Expert Systems with Applications 37 (2010) 2644–2653

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Mining from incomplete quantitative data by fuzzy rough sets

Tzung-Pei Hong [a,*], Li-Huei Tseng [b], Been-Chian Chien [c]

[a] Department of Computer Science and Information Engineering, National University of Kaohsiung, Kaohsiung 811, Taiwan, ROC
[b] Institute of Computer Science and Information Engineering, I-Shou University, Kaohsiung 840, Taiwan, ROC
[c] Department of Computer Science and Information Engineering, National University of Tainan, Tainan 700, Taiwan, ROC

ARTICLE INFO

Keywords:
Rough set
Machine learning
Fuzzy certain rule
Fuzzy possible rule
Incomplete data
Quantitative data

ABSTRACT

Machine learning can extract desired knowledge from existing training examples and ease the development bottleneck in building expert systems. Most learning approaches derive rules from complete data sets. If some attribute values are unknown in a data set, it is called incomplete. Learning from incomplete data sets is usually more difficult than learning from complete data sets. In the past, the rough-set theory was widely used in dealing with data classification problems. Most conventional mining algorithms based on the rough-set theory identify relationships among data using crisp attribute values. Data with quantitative values, however, are commonly seen in real-world applications. In this paper, we thus deal with the problem of learning from incomplete quantitative data sets based on rough sets. A learning algorithm is proposed, which can simultaneously derive certain and possible fuzzy rules from incomplete quantitative data sets and estimate the missing values in the learning process. Quantitative values are first transformed into fuzzy sets of linguistic terms using membership functions. Unknown attribute values are then assumed to be any possible linguistic terms and are gradually refined according to the fuzzy incomplete lower and upper approximations derived from the given quantitative training examples. The examples and the approximations then interact on each other to derive certain and possible rules and to estimate appropriate unknown values. The rules derived can then serve as knowledge concerning the incomplete quantitative data set.

1. Introduction

The rough-set theory was proposed by Pawlak in 1982 (Pawlak, 1982; Pawlak, 1996) and has been used in reasoning and knowledge acquisition for expert systems (Grzymala-Busse, 1988; Orlowska, 1994). It uses the concept of equivalence classes as its basic principle. Several applications and extensions of the rough-set theory have been proposed. Examples are Orlowska's reasoning with incomplete information (Orlowska, 1994), Germano and Alexandre's knowledge-base reduction (Germano & Alexandre, 1996), Lingras and Yao's data mining (Lingras & Yao, 1998), Zhong et al.'s rule discovery (Zhong, Dong, Ohsuga, & Lin, 1998). Because of the success of the rough-set theory in knowledge acquisition, many researchers in the database and machine-learning fields are very interested in this new research topic since it offers opportunities to discover useful information in training examples.

In the past, most learning approaches derive rules from complete data sets. If some attribute values are unknown in a data set, it is called incomplete. Several methods were proposed to handle the problem of incomplete data sets (Chmielewski, Grzymala-Busse, Peterson, & Than, 1993; Hong & Tseng, 1997; Klir & Folger, 1992; Slowinski & Stefanowski, 1989; Slowinski & Stefanowski, 1994). For example, incomplete data sets may first be transformed into complete data sets (such as by similarity measure) before learning programs begin (Chmielewski et al., 1993), objects with unknown values may be directly removed from data sets (Chmielewski et al., 1993), or unknown objects may be processed in a particular way (Kryszkiewicz, 1998; Liang & Xu, 2000).

Besides, training data in real-world applications sometimes consist of quantitative values, so designing a sophisticated learning algorithm able to deal with quantitative data sets presents a challenge to workers in this research field. Recently, fuzzy-set concepts have often been used to represent quantitative data expressed in linguistic terms and membership functions in intelligent systems because of its simplicity and similarity to human reasoning (Graham & Jones, 1988; Hong & Chen, 1999; Hong, Wang, & Wang, 2000; Hong, Kuo, & Chi, 1999). They have been applied to many fields such as manufacturing, engineering, diagnosis, and economics (Zadeh, 1988; Ziarko, 1993; Zimmermann, 1991; Zimmermann, 1987). Dubois and Prade combined rough sets and fuzzy sets together in order to get a more accurate account of imperfect information (Dubois & Prade, 1992). They built up a very good theoretic basis for fuzzy rough sets. Also, Nakamura predefined similarity matrices and used fuzzy rough sets to logic reasoning (Nakamura, 1992).

* Corresponding author.
E-mail addresses: tphong@nuk.edu.tw (T.-P. Hong), m893331m@isu.edu.tw (L.-H. Tseng), bcchien@mail.nutn.edu.tw (B.-C. Chien).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.08.002

In this paper, we thus deal with the problem of producing a set of certain and possible fuzzy rules from incomplete quantitative data. We combine the rough-set theory and the fuzzy-set concepts to solve this problem. A new generalized fuzzy learning algorithm based on the fuzzy incomplete equivalence classes is proposed to simultaneously derive certain and possible fuzzy rules from incomplete quantitative data sets and estimate the missing values in the learning process. Quantitative values are first transformed into fuzzy sets of linguistic terms using membership functions. Unknown attribute values are then assumed to be any possible linguistic terms and are gradually refined according to the fuzzy incomplete lower and upper approximations derived from the given quantitative training examples. The examples and the approximations then interact on each other to derive certain and possible fuzzy rules and to estimate appropriate unknown values. Rule effectiveness for future data is also derived from these membership values.

The remainder of this paper is organized as follows. The rough-set theory is reviewed in Section 2. The related fuzzy-set concepts are introduced in Section 3. Kryszkiewicz's approach for managing incomplete data sets is described in Section 4. The notation and definitions used in this paper are described in Section 5. A novel learning algorithm to simultaneously induce fuzzy rules and estimate unknown values from incomplete quantitative data sets is proposed in Section 6. An example to illustrate the proposed algorithm is given in Section 7. Conclusion and future work are finally given in Section 8.

2. Review of the rough-set theory

The rough-set theory, proposed by Pawlak in 1982 (Pawlak, 1982; Pawlak, 1996), can serve as a new mathematical tool for dealing with data classification problems. It adopts the concept of equivalence classes to partition training instances according to some criteria. Two kinds of partitions are formed in the mining process: lower approximations and upper approximations, from which certain and possible rules can easily be derived.

Formally, let U be a set of training examples (objects), A be a set of attributes describing the examples, C be a set of classes, and V_j be a value domain of an attribute A_j. Also let $v_j^{(i)}$ be the value of attribute A_j for the ith object $Obj^{(i)}$. When two objects $Obj^{(i)}$ and $Obj^{(k)}$ have the same value of attribute A_j, (that is, $v_j^{(i)} = v_j^{(k)}$), $Obj^{(i)}$ and $Obj^{(k)}$ are said to have an indiscernibility relation (or an equivalence relation) on attribute A_j. Also, if $Obj^{(i)}$ and $Obj^{(k)}$ have the same values for each attribute in subset B of $A, Obj^{(i)}$ and $Obj^{(k)}$ are also said to have an indiscernibility (equivalence) relation on attribute set B. These equivalence relations thus partition the object set U into disjoint subsets, denoted by U/B, and the partition including $Obj^{(i)}$ is denoted by $B(Obj^{(i)})$. The set of equivalence classes for subset B is referred to as *B-elementary set*.

Example 1. Table 1 shows a data set containing seven objects denoted by $U = \{Obj^{(1)}, Obj^{(2)}, \ldots, Obj^{(7)}\}$, two attributes denoted by A = {*Systolic Pressure* (*SP*), *Diastolic Pressure* (*DP*)}, and a class set *Blood Pressure* (*BP*). Assume the attributes and the class set have three possible values: {*Low* (*L*), *Normal* (*N*), *High* (*H*)}.

Table 1
The data set for Example 1.

