A PROBABILISTIC ROUGH SET APPROACH FOR INCREMENTAL LEARNING KNOWLEDGE ON THE CHANGE OF ATTRIBUTES

DUN LIU†
School of Economics and Management, Southwest Jiaotong University
Chengdu 610031, China

JUNBO ZHANG, TIANRUI LI
School of Information Science and Technology, Southwest Jiaotong University
Chengdu 610031, China

The attribute set in an information system may evolve in time when new information arrives. Approximations of a concept in probabilistic rough set theory are introduced into our studies to discover the usefulness knowledge. Then, the strategies and propositions for incremental learning knowledge are proposed when attributes are changed. An example is given to validate the feasibility of our proposed methods.

1. Introduction

Data grow at an unprecedented rate nowadays. Then the information system evolves over time. Approaches for updating knowledge incrementally are getting more and more popular. How to design the algorithms or strategies for updating knowledge from information systems has become one of hot topics in data mining.

An information system is composed by the objects, the attributes and the domain of attributes’ values. The previous work for updating knowledge from information systems are mainly focused on these three aspects: the change of objects (instances) \(^{[1-4]}\), the change of attributes (features) \(^{[5,6]}\) and the change of attributes’ values \(^{[7,8]}\). All these work may help people to obtain the updating knowledge with different viewpoints.

In this paper, we focus on the second case when the attributes change. The variation of the attributes may affect the granularity of the knowledge space decision rules induced by lower and upper approximations may also alter. In the earlier studies, Chan discussed an incremental approach for updating approximations of a concept when adding or deleting an attribute in a complete system \(^{[5]}\). Li et al. presented a method for updating approximations in an incomplete information system under the

---

1 This work is partially supported by the Chinese National Science Foundation (60873108) and the Doctoral Innovation Fund of Southwest Jiaotong University (200907).
characteristic relation when an attribute set varies over time. However, the errors in the system are not considered in their work. Our study here is to investigate the method for incremental learning knowledge under Probabilistic Rough Sets (PRS).

The remainder of the paper is organized as follows: Section 2 provides the basic concepts of PRS. In Section 3, the related propositions and strategies for incremental learning knowledge in PRS are presented when attributes change. In Section 4, examples are employed to validate the proposed approaches. The paper ends with conclusions and further research topics in Section 5.

2. Preliminaries

Some basic concepts, notations and results of PRS as well as their extensions are briefly reviewed in this section.

A complete information system is defined as a 4-tuple \( S = (U, A, V, f) \), where \( U = \{x_1, x_2, \ldots, x_m\} \) is a non-empty finite set of objects. \( A \) is an attribute set and \( A = C \cup D \). \( C = \{a_1, a_2, \ldots, a_k\} \) is the set of condition attributes and \( D \) is the set of decision attributes. \( V = \bigcup_{a \in A} V_a \), \( V_a \) is a domain of the attribute \( a \). \( f : U \times A \rightarrow V \) is an information function such that \( f(x, a) \in V_a \) for every \( x \in U \), \( a \in A \).

**Definition 2.1** Let \( S = (U, A, V, f) \) be an information system, \( \forall x \in U \), \( X \subseteq U \), let \( \Pr(X \mid \{x\}) = \frac{|X \cap \{x\}|}{|X|} \), where \( |\cdot| \) stands for the cardinal number of objects in sets, \( \Pr(X \mid \{x\}) \) is denoted the conditional probability of the classification and we denote it as \( \alpha \), \( \alpha \in (0, 1] \).

**Definition 2.2** Let \( S = (U, A, V, f) \) be an information system. \( \forall X \subseteq U \) and \( 0 \leq \beta < \alpha \leq 1 \), the \((\alpha, \beta)\)-lower approximation, upper approximation are defined as: \( R^{(\alpha, \beta)}(X) = \{x \in U \mid \Pr(X \mid \{x\}) \geq \alpha\} ; R^{(\alpha, \beta)}(X) = \{x \in U \mid \Pr(X \mid \{x\}) > \beta\} \).

