On concept reduction based on some graph properties

František Kardoš, Jozef Pócs, Jana Pócsová

HAL Id: hal-01416350

https://hal.archives-ouvertes.fr/hal-01416350

Submitted on 14 Dec 2016

On Concept Reduction Based on Some Graph Properties

František Kardoš[a], Jozef Pócs[b,c], Jana Pócsová[d]

[a]*LaBRI, University of Bordeaux, 351 cours de la Libération, 33405 Talence, France*
[b]*Palacký University Olomouc, Faculty of Science, Department of Algebra and Geometry, 17. listopadu 12, 771 46 Olomouc, Czech Republic*
[c]*Mathematical Institute, Slovak Academy of Sciences, Grešákova 6, 040 01 Košice, Slovakia*
[d]*Technical University of Košice, BERG Faculty, Institute of Control and Informatization of Production Processes, Boženy Němcovej 3, 043 84 Košice, Slovakia*

Abstract

The theory of concept lattices represents a well established and widely used conceptual data-mining method. Considering additional information represented by a graph structure on a set of objects, we propose a reduction of concepts. Using graph-theoretical point of view on FCA together with simple probabilistic arguments we derive the mean value of the cardinality of the reduced hierarchical structure.

Key words: Galois connection, concept lattice, concept reduction, mean value of concepts
2000 MSC: 06A15

1. Introduction

Formal concept analysis (FCA), with its roots in the algebraic theory of ordered sets, provides a mathematical framework for data analysis and knowledge discovery. As an effective tool, FCA has been applied to a variety of fields such as decision making, data mining, information retrieval or machine learning.

The theory of FCA is based on the notions of formal context and formal concept respectively. A formal context is a triple (O, A, I) where O, A are

non-empty sets and $I \subseteq O \times A$ is a binary relation. The formal context (O, A, I) formalizes the notion of object-attribute model, where O is interpreted as the set of objects, A the set of attributes (properties), and $(o, a) \in I$ is interpreted as the object o has the attribute (property) a. A concept can be formally seen as an extent-intent pair. The extent is understood as the collection of all objects belonging to the concept, while the intent is the multitude of all attributes common to all those objects. The set of all formal concepts of a formal context forms a complete lattice, called the concept lattice, which reflects the relationship of generalization and specialization among the formal concepts.

The notion of a formal context is essentially very simple and it also admits a graph-theoretic interpretation. A graph, among all representation forms, is a way to present data or knowledge. One of the major advantages of graph-based models is the presence of structural information embedded in the graphs. Assuming that an additional graph structure is given on a set of objects, one can consider only such concepts where extents satisfy a prescribed graph property. This can lead to a significant reduction of the concept lattice. Let us mention that the problem of the concept lattice reduction, or some kind of knowledge reduction, plays an important role in the recent research of the classical concept lattices ([7, 20]) as well as fuzzy concept lattices cf. [1, 2, 3, 4, 10, 11, 18, 19, 21, 22, 26, 24, 25, 27, 28].

A natural idea to constrain a concept lattice by a closure system appeared e.g. in [4]. We use the same idea, however for restricted closure systems given by independent systems induced by graph structure on objects. Our main aim is to study a reduction effect from the statistical point of view, i.e., to quantitatively describe a reduction effect in an average case. We provide an exact characteristic of the mean value of number of concepts in randomly graph based formal contexts. Initially, in order to describe a general situation, we will consider the so-called independent systems and concept lattices constrained by such systems. The presented construction involves a composition of a closure operator with the classical concept forming operators induced by a formal context. This leads to the so-called weak Galois connections which provide more general framework for defining concept-lattice-like structures than the Galois connections do. Also in this case, the concepts represent combinations of objects and attributes that are closed in a certain sense. After introducing the general theory, we focus on two independent systems induced by a graph structure on the set of all objects. In particular we will consider clique independent system determined by maximal cliques

and connected independent system determined by connected components.

2. Preliminaries

In this paper all considered sets are assumed to be finite. The theory of concept lattices is based on the notion of Galois connection, hence we recall this notion first, cf. [13]. Let $(P, \leq)$ and $(Q, \leq)$ be the ordered sets and let

$$\varphi \colon P \to Q \quad \text{and} \quad \psi \colon Q \to P$$

be maps between these ordered sets. Such a pair (φ, ψ) of mappings is called a Galois connection if:

(a) $p_1 \leq p_2 \implies \varphi(p_1) \geq \varphi(p_2)$,

(b) $q_1 \leq q_2 \implies \psi(q_1) \geq \psi(q_2)$,

(c) $p \leq \psi(\varphi(p))$ and $q \leq \varphi(\psi(q))$.

We note that any Galois connection (φ, ψ) between P and Q satisfies

$$\varphi(p) = \varphi\big(\psi\big(\varphi(p)\big)\big) \quad \text{and} \quad \psi(q) = \psi\big(\varphi\big(\psi(q)\big)\big)$$

for each $p \in P$ and $q \in Q$.

Galois connections are closely related to the notions of closure operator and closure system. A closure operator on a set S is a function $c \colon \mathbf{P}(S) \to \mathbf{P}(S)$ from the power set of S to itself which is extensive, increasing and idempotent, i.e., it satisfies the following conditions for all subsets $X, Y \subseteq S$:

(1) $X \subseteq c(X)$,

(2) $X \subseteq Y \implies c(X) \subseteq c(Y)$,

(3) $c(c(X)) = c(X)$.

A family of subsets $\mathcal{F} \subseteq \mathbf{P}(S)$ is called a closure system, if $\mathcal{F}$ satisfies the following two properties:

(1) $S \in \mathcal{F}$,

(2) $X, Y \in \mathcal{F} \implies X \cap Y \in \mathcal{F}$.

The sets of a closure system $\mathcal{F}$ are called closed sets of $\mathcal{F}$. A closure system $\mathcal{F}$ forms a lattice with respect to set inclusion. In this case, lattice operations $\wedge$ and $\vee$ are given by

$$X \wedge Y = X \cap Y \qquad X \vee Y = \bigcap\{Z \in \mathcal{F} : X \cup Y \subseteq Z\}.$$

Let us remark that closure operators are in one-to-one correspondence with closure systems by the map $c \mapsto \mathsf{FP}(c) = \{X \subseteq S : c(X) = X\}$ which associates to a closure operator c the family of all its fixed points which forms a closure system. Conversely, the inverse map associates to each closure system $\mathcal{F}$ the closure operator $c_{\mathcal{F}}$ defined by $c_{\mathcal{F}}(X) = \bigcap\{Y \in \mathcal{F} : X \subseteq Y\}$.

Let A, B be sets and a pair φ, ψ be a Galois connection between the power sets $\mathbf{P}(A)$ and $\mathbf{P}(B)$. It is well known that the composition of φ and ψ, $\psi \circ \varphi \colon \mathbf{P}(A) \to \mathbf{P}(A)$ forms a closure operator on the set A. Similarly, $\varphi \circ \psi \colon \mathbf{P}(B) \to \mathbf{P}(B)$ forms a closure operator on the set B. Moreover, the corresponding closure systems are dually isomorphic as posets.