Object	*Systolic Pressure* (*SP*)	*Diastolic Pressure* (*DP*)	*Blood Pressure* (*BP*)
$Obj^{(1)}$	*L*	*N*	*L*
$Obj^{(2)}$	*H*	*N*	*H*
$Obj^{(3)}$	*N*	*N*	*N*
$Obj^{(4)}$	*L*	*L*	*L*
$Obj^{(5)}$	*H*	*H*	*H*
$Obj^{(6)}$	*N*	*H*	*H*
$Obj^{(7)}$	*N*	*L*	*N*

Since $Obj^{(1)}$ and $Obj^{(4)}$ have the same attribute value (*L*) for attribute *SP*, they share an indiscernibility relation and thus belong to the same equivalence class for *SP*. The equivalence partitions (elementary sets) for singleton attributes can be derived as follows:

$$U/\{SP\} = \{\{Obj^{(2)}, Obj^{(5)}\}\{Obj^{(3)}, Obj^{(6)}, Obj^{(7)}\}\{Obj^{(1)}, Obj^{(4)}\}\}, \text{ and}$$
$$U/\{DP\} = \{\{Obj^{(1)}, Obj^{(2)}, Obj^{(3)}\}\{Obj^{(4)}, Obj^{(7)}\}\{Obj^{(5)}, Obj^{(6)}\}\}.$$
$$\text{Also, } \{SP\}(Obj^{(1)}) = \{SP\}(Obj^{(4)}) = \{Obj^{(1)}, Obj^{(4)}\}.$$

The rough-set approach analyzes data according to two basic concepts, namely the lower and the upper approximations of a set. Let X be an arbitrary subset of the universe U, and B be an arbitrary subset of attribute set A. The lower and the upper approximations for B on X, denoted $B_*(X)$ and $B^*(X)$ respectively, are defined as follows:

$$B_*(X) = \{x | x \in U, B(x) \subseteq X\}, \text{ and}$$
$$B^*(X) = \{x | x \in U \text{ and } B(x) \cap X \neq \emptyset\}.$$

Elements in $B_*(x)$ can be classified as members of set X with full certainty using attribute set B, so $B_*(x)$ is called the lower approximation of X. Similarly, elements in $B^*(x)$ can be classified as members of the set X with only partial certainty using attribute set B, so $B^*(x)$ is called the upper approximation of X.

Example 2. Continuing from Example 1, assume $X = \{Obj^{(1)}, Obj^{(4)}\}$. The lower and the upper approximations of attribute *DP* with respect to X can be calculated as follows:

$$DP_*(X) = \emptyset, \text{ and}$$
$$DP^*(X) = \{\{Obj^{(1)}, Obj^{(2)}, Obj^{(3)}\}\{Obj^{(4)}, Obj^{(7)}\}\}.$$

After the lower and the upper approximations have been found, the rough-set theory can then be used to derive both certain and uncertain information and induce certain and possible rules from them (Grzymala-Busse, 1988).

3. Review of the related fuzzy-set concepts

The fuzzy-set theory was first proposed by Zadeh in 1965 (Zadeh, 1988). It is primarily concerned with quantifying and reasoning using natural language in which words can have ambiguous meanings. This can be thought of as an extension of traditional crisp sets in which each element must either be in or not in a set.

Formally, the process by which individuals from a universal set X are determined to be either members or non-members of a crisp set can be defined by a *characteristic or discrimination function* (Zadeh, 1988). For a given crisp set A, this function assigns a value $\mu_A(x)$ to every $x \in X$ such that:

$$\mu_A(x) = \begin{cases} 1 & \text{if and only if } x \in A \\ 0 & \text{if and only if } x \notin A \end{cases}$$

This kind of functions can be generalized such that the values assigned to the elements of the universal set fall within specified ranges, referred to as the membership grades of these elements in the set, with larger values denoting higher degrees of set membership. Such a function is called the membership function, $\mu_A(x)$, by which a fuzzy set A is usually defined. This function is represented by:

$$\mu_A : X \to [0, 1],$$

where [0, 1] denotes the interval of real numbers from 0 to 1, inclusive. The function can also be generalized to any real interval and is not restricted to [0, 1].

A special notation is often used in the literature to represent fuzzy sets. Assume that x_1 to x_n are the elements in fuzzy set A, and μ_1 to μ_n are, respectively, their grades of membership in A. A is then usually represented as follows:

$$A = \mu_1/x_1 + \mu_2/x_2 + \cdots + \mu_n/x_n.$$

An *α-cut* of a fuzzy set A is a crisp set A_α that contains all elements in the universal set X with membership grades in A greater than or equal to a specified value of α. This definition can be written as:

$$A_\alpha = \{x \in X | \mu_A(x) \geqslant \alpha\}.$$

The *scalar cardinality* of a fuzzy set A defined on a finite universal set X is the summation of the membership grades of all the elements of X in A. Thus,

$$|A| = \sum_{x \in X} \mu_A(x).$$

Among operations on fuzzy sets are the basic and commonly used *complementation*, *union* and *intersection*, as proposed by Zadeh. They are defined as follows.

(1) The complementation of a fuzzy set A is denoted by $\neg A$, and the membership function of $\neg A$ is given by:

$$\mu_{\neg A}(x) = 1 - \mu_A(x), \quad \forall x \in X.$$

(2) The intersection of two fuzzy sets A and B is denoted by $A \cap B$, and the membership function of $A \cap B$ is given by:

$$\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\}, \quad \forall x \in X.$$

(3) The union of fuzzy sets A and B is denoted by $A \cup B$, and the membership function of $A \cup B$ is given by:

$$\mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\}, \quad \forall x \in X.$$

The above fuzzy operations are used in the proposed learning algorithm to find linguistic certain and possible rules.

4. Incomplete data sets

Data sets can be roughly classified into two classes: complete and incomplete data sets. All the objects in a complete data set have known attribute values. If at least one object in a data set has a missing value, the data set is incomplete. Table 2 shows an example of an incomplete data set.

Table 2
An incomplete data set.

Object	*Systolic Pressure (SP)*	*Diastolic Pressure (DP)*	*Blood Pressure (BP)*
$Obj^{(1)}$	*L*	*N*	*N*
$Obj^{(2)}$	*H*	*L*	*H*
$Obj^{(3)}$	*N*	*H*	*N*
$Obj^{(4)}$	*L*	*L*	*L*
$Obj^{(5)}$	$*$	*H*	*H*
$Obj^{(6)}$	*N*	*H*	*H*
$Obj^{(7)}$	*L*	$*$	*L*
$Obj^{(8)}$	*L*	*H*	*N*
$Obj^{(9)}$	$*$	*N*	*H*

In Table 2, the symbol '$*$' denotes an unknown attribute value. Thus, the *SP* values of $Obj^{(5)}$ and $Obj^{(9)}$ are unknown. Similarly, the *DP* value of $Obj^{(7)}$ is unknown. The data set is thus incomplete.

Learning rules from incomplete data sets is usually more difficult than from complete data sets. Designing a sophisticated learning algorithm able to deal with incomplete data sets thus presents a challenge in this research field. In the past, several methods were proposed to handle the problem of incomplete data sets (Chmielewski et al., 1993; Slowinski & Stefanowski, 1989; Slowinski & Stefanowski, 1994). For example, incomplete data sets may be transformed into complete data sets by similarity measures or by removing objects with unknown values before learning programs begin (Chmielewski et al., 1993). Incomplete data sets may also be directly processed in a particular way to get the rules (Kryszkiewicz, 1998; Liang & Xu, 2000).

Kryszkiewicz proposed a rough-set approach to directly learn rules from incomplete data sets without guessing unknown attribute values (Kryszkiewicz, 1998). They defined a similarity relation between objects for attribute subset A as follows:

$$SIM(A) = \{(x, y) \in O \times O | a \in A, a(x) = a(y) \text{ or } a(x) = * \text{ or } a(y) = *\},$$

where O is the object set. It means that two objects have a similarity relation for attribute subset A if they have the same attribute values of A except for unknown values. For each object x, the decisions (classes) of the objects having a similarity relation with x are then collected to form a generalized decision, which is thought of as a new decision (class) in the learning process. The lower approximation and the upper approximation are then derived from the generalized decisions (classes). Rules possibly with disjunctive conclusions are then derived from the lower and the upper approximations.

