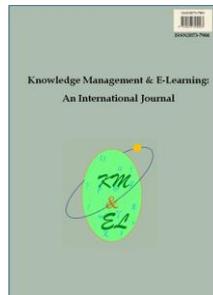


---

**A comprehensive overview on the foundations of formal  
concept analysis**

---

**K. Sumangali  
Ch. Aswani Kumar**  
VIT University, Vellore, India



**Knowledge Management & E-Learning: An International Journal (KM&EL)**  
ISSN 2073-7904

**Recommended citation:**

Sumangali, K., & Kumar, C. A. (2017). A comprehensive overview on the foundations of formal concept analysis. *Knowledge Management & E-Learning*, 9(4), 512–538.

---

## **A comprehensive overview on the foundations of formal concept analysis**

---

**K. Sumangali\***

School of Information Technology and Engineering  
VIT University, Vellore, India  
E-mail: sumivenkata@gmail.com

**Ch. Aswani Kumar**

School of Information Technology and Engineering  
VIT University, Vellore, India  
E-mail: cherukuri@acm.org

\*Corresponding author

**Abstract:** The immersion of voluminous collection of data is inevitable almost everywhere. The invention of mathematical models to analyse the patterns and trends of the data is an emerging necessity to extract and predict useful information in any Knowledge Discovery from Data (KDD) process. The Formal Concept Analysis (FCA) is an efficient mathematical model used in the process of KDD which is specially designed to portray the structure of the data in a context and depict the underlying patterns and hierarchies in it. Due to the huge increase in the application of FCA in various fields, the number of research and review articles on FCA has raised to a large extent. This review differs from the existing ones in presenting the comprehensive survey on the fundamentals of FCA in a compact and crisp manner to benefit the beginners and its focuses on the scalability issues in FCA. Further, we present the generic anatomy of FCA apart from its origin and growth at a primary level.

**Keywords:** Attribute implications; Concept lattice; Data analysis; Formal contexts; Formal concept analysis; Knowledge discovery; Scalability

**Biographical notes:** Ms. Sumangali K. is an Assistant Professor (Selection Grade) of School of Information Technology and Engineering, at VIT University, India. She holds a PhD degree in Information Technology and Engineering from VIT University, India. She also possesses Master degree in Computer Applications from University of Madras, India and another Master degree in computer Science and Engineering from VIT University, India. Her research interests include data mining, formal concept analysis, fuzzy and rough computing. She has contributed few articles in the field of formal concept analysis and knowledge discovery.

Mr. Aswani Kumar Ch. is Professor of School of Information Technology and Engineering, VIT University, Vellore, India. He holds a PhD degree in Computer Science from VIT University, India. He also possesses Bachelor's and Master's Degrees in Computer Science from Nagarjuna University, India. His current research interests are Data Mining, Formal Concept Analysis, Information Security, and Machine Intelligence. Aswani Kumar has published 90 refereed research papers so far in various national, international journals and conferences. He is a senior member of ACM and is associated with other

professional bodies including ISC, CSI, ISTE, IACSIT. He is reviewer for many reputed international journals and conferences. He is editorial board member for several international journals.

---

## 1. Introduction

The developments of information technologies and network have produced huge collection of data every year from different trades. The data flows from various fields such as information technology, agriculture, medicine, finance, markets, social science, demography, etc. This data has no direct information and it is concealed in the data. Extracting the useful information from the huge data is known as knowledge discovery and is an important task in any knowledge based system. According to Han and Kamber (2006), knowledge discovery is to discover the rules and patterns that exist in the data by which one can foretell the trends of the future in the system. So, the invention of methods and means to automatically analyse the patterns and trends of the data is an emerging necessity in order to extract and predict useful information to the society (Malzahn, Ziebarth, & Hoppe, 2013; Mattingly, Rice, & Berge, 2012; Zushi, Miyazaki, & Norizuki, 2012). This is an important issue and apparently has high priority.

To this end, several researchers have proposed various models and techniques (Huang, Yang, Chen, & Wu, 2012). Among such models, mathematical models have contributed enormously to understand the KDD (Knowledge Discovery from Data) precisely. Some of such models are: Set theory, Rough set theory, Fuzzy set theory, Probabilistic set theory, Intuitionistic set theory, Soft set theory, etc. Along these mathematical models falls the lattice theory based notion of Formal Concept Analysis (FCA) (Wille, 1982). FCA concentrates mainly on the clustering of certain objects and attributes which are termed as concepts by which the functionality of cluster analysis from knowledge discovery point of view is carried out. Under the poset relation the concepts can be presented in a form of lattice due to which the functionalities such as presentation and prediction of information can be carried out. The functionality of determining associations can be achieved by finding the implications for the given context using FCA. Thus, FCA based techniques in the practices of knowledge discovery process yield fruitful results to the users.

The extraction of knowledge using FCA from any database is of three dimensions viz., conceptual clusters, lattices (graphical representation) and association rules. Concepts express the underlying relationships between objects and attributes in the context; a concept lattice portrays the context graphically; and the association rules discover the underlying associations within the attributes of the context.

Although FCA is an important formalism for knowledge representation, extraction and analysis, one of the major issues in FCA is the issue of scalability arising due to the size of the contexts which yield bigger concept lattices. As the size of the concept lattice increases, the visualisation of concepts along with hierarchy becomes complex and impractical. This complexity issue arises due to the scalability of FCA and its extensions in various environments. According to Poelmans, Ignatov, Kuznetsov, and Dedene (2013) the scalability issue is focused in 9% of the articles on FCA. We also point out the scalability issue in FCA and review it.

In view of the growing and applicative nature of FCA in various fields, we present its fundamental notions with examples in this article as follows. In section 2, the

origin and growth of FCA is discussed. The terms and notions related to FCA are presented and illustrated in section 3. Section 4 deals with the scalability issues and the current trends on it. At last, we conclude the article in section 5.

## 2. Origin and growth of FCA

The lattice theory based framework namely Formal concept analysis (FCA) has emerged as a distinctive tool in the field of knowledge discovery. FCA has found its immense growth since its inception in to the field of data analysis and knowledge representation few decades ago. FCA is a theory of mathematics means for determining the concepts and their hierarchies that underlie in any information system (Wille, 1982). The mathematical foundations of FCA were first laid by (Birkhoff, 1948) who first bridged the partial orders and lattices. Further he also proved that any binary relation between a set of objects and a set of attributes can be depicted by means of a unique lattice which provides an insight into the structure of the original relation. FCA has emerged as a result of the attempts of a group of researchers to develop the applications of lattice theory at Darmstadt University of Technology in Germany. The research group was led by the Professor Rudolf Wille, who became the founder of FCA by publishing his first article on FCA in 1982 (Wille, 1982) in which he discussed about the approach of restructuring lattices using hierarchies of concepts.

FCA is a mathematical framework of data analysis which discovers conceptual structures among data structures. The main use of FCA is to analyse the data. In other words, FCA is used to investigate and process any data explicitly to obtain meaningful information. FCA has found several applications since its inception. Details of these applications are available in the literature (Carpineto & Romano, 2004; Priss, 2006). The major areas of its applications are (Wille, 1992; Ganter & Wille, 1999): Knowledge Extraction, Information Restructuring and Classification, Data Mining (Kumar & Sumangali, 2012; Sumangali & Kumar, 2013), Data Visualization, Decision Making. FCA is also used in numerous fields involving the above tasks. Some of the fields wherein FCA has been used are: psychology (Spangenberg, 1999; Spangenberg & Wolff, 1991; Duquenne, 1999; Duquenne, Chabert, Cherfouh, Delabar, Doyen, & Pickering, 2003), biological and social sciences (Duquenne, 1999; Ganter & Wille, 1989), civil engineering (Kollewe, Skorsky, Vogt, & Wille, 1994), experimental design (Duquenne, 1999), information retrieval (Godin, Missaoui, & April, 1993; Carpineto & Romano, 2000; Cole, 2000; Rock & Wille, 2000), memory modeling (Kumar, Ishwarya, & Loo, 2015) and software engineering (Codocedo, Taramasco, & Astudillo, 2011). The researchers have enriched the notions of descriptive logics and FCA to understand the images (Atif, Hudelot, & Bloch, 2014). The human expert knowledge is linked through different combinations such as descriptive logics, FCA, and mathematical morphology. Poelmans, Ignatov, Kuznetsov, and Dedene (2013; 2014) have presented a different strategy on the applications of FCA.

The mathematical theory of FCA has been extended into various frontiers and included with other knowledge representation schemes. Recently Singh, Kumar, and Gani (2016) provided necessary mathematical background for few extensions of FCA for various environments such as FCA with granular computing (rough set theory), fuzzy set theory, interval-valued set theory, possibility theory, triadic concepts, factor concepts and handling incomplete data. Yao (2016) interpreted the notion of rough set (RS) definable concepts and thereafter derived the Boolean algebra for RS- definable concepts. RS-definable concept is a pair of extension and intension, with extension being a set of objects and the intension being a family of sets of attribute-value pairs.

