}^{n} \{x \mid [x]_{R_{i\beta}} \subseteq A_i\},$$

and $A_1 = A$ and $A_i = A_{i-1} - \underline{R_{i-1\beta}} A_{i-1}$,for $i = 2,3,\cdots,n, \beta \in [0,1]$.

Theorem 4. Let $S = (U, P)$ be a fuzzy information system, A be a crisp set of U and $P = \{R_1, R_2, \cdots; R_n\}$ be a family of fuzzy attribute sets with $R_1 \succ R_2 \succ \cdots \succ R_n$, where $R_k (k = 1,2\cdots,n)$ is a fuzzy similarity relation. Let $P_i = \{R_1, R_2, \cdots; R_i\}$, then for $\forall P_i, (i = 1,2,\cdots,n)$, the following properties hold:

$$\underline{apr}_P(A) \subseteq A \subseteq \overline{apr}_P(A),$$

$$\underline{apr}_{P_1}(A) \subseteq \underline{apr}_{P_2}(A) \subseteq \cdots \subseteq \underline{apr}_{P_n}(A).$$

In order to describe the uncertainty of concept under a granulation order, the approximation precision is defined as follows.

Definition 3. Let $S=(U,P)$ be a fuzzy information system, A be a crisp set of U and $P=\{R_1,R_2,\cdots,R_n\}$ be a family of fuzzy attribute sets with $R_1 \succ R_2 \succ \cdots \succ R_n$, where $R_k(k=1,2\cdots,n)$ is a fuzzy similarity relation. The approximation precision $\alpha_P(A)$ is defined as:

$$\alpha_P(A)=\frac{power\ (\underline{\underline{apr}}_P(A))}{power\ (\overline{\overline{apr}}_P(A))}=\frac{\sum_{x\in U}\mu_{\underline{apr}_P(A)}(x)}{\sum_{x\in U}\mu_{\overline{apr}_P(A)}(x)},$$

Where $A\neq\phi$.

Theorem 5. Let $P_i=\{R_1,R_2,\cdots,R_i\}$, then for $\forall P_i,(i=1,2,\cdots,n)$, we have

$$\alpha_{P_1}(A)\le\alpha_{P_2}(A)\le\cdots\le\alpha_{P_n}(A).$$

Theorem 4 and 5 state that the lower approximation enlarges and the approximation precision $\alpha_P(A)$ increases as the granulation order become longer through adding fuzzy similarity relations.

3.3. Dynamic fuzzy rough sets

The idea of describing a general concept by using the variational upper and lower approximations under dynamic granulation can also be extended to FRS directly.

Definition 4. Let $S=(U,P)$ be a fuzzy information system, A be a crisp set of U and $P=\{R_1,R_2,\cdots,R_n\}$ be a family of fuzzy attribute sets with $R_1 \succ R_2 \succ \cdots \succ R_n$, where $R_k(k=1,2\cdots,n)$ is a fuzzy similarity relation. P -upper approximation $\overline{apr}_P(A)$ and P -lower approximation $\underline{apr}_P(A)$ of A are defined as:

$$\mu_{\overline{apr}_P(A)}(x)=\sup_{y\in A}\mu_{[x]_{R_n}}(y),$$

$$\mu_{\underline{apr}_P(A)}(x)=\begin{cases}\inf_{y\notin A}\{1-\mu_{[x]_{R_1}}(y)\}, x\in W_1\\ \cdots\\ \inf_{y\notin A}\{1-\mu_{[x]_{R_n}}(y)\}, x\in W_n\\ \inf_{y\notin A}\{1-\mu_{[x]_{R_{n+1}}}(y)\}, x\in U_{n+1}\end{cases},$$

Where $\mu_{[x]_{R_k}}(y)=\mu_{R_k}(x,y),(k=1,2,\cdots,n)$, $[x]_{R_k}\in U_k/R_k(k=1,2,\cdots,n)$, specially, $[x]_{R_{n+1}}\in U_{n+1}/R_n$ and $U_1=U,U_i=U_{i-1}-W_{i-1}(i=2,3,\cdots,n+1)$,

$$W_{i-1}=\{x\mid\mu_{\underline{apr}_{R_{i-1}}(A)}(x)=\inf_{y\notin A}\{1-\mu_{[x]_{R_{i-1}}}(y)\}=1\}.$$

Generally, in order to improve adaptability, W_{i-1} can be modified as

$$W_{i-1}=\{x\mid\mu_{\underline{apr}_{R_{i-1}}(A)}(x)=\inf_{y\notin A}\{1-\mu_{[x]_{R_{i-1}}}(y)\}\ge\eta\},$$

where $\eta\in(0.5,1]$ is a suitable threshold chosen according to practical requirements. ($\underline{apr}_P(A)$, $\overline{apr}_P(A)$) is called a DFRS.

