Learning Fuzzy Rules from Incomplete Quantitative Data by Rough Sets

Tzung-Pei Hong†, Li-Huei Tseng‡ and Been-Chian Chien‡

†Department of Electrical Engineering
National University of Kaohsiung
Kaohsiung, 811, Taiwan, R.O.C.

‡Graduate School of Information Engineering
I-Shou University
Kaohsiung, 840, Taiwan, R.O.C.

Abstract - In this paper, we deal with the problem of learning from incomplete quantitative data sets based on rough sets. 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.

I. INTRODUCTION

The rough-set theory was proposed by Pawlak in 1982 [16][17] and has been used in reasoning and knowledge acquisition for expert systems [5][15]. 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 [15], Germano and Alexandre's knowledge-base reduction [3], Lingras and Yao's data mining [13], Zhong et al.'s rule discovery [21]. 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 [1][18][19].

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 [4][6][7]. They have been applied to many fields such as manufacturing, engineering, diagnosis, and economics [20][22][24]. Dubois and Prade combined rough sets and fuzzy sets together in order to get a more accurate account of imperfect information [2]. 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 [14].

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. Rule effectiveness for future data is also derived from the membership values.

II. REVIEW OF THE ROUGH SET THEORY

The rough-set theory, proposed by Pawlak in 1982 [16][17], 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 $V_i$ be a value domain of an attribute $A_i$. Also let $v^{(i)}_j$ be the value of attribute $A_i$ for the $i$-th object $Obj^{(i)}$. When two objects $Obj^{(i)}$ and $Obj^{(k)}$ have the same value of attribute $A_i$ (that is, $v^{(i)}_j = v^{(k)}_j$), $Obj^{(i)}$ and $Obj^{(k)}$ are said to have an indiscernibility relation (or an equivalence relation) on attribute $A_i$. 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
\(B(Obj^0)\). The set of equivalence classes for subset \(B\) is referred to as \(B\)-elementary set.

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:

\[
B(X) = \{ x | x \in U, B(x) \subseteq X \}, \quad \text{and} \quad B'(X) = \{ x | x \in U \text{ and } B(x) \cap X \neq \emptyset \}.
\]

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 \(X\). 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 [5].

III. 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.

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 [1] [18][19]. 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 [1]. Incomplete data sets may also be directly processed in a particular way to get the rules [11][12].

IV. DEFINITIONS

When the same linguistic term \(R_{jk}\) of an attribute \(A_j\) exists in two fuzzy objects \(Obj^i\) and \(Obj^j\) with membership values \(f_{jk}^i\) and \(f_{jk}^j\) larger than zero, \(Obj^i\) and \(Obj^j\) are said to have a fuzzy indiscernibility relation (or fuzzy equivalence relation) on attribute \(A_j\) with a membership value equal to \(\min( f_{jk}^i, f_{jk}^j )\). Also, if the same linguistic terms of an attribute subset \(B\) exist in both \(Obj^i\) and \(Obj^j\) with membership values larger than zero, \(Obj^i\) and \(Obj^j\) 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 \(UB\). The set of fuzzy partitions, based on \(B\) and including \(Obj^0\), is denoted \(B(Obj^0)\). Thus, \(B(Obj^0) = \{(B(A_j(Obj^0)), \mu_B(Obj^0)), \ldots, (B(A_j(Obj^0)), \mu_B(Obj^0))\}\), where \(r\) is the number of partitions included in \(B(Obj^0)\), \(B(Obj^0)\) is the \(j\)-th partition in \(B(Obj^0)\), and \(\mu_B(Obj^0)\) is the membership value of the \(j\)-th partition.

Since an incomplete quantitative data set contains unknown attribute values, each object \(Obj^0\) is thus represented as a tuple \((Obj^0, \text{symbol})\), where the symbol may be certain (c) or uncertain (u). If an object \(Obj^0\) has an uncertain value for attribute \(A_j\), then \((Obj^0, u)\) 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.