### 3. Terms and notations in FCA

FCA is an art of describing the world in terms of the objects and attributes possessed by those objects. In FCA the adjective ‘*formal*’ is often used to emphasise the mathematisation of the notions used from those of the human mind. The terms and notions used in this article are on the basis of the text book of Ganter and Wille (1999) and also consistent with the notions dealt by Davey and Priestly (2002).

FCA is the theory of formalisation of the idea *concept*. The notion of *concept* has been already suggested from ancient times by the eminent philosophers such as Plato, Francis Bacon and John Stuart Mill in order to characterize formal logic systems. The notion of a concept from a context was first studied in (Arnauld & Nicole, 1981) and the term has been recognized in the German standard (DIN 2330, 1993). The philosophical thought of the notion ‘*concept*’ can be described by its *extension* – that is the set of all objects belonging to the concept and its *intension* – the set of all attributes possessed by those objects in common.

For example, consider the object-attribute relation: ‘All living beings need water to live.’ This relation obviously forms a *concept* since it has an extent and the corresponding *intent*. The extent is the set of objects of all living beings including mankind, animals, birds, etc., and the intent is the attribute ‘water’. The relationship covering the set of objects and a set of attributes is often represented by means of a formal context which is formally defined as below.

#### 3.1. Formal context

A *formal context* is a triplet  $K:=(G, M, I)$  where  $G$  denotes a set of formal objects,  $M$  denotes a set of attributes and  $I \subseteq G \times M$  is the incidence relation between the objects  $G$  and attributes  $M$ . The symbols  $G$  and  $M$  stand for the German words *Gegenstande* (Objects) and *Merkmale* (Attributes) respectively. For any two elements  $g \in G$  and  $m \in M$ , the binary relation  $(g, m) \in I$  has to be read as “Object  $g$  has the attribute  $m$ ” and is usually written as  $gIm$ .

A *formal context* is often represented using a cross-table in which rows correspond to the set of objects  $G$  while the columns correspond to the attributes  $M$ . The presence/absence of the incidence relation between  $G$  and  $M$  are denoted by the presence/absence of crosses. Such contexts with *yes* or *no* attribute values are known as binary contexts or one-valued contexts (possible number of attribute values in case of its presence). In order to explain the further notions of FCA, its convenient for us to consider a small *formal context*. To this end, we look at the context of Wolff (1994) on animals and their characteristics shown in Table 1.

In this example, the object set  $G$  consists of the animals– Lion, Finch, Eagle, Hare and Ostrich while the attribute set  $M$  includes the characteristics– preying, flying, bird and mammal. The symbol  $\times$  at the intersection of an object row and attribute column points out that the object possesses that attribute. For example, the animal *Lion* has the attributes *preying* and *mammal* in the given context  $K$ .

The terms *context*, *formal context* and *cross-table* are synonymous and henceforth are used interchangeably throughout the article. Furthermore, since we only deal with the mathematisation of the notions context and concept throughout the article, we don’t emphasise the use of the prefix ‘*formal*’ with them. Before formally defining a formal

concept in a context, we first need to know about the *concept-forming operators* or  $\uparrow$  (up),  $\downarrow$  (down)/*up-down operators* on any given context  $K$ .

**Table 1**  
Formal context of animals

Animals	Preying	Flying	Bird	Mammal
Lion	X			X
Finch		X	X	
Eagle	X	X	X	
Hare				X
Ostrich			X	

For any subset of objects  $A \subseteq G$ , the up operator  $A^\uparrow$  is the set of common attributes of  $A$ . Symbolically,  $A^\uparrow := \{m \in M \mid (g, m) \in I, \forall g \in A\}$ . Similarly, for any subset of attributes  $B \subseteq M$ , the down operator  $B^\downarrow$  is the set of common objects belonging to  $B$ . Symbolically,  $B^\downarrow := \{g \in G \mid (g, m) \in I, \forall m \in B\}$ .

For example, in the given cross-table shown in Table 1:

- $\{Lion\}^\uparrow := \{preying, mammal\}$
- $\{Finch, Eagle\}^\uparrow := \{flying, bird\}$
- $\{mammal\}^\downarrow := \{Lion, Hare\}$
- $\{mammal, preying\}^\downarrow := \{Lion\}$
- $\{bird, mammal\}^\downarrow := \{\}$

Clearly, every context induces the concept-forming operators. We notice that the operator  $\uparrow$  assigns the subsets of  $G$  to subsets of  $M$  and dually the operator  $\downarrow$  assigns the subsets of  $M$  to subsets of  $G$ . For brevity, the concept-forming (up-down) operators  $A^\uparrow$  and  $B^\downarrow$  are denoted as  $A'$  and  $B'$ .

### 3.2. Formal concepts

We next define the notion of a formal concept (Ganter & Wille, 1999). The formal concepts are the clusters of the given context formed as a result of attribute sharing. More formally, For any given context  $K := (G, M, I)$ , a formal concept is a duple  $(A, B)$  where  $A \subseteq G$  and  $B \subseteq M$  such that  $A' = B$  and  $B' = A$ . Plainly, in a given context, if  $A$  is a maximal set of objects sharing a maximal set of attributes  $B$  then the ordered duple  $(A, B)$  is called as a formal concept. The sets  $A$  and  $B$  are respectively known as the extent and the intent of the concept  $(A, B)$ . The following proposition on extent and intent follows directly.

*Proposition 1* (Ganter & Wille, 1999):

Let  $K := (G, M, I)$  be a context. For any subsets  $A \subseteq G$  and  $B \subseteq M$  the following are valid:  $A \subseteq A''$  and  $B \subseteq B''$ . Consequently,  $(A'', A')$  and  $(B', B'')$  are valid concepts of the context  $K$ . Further,  $A$  is an extent if and only if  $A = A''$  and dually  $B$  is an intent if and only if  $B = B''$ . Thus, combining the definition and properties of a concept we can tell that for any concept  $(A, B)$ ,  $(A, B) = (A'', A') = (B', B'')$

### 3.2.1. Properties of concepts

It is noteworthy to mention that a set  $S$  is said to be maximal/minimal set with property  $P$  if there exists no other proper superset/subset of  $S$  with property  $P$ . Furthermore,  $S$  is said to be a maximum/minimum set if its cardinality is maximum/minimum among such sets.

A rectangle in a context  $K := (G, M, I)$  is a duple  $(A, B)$  such that the cartesian product  $A \times B \subseteq I$ . i.e., for every  $x \in A$  and  $y \in B$ ,  $(x, y) \in I$ . For any two rectangles  $A_1 \times B_1$  and  $A_2 \times B_2$  we call  $A_1 \times B_1 \subseteq A_2 \times B_2$  if and only if  $A_1 \subseteq A_2$  and  $B_1 \subseteq B_2$ . Any formal concept  $(A, B)$  can also be viewed as a maximal rectangle in the context. The formal concepts remain invariant under the row or column permutations of the cross-table.

### 3.2.2. Computation of concepts

The formal concepts can be easily computed for any given context  $K := (G, M, I)$ . Though there may be several techniques to compute the formal concepts, the easiest way is to start with an object  $g \in G$  and determine its attribute-set  $B \subseteq M$ , the intent of the concept. Next, determine the set of all objects  $A \subseteq G$  which possess all the attributes in  $B$  (intents) which form the extent of the concept. Thus, the ordered pair  $(A, B)$  is the required concept. Dual approach of starting with any attribute  $m \in M$  can also be adopted in the determination process of concepts. More generally, for any subset of objects or attributes of a context the corresponding concept can be determined.

A concept  $(\{g\}'', \{g\}')$  obtained by starting with an object  $g \in G$  is called as an object concept denoted by  $\gamma(g)$ . Dually, a concept  $(\{m\}', \{m\}'')$  obtained with the start of an attribute  $m \in M$  is called as an attribute concept denoted by  $\mu(m)$ . Clearly, not all concepts of a context are object or attribute concepts. Any concept may be either object concept or attribute concept or both or neither.

We shall illustrate the above concept determination process through the context given in cross-table (Wolff, 1994). Let us start with the object Lion, its intent set is  $B = \{\text{preying, mammal}\}$ . The extent set corresponding to  $B$  is  $A = \{\text{Lion}\}$  only. So, the pair  $(\{\text{Lion}\}, \{\text{preying, mammal}\})$  is a concept for the given context. If we start with the objects Finch, Ostrich we obtain the intent set  $B = \{\text{bird}\}$  whose extent set is  $A = \{\text{Finch, Eagle, Ostrich}\}$ . So, the ordered pair  $(\{\text{Finch, Eagle, Ostrich}\}, \{\text{bird}\})$  is another concept in the given context. Exploration of the given context further yields the following 8 concepts.

