Incremental Method of Generating Decision Implication Canonical Basis

Shaoxia Zhang
Shanxi University of Finance and Economics
Deyu Li (lidysxu@163.com)
Shanxi University https://orcid.org/0000-0001-5927-3060
Yanhui Zhai
Shanxi University

Research Article

Keywords: formal concept analysis, decision premise, decision implication canonical basis, incremental method

Posted Date: May 24th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-524907/v1

Version of Record: A version of this preprint was published at Soft Computing on November 22nd, 2021. See the published version at https://doi.org/10.1007/s00500-021-06452-3.

Incremental method of generating decision implication canonical basis

Shaoxia Zhang[a], Deyu Li[b,c,*], Yanhui Zhai[b,c]

[a] College of Information, Shanxi University of Finance and Economics, Taiyuan Shanxi 030006, China
[b] School of Computer and Information Technology, Shanxi University, Taiyuan Shanxi 030006, China
[c] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan Shanxi 030006, China

Abstract

Decision implication is an elementary representation of decision knowledge in formal concept analysis. Decision implication canonical basis (DICB), a set of decision implications with completeness and non-redundancy, is the most compact representation of decision implications. The method based on true premises (MBTP) for DICB generation is the most efficient one at present. In practical applications, however, data is always changing dynamically, and MBTP has to re-generate inefficiently the whole DICB. This paper proposes an incremental algorithm for DICB generation, which obtains a new DICB just by modifying and updating the existing one. Experimental results verify that when the samples in data are much more than condition attributes, which is actually a general case in practical applications, the incremental algorithm is significantly superior to MBTP. Furthermore, we conclude that, even for the data in which samples are less than condition attributes, when new samples are continually added into data, the incremental algorithm must be also more efficient than MBTP, because the incremental algorithm just needs to modify the existing DICB, which is only a part of work of MBTP.

Keywords: formal concept analysis; decision premise; decision implication canonical basis; incremental method

1. Introduction

1.1. A brief review of formal concept analysis

Formal Concept Analysis (FCA) is an order-theoretic method for concept analysis and visualization, pioneered by Wille [38] in the mid-80s. In essence, FCA comes from a philosophical understanding of a concept, which is viewed as a unit of thought constituted by its extent and intent. The extent of a concept is a collection of all objects belonging to that concept and its intent is the set of all attributes common to all the objects of the extent. FCA is capable of presenting the relationship between intent and extent and visualizing the generalization and specialization of concepts by means of concept lattice. Because of its strengths, FCA attracts the interest of researchers in the fields of data mining [27, 1, 5, 36, 2, 29, 52], social networks [32], cognition-based concept learning [41, 40, 16, 23, 54, 14], knowledge reduction [11, 4, 35, 25, 17, 12, 13, 6] and decision making [50, 51].

1.2. A brief review of decision implication logic

The study of knowledge representation and reasoning in FCA is that of attribute implication [42, 8, 30, 43, 26, 37], which has been widely applied to areas of information acquisition, text mining, software engineering and machine learning [3, 7]. Decision implication, revealing the dependency between conditions and consequences (causes and effects), is an elementary representation of decision knowledge in

*Corresponding author

Email addresses: zhangshaoxia_sxu@163.com (Shaoxia Zhang), lidysxu@163.com (Deyu Li), chai_yanhui@163.com (Yanhui Zhai)

FCA. A decision implication is defined as a formula $A \to B$, meaning that if the conditions in A are satisfied, then the conclusions in B hold. Compared with other classifiers, decision implication has an equal or better classification ability [33, 10, 37].

In practical applications, however, a small-scale data may produce a large number of decision implications. Thus, for the sake of easy storage and efficient processing, it is widely recognized that decision implications should be deduced from decision implication basis (a complete set of decision implications) rather than being computed from data [31, 47, 46]. To achieve this, Qu et al. [31] introduced α-decision inference rule, meaning that new decision implications can be deduced from decision implications by amplifying their premises or reducing their consequences. By using α-decision inference rule, one can obtain a decision implication set with relative completeness and non-redundancy. Li et al. [19, 21, 20, 22, 18] discussed the application of α-decision inference rule in decision contexts, incomplete decision contexts and real decision contexts. Nevertheless, all the above studies fail to present an integrated logic description of decision implication. Hence, Zhai et al. [45, 15, 47, 46] researched decision implication logic from semantical aspect and syntactical aspect. The semantical aspect accounts for the soundness of decision implications and the non-redundancy and completeness of decision implication sets. In the syntactical aspect, two inference rules, namely *Augmentation* and *Combination* were proposed, which were proven to be sound, complete and non-redundant w.r.t. the semantical aspect.

1.3. A brief review of decision implications in decision contexts

Based on decision implication logic, Zhai [46, 45, 47] proposed the most compact set of decision implications, decision implication canonical basis (DICB), in decision contexts. This basis takes decision premises as its premises and the closures of decision premises as its consequences. DICB keeps all decision implications in data by the least amount of decision implications, that is because this basis is complete and non-redundant w.r.t. decision implication logic, and more importantly, it is optimal, i.e., it is of minimal cardinality among all complete sets of decision implications [46, 45, 47]. Starting from DICB, one can obtain all decision implications in data by iteratively applying *Augmentation* and *Combination*. Furthermore, DICB has been proven to have the strongest strength of knowledge representation, comparing with other models of decision implications, such as concept rule and granule rule [39, 53, 24]. Other researches about decision implications and fuzzy decision implications can be found in [34, 44, 39, 49, 48, 28].

1.4. Motivations and contributions of the paper

An efficient method for DICB generation is essential for decision implication-based decision knowledge representation and reasoning. Zhai et al. [46] proposed a minimal generator-based method to generate DICB, which is, however, of exponential time complexity [15]. Considering this shortcoming, Li et al. [15] put forward a method based on true premises (abbreviated to MBTP). MBTP has a polynomial time complexity; and by experiments, it is more efficient than the minimal generator-based method [15]. In practical applications, however, data is always changing dynamically [9]. In this case, MBTP always needs to re-generate the whole DICB, and thus does not applies to this situation.

This paper proposes an incremental method for DICB generation, which intends to obtain DICB just by updating the existing one. In this method, decision premises are clarified into four categories: unchanged decision premises, modified decision premises, invalid decision premises and new decision premises. We study their properties and renewal mechanisms, by which, the existing DICB can be modified and then a new DICB is achieved. Experimental results verify that when the samples are much more than condition attributes, which is actually a general case in practical applications, the incremental algorithm is significantly superior to MBTP. Furthermore, we conclude that, even for the data in which objects are less than condition attributes, when new samples are continually added into data (such as new purchase records being added moment by moment into the database of a supermarket), MBTP still re-generates the whole DICB; by contrast, the incremental algorithm just needs to modify the existing DICB, which is only a part of work of MBTP, and thus is also more efficient than MBTP.

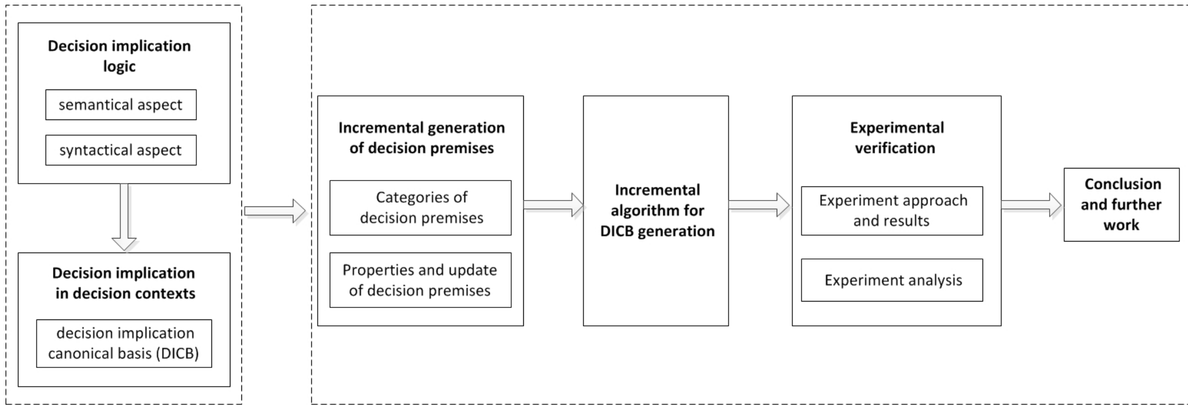

Figure 1: The structure of the article

1.5. The arrangement of the paper

This article is organized as follows. Section 2 introduces decision implication logic. Section 3 presents decision implication in decision contexts. Section 4 studies the incremental generation of decision premises. Section 5 proposes an incremental algorithm of generating DICB. Section 6 conducts an experiment to compare the performance of the incremental algorithm and MBTP. Conclusions and further work end the paper in Section 7. In what follows, we give the overall structure diagram of this article, as shown in Figure 1.

2. Decision implication logic

Decision implication [47, 49, 48, 15, 44, 45, 46], revealing the dependency between premises and consequences (causes and effects), is defined as a formula between condition attributes and decision attributes.

Definition 1 ([45])**.** Let C be a set of condition attributes and D be a set of decision attributes such that $C \cap D = \emptyset$. If $A \subseteq C$ and $B \subseteq D$, then $A \to B$ is called a decision implication, where A is the premise of $A \to B$ and B is the consequence.

Decision implication logic gives the semantical and syntactical description of decision implications. The semantical aspect studies the completeness and non-redundancy of decision implication sets [45, 47].

Definition 2 ([45])**.** Let C be a set of condition attributes, D be a set of decision attributes, and L and L_1 be sets of decision implications.

(1) For a set $T \subseteq C \cup D$ and a decision implication $A \to B$, if $A \nsubseteq T \cap C$ or $B \subseteq T \cap D$, then T respects $A \to B$ (or T is a model of $A \to B$), denoted by $T \models A \to B$. If for any $A_1 \to B_1 \in L$, $T \models A_1 \to B_1$ holds, then T respects L, denoted by $T \models L$.

(2) For a decision implication $A \to B$, if for any $T \subseteq C \cup D$, $T \models L$ implies $T \models A \to B$, then $A \to B$ can be semantically deduced from L, denoted by $L \vdash A \to B$. If for any $A_1 \to B_1 \in L_1$, $L \vdash A_1 \to B_1$ holds, then L_1 can be semantically deduced from L, denoted by $L \vdash L_1$.

(3) If for any $A \to B \in L$, $L \setminus \{A \to B\} \nvdash A \to B$ holds, then L is non-redundant.

(4) If any $A \to B$ that can be semantically deduced from L is contained in L, then L is closed.

(5) If L is closed, $L_1 \subseteq L$ and $L_1 \vdash L$, then L_1 is complete w.r.t. L.

For a given dataset, the soundness of a decision implication means that the decision implication is valid in the dataset. The completeness of a set of decision implications means that all valid decision implications can be deduced from the set. A set of decision implications is non-redundant if, in the set, no valid decision implications can be deduced from the other decision implications.

In syntactical aspect[45, 47], two inference rules *Augmentation* and *Combination* are proposed, and their soundness, completeness and redundancy of semantic compatibility are proved.

Augmentation :

$$\frac{A \to B, A_1 \supseteq A, B_1 \subseteq B}{A_1 \to B_1}$$

Combination :

$$\frac{A \to B, A_1 \to B_1}{A \cup A_1 \to B \cup B_1}$$

Theorem 1 ([45])**.** *Augmentation and Combination are sound, i.e.,*

(1) If $A \to B$, $A_1 \supseteq A$ and $B_1 \subseteq B$, then $A \to B \vdash A_1 \to B_1$;

(2) If $A \to B$ and $A_1 \to B_1$, then $\{A \to B, A_1 \to B_1\} \vdash A \cup A_1 \to B \cup B_1$.

Theorem 2 ([45])**.** *Augmentation and Combination are complete w.r.t. the semantical aspect, i.e., for any closed set of decision implication$L$ and a complete set $L_1 \subseteq L$, all decision implications in L can be obtained from L_1, by applying Augmentation and Combination.*

Theorem 3 ([45])**.** *Augmentation and Combination are non-redundancy, i.e., they cannot be replaced by each other in deduction process.*

3. Decision implication in decision contexts

Zhai et al. [46, 30, 31, 19, 20, 18, 17, 15] studied decision implication in decision contexts, and proposed the most compact set of decision implications, i.e., decision implication canonical basis [46].

Decision context in firstly introduced in the following.

Definition 3 ([45])**.** A decision context is a triple $K = (G, C \cup D, I_C \cup I_D)$, where G is the object set, C is the condition attribute set, and D is the decision attribute set such that $C \cap D = \emptyset$. In this case, $I_C \subseteq G \times C$ is the set of condition incidence relations and $I_D \subseteq G \times D$ is the set of decision incidence relations. For $g \in G$ and $m \in C \cup D$, $(g, m) \in I_C$ or $(g, m) \in I_D$ denotes "the object g has the attribute m".

A decision context can also be represented by a two-dimensional table, in which row headers are object names, column headers are attribute names, and a "1" indicates the row object g has the column attribute m, i.e., $(g, m) \in I_C$ or $(g, m) \in I_D$.

