Towards Ordinal Data Science

Gerd Stumme[1] ✉ 🏠 🆔
Knowledge & Data Engineering Group, Research Center for Information System Design &
Department of Electrical Engineering and Computer Science, University of Kassel, Germany

Dominik Dürrschnabel ✉ 🏠 🆔
Knowledge & Data Engineering Group, Research Center for Information System Design &
Department of Electrical Engineering and Computer Science, University of Kassel, Germany

Tom Hanika ✉ 🏠 🆔
Institute of Computer Science, University of Hildesheim, Germany
Berlin School of Library and Information Science, Humboldt-Universität zu Berlin, Germany

── Abstract ──

Order is one of the main instruments to measure the relationship between objects in (empirical) data. However, compared to methods that use numerical properties of objects, the amount of ordinal methods developed is rather small. One reason for this is the limited availability of computational resources in the last century that would have been required for ordinal computations. Another reason – particularly important for this line of research – is that order-based methods are often seen as too mathematically rigorous for applying them to real-world data. In this paper, we will therefore discuss different means for measuring and 'calculating' with ordinal structures – a specific class of directed graphs – and show how to infer knowledge from them. Our aim is to establish Ordinal Data Science as a fundamentally new research agenda. Besides cross-fertilization with other cornerstone machine learning and knowledge representation methods, a broad range of disciplines will benefit from this endeavor, including, psychology, sociology, economics, web science, knowledge engineering, scientometrics.

1 Introduction

Order is a predominant concept for perceiving and organizing our physical and social environment, to infer meaning and explanation from observation, and to search and rectify decisions. For instance, we admire the highest mountain on earth, observe pecking order among animals, schedule events in time, and structure our organizations, libraries, and diseases hierarchically. The notion

---

[1] corresponding author

Transactions on Graph Data and Knowledge, Vol. 1, Issue 1, Article No. 6, pp. 6:1–6:39

Transactions on Graph Data and Knowledge
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

of order is deeply embedded in our language, as every adjective gives rise to a comparative (e. g., better, more expensive, more beautiful). Specific technical and social processes have been established for dealing with ordinal structures, e. g., scheduling routines for aircraft take-offs and production planning, queuing at bus stops, deriving the succession order as depth-first linear extension of the royal family tree, or discussing only the borderline cases in scientific program committees. These processes, however, are rather task-specific – there exist only few generic data analysis and machine learning tasks that are particularly tailored for ordinal data in general.

It comes thus as little surprise that order *is one of the two main types of relations between objects in empirical data*, together with *proximity/distance* [39]. While there exists a wide range of scientific work on analysis methods for ordinal data, their amount and coverage is far less than for proximity data. As a consequence, many data analysts resort to methods designed for other data types. We therefore call interested researchers to join us in the aim to establish the new field of *Ordinal Data Science* – both by organizing existing work in a unifying framework and by extending these approaches: Contributions are sought in two specific areas. Theoretical foundations for analyzing ordinal structures, in particular with respect to knowledge extraction and representation, on the one hand; and algorithmic methods that provide the means to measure and "calculate" with ordinal structures and closely related algebraic notions on the other hand.

A wide range of applied disciplines will benefit from the new field of Ordinal Data Science, a. o. psychology, web science, knowledge engineering, scientometrics. Besides these disciplines, ordinal data are related to the large family of relational data which have received high interest of the computer science community in the last years. This is due to developments in fields such as sociology ("relational turn") [141] or genetics [84] or epidemiology [36], and in particular sociotechnical developments such as the rise of online social networks or knowledge graphs. This means that, for the analysis of ordinal data, one can benefit from all kinds of measures and methods for relational data, as for instance centrality measures and clustering algorithms for (social) network data, or inductive logic programming. The specific structure of ordinal data, however, allows additionally to tap on the rich – but up to date mostly unexploited for data science – tool-set of mathematical order theory [34] and lattice[2] theory [42].

We expect Ordinal Data Science to become a scientific field with many ramifications. In this paper, we will discuss – in an exemplary fashion and to kindle the research field – five out of many potential subfields as programmatic research tasks:

1) Ordinal Measurement Theory.   Representational Measurement Theory, the most influential theory of measurement to date, discusses how empirically observed phenomena can be measured. It considers a measurement as a mapping from some empirical relational structure to a numerical relational structure. While there seems to be agreement that the latter could be any mathematical structure that is appropriate for the purpose at hand, almost all theory is centered around the real numbers (and variations) as numerical relational structure. As order is the second main type of relation in empirical data beside proximity/distance, we will discuss, in Section 5, the vision of an order-theoretic version of measurement theory.

2) Ordered Metric Spaces.   Real-world datasets are typically heterogeneous and contain both ordinal and numerical dimensions. The latter naturally give rise to a metric. In Section 6, we will motivate the need for a theory about the compatibility (or consistency) of an order relation and a metric. To this end, we introduce ordered metric spaces (om-spaces), in the spirit of metric measure spaces. Such a theory should be able to answer a range of foundational questions about

---

[2] In Section 4.1, we will recall that lattices are specific order relations with additional structure.

the relationship between order relations and metrics, including: a) If a metric has been learned based on information about the order, how consistent is the metric with that order? This will provide a fundamental new quality criterion for order-related machine learning tasks. b) How strong do order and metrics "agree" with their view on a dataset? An answer might be highly beneficial to an investigation whether order and metrics reflect the same hidden structure. c) How justified is it to represent an order relation by means of the metric? If, for a given dataset, the answer is satisfactory, this would justify in this case the use of the rich toolkit of numerical data analysis and machine learning methods for analyzing the ordinal structure.

3) Algebraic constructions/decompositions for complexity reduction. Algebraic operations provide a large variety of methods for reducing the complexity of ordered sets, in particular lattices. However, they are very sensitive to small perturbations in the data; and approximations and heuristics in the style of data mining are not (yet) considered in algebraic research. In Section 7, we will discuss how they could be made less sensitive in order to make them applicable to data science tasks.

4) Ordinal Factor Analysis. Observed and measured data is often highly correlated and interlinked, being caused by a small amount of factors. A significant task in data analysis is the identification of these factors. In Section 8, we discuss – in analogy to the classical factor analysis which is based on linear algebra – potential ways for extracting these factors in ordinal data based on the relation product.

5) Visualization, Exploration, and Browsing Ordinal Data. The typical means of presenting ordered sets to humans is via line diagrams. Open problems include in particular the specification of HCI-founded, formal optimization criteria for graph drawing (and the development of corresponding layout algorithms), as well as the (semi-)automatic break-down of large ordered sets into smaller, visualizable parts together with suitable means for their interactive exploration. In Section 9, we will call for the development of new paradigms for exploring and browsing ordinal datasets.

Before diving in the challenges of these subfields in more detail, we will first discuss the nature of ordinal data and the role of hierarchies in data science in Sections 2 and 3, resp., and will provide some order-theoretical foundations in Section 4.

2 Order as Foundational Principle for Organizing Data

Quantification with real numbers has been boosted by different factors, including i) the development of scientific measuring instruments since the scientific revolution, ii) the claim that the social sciences (starting with psychology) should use the same numerical methods which had been successful in natural sciences [63], and iii) nowadays by the instant availability of an enormous range of datasets to almost all aspects of science and everyday life. Indeed, in many cases, entities can be ordered through real-valued valuation functions like price or size. As the real numbers constitute an ordered field, the analysis of such data benefits from the existence of the operators $+, -, \cdot, /, 0, 1$ together with total[3] comparability $\leq$. Moreover, this combination allows for measures of tendency (such as mean, variance, and skewness) and transformations. If more than one real-valued dimension is present, this yields to a real vector space $\mathbb{R}^n$, which results in additional descriptive measures and metric properties, such as volumes, angles, correlation, covariance. This is the standard setting for the majority of data analysis and machine learning

---

[3] Total refers to the property that for any two elements a, b either $a \leq b$ or $b \leq a$ is true. A total order relation is often also called *linear order*.

models, and algorithms (e. g., density-based clustering, logistic regression, SVMs, to name just a few).

However, organizing hierarchical relationships by means of numerical values is not always adequate, as this kind of organization presupposes two important conditions:

1. every pair of entities has to be comparable, and

2. the magnitudes of the differences between numerical values are meaningful and thus comparable themselves.

In many situations, however, this is not the case: (*i*) does not hold, e. g., in concept hierarchies ("mankind" is neither a subconcept nor a superconcept of "ocean") nor in organizations (a member of parliament is neither above nor below a Secretary of State), and also teaching curricula are far from being linearly ordered; (*ii*) does not hold, e. g., in school grades (In the German system is the difference between "1 – very good" and "2 – good" equal to the difference between "4 – sufficient" and "5 – insufficient/fail"?) nor in organizations (In the European Commission, is an advisor closer to a deputy director general than a head of group to a director?).

To address such variations of data types, S. S. Stevens has distinguished in [174] four *levels of measurement: nominal, ordinal, interval,* and *ratio.* For data on the *ratio level* (e. g., height), all above-mentioned operations are allowed (division, for instance, provides ratios). Data on the *interval level* (e. g., temperature measured in Celsius or Fahrenheit) do not have a meaningful zero as point of reference and thus do not allow for ratios, while the comparison of differences is still meaningful. Data on the *ordinal level* (e. g., the parent relation) can be compared hierarchically only, and on *nominal level* (e. g., eye color) only up to equality. Over time, this classification was discussed, refined and expanded, but never discarded [144, 35].

The mathematical field of Order and Lattice Theory [97, 34, 163, 21, 88, 42] is a canonical theory for modeling ordinal data. Its development started – independently of the above-mentioned quest for a suitable theory for measurements – in the second half of the 19th century as algebraic logic with the aim of George Boole [23], Charles S. Peirce [150] and Ernst Schröder [164] to formalize a calculus for concept hierarchies as cornerstone for a mathematical logic. In symbolic knowledge representation, ordered sets are extensively used for this purpose. For instance, in RDFS, the set of all `rdfs:Class`es with the `rdfs:subClassOf` relation is an ordered set, and in the field of conceptual graphs, one frequently assumes that the type hierarchy is even a lattice [168].

3   Hierarchies in Data Science

Data Science is the scientific field of analyzing data and extracting knowledge from data. This understanding of data science as an interdisciplinary field grew over the past decades [208, 152, 58, 41, 46] as a confluence of methods from statistics, data mining/knowledge discovery, machine learning, data management and big data. Here, we briefly discuss the relationship between some of these areas to ordinal data.

In *statistics*, several correlation measures exist for rankings (e. g., Kendall's τ [111], Spearman's ρ [170], Goodman and Kruskal's γ [83]) and tests for the significance of correlation (e. g., Mann–Whitney U test and Wilcoxon signed-rank test [213]). However, these measures and tests are designed only for *linear orders* (i. e., multi-inheritance and pairs of incomparable elements are not considered). In *data mining and machine learning*, most methods are designed for numerical (and some for categorical) data. When ordered sets occur, then mostly as part of the model, e. g., in decision trees [27], as topology of neural networks [114], as inclusion hierarchy for frequent item sets [3] (cf. next subsection), or as structure for coarser and finer set systems in hierarchical agglomerative clustering [212]. Often data points are organized in (linear) rankings, as a result from some numerical score, e. g., the predicted accuracy in supervised learning, the cosine similarity to a given query in information retrieval, or the result of a learning-to-rank approach. Here again,

these methods produce linear orders. A setting where the resulting order is non-linear is *ontology learning* (see below).

Ordinal Data Science will profit from and contribute to the field of *machine-learned ranking* [129, 85], particularly *preference learning* [106], which is a subfield of AI concerned with label ranking and instance ranking. Other areas in AI do also turn to ordinal data, e.g., "Deep Ordinal Reinforcement Learning" [230] attempts to adapt reinforcement learning to ordinal data. Machine learning methods for linear ordered sets are already somewhat further developed, e.g., treatment of linear ordered classes in supervised learning [32]. The same is true for the important machine learning method *metric embedding*, which was recently adapted to linear ordered data [165] and ordinal constraints inferred from metrics [115, 199]. An even more extensive study on the use of linear ordered data [109] concluded that even in the presence of metric features the study of purely ordinal features might be fruitful. Early methods of how this might be done are already being explored in a rather practical way [231]. Finally, Ordinal Data Science is strongly related to the discipline *statistical relational learning*, a research area that itself has a lot in common with many other AI areas, such as reasoning and knowledge representation. Foremost their advances in treating graph data [159] is of high relevance to Ordinal Data Science and vice versa.

4 Order-Theoretical Foundations

4.1 Ordered Sets

The key concept for studying ordinal data is an ordered set. An *ordered set* $(P, \leq)$ is a directed graph (i.e., it consists of a set P and a binary relation $\leq$ on P) such that $\leq$ is reflexive ($p \leq p$), transitive ($p \leq q$, $q \leq r \implies p \leq r$), and antisymmetric ($p \leq q$, $q \leq p \implies p = q$) [21]. This is equivalent to a transitive, directed acyclic graph with self-loops. Ordered sets could thus be seen as special cases of directed graphs or networks, and indeed some results of graph theory and network analysis are used; but order theory extends this with its own, specific flavor: For instance, the acyclicity of an ordered set simplifies its decomposition (e.g., , by filters and ideals), results in more specific structural measures such as height, width and dimension, and allows for more readable visualisations.

A typical ordered set is the set of all human beings together with the `is_ancester_of` relation. Another example are the dependencies of tasks in production planning. Note that, in general, we do not require comparability of all elements. The latter would mean that the order is *total* (also called *linear*), i.e., that, for all $p, q \in P$, one of $p < q$, $p = q$ or $p > q$ holds. To emphasize that comparability is not required, ordered sets are also called *partially ordered sets* (or *posets* for short) in the literature, especially in computer science.

While one cannot exploit numerical operations for describing and analyzing ordinal data, they come with other constructions, such as *order filters, order intervals, Pareto optima* etc.

For finite ordered sets – as studied in data science – we can reduce $\leq$ without loss of information to the *neighboring relation* $\prec$ by removing all pairs that can be deduced by reflexivity and transitivity. This is the relation predominantly used for visualization. Note that in many cases, the order is not *linear* (i.e., there may be incomparable elements and multi-inheritance).

A particularly useful kind of order relations are *lattices*, as they provide more structure and operations: In a *lattice*, any two elements p and q always have a *unique least common upper bound* $p \vee q$ and a *unique greatest common lower bound* $p \wedge q$. These two operations are often called *join* and *meet* respectively. One could say that a lattice is a hierarchy in which one can navigate up and down with these two operations. A *complete lattice* is an ordered set where every subset of P (including the empty set and infinite subsets) has meet and join. This particular property is only relevant if P is not finite, since each finite lattice is complete. Although the datasets considered in
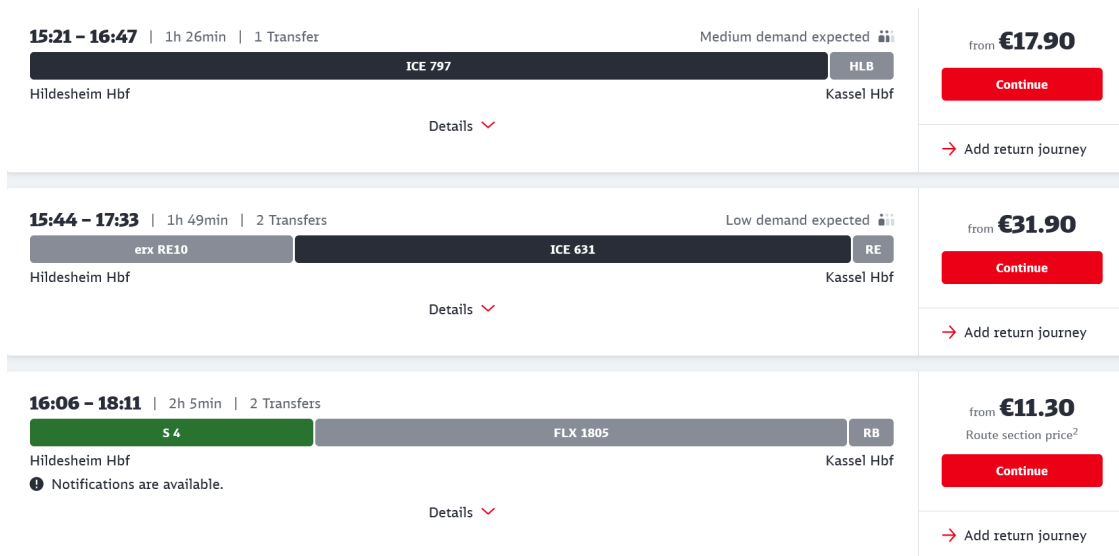
Figure 1 Pareto optima of train connections between the cities Hildesheim and Kassel as inferred from the German railway corporation Deutsche Bahn via `www.bahn.de`.

data science are always finite, a deeper understanding of the infinite case, and hence of complete lattices, can provide new insights.

A complexity measure for an ordered set or lattice $(P, \leq)$ is its *order dimension*, i.e., the number n of linear orders $\leq_1, \ldots, \leq_n$ on P with $\bigcap_{i=1}^{n} \leq_i = \leq$. In other terms, it is the smallest $n \in \mathbb{N}$ such that the ordered set is embeddable in the Cartesian product of n linear orders [206]. In a previous work [50] we explored how to reduce the complexity of a dataset by identifying parts that contribute to it having a high order dimension. However, computing (i.e., deciding) the order dimension is $\mathcal{NP}$-complete if the dimension is higher than two [228]. Even approximations are proven to be $\mathcal{NP}$-complete [99].

4.2 Ordinal Structures

As data may be ordered by different criteria, we define, more generally, an *ordinal structure* $\mathbf{P} = \langle P, (\leq_i)_{i \in I} \rangle$ as a set P of objects that is equipped with quasi-orders $\leq_i$ [176].[4] *Quasi-orders* are transitive, reflexive relations. Not insisting on anti-symmetry for $\leq_i$ means that there may exist objects that are indistinguishable in this relation (e.g., there may be different objects with identical prize).

▶ **Example 1** (Ordinal Structure). Train connections between two cities are an example for an ordinal structure. They can be ordered by different criteria, e.g., by departure time (closer to time given in the query is better), travel duration (shorter is better), number of transfers (fewer is better), expected demand (lower is better), and price (lower is better). Figure 1 shows the *pareto optima* for connections between Hildesheim and Kassel for a requested departure at 3pm, i.e., those connections for which it is impossible to improve a criterion without becoming worse in any of the other criteria.   ⌟

In our vision, we extend the notion of ordinal scale to allow for the analysis of all kinds of ordinal data. With terms like "*order*", "*ordinal*", we always refer to arbitrary orders – be they

---

[4] The entities may – and in most cases will – have additional attributes on other levels of measurement. The development of hybrid analysis methods is of particular interest here.

linear or not.[5] We thus go beyond Stevens' original understanding of *"ordinal scale"* which only referred to linear orders. Even though this seems to be only a minor modification at first glance, it has far-reaching consequences for the question of what "measurement" actually is, as well as for the complexity of computational problems.

