HAL
open science

Pattern Setups and Their Completions

Aimene Belfodil, Sergei O. Kuznetsov, Mehdi Kaytoue

To cite this version:

Aimene Belfodil, Sergei O. Kuznetsov, Mehdi Kaytoue. Pattern Setups and Their Completions. The 14th International Conference on Concept Lattices and Their Applications (CLA 2018), Jun 2018, Olomouc, Czech Republic. pp.243-253. hal-01818740

HAL Id: hal-01818740

https://hal.science/hal-01818740v1

Submitted on 19 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pattern Setups and Their Completions

Aimene Belfodil[1,4], Sergei O. Kuznetsov[2,3], and Mehdi Kaytoue[1,5]

[1] Univ Lyon, INSA Lyon, CNRS, LIRIS UMR 5205, F-69621, LYON, France
[2] National Research University Higher School of Economics, Moscow, Russia
[3] St. Petersburg Department of Steklov Mathematical Institute of Russian Academy of Sciences, Russia
[4] Mobile Devices Ingénierie, 100 Avenue Stalingrad, 94800, Villejuif, France
[5] Infologic, 99 avenue de Lyon, 26500 Bourg-Lès-Valence, France
firstname.surname@insa-lyon.fr, skuznetsov@hse.ru

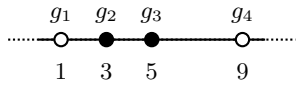
Abstract. Pattern mining consists in discovering interesting patterns in data. For that, algorithms rely on smart techniques for enumerating the pattern search space and, generally, focus on compressed collections of patterns (e.g. closed patterns), to avoid redundancy. Formal Concept Analysis offers a generic framework, called *pattern structures*, to formalize many types of patterns, such as closed itemsets, intervals, graph and sequence sets. Additionally, it provides generic algorithms to enumerate *all* closed patterns and *only* them. The only condition is that the pattern space is a meet-semilattice, which, unfortunately does not always hold (e.g., for sequential and graph patterns). In this paper, we discuss *pattern setups*, a tool that models pattern search spaces relying only on posets. Subsequently, we revisit some techniques transforming pattern setups to a pattern structure using set of patterns, namely *completions*, and we state *a necessary and sufficient condition* for a pattern setup completion using *antichains* to be a *pattern structure*.

1 Introduction

Pattern mining is the task of finding interesting patterns in a predefined search space. A generic tool for defining such a pattern search space is pattern structures [10,14] based on *Formal Concept Analysis (FCA)* [9]. Indeed, itemsets, interval [12], convex [3], partition [2] pattern spaces among others can be modeled within the pattern structure framework. However, since pattern structures rely on meet-semilattices, some pattern spaces that are only posets cannot be "directly" defined using such a tool.

Fig. 1 depicts a dataset of 4 objects described by attribute `"value"` and labeled positive or negative. Consider the task of finding "good" rules $d \to +$ in this dataset with d a description given by attribute `value`. Rather than considering the usual meet-semilattice of intervals [12]; descriptions d are restrained to intervals of the form $(v]$ and $[v)$ or singleton $\{v\} \subseteq \mathbb{R}$ (see Fig. 1 - right). Descriptions form a poset $(\mathcal{D}, \supseteq)$ but not a meet-semilattice. For instance, the set $\{\{3\}, \{5\}\}$ has not a meet, since lower bounds of $\{\{3\}, \{5\}\}$ has two maximal elements w.r.t. $\supseteq$ (*i.e.* [3) and (5]). Hence, the triple $(\mathcal{G}, (\mathcal{D}, \supseteq), \delta)$ with $\delta : g \mapsto \{\texttt{value}(g)\}$ does

object	value	label
g_1	1	$-$
g_2	3	$+$
g_3	5	$+$
g_4	9	$-$



Description language:

$\texttt{value} = v \mid v \in \mathbb{R}$

$\texttt{value} \leq v$

$\texttt{value} \geq v$

Fig. 1. Dataset (left), its representation in $\mathbb{R}$ with black dots representing positive objects (center) and the description language (right).

not form a pattern structure [10] since $\{\delta(g_2), \delta(g_3)\} = \{\{3\}, \{5\}\}$ has not a meet. It does form actually a **pattern setup** [16] which is based on a poset.

Description spaces like the one in Fig. 1 are numerous. For instance, sequential patterns [1] ordered by *"is subsequence of"* do not form a meet-semilattice [19] [7]. Sequential meet-semilattice in FCA [6] [7] refers usually to **set of sequences** rather than to the poset of **sequences**. Same holds for the graph meet-semilattice from [13]. In general, the base pattern setup is transformed to a pattern structure using sets of descriptions thus providing a richer space (language). Such transformations are naturally called *completions*. For example, in Fig. 1, restriction $\{\texttt{value} \geq 3, \texttt{value} \leq 5\}$ is equivalent to $3 \leq \texttt{value} \leq 5$ and does not belong to the base description language.

Understanding properties of pattern setups independently from their completions is fundamental for answering many practical questions. For instance, consider the question *"What are the best descriptions covering all positive instances?"*. If **better** stands for **more relevant than** as in relevance theory [11], the answer will be the two best incomparable rules $\texttt{value} \geq 3 \rightarrow +$ and $\texttt{value} \leq 5 \rightarrow +$ rather than only one in the completion $3 \leq \texttt{value} \leq 5 \rightarrow +$.

In this paper, after recalling basic facts on pattern structures in Section 2, we discuss the properties of a pattern setup in Section 3. Subsequently, Section 4 revisits pattern setup completions and states a necessary and sufficient condition on a pattern setup to augment it to a pattern structure using antichains of patterns [13,6,7]. We coin pattern setups verifying such a condition *pattern hyperstructures*. The basic order-theoretic concepts we use in this paper are well-presented in [18] and [9]. Table 1 resumes the used notations.

