

On Invariance of Concept Stability for Attribute Reduction in Concept Lattice



Fei Hao, Erhe Yang, Lantian Guo, Aziz Nasridinov, and Doo-Soon Park

Abstract Formal Concept Analysis (FCA) methodology, as an efficient knowledge representation and knowledge discovery tool and has been widely used in various fields, such as data mining, expert systems, and others. Knowledge reduction is an essential issue for knowledge discovery. This paper focuses on attribute reduction in FCA and explores the internal relation between concept stability and attribute reduction. By observing the concept stability of concepts in original concept lattice and reduced concept lattice, a theorem about the invariance of concept stability for attribute reduction in concept lattice is presented and proved mathematically. It is believed that the proposed theorem provides a novel solution for quick attribute reduction and benefit for other social system applications.

Keywords FCA · Attribute reduction · Concept stability · Invariance

This research was supported by the National Natural Science Foundation of China (Grant No. 61702317), MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2019-2014-1-00720) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation) and the National Research Foundation of Korea (No. 2017R1A2B1008421) and was also supported by the Natural Science Basic Research Plan in Shaanxi Province of China (2019JM-379).

F. Hao · E. Yang
School of Computer Science, Shaanxi Normal University, Xi'an, China
e-mail: feehao@gmail.com

L. Guo
School of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao, China
e-mail: guolt0211@gmail.com

A. Nasridinov
Department of Computer Science, Chungbuk National University, Cheongju, Korea
e-mail: aziz@chungbuk.ac.kr

D.-S. Park (✉)
Department of Computer Software Engineering, Soonchunhyang University, Asan, Korea
e-mail: parkds@sch.ac.kr

1 Introduction

Formal Concept Analysis (FCA), as an effective methodology for knowledge discovery, has attracted much attention from the artificial intelligence community. At present, FCA has been widely used in machine learning [1], pattern recognition, expert systems [2], decision-making and data mining, and other related fields [3].

Knowledge reduction is an essential direction for knowledge discovery. In particular, attribute reduction in concept lattice provides powerful data reduction solution for knowledge discovery and data mining. Besides, attribute reduction as a key issue of FCA, its goal is to find minimal attribute set to maintain the invariant property. Generally, attribute reduction in FCA can be categorized as follows in terms of different reduction purpose: (1) **concept lattice complexity based attribute reduction**: this reduction aims to reduce the number of nodes in concept lattice and to delete the redundant concepts [4]; (2) **rule-based attribute reduction**: this reduction aims to maintain the rule sets unchanged [5]; (3) **lattice structure based attribute reduction**: this type of reduction is commonly studied in literature. It targets to keep the same structure of concept lattice and deletes the redundant attributes [6]. The attribute reduction in this work falls into the category (3).

This paper focuses on investigating the relation between concept stability in the original lattice and reduced lattice. For this, we introduce the definition of concept stability, which is used to measure the strength of dependency between intent and objects of extent. The main contributions of this paper are two-fold: (1) we present a theorem to prove that the concept stability is not changed during the attribute reduction of concept lattice (2).

The rest of this paper is organized as follows. Section 2 provides the preliminaries about Formal Concept Analysis methodology and elaborates the attributes reduction in concept lattice. By observing the concept stability of concepts in original concept lattice and reduced concept lattice, a theorem about the invariance of concept stability for attribute reduction in concept lattice is presented and proved mathematically in Sect. 3. Finally, Sect. 4 concludes this work.

2 Attribute Reduction in Concept Lattice

Before the presentation of attribute reduction in concept lattice, we briefly provide the preliminaries of FCA. FCA is a powerful methodology for describing the binary relationships between object and attribute and has been applied to many areas. Formally, a formal context is formulated as $C = (O, A, I)$ where O indicates the object set and A denotes the attribute set respectively, and the relation $I \subseteq O \times A$ refers to a binary relation between object and attribute. Generally, $o \in O$ and $a \in A$, $(o, a) \in I$ is interpreted as objective o has the attribute a .

For the sake clarity of formal concept lattice and its generated formal concepts, the following two operators are given [7].

(Operator for extracting the common attribute of objects subset X) For $X \subseteq O$, we define a set of common attributes of X,

$$X^\uparrow = \{ a \in A \mid (x,a) \in I, \forall x \in X \} \quad (1)$$

(Operator for extracting the common objects of attributes subset Y) For $Y \subseteq A$, we also define a set of common objects of Y,

$$Y^\downarrow = \{ o \in O \mid (o,y) \in I, \forall y \in Y \} \quad (2)$$

In a formal context $C = (O, A, I)$, for $X \subseteq O, Y \subseteq A$, if $X^{\uparrow\downarrow} = Y$, then this pair (X, Y) is called as a concept where X, Y are the extent and intent of the concept. Let $\text{Ext}(C)$ be the set of extents w.r.t. the formal context C. With the above operators, a concept lattice $L(C)$ can be defined as concepts organized according to a special hierarchical partial order, i.e., $(X_1, Y_1) \leq (X_2, Y_2) \Leftrightarrow X_1 \subseteq X_2 (\Leftrightarrow Y_1 \supseteq Y_2)$.