1.  $(\{\}, \{\text{preying, flying, bird, mammal}\})$
2.  $(\{\text{Lion}\}, \{\text{preying, mammal}\})$
3.  $(\{\text{Eagle}\}, \{\text{preying, flying, bird}\})$

4. ( $\{Lion, Hare\}, \{mammal\}$ )
5. ( $\{Lion, Eagle\}, \{preying\}$ )
6. ( $\{Finch, Eagle\}, \{flying, bird\}$ )
7. ( $\{Finch, Eagle, Ostrich\}, \{bird\}$ )
8. ( $\{Lion, Finch, Eagle, Hare, Ostrich\}, \{ \}$ )

In this example, concepts 2, 3, 4, 6, 7 are object concepts and concepts 4, 5, 6, 7 are attribute concepts. One can easily note that only certain and not all of the subsets of objects have formed as extents of some concepts and the case of intents also similar, though up-down operation exists for any such subsets. For any given subset of objects/attributes the resulting concept is always unique. Moreover, if the extent  $A$  of any concept  $(A, B)$  is known then its intent  $B$  can be uniquely determined and vice versa.

There are several algorithms to generate the formal concepts in a context. Some of the famous algorithms serving this purpose are: Ganter, Bordat, Next neighbours, etc. Kumar and Singh (2014) have studied the performance of various concept generation algorithms.

### 3.2.3. *Hierarchy of concepts*

In order to discuss about the properties of the set of all concepts we require to know the fundamental notions associated with the lattices from set theory, a branch of mathematics. We refer the readers to Davey and Priestly's *Lattices and Order* (Davey & Priestley, 2002) for an introductory knowledge on lattices and to George Grätzer's *General Lattice Theory* (Grätzer, 2003) for an encyclopaedic knowledge on lattices. In order to make the article self-content we recall some of the basics of lattice theory.

Let  $P$  be any set in which any two elements  $x, y$  are related using some relation  $R$  denoted as  $xRy$ . Then  $P$  is said to be a partially ordered set or simply a poset if the following properties hold:

- i. Reflexivity : For any element  $x \in P$ ,  $xRx$ .
- ii. Anti-symmetry : If  $x, y \in P$  such that  $xRy$  and  $yRx$ , then  $x = y$ .
- iii. Transitivity : If  $x, y, z \in P$  such that  $xRy$  and  $yRz$ , then  $xRz$ .

The relation  $R$  by which a set  $P$  is a partially ordered set resembles the usual relation of  $\leq$  (less than or equal to) in view of the above stated three properties. Hence, conventionally, the symbol  $R$  is replaced by the symbol  $\leq$ . A set  $P$  with a partial order  $\leq$  is denoted by  $(P, \leq)$ .

Let  $(P, \leq)$  be a partially ordered set and let  $S$  be its subset ( $S \subseteq P$ ). An upper bound of  $S$  is an element  $x \in P$  such that  $s \leq x$  for all  $s \in S$ . Dually, a lower bound of  $S$  is an element  $y \in P$  such that  $y \leq s$  for all  $s \in S$ . A smallest element amongst the set of all upper bounds of  $S$  is called the supremum or least upper bound of  $S$  and is denoted by  $\vee S$ . Dually, the greatest element amongst the lower bounds of  $S$  is called the infimum or greatest lower bound of  $S$  and is denoted by  $\wedge S$ . If  $S = \{x, y\}$ , we write simply  $x \vee y$  instead of  $\vee S$  and  $x \wedge y$  instead of  $\wedge S$ . The terms supremum and infimum are also referred to as join and meet respectively.

Consider a partially ordered set  $(P, \leq)$ .

- If for any two elements  $x, y \in P$ ,  $x \vee y$  and  $x \wedge y$  exist, then  $(P, \leq)$  is called a lattice.
- If for any subset  $S \subseteq P$ ,  $\vee S$  and  $\wedge S$  exist, then  $(P, \leq)$  is called a complete lattice.

Turning back to the discussion of concepts of a context, any two concepts  $(A_1, B_1)$  and  $(A_2, B_2)$  of a context can be ordered/related by means of the subconcept-superconcept ordering relation  $\leq$ , which is defined as follows:

$(A_1, B_1) \leq (A_2, B_2)$  if and only if  $A_1 \subseteq A_2$  which otherwise also means that  $B_2 \subseteq B_1$ .

The ordering relation  $\leq$  between concepts can be identified to be a partial order. In other words, the relation  $\leq$  satisfies the three properties of set theory viz., reflexivity, anti-symmetry and transitivity. The partial order  $\leq$  between the elements of a poset is also known as the hierarchical order or lexicographical order. The definition of the object and attribute concepts achieves the following straight forward result.

*Proposition 2* (Davey & Priestley, 2002; Lambrechts, 2012):

If  $(A, B)$  is a concept of a context  $K := (G, M, I)$ , then  $g \in A$  if and only if  $\gamma(g) \leq (A, B)$  and dually,  $m \in B$  if and only if  $\mu(m) \geq (A, B)$ . Furthermore,  $gIm$  if and only if  $\gamma(g) \leq \mu(m)$ .

### 3.3. Concept lattices

If  $(A_1, B_1) \leq (A_2, B_2)$ , the concept  $(A_1, B_1)$  is called as subconcept/child/successor concept and the concept  $(A_2, B_2)$  is called as super/parent/predecessor concept. Moreover, the concept  $(A_1, B_1)$  is said to be more specific than the concept  $(A_2, B_2)$  or equivalently the concept  $(A_2, B_2)$  is more general than the concept  $(A_1, B_1)$ .

Let  $\mathbf{B}(K)$  be the set of all concepts of a context  $K := (G, M, I)$ . Then,  $(\mathbf{B}(K), \leq)$  is a partially ordered set. Moreover, for any subset of concepts in  $\mathbf{B}(K)$ , there always exists supremum as well as infimum and hence the poset  $(\mathbf{B}(K), \leq)$  forms a complete lattice. The complete lattice  $(\mathbf{B}(K), \leq)$  is often known as a concept lattice for obvious reason.

The symbol  $\mathbf{B}$  for concept lattices is attributed to the mathematician Birkhoff who initiated the theory of formal concept by proving the existence of lattices for binary relations of any context in his *Lattice Theory* (Birkhoff, 1948). A detailed study about concept lattices and their theoretical aspects can be found from (Sarmah, Hazarika, & Sinha, 2015).

The one of the main reasons for considering FCA as a powerful method in the analysis of data is due to the fact that it has the added feature of graphical visualisation of the context which explores the underlying implicit relationships in the given context.

#### 3.3.1. Graphical representation of concept lattices

Any lattice can be graphically viewed using Hasse (line) diagrams (Davey & Priestley, 2002) and so also the concept lattices. The Hasse diagram of a lattice can be easily drawn as follows.

Represent the elements of a lattice  $(P, \leq)$  by means of nodes/circles. Let  $x, y \in P$  be any two elements. Then join the nodes corresponding to the elements  $x, y \in P$  if and only if  $x \leq y$  and there exists no other element  $z \in P$  such that  $x \leq z \leq y$ . Or simply, if  $x, y \in P$  are the immediate predecessor and successor (sub-concept and super-concept) respectively, then join their nodes by a line. Another convention adopted in the drawing of Hasse diagrams is that if  $x, y \in P$  such that  $x \leq y$ , then the node corresponding to  $x$  is placed below that of  $y$ . It is interesting to note that the Hasse diagrams of a lattice need not be unique in the sense that there can be different drawings for the same lattice, since nodes can be placed as desired. However, any two Hasse diagrams of a lattice are always isomorphic graphs. Isomorphic graphs are the different drawings of the same graph. Having understood the Hasse diagrams of lattices, let us now illustrate the graphical representation of the concept lattices.

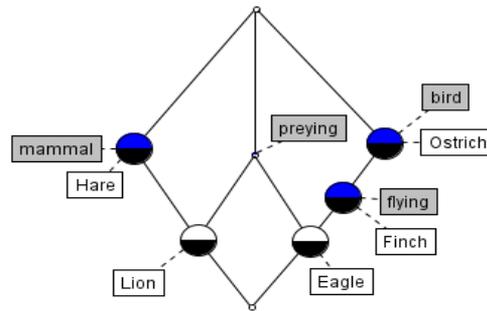
Hasse diagrams endow us to view every concept lattice graphically easily than any other representation scheme. The only part that remains for us to know is the labelling of the concepts in a concept lattice. Obviously, every concept in a concept lattice is attributed to a node/circle in a Hasse diagram. Labelling each concept over the nodes would be overkill. As an alternative method, 'Reduced Labelling' scheme is available to this end by which the concepts of a concept lattice are labelled as follows:

- A node corresponding to an object concept  $\gamma(g)$  is labelled by the object  $g \in G$ .
- A node corresponding to an attribute concept  $\mu(m)$  is labelled by the attribute  $m \in M$ .
- Object labels are written below the nodes while attribute labels above the nodes.

The remaining concepts can be retrieved using the proposition 2 stated earlier by understanding their extents and intents properly. In a concept lattice, one can determine the extent of a concept node by collecting all the object labels of the nodes that can be reached starting from the corresponding concept node by descending/downward path including the object label of the starting node if it has one. Similarly, starting from a concept node, the collection of attribute labels of the nodes which can be reached by ascending/upward path including the attribute label of the starting node if it is an attribute concept node yields the intent of the starting concept node.

