



Introducer Concepts in n-Dimensional Contexts

Giacomo Kahn, Alexandre Bazin

► To cite this version:

Giacomo Kahn, Alexandre Bazin. Introducer Concepts in n-Dimensional Contexts. 2018. hal-01706404v1

HAL Id: hal-01706404

<https://hal.science/hal-01706404v1>

Preprint submitted on 11 Feb 2018 (v1), last revised 1 Feb 2019 (v2)

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.

Introducer Concepts in n -Dimensional Contexts

Giacomo Kahn¹ and Alexandre Bazin²

¹ LIMOS & Université Clermont Auvergne, France

² Le2i - Laboratoire Electronique, Informatique et Image, France
giacomo.kahn@isima.fr, contact@alexandrebazin.com

Abstract. Concept lattices are well-known conceptual structures that organise interesting patterns—the concepts—extracted from data. In some applications, such as software engineering or data mining, the size of the lattice can be a problem, as it is often too large to be efficiently computed, and too complex to be browsed. For this reason, the Galois Sub-Hierarchy, a restriction of the concept lattice to introducer concepts, has been introduced as a smaller alternative. In this paper, we generalise the Galois Sub-Hierarchy to n -lattices, conceptual structures obtained from multidimensional data in the same way that concept lattices are obtained from binary relations.

Keywords: Formal Concept Analysis, Polyadic Concept Analysis, Introducer Concept, AOC-poset, Galois Sub-Hierarchy.

1 Introduction

Formal Concept Analysis [1] is a mathematical framework that allows, from a binary relation, to extract interesting patterns called concepts. Those patterns form a hierarchy called a concept lattice.

The size of the lattice, potentially exponential in the size of the relation, is one of the main drawbacks in its use as data representation. In some cases, it is possible to avoid using the whole lattice, as it contains redundant information [2]. The AOC-poset (or Galois Sub-Hierarchy) is a sub-order of the lattice that preserves only some of its key elements [3]. Its size is potentially much smaller than the size of the associated concept lattice [4] and it can be used in place of the concept lattice to perform certain tasks [5, 6].

The generalisation of FCA to the n -dimensional case, Polyadic Concept Analysis [7], focuses on multidimensional data, i.e. n -ary relations.

In this paper, we generalise the notion of AOC-posets to n -lattices. In Section 2, we provide the definitions and notations that we use throughout the paper. Section 3 is dedicated to the definition of introducer concepts in n -lattices, and some properties about those concepts. In Section 4, we present an algorithm to compute the introducer sub-order, and study its complexity. Finally, we conclude and discuss some future works.

2 Definitions and Notations

In this section, we introduce classical definitions and notations from Formal Concept Analysis and Polyadic Concept Analysis. They can also be found in [1] and [7].

2.1 Formal Concept Analysis

From now on, we will omit the brackets in the notation for sets when no confusion is induced by this simplification.

A (formal) context is a triple $(\mathcal{S}_1, \mathcal{S}_2, \mathcal{R})$ in which \mathcal{S}_1 and \mathcal{S}_2 are sets and $\mathcal{R} \subseteq \mathcal{S}_1 \times \mathcal{S}_2$ is a binary relation between them. The elements of \mathcal{S}_1 are called the (formal) objects and those of \mathcal{S}_2 the (formal) attributes. A pair $(x_1, x_2) \in \mathcal{R}$ means that “the object x_1 has the attribute x_2 ”. A context can be represented as a cross table, as shown in Fig. 1. For instance, object 1 has attributes a and b , and attribute b is shared by objects 1 and 2.

	a	b	c
1	×	×	
2		×	×
3	×		×

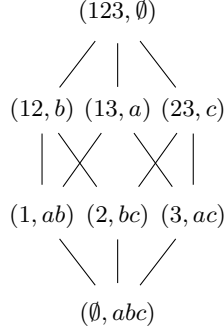
Fig. 1: An example of a context $\mathcal{C} = (\mathcal{S}_1, \mathcal{S}_2, \mathcal{R})$ with $\mathcal{S}_1 = \{1, 2, 3\}$ and $\mathcal{S}_2 = \{a, b, c\}$.

