Realizability of Rectangular Euler Diagrams

Dominik Dürrschnabel[1,2][0000−0002−0855−4185] and Uta
Priss[3][0000−0003−0375−429X]

[1] Knowledge & Data Engineering Group, University of Kassel, Kassel, Germany
[2] Interdisciplinary Research Center for Information System Design, Kassel, Germany
duerrschnabel@cs.uni-kassel.de
[3] Faculty of Computer Science, Ostfalia University of Applied Sciences, Wolfenbüttel,
Germany
u.priss@ostfalia.de

Abstract Euler diagrams are a tool for the graphical representation
of set relations. Due to their simple way of visualizing elements in the
sets by geometric containment, they are easily readable by an inexperienced reader. Euler diagrams where the sets are visualized as aligned
rectangles are of special interest. In this work, we link the existence of
such rectangular Euler diagrams to the order dimension of an associated
order relation. For this, we consider Euler diagrams in one and two dimensions. In the one-dimensional case, this correspondence provides us
with a polynomial-time algorithm to compute the Euler diagrams, while
the two-dimensional case is linked to an NP-complete problem which we
approach with an exponential-time algorithm.

Keywords: Order dimension · Rectangular Euler diagrams · Formal
concept analysis · Containment orders.

1 Introduction

Set diagrams are commonly used for the visualization of sets in set theory.
Thereby, two kinds of set diagrams have proven themselves to be useful. Venn
diagrams, named after John Venn (1834-1923), can be used to visualize elementary set theory by illustrating the sets as geometric objects. Hereby, frequently,
the shapes of circles or ellipsis are used. They show all possible logical relations
as intersections of their regions, i.e., in the case of n sets there are 2^n different
intersecting zones. Because of this exponential nature, they reach the limit of
what is perceived as readable for values of n as small as 4 or 5. Euler diagrams
however are not subject to the same restriction. Only regions which contain a
common object have to intersect. This additional degree of freedom makes them
much more suitable for the representation of more complex set relations. Diagrams of this kind were first introduced by Leonhard Euler (1707-1783). As a
second interpretation, they permit the visualization of binary datasets, such as
formal contexts from the research realm of formal concept analysis. Each set corresponds to an attribute and each element in the sets to an object. The incidence
between objects and attributes is depicted by containment.
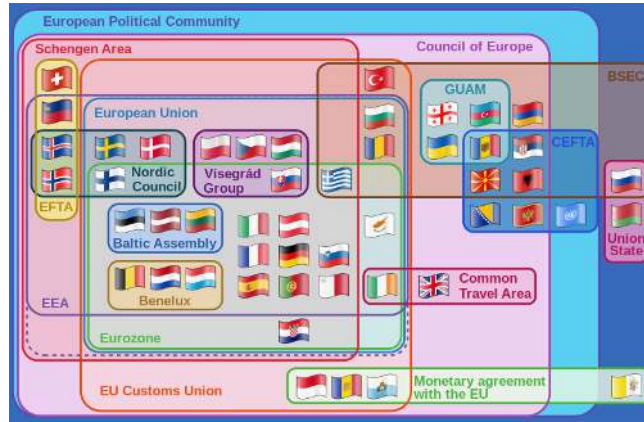
Figure 1. Supranational European bodies represented by a two-dimensional Euler diagram.[1]

One especially readable type of Euler diagrams are the ones, where the sets are represented by aligned rectangles. Because of their readability, those diagrams are often used to depict set relations without further explanation on how to read them. One common example is the membership of countries in supranational bodies, as one can see for example in Figure 1. Here, European countries are related to their supranational bodies. To see that this diagram is in fact an Euler and not a Venn diagram, note that, for example, Eurozone and EFTA do not intersect.

Even though these diagrams are commonly used in applications, their automatic generation is surprisingly not very well investigated. Until today, most rectangular Euler diagrams are still hand-crafted. To our knowledge, the first full research paper dealing with the automatic generation of rectangular Euler diagrams is due to Paetzold et al. from 2023 [10]. There, a mixed-integer optimization scheme is used to compute the diagrams. As it is not always possible to represent each family of sets using a rectangular Euler diagram, they iteratively split the diagram into multiple subdiagrams if they cannot find a suitable solution. As they do not have a reliable way to check for the existence of diagrams for a given dataset, they do a split if their algorithm does not find a solution within 50 seconds.

In this work, we give a criterion for the existence of rectangular Euler diagrams, using the mathematical theory of order dimension. In doing so, we consider the problem in two different settings. First, we consider the problem in one dimension. We show that an associated order relation, which we introduce in this work, has order dimension two if and only if the dataset has a one-dimensional Euler diagram. This characterization is given in a constructive way, which al-

---

[1] https://commons.wikimedia.org/wiki/File:Supranational_European_Bodies.svg. Accessed January 19, 2024.

lows it to be directly translated into a polynomial time algorithm. Then, we consider the two-dimensional setting, which is the same setting investigated by Paetzold et al. [10] and depicted in Figure 1. Similarly to the one-dimensional case, we construct an associated partial order and use it to give a criterion for the existence of the Euler diagrams based on its order dimension. In comparison to the one-dimensional case however, this is based on the order dimension being four, which is NP-complete to decide. Thus, the algorithm that follows from this condition does not run in polynomial time.

The rest of this work is structured as follows. In Section 2 we embed our work in previous research. Then, we introduce the necessary mathematical notations in Section 3. In Sections 4 and 5 we investigate the one- and two-dimensional Euler diagrams respectively, before we discuss the time complexity of the resulting algorithms in Section 6. Finally, in Section 7 we conclude this work.

2   Related Work

A (partially) ordered set $(X, \leq)$ is called a geometric containment order if there is a mapping from its elements into shapes of the same type in a Euclidean space of finite dimension. For each pair of elements it holds that $x < y$ if and only if the shape corresponding to x is properly included in the shape of y. Geometric containment orders are closely related to this work. In our work, attributes form an interval or rectangle type of geometric containment order. Furthermore, the objects are represented as points in the same space and the incidence between objects and attributes is represented by geometric containment. The first result relevant to our work in that regard can be found in the very first paper introducing order dimension [4]. In this work, Dushnik and Miller note that interval containment orders are exactly the orders of dimension of at most two. A generalization of this result is given in the work of Golumbic and Scheinerman [7], where they show that orders of dimension of at most $2k$ can be characterized as the geometric containment orders with aligned boxes in an at most k-dimensional Euclidean space. Following this, we relate the one-dimensional Euler diagrams to orders of dimension two, and the two-dimensional Euler diagrams to orders of dimension four. More results on the order dimension of geometric containment orders of various types can be found in a survey by Fishburn and Trotter [5].

