Computers and Mathematics with Applications 61 (2011) 1096–1106

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

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

A heuristic knowledge-reduction method for decision formal contexts

Jinhai Li[a,*], Changlin Mei[a], Yuejin Lv[b]

[a] School of Science, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, PR China

[b] School of Mathematics and Information Sciences, Guangxi University, Nanning, Guangxi 530004, PR China

ARTICLE INFO

Article history:
Received 13 March 2010
Received in revised form 24 October 2010
Accepted 27 December 2010

Keywords:
Decision formal context
Concept lattice
Knowledge reduction
Heuristic algorithm
Time complexity

ABSTRACT

Computing a minimal reduct of a decision formal context by Boolean reasoning is an NP-hard problem. Thus, it is essential to develop some heuristic methods to deal with the issue of knowledge reduction especially for large decision formal contexts. In this study, we first investigate the relationship between the concept lattice of a formal context and those of its subcontexts in preparation for deriving a heuristic knowledge-reduction method. Then, we construct a new framework of knowledge reduction in which the capacity of one concept lattice implying another is defined to measure the significance of the attributes in a consistent decision formal context. Based on this reduction framework, we formulate an algorithm of searching for a minimal reduct of a consistent decision formal context. It is proved that this algorithm is complete and its time complexity is polynomial. Some numerical experiments demonstrate that the algorithm can generally obtain a minimal reduct and is much more efficient than some Boolean reasoning-based methods.

1. Introduction

Formal concept analysis (FCA), proposed by Wille [1] in 1982, is one of the effective mathematical tools for conceptual data analysis and knowledge processing. Two important notions in FCA are formal context and formal concept. The family of all formal concepts of a formal context forms a complete lattice [2] which is termed as the concept lattice of the formal context in FCA and reflects the relationship of generalization and specialization among the formal concepts. Nowadays, FCA has been applied to a variety of fields such as data mining, machine learning, artificial intelligence and software engineering [3–11].

As is well known, much attention has been paid to the issue of knowledge reduction in rough set theory [12] and many reduction methods have been proposed for information systems and decision tables [13–16]. Since computing a minimal reduct of an information system or a decision table by Boolean reasoning is an NP-hard problem [17], some heuristic methods have been developed to find an approximate solution instead [18,19]. Similar to the case in rough set theory, knowledge reduction is also one of the key issues in FCA. In fact, these two theories often complement one another in data analysis and some studies have been devoted to combining them in a common framework [20,21].

Recently, there has been growing interest in knowledge reduction in FCA. For instance, Ganter and Wille [2] proposed a knowledge-reduction method by removing the reducible objects and attributes of a formal context. Elloumi et al. [22] put forward a multilevel reduction approach in which some rows in the initial context may be removed at a given precision level without changing the association rules derived from the reduced databases. In the sense of lattice isomorphism, Zhang et al. [23] presented a knowledge-reduction method in formal contexts and, from the viewpoint of rough set theory, Liu et al. [24] proposed two knowledge reduction approaches. Additionally, some methods for knowledge reduction in

* Corresponding author. Tel.: +86 29 82665483; fax: +86 29 82663938.
E-mail addresses: jhlixjtu@163.com (J. Li), clmei@mail.xjtu.edu.cn (C. Mei), lvyjin@126.com (Y. Lv).

0898-1221/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2010.12.060

consistent decision formal contexts were also explored. For example, Wang and Zhang [25] developed a method to compute such reducts that can make each image in the decision concept lattice have at least one preimage in the conditional concept lattice. Wei et al. [26] investigated the issue of knowledge reduction in consistent decision formal contexts by defining a strongly partial order and a weakly partial order between the conditional concept lattice and the decision concept lattice. Wu et al. [27] put forward the notion of granular reduction in consistent formal decision contexts and developed some approaches for computing granular reducts.

A minimal reduct of a decision formal context plays an important role in rule acquisition. However, like that in rough set theory, computing a minimal reduct of a consistent decision formal context by Boolean reasoning is still an NP-hard problem [27]. Therefore, the existing methods such as these in [25–27] for computing a minimal reduct are computationally expensive and they are even impossibly implemented for a large database. In this paper, we develop a heuristic method with polynomial time complexity to search for a minimal reduct of a consistent decision formal context. Numerical experiments demonstrate that this method can in general obtain a minimal reduct and is much more efficient than some Boolean reasoning-based methods.