Ordered sets in general do not provide much structure. However, every ordered set is embeddable into a complete lattice, its *Dedekind-MacNeille completion* [135] (which generalizes Dedekind's construction for embedding $\mathbb{Q}$ in $\mathbb{R}$). This allows for tapping into the rich algebraic theory of lattices and of universal algebra [20, 87, 30]. The class of lattices forms a so-called *variety*, which implies that every homomorphic image, every factor lattice and every subset of a lattice that is closed under join and meet, as well as all direct products and direct sub-products of lattices are lattices again. Hence, these operators can be exploited for composing and decomposing lattices – and thus eventually of the ordered sets that generated them. In the worst case, the size of the Dedekind-MacNeille completion is exponential in the size of the ordered set. However, this growth is expected to be tame for real-world data, due to low amounts of multi-inheritance in it. Ganter and Kuznetsov provided an algorithm for its computation [73] in $\mathrm{O}(c \cdot w \cdot n^2)$ with c being the size of the completion, w the width of the ordered set and n its size. The new nodes bear (yet unexplored) potential for identifying substructures in ordered sets.

4.3 Formal Concept Analysis and knowledge spaces

A direct application of lattice theory to data analysis has been established as *Formal Concept Analysis* (FCA) by R. Wille [216]. It follows the spirit of Boole, Peirce and Schröder by computing concept hierarchies from datasets. In its most basic form, FCA derives a *concept lattice* from a so-called *formal context*, a dataset (G, M, I) consisting of a set of objects G, a set of binary attributes M, and $I \subseteq G \times M$ indicating which objects have which attributes. *Concepts* are all maximal pairs (A, B) with $A \subseteq G$ and $B \subseteq M$ s.t. all objects in the *extent A* share all attributes in the *intent B*, and vice versa. The set of formal concepts is denoted by $\underline{\mathfrak{B}}(G, M, I)$ and the relation $\leq \subseteq \underline{\mathfrak{B}}(G, M, I) \times \underline{\mathfrak{B}}(G, M, I)$ with $(A, B) \leq (C, D) :\Longleftrightarrow A \subseteq C$ constitutes an order relation on $\underline{\mathfrak{B}}(G, M, I)$. Thereby (A, B) is designated as a subconcept of (C, D). More importantly, $\leq$ is a lattice order. This fact follows from the *basic theorem on Concept Lattices* [78, Theorem 3]. This theorem states that for any subset $\{(A_i, B_i)\}_{i \in I}$ of $\underline{\mathfrak{B}}(G, M, I)$ there exists an infimum (meet) $\bigwedge_{i \in I}(A_i, B_i)$ and a supremum (join) $\bigvee_{i \in I}(A_i, B_i)$ in $\leq$. The set of all concept extents forms a closure system, i.e., it is a set system which is closed under intersection, and the same holds for the set of all concept intents.

▶ **Example 2.** The following table depicts a real world data set about paintings by Rembrandt (the object set G) and their properties (M), taken from "Concept lattices and conceptual knowledge systems" [215]. The formal context $\mathbb{K} = (G, M, I)$ and its incidence I is represented by means of a *cross table*:

$\mathbb{K}$	Family Portrait	Group Portrait	Oak	Canvas	≥ 1660
Nightwatch		×		×	
Anatomical lessons		×		×	×
Portrait Titus	×			×	×
Staalmeesters				×	×
Mother	×		×		

---

[5] When addressing linear orders, we will state this explicitly. For sake of linguistic simplicity, we will also summaries all levels of measurement above ordinal – essentially interval and ratio – as *"numerical"*.
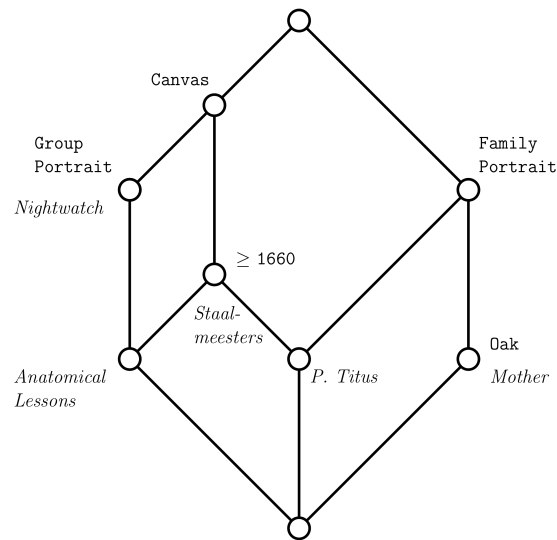
Figure 2 Line diagram of the concept lattice from Example 2. The drawing uses a short-hand notation for the labels. The nodes represent the set of formal concepts. Attributes are drawn on top of nodes and objects below. Only the first occurrence of an attribute and the last occurrence of an object is annotated (in the reading direction from top to bottom).

The line diagram of the concept lattice corresponding to $\mathbb{K}$ is shown in Figure 2. Such a line diagram displays the covering relation, i.e., the subset of $\leq$ which contains only the comparable elements that are immediate neighbors. From the lattice we can infer various types of information. For example, we find that all paintings that were painted in 1660 or later were painted on canvas. More intriguingly, family portraits that were painted on canvas occurred in 1660 or later. For a comprehensive list of how-to infer information from lattices, and in particular lattice diagrams, we refer the reader to the literature [78]. For a more in-depth discussion of questions concerning the diagrammatic representation of order structures in the specific case of the present work, please refer to Section 9.

Via *conceptual scaling* [77], FCA can be applied to any kind of data. Since its foundation in 1982, FCA has grown to an extensive theory [79, 33, 74, 64] and has seen several extensions, for instance, to fuzzy concepts [11]. It has been applied in a large variety of domains (see, e.g., [195, 68] for surveys). FCA has been connected to other disciplines, such as Software Engineering [167], Psychology [169], Scientometrics [44], and Social Choice Theory [108].

Similar to FCA is the theory of *knowledge spaces* by J.–P. Doignon and J.–C. Falmagne [45]. They define a *knowledge structure* as a pair $(Q, \mathcal{K})$ where Q is a non-empty set and $\mathcal{K}$ is a family of subsets of Q. The set Q is called *domain* and its elements are called *questions* or *items*. The subsets in $\mathcal{K}$ are called *(knowledge) states*. The intuition for this definition is that Q is a collection of questions of some school subject, and that the sets in $\mathcal{K}$ describe which combinations of questions could be solved by individual students. A *knowledge space* is then a knowledge structure that is closed under set union. This is inspired by the assumption that if two students have knowledge states K_1 and K_2, resp., then it should be possible for them to learn what the other student knows, resulting in $K_1 \cup K_2$.

A knowledge space is thus a kernel system, i.e., a set system closed under union. If $\mathcal{K}$ is a kernel system, then $\{Q \setminus K \mid K \in \mathcal{K}\}$ is a closure system, and vice versa. This means in particular that all theoretical and algorithmic results of FCA can directly be transferred to knowledge spaces, and vice versa (see in particular [72]). From a mathematical viewpoint, the structures of closure

systems, kernel systems, complete lattices and concept lattices are equivalent [216], as there exist natural 1-to-1 correspondences. These structures are also in a natural 1-to-1 correspondence to sets of implications (of propositional Horn logic) [214] and functional dependencies (of database theory). These natural correspondences mean that for the computation, analysis and (de)compositions of such structures one can switch to the representation that is most suitable for the task at hand. Formal contexts play an important role when it comes to computations, as their sizes are logarithmic in comparison to the sizes of the corresponding (concept) lattices.

4.4 Ordered Sets and Lattices in Computer Science

Specific ordered sets and lattices are frequently found in computer science: For instance, queues and lists are linear orders, (B-)trees are used for indexing data, tries for Information Retrieval, CPOs for modeling the semantics of recursion, and Boolean algebras for propositional logic implemented in logical gates. Inheritance (including multiple inheritance) in object-oriented programming induces a partial order on the set of classes. An introduction to the mathematical foundations of order theory relevant to *Data Mining* is provided in [166]. Ordered sets are also the key ingredient for modeling dependencies between tasks in scheduling problems.

FCA is used extensively in the field of *Ontology Engineering / Semantic Web*, in particular for conceptual information systems [192], knowledge acquisition [186, 187, 189], for knowledge base completion in Description Logics [6, 7], for computing probably approximately correct implication bases [24], for ontology management [185, 76, 37], [160, 14], and text corpus analysis [104]. An extensive survey on FCA based models and techniques in knowledge processing is given in [153].

4.5 Algorithms and Software for Ordered Sets and Lattices

The number of algorithms and software tools in the field of order theory and (concept) lattices is extensive. Listing all relevant works would go far beyond the scope of this article; we only mention the most important and recent tools. Ordinal Data Science may benefit from algorithms for the uniform sampling of linear extensions [29] and the uniform sampling of ordered sets [119], and the even more complicated problem for sampling binary relations in general, which was tackled in [61]. The research field of Formal Concept Analysis can contribute algorithms for the fast generation of concept lattices, e.g. In-Close4 [5] or Parallel Close-by-One [120]. Also, the computation of (inter-feature-) dependencies will be an essential asset. The computation of such dependencies can be done with proven parallel canonical base algorithms [122]. One may also resort to approximate methods, since the related computational problems become computationally intractable with increasing input size. For example, the dependencies can be computed with *probably approximately correct learning* methods [25, 24], a combination of order theory and classical machine learning model. Many of the mentioned algorithms are already available, or are currently being implemented, in the analysis software `conexp-clj` [91], which is a reimplementation of `conexp` by means of the functional programming language *Clojure*. It is currently maintained and further developed in the group of the authors. There are many further FCA software tools under active developed (and even more historic ones), most with a particular specialization. Just to name a few: *Graph-FCA*[6], an extension of FCA to knowledge graphs; *lattice miner* [142]; *conexp-ng*[7], an educational variant of *conexp*; *fcaR*[8], a package for the statistical software suite R.

---

[6] `https://bitbucket.org/sebferre/graph-fca/src/master/`
[7] `https://github.com/fcatools/conexp-ng`
[8] `https://cran.r-project.org/web/packages/fcaR/fcaR.pdf`

In the next five sections, we will now discuss exemplarily five out of many possible lines of research in the new field of Ordinal Data Science in more detail.

5   Ordinal Measurement Theory

While measurements in physics were accepted without objections, other disciplines, in particular psychology, asked in the 20th century – and still do so today – if and under which conditions non-physical sensations such as loudness, fear, or intelligence can be measured. The most widely acknowledged answer to this is the *Representational Theory of Measurement (RTM)* [121, 198, 131], which aims at identifying criteria that justify a measurement on a given level. These criteria can (at least in principle) be verified empirically without actually performing the measurement. RTM defines *measurement* as a homomorphism from an empirical relational structure $\mathbf{A} = \langle A, (R_i)_{i \in I}\rangle$ – i.e., a set A of empirical objects (such as a collection of rods) equipped with qualitative relations R_i (such as "rod a is shorter than rod b" and "when laid end to end, rods a and b together have the same extension as rod c") – to a *numerical relational structure* $\mathbf{B} = \langle B, (S_i)_{i \in I}\rangle$ [151, 121]. In the latter, B is usually the set $\mathbb{R}$ of real numbers (or $\mathbb{R}^n$ and other variations) together with relations such as $x < y$ and $x + y = z$, even though several researchers [158, p. 253] observe that in some cases it might be beneficial to allow for other algebraic – "non-numerical" – structures.

RTM puts a strong emphasis on measurements into the real numbers, and less emphasis on ordinal phenomena. As it allows for linear ordinal scales only, RTM does not allow for studying phenomena involving incomparability or multi-inheritance. Therefore, inspired by RTM, R. Wille et al. [75] developed a theory of *conceptual measurement* where $\mathbf{B}$ takes the form of specific lattices, called *conceptual scales* – a first step towards an ordinal version of RTM. In this context several studies followed [77, 224, 154, 188, 182]. Conceptual measurement provides structural statements about homomorphisms to standard scales, but it does not yet provide a full theory about the conditions under which ordinal phenomena can be measured with the different types of ordinal scales. In order to provide a *theoretical foundation to Ordinal Data Science*, we propose to extend RTM by non-numerical scales. This will allow for empirically testing on a given dataset which scales are appropriate for its analysis.

To this end, in this section, we suggest to analyze under which conditions measurements into (linear and in particular in non-linear) ordinal scales exist, and how unique they are (i.e., under which permissible transformations they remain invariant). We further suggest investigating how the – potentially high – order dimension of the original data relates to the order dimensions of a set of scales which together fully measure the dataset. At last, we consider non-linear versions of Guttman scales [89], to allow for measuring simultaneously objects and their features.

In the long run, this endeavor may open the avenue to a *"Grand Unified Theory of Measurement"* which will allow, for instance, in (the philosophy of) physics to discuss the analysis of symmetries of elementary particles (by letting $\mathbf{B}$ be a (non-ordered) symmetry group) in the same terms as the numerical measurement of their mass and charge.

5.1   State of the Art and Open Questions

Based on the successful role of measurement in physics and the demand of other disciplines – in particular psychology – , in the late 19th and early 20th century a universal definition of measurement was sought [63]. S. S. Stevens [174] provided an operational approach by distinguishing four *levels of measurement*: nominal, ordinal, interval, and ratio; and characterized them by families of automorphisms of the set of real numbers that respect the scale type ("admissible functions"). Stevens' levels of measurement have been (and still are [162, 202]) under dispute. A particular controversial question among practitioners that is discussed since Stevens' paper for over 70 years is whether computing the mean of ordinal data is an allowed operation or not (see, e.g., the

football number dispute [130, 162] as illustrative example). In practice, this is frequently done (e. g., aggregating the jury votes in gymnastics or figure skating), while in other cases this is considered bad practice (e. g., aggregating reviewer judgments in scientific peer review, where often the average is considered as bearing little meaning, and therefore consensus is sought). Several extensions were proposed for Stevens' levels (e. g., [144, 35]), see [201] for a systematic survey. All those extensions, however, consider ordinal data as a separate category; so our endeavor is widely agnostic to this discussion.

Stevens did not answer the question of how to decide, in a given setting, which level is appropriate. *Representational Theory of Measurement (RTM)* [121, 198, 131], [151, 158] provides an answer with the notion of conjoint measurement. RTM defines a *measurement* as a mapping of an empirical relational structure (ERS), i. e., a set of objects equipped with qualitative relations such as "is larger than", to a numerical relational structure (NRS). Luce and Tukey [132] proved that whether some empirical sensation is a continuous variable depends on characteristics of the ERS and is thus in principle empirically testable. How feasible this testing is in practice is one of the questions of the still ongoing discussion on RTM, in particular in psychology [204, 28]. The envisaged substantially new methods in Ordinal Data Science for measuring ordinal data may turn RTM more acceptable to practitioners, as it is easier to show that ordinal methods are legible in a given situation than it is for methods requiring interval or even ratio level.

Several researchers working on RTM (e. g., [158, p. 253]) stated that it might be beneficial not to focus solely on the ordered field of real numbers as NRS, but to allow instead for other mathematical structures. However, this approach has been followed rarely. For instance, R. Wille and his group took up this line of thought, as described above: On one hand the group studied the representation of ordinal structures in the real numbers [223, 224, 226] and weaker algebras such as ordered quasi-groups [225]. On the other hand, work on a theory of *conceptual measurement* [75, 77, 188, 182, 154] was started, where $\mathbb{R}$ is replaced by complete lattices which can be understood as *partial ontologies*.

As Dedekind-MacNeille completions preserve important properties of ordered sets (e. g., order dimension) and are isomorphic to a concept lattice, conceptual measurement is just a special – but conceptually and computationally convenient – way of formalizing a general theory of ordinal measurement. Conceptual measurement has frequently been used in applications. However, a theoretical investigation in terms of RTM (e. g., what are the equivalences of cancellation and solvability when $\mathbb{R}$ is replaced by a lattice?) is still lacking.

5.2 Promising Research Questions

The main objective in this field of research will be to *extend Relational Measurement Theory to non-numerical scales*, as it would allow to empirically test on given data which of these scales are appropriate. We will address three key aspects: The first is to establish *theorems about the existence and uniqueness of measurements into non-linear ordinal scales* and to explore how well the *theory can be extended to fundamentally different scale types* (e. g., (non-ordered) symmetry groups), as a first step towards a *"Grand Unified Theory of Measurement"*. As the order dimension is a strong indicator for the complexity of an ordered set, we will then discuss the *relationship between the order dimensions* of an ordinal structure and its scales, in order to provide guidance for suitable decompositions and factorizations in Sections 7 and 8; and extend the discussion to scales that measure simultaneously the objects and their attributes.

5.2.1   Measuring into ordinal scales

Research in Representational Measurement Theory is centered around the real numbers as "numerical relational structure" (including derived structures, such as semiorders and interval orders for modeling error and variation) and real vector spaces. Measurements with other types of ordered sets have not been studied systematically, except for the theory of conceptual measurement as described above. So a general research question will be: how does measurement theory look like if we replace $(\mathbb{R}, \leq, +, -, \cdot, 0, 1)$ by a specific ordered set $(P, \leq)$ or (complete) lattice $(L, \leq, \wedge, \vee)$? Are there equivalences to the notions of cancellation, solvability and conjointness, and how are they constituted?

Following the line of traditional RTM, we suggest to study in particular the representation problem and the uniqueness problem for specific scale types. The *representation problem* deals with the question of which conditions a dataset has to fulfil in order to permit a measurement in a given scale. For binary attributes on the empirical side and many standard ordinal scales (nominal, linear, interval, multi-ordinal, contra-nominal, contra-ordinal, and convex-ordinal scales) on the "numerical" side, this problem was addressed in [79, Chapt. 7]. In this case, the task is reduced to translating the results to the terminology of RTM. For other types of empirical data, and for other types of scales (e. g., trees) and ordinal scales in general[9] on the "numerical" side, the necessary theorems still have to be established. Once a measure has been identified, one should also consider the *uniqueness problem* by analyzing invariance under permissible transformations: How unique is the resulting measure or scale?

Our results constitute the first steps towards a *"Grand Unified Theory of Measurement"*: Once a theory for ordinal measurement as been established, one may start to explore – as a side-trip and preparation of more extensive research – how well this theory can be extended to fundamentally different scales, as for instance to non-ordered groups (such as symmetry groups).

5.2.2   Dimensionality

The complexity of an ordered set is not only determined by its cardinality, but also to a large extent by its order dimension. In contrast to numerical data, where most frequently each feature spans one dimension, the identification of the different dimensions of an ordered set is not straightforward – even determining their number is $\mathcal{NP}$-complete [228]. The question of dimensionality thus takes a completely different stance as in standard RTM (where dimensional analysis means that physical dimensions such as length, mass, time duration, speed, and their relationships are studied [121, Chapt. 10]). Breaking down a complex ordered set into smaller, less complex parts (i. e., parts with lower dimensionality) is an important – but challenging – task for data science, which we will discuss further in Sections 7 and 8. The following example shows that the development of a theory about the dependencies between the ordinal dimension of an ordered set and the dimensions of the scales it can be measured to might be of high interest.

▶ **Example 3** (Order Dimension). Table 1 depicts the final standings of the German Soccer Bundesliga season in 2022/23. In this dataset, the strength of the soccer clubs is measured using four different scales. In the columns "Won" and "Goals for", a high value implies a strong club, while the converse is true for "Lost" and "Goals against". The direct product of these four linear orders describes an order relation in which two clubs are comparable, if one dominates the other in all four categories. Two clubs may be incomparable in this ordering, as for instance Borussia Dortmund has more wins than FC Bayern München, but also more losses.