Example 1. A decision context $K = (G, C \cup D, I_C \cup I_D)$ is given in Table 1, where $G = \{g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8\}$, $C = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ and $D = \{d_1, d_2\}$.

Table 1: A decision context

	a_1	a_2	a_3	a_4	a_5	a_6	d_1	d_2
g_1	1	1	0	0	1	1	1	0
g_2	0	0	1	0	0	0	0	1
g_3	1	0	1	0	0	1	1	0
g_4	0	1	0	0	0	0	0	0
g_5	0	1	1	1	1	1	1	1
g_6	0	0	1	0	1	0	0	0

Definition 4 defines operators $(.)^C$ and $(.)^D$ in decision contexts.

Definition 4 ([45]). Let $K = (G, C \cup D, I_C \cup I_D)$ be a decision context. For sets $A_1 \subseteq C$, $A_2 \subseteq D$ and $B \subseteq G$:

$$\begin{aligned}
(1) A_1^C &= \{g \in G | (g, m) \in I_C, \forall m \in A_1\} \\
(2) A_2^D &= \{g \in G | (g, m) \in I_D, \forall m \in A_2\} \\
(3) B^C &= \{m \in C | (g, m) \in I_C, \forall g \in B\} \\
(4) B^D &= \{m \in D | (g, m) \in I_D, \forall g \in B\}
\end{aligned}$$

For $g \in G$ and $A \subseteq C$, $\{g\}^C$, $\{g\}^D$ and $(A^C)^D$ are abbreviated to g^C, g^D and A^{CD}, respectively. Actually, g^C and g^D are respectively the condition attribute set and decision attribute set of g.

Proposition 1states the properties of operators $(.)^C$ and $(.)^D$.

Proposition 1 ([38, 15]). *Let $K = (G, C \cup D, I_C \cup I_D)$ be a decision context. For sets $A, A_1, A_2 \subseteq C$ and $B, B_1, B_2 \subseteq G$, we have:*

(1)	$A_1 \subseteq A_2 \Rightarrow A_2^C \subseteq A_1^C$	(1′)	$B \subseteq B_2 \Rightarrow B_2^C \subseteq B_1^C$
(2)	$A \subseteq A^{CC}$	(2′)	$B \subseteq B^{CC}$
(3)	$A^C = A^{CCC}$	(3′)	$B^C = B^{CCC}$
(4)	$(A_1 \cup A_2)^C = A_1^C \cap A_2^C$	(4′)	$(B_1 \cup B_2)^C = B_1^C \cap B_2^C$
(5)	$A_1 \subseteq A_2 \Rightarrow A_1^{CD} \subseteq A_2^{CD}$		

Operator $(.)^D$ has the similar properties in Proposition 1. Definition 5 introduces decision implications in decision contexts.

Definition 5 ([46]). Let $K = (G, C \cup D, I_C \cup I_D)$ be a decision context. For $A \subseteq C$ and $B \subseteq D$, if $A^C \subseteq B^D$, then $A \to B$ is called a decision implication of K, where A is the premise and B is the consequence of A.

Example 2 (Continuing Example 1). Take the decision context in Example 1 as an example. It is verified that $\{a_1\} \to \{d_1\}$ is a decision implication of K, because $\{a_1\}^C = \{g_1, g_3, g_8\} \subseteq \{g_1, g_3, g_5, g_8\} = \{d_1\}^D$, i.e., $\{a_1\}^C \subseteq \{d_1\}^D$.

Proposition 2 ([46]). *Let $K = (G, C \cup D, I_C \cup I_D)$ be a decision context, $A \subseteq C$ and $B \subseteq D$. Then,*
(1) $A \to A^{CD}$ is a decision implication of K;
(2) $A \to B$ is a decision implication of K if and only if $B \subseteq A^{CD}$.

By Proposition 2, we can see that for set $A \subseteq C$, A^{CD} is the maximal consequence of A.

Zhai et al. [46, 15] defined the most compact set of decision implications, i.e., decision implication canonical basis.

Definition 6 ([46, 15]). Let $K = (G, C \cup D, I_C \cup I_D)$ be a decision context. A set $A \subseteq C$ is called a decision premise of K, if $A^{CD} \supset \Theta(A)$, where

$$\Theta(A) = \cup\{A_i^{CD} | A_i \subset A, A_i \, is \, a \, decision \, premise \, of \, K\} \tag{1}$$

Proposition 3. *Let $K = (G, C \cup D, I_C \cup I_D)$ be a decision context. A set $A \subseteq C$ is not a decision premise of K if and only if $A^{CD} = \Theta(A)$.*

Proof. It is easy to see that $A^{CD} \supseteq \Theta(A)$. Then, by Definition 6, the proof is straightforward. □

By Definition 6, we know that A is a decision premise if and only if A^{CD} contains more decisions than $\Theta(A)$. In other words, if A is a decision premise, one can only collect $\Theta(A)$ from the decision premise subsets of A, but the consequence of A is A^{CD}, which contains more conclusions than $\Theta(A)$. In this case, decision implication $A \to A^{CD}$ is indispensable to derive more decisions. On the contrary, if A is not a decision premise, by Proposition 3, A^{CD} is equal to $\Theta(A)$, meaning that A^{CD} can be collected from the decision premise subsets of A, and hence $A \to A^{CD}$ is not necessary.

Decision implication canonical basis is a set of decision implications which take decision premise A as premises and A^{CD} as consequences.

Definition 7 ([46, 15])**.** Let $K = (G, C \cup D, I_C \cup I_D)$ be a decision context. We call the set

$$O = \{A \to A^{CD} | A\, is\, a\, decision\, premise\, of\, K\}$$

the decision implication canonical basis (DICB) of K.

Decision implication canonical basis is proven to be complete, non-redundant and optimal w.r.t. decision implication logic [46], as shown in Theorem 4.

Theorem 4 ([46])**.** *Let $K = (G, C \cup D, I_C \cup I_D)$ be a decision context and O be the DICB of K. Then,:*

(1) O is complete, i.e., all decision implications in K can be obtained from O, by applying Augmentation and *Combination.*

(2) O is non-redundant, i.e., any decision implication in O cannot be obtained from others in O, by applying Augmentation and *Combination.*

(3) O is optimal, i.e., it is of minimal cardinality among all complete sets of decision implications.

Example 3 (Continuing Example 1)**.** The DICB of Table 1 is shown in Table 2.

Table 2: The DICB of Table 1

$\{a_1\} \to \{d_1\}$	$\{a_2, a_5\} \to \{d_1\}$
$\{a_6\} \to \{d_1\}$	$\{a_1, a_3, a_5\} \to \{d_1, d_2\}$
$\{a_4\} \to \{d_1, d_2\}$	$\{a_3, a_5, a_6\} \to \{d_1, d_2\}$
$\{a_2, a_3\} \to \{d_1, d_2\}$	

Starting from Table 2, all decision implications in Example 1 can be deduced by applying *Augmentation* and *Combination.*

In practical applications, however, data always changes when new objects/samples are continuously added into data, and DICB also changes simultaneously. In this paper, we study an incremental method for DICB generation, which updates the existing DICB to obtain a new one, when new objects come.

4. Incremental generation of decision premises

By Definition 7, we can see DICB is defined based on decision premise. Hence, to obtain a new DICB by updating the existing DICB, the key is to updating the existing decision premises.

4.1. Categories of decision premises

For the given decision context $K = (G, C \cup D, I_C \cup I_D)$, when a new object g is added into K, we denote the new decision context as:

$$K_g = (G \cup \{g\}, C \cup D, I_{\overline{C}} \cup I_{\overline{D}})$$

where $I_{\overline{C}} \subseteq (G \cup \{g\}) \times C$ and $I_{\overline{D}} \subseteq (G \cup \{g\}) \times D$, and write respectively the operators $(.)^C$ and $(.)^D$ in K_g as $(.)^{\overline{C}}$ and $(.)^{\overline{D}}$.

Since decision premise is defined via the operators $(.)^{CD}$ (Definition 6), in order to check the changes of decision premises from K to K_g, it is necessary to check the change from A^{CD} to $A^{\overline{CD}}$.

Proposition 4. *For decision context $K = (G, C \cup D, I_C \cup I_D)$ and K_g, let $A \subseteq C$. Then, we have the following conclusions:*

(1) $A^{\overline{CD}} \subseteq A^{CD}$;

(2) If $A \nsubseteq g^{\overline{C}}$, then $A^{\overline{CD}} = A^{CD}$;

(3) If $A \subseteq g^{\overline{C}}$, then $A^{\overline{CD}} = A^{CD} \cap g^{\overline{D}}$;

(4) $A \subseteq g^{\overline{C}}$ and $A^{CD} \nsubseteq g^{\overline{D}}$ if and only if $A^{\overline{CD}} \subset A^{CD}$.

Proof. (1) Firstly, by the definitions of $(.)^{C\overline{D}}$ and $(.)^{CD}$, it is easy to see that $A^{C\overline{D}} = A^{CD}$. There are two cases to be considered:

- $A \nsubseteq g^{\overline{C}}$. In this case, by the definitions of $(.)^{C}$ and $(.)^{\overline{C}}$, we have $A^{C} = A^{\overline{C}}$ and hence $A^{C\overline{D}} = A^{\overline{CD}}$; and considering $A^{C\overline{D}} = A^{CD}$, we have $A^{CD} = A^{\overline{CD}}$.
- $A \subseteq g^{\overline{C}}$. In this case, we have $A^{\overline{C}} = A^{C} \cup \{g\}$. By conclusion (4) of Proposition 2, we know $A^{\overline{CD}} = (A^{C} \cup \{g\})^{\overline{D}} = A^{C\overline{D}} \cap g^{\overline{D}}$, i.e., $A^{\overline{CD}} = A^{C\overline{D}} \cap g^{\overline{D}}$; and considering $A^{C\overline{D}} = A^{CD}$, we have $A^{\overline{CD}} = A^{CD} \cap g^{\overline{D}}$, which implies that $A^{\overline{CD}} \subseteq A^{CD}$.

In conclusion, $A^{C\overline{D}} = A^{CD}$ holds.

(2) and (3) have been proven in the process of proving (1).

(4) "$\Leftarrow$".Assume that $A \nsubseteq g^{\overline{C}}$. By conclusion (2), $A^{CD} = A^{\overline{CD}}$ holds, which contradicts $A^{\overline{CD}} \subset A^{CD}$, and hence $A \subseteq g^{\overline{C}}$.

Now, since $A \subseteq g^{\overline{C}}$, by conclusion (3), $A^{\overline{CD}} = A^{CD} \cap g^{\overline{D}}$ holds, and hence $A^{CD} \nsubseteq g^{\overline{D}}$ (because once $A^{CD} \subseteq g^{\overline{D}}$, $A^{CD} \cap g^{\overline{D}} = A^{CD}$ holds, i.e., $A^{\overline{CD}} = A^{CD}$, contradicting $A^{\overline{CD}} \subset A^{CD}$).

"$\Rightarrow$". Because $A \subseteq g^{\overline{C}}$, by conclusion (3), we know $A^{\overline{CD}} = A^{CD} \cap g^{\overline{D}}$; and considering $A^{CD} \nsubseteq g^{\overline{D}}$, $A^{CD} \cap g^{\overline{D}} \subset A^{CD}$ holds and thus $A^{\overline{CD}} \subset A^{CD}$. □

By Proposition 4, we conclude that:

- If $A \nsubseteq g^{\overline{C}}$, by conclusion (2), we have $A^{\overline{CD}} = A^{CD}$.
- If $A \subseteq g^{\overline{C}}$ and $A^{CD} \subseteq g^{\overline{D}}$, by conclusion (3), we have $A^{\overline{CD}} = A^{CD} \cap g^{\overline{D}} = A^{\overline{CD}}$, i.e., $A^{\overline{CD}} = A^{CD}$.
- If $A \subseteq g^{\overline{C}}$ and $A^{CD} \nsubseteq g^{\overline{D}}$, by conclusion(4), we have $A^{\overline{CD}} = A^{CD} \cap g^{\overline{D}} \subset A^{\overline{CD}}$, i.e.,$A^{\overline{CD}} \subset A^{CD}$.

To generate decision premises of K_g based on the existing decision premises of K, we classify the decision premises of K and K_g as follows:

(1) A is a decision premise of K, and A is also a decision premise of K_g. Despite this, one may not obtain the same decision implication, since the consequence may change, i.e., $A^{\overline{CD}} \neq A^{CD}$. Thus, by (1) of Proposition 4, we divide $A^{\overline{CD}} \subseteq A^{CD}$ into two cases:

(1.1) $A^{\overline{CD}} = A^{CD}$, i.e., the consequence of A is unchanged. In this case, we call A an unchanged decision premise;

(1.2) $A^{\overline{CD}} \subset A^{CD}$, i.e., the consequence of A needs to be changed into $A^{\overline{CD}}$. In this case, we call A a modified decision premise;

(2) A is a decision premise of K, but A is not a decision premise of K_g. In this case, we call A an invalid decision premise.

(3) A is a not decision premise of K, but A is a decision premise of K_g. In this case, we call A a new decision premise.

