Size of random Galois lattices and number of frequent itemsets

Richard Emilion, Gerard Levy

Size of random Galois lattices and number of closed frequent itemsets.

Richard Emilion

MAPMO, Université d'Orléans, 45100 Orléans, France

Gérard Lévy

Université Paris Dauphine, 75016 Paris, France

Abstract

Given a sample of binary random vectors with i.i.d. Bernoulli(p) components, that is equal to 1 (resp. 0) with probability p (resp. $1-p$), we first establish a formula for the mean of the size of the random Galois lattice built from this sample, and a more complex one for its variance. Then, noticing that closed α-frequent itemsets are in bijection with closed α-winning coalitions, we establish similar formulas for the mean and the variance of the number of closed α-frequent itemsets. This can be interesting for the study of the complexity of some data mining problems such as association rule mining, sequential pattern mining and classification.

Key words: association rule, Bernoulli distribution, classification, complexity, data mining, frequent itemset, Galois lattice, winning coalition.

1 Introduction

Extraction of hidden and useful information from large databases is nowadays of great interest in various application fields. This is the main purpose of data mining, a recent technology that provides tools which can answer questions that traditionally were too time consuming to resolve. An important component of data mining is rule induction, that is extraction of useful if-then rules from data, and a key step in this induction consists in mining what is usually called frequent itemsets (FI's) as introduced in the pioneering works of

Email addresses: richard.emilion@univ-orleans.fr (Richard Emilion), gerard.levy@dauphine.fr (Gérard Lévy).

Agrawal, Iemelinski, Srikant and Swamy [1], [2].
Although the study of this step has expanded considerably in the algorithmic aspects, the theory is still at its beginning. For example, it seems that there is not any result which provides estimations of the number of closed FI's when the dataset is generated by some standard probability distributions. Even so, it can be thought that such estimations are of interest for memory storage management, response time prediction and analysis of algorithms efficiency and complexity.
The present paper brings an answer when the dataset is generated by a sample of size m of n-dimensional binary random vectors with i.i.d. $Bernoulli(p)$ components, in other words, when dealing with a $m \times n$ binary matrix T with i.i.d. Bernoulli(p) entries.
Even though this model is quite simple, the computations, mainly that of the variance, are on the one hand rather non-trivial and, on the other hand, they give some indications on how to deal with more realistic and complex models.
Our method hinges on the elegant notion of Galois connection (GC) and Galois lattice (GL) built from T, using in an essential way, both Diday-Emilion formula of a GC [8] and the well-known elementary inclusion-exclusion formula of Poincaré.
We establish formulas which give the expectation and the variance of the size of a random GL and that of the number of closed FI's. This can be used to obtain confidence intervals for these numbers.
Note that some papers (see eg, [13], [19]) emphasize the influence of the density of 1 in the matrix T on the size of the GL and on some algorithms efficiency, but at our knowledge, no precise statement has been given and hence our result on GL's seems to be new. This can be of interest as GL's are popular in various applied domains such as rule mining [9], formalization of the notion of *concept* [21], learning and classification [13], bioinformatics [11], object-oriented programming [14], robotics [22], maketing [7], relational database [6] and so on.
The paper is organized as follows. Section 2 is concerned with notations and terminology on GC's and GL's. In Section 3, we consider random GC's and GL's and we state a simple but useful proposition in the case of i.i.d. Bernoulli(p) entries. In section 4 we establish the formula of the expectation of the size of a random GL and we present some simulation results in section 5. Section 6 is concerned with the more complex formula of the variance. In section 7 we show that closed FI's are in bijection with closed *winning coalitions*. This yields the mean and the variance of the number of closed FI's. We conclude in the last section by suggesting the extension of the method to more general distributions and more general descriptions.

2 Notations and terminology

Let $\mathcal{I} = \{1, ..., m\}$, any element $i \in \mathcal{I}$ representing an *object*. The lattice $(\mathcal{P}(\mathcal{I}), \subseteq, \cap, \cup)$ of all subsets of $\mathcal{I}$ will be denoted by $\mathcal{E}$. Let $\mathcal{J} = \{1, ..., n\}$, any element $j \in \mathcal{J}$ representing a *property*. The lattice $(\mathcal{P}(\mathcal{J}), \subseteq, \cup, \cap)$ of all subsets of $\mathcal{J}$ will be denoted by $\mathcal{F}$.
We are given a binary matrix T with m lines and n columns, the ith line being a binary vector $d(i) = (d_1(i), ..., d_j(i), ..., d_n(i))$ where $d_j(i) = 1$ (resp. 0) means that object $i \in \mathcal{I}$ has (resp. has not) property $j \in \mathcal{J}$. In data mining, where marketing terminology has been adopted, $d(i)$ is called a (customer) *transaction*, $j \in \mathcal{J}$ an *item* and any $F \in \mathcal{F}$ is called an *itemset* so that $d_j(i) = 1$ (resp.0) means that transaction $d(i)$ contains item j.

2.1 Intent, extent, binary Galois connection

The matrix T induces a binary relation $\mathcal{R}$ on $\mathcal{I} \times \mathcal{J}$ as follows: $i\mathcal{R}j$ iff $d_j(i) = 1$. For any non-emptyset $A \in \mathcal{E} = \mathcal{P}(\mathcal{I})$ let

$$f(A) = \{j \in \mathcal{J} : i\mathcal{R}j \text{ for all } i \in A\} \text{ and } f(\emptyset) = \mathcal{J} \tag{1}$$

be the the *intent* or the *description* of A, that is the set of properties satisfied by all objects of A. For any non-empty set $B \in \mathcal{F} = \mathcal{P}(\mathcal{J})$ let

$$g(B) = \{i \in \mathcal{I} : i\mathcal{R}j \text{ for all } j \in B\} \text{ and } g(\emptyset) = \mathcal{I} \tag{2}$$

be the *extent* of B, that is the set of objects satisfying all the properties given by B. The pair (f, g) is called a binary *Galois connection* (GC) between $\mathcal{E}$ and $\mathcal{F}$ as it satisfies the following properties:

$$f : \mathcal{E} \longrightarrow \mathcal{F} \text{ and } g : \mathcal{F} \longrightarrow \mathcal{E} \text{ are decreasing} \tag{3}$$

$$H = gof : \mathcal{E} \longrightarrow \mathcal{E} \text{ and } K = fog : \mathcal{F} \longrightarrow \mathcal{F} \text{ are extensive} \tag{4}$$

$$i.e.\ A \subseteq H(A) \text{ for any } A \in \mathcal{E} \text{ (resp. } B \subseteq K(B) \text{ for any } B \in \mathcal{F})$$

The notion of GC was early introduced by O. Ore [18], it is also mentionned in the book by G. Birkhoff [5] (chapter 5). A more elegant and tractable formula (see (7) and (8) below) will be used in our computations of f and g.
It is interesting to know that the name of Galois appears here because of the analogy with a fundamental result in the celebrated Galois theory on the one-to-one correpondance between intermediate fields of a field extension and subgroups of its Galois group (see eg, Stewart [20] page 114). Indeed a GC induces a one-to-one correspondance between closed (or invariant) elements of each lattice.