Next, we briefly describe the basics about FCA. The notion of formal context provides a framework for this theory. A formal context is a triple (O, A, I) where O, A are nonempty sets, $I \subseteq O \times A$ is a relation from O to A. In a formal context (O, A, I), O is interpreted as the set of objects, A the set of attributes (properties), and $(o, a) \in I$ reads as that the object o has the property a.

Let (O, A, I) be a formal context. There is a pair of mappings $^{\uparrow} \colon \mathbf{P}(O) \to \mathbf{P}(A)$ and $^{\downarrow} \colon \mathbf{P}(A) \to \mathbf{P}(O)$, which forms a Galois connection between power sets of O and A respectively.

$$X^{\uparrow} = \{a \in A : (o, a) \in I, \forall o \in X\},$$

$$Y^{\downarrow} = \{o \in O : (o, a) \in I, \forall a \in Y\}.$$

A formal concept of the formal context (O, A, I) is a pair (X, Y) with $X \subseteq G$, $Y \subseteq M$, $X^{\uparrow} = Y$ and $X = Y^{\downarrow}$. The set X is called the extent and Y is called the intent of the concept (X, Y). If (X_1, Y_1) and (X_2, Y_2) are concepts, then (X_1, Y_1) is a subconcept of (X_2, Y_2), if $X_1 \subseteq X_2$ (which is equivalent to $Y_1 \supseteq Y_2$). This fact is denoted by $(X_1, Y_1) \leq (X_2, Y_2)$. The relation $\leq$ is called the hierarchical order of concepts and the structure of all concepts ordered by this relation is denoted by $\underline{\mathfrak{B}}(O, A, I)$.

Let (O, A, I) be a formal context. Any extent X of the concept (X, Y) satisfies $X = Y^{\downarrow} = X^{\uparrow\downarrow}$, hence it is closed under the composition of operators

4

$\uparrow$ and $\downarrow$. Conversely, if X is closed, then the pair $(X, X^\uparrow)$ forms a concept. Hence the set of all extents of $\underline{\mathfrak{B}}(O, A, I)$ forms a closure system on the set O of all objects. The order structure of $\underline{\mathfrak{B}}(O, A, I)$ is uniquely determined by the order structure of extents, thus $\underline{\mathfrak{B}}(O, A, I)$ forms a lattice, referred as the concept lattice. In this case, the lattice operations $\wedge$ and $\vee$ are given by

$$(X_1, Y_1) \wedge (X_2, Y_2) = \left(X_1 \cap X_2, (X_1 \cap X_2)^\uparrow\right) = \left(X_1 \cap X_2, (Y_1 \cup Y_2)^{\downarrow\uparrow}\right)$$

$$(X_1, Y_1) \vee (X_2, Y_2) = \left((Y_1 \cap Y_2)^\downarrow, Y_1 \cap Y_2\right) = \left((X_1 \cup X_2)^{\uparrow\downarrow}, Y_1 \cap Y_2\right).$$

An analogy of concept lattices can be defined in more general framework than provide Galois connections. One of such possibility is described in [23], where the so-called weak Galois connections are presented.

Let P and Q be partially ordered sets. A pair of mappings $\alpha : P \to Q$, $\beta : Q \to P$ is said to be a weak Galois connection between P and Q if it satisfies:

(i) $x_1 \leq x_2$ implies $\alpha(x_1) \geq \alpha(x_2)$,

(ii) $y_1 \leq y_2$ implies $\beta(y_1) \geq \beta(y_2)$,

(iii) $\beta(\alpha(x)) = \beta(\alpha(\beta(\alpha(x))))$ and $\alpha(\beta(y)) = \alpha(\beta(\alpha(\beta(y))))$.

Similarly, as in the case of Galois connections, if α and β is a weak Galois connection between $\mathbf{P}(O)$ and $\mathbf{P}(A)$, a concept lattices like structure can be defined. Denote by $\mathcal{B}(O, A, \alpha, \beta)$ a subset of $\mathbf{P}(O) \times \mathbf{P}(A)$ such that

$$\mathcal{B}(O, A, \alpha, \beta) = \{(X, Y) : Y = \alpha(X) \text{ and } X = \beta(Y)\}.$$

A partial order $\leq$ on $\mathcal{B}(O, A, \alpha, \beta)$ is defined in the same way as in the case of concept lattices, i.e.,

$$(X_1, Y_1) \leq (X_2, Y_2) \text{ iff } X_1 \subseteq X_2 \text{ iff } Y_1 \supseteq Y_2.$$

The following proposition shows that such defined structure has almost the same properties as concept lattices defined via Galois connections.

Proposition 2.1. *The structure $\mathcal{B}(O, A, \alpha, \beta)$ forms a lattice, where the operation $\wedge$ and $\vee$ are given by*

$$(X_1, Y_1) \wedge (X_2, Y_2) = \left(\beta\left(\alpha(X_1 \cap X_2)\right), \alpha\left(\beta(Y_1 \cup Y_2)\right)\right)$$

$$(X_1, Y_1) \vee (X_2, Y_2) = \left(\beta\left(\alpha(X_1 \cup X_2)\right), \alpha\left(\beta(Y_1 \cap Y_2)\right)\right).$$

Proof. We prove the formula for infimum. Let $(X_1, Y_1), (X_2, Y_2) \in \mathcal{B}(O, A, \alpha, \beta)$ be arbitrary elements. In order to simplify notations, denote $X = \beta(\alpha(X_1 \cap X_2))$ and $Y = \alpha(\beta(Y_1 \cup Y_2))$. First, we show that $\alpha(X) = Y$ and $\beta(Y) = X$, i.e.,

$$\Big(\beta\big(\alpha(X_1 \cap X_2)\big), \alpha\big(\beta(Y_1 \cup Y_2)\big) \Big) \in \mathcal{B}(O, A, \alpha, \beta).$$

Since both mappings α and β are antitone, it follows that the compositions $\beta \circ \alpha \colon \mathbf{P}(O) \to \mathbf{P}(O)$ as well as $\alpha \circ \beta \colon \mathbf{P}(A) \to \mathbf{P}(A)$ are monotone. Consequently, $X = \beta(\alpha(X_1 \cap X_2)) \subseteq \beta(\alpha(X_1)) = \beta(Y_1) = X_1$ and $X \subseteq X_2$. Applying the mapping α we obtain $\alpha(X) \supseteq \alpha(X_1) = Y_1$ and $\alpha(X) \supseteq \alpha(X_2) = Y_2$, which yields $\alpha(X) \supseteq Y_1 \cup Y_2$. Similarly, it can be shown that $\beta(Y) \subseteq X_1 \cap X_2$. From these two inequalities and using the property iii) of weak Galois connections we obtain

$$\alpha(X) = \alpha(\beta(\alpha(X))) \supseteq \alpha(\beta(Y_1 \cup Y_2)) = Y$$

and

$$\beta(Y) = \beta(\alpha(\beta(Y))) \subseteq \beta(\alpha(X_1 \cap X_2)) = X.$$

On the other hand, $\alpha(X) \supseteq Y$ implies $X = \beta(\alpha(X)) \subseteq \beta(Y)$, while $\beta(Y) \subseteq X$ implies $Y = \alpha(\beta(Y)) \supseteq \alpha(X)$. Hence, $\alpha(X) = Y$ and $\beta(Y) = X$.