For a general overview on Euler diagrams, we refer the reader to a survey of Rodgers [16]. There, Euler diagrams are introduced using the very common definition of being "finite sets of labeled and closed curves". Thus, the Euclidean space is divided into regions or zones by the curves. This is in contrast to the definition used in this work where we refer to Euler diagrams as intervals and points on the real line or as rectangles and points in the two-dimensional Euclidean space. Our reason to deviate from the common notion is that it is easier to work with our definition in the context of geometric containment. The only setting which is similar to ours and was previously investigated is the work of Paetzold et al. [10]. Other approaches that incorporate rectangles in the generated Euler diagrams are given by Riche and Dwyer [15] and Yoghourdjian [22].

Furthermore, there are adjacent approaches for the generation of Euler diagrams
if we forgo the condition of the sets being rectangular. Thereby, an often-applied
strategy [17,8] is to use the dual graph of the diagram. This graph contains a
vertex for each intersection of regions and two vertices are connected by an edge
if their corresponding regions are neighboring.

The canonical way to represent a formal context in formal concept analysis
is the concept lattice, for which drawing approaches based on its order dimen-
sion exist [1]. An alternative approach was developed under the name ordinal
factor analysis [6]. For this approach several visualization methods based on the
order dimension have been proposed recently [2,3]. In his famous theorem [18],
Schnyder gives a condition on the existence of a planar graph based on the order
dimension of its incidence poset. From this, also an algorithm to compute a plane
embedding follows, thus our work is conceptually similar to his. One especially
noteworthy result related to our work is due to Petersen [11,12] who investigates
S-alphabets which coincide with our one-dimensional Euler diagrams. Her work
relates the realizability as an S-alphabet to the planarity of an associated lat-
tice. As for lattices planarity and two-dimensionality are equivalent, our work
can be seen as an extension of hers. In fact, the lattice proposed in her work is
the Dedekind-MacNeille completion of the order relation discussed in Section 4
on one-dimensional Euler diagrams of this work. In [14], the authors of this
paper describe connections between order theory and Euler diagrams and give
some negative results on the relation between interval containment orders and
one-dimensional Euler diagrams. Finally, in the realm of formal concept analy-
sis, there is recent interest in using formal concepts for the generation of Euler
diagrams [13].

3   Mathematical Foundations

For binary relations R the notions $(x, y) \in R$ and xRy can be used interchange-
ably. A homogeneous relation $\leq$ on a set X that is reflexive ($\forall x \in X : x \leq x$),
antisymmetric ($\forall x, y \in X : x \leq y \wedge x \leq y \Rightarrow x = y$), and transitive ($\forall x, y, z \in
X : x \leq y \wedge y \leq z \Rightarrow x \leq z$) is called an *order relation* on X. We call the tuple
$(X, \leq)$ an *ordered set*. If $x \leq y$ and $x \neq y$ we write $x < y$. Furthermore, we
sometimes use the notation $y \geq x$ instead of $x \leq y$ and $y > x$ instead of $x < y$.
Two elements $x, y \in X$ are called *comparable* if $x \leq y$ or $y \leq x$, otherwise they
are called *incomparable*. A *linear order* on X is an order relation, where all pairs
of elements in X are comparable. For an ordered set $(X, \leq)$ a linear order $\leq_l$ on
X is called a *linear extension* of $\leq$, if $\leq_l \supseteq \leq$. A family $\mathcal{R} = (\leq_1, \ldots, \leq_t)$ of linear
extensions of $\leq$ is called a *realizer* of $\leq$ if $\leq = \bigcap_{i=1}^{t} \leq_i$. The *order dimension* of
$\leq$ is defined as the least cardinality of a realizer and denoted by $\dim(X, \leq)$.

A formal context $\mathbb{K}$ is a triple $\mathbb{K} = (G, M, I)$ where G is a set called *objects*, M
a set called *attributes*, and $I \subseteq G \times M$ an *incidence relation*. For a set $A \subseteq G$, its
derivation is given by $A' = \{m \in M \mid \forall g \in A : (g, m) \in I)\}$. Dually, for $B \subseteq M$,
let $B' = \{g \in G \mid \forall m \in B : (g, m) \in I\}$. A *formal concept* is a tuple (A, B) with
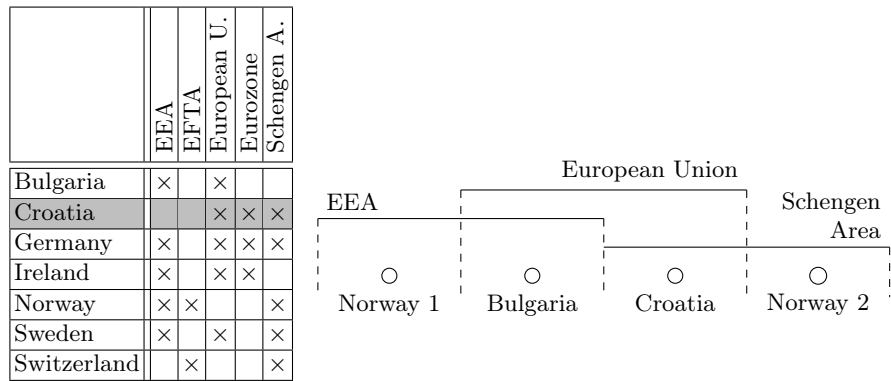$A \subseteq G$ and $B \subseteq M$ such that $A' = B$ and $B' = A$. We call A the *extent* and

