Knowledge-Based Systems xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

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

Attribute reduction: A dimension incremental strategy

Feng Wang, Jiye Liang *, Yuhua Qian

Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China

ARTICLE INFO

Article history:
Received 31 October 2011
Received in revised form 11 October 2012
Accepted 12 October 2012
Available online xxxx

Keywords:
Expansion of attributes
Rough sets
Information entropy
Attribute reduction

ABSTRACT

Many real data sets in databases may vary dynamically. With the rapid development of data processing tools, databases increase quickly not only in rows (objects) but also in columns (attributes) nowadays. This phenomena occurs in several fields including image processing, gene sequencing and risk prediction in management. Rough set theory has been conceived as a valid mathematical tool to analyze various types of data. A key problem in rough set theory is executing attribute reduction for a data set. This paper focuses on attribute reduction for data sets with dynamically-increasing attributes. Information entropy is a common measure of uncertainty and has been widely used to construct attribute reduction algorithms. Based on three representative entropies, this paper develops a dimension incremental strategy for redcut computation. When an attribute set is added to a decision table, the developed algorithm can find a new reduct in a much shorter time. Experiments on six data sets downloaded from UCI show that, compared with the traditional non-incremental reduction algorithm, the developed algorithm is effective and efficient.

1. Introduction

Rough set theory, proposed by Pawlak, is a relatively new soft computing tool to conceptualize and analyze various types of data [23–25]. It has become a popular mathematical framework for pattern recognition, image processing, feature selection, rule extraction, neuro-computing, conflict analysis, decision supporting, granular computing, data mining and knowledge discovery from given data sets [4–7,13,16,19,33,34,38,41,44].

In rough set theory, an important concept is attribute reduction which can be considered a kind of specific feature selection. In other words, based on rough set theory, one can select useful features from a given data set. Attribute reduction does not attempt to maximize the class separability but rather to keep the discernibility ability of the original ones [8,11,12,26,31,37,42]. In the last two decades, researchers have proposed many reduction algorithms [10,14,20,21,29,39,40]. However, most of these algorithms can only be applicable to static data sets. In other words, when data sets vary with time, these algorithms have to be implemented from scratch to obtain new reduct. As data sets change with time, especially at an unprecedented rate, it is very time-consuming or even infeasible to run repeatedly an attribute reduction algorithm.

To overcome this deficiency, researchers have recently proposed many new analytic techniques for attribute reduction. These techniques usually can directly carry out the computation using the existing result from the original data set [9,14,18,22,36]. A common character of these algorithms is that they were proposed to deal with dynamically-increasing data sets in an incremental manner. However, many real databases expand not only in rows (objects) but also in columns (attributes) in many applications. For example, with the development of tools in gene sequencing, the obtained segments of DNA may get longer, which results in storing more columns. So does cancer patients, there will be more clinical features as the disease progresses, which also results in expansion of attributes. Another example is about the information input of students. For a student, different departments in a school may save his various information. Merging all of his information can offers his a comprehensive evaluation. The process of merging information may also result in the expansion of attributes in databases. Moreover, there are many other examples about the expansion of attributes such as image processing, risk prediction and animal experiments. Therefore, to acquire knowledge from data sets with dynamically-increasing attributes, it is necessary to design a dimension incremental strategy for reduct computation.

Based on rough set theory, there exists some research on knowledge updating caused by the variation of attributes. In [1], an incremental algorithm was proposed to update the upper and lower approximations of a target concept in an information system. For an incomplete information system, when there are multiple attributes that are deleted from or added into it, Li et al. proposed an approach to update approximations of a target concept [15]. In addition, based on rough fuzzy set theory, two

* Corresponding author. Tel./fax: +86 0351 7018176.
E-mail addresses: sxuwangfeng@126.com (F. Wang), ljy@sxu.edu.cn (J. Liang), jinchengqyh@126.com (Y. Qian).

0950-7051/$ - see front matter © 2012 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.knosys.2012.10.010

incremental approaches to update rough fuzzy approximations were presented in [2]. One of these two approaches starts from the boundary set, and the other one is based on the cut sets of a fuzzy set. In [43], Zhang et al. proposed an incremental algorithm for updating approximations of a concept in variable precision rough set. Based on above analysis, we remark that existing dimension incremental algorithms mainly focus on updating approximations. The dimension incremental algorithms for updating reduct have not yet been discussed so far. Therefore, this paper presents a dimension incremental algorithm for redcut computation.

The information entropy from classical thermodynamics is used to measure out-of-order degree of a system. It is introduced in rough set theory to measure uncertainty of a data set, which has been widely applied to devise heuristic attribute reduction algorithms [16,17,27–29,32]. Complementary entropy [17], combination entropy [27] and Shannon's entropy [30] are three representative entropies which have been mainly used to construct reduction algorithms in rough set theory. To fully explore properties in reduct updating caused by the expansion of attributes, this paper develops a dimension incremental algorithms for dynamic data sets based on the three entropies. In view of that a key step of the development is the computation of entropy, this paper first introduces three dimension incremental mechanisms of the three entropies. These mechanisms can be used to determine an entropy by adding an attribute set to a decision table. When several attributes are added, instead of recomputation on the given decision table, the dimension incremental mechanisms derive new entropies by integrating the changes of conditional classes and decision classes into the existing entropies. With these mechanisms, a dimension incremental attribute reduction algorithm is proposed for dynamic decision tables. When an attribute set is added to a decision table, the developed algorithm can find a new reduct in a much shorter time. Experiments on six data sets downloaded from UCI show that, compared with the traditional non-incremental reduction algorithm, the developed algorithm is effective and efficient.

The rest of this paper is organized as follows. Some preliminaries in rough set theory are briefly reviewed in Section 2. Three representative entropies are introduced in Section 3. Section 4 presents the dimension incremental mechanisms of the three entropies for dynamically-increasing attributes. In Section 5, based on the dimension incremental mechanisms, a reduction algorithm is proposed to compute reducts for dynamic data sets. In Section 6, six UCI data sets are employed to demonstrate effectiveness and efficiency of the proposed algorithm. Section 7 concludes this paper with some discussions.

2. Preliminary knowledge on rough sets

This section reviews several basic concepts in rough set theory. Throughout this paper, the universe U is assumed a finite nonempty set.

An information system, as a basic concept in rough set theory, provides a convenient framework for the representation of objects in terms of their attribute values. An information system is a quadruple $S = (U,A,V,f)$, where U is a finite nonempty set of objects and is called the universe and A is a finite nonempty set of attributes, $V = \bigcup_{a\in A} V_a$ with V_a being the domain of a, and $f: U \times A \to V$ is an information function with $f(x,a) \in V_a$ for each $a \in A$ and $x \in U$. The system S can often be simplified as $S = (U,A)$.

Each nonempty subset $B \subseteq A$ determines an indiscernibility relation in the following way,

$$R_B = \{(x,y) \in U \times U | f(x,a) = f(y,a), \forall a \in B\}.$$

The relation R_B partitions U into some equivalence classes given by

$$U/R_B = \{[x]_B | x \in U\}, \quad \text{just } U/B,$$

where $[x]_B$ denotes the equivalence class determined by x with respect to B, i.e.,

$$[x]_B = \{y \in U | (x,y) \in R_B\}.$$

Given an equivalence relation R on the universe U and a subset $X \subseteq U$, one can define a lower approximation of X and an upper approximation of X by

$$\underline{R}X = \bigcup\{x \in U | [x]_R \subseteq X\}$$

and

$$\overline{R}X = \bigcup\{x \in U | [x]_R \cap X \neq \emptyset\},$$

respectively [3]. The order pair $(\underline{R}X, \overline{R}X)$ is called a rough set of X with respect to R. The positive region of X is denoted by $POS_R(X) = \underline{R}\,X$.

A partial relation $\preceq$ on the family $\{U/B | B \subseteq A\}$ is defined as follows [27]: $U/P \preceq U/Q$ (or $U/Q \succeq U/P$) if and only if, for every $P_i \in U/P$, there exists $Q_j \in U/Q$ such that $P_i \subseteq Q_j$, where $U/P = \{P_1, P_2, \ldots, P_m\}$ and $U/Q = \{Q_1, Q_2, \ldots, Q_n\}$ are partitions induced by $P, Q \subseteq A$, respectively. In this case, we say that Q is coarser than P, or P is finer than Q. If $U/P \preceq U/Q$ and $U/P \neq U/Q$, we say Q is strictly coarser than P (or P is strictly finer than Q), denoted by $U/P \prec U/Q$ (or $U/Q \succ U/P$).

It is clear that $U/P \prec U/Q$ if and only if, for every $X \in U/P$, there exists $Y \in U/Q$ such that $X \subseteq Y$, and there exist $X_0 \in U/P$ and $Y_0 \in U/Q$ such that $X_0 \subset Y_0$.

A decision table is an information system $S = (U, C \cup D)$ with $C \cap D = \emptyset$, where an element of C is called a condition attribute, C is called a condition attribute set, an element of D is called a decision attribute, and D is called a decision attribute set. Given $P \subseteq C$ and $U/D = \{D_1, D_2, \ldots, D_r\}$, the positive region of D with respect to the condition attribute set P is defined by $POS_P(D) = \bigcup_{k=1}^{r} \underline{P}D_k$.

3. Three representative entropies

In rough set theory, a given data table usually has multiple reducts, whereas it has been proved that finding its minimal is an NP-hard problem [31]. To overcome this deficiency, researchers have proposed many heuristic reduction algorithms which can generate a single reduct from a given table [10–12,16,17,28]. Most of these algorithms are of greedy and forward search type. Starting with a nonempty set, these algorithms keep adding one or several attributes of high significance into a pool at each iteration until the dependence no longer increases. Among various heuristic attribute reduction algorithms, reduction based on information entropy (or its variants) is a kind of common algorithm which has attracted much attention. The main idea of these algorithms is to keep the conditional entropy of target decision unchanged. This section reviews three representative entropies which are usually used to measure the attribute significance in a heuristic reduction algorithm.

In [16], the complementary entropy was introduced to measure uncertainty in rough set theory. Liang et al. also proposed the conditional complementary entropy to measure uncertainty of a decision table in [17]. By preserving the conditional entropy unchanged, the conditional complementary entropy was applied to construct reduction algorithms and reduce the redundant features in a decision table [28]. The conditional complementary entropy used in this algorithm is defined as follows [16,17,28].

Definition 1. Let $S = (U, C \cup D)$ be a decision table and $B \subseteq C$. Then, one can obtain the condition partitions $U/B = \{X_1, X_2, \ldots, X_m\}$ and $U/D = \{Y_1, Y_2, \ldots, Y_n\}$. Based on these partitions, a conditional entropy of B relative to D is defined as

Q1 Please cite this article in press as: F. Wang et al., Attribute reduction: A dimension incremental strategy, Knowl. Based Syst. (2012), http://dx.doi.org/10.1016/j.knosys.2012.10.010

$$E(D|B)=\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|X_i\cap Y_j|}{|U|}\frac{\left|Y_j^c-X_i^c\right|}{|U|}, \tag{1}$$

where Y_i^c and X_j^c are complement sets of Y_i and X_j respectively.

Based on the classical rough set model, Shannon's information entropy [30] and its conditional entropy were also introduced to find a reduct in a heuristic algorithm [29,32]. In [32], the reduction algorithm keeps the conditional entropy of target decision unchanged, and the conditional entropy is defined as follows [32].

