

Efficient Closed Pattern Mining in Strongly Accessible Set Systems

(extended abstract)

Mario Boley¹, Tamás Horváth¹, Axel Poigné¹, and Stefan Wrobel^{1,2}

¹ Fraunhofer IAIS, Schloss Birlinghoven, Sankt Augustin, Germany

² Dept. of Computer Science, University of Bonn, Germany

{mario.boyey,tamas.horvath,axel.poigne,stefan.wrobel}@iais.fraunhofer.de

Abstract. Many problems in data mining can be viewed as a special case of the problem of enumerating the closed elements of an independence system with respect to some specific closure operator. Motivated by real-world applications, e.g., in track mining, we consider a generalization of this problem to strongly accessible set systems and arbitrary closure operators. For this more general problem setting, the closed sets can be enumerated with polynomial delay if deciding membership in the set system and computing the closure operator can be solved in polynomial time. We discuss potential applications in graph mining.

1 Introduction

Over the past years, a large body of research has been devoted to finding efficient algorithms for the frequent itemset enumeration problem, and it has turned out that by looking at *closed* frequent itemsets, important gains can be made in the design of efficient algorithms (see, e.g., [6]). A closed frequent itemset is a frequent itemset that cannot be further enlarged without changing its support in the database. Unfortunately, similar results do not yet exist for (closed) pattern enumeration tasks in many of the more complex representations that are becoming increasingly popular due to applications in highly structured domains. Consider for example the task of finding closed frequent *connected* subgraphs of movements of people or cars in a street network given a database of GPS-based recordings of spatio-temporal movements (so called *tracks*) [5]. In mining such tracks instead of itemsets, some important properties that are true for the frequent itemset mining problem no longer hold. In particular, it is not true that all subpatterns of a frequent connected pattern must necessarily also be frequent connected, since subpatterns need not be connected.

Technically, for problems like the track mining problem mentioned above, we note that unlike for the simple frequent itemset case, where the underlying set system is an *independence system*, here we are dealing with a weaker property of set systems which is only *strongly accessible*. In this paper, we show that for this generalized problem, it is possible to design an algorithm that enumerates all

closed frequent patterns, for arbitrary closure operators, with polynomial delay (provided deciding membership in the set system and computing the closure operator can be done in polynomial time). To our knowledge, this result gives the first efficient closed pattern enumeration algorithm for this generalized and practically important task.

2 Preliminaries

In this section we define the notions and notations used in this paper. We will sometimes denote a set $\{a_1, \dots, a_n\}$ by the string $a_1 \dots a_n$.

Set Systems A *set system* is an ordered pair (E, \mathcal{F}) , where E is the ground set and $\mathcal{F} \subseteq 2^E$. A set system is called finite if its ground set is finite. An element X of \mathcal{F} is called *maximal* if there is no $Y \in \mathcal{F}$ such that X is a proper subset of Y . A set system (E, \mathcal{F}) with $\emptyset \in \mathcal{F}$ is called

- *accessible* if for all $X \in \mathcal{F} \setminus \{\emptyset\}$ there is an $e \in X$ such that $X \setminus \{e\} \in \mathcal{F}$,
- *strongly accessible* if for every $X, Y \in \mathcal{F}$ satisfying $X \subsetneq Y$, there is an $e \in Y \setminus X$ such that $X \cup \{e\} \in \mathcal{F}$, and
- an *independence system* if $Y \in \mathcal{F}$ and $X \subseteq Y$ implies $X \in \mathcal{F}$.

The definitions imply that (i) every independence system is strongly accessible and (ii) every finite strongly accessible set system is accessible. However, the converse of (i) and (ii) does not hold.

Closure Operators We now recall some notions related to closure operators. Let (E, \mathcal{F}) be a set system. A mapping $\rho : \mathcal{F} \rightarrow \mathcal{F}$ is called a *closure operator* if (i) $X \subseteq \rho(X)$ (*extensivity*), (ii) $X \subseteq Y \Rightarrow \rho(X) \subseteq \rho(Y)$ (*monotonicity*), and (iii) $\rho(X) = \rho(\rho(X))$ (*idempotence*) hold for all $X, Y \in \mathcal{F}$. A set $F \in \mathcal{F}$ satisfying $\rho(F) = F$ is called *ρ -closed*. The *family of ρ -closed sets* of a set system (E, \mathcal{F}) with respect to a closure operator ρ is denoted by $\rho(\mathcal{F})$. For a ρ -closed set $C \in \rho(\mathcal{F})$, the family of all ρ -closed *proper* subsets of C is denoted by $\lambda(C)$. Let (E, \mathcal{F}) be a set system with $\emptyset \in \mathcal{F}$. Then, because of the monotonicity, for every $C \in \rho(\mathcal{F})$ it holds that $\lambda(C) = \emptyset$ if and only if $C = \rho(\emptyset)$. A set $F \in \mathcal{F}$ is a *generator* of a ρ -closed set $C \in \rho(\mathcal{F})$ if $\rho(F) = C$.