Table 1. Notations

$\wp(E)$	Powerset of E	$\{S \mid S \subseteq E\}$
$f[S]$	Image of S by f	$\{f(s) \mid s \in S\} \subseteq F$ with $S \subseteq E$ and $f : E \rightarrow F$
$(\mathcal{D}, \sqsubseteq)$	Poset	partially ordered set. Below $S \subseteq \mathcal{D}$ is a subset
$\downarrow S$	Down closure	$\{d \in \mathcal{D} \mid (\exists s \in S)\, d \sqsubseteq s\}$. For $s \in S : \downarrow \{s\} \triangleq \downarrow s$
$\uparrow S$	Up closure	$\{d \in \mathcal{D} \mid (\exists s \in S)\, s \sqsubseteq d\}$. For $s \in S : \uparrow \{s\} \triangleq \uparrow s$
S^ℓ	Lower bounds	$\{d \in \mathcal{D} \mid (\forall s \in S)\, d \sqsubseteq s\}$. For $s \in S : \{s\}^\ell \triangleq s^\ell$
S^u	Upper bounds	$\{d \in \mathcal{D} \mid (\forall s \in S)\, s \sqsubseteq d\}$. For $s \in S : \{s\}^u \triangleq s^u$
$min(S)$	Minimal elements	$\{s \in S \mid d \in S, d \sqsubseteq s \Rightarrow d = s\}$
$max(S)$	Maximal elements	$\{s \in S \mid d \in S, s \sqsubseteq d \Rightarrow d = s\}$
$\bigsqcap S$	Meet	greatest element of S^ℓ if exists ($d \in S^\ell \Leftrightarrow d \sqsubseteq \bigsqcap S$)
$\bigsqcup S$	Join	smallest element of S^u if exists ($d \in S^u \Leftrightarrow \bigsqcup S \sqsubseteq d$)

2 Basic Definitions

In general, pattern search spaces are formalized by partially ordered sets. In this paper, we call **description space (language)** or **pattern space (language)** any *poset* $\underline{\mathcal{D}} := (\mathcal{D}, \sqsubseteq)$. Elements of $\mathcal{D}$ are called *descriptions* or *patterns*. For any $c, d \in \mathcal{D}$, $c \sqsubseteq d$ is read "*c subsumes d*" or "*c* is *less restrictive than d*".

Pattern structures is an extension of the basic FCA model [10]. Objects in a pattern structure have descriptions in a meet-semilattice. **Pattern setups** [16] generalize pattern structures by demanding only a partial order on descriptions.

Definition 1. *A **pattern setup** is a triple* $\mathbb{P} = (\mathcal{G}, \underline{\mathcal{D}}, \delta)$ *where* $\mathcal{G}$ *is a* set *of objects,* $\underline{\mathcal{D}}$ *is a description space and* $\delta : \mathcal{G} \to \mathcal{D}$ *is a mapping that takes each object* $g \in \mathcal{G}$ *to a description* $\delta(g) \in \mathcal{D}$*. We say that an object* $g \in \mathcal{G}$ ***realizes*** *a description* $d \in \mathcal{D}$ *or* d ***covers*** g *if* $d \sqsubseteq \delta(g)$*. A pattern setup* $\mathbb{P}$ *is said to be a **pattern structure** if every subset of* $\delta[\mathcal{G}] = \{\delta(g) \mid g \in \mathcal{G}\}$ *has a **meet** in* $\mathcal{D}$*.*

Note that a necessary and sufficient condition on $\underline{\mathcal{D}}$ to have a pattern structure $(\mathcal{G}, \underline{\mathcal{D}}, \sqsubseteq)$ on any finite set of objects $\mathcal{G}$ and any mapping $\delta : \mathcal{G} \to \mathcal{D}$ is that $\underline{\mathcal{D}}$ is a **meet-semilattice** with a **top element** ($\mathcal{D}$ has its greatest element).

In **pattern structures**, two derivation operators map posets $(\wp(\mathcal{G}), \subseteq)$ and $(\mathcal{D}, \sqsubseteq)$ to each other. They are usually both denoted by $(\cdot)^{\square}$. For more clarity we use *ext* and *int* notation.

Definition 2. *The **extent operator** of a pattern setup, denoted by ext, takes each description* $d \in \mathcal{D}$ *to the set of objects in* $\mathcal{G}$ ***realizing*** *it. In the case the pattern setup is a* pattern structure*, the **intent operator**, denoted by int, takes a subset of objects* $A \subseteq \mathcal{G}$ *to the largest common description in* $\mathcal{D}$ ***covering*** *them, with* $\bigsqcap$ *representing the meet in* $\underline{\mathcal{D}}$*, we have:*

$$ext : \mathcal{D} \to \wp(\mathcal{G}), d \mapsto \{g \in \mathcal{G} \mid d \sqsubseteq \delta(g)\} \qquad int : \wp(\mathcal{G}) \to \mathcal{D}, A \mapsto \bigsqcap \delta[A]$$

The size $|ext(d)|$ *is called the **support** of* d *and is denoted by* $support(d)$*.*

In a *pattern structure* $\mathbb{P}$, the pair of operators (ext, int) forms a **Galois connection** between posets $(\wp(\mathcal{G}), \subseteq)$ and $(\mathcal{D}, \sqsubseteq)$. Thus, $ext \circ int$ and $int \circ ext$ form **closure operators** on the two posets and $(ext[\mathcal{D}], \subseteq)$ is a **$\bigcap$-structure** (*i.e.* a complete lattice). Moreover, another *complete lattice* isomorphic to poset $(ext[\mathcal{D}], \subseteq)$ is used. It is called **pattern concept lattice** and is denoted by $\underline{\mathfrak{B}}(\mathbb{P}) = (\mathfrak{B}(\mathbb{P}), \leq)$. Elements of $\underline{\mathfrak{B}}(\mathbb{P})$ are called **pattern concepts** and are of the form (A, d_A) with $A = ext(d_A)$ and $d_A = int(A)$. Pattern concepts are ordered as follows: $(A, d_A) \leq (B, d_B) \Leftrightarrow A \subseteq B \Leftrightarrow d_B \sqsubseteq d_A$.

3 Pattern Setups

Whereas basic FCA considers only binary data (descriptions are sets), pattern structures can handle more complex languages. Nevertheless, it fails at considering some types of patterns, which do not make a semi-lattice. Before diving into details, consider the following example.

$\mathcal{G}$	$\delta(\cdot)$
g_1	$c \to a \to b$
g_2	$c \to b \to b \to a$
g_3	a
g_4	$b \to b \to c$

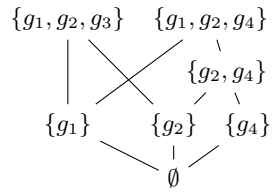


Fig. 2. The table (left) represents the pattern setup considered in example 1. The diagram (right) represents the poset $(ext[\mathcal{D}], \subseteq)$ (not making a closure system).