2.2 General Galois lattices

GC's can be defined for general lattices (Barbut and Monjardet [4], pages 13 and 25): given two general lattices $< \mathcal{E}, \leq, \vee, \wedge >$ and $< \mathcal{F}, \leq, \vee, \wedge >$, a GC between $\mathcal{E}$ and $\mathcal{F}$ is a pair (f, g) verifying properties (3) and (4), this last property meaning that:

$$X \leq H(X) \text{ and } Y \leq K(Y), \forall X \in \mathcal{E}, \forall Y \in \mathcal{F}.$$

These definitions imply that

$$foH = f,\ HoH = H,\ goK = g,\ KoK = K. \tag{5}$$

Let

$$I_H = \{X \in \mathcal{E} : H(X) = X\} \text{ (resp. } I_K = \{Y \in \mathcal{F} : K(Y) = Y\})$$

be the set of *closed* (or invariant) elements of $\mathcal{E}$ (resp. of $\mathcal{F}$). It can be seen that the restriction of f to I_H is a one-to-one mapping into I_K, its inverse being the restriction of g to I_K. The *Galois lattice* (GL) $\mathcal{G}$ induced by the GC (f, g) is defined as the set of nodes

$$\{(X, f(X)),\ X \in I_H\}$$

which has a lattice structure if $\leq, \vee$ and $\wedge$ are defined as follows:

$$(X, f(X)) \leq (X', f(X')) \text{ iff } X \leq X' \text{ and } f(X') \leq f(X)$$

$$(X, f(X)) \vee (X', f(X')) = (H(X \vee X'), f(X) \wedge f(X'))$$
$$(X, f(X)) \wedge (X', f(X')) = (X \wedge X', K(f(X) \vee f(X')))$$

It is easily seen that $\{(X, f(X)),\ X \in I_H\} = \{(g(Y), Y),\ Y \in I_K\}$ so that for finite GL's, the cardinality of G, say $L = \#\mathcal{G}$, satisfies

$$L = \#I_H = \#I_K \tag{6}$$

2.3 Explicit formulas for a general GC

Let $\mathcal{E} = \mathcal{P}(\mathcal{I})$. In most concrete situations, only the descriptions $d(i), i \in \mathcal{I}$, which belong to a *general* lattice $\mathcal{F}$, are given. These descriptions can be for example vector of real numbers, sets, functions, fuzzy sets, cumulative histograms, probability cumulative distribution functions and so on. A natural question to ask is the existence of a GC (f, g) such that $f(\{i\}) = d(i)$ with explicit fomulas generalizing formulas (1) and (2) of the binary case. The solution exists, and is unique if the GC is supposed maximal (that is not

dominated by a GC) and $\mathcal{F}$ has a greatest element denoted by $1_{\mathcal{F}}$:

Theorem (Diday - Emilion [8]) *There exists a unique maximal* GC (f, g) *between* $\mathcal{E} = \mathcal{P}(\mathcal{I})$ *and* $\mathcal{F}$ *verifying* $f(\{i\}) = d(i)$. *It is given by the formulas*:

$$f(X) = \wedge_{x \in X} d(x) \text{ for any non-empty } X \in \mathcal{E} \tag{7}$$

$$f(\emptyset) = 1_{\mathcal{F}}$$

$$g(Y) = \{i \in I : Y \leq d(i)\} \text{ for any } Y \in \mathcal{F} \tag{8}$$

Note that (7) and (8) imply

$$H(X) = g(f(X)) = \{i \in \mathcal{I} \ : f(X) \ \leq d(i)\} \text{ for any } X \in \mathcal{E} \tag{9}$$

and since $X \subseteq H(X)$ always holds we obtain a very crucial point:

Corollary

$$X \in I_H \ \Leftrightarrow \ (\nexists i \in \mathcal{I} \backslash X : f(X) \leq d(i)) \tag{10}$$

Finally also notice that

$$f(X \cup Y) = f(X) \wedge f(Y) \tag{11}$$

For sake of simplicity, $f(\{i\})$, which is equal to $d(i)$, will be denoted by $f(i)$. In the binary case, $\mathcal{F} = \mathcal{P}(\mathcal{J})$ is lattice isomorphic to $\{0,1\}^n$ where $\leq, \vee, \wedge$ are defined coordinatewisely. In particular, if $d(i) = (d_1(i), ..., d_j(i), ..., d_n(i)) \in \{0,1\}^n$, we let

$$f_j(X) = \wedge_{x \in X} d_j(x) \text{ for any non-empty } X \in \mathcal{E} \text{ and } f_j(\emptyset) = 1 \tag{12}$$

Remark 1 *Most of the above definitions and results can be extended when working with only a meet-semilattice* $(\mathcal{F}, \wedge)$.

3 Random Galois lattices

Now come back to the case $\mathcal{E} = \mathcal{P}(\mathcal{I})$ and $\mathcal{F} = \mathcal{P}(\mathcal{J})$ and working within a standard probabilistic and statistical framework that will be of great interest for huge tables as it is the case in data mining.
Let $(\Omega\ , \mathcal{B}, P)$ be a probability space and let $d(i), i \in \mathcal{I}$, be a sample of size m of a $n-$dimensional random binary vector $d : \Omega \longrightarrow \{0,1\}^n$. This means that the $d(i)'$s, $i \in \mathcal{I}$, are m i.i.d. (independent and identically distributed) random vectors, $d(i) : \Omega \longrightarrow \{0,1\}^n$, having the same probability distribution as d. We will assume below that the n components d_j of d are i.i.d. $Bernoulli(p)$

so that $T = (d_j(i))_{i\in\mathcal{I},j\in\mathcal{J}}$ is a random matrix with i.i.d. *Bernoulli*(p) entries, that is:

$$P(d_j(i) = 1) = p, P(d_j(i) = 0) = 1 - p.$$

Even though this model is quite simple, the computations in the next sections, mainly that of the variance, are rather non-trivial and, in addition, they give some indications for dealing with more complex models. This will be discussed in the conclusion Section.
If T has random entries then for any $X \in \mathcal{E}$, (resp. $F \in \mathcal{F}$) the description $f(X) = \wedge_{x\in X} d(x) : \Omega \longrightarrow \{0,1\}^n$, (resp. the extent $g(F)$) is a random binary vector (resp. a random subset of $\mathcal{I}$). This defines a random GC and a random GL $\mathcal{G}$ whose size L, that is the number of its nodes, is a random integer.
In this random setting, our aim is to estimate L (resp. the number of FI's) by first computing its mean and its variance. In the following proposition we list some properties of the random variable $f_j(X)$ that will be very useful in the coming computations. As usual events are mentioned within parenthesis, for example the event $(f_j(X) = 1)$ denotes the set $\{\omega \in \Omega : f_j(X)(\omega) = 1\}$.