Two *derivation operators* $(\cdot)' : 2^{\mathcal{S}_1} \mapsto 2^{\mathcal{S}_2}$ and $(\cdot)' : 2^{\mathcal{S}_2} \mapsto 2^{\mathcal{S}_1}$ are defined. For $X_1 \subseteq \mathcal{S}_1$ and $X_2 \subseteq \mathcal{S}_2$, $X_1' = \{a \mid \forall x \in X_1, (x, a) \in \mathcal{R}\}$ and $X_2' = \{o \mid \forall y \in X_2, (o, y) \in \mathcal{R}\}$.

A formal concept is a pair (X_1, X_2) where $X_1 \subseteq \mathcal{S}_1$, $X_2 \subseteq \mathcal{S}_2$, $X_1' = X_2$ and $X_2' = X_1$. This corresponds to a maximal set of objects that share a maximal set of attributes and can be viewed as a maximal rectangle full of crosses in the formal context, up to permutations on the elements of the rows and columns. X_1 is called the *extent* of the concept, while X_2 is called the *intent*.

The set of all concepts of a context ordered by the inclusion relation on either one of their components forms a complete lattice. Additionally, every complete lattice is isomorphic to the concept lattice of some context [1]. The concept lattice associated with the formal context from Fig. 1 is shown in Fig. 2.

In an application, the size of the concept lattice might be a drawback. For this reason, Godin et al. [3] introduced a sub-hierarchy of the lattice, the Galois sub-hierarchy. This sub-hierarchy was introduced and is most often used in the field of software engineering, but is also used in other fields, such as Relational


 Fig. 2: Concept lattice associated with \mathcal{C} .

Concept Analysis (RCA) [8] and data mining [5]. Additionally, the Galois sub-hierarchy is integrated in some FCA tools, such as Latviz [9], GALICIA [10], RCAExplore¹ or AOC-poset Builder².

Definition 1 (Introducer concept). An Object-Concept is a concept (o'', o') with $o \in \mathcal{S}_1$. We say that this concept introduces o .

An Attribute-Concept is a concept (a', a'') with $a \in \mathcal{S}_2$. We say that this concept introduces a .

A concept can introduce both attributes and objects, or it can introduce neither. We call the sub-order restricted to the introducer concepts an *Attribute-Object-Concept partially ordered set* (AOC-poset), or Galois Sub-Hierarchy (GSH). While a concept lattice can have up to $2^{\min(|\mathcal{S}_1|, |\mathcal{S}_2|)}$ concepts, the associated GSH has at most $|\mathcal{S}_1| + |\mathcal{S}_2|$ elements. Several algorithms exist to compute the GSH [11–14].

2.2 Polyadic Concept Analysis

Definition 2. An n -context is an $(n + 1)$ -tuple $\mathcal{C} = (\mathcal{S}_1, \dots, \mathcal{S}_n, \mathcal{R})$ in which \mathcal{S}_i , $i \in \{1, \dots, n\}$, is a set called a dimension and \mathcal{R} is an n -ary relation between the dimensions.

An n -context can be represented by a $|\mathcal{S}_1| \times \dots \times |\mathcal{S}_n|$ cross table as illustrated in Fig. 3.

Definition 3. An n -concept of $\mathcal{C} = (\mathcal{S}_1, \dots, \mathcal{S}_n, \mathcal{R})$ is an n -tuple (X_1, \dots, X_n) such that $\prod_{i \in \{1, \dots, n\}} X_i \subseteq \mathcal{R}$ and there are no $i \in \{1, \dots, n\}$ and $k \in \mathcal{S}_i \setminus X_i$ such that $\{k\} \times \prod_{j \in \{1, \dots, n\} \setminus \{i\}} X_j \subseteq \mathcal{R}$.

¹ <http://dolques.free.fr/rcaexplore/>

² <http://www.lirmm.fr/AOC-poset-Builder/>

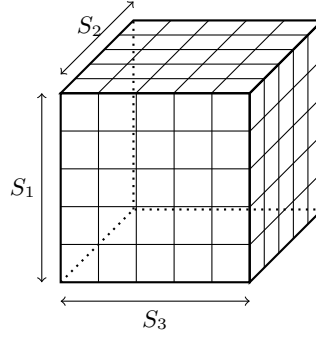


Fig. 3: Visual representation of a 3-context without its crosses.