Example 1. Let be the dataset with the set of objects $\mathcal{G} = \{g_1, g_2, g_3, g_4\}$ in Fig. 2 (left). The description space $(\mathcal{D}, \sqsubseteq)$ contains all **non empty sequences** that can be built using items in $\{a, b, c\}$. It is ordered by the relationship **"is substring of"** (i.e. is subsequence of -without gaps-) denoted by $\sqsubseteq$. Such an order does not form a meet-semilattice. Indeed, consider sequences $\delta(g_1) = c \to a \to b$ and $\delta(g_2) = c \to b \to b \to a$; clearly, their common lower bounds $\{\delta(g_1), \delta(g_2)\}^\ell = \{a, b, c\}$ form an antichain and hence $\{\delta(g_1), \delta(g_2)\}$ has no meet. It follows that $(\mathcal{G}, \underline{\mathcal{D}}, \delta)$ is a *pattern setup* but not a *pattern structure*.

Important Remark. *Unless otherwise mentioned, in this paper* $\mathbb{P} = (\mathcal{G}, \underline{\mathcal{D}}, \delta)$ *denotes a* ***pattern setup*** *where* $\mathcal{G}$ *is a* ***non-empty finite set****. Hence, theorems and propositions in this paper are guaranteed to be valid only if* $\mathcal{G}$ *is* ***finite****.*

It is clear that the *extent operator* (see Definition 2) can be used in the general case of pattern setups since it requires only the order. However, the *intent* of a set may not exist. We define here the **cover operator**.

Definition 3. *We call* ***cover operator****, denoted by cov, the operator that takes each subset* $A \subseteq \mathcal{G}$ *to the set of common descriptions in* $\mathcal{D}$ ***covering*** *them:*

$$cov : \wp(\mathcal{G}) \to \wp(\mathcal{D}), A \mapsto \delta[A]^\ell = \{d \in \mathcal{D} \mid (\forall g \in A)\ d \sqsubseteq \delta(g)\}$$

Note that *ext* and *cov* are **order reversing mappings** in a pattern setup: $(\forall A, B \subseteq \mathcal{G}) A \subseteq B \Rightarrow cov(B) \subseteq cov(A)$ and $(\forall c, d \in \mathcal{D}) c \sqsubseteq d \Rightarrow ext(d) \subseteq ext(c)$.

Some sets of objects $A \subseteq \mathcal{G}$ can have no common descriptions covering them. In other words, $cov(A) = \emptyset$. Definition 4 develops a categorization of sets $A \subseteq \mathcal{G}$.

Definition 4. *Let* A *be a subset of* $\mathcal{G}$*.* A *is said to be an* ***extent*** *iff* $\exists d \in \mathcal{D} \mid A = ext(d)$*.* A *is said to be* ***coverable*** *iff* $\exists d \in \mathcal{D} : A \subseteq ext(d)$*, otherwise it is said to be* ***non coverable****. The set of all possible* ***extents*** *is denoted by* $\mathbb{P}_{ext}$ *and given by* $\mathbb{P}_{ext} = ext[\mathcal{D}]$*. The set of* ***coverable sets*** *is given by* $\downarrow \mathbb{P}_{ext}$*.*

Example 2. Consider Fig. 2. Set $A = \{g_2, g_4\}$ has cover $cov(A) = \{b, b \to b, c\}$. Moreover, A is an **extent** since $ext(b \to b) = A$. Set $B = \{g_3\}$ is **not an extent** since the only covering description is "a" for which $ext(a) = \{g_1, g_2, g_3\}$ (B is **coverable**). Set $C = \{g_3, g_4\}$ is **non coverable**, since g_3 and g_4 do not have common symbols.

Remark 1. As in **pattern structures** [10], **pattern implication** and **object implication** can be defined thanks to extent and cover operators, respectively. For $c, d \in \mathcal{D}$, the **pattern implication** $c \to d$ holds if $ext(c) \subseteq ext(d)$. Conversely, for $A, B \subseteq \mathcal{G}$, the **object implication** $A \to B$ holds if $cov(A) \subseteq cov(B)$.

3.1 On Maximal Descriptions in a Cover

In pattern structures, **intent** operator associates the largest common description to a set of objects. However, **cover** operator takes a set of objects to the set of all descriptions covering them. Such a set can be huge and needs to be "compressed". Since we have no guarantees of the existence of the **largest** common description, a reasonable suggestion would be to consider the **maximal** ones.

Definition 5. *The set of **maximal descriptions** covering a subset $A \subseteq \mathcal{G}$, denoted by $cov^{\mathcal{I}}(A)$, is given by:*

$$cov^{\mathcal{I}}(A) = max(cov(A)) = max(\delta[A]^{\ell})$$

Remark 2. Maximal descriptions are sometimes called **support-closed descriptions** [5]. A description d is said to be *support-closed* if $\forall c \in \mathcal{D}$ such that for $d \sqsubseteq c$ and $c \neq d$ we have $support(c) < support(d)$ (*i.e.* $ext(c) \subsetneq ext(d)$).

3.2 On Upper Approximations of a Set

In pattern structures a description $d \in \mathcal{D}$ **covering** the set A will **certainly cover** $ext(int(A))$. This observation is important when it comes to search for positive or negative hypotheses in a labeled dataset [13] or in other words classification association rules [15]. Indeed, if $\mathcal{G}^+$ represents the whole set of positive objects in the dataset $\mathcal{G}$ and if we want exactly one rule which covers all positive instances, the best one (in terms of relevance [11]) will be the rule $int(\mathcal{G}^+) \rightarrow +$ with confidence $|\mathcal{G}^+|/|ext(int(\mathcal{G}^+))|$. For pattern structures $ext(int(A))$ is the closure of A, in Rough Set Theory [17], $ext(int(A))$ can be seen as the **upper approximation** of A in $\mathcal{G}$. To have a similar notion for pattern setups, we define what we call **upper-approximation extents**.

Definition 6. *The **set of upper-approximation extents** of a subset $A \subseteq \mathcal{G}$, denoted by $\overline{A}$, is given by: $\overline{A} = min(\{E \in ext[\mathcal{D}] \mid A \subseteq E\}) = min(\uparrow A \cap ext[\mathcal{D}])$.*

Example 3. Extent $A = \{g_2, g_4\}$ in Fig. 2 has a unique upper-approximation (i.e. $\overline{A} = \{\{g_2, g_4\}\}$). Coverable subset $B = \{g_1, g_2\}$ has multiple upper approximations $\overline{B} = \{\{g_1, g_2, g_3\}, \{g_1, g_2, g_4\}\}$. Non coverable subset $C = \{g_3, g_4\}$ has no upper approximations (i.e. $\overline{C} = \emptyset$). Proposition 1 formalizes these observations.