Since any lattice diagram is always Hasse diagram only, we need not emphasise the term 'Hasse' and henceforth we omit it from the discussion. As we move through the nodes of a concept lattice from the bottom/top to the top/bottom, we find that the object set increasing/decreasing and attribute set decreasing/increasing respectively. Thus, the predecessor concepts inherit the objects from their successors while the successors inherit the attributes from their predecessors. Briefly, as we traverse from bottom to top we achieve more general concepts and the reverse traversal achieves more specific concepts. The top most concept which consists of all the objects is called as the unit concept and dually the bottom most concept consisting of all attributes is called as the empty concept. The set of concepts lying on the downward path is known as the down-set or order ideal and dually that on the upward path is known as the up-set or order filter. The concept lattice reflects the relationship of generalization and specialization among concepts. It thereby is more intuitional and effective for knowledge representation and knowledge discovery.

Using the principles of the partial order of concepts and Hasse diagram we are now able to draw the concept lattice of the context given in Table 1 as shown in Fig. 1.



**Fig. 1.** Concept lattice for the formal context of Table 1

The given context Table 1 as explained earlier has eight concepts; each of them is represented by means of a node in the concept lattice. Any two immediate predecessors / successors are joined directly which on the whole yields the concept lattice as desired. The concepts corresponding to the nodes can be identified as interpreted earlier the example the node with object label FINCH corresponds to the 6<sup>th</sup> concept ({Finch, Eagle}, {flying, bird}). [Recall that extent is the collection of objects from downward paths while the intent is that of attributes from upward paths].

### 3.4. Many valued contexts and their scaling processes

Having understood the fundamental notions of FCA, we will now try to explain the FCA structures for varieties of information contexts. In general, FCA is not compatible with all types of information contexts. In such circumstances, the information context is modified using appropriate principles so that it becomes compatible to be processed by FCA.

Usually, attributes are considered to be one-valued viz., ‘yes’. But in several contexts, attributes are identified with many values. For example, the attributes such as weight, colour, grade, etc., may be characterized as low, medium, high, etc. Such contexts are known as many-valued contexts. In such cases, the usual context representation scheme is not suitable for the analysis using FCA and the context is modified to a one-valued context using methods of ‘conceptual scaling’ (Ganter & Wille, 1989; Davey & Priestley, 2002). The modified one-valued context is known as the *derived context*.

In the process of scaling, a many-valued context is first transformed into a one-valued context or binary context using *conceptual scaling* techniques. However, this transformation process is accomplished by the users. Hence, the conceptual scaling of a many-valued (MV) context is not determined uniquely.

Literature lists several research articles centered on MV contexts. Messai, Devignes, Napoli, and Smaïl-Tabbone (2008) have studied MV contexts for the first time. They observed that MV contexts yield multi-level concept lattices of have higher precision levels. In the retrieval process of valid information from complex queries, use of MV context methods brings out fruitful results. Before we proceed, let us formally define a MV context.

A many-valued (MV) context  $(G, M, W, I)$  comprises of sets of objects  $G$ , attributes  $M$ , attribute values  $W$  together with a ternary relation  $I$  between  $G$  and  $M, W$ .

Stated otherwise,  $I \subseteq G \times M \times W$  such that  $(g, m, w) \in I$  and  $(g, m, v) \in I$  imply  $w = v$ . The notation  $(g, m, w) \in I$ , means that ‘for the object  $g$ , the attribute  $m$  possesses the value  $w$ ’. If  $W$  contains  $n$  elements, then the quadruple  $(G, M, W, I)$  is called an  $n$ -valued context. Every MV attribute is a partial map  $m: G \rightarrow W$  such that  $m(g) = w$ . For any attribute  $m$ , its domain is defined as  $dom(m) = \{g \in G \mid m(g) = w \text{ for some } w \in W\}$ . If  $dom(m) = G$ , then the attribute  $m$  is said to be complete.

Concept lattices cannot be determined instantly for many valued contexts. In this case, one has to convert it into a binary valued context which is termed as *conceptual scaling* according to (Ganter & Wille, 1989). Such a modified context is known as the derived context. Normally, a conceptual scale is employed on a single attribute  $m$ , and in this case the scale forms a basis for the formal context. The standard scaling method namely *plain scaling* creates from a scaled MV context  $((G, M, W, I), (S_m \mid m \in M))$ , which is an ordered pair that consists a many valued context  $(G, M, W, I)$ , and a set of scales  $(S_m \mid m \in M)$ , the derived context namely  $K := (G, \{(m, n) \mid m \in M, n \in M_m\}, J)$ , where  $gJ(m, n)$  iff  $m(g)I_m n$  where  $g \in G, m \in M, n \in M_m$ .

We will now require an example of a MV context to interpret the forthcoming notions clearly. Let us consider a simple context of platonian bodies given by (Hitzler & Scharfe, 2016) shown in Table 2.

**Table 2**  
Many-valued context of platonian bodies

	Corners	Edges	Facets
Tetrahedron	4	6	4
Hexahedron	8	12	6
Octahedron	6	12	8
Dodecahedron	20	30	12
Isocahedron	12	30	20

### 3.4.1. Conceptual scaling

We will now discuss about the process of conceptual scaling. Every attribute of a MV context is first interpreted using a context. This context is known as conceptual scale. Theoretically, a scale for an attribute can be defined as follows.

The ‘scale’ of an attribute  $m$  in a many-valued context is a one-valued context  $S_m := (G_m, M_m, I_m)$  where  $G_m \subseteq G$ . The objects and attributes in a scale are respectively known as scale values and scale attributes. A scale for an attribute is a context which serves in the process of transformation of a many-valued context into a binary context. For example, in the given example of platonian bodies we can classify the attribute facets into simple, medium and complex using the following scale shown in Table 3.

**Table 3**  
One-valued context for attribute facets

$S_{face} :=$	Simple	Medium	Complex
$\leq 10$ facets	X		
$> 10$ facets and $\leq 15$ facets		X	
$> 15$ facets			X

Conceptual scales interpret the columns of a MV context. Conventionally, the contexts which are binary and are clear in structure are called as scales, even though every context can be regarded as a scale. The simplest of all conceptual scales are the *nominal scales* in which every attribute is subdivided by each of its values. Using the nominal scale in the context given in Table 2, the attributes corners, edges and facets are respectively subdivided into 5, 3 and 5 columns in the derived context. The derived context out of the nominal scale is shown in Table 4 which is followed by its concept lattice Fig. 2.

**Table 4**  
Formal context derived from Table 2 by nominal scaling

	Corners					Edges			Facets				
	$\leq 4$	$\leq 6$	$\leq 8$	$\leq 12$	$\leq 20$	$\leq 6$	$\leq 12$	$\leq 30$	$\leq 4$	$\leq 6$	$\leq 8$	$\leq 12$	$\leq 20$
Tetrahedron	X					X			X				
Hexahedron			X				X			X			
Octahedron		X					X				X		
Dodecahedron					X			X				X	
Isocahedron				X				X					X

The other class of conceptual scales is the ordinal scales and its variety is several. To mention few, we will glance at some basic ordinal scales viz., one-dimensional ordinal scale, inter ordinal scale, biordinal scale and dichotomic scale.

In a one-dimensional ordinal scale, the attribute values of every attribute are ordered such that some attribute values subsume other values because the former attribute values are greater or lesser than those of later ones. As a result, the extents form a chain of hierarchy. For example, the attribute values may be arranged in the order {good, better, best}. The following Table 5 is an example of one-dimensional ordinal scaling in the example context under consideration in Table 2 which is followed by its concept lattice in Fig. 3.

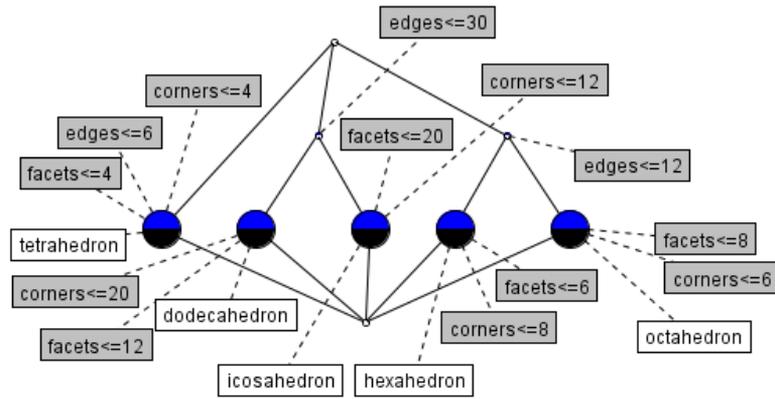


Fig. 2. Concept lattice for the formal context of Table 4

Table 5  
Formal context derived from Table 2 by ordinal scaling

	Corners					Edges			Facets				
	<=4	<=6	<=8	<=12	<=20	<=6	<=12	<=30	<=4	<=6	<=8	<=12	<=20
Tetrahedron	X	X	X	X	X	X	X	X	X	X	X	X	X
Hexahedron			X	X	X		X	X		X	X	X	X
Octahedron		X	X	X	X		X	X			X	X	X
Dodecahedron					X			X				X	X
Isocahedron				X	X			X					X