An n -concept can be viewed as a maximal n -dimensional box full of crosses up to permutations on the elements of the dimensions. We denote by $\mathcal{T}(\mathcal{C})$ the set of n -concepts of a n -context \mathcal{C} .

	a	b	c		a	b	c
1	×	×			×		
2					×		
3	×				×		×
	α				β		

Fig. 4: An example of a $2 \times 3 \times 3$ 3-context.

In the Fig. 4 example, seven 3-concepts are present: $(\alpha, 1, ab)$, $(\alpha\beta, 13, a)$, $(\beta, 3, ac)$, $(\beta, 123, a)$, $(\alpha\beta, 123, \emptyset)$, $(\alpha\beta, \emptyset, abc)$ and $(\emptyset, 123, abc)$.

Definition 4 (From [15]). $\mathcal{S} = (S, \lesssim_1, \dots, \lesssim_n)$ is an n -ordered set if for $A \in S$ and $B \in S$:

1. $A \sim_i B, \forall i \in \{1, \dots, n\} \Rightarrow A = B$ (Uniqueness Condition)
2. $A \lesssim_i B, \forall i \in (\{1, \dots, n\} \setminus j) \Rightarrow B \lesssim_j A$ (Antiordinal Dependency)

For the Antiordinal Dependency condition to be respected, it is sufficient to have $i, j \in \{1, \dots, n\}, i \neq j$ such that $A \lesssim_i B$ and $B \lesssim_j A$.

The set of all the n -concepts of an n -context together with the n quasi-orders \lesssim_i induced by the inclusion relations on the subsets of each dimension forms an n -ordered set. Additionally, the existence of some particular joins makes it a complete n -lattice. Every n -lattice can be associated with some n -context [7].

Definition 5. Let $x \in \mathcal{S}_i$ be an element of a dimension i . We denote by \mathcal{C}_x the $(n - 1)$ -context $\mathcal{C}_x = (S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_n, \mathcal{R}_x)$ where

$$\mathcal{R}_x = \{(s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n) \mid (s_1, \dots, s_{i-1}, x, s_{i+1}, \dots, s_n) \in \mathcal{R}\}$$

With the previous definition, \mathcal{C}_x is the $(n - 1)$ -context corresponding to element x , represented by the shaded area in Fig. 5.

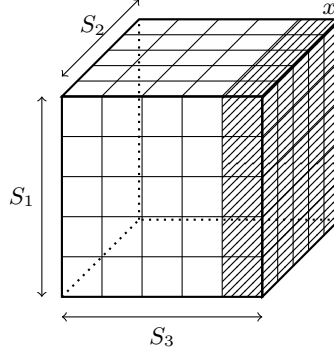


Fig. 5: If x is an element of \mathcal{S}_3 in this 3-context, then \mathcal{C}_x is the 2-context resulting from fixing x .

3 Introducer concepts in n -Lattices

In this section, we define introducer concepts in n -lattices. In the next definitions, we will call dimension i the height, while all other dimensions are called the width.

Definition 6. Let $x \in \mathcal{S}_i$ be an element of a dimension i . The concepts with maximal width such that x is in the height are the introducer concepts of x . The set of introducer concepts of x is denoted by I_x .

In the Fig. 4 example, we have $I_\alpha = \{(\alpha\beta, 13, a), (\alpha, 1, ab)\}$.

We denote by $\mathcal{I}(\mathcal{S}_i) = \bigcup_{x \in \mathcal{S}_i} I_x$ the set of concepts that introduce an element of dimension i and by $\mathcal{I}(\mathcal{C}) = \bigcup_{i \in \{1, \dots, n\}} \mathcal{I}(\mathcal{S}_i)$ the set of all introducer concepts of a context \mathcal{C} .

As in the 2-dimensional case, irreducible elements are introducer concepts. However, $\mathcal{I}(\mathcal{C})$ is not always strictly the set of irreducible elements as some applications expect the context not to be reduced.

Proposition 1. $(\mathcal{I}(\mathcal{C}), \lesssim_1, \dots, \lesssim_n)$ is an n -ordered set.

Proof. Let A and B be in $\mathcal{I}(\mathcal{C})$. We recall that $A_i \subseteq B_i \Leftrightarrow A \lesssim_i B$ and that $A_i = B_i \Leftrightarrow A \sim_i B$. Without loss of generality, let $A \in \mathcal{I}(\mathcal{S}_i)$ and $B \in \mathcal{I}(\mathcal{S}_j)$.