Proposition 1. *For any $A \subseteq \mathcal{G}$, we have:*

$$A \in \mathbb{P}_{ext} \Leftrightarrow \overline{A} = \{A\} \ \textbf{(1)} \ A \in \downarrow \mathbb{P}_{ext} \Leftrightarrow \overline{A} \neq \emptyset \ \textbf{(2)} \ \overline{A} = min(ext[cov(A)]) \ \textbf{(3)}$$

Proof. We demonstrate each property independently:

1. $A \in ext[\mathcal{D}] \Leftrightarrow \overline{A} = \{A\}$: For ($\Leftarrow$), $\overline{A} \subseteq ext[\mathcal{D}]$, thus $A \in ext[\mathcal{D}]$. For ($\Rightarrow$), we have $A \in \uparrow A \cap ext[\mathcal{D}]$ thus $\overline{A} = min(\uparrow A \cap ext[\mathcal{D}]) = \{A\}$.
2. $A \in \downarrow ext[\mathcal{D}] \Leftrightarrow \overline{A} \neq \emptyset$: For ($\Rightarrow$), we have $A \in \downarrow ext[\mathcal{D}]$, that is $\exists B \in ext[\mathcal{D}]$ s.t. $A \subseteq B$ (*i.e* $B \in \uparrow A$). Thus $\uparrow A \cap ext[\mathcal{D}]$ is not empty and so does $\overline{A}$ (since $\uparrow A \cap ext[\mathcal{D}]$ is finite). For ($\Leftarrow$), we have $\overline{A} \neq \emptyset$, that is $\uparrow A \cap ext[\mathcal{D}] \neq \emptyset$ thus $\exists B \in ext[\mathcal{D}]$ such that $A \subseteq B$ thus $A \in \downarrow ext[\mathcal{D}] = \downarrow \mathbb{P}_{ext}$.

3. We show $ext[cov(A)] = \uparrow A \cap ext[\mathcal{D}]$. Let $B \subseteq \mathcal{G}$. We have $B \in ext[cov(A)]$ $\Leftrightarrow \exists d \in cov(A)$ s.t. $B = ext(d) \Leftrightarrow \exists d \in \mathcal{D}\ \forall g \in A : d \sqsubseteq \delta(g) \Leftrightarrow \exists d \in \mathcal{D}$ $A \subseteq ext(d) = B \Leftrightarrow B \in ext[\mathcal{D}] \cap \uparrow A$. This concludes the proof. □

For a set $A \subseteq \mathcal{G}$ s.t. $\overline{A} \neq \emptyset$, any description covering A **contains at least** elements of $\bigcap \overline{A}$, the intersection of upper approximations of A.

Proposition 2. *$I : \wp(\mathcal{G}) \to \wp(\mathcal{G}), A \mapsto \bigcap \overline{A} = \bigcap \{E \in ext[\mathcal{D}] \mid A \subseteq E\}$ is a **closure operator** on $(\wp(\mathcal{G}), \subseteq)$. Moreover, we have $\overline{I(A)} = \overline{A}$ for any $A \subseteq \mathcal{G}$.*

Proof. In fact, this proposition is a direct application of the following Lemma.

Lemma 1. *Let $(P, \leq)$ be a complete lattice with $\bigwedge$ its meet and let $E \subseteq P$ be a subset. The mapping $\phi_E : P \to P, p \mapsto \phi_E(p) = \bigwedge \{e \in E \mid p \leq e\}$ is a **closure operator** on $(P, \leq)$ and $\phi_E[P]$ is a **meet-structure** in poset $(P, \leq)$.*

We prove Lemma 1. **(1)** ϕ_E **is extensive**. Trivially $p \in \{e \in E \mid p \leq e\}^\ell$ for $p \in P$. Since the meet is the greatest element, we conclude: $p \leq \phi_E(p)$. **(2)** ϕ_E **is monotone**. Let $p_1, p_2 \in P$ s.t. $p_1 \leq p_2$, we have $\{e \in E \mid p_2 \leq e\} \subseteq \{e \in E \mid p_1 \leq e\}$ thus $\{e \in E \mid p_1 \leq e\}^\ell \subseteq \{e \in E \mid p_2 \leq e\}^\ell$. We conclude that $\phi_E(p_1) \leq \phi_E(p_2)$. **(3)** ϕ_E **is idempotent**. Let us show $\{e \in E \mid \phi_E(p) \leq e\} = \{e \in E \mid p \leq e\}$. Inclusion $\subseteq$ is verified since $p \leq \phi_E(p)$. Inclusion $\supseteq$ comes from the definition since $\phi_E(p)$ is a lower bound of $\{e \in E \mid p \leq e\}$, then for any $e \in E$ such that $p \leq e$ we have $\phi_E(p) \leq e$. The idempotence is straightforward.

Proposition 2 is a corollary of Lemma 1 since $(\wp(\mathcal{G}), \subseteq)$ is a complete lattice in which the meet is the set intersection. Indeed, $I \triangleq \phi_{ext[\mathcal{D}]}$ is a closure operator on $(\wp(\mathcal{G}), \subseteq)$. $\overline{I(A)} = \overline{A}$ comes directly from the previous idempotence proof. □

Proposition 3. *For any $g \in \mathcal{G}$ we have $\overline{\{g\}} = \{ext(\delta(g))\}$. That is, any singleton set $\{g\} \subseteq \mathcal{G}$ is coverable and has a **unique upper-approximation**.*

Proof. Let $g \in \mathcal{G}$, we have $\delta(g) \in cov(\{g\})$, thus $ext(\delta(g)) \in ext[cov(\{g\})]$. Let us show that $ext(\delta(g))$ is a lower bound of $ext[cov(\{g\})]$. We have by definition: $cov(\{g\}) = \{d \in \mathcal{D} \mid d \sqsubseteq \delta(g)\}$. Thus, $\forall d \in cov(\{g\}) : d \sqsubseteq \delta(g)$. Since ext is an order reversing operator, we obtain: $\forall A \in ext[cov(\{g\})] : ext(\delta(g)) \subseteq A$. Thus $ext(\delta(g))$ is the smallest element of $ext[cov(\{g\})]$. That is $\overline{\{g\}} = \{ext(\delta(g))\}$. □

3.3 Linking Upper Approximations and Maximal Descriptions

Now that we have both upper approximations and maximal covering descriptions, a judicious question would be:

Question 1. *What is the relationship between $cov^{\mathcal{I}}(A)$ and $\overline{A}$ for $A \subseteq \mathcal{G}$?*

Before diving into more details, consider the example below.