Dubois and Prade studied a more general framework in which a fuzzy set was approximated in a fuzzy approximation space [1], for conciseness, this model is named general fuzzy rough sets (GFRS). Similarly, one can define GFRS under dynamic granulation.

Definition 5. Let $S=(U,P)$ be a fuzzy information system, A be a fuzzy set of U and $P=\{R_1,R_2,\cdots,R_n\}$ be a family of fuzzy attribute sets with $R_1 \succ R_2 \succ \cdots \succ R_n$, where $R_k(k=1,2\cdots,n)$ is a fuzzy similarity relation. P -upper approximation $\overline{apr}_P(A)$ and P -lower approximation $\underline{apr}_P(A)$ of A are defined as:

$$\mu_{\overline{apr}_P(A)}(x)=\sup_{y\in U}\min\{\mu_{[x]_{R_n}}(y),\mu_A(y)\},$$

$$\mu_{\underline{apr}_P(A)}(x)=\begin{cases}\inf_{y\in U}\max\{1-\mu_{[x]_{R_1}}(y),\mu_A(y)\}, x\in W_1\\ \cdots\\ \inf_{y\in U}\max\{1-\mu_{[x]_{R_n}}(y),\mu_A(y)\}, x\in W_n\\ \inf_{y\in U}\max\{1-\mu_{[x]_{R_{n+1}}}(y),\mu_A(y)\}, x\in U_{n+1}\end{cases}$$

Where $\mu_{[x]_{R_k}}(y)=\mu_{R_k}(x,y),(k=1,2,\cdots,n)$, $[x]_{R_k}\in U_k/R_k(k=1,2,\cdots,n)$, specially $[x]_{R_{n+1}}\in U_{n+1}/R_n$ and $U_1=U,U_i=U_{i-1}-W_{i-1}(i=2,3,\cdots,n+1)$,

$$W_{i-1}=\{x\mid\mu_{\underline{apr}_{R_{i-1}}(A)}(x)=\inf_{y\in U}\max\{1-\mu_{[x]_{R_{i-1}}}(y),\mu_A(y)\}\ge\eta\}$$

Where $\eta\in(0.5,1]$ is a suitable threshold.

Theorem 4 and Theorem 5 hold true for $\underline{apr}_P(A)$ in Definition 4 and Definition 5.

3.4. An algorithm of mining rules in decision table based on DFRS

In this section, an algorithm is designed to generating fuzzy rules based on DFRS.

Let $S=(U,C\cup D)$ be a decision table, where C is a fuzzy condition attribute set and D is a crisp(or fuzzy) decision attribute set, and $C \cap D = \phi$. The positive region of D with respect to C is defined as: $\mu_{pos_C(D)}(x) = \sup_{A \in U/D} \mu_{\underline{apr}_C(A)}(x)$ .The dependency degree $\gamma_C(D)$ of C with regard to D is defined as: $\gamma_C(D) = \frac{\sum_{x \in U} \mu_{pos_C(D)}(x)}{|U|}$.

Based on DFRS in Definition 4 or Definition 5, we propose an algorithm of mining rules.

Algorithm MRBDFRS (mining rules based on DFRS)

Input: decision table $S = (U, C \cup D)$

Output: decision rules.