Further, we show that (X, Y) is the infimum of (X_1, Y_1) and (X_2, Y_2). Obviously, (X, Y) is a lower bound of these elements. If (C, D) is a lower bound of (X_1, Y_1) and (X_2, Y_2), then $C \subseteq X_1$ and $C \subseteq X_2$ yielding $C \subseteq X_1 \cap X_2$ and dually $D \supseteq Y_1 \cup Y_2$. Consequently $C = \beta(\alpha(C)) \subseteq \beta(\alpha(X_1 \cap X_2)) = X$ and $D = \alpha(\beta(D)) \supseteq \alpha(\beta(Y_1 \cup Y_2)) = Y$, i.e., (X, Y) represents the greatest lower bound of (X_1, Y_1) and (X_2, Y_2).

Using similar arguments, one can prove the formula for supremum. Hence $\mathcal{B}(O, A, \alpha, \beta)$ is a partially ordered set in which every pair of elements has infimum and supremum, i.e., it forms a lattice. $\square$

Let us note, that extents and intents of $\mathcal{B}(O, A, \alpha, \beta)$ need not form closure systems in general. In fact, this is the only relevant difference between concept lattices and concept lattices defined via weak Galois connections, respectively.

After introducing the basics about concept lattices, we recall some definitions from graph theory used through the paper, cf. [5]. A graph G is an ordered pair $(V(G), E(G))$ consisting of a set $V(G)$ of vertices and a set $E(G)$ of 2-element subsets of $V(G)$, called edges. We will indicate the fact

that vertices u and v form an edge by $uv \in E(G)$. A graph is bipartite if its vertex set can be partitioned into two subsets A and B so that every edge has one end in A and one end in B. We denote a bipartite graph G with bipartition (A, B) by $G[A, B]$. If every vertex in A is joined to every vertex in B, then $G[A, B]$ is called a complete bipartite graph. Note that for our purpose graphs of the form $G[A, \emptyset]$ and $G[\emptyset, B]$ are considered as complete bipartite too. A clique (complete subgraph) of a graph G is a subset of the vertex set such that any two vertices are adjacent. Note, that $\emptyset$ is considered to be a clique of any graph. A graph is connected if, for every partition of its vertex set into two nonempty sets X and Y, there is an edge with one end in X and one end in Y. Every graph may be expressed uniquely (up to order) as a disjoint union of connected graphs. These graphs are called the connected components, or simply the components of the graph.

A formal context can be seen as a biadjacency matrix of the corresponding bipartite graph. In particular, if (A, B, I) is a formal context, the corresponding bipartite graph $G[A, B]$ is given by $ab \in E(G[A, B])$ if and only if $(a, b) \in I$. In the same way, each bipartite graph gives rise to the corresponding formal context. Hence, there is a one-to-one correspondence between the class of all formal contexts and the class consisting of all bipartite graphs.

The concepts of a context also permit a graph theoretical interpretation, because they can be seen as maximal complete bipartite subgraphs of the bipartite graph representing the context. More precisely, a maximal complete bipartite subgraph of $G[A, B]$ is a bipartite graph $G[X, Y]$ where $X \subseteq A$, $Y \subseteq B$, and such that

$\quad G[X, Y]$ is a complete bipartite graph,

$\quad \forall\, a \in A \smallsetminus X, \exists\, b \in Y : ab \notin E\big(G[A, B]\big),$

$\quad \forall\, b \in B \smallsetminus Y, \exists\, a \in X : ab \notin E\big(G[A, B]\big).$

Remark 2.2. Obviously, a pair (X, Y) forms a concept of a context (O, A, I) if and only if $G[X, Y]$ is a maximal complete bipartite subgraph of the bipartite graph corresponding to the context (O, A, I).

3. Concept lattices constrained by independence systems on objects

Although FCA is a very useful data analysis and visualization method, the number of concepts resulting from a relatively small object-attribute

7

models is usually large. Note, that it can be exponential w.r.t. the size of the input data. Therefore, as we already mentioned in the introduction, the reduction problem of the concept lattices represents one of the main challenges in FCA research. One of the possible approaches is to use an approach based on constraining concept lattices by additional information (a user's background knowledge). In [4], a situation when an additional background knowledge is imposed in a form of a closure system on attributes, is described. In this paper, we consider a more specific situation, i.e., when a constraint is given by a closure operator arisen from the so-called independent system. These systems occur naturally in connection with various graph-theoretical notions, moreover it enables us to characterize such constrained concept lattices in terms of concept forming operators based on weak Galois connections. Remark that this is not always possible when a general closure system as a constraint is considered, since some repeated iteration is needed for a characterization of the closure operator emerged from the intersection of two closure systems.

In the sequel, we will assume that a constraint in a form of an independence system on a set O of objects is given. Recall (see [5]), that a set system $(O, \mathcal{F})$ is called an independence system on O if $\mathcal{F}$ is nonempty and, for any member $F \in \mathcal{F}$, $H \subseteq F$ implies $H \in \mathcal{F}$. The members of $\mathcal{F}$ are referred to as independent sets and their maximal elements as bases.

If $(O, \mathcal{F})$ is an independence system, $X_1, X_2 \subseteq O$ belong to $\mathcal{F}$, then $X_1 \cap X_2 \in \mathcal{F}$ as well. This yields that the set system $\mathcal{F} \cup \{O\}$ forms a closure system on the set O.

Let (O, A, I) be a formal context and $(O, \mathcal{F})$ be an independence system on the set of all objects. The closure operator corresponding to the closure system $\mathcal{F} \cup \{O\}$ is denoted by c. From the structure of this closure system one can easily obtain that for all $X \subseteq O$

$$c(X) = \begin{cases} X, \text{ if } X \in \mathcal{F}, \\ O, \text{ otherwise.} \end{cases}$$

In order to define concept forming operators we put $\nearrow \colon \mathbf{P}(O) \to \mathbf{P}(A)$ and $\swarrow \colon \mathbf{P}(A) \to \mathbf{P}(O)$ in the following way:

$$\nearrow(X) = c(X)^{\uparrow}, \text{ for all } X \subseteq O, \tag{1}$$

$$\swarrow(Y) = c(Y^{\downarrow}), \text{ for all } Y \subseteq A. \tag{2}$$

The main features of such defined mappings are described in the following theorem.

Theorem 3.1. *The pair of mappings $\nearrow$ and $\swarrow$ has the following properties:*

(i) $X_1 \subseteq X_2 \subseteq O \implies \nearrow(X_1) \supseteq \nearrow(X_2)$,

(ii) $Y_1 \subseteq Y_2 \subseteq A \implies \swarrow(Y_1) \supseteq \swarrow(Y_2)$,

(iii) $X \subseteq \swarrow(\nearrow(X))$ *for all* $X \subseteq O$,

(iv) $\swarrow(Y) = \swarrow\big(\nearrow(\swarrow(Y))\big)$ *for all* $Y \subseteq A$,

(v) $\swarrow\big(\nearrow(X)\big) = \swarrow\big(\nearrow(\swarrow(\nearrow(X)))\big)$ *and* $\nearrow\big(\swarrow(Y)\big) = \nearrow\big(\swarrow(\nearrow(\swarrow(Y)))\big)$.