The \((\alpha, \beta)\)-probabilistic positive, boundary and negative regions are as follows.

\( POS^{(\alpha, \beta)}(X) = \{x \in U \mid \Pr(X \mid \{x\}) \geq \alpha\} \);
\( BND^{(\alpha, \beta)}(X) = \{x \in U \mid \beta < \Pr(X \mid \{x\}) < \alpha\} \);
\( NEG^{(\alpha, \beta)}(X) = \{x \in U \mid \Pr(X \mid \{x\}) \leq \beta\} \).

Clearly, the variable precision rough set is a special case of PRS. In addition, the parameters \( \alpha \) and \( \beta \) allow certain acceptable level of errors. Then, it makes the process of decision making more reasonable.

3. Approaches for incremental learning knowledge when the attributes change

The variation of the attributes in an information system may lead to two cases: the increasing of the attributes and the decreasing of the attributes. In the first
case, the lower approximation may expand and the upper approximation may contract because of the specialization of the system; the opposite situation may happen in the second case because of the generalization of the system. Since the decision rules are generated by the upper and lower approximations, the studies of incremental learning knowledge can convert to the studies of the variety of the two approximations when attributes change.

**Definition 3.1** Let \( S = (U, A, V, f) \) be an information system, the lower boundary set and upper boundary set can be denoted as \( \Delta_\pi X = X - \overline{R}_X \) and \( \overline{\Delta}_\pi X = \overline{R}_X - X \), respectively. The boundary set is \( BNG_\pi X = \overline{R}_X - R_X = \overline{\Delta}_\pi X \cup \Delta_\pi X \).

Given the acceptable level parameters \( \alpha \) and \( \beta \), we have \( \overline{R}(X) \subseteq \overline{R}^{(\alpha, \beta)}(X) \), \( \overline{R}^{(\alpha, \beta)}(X) \subseteq \overline{R}(X) \), where \( \overline{R}^{(\alpha, \beta)}(X) = \overline{R}(X) \cup \Delta_\pi^{(\alpha, \beta)} X \), \( \Delta_\pi^{(\alpha, \beta)} X = \{ x \in BNG_\pi X | \Pr(\bigcap_{\alpha \in A} [x], X) \geq \alpha \} \) and \( \overline{R}^{(\alpha, \beta)}(X) = \overline{R}(X) - \Delta_\pi^{(\alpha, \beta)} X \), \( \Delta_\pi^{(\alpha, \beta)} X = \{ x \in BNG_\pi X | \Pr(\bigcap_{\alpha \in A} [x], X) \leq \beta \} \). Note that the variation of \( \Delta_\pi^{(\alpha, \beta)} X \) and \( \overline{\Delta}_\pi^{(\alpha, \beta)} X \) may happen in two different ways, that is, the immigration and emigration of objects in \((\alpha, \beta)\)-lower approximations and upper approximations may happen simultaneously when attributes change. Therefore, the incremental strategy for updating approximations when attributes change may divide into two parts: the first one is the incremental updating of \( \overline{R}(X) \) and \( \overline{R}(X) \), the second one is the incremental updating of \( \Delta_\pi^{(\alpha, \beta)} X \) and \( \overline{\Delta}_\pi^{(\alpha, \beta)} X \).

**Proposition 3.1** Suppose \( P, Q \) are two attribute sets, \( P \cap Q = \emptyset \) and \( A = P \cup Q \), we have \( \Delta^{(\alpha, \beta)} X = P_X \cup Q_X \cup Y \cup Y' \), where \( Y = \{ x \in \Delta_\pi X \cap \Delta_\pi Q X \cap \bigcup_{\alpha \in A} [x] \} \subseteq X \}, \ Y' = \{ x \in \Delta_\pi Q X \cap \bigcup_{\alpha \in A} [x] \} \subseteq X \}, \ \Delta_\pi^{(\alpha, \beta)} X = \{ x \in \bigcup_{\alpha \in A} [x] \} \subseteq X \} \subseteq X \), \( \overline{\Delta}_\pi^{(\alpha, \beta)} X = \{ x \in \bigcup_{\alpha \in A} [x] \} \subseteq X \} \subseteq X \).

