Distances Between Formal Concept Analysis Structures

Alexandre Bazin
LIRMM, CNRS, Université de Montpellier, Montpellier, France

Giacomo Kahn
Université Lumière Lyon 2, INSA Lyon, Université Claude Bernard Lyon 1, France
Université Jean Monnet Saint-Etienne, DISP UR4570, Bron, France

Abstract

In this paper, we study the notion of distance between the most important structures of formal concept analysis: formal contexts, concept lattices, and implication bases. We first define three families of Minkowski-like distances between these three structures. We then present experiments showing that the correlations of these distances are low and depend on the distance between formal contexts.

2012 ACM Subject Classification Mathematics of computing → Discrete mathematics

Keywords and phrases Formal Concept Analysis, Implication Base, Concept Lattice, Pattern Mining, Ordinal Data Science

Digital Object Identifier 10.4230/TGDK.3.2.2

Category Research

Related Version *Previous Version*: `https://hal.science/hal-04475242`

Supplementary Material *Software (Source Code)*: `https://github.com/Authary/FCAD`
archived at `swh:1:dir:75951913d5a4222771718415a619dc6b1a97a6ed`
Software (Experiments Code): `https://github.com/Authary/experiments_distances_fca`
archived at `swh:1:dir:00363b3ec2ad2d63a8235a852392699f1ccf6688`

Funding This work was partially supported by the ANR SmartFCA project Grant ANR-21-CE23-0023 of the French National Research Agency.

Acknowledgements The authors thank the members of the SmartFCA project for their advice.

Received 2024-12-09 **Accepted** 2025-05-07 **Published** 2025-10-15

1 Introduction

Formal Concept Analysis (FCA [12]) is a mathematical framework that allows extracting patterns called concepts from data in the form of objects described by attributes, and organises them in an ordered structure called a concept lattice. Concept lattices are then used for exploratory search [15, 14], conceptual navigation [22, 1], and other applications – see [17] for a survey. The framework also handles implications between sets of attributes, that can be summarised by implication bases. In FCA, formal contexts, concept lattices and sets of implications are three representations of – or points of view on – the same entity and all three of them are well known, well studied, and well used in various fields of data mining [19, 18, 7].

We are interested in distances between these FCA structures. Given two data tables on the same objects and attributes, how far apart are the structures that are extracted from them? In this paper, we define three families of distances: one between formal contexts, one between implication bases, and one between concept lattices. For formal contexts, we consider the context as a set of pairs (the incidence relation) and use set-based analogues of Minkowski distances to define the *factual distance*. For concept lattices and implication bases, we consider the structures as representations of, respectively, the derivation operators and the closure operator of the

corresponding context and propose Minkowski-like distances: the *conceptual distance* and the *logical distance*. We show that these distances are metrics and we provide algorithms to compute them. We experimentally study the correlations between those distances on formal contexts that are *closer* or *farther apart* and observe that these correlations depend on the factual distance.

There are multiple expected applications for this work. The most direct one would be the comparison of concept lattices or implication bases, for instance to study the differences in the variability in different software product lines [3]. This contribution would then allow for the study of the trajectory of a given software collection between versions – how much a new version differs from older versions. The notion of trajectory can also be used for iterative processes such as Relational Concept Analysis [21, 2], to quantify how much variation there is between concept lattices in some given steps. In those two contexts, the objects and attributes are basically the same through time or during the process, and the labels of objects carry some significance, which is why our distances consider contexts with the same object and attributes sets. In distance-based machine learning, knowledge is often embedded in numerical vectors. This contribution allows for the direct computation of a distance between knowledge structures. More broadly, these distances could be used to define complexity indicators in triadic or polyadic datasets [24] as the relative distance of each *slice* of n-context to every other. Additionally, this is a contribution to Ordinal Data Science, as defined in its manifesto [23], and it seemed like an interesting question to be answered in itself.

The paper follows a classic structure: in Section 2 we define the necessary notions of FCA and distances, then we introduce our distances between FCA structures and the algorithms to compute them in Section 3. In Section 4 we experiment on the new distances: we study the correlations between them and compare them together and with Domenach's dissimilarity measure [9] on concept lattices.

2 Preliminaries

2.1 Formal Concept Analysis

In the following, we consider only finite sets.

Formal Concept Analysis (FCA) is a mathematical framework based on lattice theory that aims at structuring the information contained in the relation between *objects* and their *attributes* [12]. It is centered around the notion of *formal context*.

▶ **Definition 1** (Formal context). *A* formal context *is a triple* $(\mathcal{O}, \mathcal{A}, \mathcal{R})$ *in which* $\mathcal{O}$ *is a set of* objects*,* $\mathcal{A}$ *is a set of* attributes *and* $\mathcal{R} \subseteq \mathcal{O} \times \mathcal{A}$ *is a binary relation between objects and attributes. We say that* the object o is described by the attribute a *when* $(o, a) \in \mathcal{R}$*.*

Formal contexts can be represented as crosstables.

	a_1	a_2	a_3	a_4	a_5
o_1	×	×			
o_2		×	×	×	
o_3		×		×	×
o_4			×		×
o_5				×	×

Figure 1 A formal context with five objects and five attributes.

A formal context $\mathcal{C}$ gives rise to two *derivation operators*, both usually noted $\cdot'$, defined as:

$$\cdot' : \mathcal{P}(\mathcal{A}) \to \mathcal{P}(\mathcal{O})$$

$$A' = \{o \in \mathcal{O} \mid \forall a \in A, (o, a) \in \mathcal{R}\}$$

$$\cdot' : \mathcal{P}(\mathcal{O}) \to \mathcal{P}(\mathcal{A})$$

$$O' = \{a \in \mathcal{A} \mid \forall o \in O, (o, a) \in \mathcal{R}\}$$

where $\mathcal{P}(X)$ denotes the powerset of X.

For instance, in the Fig. 1 context, $\{a_2, a_4\}' = \{o_2, o_3\}$ and $\{a_1\}'' = \{a_1, a_2\}$. Both operators $\cdot'$ form a Galois connection and their compositions $\cdot''$ are closure operators. Throughout this paper, when in the presence of two different formal contexts $\mathcal{C}_1$ and $\mathcal{C}_2$, we shall use $\cdot'^i$ and $\cdot''^i$ to denote the derivation and closure operators of context $\mathcal{C}_i$.

▶ **Definition 2** (Formal concept). *In a formal context $(\mathcal{O}, \mathcal{A}, \mathcal{R})$, a* formal concept *is a pair (E, I) in which E is a set of objects, I is a set of attributes, and such that $E = I'$ and $I = E'$. As such, $I = I''$ and $E = E''$ are both closed sets. We call E the* extent *and I the* intent *of the concept.*

Visually, concepts correspond to maximal rectangles of crosses in the context's crosstable, up to permutation of rows and columns. In the Fig. 1 context, the pair $(\{o_2, o_3\}, \{a_2, a_4\})$ is a concept while the pair $(\{o_3, o_4\}, \{a_5\})$ is not as $\{a_5\}' = \{o_3, o_4, o_5\}$. Concepts can be ordered by the inclusion relation on their extents, i.e. $(E_1, I_1) \leq (E_2, I_2) \Leftrightarrow E_1 \subseteq E_2$. As per the basic theorem of formal concept analysis [12], the set of all concepts of a context $\mathcal{C}$ ordered in such a way forms a complete lattice called the *concept lattice of* $\mathcal{C}$. Additionally, all complete lattices are isomorphic to the concept lattice of some context.