Graphs An *undirected graph* is a pair (V, E) , where $V \neq \emptyset$ is a finite set of *vertices* and $E \subseteq \{e \subseteq V : |e| = 2\}$ is a set of *edges*. Unless otherwise stated, in this paper by graphs we always mean undirected graphs and denote the set of vertices and the set of edges of a graph G by $V(G)$ and $E(G)$, respectively. Let G and G' be graphs. G' is a *subgraph* of G , if $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$. Let $X \subseteq E(G)$ for some graph G . Then the graph *induced* by X , denoted $G[X]$, is the subgraph G' of G such that $V(G')$ is the set of vertices occurring in X and $E(G') = X$. A graph G is *connected* if for every $u, v \in V(G)$, there is a sequence w_0, w_1, \dots, w_ℓ of vertices such that $w_0 = u$, $w_\ell = v$, and $\{w_i, w_{i+1}\} \in E(G)$ for every i ($0 \leq i < \ell$). A *connected component* of G is a maximal subgraph of G that is connected. For the sets $Y \subseteq X \subseteq E(G)$ of edges such that $G[Y]$ is connected, $\text{CONN}_X(Y)$ denotes the connected component of $G[Y]$ in $G[X]$.

Frequent Pattern Mining We recall some notions from frequent pattern mining. Let \mathcal{D} be a transaction database over a set I , i.e., \mathcal{D} is a multiset of subsets of I . For a set $X \subseteq I$, the *support* of X with respect to \mathcal{D} , denoted $\mathcal{D}[X]$, is the multiset of transactions of \mathcal{D} containing X . The *support count* of X , denoted $\sigma(X)$, is defined by $|\mathcal{D}[X]|$. For an integer *frequency threshold* $t > 0$, a subset $X \subseteq I$ is *t-frequent* if $\sigma(X) \geq t$. The proof of the following proposition is immediate from the definitions.

Proposition 1 *Let \mathcal{D} be a transaction database over a set I and $X \subseteq Y \subseteq I$. Then $X \subseteq Y \Rightarrow \mathcal{D}[X] \supseteq \mathcal{D}[Y] \Rightarrow \bigcap \mathcal{D}[X] \subseteq \bigcap \mathcal{D}[Y]$.*

A subset $X \subseteq I$ is *closed* if for every $X \subsetneq Y \subseteq I$ it holds that $\mathcal{D}[Y] \subsetneq \mathcal{D}[X]$. We note that this notion of “closeness” does not relate necessarily to that used in the definition of closure operators. In Section 4 we present an example, where there is no closure operator defining the above notion of closeness.

For some enumeration problems, the size of the set to be enumerated can be exponential in the size of the input. In such cases, the algorithm cannot enumerate the output in time polynomial only in the size of the input. We consider enumeration with *polynomial delay* (see, e.g., [4]), i.e., the number of steps between the output of two successive elements is bounded by a polynomial in the input size.

3 The General Problem

Many problems in data mining (e.g., closed frequent itemset mining) can be considered as a special case of the following enumeration problem:

THE CLOSED SET MINING (CSM) PROBLEM: *Given a finite set E , a membership oracle $M_{\mathcal{F}} : 2^E \rightarrow \{0, 1\}$ defining a family $\mathcal{F} \subseteq 2^E$ satisfying $\emptyset \in \mathcal{F}$, and a closure operator $\rho : \mathcal{F} \rightarrow \mathcal{F}$, compute $\rho(\mathcal{F})$.*

As an instance of this problem, consider the closed frequent itemset mining problem. For this problem, E and \mathcal{F} correspond to the set of items and the family of frequent itemsets, respectively. Notice that \mathcal{F} is given implicitly by a frequency testing procedure denoted by $M_{\mathcal{F}}$ in the problem definition.