---

[9]   Reminder: In this project, ordinal scales may be any type of ordered set, either linear or non-linear. What is usually called ordinal scale (such as Likert scale or Beaufort) is a special, linear case thereof.

■ **Table 1** The final league table of the 2022–23 Bundesliga, Germany's premier soccer league. (Pos: Position, W: Won, D: Drawn, L: Lost, GF: Goals for, GA: Goals against, GD: Goal difference, Pts: Points)

Pos	Team	W	D	L	GF	GA	GD	Pts
1	FC Bayern München	21	8	5	92	38	+54	**71**
2	Borussia Dortmund	22	5	7	83	44	+39	**71**
3	RB Leipzig	20	6	8	64	41	+23	**66**
4	1. FC Union Berlin	18	8	8	51	38	+13	**62**
5	SC Freiburg	17	8	9	51	44	+7	**59**
6	Bayer 04 Leverkusen	14	8	12	57	49	+8	**50**
7	Eintracht Frankfurt	13	11	10	58	52	+6	**50**
8	VfL Wolfsburg	13	10	11	57	48	+9	**49**
9	1. FSV Mainz 05	12	10	12	54	55	−1	**46**
10	Borussia Mönchengladbach	11	10	13	52	55	−3	**43**
11	1. FC Köln	10	12	12	49	54	−5	**42**
12	TSG 1899 Hoffenheim	10	6	18	48	57	−9	**36**
13	Werder Bremen	10	6	18	51	64	−13	**36**
14	VfL Bochum	10	5	19	40	72	−32	**35**
15	FC Augsburg	9	7	18	42	63	−21	**34**
16	VfB Stuttgart	7	12	15	45	57	−12	**33**
17	FC Schalke 04	7	10	17	35	71	−36	**31**
18	Hertha BSC	7	8	19	42	69	−27	**29**

The order dimension of the domination order of the 2022/23 season is three. This means that only three linear orders are necessary to represent the combination of the four scales. There is hence some redundancy in the data – but not enough to produce a total ranking. The latter would hold if the order dimension of the domination order would be one. It would then, for every pair of clubs, provide an answer to the natural question of which of the two is better than the other. As this is not the case, further rules are used to compute the final position of each club in the league. This results in a linear extension of this domination order. The choice of these rules captures the understanding which quality criteria are considered most important.

The current rules of Bundesliga [127] are one possible way to reach such a linear extension: First, a "Points" scale is derived, where each team is awarded three Points for every win and one point for every draw. This results in a linear extension of the direct product of the "Won" and "Lost" scales, which are considered to be the most important criteria. However, this may not yet result in a total ranking as multiple teams can have the same number of points (for instance Borussia Dortmund and Bayern München). In this case, the tie is broken using the "Goal difference", which itself is a linear extension of the direct product of "Goals for" and "Goals against". Thus, from the original four scales, the two linear extensions "Points" and "Goal difference" are computed as derived scales. Then, these two give the final ranking by taking once again a linear extension of the two derived scales.[10]

Note that the rules are arbitrarily chosen (e. g., , one could assign only two points for a win, as is done in other sports (and was done earlier in Bundesliga), or consider the direct comparison of clubs in a tie as second criterion) and for every linear extension of the domination order there is a set of rules that gives rise to the ranking described by this extension. Thus, one could consider each linear extension a valid final ranking of the participating teams.   ⌐

---

[10] [127] provides further rules for cases where this still does not resolve all ties.

A next research question could be a dimensional analysis of combinations of the standard ordinal scales – for each of them studied separately, the answer is straightforward, but dimensional properties of full measures of an ordered set in a (heterogeneous) collection of scales have not yet been studied systematically. As deciding the order dimension of an ordered set is $\mathcal{NP}$-complete, one has to keep an eye on the computational tractability: Are exact algorithms fast enough for typically sized datasets? Alternatively one may resort to computing upper bounds, together with the *realizers* (i. e., the spanning linear orders) [229]. Also, the development of efficient algorithms for computing lower bounds (and eventually also the critical pairs) will be of interest.

5.2.3   Joint scales on objects and attributes

Guttman scales have been introduced as joint scales for objects and attributes. Consider, for instance, a set G of students, a set M of exercises of (assumingly) increasing difficulty, and a relation I stating which student solved which exercise. If the corresponding concept lattice is linearly ordered then it provides a linear ranking both for the difficulty of the exercises and for the ability of the students. This can be interpreted such that there is no exception to the rule: a student who is capable of solving a specific exercise can also solve all easier exercises. This is equivalent to the original definition [89] which requires the existence of maps $s\colon G \to \mathbb{R}$ and $e\colon M \to \mathbb{R}$ s.t. $(g, m) \in I \iff s(g) \leq e(m)$. In practice, Guttman scales can be observed frequently, but often only up to some "noise" or "impurity". Standard approaches are then either to assume that the assumptions hold in principle and to statistically find best-fitting functions [38], or to analyze whether the data result from more than one Guttman scale. We propose to follow two new order-theoretic approaches for analyzing these "impurities" instead. The first approach is to relax the condition of linear order and to explore if a more general scale – e. g., a tree – may be suitable as a joint scale. One can then identify necessary and sufficient conditions for the existence of such generalized Guttman scales as well as criteria for their minimality (e. g., the lowest cardinality, the lowest dimensionality, least multi-inheritance). The second approach is to analyze more closely the disturbance of the linear scale by identifying which parts of the scale increase its *local dimension* [207]. Such a tool would allow analysts to investigate whether a linear feature is disturbed by "noisy data" or whether it really shows some unexpected non-linear characteristics.

6   Ordered Metric Spaces

Real-world data is feature-heterogeneous [31], i.e., it is (apart from others) composed of ordinal and numerical dimensions. The former can be viewed as a partially ordered set, and the latter often leads naturally to a metric. In principle, both can be analyzed separately, but this approach has two disadvantages. First, the results calculated in this way can be contradictory to each other. Second, the knowledge about the connection of the two feature dimensions is not exploited.

We therefore propose to develop theoretical and practical methods that honor the compatibility (or consistency) of order relations and metrics on a set. To this end, we envision the unified representation of such data by means of *ordered metric spaces (om-spaces)* (P, R, d), where P denotes a set, R is an order relation on P, and a d a metric $d\colon P \times P \to \mathbb{R}^+$. This modeling is inspired by similar investigations in *metric measure spaces*[11] [146, 196].

---

[11] In metric measure spaces, *measures* are those of mathematical analysis, being defined over σ-algebras [90]. They model a different aspect of the concept of measurement as the homomorphisms between relational structures in RTM do, and should not be confused with them.

The corresponding highly ambitious task, due to its novelty, is to develop a *structure theory for om-spaces*. Particularly interesting in this context is the study of om-space embeddings into Euclidean space. Since the latter is a very commonly used space for machine learning, the question of how to measure *ordinal distortions* is imperative. Vice versa, it is important to understand to which extent the extraction of linear orders from data via valuation functions is meaningful with respect to om-spaces. With a new class of distortion measures – and computationally tractable approximation algorithms – which allow to *quantify the compatibility of R and d*, one might be able to answer a range of fundamental questions about the relationship between orders and metrics, including: 1) If a metric has been (machine) learned based on information about the order, how consistent is the metric with the order? This will provide a fundamental new quality criterion for order-related machine learning tasks. 2) How strong do order and metrics "agree" with their view on the dataset? An answer might be highly beneficial to an investigation whether order and metrics reflect the same hidden structure. 3) How justified is it to represent the order by means of a metric? If the answer is satisfactory for a given dataset, this would justify the use of the rich toolkit of numerical data analysis and machine learning methods for analyzing the ordinal structure. If not, this would indicate that alternative (ordinal) methods are mandatory in this particular case.

6.1 State of the Art and Open Questions

Order relations entail many geometric properties. Likewise, metric spaces, which play a dominant role in data science, have a rich geometric structure. Examples for datasets exhibiting both a metric and an order relation are numerous. For example: Humans have (distances between their) birthplaces and an ancestry order relation; scientists have an ancestry relation (doctoral advisor) and different proximities (topical, geographic, social); mountains are ranked by height and prominence [161] and have a geographic distance. A more general example for an om-space is the set of Airports together with the connecting plane routes relation and the geodesic distances.

The analysis of ordinal data is often based on a (numerical) metrification. Imposing metric structures on ordered sets[12] is a widely studied [143], yet still very active, field of research. Almost all mathematical disciplines contribute to this, e.g., algebra [60], combinatorics and geometry [128], and analysis [110]. Furthermore, the task of (distance) metric learning in the realm of (weakly) supervised machine learning addresses a similar problem, e.g., based on the generalized *Mahalanobis distance*. Contemporary developments [211] do often use deep learning settings. However, recent works claim that the reported achievements are rather marginal [145]. We contributed to the task of finding meaningful metrics in ordinal data in two ways. First, we adapted the notions of dominance and prominence from the research field of orometrics to network science [161]. This process revealed a whole class of new metrics to study. Second, we have further developed the popular *word2vec* method, which is based on a neural network model, for the special requirements of (concept) lattices [48].

A popular embedding structure for ordinal and metric data in the realm of machine learning is the Euclidean space, especially represented as $\mathbb{R}^n$. Depending on the properties of the relation (symmetric, transitive, reflexive, etc.) and the kind of embedding, e.g., *distance*, *similarity*, or *transitional*, different bounds for n [18] are required for distortion free, i.e., isometric, embeddings. Often n equals $|P|$, which may be computationally intractable. Moreover, it is a well known fact [136] that many finite metric spaces are not distortion-free embeddable into $\mathbb{R}^n$, for any $n \in \mathbb{N}$. A simple example for this is a 4-element ring with the graph distance as a metric. Therefore, low

---

[12] In the realm of machine learning a natural consequence when learning an "embedding".

distortion embeddings [1] and different means of measuring distortion [210] are widely studied. Hence, the investigation of distortions for embedding finite om-spaces into $\mathbb{R}^n$, which takes into account metric and order relational characteristics, is at hand, highly ambitious, and up to our knowledge, not conducted. We have already found initial evidence in a special case that the distortion can be extensive [92].

Understanding different metric embeddings of the same ordinal data is possible by comparing the resulting metric spaces. Amongst others, the Gromov-Hausdorff (GH) distance (by D. Edwards [55, 209], M. Gromov [86]) enables such a comparison naturally. However, since GH is computationally intractable [138] one might resort to computable lower bound methods, e.g., the modified GH distance [147], which is time polynomial.

To further mitigate the computational demands in the last paragraphs, we can refer the reader to various works we carried out, e.g., the identification of important features in relational data [94] based on entropy maximization and structure preservation, dataset size reduction using cores in formal contexts [93], and the identification of clones [43], i. e., concept lattice preserving permutations. All works are applicable to om-spaces, as every order relation gives rise to a (unique) concept lattice; and every metric on P can be lifted to the power set.

To the best of our knowledge, studies purely concerned with the interplay of order relations and metrics with respect to data analysis are rare. We can identify in the literature few results about fix points [126, 17] and contractions [2]. Related topological results for ordered metric spaces are of less interest for our envisioned task, as the considered om-spaces are almost always finite, which in turn implies that the induced topological space is discrete.

Coming full circle with our motivation from mm-spaces, we propose to consider different results on valuation functions in ordered sets. This topic is already studied for order relations and lattices [143, 124]. We already contributed to this by transferring ideas from orometrics to bounded metric spaces endowed with binary relations [179, 180]. Even more related to the realm of mm-spaces is our work on an intrinsic dimension of geometric datasets [95, 178]. Although the proposed dimension function relies on mm-spaces, it is fully capable for the dimension analysis of om-spaces.

Finally, comparing real-world data with randomly generated data often reveals new structural insights and provides benchmarks. Yet, uniform sampling of ordered sets is a difficult problem, at least for reasonable set sizes [119]. Our previous work allows to some extent for randomly generating binary relations [61] and null-model generation for (order) relations [62]. To which extent these methods can be useful for the study of om-spaces is an open question.

6.2   Promising Research Questions

Central to this research is the theoretical and experimental study of the triple (P, R, d), as introduced in the last section as *ordered metric space (om-space)*. One should distinguish between originally metric data (and therefore spaces) and those where the metric was obtained through an external (e. g., machine learning) method. Special interpretations of this structure are, for example, directed acyclic graph metric spaces (if the metric values are additive along paths), or weighted directed acyclic graphs (when considering metric values for connected pairs, i.e., relational pairs, only). One should approach the goal for a structure theory of om-spaces for ordinal data science from different angles, most importantly graph theory, order theory, and the analysis of metric spaces.

Any new result will surely build up on modelings concerned with R being a tree or forest relation and d a graph metric on it. The guiding methodology should be to examine the extensive apparatus of order theory with respect to the om-space. In particular, it is necessary to develop new concepts and methods to assess algorithmically the compatibility of R and d. A special focus

$\mathbb{K}$	Aeroflot	Air France	Lufthansa	Scandinavian	British Airways	Austrian A.	Alitalia
Hamburg	×	×	×	×	×	×	×
Lisbon		×	×		×	×	×
Madrid	×	×	×	×	×		×
Moscow	×	×		×			×
Toulouse		×	×		×		×
Budapest	×	×	×	×	×	×	×
Dresden	×	×	×				
London	×	×	×	×	×	×	×
Marseille	×	×	×		×		×
Rom	×	×	×	×	×	×	×
Palma D.M.		×	×	×	×	×	×
Leipzig/Halle			×			×	
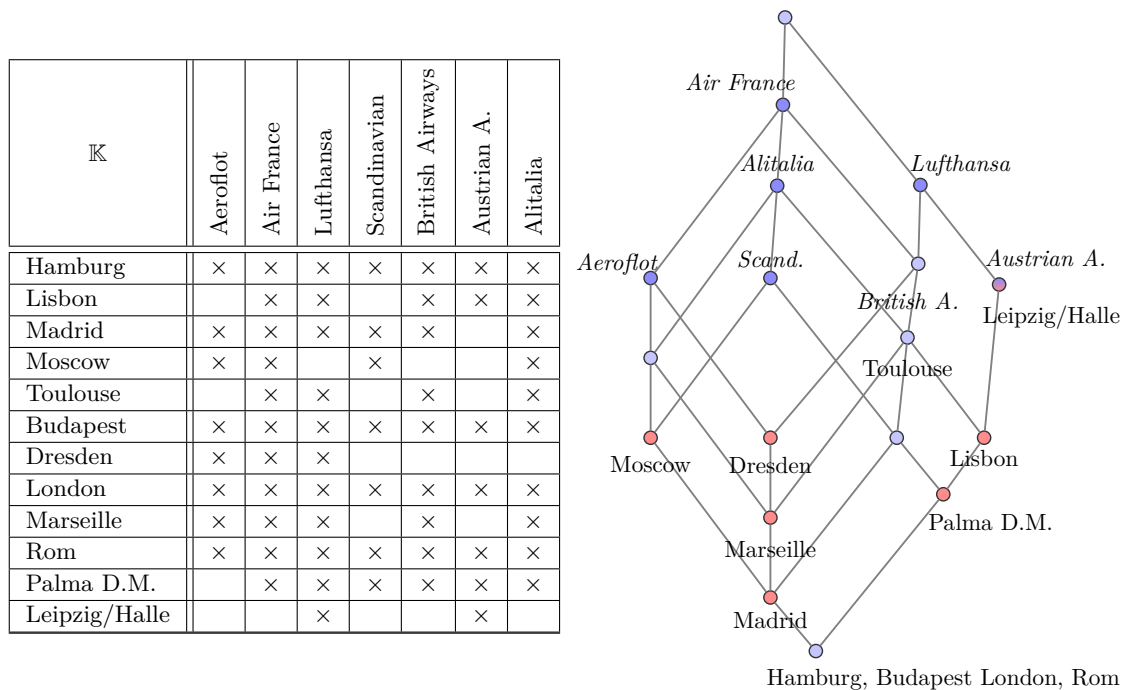
Figure 3 Left: Simplified formal context from real-world data about cities (objects) that are serviced by airlines (attributes). **Right:** Corresponding concept lattice.

should be set on understanding if and how different kinds of dimensionality (e.g., order dimension, intrinsic dimension) do inflict the compatibility of R and d as well as a possible later embedding of (P, R, d) into Euclidean space.

6.2.1 Structure Theory of om-Space Datasets

To explore the distortion of order relations on sets with respect to indigenous and externally imposed metric functions is the focus of this research task. A first attempt to assess the distortion of an order relation in this sense can be based on the distortion of a map φ between two metric spaces (X, d_X) and (Y, d_Y). In the realm of machine learning numerous distortion functions are studied [210]. A promising start may be with $dis(\varphi) := \sup_{x, \hat{x} \in X} |d_X(x, \hat{x}) - d_Y(\varphi(x), \varphi(\hat{x}))|$. To study the *distortion introduced by relation R in om-space* (P, R, d), one could then study the canonical (by means of the Galois connection) map $\varphi : P \to 2^P$ defined by $x \mapsto \{y \in P \mid (x, y) \in R\}$ from metric space (P, d) to $(2^P, d_H)$ where d_H is the Hausdorff metric derived from d. This setting allows for analytical tools to assess the relational distortion (as well as expansion, extraction, etc.) of ordered metric spaces for real world datasets. If R is a purely reflexive relation, i.e., the simplest order relation on P, there will be no distortion, i.e., $dis(\varphi) = 0$. Yet, for arbitrary relations in real-world datasets the change of distortion is unknown.

▶ **Example 4.** The data shown in Figure 3 and Table 2 can be comprehended as an om-space. The elements of P are the cities, the map d assigns any two cities their geodesic distance, and any two cities are in relation R iff there is an airline servicing both cities. ⌟

Applying this new measure on real-world data would require non-incremental research to develop efficient algorithms to compute them, as computational demands increase drastically with increasing $|P|$. One might face this highly ambitious challenge by drawing from results on closed sets, which arise naturally from the closure operator generated by R. Hence, one may employ

Table 2 Geodesic distances between cities (in nautical miles [nmi]) that are the objects of the formal context in Figure 3.

	Hamburg	Lisbon	Madrid	Moscow	Toulouse	Budapest	Dresden	London	Marseille	Rom	Palma D.M.	Leipzig/Halle
Hamburg	0	2198	1781	1765	1278	950	378	745	1186	1326	1658	289
Lisbon	2198	0	513	3892	1029	2480	2250	1564	1307	1839	1027	2174
Madrid	1781	513	0	3420	537	1975	1784	1246	797	1330	546	1718
Moscow	1765	3892	3420	0	2883	1563	1643	2508	2666	2399	3119	1722
Toulouse	1278	1029	537	2883	0	1453	1249	883	312	911	468	1188
Budapest	950	2480	1975	1563	1453	0	573	1490	1180	836	1590	674
Dresden	378	2250	1784	1643	1249	573	0	988	1070	1044	1546	111
London	745	1564	1246	2508	883	1490	988	0	989	1444	1348	879
Marseille	1186	1307	797	2666	312	1180	1070	989	0	602	479	1032
Rom	1326	1839	1330	2399	911	836	1044	1444	602	0	839	1070
Palma D.M.	1658	1027	546	3119	468	1590	1546	1348	479	839	0	1511
Leipzig/Halle	289	2174	1718	1722	1188	674	111	879	1032	1070	1511	0

fast algorithms for the generation of closed sets (or their generators), which will facilitate the computation of the Hausdorff distances for large sets P. Starting from this promising modeling, the next step would then be to test, examine and develop further distortion measures in the realm of om-spaces.

