ϕ-Generalized Concept Lattice Models: From Power Concepts to Formal Concepts, and then to Robust Concepts

Zhipeng Xie

Department of Computing and Information Technology
Fudan University, Shanghai, 200433 China
xiezp@fudan.edu.cn

Abstract. Traditionally, Formal Concept Analysis theory defines formal concept via a Galois connection formed by two derivation operators. To extend it, this paper gets started by defining formal concept in another form, which is equivalent to its classical definition. Introducing a non-negative parameter into this new definition gives rise to definition of the so-called "generalized concept" in this paper. Formal concept is a special case of the generalized concept with 1 as the parameter value. When the parameter is set to 0, the generalized concept lattice is isomorphic to the power-set lattice of the attribute set. As the parameter is larger than 1, the generalized concept lattice is a supremum-subsemilattice of the formal concept lattice. A simple algorithm is also developed for building ϕ-generalized concept lattice of $\phi>1$ from the formal concept lattice.

1 Introduction

Formal concepts embody the unification of concept intent and concept extent. The hierarchical order among them is one kind of generalization/specialization relationship. Concept lattice, consisting of the set of all formal concept and the hierarchical order, is the core data structure in Formal Concept Analysis [4][9]. Since its invention, concept lattice has been widely applied to fulfill many different tasks. It can be used to discover association rules [1][8][10] or data dependencies in databases, to index documents for information retrieval [3], to explore attributes in analyzing simple binary data structure or even the complex symbolic ones, and to perform some machine learning tasks such as clustering and classification [6][11].

However, the construction and the storage of formal concept lattices are normally of high computational and space complexity, which is a main obstacle in their practical applications. Various techniques have been proposed to reduce the size of concept lattices by eliminating part of the nodes. For Example, Iceberg concept lattices [7] represented the topmost part of a concept lattice, [2] used a closure system on the attribute set as the user constraint to define non-interesting formal concepts which could be pruned, and in [11] three different constraints were incorporated into the local classifier induction process to prune the formal concept lattice.

The research work done by this paper is to present a generalized concept lattice model. It begins to work on redefining formal concept in another equivalent form, then extends this new definition by injecting a control parameter ϕ into it, resulting in the definition of so-called "ϕ-generalized concept". It has the following properties:

- When ϕ=0, the ϕ-generalized concept lattice is isomorphic to the power set lattice of the attribute set
- When ϕ=1, the ϕ-generalized concept lattice is totally the same as the formal concept lattice
- For any given control parameter ϕ, the ϕ-generalized concept lattice is a complete lattice
- For any two control parameters ϕ_1 and ϕ_2, if $\phi_1>\phi_2$, then the ϕ_1-generalized concept lattice is an supremum-subsemilattice of the ϕ_2-generalized concept lattice

The classical definitions about formal concept lattice and another equivalent definition are given in Section 2. In Section 3, we extend this new definition, present the definitions of ϕ-generalized concept and ϕ-generalized concept lattice, and then study their properties. Section 4 develops a simple algorithm for building ϕ-generalized concept lattice of ϕ>1 by pruning formal concept lattice. Finally, we conclude the whole paper and point out some future work.

2 Another Equivalent Definition of Formal Concept

In this section, what we would like to do is to define formal concept in another new form, which is equivalent to its classical definition defined by Galois connection. This new definition will naturally and easily lead to the notion of generalized concept in this paper. Before that, let us have a review of some classical definition in Formal Concept Analysis (Ganter and Wille 1999).

Definition 1. (Formal Context) A formal context $K = (G, M, I)$ consists of two sets, G and M, and a binary relation I between G and M. The elements of G are called the objects, and the elements of M are called the attributes of the context. In order to express that an object g is in a relation I with an attribute m, we write $(g, m)\in I$ and read it as "the object g has the attribute m".

Definition 2. For a set $A\subseteq G$ of objects we define $A^K := \{ m\in M \mid (g, m)\in I$ for all $g\in A \}$ (the set of attributes common to the objects in A). Correspondingly, for a set B of attributes we define $B^K := \{ g\in G \mid (g, m)\in I$ for all $m\in B \}$ (the set of objects which have all attributes in B).

Definition 3. (Classical Definition of Formal Concept) A formal concept of the context $K=(G, M, I)$ is a pair (A, B) with $A\subseteq G$, $B\subseteq M$, $A^K=B$, and $B^K=A$. We call A the extent and B the intent of the concept (A, B). $\mathfrak{B}(G, M, I)$ denotes the set of all concepts of the context K.