Definition 2. Let $S=(U,C\cup D)$ be a decision table and $B\subseteq C$. Then, one can obtain the condition partitions $U/B=\{X_1,X_2,\ldots,X_m\}$ and $U/D=\{Y_1,Y_2,\ldots,Y_n\}$. Based on these partitions, a conditional entropy of B relative to D is defined as

$$H(D|B)=-\sum_{i=1}^{m}\frac{|X_i|}{|U|}\sum_{j=1}^{n}\frac{|X_i\cap Y_j|}{|X_i|}\log\left(\frac{|X_i\cap Y_j|}{|X_i|}\right). \tag{2}$$

Another information entropy, called combination entropy, was presented in [27] to measure the uncertainty of data tables. The conditional combination entropy was also introduced and can be used to construct the heuristic reduction algorithms [27]. This reduction algorithm can find a feature subset that possesses the same number of pairs of indistinguishable elements as that of the original decision table. The definition of the conditional combination entropy is defined as follows [27].

Definition 3. Let $S=(U,C\cup D)$ be a decision table and $B\subseteq C$. Then one can obtain the condition partitions $U/B=\{X_1,X_2,\ldots,X_m\}$ and $U/D=\{Y_1,Y_2,\ldots,Y_n\}$. Based on these partitions, a conditional entropy of B relative to D is defined as

$$CE(D|B)=\sum_{i=1}^{m}\left(\frac{|X_i|}{|U|}\frac{C_{|X_i|}^2}{C_{|U|}^2}-\sum_{j=1}^{n}\frac{|X_i\cap Y_j|}{|U|}\frac{C_{|X_i\cap Y_j|}^2}{C_{|U|}^2}\right). \tag{3}$$

where $C_{|X_i|}^2$ denotes the number of pairs of objects which are not distinguishable from each other in the equivalence class X_i.

4. Dimension incremental mechanism

Given a dynamic decision table, this section introduces the dimension incremental mechanisms for the three entropies. When an attributes set is added to a decision table, instead of recomputation on the given decision table, the dimension incremental mechanisms derive new entropies by integrating the changes of conditional classes and decision classes into the existing entropies.

For convenience, here are some explanations which will be used in the following theorems. Given a decision table $S=(U,C\cup D)$, $B\subseteq C$, $U/B=\{X_1,X_2,\ldots,X_m\}$ and $U/D=\{Y_1,Y_2,\ldots,Y_n\}$. Suppose that P is a conditional attribute set, and $U/(B\cup P)$ can be expressed as

$$U/(B\cup P)=\Big\{X_1,X_2,\ldots,X_k,X_1^{k+1},X_1^{k+1},\ldots,X_{l_{k+1}}^{k+1},X_1^{k+2},X_2^{k+2},\ldots,X_{l_{k+2}}^{k+2},\ldots,X_1^m,X_2^m,\ldots,X_{l_m}^m\Big\},$$

where $\bigcup_{j=1}^{l_i}X_j^i=X_i$ $(i=k+1,k+2,\ldots,m)$, i.e., $X_i\in U/B$ is divided into $X_1^i,X_2^i,\ldots,X_{l_i}^i$ in $U/(B\cup P)$.

Example 1. Let $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7\}$ and $U/B=\{\{x_1,x_2\},\{x_3,x_4\},\{-$ $=\{\{x_1,x_2\},\{x_3,x_4\},\{x_5,x_6,x_7\}\}$. Suppose that P is the incremental attribute set, and $U/(B\cup P)=\{\{x_1,x_2\},\{x_3\},\{x_4\},\{x_5\},\{x_6,x_7\}\}$. Hence, we have

$X_1=\{x_1,x_2\},X_2=\{x_3,x_4\};$
$X_1^2=\{x_3\},X_2^2=\{x_4\};$
$l_2=2;$
$X_2=X_1^2\cup X_2^2.$

And

$X_3=\{x_5,x_6,x_7\};$
$X_1^3=\{x_5\},X_2^3=\{x_6,x_7\};$
$l_3=2;$
$X_3=X_1^3\cup X_2^3.$

4.1. Dimension incremental mechanism of complementary entropy

Given a decision table, Theorem 1 introduces the dimension incremental mechanism based on complementary entropy (see Definition 1).

Theorem 1. *Let $S=(U,C\cup D)$ be a decision table and $B\subseteq C$. $U/B=\{X_1,X_2,\ldots,X_m\}$ and $U/D=\{Y_1,Y_2,\ldots,Y_n\}$. Suppose that P is the incremental conditional attribute set and $U/(B\cup P)=\{X_1,X_2,\ldots,X_k,X_1^{k+1},X_2^{k+1},\ldots,X_{l_{k+1}}^{k+1},X_1^{k+2},X_2^{k+2},\ldots,X_{l_{k+2}}^{k+2},\ldots,X_1^m,\ X_2^m,\ldots,X_{l_m}^m\}$. Then, the new conditional entropy becomes*
$E(D|(B\cup P))=E(D|B)-\Delta,$

where

$$\Delta=\sum_{l=k+1}^{m}\sum_{i=1}^{l_l}\sum_{j=1}^{n}\frac{\left|X_i^l\cap Y_j\right|\sum_{i'\neq i}\left|X_{i'}^l-Y_j\right|}{|U|^2}.$$

Proof. From Definition 1, we have

$$E(D|B)=\sum_{l=1}^{m}\sum_{j=1}^{n}\frac{|X_l\cap Y_j|}{|U|}\frac{\left|Y_j^c-X_l^c\right|}{|U|}=\sum_{l=1}^{m}\sum_{j=1}^{n}\frac{|X_l\cap Y_j|}{|U|}\frac{|X_l-Y_j|}{|U|}.$$

Because $X_l=\bigcup_{i=1}^{l_l}X_i^l(l=k+1,\ldots,m)$ (the specific introduction of l_l can be got from Example 1), we have

$$\begin{aligned}E(D|B)&=\sum_{l=1}^{k}\sum_{j=1}^{n}\frac{|X_l\cap Y_j|}{|U|}\frac{|X_l-Y_j|}{|U|}+\sum_{l=k+1}^{m}\sum_{j=1}^{n}\frac{|X_l\cap Y_j|}{|U|}\frac{|X_l-Y_j|}{|U|}\\&=\sum_{l=1}^{k}\sum_{j=1}^{n}\frac{|X_l\cap Y_j|}{|U|}\frac{|X_l-Y_j|}{|U|}+\sum_{l=k+1}^{m}\sum_{j=1}^{n}\frac{\sum_{i=1}^{l_l}\left|X_i^l\cap Y_j\right|}{|U|}\\&\quad\times\frac{\sum_{i=1}^{l_l}\left|X_i^l-Y_j\right|}{|U|}.\end{aligned}$$

Because that

$$\begin{aligned}\sum_{i=1}^{l_l}\left|X_i^l\cap Y_j\right|\cdot\sum_{i=1}^{l_l}\left|X_i^l-Y_j\right|&=\sum_{i=1}^{l_l}\left(\left|X_i^l\cap Y_j\right|\left|X_i^l-Y_j\right|+\left|X_i^l\cap Y_j\right|\cdot\sum_{i'\neq i}\left|X_{i'}^l-Y_j\right|\right)\\&=\sum_{i=1}^{l_l}\left|X_i^l\cap Y_j\right|\left|X_i^l-Y_j\right|+\sum_{i=1}^{l_l}\left|X_i^l\cap Y_j\right|\cdot\sum_{i'\neq i}\left|X_{i'}^l-Y_j\right|,\end{aligned}$$

we can get

$$\begin{aligned}E(D|B)&=\sum_{l=1}^{k}\sum_{j=1}^{n}\frac{|X_l\cap Y_j|}{|U|}\frac{|X_l-Y_j|}{|U|}\\&\quad+\sum\nolimits_{l=k+1}^{m}\sum_{j=1}^{n}\frac{\sum_{i=1}^{l_l}\left|X_i^l\cap Y_j\right|\left|X_i^l-Y_j\right|+\sum_{i=1}^{l_l}\left|X_i^l\cap Y_j\right|\cdot\sum_{i'\neq i}\left|X_{i'}^l-Y_j\right|}{|U|^2}\\&=\sum_{l=1}^{k}\sum_{j=1}^{n}\frac{|X_l\cap Y_j|}{|U|}\frac{|X_l-Y_j|}{|U|}+\sum\nolimits_{l=k+1}^{m}\sum\nolimits_{j=1}^{n}\sum\nolimits_{i=1}^{l_l}\\&\quad\times\frac{\left|X_i^l\cap Y_j\right|\left|X_i^l-Y_j\right|+\left|X_i^l\cap Y_j\right|\cdot\sum_{i'\neq i}\left|X_{i'}^l-Y_j\right|}{|U|^2}.\end{aligned}$$

Q1 Please cite this article in press as: F. Wang et al., Attribute reduction: A dimension incremental strategy, Knowl. Based Syst. (2012), http://dx.doi.org/10.1016/j.knosys.2012.10.010

It is obvious that $\sum_{I=1}^{k}\sum_{j=1}^{n}\frac{|X_I\cap Y_j|}{|U|}\frac{|X_I-Y_j|}{|U|}+\sum_{I=k+1}^{m}\sum_{j=1}^{n}\sum_{i=1}^{l_I}\frac{\left|X_i^I\cap Y_j\right|\left|X_i^I-Y_j\right|}{|U|^2=E(D|(B\cup P))}$. Let $\Delta=\sum_{I=k+1}^{m}\sum_{j=1}^{n}\sum_{i=1}^{l_I}\frac{\left|X_i^I\cap Y_j\right|\cdot\sum_{i'\neq i}\left|X_{i'}^I-Y_j\right|}{|U|^2}$, we have

$$E(D|B)=E(D|(B\cup P))+\Delta,$$

namely,

$$E(D|(B\cup P))=E(D|B)-\Delta.$$

This completes the proof. □

4.2. Dimension incremental mechanism of Shannon's information entropy

In this subsection, the dimension incremental mechanism based on Shannon's entropy (see Definition 2) is introduced in Theorem 2.