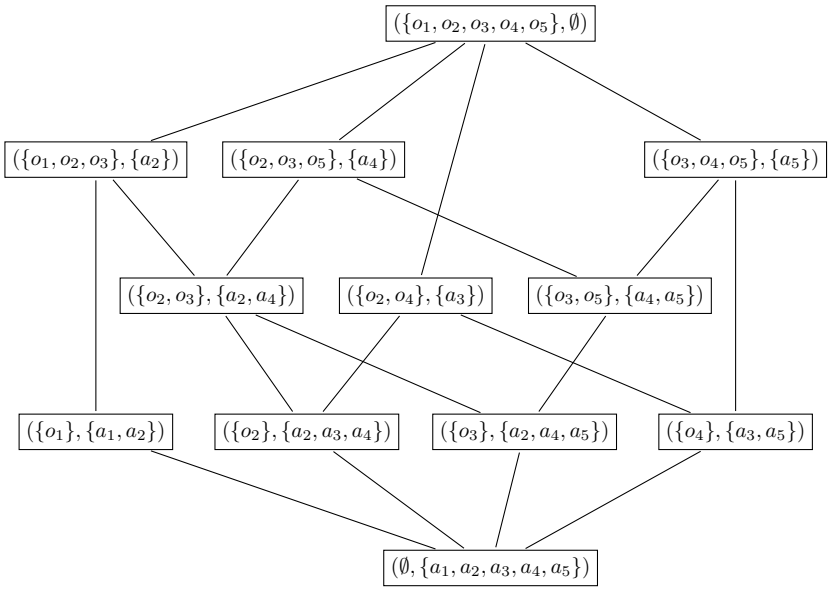


Figure 2 Concept lattice of the formal context depicted in Fig. 1.

▶ **Definition 3** (Implications). *In a formal context $(\mathcal{O}, \mathcal{A}, \mathcal{R})$, an* implication *is a pair of attribute sets (X, Y), usually noted $X \to Y$. An implication $X \to Y$* holds *in the context when $X' \subseteq Y'$ or, equivalently, $Y \subseteq X''$. In other words, the implication holds when all the objects described by X are also described by Y.*

In the Fig. 1 context, the implications $\{a_1\} \to \{a_1, a_2\}$ and $\{a_3, a_4\} \to \{a_2\}$ hold while the implication $\{a_3\} \to \{a_5\}$ does not. For simplicity's sake, we thereafter say "$X \to Y$" instead of "$X \to Y$ holds". Some implications can be inferred from others through Armstrong's axioms:

- if $Y \subseteq X$, then $X \to Y$ (*Reflexivity*)
- if $X \to Y$, then $X \cup Z \to Y \cup Z$ for all attribute sets Z (*Augmentation*)
- if $X \to Y$ and $Y \to Z$, then $X \to Z$. (*Transitivity*)

▶ **Definition 4** (Implication base). *An* implication base *of a formal context is an implication set $\mathcal{I}$ such that the set of implications that can be inferred from $\mathcal{I}$ through Armstrong's axioms is the set of all implications that hold in the context.*

Several implication bases with interesting properties exist in the literature [5, 4]. In this paper, we are interested in only one.

▶ **Definition 5** (Proper Premises). *Let $(\mathcal{O}, \mathcal{A}, \mathcal{R})$ be a formal context and a an attribute. A* proper premise of a *is an inclusion-minimal, non-closed attribute set X such that $X \to \{a\}$, i.e. there is no $Y \subset X$ such that $Y \to \{a\}$.*

In the Fig. 1 example, the set $\{a_2, a_3\}$ is a proper premise of the attribute a_4 as no proper subset of $\{a_2, a_3\}$ implies $\{a_4\}$. The set of all implications $X \to \{a\}$ where a is an attribute and X is one of its proper premises forms an implication base.

▶ **Definition 6** (Logical closure). *Let $\mathcal{I}$ be an implication base. The* logical closure *of an attribute set X by $\mathcal{I}$, denoted $X^{\mathcal{I}}$, is defined as the largest $Y \supseteq X$ such that $X \to Y$ can be inferred from $\mathcal{I}$.*

For instance, the logical closure of the attribute set $\{a_1, a_3\}$ by the implication base $\mathcal{I} = \{\{a_1\} \to \{a_2\}, \{a_2, a_3\} \to \{a_4\}\}$ is $\{a_1, a_3\}^{\mathcal{I}} = \{a_1, a_2, a_3, a_4\}$. The logical closure, as its name indicates, is a closure operator. If $\mathcal{C}$ is a formal context and $\mathcal{I}$ an implication base of $\mathcal{C}$, then $.^{\mathcal{I}} = .''$.

2.2 Metrics

A *metric* on a set S is a function of distance between the elements of S satisfying the following axioms:

- $f(x, x) = 0$
- $f(x, y) > 0$ when $x \neq y$, (positivity)
- $f(x, y) = f(y, x)$, (symmetry)
- $f(x, z) \leq f(x, y) + f(y, z)$. (triangular inequality)

In this paper, we make use of two families of metrics between vectors and sets so as to build our own metrics between FCA structures. The first is the well-known family of Minkowski distances between vectors $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$ defined as

$$D_p(X, Y) = \sqrt[p]{\sum_{i=1}^{n} |x_i - y_i|^p}.$$

The second is the family of normalised set-based analogues of Minkowski distances [13] defined, for two sets X and Y, as

$$d_{2,q}(X, Y) = \frac{\sqrt[q]{(|X| - |X \cap Y|)^q + (|Y| - |X \cap Y|)^q}}{|X \cap Y| + \sqrt[q]{(|X| - |X \cap Y|)^q + (|Y| - |X \cap Y|)^q}}.$$

In this paper, we chose to use the Minkowski distance for sets. Other, more usual distances (e.g. Hamming distance, or another edit distance) might also be interesting, and in most cases they can be plugged into the calculations in Section 3 to describe new families of distances between FCA structures. In [8], some distances are described directly for binary relations and for (semi-)lattices. However, formal contexts are special cases of binary relation (bipartite graphs) and concept lattices are more than lattices as they take two dimensions into account – objects and attributes. Therefore, we did not make use of those distances and instead chose to propose new ones.

3 Distances Between FCA Structures

3.1 Aim

We aim at proposing distances between FCA structures. This is not a brand new endeavor. Distances between formal contexts can be obtained by considering contexts as being any more widely known structures, such as bipartite graphs or hypergraphs, and using existing distances for these structures. Similarity measures between concept lattices have already been studied [9]. However, these are not sufficient. What we want is a set of three distances that can be used to compare two entities in their three different forms (context, lattice and implication base) and the knowledge of how these three distances relate to each others. In this paper, we suppose that all pairs of structures we compare use the same objects and attributes.

In this section, we define families of distances for each of the usual structures of FCA, and show that they are metrics. The three families are based on the normalised set-based analogues of Minkowski distances $d_{2,p}$ [13]. In Section 4, we provide experimental results on the interaction of those distances.

3.2 Distance Between Contexts

As we only consider contexts on the same sets of objects and attributes, the distance between the contexts depends only on their incidence relations. Hence, we define our distances between formal contexts as a distance between binary relations seen as sets of pairs.

▶ **Definition 7.** *Let* $\mathcal{C}_1 = (\mathcal{O}, \mathcal{A}, \mathcal{R}_1)$ *and* $\mathcal{C}_2 = (\mathcal{O}, \mathcal{A}, \mathcal{R}_2)$ *be two formal contexts. The* factual distance (FD) *between* $\mathcal{C}_1$ *and* $\mathcal{C}_2$ *is defined as*

$$FD_p(\mathcal{C}_1, \mathcal{C}_2) = d_{2,p}(\mathcal{R}_1, \mathcal{R}_2).$$

The two formal contexts depicted in Fig. 3 have a factual distance of ≈ 0.13.

	a_1	a_2	a_3	a_4
o_1	×	×	×	×
o_2	×	×	×	
o_3	×	×		
o_4	×			

	a_1	a_2	a_3	a_4
o_1	×	×	×	×
o_2	×	×		×
o_3	×	×		
o_4	×			

Figure 3 Two chain contexts. The two contexts have a factual distance of ≈ 0.13 with $p = 2$.

As $d_{2,p}$ is a metric, the factual distance is a metric.

