REVIEW

Seriation and Matrix Reordering Methods: An Historical Overview

Innar Liiv*

Department of Informatics, Tallinn University of Technology, Tallinn, Estonia

Received 11 October 2009; revised 18 January 2010; accepted 19 January 2010
DOI:10.1002/sam.10071

Abstract: Seriation is an exploratory combinatorial data analysis technique to reorder objects into a sequence along a one-dimensional continuum so that it best reveals regularity and patterning among the whole series. Unsupervised learning, using seriation and matrix reordering, allows pattern discovery simultaneously at three information levels: local fragments of relationships, sets of organized local fragments of relationships, and an overall structural pattern. This paper presents an historical overview of seriation and matrix reordering methods, several applications from the following disciplines are included in the retrospective review: archaeology and anthropology; cartography, graphics, and information visualization; sociology and sociometry; psychology and psychometry; ecology; biology and bioinformatics; cellular manufacturing; and operations research. © 2010 Wiley Periodicals, Inc. Statistical Analysis and Data Mining 3: 70–91, 2010

Keywords: seriation; matrix reordering; Bertin's matrices; co-clustering; biclustering; combinatorial data analysis

1. INTRODUCTION

Different traditions and disciplines struggle for the leading position toward providing the best insights about the data. Seriation, which is the focus of this paper, can be positioned at the intersection of data mining, information visualization, and network science. All those areas share similar essential goals, but have minimal overlap in mainstream research progress. Prominent authors in the discipline of information visualization [1] (p. 351) have identified that the data mining community gives minimal attention to information visualization, but believe that 'there are hopeful signs that the narrow bridge between data mining and information visualization will be expanded in the coming years'. Shneiderman [2] has pointed out that 'most books on data mining have only a brief discussion of information visualization and vice versa' and that 'the process of combining statistical methods with visualization tools will take some time because of the conflicting philosophies of the promoters'.

A seriation and matrix visualization result contains the clustering of data with additional information about how one cluster is related to another, what the bridging objects are, and what the transition of the objects is like inside the cluster. From the perspective of association rules, a result of seriation can also be interpreted as a chain of associations between objects, which is a non-redundant and optimal presentation of a possibly very lengthy list of association rules. Moreover, it also introduces structural context to those relationships.

Späth [3] (p. 212) considered such matrix permutation approaches to have a great advantage in contrast to the cluster algorithms, because 'no information of any kind is lost, and because the number of clusters does not have to be presumed; it is easily and naturally visible'. Murtagh [4], Arabie and Hubert [5,6] referred to similar advantages calling such an approach a 'non-destructive data analysis', emphasizing the essential property that no transformation or reduction of the data itself takes place.

Bertin described the procedure [7] (p. 6) as 'simplifying without destroying' and was convinced (p. 7) that simplification was 'no more than regrouping similar things'.

Seriation is closely related to clustering, although there does not exist an agreement across the disciplines about defining their distinction. In this paper, seriation is considered different from clustering as shown in Fig. 1. The linking element between those two methods is a clustering with optimal leaf ordering [8–10], which, from the perspective of clustering, is a procedure 'to order the clusters at each level so that the objects on the edge of each cluster are adjacent to that object outside the cluster to which it is nearest' [8]. From the perspective of seriation, the result is

Correspondence to: Innar Liiv (innar.liiv@ttu.ee)

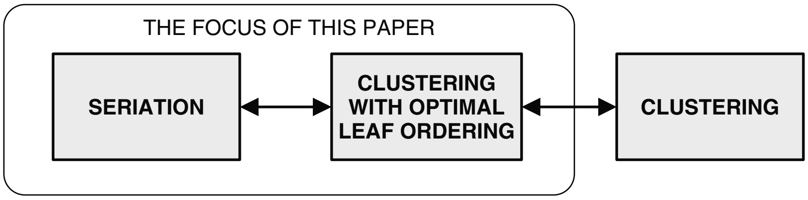


Fig. 1 Seriation and clustering.

an optimal ordering and rearrangement, but an additional grouping procedure is necessary to identify cluster boundaries and clusters [11–13].

The goal of this paper is to present an historical overview of seriation and matrix reordering methods in order to cross-fertilize and align research activities and applications in different disciplines.

In the following section, a definition and an example of seriation will be presented, along with the introduction of the three most common forms of data representations found in related literature.

1.1. Seriation: A Definition

Seriation has the longest roots among the disciplines of archaeology and anthropology, where, for the moment, it has reached also the maturest level of research. A recent monograph about seriation by O'Brien and Lyman [14] includes an extended discussion of the terminology, and a consensual definition by Marquardt [15] (p. 258):

[Seriation is] a descriptive analytic technique, the purpose of which is to arrange comparable units in a single dimension (that is, along a line) such that the position of each unit reflects its similarity to other units.

O'Brien and Lyman [14] (p. 60) point out that 'nowhere in Marquardt's definition is the term *time* mentioned', which perfectly coincides with our interpretation and focus. However, to make the definition more general and compatible with the rest of this paper and set the scene for the construction of our own definition, we suggest to:

- understand and interpret the phrase of *comparable units* as *units from the same mode* according to Tucker's terminology [16], making it less ambiguous whether it is allowed or not to arrange column-conditional and other explicitly non-comparable units along a continuum;
- give more emphasis to simultaneous pattern discovery at several information levels—from local patterns to global. It would make the definition compatible with the requirements set to such matrix permutations by Bertin [7] (p. 12). He saw information as a relationship, which can exist among elements, subsets, or sets, and was convinced that 'the eye perceives the three levels of informations spontaneously' [7] (p. 181).

Consequently, we are able to construct a definition of seriation to reflect our emphasis and focus as follows: **Seriation** is an exploratory data analysis technique to reorder objects into a sequence along a one-dimensional continuum so that it best reveals regularity and patterning among the whole series.

A higher-mode seriation can be simultaneously performed on more than one set of entities, however, entities from different sets are not mixed in the sequence and preserve a separate one-dimensional continuum.

This kind of definition is compatible with less rigorous definitions and highly subjective goals (e.g., to maximize the human visual perception of patterns and the overall trend), but encourages the interpretation from the perspective of minimum description length principle [17,18] and data compressibility [19-21], [22, p. 531].

For didactic purposes, examples in this section will only use binary values, however, we consider and discuss several common value types of the data, where applicable. The scope is additionally limited to entity-to-entity and entity-to-attribute data tables, or using Tucker's [16] terminology and Carroll–Arabie [23] taxonomy, we are concentrating on two-way one-mode ($N \times N$; square tables, where rows and columns refer to the same entities) and two-mode ($N \times M$; rectangular tables, where rows and columns refer to two different sets of entities) data tables. It should be emphasized that such a scope definition does not restrict the one-mode data table to be symmetric, or make entity-to-entity data table to be exclusively one-mode, i.e., there can be relations between entities from different sets, making such a table a two-mode matrix.

Let us look at the following example of seriation along with the introduction of the three most common forms of how data are presented in the related literature and throughout this paper: a matrix, a double-entry table with labels, and a color-coded graphical plot. We may often use the word *matrix* in the paper interchangeably to refer to all of those forms.

An example dataset is first presented as a graph in Fig. 2, where nodes (vertices) are, for example, companies and arcs

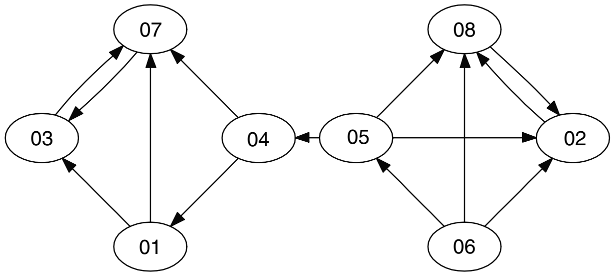


Fig. 2 An example graph.

between the nodes represent a value stream in the supply chain.

We will first construct an asymmetric adjacency matrix that reflects the structure of such a directed graph (i.e., a non-diagonal entry a_{ij} is the number of arcs from node i to node j; see refs [24–26] for a discussion about the differences between node-link diagrams (e.g., Fig. 2) and matrix-based representation):

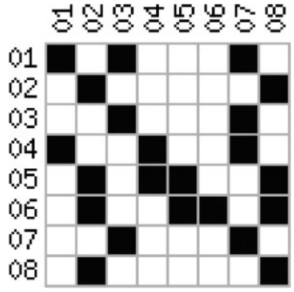
$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Another way to present the same structure is by using a double-entry table with node labels, where the positive elements have been shaded and formatted differently for better visual perception (Table 1).

Inspired by Czekanowski [27] and Bertin [28], it is often reasonable to present the matrix with a graphical plot, where numerical values are color coded. With binary data, the most typical way is to use filled and empty cells to denote 'ones' and 'zeros', respectively. Using such an approach, we can visualize the above structure as presented in Fig. 3.

Table 1. Double-entry table of the example dataset.