Usually, \mathcal{F} can be enumerated efficiently. Even then the naïve algorithm enumerating each set $S \in \mathcal{F}$ and testing whether S is ρ -closed is inefficient because $|\mathcal{F}|$ can be exponential in $|\rho(\mathcal{F})|$. There are several results on efficient enumeration of ρ -closed sets for the case that the underlying set system is finite and closed under intersection (see, e.g., [2, 3]). Among others, formal concept analysis [7] and closed frequent itemset mining (see, e.g., [6]) provide some representative applications of this case.

In contrast to these results, we do *not* require the set system to be closed under intersection. Instead, we consider finite set systems (E, \mathcal{F}) associated with closure operators $\rho : \mathcal{F} \rightarrow \mathcal{F}$ satisfying the following property: for any ρ -closed element of \mathcal{F} , there exists an *inductive generator*. An inductive generator of a ρ -closed element $C \in \rho(\mathcal{F})$ is an element $C' \cup \{e\} \in \mathcal{F}$ such that $C' \in \rho(\mathcal{F})$,

$e \in E \setminus C'$, and $C = \rho(C' \cup \{e\})$. These inductive generators can then be used to enumerate all ρ -closed sets with a DFS algorithm resulting in the following positive result:

Lemma 2 *The CSM problem can be solved with polynomial delay for instances satisfying (i) the membership oracle $M_{\mathcal{F}}$ and the closure operator ρ can be computed in polynomial time and (ii) for every ρ -closed set except $\rho(\emptyset)$, there exists an inductive generator.*

Proof (sketch). The ρ -closed sets can be enumerated by traversing the graph $(\rho(\mathcal{F}), X)$ with

$$X = \{(C, C') : C' = \rho(C \cup \{e\}) \text{ for some } e \in E \setminus C \text{ satisfying } C \cup \{e\} \in \mathcal{F}\}$$

in a depth first manner. Using prefix trees for the storage of the enumerated ρ -closed sets, one can decide in time linear in the size of a new ρ -closed set, whether it has already been visited. Condition (ii) implies that every ρ -closed set is reached when starting from $\rho(\emptyset)$. Since a new ρ -closed set is reached or the algorithm terminates after at most $|E|^2$ closure computations and membership queries, condition (i) implies polynomial delay. \square

To formulate our main result, we first give a sufficient and necessary condition for generators in arbitrary set systems and closure operators.

Proposition 3 *Let (E, \mathcal{F}) be a set system, $\rho : \mathcal{F} \rightarrow \mathcal{F}$ a closure operator, and $C \in \rho(\mathcal{F})$ a ρ -closed set. Then for every $F \in \mathcal{F}$ satisfying $F \subseteq C$ it holds that F is a generator of C if and only if there is no $C' \in \lambda(C)$ satisfying $F \subseteq C'$.*

Proof. (“if”) Since ρ is a closure operator we have $F \subseteq \rho(F) \subseteq \rho(C) = C$. Now consider the case that $\rho(F) \subsetneq C$. Then $\rho(F) \in \lambda(C)$ contradicting the assumption. Hence, $\rho(F) = C$ must hold.

(“only if”) Let $F \in \mathcal{F}$ be a generator of C and suppose for contradiction that $F \subseteq C'$ for some $C' \in \lambda(C)$. Then $F \subseteq C'$, but $\rho(F) = C \not\subseteq C' = \rho(C')$ contradicting the monotonicity of ρ . \square

Thus, if C' is a maximal element of $\lambda(C)$ for some ρ -closed set C then C' is not necessarily an *inductive* generator of C together with some element from $E \setminus C'$. For strongly accessible set systems, however, we have a different situation which allows us to state the main result of this section.

Theorem 4 *For any finite strongly accessible set system (E, \mathcal{F}) given by a polynomial membership oracle and for any polynomially computable closure operator $\rho : \mathcal{F} \rightarrow \mathcal{F}$, the family $\rho(\mathcal{F})$ of ρ -closed sets can be enumerated with polynomial delay.*

Proof. Let $C = \rho(F)$ be a ρ -closed set satisfying $C \neq \rho(\emptyset)$ (or equivalently, $\lambda(C) \neq \emptyset$). Then there is a *maximal* ρ -closed set $C' \in \lambda(C)$ because (E, \mathcal{F}) is finite. Since (E, \mathcal{F}) is strongly accessible, there is an $e \in E \setminus C'$ such that $C' \cup \{e\} \in \mathcal{F}$. The maximality of C' in $\lambda(C)$ implies that there is no set in $\lambda(C)$ containing $C' \cup \{e\}$. Thus, by Proposition 3, $C' \cup \{e\}$ is an inductive generator of C and the statement follows from Lemma 2. \square