B the *intent* of the concept, and denote the set of all concepts by $\mathfrak{B}(K)$. The ordered set $(\mathfrak{B}, \leq)$ with $(A, B) \leq (C, D)$ iff $A \subseteq B$ is called the *concept lattice* of $\mathbb{K}$. In this work, we assume all contexts to be *clarified*, i.e., if multiple attributes share the same derivation, they are grouped to a single attribute, and, if multiple objects share the same derivation, they are grouped to a single object. Such an assumption can be made, as going back from a clarified to the original context is only a matter of duplication. We denote by $\mathcal{I}$ the set of all closed intervals in the real numbers $\mathbb{R}$. A *rectangle* is a tuple (x, y) where $x, y \subseteq \mathbb{R}$ are closed intervals. We say that a point $(a, b) \in R \times R$ is *contained* in a rectangle (x, y) if $a \in x$ and $b \in y$.
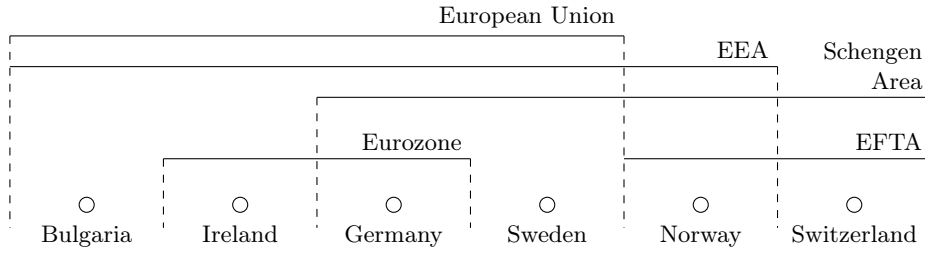
4 One-Dimensional Euler Diagrams

The first kind of Euler diagram that we are investigating in this work are *one-dimensional*, which means that their sets can be aligned along the real line. These diagrams are the essential building blocks for the more complex two-dimensional variant that we will discuss in the next section. In these diagrams, attributes are represented by intervals and objects by points. Formally, they are defined as follows.

Definition 1. *A* one-dimensional Euler diagram $\mathbb{E} = (\mathcal{J}, P)$ *is a set of closed intervals* $\mathcal{J} \subseteq \mathcal{I}$ *and a set of points* $P \subseteq \mathbb{R}$. *Let* $\mathbb{K} = (G, M, I)$ *be a formal context. We say that* $\mathbb{E}$ *corresponds to* $\mathbb{K}$ *with bijective maps* $\phi \colon G \to P$ *and* $\psi \colon M \to \mathcal{J}$ *and for all objects* g *and all attributes* m *it holds that* $(g, m) \in I$ *iff* $\phi(g) \in \psi(m)$. *We say that* $\mathbb{K}$ *can be represented by a one-dimensional Euler diagram if there is a one-dimensional Euler diagram corresponding to* $\mathbb{K}$.

One dimensional Euler diagrams are well suited for a readable representation of the data of a representable corresponding formal context. Consider for example

	EEA	EFTA	European U.	Eurozone	Schengen A.
Bulgaria	×		×		
Croatia			×	×	×
Germany	×		×	×	×
Ireland	×		×	×	
Norway	×	×			×
Sweden	×		×		×
Switzerland		×			×

Figure 2. *Left:* A formal context representing how some European states are members of supranational European bodies. *Right:* A figure supporting the argument that the whole context cannot be represented by a one-dimensional Euler diagram.

Figure 3. A graphical representation of a one-dimensional Euler diagram. The horizontal lines correspond to the intervals and the disks correspond to the points of the Euler diagram. The intervals are disposed in vertical direction for better readability, the dashed lines highlight where an interval starts or ends.

the formal context in Figure 2 without the object Croatia. We can represent this context with the corresponding Euler diagram from Figure 3. Even a reader with no prior training should be able to recognize the memberships of the different states purely from this diagram.

Note that it is however not possible for every formal context to be represented by a one-dimensional Euler diagram. In fact, the same context from Figure 2 cannot be represented if we include all objects. To see this, consider only the three attributes European Union, Schengen, and EEA and the three objects Croatia, Bulgaria, and Norway. A visualization supporting the following argument is depicted in the right part of Figure 2. Start with the interval representing the European Union. This interval has to contain Bulgaria and Croatia. Without loss of generality we can assign to Bulgaria a position left of Croatia. On the one hand, Croatia is not in the EEA and Norway not in the European Union, thus, Norway has to be positioned left of the European Union (Norway 1 in Figure 2). Furthermore, Bulgaria is not in the Schengen Area in which Norway however is, thus Norway has to be positioned on the right of the European Union (Norway 2 in Figure 2). Thus, this context cannot be represented using a one-dimensional Euler diagram.

Before we go in depth on the realizability of one-dimensional Euler diagrams we investigate a property that would be desireable for the diagrams to become more readable. That is that the intersections of the intervals should correspond to the intents of the formal concepts. To determine this, we define the property of being conceptual as follows.

Definition 2. *A one-dimensional Euler diagram* $\mathbb{E} = (\mathcal{J}, P)$ *is called* conceptual *if for all* a, b *with* a *and* b *being boundaries of intervals in* $\mathcal{J}$ *and* $a < b$ *there is some element* $p \in P$ *with* $a < p < b$.

So, being conceptual means that there is no minimal region in the diagram which does not contain a point. In Euler diagram research, such diagrams are sometimes referred to as diagrams of *existential import*. The following lemma

shows, that this property is already sufficient to guarantee that the concepts really do correspond to the interval intersections.

Lemma 1. *Let $\mathbb{K} = (G, M, I)$ be a formal context corresponding to an Euler diagram $\mathbb{E} = (\mathcal{J}, P)$ under the maps ϕ and ψ. If $\mathbb{E}$ is conceptual, the map*

$$\chi : \begin{cases} \mathfrak{B}(\mathbb{K}) & \mapsto \{\bigcap \tilde{\mathcal{J}} \mid \tilde{\mathcal{J}} \subseteq \mathcal{J}\} \\ (A, B) & \to \bigcap_{b \in B} \psi(b) \end{cases}$$

is a one-to-one correspondence, and for all concepts $(A, B) \in \mathfrak{B}(\mathbb{K})$ it holds that $\phi(a) \in \chi(A, B)$ if and only if $a \in A$.

Proof. We first show the second part of the lemma. For this let $a \in A$ for some concept (A, B). For all $b \in B$, if $(a, b) \in I$, then it holds that $\phi(a) \in \psi(b)$. Thus, $\phi(a) \in \chi(A, B)$. Suppose now that $a \notin A$. Then there is a $b \in B$ with $(a, b) \notin I$ and thus $\phi(a) \notin \psi(b)$ and hence $\phi(a) \notin \chi(A, B)$.