3.3 Distance Between Concept Lattices

We consider concept lattices as pairs of functions that map sets of objects to the set of attributes they have in common and sets of attributes to the set of objects they all describe, i.e. we see concept lattices as representations of the derivation operators $\cdot'$. If (E, I) is a concept, then all subsets of E that are not subsets of lower neighbours in the lattice are mapped to I and reciprocally. This is notationally easier to express in terms of the derivation operators associated with the formal context of the lattice: object sets O are mapped to O'. As such, we define our distance between concept lattices as a distance between the derivation operators. For this reason, our distance makes use of the distances between the derivations of every element of the powerset of objects/attributes in both contexts. This has the added benefit of facilitating the comparison of concept lattices with different extents/intents.

▶ **Definition 8.** *Let $\mathcal{L}_1, \mathcal{L}_2$ be the two concept lattices of two contexts $\mathcal{C}_1$ and $\mathcal{C}_2$ with the same sets of objects $\mathcal{O}$ and attributes $\mathcal{A}$. We define the* lattice object distance *as*

$$LOD_{p,q}(\mathcal{L}_1, \mathcal{L}_2) = \frac{\sqrt[p]{\sum_{o \in \mathcal{O}} d_{2,q}(\mathcal{P}(\{o\}'^1), \mathcal{P}(\{o\}'^2))^p}}{\sqrt[p]{|\mathcal{O}|}}$$

and the lattice attribute distance *as*

$$LAD_{p,q}(\mathcal{L}_1, \mathcal{L}_2) = \frac{\sqrt[p]{\sum_{a \in \mathcal{A}} d_{2,q}(\mathcal{P}(\{a\}'^1), \mathcal{P}(\{a\}'^2))^p}}{\sqrt[p]{|\mathcal{A}|}}.$$

The conceptual distance *(CD) between $\mathcal{L}_1$ and $\mathcal{L}_2$ is then defined as*

$$CD_{p,q}(\mathcal{L}_1, \mathcal{L}_2) = min(LOD_{p,q}(\mathcal{L}_1, \mathcal{L}_2), LAD_{p,q}(\mathcal{L}_1, \mathcal{L}_2)).$$

In this definition, we chose to use the minimum between the lattice object distance and the lattice attribute distance. One could use the maximum between those two quantities to obtain a slightly different distance.

Figure 4 depicts the two chain concept lattices of the two contexts in Fig. 3. Even though they are isomorphic, their conceptual distance is ≈ 0.33 with $p = 2$ and $q = 1$.

In the following example, as well as in the experiments section (Section 4), we chose to fix $p = 2$ and $q = 1$ in our calculations. Fixing $q = 1$ creates an analogue to the Manhattan distance. Then, a generalised mean is computed over the set of attributes (resp. objects). With $p = 2$, we are using the *root mean square* deviation.

The conceptual distance takes its values in $[0, 1]$ and is a metric, satisfying the following axioms:

1. $CD(x, x) = 0$
2. $CD(x, y) > 0$ when $x \neq y$, (positivity)
3. $CD(x, y) = CD(y, x)$, (symmetry)
4. $CD(x, z) \leq CD(x, y) + CD(y, z)$. (triangular inequality)

These follow directly from the fact that $d_{2,p}$ is a metric:

1. because $\mathcal{L}_1 = \mathcal{L}_2 \Rightarrow \mathcal{P}(\{o\}'^1) = \mathcal{P}(\{o\}'^2)$ and $d_{2,q}(\mathcal{P}(\{o\}'^1), \mathcal{P}(\{o\}'^1)) = 0$
2. because $d_{2,q}(\mathcal{P}(\{o\}'^1), \mathcal{P}(\{o\}'^2)) > 0$
3. because $d_{2,q}(\mathcal{P}(\{o\}'^1), \mathcal{P}(\{o\}'^2)) = d_{2,q}(\mathcal{P}(\{o\}'^2), \mathcal{P}(\{o\}'^1))$
4. because

$$d_{2,q}(\mathcal{P}(\{o\}'^1), \mathcal{P}(\{o\}'^3)) \leq d_{2,q}(\mathcal{P}(\{o\}'^1), \mathcal{P}(\{o\}'^2)) + d_{2,q}(\mathcal{P}(\{o\}'^2), \mathcal{P}(\{o\}'^3))$$

Computing the conceptual distance is quite easy: for each object o, find the concepts with the smallest extents that contain o in both lattices. Their intents are $\{o\}'^1$ and $\{o\}'^2$ respectively. Then, computing $d_{2,q}(\mathcal{P}(\{o\}'^1), \mathcal{P}(\{o\}'^2))$ is straightforward. Algorithm 1 follows this principle. Finding the concept with the smallest extent that contains an object in the lattice $\mathcal{L}_i$ can be done in $O(|\mathcal{L}_i|)$ so the time complexity of Algorithm 1 is in $O((|\mathcal{O}| + |\mathcal{A}|) \times max(|\mathcal{L}_1|, |\mathcal{L}_2|))$.

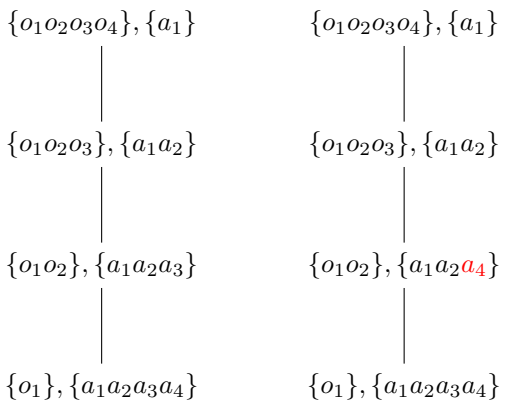


Figure 4 The two concept lattices of the Fig. 3 contexts. These have a conceptual distance of $CD_{2,1} \approx 0.33$ with $p = 2$ and $q = 1$. A small difference in the intents leads to a non-zero distance, even on isomorphic lattices with the same extents.

Algorithm 1 $CD_{p,q}$.

Input: Two concept lattices $\mathcal{L}_1$ and $\mathcal{L}_2$ with the same sets of objects $\mathcal{O}$ and attributes $\mathcal{A}$, p and q
Output: $CD_{p,q}(\mathcal{L}_1, \mathcal{L}_2)$

$LOD = 0$
foreach *object* $o \in \mathcal{O}$ **do**
  $LOD = LOD + (\sqrt[q]{(2^{|\{o\}'^1|} - 2^{|\{o\}'^1 \cap \{o\}'^2|})^q + (2^{|\{o\}'^2|} - 2^{|\{o\}'^1 \cap \{o\}'^2|})^q})^p$
$LOD = \sqrt[p]{LOD} / \sqrt[p]{|\mathcal{O}|}$
$LAD = 0$
foreach *attribute* $a \in \mathcal{A}$ **do**
  $LAD = LAD + (\sqrt[q]{(2^{|\{a\}'^1|} - 2^{|\{a\}'^1 \cap \{a\}'^2|})^q + (2^{|\{a\}'^2|} - 2^{|\{a\}'^1 \cap \{a\}'^2|})^q})^p$
$LAD = \sqrt[p]{LAD} / \sqrt[p]{|\mathcal{A}|}$
return $min(LOD, LAD)$

3.4 Distance between Implication Bases

For our distance between implication bases, we consider implication bases as functions mapping attribute sets X to attribute sets $Y = \{y \mid X \to \{y\}\}$, i.e. we see implication bases as representations of the closure operator $\cdot''$ on attributes. Note that, from Armstrong's axioms, we can infer that

$$X \to Y \Leftrightarrow \forall y \in Y, X \to \{y\}.$$

▶ **Definition 9.** *Let $\mathcal{I}_1, \mathcal{I}_2$ be two implication bases on the same attribute set $\mathcal{A}$. For an attribute $a \in \mathcal{A}$ and an implication base $\mathcal{I}$, we denote by $\mathcal{I}^a = \{X \mid a \in X^{\mathcal{I}}\}$ the set of attributes sets that imply a. The* logical distance *(LD) between $\mathcal{I}_1$ and $\mathcal{I}_2$ is then defined as*

$$LD_{p,q}(\mathcal{I}_1, \mathcal{I}_2) = \frac{\sqrt[p]{\sum_{a \in \mathcal{A}} d_{2,q}(\mathcal{I}_1^a, \mathcal{I}_2^a)^p}}{\sqrt[p]{|\mathcal{A}|}}.$$

Fig. 5 depicts the two proper premises implication bases of the contexts in Fig 3. These two implication bases have a logical distance of ≈ 0.23. Indeed, the attribute a_3 is implied by all supersets of $\{a_4\}$ only in the first context and the attribute a_4 is implied by all supersets of $\{a_3\}$ only in the second context.