(1)For $\forall c \in C$, compute the dependency degree $\gamma_c(D)$ of D to c, let $\gamma_{c_1}(D) = \max\{\gamma_c(D) \mid c \in C\}$ and $P_1 = c_1$;

(2)$U/D = \{A_1, A_2, \cdots, A_d\}$;

(3)Let $P = \{P_1\}, i = 1, U^* = U, Rule' = Rule = \phi$;

(4)Let $W_i = \{x \mid \mu_{\underline{apr}_P(A_j)}(x) = \inf_{y \notin A_j}\{1 - \mu_{[x]_{P_i}}(y)\} \geq \eta, j = 1, 2, \cdots, d\}$. If $W_i \neq \phi$, then for $\forall x \in W_i$, put $des_P(x) \rightarrow des_{A_j}(x)$ $(j = 1, 2, \cdots, d)$ into $Rule'$. Let $Rule = Rule \cup Rule'$ and $U^* = U^* - W_i$;

(5)If $C - P = \phi$ and $U^* \neq \phi$, then for $\forall x \in U^*$, put $des_P(x) \rightarrow des_{A_j}(x)$ $(j = 1, 2, \cdots, d)$ into $Rule$, go to (8);

(6)If $U^* = \phi$, go to (8);

(7)For $\forall c \in C - P$, compute $\gamma_{P \cup \{c\}}(D)$, let $\gamma_{P \cup \{c_2\}}(D) = \max\{\gamma_{P \cup \{c\}}(D) \mid c \in C - P\}$. Let $P_{i+1} = P_i \cup \{c_2\}, P = P \cup P_{i+1}, i = i + 1$,go to (4);

(8)Output $Rule$.

It's clear that the generation of decision rules is based on granulation order P and variational U^*. The time complexity to extract rules is polynomial. At the first step, we need to compute $\gamma_c(D)$ for $\forall c \in C$, the time complexity is $O(|C||U|^2)$. At step 2, the time complexity for computing U/D is $O(|U|^2)$. At step 7, the time complexity for computing $\gamma_{P \cup \{c\}}(D)$ is $O(|C - P_i||C||U|^2)$. From step 5 to step 7, |C| − 1 is the maximum value for the circle times. Therefore, the time complexity of this algorithm is $O(|C|^2|U|^2)$.

3.5. An example

Initial experimentation is carried out by MRBDFRS. The decision table $S = (U, C \cup D)$ comes from Refs.[4] and is given by Table 1, where C is a fuzzy condition attribute set that includes B, G, F, each with corresponding linguistic terms, e.g. B has terms B_1, B_2 and B_3. The decision attribute D is also fuzzy, separated into three linguistic decision X, Y and Z. According to MRBDFRS, we can obtain:

$Rule = \{r_1 : B$ is $B_1 \rightarrow D$ is X,

$r_2 : B$ is B_2 and F is $F_1 \rightarrow D$ is Z;

$r_3 : B$ is B_2 and F is $F_2 \rightarrow D$ is Y;

$r_4 : B$ is B_3 and F is $F_1 \rightarrow D$ is Z;

$r_5 : B$ is B_3 and F is $F_2 \rightarrow D$ is $Y\}$.

4. Comparison with rules induction based on attributes reduction

The existing researches of fuzzy rules induction based on FRS are mainly found in [3-6].The process usually includes two steps: attributes reduction based on FRS and fuzzy rules induction based on conventional rules mining algorithms. A representative work is found in Refs.[4], which developed an algorithm to compute a reduction employing the idea of relative reduction in Pawlak rough sets to keep the dependence degree invariant. The attributes reduction is applied as a preprocessor to an existing fuzzy rule induction algorithm (RIA). The RIA begins with organizing the dataset objects into subgroups according to their highest decision value. Within each subgroup, the fuzzy subsethood between each decision and condition attribute term is calculated. Fuzzy subsethood is defined as follows:

$$S(A, B) = \frac{M(A \cap B)}{M(A)} = \frac{\sum_{u \in U} \min\{\mu_A(u), \mu_B(u)\}}{\sum_{u \in U} \mu_A(u)}.$$

These subsethood values indicate the relatedness between condition and decisions attribute terms. A suitable threshold $\alpha \in [0,1]$ must be chosen in order to determine whether terms are close enough or not. At most, one term is selected per attribute. When there are no suitable terms for a decision, a rule is produced that classifies cases to the rest decision value. In order to entail the learned rules full cover the entire problem domain, this requires another threshold value, $\beta \in [0,1]$, which determines whether a classification is reasonable or not.