Theorem 2. *Let $S=(U,C\cup D)$ be a decision table and $B\subseteq C$. $U/B=\{X_1,X_2,\ldots,X_m\}$ and $U/D=\{Y_1,Y_2,\ldots,Y_n\}$. Suppose that P is the incremental conditional attribute set and $U/(B\cup P)=\{X_1,X_2,\ldots,X_k,X_1^{k+1},X_2^{k+1},\ldots,X_{l_{k+1}}^{k+1},X_1^{k+2},X_2^{k+2},\ldots,X_{l_{k+2}}^{k+2},\ldots,X_1^m,X_2^m,\ldots,X_{l_m}^m\}$. Then, the new Shannon's information entropy becomes*

$$H(D|(B\cup P))=H(D|B)+\Delta,$$

where,

$$\Delta=\sum_{I=k+1}^{m}\sum_{i=1}^{l_I}\sum_{j=1}^{n}\frac{\left|X_i^I\cap Y_j\right|}{|U|}log\frac{\left|X_i^I\right||X_I\cap Y_j|}{|X_I|\left|X_i^I\cap Y_j\right|}.$$

Proof. Because $X_I=\bigcup_{i=1}^{l_I}X_i^I(I=k+1,\ldots,m)$, we have

$$\begin{aligned}
&H(D|B)\\
&=-\left(\sum_{I=1}^{k}\frac{|X_I|}{|U|}\sum_{j=1}^{n}\frac{|X_I\cap Y_j|}{|X_I|}\log\frac{|X_I\cap Y_j|}{|X_I|}+\sum_{I=k+1}^{m}\frac{|X_I|}{|U|}\sum_{j=1}^{n}\frac{|X_I\cap Y_j|}{|X_I|}log\frac{|X_I\cap Y_j|}{|X_I|}\right)\\
&=-\left(\sum_{I=1}^{k}\frac{|X_I|}{|U|}\sum_{j=1}^{n}\frac{|X_I\cap Y_j|}{|X_I|}\log\frac{|X_I\cap Y_j|}{|X_I|}+\sum_{I=k+1}^{m}\frac{|X_I|}{|U|}\sum_{j=1}^{n}\frac{\sum_{i=1}^{l_I}\left|X_i^I\cap Y_j\right|}{|X_I|}\log\frac{|X_I\cap Y_j|}{|X_I|}\right)\\
&=-\left(\sum_{I=1}^{k}\frac{|X_I|}{|U|}\sum_{j=1}^{n}\frac{|X_I\cap Y_j|}{|X_I|}\log\frac{|X_I\cap Y_j|}{|X_I|}+\sum_{I=k+1}^{m}\sum_{i=1}^{l_I}\frac{\left|X_i^I\right|}{|U|}\sum_{j=1}^{n}\frac{\left|X_i^I\cap Y_j\right|}{\left|X_i^I\right|}log\left(\frac{\left|X_i^I\cap Y_j\right|}{\left|X_i^I\right|}\cdot\frac{\left|X_i^I\right||X_I\cap Y_j|}{|X_I|\left|X_i^I\cap Y_j\right|}\right)\right)\\
&=-\left(\sum_{I=1}^{k}\frac{|X_I|}{|U|}\sum_{j=1}^{n}\frac{|X_I\cap Y_j|}{|X_I|}log\frac{|X_I\cap Y_j|}{|X_I|}+\sum_{I=k+1}^{m}\sum_{i=1}^{l_I}\frac{\left|X_i^I\right|}{|U|}\sum_{j=1}^{n}\frac{\left|X_i^I\cap Y_j\right|}{\left|X_i^I\right|}\log\frac{\left|X_i^I\cap Y_j\right|}{\left|X_i^I\right|}\right.\\
&\quad\left.+\sum_{I=k+1}^{m}\sum_{i=1}^{l_I}\frac{\left|X_i^I\right|}{|U|}\sum_{j=1}^{n}\frac{\left|X_i^I\cap Y_j\right|}{\left|X_i^I\right|}\log\frac{\left|X_i^I\right||X_I\cap Y_j|}{|X_I|\left|X_i^I\cap Y_j\right|}\right)\\
&=-\left(\sum_{I=1}^{k}\frac{|X_I|}{|U|}\sum_{j=1}^{n}\frac{|X_I\cap Y_j|}{|X_I|}\log\frac{|X_I\cap Y_j|}{|X_I|}+\sum_{I=k+1}^{m}\sum_{i=1}^{l_I}\frac{\left|X_i^I\right|}{|U|}\sum_{j=1}^{n}\frac{\left|X_i^I\cap Y_j\right|}{\left|X_i^I\right|}\log\frac{\left|X_i^I\cap Y_j\right|}{\left|X_i^I\right|}\right)\\
&\quad-\sum_{I=k+1}^{m}\sum_{i=1}^{l_I}\sum_{j=1}^{n}\frac{\left|X_i^I\cap Y_j\right|}{|U|}\log\frac{\left|X_i^I\right||X_I\cap Y_j|}{|X_I|\left|X_i^I\cap Y_j\right|}.
\end{aligned}$$

Because $-\left(\sum_{I=1}^{k}\frac{|X_I|}{|U|}\sum_{j=1}^{n}\frac{|X_I\cap Y_j|}{|X_I|}log\frac{|X_I\cap Y_j|}{|X_I|}+\sum_{I=k+1}^{m}\sum_{i=1}^{l_I}\frac{\left|X_i^I\right|}{|U|}\sum_{j=1}^{n}\frac{\left|X_i^I\cap Y_j\right|}{\left|X_i^I\right|}\log\frac{\left|X_i^I\cap Y_j\right|}{\left|X_i^I\right|}\right)=H(D|(B\cup P))$, and let $\Delta=\sum_{I=k+1}^{m}\sum_{i=1}^{l_I}\sum_{j=1}^{n}\frac{\left|X_i^I\cap Y_j\right|}{|U|}\log\frac{\left|X_i^I\right||X_I\cap Y_j|}{|X_I|\left|X_i^I\cap Y_j\right|}$, we have $H(D|B)=H(D|(B\cup P))-\Delta$. Hence, $H(D|(B\cup P))=H(D|B)+\Delta$. This completes the proof. □

4.3. Dimension incremental mechanism of combination entropy

For convenience of introducing dimension incremental mechanism of combination entropy, here gives a variant of the definition of combination entropy (see Definition 3). According to $C_N^2=\frac{N(N-1)}{2}$, Definition 4 shows a variant of combination entropy. Based on this variant, the dimension incremental mechanism of combination entropy is introduced in Theorem 3.

Definition 4. Let $S=(U,C\cup D)$ be a decision table and $B\subseteq C$. One can obtain the condition partition $U/B=\{X_1,X_2,\ldots,X_m\}$ and $U/D=\{Y_1,Y_2,\ldots,Y_n\}$. Then, the conditional entropy of B relative to D is defined as

$$CE(D|B)=\sum_{i=1}^{m}\left(\frac{|X_i|^2(|X_i|-1)}{|U|^2(|U|-1)}-\sum_{j=1}^{n}\frac{|X_i\cap Y_j|^2(|X_i\cap Y_j|-1)}{|U|^2(|U|-1)}\right).\qquad(4)$$

Theorem 3. *Let $S=(U,C\cup D)$ be a decision table and $B\subseteq C$. $U/B=\{X_1,X_2,\ldots,X_m\}$ and $U/D=\{Y_1,Y_2,\ldots,Y_n\}$. Suppose that P is the incremental conditional attribute set and $U/(B\cup P)=\{X_1,X_2,\ldots,X_k,X_1^{k+1},X_2^{k+1},\ldots,X_{l_{k+1}}^{k+1},X_1^{k+2},X_2^{k+2},\ldots,X_{l_{k+2}}^{k+2},\ldots,X_1^m,X_2^m,\ldots,X_{l_m}^m\}$. Then, the new conditional entropy becomes*

$$CE(D|(B\cup P))=CE(D|B)-\Delta,$$

where

$$\begin{aligned}
\Delta=\sum_{I=k+1}^{m}&\left(\frac{\sum_{i=1}^{l_I}\sum_{i\neq i'}\left|X_i^I\right|^2\left|X_{i'}^I\right|}{|U|^2(|U|-1)}+\frac{2(|X_I|-1)\sum_{i=1}^{l_{I-1}}\sum_{i'=i+1}^{I}\left|X_i^I\right|\left|X_{i'}^I\right|}{|U|^2(|U|-1)}\right.\\
&-\sum_{j=1}^{n}\left(\frac{\sum_{i=1}^{l_I}\sum_{i\neq i'}\left|X_i^I\cap Y_j\right|^2\left|X_{i'}^I\cap Y_j\right|}{|U|^2(|U|-1)}\right.\\
&\left.\left.+\frac{2(|X_I\cap Y_j|-1)\sum_{i=1}^{l_{I-1}}\sum_{i'=i+1}^{l_I}\left|X_i^I\cap Y_j\right|\left|X_{i'}^I\cap Y_j\right|}{|U|^2(|U|-1)}\right)\right).
\end{aligned}$$

Proof. Because $X_I=\bigcup_{i=1}^{l_I}X_i^I(I=k+1,\ldots,m)$, we have

$$\begin{aligned}
CE(D|B)=&\sum_{I=1}^{k}\left(\frac{|X_I|^2(|X_I|-1)}{|U|^2(|U|-1)}-\sum_{j=1}^{n}\frac{|X_I\cap Y_j|^2(|X_I\cap Y_j|-1)}{|U|^2(|U|-1)}\right)\\
&+\sum_{I=k+1}^{m}\left(\frac{|X_I|^2(|X_I|-1)}{|U|^2(|U|-1)}-\sum_{j=1}^{n}\frac{|X_I\cap Y_j|^2(|X_I\cap Y_j|-1)}{|U|^2(|U|-1)}\right).
\end{aligned}$$

And we simplify two items in the above formula as follows

$$\begin{aligned}
&|X_I|^2(|X_I|-1)=\left(\sum_{i=1}^{l_I}\left|X_i^I\right|\right)^2\cdot\left(\sum_{i=1}^{l_I}\left|X_i^I\right|-1\right)=\\
&\left(\sum_{i=1}^{l_I}\left|X_i^I\right|^2+2\sum_{i=1}^{l_{I-1}}\sum_{i'=i+1}^{l_I}\left|X_i^I\right|\left|X_{i'}^I\right|\right)\cdot\left(\sum_{i=1}^{l_I}\left|X_i^I\right|-1\right)\\
&=\sum_{i=1}^{l_I}\left(\left|X_i^I\right|^3+\sum_{i\neq i'}\left|X_i^I\right|^2\left|X_{i'}^I\right|\right)-\sum_{i=1}^{l_I}\left|X_i^I\right|^2+2\sum_{i=1}^{l_{I-1}}\sum_{i'=i+1}^{l_I}\left|X_i^I\right|\left|X_{i'}^I\right|\\
&\times\left(\sum_{i=1}^{l_I}\left|X_i^I\right|-1\right)=\sum_{i=1}^{l_I}\left|X_i^I\right|^2\left(\left|X_i^I\right|-1\right)+\sum_{i=1}^{l_I}\sum_{i\neq i'}\left|X_i^I\right|^2\left|X_{i'}^I\right|\\
&+2(|X_I|-1)\sum_{i=1}^{l_{I-1}}\sum_{i'=i+1}^{I}\left|X_i^I\right|\left|X_{i'}^I\right|.
\end{aligned}$$

And

$$|X_l \cap Y_j|^2(|X_l \cap Y_j| - 1) = \sum_{i=1}^{l_l}\left|X_i^l \cap Y_j\right|^2\left(\left|X_i^l \cap Y_j\right| - 1\right) + \sum_{i=1}^{l_l}\sum_{i \neq i'}\left|X_i^l \cap Y_j\right|^2 |X_{i'}^l \cap Y_j| + 2(|X_l \cap Y_j| - 1)\sum_{i=1}^{l_l-1}\sum_{i'=i+1}^{l_l}\left|X_i^l \cap Y_j\right|\left|X_{i'}^l \cap Y_j\right|.$$