From the above, the decision premises of K and K_g can be classified into four categories: unchanged decision premises, modified decision premises, invalid decision premises and new decision premises.

Table 5: The categories of premises and the corresponding decision implications

Categories	Corresponding decision implications
Unchanged decision premises	$\{a_1\} \to \{d_1\}$ $\{a_2, a_5\} \to \{d_1\}$ $\{a_1, a_3, a_5\} \to \{d_1, d_2\}$ $\{a_3, a_5, a_6\} \to \{d_1, d_2\}$
Invalid decision premises	$\{a_6\} \to \{d_1\}$
Modified decision premises	$\{a_4\} \to \{d_2\}$ $\{a_2, a_3\} \to \{d_2\}$
New decision premises	$\{a_4, a_5\} \to \{d_1, d_2\}$ $\{a_5, a_6\} \to \{d_1\}$

Example 4 (Continuing Examples 1 and 3)**.** Consider the decision context K in Example 1. Its DICB is shown in Example 3.We add a new object g_7, which processes attributesa_2, a_3, a_4, a_6 and d_2, into K, and then obtain a new decision context K_{g_7}, as shown in Table 3.

Table 3: Decision context K_{g_7}

	a_1	a_2	a_3	a_4	a_5	a_6	d_1	d_2
g_1	1	1	0	0	1	1	1	0
g_2	0	0	1	0	0	0	0	1
g_3	1	0	1	0	0	1	1	0
g_4	0	1	0	0	0	0	0	0
g_5	0	1	1	1	1	1	1	1
g_6	0	0	1	0	1	0	0	0
g_7	0	1	1	1	0	1	0	1

The DICB of K_{g_7} is shown in Table 4.

Table 4: The DICB of K_{g_7}

$\{a_1\} \to \{d_1\}$	$\{a_5, a_6\} \to \{d_1\}$
$\{a_4\} \to \{d_2\}$	$\{a_4, a_5\} \to \{d_1, d_2\}$
$\{a_2, a_3\} \to \{d_2\}$	$\{a_1, a_3, a_5\} \to \{d_1, d_2\}$
$\{a_2, a_5\} \to \{d_1\}$	$\{a_3, a_5, a_6\} \to \{d_1, d_2\}$

Comparing the DICB of K (Table 2) and that of K_{g_7} (Table 4), the types of decision premises are recognized, as shown in Table 5.

4.2. Properties and update of decision premises

In this section, we will study the properties of the four types of decision premises, by which one can determine which category they belong to and how to modify or update them. We rewrite the set $\Theta(A)$

(formula 1) in K_g as:

$$\overline{\Theta}(A) = \cup\{A_i^{\overline{CD}} | A_i \subset A, A_i\, is\, a\, decision\, premise\, of\, K_g\} \tag{2}$$

Proposition 5. *For decision contexts K and K_g, let O be the DICB of K and $A \to A^{CD} \in O$. Then, A is an unchanged decision premise if and only if $A^{\overline{CD}} = A^{CD}$.*

Proof. "$\Leftarrow$". We firstly prove that if A is a modified decision premise or an invalid decision premise, then $A^{\overline{CD}} \subset A^{CD}$ holds.

(1) A is a modified decision premise, then by the definition of modified decision premise, we have $A^{\overline{CD}} \subset A^{CD}$.

(2) A is an invalid decision premise, i.e., A is a decision premise of K but not a decision premise of K_g. Assume that $A^{\overline{CD}} \not\subset A^{CD}$. By (1) of Proposition 4, we have $A^{\overline{CD}} = A^{CD}$. By Corollary 1 of [15], we have

$$\Theta(A) = \cup\{A_j^{CD} | A_j \subset A\} \tag{3}$$

and

$$\overline{\Theta}(A) = \cup\{A_j^{\overline{CD}} | A_j \subset A\} \tag{4}$$

Because for any $A_j \subset A$, $A_j^{\overline{CD}} \subseteq A_j^{CD}$ holds by (1) of Proposition 4, and hence $\overline{\Theta}(A) \subseteq \Theta(A)$ holds. Because A is not a decision premise of K_g, and hence $A^{\overline{CD}} = \overline{\Theta}(A)$ by Proposition 3. Taking $A^{CD} = A^{\overline{CD}}$, $\overline{\Theta}(A) \subseteq \Theta(A)$ and $A^{\overline{CD}} = \overline{\Theta}(A)$ into consideration, we have $A^{CD} \subseteq \Theta(A)$, i.e., A is not a decision premise of K by Proposition 3, which contradicts the fact that A is a decision premise of K. Hence, $A^{\overline{CD}} \subset A^{CD}$.

Because $A^{\overline{CD}} = A^{CD}$, then $A^{\overline{CD}} \not\subset A^{CD}$, i.e., A is neither a modified decision premise nor an invalid decision premise by Proposition 6. Because $A \to A^{CD} \in O$, A is not a new decision premise. In conclusion, A is an unchanged decision premise.

"$\Rightarrow$". It is straightforward. □

Example 5 (Continuing 4)**.** Take $\{a_2, a_5\}$, an unchanged decision premise in Example 4 as an example. We compute in K (Table 1) that $\{a_2, a_5\}^{CD} = \{d_1\}$, and compute in K_g (Table 3) that $\{a_2, a_5\}^{\overline{CD}} = \{d_1\}$. It is seen that $\{a_2, a_5\}^{\overline{CD}} = \{a_2, a_5\}^{CD}$.

Proposition 6. *For decision contexts K and K_g, let O be the DICB of K and $A \to A^{CD} \in O$. Then, A is a modified decision premise if and only if $A^{\overline{CD}} \subset A^{CD}$ and $A^{\overline{CD}} \supset \overline{\Theta}(A)$.*

Proof. "$\Rightarrow$". It has been proven in the sufficiency proof of Proposition 5 that $A^{\overline{CD}} \subset A^{CD}$. Because A is a modified decision premise, A is a decision premise of K_g, i.e., $A^{\overline{CD}} \supset \overline{\Theta}(A)$ by Definition 6.

"$\Leftarrow$". Because $A \to A^{CD} \in O$, A is a decision premise of K. Because $A^{\overline{CD}} \supset \overline{\Theta}(A)$, by Definition 6, A is a decision premise of K_g. Because $A^{\overline{CD}} \subset A^{CD}$, A is a modified decision premise. □

Example 6 (Continuing Example 4)**.** Take $\{a_4\}$, an modified decision premise, as an example. We compute that $\{a_4\}^{CD} = \{d_1, d_2\}$, $\{a_4\}^{\overline{CD}} = \{d_2\}$ and $\overline{\Theta}(a_4) = \emptyset$, satisfying that $\{a_4\}^{\overline{CD}} \subset \{a_4\}^{CD}$ and $\{a_4\}^{\overline{CD}} \supset \overline{\Theta}(a_4)$.

Proposition 7. *For decision contexts K and K_g, let O be the DICB of K and $A \to A^{CD} \in O$. Then, A is an invalid decision premise if and only if $A^{\overline{CD}} \subset A^{CD}$ and $A^{\overline{CD}} = \overline{\Theta}(A)$.*

Proof. "$\Rightarrow$". It has been proven in the sufficiency proof of Proposition 5 that $A^{\overline{CD}} \subset A^{CD}$. Because A is an invalid decision premise, A is a not decision premise of K_g, i.e., $A^{\overline{CD}} = \overline{\Theta}(A)$ by Proposition 3.

"$\Leftarrow$". Because $A \to A^{CD} \in O$, A is a decision premise of K. Because $A^{\overline{CD}} = \overline{\Theta}(A)$, by Proposition 3, A is not a decision premise of K_g. Hence, A is an invalid decision premise. □

Example 7 (Continuing Example 4)**.** As shown in Example 4, $\{a_6\}$ is an invalid decision premise. From K and K_g, we know $\{a_6\}^{CD} = \{d_1\}$, $\{a_6\}^{\overline{CD}} = \emptyset$ and $\overline{\Theta}(a_6) = \emptyset$, implying that $\{a_6\}^{\overline{CD}} \subset \{a_6\}^{CD}$ and $\{a_6\}^{\overline{CD}} = \overline{\Theta}(a_6)$.

Proposition 8. *For decision contexts K and K_g, if A is a new decision premise, then:*
(1) $A \nsubseteq g^{\overline{C}}$;
(2) $A^{\overline{CD}} = \Theta(A)$;
(3) $\overline{\Theta}(A) \subset \Theta(A)$;
(4) There exist a $A_i \to A_i^{CD} \in O$ such that $A_i \subset A$ and $A_i^{\overline{CD}} \subset A_i^{CD}$.

Proof. (1) Assume $A \subseteq g^{\overline{C}}$. If we can prove A is not a decision premise of K_g, i.e., $A^{\overline{CD}} = \overline{\Theta}(A)$ by Proposition 3, it contradicts the fact that A is a new decision premise. In this case, the assumption $A \subseteq g^{\overline{C}}$ is wrong, and hence$A \nsubseteq g^{\overline{C}}$.

We firstly prove $\Theta(A) = \overline{\Theta}(A)$. By formulas 3 and 4, we just need to prove that for any $A_i \subset A$, $A_i^{\overline{CD}} = A_i^{CD}$. Because $A_i \subset A$ and $A \subseteq g^{\overline{C}}$, we have $A_i \subseteq g^{\overline{C}}$. On the one hand, $A_i^{\overline{CD}} = A_i^{CD} \cap g^{\overline{D}}$ by (3) of Proposition 4; on the other hand, $A_i^{CD} \subseteq g^{\overline{C}CD}$ by Proposition 1. Because $g^{\overline{C}} \subseteq g^{C}$, we have $g^{\overline{C}CD} \subseteq g^{CCD} \subseteq g^{D}$ by Proposition 1, i.e., $g^{\overline{C}CD} \subseteq g^{D}$. Considering $A_i^{CD} \subseteq g^{\overline{C}CD}$ and $g^{\overline{C}CD} \subseteq g^{D}$, we have $A_i^{CD} \subseteq g^{D}$. Considering $A_i^{\overline{CD}} = A_i^{CD} \cap g^{\overline{D}}$ and $A_i^{CD} \subseteq g^{D}$, we have $A_i^{\overline{CD}} = A_i^{CD}$.

By (1) of Proposition 4, we have $A^{\overline{CD}} \subseteq A^{CD}$. Because A is a new decision premise, it is not a decision premise of K, and hence $A^{CD} = \Theta(A)$ by Proposition 3. Take $A^{\overline{CD}} \subseteq A^{CD}$, $A^{CD} = \Theta(A)$ and $\Theta(A) = \overline{\Theta}(A)$ into consideration, we have $A^{\overline{CD}} \subseteq \Theta(A)$. And it is easy to see that $A^{\overline{CD}} \supseteq \overline{\Theta}(A)$, thus, $A^{\overline{CD}} = \overline{\Theta}(A)$ holds.

(2) On the one hand, by (2) of Proposition 4, we have$A^{\overline{CD}} = A^{CD}$. On the other hand, A is not a decision premise of K, and then $A^{CD} = \Theta(A)$ by Proposition 3. Take $A^{\overline{CD}} = A^{CD}$ and $A^{CD} = \Theta(A)$ into consideration, we have $A^{\overline{CD}} = \Theta(A)$.

(3) Because A is a new decision premise, on the one hand, A is not a decision premise of K, and then $A^{CD} = \Theta(A)$ by Proposition 3. On the other hand, A is a decision premise of K_g, and then $\overline{\Theta}(A) \subset A^{\overline{CD}}$ by Definition 6. By (2) of Proposition 4, we have$A^{\overline{CD}} = A^{CD}$. Taking $\overline{\Theta}(A) \subset A^{\overline{CD}}$, $A^{\overline{CD}} = A^{CD}$ and $A^{CD} = \Theta(A)$ into consideration, we have $\overline{\Theta}(A) \subset \Theta(A)$.

(4) Assume that for any $A_j \to A_j^{CD} \in O$, $A_j \not\subset A$ or $A_j^{\overline{CD}} \not\subset A_j^{CD}$.

We will prove $\Theta(A) \subseteq \overline{\Theta}(A)$. For any $d \in \Theta(A)$, by the definition of $\Theta(A)$, there must exist a $A_k \to A_k^{CD} \in O$ such that $A_k \subset A$ and $d \in A_k^{CD}$. Because $A_k \subset A$, one the one hand, by the assumption, we know $A_k^{\overline{CD}} \not\subset A_k^{CD}$, and hence $A_k^{\overline{CD}} = A_k^{CD}$ by (1) of Proposition 4. On the other hand, because $A_k \subset A$, by formula 3, we have $A_k^{\overline{CD}} \subseteq \overline{\Theta}(A)$. Take $d \in A_k^{CD}$, $A_k^{CD} = A_k^{\overline{CD}}$ and $A_k^{\overline{CD}} \subseteq \overline{\Theta}(A)$ into consideration, we have $d \in \overline{\Theta}(A)$. In conclusion, $d \in \Theta(A)$ implies $d \in \overline{\Theta}(A)$, i.e., $\Theta(A) \subseteq \overline{\Theta}(A)$. By (3), A is not a new decision premise, contradicting the fact that A is a new decision premise. Thus, the assumption is wrong, and then there must exist a $A_i \to A_i^{CD} \in O$ such that $A_i \subset A$ and $A_i^{\overline{CD}} \subset A_i^{CD}$. $\square$

Example 8 (Continuing Example 4)**.** Take $\{a_4, a_5\}$, a new decision premise, as an example. It is verified that $\{a_4, a_5\}$ satisfied the conditions in Proposition 8:

(1) $\{a_4, a_5\} \nsubseteq g_7^{\overline{C}} = \{a_2, a_3, a_4, a_6\}$.