	01	02	03	04	05	06	07	08
01	**1**	0	**1**	0	0	0	**1**	0
02	0	**1**	0	0	0	0	0	**1**
03	0	0	**1**	0	0	0	**1**	0
04	**1**	0	0	**1**	0	0	**1**	0
05	0	**1**	0	**1**	**1**	0	0	**1**
06	0	**1**	0	0	**1**	**1**	0	**1**
07	0	0	**1**	0	0	0	**1**	0
08	0	**1**	0	0	0	0	0	**1**

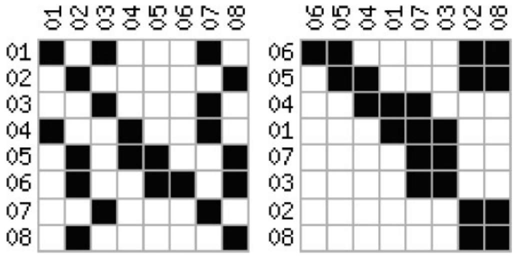


	01	02	03	04	05	06	07	08
01	■		■				■	
02		■						■
03			■				■	
04	■			■			■	
05		■		■	■			■
06		■			■	■		■
07			■				■	
08		■						■

Fig. 3 A graphical 'Bertin' plot for the example dataset.

From such plain matrices, tables and plots, it is still rather complicated to identify the underlying relationships in the data, find patterns and an overall trend. Objects in such an adjacency matrix are ordered arbitrarily, typically in the order of data acquisition/generation, or just sorted alphabetically by labels or names. Changing the order of rows and columns, therefore, does not change the structure: there are *n!* (or *n!m!* in case of a two-mode matrix) permutations of the same matrix that will explicitly reflect the identical structure of the system under observation. The goal of seriation is to find such a permutation, i.e., to reorder the objects from the same mode in a sequence so that it best reveals regularity and patterning among the whole series. This does not, by any means, exclude the chance that data acquisition or alphabetical ordering actually lead to structurally best ordering, but it should never be assumed *a priori*. We can look at this also from the perspective of a single element (cell), the position of which can be changed arbitrarily with the constraint that it must always be moved together with the whole row or column—making it somewhat similar to the classical game of Rubik's cube. An example of the seriation procedure is demonstrated in Fig. 4.

Clearly, from the right plot of Fig. 4, the underlying structure and relationships can be far more easily perceived. However, this is exactly where the challenge of this problem is hidden—how to develop algorithms to perform seriation without exhaustive search of all permutations and how to evaluate the goodness of the result. The new order for rows and columns on the right plot of Fig. 4



	01	02	03	04	05	06	07	08
01	■		■				■	
02		■						■
03			■				■	
04	■			■			■	
05		■		■	■			■
06		■			■	■		■
07			■				■	
08		■						■

	06	05	04	01	07	03	02	08
06	■	■					■	■
05		■	■				■	■
04			■	■	■			
01				■	■	■		
07					■	■		
03					■	■		
02							■	■
08							■	■

Fig. 4 An example of the seriation procedure.

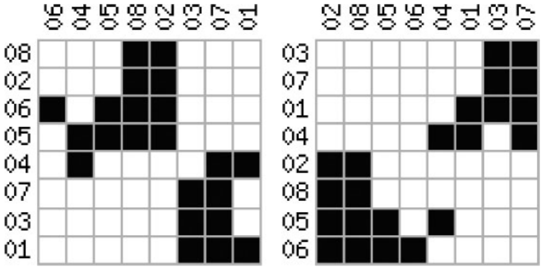


Fig. 5 Alternative permutations for the same dataset.

was reached manually by the author with a highly subjective on-the-fly evaluation of the goodness using visual perception. Actually, this is exactly how it was done in the 1960s and 1970s by a research group directed by a French cartographer Jacques Bertin [7] (p. 47), who stated that, with assistants and mechanical devices, ‘it only takes three days to construct a matrix and three weeks to process and interpret it more deeply’, which was hoped to become easier using computers. At the same time, several algorithms for automatic seriation already existed, but a quick propagation of such developments and results was restrained and muted by the barriers of different scientific traditions and disciplines.

An example of two valid alternative permutations found for the investigated dataset is presented in Fig. 5. Those results are achieved with algorithms called a bond energy algorithm (BEA) [29,30] and ‘minus’ technique [31–35]. The former optimizes an objective function and the latter, if we use the terminology proposed by Van Mechelen *et al.* [36] for similar algorithms, does modeling at a procedural level—a specific heuristical strategy is followed and an overall loss or objective function to be optimized is not implied.

One might notice that both of those matrices (Fig. 5) have different orders for rows and columns. Although we are dealing with one-mode data, such a treatment is reasonable if the graph is directed and, therefore, the adjacency matrix is asymmetric. Finding only a single permutation is possible (as seen in Fig. 4) in such a scenario, but it could result in less structured output due to the reordering restrictions and require some extra data processing (e.g., making the adjacency matrix temporarily symmetric for the duration of the seriation procedure).

Another challenging and focal question concerning the problem of seriation is defining and evaluating which permutation is the best. For the example presented as a graph in Fig. 2 and as a Bertin plot in Fig. 3, we have already proposed three (one on the right of Fig. 4 and two in Fig. 5) relatively good and subjectively interesting permutations. But which one of those opens up the natural inner structure, patterns, regularity, and the overall trend the most? One could subjectively argue that the manual reordering result (Fig. 4, right) offers the best seamless structural transformation, or that the result of the minus technique (Fig. 5, left) illustrates clearly the decomposition of the system and identifies the bridging elements between the two groups. The consensual seriation goal is to maximize the similarity between neighboring objects. However, it still leaves a lot of ambiguity and vagueness for the exact objective function formulation due to virtually hundreds of ways to define sameness and similarity.

In the following retrospective overview, the background, goals and applications of seriation, and matrix reordering methods in different disciplines will be presented.

2. AN HISTORICAL OVERVIEW

An interested reader of all related work on the topic should probably start from the *Organon* collection of the works by Aristotle (384 BC–322 BC), especially the *Categories* [an interesting discussion from the perspective of classification and clustering has been published by Mirkin [37] (Section 1.1)] and the *Topics*. However, in order to keep the specific and incisive focus on the problem of seriation, we will begin with the works of Petrie [38] and Czekanowski [27]. Those works represent a recognized and a systematic start of seriation and matrix permutation visualization, respectively.

Even within the area of seriation, this overview clearly cannot be an exhaustive one, but it should give a coherent view of the related work on the problem and not overlap existing reviews and survey papers. The main emphasis is given to the motivation and the incentive to use seriation, commenting on the developments and examples from the perspective of the taxonomy developed by Carroll and Arabie [23] and making suggestions of minor modifications for implementation steps, where necessary, to make the approaches compatible with others and cross-applicable.

Where possible, related contributions are categorized by disciplines. This enables to highlight the domain-specific motivations, peculiarities, and traits of character, which could possibly have other interesting interpretations across disciplines.

Most of the redundancy of research comes from calling the same thing with different names and calling different things with the same name. This also applies to seriation; therefore, we also try to identify the common terminology in different disciplines.

In addition, two overview sketches (Fig. 6a and b) have been compiled to summarize the connectedness of related work in different disciplines. Relations between research groups and contributions in Fig. 6a are broadly defined as combinations of implicit and explicit references in those works, together with the author’s subjective judgment of