Liang and Xu modified Kryszkiewicz's approach by introducing the rough entropy to distinguish the power of the attribute subsets that have the same partition for similarity relations. The attribute subset with the minimum rough entropy is selected for partition in incomplete information systems.

In this paper, we propose a new learning algorithm based on rough sets to simultaneously derive certain and possible fuzzy rules from incomplete quantitative data sets and estimate the missing values.

5. Notation and definitions

The following notation is used in this paper.

U	the universe of all objects
n	the total number of objects in U
$Obj^{(i)}$	the ith object, $1 \leqslant i \leqslant n$
A	the set of all attributes describing U
m	the total number of attributes in A
B	an arbitrary subset of A
A_j	the jth attribute, $1 \leqslant j \leqslant m$
$\|A_j\|$	the number of possible attribute values for A_j
R_{jk}	the kth fuzzy region of A_j, $1 \leqslant k \leqslant \|A_j\|$
R_B^k	the kth fuzzy region combination of B
$v_j^{(i)}$	the value of A_j for $Obj^{(i)}$
$f_j^{(i)}$	the fuzzy set converted from $v_j^{(i)}$
$f_{jk}^{(i)}$	the membership value of $v_j^{(i)}$ in region R_{jk}
C	the set of classes to be determined
$\|C\|$	the total number of classes in C
x_l	the lth class, $1 \leqslant l \leqslant \|C\|$
$*$	a missing attribute value
c	a symbol attached to a certain object
u	a symbol attached to an uncertain object

$B(Obj^{(i)})$ the fuzzy incomplete equivalence classes in which $Obj^{(i)}$ exists
$B_k^c(Obj^{(i)})$ the certain part of the kth fuzzy incomplete equivalence class in $B(Obj^{(i)})$
$B_*(X)$ the fuzzy incomplete lower approximation for B on X
$B^*(X)$ the fuzzy incomplete upper approximation for B on X

When the same linguistic term R_{jk} of an attribute A_j exists in two fuzzy objects $Obj^{(i)}$ and $Obj^{(r)}$ with membership values $f_{jk}^{(i)}$ and $f_{jk}^{(r)}$ larger than zero, $Obj^{(i)}$ and $Obj^{(r)}$ are said to have a fuzzy indiscernibility relation (or fuzzy equivalence relation) on attribute A_j with a membership value equal to $min\left(f_{jk}^{(i)} \cap f_{jk}^{(r)}\right)$. Also, if the same linguistic terms of an attribute subset B exist in both $Obj^{(i)}$ and $Obj^{(r)}$ with membership values larger than zero, $Obj^{(i)}$ and $Obj^{(r)}$ are said to have a fuzzy indiscernibility relation (or a fuzzy equivalence relation) on attribute subset B with a membership value equal to the minimum of all the membership values. These fuzzy equivalence relations thus partition the fuzzy object set U into several fuzzy subsets that may overlap, and the result is denoted by U/B. The set of fuzzy partitions, based on B and including $Obj^{(i)}$, is denoted $B(Obj^{(i)})$. Thus, $B(Obj^{(i)}) = \{((B_1(Obj^{(i)}), \mu_{B_1}(Obj^{(i)})), \ldots, ((B_r(Obj^{(i)}), \mu_{B_r}(Obj^{(i)}))\}$, where r is the number of partitions included in $B(Obj^{(i)}), B_j(Obj^{(i)})$ is the jth partition in $B(Obj^{(i)})$, and $\mu_{B_j}(Obj^{(i)})$ is the membership value of the jth partition.

Since an incomplete quantitative data set contains unknown attribute values, each object $Obj^{(i)}$ is thus represented as a tuple $(Obj^{(i)}, symbol)$, where the symbol may be *certain* (c) or *uncertain* (u). If an object $Obj^{(i)}$ has an uncertain value for attribute A_j, then $(Obj^{(i)}, u)$ is put in each fuzzy equivalence class of attribute A_j. The fuzzy sets formed in this way are called fuzzy incomplete equivalence classes, which are not necessarily equivalence classes. The above definition of fuzzy incomplete equivalence classes for single attributes can easily be extended to attribute subsets. The set of fuzzy incomplete equivalence classes for subset B is referred to as *B-elementary fuzzy set*.

Example 3. Consider the following three incomplete fuzzy objects shown in Table 3. Assume the linguistic terms in the objects are transformed from quantitative values by membership functions. $Obj^{(1)}$ has a normal systolic pressure with a membership value of 0.1 and a high systolic pressure with a membership value of 0.75. $Obj^{(1)}$ has also a normal diastolic pressure with a membership value of 0.4 and a high diastolic pressure with a membership value of 0.8. Furthermore, $Obj^{(1)}$ is classified as having a high blood pressure. $Obj^{(2)}$ and $Obj^{(3)}$ are classified similarly, but $Obj^{(2)}$ has a missing value of attribute *SP*.

Assume the attribute SP has three possible linguistic terms $\{L, H, N\}$. Three fuzzy incomplete equivalence classes are then formed. Each object with a symbol c or u is put into appropriate incomplete equivalence classes. Take the fuzzy incomplete equivalence class for the linguistic term (N) as an example. The linguistic term (N) for attribute *SP* appears in $Obj^{(1)}$ and $Obj^{(3)}$, they thus have a fuzzy indiscernibility relation on the fuzzy term *SP.N* and thus form a fuzzy equivalence class with a membership value of min(0.1, 0.3). Also, since $Obj^{(2)}$ has a missing value for attribute *SP*, it is then put into each fuzzy incomplete equivalence class for *SP* with symbol u. The fuzzy incomplete equivalence class from

Table 3
The three incomplete fuzzy objects.

Object	*Systolic Pressure (SP)*	*Diastolic Pressure (DP)*	*Blood Pressure (BP)*
$Obj^{(1)}$	$0.1/N + 0.75/H$	$0.4/N + 0.8/H$	H
$Obj^{(2)}$	$*$	$0.16/N + 0.6/H$	H
$Obj^{(3)}$	$0.5/L + 0.3/N$	$0.4/N + 0.3/L$	L

$SP = N$ is then formed as $\{(Obj^{(1)}, c)(Obj^{(3)}, c)(Obj^{(2)}, u), 0.1\}$. The other fuzzy incomplete indiscernibility relations can be similarly derived. $U/\{SP\}$ has thus been found as follows:

$$U/\{SP\} = \{\{(Obj^{(1)}, c)(Obj^{(3)}, c)(Obj^{(2)}, u), 0.1\},\\ \{(Obj^{(1)}, c)(Obj^{(2)}, u), 0.75\}, \{(Obj^{(3)}, c)(Obj^{(2)}, u), 0.5\}\}.$$

Similarly,

$$U/\{DP\} = \{\{(Obj^{(1)}, c)(Obj^{(2}, c)(Obj^{(3)}, c), 0.16),\\ \{(Obj^{(1)}, c)(Obj^{(2)}, c), 0.6\}, \{(Obj^{(3)}, c), 0.3\}\}.$$

It is easily observed that an object may exist in more than one fuzzy incomplete equivalence class of an attribute. In the above example, $Obj^{(1)}$ exists in two fuzzy incomplete equivalence classes for attribute *SP*. Also, $SP(Obj^{(1)})$ represents the fuzzy incomplete equivalence classes in which $Obj^{(1)}$ exists. Thus:

$$SP(Obj^{(1)}) = \{\{(Obj^{(1)}, c)(Obj^{(3)}, c)(Obj^{(2)}, u), 0.1\},\\ \{(Obj^{(1)}, c)(Obj^{(2)}, u), 0.75\}\}.$$

$B_k(Obj^{(i)})$ represents the kth fuzzy incomplete equivalence class in $B(Obj^{(i)})$. $B_k^c(Obj^{(i)})$ then represents the certain part of $B_k(Obj^{(i)})$. In the above example, $SP_1^c(Obj^{(1)})$ includes $(Obj^{(1)}, c)$ and $(Obj^{(3)}, c)$.