(2) From Table 3, we have $\{a_4, a_5\}^{\overline{CD}} = \{d_1, d_2\}$; and from Table 2, we have $\Theta(\{a_4, a_5\}) = \{d_1, d_2\}$, implying that $\{a_4, a_5\}^{\overline{CD}} = \Theta(\{a_4, a_5\})$.

(3) From Table 4, we have $\overline{\Theta}(\{a_4, a_5\}) = \{d_2\}$, and then $\overline{\Theta}(\{a_4, a_5\}) \subset \Theta(\{a_4, a_5\})$.

(4) For $\{a_4, a_5\}$, there exists $\{a_4\} \to \{d_1, d_2\}$ in the DICB of K (Table 2) satisfying that $\{a_4\} \subset \{a_4, a_5\}$ and $\{a_4\}^{\overline{CD}} \subset \{a_4\}^{CD}$ ($\{a_4\}^{\overline{CD}} = \{d_2\}$, $\{a_4\}^{CD} = \{d_1, d_2\}$).

By (4) of Proposition 8, we know that for any new decision premise A, there must exist a decision implication such that $A_i \to A_i^{CD} \in O$, $A_i \subset A$ and $A_i^{\overline{CD}} \subset A_i^{CD}$. Inspired by this, we define the generator of new decision premises, and intend to generate new premises based on their generators.

Definition 8. For decision contexts K and K_g, let O be the DICB of K and A be a new decision premise. Let

$$\Omega(A) = \{A_i | A_i \to A_i^{CD} \in O, A_i \subset A, A_i^{\overline{CD}} \subset A_i^{CD}\}$$

For $A_m \in \Omega(A)$, if $|A_m| = \max\{|A_i||A_i \in \Omega(A)\}$, we call A_m a generator of A.

By Definition 8, we know that there may be more than one generator of A.

Proposition 9 further clarifies the relationship between A and its generators.

Proposition 9. *For decision contexts K and K_g, let A be a new decision premise. For any generator A_m of A, A has exactly one more attribute m than A_m and m does not belong to $g^{\overline{C}}$, i.e., $m \in C - g^{\overline{C}}$.*

Proof. Because A_m is a generator of A, we have $A_m \subset A$ and $A_m^{\overline{CD}} \subset A_m^{CD}$ by Definition 8; and $A_m \subseteq g^{\overline{C}}$ by (4) of Proposition 4. Considering $A_m \subseteq g^{\overline{C}}$ and $A_m \subset A$, A can be written as:

$$A = A_m \cup S \cup T$$

where $S = (g^{\overline{C}} - A_m) \cap A$ and $T = (C - g^{\overline{C}}) \cap A$. It is clear that $A_m \cup S = A_m \cup ((g^{\overline{C}} - A_m) \cap A) = g^{\overline{C}} \cap A$, i.e., $A_m \cup S = g^{\overline{C}} \cap A$.

(1) We prove $|T| = 1$ in the following.

Assume $|T| = 0$. We have $A = A_m \cup S \cup T = A_m \cup S$, and by $A_m \cup S = g^{\overline{C}} \cap A$, we have $A = g^{\overline{C}} \cap A \subseteq g^{\overline{C}}$, i.e., $A \subseteq g^{\overline{C}}$. Because A is a new decision premise, by (1) of Proposition 8, $A \nsubseteq g^{\overline{C}}$ holds, which contradicts $A \subseteq g^{\overline{C}}$. Thus, $|T| \neq 0$ holds.

Assume $|T| > 1$. For any $A_i \subset A$, there are two cases to be considered:

- $A_i \nsubseteq g^{\overline{C}}$. By (2) of Proposition 4, we have $A_i^{CD} = A_i^{\overline{CD}}$. Because $A_i \subset A$, by formula 4, we have$A_i^{\overline{CD}} \subseteq \overline{\Theta}(A)$, and hence $A_i^{CD} \subseteq \overline{\Theta}(A)$ because $A_i^{CD} = A_i^{\overline{CD}}$.

- $A_i \subseteq g^{\overline{C}}$. Because $|T| > 1$, we can define $A_j = A_i \cup \{a\}$ where $\{a\} \subset T$. It is clear that $A_i \subset A_j \subset A$. By $A_i \subset A_j$ and Proposition 1, we have $A_i^{CD} \subseteq A_j^{CD}$. Because $T \cap g^{\overline{C}} = \emptyset$ and $a \in T$, then $a \notin g^{\overline{C}}$, and hence $A_j = A_i \cup \{a\} \nsubseteq g^{\overline{C}}$; and by (2) of Proposition 4, we have $A_j^{CD} = A_j^{\overline{CD}}$. Because $A_j \subset A$, by formula 4, we have $A_j^{\overline{CD}} \subseteq \overline{\Theta}(A)$. Considering $A_i^{CD} \subseteq A_j^{CD}$, $A_j^{CD} = A_j^{\overline{CD}}$ and $A_j^{\overline{CD}} \subseteq \overline{\Theta}(A)$, we have $A_i^{CD} \subseteq \overline{\Theta}(A)$.

In conclusion, for any $A_i \subset A$, $A_i^{CD} \subseteq \overline{\Theta}(A)$ holds, i.e., $\Theta(A) \subseteq \overline{\Theta}(A)$ by formula 3, implying that A is not a new decision premise by (3) of Proposition 8, which contradicts the fact that A is a new decision premise. Thus, the assumption $|T| > 1$ is wrong; and considering $|T| \neq 0$, we have $|T| = 1$.

(2) We prove $|S| = 0$. Assume that $|S| > 0$. Because $|S| > 0$ and $|T| = 1$, by the definitions of S and T, it is clear that $A_m \subset A_m \cup S \subset A$.

Prove $(A_m \cup S)^{\overline{CD}} \subset (A_m \cup S)^{CD}$. By (4) of Proposition 4, it is equivalent to prove $A_m \cup S \subseteq g^{\overline{C}}$ and $(A_m \cup S)^{CD} \nsubseteq g^{\overline{D}}$. Because, $A_m \cup S = g^{\overline{C}} \cap A$, $A_m \cup S \subseteq g^{\overline{C}}$ holds. Because A_m is a generator of A, by Definition 8, we have $A_m^{\overline{CD}} \subset A_m^{CD}$, which implies $A_m^{CD} \nsubseteq g^{\overline{D}}$ by (4) of Proposition 4. Because $A_m \subset A_m \cup S$, $A_m^{CD} \subseteq (A_m \cup S)^{CD}$ holds, and hence $(A_m \cup S)^{CD} \nsubseteq g^{\overline{D}}$ because $A_m^{CD} \nsubseteq g^{\overline{D}}$.

Assume that $A_m \cup S$ is a decision premise of K. Because $A_m \cup S \subset A$ and $(A_m \cup S)^{\overline{CD}} \subset (A_m \cup S)^{CD}$, by Definition 8,we have $A_m \cup S \in \Omega(A)$, and hence $|A_m| < |A_m \cup S| \leq \max\{|A_i||A_i \in \Omega(A)\}$, implying that A_m is not a generator of A, which contradicts the fact that A_m is a generator of A. Thus, the assumption $A_m \cup S$ is a decision premise of K is wrong, i.e., $A_m \cup S$ is not a decision premise of K.

For any decision premise A_i of K such that$A_i \subset A$, there are two cases to be considered:

- $A_i \nsubseteq g^{\overline{C}}$. By (2) of Proposition 4, we have $A_i^{CD} = A_i^{\overline{CD}}$. Because $A_i \subset A$, by formula 4, we have$A_i^{\overline{CD}} \subseteq \overline{\Theta}(A)$, and hence $A_i^{CD} \subseteq \overline{\Theta}(A)$ because $A_i^{CD} = A_i^{\overline{CD}}$.

- $A_i \subseteq g^{\overline{C}}$. Let $A_j = A_i \cup T$. It is clearly that $A_i \subseteq A_j$, and then $A_i^{CD} \subseteq A_j^{CD}$ by Proposition 1. Because $T \cap g^{\overline{C}} = \emptyset$, then $T \nsubseteq g^{\overline{C}}$, and hence $A_j = A_i \cup T \nsubseteq g^{\overline{C}}$, and then $A_j^{CD} = A_j^{\overline{CD}}$ by (2) of Proposition 4. Because $A_i \subset A$ and $A_i \subseteq g^{\overline{C}}$, then $A_i \subseteq A \cap g^{\overline{C}}$, i.e., $A_i \subseteq A_m \cup S$ because $A \cap g^{\overline{C}} = A_m \cup S$. Because$A_i$ is a decision premise of K and we have proven that $A_m \cup S$ is not a decision premise of K, we can ensure$A_i \neq A_m \cup S$, and hence $A_i \subset A_m \cup S$. Because $A = A_m \cup S \cup T$ and $A_m \cup S \supset A_i$, then $A \supset A_i \cup T = A_j$, i.e., $A \supset A_j$, and hence $A_j^{\overline{CD}} \subseteq \overline{\Theta}(A)$ by formula 4. Considering $A_i^{CD} \subseteq A_j^{CD}$, $A_j^{CD} = A_j^{\overline{CD}}$ and $A_j^{\overline{CD}} \subseteq \overline{\Theta}(A)$, we have $A_i^{CD} \subseteq \overline{\Theta}(A)$.

In conclusion, for any decision premise A_i of K satisfying that$A_i \subset A$ and $A_i^{CD} \subseteq \overline{\Theta}(A)$, $\Theta(A) \subseteq \overline{\Theta}(A)$ holds by formula 1, implying that A is not a new decision premise by (3) of Proposition 8, contradicting the fact that A is a new decision premise. Thus, the assumption $|S| > 0$ is wrong, and hence $|S| = 0$. □

Example 9 (Continuing Example 4)**.** In Example 4, an object g_7, whose condition attribute set is $\{a_2, a_3, a_4, a_6\}$ and decision attribute set is $\{d_2\}$, is added into the decision context.

Take the new decision premise $\{a_4, a_5\}$ as an example. By calculation, it is known that $\Omega(A) = \{\{a_4\}\}$, implying that $\{a_4\}$ is a generator of $\{a_4, a_5\}$ by Definition 8. Comparing $\{a_4\}$ with $\{a_4, a_5\}$, it is seen that $\{a_4, a_5\}$ has exactly one more attribute $\{a_5\}$ than $\{a_4\}$, and $\{a_5\}$ does not belong to g_7.

By Theorem 5, one can obtain all candidate new decision premises.

Theorem 5. *For decision contexts K and K_g, let O be the DICB of K. If A is a new decision premise, then there must exist $A_i \to A_i^{CD} \in O$ satisfying the following conditions:*
(1) $A_i^{\overline{CD}} \subset A_i^{CD}$;
(2) A has exactly one more attribute than A_i, and the extra attribute does not belong to $g^{\overline{C}}$.

Proof. By (4) of Proposition 8 and Proposition 9, they are straightforward. □

By Theorem 5, for a new decision premise A, there must exist a set A_i that satisfies conditions (1) and (2) in Theorem 5. Thus, we can obtain all the candidate new decision premises $A' \cup \{d'\}$, where $d' \in C - g^{\overline{C}}$, by searching all $A' \to A'^{CD}$ in O that satisfy condition (1) and filtering those A' that do not satisfy condition (2). By Theorem 5, there must exist a premise A_i and condition attribute $d' \in C - g^{\overline{C}}$ that satisfy $A = A_i \cup \{d\}$.

5. Incremental algorithm for DICB generation

In this section, we will give the incremental algorithm for generating DICB. The incremental algorithm starts with an empty decision context, i.e., there is no objects in this decision context.

Proposition 10. *Let $K = (G, C \cup D, I_C \cup I_D)$ be a decision context where $D \neq \emptyset$. If $G = \emptyset$, then the DICB of K is $\{\emptyset \to \emptyset^{CD}\}$.*

Proof. We firstly prove $\emptyset$ is a decision premise of K. By Definition 4, we have $\emptyset^C = G = \emptyset$, and hence $\emptyset^{CD} = \emptyset^D = D \supset \emptyset = \Theta(\emptyset)$, i.e., $\emptyset^{CD} \supset \Theta(\emptyset)$, and hence $\emptyset$ is a decision premise by Definition 6.

We then prove for any $\emptyset \subset A \subseteq C$, A is not a decision premise of K. Because $\emptyset \subset A$ and $\emptyset$ is a decision premise, by the definition of $\Theta(A)$, we have $\emptyset^{CD} \subseteq \Theta(A)$, and hence $D \subseteq \Theta(A)$ because $\emptyset^{CD} = D$. Considering $D \subseteq \Theta(A)$ and $A^{CD} \subseteq D$, we have $A^{CD} \subseteq \Theta(A)$, i.e., A is not a decision premise by Definition 6.