We now show that χ is injective. Let (A, B) and (C, D) be two different formal concepts, i.e., $A \neq C$. Without loss of generality, there is an element $a \in A \setminus C$. Then $\phi(a) \in \chi(A, B)$ and $\phi(a) \notin \chi(C, D)$ which shows injectivity.

Finally, we show that χ is surjective. For this, let $K = \bigcap \tilde{\mathcal{J}}$ for some $\tilde{\mathcal{J}} \subseteq \mathcal{J}$. Let $L = P \cap K$ which is, by assumption, not empty. Let A be the set of objects corresponding to L. We now show, that for the concept (A'', A') it holds that $\chi(A'', A') = K$. We know, that for all $a \in A$ it holds that $\phi(a) \in L \subseteq K$. Thus, $K \subseteq \bigcap_{b \in A'} \psi(b)$ and therefore, $\chi(A'', A') \subseteq K$. Assume that $\chi(A'', A') \subsetneq K$. As $\mathbb{E}$ is conceptual, $(K \setminus \chi(A'', A')) \cap P \neq \emptyset$. Let $a \in A$ be such that $\phi(a) \in K \setminus \chi(A'', A')$. But if $\phi(a) \in K$, then also $\phi(a) \in L$ and $a \in A \subseteq A''$, a contradiction. $\qquad\square$

Note that in the Euler diagram, there will be no geometric representation for the empty attribute set, which might correspond to the bottom concept. However, the empty set is in any case a subset of the real line, thus, the previous lemma still holds. It is also noteworthy that for the definition of conceptual one-dimensional Euler diagrams, we do not need a corresponding formal context.

We are going to provide a necessary and sufficient condition to check whether a formal context can be represented by a (conceptual) one-dimensional Euler diagram. For this we require the notion of Euler-posets, which we define as follows.

Definition 3. *The* Euler-poset *of a formal context $\mathbb{K} = (G, M, I)$ is defined as the ordered set $(G \cup M, \leq)$ such that*

i) $\forall g \in G, m \in M : g < m \iff (g, m) \in I$,
ii) $\forall m_1, m_2 \in M : m_1 \leq m_2 \iff m_1' \subseteq m_2'$,
iii) $\forall g_1, g_2 \in G : g_1 \not< g_2 \wedge g_1 \not> g_2$.

Recall that we are working on clarified contexts, thus the Euler-poset is well defined. We give an order diagram of the Euler-poset of a formal context from Figure 2 in Figure 4. The order dimension of this Euler-poset is connected to the realizability of a formal context as we are going to show in the following lemmas and theorem.

Algorithm 1 Euler1D

```
def Euler1D(G, M, I):
    (X, ≤) = EulerPoset(G, M, I)
    L₁, L₂ = 2D_Realizer(X, ≤)
    return EulerFromLinearExtension(L₁, L₂)

def EulerFromLinearExtension(L₁, L₂):
    i₁, i₂, k = 0
    while i₁ < |X| or i₂ < |X|:
        if L₁[i₁] ∈ M:
            start[L₁[i₁]] = k
            i₁ = i₁ + 1
        if L₂[|X| − i₂ − 1] ∈ M:
            end[L₂[|X| − i₂ − 1]] = k
            i₂ = i₂ + 1
        if L₁[i₁] ∈ G and L₂[i₂] ∈ G:
            position[L₁[i₁]] = k + 1
            k = k + 2
            i₁ = i₁ + 1
            i₂ = i₂ + 1
    return (start, end, position)
```

Lemma 2. *If $\mathbb{K}$ is a formal context with an Euler-poset of order dimension two and realizer $\{L_1, L_2\}$, then the order of the objects of $\mathbb{K}$ in L_1 is reversed in L_2.*

Proof. For all objects $g, h \in G$ it holds that they are incomparable in the Euler-poset, if and only if $g \neq h$. Assume that the statement is not true, i.e., there is a pair $g \neq h$ with $g < h$ in both linear extensions. As the Euler-poset is two-dimensional, there is no way to break this comparability and thus $g < h$ in the Euler-poset, which is a contradiction.

Given a two-dimensional Euler-poset, we define an algorithm to compute a one-dimensional Euler diagram in Algorithm 1. This algorithm iterates through both linear extensions L_1 and L_2 of a realizer simultaneously, one from the smallest element to the largest and the other from the largest element to the smallest. By Lemma 2, the objects appear in the same order. Each attribute in the first linear extension marks the start-point of an interval while the same attribute in the second order marks its end-point. In the following we show that this algorithm does in fact always compute a corresponding Euler diagram if the Euler-poset is two-dimensional.

Lemma 3. *If the Euler-poset is two-dimensional, Algorithm 1 will compute a conceptual order diagram.*

Proof. In both L_1 and L_2 it holds for all $(g, m) \in I$ that $g < m$. As L_1 is iterated from the smallest to the largest element, the start point of m is positioned before
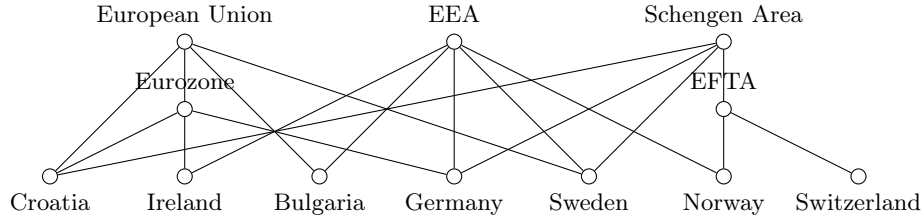
Figure 4. An order diagram of the Euler-poset of the formal context from Figure 2.

g. As L_2 is iterated from the largest to the smallest element, the end point of m is positioned after g. On the other hand if $(g, m) \notin I$, they are incomparable in the Euler-poset. Thus, in one of the two linear extensions $g \leq m$ and in the other $m \leq g$. Let $g \leq m$ in L_1 and $m \leq g$ in L_2. Thus, the start and end point of the interval of m get assigned smaller values than the position of g. In the other case with $m \leq g$ in L_1 and $g \leq m$ in L_2 both the start and end point of the interval corresponding to m is assigned values larger than the position of g.