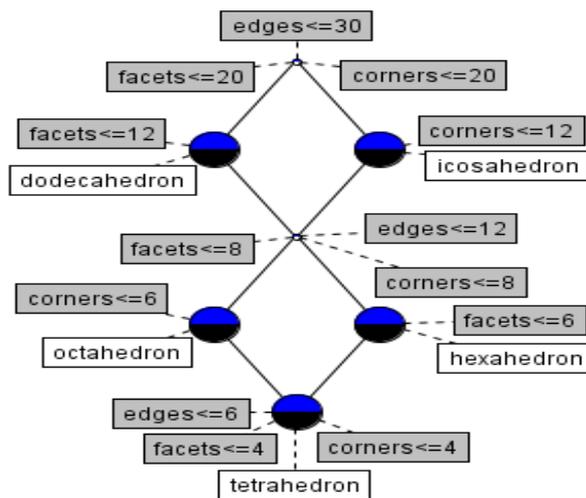


Fig. 3. Concept lattice for the formal context of Table 5

'Interordinal scales' are used in the representation of contexts of mixed attribute values. For example, the contexts such as the answers of a questionnaire contain bipolar attributes which are mixed, and it can be efficiently scaled using interordinal scales. For instance the attribute values  $\{\leq 1, \leq 2, \leq 3, \geq 1, \geq 2, \geq 3\}$  yield extents which fall on the attribute interval values. Another example for application of biordinal scales can be in a marking scheme having values  $\{poor, middle\ class, rich, very\ rich\}$ , in which the attribute 'rich' can belong to both attributes 'middle class' and 'very rich'.

The 'dichotomic scale' context of binary attributes contain the values of the kind {yes, no} shown in the following Table 6.

**Table 6**  
Dichotomic scale context

	0	1
0	X	
1		X

Readers may refer to (Ganter & Wille, 1999; Hitzler & Scharfe, 2016; Carpineto & Romano, 2004) to know the various scaling techniques in detail.

Having understood various real-life contexts and their scales, one may now be able to construct concept lattices for any given context. Apart from the benefit of understanding the contexts by concepts and their graphical view of line diagrams FCA also empowers the users to explore the hidden rule patterns present in the formal context and we present some fundamental aspects of the same subsequently.

### 3.5. Attribute implications

The quest of understanding dependencies between attributes leads to the study of attribute exploration in contexts. The attribute logic is the underlying rules between the sets of attributes in a context. Attribute implications portray the data dependencies. For example, the following are some attribute implications.

- Every number divisible by 2 and 5 is also divisible by 10.
- Every patient with symptoms head ache and fever also gets vomiting symptom.

From the attribute hierarchy of concept lattices, we infer that in any intent, the attributes always occur along with those above them. This mathematical property of lattices paves the way to another broad area of knowledge discovery in FCA viz., 'Attribute Exploration'. Let us explore some of its associated basics.

According to mathematical logic, an implication  $X \rightarrow Y$  is a logical statement that relates a set of formulas  $X$  with another set of formulas  $Y$  such that  $Y$  is a logical consequence of  $X$ . The implication  $X \rightarrow Y$  literally means, 'if  $X$  then  $Y$ ' and hence can be thought of as 'if-then' dependencies between attributes. In this angle, the implication formulas are also known as 'functional / attribute dependency (AD) formulas' or association rules. In view this, the study of rules or implications viz., Attribute Exploration in FCA is also referred to as 'association rule mining'.



By observing the above context, one can easily infer the existence of the implication  $\{1, 2, 3\} \rightarrow \{6\}$ , since all the objects (numbers) having the attributes (divisors) ‘ $\{1, 2, 3\}$ ’ also have the attribute (divisor) ‘ $\{6\}$ ’. In this case, the converse of the implication viz,  $\{6\} \rightarrow \{2, 3\}$  also holds. Note that since the divisor 1 is present with all objects it is a redundant attribute and hence can be ignored. Not all the converse implications are valid. For example, the converse of the implication  $\{1, 3, 5\} \rightarrow \{15\}$  does not hold. Perhaps all the implications of the divisors context presented below are easy to understand because of the logical division relation which is familiar to us. But in general, it is not always possible to examine the validity of implication formulas directly by observing the context. To this end, the following proposition helps us to verify the validity of implications.

*Proposition 3* (Lambrechts, 2012):

Let  $X, Y \subseteq M$  in a context  $(G, M, I)$ . Then the implication  $X \rightarrow Y$  is valid if and only if  $Y \subseteq X''$ . Furthermore, it is directly valid in the set of all intents of the formal concepts  $\mathbf{B}(G, M, I)$ .

In the given context shown in Table 7 consider the possibility of the implication  $\{1, 3, 5\} \rightarrow \{15\}$ .

Let us validate this implication using the above proposition.

$$\{1, 3, 5\} \rightarrow \{15\}$$

$$\{1, 3, 5\}'' = \{15, 30, 60\}' = \{1, 3, 5, 15\} \supset \{15\}.$$

Hence the above proposition holds good. Similarly, one can verify the validity of the following propositions. The use of DG basis (discussed subsequently) yields the following set of implications to the divisors context for 60 as shown in Table 8.

From the perspective of data mining, a formal context  $(G, M, I)$  is replaced by  $(T, I, R)$  whose symbols stand for Transactions (Objects), Itemsets (Attributes), and Relations (Incidence Relation) respectively. Any subset of  $k$  attributes is called as a ‘ $k$ -itemset’. An ‘intent’ is referred to as a ‘closed itemset’. The detailed discussions on discovery of association rules in data mining can be found in (Agrawal, Imielinski, & Swami, 1993; Agrawal & Srikant, 1994). The following measures are often used in the mining of association rules.

The support of an itemset  $X \subseteq M$  is defined as,  $supp(X) = \frac{|X'|}{|G|}$

An itemset is said to be a frequent itemset if its support is greater than or equal to some user specified threshold value. For any implication/ rule,  $X \rightarrow Y$ , its degree of association is measured using support and confidence measures which are defined as follows,

$$\text{Support: } supp(X \rightarrow Y) = supp(X \cup Y)$$

$$\text{Confidence: } \text{conf}(X \rightarrow Y) = \frac{\text{supp}(X \cup Y)}{\text{supp}(X)}$$

We infer that the support of an implication  $X \rightarrow Y$  is merely a statistical significance while confidence is a statistical measure of conditional probability that an object contains  $Y$  given that it already contains  $X$ . An intent which is frequent that is, its support exceeds or equal to some user specified value is called as a frequent closed itemset. The concept lattices can be pruned using closed itemsets as well (Saquer, 2000).

**Table 8**

Attribute implications derived from a formal context of Table 7

S.No	Implications	support
1	$\{ \} \rightarrow \{1\}$	12
2	$\{1, 2, 3\} \rightarrow \{6\}$	4
3	$\{1, 4\} \rightarrow \{2\}$	4
4	$\{1, 2, 5\} \rightarrow \{10\}$	4
5	$\{1, 3, 5\} \rightarrow \{15\}$	3
6	$\{1, 6\} \rightarrow \{2, 3\}$	4
7	$\{1, 2, 3, 4, 6\} \rightarrow \{12\}$	2
8	$\{1, 10\} \rightarrow \{2, 5\}$	4
9	$\{1, 2, 4, 5, 10\} \rightarrow \{20\}$	2
10	$\{1, 12\} \rightarrow \{2, 3, 4, 6\}$	2
11	$\{1, 15\} \rightarrow \{3, 5\}$	3
12	$\{1, 2, 3, 5, 6, 10, 15\} \rightarrow \{30\}$	2
13	$\{1, 20\} \rightarrow \{2, 4, 5, 10\}$	2
14	$\{1, 30\} \rightarrow \{2, 3, 5, 6, 10, 15\}$	2
15	$\{1, 60\} \rightarrow \{2, 3, 4, 5, 6, 10, 12, 15, 20, 30\}$	1
16	$\{1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30\} \rightarrow \{60\}$	1

FCA facilitates the discovery of attribute dependencies known as implication basis out of which one can derive all other rules. In practice there are two types of basis in determining the implications namely Duquenne-Guigues (DG) basis and Luxenburger

basis whose confidence levels are 100% and  $< 100\%$  respectively (Stumme, 2002; Zhang & Wu, 2011). Implications obey Armstrong rules namely,

Reflexivity: (if  $Y \subseteq X$  then  $X \rightarrow Y$ ),

Augmentation: (if  $X \rightarrow Y$  then  $X \cup Z \rightarrow Y \cup Z$ ),

Transitivity: (if  $X \rightarrow Y, Y \rightarrow Z$  then  $X \rightarrow Z$ ).

A DG basis is a minimal subset of implications/rules which can derive all implications with Armstrong rules. The main advantage of DG base of attribute implications is that, it produces a minimal possible number of implications among all other bases of implications, which hold in context. In our article, we treat with the implications derived out of DG basis.