Proof. Suppose that $X_1 \subseteq X_2$. Since the closure operator c is isotone, we obtain $c(X_1) \subseteq c(X_2)$. Applying the antitone mapping $^\uparrow$ we obtain $\nearrow(X_1) = c(X_1)^\uparrow \supseteq c(X_2)^\uparrow = \nearrow(X_2)$, what proves that the mapping $\nearrow$ is antitone. The property (ii) can be proved analogously.

Since $\swarrow(\nearrow(X)) = \swarrow\big(c(X)^\uparrow\big) = c\big(c(X)^{\uparrow\downarrow}\big)$, using the properties of Galois connections and closure operators we obtain

$$X \subseteq c(X) \subseteq c(X)^{\uparrow\downarrow} \subseteq c\big(c(X)^{\uparrow\downarrow}\big) = \swarrow(\nearrow(X)).$$

In order to prove property (iv), it suffices to consider two possibilities, i.e., whether $c(Y^\downarrow) = Y^\downarrow \neq O$ (in this case $Y^\downarrow \in \mathcal{F}$, i.e., $Y^\downarrow$ is independent) or $Y^\downarrow = O$.

$$\swarrow\big(\nearrow(\swarrow(Y))\big) = c\big(c\big(c(Y^\downarrow)\big)^{\uparrow\downarrow}\big) = c\big(c(Y^\downarrow)^{\uparrow\downarrow}\big) =$$
$$= \begin{cases} c(Y^{\downarrow\uparrow\downarrow}) = c(Y^\downarrow) = \swarrow(Y), & \text{if } c(Y^\downarrow) = Y^\downarrow \neq O, \\ c(O^{\uparrow\downarrow}) = c(O) = O = \swarrow(Y), & \text{if } c(Y^\downarrow) = O. \end{cases}$$

Property (v) follows directly from (iv). $\square$

Corollary 3.2. *The mappings $\nearrow$ and $\swarrow$ form a weak Galois connection between $\mathbf{P}(O)$ and $\mathbf{P}(A)$. Moreover, the composition of mappings $\swarrow \circ \nearrow \colon \mathbf{P}(O) \to \mathbf{P}(O)$ forms a closure operator on the set of objects O.*

Applying Proposition 2.1, we obtain that the structure $\mathcal{B}(O, A, \nearrow, \swarrow)$ of fixed points (X, Y) satisfying $X \subseteq O$, $Y \subseteq A$, $\nearrow(X) = Y$ and $\swarrow(Y) = X$ with the hierarchical order forms a lattice. The following theorem characterizes the corresponding lattice operations.

9

Theorem 3.3. *The lattice operations $\wedge$ and $\vee$ in $\mathcal{B}(O, A, \nearrow, \swarrow)$ satisfy*

$$(X_1, Y_1) \wedge (X_2, Y_2) = \Big(X_1 \cap X_2, \nearrow\big(\swarrow(Y_1 \cup Y_2)\big)\Big)$$

$$(X_1, Y_1) \vee (X_2, Y_2) = \Big(\swarrow\big(\nearrow(X_1 \cup X_2)\big), \nearrow\big(\swarrow(Y_1 \cap Y_2)\big)\Big)$$

Proof. Since the composition $\swarrow \circ \nearrow$ forms a closure operator on the set of objects O, the intersection of closed sets X_1 and X_2 satisfies $X_1 \cap X_2 = \swarrow\big(\nearrow(X_1 \cap X_2)\big)$. Consequently, the expressions for meet and join of concepts follow from Proposition 2.1. $\square$

In the following lemma, we provide a characterization of the extents of the lattice $\mathcal{B}(O, A, \nearrow, \swarrow)$.

Lemma 3.4. *Let $X \subset O$ be a proper subset of objects. Then $X = \swarrow(\nearrow(X))$ if and only if $X = X^{\uparrow\downarrow}$ and X is independent.*

Proof. Assume that $X = \swarrow(\nearrow(X)) = c\big(c(X)^{\uparrow\downarrow}\big)$. Since c is idempotent we obtain
$$c(X) = c\big(c\big(c(X)^{\uparrow\downarrow}\big)\big) = c\big(c(X)^{\uparrow\downarrow}\big) = X,$$
which yields $X \in \mathcal{F}$, i.e., the subset X is independent. Further, $X \subseteq X^{\uparrow\downarrow}$. Since $c(X) = X$ and c is extensive, $X \subset X^{\uparrow\downarrow}$ would imply $X \subset \swarrow(\nearrow(X))$, which is a contradiction.

Conversely, suppose that X is independent and $X = X^{\uparrow\downarrow}$. Then

$$\swarrow(\nearrow(X)) = c\big(c(X)^{\uparrow\downarrow}\big) = c(X^{\uparrow\downarrow}) = c(X) = X.$$

$\square$

This lemma shows that the lattice $\mathcal{B}(O, A, \nearrow, \swarrow)$ can be seen as a reduction of the concept lattice $\underline{\mathfrak{B}}(O, A, I)$ where extents are constrained by the independence system $(O, \mathcal{F})$. From an algebraic point of view, the lattice $\mathcal{B}(O, A, \nearrow, \swarrow)$ forms a meet-subsemilattice of the lattice $\underline{\mathfrak{B}}(O, A, I)$. This fact follows directly from the expression for infima in Theorem 3.3.

At the end of this section we provide a small illustrative example concerning social networks, cf. also [8, 9] which deal with social networks from the hierarchical structures point of view and FCA perspective respectively.

10

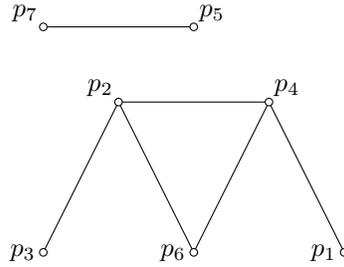| I	a	b	c	d	e
p_1	×				×
p_2	×	×	×		
p_3	×			×	
p_4	×	×	×		×
p_5					×
p_6			×	×	
p_7					×

Table 1: The incidence relation of Example 3.5.

Fig. 1: The interaction graph between people of the context (O, A, I).

Example 3.5. Consider the following data, represented by a formal context (O, A, I), obtained from a social network. The set of objects consists of seven persons, denoted by p_i, $i \in \{1, \ldots, 7\}$, the attributes $\{a, b, c, d, e\}$ represent some posts shared by the persons and the incidence relation I (represented by Table 1) is given by $(p, x) \in I$ if and only if a person p "likes" a post x.

Further, we have additional information in the form of a graph, where edges represent interactions between people (Figure 1).

With this graph, two independence systems can be associated. The first one consists of all cliques, i.e., sets of nodes in the graph connected with each other. In this case, a clique represents a group of people who interact directly with each other. Remark that each one-element set is considered as a clique. The second one is determined by connected components of the graph, i.e., a subset of nodes is independent, provided that for any two nodes there is a path in the graph connecting them.