Thus, the new combination is

$$\begin{aligned} CE(D|B) &= \sum_{l=1}^{k}\left(\frac{|X_l|^2(|X_l|-1)}{|U|^2(|U|-1)} - \sum_{j=1}^{n}\frac{|X_l \cap Y_j|^2(|X_l \cap Y_j|-1)}{|U|^2(|U|-1)}\right) + \sum_{l=k+1}^{m}\left(\frac{\sum_{i=1}^{l_l}\left|X_i^l\right|^2\left(\left|X_i^l\right|-1\right)}{|U|^2(|U|-1)} + \frac{\sum_{i=1}^{l_l}\sum_{i \neq i'}\left|X_i^l\right|^2|X_{i'}^l| + 2(|X_l|-1)\sum_{i=1}^{l_l-1}\sum_{i'=i+1}^{l_l}\left|X_i^l\right|\left|X_{i'}^l\right|}{|U|^2(|U|-1)} - \right. \\ &\left. \sum_{j=1}^{n}\left(\frac{\sum_{i=1}^{l_l}\left|X_i^l \cap Y_j\right|^2\left(\left|X_i^l \cap Y_j\right|-1\right)}{|U|^2(|U|-1)} + \frac{\sum_{i=1}^{l_l}\sum_{i \neq i'}\left|X_i^l \cap Y_j\right|^2\left|X_{i'}^l \cap Y_j\right|}{|U|^2(|U|-1)} + \frac{2(|X_l \cap Y_j|-1)\sum_{i=1}^{l_l-1}\sum_{i'=i+1}^{l}\left|X_i^l \cap Y_j\right|\left|X_{i'}^l \cap Y_j\right|}{|U|^2(|U|-1)}\right)\right) \\ &= \sum_{l=1}^{k}\left(\frac{|X_l|^2(|X_l|-1)}{|U|^2(|U|-1)} - \sum_{j=1}^{n}\frac{|X_l \cap Y_j|^2(|X_l \cap Y_j|-1)}{|U|^2(|U|-1)}\right) + \sum_{l=k+1}^{m}\left(\frac{\sum_{i=1}^{l_l}\left|X_i^l\right|^2\left(\left|X_i^l\right|-1\right)}{|U|^2(|U|-1)} - \sum_{j=1}^{n}\frac{\sum_{i=1}^{l_l}\left|X_i^l \cap Y_j\right|^2\left(\left|X_i^l \cap Y_j\right|-1\right)}{|U|^2(|U|-1)}\right) + \\ &\sum_{l=k+1}^{m}\left(\frac{\sum_{i=1}^{l_l}\sum_{i \neq i'}\left|X_i^l\right|^2\left|X_{i'}^l\right| + 2(|X_l|-1)\sum_{i=1}^{l_l-1}\sum_{i'=i+1}^{l}\left|X_i^l\right|\left|X_{i'}^l\right|}{|U|^2(|U|-1)} - \sum_{j=1}^{n}\left(\frac{\sum_{i=1}^{l_l}\sum_{i \neq i'}\left|X_i^l \cap Y_j\right|^2\left|X_{i'}^l \cap Y_j\right|}{|U|^2(|U|-1)} + \frac{2(|X_l \cap Y_j|-1)\sum_{i=1}^{l_l-1}\sum_{i'=i+1}^{l_l}\left|X_i^l \cap Y_j\right|\left|X_{i'}^l \cap Y_j\right|}{|U|^2(|U|-1)}\right)\right). \end{aligned}$$

Obviously, let

$$\Delta = \sum_{l=k+1}^{m}\left(\frac{\sum_{i=1}^{l_l}\sum_{i \neq i'}\left|X_i^l\right|^2|X_{i'}^l|}{|U|^2(|U|-1)} + \frac{2(|X_l|-1)\sum_{i=1}^{l_l-1}\sum_{i'=i+1}^{l_l}\left|X_i^l\right|\left|X_{i'}^l\right|}{|U|^2(|U|-1)} - \sum_{j=1}^{n}\left(\frac{\sum_{i=1}^{l_l}\sum_{i \neq i'}|X_i^l \cap Y_j|^2\left|X_{i'}^l \cap Y_j\right|}{|U|^2(|U|-1)} + \frac{2(|X_l \cap Y_j|-1)\sum_{i=1}^{l_l-1}\sum_{i'=i+1}^{l_l}\left|X_i^l \cap Y_j\right|\left|X_{i'}^l \cap Y_j\right|}{|U|^2(|U|-1)}\right)\right),$$

we have

$$CE(D|B) = CE(D|(B \cup P)) + \Delta,$$

namely,

$$CE(D|(B \cup P)) = CE(D|B) - \Delta.$$

This completes the proof. □

5. Dimension incremental algorithms

In rough set theory, core is also a key concept [23,24]. Given a decision table, core is the intersection of all reducts, and includes all indispensable attributes in a reduct. Based on the dimension incremental mechanisms, this section introduces dimension incremental algorithms for core and reduct.

For convenience, a uniform notation $ME(D|B)$ is introduced to denote the above three entropies. For example, if one adopts Shannon's conditional entropy to define the attribute significance, then $ME(D|B) = H(D|B)$. In [16,28,32], the attribute significance is defined as follows.

Definition 5. Let $S = (U, C \cup D)$ be a decision table and $B \subseteq C$. $\forall a \in B$, the significance measure (inner significance) of a in B is defined as

$$Sig^{inner}(a, B, D) = ME(D|B - \{a\}) - ME(D|B). \quad (5)$$

Definition 6. Let $S = (U, C \cup D)$ be a decision table and $B \subseteq C$. $\forall a \in C - B$, the significance measure (outer significance) of a in B is defined as

$$Sig^{outer}(a, B, D) = ME(D|B) - ME(D|B \cup \{a\}). \quad (6)$$

Given a decision table $S = (U, C \cup D)$ and $a \in C$. From the literatures [23,16,28,27], one can get that if $Sig^{inner}(a, C, D) > 0$, then the attribute a is indispensable, i.e., a is a core attribute of S. Based on the core attributes, a heuristic attribute reduction algorithm can find an attribute reduct by gradually adding selected attributes to the core. The definition of reduct based on information entropy is defined as follows.

Definition 7. Let $S = (U, C \cup D)$ be a decision table and $B \subseteq C$. Then the attribute set B is a relative reduct if B satisfies:

(1) $ME(D|B) = ME(D|C)$;
(2) $\forall a \in B$, $ME(D|B) \neq ME(D|B - \{a\})$.

The first condition guarantees that the reduct has the same distinguish power as the whole attribute set, and the second condition guarantees that there is no redundant attributes in the reduct.

Based on Definition 5, when a conditional attribute set is added to a decision table, we propose in the following a dimension incremental algorithm for core computation. In this algorithm, there are two key problems need to be considered. The first one is removing non-core attributes from the original core. And the second one is finding new core attributes from the incremental attribute set.

Algorithm 1. A dimension incremental algorithm for core computation (*DIA_CORE*)

Input: A decision table $S = (U, C \cup D)$, core attributes $CORE_C$ on C and the new condition attribute set P.
Output: Core attribute $CORE_{C \cup P}$ on $C \cup P$.
Step 1: Compute $U/(C \cup P) = \{X_1, X_2, \ldots, X_k, X_1^{k+1}, X_1^{k+1}, \ldots, X_{l_{k+1}}^{k+1}, X_1^{k+2}, X_2^{k+2}, \ldots, X_{l_{k+2}}^{k+2}, \ldots, X_1^m, X_2^m, \ldots, X_{l_m}^m\}$ and $U/D = \{Y_1, Y_2, \ldots, Y_n\}$.
Step 2: Compute $ME(D|(C \cup P))$(according to Theorems or1–3).
Step 3: $CORE_{C \cup P} \leftarrow CORE_C$.
For each $a \in CORE_{C \cup P}$ do
{
If $ME(D|(C - \{a\}) \cup P) = ME(D|C \cup P)$, then $CORE_{C \cup P} \leftarrow CORE_{C \cup P} - \{a\}$.
}
Step 4: For each $a \in P$ do
{
If $ME(D|C \cup (P - \{a\})) \neq ME(D|C \cup P)$, then $CORE_{C \cup P} \leftarrow CORE_{C \cup P} \cup \{a\}$.
}
Step 5: Return $CORE_{C \cup P}$ and end.

In rough set theory, as mentioned above, attribute reduct is a very important issue. Algorithm 2 introduces a dimension incremental algorithm for reduct computation. Supposed that P is an incremental conditional attribute set. In this algorithm, new core attributes are found from P firstly, and then attributes with highest significance are selected from P and added to the reduct gradually. At last, the redundant attributes in the reduct are deleted.

Algorithm 2. A dimension incremental algorithm for reduction computation (*DIA_RED*)

Input: A decision table $S = (U, C \cup D)$, reduct RED_C on C and the incremental conditional attribute set P.
Output: Reduct $RED_{C \cup P}$ on $C \cup P$.
Step 1: Compute
$U/(C \cup P) = \{X_1, X_2, \ldots, X_k, X_1^{k+1}, X_1^{k+1}, \ldots, X_{l_{k+1}}^{k+1}, X_1^{k+2}, X_2^{k+2}, \ldots, X_{l_{k+2}}^{k+2}, \ldots, X_1^m, X_2^m, \ldots, X_{l_m}^m\}$ and $U/D = \{Y_1, Y_2, \ldots, Y_n\}$.
Step 2: Compute $ME(D|(C \cup P))$(according to Theorems or1–3).
Step 3: $Core_P \leftarrow \emptyset$, for each $a \in P$ do
{
If $ME(D|C \cup (P - \{a\})) \neq ME(D|C \cup P)$, then $Core_P \leftarrow Core_P \cup \{a\}$.
}
Step 4: $B \leftarrow RED_C \cup Core_P$, if $ME(D|B) = ME(D|C \cup P)$, then turn to *Step 6*; else turn to *Step 5*.
Step 5: *while* $ME(D|B) \neq ME(D|C \cup P)$ do
{For each $a \in P - Core_P$, compute $Sig^{outer}(a, B, D)$ (according to Theorems or1–3 and Definition 6);
Select $a_0 = max\{Sig^{outer}(a, B, D) : a \in P - Core_P\}$;
$B \leftarrow B \cup \{a\}$.
}
Step 6: For each $a \in RED_C$ do
{
If $Sig^{inner}(a, B, D) = 0$, then $B \leftarrow B - \{a\}$.
}
Step 7: $RED_{C \cup P} \leftarrow B$, return $RED_{C \cup P}$ and end.

In addition, time complexities of above two algorithms are discussed as follows. The time complexity of a traditional non-incremental heuristic reduction algorithm based on information entropy given in [28] is $O(|U||C|^2)$. However, this time complexity does not include the computational time of entropies. For a given decision table, computing entropies is a key step in above reduction algorithm, which is not computationally costless. Thus, to analyze the exact time complexity of above algorithm, the time complexity of computing entropies is given as well.

Given a decision table, according to Definitions 1–3, it first needs to compute the conditional classes and decision classes, respectively, and then computes the value of entropy. Xu et al. in [35] gave a fast algorithm for partition with time complexity being $O(|U||C|)$. So, the time complexity of computing entropy is

$$O(|U||C| + |U| + \sum_{i=1}^{m}|X_i| \cdot \sum_{j=1}^{n}|Y_j|) = O(|U||C| + |U| + |U||U|) = O(|U||C| + |U|^2),$$

where the specific introduction of m, n, X_i and Y_j is shown in Definitions 1–3. Hence, when P is added to the table, the time complexity of computing entropy is

$$\Theta = O(|U||C \cup P| + |U|^2) = O(|U|(|C| + |P|) + |U|^2).$$

By using the dimension incremental formulas shown in Theorems 1–3, one can also get the entropy. According to Theorems 1–3, the time complexity of computing entropy is

$$\Theta' = O(|U|(|C| + |P|) + |X||U|),$$

where X denotes the union of changed conditional classes in the universe before and after adding P to the table.

Table 1
Comparison of time complexity.