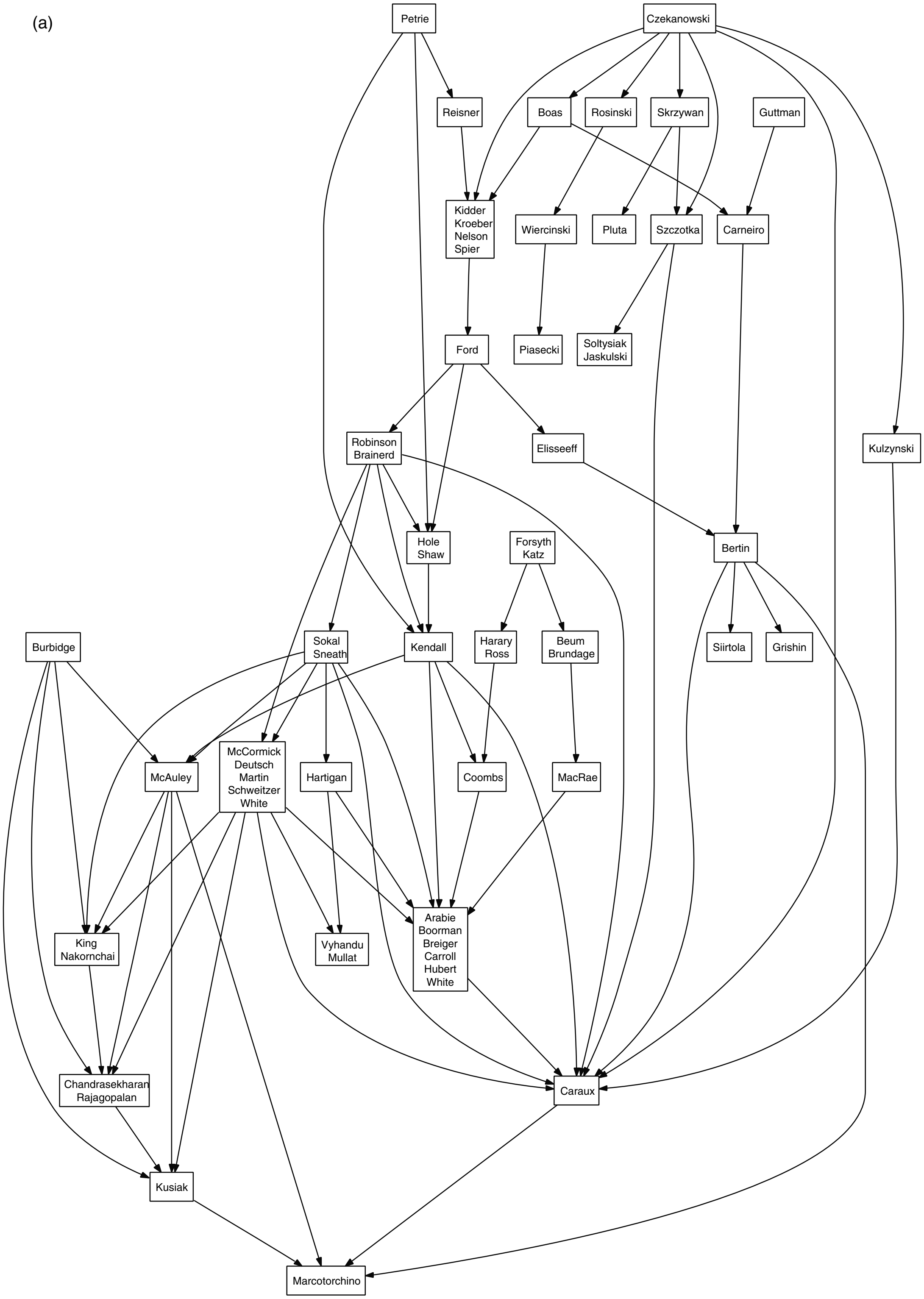


Fig. 6 A visual abstract of the related work in different disciplines.

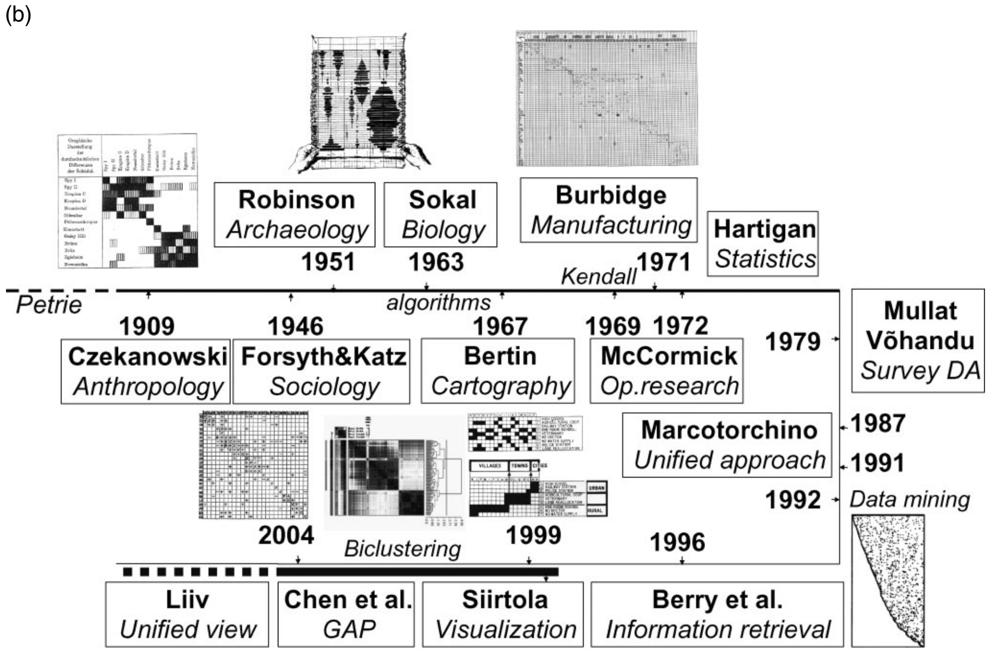


Fig. 6 Continued.

influences, similar approaches and descendence of methods. Both of the figures also serve as visual abstracts of the insights found in the subsequent sections.

2.1. Archaeology and Anthropology

Regardless of the geographic differences in understanding and the exact classification of the fields of archaeology and anthropology, in this section, we are interested in something they have in common—scientific approaches for understanding and reconstructing the past upon partial and incomplete information. Time and sequential arrangement of events, cultures, and traditions play an important role in achieving that goal. There is a wide range of methods for dating in the field of archaeology. According to our focus, we are considering only *seriation*, which belongs to the branch of relative dating. Although the meaning of the word is strongly associated with chronological ordering and time according to most definitions and among field practitioners, there exist several accepted and more general definitions that fit our focus better. According to the definition proposed by Marquardt [15] (p. 258), seriation is ‘a descriptive analytic technique, the purpose of which is to arrange comparable units in a single dimension (that is, along a line) such that the position of each unit reflects its similarity to other units’. O’Brien and Lyman [14] (p. 60) emphasized that such a definition did not narrow down the range or characteristics of units to be seriated, nor did it mention the time as the only or preferred resulting continuum of the linear order attained.

As far as the author knows, the first systematic method for seriation was developed by an English Egyptologist W. M. Flinders Petrie [38], who called it *sequence dating*. His approach was different from others for depending exclusively on the information and the similarity of findings versus professional human judgment of evolutionary and development complexity of artifacts. This type of distinction and classification is also supported by the seriation taxonomy presented by Lyman, Wolverton, and O’Brien [39], who called those two fundamentally different branches as *similiary* and *evolutionary* seriation, respectively. They also distinguished between three types of similiary seriation—phyletic, frequency, and occurrence. However, the goal and the structure to be found in those three distinctions coincided with the only difference in the types of the underlying data. Therefore, from the perspective of this paper, we will discuss those methods interchangeably and not highlight such distinction.

Observations and methods presented by Petrie [38] were not written down using classical mathematical notations, but are nevertheless recognized [40] for being the first to clearly formulate the idea of sequencing objects on the basis of their incidence or abundance. Petrie examined about 900 graves, ‘representing the best selected graves from among over 4000’ [38] and assigned them *sequence dates* using mainly the characteristics of the found pottery. Hole and Shaw [41] (p. 4) describe Petrie being able to ‘seriate the pottery chronologically by merely looking at the characteristics of the handles’. However, there is another way to look at his data, which makes it more systematic.

Table 2. Number of the types of pottery which pass through stages.

Sequence dates	30	31–34	35–42	43–50	51–62	63–71	72–80
30	6	2	0	0	0	0	0
31–34	2	8	2	0	0	0	0
35–42	0	2	8	2	0	0	0
43–50	0	0	2	7	2	0	0
51–62	0	0	0	2	7	2	0
63–71	0	0	0	0	2	7	2
72–80	0	0	0	0	0	2	6

From the figure presented by Petrie [38] (p. 301), we have compiled Table 2, showing the enumerations of the types of pottery that pass through into an adjacent stage. Such a transformation makes the results compatible with the current, generally acceptable seriation formats and would classify as a two-way one-mode data table.

We are able to construct several matrices based on the data from Table 2, depending on the required input of our analysis. Matrices can either take into consideration the numerical values of enumerations (e.g., $A^{(1)}$ and $A^{(2)}$) or present only the occurrence or absence (e.g., $A^{(3)}$) of pottery forms passing through into an adjacent stage:

$$A^{(1)} = \begin{pmatrix} 6 & 2 & 0 & 0 & 0 & 0 & 0 \\ 2 & 8 & 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 8 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 7 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 7 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 6 \end{pmatrix}$$

$$A^{(2)} = \begin{pmatrix} \times & 2 & 0 & 0 & 0 & 0 & 0 \\ 2 & \times & 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & \times & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & \times & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & \times & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 & \times & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & \times \end{pmatrix}$$

$$A^{(3)} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Petrie demonstrated using his visual representation that 'it will be readily seen how impossible it would be to invert the order of any of these stages without breaking up the links between them' and stated that 'the degradation of this type was the best clue to the order of the whole period'. Those statements are also intuitively true for such a matrix representation, with the reservation that complete inversions of the order (e.g., flipping vertically or horizontally) or other operations that would not change the adjacent links, will preserve the regularity, and will be considered with equal regularity throughout this paper.

The interpretation of the work done by Petrie will remain largely subjective, as he did not explicitly describe all the details in the papers and according to ref. [42], his notes and records were destroyed.

Petrie's work influenced several prominent American anthropologists and archaeologists like George Andrew Reisner, Alfred Vincent Kidder, Alfred Louis Kroeber, Nels Nelson, Leslie Spier, James A. Ford (several good reviews and a discussion of such a methodological lineage include refs [14,39,42,43]), who applied, popularized, and further developed the methodology to better suit the practical needs for relative dating.