Definition 4. (Hierarchical Order and Formal Concept Lattice) If (A_1, B_1) and (A_2, B_2) are concepts of a context, (A_1, B_1) is called a subconcept of (A_2, B_2), provided that $B_1 \supseteq B_2$. In this case, (A_2, B_2) is a superconcept of (A_1, B_1), and we write $(A_1, B_1) \leq (A_2, B_2)$. The relation $\leq$ is called the hierarchical order of the concepts. The set of all concepts of (G, M, I) ordered in this way is denoted by $\underline{\mathbf{B}}(G, M, I)$ and is called the formal concept lattice of the context (G, M, I).

The hierarchical order is a partial order relation on the set of formal concepts, and the formal concept lattice is actually a partially ordered set. To facilitate the narration, we have to review several notations in partial order. Let (P, $\leq$) be a partially ordered set. For $x \in P$ and $y \in P$, we write $x < y$ for $x \leq y$ and $x \neq y$. Element $x \in P$ is called a **lower cover** of y, if $x < y$ and there is no other element $z \in P$ fulfilling $x < z < y$. To visualize a partially ordered set (P, $\leq$), we normally use small circles to represent the elements in P. If element x is a lower cover of element y, the circle of y is place above that of x, and the two circles are connected by a line segment. Such a diagram is called a line diagram or Hasse diagram.

To better understand the robust concept presented in this paper, let us first define formal concept in another form.

Definition 5. (Another Definition of Formal Concept) A formal concept $c = (A, B)$ of the context $K=(G, M, I)$ is a pair (A, B) that satisfies

(1) $A \subseteq G$, $B \subseteq M$; and
(2) $A = B^K$; and
(3) For any $m \in M-B$, $|B^K| > |(B \cup \{m\})^K|$.

These two definitions (Definition 4 and Definition 5) about formal concept are actually equivalent. We could prove this equivalence by proving the following theorem.

Theorem 6. If a pair (A, B) satisfies $A \subseteq G$, $B \subseteq M$, and $B^K=A$, then $A^K=B \Leftrightarrow \forall m \in M-B$ ($|B^K| > |(B \cup \{m\})^K|$).

Proof.

It is evident that $B^K \supseteq (B \cup \{m\})^K$, thus we have $|B^K| > |(B \cup \{m\})^K| \Leftrightarrow B^K \supset (B \cup \{m\})^K$.

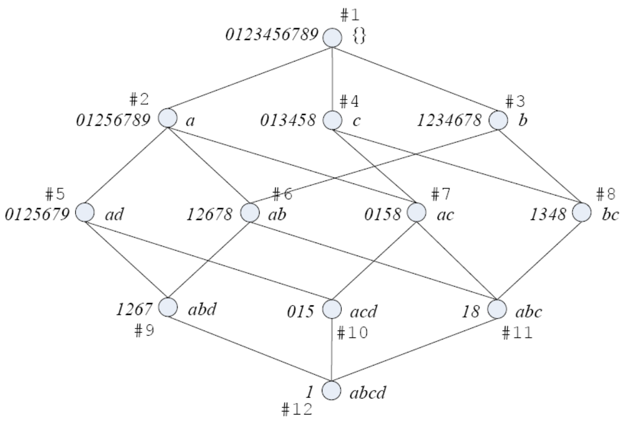
Firstly, we prove $A^K=B \Rightarrow \forall m \in M-B$ ($B^K \supset (B \cup \{m\})^K$). Assume that there exists some m∈M-B such that $B^K = (B \cup \{m\})^K$. Due to the fact that $B^K=A$, it holds that $A=(B \cup \{m\})^K \Rightarrow A^K= (B \cup \{m\})^{KK} \supseteq B \cup \{m\} \supset B$, which contradicts to $A^K=B$. Therefore, the assumption is wrong.

Secondly, let us prove that $\forall m \in M-B$ ($B^K \supset (B \cup \{m\})^K$) $\Rightarrow A^K=B$. Because of $B^K=A$, we have $B^{KK}=A^K \supseteq B$. Assume that $A^K \neq B$, i.e. $A^K \supset B$, which means that there exists some $m \in M-B$ such that $A^K \supseteq B \cup \{m\}$. Then, it follows that $A^{KK}=B^{KKK}=B^K \subseteq (B \cup \{m\})^K$, which contradicts to the condition "$\forall m \in M-B$ ($B^K \supset (B \cup \{m\})^K$)". As a result, the assumption is wrong.