The fuzzy incomplete lower and upper approximations for \(B\) on \(X\), denoted \(B(X)\) and \(B'(X)\) respectively, are defined as follows:

\[
B(X) = \{ (B(A_j(Obj^0)), \mu_B(Obj^0)) : 1 \leq i \leq n, Obj^0 \in X, B(A_j(Obj^0)) \subseteq X_i, 1 \leq k \leq |B(Obj^0)| \},
\]

\[
B'(X) = \{ (B(A_j(Obj^0)), \mu_B(Obj^0)) : 1 \leq i \leq n, B(A_j(Obj^0)) \cap X_i \neq \emptyset, B(A_j(Obj^0)) \subseteq X_i, 1 \leq k \leq |B(Obj^0)| \},
\]

Here, the definition of the fuzzy incomplete upper approximation has the constraint \(B(Obj^0) \subseteq X\) to exclude the objects in the fuzzy incomplete lower approximation for avoiding redundant calculation.

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.

V. 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 certain and possible rules from fuzzy
incomplete quantitative data sets. The proposed fuzzy learning algorithm first transforms each quantitative value into a fuzzy set of linguistic terms using membership functions. The algorithm then calculates fuzzy incomplete lower approximations and tries to estimate missing values from them. Next, the algorithm calculates fuzzy incomplete upper approximations and tries to estimate the remaining missing values from them. The details of the proposed 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 \( C_i \) as \( X_i \).

**STEP 2:** Transform the quantitative value \( v^{(i)} \) 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:

\[
\left( \frac{f_j^{(i)}}{R_k} + \frac{f_j^{(i)}}{R_{k1}} + \ldots + \frac{f_j^{(i)}}{R_{kn}} \right),
\]

using the given membership functions, where \( R_k \) is the \( k \)-th fuzzy region of attribute \( A_j \). \( f_j^{(i)} \) is \( v^{(i)} \)'s fuzzy membership value in region \( R_k \), 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 \( f_j^{(i)} \) for attribute \( A_j \), put \((Obj^{(i)}, c)\) into the fuzzy incomplete equivalence class from \( A_j = R_k \). 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 \( \mu_{A_j} \) of a fuzzy incomplete class for \( A_j = R_k \) is calculated as:

\[
\mu_{A_j} = \text{Min} \ f_j^{(i)} ,
\]

where \( Obj^{(i)} \) is certain and \( f_j^{(i)} \neq 0 \).

**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 \( X_i \) as:

\[
B(\{\text{attributes}\}) = \{B_i(\text{attributes}) \mid 1 \leq i \leq n, B_i(\text{attributes}) \in X_i, B_i(\text{attributes}) \subseteq X_i, 1 \leq k \leq |B(\text{attributes})|\},
\]

where \( B(\text{attributes}) \) is the set of fuzzy incomplete equivalence classes including \( Obj^{(i)} \) and derived from attribute subset \( B \), \( B_i(\text{attributes}) \) is the certain part of the \( k \)-th fuzzy incomplete equivalence class in \( B(\text{attributes}) \).

**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 \( B_i(Obj^{(i)}) \) of the \( k \)-th region combination \( R_k \) from attribute subset \( B \) in a fuzzy incomplete lower approximation, assign the uncertain value of \( Obj^{(i)} \) as:

\[
\frac{\sum_{r \in B_i(Obj^{(i)})} f_{jk}^{(r)} \times v_{jk}^{(r)}}{\sum_{r \in B_i(Obj^{(i)})} f_{jk}^{(r))}} ,
\]

where \( v_{jk}^{(r)} \) is the quantitative value of \( Obj^{(i)} \) for attribute \( A_j \), and \( f_{jk}^{(r)} \) is \( v_{jk}^{(r)} \)'s fuzzy membership value in \( R_k \). 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 a 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 to 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 \( X_i \) as:

\[
B(\{\text{attributes}\}) = \{B_i(\text{attributes}) \mid 1 \leq i \leq n, B_i(\text{attributes}) \in X_i, B_i(\text{attributes}) \subseteq X_i, 1 \leq k \leq |B(\text{attributes})|\},
\]

where \( B(\text{attributes}) \) is the set of fuzzy incomplete equivalence classes including \( Obj^{(i)} \) and derived from attribute subset \( B \), \( B_i(\text{attributes}) \) is the certain part of the \( k \)-th fuzzy incomplete equivalence class in \( B(\text{attributes}) \).