The rest of this paper is organized as follows. In Section 2, we briefly introduce some basic notions and results related to formal contexts and discuss the relationship between the concept lattice of a formal context and those of its subcontexts. We construct in Section 3 a new framework of knowledge reduction for consistent decision formal contexts. In Section 4, we formulate a heuristic algorithm with polynomial time complexity to search for a minimal reduct of a consistent decision formal context. Some numerical experiments are conducted in Section 5 to access the performance of the proposed algorithm. The paper is concluded with a brief summary.

2. Preliminaries

In this section, we briefly introduce some basic notions and results about formal contexts and further investigate the relationship between the concept lattice of a formal context and those of its subcontexts.

2.1. Formal contexts and concept lattices

Definition 1 ([1]). A formal context is a triple (U, A, I), where $U = \{x_1, x_2, \ldots, x_n\}$, called the universe of discourse, is a nonempty and finite set of objects, $A = \{a_1, a_2, \ldots, a_m\}$ is a nonempty and finite set of attributes, and $I \subseteq U \times A$ is a binary relation between U and A with $(x, a) \in I$ indicating that the object x owns the attribute a.

In this paper, we assume that the binary relation I is *regular*. That is, for any $(x, a) \in U \times A$, it satisfies the following conditions:

(1) There exist $a_1, a_2 \in A$ such that $(x, a_1) \in I$ and $(x, a_2) \notin I$, and
(2) there exist $x_1, x_2 \in U$ such that $(x_1, a) \in I$ and $(x_2, a) \notin I$.

For $X \subseteq U$ and $B \subseteq A$, define

$$\begin{aligned} X^* &= \{a \in A \mid \forall x \in X, (x, a) \in I\}, \\ B^* &= \{x \in U \mid \forall a \in B, (x, a) \in I\}. \end{aligned} \tag{1}$$

That is, X^* is the maximal family of the attributes that all the objects in X have in common and B^* is the maximal family of the objects shared by all the attributes in B.

Definition 2 ([1]). Let $\mathbb{K} = (U, A, I)$ be a formal context. For $X \subseteq U$ and $B \subseteq A$, the ordered pair (X, B) is called a formal concept (or simply a concept) of $\mathbb{K}$ if it satisfies $X^* = B$ and $B^* = X$. Here, X and B are termed, respectively, as the extension and the intension of the formal concept (X, B). The sets of all the formal concepts, all the extensions, and all the intensions of (U, A, I) are denoted by $\mathfrak{B}(U, A, I)$, $\mathfrak{U}(U, A, I)$, and $\mathfrak{I}(U, A, I)$, respectively.

Proposition 1 ([1]). *Let $\mathbb{K} = (U, A, I)$ be a formal context. For $X_1, X_2, X \subseteq U$ and $B_1, B_2, B \subseteq A$, we have the following conclusions:*

(1) $X_1 \subseteq X_2 \Rightarrow X_2^* \subseteq X_1^*$; $B_1 \subseteq B_2 \Rightarrow B_2^* \subseteq B_1^*$.
(2) $X \subseteq X^{**}$ *and* $B \subseteq B^{**}$.
(3) (X^{**}, X^*) *and* (B^*, B^{**}) *are two formal concepts of* $\mathbb{K}$.

The set of all formal concepts of a formal context (U, A, I) forms a complete lattice [2], called the *concept lattice* of (U, A, I) and denoted by $\mathfrak{B}(U, A, I)$. The *meet* and *join* in $\mathfrak{B}(U, A, I)$ are defined by

$$\begin{aligned} &(X_1, B_1) \wedge (X_2, B_2) = (X_1 \cap X_2, (B_1 \cup B_2)^{**}) \quad \text{and} \\ &(X_1, B_1) \vee (X_2, B_2) = ((X_1 \cup X_2)^{**}, B_1 \cap B_2), \end{aligned} \tag{2}$$

respectively. The partial order relation $\preceq$ in $\mathfrak{B}(U, A, I)$ is defined as follows: For $(X_1, B_1), (X_2, B_2) \in \mathfrak{B}(U, A, I)$,

$$(X_1, B_1) \preceq (X_2, B_2) \Longleftrightarrow X_1 \subseteq X_2 \Longleftrightarrow B_2 \subseteq B_1. \tag{3}$$