The fuzzy incomplete lower and upper approximations for B on X, denoted $B_*(X)$ and $B^*(X)$ respectively, are defined as follows:

$$B_*(X_l) = \{(B_k(Obj^{(i)}), \mu_{B_k}(Obj^{(i)}))|1 \leqslant i \leqslant n, obj^{(i)} \in X_l,\\ B_k^c(Obj^{(i)}) \subseteq X_l, 1 \leqslant k \leqslant |B(Obj^{(i)})|\},$$

$$B^*(X_l) = \{(B_k(Obj^{(i)}), \mu_{B_k}(Obj^{(i)}))|1 \leqslant i \leqslant n, B_k^c(Obj^{(i)}) \cap X_l \neq \emptyset,\\ B_k^c(Obj^{(i)}) \not\subset X_l, 1 \leqslant k \leqslant |B(Obj^{(i)})|\}.$$

Here, the definition of the fuzzy incomplete upper approximation has the constraint $B_k^c(Obj^{(i)}) \not\subset X$ to exclude the objects in the fuzzy incomplete lower approximation for avoiding redundant calculation.

Example 4. Continuing from Example 3, assume $X = \{Obj^{(1)}, Obj^{(2)}\}$. Since only the certain part of the second fuzzy incomplete equivalence class in $U/\{SP\}$ is included in X, the fuzzy incomplete lower approximation for attribute *SP* on X is thus:

$$SP_*(X) = \{((Obj^{(1)}, c)(Obj^{(2)}, u), 0.75)\}.$$

The certain parts of the first and second incomplete equivalence class in $U/\{SP\}$ have non-empty intersections with X. Since the second equivalence class has been included in the fuzzy incomplete lower approximation, the fuzzy incomplete upper approximation for attribute SP on X is thus:

$$SP^*(X) = \{((Obj^{(1)}, c)(Obj^{(3)}, c)(Obj^{(2)}, u), 0.1)\}.$$

The fuzzy incomplete lower and upper approximations for attribute *DP* on X can be similarly derived.

Elements in $B_*(x)$ can be classified as members of set X with full certainty using attribute set B. Also, the membership values of the fuzzy incomplete lower approximations may be considered effectiveness measures for future data. A low membership value with a fuzzy incomplete lower approximation means the lower approximation will have a low tolerance (or effectiveness) on future data. In this case, the fuzzy partitions from the fuzzy incomplete lower approximation have a high probability to be removed when future data are considered. All of the partitions are, however, valid for the current data set and can be used to correctly classify its elements.

On the other hand, elements in $B^*(x)$ can be classified as members of set X with only partial certainty using attribute set B, and their certainty degrees can be calculated from the membership values of elements in the upper approximations.

After the fuzzy lower and the fuzzy upper approximations have been found, certain and uncertain information can be analyzed, and fuzzy rules can then be derived.

6. The proposed algorithm for incomplete quantitative data sets

In the section, a learning algorithm based on rough sets is proposed, which can simultaneously estimate the missing values and derive fuzzy certain and possible rules from incomplete quantitative data sets. The proposed fuzzy learning algorithm first transforms each quantitative value into a fuzzy set of linguistic terms using membership functions. As mentioned before, each object is represented as a tuple $((Obj^{(i)}), symbol)$, where the symbol may be certain (c) or uncertain (u). If the object has a missing value of an attribute, it is first put into each fuzzy incomplete equivalence class from that attribute.

The algorithm then calculates fuzzy incomplete lower approximations and tries to estimate missing quantitative values from them. If an uncertain object exists in only one fuzzy incomplete equivalence class in a fuzzy incomplete lower approximation, its unknown values can easily be estimated by the values representing the equivalence class, and its symbol is changed into c to reflect the estimation; otherwise, the value estimation is postponed until fuzzy incomplete lower approximations from more attributes can determine it. If an unknown object still exists in more than one fuzzy incomplete equivalence class after all attributes are processed, it is heuristically assigned. Note that any value heuristically assigned will not actually affect the certain rules derived.

Next, the algorithm calculates fuzzy incomplete upper approximations and tries to estimate the other quantitative missing values from them. Similarly, if an uncertain object exists in only one fuzzy incomplete equivalence class in a fuzzy incomplete upper approximation, its unknown values can easily be estimated by the values representing the equivalence class, and its symbol is changed into c to reflect the estimation; otherwise, the value estimation is postponed until fuzzy incomplete upper approximations from more attributes can determine it. If an unknown object still exists in more than one fuzzy incomplete equivalence class after all attributes are processed, it is heuristically assigned. The details of the proposed fuzzy learning algorithm are described as follows.

The algorithm:

Input: An incomplete quantitative data set U with n objects, each of which has m attribute values and belongs to one of a class set C, and a set of membership functions.

Output: A set of certain and possible fuzzy rules.

Step 1: Partition the object set into disjoint subsets according to class labels. Denote each set of objects belonging to the same class C_l as X_l.

Step 2: Transform the quantitative value $v_j^{(i)}$ of each object $Obj^{(i)}, i = 1$ to n, for each attribute A_j, $j = 1$ to m, into a fuzzy set $f_j^{(i)}$, represented as $\left(\frac{f_{j1}^{(i)}}{R_{j1}} + \frac{f_{j2}^{(i)}}{R_{j2}} + \cdots + \frac{f_{jl}^{(i)}}{R_{jl}}\right)$, using the given membership functions, where R_{jk} is the kth fuzzy region of attribute A_j, $f_{jk}^{(i)}$ is $v_j^{(i)}$'s fuzzy membership value in region R_{jk}, and $l\ (= |A_j|)$ is the number of fuzzy regions for A_j. If $Obj^{(i)}$ has a missing value for A_j, keep it with a missing value $(*)$.

Step 3: Find the fuzzy incomplete elementary sets of singleton attributes; That is, if an object $Obj^{(i)}$ has a certain fuzzy membership value $f_{jk}^{(i)}$ for attribute A_j, put $(Obj^{(i)}, c)$ into the fuzzy incomplete equivalence class from $A_j = R_{jk}$; If $Obj^{(i)}$ has a missing value for A_j, put $(Obj^{(i)}, u)$ into each fuzzy incomplete equivalence class from A_j; The membership value $\mu_{A_{jk}}$ of a fuzzy incomplete class for $A_j = R_{jk}$ is calculated as:

$$\mu_{A_{jk}} = \underset{i}{Min} f_{jk}^{(i)},$$

where $Obj^{(i)}$ is certain and $f_{jk}^{(i)} \neq 0$.

Step 4: Initialize $q = 1$, where q is used to count the number of attributes currently being processed for fuzzy incomplete lower approximations.

Step 5: Compute the fuzzy incomplete lower approximations of each subset B with q attributes for each class X_l as:

$$B_*(X_l) = \{(B_k(Obj^{(i)}), \mu_{B_k}(Obj^{(i)}))| 1 \leqslant i \leqslant n, obj^{(i)} \in X_l, B_k^c(Obj^{(i)}) \subseteq X_l, 1 \leqslant k \leqslant |B(Obj^{(i)})|\},$$

where $B(Obj^{(i)})$ is the set of fuzzy incomplete equivalence classes including $Obj^{(i)}$ and derived from attribute subset $B, B_k^c(Obj^{(i)})$is the certain part of the kth fuzzy incomplete equivalence class in $B(Obj^{(i)})$.

Step 6: Do the following substeps for each uncertain instance $Obj^{(i)}$ in the fuzzy incomplete lower approximations:

(a) If $Obj^{(i)}$ exists in only one fuzzy incomplete equivalence class $B_k^c(Obj^{(i)})$ of the kth region combination R_B^k from attribute subset B in a fuzzy incomplete lower approximation, assign the uncertain value of $Obj^{(i)}$ as:

$$\frac{\sum_{Obj^{(r)} \in B_k^c(Obj^{(i)})} v_j^{(r)} \times f_{jk}^{(r)}}{\sum_{Obj^{(r)} \in B_k^c(Obj^{(i)})} f_{jk}^{(r)}},$$

where $v_j^{(r)}$ is the quantitative value of $Obj^{(r)}$ for attribute A_j and $f_{jk}^{(r)}$ is $v_j^{(r)}$'s fuzzy membership value in R_B^k. Also transform the estimated $Obj^{(i)}$ value into a fuzzy set, remove $(Obj^{(i)}, u)$'s with membership values equal to zero from the fuzzy incomplete equivalence classes, change $(Obj^{(i)}, u)$'s with membership values not equal to zero into $(Obj^{(i)}, c)$'s and re-calculate the membership values of the fuzzy incomplete equivalence classes including them by the minimum operation. Besides, backtrack to the previously found fuzzy incomplete lower approximations for doing the same actions on $Obj^{(i)}$.