Proposition 2 *Let T be a $m \times n$ matrix with i.i.d. Bernoulli(p) entries. If $X_1, ..., X_k \in \mathcal{E}$ are disjoint sets, then for any $j \in \mathcal{J}$, $f_j(X_1), ...,$ and $f_j(X_k)$ are independent. For any $X, Y \in \mathcal{E}$ we have*

$$P(f_j(X) = 1) = p^{\#X} \tag{13}$$

$$P(f_j(X) = 1, f_j(Y) = 1) = p^{\#(X\cup Y)} \tag{14}$$

$$P(f_j(X) = 0, f_j(Y) = 1) = p^{\#Y} - p^{\#(X\cup Y)} \tag{15}$$

$$P(f_j(X) = 0, f_j(Y) = 0) = 1 - p^{\#X} - p^{\#Y} + p^{\#(X\cup Y)} \tag{16}$$

Proof. By (12), $f_j(X_l) = \wedge_{i\in X_l} d_j(i)$, so that the independence of the rows of T implies that of $f_j(X_1), ..., f_j(X_l)$ for disjoint sets $X_1, ..., X_k$.
Since $f_j(X) = 1$ iff $\forall i \in X : d_j(i) = 1$, the independence of the Bernoulli r.v. implies (13).
By (11), both $f_j(X) = 1$ and $f_j(Y) = 1$ hold iff $f_j(X \cup Y) = 1$. Applying (13) to the set $X \cup Y$, we then obtain (14):

$$P(f_j(X) = 1, f_j(Y) = 1) = P(f_j(X \cup Y) = 1) = p^{\#(X\cup Y)}$$

Since the event $(f_j(Y) = 1)$ is the disjoint union of $(f_j(Y) = 1, f_j(X) = 1)$ and $(f_j(Y) = 1, f_j(X) = 0)$, we have

$$P(f_j(Y) = 1) = P(f_j(X) = 1, f_j(Y) = 1) + P(f_j(X) = 0, f_j(Y) = 1)$$

Equalities (13) and (14) then imply (15):

$$P(f_j(X) = 0, f_j(Y) = 1) = p^{\#Y} - p^{\#(X\cup Y)}$$

Since the event $(f_j(X) = 0)$ is the disjoint union of $(f_j(X) = 0, f_j(Y) = 1)$ and $(f_j(X) = 0, f_j(X) = 0)$, we have

$$P(f_j(X) = 0) = P(f_j(X) = 0, f_j(Y) = 1) + P(f_j(X) = 0, f_j(Y) = 0)$$

As $P(f_j(X) = 0) = 1 - P(f_j(X) = 1) = 1 - p^{\#X}$, equality (15) implies (16):

$$P(f_j(X) = 0, f_j(Y) = 0) = 1 - p^{\#X} - p^{\#Y} + p^{\#(X \cup Y)} \quad \square$$

4 Expectation of the size

For any $X \subseteq \mathcal{I}$ consider the probability $\pi(X)$ that X is a closed set:

$$\pi(X) = P(X \in I_H)$$

The following theorem evaluates $\pi(X)$ and the mean size of a random GL.

Theorem 3 *Let T be a $m \times n$ binary matrix with i.i.d. Bernoulli(p) random entries. For any $X \subseteq \mathcal{I}$ such that $\#X = k$ we have*

$$\pi(X) = \sum_{l=0}^{m-k} (-1)^l \binom{m-k}{l} [1 - p^k(1-p^l)]^n$$

and the mean $E(L)$ of the size L of the random Galois lattice built from T is given by:

$$E(L) = \sum_{k=0}^{m} \binom{m}{k} [\sum_{l=0}^{m-k} (-1)^l \binom{m-k}{l} (1 - p^k(1-p^l))^n]$$

Proof. Using (10) we have, for any $X \subseteq \mathcal{I}$ and any $i \in \mathcal{I}$:

$$\pi(X) = P(X \in I_H) = P(\nexists i \in \mathcal{I} \backslash X : f(X) \leq d(i))$$

Let define

$$A_{i,X} = \{\omega \in \Omega : f(X)(\omega) \leq d(i)(\omega)\} = \cap_{j \in \mathcal{J}} (f_j(X) \leq d_j(i)) \qquad (17)$$

and, if X^c denotes the complementary $\mathcal{I} \backslash X$, let

$$\rho(X) = \pi(X^c) = 1 - \pi(X)$$

We have $\rho(X) = P(\exists i \in X^c : f(X) \leq d(i)) = P(\cup_{i \in X^c} A_{i,X})$ so that the well-known inclusion-exclusion rule of Poincaré implies

$$\rho(X) = \sum_{\emptyset \neq R \subseteq X^c} (-1)^{\#R-1} P(\cap_{i \in R} A_{i,X}) \qquad (18)$$

Using (17), observe now that

$$\begin{aligned} P(\cap_{i\in R} A_{i,X}) &= P(\cap_{i\in R}\cap_{j\in\mathcal{J}}\,(f_j(X)\le d_j(i)) = P(\cap_{j\in\mathcal{J}}\cap_{i\in R}\,(f_j(X)\le d_j(i))) \\ &= [P(\cap_{i\in R}(f_1(X)\le d_1(i)))]^n \end{aligned}$$

since the columns of T are i.i.d..
Further, we have

$$\begin{aligned} P(\cap_{i\in R}(f_1(X)\le d_1(i))) &= P(\cap_{i\in R}(f_1(X)\le d_1(i), f_1(X)=0) \\ &\quad + P(\cap_{i\in R}(f_1(X)\le d_1(i), f_1(X)=1) \\ &= P(f_1(X)=0) + P(f_1(X)=1, \forall i\in R: d_1(i)=1) \\ &= P(f_1(X)=0) + P(f_1(X)=1, f_1(R)=1) \end{aligned}$$

and since X and R are disjoint, Proposition 2 yields

$$\begin{aligned} P(\cap_{i\in R}(f_1(X)\le d_1(i))) &= 1 - P(f_1(X)=1) + P(f_1(X)=1)P(f_1(R)=1) \\ &= 1 - p^{\#X} + p^{\#X}p^{\#R} = 1 - p^{\#X}(1-p^{\#R}) \end{aligned}$$