Example 1. To illustrate the idea throughout the whole paper, we will use the following simple formal context $K=(G, M, I)$, where $M=\{a, b, c, d\}$ consists of only 4 attributes, and $G=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ contains 10 objects. This context has 12 formal concepts. The line diagram in **Fig 1** depicts the concept lattice of this context.

Fig. 1. A simple formal context and its corresponding formal concept lattice

	a	*b*	*c*	*d*
0	×		×	×
1	×	×	×	×
2	×	×		×
3		×	×	
4		×	×	
5	×		×	×
6	×	×		×
7	×	×		×
8	×	×	×	
9	×			



3 Generalized Concept and Generalized Concept Lattice

3.1 Generalized Concept

The condition (3) in definition 5 can be rephrased as "For any $m\in M-B$, $|B^K| - |(B\cup\{m\})^K| \geq \phi$ where $\phi=1$". Now, the underlying reason why concept lattice is so large is ready to show up. It rests with the restriction of "$\phi=1$". If this restriction gets relaxed, a more flexible and general definition would be developed as follows.

Definition 7. (Generalized Concept) Given an non-negative integer ϕ, a ϕ-generalized concept of the context $K=(G, M, I)$ is a pair (A, B) that satisfies

(1) $A\subseteq G$, $B\subseteq M$; and
(2) $A = B^K$; and
(3) For any $m\in M-B$, $|B^K| - |(B\cup\{m\})^K| \geq \phi$.

Here, A and B are called the extent and intent of this ϕ-generalized concept, while ϕ is called the control parameter, respectively. $\boldsymbol{GB_\phi}(G, M, I)$ denotes the set of all ϕ-generalized concepts of the context K.

Since the control parameter can take different values, we would like to give alias to some special cases of ϕ-generalized concept.

Definition 8. (Power Concept and Robust Concept) If the given non-negative integer $\phi = 0$, a ϕ-generalized concept is called a power concept. If the given integer $\phi > 1$, a ϕ-generalized concept is called a ϕ-robust concept. In addition, formal concept is just a special case of ϕ-generalized concept with $\phi = 1$.

Due to the fact that $B_1 \subseteq B_2 \Rightarrow B_2^K \subseteq B_1^K$ and $B \subseteq B\cup\{m\}$ for any $m\in M-B$, it holds that $(B\cup\{m\})^K \subseteq B^K$ and thus $|B^K| - |(B\cup\{m\})^K| \geq 0$. Therefore, given that $\phi = 0$, (B^K, B) is a power concept for any subset $B\subseteq M$. As a result, a context (G, M, I) has $2^{|M|}$ different power concepts. In this case, the lattice is isomorphic to the power set lattice of M. A ϕ-generalized concept of $\phi = 0$ is also called a power concept, just because of this fact.

Property 9.

(1) Any ϕ_1-generalized concept is also a ϕ_2-generalized concept if $\phi_1 \geq \phi_2 \geq 0$.

(2) If (A_1, B_1) and (A_2, B_2) are two ϕ-generalized concepts, then we have $B_2\subseteq B_1 \Rightarrow A_1\subseteq A_2$.

(3) If (A_1, B_1) and (A_2, B_2) are two ϕ-generalized concepts, then the pair $((B_1\cap B_2)^K, (B_1\cap B_2))$ is also a ϕ-generalized concept.

Proof.

(1) It is evident, according to the definition of generalized concept.

(2) $B_2\subseteq B_1 \Rightarrow B_2^K\supseteq B_1^K \Rightarrow A_2\supseteq A_1$.

(3) For each attribute $m\in M-(B_1\cap B_2)$, there are totally two possible situations, $m\notin B_1$ or $m\notin B_2$.

If $m\notin B_1$, then we have $|B_1^K| - |(B_1\cup\{m\})^K| > \phi$, and thus

$$|(B_1\cap B_2)^K| - |((B_1\cap B_2)\cup\{m\})^K| \geq |B_1^K| - |(B_1\cup\{m\})^K| > \phi$$ (because of $(B_1\cap B_2) \subseteq B_1$).

If $m\notin B_2$, similar conclusion can be inferred.

From the Property 9 (1), we get to know that the set of all ϕ_1-generalized concepts are a subset of the set of all ϕ_2-generalized concept, given that $\phi_1 > \phi_2$. On the contrary, a ϕ_2-generalized concept (A, B) is not necessarily a ϕ_1-generalized concept. In this case, the concept (A, B) is called a collapsed concept in the transformation from control parameter of ϕ_2 to control parameter of ϕ_1.