From the above, the DICB of K is $\{\emptyset \to \emptyset^{CD}\}$. □

The incremental algorithm for generating DICB is shown in Algorithm 1.

Algorithm 1 Incremental algorithm for DICB generation

Input: Decision context $K = (G, C \cup D, I_C \cup I_D)$ ($G \neq \emptyset$, $C \neq \emptyset$ and $D \neq \emptyset$)

Output: The DICB O of K

1. $K_{current} = (\emptyset, C \cup D, \emptyset \cup \emptyset)$
2. $O = \{\emptyset \to \emptyset^{CD}\}$
3. for all $g \in G$ do
4. $O = Update\_CanoBasis(K_{current}, g, O)$
5. $K_{current} = K_g$
6. end for
7. return O

For the given decision context $K = (G, C \cup D, I_C \cup I_D)$, starting with $G = \emptyset$ and $D \neq \emptyset$, by Proposition 10, we initialize O with $\{\emptyset \to \emptyset^{CD}\}$ (steps 1-2). At each iteration (steps 3-6), we add an object g to the existing decision context $K_{current}$, and update the existing DICB by the function $Update\_CanoBasis(K_{current}, g, O)$ (Algorithm 2).

Algorithm 2 $Update\_CanoBasis$ function

Input: $K_{current}$, g and O

Output: The DICB O of K_g

for all $A_i \to A_i^{CD} \in O$ do
    modify $A_i \to A_i^{CD}$ to $A_i \to A_i^{\overline{CD}}$
    if $A_i^{\overline{CD}} \subset A_i^{CD}$ then
        for all $a \in C - g^{\overline{C}}$ do
            add $A_i \cup \{a\} \to (A_i \cup \{a\})^{\overline{CD}}$ to O //by (2) of Proposition 8, $(A_i \cup \{a\})^{\overline{CD}} = \Theta(A_i \cup \{a\})$
            sort the decision implications in O by the cardinality of the premises
        end for
    end if
end for
for all $A \to A^{\overline{CD}} \in O$ do
    $\Theta(A) = \emptyset$
    for all $A_i \to A_i^{\overline{CD}} \in O$ do
        if $A_i \subset A$ then
            $\Theta(A) = \Theta(A) \cup A_i^{\overline{CD}}$
        end if
    end for
    if $\Theta(A) = A^{\overline{CD}}$ then
        remove $A \to A^{\overline{CD}}$ from O
    end if
end for
return O

In Algorithm 2, steps 1-9 are used to generate all the candidate new decision premises: for each decision implication $A_i \to A_i^{CD} \in O$, we firstly modify its consequence from A_i^{CD} to $A_i^{\overline{CD}}$ (step 2), and then generate all the candidate new decision premises (steps 3-8). This is reasonable, because by Theorem 5, if A_i satisfies condition (1) and $A = A_i \cup \{d\}$ satisfies condition (2), then $A = A_i \cup \{d\}$ is a candidate new decision premise. In this case, the premises of decision implications in O are divided into four categories: candidate new decision premises, unchanged decision premises, invalid decision premises, and modified decision premises (the consequences of the modified decision premises have been modified in step 2). Steps 10-20 remove the invalid decision premises and the candidate new decision premises that are not new decision premises (steps 17-19). Note that, because the decision implications in O are stored in the increasing cardinality of decision premises (step 6) and O is traversed in order (step 10), we can compute $\overline{\Theta}(A)$ (steps 11-16).

Example 10 illustrates Algorithm 2.

Example 10 (Continuing Example 4)**.** In Example 4, object g_7, processing attributes a_2, a_3, a_4, a_6 and d_2, is added into the existing decision context $K_{current}$. The DICB O of $K_{current}$ is shown in Table 2.

Take the decision premise $\{a_4\}$ of $K_{current}$ as an example. It is verified that $\{a_4\}$ satisfies condition (1) of Theorem 5, and both $\{a_1, a_4\}$ and $\{a_4, a_5\}$ satisfy condition (2) of Theorem 5, and thus the two decision implications $\{a_1, a_4\} \to \{a_1, a_4\}^{\overline{CD}}$ and $\{a_4, a_5\} \to \{a_4, a_5\}^{\overline{CD}}$ are added into O. The set O after steps 1-9 is shown in Table 6.

Table 6: Set O

Categories	Corresponding decision implications
Unchanged decision premises	$\{a_1\} \to \{d_1\}$ $\{a_2, a_5\} \to \{d_1\}$ $\{a_1, a_3, a_5\} \to \{d_1, d_2\}$ $\{a_3, a_5, a_6\} \to \{d_1, d_2\}$
Invalid decision premises	$\{a_6\} \to \{d_1\}$
Modified decision premises	$\{a_4\} \to \{d_2\}$ $\{a_2, a_3\} \to \{d_2\}$
Candidate new decision premises	$\{a_4, a_5\} \to \{d_1, d_2\}$ $\{a_5, a_6\} \to \{d_1\}$ $\{a_1, a_4\} \to \{d_1, d_2\}$ $\{a_1, a_6\} \to \{d_1\}$ $\{a_1, a_2, a_3\} \to \{d_1, d_2\}$ $\{a_2, a_3, a_5\} \to \{d_1, d_2\}$

Take $\{a_1, a_4\}$ in Table 6 as an example. We have

$$T = \{a_1\}^{\overline{CD}} \cup \{a_4\}^{\overline{CD}} = \{d_1, d_2\} = \{a_1, a_4\}^{\overline{CD}}$$

i.e., $\{a_1, a_4\}$ is not a decision premise of K_{g_7}, and then $\{a_1, a_4\} \to \{d_1, d_2\}$ is removed. In the same way, we remove all those invalid decision premises and the candidate new decision premises that are not decision premises of K_{g_7}.

The DICB of K_{g_7} is shown in Table 7.

Table 7: The DICB of K_{g7}

$\{a_1\} \rightarrow \{d_1\}$	$\{a_2, a_5\} \rightarrow \{d_1\}$
$\{a_4\} \rightarrow \{d_2\}$	$\{a_4, a_5\} \rightarrow \{d_1, d_2\}$
$\{a_2, a_3\} \rightarrow \{d_2\}$	$\{a_1, a_3, a_5\} \rightarrow \{d_1, d_2\}$
$\{a_5, a_6\} \rightarrow \{d_1\}$	$\{a_3, a_5, a_6\} \rightarrow \{d_1, d_2\}$

Now let us analyze the time complexity of Algorithm 2. Let $|C|$ be the number of condition attributes, $|D|$ be the number of decision attributes, and o_n be the number of decision implications in DICB of $K_{current}$. The time complexity of steps 1-9 is $O(o_n \cdot |C| \cdot o_n \cdot (|C| + |D|))$. Denote the number of decision implications after step 9 by $\bar{o}_n$. Then, the time complexity of steps 10-20 is $O(\bar{o}_n^2 \cdot (|C| + |D|))$. Thus, the time complexity of Algorithm 2 is:

$$O(o_n \cdot |C| \cdot o_n \cdot (|C| + |D|) + \bar{o}_n^2 \cdot (|C| + |D|))$$

Let $|G|$ be the number of objects in K. Then, the time complexity of Algorithm1 is:

$$O(|G| \cdot |C| \cdot o_n^2 \cdot (|C| + |D|) + |G| \cdot \bar{o}_n^2 \cdot (|C| + |D|))$$

Let $|O|$ be the number of decision implications in DICB of K. We approximate o_n by $|O|$ and $\bar{o}_n$ by $|O| + |O| \cdot |C|$. The time complexity of Algorithm 1 becomes:

$$O(|G| \cdot |C|^2 \cdot |O|^2 + |G| \cdot |C| \cdot |D| \cdot |O|^2 + |G| \cdot |C|^3 \cdot |O|^2 + |G| \cdot |C|^2 \cdot |D| \cdot |O|^2)$$

i.e.,

$$O(|G| \cdot |C|^3 \cdot |O|^2 + |G| \cdot |C|^2 \cdot |D| \cdot |O|^2) \tag{5}$$

[15] put forward a true premise-based algorithm (abbreviated to MBTP below), whose time complexity is[1]

$$O(|G|^2 \cdot |C| \cdot |D| + |G| \cdot |C|^3 \cdot |D| \cdot |O|^2 + |G| \cdot |D|^2 \cdot |O|) \tag{6}$$

and proved its absolute advantages compared with the minimal generator-algorithm in [46]. Section 6 makes a further comparative experiment between our proposed algorithm and MBTP algorithm.

6. Experimental verification

An advantage of the incremental algorithm is that it is able to update, instead of re-generating, DICB when new objects are added to decision contexts. Taking this advantage into account, one can compare the performance of MBTP and the incremental algorithm in two ways:

1. new objects are added to decision contexts and we compare the time consumption of MBTP in generating the whole DICB and the time consumption of the incremental algorithm in updating the existing DICB;
2. a decision context is given and we compare the time consumption of MBTP and the incremental algorithm in generating the whole DICB.

Obviously, the time consumption of the incremental algorithm in the first way is only a part of that in the second way, whereas the time consumption of MBTP in the first way is the same as that in the second way. Thus, if the incremental algorithm is more efficient than MBTP in the second way, it must be also

[1]The time complexity of MBTP obtained in the paper is slightly different with that of [15], where Li et al. made a mistake when analyzing o'_i, i.e., setting $o'_i = |O|$ instead of $o'_i = |O| + |C| \cdot |O|$.

more efficient than MBTP in the first way. In the section, we compare their performances in the second way.

6.1. Experiment data

We selected 8 UCI datasets with different scales, carried out some necessary pre-processing, such as removing missing values and normalizing the continuous attributes, and finally obtained 8 formal contexts according to the threshold value 0.5. The summary information of the formal contexts is shown in Table 8, in which $|G|$ and $|M|$ denote the numbers of objects and attributes respectively.

Table 8: Data sets

No	Data sets	$\|G\|$	$\|M\|$
1	cloud	108	21
2	hou	506	14
3	ion	351	34
4	triazines	186	181
5	bank8FM	8192	27
6	dplanes	40767	33
7	bank32nh	8192	99
8	waveform	5000	123

6.2. Experiment approach and results

For a dataset, we generated the first decision context by randomly selecting one condition attribute from all the attributes and taking the remainder attributes as the decision attributes. Subsequently, we equably increased the number of condition attributes, which were also randomly selected from all the attributes, with the rest being taken as the decision attributes. This process was repeated 10 times. It is noted that when the number of condition attributes was not an integer, we rounded it to the nearest integer. Take the dataset "cloud" as an example. We obtained 10 decision contexts by randomly choosing 1, 3, 5, 7, 9, 11, 13, 16, 18, 20 condition attributes from all the attributes and taking the rest as decision attributes.

The results are shown in Tables 9-16, in which:

1. "Incre" represents the incremental algorithm;
2. If the time consumption of the incremental algorithm is less than that of MBTP, we highlight it bold;
3. If the time consumption of one algorithm is more than 3600 seconds and 10 times that of the other one, the algorithm is terminated and its time consumption is appended with the symbol "+", e.g. "9922.8+";
4. If an algorithm has ran more than 24 hours, it is also terminated and its time consumption is denoted as "24h+". It is noted that when both the algorithms are determined, $|O|$ is unknown, and then it is denoted as "$*$".