The evaluation of seriation results remained primarily intuitive and subjective until Brainerd [44] and Robinson [45] proposed a desired final form for the matrix: the highest values in the matrix should be along the diagonal and monotonically decrease when moving away from the diagonal. The paper included a description of an agreement coefficient (basically a similarity coefficient customized for percentage calculations) and a manual procedure and guideline for reordering. In addition, an external relative dating and archaeology-specific test for reordering validation was also introduced. The most influential contribution, however, was the mathematical property of the desired matrix form, which has remained popular along other authors and is often referred to as *Robinsonian matrix, Robinson Matrix* or *R-matrix*.

About a decade later, several algorithms for chronological ordering were proposed [46,47], following a comprehensive monograph by Hole and Shaw [41], covering and evaluating the state-of-the-art techniques for automatic seriation. Hole and Shaw also presented an efficient algorithm called the *permutation search*, which requires $n(n-1)/2 + n^2$ evaluations instead of the exhaustive evaluation of all possible orderings.

Besides the algorithmic enhancements, several papers [48,49] were published about the assumptions, requirements, and conditions under which seriation results can be considered to be approximating the chronological order, not some other underlying regularity.

The approach of Brainerd and Robinson faced some immediate criticism by an anthropologist Lehmer [50] for being too dependent on exact numbers of frequencies and for not taking into account the differences in the size of the collections. Dempsey and Baumhoff [51] proposed a *contextual analysis* method to cope with such a problem, which would merely use the information, irrespective of whether a specific type of artifact was present or absent. They justified their approach for being less sensitive to sampling variations and emphasized that 'types that occur with low frequency may be among the best time-indicators [and] the presence of single specimens of certain types may be crucial in establishing chronologies'. Their approach was classified as occurrence seriation by O'Brein and Lyman [14], together with an extensive discussion on the differences between frequency and occurrence seriation. We consider the progress from frequency seriation to occurrence seriation favorable due to being directly compatible with our definition of seriation.

The dialogue between archaeologists and statisticians was pioneered by Kendall [52,53], who contributed several papers on the research of mathematical properties of the matrix analysis used in archaeology. A similar cooperation among mathematicians, statisticians, and archaeologists eventually led to a dedicated joint conference on *Mathematics in the Archaeological and Historical Sciences*. Those proceedings published in a volume edited by Hodson, Kendall, and Tautu [54], serve to date as one of the most comprehensive collections on research done on archaeological seriation. The research mainly focused on one-mode two-way seriation methods, but there were also examples of two-mode two-way seriation (e.g., ref. [55], p. 7) where artifacts and their variables were directly analyzed without transformation to a similarity matrix format.

Regardless of the classical retrospective look on matrix reordering techniques in archaeology and anthropology, there is another important branch of research that is seldom if ever mentioned in the context of previous methods. It is the work of Jan Czekanowski [27] on matrix reordering and visualization, which is probably the first published work on one-mode data analysis that was based on the permutation of the rows and columns, complemented with color (pattern) coding for better visual perception. He did not have the goal of chronological ordering, but aimed to develop a differential diagnosis of the Neanderthal groups. Differential diagnosis as a term is mainly used in medicine as a systematic method to identify the disease based on an analysis of the clinical data. However, Czekanowski used it in a wider sense—as a systematic classification method to identify and describe groups and their formation in the data. He defined the (dis)similarity coefficient as the average difference of the characteristics of two individuals—the average of the absolute values of differences in characteristics. The results of the difference calculations helped to form a similarity matrix, where the elements/cells were shaded in five different (visual) patterns (as shown in Fig. 7). Czekanowski did not have any formal procedure for rearrangement of the elements in the matrix; therefore, probably, visual inspection and intuition was used because the size of the dataset was also considerably small.

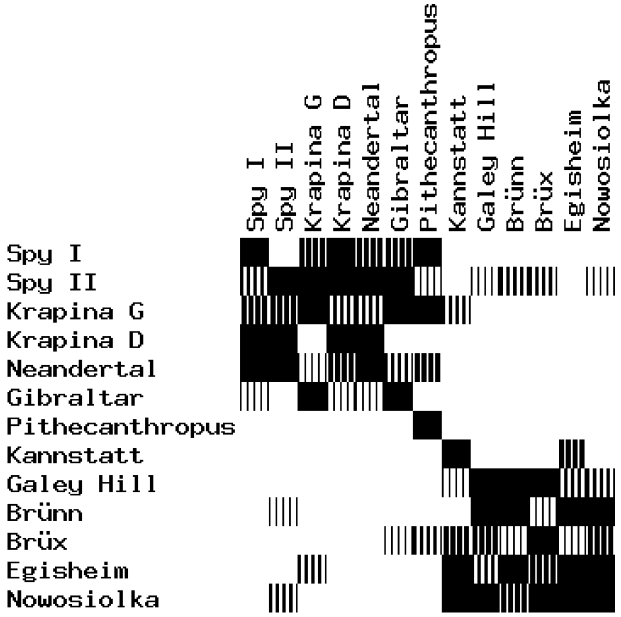


Fig. 7 Czekanowski's diagram [27] of differences and groups of skulls.

Methods developed by Czekanowski have suffered partly from being isolated from Western science. However, the method was widely used by Polish anthropologists Boleslaw Rosinski, Andrzej Wiercinski, and Karol Piasecki (Soltysiak, personal communication, March 12, 2007), who were disciples of Czekanowski's tradition and secured the methodological continuity. According to Soltysiak (personal communication, March 12, 2007), rearrangement of the objects was done visually up to matrices with 50 or more objects and 'first attempts to find a less intuitive ordering procedure were made in early 1950s by Skrzywan [56] and in the Wroclaw school of math (so-called "Wroclaw dendrit", a kind of graph accompanying the Czekanowski's diagram)'. Decades later, Szczotka published a method [57] and developed a computer program for that purpose, and several applications to economics were reported by Pluta [58]. Recent algorithmic advances on the research of Czekanowski's diagram include a genetic algorithm proposed by Soltysiak and Jaskulski [59].

Another interesting methodological lineage exception in the field of anthropology was the work of Carneiro [60], who performed a seriation of a two-way two-mode data table of nine societies and eight culture traits. He developed a scale analysis method for the study of cultural evolution, based on a renowned concept of Guttman scale, applied typically to statistical surveys. An example of the initial data used by Carneiro and the rearranged 'scalogram' is presented on Tables 3 and 4, respectively.

Using the rearrangement procedure proposed in Carneiro's paper, it is clear that, one does not need to evaluate all $n! \bullet m!$ permutations of the table. However, the

Table 3. Initial data used by Carneiro.

	Kuikuru	Anserma	Jivaro	Tupinamba	Inca	Sherente	Chibcha	Yahgan	Cumana
Social stratification	−	+	−	−	+	−	+	−	+
Pottery	+	+	+	+	+	−	+	−	+
Fermented beverages	−	+	+	+	+	−	+	−	+
Political state	−	−	−	−	+	−	+	−	−
Agriculture	+	+	+	+	+	+	+	−	+
Stone architecture	−	−	−	−	+	−	−	−	−
Smelting of metal ores	−	+	−	−	+	−	+	−	−
Loom weaving	−	+	+	−	+	−	+	−	+

Table 4. Rearranged Carneiro's 'scalogram'.

	Yahgan	Sherente	Kuikuru	Tupinamba	Jivaro	Cumana	Anserma	Chibcha	Inca
Stone architecture	−	−	−	−	−	−	−	−	+
Political state	−	−	−	−	−	−	−	+	+
Smelting of metal ores	−	−	−	−	−	−	+	+	+
Social stratification	−	−	−	−	−	+	+	+	+
Loom weaving	−	−	−	−	+	+	+	+	+
Fermented beverages	−	−	−	+	+	+	+	+	+
Pottery	−	−	+	+	+	+	+	+	+
Agriculture	−	+	+	+	+	+	+	+	+

solution is far from trivial in case of larger tables with noisy and missing data. An instructive discussion on non-perfect scales and unilinear evolution was included in the paper.

Carneiro's work serves as an interesting example of how fundamentally different methodological foundations can lead to methods with similar goals and results.

2.2. Cartography and Graphics

From the perspective of this paper, it would be hard to overestimate the importance of a monograph, *Semiology of Graphics,* published by a French cartographer, Jacques Bertin [28]. The main arguments and statements of the presented methodology are accompanied by fine-grained illustrative examples. Despite his main area of expertise, his goal was to propose a concept of a 'reorderable matrix' (*matrice ordonnable*) as a convenient generic tool for analyzing different structures and systems. Reordering of the rows and columns of matrices was performed on two-mode data tables, with a strong emphasis on visualization and value encoding aspects. He stressed the importance of simultaneous availability of three information levels in every effective visual display of data, e.g., a classical Bertin's [7] (p. 33) example of townships in Fig. 8. One should be able to find an immediate visual answer to:

- questions asked specifically about the details of data presented in rows and columns (e.g., Does township *'08'* have a railway station? Which townships have police stations?);
- local patterns found in the data (e.g., Where there is no water supply, there are no high schools);
- global patterns and trends found in the data (e.g., We are able to identify the transformation of rural areas to urban and what changes take place in the characteristics supporting such a transition).