Example 2. The context in Example 1 has 16 power concepts, biz. 16 ϕ-generalized concept of $\phi < 0$. Among these 16 power, 4 of them are collapsed in the transformation from $\phi = 0$ to $\phi = 1$. These 4 collapsed power concepts are ({0, 1, 2, 5, 6, 7, 9}, {d}), ({1, 2, 6, 7}, {b, d}), ({0, 1, 5}, {c, d}), and ({1}, {b, c, d}), respectively. All the other 12 power concepts are also formal concept, as illustrated in Example 1.

3.2 Generalized Concept Lattice

Definition 10. (Hierarchical Order) If (A_1, B_1) and (A_2, B_2) are two ϕ-generalized concepts of a given context, (A_1, B_1) is a subconcept of (A_2, B_2), provided that $B_2 \subseteq B_1$. In this case, we write $(A_1, B_1) \leq (A_2, B_2)$. The set of all ϕ-generalized concepts of (G, M, I) ordered in this way is denoted by $\underline{\mathbf{GB}}_\phi(G, M, I)$ and is called the ϕ-generalized concept lattice of the context (G, M, I).

Similarly with the name of power concept and robust concept, a ϕ-generalized concept lattice is called power concept lattice when $\phi = 0$, it is called robust concept lattice when $\phi > 1$. Certainly, it is a formal concept lattice when $\phi = 1$.

Property 11. If $\phi \geq 1$, and (A_1, B_1) and (A_2, B_2) be two ϕ-generalization concepts, then we have $B_2 \subseteq B_1 \Leftrightarrow A_1 \subseteq A_2$.
Proof. It is evident, because (A_1, B_1) and (A_2, B_2) are also formal concepts.

When $\phi = 0$, the concept lattice constructed from (G, M, I) and the one from (M, G, I^{-1}) are isomorphic. However, this does not hold when $\phi \neq 0$.

Property 12.
(1) For any two subsets, $B_1 \subseteq M$ and $B_2 \subseteq M$, of attributes, if $B_1 \subseteq B_2$, then it holds that: $|B_1^K| - |(B_1 \cup \{m\})^K| \geq |B_2^K| - |(B_2 \cup \{m\})^K|$ for any attribute $m \in M$.
(2) For a formal concept $c_1 = (A_1, B_1)$, if $m \in M - B_1$ satisfies $|B_1^K| - |(B_1 \cup \{m\})^K| \leq \phi$, then for each sub-concept $c_2 = (A_2, B_2)$ of c, we have $|B_2^K| - |(B_2 \cup \{m\})^K| \leq \phi$.
Proof. (1) It is evident that $(B_1 \cup \{m\})^K \subseteq B_1^K$ and $(B_2 \cup \{m\})^K \subseteq B_2^K$. $B_1^K \supseteq B_2^K$ can be inferred from $B_1 \subseteq B_2$.
For each object $g \in B_2^K - (B_2 \cup \{m\})^K$, it is clear that $g \in B_2^K \Rightarrow g \in B_1^K$, while $g \notin (B_2 \cup \{m\})^K \Rightarrow g \notin (B_1 \cup \{m\})^K$. Therefore, it holds that $g \in B_1^K - (B_1 \cup \{m\})^K$. $B_2^K - (B_2 \cup \{m\})^K \subseteq B_1^K - (B_1 \cup \{m\})^K$.
(2) It can be easily inferred from Property 12(1).

3.3 The Basic Theorem on Generalized Concept Lattices

For each subset, B, of attributes, function $\tau_\phi: 2^M \rightarrow 2^M$ is defined as

$$\tau_\phi(B) = \{ m \in M : |B^K| - |(B \cup \{m\})^K| < \phi \}.$$

Evidently, if $B_1 \subseteq B_2$, then for each $m \in \tau_\phi(B_1)$, we have $|B_1^K| - |(B_1 \cup \{m\})^K| < \phi \Rightarrow |B_2^K| - |(B_2 \cup \{m\})^K| < \phi$ (according to the Property 12), which means that $\tau_\phi(B_1) \subseteq \tau_\phi(B_2)$. Therefore, the operator τ_ϕ is order-preserving.