We now show that the computed Euler diagram $\mathbb{E}$ is conceptual. Assume not, i.e., there are two elements a, b with $a < b$ which are boundaries of some intervals and there is no point p with $a < p < b$. But note that the positioning of the start and end-points counter of all intervals is set to the counter k, which only increases in size if an object is positioned in-between. Thus, $a = b$, a contradiction.

We now know that it is always possible to compute the Euler diagram if the Euler-poset is two-dimensional. The other direction, i.e., that the Euler-poset is also two-dimensional if a one-dimensional order diagram exists and thus a full characterization of these Euler diagrams is provided by the following theorem.

Theorem 1. *Let* $\mathbb{K} = (G, M, I)$. *Then the following statements are equivalent.*

 i) $\mathbb{K}$ *can be represented by a conceptual one-dimensional Euler diagram.*
 ii) $\mathbb{K}$ *can be represented by a one-dimensional Euler diagram.*
iii) *The Euler-poset of* $\mathbb{K}$ *is two-dimensional.*

Proof. $i) \Rightarrow ii)$: By definition.
$ii) \Rightarrow iii)$: Let $\mathbb{E} = (\mathcal{J}, P)$ be the one-dimensional Euler diagram. Consider the set of intervals $K = \mathcal{J} \cup \{[p, p] \mid p \in P\}$. Note that the ordered set $(K, \subseteq)$ is isomorphic to the Euler-poset of $\mathbb{K}$. A result of [4] is, that ordered sets that can be represented by containment of intervals are two-dimensional.
$iii) \Rightarrow i)$: Lemma 3 demonstrates that you can always compute the conceptual one-dimensional Euler diagram from a two-dimensional Euler-poset.

This theorem does not only relate the existence of one-dimensional Euler diagrams to the existence of one-dimensional conceptual Euler diagrams, it also gives an easy-to-check criterion for their existence based on the two-dimensionality of the Euler-poset. Finally, Algorithm 1 guarantees the construction of Euler diagrams if they exist.

	m_1	m_2	m_3	m_4
g_1	×	×		×
g_2		×	×	×
g_3				×
g_4	×			
g_5			×	

Figure 5. A formal context with a concept lattice of order dimension two that cannot be realized by a one-dimensional Euler diagram.

Corollary 1. *Algorithm 1 computes an Euler diagram if it exists.*

Furthermore, as we will discuss in Section 6, the algorithm runs in polynomial time. Note that we are working on the clarified context and thus, to achieve an Euler diagram of the not-clarified context, we might have to duplicate the intervals or points. Consider once more the Euler-poset in Figure 4 of the previously discussed Europe dataset. A subposet of this poset is what is called the standard example S_3 in order theory, a poset that is well known to be three-dimensional. Here, the standard example can be found on the objects Croatia, Bulgaria, Norway, and the attributes EEA, European Union, Schengen Area. This confirms our previous observation that we cannot represent this formal context by a one-dimensional order diagram.

Note that the definition of Euler diagrams allows for regions where exactly the same set of intervals intersect. If the diagram is conceptual, both of these regions have to contain a point corresponding to an object. But then, both of these objects have the same derivations. As we assumed that we are working on clarified contexts, our algorithm will not produce such a diagram.

Finally, we are interested in how the dimensionality of the Euler-poset is related to the dimensionality of the concept lattice. One might expect that they coincide. However, we now show that this is not always the case.

Theorem 2. *For a formal context $\mathbb{K} = (G, M, I)$, let e be the order dimension of the Euler-poset E and d be the order dimension of its concept lattice. Then it holds that $d \leq e \leq d + 1$ and these bounds are tight.*

Proof. Let $\mathbb{E}$ be the Euler-poset of $\mathbb{K}$. Let $(\mathfrak{B}_{AOC}, \leq)$ be the ordered set that emerges if we restrict the concept lattice to concepts that can be generated by the derivation of a single attribute or object. It is well known, that the Dedekind-MacNeille completion of this poset is exactly the concept lattice $(\mathfrak{B}, \leq)$. Because the Dedekind-MacNeille completion preserves order dimension, as shown in [9], we know that $dim(\mathfrak{B}_{AOC}) = d$.

For the lower bound, note that $\mathfrak{B}_{AOC} \subseteq E$ and thus if we restrict the realizer of $\mathfrak{B}_{AOC}$ to the elements of E we receive a realizer of E with the same number of linear extensions which proves the lower bound. To see that the lower bound is tight, consider the contranominal scale with d objects and attributes. The concept lattice of this contranominal scale has order dimension d. Its Euler-poset is exactly the standard example S_d which also has order dimension d.

We now show the upper bound. Take a realizer of $\mathfrak{B}_{AOC}$ and modify it as follows. Replace all attribute concepts by the attributes they correspond to and all object concepts by the objects they correspond to. If there are some concepts in $\mathfrak{B}_{AOC}$ which are attribute-concepts and object-concepts at the same time, we can replace their element in the realizer by two consecutive elements. The smaller of these two elements will correspond to the attribute and the larger one to the object of the concept. The linear extensions $L_1, \ldots, L_d$ generated by the procedure will correspond to the Euler-poset, except that there still might be comparable objects. To fix this, generate an additional linear extension as follows. Take a linear extension of the so far generated realizer, without loss of generalization take L_1 and generate the modified linear order L_{d+1} as follows. In L_{d+1} we first have the attributes in the order as they appear in L_1 and above all attributes the objects in reversed order to L_1. Then $L_1, \ldots, L_{d+1}$ will be a linear extension of the Euler-poset. Tightness of the upper bound follows from the example in Figure 5. The concept lattice of the formal concept has order dimension two, however its Euler-poset contains a "spider" on the elements $\{m_1, m_2, m_3, m_4, g_1, g_2, g_3\}$, which is well known to be three-dimensional.

Thus, we cannot even compute one-dimensional Euler diagrams for all two-dimensional concept lattices. Therefore, the investigation of higher dimensional Euler diagrams which we are going to do in the following section is of interest.

5 Two-Dimensional Euler Diagrams

The natural extension of one-dimensional Euler diagrams is two-dimensional Euler diagrams, where the intervals are replaced by rectangles and the one- by two-dimensional points. Formally, they are defined as follows.