Hence

$$P(\cap_{i\in R} A_{i,X}) = [1 - p^{\#X}(1-p^{\#R})]^n \tag{19}$$

showing that $P(\cap_{i\in R} A_{i,X})$ only depends on the cardinality of the sets X and R.
If $\#X = k$, then $\#X^c = m-k$ and there are $\binom{m-k}{l}$ subsets R such that $\#R = l$. Thus (18) and (19) yield

$$\begin{aligned} \rho(X) &= \sum_{\emptyset\neq R\subseteq X^c} (-1)^{\#R-1}[1-p^{\#X}(1-p^{\#R})]^n \\ &= \sum_{l=1}^{m-k}(-1)^{l-1}\binom{m-k}{l}[1-p^k(1-p^l)]^n \end{aligned}$$

and

$$\pi(X) = 1-\rho(X) = \sum_{l=0}^{m-k}(-1)^l\binom{m-k}{l}[1-p^k(1-p^l)]^n \tag{20}$$

Finally, 1_A denoting the indicator function of event A, we have by (6),

$$L = \sum_{X\in\mathcal{P}(\mathcal{I})} 1_{X\in I_H}$$

and since there are $\binom{m}{k}$ subsets X of $\mathcal{I}$ such that $\#X = k$, we have by (20):

$$
\begin{aligned}
E(L) &= \sum_{X\in\mathcal{P}(\mathcal{I})} E(1_{X\in I_H}) = \sum_{X\in\mathcal{P}(\mathcal{I})} P(X\in I_H) \\
&= \sum_{X\in\mathcal{P}(\mathcal{I})} \pi(X) = \sum_{k=0}^{m} \sum_{X\in\mathcal{P}(\mathcal{I}),\#X=k} \pi(X) \\
&= \sum_{k=0}^{m} \binom{m}{k} [\sum_{l=0}^{m-k} (-1)^l \binom{m-k}{l} (1-p^k(1-p^l))^n] \quad \square
\end{aligned}
$$

4.1 Symmetry

Since m and n clearly play a symmetric role, the above expression of $E(L)$ is symmetric w.r.t. m and n :

$$
E(L) = \sum_{l=0}^{n} \binom{n}{l} \sum_{k=0}^{n-l} (-1)^k \binom{n-l}{k} (1-p^l(1-p^k))^m
$$

This is of interest since, in practice, we often have much less items than transactions: $n \ll m$. Another consequence is that for fixed n and fixed $0 < p < 1 : \lim_{m\to\infty} E(L) = \sum_{l=0}^{n} \binom{n}{l} = 2^n$.
Note that the symmetry can be proved directly as mentionned to us by J.-P. Schreiber of University of Orléans:

$$
\begin{aligned}
(1-p^k(1-p^l))^n &= (1-p^k+p^{k+l})^n \\
&= \sum_{i=0}^{n} \sum_{j=0}^{n-i} \binom{n}{i} \binom{n-i}{j} (-p^k)^j (p^{k+l})^i \\
&= \sum_{i=0}^{n} \sum_{j=0}^{n-i} (-1)^j \binom{n}{i} \binom{n-i}{j} p^{kj} p^{ki} p^{li}
\end{aligned}
$$

by using twice the binomial formula. This implies that

$$
\begin{aligned}
E(L) &= \sum_{k=0}^{m} \sum_{l=0}^{m-k} \sum_{i=0}^{n} \sum_{j=0}^{n-i} (-1)^l (-1)^j \binom{m}{k} \binom{m-k}{l} \binom{n}{i} \binom{n-i}{j} p^{kj} p^{ki} p^{li} \\
&= \sum_{i=0}^{n} \sum_{j=0}^{n-i} (-1)^j \binom{n}{i} \binom{n-i}{j} \sum_{k=0}^{m} \sum_{l=0}^{m-k} (-1)^l p^{li} (p^{i+j})^k \\
&= \sum_{i=0}^{n} \sum_{j=0}^{n-i} (-1)^j \binom{n}{i} \binom{n-i}{j} (1-p^i-p^{i+j})^m \\
&= \sum_{l=0}^{n} \binom{n}{l} \sum_{k=0}^{n-l} (-1)^k \binom{n-l}{k} (1-p^l(1-p^k))^m
\end{aligned}
$$

5 Simulation experiments

In our simulation experiments, $m = n = 15$. For each $p = 0, 0.05, 0.1, 0.15, ..., 1$, 50 matrices with i.i.d. Bernoulli(p) entries are drawn. A fast algorithm [10] based on (7), (8) and (10) is performed to build the Galois lattices. As shown in [3], this algorithm outperforms well-known algorithms such as [12] and [19]. In figure 1 below, we see that the empirical mean of L, that is the average of L over the 50 simulations, is very closed to the theoretical mean stated in Theorem 1. Note that the mean number of closed sets is neither increasing with p nor symmetric wrt $\frac{1}{2}$, the maximum seems reached for p closed to $1-\frac{1}{n}$. Note that if $m = n$ and if all the entries of T are equal to 1 except those on the diagonal equal to 0, then the percentage of 1 is $1-\frac{1}{n}$ and $L = 2^n$ is maximum. Finally observe that the number of closed sets tends to 1 (resp. 2) as p tends to 1 (resp. to 0).

p	0	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95	1
Th. *E(L)*	2	10	17	24	33	45	60	80	106	140	186	245	323	421	538	661	750	723	489	136	1
Sim. *E(L)*	2	10	16	24	33	44	60	80	102	136	185	239	310	411	559	636	768	869	537	80	1

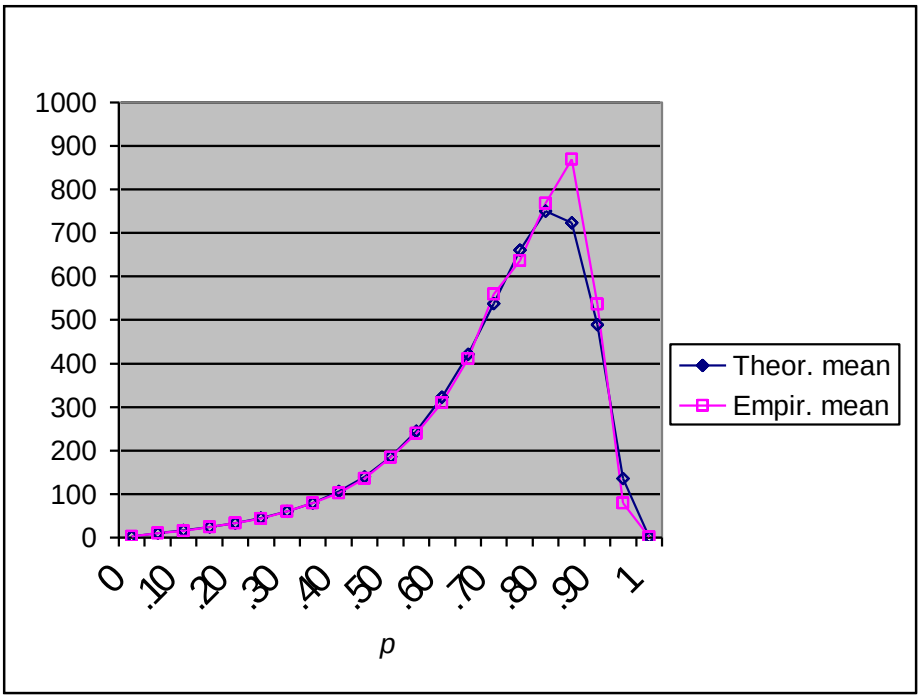

Fig. 1. Theoretical and simulated mean size of a Galois lattice

6 Variance of the size

The second-order moment and the variance of L are given by the following formulas which are more complex than the above first-order moment formula.