If $\forall k \in \{1, \dots, n\}$, $A \sim_k B$, then $\forall k \in \{1, \dots, n\}$, $A_k = B_k$, so $A = B$ (Uniqueness Condition).

If A and B are distinct, $\exists k \in \{1, \dots, n\}$ such that $A \lesssim_k B$ or $B \lesssim_k A$. Without loss of generality, suppose $A \lesssim_k B$. Suppose that there is no $\ell \in \{1, \dots, n\}$ such that $B \lesssim_\ell A$. That implies that all the components of A are included in the components of B . Then this is in contradiction with the maximality condition implied by A being a concept. Thus $\exists j \in \{1, \dots, n\} \setminus i$ such that $B \lesssim_j A$ (Antiordinal Dependency). \square

As in the 2-dimensional case where concept lattices and GSH are respectively complete lattices and partially ordered sets, in the n -dimensional case we have complete n -lattices and n -ordered sets.

Proposition 2. Let $x \in \mathcal{S}_i$. If $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$ is an $(n-1)$ -concept of \mathcal{C}_x , then $(X_1, \dots, \{x\} \cup X_i, \dots, X_n)$ is an introducer of x .

If $(X_1, \dots, \{x\} \cup X_i, \dots, X_n)$ is an introducer of x , then there exists an $(n-1)$ -concept $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$ in \mathcal{C}_x .

Proof. We suppose, without loss of generality, that $x \in \mathcal{S}_1$. The $(n-1)$ -concepts of \mathcal{C}_x are of the form (X_2, \dots, X_n) . If (x, X_2, \dots, X_n) is an n -concept of \mathcal{C} , then it is minimum in height and maximal in width and is thus an introducer of x .

If (x, X_2, \dots, X_n) is not an n -concept of \mathcal{C} , as (X_2, \dots, X_n) is a $(n-1)$ -concept of \mathcal{C}_x , then (x, X_2, \dots, X_n) can be augmented only on the first dimension. As such, there exists an n -concept $(\{x\} \cup X_1, X_2, \dots, X_n)$ that is maximal in width and is thus an introducer of x .

Suppose that there is an $X = (X_1, \dots, X_n) \in I_x$, $x \in X_1$, that is not obtained from an $(n-1)$ -concept of \mathcal{C}_x by extending X_1 . It means that (X_2, \dots, X_n) is not maximal in \mathcal{C}_x (else it would be an $(n-1)$ -concept). Then, there exists an n -concept $Y = (Y_1, Y_2, \dots, Y_n)$ with $Y_1 \subseteq X_1$ and $X_i \subseteq Y_i$, for $i \in \{2, \dots, n\}$. This is a contradiction with the fact that X is an introducer of x . \square

Proposition 2 states that every $(n-1)$ -concept of \mathcal{C}_x maps to an introducer of x in \mathcal{C} , and that every introducer of x is the image of an $(n-1)$ -concept of \mathcal{C}_x .

4 Algorithm

In this section, we present an algorithm to compute the introducer concepts in an n -context.

Algorithm 1 computes the introducers for each element of a dimension i . For a given element $x \in \mathcal{S}_i$, we compute $\mathcal{T}(\mathcal{C}_x)$. Then, for each $(n-1)$ -concept

$X \in \mathcal{T}(\mathcal{C}_x)$, we build the set X_i needed to extend X into an n -concept. An element y is added to X_i when $y \times \prod_{j \neq i} X_j \subseteq \mathcal{R}$, that is if there exists an $(n-1)$ -dimensional box full of crosses (but not necessarily maximal) in \mathcal{R} , at level y . The final set X_i always contains at least x .

Algorithm 1: INTRODUCERDIM(\mathcal{C}, i)

Input: \mathcal{C} an n -context, $i \in \{1, \dots, n\}$ a dimension
Output: $\mathcal{I}(\mathcal{S}_i)$ the set of introducer concepts of elements of dimension i

```

1  $I \leftarrow \emptyset$ 
2 foreach  $x \in \mathcal{S}_i$  do
3    $C \leftarrow \emptyset$ 
4   foreach  $X = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n) \in \mathcal{T}(\mathcal{C}_x)$  do
5      $X_i \leftarrow \emptyset$ 
6     foreach  $y \in \mathcal{S}_i$  do
7       if  $\prod_{j \neq i} X_j \times y \subseteq \mathcal{R}$  then
8          $X_i \leftarrow X_i \cup y$ 
9      $C \leftarrow C \cup (X_1, \dots, X_i, \dots, X_n)$ 
10   $I \leftarrow I \cup C$ 
11 return  $I$ 
    
```

Algorithm 2 calls Algorithm 1 on each dimension. This ensure that each element of each dimension will be scanned for its introducer concepts. Algorithm 2 computes the introducer set $\mathcal{I}(\mathcal{C})$ for n -context \mathcal{C} .