Classic	Incremental
Entropy	
$O(|U|(|C| + |P|) + |U|^2)$	$O(|U|(|C| + |P|) + |X||U|)$
TA_CORE	*DIA_CORE*
Core	
$O((|C| + |P|)^2|U| + (|C| + |P|)|U|^2)$	$O((|C| + |P|)^2|U| + (|C| + |P|)|X||U|)$
TA_RED	*DIA_RED*
Reduct	
$O((|C| + |P|)^2|U| + (|C| + |P|)|U|^2)$	$O((|C| + |P|)^2|U| + (|C| + |P|)|X||U|)$

In a traditional heuristic algorithm based on entropy, the time complexity of core computation is $O(|C|(|U||C| + |U|^2)) = O(|C|^2|U| + |C||U|^2)$. Hence, when P is added to a decision table, the time complexity of core computation is $O(|C \cup P|^2|U| + |C \cup P||U|^2) = (|C \cup P|^2|U| + |C \cup P||U|^2) = O((|C| + |P|)^2|U| + (|C| + |P|)|U|^2)$. In the algorithm *DIA_CORE*, the time complexity of *Step 1–2* is Θ'; in *Step 3*, the time complexity of deleting non-core attributes is $O(|CORE_C||\Theta') = O(|C|\Theta')$; new core attributes are selected in *Step 4* and its time complexity is $O(|P|\Theta')$. Hence, the total time complexity of *DIA_CORE* is

$$O(\Theta' + |C|\Theta' + |P|\Theta') = O((|C| + |P|)^2|U| + (|C| + |P|)|X||U|).$$

In a traditional heuristic reduct algorithm based on entropy, the time complexity of reduct computation is $O(|C|^2|U| + |C||U|^2 + (|C|^2|U| + |C||U|^2 + |C|\Theta) = O(|C|^2|U| + |C||U|^2)$. Hence, when P is added to a decision table, the time complexity of reduct computation is $O((|C| + |P|)^2|U| + (|C| + |P|)|U|^2)$. In the algorithm *DIA_RED*, the time complexity of *Step 1–2* is Θ'; the time complexity of *Step 3* is $O(|P|\Theta')$; in *Step 5*, the time complexity of adding attributes is also $O(|P|\Theta')$; and in *Step 6*, the time complexity of deleting reductant attributes is $O(|C|\Theta')$. Thus, the total time complexity of the algorithm *DIA_RED* is

$$O(\Theta' + |P|\Theta' + |C|\Theta') = O((|C| + |P|)^2|U| + (|C| + |P|)|X||U|).$$

To stress above findings, the time complexities of computing entropy, core and reduct are shown in Table 1. *TA_CORE* and *TA_RED* denote the traditional algorithm for core and reduct, respectively.

From Table 1, because of that $|X|$ is usually much smaller than $|U|$, we can conclude that the computational time of new dimension incremental algorithms are usually much smaller than that of the traditional ones.

6. Experimental analysis

The objective of the following experiments is to show effectiveness and efficiency of the proposed dimension incremental algorithms. The data sets used in the experiments are outlined in Table 2, which are all downloaded from UCI repository of machine learning databases. All the experiments have been carried out on a personal computer with Windows 7, Inter (R) Core (TM) i7-2600 CPU (2.66 GHz) and 4.00 GB memory. The software being used is Microsoft Visual Studio 2005 and the programming language is

Table 2
Description of data sets.

	Data sets	Samples	Attributes	Classes
1	Backup-large	307	35	19
2	Dermatology	366	33	6
3	Splice	3910	60	3
4	Kr-vs-kp	3196	36	2
5	Mushroom	5644	22	2
6	Ticdata2000	5822	85	2

Table 3
Comparison of algorithms for core computation based on complementary entropy.

Data sets	SIA (%)	TA_CORE		DIA_CORE		PIT (%)
		Core	Time/s	Core	Time/s	
Backup-large	20	7, 16	0.4240	7, 16	0.0330	92.21
	40	7, 16	0.4910	7, 16	0.0350	92.87
	60	7, 16	0.5800	7, 16	0.0440	92.41
	80	7, 16	0.6670	7, 16	0.0615	90.78
	100	7, 16	0.6970	7, 16	0.0785	88.74
Dermatology	20	16, 18	0.4650	16, 18	0.0355	92.37
	40	16, 18	0.5570	16, 18	0.0320	94.25
	60	∅	0.6620	∅	0.0453	93.15
	80	∅	0.7780	∅	0.0685	91.20
	100	∅	0.8140	∅	0.0714	91.23
Splice	20	∅	55.361	∅	3.5819	93.53
	40	∅	66.870	∅	7.2015	89.23
	60	∅	78.369	∅	13.082	83.31
	80	∅	89.682	∅	16.000	82.16
	100	∅	93.364	∅	18.792	79.87
Kr-vs-kp	20	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	13.401	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	6.2631	53.26
	40	2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23	23.768	2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23	9.5251	59.93
	60	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	30.601	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	12.881	57.91
	80	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	47.492	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	13.098	72.42
	100	1, 3, 4, 5, 6, 7, 10, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	61.341	1, 3, 4, 5, 6, 7, 10, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	15.756	74.31
Mushroom	20	1, 2, 3, 9	11.321	1, 2, 3, 9	1.2912	88.59
	40	1, 2, 3, 9	19.677	1, 2, 3, 9	2.0521	89.57
	60	20	35.650	20	3.0354	91.49
	80	∅	90.832	∅	5.2079	94.27
	100	∅	120.34	∅	8.9416	92.57
Ticdata2000	20	2, 5, 43, 44, 45, 46, 47, 48, 49, 51	228.97	2, 5, 43, 44, 45, 46, 47, 48, 49, 51	15. 215	93.36
	40	2, 5, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	338.71	2, 5, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	33.163	90.21
	60	2, 5, 43, 44, 47, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	424.13	2, 5, 43, 44, 47, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	51.768	87.79
	80	2, 5, 43, 44, 47, 55, 58, 59, 61, 62, 63, 64, 68	494.39	2, 5, 43, 44, 47, 55, 58, 59, 61, 62, 63, 64, 68	73.063	85.22
	100	2, 5, 43, 44, 47, 55, 59, 68, 80, 83	563.78	2, 5, 43, 44, 47, 55, 59, 68, 80, 83	81.231	85.59

Table 4
Comparison of algorithms for core computation based on combination entropy.

Data sets	SIA (%)	TA_CORE		DIA_CORE		PIT (%)
		Core	Time/s	Core	Time/s	
Backup-large	20	7, 16	0.4180	7, 16	0.0462	88.95
	40	7, 16	0.4870	7, 16	0.0302	93.80
	60	7, 16	0.5720	7, 16	0.0307	94.64
	80	7, 16	0.6610	7, 16	0.0420	93.64
	100	7, 16	0.6840	7, 16	0.0650	90.50
Dermatology	20	16, 18	0.4970	16, 18	0.0921	81.47
	40	16, 18	0.5880	16, 18	0.2001	65.96
	60	∅	0.6770	∅	0.2066	69.49
	80	∅	0.7980	∅	0.3098	61.18
	100	∅	0.8460	∅	0.3208	62.08
Splice	20	∅	53.071	∅	5.0801	90.43
	40	∅	63.867	∅	7.5008	88.26
	60	∅	75.333	∅	10.201	86.46
	80	∅	86.767	∅	11.801	86.40
	100	∅	91.151	∅	12.780	85.98
Kr-vs-kp	20	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	13.073	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	3.9028	70.15
	40	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	23.026	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	8.0024	65.25
	60	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	29.858	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	10.302	65.50

(continued on next page)

Table 4 (*continued*)

Data sets	*SIA* (%)	*TA_CORE* Core	Time/s	*DIA_CORE* Core	Time/s	*PIT* (%)
	80	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	44.897	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	13.922	68.99
	100	1, 3, 4, 5, 6, 7, 10, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	57.954	1, 3, 4, 5, 6, 7, 10, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	15.420	73.40
Mushroom	20	1, 2, 3, 9	11.216	1, 2, 3, 9	1.8140	83.83
	40	1, 2, 3, 9	19.359	1, 2, 3, 9	3.0071	84.47
	60	20	46.080	20	4.0093	88.46
	80	∅	88.483	∅	5.3102	93.99
	100	∅	98.337	∅	6.8436	93.04
Ticdata2000	20	2, 5, 43, 44, 45, 46, 47, 48, 49, 51	226.20	2, 5, 43, 44, 45, 46, 47, 48, 49, 51	8.3150	96.32
	40	2, 5, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	340.05	2, 5, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	14.302	95.79
	60	2, 5, 43, 44, 47, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	409.33	2, 5, 43, 44, 47, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	35.509	91.32
	80	2, 5, 43, 44, 47, 55, 58, 59, 61, 62, 63, 64, 68	470.11	2, 5, 43, 44, 47, 55, 58, 59, 61, 62, 63, 64, 68	70.147	85.07
	100	2, 5, 43, 44, 47, 55, 59, 68, 80, 83	527.01	2, 5, 43, 44, 47, 55, 59, 68, 80, 83	91.437	82.65

Table 5
Comparison of algorithms for core computation based on Shannon's entropy.

Data sets	*SIA* (%)	*TA_CORE* Core	Time/s	*DIA_CORE* Core	Time/s	*PIT* (%)
Backup-large	20	7, 16	0.4290	7, 16	0.0300	93.00
	40	7, 16	0.5130	7, 16	0.0330	93.56
	60	7, 16	0.6110	7, 16	0.0450	92.64
	80	7, 16	0.6900	7, 16	0.0590	91.45
	100	7, 16	0.7110	7, 16	0.0650	90.86
Dermatology	20	16, 18	0.4730	16, 18	0.0350	92.60
	40	16, 18	0.5810	16, 18	0.0440	92.43
	60	∅	0.7060	∅	0.0590	91.64
	80	∅	0.8110	∅	0.0790	90.26
	100	∅	0.8250	∅	0.0870	89.45
Splice	20	∅	57.105	∅	4.6803	91.80
	40	∅	68.741	∅	8.3805	87.81
	60	∅	80.573	∅	12.441	84.56
	80	∅	91.010	∅	15.581	82.88
	100	∅	105.96	∅	18.081	82.94
Kr-vs-kp	20	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	12.511	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	4.3630	65.12
	40	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	22.245	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	6.9065	68.95
	60	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	28.751	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	12.033	58.15
	80	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	43.165	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	14.093	67.35
	100	1, 3, 4, 5, 6, 7, 10, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	55.864	1, 3, 4, 5, 6, 7, 10, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	15.810	71.70
Mushroom	20	1, 3, 9	10.795	1, 3, 9	1.0122	90.62
	40	1, 3, 9	18.689	1, 3, 9	1.9027	89.82
	60	20	33.867	20	2.0367	93.98
	80	∅	87.452	∅	4.0056	95.42
	100	∅	102.37	∅	8.3276	91.87
Ticdata2000	20	2, 5, 43, 44, 45, 46, 47, 48, 49, 51	240.53	2, 5, 43, 44, 45, 46, 47, 48, 49, 51	15.012	93.76
	40	2, 5, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	362.86	2, 5, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	17.181	95.26
	60	2, 5, 43, 44, 47, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	428.77	2, 5, 43, 44, 47, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	23.533	94.51
	80	2, 5, 43, 44, 47, 55, 58, 59, 61, 62, 63, 64, 68	487.45	2, 5, 43, 44, 47, 55, 58, 59, 61, 62, 63, 64, 68	40.019	91.79
	100	2, 5, 43, 44, 47, 55, 59, 68, 80, 83	555.83	2, 5, 43, 44, 47, 55, 59, 68, 80, 83	50.610	90.89

C#. And in the data sets, *Mushroom* is a data set with missing values, and for a uniform treatment of all data sets, we remove the objects with missing values. Moreover, *Ticdata2000* is preprocessed using the data tool Rosetta.

Table 6
Comparison of algorithms for reduct computation based on complementary entropy.