Definition 4. *A* two-dimensional Euler diagram $\mathbb{E} = (\mathcal{J}, P)$ *is a set of closed rectangles* $\mathcal{J} \subseteq \mathcal{I} \times \mathcal{I}$ *and a set of points* $P \subseteq \mathbb{R}^2$. *Let* $\mathbb{K} = (G, M, I)$ *be a formal context. We say that* $\mathbb{E}$ *corresponds to* $\mathbb{K}$ *if there is a pair of bijective maps* $\phi : M \to \mathcal{J}$ *and* $\psi : G \to P$, *such that for all objects* g *and all attributes* m *it holds, that* $(g, m) \in I$ *iff* $\phi(g)$ *is contained in* $\psi(m)$. *We say that* $\mathbb{K}$ *can be represented by a two-dimensional Euler diagram* if there is a two-dimensional Euler diagram corresponding to $\mathbb{K}$.

Every two-dimensional Euler diagram can be seen as the direct product of two one-dimensional diagrams. That is, as the first interval of the rectangle together with the first dimension of the point forms a one-dimensional Euler diagram and the second interval together with the second dimension of the point the other one. An object is exactly incident to an attribute, if in both diagrams the corresponding point is in the attribute interval. So, in summary, the following corollary holds.

Corollary 2. *A formal context* $\mathbb{K} = (G, M, I)$ *can be represented by a two-dimensional Euler diagram if and only if there are two one-dimensional Euler diagrams corresponding to two formal contexts* $\mathbb{K}_1 = (G, M, I_1)$ *and* $\mathbb{K}_1 = (G, M, I_2)$ *such that* $I = I_1 \cap I_2$.

We will now give a characterization on the existence of these two-dimensional Euler diagrams using the same concepts as in the one-dimensional case. We once again define an ordered set and then show a connection between the order dimension of this order and the existence of the Euler diagram.

Definition 5. *Let $S = \{a_i \mid i \in [4]\} \cup \{b_i \mid i \in [4]\}$. The* extended Euler-poset *of a formal context $\mathbb{K} = (G, M, I)$ is the ordered set $(G \cup G_1 \cup G_2 \cup M \cup S, \leq)$ with $G_i = \{g_i \mid g \in G\}$ for $i \in \{1, 2\}$ where the following conditions hold:*

i) $\forall g \in G, m \in M : g < m \iff (g, m) \in I$,
ii) $\forall m_1, m_2 \in M : m_1 \leq m_2 \iff m_1' \subseteq m_2'$,
iii) $\forall g \in G : (g < g_1) \wedge (g < g_2)$,
iv) $\forall i, j \in [4] : a_i \leq b_j \iff i \neq j$,
v) $\forall x \in (G \cup M), i \in [4] : x < b_i$,
vi) $\forall g \in G : (g_1 > a_1) \wedge (g_1 > a_2) \wedge (g_2 > a_3) \wedge (g_2 > a_4)$
*vii)* *all other pairs are incomparable.*

Note that for the elements of G and M the same comparison conditions hold as for the Euler-poset from the one-dimensional case. Thus, the extended Euler-poset contains the Euler-poset as a suborder. Now, we can characterize the existence of a two-dimensional Euler diagram using this extended Euler-poset.

Theorem 3. *A formal context $\mathbb{K} = (G, M, I)$ can be represented by a two-dimensional Euler diagram if and only if its extended Euler-poset has order dimension four.*

Proof. "$\Rightarrow$": Let $\mathbb{K} = (G, M, I)$ be a formal context that can be represented by a two-dimensional Euler diagram. Projecting this two-dimensional Euler diagram on the first and on the second coordinate produces two one-dimensional Euler diagrams, call them $\mathbb{E}_1$ and $\mathbb{E}_2$. By Corollary 2 it holds that $(g, m) \in I$ for the corresponding point is in the corresponding interval for both of these diagrams. By Theorem 1, the Euler-poset of $\mathbb{E}_1$ and $\mathbb{E}_2$ have order dimension two. Let L_1, L_2 be the realizer of E_1 and L_3, L_4 the realizer of E_2. We now construct a realizer of the extended Euler-poset. Consider the following linear extensions.

$$\hat{L}_1\colon \; a_1 < a_2 < a_3 < \tilde{L}_1 < b_4 < a_4 < \tilde{G}_1 < b_3 < b_2 < b_1$$
$$\hat{L}_2\colon \; a_1 < a_2 < a_4 < \tilde{L}_2 < b_3 < a_3 < \tilde{G}_2 < b_4 < b_2 < b_1$$
$$\hat{L}_3\colon \; a_1 < a_3 < a_4 < \tilde{L}_3 < b_2 < a_2 < \tilde{G}_3 < b_4 < b_3 < b_1$$
$$\hat{L}_4\colon \; a_4 < a_3 < a_2 < \tilde{L}_4 < b_1 < a_1 < \tilde{G}_4 < b_2 < b_3 < b_4$$