(b) If $Obj^{(i)}$ exists in more than one fuzzy incomplete equivalence class in an fuzzy incomplete lower approximation from attribute subset B, postpone the estimation of its uncertain value until more attributes can determine them.

Step 7: Set $q = q + 1$ and repeat Steps 5–7 until $q > m$.

Step 8: If an object $Obj^{(i)}$ still exists in more than one fuzzy incomplete equivalence class in a fuzzy incomplete lower approximation, use the equivalence class with the maximum scalar cardinality for certain objects to estimate the uncertain values of $Obj^{(i)}$. The estimation and processing are the same as those stated in Step 6(a).

Step 9: Derive the certain fuzzy rules from the fuzzy incomplete lower approximation of each subset B, and set the membership values of equivalence classes in the lower approximation as effectiveness measures for future data.

Step 10: Remove certain fuzzy rules with condition parts more specific and effectiveness measure equal to or smaller than those of some other certain fuzzy rules.

Step 11: Reset $q = 1$, where q is used to count the number of attributes currently being processed for fuzzy incomplete upper approximations.

Step 12: Compute the fuzzy incomplete upper approximations of each subset B with q attributes for each class X_l as:

$$B^*(X_l) = \{(B_k(Obj^{(i)}), \mu_{B_k}(Obj^{(i)}))|1 \leqslant i \leqslant n, B_k^c(Obj^{(i)}) \cap X_l \neq \varnothing, B_k^c(Obj^{(i)}) \not\subset X_l, 1 \leqslant k \leqslant |B(Obj^{(i)})|\},$$

where $B(Obj^{(i)})$ is the set of fuzzy incomplete equivalence classes including $Obj^{(i)}$ and derived from attribute subset B, $B_k^c(Obj^{(i)})$ is the certain part of the kth fuzzy incomplete equivalence class in $B(Obj^{(i)})$.

Step 13: Do the following substeps for each uncertain instance $Obj^{(i)}$ in the fuzzy incomplete upper approximations:

(a) If $Obj^{(i)}$ exists in only one fuzzy incomplete equivalence class $B_k^c(Obj^{(i)})$ of the kth region combination R_B^k from attribute subset B in a fuzzy incomplete upper approximation, assign the uncertain value of $Obj^{(i)}$ as:

$$\frac{\sum_{\substack{Obj^{(r)} \in B_k^c(Obj^{(i)}) \\ \& Obj^{(r)} \in X_l}} v_j^{(r)} \times f_{jk}^{(r)}}{\sum_{\substack{Obj^{(r)} \in B_k^c(Obj^{(i)}) \\ \& Obj^{(r)} \in X_l}} f_{jk}^{(r)}},$$

where $v_j^{(r)}$ is the quantitative value of $Obj^{(r)}$ for attribute A_j and $f_{jk}^{(r)}$ is $v_j^{(r)}$'s fuzzy membership value in R_B^k.Also transform the estimated $Obj^{(i)}$ value into a fuzzy set, remove $(Obj^{(i)}, u)$'s with membership values equal to zero from the fuzzy incomplete equivalence classes, change $(Obj^{(i)}, u)$'s with membership values not equal to zero into $(Obj^{(i)}, c)$'s and re-calculate the membership values of the fuzzy incomplete equivalence classes including them by the minimum operation. Besides, backtrack to the previously found fuzzy incomplete upper approximations for doing the same actions on $Obj^{(i)}$.

(b) If $Obj^{(i)}$ exists in more than one fuzzy incomplete equivalence class in a fuzzy incomplete upper approximation from attribute subset B, postpone the estimation of its uncertain value until more attributes can determine them.

Step 14: Set $q = q + 1$ and repeat Steps 12–14 until $q > m$.

Step 15: Calculate the plausibility measures of each fuzzy incomplete equivalence class in an upper approximation for each class X_l as:

$$P\left(B_k^c(Obj^{(i)})\right) = \frac{\sum_{\substack{Obj^{(r)} \in B_k^c(Obj^{(i)}) \\ \& Obj^{(r)} \in X_l}} f_{jk}^{(r)}}{\sum_{Obj^{(r)} \in B_k^c(Obj^{(i)})} f_{jk}^{(r)}},$$

where $f_{jk}^{(r)}$ is the fuzzy membership value of the quantitative value of $Obj^{(r)}$ for attribute A_j in R_B^k.

Step 16: If an object $Obj^{(i)}$ still exists in more than one fuzzy incomplete equivalence class in a fuzzy incomplete upper approximation, use the equivalence class with the maximum plausibility measure to estimate the uncertain value of $Obj^{(i)}$. The estimation and processing are the same as those stated in Step 13(a).

Step 17: Derive the possible fuzzy rules from the fuzzy incomplete upper approximation of each subset B, with the plausibility measure recalculated due to the estimated objects. Besides, set the membership values of equivalence classes in the upper approximation as effectiveness measures for future data.

Step 18: Remove possible fuzzy rules with condition parts more specific and both the effectiveness measure and plausibility equal to or smaller than those of some other possible fuzzy rules or certain fuzzy rules.

Step 19: Output the certain and possible fuzzy rules.

The fuzzy rules output after Step 19 can then sever as meta-knowledge concerning the given incomplete quantitative data set. Also, the missing values in the data set are derived at the same time in the learning process.

7. An example

In this section, an example is given to show how the proposed algorithm can be used to generate certain and possible fuzzy rules from incomplete quantitative data. Table 4 shows an incomplete quantitative data set, which is similar to that shown in Table 1 except that the data attributes are represented as incomplete quantitative values.

Assume the membership functions for each attribute are given by experts as shown in Fig. 1.

The proposed learning algorithm processes this incomplete quantitative data set as follows.

Step 1: Since three classes exist in the data set, three partitions are formed as follows:

$$X_L = \{Obj^{(4)}, Obj^{(7)}\},$$
$$X_N = \{Obj^{(1)}, Obj^{(3)}, Obj^{(8)}\}, \text{ and}$$
$$X_H = \{Obj^{(2)}, Obj^{(5)}, Obj^{(6)}, Obj^{(9)}\}.$$

Step 2: The quantitative values of each object are transformed into fuzzy sets. Take the attribute *Systolic Pressure* (*SP*) in $Obj^{(2)}$ as an example. The value "155" is converted into a fuzzy set $(0.1/N + 0.75/H)$ using the given membership functions. Results for all the objects are shown in Table 5. The missing values are not transformed.

Step 3: Since the attribute *SP* has three possible linguistic terms, $\{L, H, N\}$, three fuzzy incomplete equivalence classes are formed. The fuzzy incomplete elementary set of attribute *SP* is found as follows:

$$U/\{SP\} = \{((Obj^{(2)}, c)(Obj^{(5)}, u)(Obj^{(9)}, u), 0.75), ((Obj^{(2)}, c)(Obj^{(3)}, c)(Obj^{(6)}, c)(Obj^{(7)}, c)(Obj^{(8)}, c)(Obj^{(5)}, u)(Obj^{(9)}, u), 0.1), ((Obj^{(1)}, c)(Obj^{(4)}, c)(Obj^{(7)}, c)(Obj^{(8)}, c)(Obj^{(5)}, u)(Obj^{(9)}, u), 0.5)\}.$$

Similarly, the fuzzy incomplete elementary set of attribute *DP* is found as follows:

Table 4
An incomplete quantitative data set as an example.