The classical concept lattice corresponding to the context (O, A, I) of Table 1 is depicted in Figure 2.
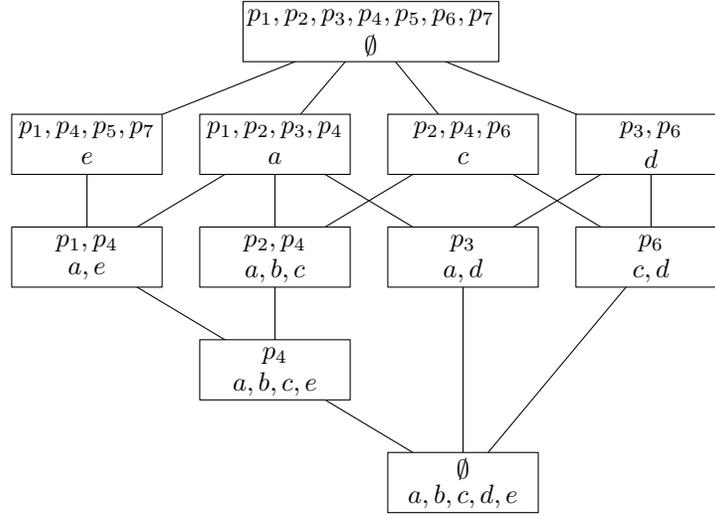
11

Fig. 2: The classical concept lattice corresponding to the context (O, A, I).

For comparison, Figure 3 shows the lattice structure when the operators $\nearrow, \swarrow$ are applied to the context (O, A, I) together with the clique independence system as some kind of a background knowledge.

According to Lemma 3.4, extents of $\mathcal{B}(O, A, \nearrow, \swarrow)$ are precisely independent fixed points of the concept forming operators $^{\uparrow}$ and $^{\downarrow}$, i.e., extent is represented by a group of people interacting with each other, while "liking" some common posts.

The structure of $\mathcal{B}(O, A, \nearrow, \swarrow)$ when the independence system based on connected components is considered, is depicted in Fig 4.

In this case, extents are precisely groups of people forming extent of the context (O, A, I) such that for any two persons of that group there is a connection between them, possibly going through some other people not belonging to the group. This kind of the independence system can be used, e.g., when the sharing of some messages (information) between people within a group is desired. Obviously, the existence of connection between any two persons can fulfill this requirement.

In practice, a social network may consists of hundreds or even thousands of particular users, which usually share several thousands of posts. In general, as the size of sample increases, an investigation with the statistical or probabilistic nature becomes more and more important. In the next section
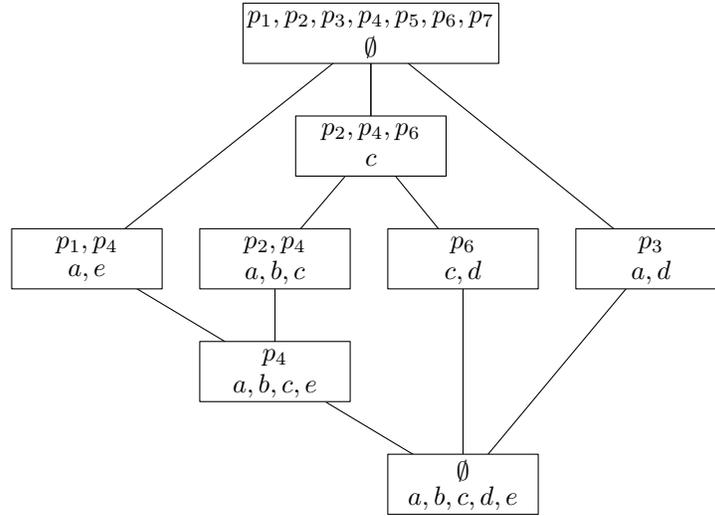
Fig. 3: The lattice $\mathcal{B}(O, A, \nearrow, \swarrow)$ corresponding to the context (O, A, I) constrained by the clique independence system induced by the graph from Figure 1.
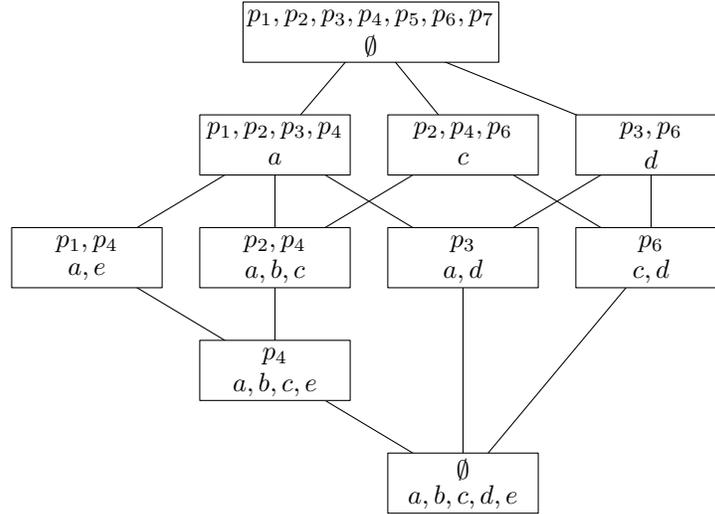

Fig. 4: The lattice $\mathcal{B}(O, A, \nearrow, \swarrow)$ corresponding to the context (O, A, I) constrained by the connected independence system induced by the graph from Figure 1.

13

we try to provide an answer to the following types of problems: Given a formal context with "typically" distributed attributes, what is the number of concepts in the resulting concept lattice? How this number is affected, provided some constraint imposed by a clique (or connected) independence system induced by a random graph is considered?

4. Expected number of concepts in a graph based model

Let (O, A, I) be a formal context. Similarly as in Example 3.5, we will assume that additional information in the form of a graph G on the object set O is provided. Such graph structure on the set of objects can represent a similarity relation where edges link similar objects together. We focus on the concept reduction, provided some of two independence systems induced by the graph structure on the set of objects is considered. Particularly, we will consider clique independence system and connected independent system respectively. Recall, that the clique independence system possesses all cliques of a graph. In this case, the basis of this independence system is formed by maximal cliques. The connected independent system consists of all subsets which belong to some connected component of a graph. Obviously, the connected components form a basis of this system. Let us note that these two graph independence systems are closely related to the complete linkage hierarchical clustering and single linkage hierarchical clustering respectively, cf. [16].

As the size of the object set increases, a graph structure on the objects can be very complex. However, we will not assume any special structural configurations (properties) of this graph model. Instead, we take into the account only the following two assumptions: there is a uniform parameter p denoting the probability that a given edge occurs in the graph and that the edges occur in the graph independently.

Under these assumptions, the well-known random graph model $G(n, p)$ will be considered (see the textbook [6], and also [14] where this model appeared explicitly for the first time). In this model, a graph is constructed by connecting n vertices randomly, in which each possible edge between two vertices is present with independent probability p, and absent with probability $1 - p$. Formally, the underlying probabilistic space, also denoted by $G(n, p)$, consists of all graphs on a fixed set of n nodes, where the probability

of a graph G with exactly k edges, $0 \leq k \leq \binom{n}{2}$, is given by

$$\mathsf{P}(G) = p^k(1-p)^{\binom{n}{2}-k}.$$

The formula follows from the fact that there are $\binom{n}{2}$ of possible edges between n vertices. If G is a graph with exactly k edges, there are $\binom{n}{2} - k$ non-edges and since these events are mutually independent, the resulting probability $\mathsf{P}(G)$ is the product of $\binom{n}{2}$ numbers, from which k numbers equal to p, while remaining $\binom{n}{2} - k$ numbers are equal to $1 - p$.