Theorem 4 *Let L be the size of a random Galois lattice $\mathcal{G}$ built from a $m \times n$ binary matrix with i.i.d. Bernoulli(p) random entries. Then*

$$E(L^2) = E(L) + \sum_{k=0}^{m} \sum_{l=0}^{m} \sum_{s=\max(0,k+l-m)}^{\min(k,l)} N_{k,l,s} Q_{k,l,s}$$

$$var(L) = E(L^2) - E(L)^2$$

where $N_{k,l,s}$ and $Q_{k,l,s}$ are defined below in (32) and (31) respectively and $E(L)$ is given by theorem 1.

Proof. For any $X, Y \subseteq \mathcal{I}$, $X \neq Y$, let us define

$$\pi(X,Y) = P(X \in I_H, Y \in I_H)$$

and

$$r(X,Y) = 1 - \pi(X,Y)$$

Using (10) and (17), we have

$$r(X,Y) = P(X \notin I_H \text{ or } Y \notin I_H) = P((\cup_{i \in X^c} A_{i,X}) \cup (\cup_{i' \in Y^c} A_{i',Y}))$$

In order to properly apply the inclusion-exclusion formula, introduce the set

$$U_{X,Y} = \{(i,X), i \in X^c\} \cup \{(i',Y), i' \in Y^c\}$$

Note that if $X \neq Y, \#X = k, \#Y = l$, then all the pairs appearing in the definition of $U_{X,Y}$ are distinct so that $\#U_{X,Y} = m - k + m - l$.
Let define the sets B_u and $B_{j,u}$, $u \in U_{X,Y}, j \in \mathcal{J}$, as follows:

$$\text{if } u = (i,X) \; B_u = A_{i,X} = (f(X) \leq d(i)) \text{ and } B_{j,u} = (f_j(X) \leq d_j(i))$$

$$\text{if } u = (i',Y) \; B_u = A_{i',Y} = (f(Y) \leq d(i')) \text{ and } B_{j,u} = (f_j(Y) \leq d_j(i'))$$

so that we have

$$r(X,Y) = P(\cup_{u \in U_{X,Y}} B_u)$$

Since $B_u = \cap_{j \in J} B_{j,u}$ for any $u \in U_{X,Y}$, the inclusion-exclusion rule of Poincaré yields

$$r(X,Y) = P(\cup_{u\in U_{X,Y}} B_u) = \sum_{\emptyset\neq U\subseteq U_{X,Y}} (-1)^{\#U-1} P(\cap_{u\in U} B_u)$$
$$= \sum_{\emptyset\neq U\subseteq U_{X,Y}} (-1)^{\#U-1} P(\cap_{u\in U} \cap_{j\in J} B_{j,u})$$
$$= \sum_{\emptyset\neq U\subseteq U_{X,Y}} (-1)^{\#U-1} P(\cap_{j\in J} \cap_{u\in U} B_{j,u})$$

Further, by the independence of the columns of T, we get

$$r(X,Y) = \sum_{\emptyset\neq U\subseteq U_{X,Y}} (-1)^{\#U-1} P(\cap_{u\in U} B_{1,u})^n \tag{21}$$

Suppose that

$$X \neq Y, \#X = k, \#Y = l, \#(X\cap Y) = s. \tag{22}$$

Let us compute $P(\cap_{u\in U} B_{1,u})$ by examining the four possible values of the pair $(f_1(X), f_1(Y))$:

$$\begin{array}{l} P\left(\cap_{u\in U} B_{1,u}\right) = \\ P\left(\cap_{u\in U} B_{1,u}, f_1(X)=0, f_1(Y)=0\right) + P(\cap_{u\in U} B_{1,u}, f_1(X)=1, f_1(Y)=0) + \\ P\left(\cap_{u\in U} B_{1,u}, f_1(X)=0, f_1(Y)=1\right) + P(\cap_{u\in U} B_{1,u}, f_1(X)=1, f_1(Y)=1) \end{array} \tag{23}$$

The definition of $B_{1,u}$ shows that the first term in (23) satisfies

$$P(\cap_{u\in U} B_{1,u}, f_1(X)=0, f_1(Y)=0) = P(f_1(X)=0, f_1(Y)=0)$$

Then applying (16), we get:

$$P(\cap_{u\in U} B_{1,u}, f_1(X)=0, f_1(Y)=0) = 1 - p^k - p^l + p^{k+l-s} \tag{24}$$

To evaluate the three other terms in (23), let define the following sets which depend on the set U:

$$R_1 = \{i \in Y\backslash X :\ (i,X)\in U\},\ R_1^{'} = \{i \in (X\cup Y)^c :\ (i,X)\in U\}$$

$$R_2 = \{i' \in X\backslash Y :\ (i,Y)\in U\},\ R_2^{'} = \{i' \in (X\cup Y)^c :\ (i,Y)\in U\}$$

Since $X \neq Y$, the cardinality of U, which is the number of distinct pairs (i,X), (i',Y), satisfies

$$\#U = \#R_1 + \#R_1^{'} + \#R_2 + \#R_2^{'}$$

However as the sets $R_1^{'}$ and $R_2^{'}$ need not be disjoint as required in Proposition 2, let us introduce the following disjoint sets

$$R_3 = R_1^{'}\backslash R_2^{'},\ R_4 = R_2^{'}\backslash R_1^{'},\ R_5 = R_1^{'}\cap R_2^{'}.$$