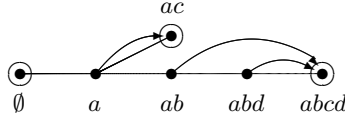


Fig. 1. An accessible set system with the closure operator $\rho(\emptyset) = \emptyset$, $\rho(a) = \rho(ac) = ac$, and $\rho(ab) = \rho(abd) = \rho(abcd) = abcd$. The ρ -closed set $abcd$ has no inductive generator.

We note that accessibility alone is not enough to guarantee the existence of inductive generators. In Figure 1 we give such an example.

4 Applications

As a first application, we state a positive result on *efficient* mining of closed frequent itemsets, which, although it is well-known (see, e.g., [1]), demonstrates the power of Theorem 4.

Theorem 5 *The family of closed frequent itemsets can be enumerated with polynomial delay.*

Proof. Let \mathcal{D} be a transaction database over a set I of items and $t > 0$ be an integer. Let $\mathcal{F}_{\mathcal{D},t}$ denote the family of t -frequent itemsets and $\rho : \mathcal{F}_{\mathcal{D},t} \rightarrow \mathcal{F}_{\mathcal{D},t}$ be the function $\rho : F \mapsto \bigcap \mathcal{D}[F]$. Clearly, ρ is a closure operator on $\mathcal{F}_{\mathcal{D},t}$.

Since any subset of a frequent itemset is also frequent, $(I, \mathcal{F}_{\mathcal{D},t})$ is an independence system and thus, it is strongly accessible. It can be decided in time polynomial in the size of \mathcal{D} whether a subset of I is frequent. Finally, for all $F \in \mathcal{F}_{\mathcal{D},t}$, $\rho(F)$ can be computed in time polynomial in the size of \mathcal{D} . The statement then follows by Theorem 4. \square

For the rest of this section, let $G = (V, E)$ be an undirected graph and \mathcal{D} be a database of subgraphs of G . That is, for every graph $G' \in \mathcal{D}$ we have that $G' = G[E']$ for some $E' \subseteq E$. Notice that the graphs in \mathcal{D} can be represented by subsets of E . We make use of this fact and consider \mathcal{D} as a transaction database over E . Transaction datasets of this type occur e.g. in *track mining* applications. Indeed, consider the application scenario, where we have a network represented by an undirected graph $G = (V, E)$ and points moving in the network within a time interval T (e.g., the network and the points could represent a city and persons moving in the city, respectively). For each point i , let $E_i \subseteq E$ be the set of edges of G visited by point i in T and let $G_i = G[E_i]$ be the subgraph of G induced by E_i . The collection of graphs G_i for every i forms the database \mathcal{D} . In contrast to other frequent subgraph mining problems defining the embedding operator by *subgraph isomorphism*, we define it by the *subset* relation.

Given some frequency threshold $t > 0$, one can consider different types of *t*-frequent patterns corresponding to appropriately chosen set systems and closure

operators defined by \mathcal{D} and t . As an example, for the set system $(E, \mathcal{F}_{\mathcal{D},t})$ with

$$\mathcal{F}_{\mathcal{D},t} = \{X \subseteq E : \sigma(X) \geq t\}$$

and closure operator $\rho : F \mapsto \bigcap \mathcal{D}[F]$, $\rho(\mathcal{F}_{\mathcal{D},t})$ corresponds to the family of *closed t -frequent subgraphs* of \mathcal{D} . Notice that in this case, the underlying set system $(E, \mathcal{F}_{\mathcal{D},t})$ is an independence system and the problem of enumerating $\rho(\mathcal{F}_{\mathcal{D},t})$ is equivalent to the closed frequent itemset mining problem. Hence, by Theorem 5, $\rho(\mathcal{F}_{\mathcal{D},t})$ can be enumerated with polynomial delay.