Though the determination of all the implications of a context may seem to be an easy task, it is not so in general due to huge size of the context and sometimes implications also. To this end, for any formal context, its concept lattice and the set of implications can be produced by the use of software tools. One such software tool developed by Dr. Serhiy Yevtushenko is given in (Yevtushenko, 2000).

In the next section we discuss about the scalability issues in FCA and briefly review some of the articles with this interest.

#### 4. Scalability issue in FCA and its improvements

Though, FCA is considered as an important formalism to represent, extract and analyse any information system, it faces few problems which are to be addressed. Contexts are in general huge, complicated and contain much redundant knowledge. So, a main problem identified in practical applications of FCA is that the computational cost in processing the information system with FCA is high and the visualisation of lattice structure is difficult to perceive. This complexity issue arises due to the scalability of FCA.

The number of formal concepts grows exponentially to the size of the context and it is found to be computationally #P-complete (Kuznetsov, 2001). In addition, the number of implications grows exponentially, as attribute size increases in formal contexts and it is computationally #P-hard (Kuznetsov, 2004). In ICFCA 2006 (International Conference on FCA) handling large context was discussed as an open problem. After this conference several researchers concentrated on the scalability issues in FCA.

Literature, describes variety of approaches to control the complexity and size of contexts, concepts, concept lattices and rules. Popular research methods for improving scalability of FCA often involve: conceptual scaling for many-valued contexts, matrix decompositions, iceberg concept lattices, clustering approach, computing granular concepts, concept similarity indices, objective functions, attribute reduction, other filtration strategies, etc.

Recently, Dias and Vieira (2015) have classified concept lattice reduction techniques into three classes. In the first class of reduction techniques the redundant information is removed from the context and thereby a minimal concept lattice is obtained. This class of techniques is useful when the context has much redundant knowledge. The second class of reduction methods is the simplification of contexts/concept lattices. This class of techniques is useful to identify very important aspects in a context/concept lattices. Finally, third class of reduction techniques is the

selection of formal concepts, objects/attributes. When the context possesses some standard applicable principles this class of reduction techniques is more useful to obtain meaningful information.

We next summarize some of the improvements in literature on the scalability issues under the stated three categories in Table 9 as shown below and describe the contribution of each work briefly.

**Table 9**

Some important contributions on FCA scalability issues

Paper	Redundant Information Removal/Context Pre-processing
Ganter & Wille (1999)	Authors obtained the clarified context by removing reducible objects and attributes, and the resulting concept lattice preserves the isomorphism with the original one.
Wu, Leung, & Mi (2009)	Granular structure of concept lattices with application in knowledge reduction in formal concept analysis is examined in this paper. Information granules and their properties in a formal context are first discussed. Concepts of a granular consistent set and a granular reducts in the formal context are then introduced.
Wei & Qi (2010)	The relation between the reduction methods using concept lattices and rough sets was discussed based on classical formal context. The method unravels the relation research between these two theories.
Pei & Mi (2011)	Authors have reduced the attributes in a decision formal context based on a homomorphism consistent set from the concept lattice.
Medina (2012)	Attribute reduction in the three frameworks namely formal, object-oriented, and property-oriented concept lattices were studied in this article. Irrespective of the frameworks, it has been found that the attributes can be classified into three levels of necessity and in any level the attribute reducts are identical.
Li, Mei, Kumar, & Zhang (2013)	The author has proposed a framework for knowledge reduction from decision formal context using the idea of rule acquisition to discover a new set of non-redundant decision rules.
Li, Mei, & Lv (2013)	This article concentrates some of the issues in incomplete decision contexts such as approximate concept construction, rule acquisition and knowledge reduction. A method is proposed to build an approximate concept lattice with an incomplete context. The notion of an approximate decision rule is defined, and a method is developed to extract non-redundant approximate decision rules from an incomplete decision context. These rules are again reduced by constructing a discernibility matrix and its associated Boolean function.
Li & Wang (2016)	This paper deals with knowledge discovery in incomplete contexts. It concentrates on two issues namely concept determination with three-way decisions and attribute reduction with incomplete contexts. The notions of acceptance, rejections and non-commitment are used in the formulation of 3-way decisions.

Xu & Li (2016) The important task of granular computing (GrC) is to represent, construct, and process information granules. The authors propose a novel GrC method using FCA description of information granules. This method organizes arbitrary fuzzy information granules to become necessary and sufficient fuzzy information granules. The method is presented along with an algorithm.

Qian, Wei, & Qi (2017) In this paper, a three-way concept lattice of a given formal context is proposed. Type-I and Type-II combinatorial contexts are constructed with original and complementary formal contexts. From these two contexts, three-way concepts are constructed by two-way operators. And then the relationships between three-way concept lattices and classical concept lattices are achieved.

Cornejo, Medina, & Ramirez-Poussa (2017) Two important research topics in FCA viz., Attribute reduction and size reduction in concept lattices were handled in this article. Particularly the authors present a procedure which simultaneously reduces the attributes and concept lattice size through irreducible  $\alpha$ -cut concept lattice.

Li, Kumar, Mei, & Wang (2017) Recently authors have studied a comparison of reduction in formal decision contexts.

Paper	Concept Lattice Simplification/Modification of Closure Operators
Belohlávek & Sklenar (2005)	Various levels of granularity (scales) are set on the attributes. The granularity of an attribute increases with multi-valued attributes. Increasing or decreasing an attribute's level of granularity can adjust its importance.
Snásel, Polovincak, Abdulla, & Horak (2008)	Authors have studied the behaviour of concept lattices which are reduced using SVD matrix and NMF decomposition techniques. They also have focused on rule reduction after the context compression.
Dias & Vieira (2010)	Junction based on objects similarity (JBOS) uses the background knowledge in order to replace similar objects by representative elements using certain degree of similarity.
Kumar & Srinivas (2010)	Reduced the size of the concept lattices using fuzzy k-means clustering (FKM). Context matrix is reduced, and quotient lattices are obtained using equivalence relations derived by means of FKM Clustering. In which each record can belong to more than one cluster, and a set of membership levels is associated with each element.
Kumar (2012)	The same technique has been adopted in association rule mining of concept lattices in (Kumar, 2012) from the healthcare item set.
Kauer & Krupka (2014)	The reduction of incidence relations from the formal context also controls the complexity of the concept lattice.
Kumar, Dias, & Vieira (2015)	Compressed the original context based on non-negative matrix factorization (NMF). Context matrix is decomposed using NMF and formal context is obtained using threshold value. The non-

- negative constraint suits the context better as attributes values are always non-negative NMF permits only additive combinations but not subtractive combinations of the original vectors.
- Li, Shao, & Wu (2017) The authors introduced the three-way decision theory viz., acceptance, rejection, non-commitment in FCA recently. An axiomatic approach is proposed to generalize the three-way concepts learning through granular computing.
- Li, Huang, Qi, Qian, & Liu (2017) Fuzzy formal contexts are reduced using attribute reduction on the basis of their characteristics. The notion of 'one-sided fuzzy concept' is introduced. The attributes are classified according to three characteristics viz., core attributes, relatively necessary attributes and unnecessary attributes. Attribute reduction method is presented by virtue of attribute characteristics.
- Singh, Kumar, & Li (2016) The authors have studied concept lattices under fuzzy environments. They analysed the fuzziness in a many-valued context which is transformed into a fuzzy formal contexts and fuzzy formal concepts.
- Singh & Kumar (2014) They have reduced the number of fuzzy formal concepts by simplifying the corresponding fuzzy concept lattice structure. An algorithm is also presented for the method.
- They also have introduced the notion of bipolar fuzzy setting in FCA. They have devised a method for investigating the bipolar fuzzy formal concepts. They also produced lattice representation using bipolar fuzzy graph.

Paper	Selection/Concept filtration
Stumme (2002)	Large databases can be analysed using iceberg lattices introduced by Stumme (2002) which uses the variant 'support'. The main drawback of this approach is that the iceberg concept lattice only denotes the most upper part of the concept lattice. As such, it may not be an extraction of all the concepts of the large context.
Belohlávek, Sklenar, & Zaczal (2004)	Authors proposed a method that reduces the number of concepts using certain constraints, which are derived from attribute dependency formulas (ADF) that are additionally inputted along with the formal context. The set of concepts, which are compatible with the given set of ADFs, are reduced as important concepts.
Pernelle, Rousset, Soldano, & Ventos (2002)	Similar to iceberg concept lattices, another type of lattices viz., alpha concept lattices were introduced by the authors. Some class restraints are constructed in a formal context with attributes. The resulting concept lattice is known as alpha concept lattice by Pernelle et al. (2002). An unrestricted lattice results in an iceberg concept lattice, i.e. one having only frequent formal concepts.
Soldano, Ventos, Champesme, &	Soldano et al. (2010) have discussed the construction of alpha lattices. The extent of a term in Alpha lattices is restricted according to constraints based on an apriori categorization of