Then, observe that,

$$U = \#R_1 + \#R_3 + \#R_5 + \#R_2 + \#R_4 + \#R_5 = k_1 + k_2 + k_3 + k_4 + 2k_5$$

where k_v denotes $\#R_v$, $v = 1, 2, 3, 4, 5$.
Now, using again (11) and Proposition 2, we see that

$$\begin{aligned}
&P\left(\cap_{u\in U} B_{1,u}, f_1(X) = 1, f_1(Y) = 0\right)\\
&= P(f_1(X) = 1, f_1(Y) = 0, \forall i \in R_1 \cup R_3 \cup R_5 : d_1(i) = 1)\\
&= P(f_1(X) = 1, f_1(Y) = 0, f_1(R_1) = 1, f_1(R_3) = 1, f_1(R_5) = 1)\\
&= P(f_1(X) = 1, f_1((Y\backslash X)\backslash R_1) = 0, f_1(R_1) = 1, f_1(R_3) = 1, f_1(R_5) = 1)\\
&= P(f_1(X) = 1)P(f_1((Y\backslash X)\backslash R_1) = 0)P(f_1(R_1) = 1)P(f_1(R_3) = 1)P(f_1(R_5) = 1)\\
&= p^{\#X}(1 - p^{\#(Y\backslash X)\backslash R_1)})p^{k_1}p^{k_3}p^{k_5}\\
&= p^{\#X+k_1+k_3+k_5} - p^{\#(X\cup Y)+k_3+k_5}
\end{aligned}$$

and since $\#X = k$, $\#Y = l$, $\#(X \cup Y) = k + l - s$,

$$P(\cap_{u\in U} B_{1,u}, f_1(X) = 1, f_1(Y) = 0) = p^{k+k_1+k_3+k_5} - p^{k+l-s+k_3+k_5} \tag{25}$$

Interverting X and Y, the third term in (23) is given by

$$P(\cap_{u\in U} B_{1,u}, f_1(X) = 0, f_1(Y) = 1) = p^{l+k_2+k_4+k_5} - p^{k+l-s+k_4+k_5} \tag{26}$$

Now, observe that since $R_2 \subseteq X$ and $R_1 \subseteq Y$ we have $f_1(X) = 1 \Rightarrow f_1(R_2) = 1$ and $f_1(Y) = 1 \Rightarrow f_1(R_1) = 1$. So, using Proposition 2, the last term in (23) can be computed as follows:

$$\begin{aligned}
&P\left(\cap_{u\in U} B_{1,u}, f_1(X) = 1, f_1(Y) = 1\right) =\\
&P\left(f_1(X) = 1, f_1(Y) = 1, \forall i \in R_1 \cup R_2 \cup R_3 \cup R_4 \cup R_5 : d_1(i) = 1\right) =\\
&P\left(f_1(X \cup Y) = 1, f_1(R_3) = 1, f_1(R_4) = 1, f_1(R_5) = 1\right) =\\
&p^{\#(X\cup Y)}p^{k_3}p^{k_4}p^{k_5}
\end{aligned}$$

Hence

$$P(\cap_{u\in U} B_{1,u}, f_1(X) = 1, f_1(Y) = 1) = p^{k+l-s+k_3+k_4+k_5} \tag{27}$$

Adding the four evaluations (24), (26), (25) and (27) yields

$$P(\cap_{u\in U} B_{1,u}) = 1 - p^k - p^l + p^{k+l-s} + p^{k+k_1+k_3+k_5} - p^{k+l-s+k_3+k_5} +$$

$$p^{l+k_2+k_4+k_5} - p^{k+l-s+k_4+k_5} + p^{k+l-s+k_3+k_4+k_5}$$

that is

$$P(\cap_{u\in U} B_{1,u}) = Q_{k,l,s,k_1,k_2,k_3,k_4,k_5} \tag{28}$$

where

$$Q_{k,l,s,k_1,k_2,k_3,k_4,k_5} = 1 - p^k(1 - p^{k_1+k_3+k_5}) - p^l(1 - p^{k_2+k_4+k_5}) +$$

$$p^{k+l-s}(1 - p^{k_3+k_5} - p^{k_4+k_5} + p^{k_3+k_4+k_5})$$

(29)

This shows that if X, Y satisfy (22) then for any $U \subseteq U_{X,Y}$, the number $P(\cap_{u\in U} B_{1,u})$ only depends on the cardinality of the sets R_v.

Now, $k_1 = \#R_1 \leq \#(Y\backslash X) = l - s$ and the number of possible such sets R_1 is $\sum_{k_1=0}^{l-s} \binom{l-s}{k_1}$. Similarly $k_2 = \#R_2 \leq \#(X\backslash Y) = k - s$ and the number of possible such sets R_2 is $\sum_{k_2=0}^{m-s} \binom{k-s}{k_2}$. Moreover

$$0 \leq \#(R_3 \cup R_4 \cup R_5) = \#R_3 + \#R_4 + \#R_5 = k_3 + k_4 + k_5 \leq \#(X \cup Y)^c = m - k - l + s$$

and the number of possible 3-uples (R_3, R_4, R_5) such that $\#R_v = k_v$, $v = 3, 4, 5$ is equal to

$$\binom{m-k-l+s}{k_3}\binom{m-k-l+s-k_3}{k_4}\binom{m-k-l+s-k_3-k_4}{k_5}.$$

Thus, the number $c_{k,l,s,k_1,k_2,k_3,k_4,k_5}$ of possible 5-uples $(R_1, R_2, R_3, R_4, R_5)$ is

$$c_{k,l,s,k_1,k_2,k_3,k_4,k_5} =$$

$$\binom{l-s}{k_1}\binom{k-s}{k_2}\binom{m-k-l+s}{k_3}\binom{m-k-l+s-k_3}{k_4}\binom{m-k-l+s-k_3-k_4}{k_5}$$

(30)

Thus, the preceding formulas (21) and (28) show that if X, Y satisfy (22) then

$$\begin{aligned}\pi(X,Y) &= 1 - r(X,Y) = 1 - \sum_{\emptyset \neq U \subseteq U_{X,Y}} (-1)^{\#U-1} P(\cap_{u\in U} B_{1,u})^n \\ &= \sum_{U \subseteq U_{X,Y}} (-1)^{\#U} P(\cap_{u\in U} B_{1,u})^n = Q_{k,l,s}\end{aligned}$$

where

$$Q_{k,l,s} =$$

$$\sum_{k_1=0}^{l-s} \sum_{k_2=0}^{k-s} \sum_{0 \leq k_3+k_4+k_5 \leq m-k-l+s} (-1)^{k_1+k_2+k_3+k_4+2k_5} c_{k,l,s,k_1,k_2,k_3,k_4,k_5} Q^n_{k,l,s,k_1,k_2,k_3,k_4,k_5}$$

(31)

the coefficients $c_{k,l,s,k_1,k_2,k_3,k_4,k_5}$ and $Q_{k,l,s,k_1,k_2,k_3,k_4,k_5}$ being defined in (30) and (29) respectively.
Hence if X, Y satisfy (22), then the number $\pi(X, Y)$ only depends on k, l and

s. Moreover, observing that $X = Y$ if and only if $k = l = s$, we see that the number $N_{k,l,s}$ of ordered pairs (X, Y) that satisfy (22) is

$$N_{k,l,s} = 0 \text{ if } k = l = s, \text{ otherwise}$$
$$N_{k,l,s} = \binom{m}{s}\binom{m-s}{k-s}\binom{m-k}{l-s} = \frac{m!}{s!(k-s)!(l-s)!(m-k-l+s)!} \tag{32}$$