**Proof.** On the one hand, due to \( P \subseteq A \) and \( Q \subseteq A \), we have \( P_X \subseteq A_X \) and \( Q_X \subseteq A_X \), that is \( P_X \cup Q_X \subseteq A_X \). In addition, \( Y \subseteq A_X \). So \( P_X \cup Q_X \cup Y \subseteq A_X \). On the other hand, let \( X \subseteq U \), \( x \in U \) and \( x \in A_X \). \( \forall x \in X \), \( x = P_X + \Delta_\pi X = \overline{Q}_X \). Furthermore, due to \( x \in A_X \), \( \bigcup_{\alpha \in A} [x] \subseteq X \), we have \( x \in Y, A_X \subseteq Y \). So \( A_X \subseteq P_X \cup Q_X \cup Y \). Therefore, we have \( A_X = P X \cup Q X \cup Y \). Due to \( \Delta^{(\alpha, \beta)} X = \Delta(X) \cup \Delta^{(\alpha, \beta)} \), we have \( \Delta^{(\alpha, \beta)} X = \).
Proposition 3.2 Suppose \( P', Q' \) are two attribute sets, \( Q' \subseteq P' \). Denote \( B = P' \setminus Q' \), we have \( B^{(a,b)} X \subseteq P'X - \Delta_{P'Q'}X \cup Y^* \), where \( \Delta_{P'Q'}X = \{ x \in \bigcap_{x \in \Delta_{P'Q'}} \Delta_x X \cap \bigcap_{x \in \Delta_{P'Q'}} \{ x \in \Delta_{P'Q'} \cap \Delta_x X \} \} \), \( \Delta_{P'Q'}X = \{ x \in \bigcap_{x \in \Delta_{P'Q'}} \Delta_x X \cap \bigcap_{x \in \Delta_{P'Q'}} \{ x \in \Delta_{P'Q'} \cap \Delta_x X \} \} \). \( Y^* = \{ x \in \Delta_{P'Q'}(X) \cup \Delta_{P'Q'}(X) \cap \Pr(\Delta_{P'Q'}(x), X) \geq \alpha \}. \)

Proof. Suppose \( Q' \) is non-redundant attribute, we have \( BX \subseteq P'X \). Due to \( X = P'X + \Delta_p X = BX + \Delta_p X \), so \( \Delta_p X \subseteq \Delta_p X \). In addition, \( \Delta_p X - \Delta_p X = \{ x \in U \cap x \in \Delta_p X - x \notin \Delta_p X \} \). So \( BX = P'X - (\Delta_p X - \Delta_p X) = P'X - \Delta_p X + \Delta_p X \). Furthermore, due to \( P'X \cap \Delta_p X = \emptyset \), we have \( BX = P'X - \Delta_p X \). Since \( B^{(a,b)} X = B(X) \cup \Delta_{\Delta_p X} \), then \( B^{(a,b)} X = X - \Delta_{P'Q'}X \cup Y^* \).

Specially, if \( Q' \) is \( (\alpha, \beta) \)-redundant attribute, we have \( B^{(a,b)} X = P^{(a,b)} X \). \( \square \)

Proposition 3.3 Suppose \( P, Q \) are two attribute sets, \( P \cap Q = \emptyset \) and \( A = P \cup Q \), we have \( A^{(a,b)} X = X \cup (\Delta_p X - Z) - M_d \) \( Z = \{ x \in \bigcap_{x \in \Delta_p} \Delta_x X \cap \bigcap_{x \in \Delta_p} \{ x \in \Delta_p \cap \Delta_x X \} \} \), \( M_d = \{ x \in \bigcup_{\Delta_p \cup Q} \Delta_x X \cap \Pr(\bigcap_{x \in \Delta_p \cup Q} [x], X) \leq \beta \} \). \( \Delta_p \cup Q X = \{ x \in \bigcap_{x \in \Delta_p \cup Q} \Delta_x X \cap \bigcap_{x \in \Delta_p \cup Q} \{ x \in \Delta_p \cup Q \cap \Delta_x X \} \} \).