Bertin had some influences from the works of Carneiro [28] (p. 196) and Elisseeff's scalogram [7, p. 58], [61], but the systematic principles he developed for matrices were far more advanced—categorical and continuous values of data were supported and the emphasis was to manipulate the matrices for maximizing the perception of regularity and relationships, regardless of what the final structural pattern would look like. However, the only thing that was not there was the mathematical treatment of matrix permutations and an automatic procedure to evaluate and perform the matrix reordering. Bertin was prepared to believe that after defining the problem, i.e., composing the matrix, data could be processed by machines, but he himself was performing it manually using visual perception. One can find an overview of several mechanical tools and equipment to aid an analyst

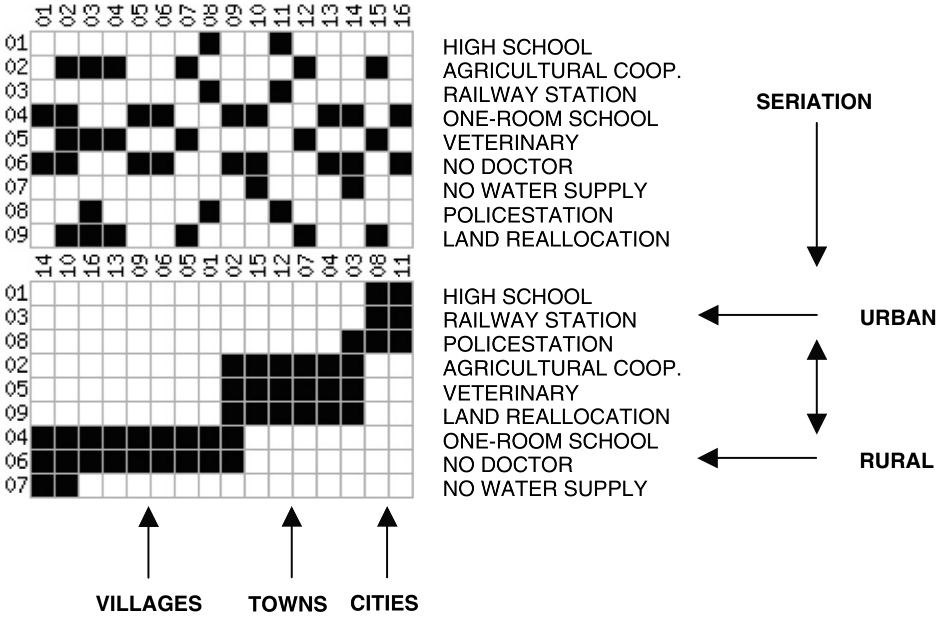


Fig. 8 Bertin's [7,28] example of matrix reordering.

to perform those datasets from the two subsequent monographs [7,28]. Bertin [7] (p. 31) considered the comfortable limits of the proposed graphic information processing to be 120 × 120 with reorderable matrix, 500 × 100 with experimental equipment, and 1000 × 30 with the 'matrix-file' approach, where one dimension (mode) is fixed to be non-permutable.

Bertin's work is highly recognized within the communities of human–computer interaction (HCI) and information visualization, however, with the main emphasis not on reorderable matrices, but fundamentals of visual perception and graphic information processing. One of the most cited recent applications and enhancement of Bertin's ideas of reorderable matrices and especially the 'matrix-file' approach is the *Table Lens* [62], which incorporated interactive elements for better usability together with the general focus + context mantra of the information visualization community.

Bertin [7] (p. 15) was convinced that a set of pie charts is one of the most useless graphical constructions. However, Friendly and Kwan [63] and Friendly [64] demonstrated that combining matrix cell shading with small pie charts to present symmetric correlation matrices can result an interesting visualization. The idea of the *correlogram* or *corrgram* was to use color and intensity of shading in the lower triangle of a symmetric matrix and circle symbols in the corresponding cells of the upper triangle.

Siirtola *et al*. have recently published, in the information visualization community, several discussions on the interaction [65–67] and algorithmic [68,69] aspects of Bertin's reorderable matrices and developed a tool for combining visualization of parallel coordinates with the reorderable matrix [70]. Several interesting papers present experiments and comparisons regarding the readability and interpretation of the matrix-based representation [25,71]. Mueller *et al* [72] have extended the work of Ling [73] on visualizing similarity matrices to large-scale graphs and evaluate the interpretability of results from different one-mode vertex-ordering algorithms, including sensitivity to the initial order of rows and columns.

Berry *et al*. [74] proposed a new information retrieval strategy for browsing hypertext documents by reordering and visualizing term-by-document matrices. Recently, this kind of approach has become very common for co-clustering research [20].

Chen *et al*. [75] most recently published a *Handbook of Data Visualization*, with several chapters presenting discussions and examples about matrix reordering and visualization. It reflects, among others, his own contributions on the *generalized association plots* [76,77], which was based on the idea of visualizing the two-way one-mode matrix after seriation, using different shading to represent the values of proximity. An interested reader is also referred to an excellent review [78] presenting the background and history of seriation and matrix reordering from the perspective of graphical cluster heat maps.

2.3. Sociology and Sociometry

One of the first influential attempts in sociology to introduce a rigorously measurable and new way of thinking was by Jacob L. Moreno with his classic *Who Shall Survive?* (ref [79]; interested readers are directed to the revised edition which is a strongly enhanced version with more background information and available online free of charge [80]). It started a new branch in sociology now known as sociometry, which stressed the importance of quantitative and mathematical methods for understanding social relationships and catalyzed several works interesting in the context of the current paper.

Forsyth and Katz [24] were the first to introduce an approach of rearranging the rows and columns of the *sociomatrix* for a better presentation of the results of sociometric tests. There seems to be neither an obvious nor an implicit influence of previous works with rearranging the matrices, and the motivation for method development seems to descend directly from Moreno's work on sociograms. Forsyth and Katz credited the sociogram as clearly advantageous over verbal descriptions and relationship listings, but 'confusing to the reader, especially if the number of subjects is large'. Katz [81] also argued that 'the sociometric art has simply progressed to the point where pictorial representation of relationships is not enough' and quantifications of the data should be sought. It was hoped that the sociomatrix and the development of methods for analyzing the matrices would fill that gap. Sociogram drawing was a manual process and there were still decades until automatic graph drawing algorithms started to emerge in computer science, and be used across disciplines.

The concept and construction of the sociomatrices (also interrelation matrices) was already an accepted research practice in sociometry. Jennings [82] analyzed leadership and isolation structures, their variations and illustrated choices of preferences between individuals with an adjacency matrix. Dodd [83] wrote a paper about interrelation matrices with a purpose 'to apply algebra to the data of inter-personal relation in order to increase both the precision and the generality of any analyses or syntheses of those data'.

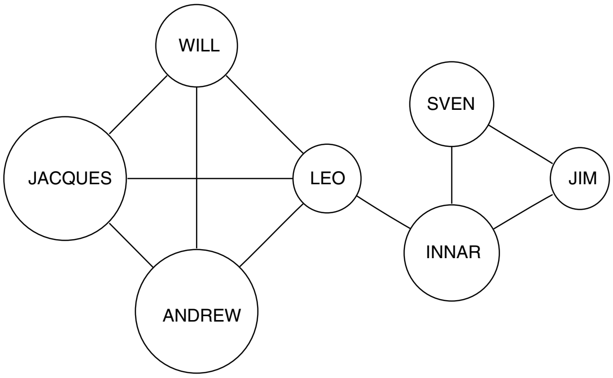


Fig. 9 Sociogram with undirected relations.

A sociomatrix is an asymmetric $N \times N$ one-mode two-way adjacency matrix reflecting the underlying structure of a directed or undirected graph, which is called a sociogram (see Fig. 9, where the links represent acquaintances between those people in a social network) in this context. According to Wasserman and Faust [84], sociomatrix is the most common form for presenting social network data.

The essence of the method which Forsyth and Katz [24] built upon the sociomatrix consisted of 'rearranging the rows and columns in a systematic manner to produce a new matrix which exhibits the group structure graphically in a standard form'. We constructed a simple example to demonstrate the concordance between a sociogram (Fig. 9) and a corresponding sociomatrix before (Fig. 10, left) and after (Fig. 10, right) row and column permutation, where one can directly identify two distinct groups of people and a seamless transformation from one cluster to another. Instead of a binary sociomatrix with undirected single relations, Forsyth and Katz [24] proposed a matrix permutation (see Fig. 11 for reconstruction of their results) with multiple directional relations, denoting positive choices with '+' and negative choices or rejections with '−'. However, those relations were mutually exclusive, so they can be considered as different values for a single relation.

Self-relations or 'self-choices' for a particular relation are usually (e.g., refs [24,82,85,86]) undefined, serve as an identifier of rows/columns [81] or, like in this case, marked with 'X' along the main diagonal of the sociomatrix. Such a common practice is clearly not accidental—sociometry is, after all, a study focusing on *inter-human* relations. Wasserman and Faust [84] point out that 'there are situations in which self-choices do make sense', but they are typically assumed to be undefined 'since most methods ignore these elements' [84] or not to relate to themselves [86]. However, from the unification point of view, there seems to be no explicit reason why self-relations should be undefined or excluded from the analysis and we suggest to replace them with positive choices, at least during the reordering procedure.