In order to compare **MRBDFRS** with method in Refs.[4], the dataset given in Table 1 is reused. Firstly,

Table 1: Decision table with fuzzy attributes

U	B			G			F		D		
	B_1	B_2	B_3	G_1	G_2	G_3	F_1	F_2	X	Y	Z
1	0.3	0.7	0.0	0.2	0.7	0.1	0.3	0.7	0.1	0.9	0.0
2	1.0	0.0	0.0	1.0	0.0	0.0	0.7	0.3	0.8	0.2	0.0
3	0.0	0.3	0.7	0.0	0.7	0.3	0.6	0.4	0.0	0.2	0.8
4	0.8	0.2	0.0	0.0	0.7	0.3	0.2	0.8	0.6	0.3	0.1
5	0.5	0.5	0.0	1.0	0.0	0.0	0.0	1.0	0.6	0.8	0.0
6	0.0	0.2	0.8	0.0	1.0	0.0	0.0	1.0	0.0	0.7	0.3
7	1.0	0.0	0.0	0.7	0.3	0.0	0.2	0.8	0.7	0.4	0.0
8	0.1	0.8	0.1	0.0	0.9	0.1	0.7	0.3	0.0	0.0	1.0
9	0.3	0.7	0.0	0.9	0.1	0.0	1.0	0.0	0.0	0.0	1.0

according to attributes reduction algorithm based on the dependency degree, the attribute G is removed. Using this reduced dataset, the RIA generates the rules as follows[4]:

Rule 1:IF B is B_1 THEN D is X;

Rule 2:IF F is F_2 THEN D is Y;

Rule 3:IF MF(Rule 1) $<\beta$ AND MF(Rule 2) $<\beta$ THEN D is Z,

where MF(Rule i)=MF(condition part of Rule i) and MF means the membership function value.

By comparing above rules with the ones in section 3.5, one can easy see that the Rule 1 is the same as Rule r_1 (in section 3.5), the others are different. It seems as if Rule 2 and Rule 3 were more simple than rules in section 3.5, however, a conflict may generate by using Rule 1 and Rule 2. For case 4, which satisfies B is B_1, then due to Rule 1, D should be X. At the same time, attribute F is F_2, by using Rule 2, we get D is Y. Hence, an inconsistency produces, which may ascribe to too simple rules rooted in attribute reduction process. Again, for case 7, the same situation happens.

Most significant, however, is the fact that the complexity of computing the dependency degree in **MRBDFRS** is less than the one in attributes reduction algorithm due to the universe dwindles gradually. For larger dataset, the effect can be expected to be greater.

5. Conclusions

Today, the grand challenge is to generate useful rules from a mass of data in a database for users to make decisions. This paper has presented such an approach for fuzzy rules induction, which different from most known fuzzy rules induction, is not based on attributes reduction but DFRS. The experimentation shows the algorithm is effective. The generality of this approach should enable it to be applied to broad domains.

Acknowledgment

This work has been supported by the national natural science foundation of China (No.60475019), the research fund for the doctoral program of higher education (No.20060247039).

References

[1] Dubois,D., Prade,H.. Putting Rough Sets and Fuzzy Sets Together. In: Slowinski, R. (ed.): Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets. Kluwer Academic Publishers, Boston Dordrecht London (1992).

[2] Radzikowska,A.M., Kerre, E.E.. A comparative study of fuzzy rough sets.Fuzzy Sets and Systems, 126(2002) 137-155.

[3] Wang, Y.F.. Mining stock price using fuzzy rough set system. Expert Systems with Applications, 24 (2003) 13–23.

[4] Shen,Q., Jensen,R.. Selecting informative features with fuzzy-rough sets and its application for complex systems monitoring. Pattern Recognition, 37(2004)1351-1363.

[5] Rajen,B.B., Gopal,M.. On fuzzy-rough sets approach to feature selection. Pattern Recognition Letters, 26(2005)965-975.

[6]Miao,D.Q., Li,D.G., Fan,S.D.. Fuzzy rough set and its improvement. Proceedings of The First IEEE International Conference on Granular Computing. Beijing, China, IEEE Press, 2005.1247-1251.

[7] Yao,Y.Y.. Combination of rough and fuzzy sets based on $\alpha -$ level sets. Rough Sets and Data Mining: Analysis for Imprecise Data. Kluwer Academic Publishers,1997, 301-321.

[8] Liang,J.Y., Qian,Y.H.,etc..Rough set approximation based on dynamic granulation. D.Slezak et al.(Eds.): RSFDGrC 2005, LNAI 3641,(2005)701-708.