Example 4. Consider Fig. 2. **Extent** $A = \{g_2, g_4\}$ has $cov(A) = \{b, b \to b, c\}$. Hence, $cov^{\mathcal{I}}(A) = \{b \to b, c\}$. Therefore, $ext[cov^{\mathcal{I}}(A)] = \{\{g_2, g_4\}, \{g_1, g_2, g_4\}\}$. Besides, $\overline{A} = \{\{g_2, g_4\}\}$. Thus, counter-intuitively, $\overline{A}$ is not equal to $ext[cov^{\mathcal{I}}(A)]$.

Furthermore, when $\underline{\mathcal{D}}$ is an **infinite poset**, the set $cov^{\mathcal{I}}(A)$ might not "hold" all the information contained in $cov(A)$. That is to say: "knowing only maximal covering descriptions does not allow us to deduce the set of covering ones". Speaking formally, $cov^{\mathcal{I}}(A)$ does not necessarily verify $cov(A) = \downarrow cov^{\mathcal{I}}(A)$.

4 Pattern Hyper-Structures and Completions

There is a standard construction how an ordered set of descriptions can be turned in a semilattice of descriptions. For example, see [13], where a semilattice on sets of graphs with labeled vertices and edges is constructed from the order given by subgraph isomorphism relation. Let $\underline{\mathcal{D}} = (\mathcal{D}, \sqsubseteq)$ be a poset and let $\mathcal{A}(\mathcal{D})$ be the set of its antichains. It is possible to define a *partial order* by letting $S_1, S_2 \in \mathcal{A}(\mathcal{D})$ [4]: $S_1 \leq S_2$ iff $\forall s_1 \in S_1\ \exists s_2 \in S_2 : s_1 \sqsubseteq s_2$ (i.e. $S_1 \subseteq \downarrow S_2$). Please note that this relation does not define an order on $\wp(\mathcal{D})$. It does define only a **pre-order** since *anti-symmetry* does not hold (see [8]).

In fact, when $\mathcal{D}$ is **finite**, $(\mathcal{A}(\mathcal{D}), \leq)$ is a **distributive lattice** where the meet and the join are given by $S_1 \wedge S_2 = max(\downarrow S_1 \cap \downarrow S_2)$ and $S_1 \vee S_2 = max(S_1 \cup S_2)$, respectively. Thus, one can build a pattern structure with $(\mathcal{A}(\mathcal{D}), \leq)$ which embeds the pattern setup $\mathbb{P} = (\mathcal{G}, \underline{\mathcal{D}}, \delta)$ for any finite set of objects $\mathcal{G}$. We call such a **pattern structure** the **antichain completion** of $\mathbb{P}$.

Definition 7. *Let $\mathbb{P} = (\mathcal{G}, \underline{\mathcal{D}}, \delta)$ be a **pattern setup**, the **antichain completion** of $\mathbb{P}$ is the **pattern setup** denoted by $\mathbb{P}^{\nabla}$ and given by:*

$$\mathbb{P}^{\nabla} = (\mathcal{G}, (\mathcal{A}(\mathcal{D}), \leq), \sigma : g \mapsto \{\delta(g)\})$$

Consider the following question for the more general case where $\underline{\mathcal{D}}$ is **infinite**:

Question 2. *What is a necessary and sufficient condition on $\mathbb{P}$ that makes $\mathbb{P}^{\nabla}$ a **pattern structure**?*

We have seen that the finiteness of $\underline{\mathcal{D}}$ is a **sufficient** condition (i.e. $(\mathcal{A}(\mathcal{D}), \leq)$ is a distributive lattice), but not a necessary one. Definition 1 requires for $\mathbb{P}^{\nabla}$ that for any $A \subseteq \mathcal{G}$ we have $\sigma[A] = \{\{\delta(g)\} \mid g \in A\}$ has a meet in $(\mathcal{A}(\mathcal{D}), \leq)$. This condition is verified **iff** pattern setup $\mathbb{P}$ satisfies the following condition:

$$\forall A \subseteq \mathcal{G} : cov(A) = \delta[A]^{\ell} \ = \ \downarrow max(\delta[A]^{\ell}) = \downarrow cov^{\mathcal{I}}(A) \tag{1}$$

and then $cov^{\mathcal{I}}(A)$ is the meet of $\sigma[A]$ in $(\mathcal{A}(\mathcal{D}), \leq)$ (see Theorem 2).

Definition 8. *A pattern setup $(\mathcal{G}, \underline{\mathcal{D}}, \delta)$ is said to be a **pattern hyper-structure** if condition (1) holds.*

Note that the term **hyper** in **pattern hyper-structure** comes from the notion of hyper-lattices briefly introduced in [19]. Note also that graphs ordered by subgraph isomorphism relation [13] induces a pattern hyper-structure but not a pattern structure. Same remark holds for sequential patterns [6,7] under the assumption of the existence of the largest sequence **1** subsumed by all sequences.

Remark 3. When $\underline{\mathcal{D}}$ is **infinite**, poset $(\mathcal{A}(\mathcal{D}), \leq)$ remains to be a join-semilattice but not necessarily has *finite meets* (as it is when $\underline{\mathcal{D}}$ is **finite**). It was shown in [4] that a **necessary and sufficient condition** on $(\mathcal{A}(\mathcal{D}), \leq)$ to have *finite meets* is as follows:

$$\forall S_1, S_2 \in \mathcal{A}(\mathcal{D}) \; \exists S \in \mathcal{A}(\mathcal{D}) : \downarrow S_1 \cap \downarrow S_2 = \downarrow S \tag{2}$$

In this case, S represents the meet (infimum) of $\{S_1, S_2\}$ in $\mathcal{A}(\mathcal{D})$; moreover, $(\mathcal{A}(\mathcal{D}), \leq)$ becomes a distributive lattice. Requiring $(\mathcal{A}(\mathcal{D}), \leq)$ to have a meet is a **sufficient condition**, but still **not necessary** for the *antichain completion* to be a pattern structure. In fact, condition (2) implies condition (1) (case $\mathcal{G}$ finite). Moreover, an example from [4] shows that condition (2) does not always hold even if $(\mathcal{D}, \sqsubseteq)$ is a meet-semilattice. Example 5 shows that condition (1) does not always hold.