Now, by (10) we have

$$L^2 = \sum_{X \in P(I)} 1_{X \in I_H} \sum_{Y \in P(I)} 1_{Y \in I_H}$$

and hence, the second-order moment is equal to

$$\begin{aligned} E(L^2) &= \sum_{X,Y \in P(I)} E(1_{X \in I_H} 1_{Y \in I_H}) = \sum_{X,Y \in P(I)} P(X \in I_H, Y \in I_H) \\ &= \sum_{X,Y \in P(I), X = Y} P(X \in I_H, Y \in I_H) + \sum_{X,Y \in P(I), X \neq Y} P(X \in I_H, Y \in I_H) \\ &= \sum_{X \in P(I)} P(X \in I_H) + \sum_{X,Y \in P(I), X \neq Y} \pi(X, Y) \\ &= E(L) + \sum_{X,Y \in P(I), X \neq Y} \pi(X, Y) \end{aligned}$$

Noticing that if X, Y satisfy (22), then $\max(0, k + l - m) \leq s \leq \min(k, l)$ and grouping such ordered pairs (X, Y), we arrive finally at

$$E(L^2) = E(L) + \sum_{k=0}^{m} \sum_{l=0}^{m} \sum_{s=\max(0,k+l-m)}^{\min(k,l)} N_{k,l,s} Q_{k,l,s}$$

$$var(L) = E(L^2) - E(L)^2$$

where $E(L)$ is given by Theorem 1. $\square$

7 Frequent itemsets and winning coalitions

For any itemset $F \in \mathcal{F}$ let 1_F denotes the n-dimensional binary vector with all components equal to 0 except those at position j for all $j \in F$, which are equal to 1. Conversely to any n-dimensional binary vector $v = (v_1, ..., v_j, ..., v_n)$, is associated its support $F \in \mathcal{F}$ which is defined as $\{j \in \mathcal{J} : v_j = 1\}$. This

obviously defines a lattice isomorphism between $(\mathcal{F}, \subseteq, \cup, \cap)$ and $(\{0,1\}^n, \leq, \vee, \wedge)$, the order relation $\leq$, the infimum $\wedge$ (resp. the supremum $\vee$) being defined coordinatewisely. In particular, we see that transaction $d(i)$ contains itemset F iff $1_F \leq d(i)$. Let α be a real number such that $0 \leq \alpha \leq 1$. An itemset F is an *α-frequent itemset* (α-FI) if the proportion of transactions that contain F is greater than α, that is

$$\frac{\#\{i \in \mathcal{I} : 1_F \leq d(i)\}}{m} \geq \alpha$$

Since Diday-Emilion formula (8) of the extent g of a GC, yields $g(F) = g(1_F) = \{i \in \mathcal{I} : 1_F \leq d(i)\}$, we see that $F \in \mathcal{F}$ is an α-FI iff

$$\frac{\#g(F)}{m} \geq \alpha$$

More generally, given any arbitrary probability measure Q on the set $\mathcal{I}$ we can define F as a (Q, α)-FI iff

$$Q(g(F)) \geq \alpha$$

the previous definition corresponding to the special case of the uniform probability on $\mathcal{I}$, that is $Q(\{i\}) = \frac{1}{m}$ for any $i \in \mathcal{I}$.
For any $G : G \subseteq F$, we have $g(G) \supseteq g(F)$ since g is decreasing, so, any subset G of an α-FI is also an α-FI. As $\mathcal{J}$ is finite, any α-FI is contained in a maximal α-FI. It thus suffices to search all the maximal α-FI's to get all the α-FI's.
Now, let us call a subset $A \in \mathcal{E}$ a (Q, α)-*winning coalition*, shortly α-*WC*, if $Q(A) \geq \alpha$.
The following proposition shows that only (maximal) closed FI's are of interest and that they are in bijection with (minimal) closed α-WC.
This can be of interest in data mining since algorithms which provide minimal α-WC's have been widely studied in games theory (see a survey in [17]).

Proposition 5 *i) Let F be an itemset, then $g(K(F)) = g(F)$ so that F is an $\alpha-$FI iff $K(F)$ is an α-FI*
ii) If F is a maximal α-FI then F is K-closed
iii) The restriction of g to the closed α-FI's is a one-to-one mapping into the set of closed α-WC's, its inverse being the restriction of f to this set.
iv) The restriction of g to the maximal (K-closed) α-FI's is a one-to-one mapping into the set of minimal closed α-WC's, its inverse being the restriction of f to this set.

Proof. i) Since g is decreasing, $F \subseteq K(F)$ implies $g(K(F)) \subseteq g(F)$. On the other side $goK = gofog = Hog$ and since H is extensive, we have $g(F) \subseteq H(g(F)) = g(K(F))$. Hence $g(K(F)) = g(F)$ and F is a α-FI iff $K(F)$ is a α-FI.

ii) If F is a maximal α-FI, $K(F)$ is also a α-FI. As $F \subseteq K(F)$ and F is maximal, $K(F) = F$.
iii) Let F be a $K-$closed α-FI, then $H(g(F)) = g(K(F)) = g(F)$ shows that $g(F)$ is closed and, obviously, a α-WC by definition.
Conversely let $A \in \mathcal{E}$ be a closed α-WC. Then $K(f(A)) = (fog)(f(A)) = f(H(A)) = f(A)$ since $H(A) = A$. This shows that $f(A)$ is K-closed with $Q(g(f(A))) = Q(H(A)) = Q(A) \geq \alpha$. Hence $f(A)$ is a K-closed FI.
The one-to-one mapping and its inverse are now obvious since $f(g(F)) = F$ (resp. $g(f(A)) = A$) whenever F (resp. A) is K-closed (resp. H-closed).
iv) Let F be a $K-$closed α-FI which is maximal among the closed α-FI's. Then $g(F)$ is H-closed and $Q(g(F)) \geq \alpha$. Let $A \in \mathcal{E} : A \subseteq g(F)$ with A closed and $Q(A) \geq \alpha$, then we have $f(g(F)) = F \subseteq f(A)$ and $f(A) = F$ by the maximality of F, and thus $A = g(f(A)) = g(F)$, showing that $g(F)$ is a minimal closed α-WC.
Conversely let $A \in \mathcal{E}$ be a minimal closed α-WC. Then $f(A)$ is a K-closed α-FI. Let $B \in \mathcal{F} : f(A) \subseteq B$ where B is a closed FI. Then $g(B) \subseteq g(f(A)) = A$ implies by the minimality of A that $g(B) = A$ and thus $B = f(g(B)) = f(A)$. Hence $f(A)$ is a maximal α-FI.