Data sets	*SIA* (%)	*TA_RED*		*DIA_RED*		*PIT* (%)
		Reduct	Time/s	Reduct	Time/s	
Backup-large	20	1, 4, 6, 7, 8, 9, 15, 16, 22	1.7271	1, 4, 6, 7, 8, 9, 15, 16, 22	0.0430	97.51
	40	1, 4, 6, 7, 8, 9, 15, 16, 22	2.0261	1, 4, 6, 7, 8, 9, 15, 16, 22	0.0360	98.22
	60	1, 4, 6, 7, 8, 9, 15, 16, 22	2.3851	1, 4, 6, 7, 8, 9, 15, 16, 22	0.0490	97.94
	80	1, 4, 6, 7, 8, 9, 15, 16, 22	2.7512	1, 4, 6, 7, 8, 9, 15, 16, 22	0.0690	97.49
	100	1, 4, 6, 7, 8, 9, 15, 16, 22	2.8142	1, 4, 6, 7, 8, 9, 15, 16, 22	0.0830	97.05
Dermatology	20	1, 2, 3, 4, 5, 14, 16, 18, 19	1.8261	1, 2, 3, 4, 5, 14, 16, 18, 19	0.2070	88.66
	40	1, 2, 3, 4, 5, 14, 16, 18, 19	2.2561	1, 2, 3, 4, 5, 14, 16, 18, 19	0.2020	91.05
	60	2, 3, 4, 7, 9, 16, 17, 19, 28	2.4941	1, 2, 3, 4, 5, 14, 16, 18, 19	0.2560	89.73
	80	1, 2, 3, 4, 5, 16, 19, 28, 31, 32	3.3382	1, 2, 3, 4, 5, 14, 16, 18, 19	0.4080	87.78
	100	1, 2, 3, 4, 5, 16, 19, 28, 31, 32	3.4612	1, 2, 3, 4, 5, 14, 16, 18, 19	0.4040	88.33
	20	1, 5, 10, 11, 16, 18, 21, 30, 32, 35	260.76	1, 5, 10, 11, 16, 18, 21, 30, 32, 35	5.5803	97.86
	40	1, 5, 10, 11, 16, 18, 21, 30, 32, 35	316.93	1, 5, 10, 11, 16, 18, 21, 30, 32, 35	9.3805	97.04
Splice	60	1, 5, 10, 11, 18, 21, 30, 32, 35, 46	377.33	1, 5, 10, 11, 16, 18, 21, 30, 32, 35	14.531	96.15
	80	1, 5, 10, 11, 18, 21, 30, 32, 35, 46	430.35	1, 5, 10, 11, 16, 18, 21, 30, 32, 35	20.661	95.20
	100	1, 5, 10, 11, 18, 21, 30, 32, 35, 46	448.15	1, 5, 10, 11, 16, 18, 21, 30, 32, 35	22.701	94.93
Kr-vs-kp	20	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	13.191	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	6.5604	50.27
	40	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	23.361	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	9.7206	58.39
	60	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	30.904	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	13.011	57.90
	80	1, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	57.224	1, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	17.101	70.12
	100	1, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	88.898	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	18.701	78.96
Mushroom	20	1, 2, 3, 5, 9	15.322	1, 2, 3, 5, 9	1.8410	87.98
	40	1, 2, 3, 5, 9	24.901	1, 2, 3, 5, 9	2.9520	88.14
	60	3, 5, 20	37.889	3, 5, 20	4.5300	88.04
	80	3, 5, 16, 20	94.591	3, 5, 20	7.7205	91.84
	100	3, 5, 16, 20	159.75	3, 5, 20	10.079	93.69
Ticdata2000	20	2, 5, 7, 15, 17, 30, 38, 43, 44, 45, 46, 47, 48, 49, 51	867.06	2, 5, 7, 15, 17, 30, 38, 43, 44, 45, 46, 47, 48, 49, 51	19.251	97.78
	40	2, 3, 5, 15, 31, 37, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	1283.7	2, 5, 7, 15, 17, 30, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	37.922	97.05
	60	2, 5, 9, 14, 18, 31, 39, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	1993.7	2, 5, 7, 15, 17, 30, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	60.763	96.95
	80	2, 3, 5, 15, 31, 38, 39, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	3156.9	2, 5, 7, 15, 17, 30, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	112.08	96.45
	100	2, 5, 7, 15, 17, 31, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 63, 64, 68, 80, 83	4886.8	2, 5, 7, 15, 17, 30, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 63, 64, 68, 80, 83	213.90	95.62

Table 7
Comparison of algorithms for reduct computation based on combination entropy.

Data sets	*SIA* (%)	*TA_RED*		*DIA_RED*		*PIT* (%)
		Reduct	Time/s	Reduct	Time/s	
Backup-large	20	1, 4, 5, 7, 8, 10, 13, 16, 22	1.6651	1, 4, 5, 7, 8, 10, 13, 16, 22	0.0624	96.25
	40	1, 4, 5, 7, 8, 10, 13, 16, 22	1.9911	1, 4, 5, 7, 8, 10, 13, 16, 22	0.0312	98.43
	60	1, 4, 5, 7, 8, 10, 13, 16, 22	2.3321	1, 4, 5, 7, 8, 10, 13, 16, 22	0.0468	97.99
	80	1, 4, 5, 7, 8, 10, 13, 16, 22	2.6782	1, 4, 5, 7, 8, 10, 13, 16, 22	0.0624	97.67
	100	1, 4, 5, 7, 8, 10, 13, 16, 22	2.7682	1, 4, 5, 7, 8, 10, 13, 16, 22	0.0780	97.18
Dermatology	20	1, 2, 3, 4, 5, 14, 16, 18, 19	1.8311	1, 2, 3, 4, 5, 14, 16, 18, 19	0.1212	93.38
	40	1, 2, 3, 4, 5, 14, 16, 18, 19	2.2401	1, 2, 3, 4, 5, 14, 16, 18, 19	0.2480	88.93
	60	1, 2, 3, 4, 5, 7, 14, 16, 18, 19	2.9262	1, 2, 3, 4, 5, 14, 16, 18, 19	0.2568	91.22
	80	1, 2, 3, 4, 14, 16, 18, 19, 31, 32	3.4372	1, 2, 3, 4, 5, 14, 16, 18, 19	0.3980	88.42
	100	1, 2, 3, 4, 14, 16, 18, 19, 31, 32	3.5472	1, 2, 3, 4, 5, 14, 16, 18, 19	0.3980	88.78
Splice	20	2, 4, 6, 8, 10, 18, 22, 30, 33, 35	249.96	2, 4, 6, 8, 10, 18, 22, 30, 33, 35	5.1480	97.94
	40	2, 9, 10, 12, 19, 22, 25, 30, 39, 43	306.59	2, 4, 6, 8, 10, 18, 22, 30, 33, 35	8.5800	97.20
	60	1, 3, 8, 10, 18, 19, 30, 34, 40, 50	363.03	2, 4, 6, 8, 10, 18, 22, 30, 33, 35	13.260	96.35
	80	1, 3, 4, 10, 18, 26, 30, 35, 50, 57	420.22	2, 4, 6, 8, 10, 18, 22, 30, 33, 35	19.188	95.43
	100	1, 4, 9, 10, 14, 20, 26, 30, 37, 59	435.62	2, 4, 6, 8, 10, 18, 22, 30, 33, 35	20.748	95.24
Kr-vs-kp	20	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	13.042	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	5.9280	54.55

(*continued on next page*)

Table 7 (*continued*)

Data sets	*SIA* (%)	*TA_RED* Reduct	Time/s	*DIA_RED* Reduct	Time/s	*PIT* (%)
	40	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	22.963	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	9.2040	59.92
	60	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	29.780	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	12.324	58.62
	80	1, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	54.460	1, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	15.912	70.78
	100	1, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	85.254	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	17.472	79.51
Mushroom	20	1, 2, 3, 5, 9	14.789	1, 2, 3, 5, 9	2.1840	85.23
	40	1, 2, 3, 5, 9	24.383	1, 2, 3, 5, 9	3.2760	86.56
	60	3, 5, 20	36.395	3, 5, 20	4.9920	86.28
	80	3, 5, 16, 20	91.947	3, 5, 20	8.1120	91.18
	100	3, 5, 16, 20	110.30	3, 5, 20	9.8335	91.08
Ticdata2000	20	2, 5, 15, 23, 26, 27, 29, 30, 43, 44, 45, 46, 47, 48, 49, 51	1022.6	2, 5, 15, 23, 26, 27, 29, 30, 43, 44, 45, 46, 47, 48, 49, 51	17.316	98.31
	40	2, 3, 5, 15, 31, 37, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	1350.1	2, 5, 15, 23, 26, 27, 29, 30, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	34.320	97.46
	60	2, 5, 14, 15, 18, 19, 23, 31, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	2297.5	2, 5, 15, 23, 26, 27, 29, 30, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	98.124	97.58
	80	2, 3, 5, 15, 31, 38, 39, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	3233.7	2, 5, 15, 23, 26, 27, 29, 30, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	112.08	96.97
	100	2, 5, 7, 15, 17, 31, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 63, 64, 68, 80, 83	5025.7	2, 5, 15, 23, 26, 27, 29, 30, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 63, 64, 68, 80, 83	191.72	96.19

Table 8
Comparison of algorithms for reduct computation based on Shannon's entropy.

Data sets	*SIA* (%)	*TA_RED* Reduct	Time/s	*DIA_RED* Reduct	Time/s	*PIT* (%)
Backup-large	20	1, 2, 4, 6, 7, 9, 13, 16, 22	1.7092	1, 2, 4, 6, 7, 9, 13, 16, 22	0.0310	98.18
	40	1, 2, 4, 6, 7, 9, 13, 16, 22	2.0100	1, 2, 4, 6, 7, 9, 13, 16, 22	0.0380	98.11
	60	1, 3, 4, 6, 7, 8, 10, 16, 29	2.3940	1, 3, 4, 6, 7, 8, 10, 16, 29	0.0510	97.87
	80	1, 3, 4, 6, 7, 8, 10, 16, 29	2.7556	1, 3, 4, 6, 7, 8, 10, 16, 29	0.0690	97.50
	100	1, 3, 4, 6, 7, 8, 10, 16, 29	2.9168	1, 3, 4, 6, 7, 8, 10, 16, 29	0.0730	97.50
Dermatology	20	1, 2, 3, 4, 5, 14, 16, 18, 19	1.8900	1, 2, 3, 4, 5, 14, 16, 18, 19	0.1550	91.80
	40	1, 2, 3, 4, 5, 14, 16, 18, 19	2.3100	1, 2, 3, 4, 5, 14, 16, 18, 19	0.2140	90.74
	60	3, 4, 5, 7, 9, 13, 15, 21, 26, 27, 28	2.4900	1, 2, 3, 4, 5, 14, 16, 18, 19	0.3590	85.58
	80	1, 2, 4, 5, 15, 21, 26, 27, 28, 31, 32	3.3400	1, 2, 3, 4, 5, 14, 16, 18, 19	0.3590	89.25
	100	1, 2, 4, 5, 15, 21, 26, 28, 31, 32, 33	3.4400	1, 2, 3, 4, 5, 14, 16, 18, 19	0.4170	87.88
Splice	20	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	282.79	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	8.6325	96.95
	40	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	337.59	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	13.358	96.04
	60	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	400.81	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	19.408	95.16
	80	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	465.94	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	28.512	93.88
	100	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	479.89	3, 5, 6, 13, 21, 28, 29, 30, 31, 32, 35	40.313	91.60
Kr-vs-kp	20	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	13.485	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22	6.4304	52.31
	40	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	23.648	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26	9.6906	59.02
	60	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	30.904	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30	13.001	57.93
	80	1, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	57.525	1, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34	17.261	70.00
	100	1, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	90.984	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36	18.831	79.30
Mushroom	20	1, 3, 5, 9	13.437	1, 3, 5, 9	2.2301	83.40
	40	1, 3, 5, 9	22.256	1, 3, 5, 9	3.6802	83.46
	60	3, 5, 20	36.703	3, 5, 20	5.7603	84.31
	80	3, 5, 16, 20	97.646	3, 5, 20	9.0605	90.72
	100	3, 5, 16, 20	131.53	3, 5, 20	12.324	90.63
Ticdata2000	20	2, 5, 15, 18, 25, 30, 38, 43, 44, 45, 46, 47, 48, 49, 51	868.36	2, 5, 15, 18, 25, 30, 38, 43, 44, 45, 46, 47, 48, 49, 51	20.901	97.59
	40	2, 5, 9, 14, 15, 18, 27, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	1292.5	2, 5, 15, 18, 25, 30, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59	37.182	97.12
	60	2, 5, 7, 14, 18, 30, 40, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	1982.9	2, 5, 15, 18, 25, 30, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	63.544	96.79
	80	2, 5, 7, 14, 15, 18, 39, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	3082.4	2, 5, 15, 18, 25, 30, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 62, 63, 64, 68	110.02	96.43
	100	2, 5, 9, 18, 31, 37, 40, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 63, 64, 68, 80, 83	4708.1	2, 5, 15, 18, 25, 30, 38, 43, 44, 45, 47, 48, 49, 54, 55, 57, 58, 59, 61, 63, 64, 68, 80, 83	211.66	95.50