Table 9: cloud

No	$\lvert G\rvert$	$\lvert C\rvert$	$\lvert D\rvert$	$\lvert O\rvert$	$\lvert G\rvert/\lvert C\rvert$	MBTP (s)	Incre (s)
1	108	1	20	0	108.00	0.39	**0.02**
2	108	3	18	2	36.00	0.39	**0.02**
3	108	5	16	13	21.60	0.16	**0.03**
4	108	7	14	22	15.43	0.17	**0.09**
5	108	9	12	27	12.00	0.17	0.17
6	108	11	10	55	9.82	0.23	0.59
7	108	13	8	76	8.31	0.30	0.73
8	108	15	6	132	7.20	0.50	2.01
9	108	17	4	180	6.35	0.92	5.63
10	108	20	1	208	5.40	0.73	9.47

Table 10: hou

No	$\lvert G\rvert$	$\lvert C\rvert$	$\lvert D\rvert$	$\lvert O\rvert$	$\lvert G\rvert/\lvert C\rvert$	MBTP (s)	Incre (s)
1	506	1	13	0	506.00	4.34	**0.03**
2	506	2	12	1	253.00	3.90	**0.03**
3	506	3	11	1	168.67	2.63	**0.03**
4	506	4	10	9	126.50	1.88	**0.03**
5	506	6	8	7	84.33	2.06	**0.08**
6	506	7	7	19	72.329	1.92	**0.09**
7	506	8	6	17	63.25	1.61	**0.11**
8	506	10	4	14	50.60	0.69	**0.09**
9	506	11	3	20	46.00	0.92	**0.14**
10	506	13	1	17	38.92	0.27	**0.08**

Table 11: ion

No	$\lvert G\rvert$	$\lvert C\rvert$	$\lvert D\rvert$	$\lvert O\rvert$	$\lvert G\rvert/\lvert C\rvert$	MBTP (s)	Incre (s)
1	351	1	33	0	351.00	2.51	**0.02**
2	351	4	30	1	87.75	1.42	**0.08**
3	351	8	26	14	43.88	0.95	**0.23**
4	351	11	23	166	31.90	1.67	5.07
5	351	15	19	705	23.40	5.80	147.11
6	351	18	16	2182	19.50	37.21	1483.88
7	351	22	12	2657	15.95	56.71	2261.87
8	351	25	9	5246	14.04	236.39	3600+
9	351	29	5	11138	12.10	992.28	9922.8+
10	351	33	1	1	10.64	0.06	0.06

Table 12: triazines

No	$\lvert G\rvert$	$\lvert C\rvert$	$\lvert D\rvert$	$\lvert O\rvert$	$\lvert G\rvert/\lvert C\rvert$	MBTP (s)	Incre (s)
1	186	1	180	2	186.00	11.73	**0.09**
2	186	20	161	94	9.30	7.59	**2.31**
3	186	40	141	466	4.65	62.46	86.19
4	186	60	121	1223	3.10	497.39	657.11
5	186	80	101	2520	2.33	1736.10	2987.83
6	186	100	81	5664	1.86	8545.09	36192.73
7	186	120	61	14263	1.55	85422.69	**63303.00**
8	186	140	41	*	1.33	24h+	24h+
9	186	160	21	*	1.16	24h+	24h+
10	186	180	1	*	1.03	24h+	24h+

Table 13: bank8FM

No	$\lvert G\rvert$	$\lvert C\rvert$	$\lvert D\rvert$	$\lvert O\rvert$	$\lvert G\rvert/\lvert C\rvert$	MBTP (s)	Incre (s)
1	8192	1	26	0	8192.00	2975.56	**0.58**
2	8192	3	24	1	2730.67	1288.52	**0.98**
3	8192	6	21	3	1365.33	1137.56	**4.06**
4	8192	9	18	5	910.22	969.14	**9.83**
5	8192	12	15	79	682.67	847.65	**23.46**
6	8192	14	13	179	585.14	665.70	**37.60**
7	8192	17	10	545	481.88	443.73	**427.19**
8	8192	20	7	360	409.60	525.86	**168.96**
9	8192	23	4	1419	356.17	356.54	3713.78
10	8192	26	1	2325	315.08	386.76	3867.6+

Table 14: dplanes

No	$\lvert G\rvert$	$\lvert C\rvert$	$\lvert D\rvert$	$\lvert O\rvert$	$\lvert G\rvert/\lvert C\rvert$	MBTP (s)	Incre (s)
1	40767	1	32	0	40767.00	3600+	**4.38**
2	40767	4	29	1	10191.75	41349.31	**11.19**
3	40767	8	25	3	5095.88	32780.72	**44.04**
4	40767	11	22	18	3706.09	29970.98	**215.20**
5	40767	15	18	49	2717.80	24646.89	**1882.13**
6	40767	18	15	187	2264.83	20602.41	**2356.20**
7	40767	22	11	2785	1853.05	19932.35	84995.86
8	40767	25	8	4746	1630.68	19782.78	24h+
9	40767	29	4	13344	1405.76	74936.36	24h+
10	40767	32	1	*	1273.97	24h+	24h+

Table 15: bank32nh

No	$|G|$	$|C|$	$|D|$	$|O|$	$|G|/|C|$	MBTP (s)	Incre (s)
1	8192	1	98	0	8192.00	8790.00	**1.84**
2	8192	11	88	271	744.73	5675.13	**125.66**
3	8192	22	77	9149	372.36	8494.20	24h+
4	8192	33	66	*	248.24	24h+	24h+
5	8192	44	55	*	186.18	24h+	24h+
6	8192	54	45	*	151.70	24h+	24h+
7	8192	65	34	*	126.03	24h+	24h+
8	8192	76	23	*	107.79	24h+	24h+
9	8192	87	12	*	94.16	24h+	24h+
10	8192	98	1	*	83.59	24h+	24h+

Table 16: waveform

No	$|G|$	$|C|$	$|D|$	$|O|$	$|G|/|C|$	MBTP (s)	Incre (s)
1	5000	1	122	0	5000.00	3332.47	**1.82**
2	5000	14	109	1465	357.14	1808.45	3630.22
3	5000	27	96	50619	185.19	61655.25	24h+
4	5000	41	82	*	121.95	24h+	24h+
5	5000	54	69	*	92.59	24h+	24h+
6	5000	68	55	*	73.53	24h+	24h+
7	5000	81	42	*	61.73	24h+	24h+
8	5000	95	28	*	52.63	24h+	24h+
9	5000	108	15	*	46.30	24h+	24h+
10	5000	122	1	*	40.98	24h+	24h+

6.3. Experiment analysis

The time complexity of MBTP and the incremental algorithm are determined by $|G|$, $|C|$, $|D|$ and $|O|$ (Equations 5 and 6), where $|G|$, $|C|$ and $|D|$ are known for a given decision context but $|O|$ is not. To explain the results in Tables 9-16, we firstly explore the factors determining $|O|$.

From Tables 9-16, it can be seen that as $|C|$ increases, $|O|$ grows in most cases, meaning that $|O|$ is largely determined by $|C|$. Although the result is obtained as $|D|$ decreases, we will explain in the following that, compared with $|C|$, the impact of $|D|$ on $|O|$ can be negligible in most cases.

For a decision context $K = (G, C \cup D, I_C \cup I_D)$, by Definition 7, $|O|$ is equal to the number of decision premises; by Theorem 3 in [15], A is a decision premise if and only if A is a true premise of a decision attribute, and thus the number of decision premises is equal to the number of true premises, i.e., $|O|$ is equal to the number of true premises. Now, we assume that for a decision context with $|C|$ condition attributes, each decision attribute has n true premises on average. Then, K, which has $|C|$ condition attributes and $|D|$ decision attributes, has $|D| \cdot n$ true premises, i.e., $|O| = |D| \cdot n$.

When $|C|$ increases to $|C| + 1$, there are $\mathbf{C}_{|C|+1}^{|C|} = |C| + 1$ sub-decision contexts of K, each of which has $|C|$ condition attributes and $|D|$ decision attributes and hence has $|D| \cdot n$ true premises on average, as assumed above. Now, K has $(|C| + 1) \cdot |D| \cdot n$ true premises and the increment of $|O|$ is $(|C| + 1) \cdot |D| \cdot n - |D| \cdot n = |C| \cdot |D| \cdot n$.

When $|D|$ increases to $|D|+1$, as assumed above, we have $(|D|+1)\cdot n$ true premises and the increment of $|O|$ is $(|D|+1)\cdot n - |D|\cdot n = n$. Thus, when $|D|$ decreases to $|D|-1$, the decrement of $|O|$ will be n.

Thus, when $|C|$ increases to $|C|+1$ and $|D|$ decreases to $|D|-1$, the increase of $|C|$ leads to an increment $|C|\cdot|D|\cdot n$ of $|O|$ and the decrease of $|D|$ leads to a decrement n of $|O|$, thus leading to an increment $|C|\cdot|D|\cdot n - n$ of $|O|$. Thus, in our experiments, when $|C|$ increases to $|C|+i$ and $|D|$ decreases to $|D|-i$, $|O|$ will get the increment of about $n\cdot i\cdot(|C|\cdot|D|-1)$.

We then evaluate the effect of $|G|$ on $|O|$ by selecting four datasets from Tables 9-16. For each decision context listed in these tables, we divided the objects into 10 equal groups and incrementally added one group to the original decision context. Figure 2 records the change of $|O|$ as $|G|$ increases.

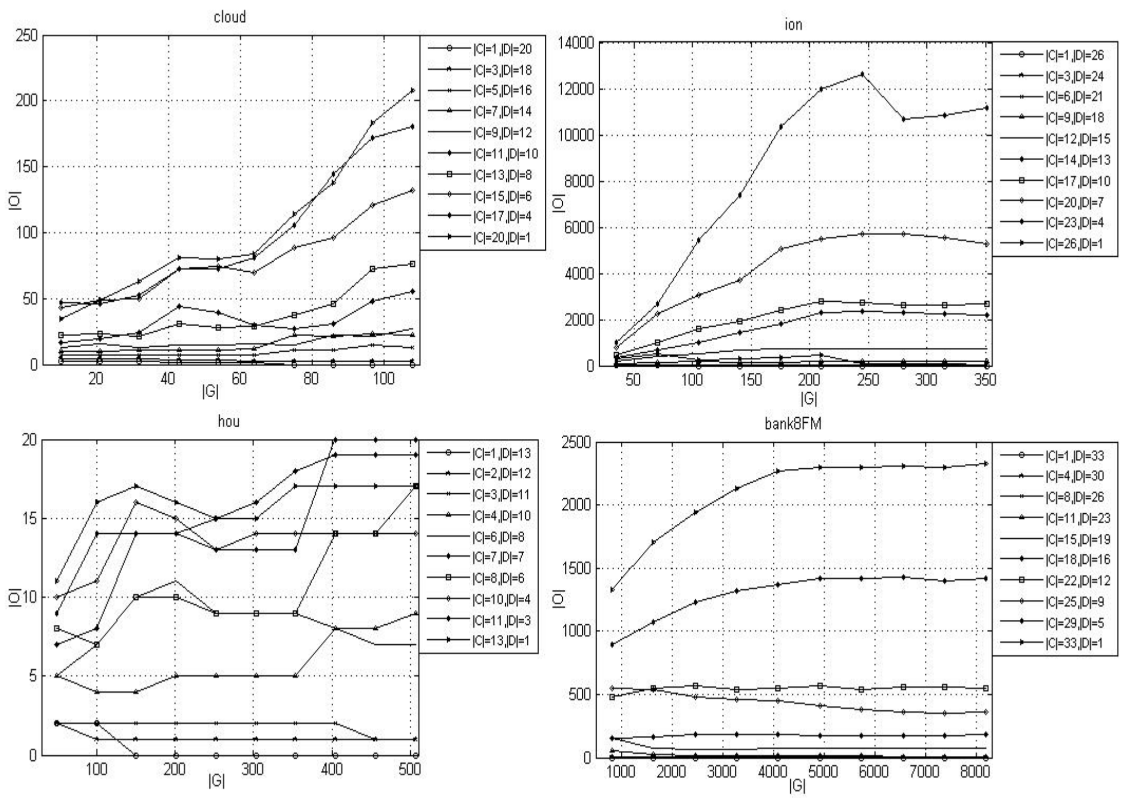

Figure 2: The change of $|O|$ as $|G|$ increases

Figure 2 shows that as $|G|$ increases, $|O|$ grows slowly in most cases; and when $|G|$ is large enough, $|O|$ holds steady.

By the above experiments, we conclude that $|C|$ has a major impact on $|O|$, whereas both $|D|$ and $|G|$ have a limited impact on $|O|$.

We then discuss the key factors that affect the performances of the two algorithms. From Tables 9-16, we can see that, for each dataset, as $|C|$ increases, with $|D|$ decreasing and $|G|$ unchanged, the time consumption of the incremental algorithm grows in most cases, which means that the incremental algorithm is mainly affected by $|C|$.

However, this is not the case with MBTP. For example, for "cloud","ion" and "triazines" (see Tables 9, 11 and 12), as $|C|$ increases, the time consumption of MBTP grows in most cases; but for "hou", "bank8FM" and "dplanes" (see Tables 10, 13 and 14), as $|C|$ increases, it drops in most cases. We then find that, for "cloud","ion" and "triazines", $|G|/|C|$ is small (≤ 50 in most cases).Comparatively, for "hou", "bank8FM" and "dplanes", $|G|/|C|$ is large (> 50 in most cases); and especially for "bank8FM" and "dplanes", $|G|/|C| > 300$ and $|G|/|C| > 1000$ respectively. In this case, we assert that the time consumption of MBTP is related to $|G|/|C|$, and give an explanation for this assertion as follows.

By [15], MBTP mainly includes two sub-functions $getAllgd$ and $DPgenerator$, whose time complexity is respectively $O(|G|^2 \cdot |C| \cdot |D|)$ and $O(|G| \cdot |C|^3 \cdot |D| \cdot |O|^2)$, meaning that $getAllgd$ is mainly affected by $|G|$, and $DPgenerator$ is mainly affected by $|C|$ and $|O|$. As analyzed previously, $|O|$ is mainly affected by $|C|$, and hence $DPgenerator$ is also mainly affected by $|C|$. In this sense, when $|G|/|C|$ is small, i.e., $|G|$ is small but $|C|$ is large, $DPgenerator$ will take more time than $getAllgd$. Thus, as $|C|$ increases, the time consumption of $DPgenerator$ increases, and hence that of MBTP also grows. When $|G|/|C|$ is large, i.e., $|G|$ is large but $|C|$ is small, $getAllgd$ will take more time than $DPgenerator$. Figure 3 further shows that when $|G|/|C|$ is large, such as for "hou", "bank8FM" and "dplanes", as $|C|$ increases, the time consumption of $getAllgd$ is very close to that of the whole algorithm MBTP, and hence they will keep a coincident variation, i.e., as the time consumption of $getAllgd$ decreases, the time consumption of MBTP also drops.