6.2.2 Formal Context Mediated Metrics and Orders by Valuation

In a more general setting one might also want to study arbitrary binary relations, as often investigated in the form of a formal context (G, M, I). In this regard, the following question is directly obvious. How does the binary incidence relation $I \subseteq G \times M$, and therefore the inherent lattice structure, mediate a metric from one set to another? In detail, given the incidence I between sets G and M, where G is equipped with a metric d, one can induce a metric on M using the afore mentioned Hausdorff metric on (G, d) and the map $m \mapsto \{g \in G \mid (g, m) \in I\}$. What properties are expected of the resulting metric space (M, d_M)? What alterations of I, i.e., operations based on adding or removing relational pairs, do change the properties of (M, d_M)? Based on this one might also want to investigate the special case in which both sets G and M are equipped with an indigenous metric. Then, the natural question arises: how compatible are these indigenous metrics with the mutually induced ones mediated by the incidence I? For this, the computation of distance between metric spaces is essential, for which it might be necessary to employ the modified GH distance [147].

▶ **Example 5.** It is clear that in Example 4 only the object set G is equipped with a distance function (actually a metric) d_G. In order to obtain a distance function between the airlines, i.e. the attributes, the most obvious way is to calculate a Hamming distance between them. This distance, however, would not consider the existing data for G. Based on the proposition above, we can infer a distance function that incorporates d_G. For example, $d_M(\text{Scandinavian}, \text{Austrian})$ would be mapped to the Hausdorff distance $d_H(M_1, M_2)$ where

$M_1 = \{\text{Hamburg, Madrid, Moscow, Budapest, London, Rom, Palma D.M.}\}$

$M_2 = \{\text{Hamburg, Lisbon, Budapest, London, Rom, Palma D.M., Leipzig/Halle}\}$.

Hence, one has to compute

$$\max \left\{ \sup_{x \in M_1} d(x, M_2), \sup_{y \in M_2} d(M_1, y) \right\},$$

where $d(u, V) := \inf_{v \in V} d(u, v)$, i.e., the distance from u to V. In our example the computation of the Hausdorff distance results in $\max\{836, 1563\} = 1563$ nmi. ◣

Closely related to the setting above is the application of valuation functions to generate (linear) order relations. An example is to rank objects from G based on sizes of their shared related subsets from M, i.e., every $g \in G$ is valued by $|\{m \in M \mid (g, m) \in I\}|$. Since valuation functions almost always map into a numerical scale, e.g., the real line, this results in a metric and in turn into a linear order. Dual to the question in the first paragraph one can investigate how a valuation map distorts an indigenous ordinal property of the set, for example extracted through the notion of concept lattices. As a start, one may count how many wrongly ordered relation tuples are present. Furthermore, it might be fruitful to investigate to which extent the order created by the valuation differs from an arbitrarily generated linear extension [29]. Finally, one might re-evaluate our preliminary work in this area [161, 179, 180]. This can shed light on the question whether the valuation functions we transferred from orometry are superior when dealing with indigenous ordinal properties.

6.2.3 Intrinsic Dimension of om-Spaces and Randomness

There are numerous notions for measuring the complexity of a dataset. For ordinal data the order dimension is a natural candidate, as described in Section 4.1. Another approach is the notion for *intrinsic dimension*. Put simply, this value reflects the number of dimensions required to accurately represent the data. For incidence structures, such as formal contexts, we proposed in our previous work "Intrinsic Dimension of Geometric Data Sets" [95] a novel intrinsic dimension function that incorporates measure- and metric properties. This function is also applicable to ordinal data. This is in particular true, if the ordinal data is scaled via the *general ordinal scale*.

Yet, it is unknown if the intrinsic dimension does relate to the order dimension in any way. Still, it is to be expected that the intrinsic dimension can be used to capture ordinal complexity to some extent. To substantiate the hypothesis, an extensive experimental evaluation is needed and should be conducted, using real-world as well as randomly generated datasets. For the latter we might refer the reader to our work on random relational data generation [61]. Moreover, one can compare our results to recent notions of estimating dimensionality in the absence of distance functions [116]. Building up on this evaluation, and even more progressive, is the further development of an extension of the intrinsic dimension to om-spaces. This extension can open up a new research avenue for dataset dimensionality – however, it is dependent on highly ambitious theoretical and algorithmic results from all paragraphs above.

7 Algebraic constructions/decompositions for complexity reduction

For algebraic structures such as lattices, there exist many ways of reducing or decomposing them to smaller parts, for instance factorizations or sub-direct decompositions. However, these constructions are very sensitive to small perturbations in the data; and approximations and heuristics in the style of data mining are not (yet) considered in algebraic research.[13] The challenge is thus to *equip the algebraic constructions with approximations and heuristics* s.t. they become *suitable for data analysis.*

We propose to develop methods for clustering, aggregating and factorizing ordered sets and lattices. For ordered sets this may include order-based modifications of clustering algorithms such as k-modes, hierarchical agglomerative clustering and density-based clustering; and one might explore which signals for clustering can be obtained from their Dedekind-MacNeille completions.

---

[13] There are other approximations in universal algebra, but of a different nature, which is not relevant here: they deal with approximating infinite algebras [137] and serve, e.g., for defining the semantics of programming languages.

For lattices, we suggest exploiting constructions of universal algebra and lattice theory (such as factorizations, tolerance relations, atlas decompositions etc.) [42] [190]. These methods, however, have not been developed with real-world data with noise or disturbances in mind – up to now there are no notions such as "the lattice is decomposable to 95 %", and no algorithms dealing with them.

## 7.1   State of the Art and Open Questions

In universal algebra [20, 87, 30] it is a well-known fact that every homomorphic image and every subalgebra of a lattice or of a finite complete lattice and every direct product thereof is a lattice again [19].[14] This means that universal algebra provides us already with a rich toolkit for decomposing and aggregating lattices. In particular, every congruence relation of the lattice (i. e., an equivalence relation that is compatible with meet and join) provides a clustering of the lattice that additionally respects the lattice operations. The interpretation as clustering works much better for lattices than for other types of algebras (e. g., groups), as each congruence class is a convex set ("cluster") and is therefore not shattered throughout the lattice. The corresponding factor lattice can be understood as an aggregate of the original lattice. Several approaches exist that weaken the requirement of a congruence relation. Tolerance relations [217], for example, allow for overlapping clusters, while our newly introduced interval relations [118] produce smaller clusters (on the cost of not respecting all meets and joins any more). Another way of reducing the complexity of a lattice is by exploiting the fact that every lattice can be decomposed into smaller, irreducible factors.

Many of these algebraic operations were linked to datasets (in the form of a formal context) within the field of FCA [69]; in particular for subdirect decompositions [155, 156], for direct products of convex-ordinal scales [175], for substitution decomposition [133, 172] and substitution product [220, 173], as well as for tensorial decomposition [219] and tensor products [222].

▶ **Example 6.** Figure 4 shows the subdirect product of two conceptual scales for a conceptual information system about flight movements at Frankfurt Airport. With this composition of scales, one can analyze the distribution of the flight movements over different dimensions. In the figure, the outer scale indicates the location of the baggage conveyor assigned to an aircraft. For instance, we can see at the left of the diagram that 2503 times a baggage conveyor in Halle A was assigned to an aircraft. The inner scale shows the position of the aircraft. Of the 2503 aircraft mentioned before, 78 were positioned on the apron (V = Vorfeld) and 2425 at Terminal 1 (T1). Both assignments are reasonable because Halle A is part of Terminal 1. However, the system also helps us to discover apparent mismatches. For instance, we find that there are 180 aircraft that had a baggage conveyor assigned at Terminal 2 but which were positioned at Terminal 1 (T1). Focussing on these 180 flight movements with further scales would support us to discover reasons for this apparent mismatch.   ⌟

Complete congruence relations of concept lattices [157, 117] are a good starting point for clustering and aggregating. Complete tolerance relations and atlas decompositions [219] have been developed explicitly for allowing large lattices to be visualized in smaller parts just like several map pages in a road atlas.   Decompositions and factorizations are more straightforward the more structure the lattice shows. A particular strong structure is distributivity ($x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ and $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$) [19, 218, 42]. In this case, the lattice can be subdirectly decomposed into linear orders (which is beneficial for ordinal factor analysis as is discussed in Section 8).

---

[14] The same holds for all classes of algebras that are defined solely by equations, e. g., for groups and rings.
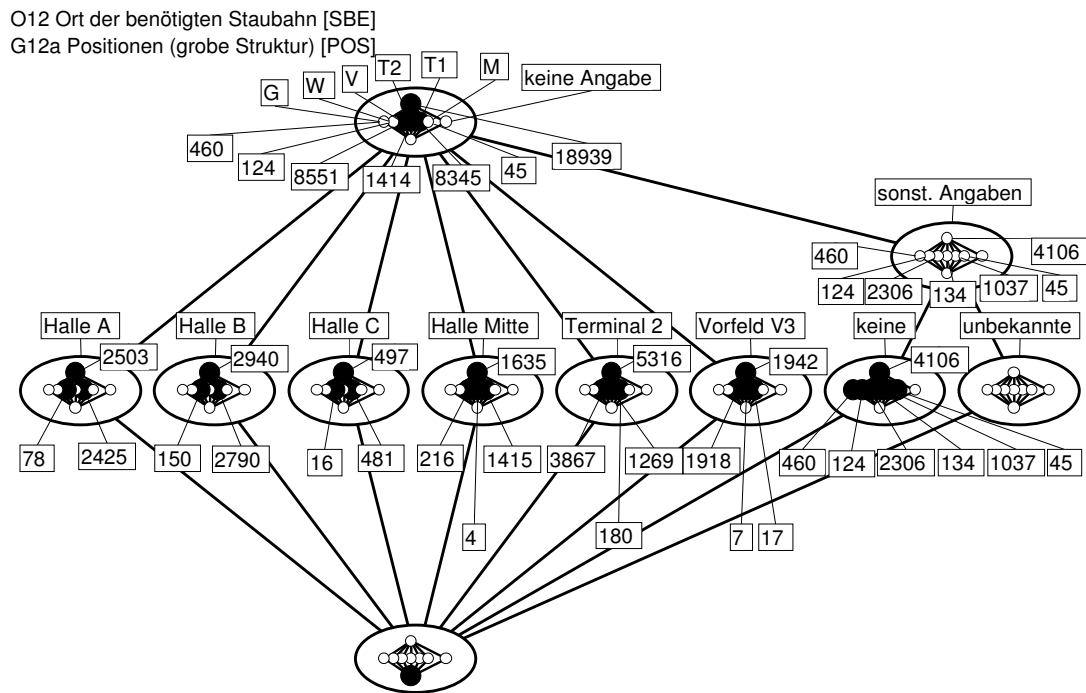
Figure 4 Subdirect product of the scales *position of baggage conveyor* and *position of aircraft* at Frankfurt Airport.

However, many of these constructions have never been used for larger data analysis tasks. This is due to the fact that lattices resulting from real-world data often do not precisely fulfil the structural conditions required for applying a decomposition or a factorization. In such a case, a lattice is considered as irreducible; there is no such notion as "95 % distributive" or "reducible with an error of x". Techniques for dealing with partly imperfect data (as they exist in numerical settings, as for instance soft margins for support vector machines) are not in the spirit of universal algebra, and have thus not been developed so far. The same holds for distributivity, which is rarely observed in the whole concept lattice of real-world data. Nevertheless, empirical evidence shows that many lattices contain "a large distributive part" – a still to-be-defined concept.

The main scientific challenge in this field is thus to *develop methods for applying algebraic constructions to lattices even if the necessary structural conditions are not satisfied.* This will include the definition of measures for the degree of satisfaction of these conditions, the provision of preprocessing methods for "cleaning" the data appropriately, the establishment of a theory for compositions, decompositions and aggregations with structural impurity, and the development of algorithms and heuristics for their computation on large datasets.

7.2 Promising Research Questions

We propose three specific research questions in this area. The first two consider clustering in ordered sets and lattices, resp., while the third one focuses on decompositions of lattices. Considering the required research approaches, it is the latter two questions that resemble more, as both can draw on the richer structure of lattices. In those two, the challenge is how to either "structurally clean up" the data before applying the algebraic operations, or how to turn the operations more "fault-tolerant".

Figure 5 Minard's map of Napoleon's campaign in Russia as inspiration for a possible density-based clustering algorithm of an ordered set.

7.2.1   Clustering in Ordered Sets

There are (at least) two straightforward approaches for clustering algorithms for ordered sets, the first being the adaptation of classical clustering techniques to ordered sets, and the second being the direct exploitation of order-theoretic constructs. For the adaption approach, we suggest to start with exploring modifications of standard clustering algorithms that rely on the comparability of pairs of objects, as opposed to numerical distances. These include for instance k-modes [105] and hierarchical agglomerative clustering [212] (where "hierarchical" does refer to the set inclusion hierarchy of the clusters). We also consider it extremely promising to develop ordinal versions of modularity-based clustering of networks [22], and – inspired by Ch. Minard's seminal map of Napoleon's campaign in Russia (Figure 5, [140]) – ordinal versions of density-based clustering [59].

In the second approach, one might investigate how to exploit order-theoretic concepts (such as common upper and lower bounds, order filters and ideals, and intervals) for new kinds of ordinal clustering. Particularly promising is to consider the new nodes in a Dedekind-MacNeille completion as cluster nuclei. This is due to the observation that, in the `ancestor_of` hierarchy of a human genealogy, a couple with two or more children always generates a new node in the Dedekind-MacNeille completion, which may be interpreted as a node representing the family.

7.2.2   Clustering and Factorizations in Lattices

If an ordered set is a lattice (including the case when we enforce this by the Dedekind-MacNeille completion), one can benefit from more structure. As stated above, in theory congruence relations and tolerance relations already provide clusterings that are even compatible with the lattice operations. However, they may not always be applicable. Again, there are (at least) two promising approaches to remediate this: The first is to assess the proportion of a lattice that has to be modified such that a congruence relation or a tolerance relation with pre-defined requirements exists; and derive methods to efficiently identify such modifications. The second is to weaken the notions of congruence relations and tolerance relations, so that they may respect only most of the operations. The questions here are again how to define them, how to compute (or approximate) them efficiently, and how to interpret the results. This may result in a new notion of "operator-

aware modularity". An expected problem case in modularity based graph clustering is a cluster that is a fully connected subgraph. In lattices, this corresponds to (convex) Boolean sublattices. This correspondence could give rise to a theory and algorithms for their compression in a cluster.

Another line of research one may draw from is employing other representation modes of the lattice: One could for instance study if there are meaningful concepts for clustering sets of implications, and how they translate back again to the clustering problem for lattices. In particular in the case of a large dataset one could also resort to probabilistic methods to derive the propositional logic and dependencies [181]. One might also employ results from preliminary work on probably approximately correct learning [24, 25] of those dependencies.

7.2.3 Decompositions of lattices

Concerning the complexity reduction of lattices by decompositions, one can follow the same research methodology as for clusterings and factorizations: There exists a large body of algebraic decomposition constructions for lattices, such as direct and sub-direct decompositions, or tensor decompositions [79], but these do not admit for minor structural disturbances. Hence, a natural task would be to adapt them to more robust versions, either by suitable preprocessing or by modifying the decomposition mechanism directly.
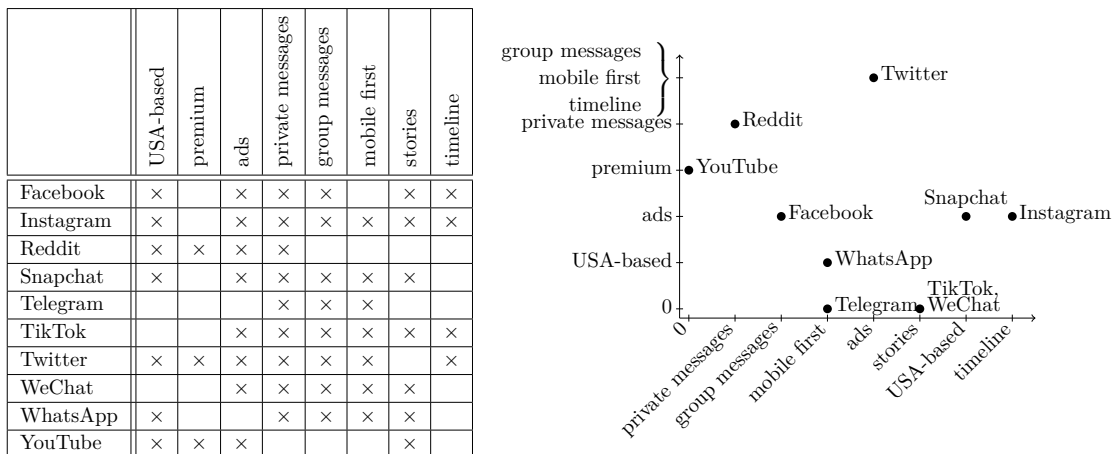
A promising observation with respect to preprocessing is that many real-world lattices are "to a large extent" distributive at their top [177], and that distributive lattices allow for particularly simple decompositions (e.g., subdirect decomposition into linear orders). Hence, the question for how to "repair" a lattice s.t. it becomes distributive is meaningful. In the case of "almost distributive" lattices, one could resort to the study of how (potentially non-distributive) ordered sets can be freely completed (in the sense of universal algebra) to lattices [190]. A potential means to identify the non-distributive part of a lattice might then be to establish a way to reverse this construction and develop efficient computation methods.

Additionally, one might want to explore – as described above – how other equivalent representations (in particular sets of implications) can be decomposed, and how this can be translated into meaningful decompositions of lattices.

8 Ordinal Factor Analysis

Factor analysis (in its traditional sense) is a method of multi-variate statistics to reduce data consisting of observations in different manifest variables to a lower-dimensional space spanned by "latent variables" with as little loss as possible. As the variables are considered to range in the real numbers, we will call this task numerical factor analysis in the sequel. A variety of approaches exists, the most prominent being principal component analysis. However, all these approaches require the data to be at least on interval scale level. Ordinal factor analysis, on the other hand, has been developed with the same intention as its numerical counterpart, but focussing on data of ordinal scale type.