Object	*Systolic Pressure* (*SP*)	*Diastolic Pressure* (*DP*)	*Blood Pressure* (*BP*)
$Obj^{(1)}$	80	80	*N*
$Obj^{(2)}$	155	70	*H*
$Obj^{(3)}$	130	92	*N*
$Obj^{(4)}$	87	68	*L*
$Obj^{(5)}$	*	93	*H*
$Obj^{(6)}$	140	100	*H*
$Obj^{(7)}$	95	*	*L*
$Obj^{(8)}$	95	93	*N*
$Obj^{(9)}$	*	78	*H*

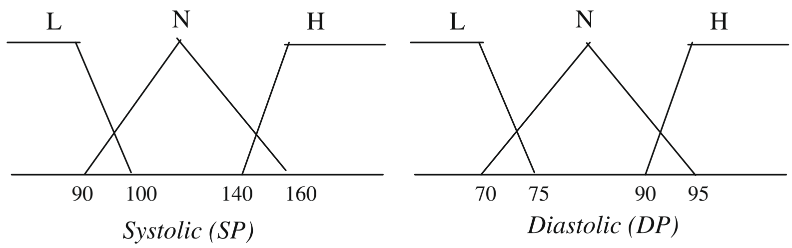

Fig. 1. The given membership functions of each attribute.

$$U/\{DP\} = \{((Obj^{(2)}, c)(Obj^{(4)}, c)(Obj^{(7)}, u), 1), ((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c)(Obj^{(8)}, c)(Obj^{(7)}, u), 0.4), ((Obj^{(1)}, c)(Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c)(Obj^{(7)}, u), 0.2)\}.$$

Step 4: q is initially set at 1, where q is used to count the number of the attributes currently being processed for fuzzy incomplete lower approximations.

Step 5: The fuzzy incomplete lower approximation of one attribute for $X_H(\{Obj^{(2)}, Obj^{(5)}, Obj^{(6)}, Obj^{(9)}\})$ is first calculated. Since only the certain part $(Obj^{(2)}, c)$ of the first incomplete equivalence class for attribute SP is included in X_H and the uncertain instances $Obj^{(5)}$ and $Obj^{(9)}$ belong to X_H, thus:

$$SP_*(X_H) = \{\{(Obj^{(2)}, c)(Obj^{(5)}, u)(Obj^{(9)}, u), 0.75\}\}.$$

Since the certain part of each fuzzy incomplete equivalence class for attribute DP is not included in X_H, thus:

$$DP_*(X_H) = \emptyset.$$

Similarly, the fuzzy incomplete lower approximations of single attributes for X_N and X_L are found as follows:

$SP_*(X_N) = \emptyset$,
$DP_*(X_N) = \emptyset$,
$SP_*(X_L) = \emptyset$, and
$SP_*(X_L) = \emptyset$.

Step 6: Each uncertain object in the above fuzzy incomplete lower approximations is checked for change to certain objects. For example in $SP_*(X_H)$, since $Obj^{(5)}$ and $Obj^{(9)}$ exist in only one fuzzy incomplete equivalence class of $SP = H$, their values can then be estimated from the certain objects in the same equivalence class. Since only one certain object $Obj^{(2)}$ exists in the fuzzy incomplete equivalence class of $SP = H$, the estimated value of $Obj^{(5)}$ is then $(155 * 0.75)/0.75$ (=155), where 155 is the quantitative value of $Obj^{(2)}$ for attribute SP and 0.75 is its fuzzy membership value for the region of $SP = H$. Similarly, the estimated value of $Obj^{(9)}$ is 155.
The estimated values of $Obj^{(5)}$ and $Obj^{(9)}$ are then transformed as the fuzzy set $(0.1/N + 0.75/H)$. $(Obj^{(5)}, u)$ and $(Obj^{(9)}, u)$ are then changed as $(Obj^{(5)}, c)$ and $(Obj^{(9)}, c)$. The modified $SP_*(X_H)$ is then:

$$SP_*(X_H) = \{((Obj^{(2)}, c)(Obj^{(5)}, c)(Obj^{(9)}, c), 0.75)\}.$$

The membership value of each fuzzy incomplete equivalence class including $Obj^{(5)}$ and $Obj^{(9)}$ is then re-calculated by the minimum operation. The fuzzy incomplete elementary set of attribute SP is then modified as:

$$U/\{SP\} = \{((Obj^{(2)}, c)(Obj^{(5)}, c)(Obj^{(9)}, c), 0.75), ((Obj^{(2)}, c)(Obj^{(3)}, c)(Obj^{(6)}, c)(Obj^{(7)}, c)(Obj^{(8)}, c)(Obj^{(5)}, c)(Obj^{(9)}, c), 0.1), ((Obj^{(1)}, c)(Obj^{(4)}, c)(Obj^{(7)}, c)(Obj^{(8)}, c), 0.5)\}.$$

The incomplete quantitative data set is then modified as shown in Table 6.

Step 7: $q = q + 1 = 2$, and Steps 5–7 are repeated. The fuzzy incomplete elementary set of attributes {SP, DP} is found as follows:

$$U/\{SP, DP\} = \{((Obj^{(1)}, c)(Obj^{(8)}, c)(Obj^{(7)}, u), 0.2), ((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c)(Obj^{(7)}, u), 0.1), ((Obj^{(5)}, c)(Obj^{(9)}, c), 0.2), ((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c)(Obj^{(8)}, c)(Obj^{(7)}, u), 0.1), ((Obj^{(4)}, c)(Obj^{(7)}, u), 0.5), ((Obj^{(5)}, c), 0.6), ((Obj^{(8)}, c)(Obj^{(7)}, u), 0.5), ((Obj^{(2)}, c)(Obj^{(7)}, u), 0.1), ((Obj^{(2)}, c), 0.75)\}.$$

The fuzzy incomplete lower approximations of {SP, DP} for the three classes are found as:

$$SP, DP_*(X_H) = \{((Obj^{(2)}, c), 0.75), ((Obj^{(2)}, c), 0.1), ((Obj^{(5)}, c)(Obj^{(9)}, c), 0.2), ((Obj^{(5)}, c), 0.6)\},$$
$$SP, DP_*(X_N) = \{((Obj^{(1)}, c)(Obj^{(8)}, c), 0.2), ((Obj^{(8)}, c), 0.5)\}, \text{ and}$$
$$SP, DP_*(X_L) = \{((Obj^{(4)}, c)(Obj^{(7)}, u), 0.5)\}.$$

Since the uncertain object $Obj^{(7)}$ in $SP, DP_*(X_L)$ exists in only the fuzzy incomplete equivalence class of $SP = L$ and $DP = L$, the estimated value of $Obj^{(7)}$ is then $(68 * 1)/1$ (=68), where 68 is the quantitative value of $Obj^{(4)}$ for attribute DP and 1 is its fuzzy membership value of $DP = L$. The estimated value of $Obj^{(7)}$ is then transformed as the fuzzy set $(1/L)$ for attribute DP. Also, $(Obj^{(7)}, u)$ is

Table 5
The fuzzy sets transformed from the data in Table 4.

Object	Systolic Pressure (SP)	Diastolic Pressure (DP)	Blood Pressure (BP)
$Obj^{(1)}$	$1/L$	$0.7/N$	N
$Obj^{(2)}$	$0.1/N + 0.75/H$	$1/L$	H
$Obj^{(3)}$	$0.75/N$	$0.3/N + 0.4/H$	N
$Obj^{(4)}$	$1/L$	$1/L$	L
$Obj^{(5)}$	*	$0.2/N + 0.6/H$	H
$Obj^{(6)}$	$0.5/N$	$1/H$	H
$Obj^{(7)}$	$0.5/L + 0.2/N$	*	L
$Obj^{(8)}$	$0.5/L + 0.2/N$	$0.2/N + 0.6/H$	N
$Obj^{(9)}$	*	$0.5/N$	H

Table 6
The modified incomplete quantitative data set.