Example 5. Consider the pattern setup $(\mathcal{G}, (\mathcal{D}, \supseteq), \delta)$ with $\mathcal{G} = \{g_1, g_2\}$, $\mathcal{D} = \{[a, b] \subseteq \mathbb{R} \mid a, b \in \mathbb{R}\} \backslash [1, 3]$, $\delta(g_1) = \{1\}$ and $\delta(g_2) = \{3\}$. The considered pattern setup does not verify condition (1) since $\delta[\mathcal{G}]^\ell$ is clearly not empty while $max(\delta[\mathcal{G}]^\ell]) = \emptyset$. $(\mathcal{G}, (\mathcal{D}, \supseteq), \delta)$ is thus not a pattern hyper-structure.

Theorem 1. *For any* ***pattern hyper-structure*** $(\mathcal{G}, \underline{\mathcal{D}}, \delta)$*, we have:*

$$\forall A \subseteq \mathcal{G} : \overline{A} = min\left(ext\left[cov^{\mathcal{I}}(A)\right]\right)$$

Proof. Let $A \in \wp(\mathcal{G})$. If $cov(A) = \emptyset$, then the property trivially holds since $cov^{\mathcal{I}}(A) = cov(A) = \emptyset$. Let $A \in \downarrow ext[\mathcal{D}]$ (i.e. $cov(A) \neq \emptyset$). We need to show the following equivalent proposition $min(ext[cov(A)]) = min(ext[max(cov(A))])$.

We start by showing $min(ext[cov(A)]) \subseteq min(ext[max(cov(A))])$. Let $B \in min(ext[cov(A)])$. Since $B \in ext[cov(A)]$, $\exists d \in cov(A)$ s.t. $B = ext(d)$. Since $cov(A) = \downarrow max(cov(A))$ ($(\mathcal{G}, \underline{\mathcal{D}}, \delta)$ is a pattern hyper-structure), there exists $c \in max(cov(A))$ s.t. $d \sqsubseteq c$. Since extent is an order revering mapping, $C = ext(c) \subseteq ext(d) = B$. Supposing that $C \in ext[max(cov(A))]$ s.t. $C \subsetneq B$ contradicts the fact that $B \in min(ext[cov(A)])$ since $C \subsetneq B$ and $C \in ext[cov(A)]$. It follows that $C = B \in ext[max(cov(A))]$. Again, supposing that $B \notin min(ext[max(cov(A)])$ leads to a contradiction ($\exists D \in ext[max(cov(A)] \subseteq ext[cov(A)]$ s.t. $D \subsetneq B$ while $B \in min(ext[cov(A)])$). We conclude that $B \in min(ext[max(cov(A))])$.

It remains to show that $min(ext[cov(A)]) \supseteq min(ext[max(cov(A))])$. Suppose the converse: $\exists E \in min(ext[max(cov(A))])$ such that $E \notin min(ext[cov(A)])$, we have $E \in ext[max(cov(A))] \subseteq ext[cov(A)]$. Since $E \notin min(ext[cov(A)])$ and $ext[cov(A)]$ is finite, we obtain that $\exists F \in min(ext[cov(A)])$ such that $F \subsetneq E$. The first inclusion implies $F \in min(ext[max(cov(A))])$. This is a contradiction, since at the same time $F \subsetneq E$ and $E \in min(ext[max(cov(A))])$.

We conclude that $\overline{A} = min(ext[cov(A)]) = min(ext[cov^{\mathcal{I}}(A)])$. □

Theorem 1 answers question 1. It says that in a **pattern hyper-structure** rather than considering all covering descriptions to compute $\overline{A}$, it is sufficient to consider only maximal covering ones. Theorem 2 answers question 2:

Theorem 2. *Let $\mathbb{P} = (\mathcal{G}, \underline{\mathcal{D}}, \delta)$ be a pattern setup, the* ***antichain completion*** *of $\mathbb{P}$ is a* ***pattern structure*** iff $\mathbb{P}$ *is a* ***pattern hyper-structure****. Operators $ext^{\triangledown}$ and $int^{\triangledown}$ denote extent and intent of $\mathbb{P}^{\triangledown}$ and are given by*

$$(\forall S \in \mathcal{A}(\mathcal{D}))\, ext^{\triangledown}(S) = \bigcap ext[S] \qquad (\forall A \subseteq \mathcal{G})\, int^{\triangledown}(A) = cov^{\mathcal{I}}(A)$$

Moreover, we have $\mathbb{P}^{\triangledown}_{ext} = \{\bigcap S \mid S \subseteq \mathbb{P}_{ext}\}$. Note that $\mathcal{G} \in \mathbb{P}^{\triangledown}_{ext}$.

Proof. Let us show that if $\mathbb{P}$ is a pattern hyper-structure then $\mathbb{P}^{\triangledown}$ is a pattern structure. $\mathbb{P}^{\triangledown}$ is a pattern structure **iff** every subset of $\sigma[\mathcal{G}]$ has a meet in $(\mathcal{A}(\mathcal{D}), \leq)$. For $A \subseteq \mathcal{G}$ we have $\sigma[A]^{\ell} = \{S \in \mathcal{A}(\mathcal{D}) \mid (\forall g \in A)\, S \subseteq \downarrow \delta(g)\} = \{S \in \mathcal{A}(\mathcal{D}) \mid S \subseteq \delta[A]^{\ell}\}$ where $\delta[A]^{\ell}$ and $\sigma[A]^{\ell}$ denote respectively the lower bounds of $\delta[A]$ w.r.t. $\sqsubseteq$ and the lower bounds of of $\sigma[A]$ w.r.t. $\leq$ (recall that $\delta[A]^{\ell} = \bigcap_{g \in A} \downarrow \delta(g)$). In this proof $\downarrow$ refers to the down-closure related to $\sqsubseteq$.