Suppose that $C = (O, A, I)$ is a formal context, $D \subseteq A, I_D = I \cap (O \times D)$, then $C_D = (O, D, I_D)$ is a formal context as well and regarded as a sub-formal context of C.

Theorem 1 Suppose $C = (O, A, I)$ as a formal context, if $D \subseteq A$, then $\text{Ext}(C) \supseteq \text{Ext}(C_D)$ holds.

Theorem 1 demonstrates that the set of extents of sub-formal context is contained with the set of extents of original formal context.

Definition 1 Let $C_1 = (O, A_1, I_1)$ and $C_2 = (O, A_2, I_2)$ be two formal contexts, $L(C_1)$ and $L(C_2)$ be the corresponding concept lattices. For any concept $(X_2, Y_2) \in L(C_2)$, there exists $(X_1, Y_1) \in L(C_1)$ that satisfies $X_1 = X_2$, then we say that $L(C_1)$ is a refined version of $L(C_2)$, denoted as $L(C_1) \leq L(C_2)$.

Particularly, if $L(C_1) \leq L(C_2)$ and $L(C_2) \leq L(C_1)$ hold simultaneously, then the concept lattices of $C_1 = (O, A_1, I_1)$ and $C_2 = (O, A_2, I_2)$ are isomorphism, i.e., $L(C_1) \cong L(C_2)$.

Definition 2 Let $C = (O, A, I)$ be a formal context, $D \subseteq A$. If $L(C) \cong L(C_D)$, then D is the consistent set of C. Additionally, for any $d \in D, L(C_D) \not\cong L(C_{D-\{d\}})$, then D is the reduction of C.

For the sake of better presentation, this paper takes the above consistent set/reduction as the consistent/reduction of the concept lattice. As matter of fact, a consistent set D of a formal context C is a type of attributes set that maintains the invariant of extent set, i.e., $\text{Ext}(C) = \text{Ext}(C_D)$.

Example 1 Table 1 shows a formal context $C = (O, A, I)$, $O = \{1, 2, 3, 4\}$, $A = \{a, b, c, d, e\}$, the corresponding concept lattice and its reduction are presented as follows.

Table 1 A formal context $C = (O, A, I)$

| | a | b | c | d | e |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 1 | 1 |
| 2 | 1 | 1 | 1 | 0 | 0 |
| 3 | 0 | 0 | 0 | 1 | 0 |
| 4 | 1 | 1 | 1 | 0 | 0 |

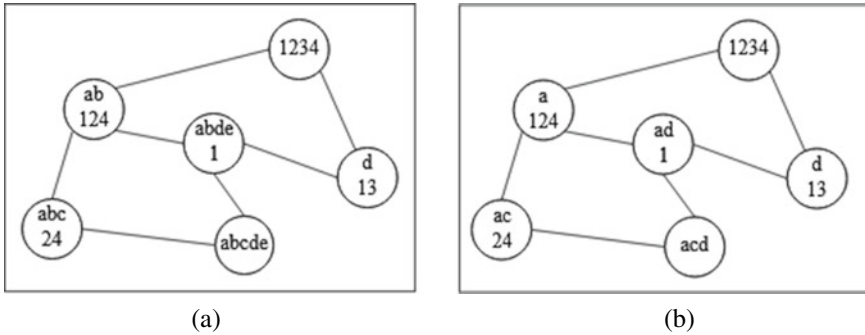


Fig. 1 Concept lattices $L(O, A, I)$ and $L(O, R_1, I_{R1})$

By invoking the concept lattice generation algorithm, the corresponding concept lattice for $C = (O, A, I)$ is shown in Fig. 1(a). Particularly, there are 6 concepts: $(1, abde)$, $(24, abc)$, $(13, d)$, $(124, ab)$, $(1234, \phi)$, $(\phi, abcde)$, denoted as $FC_i (i = 1, 2 \dots 6)$.

There are two reductions of this formal context, i.e., $R_1 = \{a, c, d\}$, $R_2 = \{b, c, d\}$. In this example, the attributes c and d are core attributes, attributes a and b are relatively necessary attributes, while attribute e is the redundant one. Figure 1(b) is the concept lattice of the formal context (O, R_1, I_{R1}) . Obviously, $L(O, R_1, I_{R1}) \cong L(O, A, I)$.

As can be seen from Fig. 1a, b, the extent of concepts are not changed, and the intent (i.e., attribute) are reduced from Fig. 1a. The structure of the lattices is an isomorphism.

3 On Invariance of Concept Stability for Attribute Reduction in Concept Lattice

In this section, we attempt to explore the relation between concept stability and attribute reduction in concept lattices. Firstly, we define concept stability. Further, the relation between concept stability and attribute reduction in concept lattices is figured out.