Object	Systolic Pressure (SP)	Diastolic Pressure (DP)	Blood Pressure (BP)
$Obj^{(1)}$	80	80	N
$Obj^{(2)}$	155	70	H
$Obj^{(3)}$	130	92	N
$Obj^{(4)}$	87	68	L
$Obj^{(5)}$	155	93	H
$Obj^{(6)}$	140	100	H
$Obj^{(7)}$	95	*	L
$Obj^{(8)}$	95	93	N
$Obj^{(9)}$	155	78	H

then changed as $(Obj^{(7)}, c)$. The modified $SP, DP_*(X_L)$ is then:

$$SP, DP_*(X_L) = \{\{(Obj^{(4)}, c)(Obj^{(7)}, c)\}, 0.5\}.$$

The membership value of each fuzzy incomplete equivalence class including $Obj^{(7)}$ is then re-calculated by the minimum operation. The fuzzy incomplete elementary set of attributes *SP* and *DP* is then modified as:

$$\begin{aligned} U/\{SP, DP\} = \{&((Obj^{(1)}, c)(Obj^{(8)}, c), 0.2), \\ &((Obj^{(2)}, c)(Obj^{(7)}, c), 0.1), \\ &((Obj^{(2)}, c), 0.75), \\ &((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c), 0.1), \\ &((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c)(Obj^{(8)}, c), 0.1), \\ &((Obj^{(4)}, c)(Obj^{(7)}, c)), 0.5\}, \\ &\{(Obj^{(5)}, c)(Obj^{(9)}, c), 0.2\}, \\ &((Obj^{(5)}, c), 0.6), \\ &((Obj^{(8)}, c), 0.5)\}. \end{aligned}$$

The incomplete quantitative data set is then modified as shown in Table 7.
All unknown values in the data set have been estimated after this step. The fuzzy incomplete elementary set of attributes *DP* is then backtracked and modified as:

$$\begin{aligned} U/\{DP\} = \{&((Obj^{(2)}, c)(Obj^{(4)}, c)(Obj^{(7)}, c), 1), ((Obj^{(3)}, c) \\ &(Obj^{(5)}, c)(Obj^{(6)}, c)(Obj^{(8)}, c), 0.4), ((Obj^{(1)}, c) \\ &(Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c), 0.2)\}. \end{aligned}$$

Step 8: Since all objects in the fuzzy incomplete lower approximations have become certain, the next step is executed.
Step 9: The certain fuzzy rule derived from the fuzzy incomplete lower approximation of *SP* is:

1. If *Systolic Pressure = High* Then *Blood Pressure = High*, *with future effectiveness* = 0.75.
   The certain fuzzy rules from the fuzzy incomplete lower approximation of attributes {*SP*, *DP*} are:
2. If *Systolic Pressure = High* and *Diastolic Pressure = Normal* Then *Blood Pressure = High*, *with future effectiveness* = 0.2.
3. If *Systolic Pressure = High* and *Diastolic Pressure = High* Then *Blood Pressure = High*, *with future effectiveness* = 0.6.
4. If *Systolic Pressure = High* and *Diastolic Pressure = Low* Then *Blood Pressure = High*, *with future effectiveness* = 0.75.
5. If *Systolic Pressure = Normal* and *Diastolic Pressure = Low* Then *Blood Pressure = High*, *with future effectiveness* = 0.1.

Table 7
The modified incomplete quantitative data set.

Object	Systolic Pressure (SP)	Diastolic Pressure (DP)	Blood Pressure(BP)
$Obj^{(1)}$	80	80	N
$Obj^{(2)}$	155	70	H
$Obj^{(3)}$	130	92	N
$Obj^{(4)}$	87	68	L
$Obj^{(5)}$	155	93	H
$Obj^{(6)}$	140	100	H
$Obj^{(7)}$	95	68	L
$Obj^{(8)}$	95	93	N
$Obj^{(9)}$	155	78	H

6. If *Systolic Pressure = Low* and *Diastolic Pressure = High* Then *Blood Pressure = Normal*, *with future effectiveness* = 0.5.
7. If *Systolic Pressure = Low* and *Diastolic Pressure = Normal* Then *Blood Pressure = Normal*, *with future effectiveness* = 0.2.
8. If *Systolic Pressure = Low* and *Diastolic Pressure = Low* Then *Blood Pressure = Low*, *with future effectiveness* = 0.5.

Step 10: Since the condition parts and the effectiveness measures of the certain rules 2, 3 and 4 are more specific and smaller than those of the first rule, the three certain rules are removed from the certain rule set.
Step 11: q is reset to 1, where q is used to count the number of attributes currently being processed for fuzzy incomplete upper approximations.
Step 12: The fuzzy incomplete upper approximations of single attributes for the three classes are calculated from the modified fuzzy incomplete elementary sets. Thus:

$$\begin{aligned} SP^*(X_H) = \{&((Obj^{(2)}, c)(Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c) \\ &(Obj^{(7)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c), 0.1)\}, \\ DP^*(X_H) = \{&((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c)(Obj^{(8)}, c), 0.4), \\ &((Obj^{(1)}, c)(Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c), 0.2), \\ &((Obj^{(2)}, c)(Obj^{(4)}, c)(Obj^{(7)}, c), 1)\}, \\ SP^*(X_N) = \{&((Obj^{(1)}, c)(Obj^{(4)}, c)(Obj^{(7)}, c)(Obj^{(8)}, c), 0.5), \\ &((Obj^{(2)}, c)(Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c)(Obj^{(7)}, c) \\ &(Obj^{(8)}, c)(Obj^{(9)}, c), 0.1)\}, \\ DP^*(X_N) = \{&((Obj^{(1)}, c)(Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c), 0.2), \\ &((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c)(Obj^{(8)}, c), 0.4)\}, \\ SP^*(X_L) = \{&((Obj^{(1)}, c)(Obj^{(4)}, c)(Obj^{(7)}, c)(Obj^{(8)}, c), 0.5), \\ &((Obj^{(2)}, c)(Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c) \\ &(Obj^{(7)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c), 0.1)\}, \text{ and} \\ DP^*(X_L) = \{&((Obj^{(2)}, c)(Obj^{(4)}, c)(Obj^{(7)}, c), 1)\}. \end{aligned}$$

Step 13: Since no uncertain objects exist in the above fuzzy incomplete upper approximations, the next step is done.
Step 14: $q = q + 1 = 2$, and Steps 12–14 are repeated. The modified fuzzy incomplete upper approximations of {*SP*, *DP*} for the three classes are found as:

$$\begin{aligned} SP, DP^*(X_H) = \{&((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c), 0.1), \\ &((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c)(Obj^{(8)}, c), 0.1), \\ &((Obj^{(2)}, c)(Obj^{(7)}, c), 0.1)\}, \end{aligned}$$

$$SP, DP^*(X_N) = \{((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c), 0.1), ((Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(6)}, c)(Obj^{(8)}, c), 0.1)\},$$

and

$$SP, DP^*(X_L) = \{((Obj^{(2)}, c)(Obj^{(7)}, c), 0.1)\}.$$

Step 15: The plausibility measures of the fuzzy incomplete equivalence classes in the fuzzy incomplete upper approximations are calculated. For example, the plausibility measure of the fuzzy incomplete equivalence class $\{(Obj^{(3)}, c)(Obj^{(5)}, c)(Obj^{(8)}, c)(Obj^{(9)}, c)\}$ for class X_H is calculated as:

$$p(SP, DP^c_H(Obj^{(3)} \text{ or } Obj^{(5)} \text{ or } Obj^{(8)} \text{ or } Obj^{(9)})) = \frac{0.1 + 0.1}{0.3 + 0.1 + 0.2 + 0.1} = 0.29.$$

Step 16: Since all the objects in the fuzzy incomplete upper approximations are certain, the next step is executed.