The parameter p, denoting the probability of an edge occurrence in this random graph model can be thought of as some kind of weighting function. Obviously, as p increases from 0 to 1, graphs with more edges become more and more likely, while it becomes less and less likely to find a set including only graphs with few edges.

In connection with formal contexts, we will assume that particular attributes are distributed independently with a probability $p_a \in [0,1]$, i.e., for an object $o \in O$, an attribute $a \in A$, $(o,a) \in I$ with probability p_a. The parameters p_a can be seen as some global characterizations of an analyzed population, denoting frequencies with which particular attributes are distributed over the population. These probabilities, as well as the probability p used in the random graph model are assumed to be given a priori. They can be approximated by sampling randomly from the set of objects and attributes.

Firstly, we derive the mean value of the number of concepts in the classical concept lattices under the above mentioned assumptions.

Theorem 4.1. *Let n, m be positive integers, O be an n-element set of objects and A be an m-element set of attributes. If each attribute $a \in A$ is assigned to an object with a constant probability p_a independently, then the mean value of the number of concepts $E(|\underline{\mathfrak{B}}|)$ is given by*

$$E\big(|\underline{\mathfrak{B}}|\big) = \sum_{k=0}^{n}\sum_{Y \subseteq A}\binom{n}{k}\Big(\prod_{a \in Y}p_a\Big)^k\Big(1 - \prod_{a \in Y}p_a\Big)^{n-k}\prod_{a \in A \smallsetminus Y}\big(1 - p_a{}^k\big). \qquad (3)$$

Before providing a proof of the theorem, we make some remarks concerning the underlying probability space. Formally, it consists of all contexts, where particular attributes are distributed in accordance with the assumptions of the theorem. From the graph-theoretical point of view, all bipartite

15

graphs are considered, where edges between O and A are added independently with probability p_a provided an edge is incident to the attribute a. In particular, the probability of a bipartite graph $G[O, A]$ is given by

$$\mathsf{P}\big(G[O, A]\big) = \prod_{a \in A} p_a{}^{\deg(a)} (1 - p_a)^{n - \deg(a)}.$$

It follows from the fact that for a given vertex $a \in A$ there are $\deg(a)$ edges incident to the vertex a and $n - \deg(a)$ non-incident to a. Since the presence and absence of an edge represent independent events, the resulting probability is the product of the formulae $p_a{}^{\deg(a)}(1 - p_a)^{n - \deg(a)}$, $a \in A$.

In the sequel, this (bipartite graph-based) probabilistic space will be denoted by $B(n, A)$.

Proof. Obviously, $|\mathfrak{B}|$ is an integer-valued random variable. In order to compute its mean value, we use simpler random variables, the so-called indicators. For $X \subseteq O$ define $I_X = 1$ if X is an extent, $I_X = 0$ otherwise. Since $|\mathfrak{B}| = j$ for some $1 \le j \le 2^n$ if and only if there are j different extents, it follows that

$$|\mathfrak{B}| = \sum_{X \subseteq O} I_X.$$

By linearity of expectation

$$E\big(|\mathfrak{B}|\big) = E\Big(\sum_{X \subseteq O} I_X\Big) = \sum_{X \subseteq O} E(I_X) = \sum_{X \subseteq O} \mathsf{P}(I_X = 1),$$

where $\mathsf{P}(I_X = 1)$ denotes the probability that the set X forms an extent.

Let X be a fixed k-element subset of objects. Then X is an extent if and only if there is a subset $Y \subseteq A$ such that (X, Y) forms a concept. For a subset $Y \subseteq A$ denote by $\mathcal{A}_{X,Y}$ the event that the pair (X, Y) is a concept. Since, given a context, intents are determined uniquely, it follows that $\{\mathcal{A}_{X,Y} : Y \subseteq A\}$ forms a disjoint family of events. Due to additivity of the probabilistic measure, denoting by $\mathsf{P}_Y(k)$ the probability that an arbitrary k-element set of objects together with the set Y of attributes forms a concept, we obtain

$$\mathsf{P}(I_X = 1) = \sum_{Y \subseteq A} \mathsf{P}_Y(k).$$

Hence, in order to determine $\mathsf{P}_Y(k)$, assume that (X, Y) forms a concept. This happens if and only if each object in X has all the attributes from Y,

16

none of the remaining $n-k$ objects has all the attributes from Y, and none of the remaining attributes in $A \smallsetminus Y$ is matched to all the objects in X. Because presence (absence) of different attributes represents independent events, the three mentioned events are independent as well. Hence,

$$\mathsf{P}_Y(k) = \Big(\prod_{a\in Y} p_a\Big)^k \Big(1 - \prod_{a\in Y} p_a\Big)^{n-k} \prod_{a\in A\smallsetminus Y}\big(1 - p_a{}^k\big). \tag{4}$$

For each $k = 0, \ldots, n$ we may choose any k-element set X of objects, which yields

$$E\big(|\mathfrak{B}|\big) = \sum_{X\subseteq O}\sum_{Y\subseteq A}\mathsf{P}_Y(k) = \sum_{k=0}^{n}\binom{n}{k}\sum_{Y\subseteq A}\mathsf{P}_Y(k) = \sum_{k=0}^{n}\sum_{Y\subseteq A}\binom{n}{k}\mathsf{P}_Y(k).$$

$\square$

Let us note, that if $|A| = m$ and $p_a = p_0$ for all attributes $a \in A$, then $E\big(|\mathfrak{B}|\big)$ can be expressed as

$$E\big(|\mathfrak{B}|\big) = \sum_{k=0}^{n}\sum_{l=0}^{m}\binom{n}{k}\binom{m}{l}p_0{}^{kl}(1-p_0^l)^{n-k}(1-p_0^k)^{m-l}.$$

Hence, the expression (3) can be seen as a generalization of the result presented in [15].

Further, we provide a similar formula for computing the expected number of concepts constrained by cliques. In the sequel, a lattice of formal concepts restricted by a clique independence system will be denoted by $\mathfrak{B}_1$.

Theorem 4.2. *Let n, m be positive integers, O be an n-element set of objects and A be an m-element set of attributes. If each attribute $a \in A$ is assigned to an object with a constant probability p_a and the probability of independence of each pair of objects is p, then the mean value of the number of clique-constrained concepts $E(|\mathfrak{B}_1|)$ is given by*

$$E\big(|\mathfrak{B}_1|\big) = 1 + \sum_{k=0}^{n-1}\sum_{Y\subseteq A}\binom{n}{k}p^{\binom{k}{2}}\Big(\prod_{a\in Y} p_a\Big)^k\Big(1-\prod_{a\in Y} p_a\Big)^{n-k}\prod_{a\in A\smallsetminus Y}\big(1-p_a{}^k\big). \tag{5}$$

17

Proof. Let us remark that in this case the underlying probabilistic space is equal to the direct product $B(n, A) \times G(n, p)$, where $B(n, A)$ denotes the graph-based probabilistic space defined before the proof of Theorem 4.1.