4.1 Mining Closed Frequent Connected Subgraphs

In the last example, we considered closed frequent subgraphs without any structural restriction, such as, for example, *connectivity*. Connectivity is perhaps the most natural structural property of graphs. In track mining (see, e.g., [5]) for instance, closed frequent connected subgraphs of a network can be considered as *homogeneous* connected subnetworks. From an algorithmic point of view, we consider the following problem:

CLOSED FREQUENT CONNECTED SUBGRAPH MINING PROBLEM: *Given an undirected graph $G = (V, E)$, a transaction database \mathcal{D} of subgraphs of G , and an integer $t > 0$, enumerate the family of closed t -frequent connected subgraphs of \mathcal{D} .*

Applying Theorem 4, we show that the above problem can be solved with polynomial delay. Let the set system $(E, \mathcal{F}_{\mathcal{D},t})$ and the function $\rho : \mathcal{F}_{\mathcal{D},t} \rightarrow \mathcal{F}_{\mathcal{D},t}$ be defined by

$$\mathcal{F}_{\mathcal{D},t} = \{X \subseteq E : \sigma(X) \geq t \wedge G[X] \text{ is connected}\} \quad (1)$$

and

$$\rho : X \mapsto \begin{cases} \emptyset & \text{if } X = \emptyset \\ \text{CONN}_{\bigcap \mathcal{D}[X]}(X) & \text{otherwise.} \end{cases} \quad (2)$$

That is, ρ maps a frequent connected subgraph X to the largest connected supergraph Y of X such that Y is a subgraph of each supergraph of X in \mathcal{D} . We note that the set system $(E, \mathcal{F}_{\mathcal{D},t})$ is *not* an independence system because a subgraph of a frequent connected graph is not necessarily connected. In fact, it is *not* even closed under intersection. But, as we note without proof due to space limitation, it is strongly accessible.

Lemma 6 *Let $G = (V, E)$ be an undirected graph, \mathcal{D} be a transaction database over E , and $t > 0$ be an integer. Let $\mathcal{F}_{\mathcal{D},t}$ and ρ be defined as in Equations (1) and (2), respectively. Then $(E, \mathcal{F}_{\mathcal{D},t})$ is a strongly accessible set system and ρ is a closure operator.*

One can easily see that deciding the membership in $\mathcal{F}_{\mathcal{D},t}$ and computing ρ can both be solved efficiently. Combining these properties with the results of Lemma 6, we can apply Theorem 4 and state the main result of this subsection:

Theorem 7 *The closed frequent connected subgraph mining problem can be solved with polynomial delay.*

4.2 Closed Frequent Subpath Mining

The setting defined in this paper actually goes beyond the standard definition of “closeness” usually employed in data mining, as it can be used to resolve an anomaly with this notion described as follows. In the standard case as discussed above, the design of a closure operator was straightforward. Since for all $X \in \mathcal{F}_{\mathcal{D},t}$, the family

$$\max\{X' \in \mathcal{F}_{\mathcal{D},t} : X \subseteq X' \subseteq \bigcap \mathcal{D}[X]\}$$

of *maximal* sets had always exactly one element, the closure operator could just assign this unique maximum element to X . However, it does not hold in general that such a unique maximum element exists. In this subsection, as an illustrative example of how to define a closure operator for such cases, we consider the problem of mining closed frequent *paths*.

We again have a graph $G = (V, E)$ and a database \mathcal{D} of transactions containing subsets of E . The set system of interest is now $(E, \mathcal{F}_{\mathcal{D},t})$ with

$$\mathcal{F}_{\mathcal{D},t} = \{P \subseteq E : \sigma(P) \geq t \wedge P \text{ is a path}\} .$$

As in the previous case, $\mathcal{F}_{\mathcal{D},t}$ is not an independence system and also not closed under intersection, but it is strongly accessible. Notice that the membership problem can be solved efficiently.

In data mining, a common definition for closed frequent subpaths is given by “A path P is *closed frequent* if it is frequent and $\mathcal{D}[P'] \subsetneq \mathcal{D}[P]$ for every path P' containing P .” Using this definition, let $\mathcal{C}_{\mathcal{D},t}$ denote the set of closed frequent paths in \mathcal{D} . However, there is a problem with this definition: $\mathcal{C}_{\mathcal{D},t}$ is *not* induced by a closure operator. In fact, there are cases, for which there is no closure operator $\rho : \mathcal{F}_{\mathcal{D},t} \rightarrow \mathcal{F}_{\mathcal{D},t}$ with $\mathcal{C}_{\mathcal{D},t} = \rho(\mathcal{F}_{\mathcal{D},t})$. In the example below, we give such a case.