Based on this foundation, we are looking forward to a *comprehensive theory and algorithms for an ordinal version of factor analysis*, that will keep track of all operations and thus avoid the problem of numerical factor analysis where distances may be distorted. For ordinal data the factors are known to be ordered sets. One might continue research in this direction by focussing on linear ordered sets as factors, as they provide an intuitive representation, in particular when their number is low. This constitutes an instance of dimension reduction and is computationally hard for ordinal data. Hence, one may have to resort to approximate methods for real-world data applications. In a second step, the theoretical foundation of ordinal factor analysis might then be adapted to non-linear factors. Our ambition is to establish ordinal factor analysis as a tool for ordinal data analysis, that will provide *a more intuitive representation of complex ordinal data.*

	USA-based	premium	ads	private messages	group messages	mobile first	stories	timeline
Facebook	×		×	×	×		×	×
Instagram	×		×	×	×	×	×	×
Reddit	×	×	×	×				
Snapchat	×		×	×	×	×	×	
Telegram				×	×	×		
TikTok			×	×	×	×	×	×
Twitter	×	×	×	×	×	×		×
WeChat			×	×	×	×	×	
WhatsApp	×			×	×	×	×	
YouTube	×	×	×					×

Figure 6 A dataset on social networks together with a visualization of a factorization into its two largest ordinal factors. All incidences can be deduced from the projection except for (TikTok, timeline), (WhatsApp, stories), (Facebook, timeline), (YouTube, stories), (Facebook, stories). The ordinal projection does not contain false data.

8.1 State of the Art and Open Questions

In multivariate statistics empirical observations are used to draw conclusions about underlying independent random variables. This was first developed by Spearman [170] in 1904 for evaluating intelligence tests, tracing them back to a single "general factor", and was generalized to multiple variables in 1919 by Garnett [80]. Building on this notion, nowadays two variants of factor analysis are employed. On one hand, *confirmatory factor analysis* [96] is used in social research to test whether the data fits to hypothesized measurement models. On the other hand, *explorative factor analysis* [149] is used to discover hidden structures in underlying data without the numbers of factors being known beforehand. This explorative factor analysis is closely related and often confused with principal component analysis as noted in [107], even though both techniques differ in execution as well as in their aim. In the area of recommender systems, factor analysis came to prominence through the "Netflix Prize" that was held between 2006 and 2009 where Funk [67] showed how to use factor analysis techniques to predict users preferences of movies. This approach was improved and refined in [123, 98]. A similar but not equal way of doing factor analysis is called *non-negative matrix factorization* and applied in cluster analysis of documents [171] or in astronomy [15]. One problem of these approaches is that the resulting factors are difficult (or even impossible) to interpret for a data analyst.

To avoid this problem, steps towards an explainable factor analysis have been made in the field of FCA. This goes back to Kerpt and Snásel [112, 113] and was further developed by Belohlavek and Vychodil [12, 13] under the name *Boolean factor analysis*. They compute a factorization of a Boolean matrix into two binary matrices, such that their Boolean matrix product results in the original matrix.

For such a factorization there is a corresponding factorizing family in the lattice, consisting of k formal concepts, called Boolean factors. Even though it is known to be $\mathcal{NP}$-complete to decide whether a Boolean matrix has a factorization into k Boolean factors, they introduce a set of algorithms for Boolean factor analysis which optimize k. Building on these notions, *ordinal factor analysis* was introduced by Ganter and Glodeanu in [71, 70], as a technique to meaningfully group Boolean factors, together with a visualization technique to depict ordinal factors in a biplot.

▶ **Example 7.** Figure 6 shows, on the left, a small dataset about prominent social networking platforms. On the right, a factorization of the dataset into its two largest ordinal data is shown. In the plot, an object contains all attributes that are listed below on the horizontal axis and all attributes that are listed left of it on the horizontal axis. The plot contains all information of the original dataset, except the four incidences listed in the caption of the figure. In order to cover these relations as well, the third factor would have to be included as well, leading to a 3D visualisation.   ⌟

The authors restrict their work to ordinal factors that form a chain in the concept lattice. This *linear factor analysis* is closely related to dimension theory of ordered sets [206], as in both cases a minimal number of chains is computed that covers an ordered set. In [82] the authors show that linear factor analysis can be applied to data sets from real-world applications and in [81] the theory of linear factor is transferred to triadic concept analysis. A modification of the visualization technique for the linear factors and the idea to extract two maximal linear factors that cover most of the data is proposed in [70]. A way that tries to cover most of a concept lattice with two chains and can thus be seen as a variant of linear factor analysis is described in our work [47].

While linear factor analysis seems to be a promising tool for a data analyst, it is still not yet fleshed out enough to be applied in practice. This is mainly due to the fact of being computationally expensive, and thus algorithms have not yet been developed (except for the Boolean case). Furthermore, the theoretical backbone of the theory is not yet broad enough to provide a data analyst with a robust toolkit of methods and variants.

8.2   Promising Research Questions

The overarching research question here is *how to develop ordinal factor analysis to a mature (visual) data analysis tool, that can be applied using little manual effort.* In particular this calls for enriching the theoretical backbone and the development of algorithms to make ordinal factor analysis applicable in practice. The algorithms should be able to extract structure that is already encapsulated in data and to deal with incremental changes in the data. Finally, a variant that also deals with non-linear factors (e. g., trees) might be of interest.

8.2.1   Computing Linear Factorizations in Two and More Dimensions

A first step in this direction is our greedy algorithm [52] that iteratively extracts the largest remaining ordinal factor from a formal context. Furthermore, our second work [53] builds on an idea proposed by Ganter [70] to cover a maximum subset of the dataset with a small set of linear factors. This proposed method allows for the incorporation of heuristics to approximate a factorization into two dimensions, and examines the computational boundaries.

We assume two-dimensional linear factor analysis to perform better on smaller datasets, as large datasets tend to comprise more complex dependencies. Thus, a second step may leap significantly beyond the state of the art by exploring how complex data can be represented with factor analysis in higher dimensions. As the two-dimensional case is already known to be of high computational complexity, we do expect this to be an even more complicated problem, which will definitely require the development of heuristics.

As the aim of these factorizations is to improve the understanding of correlations in data sets for a data analyst, novel visualization techniques are required. The practicability of resulting breakthrough approaches and algorithms of ordinal factor analysis in higher dimensions will have to be demonstrated. As it hardly seems possible to visualize factors far beyond dimension three on a static two-dimensional document, interactive navigation methods will have to be developed as well.

8.2.2 Background Orders and Evolving Data

In real-world settings, parts of the data may already be (linearly) ordered because of additional background information. Such data often appears in practice, as especially data containing numerical values (including time) are already ordered. An ordinal factor analysis should respect such *background orders* as it is unnatural for an analyst to have concepts such as time divided and distributed over different factors. One could follow (at least) two possible routes to avoid this problem. The first one is to disregard the ordered part of the data from for the computation of an ordinal factorization. Then in a second step this order would be added back as an additional factor. While this guarantees the purity of this additional factor, it might not depict relationships to a sufficient degree. Thus, one might want to follow a more ambitious path by developing algorithms which do not split such an ordinal structure into different factors but are completely contained in a single factor. This factor is however still allowed to contain additional information and is contrary to the first approach not restricted to the background order.

Another typical requirement in practice is to update the factorization when data is evolving. Computing an ordinal factorization of a large dataset will not be (computationally) cheap because of the high computational complexity. Thus, if a small detail changes in the data it should be possible to integrate it in an existing ordinal factorization without having to repeat the whole computational effort. Furthermore, the structure of the factorization should be stable to a sufficient degree, as a human data analyst looking a second time at slightly modified data should be presented with familiar ordinal factors. This proposes two research questions, the first being if one already has to follow additional restrictions when computing the first factorization to allow a later change in the data. The second is how one can insert the changes into an existing factorization.

8.2.3 Non-Linear Factor Analysis

As often complex data is not correlated linearly, further research on non-linear factorizations seems to be promising. The definition of ordinal factors stemming from a linear order of Boolean factors allows the definition to be extended to factorization into arbitrarily ordered sets. However, these non-linear factors have to be further restricted as otherwise every lattice is its own, trivial factorization. We propose to investigate several restrictions for non-ordinal factors that seem reasonable, such as planarity, two-dimensionality, distributivity, and trees. Another possible approach to gain a reasonable factorization is to fix the number of factors beforehand and to allow arbitrary non-ordinal factors that should approximately be of similar cardinality. For all those restrictions no preliminary work is available, making this endeavor highly challenging because of anticipated potential complexity as well as visualization problems.

9 Visualizing, Exploring and Explaining Ordinal Data

Real-world ordinal data are usually too large to be analyzed by simply browsing the set along the edges of the order relation. The typical means of presenting ordered sets to humans is via line diagrams. A surprisingly hard conceptual problem that we will discuss in this section is the specification of HCI-founded, formal optimization criteria for graph drawing, for which one can then develop *efficient layout algorithms* [203, 47]. A second conceptually demanding problem – still addressed manually today – is the automatic break-down of large ordered sets into smaller, visualizable parts together with suitable means for their *interactive exploration* [182]. We will motivate the development of new interaction paradigms based on the decomposition and factorization methods discussed in Sections 7 and 8. These are particularly challenging as they also have to transmit the information about the effects caused by approximations.

9.1 State of the Art and Open Questions

Graph Drawing has a long research history, represented a. o. by the series of Graph Drawing Symposia[15] since 1992. Partial orders and lattices are specific types of graphs, and their drawings can benefit from cycle freeness. Surprisingly, though, there exist only a few criteria for formalizing readability (e. g., maximizing distances between nodes and lines, maximizing angles of crossing lines, minimizing the number of different edge directions, organizing nodes in layers). There is agreement in the community that further criteria are needed, and that they have to be empirically evaluated from the HCI perspective.

In order theory, a common tool for visualizing and investigating ordered set is the *order diagram*, that is sometimes also called *line diagram* or *Hasse diagram*. The order diagram of an ordered set $(X, \leq)$ is a directed graph, where the edge $(a, b) \in X \times X$ exists if and only if $a < b$ and there is no $c \in X$ such that $a < c < b$. In a drawing of this graph the direction of an edge is not signaled by an arrow but by the y-coordinates, with the dot marking the position of b being above the dot marking the position of a. The relative horizontal positions are thus used as visual variable [16]. While readable order diagrams can be drawn from an experienced expert by hand, this is not viable in practice, as it is a time-consuming task. The automatic generation of well interpretable graph drawings is a *surprisingly hard task*. The problem starts with the fact that the readability criteria described above are partly conflicting and that their relative importance varies in different settings. In the standard work about graph drawing [203, Sect. 3.1] several divide-and-conquer algorithms for drawing trees, which are closely related to order diagrams, are described.

In the case of a planar graph (which can be checked in linear time [103]), a drawing only consisting of straight lines without bends or curves always exists [148, Sect. 4.2 & 4.3] and should be preferred. For an order diagram with a unique maximum and minimum it can be checked whether it is planar and in this case an upward planar drawing can be computed in linear time [203, Sect. 6]. However, while lattices by definition contain a maximum and minimum element, they are usually not planar if they are derived from real world data [4]. A drawing algorithm [203, Sect. 3.2] for "serial parallel graphs", a special family of planar, acyclic graphs straight-line drawings can be equipped with a modification to produce symmetries based on the automorphism group [102]. Sugiyama et al. [197] published in 1981 an algorithm framework to compute upwards layered drawings of acyclic graphs. Force directed algorithms were introduced in [54] and further refined in [66]. In [65, 51] this idea is applied to order diagrams. Our work [194] proposes a geometrical heuristic to support a human to draw order diagrams manually. By trying to emphasize the intrinsic structure of ordered sets our work [47] proposes a drawing algorithm based on the computation of two linear extensions. The algorithm is included as a tool in the software-suite `conexp-clj`[16] under the name `DimDraw` [49].

▶ **Example 8.** The example in Figure 7 shows that it is possible to achieve competitive order diagrams using this approach. The concept lattice was created for structuring an educational film about living beings and water [200, 79]. The left diagram was hand-drawn by an FCA expert. The diagram in the middle was generated with the algorithm of Sugiyama et al. [197] while the right one was drawn with the `DimDraw`-algorithm [47] which leverages the dimensional structure of the order. Experiments have shown that the DimDraw results are often rather close to hand-drawn diagrams. ↵

---

[15] https://dblp.uni-trier.de/db/conf/gd/
[16] https://github.com/tomhanika/conexp-clj
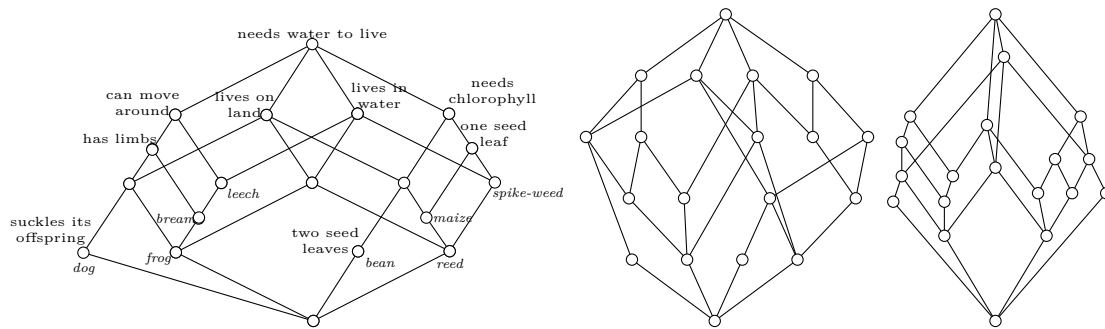
Figure 7 Three order diagrams for the visualization of the same ordered set: hand-drawn by an expert, by the Sugiyama algorithm and by the DimDraw algorithm.

As ordered sets and lattices even of modest size (from, say, 30-50 elements upwards) are hard to visualize such that they are still beneficial for a data analyst, further means of interaction are necessary for larger ordered sets. One way is to compute a diagram in 3D, which can be done using [65] or a modification of [47], and then allow to rotate it on the screen, but this does not increase the size of presentable nodes significantly. Another way is to display only a single node together with its upper and lower neighbors at a time, and to shift the focus when the user selects one of these neighbors. This has for instance been used by Carpineto and Romano for Information Retrieval [33]. A third approach follows the *divide & conquer* paradigm. Its first realization for concept lattices decomposed the lattice in two or more factors of a sub-direct semi-product, and combined their smaller diagrams in a *nested line diagram* [221, 188, 182]. This is also the archetype for the kind of algebraic constructions that we suggested studying in Section 7. This approach became part of a model for *Conceptual Information Systems* [8] and has been implemented in the *ToscanaJ tool suite*[17] [9, 10, 40]. ToscanaJ also allows, by means of conceptual scaling, to visualize concept lattices derived from data types of any kind, and to navigate intuitively through the data by switching between the conceptual scales and zooming into concepts. However, the initialization of the system requires conceptual and manual work, in particular a meaningful selection of the factors and a manual layout of the order diagram for each factor. The model of Conceptual Information Systems has been generalized in various ways: as ordinal version of *Online-Analytical Processing* [191, 183, 56, 100, 193], for accessing databases and conceptual graphs [125] and knowledge bases [64]. A good survey over navigation paradigms and layout algorithms for concept lattices with hybrid data is provided by [57].

9.2 Promising Research Questions

The main objective of the visualization of ordinal structures is to *support human analysts in analyzing and exploring large ordinal data*. In particular, we emphasize on the urgent need of new layout algorithms *that allow for a completely automatic, well interpretable visualization of medium-sized ordered sets and lattices*. These might include *new interaction paradigms based on the decomposition and factorization methods* as discussed in Sections 7 and 8. This approach is particularly challenging as it also has to *transmit the information about which aspects of the data are not represented in the visualization*, due to an approximation.

These (static) visualizations may pave the way for (dynamic) interactive browsing and exploring of complete datasets. Again, the opportunities for new browsing paradigms may benefit from advances in the field of algebraic decompositions and ordinal factor analysis as discussed in Sections 7 and 8.

---

[17] http://toscanaj.sourceforge.net/

9.2.1 Visualization of Ordered Sets and Lattices

Ordinal factor analysis, as discussed in Section 8, will provide a set of chains that cover all points represented in a dataset. Using ideas from [47], this could be applied for a novel drawing algorithm, which, however, would not reduce the – typically too large – order dimension. One could hence exploit means for order dimension reduction (towards dimension 2) as discussed in Section 7, or employ force directed drawing algorithms in three or more dimensions. The challenge is to incorporate the order constraints such that they are all reflected properly in the resulting diagram. Another potential approach might be machine-learning-based dimension reduction. This is challenging since the requirement that the y-coordinates of data points have to obey the order relation is not encoded in non-linear methods such as t-SNE [134], in contrast to linear methods such as PCA [227]. This is also true for deep learning procedures, which nevertheless might be worth being adapted to ordinal data.

For some special families of lattices (such as distributive lattices, two-dimensional lattices or Boolean lattices), methods exist for generating readable drawings. Thus, in a third line of research one could explore modifications (removing or adding vertices and/or edges) such that the lattice "partly fulfills" these properties – which first needs to be defined. In a second step, the drawing of this structurally simpler part could then be used to derive a readable drawing for the whole lattice. A supporting study of this endeavor would be to empirically investigate how large the two-dimensional, distributive or Boolean parts of real-world lattices have to be to justify the applicability of such an algorithm.

9.2.2 Browsing and Exploring Large Orders

There may be many ways for future approaches to browsing and exploring ordered sets. We assume that some of them will be based on the outcomes of research along the lines described in Sections 7 and 8. From today's perspective we anticipate these approaches as follows. The *slice & dice* paradigm of On-Line Analytical Processing (OLAP) provides a good scheme for analyzing multidimensional numerical data, and first applications to ordinal data exist [191, 183, 56, 100, 193]. A similar approach might be followed for the clusterings, decompositions and factorizations developed as described in Sections 7 and 8. To date it is not obvious, though, whether the same visualization and interaction paradigms can also be applied to the new constructions. What is definitely missing in the preliminary work is the handling of disturbances. In a second step, we might therefore address the visualization of and interaction with algebraic clusterings, decompositions and factorizations *that have been approximated*. As there does not exist any preliminary work on this, this part of the research will be of non-incremental nature. Our assumption is that the analyst will be informed about the existence of any non-covered and eventually wrongly assigned data, at least on request. Following the paradigm of *Explainable AI*, the analyst should be enabled to drill down to the original data if necessary, and to obtain explanations in case the construction could not fully respect the data.

10 Challenges and Long-Term Perspective

While a wide range of methods for Data Analysis and Knowledge Discovery has been developed for data that are on ratio level (i. e., can be adequately modeled with real numbers), there are fewer methods for data on ordinal level. In this paper, we call for a joint activity to bring these methods together and to extend them with new ones, and to establish thus the new research field of Ordinal Data Analysis. The amount of ordinal data in the real world is large, and so is the requirement for adequate methods and algorithms. There are many research opportunities in this

field, and we can provide only some teasers in this paper. We are looking forward to all kinds of contributions to the new field of Ordinal Data Science, both within and outside the five subfields discussed in the previous sections.

We see Ordinal Data Science in a long tradition of Conceptual Knowledge Processing, where we aim at a highly human-centered process of asking, exploring, analyzing, interpreting, and learning about the data in interaction with the underlying database [101, 184]. In this line, we follow R. S. Brachman et al. [26], who introduced the notion of Data Archaeology for knowledge discovery tasks in which a precise specification of the discovery strategy, the crucial questions, and the basic goals of the task have to be elaborated during such an unpredictable interactive exploration of the data.

We expect that all future contributions to this new field will share a range of challenges that are specific for research on the intersection of computer science and mathematics, which both come with different philosophical foundations, research aims and working paradigms. One challenge will be to develop sound theoretical theories, another one to find computationally feasible solutions and/or approximations. Last but not least the algorithms and methods have to be made accessible to their users, e. g., by suitable visualizations, interaction paradigms and human-computer interfaces.