$\{a_4\} \to \{a_2, a_3\}$ $\quad$ $\{a_4\} \to \{a_2\}$
$\{a_3\} \to \{a_2\}$ $\quad$ $\{a_3\} \to \{a_2, a_4\}$
$\emptyset \to \{a_1\}$ $\quad$ $\emptyset \to \{a_1\}$

Figure 5 The two proper premises bases of the Fig. 3 contexts. The logical distance, with parameters $p = 2$ and $q = 1$, between these two bases is ≈ 0.23.

The logical distance takes its values in $[0, 1]$ and is a metric, satisfying the following axioms:

1. $LD(x, x) = 0$
2. $LD(x, y) > 0$ when $x \neq y$, (positivity)
3. $LD(x, y) = LD(y, x)$, (symmetry)
4. $LD(x, z) \leq LD(x, y) + LD(y, z)$. (triangular inequality)

Just as those for the conceptual distance, these axioms follow from the fact that $d_{2,q}$ is a metric.

To compute the logical distance, one requires the knowledge of all the attribute sets X that imply a given attribute a. This is not explicitly contained in implication bases and retrieving it is the computationally most expensive part of computing the distance. We propose Algorithm 3 to compute the logical distance. We assume that the implication bases are proper premises bases. If this is not the case, other bases can be converted to proper premises bases in output-polynomial time [16].

The algorithm treats each attribute a separately. The first step is to compute the cardinalities of $\mathcal{I}_1^a \cap \mathcal{I}_2^a$, $\mathcal{I}_1^a$ and $\mathcal{I}_2^a$. To do so, we start with computing the attribute sets P that are minimal such that $P \to \{a\}$ in both implication bases ($commonPremises$). The cardinality of $\mathcal{I}_1^a \cap \mathcal{I}_2^a$ is then the number of attributes sets that contain one of the elements of $commonPremises$. To obtain it, we compute the union closure U_c of the set $commonPremises$, i.e. the minimal sets of attributes sets such that $X, Y \in U_c \Rightarrow X \cup Y \in U_c$. The set U_c ordered by set-inclusion forms a lattice. We use Algorithm 2 to associate to each element x of U_c the number of attribute sets that contain x but not its supersets in U_c, *i.e.* the size of the equivalence classes in the union-closed lattice. Algorithm 3 then sums those numbers (sum_c) to obtain the numbers of attribute sets containing one of the corresponding premises. The same approach is applied to compute the cardinalities of $\mathcal{I}_1^a$ and $\mathcal{I}_2^a$. As the size of the union closure is bounded by $2^{|\mathcal{A}|}$ (when all singletons are premises), the worst case complexity of Algorithm 3 is in $O(|\mathcal{A}| \times 2^{|\mathcal{A}|})$.

4 Experiments

In all these experiments, we used parameters $p = 2$ and $q = 1$ for all distances. A Python module[1] containing the three distances, as well as the script for the experiments themselves[2], are publicly available.

4.1 Correlation Between distances

The first question that may come to mind is "how do these distances compare to each others?". Let us consider the Fig. 6 example representing three contexts $\mathcal{C}_{B_3}$, $\mathcal{C}_{M3}$ and $\mathcal{C}_{N5}$ corresponding respectively to Boolean ($\mathcal{B}_3$), M_3 and N_5 concept lattices, and their associated proper premises implication bases $\mathcal{I}_{B_3}$, $\mathcal{I}_{M_3}$ and $\mathcal{I}_{N_5}$. We compute the factual, conceptual and logical distances between $\mathcal{C}_{\mathcal{B}_3}$ and the other two and obtain the following results:

[1] `https://github.com/Authary/FCAD`
[2] `https://github.com/Authary/experiments_distances_fca`

Algorithm 2 *sizeEQ.*

Input: A set U of premises
Output: $sizeEQ(U)$
Build a dictionary D mapping each premise P in U to the set of premises $P_2 \supset P$
$sum \leftarrow 0$
$over \leftarrow false$
while $over = false$ **do**
  $over \leftarrow true$
  foreach *premise* P *in* U **do**
    if *all* $P_2 \in D(P)$ *have been tagged* **then**
      $|P^{\equiv}| \leftarrow 2^{|\mathcal{A}|-|P|-1} - \sum_{P_2 \in D(P)} |P_2^{\equiv}|$
      Tag P
      $sum \leftarrow sum + |P^{\equiv}|$
      $over \leftarrow false$
return sum

Algorithm 3 *LD.*

Input: Two implication bases $\mathcal{I}_1$ and $\mathcal{I}_2$, p, q
Output: $LD_{p,q}(\mathcal{I}_1, \mathcal{I}_2)$
$Result \leftarrow 0$
foreach *attribute* a **do**
  $U_1 = unionClosure(\{P \mid P \to \{a\} \in \mathcal{I}_1\})$
  $U_2 = unionClosure(\{P \mid P \to \{a\} \in \mathcal{I}_2\})$
  $commonPremises = min(\{P_1 \cup P_2 \mid P_1 \to \{a\} \in \mathcal{I}_1, P_2 \to \{a\} \in \mathcal{I}_2\})$
  $U_c = unionClosure(commonPremises)$
  $sum_c = sizeEQ(U_c)$
  $sum_1 = sizeEQ(U_1)$
  $sum_2 = sizeEQ(U_2)$
  $Result = Result + (\sqrt[q]{(sum_1 - sum_c)^q + (sum_2 - sum_c)^q})^p$
return $\sqrt[p]{Result} / \sqrt[p]{|\mathcal{A}|}$

$$FD_{2,1}(\mathcal{C}_{\mathcal{B}_3}, \mathcal{C}_{M_3}) = 1 > FD_{2,1}(\mathcal{C}_{\mathcal{B}_3}, \mathcal{C}_{N_5}) = 0.51$$

$$CD_{2,1}(\mathcal{B}_3, M_2) = 0.75 > CD_{2,1}(\mathcal{B}_3, N_5) = 0.54$$

$$LD_{2,1}(\mathcal{I}_{\mathcal{B}_3}, \mathcal{I}_{M_3}) = 0.20 < LD_{2,1}(\mathcal{I}_{\mathcal{B}_3}, \mathcal{I}_{N_5}) = 0.22$$

We observe that $\mathcal{C}_{N_5}$ is factually and conceptually closer to $\mathcal{C}_{\mathcal{B}_3}$ while $\mathcal{C}_{M_3}$ is logically closer. The three distances therefore do not always agree. Are they at least correlated? In order to answer this question, we generated structures in different ways. As contexts are the easiest structure to manipulate, we explored the following approaches:

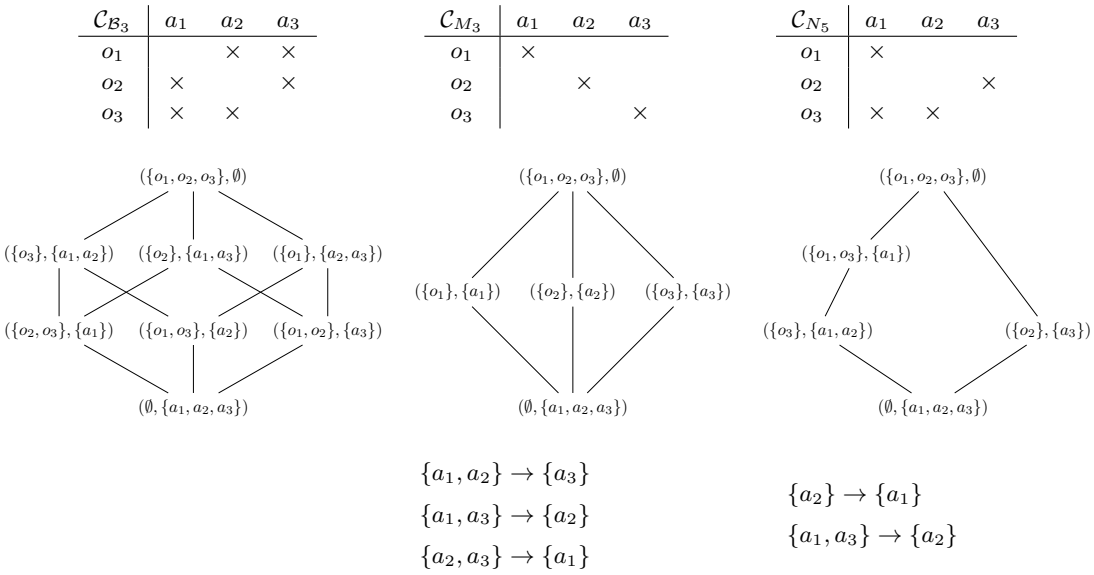


Figure 6 Three formal contexts $\mathcal{C}_{\mathcal{B}_3}$, $\mathcal{C}_{M_3}$ and $\mathcal{C}_{N_5}$ corresponding respectively to the Boolean, $N5$ and $M3$ concept lattices, and their associated implication bases (proper premises).

- starting from a full context and iteratively removing crosses either row by row or randomly chosen
- randomly generating contexts
  - by having each cross with a probability p
  - by randomly flipping each cross of a reference context with a probability p
- generating *pseudo-real* contexts by sampling real data.

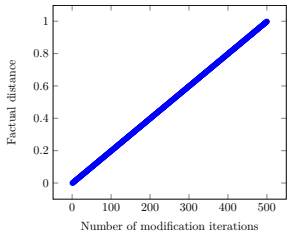
In each case, we computed three correlation coefficients, Pearson, Spearman and Kendall's τ.

4.2 Iterative Emptying of a Full Context

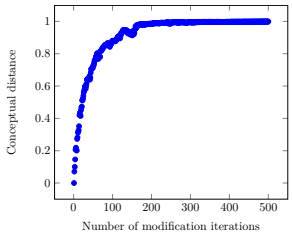
In a first series of experiment, we explored whether we could control the variation of the distances by purposefully modifying a context.

4.2.1 Random Removal of Crosses

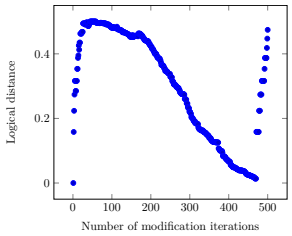
In a first experiment, we started with a full formal context $(\mathcal{O}, \mathcal{A}, \mathcal{R})$ with $|\mathcal{O}| = 50$, $|\mathcal{A}| = 10$ and $\mathcal{R} = \mathcal{O} \times \mathcal{A}$. We iteratively removed random crosses one by one and, at each step, computed the three distances between the current context and the initial one. Fig. 7 presents the progression of the three distances. We observe that while the factual distance increases linearly (which was expected), the other two distances behave very differently. In particular, while the factual and conceptual distances are increasing, the logical distance varies cyclically. This is because in both the full and the empty contexts each attribute implies all the others. Fig. 7 also depicts three diagrams illustrating respectively the relation between the factual (x-axis) and conceptual (y-axis) distances, the relation between the factual (x-axis) and the logical (y-axis) distances and the relation between the conceptual (x-axis) and the logical (y-axis) relations. Additionally, the figure also presents the values of the three correlation coefficients, Pearson, Spearman and Kendall's τ. In this experiment, the values of the three distances are fairly correlated.



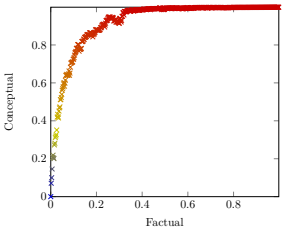
(a) Progression of the factual distance.



(b) Progression of the conceptual distance.

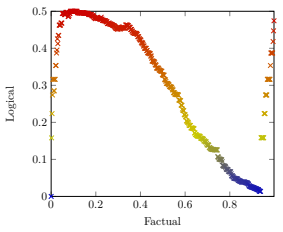


(c) Progression of the logical distance.



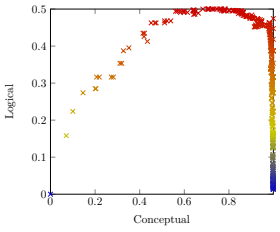
Pearson	Spearman	Kendall
0.70	0.97	0.92

(d) Factual (x-axis) and conceptual (y-axis) distances to the full context of contexts obtained by iteratively removing random crosses.



Pearson	Spearman	Kendall
−0.88	−0.88	−0.81

(e) Factual (x-axis) and logical (y-axis) distances to the full context of contexts obtained by iteratively removing random crosses.



Pearson	Spearman	Kendall
−0.43	−0.88	−0.75

(f) Conceptual (x-axis) and logical (y-axis) distances to the full context of contexts obtained by iteratively removing random crosses.

Figure 7 Iterative removal of crosses from a full context: random crosses.

4.2.2 Iterative Removal of Specific Crosses

In the second set of experiments, we tried to control the variation of the conceptual and logical distances by removing specific crosses in the starting context instead of random ones. We removed crosses row by row so that, *i.e.*, at every step, at most one row is neither full nor empty. Fig. 8 depicts the progression of the three distances and the relations between pairs of distances. By modifying the contexts in this way, the factual distance to the original full context still increases linearly but the conceptual distance increases more slowly. More interestingly, the values of the logical distance repeat every 10 iterations, *i.e.* the number of attributes:

0.1581 0.1118 0.0684 0.0395 0.0220 0.0121 0.0065 0.0034 0.0018 0.0009

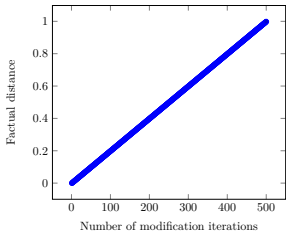
This phenomenon is due to the fact that empty rows do not impact the implications, so every 10 iterations results in a context equivalent to the full context with a single row missing some crosses. In this experiment, only the factual and conceptual distances are correlated.

4.3 Random Contexts

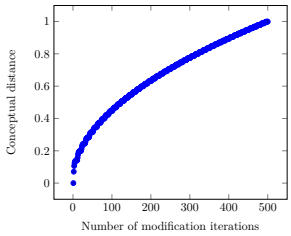
In a second series of experiments, we explored how random generations and modifications of contexts impact the distances, with a focus on how correlated the distances are.

4.3.1 Randomly Generated Contexts

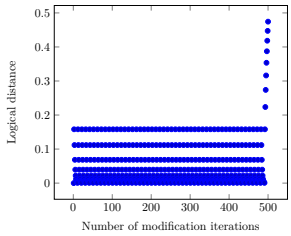
We randomly generated 1500 pairs (A, B) of formal contexts with 50 objects and 10 attributes, with a pair (o, a) having a probability 0.3 of being in the incidence relation. We then computed the distances between the contexts (resp. their associated lattices and implication bases) in each



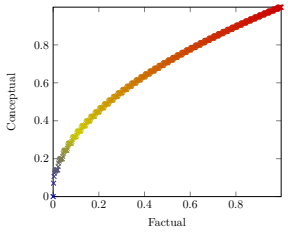
(a) Progression of the factual distance.