Proof. Let \( x \in \Delta X \) and \( x \notin X \). Since \( \Delta X = X + \Delta_x X \), we have \( x \in \Delta_x X \). On the one hand, due to \( \Delta X \subseteq \Delta_p X \), we have \( x \in \Delta_p X \) and \( \bigcap_{x \in \Delta_p \cup Q} \{ x \in \Delta_p \cup Q \cap \Delta_x X \} \). On the other hand, since \( \bigcap_{x \in \Delta_p \cup Q} \Delta_x X \cap \Delta_p X \), then \( \bigcap_{x \in \Delta_p \cup Q} \Delta_x X \subseteq \bigcap_{x \in \Delta_p \cup Q} \Delta_x X \). If \( x \notin X \), \( x \in \Delta_p X - Z \), that is, \( \Delta X \subseteq X \cup (\Delta_p X - Z) \). In addition, \( X \subseteq \Delta X \), \( \Delta_p X = 0 \). \( \square \)

Proposition 3.4 Suppose \( P', Q' \) are two attribute sets, \( Q' \subseteq P' \). Denote \( B = P' \setminus Q' \), we have \( B^{(a,b)} X = X \cup \Delta_p X - Z - M_d \), where \( Z = \{ x \in \bigcap_{x \in \Delta_p} \Delta_x X \} \).
\[ \cap_{x \in P} \{ x \in \Delta_{\bar{P}} \bar{P} X \} \leq \cap_{x \in \Delta_{\bar{P}} \bar{P} X} \{ x \in \Delta_{\bar{P}} \bar{P} X \} \leq \{ x \in \Delta_{\bar{P}} \bar{P} X \} \leq \{ x \in \Delta_{\bar{P}} \bar{P} X \} \leq \{ x \in \Delta_{\bar{P}} \bar{P} X \} \]

**Proof.** Let \( x \in \bar{B} X \) and \( x \notin X \). Since \( \bar{B} X = X + \bar{A}_a X \), then \( x \notin \bar{A}_a X \). On the one hand, suppose \( Q' \) is non-redundant attribute. \( \bar{A}_a X \leq \bar{A}_a X \). If \( x \in \bar{B} X \), \( x \notin X \) and \( x \notin \bar{A}_a X \), we have \( x \in \bar{A}_a X - \bar{A}_a X \), that is, \( x \notin \cap_{x \in \bar{A}_a X} \). In the other hand, when \( x \in \cap_{x \in \bar{A}_a X} \), due to \( \cap_{x \in \bar{A}_a X} \leq \cap_{x \in \bar{A}_a X} \), we have \( x \in Z \leq \bar{B} X \). Furthermore, due to \( B \subseteq P' \), we have \( \bar{P} X \subseteq \bar{B} X \). Since \( \bar{P} X \subseteq \bar{B} X \), then \( \bar{P} X \subseteq \bar{B} X \). Since \( \bar{P} X \subseteq \bar{B} X \), we have \( \bar{P} X \subseteq \bar{B} X \) and \( \cap_{x \in \bar{A}_a X} \). Specially, if \( Q' \) is \((\alpha, \beta)\)-redundant, we have \( \bar{P} X \subseteq \bar{B} X \).

However, with the insight gain from the above discussion, we can directly obtain the updating lower and upper approximations by only using the original database and need not recalculate the new information system when attributes change, which gives a quicker algorithm for incremental learning.