In addition, we also define that $\tau_\phi^1(B) = \tau_\phi(B)$ and $\tau_\phi^n(B) = \tau_\phi^{n-1}(\tau_\phi^1(B))$. Thus for any given set, $B \subseteq M$, of attributes, we form the set $\tau_\phi^1(B)$, $\tau_\phi^2(B)$, $\tau_\phi^3(B)$, … until we obtain a set $\tau_\phi{}^*(B) := \tau_\phi^t(B)$ with $\tau_\phi^t(B) = \tau_\phi^{t+1}(B)$. As a special case of $\phi = 1$, $\tau_\phi{}^*(B) = \tau_1{}^1(B) = B^{KK}$ for any subset $B \subseteq M$; while for the special case of $\phi = 0$, $\tau_\phi{}^*(B) = \tau_1{}^1(B) = B$.

Proposition 13. The function τ_ϕ^* is a closure operator on 2^M.

Proof. (1) for any two subsets, B_1 and B_2, of M, if $B_1 \subseteq B_2$, then $\tau_\phi^1(B_1) \subseteq \tau_\phi^1(B_2)$. By natural induction, it is evident that $\tau_\phi^t(B_1) \subseteq \tau_\phi^t(B_2)$ for any positive integer t.

(2) $B \subseteq \tau_\phi^*(B)$ holds evidently.

(3) Because $\tau_\phi(\tau_\phi^*(B)) = \tau_\phi^*(B)$, we have $\tau_\phi^*(\tau_\phi^*(B)) = \tau_\phi^*(B)$.

With the help of the closure operator τ^*, any formal concept (A, B) can be uniquely mapped to a robust concept $\tau_\phi^*((A, B)) = ((\tau_\phi^*(B))^K, \tau_\phi^*(B))$. Furthermore, according to the definitions of function τ_ϕ and function τ_ϕ^*, it holds that $\tau_\phi(M) = \tau_\phi^*(M) = M$, and thus (M^K, M) is smallest ϕ-generalized concept in the ϕ-generalized concept lattices for any given context $K=(G, M, I)$.

Theorem 14.[1] **(The basic theorem on ϕ-generalized concept lattices)** The ϕ-generalized concept lattice is a complete lattice in which infimum and supremum are given by:

$$\bigvee_{t\in T}(A_t, B_t) = \left((\bigcap_{t\in T} B_t)^K, \bigcap_{t\in T} B_t \right)$$

$$\bigwedge_{t\in T}(A_t, B_t) = \left((\tau_\phi * (\bigcup_{t\in T} B_t))^K, \tau_\phi * (\bigcup_{t\in T} B_t) \right)$$

Proof.

(1) Let us first prove the formula for the supremum. For each attribute $m \in M - \bigcap_{t\in T} B_t$, there must be some $i \in T$ such that $m \in M - B_i$. Due to the fact that (A_i, B_i) is a robust concept, it holds that $|B_i^K| - |(B_i \cup \{m\})^K| \geq \phi$. Therefore, because of $\bigcap_{t\in T} B_t \subseteq B_i$, it can be derived that $|\bigcap_{t\in T} B_t{}^K| - |(\bigcap_{t\in T} B_t \cup \{m\})^K| \geq |B_i^K| - |(B_i \cup \{m\})^K| \geq \phi$. It means that the pair $\left((\bigcap_{t\in T} B_t)^K, \bigcap_{t\in T} B_t \right)$ is a robust concept and it is a super concept of the robust concept (A_t, B_t) for each $t \in T$. Furthermore, it is evident that this pair is also the smallest common superconcept, because its intent is exactly the intersection of the extents of these (A_t, B_t).

(2) Next, we shall prove the formula for the infimum. It is evident that pair $\left((\tau * (\bigcup_{t\in T} B_t))^K, \tau * (\bigcup_{t\in T} B_t) \right)$ is a robust concept. Since $\tau * (\bigcup_{t\in T} B_t) \supseteq \bigcup_{t\in T} B_t$, this pair is also a common subconcept of all the robust concepts (A_t, B_t), $t \in T$. Assume that it is not the largest common subconcept, which means that there is another common subconcept (A, B) such that $\bigcup_{t\in T} B_t \subseteq B \subset \tau * (\bigcup_{t\in T} B_t)$.

[1] Strictly speaking, this theorem only gives out the first part of the basic theorem.

From the Theorem 14, we can get the following property about ϕ-generalized concept lattices: "For a given context (G, M, I) and two control parameters ϕ_1 and ϕ_2, if $\phi_1>\phi_2$, then the ϕ_1-generalized concept lattice is a supremum-subsemilattice of the ϕ_2-generalized concept lattice".