(b) Progression of the conceptual distance.

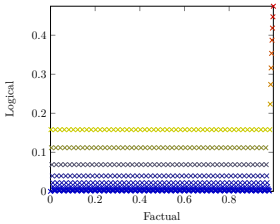


(c) Progression of the logical distance.



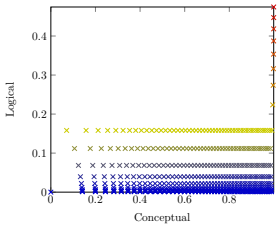
Pearson	Spearman	Kendall
0.98	1	0.99

(d) Factual (x-axis) and conceptual (y-axis) distances to the full context of contexts obtained by iteratively removing minimally aligned crosses.



Pearson	Spearman	Kendall
0.12	0.04	0.02

(e) Factual (x-axis) and logical (y-axis) distances to the full context of contexts obtained by iteratively removing minimally aligned crosses.



Pearson	Spearman	Kendall
0.10	0.04	0.02

(f) Conceptual (x-axis) and logical (y-axis) distances to the full context of contexts obtained by iteratively removing minimally aligned crosses.

Figure 8 Iterative removal of crosses from a full context: minimally aligned crosses.

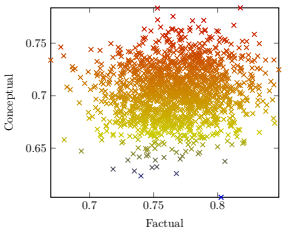
pair. Fig. 9 depicts three diagrams illustrating the relation between the factual (x-axis) and logical (y-axis) distances, the relation between the factual (x-axis) and the conceptual (y-axis) distances and the relation between the conceptual (x-axis) and the logical (y-axis) relations.

We observe that the three distances appear to be pairwise independent when the contexts are randomly generated in such a way. Fig. 9 also depicts the values of the three correlation coefficients, Pearson, Spearman and Kendall's τ. Their values confirm the independence, with the factual and conceptual distances being very slightly less independent. Note that Pearson measures linear correlation, Spearman assesses monotonic relationships and Kendall's τ measures rank correlation.

Interestingly, all factual distances are between 0.67 and 0.85, suggesting that random generation produces contexts that are far apart.

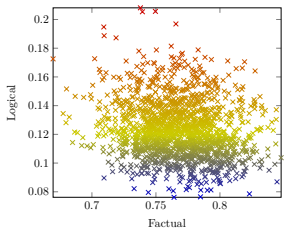
4.3.2 Randomly Modified Contexts

In a second batch of experiments, we generated 1500 other pairs of contexts such that A is a randomly generated context and B is obtained by randomly modifying A. All contexts contain 50 objects and 10 attributes. The contexts A were generated with a probability 0.3 for each cross. The modified contexts were obtained through the following algorithm: for each $(object, attribute)$ pair, with a probability 0.05, remove the pair from the incidence relation if it belongs to it or add it if it does not. We then computed the distances between the contexts (resp. their associate lattices and implication bases) in each pair. Fig. 10 depicts three diagrams illustrating the relation between the factual (x-axis) and logical (y-axis) distances, the relation between the factual (x-axis) and the conceptual (y-axis) distances and the relation between the conceptual (x-axis) and the logical (y-axis) relations. Fig. 10 also depicts the values of the three correlation coefficients, Pearson, Spearman and Kendall's τ.



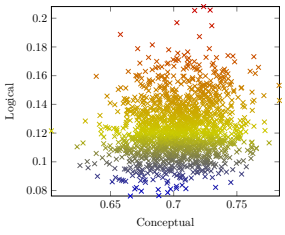
Pearson	Spearman	Kendall
0.09	0.10	0.06

(a) Factual (x-axis) and conceptual (y-axis) distances of random contexts.



Pearson	Spearman	Kendall
−0.16	−0.15	−0.10

(b) Factual (x-axis) and logical (y-axis) distances of random contexts.

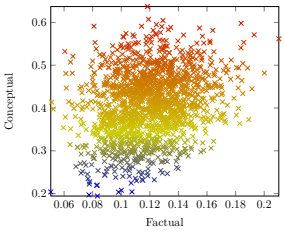


Pearson	Spearman	Kendall
0.14	0.13	0.09

(c) Conceptual (x-axis) and logical (y-axis) distances of random contexts.

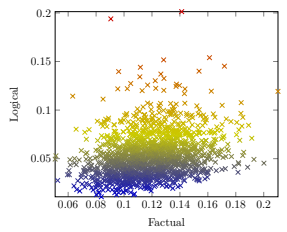
Figure 9 Randomly generated contexts : correlation between the distance measures.

Visually, we observe some slight positive correlation between the factual and conceptual distances and between the factual and logical distances. This is in opposition to the previous experiment with randomly generated contexts. All factual distances are below 0.2, suggesting that our modification algorithms successfully produces contexts that are close together. This result, together with the previous one on randomly generated contexts, hints at the correlations between the factual distance and the others being stronger for very close contexts.



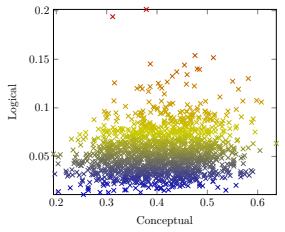
Pearson	Spearman	Kendall
0.25	0.24	0.16

(a) Factual (x-axis) and conceptual (y-axis) distances to a reference context of contexts obtained by randomly modifying the reference.



Pearson	Spearman	Kendall
0.26	0.28	0.19

(b) Factual (x-axis) and logical (y-axis) distances to a reference context of contexts obtained by randomly modifying the reference.



Pearson	Spearman	Kendall
0.14	0.13	0.09

(c) Conceptual (x-axis) and logical (y-axis) distances to a reference context of contexts obtained by randomly modifying the reference.

Figure 10 Randomly modified contexts: correlation between the distance measures (factual,logical), (factual,conceptual) and (conceptual,logical).

4.3.3 Variation of Correlation Relative to the Factual Distance

In the previous experiments, the distances seemed to be more correlated for low factual distances. This hinted at differences in correlations depending on factual distances. Let us check whether this is really the case. For this experiment, we generated pairs (A, B) of contexts such that A is a 50×10 randomly generated context and B is obtained by randomly flipping the truth value of each $(object, attribute)$ pair in A with a probability p. We made p vary from 0.025 to 0.5 with 0.025 increments. For each value of p, we generated 1000 pairs of contexts and computed the three distances between A and B. We then computed the three correlation coefficients for each pair of distances. Fig. 11 presents these correlation values (y-axes) for the different values of p (x-axes).

As a higher p results in factually more distant contexts, we observe that the correlations between the factual distance and the other two decrease when p increases. This confirms our previous observation that these distances are more correlated for low factual distances.