Step 17: The possible fuzzy rules derived from the fuzzy incomplete upper approximations of *SP* and *DP* are then derived as follows:

1. If *Systolic Pressure* = *Normal* Then *Blood Pressure* = *High with plausibility* = 0.41, *with future effectiveness* = 0.1.
2. If *Diastolic Pressure* = *Normal* Then *Blood Pressure* = *High with plausibility* = 0.37, *with future effectiveness* = 0.2.
3. If *Diastolic Pressure* = *Low* Then *Blood Pressure* = *High with plausibility* = 0.33, *with future effectiveness* = 1.
4. If *Diastolic Pressure* = *High* Then *Blood Pressure* = *High with plausibility* = 0.62, *with future effectiveness* = 0.4.
5. If *Systolic Pressure* = *Normal* Then *Blood Pressure* = *Normal with plausibility* = 0.49, *with future effectiveness* = 0.1.
6. If *Systolic Pressure* = *Low* Then *Blood Pressure* = *Normal with plausibility* = 0.5, *with future effectiveness* = 0.5.
7. If *Diastolic Pressure* = *Normal* Then *Blood Pressure* = *Normal with plausibility* = 0.63, *with future effectiveness* = 0.2.
8. If *Diastolic Pressure* = *High* Then *Blood Pressure* = *Normal with plausibility* = 0.38, *with future effectiveness* = 0.4.
9. If *Systolic Pressure* = *Normal* Then *Blood Pressure* = *Low with plausibility* = 0.1, *with future effectiveness* = 0.1.
10. If *Systolic Pressure* = *Low* Then *Blood Pressure* = *Low with plausibility* = 0.5, *with future effectiveness* = 0.5.
11. If *Diastolic Pressure* = *Low* Then *Blood Pressure* = *Low with plausibility* = 0.67, *with future effectiveness* = 1.

    The possible fuzzy rules derived from the upper approximations of attributes {*SP, DP*} are:

12. If *Systolic Pressure* = *Normal* and *Diastolic Pressure* = *Normal* Then *Blood Pressure* = *High with plausibility* = 0.29, *with future effectiveness* = 0.1.
13. If *Systolic Pressure* = *Normal* and *Diastolic Pressure* = *Low* Then *Blood Pressure* = *High with plausibility* = 0.33, *with future effectiveness* = 0.1.
14. If *Systolic Pressure* = *Normal* and *Diastolic Pressure* = *High* Then *Blood Pressure* = *High with plausibility* = 0.5, *with future effectiveness* = 0.1.
15. If *Systolic Pressure* = *Normal* and *Diastolic Pressure* = *Normal* Then *Blood Pressure* = *Normal with plausibility* = 0.71, *with future effectiveness* = 0.1.
16. If *Systolic Pressure* = *Normal* and *Diastolic Pressure* = *Low* Then *Blood Pressure* = *Low with plausibility* = 0.67, *with future effectiveness* = 0.1.
17. If *Systolic Pressure* = *Normal* and *Diastolic Pressure* = *High* Then *Blood Pressure* = *Normal with plausibility* = 0.5, *with future effectiveness* = 0.1.

Step 18: Since the condition parts, plausibility measures and effectiveness measures of the possible fuzzy rules 12, 13, 14 and 16 are more specific and smaller than those of the fuzzy rules 1, 2, 4 and 11, rules 12, 13, 14 and 16 are thus removed from the possible fuzzy rule set.

Step 19: All the certain fuzzy rules and possible fuzzy rules are then output as knowledge about the given fuzzy incomplete quantitative data set.

8. Conclusions and future work

In this paper, we have proposed a new learning approach to derive fuzzy rules from incomplete quantitative data sets based on the rough-set theory. The proposed approach is different from others in that it can derive fuzzy rules and estimate the missing quantitative values at the same time. The fuzzy incomplete lower and upper approximations have been defined for managing uncertain objects in fuzzy incomplete data sets. The interaction between data and approximations helps derive certain and possible rules from fuzzy incomplete data sets and estimate appropriate unknown values. An example has been given to illustrate the proposed algorithm in details. The fuzzy rules derived in this way can then serve as fuzzy knowledge concerning incomplete data sets.

In addition to unknown attribute values, incorrect attribute values are also commonly seen in real-world applications. Although possible fuzzy rules can still be generated from an incorrect data set, other approaches can be adopted to reduce the bad effects. Ziarko proposed the variable precision rough set model (Ziarko, 1993) to allow for a controlled degree of misclassification. One aspect of our future research is thus to extend our method with Ziarko's model for managing unknown and incorrect data sets.

References

Chmielewski, M. R., Grzymala-Busse, J. W., Peterson, N. W., & Than, S. (1993). The rule induction system LERS – A version for personal computers. *Foundations of Computing and Decision Sciences, 18*, 181–212.

Dubois, D., & Prade, H. (1992). Putting rough sets and fuzzy sets together. *Intelligent Decision Support Handbook of Applications and Advances of the Rough Sets Theory*, 203–232.

Germano, L. T., & Alexandre, P. (1996). Knowledge-base reduction based on rough set techniques. *The Canadian Conference on Electrical and Computer Engineering*, 278–281.

Graham, I., & Jones, P. L. (1988). *Expert systems – Knowledge, uncertainty and decision*. Boston: Chapman and Computing. pp. 117–158.

Grzymala-Busse, J. W. (1988). Knowledge acquisition under uncertainty: A rough set approach. *Journal of Intelligent Robotic Systems, 1*, 3–16.

Hong, T. P., & Chen, J. B. (1999). Finding relevant attributes and membership functions. *Fuzzy Sets and Systems, 103*(3), 389–404.

Hong, T. P., Kuo, C. S., & Chi, S. C. (1999). Mining association rules from quantitative data. *Intelligent Data Analysis, 3*(5), 363–376.

Hong, T. P., & Tseng, S. S. (1997). A generalized version space learning algorithm for noisy and uncertain data. *IEEE Transactions on Knowledge and Data Engineering, 9*(2), 336–340.

Hong, T. P., Wang, T. T., & Wang, S. L. (2000). Knowledge acquisition from quantitative data using the rough-set theory. *Intelligent Data Analysis, 4*, 289–304.

Klir, G. J., & Folger, T. A. (1992). *"Fuzzy sets", uncertainty and information*. New Jersey: Prentice Hall. pp. 4–14.

Kryszkiewicz, M. (1998). Rough set approach to incomplete information systems. *Information Science, 112*, 39–49.

Liang, J., & Xu, Z. (2000). Uncertainty measures of roughness of knowledge and rough sets in incomplete information systems. In *The third world congress on intelligent control and automation* (pp. 2526–2529).

Lingras, P. J., & Yao, Y. Y. (1998). Data mining using extensions of the rough set model. *Journal of the American Society for Information Science, 49*(5), 415–422.

Nakamura, A. (1992). Applications of fuzzy-rough classifications to logics Intelligent Decision Support. *Handbook of Applications and Advances of the Rough Sets Theory*, 233–250.

Orlowska, E. (1994). *Reasoning with incomplete information: Rough set based information logics. Incompleteness and uncertainty in information systems*. New York: Springer-Verlag. pp. 16–33.

Pawlak, Z. (1982). Rough set. *International Journal of Computer and Information Sciences*, 341–356.
Pawlak, Z. (1996). Why rough sets? *The Fifth IEEE International Conference on Fuzzy Systems, 2*, 738–743.
Slowinski, R., & Stefanowski, J. (1989). Rough classification in incomplete information systems. *Mathematical and Computer Modelling, 12*, 1347–1357.
Slowinski, R., & Stefanowski, J. (1994). Handling various types of uncertainly in the rough set approach. In W. Ziarko (Ed.), *Rough sets fuzzy sets and knowledge discovery (RSKD'93)*. Berlin: Springer.
Zadeh, L. A. (1988). Fuzzy logic. *IEEE Computer*, 83–93.
Zhong, N., Dong, J. Z., Ohsuga, S., & Lin, T. Y. (1998). An incremental probabilistic rough set approach to rule discovery. *The IEEE International Conference on Fuzzy Systems, 2*, 933–938.
Ziarko, W. (1993). Variable precision rough set model. *Journal of Computer and System Sciences, 46*, 39–59.
Zimmermann, H. J. (1987). *Fuzzy sets decision making and expert systems*. Boston: Kluwer Academic Publishers.
Zimmermann, H. J. (1991). *Fuzzy set theory and its applications*. Boston: Kluwer Academic Publisher.