Similarly as in the proof of Theorem 4.1, $E(|\mathfrak{B}_1|)$ can be expressed as

$$E(|\mathfrak{B}_1|) = \sum_{X \subseteq O} \mathsf{P}(I_X = 1).$$

However, in this case the indicator random variable I_X for $X \subseteq O$ is defined by $I_X = 1$ if X is an extent and at the same time it forms a clique; $I_X = 0$ otherwise. Note that the whole object set O is considered being clique. Again, being an extent and forming clique represent independent events. Obviously, for $k < n$ a fixed k-element set X of objects forms a clique with probability $p^{\binom{k}{2}}$, hence

$$\mathsf{P}(I_X = 1) = \sum_{Y \subseteq A} p^{\binom{k}{2}} \mathsf{P}_Y(k).$$

The set O itself of size n is considered clique-closed, thus always forming a single concept with some $Y \subseteq A$, which yields the following formula

$$E(|\mathfrak{B}_1|) = 1 + \sum_{k=0}^{n-1} \sum_{Y \subseteq A} \binom{n}{k} p^{\binom{k}{2}} \mathsf{P}_Y(k).$$

$\square$

Before we state the next theorem let us recall some known results about connectivity of random graphs [12].

Considering the random graph model $G(n, p)$, let $\mathsf{Q}(j)$ denote the probability that a random j-element subset of vertices induces a connected subgraph (for $j = 1, \ldots, n$). Then $\mathsf{Q}(1) = 1$ and for $j \geq 2$ we have

$$\mathsf{Q}(j) = 1 - \sum_{i=1}^{j-1} \binom{j-1}{i-1} \mathsf{Q}(i)(1-p)^{i(j-i)}. \tag{6}$$

To observe this, consider a j-element set X. Fix a vertex $v \in X$. The set X does not induce a connected subgraph if and only if the component containing v of the graph induced by X has exactly i vertices for some $i =$

18

$1, \ldots, j - 1$. For a fixed i, there is $\binom{j-1}{i-1}$ subsets of X containing v, each of them is connected with probability $Q(i)$, and maximal with respect to connectivity (no edges towards the remaining vertices of X) with probability $(1 - p)^{i(j-i)}$.

Let $R(k)$ denote the probability that a random k-element subset of vertices belong to the same connected component of G. Then

$$R(k) = \sum_{j=k}^{n} \binom{n-k}{j-k} Q(j)(1 - p)^{j(n-j)}. \tag{7}$$

To observe this, consider a k-element set X of vertices. All the vertices of X belong to the same connected component of G if and only if there exists a connected component C of size $j \geq k$ containing all of them, which means there exists a set X' of $j-k$ vertices disjoint from X such that $X \cup X'$ induces a maximal connected graph.

Denote by $\mathfrak{B}_2$ a lattice of formal concepts constrained by a connected independence system on an object set O. Recall that the basis of such system consists of connected components and the whole object set is considered as connected as well. Knowing the expression (7) for the probability $R(k)$ and using similar arguments as in the proof of Theorem 4.2, one can prove the following:

Theorem 4.3. *Let n, m be positive integers, O be an n-element set of objects and A be an m-element set of attributes. If each attribute $a \in A$ is assigned to an object with a constant probability p_a and the probability of independence of each pair of objects is p, then the mean value of the number of connectivity constrained concepts $E(|\mathfrak{B}_2|)$ is given by*

$$E\big(|\mathfrak{B}_2|\big) = 1 + \sum_{k=0}^{n-1} \sum_{Y \subseteq A} \binom{n}{k} R(k) P_Y(k). \tag{8}$$

In order to obtain at least some minimal idea concerning the behavior of the expected number of concepts, we computed some of these values.

For the computations, since the standard techniques are insufficient to calculate large binomial coefficients as well as high powers of fractions, we used the arbitrary precision calculator language bc (bench calculator) with

19

		$p = 0.01$		$p = 0.1$	
n	$E(|\mathfrak{B}|)$	$E(|\mathfrak{B}_1|)$	$E(|\mathfrak{B}_2|)$	$E(|\mathfrak{B}_1|)$	$E(|\mathfrak{B}_2|)$
10	542.71	12.45	12.53	16.62	50.70
20	12572.85	23.90	24.92	42.12	4403.27
50	660373.54	64.24	116.19	193.19	638078.43
100	11027045.65	151.64	16972.57	745.80	11025031.86
200	160274501.06	402.07	40301813.46	3336.20	160274500.24
300	723254395.07	753.96	485319530.50	8504.18	723254395.07
400	2055489986.29	1207.90	1798190757.95	16899.07	2055489986.29
500	4558230784.22	1764.43	4347302593.59	29113.79	4558230784.22

Table 2: Expected number of concepts for $p_0 = 0.5$, $m = 100$, and various values of n and p.

the precision of 3000 decimals. For the formulae (6) and (7) we used recurrence relations between consecutive terms of the sums to facilitate the evaluation.

Let us note, that for computed values probabilities for attribute distributions are equal each other, i.e., $p_a = p_0$ for all $a \in A$. In Table 2 the calculated expectations for different combinations of values of parameters p, p_0, m and n are described.

The results are in a good match with the theoretical predictions. If the probability p is big enough compared to $1/n$, the expectation of the number of connectivity-constrained concepts is almost the same as for all concepts. This is due to the fact that if the number of edges in a random graph $G(n, p)$ exceeds $n/2$, it is very probable to find a single big component containing almost all vertices [17], hence, it is very probable for a concept to belong to this component, and thus to be connected-closed. Therefore, in this case, restriction by cliques is a more powerful way for concept reduction.

On the other hand, if the probability p is small enough, it is rather improbable to find a clique different from singletons and edges in the graph, hence, the expectation of the number of clique-constrained concepts drops dramatically compared to the number of all concepts. However, the connected components are still unbounded in size. It turns out that in this case the connectedness provides a more refined tool for concept reduction.

Example 4.4. Continuing in the framework described in Example 3.5, suppose that the data of 500 users are to be analyzed. In particular, as an

input a formal context (O, A, I) will be available, with $n = |O| = 500$, $m = |A| = 100$. Obviously, the detailed information about the structure of the incidence relation I is not available yet. Since the relatively large data collection will be available, we can assume a random distribution of the considered attributes. In order to use the results presented in this section, we have to estimate the parameters p_a for $a \in A$ and the parameter p.

Assume that we have some previous experience with the behaviour of the analyzed population and that the formal context of Example 3.5 represents the "typical" example of a small context. As one can see, there are two attributes with 4 occurrences, two attributes with 2 occurrences and one attribute with 3 occurrences. As the average occurrence rate of an attribute is 3 objects out of 7, we may assume that in the analyzed context the attributes occur with probability $p_a = 3/7$. Note, that the need for an estimation of the parameters p_a may lead to less precise results, while knowledge of the exact values of these parameters could probably lead to more precise results.