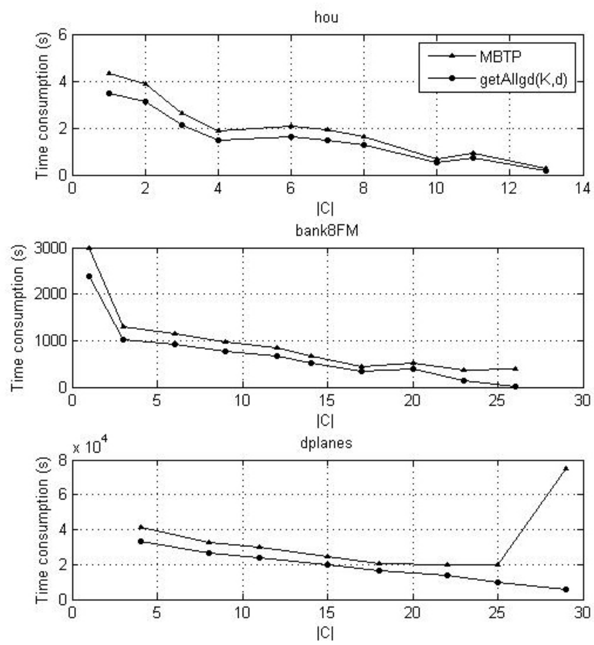

Figure 3: The time consumption of $getAllgd$ and that of MBTP as $|C|$ increases

Based on the above analysis, we can further compare the performances of the two algorithms by Tables 9-16.

First, we notice that, when $|G|/|C|$ is large, the incremental algorithm is more efficient than MBTP. Examplary datasets are "bank8FM" and "dplanes" (Tables 13 and 14). When $|G|/|C| > 300$ in "bank8FM" and $|G|/|C| > 1000$ in "dplanes", compared with MBTP, the incremental algorithm has a remarkable advantage. Especially for "dplanes", where $|G| = 4076$, $|C| = 4$ and $|D| = 29$, MBTP takes 41349.31s but the incremental algorithm just takes 11.19s, with the efficiency being increased by more than 99%. Another example is "hou" (Table 10), where $|G|/|C| > 30$ and both the two algorithms are of low time-consumption, with the incremental algorithm being more efficient.

When $|G|/|C|$ is small, however, the incremental algorithm loses this advantage. Take "ion" and "triazines" (Tables 11 and 12) for example. It is observed that while the incremental algorithm is more efficient than MBTP in the beginning, when $|G|/|C| < 40$ or $|G|/|C| < 5$, MBTP will be more efficient. Another example is "cloud" (Table 9). When $|G|/|C| < 30$, both the two algorithms are of low time-consumption, with MBTP being more efficient.

In addition, we conclude that when $|C|$ is large, both algorithms are time-consuming. For example, in "bank32nh" and "waveform" (Tables 15 and16), when $|C| \geq 33$ or $|C| \geq 41$, both the two algorithms take more than 24 hours and were terminated. That is because when $|C|$ is large, there are a great number of decision implications in O, as analyzed before; and as $|C|$ increases, $|O|$ may have an explosive growth. For example, in "bank32nh", when $|C|$ increases from 11 to 22, $|O|$ increases from 271 to 9149; and in "waveform", when $|C|$ increases from 14 to 27, $|O|$ increases from 1465 to 50619. It is known that, when $|O|$ is large, traversing O will be time-consuming for the two algorithms.

It is noted that, in most practical cases, objects are much more than condition attributes, i.e., $|G|/|C|$ is large, and as analyzed before, the incremental algorithm will be superior to MBTP. Actually, even for the data in which objects are less than condition attributes, new objects may be continually added into data. For example, new purchase records are added moment by moment into the database of a supermarket. In this case, $|G|/|C|$ increases and remains large. MBTP, however, still generates the whole DICB; and by contrast, the incremental algorithm just needs to modify the existing DICB to obtain a new one, which is only a part of work of MBTP. Hence, the incremental algorithm must be also more efficient than MBTP.

7. Conclusion and further work

In this paper, we proposed an incremental algorithm, which produces a new DICB by modifying and updating the existing DICB in the case of new objects being continually added into data. Experimental results verified that when samples in data are much more than condition attributes, the time consumption of generating the whole DICB by incremental algorithm will be remarkably less than that by MBTP. In addition, we conclude that, even for the data in which samples are less than condition attributes, when new objects are continually added into data, the incremental algorithm will also be more efficient than MBTP.

In practice cases, multiple samples but not a single one may be added simultaneously into data. Hence, when a bunch of samples come, how to modify the existing DICB to obtain a new one deserves further exploration. Furthermore, an improved distributed algorithm for DICB generation will be designed in our future study.

DICB is a complete and extremely compact representation of decision information in data. Hence, the DICB-based knowledge representation and reasoning is a valuable topic, which includes studying the effects of inference rules when they are applied to knowledge inference in different orders and different times, designing optimal inference strategies and constructing a system of knowledge representation and inference that takes DICB as its knowledge base and inference rules as its inference engine.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (62072294, 61972238, 61806116), the Key R&D program of Shanxi Province (International Cooperation, 201903D421041), the Natural Science Foundation of Shanxi Province (201801D221175), the Cultivate Scientific Research Excellence Programs of Higher Education Institutions in Shanxi (CSREP) (2019SK036), the Training Program for Young Scientific Researchers of Higher Education Institutions in Shanxi, the Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (2019L0500), and the Shanxi Application Basic Research Plan (201901D211414).

Compliance with Ethical Standards

- **Conflict of interest.** The authors declare no conflict of interest.
- **Ethical approval.** We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. We confirm that we have followed the regulations of our institutions concerning intellectual property.
- **Human and animal studies.** This article does not contain any studies with human participants or animals performed by any of the authors.
- **Data availability.** The datasets generated during the current study are available from the corresponding author on reasonable request.

Reference

[1] Eduard Bartl and Jan Konecny. L-concept lattices with positive and negative attributes: Modeling uncertainty and reduction of size. *Information Sciences*, 472:163–179, 2019.

[2] Radim Belohlavek, Jan Outrata, and Martin Trnecka. Factorizing boolean matrices using formal concepts and iterative usage of essential entries. *Information Sciences*, 489:37–49, 2019.

[3] Claudio Carpineto and Giovanni Romano. *Concept data analysis: Theory and applications*. John Wiley & Sons, 2004.

[4] Aswani Kumar Cherukuri, Sérgio M. Dias, and Newton Josě Vieira. Knowledge reduction in formal contexts using non-negative matrix factorization. *Mathematics and Computers in Simulation*, 109:46–63, 2015.

[5] Gabriel Ciobanu and Cristian Văideanu. A note on similarity relations between fuzzy attribute-oriented concept lattices. *Information Sciences*, 460-461:254–263, 2018.

[6] M. Eugenia Cornejo, Jesús Medina, and EloÍsa RamÍrez-Poussa. Characterizing reducts in multi-adjoint concept lattices. *Information Sciences*, 422:364–376, 2018.

[7] Bernhard Ganter, Gerd Stumme, and Rudolf Wille. *Formal concept analysis: foundations and applications*. Springer, Berlin, 2005.

[8] Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis: Mathematical Foundations*. Springer Science & Business Media, 1999.

[9] Robert Godin, Rokia Missaoui, and Hassan Alaoui. Incremental concept formation algorithms based on galois (concept) lattices. *Computational intelligence*, 11(2):246–267, 1995.

[10] Keyun Hu, Yuchang Lu, and Chunyi Shi. An integrated mining approach for classification and association rule based on concept lattice. *Journal of software*, 11(11):1479–1484, 2000.

[11] Hayato Ishigure, Atsuko Mutoh, Tohgoroh Matsui, and Nobuhiro Inuzuka. Concept lattice reduction using attribute inference. In *IEEE Global Conference on Consumer Electronics*, pages 108–111, 2015.

[12] Jan Konecny and Petr Krajča. On attribute reduction in concept lattices: Experimental evaluation shows discernibility matrix based methods inefficient. *Information Sciences*, 467:431–445, 2018.

[13] Jan Konecny and Petr Krajča. On attribute reduction in concept lattices: the polynomial time discernibility matrix-based method becomes the cr-method. *Information Sciences*, 491:48–62, 2019.

[14] S. O. Kuznetsov and T. Makhalova. On interestingness measures of formal concepts. *Information Sciences*, 442-443:202–219, 2018.

[15] Deyu Li, Shaoxia Zhang, and Yanhui Zhai. Method for generating decision implication canonical basis based on true premises. *International journal of machine learning and cybernetics*, 8(1):57–67, 2017.

[16] Jinhai Li, Chenchen Huang, Jianjun Qi, Yuhua Qian, and Wenqi Liu. Three-way cognitive concept learning via multi-granularity. *Information Sciences*, 378(1):244–263, 2017.

[17] Jinhai Li, Cherukuri Aswani Kumar, Changlin Mei, and Xizhao Wang. Comparison of reduction in formal decision contexts. *International Journal of Approximate Reasoning*, 80:100–122, 2017.

[18] Jinhai Li, Changlin Mei, Cherukuri Aswani Kumar, and Xiao Zhang. On rule acquisition in decision formal contexts. *International journal of machine learning and cybernetics*, 4(6):721–731, 2013.

[19] Jinhai Li, Changlin Mei, and Yuejin Lv. Knowledge reduction in decision formal contexts. *Knowledge-Based Systems*, 24(5):709–715, 2011.

[20] Jinhai Li, Changlin Mei, and Yuejin Lv. Knowledge reduction in formal decision contexts based on an order-preserving mapping. *International Journal of General Systems*, 41(2):143–161, 2012.

[21] Jinhai Li, Changlin Mei, and Yuejin Lv. Knowledge reduction in real decision formal contexts. *Information Sciences*, 189:191–207, 2012.

[22] Jinhai Li, Changlin Mei, and Yuejin Lv. Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction. *International Journal of Approximate Reasoning*, 54(1):149–165, 2013.

[23] Jinhai Li, Changlin Mei, Weihua Xu, and Yuhua Qian. Concept learning via granular computing:a cognitive viewpoint. *Information Sciences*, 298:447–467, 2015.

[24] Jinhai Li and Weizhi Wu. Granular computing approach for formal concep analysis and its research outlooks. *Journal of Shandong University (Natural Science)*, 52(7):1–12, 2017.

[25] Junyu Li, Zeng Qiang, and Xia Wang. Attribute reduction in property oriented concept lattices of formal contexts and formal decision contexts. *International Journal of Database Theory and Application*, 9(9):37–48, 2016.

[26] Yuan Ma, Xuedong Zhang, and Chengying Chi. Compact dependencies and intent waned values. *Journal of Software*, 22(5):962–971, 2011.

[27] A. Mouakher and S. Ben Yahia. On the efficient stability computation for the selection of interesting formal concepts. *Information Sciences*, 472:15–34, 2019.

[28] Nanjia, Shaoxia Zhang, Yanhui Zhai, and Deyu Li. Study of inference rules and deduction process based on decision implications. *Journal of Frontiers of Computer Science and Technology*, 2019.

[29] Sebastiao M. Neto, Luis E. Zárate, and Mark A. J. Song. Handling high dimensionality contexts in formal concept analysis via binary decision diagrams. *Information Sciences*, 429:361–376, 2018.

[30] Kaishe Qu and Yanhui Zhai. Generating complete set of implications for formal contexts. *Knowledge-Based Systems*, 21(5):429–433, 2008.

[31] Kaishe Qu, Yanhui Zhai, Jiye Liang, and Min Chen. Study of decision implications based on formal concept analysis. *International journal of general systems*, 36(2):147–156, 2007.

[32] Camille Roth, Sergei A. Obiedkov, and Derrick G. Kourie. Towards concise representation for taxonomies of epistemic communities. In *Concept Lattices and Their Applications*, pages 240–255. Springer, 2008.

[33] Mehran Sahami. Learning classification rules using lattices. In *Machine Learning: ECML-95*, pages 343–346. Springer, 1995.

[34] Mingwen Shao. Knowledge acquisition in decision formal contexts. In *Proceedings of the sixth international conference on machine learning and cybernetics*, pages 4050–4054, Hong Kong, 2007.

[35] Mingwen Shao, Yee Leung, and Weizhi Wu. Rule acquisition and complexity reduction in formal decision contexts. *International Journal of Approximate Reasoning*, 55(1):259–274, 2014.

[36] Francisco J. Valverde-Albacete and CarmenPeláez-Moreno. K-formal concept analysis as linear algebra over idempotent semifields. *Information Sciences*, 467:579–603, 2018.

[37] Yun Wang, ZhiHai Hu, Ke Hu, Xuegang Hu, Zongtian Liu, and Diancheng Zhang. General and incremental algorithms of rule extraction based on concept lattice. *Chinese Journal of Computers*, 22(01):66–70, 1999.

[38] Rudolf Wille. *Restructuring lattice theory: an approach based on hierarchies of concepts*. Springer, 2009.

[39] Weizhi Wu, Leung Yee, and Jusheng Mi. Granular computing and knowledge reduction in formal contexts. *IEEE Transactions on Knowledge and Data Engineering*, 21(10):1461–1474, 2009.