We will conclude this paper by illustrating, for the five fields that we discussed in the previous sections, the challenges and long-term perspectives that we expect to face:

We expect to meet the largest theoretical challenges in the endeavor of developing an ordinal version of the Representational Theory of Measurement. The replacement of the numerical relational structure in RTM by ordinal relational structures (and in the long run by other, non-ordinal relational structures such as symmetry groups) implies the loss of focus on additivity. A generalization of RTM thus requires a much more flexible framework that allows for a large variety of (numerical and) non-numerical scales, and it is not obvious a priori how such a framework might look like. A major criticism to RTM is that in practice the pre-conditions for applicability of measures on interval or ratio scale are hard to verify, and that nominal measurements (whose pre-conditions are easier to verify) do not provide enough insight. When we succeed, we expect a new ordinal theory of measurement to fill the gap between these extremes and to become a significant contribution to the still intense discussion about the applicability of RTM for psychological studies ([139, 204, 205], a good survey is [201]).

Generalized ordinal scales are just a special case of allowing arbitrary algebraic structures as ranges of measurements. In the long run we have the ambition to establish a *"Grand Unified Theory of Measurement"*, which is likely to be of high interest to (the philosophical foundations of) physics (where traditionally the adaptation of measurement to psychological phenomena has been eyed with some suspicion), as it might allow for describing the study of, e. g., the symmetries of elementary particles in the same way as their numerical properties.

In Section 6, we proposed to develop a joint theory for the interplay of metric functions and ordinal relations, based on the archetype of metric measure spaces. While, in the latter, both metrics and measures operate on the set of real numbers, our proposed approach tries to marry two structurally different domains. We might thus face the task of developing a new paradigm of interaction which has no precedents that can be followed.

In Sections 7 to 9, we argued for the exploitation of algebraic structures for data analysis tasks, which unfortunately go along with large size and computationally hard problems. E.g., the size of concept lattices grows in the worst case exponentially in the size of the data, which in the Information Age are already large by themselves; and determining the order dimension of an

ordered set is $\mathcal{NP}$-hard. In Data Science, the typical approach is to use projections, decompositions, approximations, and heuristics. However, these kinds of approximations and heuristics are no subject of study in classical universal algebra, and still have to be established. In particular, we suggest to make use of algebraic composition and decomposition techniques for a divide & conquer approach for browsing and exploring the data. The transformation of the (approximated) algebraic constructions to human–computer interaction paradigms that support users without knowledge of universal algebra and order theory is challenging, as there is little precedence. Nonetheless, if this works out, it might provide a whole new family of analysis techniques for ordinal data.

References

1 Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. *Advances in Mathematics*, 228(6):3026–3126, 2011. `doi:10.1016/j.aim.2011.08.003`.

2 Ravi P Agarwal, MA El-Gebeily, and Donal O'Regan. Generalized contractions in partially ordered metric spaces. *Applicable Analysis*, 87(1):109–116, 2008. `doi:10.1080/00036810701556151`.

3 Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. Mining association rules between sets of items in large databases. *SIGMOD Rec.*, 22(2):207–216, jun 1993. `doi:10.1145/170035.170072`.

4 Alexandre Albano and Bogdan Chornomaz. Why concept lattices are large: extremal theory for generators, concepts, and vc-dimension. *International Journal of General Systems*, 46(5):440–457, 2017. `doi:10.1080/03081079.2017.1354798`.

5 Simon Andrews. Making use of empty intersections to improve the performance of cbo-type algorithms. In *International Conference on Formal Concept Analysis*, pages 56–71. Springer, 2017. `doi:10.1007/978-3-319-59271-8_4`.

6 Franz Baader. Computing a minimal representation of the subsumption lattice of all conjunctions of concepts defined in a terminology. In *Proceedings of the International Symposium on Knowledge Retrieval, Use, and Storage for Efficiency, KRUSE 95*, pages 168–178, Santa Cruz, USA, 1995.

7 Franz Baader, Bernhard Ganter, Baris Sertkaya, and Ulrike Sattler. Completing description logic knowledge bases using formal concept analysis. In Manuela M. Veloso, editor, *IJCAI*, pages 230–235, 2007. `doi:10.25368/2022.155`.

8 K. Becker, G. Stumme, R. Wille, U. Wille, and M. Zickwolff. Conceptual information systems discussed through an it-security tool. In R. Dieng and O. Corby, editors, *Knowledge Engineering and Knowledge Management. Methods, Models, and Tools.*, volume 1937 of *LNAI*, pages 352–365, Heidelberg, 2000. Springer. `doi:10.1007/3-540-39967-4_27`.

9 P. Becker, J. Hereth, and G. Stumme. ToscanaJ: An open source tool for qualitative data analysis. In V. Duquenne, B. Ganter, M. Liquiere, E. M. Nguifo, and G. Stumme, editors, *Advances in Formal Concept Analysis for Knowledge Discovery in Databases.*, pages 1–2, Lyon, France, 7 23 2002.

10 Peter Becker and Joachim Hereth Correia. The toscanaj suite for implementing conceptual information systems. In *Formal Concept Analysis*, pages 324–348. Springer, 2005. `doi:10.1007/11528784_17`.

11 Radim Belohlavek. *Fuzzy relational systems: foundations and principles*, volume 20. Springer Science & Business Media, 2012. `doi:10.1007/978-1-4615-0633-1`.

12 Radim Belohlávek and Vilém Vychodil. Formal concepts as optimal factors in boolean factor analysis: Implications and experiments. In Peter W. Eklund, Jean Diatta, and Michel Liquiere, editors, *Proceedings of the Fifth International Conference on Concept Lattices and Their Applications, CLA 2007, Montpellier, France, October 24–26, 2007*, volume 331 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2007. URL: `https://ceur-ws.org/Vol-331/Belohlavek1.pdf`.

13 Radim Belohlavek and Vilem Vychodil. Discovery of optimal factors in binary data via a novel method of matrix decomposition. *Journal of Computer and System Sciences*, 76(1):3–20, 2010. `doi:10.1016/J.JCSS.2009.05.002`.

14 R. Bendaoud, A. Napoli, and Y. Toussaint. Formal concept analysis: A unified framework for building and refining ontologies. In *EKAW*, 2008. `doi:10.1007/978-3-540-87696-0_16`.

15 Olivier Berne, C. Joblin, Y. Deville, J. D. Smith, M. Rapacioli, J. P. Bernard, J. Thomas, W. Reach, and A. Abergel. Analysis of the emission of very small dust particles from spitzer spectro-imagery data using blind signal separation methods. *Astronomy & Astrophysics*, 469(2):575–586, 2007. `doi:10.1051/0004-6361:20066282`.

16 J. Bertin and M. Barbut. *Sémiologie graphique: les diagrammes, les réseaux, les cartes*. Gauthier Villars, 1967.

17 T Gnana Bhaskar and V Lakshmikantham. Fixed point theorems in partially ordered metric spaces and applications. *Nonlinear Analysis: Theory, Methods & Applications*, 65(7):1379–1393, 2006. `doi:10.1016/j.na.2005.10.017`.

18 Robi Bhattacharjee and Sanjoy Dasgupta. What relations are reliably embeddable in euclidean space? In Aryeh Kontorovich and Gergely Neu, editors, *Proceedings of the 31st International Conference on Algorithmic Learning Theory*, volume 117 of *Proceedings of Machine Learning Research*, pages 174–195, San Diego, California, USA, 08 February–11 February 2020. PMLR. URL: `http://proceedings.mlr.press/v117/bhattacharjee20a.html`.

19 G. Birkhoff. On the structure of abstract algebras. *Proceedings of the Cambridge Philosoph-*

ical Society, 31(4):433–454, 1935. `doi:10.1017/s0305004100013463`.

20 Garrett Birkhoff. Universal algebra. In *Comptes Rendus du Premier Congrès Canadien de Mathématiques*, pages 310–326, Toronto, 1946. University of Toronto Press.

21 Garrett Birkhoff. *Lattice theory. Rev. ed*, volume 25 of *Colloq. Publ., Am. Math. Soc.* American Mathematical Society (AMS), 1948.

22 Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008:10008, 2008. `doi:10.1088/1742-5468/2008/10/p10008`.

23 George Boole. *Investigation of The Laws of Thought On Which Are Founded the Mathematical Theories of Logic and Probabilities*. Walton and Maberly, 1853. Also available from Dover, New York 1958, ISBN 0-486-60028-9.

24 Daniel Borchmann, Tom Hanika, and Sergei Obiedkov. On the usability of probably approximately correct implication bases. In Karell Bertet, Daniel Borchmann, Peggy Cellier, and Sébastien Ferré, editors, *ICFCA*, volume 10308 of *LNCS*, pages 72–88. Springer, 2017. `doi:10.1007/978-3-319-59271-8_5`.

25 Daniel Borchmann, Tom Hanika, and Sergei Obiedkov. Probably approximately correct learning of horn envelopes from queries. *Discrete Applied Mathematics*, 273:30–42, 2020. Advances in Formal Concept Analysis: Traces of CLA 2016. `doi:10.1016/J.DAM.2019.02.036`.

26 Ronald J. Brachman, Peter G. Selfridge, Loren G. Terveen, Boris Altman, Alex Borgida, Fern Halper, Thomas Kirk, Alan Lazar, Deborah L. McGuinness, and Lori Alperin Resnick. Integrated support for data archaeology. In *Proceedings of the 2nd International Conference on Knowledge Discovery in Databases*, AAAIWS'93, pages 197–211. AAAI Press, 1993. `doi:10.1142/S0218215793000083`.

27 L. Breiman, J. Friedman, C.J. Stone, and R.A. Olshen. *Classification and Regression Trees*. The Wadsworth and Brooks-Cole statistics-probability series. Taylor & Francis, 1984.

28 Laura F. Bringmann and Markus I. Eronen. Heating up the measurement debate: What psychologists can learn from the history of physics. *Theory & Psychology*, 26(1):27–43, 2016. `doi:10.1177/0959354315617253`.

29 Russ Bubley and Martin Dyer. Faster random generation of linear extensions. *Discrete mathematics*, 201(1-3):81–88, 1999. `doi:10.1016/S0012-365X(98)00333-1`.

30 S. Burris and H. P. Sankappanavar. *A Course in Universal Algebra*. Springer, New York, 1981.

31 Xiao Cai, Feiping Nie, and Heng Huang. Multiview k-means clustering on big data. In *Twenty-Third International Joint conference on artificial intelligence*. Citeseer, 2013. URL: `http://ijcai.org/Abstract/13/383`.

32 Jaime S Cardoso, Joaquim F Pinto da Costa, and Maria J Cardoso. Modelling ordinal relations with svms: An application to objective aesthetic evaluation of breast cancer conservative treatment. *Neural Networks*, 18(5-6):808–817, 2005. `doi:10.1016/J.NEUNET.2005.06.023`.

33 Claudio Carpineto and Giovanni Romano. *Concept Data Analysis: Theory and Applications*. John Wiley & Sons, 2004.

34 N. Caspard, B. Leclerc, and B. Monjardet. *Finite Ordered Sets: Concepts, Results and Uses*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2012. `doi:10.1017/cbo9781139005135`.

35 Nicholas R. Chrisman. Rethinking levels of measurement for cartography. *Cartography and Geographic Information Systems*, 25(4):231–242, 1998. `doi:10.1559/152304098782383043`.

36 Constanze Ciavarella, Laura Fumanelli, Stefano Merler, Ciro Cattuto, and Marco Ajelli. School closure policies at municipality level for mitigating influenza spread: a model-based evaluation. *BMC Infectious Diseases*, 16(1):576, 2016. `doi:10.1186/s12879-016-1918-z`.

37 Philipp Cimiano, Andreas Hotho, Gerd Stumme, and Julien Tane. Conceptual knowledge processing with formal concept analysis and ontologies. In Peter Eklund, editor, *Concept Lattices*, volume 2961 of *LNAI*, pages 189–207, Heidelberg, 2004. Second International Conference on Formal Concept Analysis, ICFCA 2004, Springer. `doi:10.1007/978-3-540-24651-0_18`.

38 C. H. Coombs, L. C. Coombs, and J. C. Lingoes. Stochastic cumulative scales. In S. Shye, editor, *Theory construction and data analysis in the behavioral sciences*. Jossey-Bass., San Francisco, 1978.

39 Clyde H. Coombs. *A Theory of Data*. Wiley, New York, 1964.

40 Frithjof Dau and Baris Sertkaya. An extension of ToscanaJ for FCA-based data analysis over triple stores. In *CUBIST Workshop*, 2011. URL: `https://ceur-ws.org/Vol-753/paper02_cubistws11.pdf`.

41 Thomas H. Davenport and Jeanne G. Harris. *Competing on Analytics: The New Science of Winning*. Harvard Business School Press, 2007. `doi:10.5860/choice.44-6322`.

42 B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge mathematical textbooks. Cambridge University Press, 2002. `doi:10.1017/CBO9780511809088`.

43 Stephan Doerfel, Tom Hanika, and Gerd Stumme. Clones in graphs. In Michelangelo Ceci, Nathalie Japkowicz, Jiming Liu, George A. Papadopoulos, and Zbigniew W. Ras, editors, *ISMIS*, volume 11177 of *LNCS*, pages 56–66. Springer, 2018. `doi:10.1007/978-3-030-01851-1_6`.

44 Stephan Doerfel, Robert Jäschke, and Gerd Stumme. Publication analysis of the formal concept analysis community. In F. Domenach, D.I. Ignatov, and J. Poelmans, editors, *ICFCA 2012*, volume 7278 of *Lecture Notes in Artificial Intelligence*, pages 77–95, Berlin/Heidelberg, 2012. Springer. `doi:10.1007/978-3-642-29892-9_12`.

45 Jean-Paul Doignon and Jean-Claude Falmagne. *Knowledge Spaces*. Imprint: Springer, Berlin, Heidelberg, 1999. `doi:10.1007/978-3-642-58625-5`.

46 Michel Dumontier and Tobias Kuhn. Data science – methods, infrastructure, and applications. *Data Science*, 1:1–5, 2017. `doi:10.3233/DS-170013`.

47 Dominik Dürrschnabel, Tom Hanika, and Gerd Stumme. Drawing order diagrams through two-dimension extension, 2019. `arXiv:1906.06208`.

48 Dominik Dürrschnabel, Tom Hanika, and Maximilian Stubbemann. FCA2VEC: Embedding techniques for formal concept analysis. In Rokia Missaoui, Léonard Kwuida, and Talel Abdessalem, editors, *Complex Data Analytics with Formal Concept Analysis*, pages 47–74. Springer International Publishing, 2022. `doi:10.1007/978-3-030-93278-7_3`.

49 Dominik Dürrschnabel, Tom Hanika, and Gerd Stumme. Dimdraw – a novel tool for drawing concept lattices. In Diana Cristea, Florence Le Ber, Rokia Missaoui, Léonard Kwuida, and Baris Sertkaya, editors, *ICFCA (Supplements)*, volume 2378 of *CEUR Workshop Proceedings*, pages 60–64. CEUR-WS.org, 2019. URL: `https://ceur-ws.org/Vol-2378/shortAT6.pdf`.

50 Dominik Dürrschnabel, Maren Koyda, and Gerd Stumme. Attribute selection using contranominal scales. In Tanya Braun, Marcel Gehrke, Tom Hanika, and Nathalie Hernandez, editors, *Graph-Based Representation and Reasoning - 26th International Conference on Conceptual Structures, ICCS 2021, Virtual Event, September 20-22, 2021, Proceedings*, volume 12879 of *Lecture Notes in Computer Science*, pages 127–141. Springer, 2021. `doi:10.1007/978-3-030-86982-3_10`.

51 Dominik Dürrschnabel and Gerd Stumme. Force-directed layout of order diagrams using dimensional reduction. In Agnès Braud, Aleksey Buzmakov, Tom Hanika, and Florence Le Ber, editors, *Formal Concept Analysis - 16th International Conference, ICFCA 2021, Strasbourg, France, June 29 - July 2, 2021, Proceedings*, volume 12733 of *Lecture Notes in Computer Science*, pages 224–240. Springer, 2021. `doi:10.1007/978-3-030-77867-5_14`.

52 Dominik Dürrschnabel and Gerd Stumme. Greedy discovery of ordinal factors. *CoRR*, abs/2302.11554, 2023. `doi:10.48550/ARXIV.2302.11554`.

53 Dominik Dürrschnabel and Gerd Stumme. Maximal ordinal two-factorizations. In *Proceedings of the 27th International Conference on Conceptual Structures, ICCS 2022*, 2023. `doi:10.1007/978-3-031-40960-8_5`.

54 Peter Eades. A heuristic for graph drawing. *Congressus numerantium*, 42:149–160, 1984.

55 David A Edwards. The structure of superspace. In *Studies in topology*, pages 121–133. Elsevier, 1975. `doi:10.1016/b978-0-12-663450-1.50017-7`.

56 P. Eklund, B. Groh, G. Stumme, and R. Wille. Contextual-logic extension of toscana. In B. Ganter and G. W. Mineau, editors, *Conceptual Structures: Logical, Linguistic, and Computational*, volume 1867 of *LNAI*, pages 453–467, Heidelberg, 2000. Springer. `doi:10.1007/10722280_31`.

57 Peter W. Eklund and Jean Villerd. A survey of hybrid representations of concept lattices in conceptual knowledge processing. In Léonard Kwuida and Baris Sertkaya, editors, *ICFCA*, volume 5986 of *LNCS*, pages 296–311. Springer, 2010. `doi:10.1007/978-3-642-11928-6_21`.

58 Yves Escoufier, Bernard Fichet, Edwin Diday, Ludovic Lebart, Chikio Hayashi, Noboru Ohsumi, and Yasumasa Baba, editors. *Data Science and its Applications*. Acadamic Press/Harcourt Brace, Tokyo, 1995.

59 Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In Evangelos Simoudis, Jiawei Han, and Usama M. Fayyad, editors, *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, Oregon, USA*, pages 226–231. AAAI Press, 1996. URL: `http://www.aaai.org/Library/KDD/1996/kdd96-037.php`.

60 Ernesto Estrada. The communicability distance in graphs. *Linear Algebra and its Applications*, 436(11):4317–4328, 2012. `doi:10.1016/j.laa.2012.01.017`.

61 Maximilian Felde and Tom Hanika. Formal context generation using dirichlet distributions. In Dominik Endres, Mehwish Alam, and Diana Sotropa, editors, *ICCS*, volume 11530 of *LNCS*, pages 57–71. Springer, 2019. `doi:10.1007/978-3-030-23182-8_5`.

62 Maximilian Felde, Tom Hanika, and Gerd Stumme. Null models for formal contexts. *J. Information*, 11(3, 135), 2020. `doi:10.3390/INFO11030135`.