Algorithm 2: INTRODUCERS(\mathcal{C})

Input: \mathcal{C} an n -context
Output: $\mathcal{I}(\mathcal{C})$ the set of all introducer concepts for \mathcal{C}

```

1  $R \leftarrow \emptyset$ 
2 foreach dimension  $i$  do
3    $R \leftarrow R \cup \text{INTRODUCERDIM}(\mathcal{C}, i)$ 
4 return  $R$ 
    
```

Algorithm 1 requires the computation of the $(n-1)$ -concepts from an $(n-1)$ -context. Several algorithms exist to complete this task [16–18].

Proposition 3. *Algorithm 1 ends and returns all the introducer concepts of elements of the dimension \mathcal{S}_i .*

Proof. The \mathcal{S}_i are finite. The set of $(n-1)$ -concepts of an $(n-1)$ -context resulting from fixing an element is also finite. Algorithm 1 passes through each element $x \in \mathcal{S}_i$ and on each concept of $\mathcal{T}(\mathcal{C}_x)$ exactly once. The maximality test on dimension i looks at the elements of \mathcal{S}_i , which is finite. Thus, the algorithm ends.

Proposition 2 ensures that every introducer of x can be computed from the concepts of \mathcal{C}_x . Thus, every introducer of an element of the dimension \mathcal{S}_i is returned. \square

At the time of writing, the only known bound for the number of n -concepts of an n -context (S_1, \dots, S_n, R) is $\prod_{i \in \{1, \dots, n\} \setminus k} 2^{|\mathcal{S}_i|}$ with $k = \operatorname{argmax}_{k \in \{1, \dots, n\}} |\mathcal{S}_k|$. Let \mathbb{K}_n be the maximal number of n -concepts in an n -context. Computing \mathcal{C}_x from \mathcal{C} is in $O(|\mathcal{R}|)$. Building the set X_i that extends an $(n-1)$ -concept of \mathcal{C}_x into an introducer of x can be done in $O(|\mathcal{S}_i| \times \prod_{j \neq i} |X_j|)$. We denote by T the complexity of computing $\mathcal{T}(\mathcal{C}_x)$ from \mathcal{C}_x .

Thus the complexity of Algorithm 1 for context $\mathcal{C} = (S_1, \dots, S_n, \mathcal{R})$ and dimension i is $O\left(|\mathcal{S}_i| \times \left(T + \mathbb{K}_{n-1} \times \prod_{j \in \{1, \dots, n\}} |\mathcal{S}_j|\right)\right)$ and the complexity of Algorithm 2 is $O\left(\sum_{i \in \{1, \dots, n\}} \left(|\mathcal{S}_i| \times \left(T + \mathbb{K}_{n-1} \times \prod_{j \in \{1, \dots, n\}} |\mathcal{S}_j|\right)\right)\right)$.

5 Conclusion

In this paper, we introduced the n -dimensional equivalent of Galois Sub-Hierarchies or AOC-posets. We showed that the set of introducer concepts, together with the n quasi-orders induced by the inclusion on each dimension, forms an n -ordered set. We provided an algorithm to compute the set of introducer concepts from an n -context.

Although our approach was not initially motivated by an applicative problem, it would be interesting to use the notion of introducer concepts in n -dimensions to address some specific problems in software engineering or data mining.

It would be interesting to experiment on datasets (real and generated) to evaluate the gains (in term of number of concepts) of the restriction to introducer concepts.

Acknowledgements

This research was partially supported by the European Union's “*Fonds Européen de Développement Régional (FEDER)*” program.

References

1. Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis - Mathematical Foundations*. Springer, 1999.