Table 9
Comparison of evaluation measures based on complementary entropy.

Data sets	*SIA* (%)	*TA_RED*			*DIA_RED*		
		Entropy	AQ	AP	Entropy	AQ	AP
Backup-large	20	0.0000	0.9055	0.8274	0.0000	0.9055	0.8274
	40	0.0000	0.9055	0.8274	0.0000	0.9055	0.8274
	60	0.0000	0.9055	0.8274	0.0000	0.9055	0.8274
	80	0.0000	0.9055	0.8274	0.0000	0.9055	0.8274
	100	0.0000	0.9055	0.8274	0.0000	0.9055	0.8274
Dermatology	20	0.0000	0.9727	0.9468	0.0000	0.9727	0.9468
	40	0.0000	0.9727	0.9468	0.0000	0.9727	0.9468
	60	0.0000	0.9645	0.9314	0.0000	0.9727	0.9468
	80	0.0000	0.9863	0.9730	0.0000	0.9727	0.9468
	100	0.0000	0.9863	0.9730	0.0000	0.9727	0.9468
Splice	20	1.3082E−07	0.9912	0.9826	1.3082E−07	0.9912	0.9826
	40	1.3082E−07	0.9912	0.9826	1.3082E−07	0.9912	0.9826
	60	1.3082E−07	0.9940	0.9882	1.3082E−07	0.9912	0.9826
	80	1.3082E−07	0.9940	0.9882	1.3082E−07	0.9912	0.9826
	100	1.3082E−07	0.9940	0.9882	1.3082E−07	0.9912	0.9826
Kr-vs-kp	20	0.0006	0.6439	0.4749	0.0006	0.6439	0.4749
	40	0.0002	0.7003	0.5388	0.0002	0.7003	0.5388
	60	0.0001	0.7412	0.5889	0.0001	0.7412	0.5889
	80	7.8321E−06	0.9712	0.9440	7.8321E−06	0.9712	0.9440
	100	0.0000	0.9994	0.9987	0.0000	0.9994	0.9987
Mushroom	20	5.0228E−07	0.9848	0.9700	5.0228E−07	0.9848	0.9700
	40	5.0228E−07	0.9848	0.9700	5.0228E−07	0.9848	0.9700
	60	0.0000	0.9433	0.8927	0.0000	0.9433	0.8927
	80	0.0000	0.9433	0.8927	0.0000	0.9433	0.8927
	100	0.0000	0.9433	0.8927	0.0000	0.9433	0.8927
Ticdata2000	20	1.6226E−05	0.9304	0.8699	1.6226E−05	0.9304	0.8699
	40	6.4315E−06	0.9425	0.8912	6.4315E−06	0.9425	0.8912
	60	4.3663E−06	0.9753	0.9517	4.3663E−06	0.9756	0.9524
	80	4.3663E−06	0.9756	0.9524	4.3663E−06	0.9756	0.9524
	100	4.1893E−06	0.9766	0.9543	4.1893E−06	0.9766	0.9543

Table 10
Comparison of evaluation measures based on combination entropy.

Data sets	*SIA* (%)	*TA_RED*			*DIA_RED*		
		Entropy	AQ	AP	Entropy	AQ	AP
Backup-large	20	0.0000	0.9023	0.8174	0.0000	0.9023	0.8174
	40	0.0000	0.9023	0.8174	0.0000	0.9023	0.8174
	60	0.0000	0.9023	0.8174	0.0000	0.9023	0.8174
	80	0.0000	0.9023	0.8174	0.0000	0.9023	0.8174
	100	0.0000	0.9023	0.8174	0.0000	0.9023	0.8174
Dermatology	20	0.0000	0.9727	0.9468	0.0000	0.9727	0.9468
	40	0.0000	0.9727	0.9468	0.0000	0.9727	0.9468
	60	0.0000	0.9781	0.9572	0.0000	0.9727	0.9468
	80	0.0000	0.9945	0.9891	0.0000	0.9727	0.9468
	100	0.0000	0.9945	0.9891	0.0000	0.9727	0.9468
Splice	20	6.6933E−11	0.9940	0.9882	6.6933E−11	0.9940	0.9882
	40	6.6933E−11	0.9950	0.9900	6.6933E−11	0.9940	0.9882
	60	6.6933E−11	0.9937	0.9875	6.6933E−11	0.9940	0.9882
	80	6.6933E−11	0.9909	0.9820	6.6933E−11	0.9940	0.9882
	100	6.6933E−11	0.9900	0.9801	6.6933E−11	0.9940	0.9882
Kr-vs-kp	20	5.3831E−06	0.6439	0.4749	5.3831E−06	0.6439	0.4749
	40	3.5955E−07	0.7003	0.5388	3.5955E−07	0.7003	0.5388
	60	1.9341E−07	0.7412	0.5889	1.9341E−07	0.7412	0.5889
	80	4.9027E−09	0.9712	0.9440	4.9027E−09	0.9712	0.9440
	100	0.0000	0.9994	0.9987	0.0000	0.9994	0.9987
Mushroom	20	4.4505E−10	0.9848	0.9700	4.4505E−10	0.9848	0.9700
	40	4.4505E−10	0.9848	0.9700	4.4505E−10	0.9848	0.9700
	60	0.0000	0.9433	0.8927	0.0000	0.9433	0.8927
	80	0.0000	0.9433	0.8927	0.0000	0.9433	0.8927
	100	0.0000	0.9433	0.8927	0.0000	0.9433	0.8927
Ticdata2000	20	1.3573E−08	0.9304	0.8699	1.3573E−08	0.9304	0.8699
	40	3.4870E−09	0.9425	0.8912	3.4870E−09	0.9425	0.8912
	60	2.2300E−09	0.9756	0.9524	2.2300E−09	0.9756	0.9524
	80	2.2300E−09	0.9756	0.9524	2.2300E−09	0.9756	0.9524
	100	2.1692E−09	0.9766	0.9543	2.1692E−09	0.9766	0.9543

Q1 Please cite this article in press as: F. Wang et al., Attribute reduction: A dimension incremental strategy, Knowl. Based Syst. (2012), http://dx.doi.org/10.1016/j.knosys.2012.10.010

As mentioned in Section 1 (Introduction), existing research on knowledge updating caused by the variation of attributes mainly focuses on updating approximation operators. However, dimension incremental algorithms for reduct (or core) computation have not yet been discussed so far. Hence, to illustrate effectiveness and efficiency of the proposed algorithms, we compare them with the traditional algorithms based on information entropy for core and reduct. Section 6.1 introduces the comparison of algorithms for core computation, and the comparison of algorithms for reduct computation is shown in Section 6.2.

6.1. Effectiveness and efficiency for core computation

This subsection is to illustrate effectiveness and efficiency of the incremental algorithm *DIA_CORE* by comparing it with the traditional algorithm for core computation (*TA_CORE*). For each data set in Table 2, 50% conditional attributes and the decision attribute are selected as the basic table. Then, from the remaining 50% conditional attributes, 20%, 40%, ... , 100% are selected, in order, as incremental attribute sets. When each incremental attribute set is added to the basic table, algorithms *TA_CORE* and *DIA_CORE* are used to update the core respectively. The effectiveness and efficiency of *TA_CORE* and *DIA_CORE* are demonstrated by comparing their computational time and found core. Experimental results are shown in Tables 3–5. For simplicity, *Size of Incremental Attribute Set* is written as *SIA*, and *Percentage Improvement of Computational Time* is written as *PIT* in these tables.

Based on the three entropies, experimental results in Tables 3–5 show that core attributes of each data set found by the two algorithms (*DIA_CORE* and *TA_CORE*) are identical to each other. However, the computational time of *DIA_CORE* is much smaller than that of *TA_CORE*. In other words, comparing with *TA_CORE*, the incremental algorithm *DIA_CORE* can find the correct core of a given data set in a much shorter time. Hence, experimental results show that the proposed incremental algorithm for core computation is effective and efficient.

6.2. Effectiveness and efficiency for reduct computation

In this subsection, to illustrate effectiveness and efficiency of the incremental algorithm *DIA_RED*, we compare it with the traditional reduction algorithms (*TA_RED*) based on the three entropies. For each employed data set, 50% conditional attributes and the decision attribute are selected as the basic table. Then, from the remaining 50% conditional attributes, 20%, 40%, ... , 100% are selected as incremental attribute sets. When each incremental attribute set is added to the basic table, algorithms *TA_RED* and *DIA_RED* are used to update the reduct respectively. The effectiveness and efficiency of the incremental algorithm are demonstrated by comparing the their computational time and found reduct. Experimental results are shown in Tables 6–8. Similarly, *Size of Incremental Attribute Set* is written as *SIA*, and *Percentage Improvement of Computational Time* is written as *PIT* in these tables.

Experimental results in Tables 6–8 show that, compared with *TA_RED*, algorithm *DIA_RED* is much more efficiency. Especially, the percentage improvement of computational time better illustrates this conclusion. In view of that there are some difference between the reducts found by the two algorithms, two common evaluation measures in rough set are employed to evaluate the decision performance of reducts. The two measures are approximate classified precision and approximate classified quality, which are defined by Pawlak to describe the precision of approximate classification [23,24]. Evaluated results and entropies induced by the reducts are given in Tables 9–11.

Table 11
Comparison of evaluation measures based on Shannon's entropy.