Forge (2010)	instances in classes, and on a degree $\alpha$ , which results in a smaller lattice.
Belohlavek & Vychodil (2009)	Authors determine the Attribute dependency (AD) formulas from the background knowledge. Those concepts which do not obey these AD formulas are removed.
Cheung & Vogel (2005)	Authors reduced the dimensionality of the concept lattice using the equivalence classes of objects in the process of information retrieval in which the matrix reduction technique was adopted.
Codoceo, Taramasco, & Astudillo (2011)	In these works, the selection of formal concepts is based on the notion of distance or similarity. The concepts of equivalence classes and similarity of objects or attributes are used in the process of selecting important concepts.
Dias & Vieira (2010)	
Belohlavek & Macko (2011)	In this article, a weight is assigned to each attribute to express its relevance, and then selects formal concepts considered relevant. To facilitate the application of weights, assign equal weights are assigned to attributes derived from multivalued attributes. The importance of a formal concept is measured by the sum of the weights of its attributes intention divided by the cardinality of its intention.
Li, Li, & He (2014)	Compressed a concept lattice arising from incomplete contexts using k-medoids clustering. In this process, Accuracy and similarity measures of approximate concepts are obtained and then K-medoids clustering is performed and concept lattice is compressed.
Singh, Cherukuri, & Li (2015)	Few studies have recently utilized the notion of entropy based FCA. Singh et al. (2015) have concentrated on decreasing the number of formal concepts in FCA with fuzzy attributes using entropy. Further, the number of fuzzy formal concepts is reduced at chosen granulation of the entropy based attribute intent weight.
Sumangali, Kumar, & Li (2017)	
Singh & Kumar (2016)	Recently, authors have concentrated to reduce a concept lattice using different subset of attributes as information granules.

---

## 5. Conclusion

In this paper we have presented an overview on the foundations of FCA and its historical growth to fulfil the thirst of the beginner towards FCA. The terms and notions relevant to FCA are recalled and illustrated by means of examples. FCA extracts the knowledge from any data par excellence in three dimensions viz., conceptual clusters, lattices (graphical representation) and association rules. The main advantages of the use of FCA are its simplicity, diagrammatical representation, and hierarchical overview of the underlying patterns and rules from the formal context. The common issue arising in FCA is the scalability owing to huge size contexts. We have reviewed some of the recent works on scalability.

## References

- Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data* (pp. 207–216). ACM.
- Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. In *Proceedings of 20th International Conference on Very Large Data Bases (VLDB)* (Vol. 1215, pp. 487–499).
- Arnould, A., & Nicole, P. (1981). *La logique, ou, L'Art de penser: countenant, outre les règles communes, plusieurs observations nouvelles, propres à former le jugement*. Librairie Philosophique J. VRIN. ISBN: 2-7116-0027-0
- Atif, J., Hudelot, C., & Bloch, I. (2014). Explanatory reasoning for image understanding using formal concept analysis and description logics. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(5), 552–570.
- Belohlávek, R., & Sklenar, V. (2005). Formal concept analysis over attributes with levels of granularity. In *Proceedings of the International Conference on Communicational Intelligence for Modelling, Control and Automation & Inter. Conf. on Intelligent Agents, Web Technologies and Internet Commerce* (pp. 619–624). IEEE.
- Belohlávek, R., Sklenar, V., & Zaczal, J. (2004). Concept lattices constrained by attribute dependencies. In *Proceedings of the Dateso 2004 Workshop* (pp. 63–73).
- Belohlavek, R., & Macko, J. (2011). Selecting important concepts using weights. In *Proceedings of the 9th International Conference on Formal Concept Analysis* (pp. 65–80). Springer.
- Belohlavek, R., & Vychodil, V. (2009). Formal concept analysis with background knowledge: Attribute priorities. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 39(4), 399–409.
- Birkhoff, G. (1948). *Lattice theory* (2nd ed.) (American Mathematical Society Colloquium Publications, Vol. 25). New York, NY: American Mathematical Society.
- Burmeister, P. (2003). *Formal concept analysis with ConImp: Introduction to the basic features*. Fachbereich Mathematik, Technische Universität, Darmstadt.
- Carpineto, C., & Romano, G. (2000). Order-theoretical ranking. *Journal of the American Society for Information Science*, 51(7), 587–601.
- Carpineto, C., & Romano, G. (2004). *Concept data analysis: Theory and applications*. John Wiley & Sons.
- Cheung, K. S. K., & Vogel, D. (2005). Complexity reduction in lattice-based information retrieval. *Information Retrieval*, 8(2), 285–299.
- Codocedo, V., Taramasco, C., & Astudillo, H. (2011). Cheating to achieve formal concept analysis over a large formal context. In *Proceedings of the Eighth International Conference on Concept Lattices and their Applications (CLA)* (pp. 349–362).
- Cole, R. J. (2000). *The management and visualisation of document collections using formal concept analysis*. Ph. D. Thesis, Griffith University, Australia.
- Cornejo, M. E., Medina, J., & Ramírez-Poussa, E. (2017). Attribute and size reduction mechanisms in multi-adjoint concept lattices. *Journal of Computational and Applied Mathematics*, 318, 388–402.
- Davey, B. A., & Priestley, H. A. (2002). *Introduction to lattices and order*. Cambridge university press.
- Dias, S. M., & Vieira, N. J. (2010). Reducing the size of concept lattices: The JBOS approach. In *Proceedings of the 7th International Conference on Concept Lattices and their Applications (CLA)* (Vol. 672, pp. 80–91).
- Dias, S. M., & Vieira, N. J. (2015). Concept lattices reduction: Definition, analysis and classification. *Expert Systems with Applications*, 42(20), 7084–7097.

- DIN 2330. (1993). *Begriffe und Benennungen – Allgemeine Grundsätze (Concepts and designations – General principles)*. Beuth, Berlin.
- Duquenne, V. (1999). Latticial structures in data analysis. *Theoretical Computer Science*, 217(2), 407–436.
- Duquenne, V., Chabert, C., Cherfouh, A., Doyen, A. L., Delabar, J. M., & Pickering, D. (2003). Structuration of phenotypes and genotypes through Galois lattices and implications. *Applied Artificial Intelligence*, 17(3), 243–256.
- Ganter, B. (1999). Attribute exploration with background knowledge. *Theoretical Computer Science*, 217(2), 215–233.
- Ganter, B., & Wille, R. (1989). Conceptual scaling. In F. Roberts (Ed.), *Applications of Combinatorics and Graph Theory to the Biological and Social Sciences* (pp. 139–167). New York, NY: Springer-Verlag.
- Ganter, B., & Wille, R. (1999). *Formal concept analysis: Mathematical foundations*. Berlin: Springer.
- Godin, R., Missaoui, R., & April, A. (1993). Experimental comparison of navigation in a Galois lattice with conventional information retrieval methods. *International Journal of Man-Machine Studies*, 38(5), 747–767.
- Gratzer, G. (2003). *General lattice theory*. Birkhauser Verlag.
- Groh, B., & Eklund, P. (1999). Algorithms for creating relational power context families from conceptual graphs. In *Proceedings of the 7th International Conference on Conceptual Structures* (pp. 389–400). Springer.
- Han, J., & Kamber, M. (2006). *Data mining: Concept and techniques* (2nd ed.). Munich, Germany: Morgan Kaufman.
- Hitzler, P., & Scharfe, H. (Eds.). (2016). *Conceptual structures in practice*. CRC Press.
- Huang, J. J. S., Yang, S. J. H., Chen, Z. S. C., & Wu, F. C. C. (2012). Web content adaptation for mobile device: A fuzzy-based approach. *Knowledge Management & E-Learning*, 4(1), 102–122.
- Kauer, M., & Krupka, M. (2014). Removing an incidence from a formal context. In *Proceedings of the International Conference on Concept Lattices and their Applications (CLA)* (pp. 195–206).
- Kollewe, W., Skorsky, M., Vogt, F., & Wille, R. (1994). TOSCANA—ein werkzeug zur begrifflichen analyse und erkundung von daten. In R. Wille & M. Zickwolff, (Eds.), *Begriffliche Wissensverarbeitung: Grundfragen und Aufgaben* (pp. 267–288). Mannheim: B.I.-Wissenschaftsverlag.
- Kumar, C. A. (2012). Fuzzy clustering-based formal concept analysis for association rules mining. *Applied Artificial Intelligence*, 26(3), 274–301.
- Kumar, C. A., Dias, S. M., & Vieira, N. J. (2015). Knowledge reduction in formal contexts using non-negative matrix factorization. *Mathematics and Computers in Simulation*, 109, 46–63.
- Kumar, C. A., Ishwarya M. S., & Loo, C. K. (2015). Formal concept analysis approach to cognitive functionalities of bidirectional associative memory. *Biologically Inspired Cognitive Architectures*, 12, 20–33.
- Kumar, C. A., & Singh, P. K. (2014). Knowledge representation using formal concept analysis: A study on concept generation. In B. K. Tripathy & D. P. Acharjya (Eds.), *Global Trends in Knowledge Representation and Computational Intelligence* (pp. 306–336). IGI Global Publishers.
- Kumar, C. A., & Srinivas, S. (2010). Concept lattice reduction using fuzzy K-means clustering. *Expert Systems with Applications*, 37(3), 2696–2704.
- Kumar, C. A., & Sumangali, K. (2012). Performance evaluation of employees of an organization using formal concept analysis. In *Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME)*