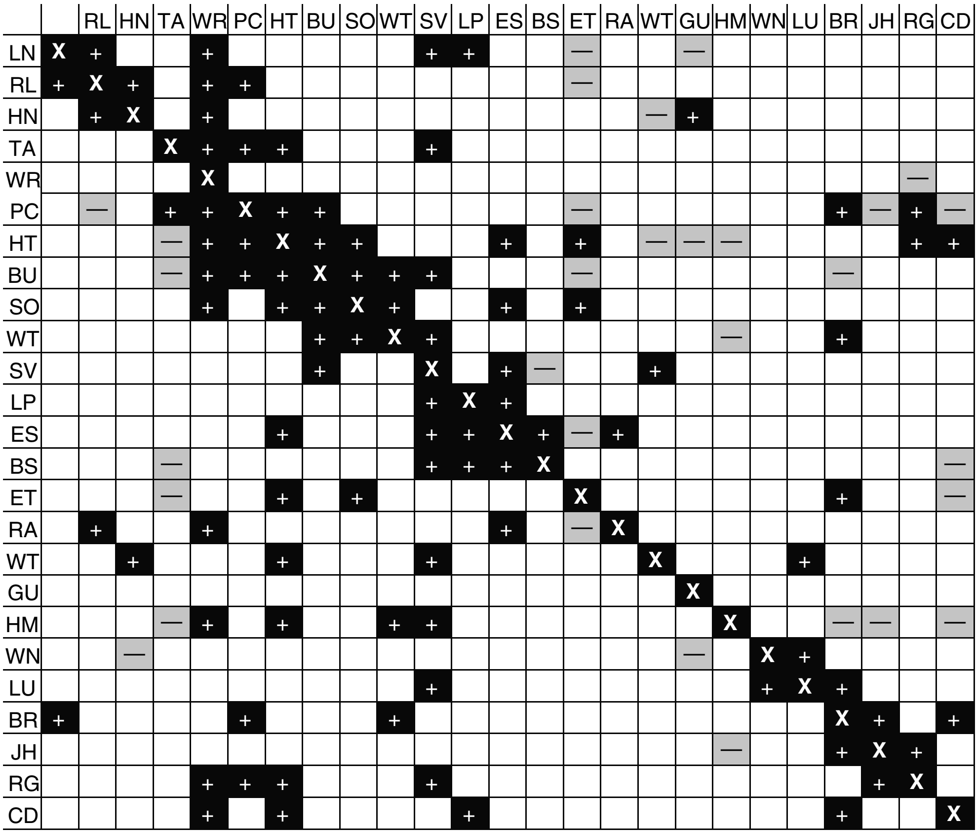
Moreno [87] agreed that a sociogram and a sociomatrix both offered certain advantages and supplemented each other, but did not agree with the claim by Forsyth and Katz that their sociomatrix was superior and more objective in its presentation than the sociogram. He emphasized that 'already pair relations are hard to find (from the sociomatrix), but when it comes to more complex structures as triangles, chain relations, and stars, the sociogram offers many advantages'. Several of those shortcomings are perfectly justified criticism even today. However, a chain of relations is actually *the* most important thing that such matrix shuffling procedures bring out, which might even be difficult to detect on a sociogram with thousands of objects. Katz

	JIM	LEO	ANDREW	SVEN	WILL	JACQUES	INNAR
JIM	X	0	0	1	0	0	1
LEO	0	X	1	0	1	1	1
ANDREW	0	1	X	0	1	1	0
SVEN	1	0	0	X	0	0	1
WILL	0	1	1	0	X	1	0
JACQUES	0	1	1	0	1	X	0
INNAR	1	1	0	1	0	0	X

→

	JACQUES	ANDREW	WILL	LEO	INNAR	SVEN	JIM
JACQUES	X	1	1	1	0	0	0
ANDREW	1	X	1	1	0	0	0
WILL	1	1	X	1	0	0	0
LEO	1	1	1	X	1	0	0
INNAR	0	0	0	1	X	1	1
SVEN	0	0	0	0	1	X	1
JIM	0	0	0	0	1	1	X

Fig. 10 Symmetric sociomatrix before (left) and after (right) permutation.



		RL	HN	TA	WR	PC	HT	BU	SO	WT	SV	LP	ES	BS	ET	RA	WT	GU	HM	WN	LU	BR	JH	RG	CD
LN	X	+			+						+	+			—			—							
RL	+	X	+		+	+									—										
HN		+	X		+												—	+							
TA				X	+	+	+				+														
WR					X																			—	
PC		—		+	+	X	+	+							—							+	—	+	—
HT				—	+	+	X	+	+				+		+		—	—	—					+	+
BU				—	+	+	+	X	+	+	+				—							—			
SO					+		+	+	X	+			+		+										
WT								+	+	X	+								—			+			
SV								+			X		+	—			+								
LP											+	X	+												
ES							+				+	+	X	+	—	+									
BS				—							+	+	+	X											—
ET				—			+		+						X							+			—
RA		+			+								+		—	X									
WT			+				+				+						X				+				
GU																		X							
HM				—	+		+			+	+								X			—	—		—
WN			—															—		X	+				
LU											+									+	X	+			
BR	+					+				+												X	+		+
JH																			—			+	X	+	
RG					+	+	+				+												+	X	
CD					+		+					+										+			X

Fig. 11 Shaded one-mode two-way asymmetric permutated Forsyth–Katz [24] sociomatrix.

[81] continued with the work of simultaneous reordering of the rows and columns of a sociomatrix with the augmentation of quantitative approaches and better mathematical formulation of the problem. A sociomatrix was formalized with bringing in the notation of zeros for indifference/no response and introducing permutation matrices. It seems there was, however, one contradiction. He used the permutation matrix and its transpose for multiplying the original matrix on the left and right, but even with an $N \times N$ matrix, if it contains asymmetric relations (which is typical for sociometric tests), it would be far more reasonable to find two different permutation matrices to maximize the concentration of positive choices. Furthermore, considering the overall notations of this paper, asymmetric sociomatrices should be taken as two-mode two-way matrices for direct compatibility reasons.

The first method for systematically rearranging a sociomatrix was presented half a decade later by Beum and

Brundage [85]. By systematically, we consider methods which have single interpretations on every step of the procedure and do not depend on human visual perception or decision making. Forsyth and Katz [24] also proposed a simple set of rules for iterative enhancement of the matrix reordering and approximate maximization of similarities, but included several abstract and intangible steps. Borgatta and Stolz [88] implemented the Beum–Brundage procedure in FORTRAN-II, which handled matrices up to 145 × 145 variables and included several interesting additional features like de-emphasizing smaller values in a matrix.

Nowadays, quantitative approaches in sociology have advanced enormously, with a strong community and a vast number of contributions in the area of social network analysis (SNA). However, as far as the author is aware, a matrix reordering paradigm is not yet very commonly used. One of the leading software packages in (large) network analysis and visualization, Pajek [89] (Section 12.2), has implemented two Murtagh's seriation algorithms from [90].

In addition, there is a reasonable amount of research originating from the seminal work on blockmodels [91], which include alike matrix reordering procedures, but with the goal of structural aggregation. Therefore, blocks models are more concerned with clustering without the optimal leaf ordering and less concerned about the specific intra-cluster behavior.

2.4. Psychology and Psychometrics

It is quite hard to draw a rigid line between the research of psychometry and sociometry, as several authors have published in both fields and there has been significant cross-influence from both communities. However, for the moment, the community of psychometry has developed a compact and focused track of research on the problem of seriation with a strong consensus on common terminology and a general understanding of the problem.

Hubert [92,93] was one of the early adaptors of seriation techniques in psychology, considering a subjects-by-item response matrix. He performed analyses on both, one-mode and two-mode matrices filled with zero-one and integer values and used permutation procedures based on the algorithms and approaches developed for archaeological seriation.

Besides the archaeological background, Hubert's work was influenced by psychological scaling research carried out by Coombs [94], who proposed the *parallelogram analysis* for searching the patterns in matrices. Coombs [94] (p. 75) called the concept of reordering objects the 'order k/n' analysis, which was a natural extension to the procedure, what he referred to as 'pick k/n'.

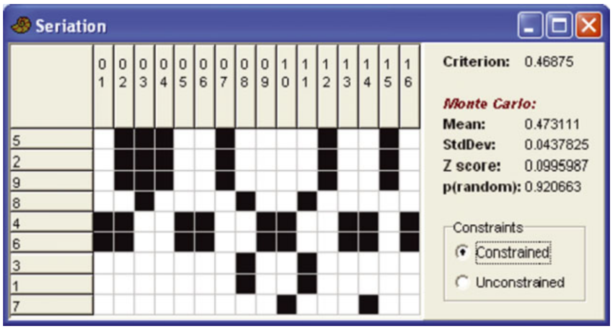
Having influences from the taxonomy of data developed by Coombs and terminology proposed by Tucker [16], Carroll and Arabie [23] proposed a taxonomy of data and models for multidimensional scaling, where the taxonomies of data and models were treated separately. To date, it can still be considered a *de facto* taxonomy to use with multidimensional scaling and related methods, e.g., seriation. From the perspective of multidimensional scaling, seriation is just a one-dimensional special case of the problem.