63 A. Ferguson, C. S. Myers, R. J. Bartlett, H. Banister, F. C. Bartlett, W. Brown, N. R. Campbell, K. J. W. Craik, J. Drever, J. Guild, R. A. Houstoun, J. O. Irwin, G. W. C. Kaye, S. J. F. Philpott, L. F. Richardson, J. H. Shaxby, T. Smith, R. H. Thouless, and W. S. Tucker. Quantitative estimates of sensory events. *Advancement of Science*, 2:331–349, 1940. (Final report of a committee appointed by the British Association for the Advancement of Science in 1932 to consider the possibility of measuring intensities of human sensation.). `doi:10.1038/130334a0`.

64 Sébastien Ferré, Marianne Huchard, Mehdi Kaytoue, Sergei O Kuznetsov, and Amedeo Napoli. Formal concept analysis: from knowledge discovery to knowledge processing. *A Guided Tour of Artificial Intelligence Research: Volume II: AI Algorithms*, pages 411–445, 2020. `doi:10.1007/978-3-030-06167-8_13`.

65 Ralph Freese. Automated lattice drawing. In *International Conference on Formal Concept Analysis*, pages 112–127. Springer, 2004. `doi:10.1007/978-3-540-24651-0_12`.

66 Thomas MJ Fruchterman and Edward M Reingold. Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164, 1991. `doi:10.1002/SPE.4380211102`.

67 Simon Funk. Netflix update: Try this at home, 2006. [Online; accessed 11-August-2020].

68 B. Ganter, G. Stumme, and R. Wille, editors. *Formal Concept Analysis: Foundations and Applications*, volume 3626 of *LNAI*, Heidelberg, 2005. Springer. `doi:10.1007/978-3-540-31881-1`.

69 Bernhard Ganter. Composition and decomposition in formal concept analysis. In Hans-Herrmann

Bock, editor, *Classification and related methods of data analysis*, pages 561–566, Amsterdam, 1988. North–Holland.

70 Bernhard Ganter. *Diskrete mathematik: Geordnete mengen.* Springer, 2013. `doi:10.1007/978-3-642-37500-2`.

71 Bernhard Ganter and Cynthia Vera Glodeanu. Ordinal factor analysis. In *International Conference on Formal Concept Analysis*, pages 128–139. Springer, 2012. `doi:10.1007/978-3-642-29892-9_15`.

72 Bernhard Ganter and Cynthia Vera Glodeanu. Factors and skills. In Cynthia Vera Glodeanu, Mehdi Kaytoue, and Christian Sacarea, editors, *Formal Concept Analysis - 12th International Conference, ICFCA 2014, Cluj-Napoca, Romania, June 10-13, 2014. Proceedings*, volume 8478 of *Lecture Notes in Computer Science*, pages 173–187. Springer, 2014. `doi:10.1007/978-3-319-07248-7_13`.

73 Bernhard Ganter and Sergei O. Kuznetsov. Stepwise construction of the dedekind-macneille completion. In Marie-Laure Mugnier and Michel Chein, editors, *Conceptual Structures: Theory, Tools and Applications*, pages 295–302, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg. `doi:10.1007/BFB0054922`.

74 Bernhard Ganter and Sergei A. Obiedkov. *Conceptual Exploration.* Springer, 2016. `doi:10.1007/978-3-662-49291-8`.

75 Bernhard Ganter, Jürgen Stahl, and Rudolf Wille. Conceptual measurement and many–valued contexts. In W. Gaul and Martin Schader, editors, *Classification as a tool of research*, pages 169–176, Amsterdam, 1986. North–Holland.

76 Bernhard Ganter and Gerd Stumme. Creation and merging of ontology top-levels. In Aldo de Moor, Wilfried Lex, and Bernhard Ganter, editors, *Conceptual Structures for Knowledge Creation and Communication.*, volume 2746 of *LNAI*, pages 131–145, Heidelberg, 2003. Springer. `doi:10.1007/978-3-540-45091-7_9`.

77 Bernhard Ganter and Rudolf Wille. Conceptual scaling. In Frank Roberts, editor, *Applications of combinatorics and graph theory to the biological and social sciences*, pages 139–167. Springer–Verlag, New York, 1989. `doi:10.1007/978-1-4684-6381-1_6`.

78 Bernhard Ganter and Rudolf Wille. *Formale Begriffsanalyse: Mathematische Grundlagen.* Springer, Heidelberg, 1996. `doi:10.1007/978-3-642-61450-7`.

79 Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis: Mathematical Foundations.* Springer, Heidelberg, 1999. `doi:10.1007/978-3-642-59830-2`.

80 J. C. Maxwell Garnett. General ability, cleverness and purpose. *British Journal of Psychology*, 9(3):345, 1919.

81 Cynthia Vera Glodeanu. Tri-ordinal factor analysis. In *International Conference on Formal Concept Analysis*, pages 125–140. Springer, 2013. `doi:10.1007/978-3-642-38317-5_8`.

82 Cynthia Vera Glodeanu and Bernhard Ganter. Applications of ordinal factor analysis. In *International Conference on Formal Concept Analy-*

sis, pages 109–124. Springer, 2013. `doi:10.1007/978-3-642-38317-5_7`.

83 Leo A. Goodman and William H. Kruskal. *Measures of Association for Cross Classifications.* Springer Verlag, New York, 1979. `doi:10.1007/978-1-4612-9995-0`.

84 S. Goto, H. Bono, H. Ogata, W. Fujibuchi, T. Nishioka, K. Sato, and M. Kanehisa. Organizing and computing metabolic pathway data in terms of binary relations. *Pac Symp Biocomput*, pages 175–186, 1997.

85 Olivier Gouvert, Thomas Oberlin, and Cédric Févotte. Ordinal non-negative matrix factorization for recommendation. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, pages 3680–3689. PMLR, 2020. URL: `http://proceedings.mlr.press/v119/gouvert20a.html`.

86 Mikhael Gromov. Groups of polynomial growth and expanding maps. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 53(1):53–78, 1981. `doi:10.1007/bf02698687`.

87 George Grätzer. *Universal Algebra.* Springer, 1968. `doi:10.1007/978-0-387-77487-9`.

88 George Grätzer. *General Lattice Theory.* Birkhäuser Verlag, 2. edition, 1998. `doi:10.1016/s0079-8169(08)x6085-4`.

89 L. Guttman. A basis for scaling qualitative data. *American Sociological Review*, 9:139–150, 1944. `doi:10.2307/2086306`.

90 Paul R. Halmos. *Measure theory*, volume 18. Springer, 2013. `doi:10.1007/978-1-4684-9440-2`.

91 Tom Hanika and Johannes Hirth. Conexp-clj – a research tool for fca. In Diana Cristea, Florence Le Ber, Rokia Missaoui, Léonard Kwuida, and Baris Sertkaya, editors, *ICFCA (Supplements)*, volume 2378 of *CEUR Workshop Proceedings*, pages 70–75. CEUR-WS.org, 2019. URL: `https://ceur-ws.org/Vol-2378/shortAT8.pdf`.

92 Tom Hanika and Johannes Hirth. Quantifying the conceptual error in dimensionality reduction. In Tanya Braun, Marcel Gehrke, Tom Hanika, and Nathalie Hernandez, editors, *Graph-Based Representation and Reasoning - 26th International Conference on Conceptual Structures, ICCS 2021, Virtual Event, September 20-22, 2021, Proceedings*, volume 12879 of *Lecture Notes in Computer Science*, pages 105–118. Springer, 2021. `doi:10.1007/978-3-030-86982-3_8`.

93 Tom Hanika and Johannes Hirth. Knowledge cores in large formal contexts. *Ann. Math. Artif. Intell.*, 90(6):537–567, 2022. `doi:10.1007/S10472-022-09790-6`.

94 Tom Hanika, Maren Koyda, and Gerd Stumme. Relevant attributes in formal contexts. In Dominik Endres, Mehwish Alam, and Diana Sotropa, editors, *ICCS*, volume 11530 of *LNCS*, pages 102–116. Springer, 2019. `doi:10.1007/978-3-030-23182-8_8`.

95 Tom Hanika, Friedrich Martin Schneider, and Gerd Stumme. Intrinsic dimension of geometric data sets. *Tohoku Mathematical Journal*, 74(1):23–52, 2022. `doi:10.2748/tmj.20201015a`.

96 D. Harrington. *Confirmatory Factor Analysis*. Oxford scholarship online: Social Work module. Oxford University Press, USA, 2009.

97 Egbert Harzheim. *Ordered Sets*. Springer, 2005. `doi:10.1007/b104891`.

98 Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*, pages 173–182, 2017. `doi:10.1145/3038912.3052569`.

99 Rajneesh Hegde and Kamal Jain. The hardness of approximating poset dimension. *Electronic Notes in Discrete Mathematics*, 29:435–443, 2007. European Conference on Combinatorics, Graph Theory and Applications. `doi:10.1016/J.ENDM.2007.07.084`.

100 J. Hereth and G. Stumme. Reverse pivoting in conceptual information systems. In H. Delugach and G. Stumme, editors, *Conceptual Structures: Broadening the Base.*, volume 2120 of *LNAI*, pages 202–215, Heidelberg, 2001. Springer. `doi:10.1007/3-540-44583-8_15`.

101 Joachim Hereth, Gerd Stumme, Rudolf Wille, and Uta Wille. Conceptual knowledge discovery – a human-centered approach. *Journal of Applied Artificial Intelligence (AAI)*, 17(3):281–301, 2003. `doi:10.1080/713827122`.

102 Seok-Hee Hong, Peter Eades, and Sang-Ho Lee. Drawing series parallel digraphs symmetrically. *Computational Geometry*, 17(3-4):165–188, 2000. `doi:10.1016/S0925-7721(00)00020-1`.

103 John Hopcroft and Robert Tarjan. Efficient planarity testing. *Journal of the ACM (JACM)*, 21(4):549–568, 1974. `doi:10.1145/321850.321852`.

104 Andreas Hotho, Steffen Staab, and Gerd Stumme. Explaining text clustering results using semantic structures. In Nada Lavrač, Dragan Gamberger, and Hendrik BlockeelLjupco Todorovski, editors, *Knowledge Discovery in Databases: PKDD 2003, 7th European Conference on Principles and Practice of Knowledge Discovery in Databases*, volume 2838 of *LNAI*, pages 217–228, Heidelberg, 2003. Springer. `doi:10.1007/978-3-540-39804-2_21`.

105 Zhexue Huang. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Min. Knowl. Discov.*, 2(3):283–304, 1998. `doi:10.1023/A:1009769707641`.

106 Eyke Hüllermeier, Johannes Fürnkranz, Weiwei Cheng, and Klaus Brinker. Label ranking by learning pairwise preferences. *Artificial Intelligence*, 172(16-17):1897–1916, 2008. `doi:10.1016/J.ARTINT.2008.08.002`.

107 IT Joliffe and BJT Morgan. Principal component analysis and exploratory factor analysis. *Statistical methods in medical research*, 1(1):69–95, 1992. `doi:10.1177/096228029200100105`.

108 Toshihiro Kamishima and Jun Fujiki. Clustering orders. In Gunter Grieser, Yuzuru Tanaka, and Akihiro Yamamoto, editors, *Discovery Science*, pages 194–207, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg. `doi:10.1007/978-3-540-39644-4_17`.

109 K Keller and E Petrov. Ordinal spaces. *Acta Mathematica Hungarica*, 160(1):119–152, 2020. `doi:10.1007/s10474-019-00972-z`.

110 Matthias Keller. Intrinsic metrics on graphs: A survey. In Delio Mugnolo, editor, *Mathematical Technology of Networks*, pages 81–119, Cham, 2015. Springer International Publishing. `doi:10.1007/978-3-319-16619-3_7`.

111 Maurice G. Kendall. *Rank correlation methods*. C. Griffin, 1948.

112 Ales Keprt and Václav Snásel. Binary factor analysis with help of formal concepts. In *CLA*, volume 110, pages 90–101, 2004. URL: `https://ceur-ws.org/Vol-110/paper10.pdf`.

113 Aleš Keprt. *Algorithms for Binary Factor Analysis*. PhD thesis, PhD thesis, 2006.

114 S.C. Kleene. Representation of events in nerve nets and finite automata. *Annals of Mathematics Studies*, 34:3–41, 1956. `doi:10.1515/9781400882618-002`.

115 Matthäus Kleindessner and Ulrike Luxburg. Uniqueness of ordinal embedding. In Maria Florina Balcan, Vitaly Feldman, and Csaba Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 40–67, Barcelona, Spain, 13–15 June 2014. PMLR. URL: `http://proceedings.mlr.press/v35/kleindessner14.html`.

116 Matthäus Kleindessner and Ulrike Luxburg. Dimensionality estimation without distances. In *Artificial Intelligence and Statistics*, pages 471–479, 2015. URL: `http://proceedings.mlr.press/v38/kleindessner15.html`.

117 Sigrid Knecht and Rudolf Wille. Congruence lattices of finite lattices as concept lattices. In J. Almeida, G. H. Bordalo, and P. Dwinger, editors, *Lattices, semigroups, and universal algebra*, pages 323–325. Plenum Press, New York–London, 1990. `doi:10.1007/978-1-4899-2608-1_32`.

118 Maren Koyda and Gerd Stumme. Factorizing lattices by interval relations. *International Journal of Approximate Reasoning*, 157:70–87, 2023. `doi:10.1016/J.IJAR.2023.03.003`.

119 Patryk Kozieł and Małgorzata Sulkowska. Uniform random posets. *Information Sciences*, 515:294–301, 2020. `doi:10.1016/J.INS.2019.12.018`.

120 Petr Krajca, Jan Outrata, and Vilem Vychodil. Parallel algorithm for computing fixpoints of galois connections. *Annals of Mathematics and Artificial Intelligence*, 59(2):257–272, 2010. `doi:10.1007/S10472-010-9199-5`.

121 David H. Krantz, R. Duncan Luce, Patrick Suppes, and Amos Tversky. *Foundations of Measurement – Additive and Polynomial Representations*, volume 1. Academic Press, 1971. `doi:10.1016/C2013-0-11004-5`.

122 Francesco Kriegel and Daniel Borchmann. Nextclosures: parallel computation of the canonical base with background knowledge. *International Journal of General Systems*, 46(5):490–510, 2017. `doi:10.1080/03081079.2017.1349570`.

123 Rajeev Kumar, BK Verma, and Shyam Sunder Rastogi. Social popularity based svd++ recommender system. *International Journal of Computer Applications*, 87(14), 2014. `doi:10.5120/15279-4033`.

124 Léonard Kwuida and Stefan E. Schmidt. Valuations and closure operators on finite lattices.

Discrete Applied Mathematics, 159(10):990–1001, 2011. `doi:10.1016/J.DAM.2010.11.022`.

125 Jens Kötters and Peter W. Eklund. Conjunctive query pattern structures: A relational database model for formal concept analysis. *Discrete Applied Mathematics*, 273:144–171, 2020. Advances in Formal Concept Analysis: Traces of CLA 2016. `doi:10.1016/J.DAM.2019.08.019`.

126 Abdul Latif, Antonio-Francisco Roldán-López-de Hierro, and Wutiphol Sintunaravat. On common α-fuzzy fixed points with applications. *Fixed Point Theory Appl.*, 2014:22, 2014. Id/No 234. `doi:10.1186/1687-1812-2014-234`.

127 Deutsche Fußball Liga. Spielordnung, 2019. URL: `https://media.dfl.de/sites/2/2019/09/Spielordnung-SpOL-2019-08-22-Stand.pdf`.

128 Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995. `doi:10.1007/BF01200757`.

129 Tie-Yan Liu. Learning to rank for information retrieval. *Foundations and Trends® in Information Retrieval*, 3(3):225–331, 2009. `doi:10.1561/1500000016`.

130 F. M. Lord. On the statistical treatment of football numbers. *American Psychologist*, 8(12):750–751, 1953. `doi:10.1037/h0063675`.

131 R. Duncan Luce, David H. Krantz, Patrick Suppes, and Amos Tversky. *Foundations of Measurement – Representation, Axiomatization, and Invariance*, volume 3. Academic Press, 1990. `doi:10.1016/C2009-0-21635-7`.

132 R. Duncan Luce and John W. Tukey. Simultaneous conjoint measurement: a new type of fundamental measurement. *Journal of Mathematical Psychology*, page 27, 1964. `doi:10.1016/0022-2496(64)90015-x`.

133 Peter Luksch and Rudolf Wille. Substitution decomposition of concept lattices. In G. Eigenthaler, H. K. Kaiser, W. B. Müller, and W. Nöbauer, editors, *Contributions to general algebra*, volume 5, pages 213–220, Wien, 1987. Hölder–Pichler–Tempsky.

134 Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605, 2008. URL: `http://jmlr.org/papers/v9/vandermaaten08a.html`.

135 H. M. MacNeille. Partially ordered sets. *Trans. Am. Math. Soc.*, 42:416–460, 1937. `doi:10.1090/s0002-9947-1937-1501929-x`.

136 Hiroshi Maehara. Euclidean embeddings of finite metric spaces. *Discrete Mathematics*, 313(23):2848–2856, 2013. `doi:10.1016/J.DISC.2013.08.029`.

137 K. Meinke and J. V. Tucker. *Universal Algebra*, pages 189–368. Oxford University Press, Inc., USA, 1993. URL: `https://dl.acm.org/doi/10.5555/162573.162534`.

138 Facundo Memoli. On the use of gromov-hausdorff distances for shape comparison. In M. Botsch, R. Pajarola, B. Chen, and M. Zwicker, editors, *Eurographics Symposium on Point-Based Graphics*. The Eurographics Association, 2007. `doi:10.2312/SPBG/SPBG07/081-090`.

139 Joel Michell. Quantitative science and the definition of measurement in psychology. *British Journal of Psychology*, 88(3):355–383, 1997. `doi:10.1111/j.2044-8295.1997.tb02641.x`.

140 Charles Minard. Carte figurative des pertes successives en hommes de l'armée française dans la campagne de russie 1812–1813, 1869. URL: `https://commons.wikimedia.org/wiki/File:Minard.png`.

141 Ann Mische. Relational sociology, culture, and agency. *The SAGE handbook of social network analysis*, pages 80–97, 2011. `doi:10.4135/9781446294413.n7`.

142 Rokia Missaoui and Kevin Emamirad. Lattice miner-a formal concept analysis tool. In *14th International Conference on Formal Concept Analysis, Supplementary Proceedings*, page 91, 2017.

143 B. Monjardet. Metrics on partially ordered sets—a survey. *Discrete Mathematics*, 35(1):173–184, 1981. Special Volume on Ordered Sets. `doi:10.1016/0012-365X(81)90206-5`.

144 Frederick Mosteller and John W Tukey. Data analysis and regression. a second course in statistics. *Addison-Wesley series in behavioral science: quantitative methods*, 1977.

145 Kevin Musgrave, Serge Belongie, and Ser-Nam Lim. A metric learning reality check, 2020. `doi:10.1007/978-3-030-58595-2_41`.

146 Facundo Mémoli. Metric structures on datasets: stability and classification of algorithms. In *International Conference on Computer Analysis of Images and Patterns*, pages 1–33. Springer, 2011. `doi:10.1007/978-3-642-23678-5_1`.

147 Facundo Mémoli. Some properties of gromov–hausdorff distances. *Discrete & Computational Geometry*, 48(2):416–440, 2012. `doi:10.1007/S00454-012-9406-8`.