- ($\Rightarrow$) Let $A \subseteq \mathcal{G} : \delta[A]^{\ell} = \downarrow max(\delta[A]^{\ell})$. Thus $\sigma[A]^{\ell} = \{S \in \mathcal{A}(\mathcal{D}) \mid S \subseteq \downarrow max(\delta[A]^{\ell})\} = \{S \in \mathcal{A}(\mathcal{D}) \mid S \leq max(\delta[A]^{\ell})\}$. Since $max(\delta[A]^{\ell}) \in \mathcal{A}(\mathcal{D})$, so $max(\delta[A]^{\ell})$ is the meet of $\sigma[A]$ in $\mathcal{A}(\mathcal{D})$.
- ($\Leftarrow$) $\mathbb{P}^{\triangledown}$ is pattern structure is **equivalent** to say: $\forall A \subseteq \mathcal{G}$, $\sigma[A]$ has a meet $M \in \mathcal{A}(\mathcal{D})$. That is, for $A \subseteq \mathcal{G}$: $\forall S \in \mathcal{A}(\mathcal{D}) : S \subseteq \delta[A]^{\ell} \Leftrightarrow S \subseteq \downarrow M$. Particularly, for $S = \{d\}$ with $d \in \mathcal{D}$, we deduce that: $\forall d \in \delta[A]^{\ell} : d \in \downarrow M$. Thus, $\delta[A]^{\ell} \subseteq \downarrow M$. Moreover, since $M \subseteq \delta[A]^{\ell}$ ($M \in \sigma[A]^{\ell}$) and $\downarrow$ is a closure operator on $(\wp(\mathcal{D}), \subseteq)$ we have by monotony $\downarrow M \subseteq \delta[A]^{\ell} \subseteq \downarrow M$. We conclude that $\delta[A]^{\ell} = \downarrow M$ (note that $\downarrow \delta[A]^{\ell} = \delta[A]^{\ell}$). It follows that $M = max(\delta[A]^{\ell})$ that is $\delta[A]^{\ell} = \downarrow max(\delta[A]^{\ell})$. This concludes the equivalence.

Let us now define $int^{\triangledown}$ and $ext^{\triangledown}$. The previous proof shown that for $A \subseteq \mathcal{G}$ we have $int^{\triangledown}(A) = max(\delta[A]^{\ell}) = cov^{\mathcal{I}}(A)$ (The meet of $\sigma[A]$ is $max(\delta[A]^{\ell})$). For $ext^{\triangledown}$ operator, let $S \in \mathcal{A}(\mathcal{D})$. We have: $ext^{\triangledown}(S) = \{g \in \mathcal{G} \mid S \leq \sigma(g)\} = \{g \in \mathcal{G} \mid S \subseteq \downarrow \delta(g)\} = \{g \in \mathcal{G} \mid (\forall d \in S)\, d \sqsubseteq \delta(g)\} = \bigcap_{d \in S} ext(d) = \bigcap ext[S]$.

Let us show that $\mathbb{P}^{\triangledown}_{ext} = \{\bigcap S \mid S \subseteq \mathbb{P}_{ext}\}$. By definition of $ext^{\triangledown}$, the property $\mathbb{P}^{\triangledown}_{ext} \subseteq \{\bigcap S \mid S \subseteq \mathbb{P}_{ext}\}$ holds. For the inverse inclusion, it is sufficient to show that $\mathbb{P}_{ext} \subseteq \mathbb{P}^{\triangledown}_{ext}$ (since $(\mathbb{P}^{\triangledown}_{ext}, \subseteq)$ is closed under intersection). Let $A \in \mathbb{P}_{ext}$. $\exists d \in \mathcal{D}$ s.t. $A = ext(d)$. Since $\{d\} \in \mathcal{A}(\mathcal{D})$, and $ext^{\triangledown}(\{d\}) = ext(d) = A$. We conclude that $A \in \mathbb{P}^{\triangledown}_{ext}$ and $\mathbb{P}^{\triangledown}_{ext} = \{\bigcap S \mid S \subseteq \mathbb{P}_{ext}\}$. □

There is a completion that transforms any pattern setup to a pattern structure. This completion relies on the **Dedekind-MacNeille completion**.

Definition 9. *The family of subsets of $\mathcal{D}$ given by $DM(\mathcal{D}) = \{A^{\ell} \mid A \subseteq \mathcal{D}\}$ is a $\bigcap$-structure and is called the* ***Dedekind-MacNeille Completion*** *of $\underline{\mathcal{D}}$.*

Theorem 3. *Let $\mathbb{P} = (\mathcal{G}, \underline{\mathcal{D}}, \delta)$ be a* ***pattern setup****. The* ***Direct Completion*** *of $\mathbb{P}$ is the* ***pattern structure*** *denoted by $\mathbb{P}^{\blacktriangledown}$ and given by:*

$$\mathbb{P}^{\blacktriangledown} = (\mathcal{G}, (DM(\mathcal{D}), \subseteq), \gamma : g \mapsto \downarrow \delta(g))$$

Operators $ext^{\blacktriangledown}$ and $int^{\blacktriangledown}$ denote extent and intent of $\mathbb{P}^{\blacktriangledown}$ and are given by

$$(\forall S \in DM(\mathcal{D}))\, ext^{\blacktriangledown}(S) = \bigcap ext[S] \qquad (\forall A \subseteq \mathcal{G})\, int^{\blacktriangledown}(A) = cov(A) = \delta[A]^{\ell}$$

Moreover, we have $\mathbb{P}^{\blacktriangledown}_{ext} = \{\bigcap S \mid S \subseteq \mathbb{P}_{ext}\}$.

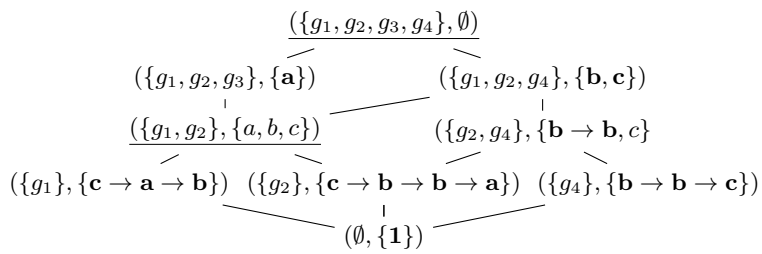


Fig. 3. Concept lattice $\underline{\mathfrak{B}}(\mathbb{P}^{\triangledown})$.

Proof. According to definition 9, $(DM(\mathcal{D}), \subseteq)$ is a complete lattice closed under intersection. Thus, the pattern setup $\mathbb{P}^{\blacktriangledown}$ is a pattern structure.

By definition, we have $int^{\blacktriangledown}(A) = \bigcap_{g \in A} \downarrow \delta(g) = \delta[A]^{\ell} = cov(A)$ since the meet in $\mathbb{P}^{\blacktriangledown}$ is $\bigcap$. For the extent operator $ext^{\blacktriangledown}$. Let $S \in DM(\mathcal{D})$, we have $ext^{\blacktriangledown}(S) = \{g \in \mathcal{G} \mid S \subseteq \downarrow \delta(g)\} = \{g \in \mathcal{G} \mid (\forall d \in S)\, d \sqsubseteq \delta(g)\} = \bigcap ext[S]$.