Comprehensive reviews and references for recent advances can be found from two subsequent monographs on combinatorial data analysis [95,96] and from a structured overview of two-mode clustering [36]. It is interesting to observe that main contributions toward taxonomization of the methods of seriation and the methods related to seriation have come from scholars working in the area of psychology.

2.5. Ecology

Traditions of seriation (which is often referred to as *ordination*) and clustering methods in the disciplines of ecology have strong roots in and descendance from the works of the Polish botanist and politician Kulczynski [97], who studied plant associations using the matrix coding and visualization approach developed by Czekanowski [27]. Kulczynski replaced the values of the upper triangle of a symmetric similarity matrix with different shadings and patterns of a cell (for a reprint of the diagram with a discussion of the ecological application, the reader is referred to ref [98]). This kind of approach for shading was somewhat different from the first visual coding proposed by Czekanowski, who did not preserve the initial similarity values and transformed the (dis)similarity matrix to an asymmetric form after recoding and shading the values.

In ecology, seriation was often considered the best practice to perform clustering without explicitly distinguishing between those two techniques. The application of seriation was also more far-spread than 'classical' clustering techniques used in other disciplines. This may also be the reason why the tools for researcher to perform seriation had the highest representation in the packages developed for ecological studies. It is a significant sign of the maturity level of the discipline from the perspective of seriation methodology development and distinguishes it strongly from other fields. We have performed an illustrative (Figs. 12 and 13) experiment using the PAST [99] software for data analysis, which was 'originally aimed at paleontology but is now also popular in ecology and other fields' [100], using the classical township dataset introduced in another field by Bertin [28] to depict the universality and cross-applicability of seriation methods.



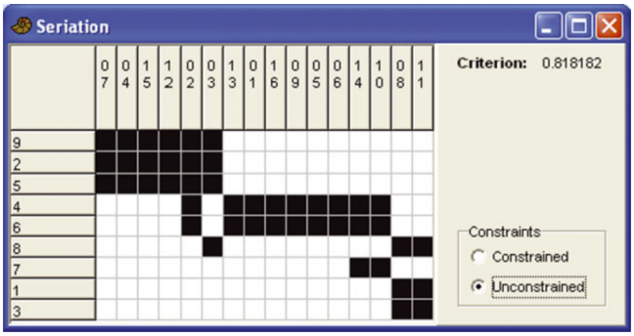
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
5		■	■	■			■					■			■	
2		■	■	■			■					■			■	
9		■	■	■			■					■			■	
8			■					■			■					
4	■	■			■	■			■	■			■	■		■
6	■	■			■	■			■	■			■	■		■
3								■			■					
1								■			■					
7										■				■		

Criterion: 0.46875

Monte Carlo:
Mean: 0.473111
StdDev: 0.0437825
Z score: 0.0995987
p(random): 0.920663

Constraints
● Constrained
○ Unconstrained

Fig. 12 PAST results for row seriation (option: constrained).



	07	04	15	12	02	03	13	01	16	09	05	06	14	10	08	11
9	■	■	■	■	■	■										
2	■	■	■	■	■	■										
5	■	■	■	■	■	■										
4					■		■	■	■	■	■	■	■	■		
6					■		■	■	■	■	■	■	■	■		
8						■									■	■
7													■	■		
1															■	■
3															■	■

Criterion: 0.818182

Constraints
○ Constrained
● Unconstrained

Fig. 13 PAST results for row and column seriation (option: unconstrained).

Besides the above-mentioned PAST software, seriation is also available in the clustering package Clustan [98] (p. 372) and described by the authors of the software package Primer-E (Plymouth Routines In Multivariate Ecological Research) as 'a simple reordering of columns (samples) and rows (species)' that can be an effective way to display 'groupings or gradual changes in species composition' [101] (p. 7/2). The developers of Primer-E also introduced an interesting method for textual coding of the abundance values.

The abundance codes proposed by Clarke and Warwick [101] (p. 7/2) are: '$1 = 1$, $2 = 2, \ldots, 9 = 9$, *a = 10–11, b = 12–13, c = 14–16, . . ., s = 100–116 ,. . ., z = 293–341, A = 342–398, . . ., G = 1000–1165, . . ., Y = 8577–9999,* $Z \geq 10\,000$ *(For* $X \geq 4$ *this is the logarithmic scale* $\text{int}[15(\log_{10}X)-5]$ *assigned to 4–9, a–z, A–Z, omitting i, l, o, I, L, O)*'. The result of this coding is an interesting visualization of a two-way two-mode matrix, using ASCII graphics instead of plotting the pixels, which is obviously a reasonable workaround for non-binary datasets on text-only displays and printers, but infeasible to apply with bigger matrices. Pioneering examples of using ASCII graphics to display one-mode two-way reordered matrices can be traced back to at least 1970s [73].

Legendre and Legendre [98] have published a monograph about *numerical ecology*, which includes a comprehensive overview of data analysis methods in ecology, including discussions, examples, and applications of seriation. Miklos *et al.* have recently published a paper [102] about rearrangement of ecological data matrices, using Markov chain Monte Carlo simulation, which also includes a representative set of references to seriation methods in ecology. Mannila *et al.* [103,104] have recently presented several seriation problem examples with special emphasis on the application in paleontology.

2.6. Biology and Bioinformatics

Seriation methods in biology have similar methodological roots to those discussed in the previous section covering the discipline of ecology. The paradigm of data analysis using the reordering of rows and columns was introduced to the community of biologists by the famous monograph of *Numerical Taxonomy* by Sokal and Sneath [105], which created a lot of controversy for the strong statements and criticism against the traditional way of creating taxonomies in biology. Sokal and Sneath [105] (p. 176) introduced matrix reordering techniques, using the name 'differential shading of the similarity matrix', and referred to the result of the seriation procedure as 'a *cleaned up* diagram'. They saw the purpose of rearranging the rows and columns in the search for the 'optimum structure in the system' and proposed a procedure suggested by Robinson [45] to be suitable for this goal. It is interesting to observe, that, while the systematic approach followed the methodological lineage of Petrie [38], the shading and general matrix visualization approach has rather a strong resemblance to the traditions of Polish scholars Czekanowski and Kulczynski. Their works were acknowledged and cited by Sokal and Sneath, yet, not in the context of matrix reordering but due to similarity coefficient contributions.

Recently, a related concept of *biclustering* (also *co-clustering* or *two-mode clustering* [106,107]) has gained acceptance in experimental molecular biology, mainly, to cope with the latest developments in microarray and gene expression research. The goal of biclustering is to find biclusters—co-occurring subsets of genes (rows) and subsets of conditions (columns) [108]. A typical dataset in bioinformatics for reordering rows and columns is a two-way two-mode matrix with continuous data. In fact, matrix reordering techniques have been introduced to gene analysis decades ago [37,109], but did not attract greater attention before the prominent contribution by Eisen *et al.* [110], who proposed a visual display for genome-wide expression patterns by combining the dendrogram resulting from hierarchical clustering with the initial data matrix from DNA microarray hybridization. The data matrix was reordered using the order of the leaves in the clustering dendrogram acquired separately for both modes of the matrix. A decade later, the paper by Eisen *et al.* [110] had accumulated well over 7000 citations, which gives a good impression of the influence of such an approach. Another important publication toward making data analysis and visualization of gene expression data popular was published by Cheng and Church [108], who introduced the term 'biclustering' to gene expression analysis and proposed a node-deletion algorithm to search for biclusters. Such mathematical treatment and introduction of two-mode clustering to the bioinformatics community attracted hundreds of follow-up articles, discussions, and algorithms.

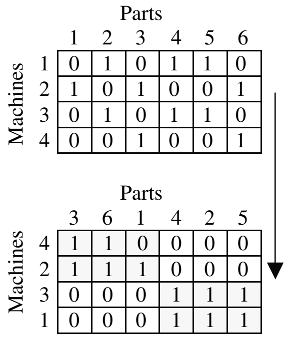
However, from the overall picture of the seriation research, it seems that the community of bioinformatics has not yet established a general consensus concerning the goals and focal emphases of biclustering results. Several surveys and evaluations of biclustering methods [111–113] have been published lately, but there seems to be little work done toward taxonomization of the contributions and the present reviews rather serve as bibliographical lists with brief comments and hubs of references. The most important open question is whether the essential emphasis of the goals is on clustering (objects are assigned to groups) or on seriation (objects are optimally rearranged and assigned to a position within a sequence). If the goal is to perform clustering simultaneously (sequentially) over two sets of objects with the motivation of finding local clusters that could otherwise be left unnoticed, the community should strongly head toward collaboration and consolidation with diclique decomposition [114] and formal concept analysis [115,116] research. If the goal is to augment the human analyst to enable better visual perception of relationships within the data for better biological insight, the use of classical results of hierarchical clustering are not efficient in establishing a seriation of the rows and columns which introduces most of the regularity within the data. A hierarchical clustering and dendrograms choose the order of succeeding elements at every split of the tree arbitrarily or according to the order of appearance in the data source. However, there are '2^{n-1} linear orderings consistent with the structure of the tree' [9] generated by hierarchical clustering. To remedy this situation, several authors have proposed additional procedures to perform optimal leaf ordering of the dendrogram [8–10,117–119].