148 Takao Nishizeki and Md Saidur Rahman. *Planar graph drawing*, volume 12. World Scientific Publishing Company, 2004. `doi:10.1142/5648`.

149 Megan Norris and Luc Lecavalier. Evaluating the use of exploratory factor analysis in developmental disability psychological research. *Journal of autism and developmental disorders*, 40(1):8–20, 2010. `doi:10.1007/s10803-009-0816-2`.

150 Charles S. Peirce. *Collected Papers*. Harvard Universit Press, Cambridge, 1931–1935. `doi:10.5860/choice.37-1485`.

151 J. Pfanzagl. *Theory of Measurement*. Physica, Heidelberg, 1971. `doi:10.1007/978-3-662-41488-0`.

152 Gregory Piatetsky-Shapiro. Knowledge discovery in real databases: A report on the ijcai-89 workshop. *AI Magazine*, 11(5):68–70, 1991. `doi:10.1609/AIMAG.V11I4.873`.

153 Jonas Poelmans, Sergei O. Kuznetsov, Dmitry I. Ignatov, and Guido Dedene. Formal concept analysis in knowledge processing: A survey on models and techniques. *Expert Syst. Appl.*, 40(16):6601–6623, 2013. `doi:10.1016/J.ESWA.2013.05.007`.

154 Silke Pollandt and Rudolf Wille. Functorial scaling of ordinal data. *Discret. Appl. Math.*, 147(1):101–111, 2005. `doi:10.1016/J.DAM.2004.06.024`.

155 Klaus Reuter. A note on the number of irreducibles of subdirect products. *Algebra Universalis*, 22:304–305, 1986. `doi:10.1007/bf01224035`.

156 Klaus Reuter. Matchings of linearly indecomposable modular lattices. *Discrete Mathematics*, 63:245–247, 1987. `doi:10.1016/0012-365X(87)90013-6`.

157 Klaus Reuter and Rudolf Wille. Complete congruence relations of concept lattices. *Acta Sci. Math.*, 51:319–327, 1987. URL: `http://acta.bibl.u-szeged.hu/id/eprint/15144`.

158 Fred S. Roberts. *Measurement Theory*. Cambridge University Press, 1984. `doi:10.1017/CBO9780511759871`.

159 Ryan A Rossi, Luke K McDowell, David William Aha, and Jennifer Neville. Transforming graph data for statistical relational learning. *Journal of Artificial Intelligence Research*, 45:363–441, 2012. `doi:10.1613/JAIR.3659`.

160 Camille Roth and Paul Bourgine. Lattice-based dynamic and overlapping taxonomies: The case of epistemic communities. *Scientometrics*, 69(2):429–447, 2006. `doi:10.1007/S11192-006-0161-6`.

161 Andreas Schmidt and Gerd Stumme. Prominence and dominance in networks. In Catherine Faron Zucker, Chiara Ghidini, Amedeo Napoli, and Toussaint Yannick, editors, *Proceedings of the 21th International Conference on Knowledge Engineering and Knowledge Management (EKAW)*, LNCS, pages 370–385. Springer, 2018. `doi:10.1007/978-3-030-03667-6_24`.

162 Annemarie Zand Scholten and Denny Borsboom. A reanalysis of lord's statistical treatment of football numbers. *Journal of Mathematical Psychology*, 53(2):69–75, 2009. `doi:10.1016/j.jmp.2009.01.002`.

163 Bernd Schröder. *Ordered sets. An introduction with connections from combinatorics to topology*. Basel: Birkhäuser/Springer, 2nd edition edition, 2016. `doi:10.1007/978-3-319-29788-0`.

164 Ernst Schröder. *Algebra der Logik I, II, III. 1890, 1891, 1895*. Thoemmes Press, Bristol, 2001.

165 Yuan Shi, Wenzhe Li, and Fei Sha. Metric learning for ordinal data. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, pages 2030–2036, 2016. `doi:10.1609/AAAI.V30I1.10280`.

166 Dan A. Simovici and Chabane Djeraba. *Mathematical Tools for Data Mining - Set Theory, Partial Orders, Combinatorics. Second Edition*. Advanced Information and Knowledge Processing. Springer, 2014. `doi:10.1007/978-1-4471-6407-4`.

167 Gregor Snelting and Frank Tip. Understanding class hierarchies using concept analysis. *ACM Trans. Program. Lang. Syst.*, 22(3):540–582, 2000. `doi:10.1145/353926.353940`.

168 John Sowa. *Conceptual Structures: Information Processing in Mind and Machine*. The Systems Programming Series. Addison-Wesley, 1984.

169 N Spangenberg and KE Wolff. Conceptual grid evaluation, classification and related methods of data analysis, 1988.

170 C. Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904. `doi:10.2307/1422689`.

171 Suvrit Sra and Inderjit S. Dhillon. Generalized nonnegative matrix approximations with bregman divergences. In *Advances in neural information processing systems*, pages 283–290, 2006. URL: `https://proceedings.neurips.cc/paper/2005/hash/d58e2f077670f4de9cd7963c857f2534-Abstract.html`.

172 Jörg Stephan. *Substitution decomposition of lattices*. PhD thesis, TH Darmstadt, 1991.

173 Jörg Stephan. Substitution products of lattices. In D. Dorninger, G. Eigenthaler, H. K. Kaiser, and W. B. Müller, editors, *Contributions to general algebra*, volume 7, pages 321–336, Wien, 1991. Hölder–Pichler–Tempsky.

174 S. S. Stevens. On the theory of scales of measurement. *Science*, 103(2684):677–680, 1946. `doi:10.1126/science.103.2684.677`.

175 Selma Strahringer. Direct products of convex–ordinal scales. *Order*, 11:361–383, 1994. `doi:10.1007/bf01108768`.

176 Selma Strahringer and Rudolf Wille. Towards a structure theory for ordinal data. In M. Schader, editor, *Analysing and modelling data and knowledge*, pages 129–139, Heidelberg, 1992. Springer–Verlag. `doi:10.1007/978-3-642-46757-8_14`.

177 Selma Strahringer, Rudolf Wille, and Uta Wille. Mathematical support for empirical theory building. In Harry S. Delugach and Gerd Stumme, editors, *Conceptual Structures: Broadening the Base, 9th International Conference on Conceptual Structures, ICCS 2001, Stanford, CA, USA, July 30-August 3, 2001, Proceedings*, volume 2120 of *LNCS*, pages 169–186. Springer, 2001. `doi:10.1007/3-540-44583-8_13`.

178 Maximilian Stubbemann, Tom Hanika, and Friedrich Martin Schneider. Intrinsic dimension for large-scale geometric learning. *Transactions on Machine Learning Research*, 2023. URL: `https://openreview.net/forum?id=85BfDdYMBY`.

179 Maximilian Stubbemann, Tom Hanika, and Gerd Stumme. Orometric methods in bounded metric data. In Michael R. Berthold, Ad Feelders, and Georg Krempl, editors, *Advances in Intelligent Data Analysis XVIII – 18th International Symposium on Intelligent Data Analysis, IDA 2020, Konstanz, Germany, April 27-29, 2020, Proceedings*, volume 12080 of *LNCS*, pages 496–508. Springer, 2020. `doi:10.1007/978-3-030-44584-3_39`.

180 Maximilian Stubbemann and Gerd Stumme. The mont blanc of twitter: Identifying hierarchies of outstanding peaks in social networks. In *Proc. 34th European Conference on Machine Learning / 27th Intl. Conference on Prinicples of Knowledge Discovery in Databases*, 2023. `doi:10.1007/978-3-031-43418-1_11`.

181 Milan Studený. *Probabilistic Conditional Independence Structures*. Springer Publishing Company, Incorporated, 1st edition, 2010. `doi:10.1007/b138557`.

182 G. Stumme. Hierarchies of conceptual scales. In T. B. Gaines, R. Kremer, and M. Musen, editors, *Proc. Workshop on Knowledge Acquisition, Modeling and Management (KAW'99)*, volume 2, pages 78–95. Banff, October 16-22 1999.

183 G. Stumme. Conceptual on-line analytical processing. In K. Tanaka, S. Ghandeharizadeh, and Y. Kambayashi, editors, *Information Organization and Databases*, chapter 14, pages 191–203. Kluwer, Boston-Dordrecht-London, 2000. `doi:10.1007/978-1-4615-1379-7_14`.

184 G. Stumme. Off to new shores – conceptual knowledge discovery and processing. *Intl. J. Human-*

Comuter Studies (IJHCS), 59(3):287–325, sep 2003. `doi:10.1016/S1071-5819(03)00044-2`.

185 G. Stumme and A. Maedche. Fca-merge: Bottom-up merging of ontologies. In B. Nebel, editor, *Proc. 17th Intl. Conf. on Artificial Intelligence (IJCAI '01)*, pages 225–230, Seattle, WA, USA, 2001. URL: `https://www.ijcai.org/Proceedings/01/IJCAI-2001-b.pdf`.

186 Gerd Stumme. Knowledge acquisition by distributive concept exploration. In G. Ellis, R. Levinson, W. Rich, and J. F. Sowa, editors, *Conceptual structures: applications, implementation and theory*, number 954 in Lecture Notes in Artificial Intelligence, Berlin–Heidelberg–New York, 1995. Springer–Verlag. (supplementary proceedings). `doi:10.1007/3-540-60161-9`.

187 Gerd Stumme. Attribute exploration with background implications and exceptions. In H.-H. Bock and W. Polasek, editors, *Data analysis and information systems*, pages 457–469, Berlin–Heidelberg–New York, 1996. Springer–Verlag. `doi:10.1007/978-3-642-80098-6_39`.

188 Gerd Stumme. Local scaling in conceptual data systems. In P. W. Eklund, G. Ellis, and G. Mann, editors, *Conceptual Structures: Knowledge Representation as Interlingua. Proc. ICCS'96*, volume 1115 of *LNAI*, pages 308–320, Heidelberg, 1996. Springer. `doi:10.1007/3-540-61534-2_20`.

189 Gerd Stumme. Concept exploration — a tool for creating and exploring > conceptual hierarchies. In D. Lukose, H. Delugach, M. Keeler, L. Searle, and J. F. Sowa, editors, *Conceptual structures: Fulfilling Peirce's dream*, volume 1257 of *Lecture Notes in Artificial Intelligence*, Berlin–Heidelberg–New York, 1997. Springer–Verlag. `doi:10.1007/BFB0027880`.

190 Gerd Stumme. Free distributive completions of partial complete lattices. *In: Order*, 14:179–189, 1998. `doi:10.1023/A:1006002507294`.

191 Gerd Stumme. On-line analytical processing with conceptual information systems. In K. Tanaka and S. Ghandeharizadeh, editors, *Proc. 5th Intl. Conf. on Foundations of Data Organization (FODO'98)*, pages 117–126, November 12-13 1998.

192 Gerd Stumme. Acquiring expert knowledge for the design of conceptual information systems. In D. Fensel and R. Studer, editors, *Knowledge Acquisition, Modeling, and Management. Proc. 11th European*, volume 1621 of *LNAI*, pages 275–290, Heidelberg, 1999. Springer. `doi:10.1007/3-540-48775-1_17`.

193 Gerd Stumme. A finite state model for on-line analytical processing in triadic contexts. In Bernhard Ganter and Robert Godin, editors, *Proc. 3rd Intl. Conf. on Formal Concept Analysis*, volume 3403 of *LNCS*, pages 315–328, Heidelberg, 2005. Springer. `doi:10.1007/978-3-540-32262-7_22`.

194 Gerd Stumme and Rudolf Wille. A geometrical heuristic for drawing concept lattices. In R. Tamassia and I.G. Tollis, editors, *Graph Drawing*, volume 894 of *LNCS*, pages 452–459, Heidelberg, 1995. Springer. `doi:10.1007/3-540-58950-3_399`.

195 Gerd Stumme and Rudolf Wille, editors. *Begriffliche Wissensverarbeitung – Methoden und Anwendungen*, Heidelberg, 2000. Springer. `doi:10.1007/978-3-642-57217-3`.

196 Karl-Theodor Sturm et al. On the geometry of metric measure spaces. *Acta mathematica*, 196(1):65–131, 2006. `doi:10.1007/s11511-006-0002-8`.

197 Kozo Sugiyama, Shojiro Tagawa, and Mitsuhiko Toda. Methods for visual understanding of hierarchical system structures. *IEEE Transactions on Systems, Man, and Cybernetics*, 11(2):109–125, 1981. `doi:10.1109/TSMC.1981.4308636`.

198 Patrick Suppes, David H. Krantz, R. Duncan Luce, and Amos Tversky. *Foundations of Measurement – Geometrical, Threshold, and Probabilistic Representations*, volume 2. Academic Press, 1989. `doi:10.1016/C2009-0-21665-5`.

199 Atsushi Suzuki, Atsushi Nitanda, Jing Wang, Linchuan Xu, Kenji Yamanishi, and Marc Cavazza. Generalization error bound for hyperbolic ordinal embedding. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 10011–10021. PMLR, 18–24 July 2021. URL: `http://proceedings.mlr.press/v139/suzuki21a.html`.

200 Viola Takàcs. Two applications of Galois graphs in pedagogical research. Manuscript of a lecture given at TH Darmstadt, 1984.

201 Eran Tal. Measurement in science. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2017 edition, 2017. URL: `https://plato.stanford.edu/Entries/measurement-science/`.

202 M. Thomas. Mathematization, not measurement: A critique of stevens' scales of measurement. *Journal of Methods and Measurement in the Social Sciences*, 10(2):76–94, 2020. `doi:10.2458/v10i2.23785`.

203 Ioannis Tollis, Giuseppe Di Battista, Peter Eades, and Roberto Tamassia. *Graph drawing. Algorithms for the visualization of graphs.* Upper Saddle River, NJ: Prentice Hall, 1999.

204 Günter Trendler. Measurement theory, psychology and the revolution that cannot happen. *Theory & Psychology*, 19:579–599, oct 2009. `doi:10.1177/0959354309341926`.

205 Günter Trendler. Conjoint measurement undone. *Theory & Psychology*, 29(1):100–128, 2019. `doi:10.1177/0959354318788729`.

206 William T. Trotter. *Combinatorics and Partially Ordered Sets: Dimension Theory.* Johns Hopkins University Press, 2001.

207 William T. Trotter and Bartosz Walczak. Boolean dimension and local dimension. *Electron. Notes Discret. Math.*, 61:1047–1053, 2017. `doi:10.1016/J.ENDM.2017.07.071`.

208 John W. Tukey. The future of data analysis. *Annals of Mathematical Statistics*, 33:1–67, 1962. `doi:10.1214/aoms/1177704711`.

209 Alexey A. Tuzhilin. Who invented the gromov-hausdorff distance?, 2016. cite arxiv:1612.00728Comment: 7 pages, 10 references. `doi:10.48550/arXiv.1612.00728`.

210 Leena Chennuru Vankadara and Ulrike von Luxburg. Measures of distortion for machine learning. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *Advances in Neural Information Processing*

Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, 3-8 December 2018, Montréal, Canada, pages 4891–4900, 2018. URL: `https://proceedings.neurips.cc/paper/2018/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html`.

211 Fei Wang and Jimeng Sun. Survey on distance metric learning and dimensionality reduction in data mining. *Data Min. Knowl. Discov.*, 29(2):534–564, 2015. `doi:10.1007/S10618-014-0356-Z`.

212 Joe H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963. `doi:10.1080/01621459.1963.10500845`.

213 Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945. `doi:10.1007/978-1-4612-4380-9_16`.

214 Marcel Wild. A theory of finite closure spaces based on implications. *Advances in Mathematics*, 108(1):118–139, 1994. `doi:10.1006/aima.1994.1069`.

215 R. Wille. Concept lattices and conceptual knowledge systems. *Computers and Mathematics with Applications*, 23:493–515, 1992. `doi:10.1016/0898-1221(92)90120-7`.

216 Rudolf Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. In Ivan Rival, editor, *Ordered sets*, pages 445–470. Reidel, 1982. `doi:10.1007/978-94-009-7798-3_15`.

217 Rudolf Wille. Complete tolerance relations of concept lattices. In G. Eigenthaler, H. K. Kaiser, W. B. Müller, and W. Nöbauer, editors, *Contributions to general algebra*, volume 3, pages 397–415. Hölder–Pichler–Tempsky, Wien, 1985.

218 Rudolf Wille. Finite distributive lattices as concept lattices. *Atti Inc. Logica Mathematica*, 2:635–648, 1985.

219 Rudolf Wille. Tensorial decomposition of concept lattices. *Order*, 2:81–95, 1985. `doi:10.1007/bf00337926`.

220 Rudolf Wille. Subdirect product construction of concept lattices. *Discrete Mathematics*, 63:305–313, 1987. `doi:10.1016/0012-365X(87)90019-7`.

221 Rudolf Wille. Lattices in data analysis: how to draw them with a computer. In Ivan Rival, editor, *Algorithms and order*, pages 33–58, Dordrecht–Boston, 1989. Kluwer. `doi:10.1007/978-94-009-2639-4_2`.

222 Rudolf Wille. Tensor products of complete lattices as closure systems. In D. Dorninger, G. Eigenthaler, H. K. Kaiser, and W. B. Müller, editors, *Contributions to general algebra*, volume 7, pages 381–385. Hölder–Pichler–Tempsky, Wien, 1991.

223 Rudolf Wille and Uta Wille. Uniqueness of coordinatizations of ordinal structures. In *Contributions to general algebra*, volume 9, pages 321–324, Wien, 1995. Hölder–Pichler–Tempsky.

224 Rudolf Wille and Uta Wille. Coordinatization of ordinal structures. *Order*, 13, 1996. `doi:10.1007/bf00338747`.

225 Uta Wille. Representation of ordinal contexts by ordered n-quasigroups. *Eur. J. Comb.*, 17(2-3):317–333, 1996. `doi:10.1006/EUJC.1996.0027`.

226 Uta Wille. Linear measurement models — axiomatizations and axiomatizability. *Journal of Mathematical Psychology*, 44(4):617–650, 2000. `doi:10.1006/jmps.2000.1326`.

227 Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987. `doi:10.1016/0169-7439(87)80084-9`.

228 Mihalis Yannakakis. The complexity of the partial order dimension problem. *SIAM Journal on Algebraic Discrete Methods*, 3(3):351–358, 1982. `doi:10.1137/0603036`.

229 Javier Yáñez and Javier Montero. A poset dimension algorithm. *J. Algorithms*, 30(1):185–208, 1999. `doi:10.1006/JAGM.1998.0974`.

230 Alexander Zap, Tobias Joppen, and Johannes Fürnkranz. Deep ordinal reinforcement learning. In Ulf Brefeld, Élisa Fromont, Andreas Hotho, Arno J. Knobbe, Marloes H. Maathuis, and Céline Robardet, editors, *Machine Learning and Knowledge Discovery in Databases – European Conference, ECML PKDD 2019, Würzburg, Germany, September 16-20, 2019, Proceedings, Part III*, volume 11908 of *LNCS*, pages 3–18. Springer, 2019. `doi:10.1007/978-3-030-46133-1_1`.

231 Yiqun Zhang and Yiu ming Cheung. An ordinal data clustering algorithm with automated distance learning. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 6869–6876. AAAI Press, 2020. `doi:10.1609/AAAI.V34I04.6168`.