Applying formula (3) on these parameters, we obtain the value $E\big(|\mathfrak{B}|\big) = 315338577.67$ for the expected number of concepts. For someone this number of concept can be unacceptable for analysis, also with respect to the time complexity of algorithms, since it is affected mainly by the number of concepts to be created. From this point of view, even with respect that the creation of one concept is done in a constant time C, the time complexity for the creation of the concept lattice will be $C \cdot E\big(|\mathfrak{B}|\big) = C \cdot 315338577.67$ of time units.

In order to obtain a concept reduction, one can consider a graph based model on users. Again, the parameter p denoting the probability of an edge occurrence in a graph should be given. It can represents some internal characteristic of the analyzed social network, or it could be estimated from available data. If the sample of the data is the graph from Example 3.5, which contains 6 edges out of 21 available, for our purposes we use the value $p = 2/7$. Using formula (5) we obtain $E\big(|\mathfrak{B}_1|\big) = 300982.26$, while formula (8) yields $E\big(|\mathfrak{B}_2|\big) = 315338577.67$. On the first sight, this result seems a little bit strange. The computation shows, that using the connected independence system based on the graph model with the described parameters yields practically the same result as for the concept lattice without this particular constraint (in fact these two values differ on some position behind the decimal separator). However this should be analyzed from the statistical point of view. As pointed out before, since $p \gg \frac{1}{n}$, a giant component will appear almost surely. Hence given two vertices, they belong to different components

21

with probability practically equal to zero. Therefore, restriction by connectedness does not decrease significantly the expected number of concepts, compared to the effect of the restriction to cliques.

In general, the values given by formulae (3), (5), (8) can be used to quantify the expected effect of concept reduction using different constrains before the analysis of the whole data collection, provided some sample data allowing to estimate the input parameters for the formulae are available.

5. Conclusion

In this paper a quantitative effect of concept reduction in graph-constrained concept lattices was studied from the probabilistic point of view. We derived the mean values of the number of concepts for classical concept lattices as well as for concept lattices constrained by clique and connected independent system respectively. As our computations suggest, a choice of a considered constraint together with the parameter p has strong quantitative impact on the resulting reduction.

Acknowledgments

References

[1] Ch. Aswani Kumar, S. Srinivas, Concept lattice reduction from fuzzy K-means clustering, Expert Systems with Applications 37(3), (2010), pp. 2696-2704.

[2] R. Belohlavek, V. Vychodil, Fuzzy concept lattices constrained by hedges, Journal of Advanced Computational Intelligence and Intelligent Informatics 11(6), (2007), pp. 536-545.

[3] R. Belohlavek, V. Vychodil. Formal concept analysis and linguistic hedges. Int. J. General Systems, 41(5), 2012, pp. 503-532.

[4] R. Belohlavek, V. Vychodil, Closure-based constraints in formal concept analysis, Discrete Applied Mathematics 161, (2013), pp. 1894-1911.

[5] J.A. Bondy, U.S.R. Murty, *Graph Theory*, Springer, (2008).

[6] B. Bollobás, *Random Graphs*, Cambridge University Press, (2001).

[7] A. Buzmakov, S.O. Kuznetsov, A. Napoli, Is concept stability a measure for pattern selection?, Information Technology and Quantitative Management 2014, Procedia Computer Science 31, (2014), pp. 918-927.

[8] F. Chen , K. Li, Detecting hierarchical structure of community members in social networks, Knowledge-Based Systems 87, (2015), pp. 3-15.

[9] P. Cordero , M. Enciso , A. Mora , M. Ojeda-Aciego, C. Rossi, Knowledge discovery in social networks by using a logic-based treatment of implications, Knowledge-Based Systems 87, (2015), pp. 16-25.

[10] M. E. Cornejo, J. Medina, E. Ramırez-Poussa, On the use of irreducible elements for reducing multi-adjoint concept lattices, Knowledge-Based Systems, (2015), doi:10.1016/j.knosys.2015.07.003

[11] S. Elloumi, J. Jaam, A. Hasnah,A. Jaoua, A multi-level conceptual data reduction approach based on the Lukasiewicz implication. Information Sciences 163(4), (2004), pp. 253-262.

[12] P. Erdős, A. Rényi, On the Evolution of Random Graphs, Magyar Tud. Akad. Mat. Kutató lnt. Közl. 5, (1960), pp. 17-61.

[13] B. Ganter, R. Wille, *Formal concept analysis. Mathematical foundations.* Springer, Berlin, (1999).

[14] E. N. Gilbert, Random graphs, The Annals of Mathematical Statistics 30(4), (1959), pp. 1141-1144.

[15] R. Godin, E. Saunders, J. Gecsei, Lattice Model of Browsable Data Spaces, Information Sciences 40, (1986), pp. 89-116.

[16] M.F. Janowitz, *Ordinal and Relational Clustering*, World Scientific Publishing, (2010).

[17] S. Janson, D.E. Knuth, T. Luczak, B. Pittel, The birth of the giant component, Random Structures and Algorithms 4 (3), (1993), pp. 233-358.

[18] J. Konecny, J. Medina, M. Ojeda-Aciego, Intensifying hedges and the size of multi-adjoint concept lattices with heterogeneous conjunctors, In Proc. of Concept Lattices and Applications CLA 2012, , (2012), pp. 245-256.

[19] J. Konecny, J. Medina, M. Ojeda-Aciego, Multi-adjoint concept lattices with heterogeneous conjunctors and hedges, Annals of Mathematics and Artificial Intelligence 72, (2014), pp. 73-89.

[20] S.O. Kuznetsov, On stability of a formal concept, Annals of Mathematics and Artificial Intelligence 49, (2007), pp. 101-115.

[21] J. Li, C. Mei, J. Wang, X. Zhang, Rule-preserved object compression in formal decision contexts using concept lattices, Knowledge-Based Systems 71, (2014), pp. 435-445.

[22] J. Medina, M. Ojeda-Aciego, Towards attribute reduction in multi-adjoint concept lattices. In Proc. of Concept Lattices and Applications CLA 2010, (2010), pp. 92-103.

[23] J. Pócs, On possible generalization of fuzzy concept lattices using dually isomorphic retracts. Information Sciences 210, (2012), pp. 89-98.

[24] M.-W. Shao, Y. Leung, Relations between granular reduct and dominance reduct in formal contexts, Knowledge-Based Systems 65, (2014), pp. 1-11.

[25] M.-W. Shao, H.-Z. Yang, W.-Z. Wu, Knowledge reduction in formal fuzzy contexts, Knowledge-Based Systems 73, (2015), pp. 265-275.

[26] P.K. Singh, A.K. Cherukuri, J. Li, Concepts reduction in formal concept analysis with fuzzy setting using Shannon entropy, International Journal of Machine Learning and Cybernetics (2014), DOI 10.1007/s13042-014-0313-6.

[27] Q. Wan , L. Wei, Approximate concepts acquisition based on formal contexts, Knowledge-Based Systems 75, (2015), pp. 78-86.

[28] L. Wei, J.J. Qi, Relation between concept lattice reduction and rough set reduction, Knowledge-Based Systems 23(8), (2010), pp. 934-938.