[40] Weihua Xu and Wentao Li. Granular computing approach to two-way learning based on formal concept analysis in fuzzy datasets. *IEEE transactions on cybernetics*, 46(2):366–379, 2016.

[41] Weihua Xu, Jinzhong Pang, and Shuqun Luo. A novel cognitive system model and approach to transformation of information granules. *International Journal of Approximate Reasoning*, 55(3):853–866, 2014.

[42] Yanhui Zhai and Deyu Li. Knowledge structure preserving fuzzy attribute reduction in fuzzy formal context. *International Journal of Approximate Reasoning*, 115:209–220, 2019.

[43] Yanhui Zhai, Deyu Li, and Kaishe Qu. Probability fuzzy attribute implications for interval-valued fuzzy sets. *International Journal of Database Theory and Application*, 5(4):95–108, 2012.

[44] Yanhui Zhai, Deyu Li, and Kaishe Qu. Fuzzy decision implications. *Knowledge-Based Systems*, 37:230–236, 2013.

[45] Yanhui Zhai, Deyu Li, and Kaishe Qu. Decision implications: a logical point of view. *International Journal of Machine Learning and Cybernetics*, 5(4):509–516, 2014.

[46] Yanhui Zhai, Deyu Li, and Kaishe Qu. Canonical basis for decision implications. *Chinese Journal of Electronics*, 43(1):18–23, 2015.

[47] Yanhui Zhai, Deyu Li, and Kaishe Qu. Decision implication canonical basis: a logical perspective. *Journal of Computer and System Sciences*, 81(1):208–218, 2015.

[48] Yanhui Zhai, Deyu Li, and Kaishe Qu. Fuzzy decision implication canonical basis. *International Journal of Machine Learning and Cybernetics*, 9(11):1909–1917, 2018.

[49] Yanhui Zhai, Deyu Li, and Jing Zhang. Variable decision knowledge representation: A logical description. *Journal of Computational Science*, 25:161–169, 2018.

[50] Chao Zhang, Deyu Li, and Jiye Liang. Hesitant fuzzy linguistic rough set over two universes model and its applications. *International Journal of Machine Learning and Cybernetics*, 9(4):577–588, 2018.

[51] Chao Zhang, Deyu Li, and Jiye Liang. Multi-granularity three-way decisions with adjustable hesitant fuzzy linguistic multigranulation decision-theoretic rough sets over two universes. *Information Sciences*, 507:665–683, 2020.

[52] Li Zhang, Jianming Zhan, and Zeshui Xu. Covering-based generalized if rough sets with applications to multi-attribute decision-making. *Information Sciences*, 478:275–302, 2019.

[53] Shaoxia Zhang, Deyu Li, Yanhui Zhai, and Xiangping Kang. A comparative study of decision implication, concept rule and granular rule. *Information Sciences*, 508:33–49, 2020.

[54] Huilai Zhi and Jinhai Li. Granule description based knowledge discovery from incomplete formal contexts via necessary attribute analysis. *Information Sciences*, 485:347–361, 2019.

Figures

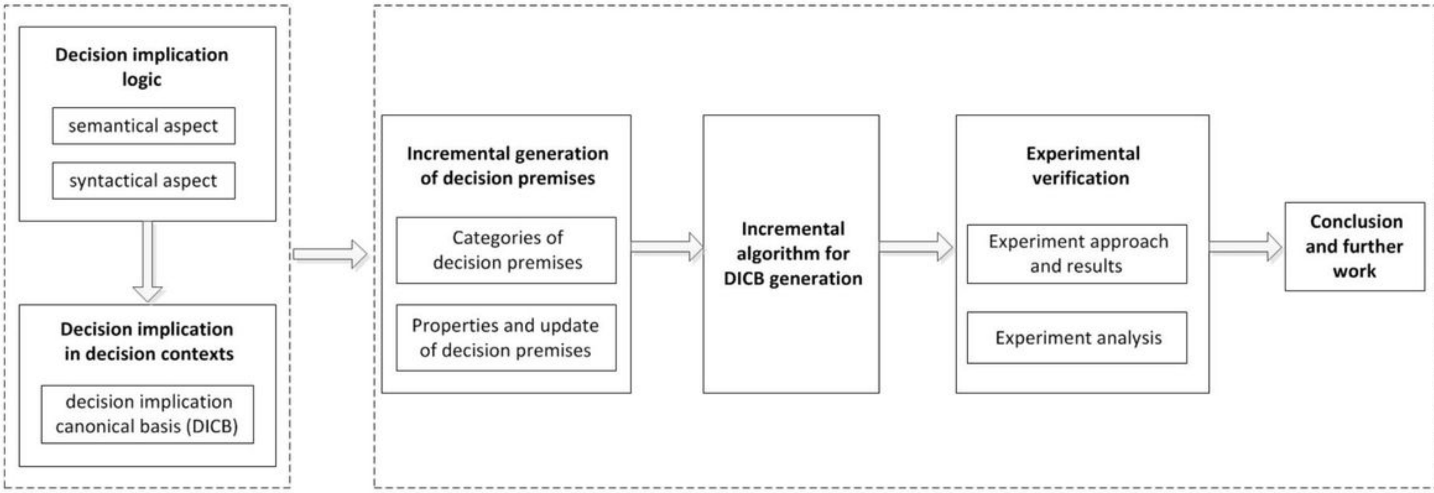

Figure 1

The structure of the article

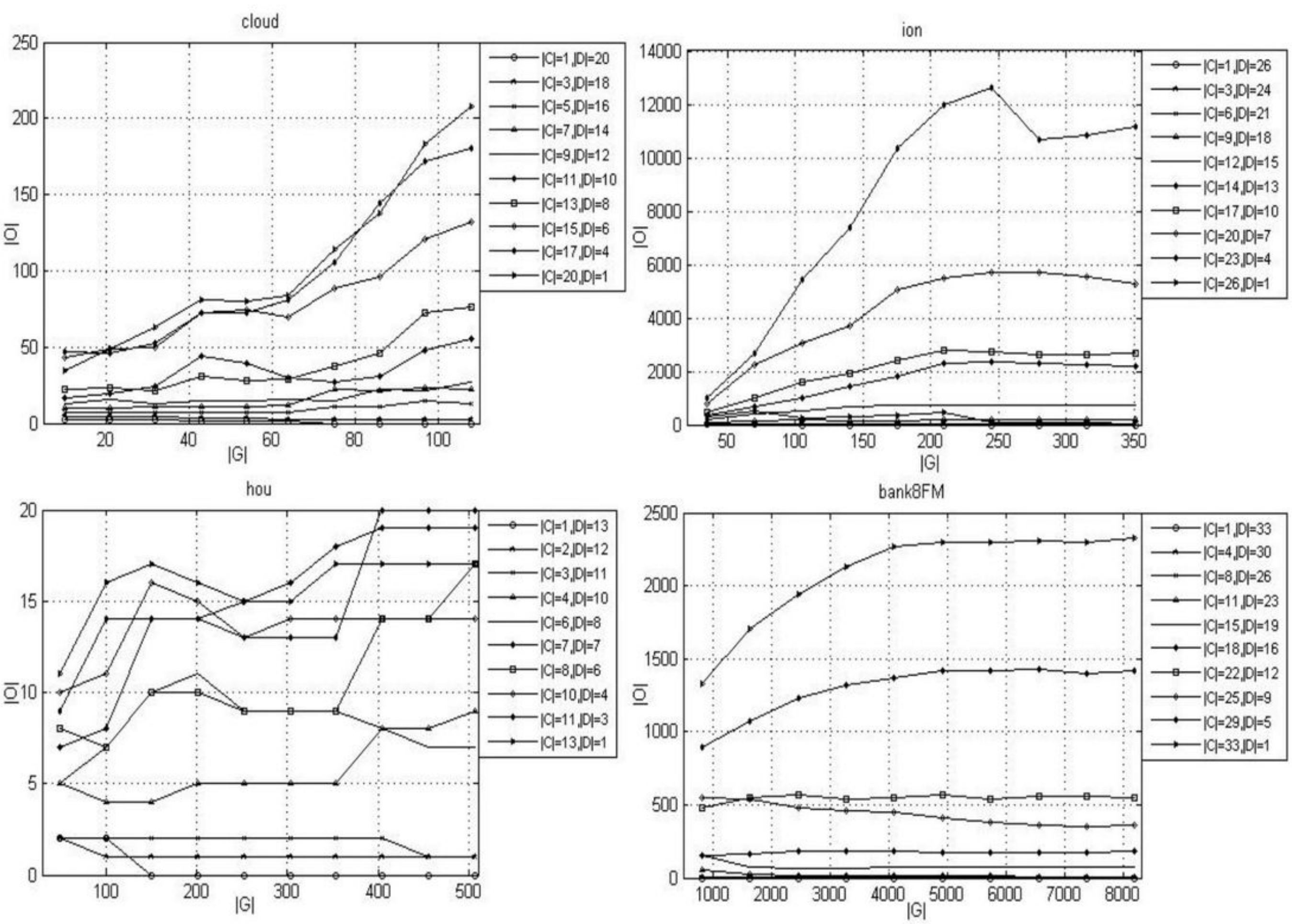

Figure 2

The change of |O| as |G| increases

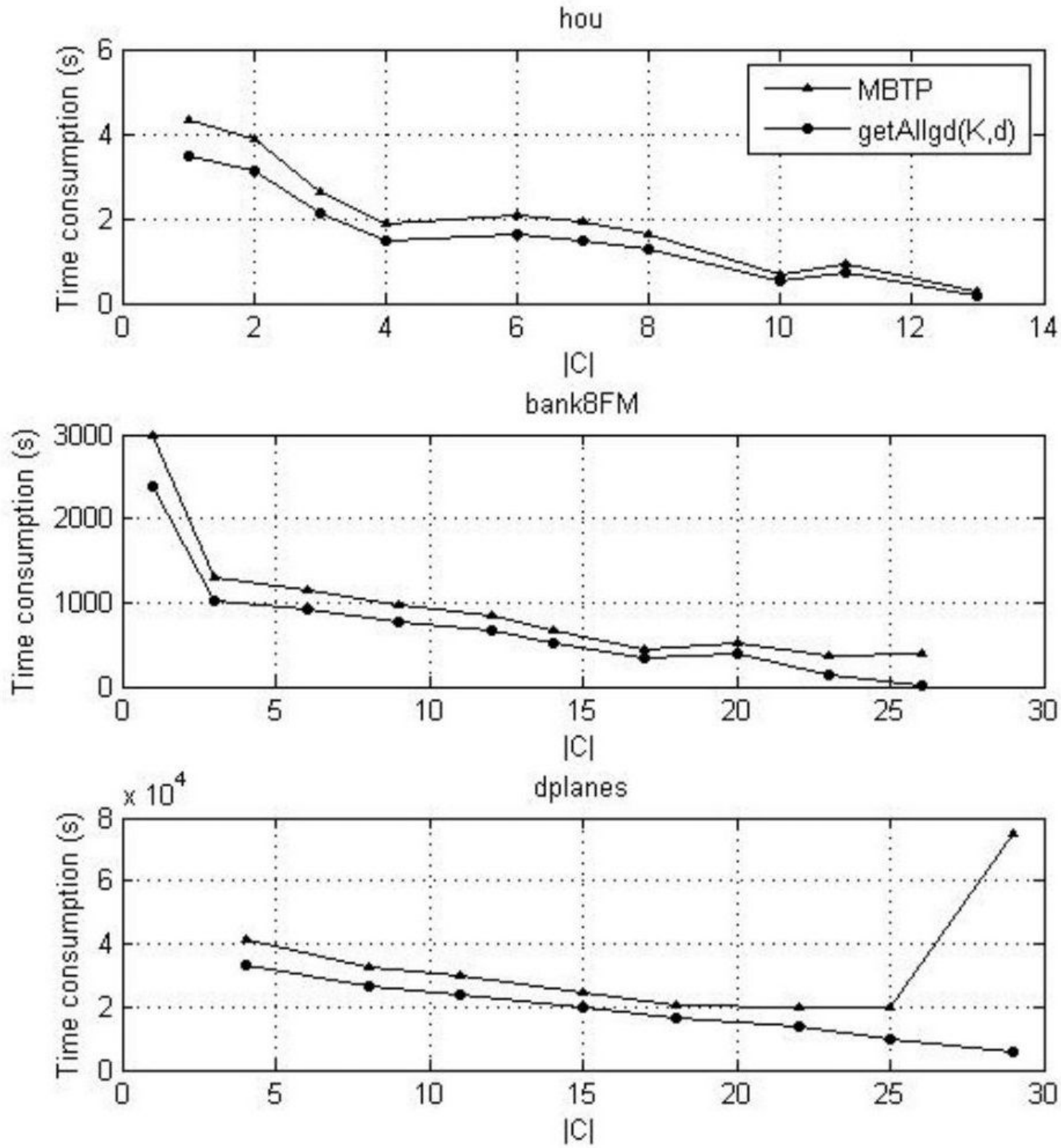

Figure 3

The time consumption of getAllgd and that of MBTP as |C| increases