Data sets	*SIA* (%)	*TA_RED*			*DIA_RED*		
		Entropy	AQ	AP	Entropy	AQ	AP
Backup-large	20	0.0000	0.9088	0.8328	0.0000	0.9088	0.8328
	40	0.0000	0.9088	0.8328	0.0000	0.9088	0.8328
	60	0.0000	0.9511	0.9068	0.0000	0.9511	0.9068
	80	0.0000	0.9511	0.9068	0.0000	0.9511	0.9068
	100	0.0000	0.9511	0.9068	0.0000	0.9511	0.9068
Dermatology	20	0.0000	0.9727	0.9468	0.0000	0.9727	0.9468
	40	0.0000	0.9727	0.9468	0.0000	0.9727	0.9468
	60	0.0000	0.9727	0.9468	0.0000	0.9727	0.9468
	80	0.0000	0.9918	0.9837	0.0000	0.9727	0.9468
	100	0.0000	0.9672	0.9365	0.0000	0.9727	0.9468
Splice	20	0.0002	0.9878	0.9758	0.0002	0.9878	0.9758
	40	0.0002	0.9878	0.9758	0.0002	0.9878	0.9758
	60	0.0002	0.9878	0.9758	0.0002	0.9878	0.9758
	80	0.0002	0.9878	0.9758	0.0002	0.9878	0.9758
Kr-vs-kp	100	0.0002	0.9878	0.9758	0.0002	0.9878	0.9758
	20	0.0917	0.6439	0.4749	0.0917	0.6439	0.4749
	40	0.0816	0.7003	0.5388	0.0816	0.7003	0.5388
	60	0.0701	0.7412	0.5889	0.0701	0.7412	0.5889
	80	0.0075	0.9712	0.9440	0.0075	0.9712	0.9440
	100	0.0000	0.9994	0.9987	0.0000	0.9994	0.9987
Mushroom	20	0.0004	0.9720	0.9455	0.0004	0.9720	0.9455
	40	0.0004	0.9720	0.9455	0.0004	0.9720	0.9455
	60	0.0000	0.9433	0.8927	0.0000	0.9433	0.8927
	80	0.0000	0.9433	0.8927	0.0000	0.9433	0.8927
	100	0.0000	0.9433	0.8927	0.0000	0.9433	0.8927
Ticdata2000	20	0.0183	0.9304	0.8699	0.0183	0.9304	0.8699
	40	0.0090	0.9421	0.8906	0.0090	0.9425	0.8912
	60	0.0063	0.9756	0.9524	0.0063	0.9756	0.9524
	80	0.0063	0.9756	0.9524	0.0063	0.9756	0.9524
	100	0.0060	0.9763	0.9537	0.0060	0.9766	0.9543

Definition 8. Let $S = (U, C \cup D)$ be a decision table and $U/D = \{X_1, X_2, \ldots, X_r\}$. The approximate classified precision of C with respect to D is defined as

$$AP_C(D) = \frac{|POS_C(D)|}{\sum_{i=1}^{r} |\overline{C}X_i|}. \tag{7}$$

Definition 9. Let $S = (U, C \cup D)$ be a decision table. The approximate classified quality of C with respect to D is defined as

$$AQ_C(D) = \frac{|POS_C(D)|}{|U|}. \tag{8}$$

In Tables 9–11, for each employed data set, entropies induced by the reducts found by the two algorithms are identical to each other. This indicates that *DIA_RED* can also find a reduct in the context of entropies. In these tables, evaluated results of the reducts found by the two algorithms are very close to each other, even identical on some data sets. For data sets *Dermatology* and *Splice* in Table 10, the evaluated results of *DIA_RED* are smaller than that of *TA_RED*. And for data sets *Ticdata2000* in Table 9, *Splice* in Table 10 and *Dermatology* in Table 11, the evaluated results of *DIA_RED* are bigger than that of *TA_RED*. Hence, experimental results show that, more commonly, algorithm *DIA_RED* can find a same reduct with *TA_RED*, and saves lots of computational time. In some cases, *DIA_RED* can efficiently find another reduct in the context of entropy, and the decision performance of this reduct is very close that of the one found by *TA_RED* without obvious superiority and inferiority.

7. Conclusions

In practices, many real data sets in databases may increase quickly not only in rows but also in columns. This paper developed a dimension incremental reduction algorithm based on information entropy for data sets with dynamically increasing attributes. Theoretical analysis and experimental results have shown that, compared with the traditional non-incremental reduction algorithm based on entropy, the proposed algorithm is effective and efficient. It is our wish that this study provides new views and thoughts on dealing with dynamic data sets in applications.

Acknowledgments

This work was supported by National Natural Science Fund of China (Nos. 71031006, 60903110, 70971080), Initial Special Research for 973 Program of China (973) (No. 2011CB311805), The Research Fund for the Doctoral Program of Higher Education (20101401110002), Soft Science Research Project of Shanxi Province 791 of China (No. 2010041054-01).

References

[1] C.C. Chan, A rough set approach to attribute generalization in data mining, Information Science 107 (1998) 169–176.
[2] Y. Cheng, The incremental method for fast computing the rough fuzzy approximations, Data Knowledge Engineering 70 (2011) 84–100.
[3] M. Dash, H. Liu, Consistency-based search in feature selection, Artificial Intelligence 151 (2003) 155–176.
[4] I. Düntsch, G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence 106 (1) (1998) 109–137.
[5] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems 17 (1990) 191–209.
[6] G. Gediga, I. Düntsch, Rough approximation quality revisited, Artificial Intelligence 132 (2001) 219–234.
[7] J.W. Guan, D.A. Bell, Rough computational methods for information systems, Artificial Intelligence 105 (1998) 77–103.
[8] T. Herawan, M.M. Deris, J.H. Abawajy, A rough set approach for selecting clustering attribute, Knowledge-Based Systems 23 (2010) 220–231.
[9] F. Hu, G.Y. Wang, H. Huang, Y. Wu, Incremental attribute reduction based on elementary sets, in: Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, Regina, Canada, 2005, pp. 185–193.
[10] X.H. Hu, N. Cercone, Learning in relational databases: a rough set approach, International Journal of Computational Intelligence 11 (2) (1995) 323–338.
[11] Q.H. Hu, Z.X. Xie, D.R. Yu, Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation, Pattern Recognition 40 (2007) 3509–3521.
[12] Q.H. Hu, D.R. Yu, J.F. Liu, C.X. Wu, Neighborhood rough set based heterogeneous feature subset selection, Information Science 178 (2008) 3577–3594.
[13] R. Jensen, Q. Shen, Fuzzy-rough sets assisted attribute selection, IEEE Transactions on Fuzzy Systems 15 (1) (2007) 73–89.
[14] M. Kryszkiewicz, P. Lasek, FUN: fast discovery of minimal sets of attributes functionally determining a decision attribute, Transactions on Rough Sets 9 (2008) 76–95.
[15] T.R. Li, D. Ruan, W. Geert, J. Song, Y. Xu, A rough sets based characteristic relation approach for dynamic attribute generalization in data mining, Knowledge-Based Systems 20 (5) (2007) 485–494.
[16] J.Y. Liang, C.Y. Dang, K.S. Chin, C.M. Yam Richard, A new method for measuring uncertainty and fuzziness in rough set theory, International Journal of General Systems 31 (4) (2002) 331–342.
[17] J.Y. Liang, Z.Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (1) (2004) 37–46.
[18] J.Y. Liang, W. Wei, Y.H. Qian, An incremental approach to computation of a core based on conditional entropy, Chinese Journal of System Engineering Theory and Practice 4 (2008) 81–89.
[19] J.S. Mi, Y. Leung, W.Z. Wu, An uncertainty measure in partition-based fuzzy rough sets, International Journal General Systems 34 (2005) 77–90.
[20] J.S. Mi, W.Z. Wu, X.W. Zhang, Approaches to knowledge reduction based on variable precision rough set model, Information Sciences 159 (3–4) (2004) 255–272.
[21] H.S. Nguyen, D.W. Slezak, Approximate reducts and association rules correspondence and complexity results, Lecture Notes in Artificial Intelligence 1711 (1999) 137C145.
[22] M. Orlowska, Maintenance of knowledge in dynamic information systems, in: R. Slowinski (Ed.), Proceeding of The Intelligent Decision Support, Handbook of Applications and Advances of the Rough Set Theory, Kluwer Academic Publishers, Dordrecht, 1992, pp. 315–330.
[23] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Boston, 1991.
[24] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (1) (2007) 3–27.
[25] Z. Pawlak, A. Skowron, Rough sets: some extensions, Information Sciences 177 (2007) 28–40.
[26] W. Pedrycz, G. Vukovich, Feature analysis through information granulation and fuzzy sets, Pattern Recognition 35 (2002) 825–834.
[27] Y.H. Qian, J.Y. Liang, Combination entropy and combination granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 16 (2) (2008) 179–193.
[28] Y.H. Qian, J.Y. Liang, W. Pedrycz, C.Y. Dang, Positive approximation: an accelerator for attribute reduction in rough set theory, Artificial Intelligence 174 (2010) 597–618.
[29] D. Slezak, Approximate entropy reducts, Fundamenta Informaticae 53 (3–4) (2002) 365–390.
[30] C.E. Shannon, The mathematical theory of communication, The Bell System Technical Journal 27 (3–4) (1948) 373–423.
[31] A. Skowron, C. Rauszer, The discernibility matrices and functions in information systems, in: R. Słowiński (Ed.), Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publisher, Dordrecht, 1992.
[32] G.Y. Wang, H. Yu, D.C. Yang, Decision table reduction based on conditional information entropy, Chinese Journal of Computer 25 (7) (2002) 759–766.
[33] W.Z. Wu, W.X. Zhang, Constructive and axiomatic approaches of fuzzy approximation operators, Information Sciences 159 (2004) 233–254.
[34] Z.B. Xu, J.Y. Liang, C.Y. Dang, K.S. Chin, Inclusion degree: a perspective on measures for rough set data analysis, Information Sciences 141 (2002) 227–236.
[35] Z.Y. Xu, Z.P. Liu, B.R. Yang, W. Song, A quick attribute reduction algorithm with complexity of max $(O(|C||U|), O(|C|^2|U/C|))$, Chinese Journal of Computer 29 (3) (2006) 391–398.
[36] M. Yang, An incremental updating algorithm for attributes reduction based on the improved discernibility matrix, Chinese Journal of Computers 30 (5) (2007) 815–822.
[37] X.B. Yang, M. Zhang, H.L. Dou, J.Y. Yang, Neighborhood systems-based rough sets in incomplete information system, Knowledge-Based Systems 24 (6) (2011) 858–867.
[38] Y.Y. Yao, Information granulation and rough set approximation, International Journal of Intelligent Systems 16 (2001) 87–104.
[39] Y.Y. Yao, Y. Zhao, Attribute reduction in decisiontheoretic rough set models, Information Sciences 178 (2008) 3356–3373.

[40] Y.Y. Yao, Decision-theoretic rough set models, Lecture Notes in Artificial Intelligence 4481 (2007) 1–12.
[41] A. Zeng, D. Pan, Q.L. Zheng, H. Peng, Knowledge acquisition based on rough set theory and principal component analysis, IEEE Intelligent Systems (2006) 78–85.
[42] W. Ziarko, Variable precision rough set model, Journal of Computer and System Science 46 (1) (1993) 39–59.
[43] J.B. Zhang, T.R. Li, D. Liu, An approach for incremental updating approximations in variable precision rough sets while attribute generalizing, in: Proc. 2010 IEEE Int. Conf. on Intelligent Systems and Knowledge Engineering, Hangzhou, China, 2010, pp. 77–81.
[44] W.X. Zhang, J.S. Mi, W.Z. Wu, Approaches to knowledge reductions in inconsistent systems, International Journal of Intelligent Systems 21 (9) (2003) 989–1000.

Q1 Please cite this article in press as: F. Wang et al., Attribute reduction: A dimension incremental strategy, Knowl. Based Syst. (2012), http://dx.doi.org/10.1016/j.knosys.2012.10.010