4 Building ϕ-Generalized Concept Lattice by Pruning Formal Concept Lattice or Power Concept Lattice

For a given context (G, M, I), power concept lattice is the simplest generalized concept lattice, but it has the largest number ($2^{|M|}$) of generalized concepts. It can be easily constructed according to the fact that it is isomorphic to the power set lattice of the attribute set. As the first step, for all subsets $B\subseteq M$, all the power concepts are generated as (B^K, B). Then, each power concept (B^K, B) is linked to its lower cover $((B\cup\{m\})^K, B\cup\{m\})$ for each attribute $m\in M-B$. The efficient construction of formal concept lattice is much more tricky, but luckily, a lot of algorithms have already been developed in the last two decades, which makes it an easy job to construct formal concept lattice. This section is to show how to get a ϕ-generalized concept lattice from the constructed formal concept lattice or power concept lattice.

Theorem 15. A formal concept (A, B) is a ϕ-generalized concept if and only if it holds that $|A|-|A_l|\geq\phi$ for each lower cover (A_l, B_l) of (A, B).

Proof.

(1) We first prove that a formal concept (A, B) is a ϕ-generalized concept if it holds that $|A|-|A_l|\geq\phi$ for each lower cover (A_l, B_l) of (A, B).

For any attribute $m\in M-B$, there is a corresponding formal concept $(A_1, B_1) = ((B\cup\{m\})^{KK}, (B\cup\{m\})^K)$, which satisfies $(A_1, B_1) < (A, B)$ because of $B_1 = (B\cup\{m\})^K \supseteq B\cup\{m\} \supset B$. Thus, there must be some lower cover (A_2, B_2) of (A, B) such that $(A, B) > (A_2, B_2) \geq (A_1, B_1)$. Therefore, we have $|A|-|A_1| = |B^K|-|(B\cup\{m\})^K| \geq |A|-|A_2| \geq \phi$, which results in that (A, B) is a ϕ-generalized concept.

(2) Next, let us prove that $|A|-|A_l| \geq \phi$ for each lower cover (A_l, B_l) of (A, B) if a formal concept (A, B) is a ϕ-generalized concept.

For each lower cover (A_l, B_l) of (A, B), and for each attribute $m\in B_l-B$, we have $m\in M-B$, and thus $|B^K|-|(B\cup\{m\})^K| \geq \phi$, because ($A$, B) is a ϕ-generalized concept. In addition, we also have $(A_l, B_l) = ((B\cup\{m\})^{KK}, (B\cup\{m\})^K)$ (*if not*, (A_l, B_l) *would not be a lower cover of* (A, B), *which is a contradiction*). It follows directly that $|A|-|A_l| = |B^K|-|(B\cup\{m\})^K| \geq \phi$.

Theorem 15 teaches us how to judge whether a formal concept is a ϕ-generalized concept for a given positive integer ϕ. That is, A formal concept (A, B) is a ϕ-generalized concept if and only if it has at least ϕ more objects in the extent than each of its lower covers in the formal concept lattice. In addition, it is also an easy job to judge whether a power concept is a formal concept according to the generated power concept lattice, which is similar to the method described in Theorem 15 and goes as follows. A power concept (A, B) is a formal concept if and only if it holds that $|A|-|A_l| \geq 1$ for each lower cover (A_l, B_l) of (A, B) in the power concept lattice. Put another way, a power concept (A, B) is a formal concept if and only if all its lower covers in the power concept lattice have larger extent than itself.

Example 3. For the formal context in Fig 1, the corresponding power concept lattice is illustrated in Fig 2. Four power concepts #13, #14, #15, and #16 are collapsed

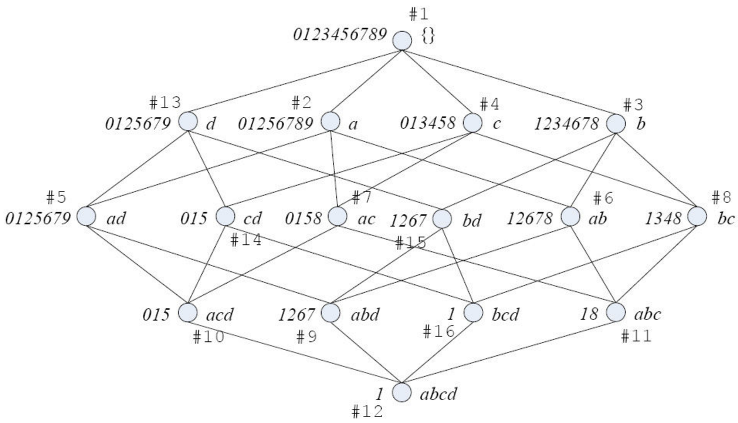


Fig. 2. A simple power concept lattice

when transforming from control parameter of 0 to control parameter of 1. For the power concept #13, it has the same extent {0125679} as its lower cover #5. In addition, #14 has the same extent {015} as its lower cover #10, #15 has the same extent {1267} as its lower cover #9, and #16 has the same extent {1} as its lower cover #12. That is why the formal concept lattice consists of only 12 formal concepts (12=16–4).