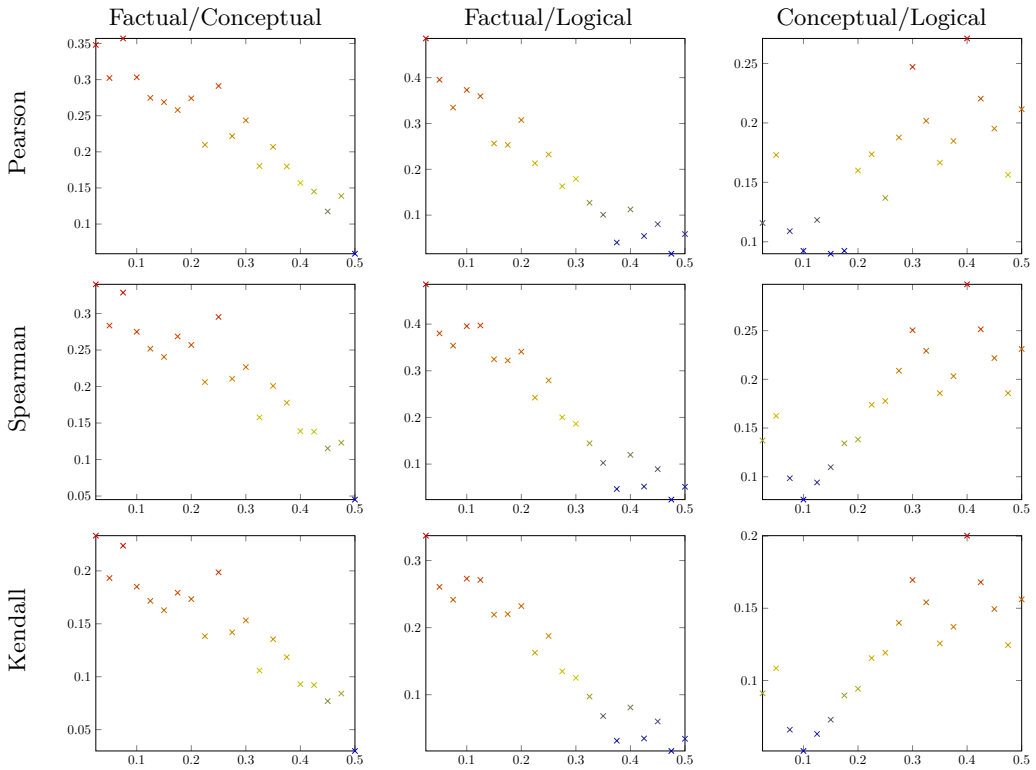


Figure 11 Correlation of the three distance measures for different values of probabilities used in the modification of contexts. Higher probabilities means higher factual distances.

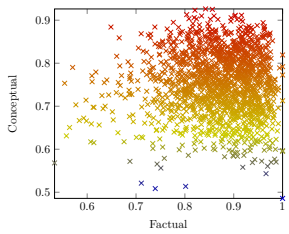
4.4 Pseudo-real data

In this experiment, we sampled real datasets to create pseudo-real contexts. We used the Mushroom dataset [20] and the data from *A statistical analysis of the work of Bob Ross*[3]. We generate pseudo-real data to minimise the risk of interference from the generation method, since generating random formal context is not a trivial task, as shown in [6]. Our sample strategy is the following: we uniformly sample a number n_o of objects along the dimension of objects and a number n_a of attributes along the dimension of attributes. Then, we keep the portion of the incidence relation that correspond to those objects and attributes. Finally, we rename the objects and attributes from 1 to, respectively, n_o and n_a so that two sampled contexts may have the same object and attribute sets. For this set of experiments, our goal is, again, to study the correlations between pairs of distances.

We sampled the Mushroom dataset (8124 objects and 119 attributes) into 1500 smaller contexts (56 objects and 11 attributes) and computed the three distances between pairs of such sampled contexts. The average density of the sampled contexts is 0.19.

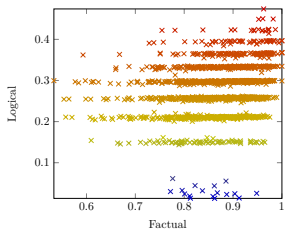
[3] `https://fivethirtyeight.com/features/a-statistical-analysis-of-the-work-of-bob-ross/`

The results are shown in Fig. 12. They show no strong correlation between the distances, except between the factual distance and the logical distance. In this experiment, the logical distance only takes values around a few quantities, unlike in the case of the removal of crosses where the logical distance had exactly the same 10 values.



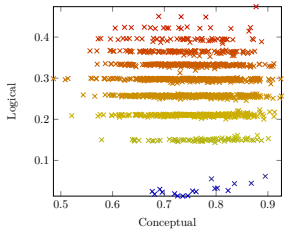
Pearson	Spearman	Kendall
0.016	-0.02	-0.01

(a) Factual (x-axis) and conceptual (y-axis) distances of random contexts.



Pearson	Spearman	Kendall
0.26	0.34	0.23

(b) Factual (x-axis) and logical (y-axis) distances of random contexts.



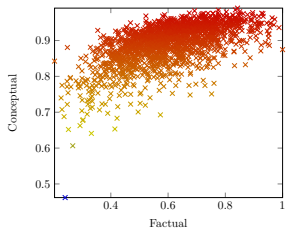
Pearson	Spearman	Kendall
-0.10	-0.12	-0.08

(c) Conceptual (x-axis) and logical (y-axis) distances of random contexts.

Figure 12 Randomly sampled contexts from the Mushroom dataset : correlation between the distance measures.

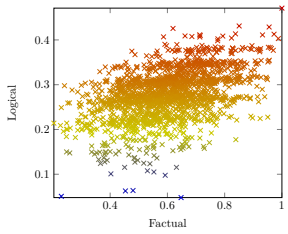
We sampled the Bob Ross dataset as well. It is a 433×133 context based on the apparition, or not, of some features in Bob Ross' paintings, for each episode of The Joy Of Painting. We sampled it into 1500 smaller, 44×10 contexts. In this case, the average density is 0.50.

The results are shown in Fig. 13. In this case, each pair of distances shows a stronger correlation, especially between the factual and conceptual distances. The difference between the two experiments might be related to the density of the contexts, or to intrinsic differences between the data sets, as Bob Ross' paintings usually had the same few elements appearing together in most episodes.



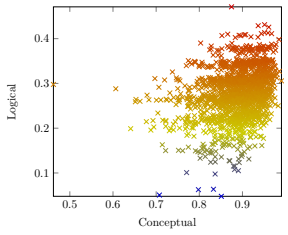
Pearson	Spearman	Kendall
0.60	0.60	0.42

(a) Factual (x-axis) and conceptual (y-axis) distances of random contexts.



Pearson	Spearman	Kendall
0.46	0.45	0.31

(b) Factual (x-axis) and logical (y-axis) distances of random contexts.



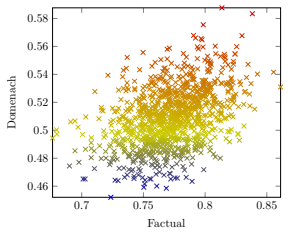
Pearson	Spearman	Kendall
0.30	0.31	0.21

(c) Conceptual (x-axis) and logical (y-axis) distances of random contexts.

Figure 13 Randomly sampled contexts from the Bob Ross dataset : correlation between the distance measures.

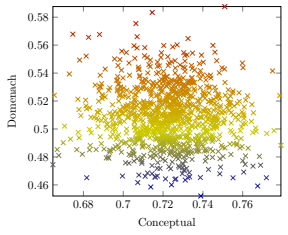
4.5 Comparison with Domenach's Dissimilarity Measure

Domenach's dissimilary measure is based on the *overhanging* relation [10] between sets of objects. Two sets are overhanged if one is a subset of the other and their closures are different. To compute a distance between concept lattices, Domenach defines two matrices, M_1 and M_2, based on the overhanging relation of pairs of objects in each concept lattice. The distance is then based on the L_1 norms of those matrices: $\frac{||M_1 - M_2||}{||M_1|| + ||M_2||}$.



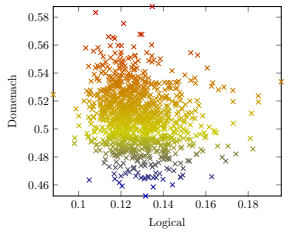
Pearson	Spearman	Kendall
0.41	0.40	0.27

(a) Factual distance (x-axis) and Domenach's dissimilarity (y-axis) of random contexts.



Pearson	Spearman	Kendall
−0.05	−0.03	−0.02

(b) Conceptual distance (x-axis) and Domenach's dissimilarity (y-axis) of random contexts.



Pearson	Spearman	Kendall
−0.17	−0.19	−0.13

(c) Logical distance (x-axis) and Domenach's dissimilarity (y-axis) of random contexts.