2. Robert Godin, Hamed Mili, Guy W. Mineau, Rokia Missaoui, Amina Arfi, and Thuy-Tien Chau. Design of Class Hierarchies Based on Concept (Galois) Lattices. *TAPOS*, 4(2):117–134, 1998.
3. Robert Godin and Hamed Mili. Building and Maintaining Analysis-Level Class Hierarchies Using Galois Lattices. In *8th Conference on Object-Oriented Programming Systems, Languages, and Applications (OOPSLA)*, pages 394–410, 1993.
4. Jessie Carbonnel, Marianne Huchard, and Alain Gutierrez. Variability Representation in Product Lines using Concept Lattices: Feasibility Study with Descriptions from Wikipedia’s Product Comparison Matrices. In *Proceedings of the International Workshop on Formal Concept Analysis and Applications, FCA&A 2015, co-located with 13th International Conference on Formal Concept Analysis (ICFCA 2015), Nerja, Málaga, Spain, June 23-26, 2015.*, pages 93–108, 2015.
5. Xavier Dolques, Florence Le Ber, Marianne Huchard, and Corinne Grac. Performance-Friendly Rule extraction in Large Water Data-sets with AOC Posets and Relational Concept Analysis. *Int. J. General Systems*, 45(2):187–210, 2016.
6. Alexandre Bazin, Jessie Carbonnel, and Giacomo Kahn. On-Demand Generation of AOC-Posets: Reducing the Complexity of Conceptual Navigation. In *Foundations of Intelligent Systems - 23rd International Symposium, ISMIS 2017, Warsaw, Poland, June 26-29, 2017, Proceedings*, pages 611–621, 2017.
7. George Voutsadakis. Polyadic Concept Analysis. *Order*, 19(3):295–304, 2002.
8. Marianne Huchard, Mohamed Rouane Hacene, Cyril Roume, and Petko Valtchev. Relational Concept Discovery in Structured Datasets. *Ann. Math. Artif. Intell.*, 49(1-4):39–76, 2007.
9. Mehwish Alam, Thi Nhu Nguyen Le, and Amedeo Napoli. Latviz: A New Practical Tool for Performing Interactive Exploration over Concept Lattices. In *Proceedings of the Thirteenth International Conference on Concept Lattices and Their Applications, Moscow, Russia, July 18-22, 2016.*, pages 9–20, 2016.
10. Petko Valtchev, David Grosser, Cyril Roume, and M Rouane Hacene. Galicia: an open platform for lattices. In *Using Conceptual Structures: Contributions to the 11th Intl. Conference on Conceptual Structures (ICCS03)*, pages 241–254, 2003.
11. Hervé Dicky, Christophe Dony, Marianne Huchard, and Thérèse Libourel. Ares, Adding a Class and Restructuring Inheritance Hierarchy. In *11èmes Journées Bases de Données Avancées*, pages 25–42, 1995.
12. Marianne Huchard, Hervé Dicky, and Hervé Leblanc. Galois Lattice as a Framework to Specify Building Class Hierarchies Algorithms. *ITA*, 34(6):521–548, 2000.
13. Gabriela Arévalo, Anne Berry, Marianne Huchard, Guillaume Perrot, and Alain Sigayret. Performances of Galois Sub-Hierarchy-Building Algorithms. In *Formal Concept Analysis, 5th International Conference, ICFCA 2007, Clermont-Ferrand, France, February 12-16, 2007, Proceedings*, pages 166–180, 2007.
14. Anne Berry, Alain Gutierrez, Marianne Huchard, Amedeo Napoli, and Alain Sigayret. Hermes: a Simple and Efficient Algorithm for Building the AOC-Poset of a Binary Relation. *Annals of Mathematics and Artificial Intelligence*, 72(1-2):45–71, 2014.
15. George Voutsadakis. Dedekind-Macneille Completion of n -Ordered sets. *Order*, 24(1):15–29, 2007.
16. Loïc Cerf, Jérémy Besson, Céline Robardet, and Jean-François Boulicaut. Closed Patterns Meet n -ary Relations. *TKDD*, 3(1):3:1–3:36, 2009.
17. Tatiana Makhalova and Lhouari Nourine. An Incremental Algorithm for Computing n -Concepts.
18. Alexandre Bazin, Giacomo Kahn, and Olivier Raynaud. An Incremental Algorithm for Computing n -Dimensional Concepts. 2017.