Example 8 Let $G = (\{1, 2, 3, 4\}, \{12, 23, 24\})$, $\mathcal{D} = \{\{12, 23, 24\}\}$ (i.e., \mathcal{D} is a singleton consisting of G), and $t = 1$. Then $\mathcal{C}_{\mathcal{D},1} = \{\{12, 23\}, \{12, 24\}, \{23, 24\}\}$. Consider the set system $(\{12, 23, 24\}, \mathcal{F}_{\mathcal{D},1})$, where $\mathcal{F}_{\mathcal{D},1}$ denotes the set of frequent paths, i.e., $\mathcal{F}_{\mathcal{D},1} = \{\emptyset, \{12\}, \{23\}, \{24\}, \{12, 23\}, \{12, 24\}, \{23, 24\}\}$. Assume that there is a closure operator ρ such that $\rho(\mathcal{F}_{\mathcal{D},1}) = \mathcal{C}_{\mathcal{D},1}$. Then $\rho(\{12\})$ must be either $\{12, 23\}$ or $\{12, 24\}$, say $\{12, 23\}$. But then $\rho(\{12\}) \not\subseteq \rho(\{12, 24\})$ contradicting the monotonicity.

To resolve this anomaly, we consider another natural notion of “closeness” which is induced by a closure operator. As stated above, we have the problem that a path X can be contained in more than one maximal path in $\bigcap \mathcal{D}[X]$. This prevents a closure operator definition in a fashion similar to the connectivity case. A canonical way to overcome this problem is to define it as the intersection of all such maximal elements. For paths, in particular, this results in the definition

$$\rho : P \mapsto \bigcap \max\{M : M \text{ is a path in } \bigcap \mathcal{D}[P] \text{ such that } P \subseteq M\}$$

for every $P \in \mathcal{F}_{\mathcal{D},t}$.

One can show that ρ is a closure operator that can be computed efficiently and thus, by Theorem 4, $\rho(\mathcal{F}_{\mathcal{D},t})$ can be enumerated with polynomial delay. Although $\rho(\mathcal{F}_{\mathcal{D},t})$ is a superset of $\mathcal{C}_{\mathcal{D},t}$, it still can reduce the output significantly and is a semantically more meaningful set of patterns than $\mathcal{F}_{\mathcal{D},t}$. As an example, let G be a path of length n and $\mathcal{D} = \{G\}$. For frequency threshold $t = 1$ we have $|\mathcal{F}_{\mathcal{D},1}| = (n+1)n/2$ and $|\rho(\mathcal{F}_{\mathcal{D},1})| = 1$.

5 Conclusion

In this paper, we have presented a positive result on efficient enumeration of the family of closed sets of strongly accessible set systems with respect to arbitrary closure operators. The significance of our result in the context of data mining is that most of the closed frequent pattern mining algorithms are restricted to the case that the underlying set system corresponding to the set of frequent patterns is an independence system or at least closed under intersection. Strongly accessible set systems, however, are not necessarily independence systems or closed under intersection. We have presented graph mining applications motivated by track mining, where the underlying set systems are strongly accessible, but not closed under intersection.

Although the applications of this paper have resorted to strongly accessible set systems, we note that the algorithm works also for set systems satisfying weaker requirements on accessibility. An interesting question is whether the positive result holds for accessible set systems as well.

References

1. E. Boros, V. Gurvich, L. Khachiyan, and K. Makino. On maximal frequent and minimal infrequent sets in binary matrices. *Annals of Mathematics and Artificial Intelligence*, 39(3):211–221, 2003.
2. B. Ganter and K. Reuter. Finding all closed sets: A general approach. *Order*, 8(3):283–280, 1991.
3. M. Habib, R. Medina, L. Nourine, and G. Steiner. Efficient algorithms on distributive lattices. *Discrete Applied Mathematics*, 110(2-3):169–187, 2001.
4. D. S. Johnson, M. Yannakakis, and C. H. Papadimitriou. On generating all maximal independent sets. *Information Processing Letters*, 27(3):119–123, 1988.
5. B. Kuijpers, M. Nanni, C. Körner, M. May, and D. Pedreschi. Spatio-temporal data mining. In F. Giannotti and D. Pedreschi, editors, *Geography, mobility, and privacy: a knowledge discovery vision*. (to appear)
6. N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal. Efficient mining of association rules using closed itemset lattices. *Information Systems*, 24(1):25–46, 1999.
7. R. Wille. Restructuring lattice theory: an approach based on hierarchies of concept. In I. Rival, editor, *Ordered Sets*, pp. 445–470. Reidel, Dordrecht/Boston, 1982.