Definition 3 (*Stability Index*) [7] Let $c = (X, Y)$ be a concept of formal context $C = (O, A, I)$, then the intensional stability of (X, Y) is defined as follows.

$$\sigma(c) = \frac{|\{e \in \wp(X) | e' = Y\}|}{2^{|X|}} \tag{3}$$

In Eq. (3), intensional stability $\sigma(c)$ is used to measure the strength of dependency between the intent Y and the objects of the extent X . Specifically, it expresses the probability to maintain Y closed when a subset of noisy objects in X are deleted with equal probability. In other words, this index quantifies the amount of noise in the extent X and overfitting in the intent Y .

Inspired by this stability index, we think the redundant attributes will not affect the intensional stability $\sigma(c)$ when we delete some of them. Therefore, to answer our guess, it is necessary to investigate the relation between concept stability and attribute reduction in concept lattices.

Let us continue the Example 1, the stability for each concept in $L(O, A, I)$ can be obtained (as shown in Table 2) according to Definition 3.

Similarly, the stability for each concept in $L(O, R_1, I_{R_1})$ are also obtained as shown in Table 3.

By observing the above concept stability, we found that there is no any change from the original concept lattice to the reduced concept lattice which is obtained after deleting the redundant attributes. Therefore, the following theorem can be derived.

Theorem 2 Given a formal context $C = (O, A, I)$, its corresponding concept lattice is represented as $L(O, A, I)$. After deleting the redundant attributes, the reduced

Table 2 Concept stability of each concept in $L(O, A, I)$

| Concept | Extent | Intent | Stability |
|---------|---------|-------------|---------------|
| FC_1 | 1 | a,b,d,e | 0.5 |
| FC_2 | 2,4 | a,b,c | 0.75 |
| FC_3 | 1,3 | D | 0.5 |
| FC_4 | 1,2,4 | a,b | 0.375 |
| FC_5 | 1,2,3,4 | ϕ | 0.0625 |
| FC_6 | ϕ | a,b,c,d,e | 1 |

Table 3 Concept stability of each concept in $L(O, R_1, I_{R_1})$

| Concept | Extent | Intent | Stability |
|---------|---------|---------|---------------|
| FC_1 | 1 | a, d | 0.5 |
| FC_2 | 2,4 | a, c | 0.75 |
| FC_3 | 1,3 | d | 0.5 |
| FC_4 | 1,2,4 | a | 0.375 |
| FC_5 | 1,2,3,4 | ϕ | 0.0625 |
| FC_6 | ϕ | a,c,d | 1 |

concept lattice is $L(O, R_1, I_{R_1})$. The stability of concepts is not changed during the reduction.

Proof Since the extracted concepts $(X, R), R \in R_1$ in reduced concept lattice $L(O, R_1, I_{R_1})$ have the same extent with original concept lattice $L(O, A, I)$. In addition, the redundant attributes are deleted from original attributes, that is to say, the $|\{e \in \wp(X) | e' = Y\}|$ equals to $|\{e \in \wp(X) | e' = R\}|$. Due to these, redundant attributes cannot affect the discernibility for a given extent X . Hence, we have the following equation:

$$\sigma(c) = \frac{|\{e \in \wp(X) | e' = R\}|}{2^{|X|}} = \frac{|\{e \in \wp(X) | e' = Y\}|}{2^{|X|}}$$

The above Theorem 1 holds.

4 Conclusions

This paper is the first work to explore the internal relation between concept stability and attribute reduction. First, this paper introduced the definition of stability index, then the concept stability of concepts in original concept lattice and reduced concept lattice are observed respectively; further, a theorem about the invariance of concept stability for attribute reduction in concept lattice is presented and proved mathematically. In the future, we will fully adopt the proposed theorem and devise a quick attribute reduction algorithm that can be used in other social system applications.

References

1. Shao M, Guo L, Wang C (2018) Connections between two-universe rough sets and formal concepts. *Int J Mach Learn Cybern* 9(11):1869–1877
2. Loia V, Orciuoli F, Pedrycz W (2018) Towards a granular computing approach based on formal concept analysis for discovering periodicities in data. *Knowl-Based Syst* 146:1–11
3. Hao F, Min G, Pei Z et al (2015) k-clique community detection in social networks based on formal concept analysis. *IEEE Syst J* 11(1):250–259
4. Wu W, Leung Y, Mi J (2008) Granular computing and knowledge reduction in formal contexts. *IEEE Trans Knowl Data Eng* 21(10):1461–1474
5. Li L, Mi J, Xie B (2014) Attribute reduction based on maximal rules in decision formal context. *Int J Comput Intell Syst* 7(6):1044–1053
6. Li M, Wang G (2016) Approximate concept construction with three-way decisions and attribute reduction in incomplete contexts. *Knowl-Based Syst* 91:165–178
7. Hao F, Sim DS, Park DS et al (2017) Similarity evaluation between graphs: a formal concept analysis approach. *J Inform Process Syst* 13(5):1158–1167
8. Ibrahim M, Missaoui R, Messaoudi A (2018) Detecting communities in social networks using concept interestingness. In: *Proceedings of the 28th annual international conference on computer science and software engineering*. IBM Corp, pp 81–90