For those, $\tilde{L}_i$ is composed of the elements of L_i with the following modification. Directly after each element g, its respective element g_1 is positioned if $i = 1$ or $i = 2$ and g_2 if $i = 3$ or $i = 4$. The sets $\tilde{G}_i$ are composed of the elements G_2 if $i = 1$ or $i = 2$ and of G_1 if $i = 3$ or $i = 4$. The order in $\tilde{G}_i$ is the same as the one of the set G in L_i. Note that the constructed four linear extensions do in fact realize the extended Euler-poset which proofs the claim that the extended Euler-poset is at most four-dimensional. To see that it is at least four-dimensional, note that the a_i and b_i elements form a standard example on four elements which is well known to be four-dimensional.
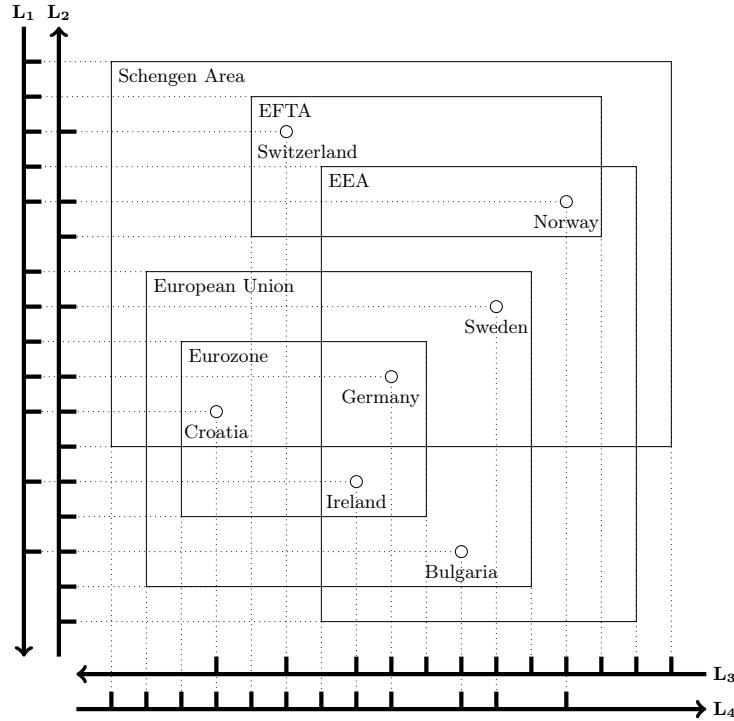
Figure 6. The Euler diagram corresponding to the dataset of Figure 2 together with the four linear extensions that give rise to this visualization.

"$\Leftarrow$": For the other direction take a realizer of the extended Euler-poset with the four linear extensions L_1, L_2, L_3 and L_4. As the elements a_i and b_i for $i \in [4]$ form a standard example of dimension four, there is exactly one linear extension for each i, where $a_i > b_i$. Without loss of generality, let L_i be the linear extension where $a_i > b_i$.

Claim: For any two objects g and h it holds that $g < h$ in L_1 iff $h < g$ in L_2 and $g < h$ in L_3 iff $h < g$ in L_4.

Note that for $i = 3$ and $i = 4$ it holds in L_i that $g < b_i < a_i < h_2$ and $h < b_i < a_i < g_2$. But the pairs (g, h_2) and (h, g_2) are incomparable, i.e., there has to be a linear extension in L_1, L_2 with $h2 < g$ and a linear extension in L_1, L_2 with $g_2 < h$. Assume that both of these inequalities are true in the same linear extension. Then $h_2 < g < g2 < h < h_2$ in this linear extension which is a contradiction. Thus, in one of the two linear extensions L_1 and L_2, it holds that $h < h_2 < g$ and in the other it holds that $g < g_2 < h$ what proofs the claim. By symmetry, the same is true for the linear extensions L_3 and L_4.

Thus, the order of the elements of G are in opposite order in L_1 and L_2 and the same is true for L_3 and L_4. Therefore, the linear extensions of L_1 and L_2 restricted to the elements G and M form a linear extension of an Euler-poset

---

Algorithm 2 Euler2D

```
def Euler2D(G, M, I):
    (X, ≤) = ExtendedEulerPoset(G, M, I)
    R = 4D_Realizer(X, ≤)
    for i in [4]:
        Lᵢ = R[bᵢ < aᵢ]
    E₁ = EulerFromLinearExtension(L₁, L₂)
    E₂ = EulerFromLinearExtension(L₁, L₂)
    return E₁, E₂
```

---

and thus, a one-dimensional Euler diagram by Theorem 1. The same is true for the linear extensions L_3 and L_4. Thus, we receive the two one-dimensional Euler-posets of Corollary 2 which shows, that the Euler-poset is two-dimensional.

As this proof is constructive, it composes an algorithm to compute two-dimensional Euler diagrams. We first compute the extended Euler-poset and then a four-dimensional realizer. Based on the order dimension we can then compute the linear extensions L_1, L_2, L_3 and L_4, which, if restricted to G and M, are the linear extensions of two Euler-posets. Finally, we can use the procedure for one-dimensional Euler diagrams to compute two one-dimensional Euler diagrams which are exactly the first and second component of the two-dimensional Euler diagram respectively. The algorithm is given in Algorithm 2. A visualization of how the four computed linear extensions, restricted to the attributes and objects, are related to the positions of the rectangles is depicted in Figure 6. The linear extensions L_1 and L_2 give the vertical dimension of the order diagram and L_3 and L_4 the horizontal dimension. Once again, in each dimension one linear extension gives the start-point of the intervals and the other the end-points.

6   Time Complexity of the Algorithms

In this paper we propose two algorithms. Both of them have in common that they use a characterization based on the order dimension of an associated poset. In both cases, the size of this poset is in $O(|M| + |G|)$. To compute these posets, we have to compare an order based on the attribute derivations, here we have to do potentially $O(|M|^2 + |M| \cdot |G|)$ comparisons.

The main difference is that the algorithm for one-dimensional Euler diagrams has to compute a two-dimensional realizer, while a realizer of size four is required in the two-dimensional case. It is possible to compute a realizer of size two by the transitive orientation of its cocomparability graph. This can be done in $(O(n)^2)$ [19]. Thus, the overall time complexity of the resulting algorithm is in $O((|M| + |G|)^2)$. The algorithm that computes a two-dimensional Euler diagram on the other hand relies on a subroutine that computes a four-dimensional realizer of an ordered set. Unfortunately, it is NP-complete to decide, if the order dimension of an ordered set is $k \geq 3$ [21]. Thus, there are no efficient algorithms

known to solve this step in general. An exact way to compute the order dimension is to solve a hypergraph coloring problem on an associated hypergraph. However, in our experience, even for small datasets the number of hyperedges gets out of hand. We thus used an algorithm by Yáñez and Montero [20] that constructs a graph coloring by restricting the hypergraph to its two-element edges. Unfortunately, this algorithm is not guaranteed to compute a realizer of minimal size. As the algorithm however reports, whether the realizer it finds is minimal, we can go back to the hypergraph coloring problem if it fails. In any case, the algorithm will have an exponential runtime with respect to the input size. Still, using the graph coloring algorithm, we were able to compute order diagrams for all datasets that we were interested in within seconds.