Example 4. Let us still take the formal concept lattice in Fig 1 for example. Four formal concepts (#2, #6, #7, and #11) would be collapsed in the transformation from control parameter of 1 to control parameter of 2. Formal concept #2 is collapsed because it has only 1 more object in the extent than its lower cover #5. #6 has only 1 more object in the extent

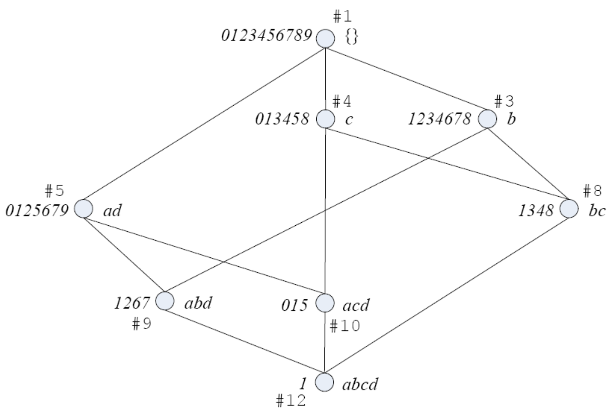


Fig. 3. A simple 2-robust concept lattice

than its lower cover #9, #7 has only 1 more object in the extent than its lower cover #10, and #11 has only 1 more object in the extent than its lower cover #12. Thus, that is why the 2-robust concept lattice (as shown in Fig 3) consists of 8 2-robust concepts.

To build a ϕ-generalized concept lattice with ϕ>1, the simplest way is to first construct a formal concept lattice, and then prune it to get the result by using the following algorithm. This algorithm judges, for each formal concept in ascending order of the cardinality of its extent, whether it should be collapsed. If a formal concept should be collapsed, we shall first record all its upper covers and all its lower covers, then remove this formal concept (delete this concept itself and remove all the links to or from it), and then recalculate the links between the recorded upper covers and the recorded lower covers.

```
Procedure Prune(L, ϕ)
INPUT: a formal concept lattice L and a positive integer ϕ
OUTPUT: the ϕ-generalized concept lattice stored in L
BEGIN
  FOR each formal concept (A, B)∈L in ascending order of |B| DO
    toCollapse:=false;
    FOR each lower cover (A_l, B_l) of (A, B) DO
      IF |A|–|A_l| < ϕ THEN
        toCollapse:=true;
        break;
      ENDIF;
    ENDFOR;
    IF toCollaps==true THEN
      Upperset:= the set of upper covers of (A, B);
      Lowerset:= the set of lower covers of (A, B);
      FOR each formal concept (A_l, B_l) in Lowerset DO
        Remove the link (A, B)→(A_l, B_l);
      ENDFOR;
      FOR each formal concept (A_u, B_u) in Upperset DO
        Remove the link (A_u, B_u)→(A, B);
      ENDFOR;
      delete the formal concept (A, B);
      FOR each formal concept (A_u, B_u) in Upperset DO
        FOR each formal concept (A_l, B_l) in Lowerset DO
          IF there does not exists one lower cover (A_ul, B_ul) of (A_u, B_u)
                        such that (A_ul, B_ul) ≥(A_l, B_l) THEN
            add a link (A_u, B_u)→(A_l, B_l);
          ENDIF;
        ENDFOR;
      ENDFOR;
    ENDIF;
  ENDFOR;
END.
```

A Comment: At last, we would like to give out a simple necessary condition on which a ϕ-generalized concept lattice possibly consists of only one generalized concept, for a given context (*G*, *M*, *I*) that there is no object which has all the attributes: "If a ϕ-generalized concept lattice consists of only one generalized concept, then it must hold that $|G|/|M|\leq\phi$". The proof is simple, which is omitted here.