**STEP 13:** Do the following substeps for each uncertain instance \( Obj^{(i)} \) in the fuzzy incomplete upper approximations:
(a) If Obj\textsuperscript{0} exists in only one fuzzy incomplete equivalence class \( B_k^i (\text{Obj}^{(l)}_j) \) of the \( k \)-th region combination \( R_k^i \) from attribute subset \( B \) in a fuzzy incomplete upper approximation, assign the uncertain value of \( \text{Obj}^{(l)}_j \) as:

\[
\sum_{\text{Obj}^{(l)}_j \in B_k^i (\text{Obj}^{(l)}_j) \text{ and } \text{Obj}^{(l)}_j \in X_l} v^{(r)}_{jl} \times f^{(r)}_{jk}
\]

\[
\sum_{\text{Obj}^{(l)}_j \in B_k^i (\text{Obj}^{(l)}_j) \text{ and } \text{Obj}^{(l)}_j \in X_l} f^{(r)}_{jk}
\]

where \( v^{(r)}_{jl} \) is the quantitative value of \( \text{Obj}^{(l)}_j \) for attribute \( A_j \) and \( f^{(r)}_{jk} \) is \( v^{(r)}_{jl} \)'s fuzzy membership value in \( R_k^i \).

Also transform the estimated \( \text{Obj}^{(l)}_j \) value into a fuzzy set, remove \( (\text{Obj}^{(l)}_j, u) \)'s with membership values equal to zero from the fuzzy incomplete equivalence classes, change \( (\text{Obj}^{(l)}_j, u) \)'s with membership values not equal to zero into \( (\text{Obj}^{(l)}_j, c) \)'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 \( \text{Obj}^{(l)}_j \).

(b) If \( \text{Obj}^{(l)}_j \) 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 to 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:

\[
P(B_k^i (\text{Obj}^{(l)}_j)) = \frac{\sum_{\text{Obj}^{(l)}_j \in B_k^i (\text{Obj}^{(l)}_j) \text{ and } \text{Obj}^{(l)}_j \in X_l} f^{(r)}_{jk}}{\sum_{\text{Obj}^{(l)}_j \in B_k^i (\text{Obj}^{(l)}_j) \text{ and } \text{Obj}^{(l)}_j \in X_l} f^{(r)}_{jk}}
\]

where \( f^{(r)}_{jk} \) is the fuzzy membership value of the quantitative value of \( \text{Obj}^{(l)}_j \) for attribute \( A_j \) in \( R_k^i \).

STEP 16: If an object \( \text{Obj}^{(l)}_j \) 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 \( \text{Obj}^{(l)}_j \). 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.

### VI. 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 1 shows an incomplete quantitative data set containing nine objects, two attributes denoted by \( A = \{\text{Systolic Pressure (SP)}, \text{Diastolic Pressure (DP)}\} \), and a class set Blood Pressure (BP). The symbol ‘*’ denotes an unknown attribute value.

<table>
<thead>
<tr>
<th>Object</th>
<th>Systolic Pressure (SP)</th>
<th>Diastolic Pressure (DP)</th>
<th>Blood Pressure (BP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obj\textsuperscript{0} ((1))</td>
<td>90</td>
<td>140</td>
<td>H</td>
</tr>
<tr>
<td>Obj\textsuperscript{0} ((2))</td>
<td>130</td>
<td>70</td>
<td>H</td>
</tr>
<tr>
<td>Obj\textsuperscript{0} ((3))</td>
<td>80</td>
<td>68</td>
<td>L</td>
</tr>
<tr>
<td>Obj\textsuperscript{0} ((4))</td>
<td>95</td>
<td>93</td>
<td>*</td>
</tr>
<tr>
<td>Obj\textsuperscript{0} ((5))</td>
<td>95</td>
<td>78</td>
<td>N</td>
</tr>
</tbody>
</table>

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

![Figure 1: The given membership functions of each attribute](image)