4. An Illustration

Considering a complete information system \( S = (U, C, V, f) \) (see Table 1), where, \( U = \{ x_1, x_2, \ldots, x_{12} \} \), \( C = \{ a_1, a_2, a_3, a_4 \} \). Let \( P = \{ a_1, a_2 \} \), \( Q = \{ a_1, a_4 \} \), \( A = P \cup Q \), \( P' = \{ a_1, a_2, a_3, a_4 \} \), \( Q' = \{ a_1, a_4 \} \), \( B = P' \setminus Q' \). \( X = \{ x_1, x_4, x_5, x_8, x_9, x_{11}, x_{12} \} \). We set \( \alpha = 0.6 \) and \( \beta = 0.4 \).

\[ \begin{array}{cccccccc}
| U | a_1 | a_2 | a_3 | a_4 | & | U | a_1 | a_2 | a_3 | a_4 |
|---|---|---|---|---|---|---|---|---|---|
| x_1 | 1 | 1 | 1 | 0 | x_7 | 2 | 2 | 2 | 1 |
| x_2 | 1 | 2 | 1 | 0 | x_8 | 2 | 3 | 3 | 1 |
| x_3 | 1 | 2 | 1 | 0 | x_9 | 2 | 3 | 3 | 1 |
| x_4 | 1 | 2 | 1 | 0 | x_{10} | 2 | 3 | 3 | 1 |
| x_5 | 1 | 2 | 2 | 1 | x_{11} | 3 | 3 | 3 | 2 |
| x_6 | 2 | 2 | 1 | 1 | x_{12} | 3 | 3 | 3 | 2 |
\end{array} \]
(1) Due to $PX = \{x_1\}$, $QX = \{x_6\}$, $\Delta_P X \cap \Delta_Q X = \{x_2, x_3, x_7, x_{12}\}$, $Y = \{x_4\}$, $\Delta_{P \cup Q} X = \{x_4, x_5, x_6, x_9, x_9, x_{12}\}$, $\overline{\Delta_{P \cup Q}} X = \{x_2, x_3, x_7, x_{12}\}$, $Y' = \{x_5, x_6, x_8, x_9, x_{10}\}$, we have: $A^{(0.6,0.4)} X = \{x_1, x_5, x_6, x_9, x_{10}\}$. 

(2) Due to $P'X = \{x_1, x_2, x_6\}$, $\Delta_{P' \cup Q'}X = \{x_4, x_5, x_6, x_9, x_9, x_{12}\}$, $\overline{\Delta_{P' \cup Q'}} X = \{x_2, x_3, x_7, x_{12}\}$, $Y^* = \{x_8, x_9, x_{10}, x_{11}, x_{12}\}$, we have: $B^{(0.6,0.4)} X = P'X - \Delta_{P' \cup Q'} X \cup Y^* = \{x_1, x_2, x_3, x_6, x_9, x_{11}, x_{12}\}$.

(3) Due to $\overline{A_{P}} X = U$, $Z = \emptyset$, $A_{P} X - Z = \emptyset$. In addition, $\Delta_{P \cup Q} X = \{x_4, x_5, x_6, x_9, x_9, x_{11}\}$, $\overline{\Delta_{P \cup Q}} X = \{x_2, x_3, x_7, x_{12}\}$, $M_{A} = \{x_2, x_3, x_4, x_7\}$. Then, we have $A^{(0.6,0.4)} X = X \cup (\overline{A_{P}} X - Z) - M_{A} = \{x_1, x_3, x_4, x_6, x_9, x_{10}, x_{11}, x_{12}\}$.

(4) Due to $\overline{A_{P}} X = U \setminus \{x_7\}$, $Z = \{x_2, x_3, x_7, x_{10}, x_{12}\}$, $\Delta_{P' \cup Q'}X = \{x_4, x_5, x_6, x_9, x_9, x_{11}\}$, $\overline{\Delta_{P' \cup Q'}}X = \{x_2, x_3, x_7, x_{10}, x_{12}\}$, $M_{B} = \emptyset$, then, we have $B^{(0.6,0.4)} X = X \cup \overline{A_{P}} X \cup Z - M_{B} = U$.

5. Conclusions

By considering the error classification in real information systems, PRS is induced into the incremental learning of attribute change at first. Then, the propositions for incremental learning knowledge in PRS as well their case studies are proposed respectively when coarsening or refining the attribute set. Our future research work will focus on the design of a heuristic algorithm and implementation of our approach. Another work will try to extend our approach to incomplete systems and other generalized relations.

References