Figure 14 Randomly generated contexts : correlation between our distance measures and Domenach's dissimilarity.

We compared our distances with Domenach's dissimilarity measure on 1000 pairs of randomly generated contexts. Fig. 14 depicts the results. We observe that Domenach's dissimilarity measure is independent of our conceptual distance and slightly correlated with our factual distance.

5 Conclusion and Perspectives

We presented three distance families between the most important structures in formal concept analysis, *i.e.* formal contexts, concept lattices and implication bases. These structures represent three complementary points of view on the information contained in formal context: the factual, conceptual and logical points of views. We see the distances we studied in this paper as a first step towards the simultaneous exploitation of the three points of view in the analysis of data.

The applications could be distance-based machine learning, both supervised and unsupervised, or the measurement of the complexity of multidimensional data.

From our point of view, the applications for the work can be multiple. For example, one can study the trajectory and dynamics of a process such as Relational Concept Analysis [21] or Attribute Exploration [11], where the concept lattice and the implication base are built iteratively. Note that since our distances consider contexts with identical dimensions, studying the trajectory of the process would have to be *a posteriori*, and not in an online manner.

Our experiments indicate that, of our distances, only the factual distance is (barely) correlated with the other two and that their correlations depend on the factual distance. In particular, there is no correlation when generating contexts by coin-flipping. As this method has been shown not to produce truly random lattices [6], and the correlation in pseudo-real data is slightly higher, it remains to be seen whether our results are a product of the generation method. In any case, we believe that this is interesting because the concept lattice is supposed to be a lossless representation of the information contained in the formal context but small variations in the context can produce large variations in the lattice. The variation of the correlation w.r.t. other distances should also be studied once we better understand how to control the conceptual and logical distances in the generation of data. Our experiments also highlight the need to study the metric spaces induced by the distances, and their relations, as experimental results are insufficient.

Future work includes the extension of these distances to contexts defined on different sets of objects and attributes, and to the polyadic concept analysis framework.

Resource Availability Statement

The source code for the computation of the distances is hosted at `https://github.com/Authary/FCAD`. The source code for the experiments if hosted at `https://github.com/Authary/experiments_distances_fca`.

References

1 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 23*, pages 611–621. Springer, 2017. doi:10.1007/978-3-319-60438-1_60.

2 Alexandre Bazin, Jessie Galasso, and Giacomo Kahn. Polyadic relational concept analysis. *International Journal of Approximate Reasoning*, 164:109067, 2024. doi:10.1016/j.ijar.2023.109067.

3 Alexandre Bazin, Marianne Huchard, and Pierre Martin. Towards analyzing variability in space and time of products from a product line using triadic concept analysis. In *Proceedings of the 27th ACM International Systems and Software Product Line Conference-Volume B*, pages 85–89, 2023. doi:10.1145/3579028.3609019.

4 Karell Bertet, Christophe Demko, Jean-François Viaud, and Clément Guérin. Lattices, closures systems and implication bases: A survey of structural aspects and algorithms. *Theoretical Computer Science*, 743:93–109, 2018. doi:10.1016/j.tcs.2016.11.021.

5 Karell Bertet and Bernard Monjardet. The multiple facets of the canonical direct unit implicational basis. *Theoretical Computer Science*, 411(22-24):2155–2166, 2010. doi:10.1016/j.tcs.2009.12.021.

6 Daniel Borchmann and Tom Hanika. Some experimental results on randomly generating formal contexts. In *CLA*, volume 1624, pages 57–69, 2016. URL: https://ceur-ws.org/Vol-1624/paper5.pdf.

7 Victor Codocedo and Amedeo Napoli. Formal concept analysis and information retrieval–a survey. In *International Conference on Formal Concept Analysis*, pages 61–77. Springer, 2015. doi:10.1007/978-3-319-19545-2_4.

8 Elena Deza, Michel Marie Deza, Michel Marie Deza, and Elena Deza. *Encyclopedia of distances*. Springer, 2009.

9 Florent Domenach. Similarity measures of concept lattices. In *Data Science, Learning by Latent Structures, and Knowledge Discovery*, pages 89–99. Springer, 2015. doi:10.1007/978-3-662-44983-7_8.

10 Florent Domenach and Bruno Leclerc. Closure systems, implicational systems, overhanging relations and the case of hierarchical classification. *Mathematical Social Sciences*, 47(3):349–366, 2004. doi:10.1016/j.mathsocsci.2003.09.008.

11 Bernhard Ganter, Sergei Obiedkov, Bernhard Ganter, and Sergei Obiedkov. Attribute exploration. *Conceptual exploration*, pages 125–185, 2016. doi:10.1007/978-3-662-49291-8_4.

12 Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis: Mathematical Foundations*. Springer Science & Business Media, 1999.

13 Kathy J Horadam and Michael A Nyblom. Distances between sets based on set commonality. *Discrete Applied Mathematics*, 167:310–314, 2014. doi:10.1016/j.dam.2013.10.037.

14 Marianne Huchard, Pierre Martin, Emile Muller, Pascal Poncelet, Vincent Raveneau, and Arnaud Sallaberry. Rcaviz: Exploratory search in multi-relational datasets represented using relational concept analysis. *International Journal of Approximate Reasoning*, page 109123, 2024. doi:10.1016/j.ijar.2024.109123.

15 Priscilla Keip, Alain Gutierrez, Marianne Huchard, Florence Le Ber, Samira Sarter, Pierre Silvie, and Pierre Martin. Effects of input data formalisation in relational concept analysis for a data model with a ternary relation. In *International Conference on Formal Concept Analysis*, pages 191–207. Springer, 2019. doi:10.1007/978-3-030-21462-3_13.

16 Claudio L Lucchesi and Sylvia L Osborn. Candidate keys for relations. *Journal of Computer and System Sciences*, 17(2):270–279, 1978. doi:10.1016/0022-0000(78)90009-0.

17 Jonas Poelmans, Dmitry I Ignatov, Sergei O Kuznetsov, and Guido Dedene. Formal concept analysis in knowledge processing: A survey on applications. *Expert systems with applications*, 40(16):6538–6560, 2013. doi:10.1016/j.eswa.2013.05.009.

18 Jonas Poelmans, Dmitry I Ignatov, Sergei O Kuznetsov, and Guido Dedene. Fuzzy and rough formal concept analysis: a survey. *International Journal of General Systems*, 43(2):105–134, 2014. doi:10.1080/03081079.2013.862377.

19 Jonas Poelmans, Sergei O Kuznetsov, Dmitry I Ignatov, and Guido Dedene. Formal concept analysis in knowledge processing: A survey on models and techniques. *Expert systems with applications*, 40(16):6601–6623, 2013. doi:10.1016/j.eswa.2013.05.007.

20 UCI Machine Learning Repository. Mushroom. UCI Machine Learning Repository, 1981. DOI: https://doi.org/10.24432/C5959T.

21 Mohamed Rouane-Hacene, Marianne Huchard, Amedeo Napoli, and Petko Valtchev. Relational concept analysis: mining concept lattices from multi-relational data. *Annals of Mathematics and Artificial Intelligence*, 67:81–108, 2013. doi:10.1007/s10472-012-9329-3.

22 Sebastian Rudolph, Christian Săcărea, and Diana Troancă. Conceptual navigation for polyadic formal concept analysis. In *IFIP International Workshop on Artificial Intelligence for Knowledge Management*, pages 50–70. Springer, 2016. doi:10.1007/978-3-319-92928-6_4.

23 Gerd Stumme, Dominik Dürrschnabel, and Tom Hanika. Towards ordinal data science. *Transactions on Graph Data and Knowledge (TGDK)*, 2023. doi:10.4230/TGDK.1.1.6.

24 George Voutsadakis. Polyadic concept analysis. *Order*, 19:295–304, 2002. doi:10.1023/A:1021252203599.