We are at last in a position to prove our main result on closed FI's.

Theorem 6 *Let L_α be the number of K-closed α-FI's (resp. of H-closed α-WC's) induced by a $m \times n$ binary matrix with i.i.d. Bernoulli(p) random entries. Then the mean $E(L_\alpha)$ of L_α is given by:*

$$E(L_\alpha) = \sum_{k \geq m\alpha}^{m} \binom{m}{k} [\sum_{l=0}^{m-k} (-1)^l \binom{m-k}{l} (1 - p^k(1-p^l))^n].$$

Proof. By the preceding proposition, we have

$$\begin{aligned} L_\alpha &= \sum_{F \in \mathcal{P}(\mathcal{J}), \#g(F) \geq \alpha} 1_{F \in I_K} \\ &= \sum_{A \in \mathcal{P}(\mathcal{I}), \#A \geq m\alpha} 1_{A \in I_H}. \end{aligned}$$

Hence, equality (20) in the proof of theorem 1 yields

$$\begin{aligned} E(L_\alpha) &= \sum_{A \in \mathcal{P}(\mathcal{I}), \#A \geq m\alpha} P(A \in I_H) = \sum_{A \in \mathcal{P}(\mathcal{I}), \#A \geq m\alpha} \pi(A) \\ &= \sum_{k \geq m\alpha}^{m} \binom{m}{k} [\sum_{l=0}^{m-k} (-1)^l \binom{m-k}{l} (1 - p^k(1-p^l))^n] \quad \square \end{aligned}$$

Remark 7 *As $(L_\alpha)^2 = \sum_{X,Y \in \mathcal{P}(\mathcal{I}), Q(X), Q(Y) \geq \alpha} 1_{X \in I_H} 1_{Y \in I_H}$, the variance of*

$(L_\alpha)^2$ *can be computed as that of* L. *We omit the statement.*

8 Conclusion

We have presented a framework for computing the mean and the variance of the size of a random Galois lattice and that of the number of closed frequent itemsets in the case of a sample of binary random vectors with i.i.d. Bernoulli(p) components. The results hinges on Diday-Emilion formulation of a general Galois connection [8]. Even in this simple case, the computations, mainly that of the variance, are rather non-trivial.
It is easily seen that our proofs work staightforward for a sample of a binary vector d whose components d_j are independent Bernoulli(p_j), but not identically distributed. For example, (20) and formulas in Theorem 1 are to be replaced by

$$
\begin{aligned}
&P(f_j(X)=1) = p_j^{\#X} \\
&\pi(X) = \sum_{l=0}^{m-k} (-1)^l \binom{m-k}{l} \prod_{j=1}^{n} (1 - p_j^k(1-p_j^l)) \\
&E(L) = \sum_{k=0}^{m} \tbinom{m}{k} [\sum_{l=0}^{m-k} (-1)^l \tbinom{m-k}{l} \prod_{j=1}^{n} (1 - p_j^k(1-p_j^l))]
\end{aligned}
$$

On the other hand, the method could be studied in the case of non-independent components and non-independent rows if conditional probabilities are given. We think of interesting Markov models with very large state-space, namely the lattice $\mathcal{F}$.
Finally notice that the generalization in case of non-binary entries can be examined by using the above mentionned formulation since it holds for very general lattices $\mathcal{F}$.
All these points seem of interest to future research.

References

[1] R. Agrawal, T. Imielinski, A. Swamy, Mining association rules between sets of items in large databases, Proceed. ACM SIGMOD, Int'l Conf. on managment of data (1993), 207-216.

[2] R. Agrawal, R. Srikant, Fast algorithm for mining association, Proceed. 20th. Intl'. conf. VLDB (1994), 478-499

[3] F. Baklouti, G. Lévy, R. Emilion, A fast algorithm for Galois lattice building, Elec. J. Symb. Data Anal., 3 (2005), 19-31.

[4] M. Barbut, B. Monjardet, Ordre et classification, Hachette, Paris, 1970.

[5] G. Birkhoff, Lattice theory, AMS Colloq. Public. Vol. XXV, 1967.

[6] N. Caspard, B. montjardet, The lattice of closure systems, Disc. Appl. Math., 127 (2003), 241-269.

[7] M. Courtine, I. Bournaud, Building a pruned inheritance lattice structure for relational description , Workshop on concept lattice-baased, IICS (2001).

[8] E. Diday, R. Emilion, Maximal and stochastic Galois lattices, C. R. Acad. Sci. Paris, 325, I (1997), 261-266, and Disc. Applied Math., 27-2 (2003) 271-284.

[9] V. Duquenne, Contextual implications, Rapport CAMS, Paris, 1986.

[10] R. Emilion, G. Lambert, G. Lévy, Algorithms for Galois lattices, Indo-French Worksh., Univ. Paris IX-Dauphine, lise-ceremade (1997).

[11] S. Ferré, R. D. King, BLID: an Application of Logical Information Systems to Bioinformatics. In P. Eklund, editor, Int. Conf. Formal Concept Analysis, LNCS 2961, Springer (2004), 47-54, .

[12] B. Ganter, Two basic algorithms in Concept analysis, 831 (1984), Technische Hoschscschule Darmstad.

[13] R. Godin, R. Missaoui, H. Aloui, Learning algorithms using Galois lattice structure, Proc. of the IEEE Int. Conf. on Tools for AI, San Jose, CA (1991), 22-29 .

[14] R. Godin, H. Mili, Building and maintaining analysis-level class hierarchies using Galois Lattices, Proceed. eighth conf. on Object-oriented programming systems, languages, and applications, 1993, 394 - 410 .

[15] R. Godin, R. Missaoui, H. Aloui, Incremental concept formation algorithms based on Galois (concept) lattices, Comp. Intelligence, 11-2 (1991), 246-267.

[16] J.-L. Guigues, V. Duquenne, Famille minimale d'implications informatives d'un tableau de donnes binaires, Math. Sc. Humaines, 95 (1998), 5-8.

[17] T. Matsui, Y. Matsui, A survey of algorithms for calculating power indices of weighted majority games, Journal of the Operations Research Society of Japan 43 (2000).

[18] O. Ore, Galois connections, Trans. Amer. Math. Soc. 55 (1944), 494-513.

[19] S. Kuznetsov, S. Ob'edkov, Comparing the performance of algorithms for generating concept lattices. J. experim. theoret. art. int. 14 (2002), 189-216.

[20] J. Stewart, Galois theory, Chapman and Hall, 1975, New York.

[21] R. Wille, Restructuring lattice theory, Ordered sets I, Rival ed. 1980, Reider.

[22] E. Zenou, M. Samuelides, Galois lattice theory for probabilistic visual landmarks, J. Univ. Comp. Sci, 10 (8), (2004), 1014.