7   Conclusion

In this work, we investigated the realizability of Euler diagrams. To be exact, we distinguished between two different types of Euler diagrams. The first kind of diagrams can be visualized as one-dimensional intervals and points on the intervals, in the second kind have attributes as aligned rectangles in two-dimensional Euclidean space. We related the two diagram types to each other. For both diagrams, we were able to give a condition on their existence for a given dataset based on the order dimension of an associated order relation. In the case of the one-dimensional order-diagrams, this characterization can be directly translated into an efficient polynomial-time algorithm to compute Euler diagrams. The algorithm that follows for two-dimensional Euler diagrams is not polynomial. Future work should focus on the development of heuristics to compute the two-dimensional diagrams. In this context and for non-realizable datasets, it would also be helpful to weaken the definition of an Euler diagram. One possibility to do so is to allow an interval or rectangle to appear multiple times or to not depict all incidences of the dataset.

References

1. Dürrschnabel, D., Hanika, T., Stumme, G.: Drawing order diagrams through two-dimension extension. J. Graph Algorithms Appl. **27**(9), 783–802 (2023). https://doi.org/10.7155/JGAA.00645, https://doi.org/10.7155/jgaa.00645
2. Dürrschnabel, D., Stumme, G.: Greedy discovery of ordinal factors. CoRR **abs/2302.11554** (2023). https://doi.org/10.48550/ARXIV.2302.11554
3. Dürrschnabel, D., Stumme, G.: Maximal ordinal two-factorizations. In: Ojeda-Aciego, M., Sauerwald, K., Jäschke, R. (eds.) Graph-Based Representation and Reasoning - 28th International Conference on Conceptual Structures, ICCS 2023, Proceedings. Lecture Notes in Computer Science, vol. 14133, pp. 41–55. Springer (2023). https://doi.org/10.1007/978-3-031-40960-8_5
4. Dushnik, B., Miller, E.W.: Partially ordered sets. Am. J. Math. **63**(3), 600–610 (1941), http://www.jstor.org/stable/2371374
5. Fishburn, P.C., Trotter, W.T.: Geometric containment orders: A survey. Order **15**(2), 167–182 (Jun 1998). https://doi.org/10.1023/A:1006110326269, https://doi.org/10.1023/A:1006110326269

6. Ganter, B., Glodeanu, C.V.: Ordinal factor analysis. In: Domenach, F., Ignatov, D.I., Poelmans, J. (eds.) Formal Concept Analysis - 10th International Conference, ICFCA 2012, Proceedings. Lecture Notes in Computer Science, vol. 7278, pp. 128–139. Springer (2012). https://doi.org/10.1007/978-3-642-29892-9_15

7. Golumbic, M.C., Scheinerman, E.R.: Containment graphs, posets, and related classes of graphs. Ann. N. Y. Acad. Sci. **555**, 192–204 (1989). https://doi.org/10.1111/j.1749-6632.1989.tb22452.x

8. Kehlbeck, R., Görtler, J., Wang, Y., Deussen, O.: SPEULER: semantics-preserving euler diagrams. IEEE Trans. Vis. Comput. Graph. **28**(1), 433–442 (2022). https://doi.org/10.1109/TVCG.2021.3114834

9. Novák, V.: Über eine eigenschaft der dedekind-macneilleschen hülle. Math. Ann. **179**(4), 337–342 (1969). https://doi.org/10.1007/BF01350778

10. Paetzold, P., Kehlbeck, R., Strobelt, H., Xue, Y., Storandt, S., Deussen, O.: Recteuler: Visualizing intersecting sets using rectangles. Comput. Graph. Forum **42** (2023)

11. Petersen, W.: A mathematical analysis of panini's sivasutras. J. Log. Lang. Inf. **13**(4), 471–489 (2004). https://doi.org/10.1007/S10849-004-2117-7

12. Petersen, W.: Zur Minimalität von Pāṇinis Śivasūtras: eine Untersuchung mit Methoden der formalen Begriffsanalyse. Ph.D. thesis, Heinrich-Heine-Universität Düsseldorf (2008)

13. Priss, U.: Representing concept lattices with euler diagrams. In: Dürrschnabel, D., López-Rodríguez, D. (eds.) Formal Concept Analysis - 17th International Conference, ICFCA 2023, Proceedings. Lecture Notes in Computer Science, vol. 13934, pp. 183–197. Springer (2023). https://doi.org/10.1007/978-3-031-35949-1_13

14. Priss, U., Dürrschnabel, D.: Rectangular euler diagrams and order theory. In: Diagrammatic Representation and Inference - 14th International Conference, Diagrams 2024, Proceedings (In Print). Lecture Notes in Computer Science, Springer (2024)

15. Riche, N.H., Dwyer, T.: Untangling euler diagrams. IEEE Trans. Vis. Comput. Graph. **16**(6), 1090–1099 (2010). https://doi.org/10.1109/TVCG.2010.210

16. Rodgers, P.: A survey of euler diagrams. J. Vis. Lang. Comput. **25**(3), 134–155 (2014). https://doi.org/10.1016/J.JVLC.2013.08.006

17. Rodgers, P., Zhang, L., Fish, A.: General euler diagram generation. In: Stapleton, G., Howse, J., Lee, J. (eds.) Diagrammatic Representation and Inference, 5th International Conference, Diagrams 2008, Proceedings. Lecture Notes in Computer Science, vol. 5223, pp. 13–27. Springer (2008). https://doi.org/10.1007/978-3-540-87730-1_6

18. Schnyder, W.: Planar graphs and poset dimension. Order **5**, 323–343 (1989)

19. Spinrad, J.P.: On comparability and permutation graphs. SIAM J. Comput. **14**(3), 658–670 (1985). https://doi.org/10.1137/0214048

20. Yáñez, J., Montero, J.: A poset dimension algorithm. J. Algorithms **30**(1), 185–208 (1999). https://doi.org/10.1006/JAGM.1998.0974

21. Yannakakis, M.: The complexity of the partial order dimension problem. SIAM J. Discrete Math. **3**(3), 351–358 (1982). https://doi.org/10.1137/0603036

22. Yoghourdjian, V., Dwyer, T., Gange, G., Kieffer, S., Klein, K., Marriott, K.: High-quality ultra-compact grid layout of grouped networks. IEEE Trans. Vis. Comput. Graph. **22**(1), 339–348 (2016). https://doi.org/10.1109/TVCG.2015.2467251