The fuzzy incomplete elementary set of attribute SP is found as follows:

\[
\{ ((\text{Obj}^{(2)}, c)(\text{Obj}^{(5)}, u)(\text{Obj}^{(6)}, u), 0.75), ((\text{Obj}^{(2)}, c)(\text{Obj}^{(3)}, c)(\text{Obj}^{(4)}, c)(\text{Obj}^{(5)}, u)(\text{Obj}^{(6)}, u), 0.1), ((\text{Obj}^{(2)}, c)(\text{Obj}^{(4)}, c)(\text{Obj}^{(5)}, c)(\text{Obj}^{(6)}, c)(\text{Obj}^{(9)}, u)(\text{Obj}^{(9)}, u), 0.5)) \}
\]

The fuzzy incomplete elementary sets for attribute DP
and attribute combination \( \{SP, DP\} \) can be similarly derived. The fuzzy incomplete lower approximations are then derived. For example, \( SP \{X_0\} = \{(O_{b2}^{(2)}, c) (O_{b5}^{(3)}, c) (O_{b9}^{(9)}, c), 0.75\} \).

The proposed learning algorithm then estimates the incomplete quantitative data set as shown in Table 2.

### Table II

<table>
<thead>
<tr>
<th>Object</th>
<th>Systolic Pressure ( (SP) )</th>
<th>Diastolic Pressure ( (DP) )</th>
<th>Blood Pressure ( (BP) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obj'(1)</td>
<td>80</td>
<td>80</td>
<td>N</td>
</tr>
<tr>
<td>Obj'(2)</td>
<td>150</td>
<td>70</td>
<td>H</td>
</tr>
<tr>
<td>Obj'(3)</td>
<td>130</td>
<td>92</td>
<td>N</td>
</tr>
<tr>
<td>Obj'(4)</td>
<td>87</td>
<td>68</td>
<td>L</td>
</tr>
<tr>
<td>Obj'(5)</td>
<td>155</td>
<td>93</td>
<td>H</td>
</tr>
<tr>
<td>Obj'(6)</td>
<td>140</td>
<td>100</td>
<td>H</td>
</tr>
<tr>
<td>Obj'(7)</td>
<td>95</td>
<td>68</td>
<td>L</td>
</tr>
<tr>
<td>Obj'(8)</td>
<td>95</td>
<td>93</td>
<td>N</td>
</tr>
<tr>
<td>Obj'(9)</td>
<td>155</td>
<td>78</td>
<td>H</td>
</tr>
</tbody>
</table>

The final fuzzy incomplete elementary sets are then derived. For example:

\[
U/\{(SP), c\}(\{O_{b2}^{(2)}, c) (O_{b5}^{(3)}, c) (O_{b9}^{(9)}, c), 0.75\}) = \{(O_{b2}^{(2)}, c), (O_{b5}^{(3)}, c), (O_{b9}^{(9)}, c), 0.75\}\.
\]

The final fuzzy incomplete lower and upper approximations can be derived from them. For example, \( SP \{X_0\} = \{(O_{b2}^{(2)}, c) (O_{b5}^{(5)}, c) (O_{b9}^{(9)}, c), 0.75\}\), and \( SP \{X_0\} = \{(O_{b2}^{(2)}, c) (O_{b5}^{(5)}, c) (O_{b9}^{(9)}, c) (O_{b7}^{(9)}, c) (O_{b8}^{(8)}, c) (O_{b9}^{(9)}, c), 0.1\}\).

The certain fuzzy rules and the possible fuzzy rules can then be derived from the fuzzy incomplete lower and upper approximations. For example, the certain fuzzy rule derived from the fuzzy incomplete lower approximation of \( SP \) is:

**If Systolic Pressure = High**

**Then Blood Pressure = High, with future effectiveness = 0.75.**

One of the possible fuzzy rules derived from the fuzzy incomplete upper approximations of \( SP \) is:

**If Systolic Pressure = Normal**

**Then Blood Pressure = High with plausibility = 0.41, with future effectiveness = 0.1.**

### VII. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a new learning approach to derive 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 values at the same time. An example has been given to illustrate the proposed algorithm. The fuzzy rules derived in this way can then serve as knowledge concerning incomplete quantitative 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 [22] 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 quantitative data sets.

**References**

14. A. Nakamura, “Applications of fuzzy-rough classifications to logics,”