- (pp. 94–98). IEEE.
- Kuznetsov, S. O. (2001). On computing the size of a lattice and related decision problems. *Order*, 18(4), 313–321.
- Kuznetsov, S. O. (2004). On the intractability of computing the Duquenne-Guigues base. *Journal of Universal Computer Science*, 10(8), 927–933.
- Lambrechts, L. (2012). *Formal concept analysis*. Master Thesis, Vrije Universiteit, Brussel, Belgium.
- Li, C., Li, J., & He, M. (2014). Concept lattice compression in incomplete contexts based on K-medoids clustering. *International Journal of Machine Learning and Cybernetics*, 7(4), 539–552.
- Li, J., Huang, C., Qi, J., Qian, Y., & Liu, W. (2017). Three-way cognitive concept learning via multi-granularity. *Information Sciences*, 378, 244–263.
- Li, J., Kumar, C. A., Mei, C., & Wang, X. (2017). Comparison of reduction in formal decision contexts. *International Journal of Approximate Reasoning*, 80, 100–122.
- Li, J., Mei, C., Kumar, C. A., & Zhang, X. (2013). On rule acquisition in decision formal contexts. *International Journal of Machine Learning and Cybernetics*, 4(6), 721–731.
- Li, J., Mei, C., & Lv, Y. (2013). Incomplete decision contexts: Approximate concept construction, rule acquisition and knowledge reduction. *International Journal of Approximate Reasoning*, 54(1), 149–165.
- Li, K., Shao, M. W., & Wu, W. Z. (2017). A data reduction method in formal fuzzy contexts. *International Journal of Machine Learning and Cybernetics*, 8(4), 1145–1155.
- Li, M., & Wang, G. (2016). Approximate concept construction with three-way decisions and attribute reduction in incomplete contexts. *Knowledge-Based Systems*, 91, 165–178.
- Malzahn, N., Ziebarth, S., & Hoppe, H. U. (2013). Semi-automatic creation and exploitation of competence ontologies for trend aware profiling, matching and planning. *Knowledge Management & E-Learning*, 5(1), 84–103.
- Mattingly, K. D., Rice, M. C., & Berge, Z. L. (2012). Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management & E-Learning*, 4(3), 236–247.
- Medina, J. (2012). Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6), 1992–2002.
- Messai, N., Devignes, M. D., Napoli, A., & Smail-Tabbone, M. (2008). Many-valued concept lattices for conceptual clustering and information retrieval. In *Proceedings of the 18th European Conference on Artificial Intelligence* (pp. 127–131).
- Pei, D., & Mi, J. S. (2011). Attribute reduction in decision formal context based on homomorphism. *International Journal of Machine Learning and Cybernetics*, 2(4), 289–293.
- Pernelle, N., Rousset, M. C., Soldano, H., & Ventos, V. (2002). Zoom: A nested Galois lattices-based system for conceptual clustering. *Journal of Experimental & Theoretical Artificial Intelligence*, 14(2/3), 157–187.
- Poelmans, J., Ignatov, D. I., Kuznetsov, S. O., & Dedene, G. (2013). Formal concept analysis in knowledge processing: A survey on applications. *Expert Systems with Applications*, 40(16), 6538–6560.
- Poelmans, J., Ignatov, D. I., Kuznetsov, S. O., & Dedene, G. (2014). Fuzzy and rough formal concept analysis: A survey. *International Journal of General Systems*, 43(2), 105–134.
- Priss, U. (2006). Formal concept analysis in information science. *Annual Review of Information Science & Technology*, 40(1), 521–543.
- Qian, T., Wei, L., & Qi, J. (2017). Constructing three-way concept lattices based on

- apposition and sub position of formal contexts. *Knowledge-Based Systems*, 116, 39–48.
- Rock, T., & Wille, R. (2000). Ein toscana-erkundungssystem zur literatursuche. In G. Stumme & R. Wille (Eds.), *Begriffliche Wissensverarbeitung* (pp. 239–253). Springer.
- Saquer, J. M. (2000). *Formal concept analysis and applications*. Ph.D. dissertation, University of Nebraska, Lincoln, USA. Retrieved from <http://digitalcommons.unl.edu/dissertations/AAI9977020/>
- Sarmah, A. K., Hazarika, S. M., & Sinha, S. K. (2015). Formal concept analysis: Current trends and directions. *Artificial Intelligence Review*, 44(1), 47–86.
- Singh, P. K., Cherukuri, A. K., & Li, J. (2015). Concepts reduction in formal concept analysis with fuzzy setting using Shannon entropy. *International Journal of Machine Learning and Cybernetics*, 8(1), 179–189.
- Singh, P. K., & Kumar, C. A. (2014). Bipolar fuzzy graph representation of concept lattice. *Information Sciences*, 288, 437–448.
- Singh, P. K., & Kumar, C. A. (2016). Concept lattice reduction using different subset of attributes as information granules. *Granular Computing*, 2(3), 159–173.
- Singh, P. K., Kumar, C. A., & Gani, A. (2016). A comprehensive survey on formal concept analysis, its research trends and applications. *International Journal of Applied Mathematics and Computer Science*, 26(2), 495–516.
- Singh, P. K., Kumar, C. A., & Li, J. (2016). Knowledge representation using interval-valued fuzzy formal concept lattice. *Soft Computing*, 20(4), 1485–1502.
- Snásel, V., Polovincak, M., Abdulla, H. M. D., & Horak, Z. (2008). On concept lattices and implication bases from reduced contexts. In *Proceedings of the ICCS Supplement* (pp. 83–90).
- Soldano, H., Ventos, V., Champesme, M., & Forge, D. (2010). Incremental construction of alpha lattices and association rules. In *Proceedings of the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems* (pp. 351–360). Springer.
- Spangenberg, N. (1999). The conceptual structure of countertransference associations: An examination of diagnostic associations through an analysis of their semantic features. In N. Spangenberg & K. E. Wolff (Eds.), *Psychoanalytic Research by Means of Formal Concept Analysis*. Special des Sigmund-Freud-Instituts. Münster: Lit Verlag.
- Spangenberg, N., & Wolff, K. E. (1991). Comparison of biplot analysis and formal concept analysis in the case of a repertory grid. In H.-H. Bock & P. Ihm (Eds.), *Classification, Data Analysis, and Knowledge Organization* (pp. 104–112). Berlin Heidelberg: Springer.
- Stumme, G. (2002). Efficient data mining based on formal concept analysis. In *Proceedings of the International Conference on Database and Expert Systems Applications* (pp. 534–546). Berlin Heidelberg: Springer.
- Sumangali, K., & Kumar, C. A. (2013). Critical analysis on open source LMSs using FCA. *International Journal of Distance Education Technologies (IJDET)*, 11(4), 97–111.
- Sumangali, K., & Kumar, C. A. (2014). Determination of interesting rules in FCA using information gain. In *Proceedings of the First International Conference on Networks & Soft Computing (ICNSC)* (pp. 304–308). IEEE.
- Sumangali, K., Kumar, C. A., & Li, J. (2017). Concept compression in formal concept analysis using entropy-based attribute priority. *Applied Artificial Intelligence*, 31(3), 251–278.
- Wei, L., & Qi, J. J. (2010). Relation between concept lattice reduction and rough set reduction. *Knowledge-Based Systems*, 23(8), 934–938.

- Wille, R. (1982). Restructuring lattice theory: An approach based on hierarchies of concepts. In I. Rival (Ed.), *Ordered sets* (pp. 445–470). Springer.
- Wille, R. (1992). Concept lattices and conceptual knowledge systems. *Computers & Mathematics with Applications*, 23(6/9), 493–515.
- Wolff, K. E. (1994). A first course in formal concept analysis: How to understand line diagram. In F. Faulbaum (Ed.), *SoftStat '93, Advances in Statistical Software 4* (pp. 429–438). Stuttgart: Gustav Fischer Verlag.
- Wu, W. Z., Leung, Y., & Mi, J. S. (2009). Granular computing and knowledge reduction in formal contexts. *IEEE Transactions on Knowledge and Data Engineering*, 21(10), 1461–1474.
- Xu, W., & Li, W. (2016). Granular computing approach to two-way learning based on formal concept analysis in fuzzy datasets. *IEEE Transactions on Cybernetics*, 46(2), 366–379.
- Yao, Y. (2016). Rough-set concept analysis: interpreting RS-definable concepts based on ideas from formal concept analysis. *Information Sciences*, 346/347, 442–462.
- Yevtushenko, S. A. (2000). System of data analysis “Concept Explorer”. In *Proceedings of the 7th National Conference on Artificial Intelligence KII-2000* (pp. 127–134).
- Zhang, S., & Wu, X. (2011). Fundamentals of association rules in data mining and knowledge discovery. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(2), 97–116.
- Zushi, M., Miyazaki, Y., & Norizuki, K. (2012). Web application for recording learners’ mouse trajectories and retrieving their study logs for data analysis. *Knowledge Management & E-Learning*, 4(1), 37–50.