Let us show that $\mathbb{P}^{\blacktriangledown}_{ext} = \{\bigcap S \mid S \subseteq \mathbb{P}_{ext}\}$. Thanks to $ext^{\blacktriangledown}$ definition, property $\mathbb{P}^{\triangledown}_{ext} \subseteq \{\bigcap S \mid S \subseteq \mathbb{P}_{ext}\}$ holds. For the inverse inclusion, it is sufficient to show that $\mathbb{P}_{ext} \subseteq \mathbb{P}^{\blacktriangledown}_{ext}$ (since $(\mathbb{P}^{\blacktriangledown}_{ext}, \subseteq)$ is closed under intersection). Let $A \in \mathbb{P}_{ext}$, $\exists d \in \mathcal{D}$ such that $A = ext(d)$. We have $ext^{\blacktriangledown}(\{d\}^{\ell}) = ext^{\blacktriangledown}(\downarrow d) = \{g \in \mathcal{G} \mid \downarrow d \subseteq \downarrow \delta(g)\} = \{g \in \mathcal{G} \mid d \sqsubseteq \delta(g)\} = ext(d) = A$. We conclude that $A \in \mathbb{P}^{\blacktriangledown}_{ext}$ and $\mathbb{P}^{\blacktriangledown}_{ext} = \{\bigcap S \mid S \subseteq \mathbb{P}_{ext}\}$, which completes the proof. □

Example 6. Fig. 3 depicts the concept lattice $\underline{\mathfrak{B}}(\mathbb{P}^{\triangledown})$ associated to the *antichain completion* of the pattern hyper-structure $\mathbb{P}$ considered in Fig. 2 (*i.e.,* the description space is augmented with the top element $\mathbf{1}$). One can see that there are two new (underlined) extents $\{g_1, g_2\}$ and $\{g_1, g_2, g_3, g_4\}$ in $\mathbb{P}^{\triangledown}_{ext} \backslash \mathbb{P}_{ext}$. For instance, consider the intent of $\{g_1, g_2\}$ in the completion, each pattern d has extent $ext(d) \supsetneq \{g_1, g_2\}$. Extent $\{g_1, g_2, g_3, g_4\}$ is **non coverable** in $\mathbb{P}$ and thus $int^{\triangledown}(\{g_1, g_2, g_3, g_4\}) = max(cov(\{g_1, g_2, g_3, g_4\})) = max(\emptyset) = \emptyset$.

5 Conclusion

In this paper, we have developed a better understanding of pattern setups, a framework that models pattern spaces relying only on a poset. Next, we studied the usual transformation of pattern setups to pattern structures using antichains. We have shown that such a completion does not always produce a pattern structure unless the pattern setup is a pattern hyper-structure. Finally, we have shown that a natural completion of a pattern setup to a pattern structure exists thanks to the Dedekind-MacNeille completion. This work paves the way to answer an important question: *How to enumerate extents of a pattern setup without "visiting" the whole set of extents of its associated completion to a pattern structure?*

Acknowledgments

This work has been partially supported the Association Nationale Recherche Technologie (ANRT) French program. Sections 1, 2 and 3.1 were written by

Sergei O. Kuznetsov supported by the Russian Science Foundation under grant 17-11-01276 and performed at St. Petersburg Department of Steklov Mathematical Institute of Russian Academy of Sciences, Russia.

References

1. Agrawal, R., Srikant, R.: Mining sequential patterns. In: Data Engineering. pp. 3–14. IEEE (1995)
2. Baixeries, J., Kaytoue, M., Napoli, A.: Characterizing functional dependencies in formal concept analysis with pattern structures. Ann. Math. Artif. Intell. **72**(1-2), 129–149 (2014)
3. Belfodil, A., Kuznetsov, S.O., Robardet, C., Kaytoue, M.: Mining convex polygon patterns with formal concept analysis. In: IJCAI. pp. 1425–1432 (2017)
4. Boldi, P., Vigna, S.: On the lattice of antichains of finite intervals. CoRR **abs/1510.03675** (2016), `https://arxiv.org/abs/1510.03675`
5. Boley, M., Horváth, T., Poigné, A., Wrobel, S.: Listing closed sets of strongly accessible set systems with applications to data mining. Theor. Comput. Sci. **411**(3), 691–700 (2010)
6. Buzmakov, A., Egho, E., Jay, N., Kuznetsov, S.O., Napoli, A., Raïssi, C.: On mining complex sequential data by means of FCA and pattern structures. Int. J. General Systems **45**(2), 135–159 (2016)
7. Codocedo, V., Bosc, G., Kaytoue, M., Boulicaut, J., Napoli, A.: A proposition for sequence mining using pattern structures. In: ICFCA. LNCS (10308). 106–121. (2017)
8. Crampton, J., Loizou, G.: The completion of a poset in a lattice of antichains. International Mathematical Journal **1**(3), 223–238 (2001)
9. Ganter, B., Wille, R.: Formal Concept Analysis. Springer (1999)
10. Ganter, B., Kuznetsov, S.O.: Pattern structures and their projections. In: ICCS. LNCS (2120). 129–142. (2001)
11. Garriga, G.C., Kralj, P., Lavrac, N.: Closed sets for labeled data. Journal of Machine Learning Research **9**, 559–580 (2008)
12. Kaytoue, M., Kuznetsov, S.O., Napoli, A.: Revisiting Numerical Pattern Mining with Formal Concept Analysis. In: IJCAI. pp. 1342–1347 (2011)
13. Kuznetsov, S.O.: Learning of simple conceptual graphs from positive and negative examples. In: PKDD. LNCS (1704). 384–391., vol. 1704, pp. 384–391 (1999)
14. Kuznetsov, S.O.: Pattern structures for analyzing complex data. In: RSFDGrC. LNCS (5908). 33–44. (2009)
15. Liu, B., Hsu, W., Ma, Y.: Integrating classification and association rule mining. In: KDD. pp. 80–86 (1998)
16. Lumpe, L., Schmidt, S.E.: Pattern structures and their morphisms. In: CLA. CEUR Workshop Proceedings, vol. 1466, pp. 171–179 (2015)
17. Pawlak, Z.: Rough sets. International Journal of Parallel Programming **11**(5), 341–356 (1982)
18. Roman, S.: Lattices and Ordered Sets. Springer New York (2008)
19. Zaki, M.J.: SPADE: an efficient algorithm for mining frequent sequences. Machine Learning **42**(1/2), 31–60 (2001)