5 Summary and Future Work

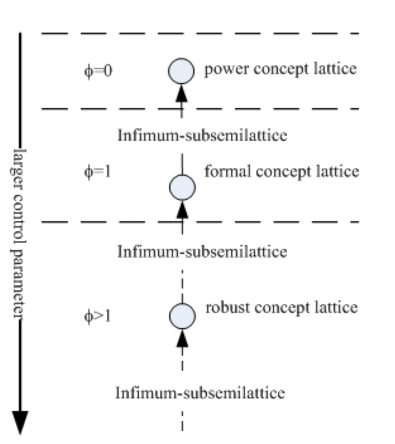


Fig 4. The family of ϕ-generalized concept lattices

Here, we would like to use the diagram in Fig. 4 to summarize the whole paper. The so-called "ϕ-generalized concept" has a control parameter. When the parameter is set to 0, the generalized concept lattice is isomorphic to the power-set lattice of the attribute set. With ϕ=1, the generalized concept lattice is totally the same as the formal concept lattice. In addition, for a given formal context, the ϕ_2-generalized concept lattice is a supremum-subsemilattice of ϕ_1-generalized concept lattice if $\phi_1\leq\phi_2$, which also means that the ϕ-generalized concept lattice with larger ϕ value normally has more generalized concepts than the one with smaller ϕ value. Furthermore, to build ϕ-generalized concept lattice for ϕ>1, this paper gives out a simple algorithm to fulfill it by pruning the formal concept lattice.

There is still much work to be done for this new type of generalized concept lattice. One important task is to design efficient algorithms for building generalized concept lattice directly, instead of the pruning method proposed here. Another is to do some experiments to compare the sizes of ϕ-generalized concept lattices with different ϕ values, on some formal contexts of real domains, which is lack in this paper. Yet another job is to apply this new model to applications to solve some problems such as information retrieval, clustering, and classification, in order to check its applied value in practice.

Acknowledgements: This work was funded in part by National Natural Science Foundation of China under grant number 60503025, and the Science & Technology Commission of Shanghai Municipality under grant number 03ZR14014. Thanks are given to the anonymous reviewers for their helpful comments.

References

[1] Ben Yahia, S., and Mephu Nguifo, E., 2004, Revisiting Generic Bases of Association Rules. DaWaK 2004: 58-67

[2] Belohlavek, R., and Vychodil, V., 2006, Formal concept analysis with constraints by closure operators. In: H. Scharfe, P. Hitzler, and P. Ohrstrom (Eds.): Proc. ICCS 2006, Lecture Notes in Artificial Intelligence 4068, Springer-Verlag, Berlin Heidelberg, pp. 131-143

[3] Carpineto, C., and Romano, G., 1996, A Lattice Conceptual Clustering System and Its Application to Browsing Retrieval. Machine Learning, 24(2): 95-122

[4] Ganter, B., and Wille, R., 1999, Formal Concept Analysis: Mathematical Foundations (Heidelberg: Springer).

[5] Kuznetsov, S. O., and Obiedkov, S. A., 2002, Comparing performance of algorithms for generating concept lattices. Journal of Experimental and Theoretical Artificial Intelligence, 2002, vol. 14(2-3): 189-216

[6] Mephu Nguifo, E., and Njiwoua, P., 2001, IGLUE: a lattice-based constructive induction system. Intelligent Data Analysis, 5(1): 73-91

[7] Stumme, G., Taouil, R., Bastide, Y., Pasquier, N., and Lakhal, L., 2002, Computing iceberg concept lattices with titanic. Data and Knowledge Engineering, 42(2): 189–222

[8] Valtchev, P., Missaoui, R., Godin, R., and Meridji, M., 2002, Generating frequent itemsets incrementally: two novel approaches based on Galois lattice theory, Journal of Experimental and Theoretical Artificial Intelligence, 2002, vol. 14(2-3): 115-142

[9] Wille, R., 1982, Restructuring lattice theory: an approach based on hierarchies of concepts. In I. Rival (ed.) Ordered Sets (Dordrecht & Boston: Reidel), pp. 445-470

[10] Xie, Z., and Liu, Z., 2005, From Intent Reducts for Attribute Implications to Approximate Intent Reducts for Association Rules, in Proceedings of the 5th International Conference on Computer and Information Technology, IEEE Computer Society, September 2005, pp. 162-169

[11] Xie, Z., Hsu, W., Liu, Z., and Lee M.-L., 2002, Concept Lattice based Composite Classifiers for High Predictability, Journal of Experimental and Theoretical Artificial Intelligence, 2002, vol. 14(2-3): pp. 143-156.