Caraux and Pinloche [120] have developed a bioinformatics software package *PermutMatrix*, where data analysis of gene expression profiles is performed using several seriation methods. The reader is also referred to an earlier comprehensive overview of matrix reordering techniques by Caraux [121].

2.7. Group Technology and Cellular Manufacturing

The machine-group formation problem and cellular manufacturing represents a community, applying block diagonal seriation with definitely the largest number of technical and algorithmic contributions toward a more optimal solution and formal definition. Machine-group formation is one of the essential steps in Burbidge's analytic 'new approach' to production [122], which later became known as the *production flow analysis* [123]. Production flow analysis is a manufacturing philosophy and technique for finding families of components and groups of machines. It was initially considered [124] a technological enabler for group technology, but those terms were later often used interchangeably. Burbidge [125] emphasized that it is 'concerned solely with methods of manufacture, and does not consider the design features or shape of components at all'. Although the general idea of product flow analysis to classify the components into product families was introduced already in the early 1960s by Burbidge [122] and independently by Mitrofanov [126,127], the machine/part incidence matrix was first explicitly presented by Burbidge [125]. The results were obtained manually [128], the first attempt to develop a non-intuitive algorithm was by McAuley [124], who also stated that 'at present, as far as is known, the only way of finding the groups of machines and families of parts is to rearrange the rows and columns of the matrix, by hand, until the pattern [. . .] is obtained'. McAuley's solution was influenced and based on the works of Sokal and Sneath [105] and Kendall [40]. However, most algorithmic approaches started to appear after the rank order algorithm was proposed [12,129], which worked directly on the initial matrix. This algorithm, among other similar approaches not requiring the conversion of two-mode matrix to one-mode matrix, is classified as *an array-based clustering method* within the cell formation research community.

The machine-group formation (or machine-part cell formation) problem is formulated as a binary part-machine incidence matrix A, where $a_{i,j} = 1$ means that the machine i is required to process part j and $a_{i,j} = 0$ otherwise. We have chosen a simple example (Fig. 14) from McAuley's paper [124] to demonstrate a typical machine/part matrix and how groups emerge after reordering the rows and columns. A zero (often referred as *void*) element ($a_{4,1}$) within an emerged group depicts the reason why reordering and matrix display is used rather than clustering or any other classical partitioning method—groups often have irregular shapes and boundaries and the goal is not only to find groups but also minimize the void elements, exceptional elements (elements which lie outside the blocks on the diagonal) and bottleneck machines [elements which obstruct (the most) the decomposition into independent blocks and subsystems]. This, however, makes the order within the cluster important and excludes, therefore, the possibility of



Machines \ Parts	1	2	3	4	5	6
1	0	1	0	1	1	0
2	1	0	1	0	0	1
3	0	1	0	1	1	0
4	0	0	1	0	0	1

Machines \ Parts	3	6	1	4	2	5
4	1	1	0	0	0	0
2	1	1	1	0	0	0
3	0	0	0	1	1	1
1	0	0	0	1	1	1

Fig. 14 McAuley's example [124] of a machine/part matrix.

using classical clustering methods. This kind of additional domain-specific structural properties might have other interesting semantic interpretations in other fields as well.

Grouping *efficiency* [130] and grouping *efficacy* [131] are the two most frequently used measures to evaluate the quality of the cell formation solution and have had a strong impact on introducing rigorous benchmarking expectations to all new contributions in the field. They are based on measuring the quality of diagonalization by enumerating the number of void elements (zeros) in the block along the diagonal and exceptional elements (ones) away from the formed cells. Such a judgment, however, presupposes that we know (or have predicted) the number of two-mode clusters (there is a similar common problem of finding the correct *k* in the classical clustering paradigm as well) and have identified the cluster boundaries correctly. In addition to those pretentious assumptions, the measures are extremely sensitive to the format of the solution. For example, inversions (flipping horizontally or vertically) and rotations of the matrix, which essentially do not alter the structure in the matrix, could be considered equal solutions, can, however, have a fatal effect toward the ability of those measures to detect a good solution. Other, more universal measures that can find any structural pattern amenable to the dataset are, therefore, more favorable in our context [e.g., the measure of effectiveness (ME) proposed by McCormick *et al.* [29], which will be discussed in the subsequent section].

The machine-part cell formation problem has been solved, using, among others, Hamiltonian path and other graph theoretic approaches [132–134], integer programming [135], fuzzy clustering [136], evolutionary approaches [13,137,138], traveling salesman problem (TSP) [139], neural networks [140–143], branch-and-bound methods [144–146] and with simulated annealing [147]. However, it seems, unfortunately, that most of the algorithms are not known and acknowledged in the other fields applying two-mode clustering and other methods to analyze binary datasets. In addition, Park and Wemmerlöv have developed an artificial shop structure generator [148] for cell formation research, which is also usable and well-applicable in all other fields doing research on seriation and two-mode clustering.

Marcotorchino published a unified approach to block seriation problems for group technology [149] and proposed a unified objective function for block seriation, which, however, required the *k* number of clusters (blocks) to be given as an external input. He also published a subsequent overview paper on seriation problems [150] with a wider span over different disciplines. Although several structural types of seriation were identified, he was mainly concerned with the block diagonal seriation.

The reader is referred to dedicated surveys [151–153] and results of comparisons [154–156] for a further analysis of the contributions in the field. A comprehensive overview and discussion of the research issues in cellular manufacturing is available in ref [157], where the applicability, justification, and implementations of cellular manufacturing systems are discussed.

2.8. Operations Research

Operations research is an interdisciplinary branch of applied mathematics and other scientific methods for determining optimization strategies for the efficient management of organizations. Potential contributions of seriation and matrix reordering techniques originating from this discipline are, therefore, inherently more abstract and contain less domain-specific insights than the ones presented in previous sections. Such settings, on the other hand, enabled McCormick *et al.* [29,30] to contribute, what retrospectively represents an important milestone toward making seriation methods universally applicable and less sensitive to structural pattern assumptions. McCormick *et al.* developed a seriation approach for matrix reordering called BEA to identify natural groups in complex data arrays. It was a nearest-neighbor sequential-selection suboptimal algorithm with the main intention [29] to assist 'the analyst who wishes to begin understanding the interactions in a complex system'. This algorithm can be considered a breakthrough for matrix reordering techniques. As far as the author of this paper is aware, no algorithms were published before 1969 [30] that could perform such a universal reordering of the initial dataset for both one-mode (object-by-object or $N \times N$) and two-mode (object-by-variable or $N \times M$) datasets. One of the strongest properties of the BEA algorithm is not having any assumptions of the underlying structure and being less sensitive to noise in the data than its precedents, making the approach more practical in real-world scenarios.

McCormick *et al.* [29] proposed a ME of an array as 'the sum of bond strengths in the array, where the bond strength between two nearest-neighbor elements is defined as their product'. For any non-negative two-mode matrix A, the ME is given by:

$$\mathrm{ME}(A) = \frac{1}{2}\sum_{i=1}^{i=M}\sum_{j=1}^{j=N} \alpha_{i,j}[\alpha_{i,j+1} + \alpha_{i,j-1} + \alpha_{i+1,j} + \alpha_{i-1,j}]$$

(with the convention $\alpha_{0,j} = \alpha_{M+1,j} = \alpha_{i,0} = \alpha_{i,N+1} = 0$)

As noted by McCormick *et al.* [29,30] and Lenstra [158], the given problem can be reduced into two separate optimizations (one for finding the order for columns, the other for rows; we have slightly modified the notation):

Let $\prod = \{\pi(1), \pi(2), \ldots, \pi(M)\}$ denote all $M!$ permutations of (1,2,. . .,M) and $\Phi = \{\phi(1), \phi(2), \ldots, \phi(N)\}$, respectively over all $N!$ permutations of $(1, 2, \ldots, N)$ with the conventions $\pi(0) = \pi(M + 1) = \alpha_{i,0} = 0$ and $\phi(0) = \phi(N + 1) = \alpha_{0,j} = 0$:

$$\arg\max\nolimits_{\prod} = \sum_{i=1}^{i=M}\sum_{j=1}^{j=N} \alpha_{\pi(i),j}[\alpha_{\pi(i-1),j} + \alpha_{\pi(i+1),j}]$$

$$\arg\max\nolimits_{\Phi} = \sum_{i=1}^{i=M}\sum_{j=1}^{j=N} \alpha_{i,\phi(j)}[\alpha_{i,\phi(j-1)} + \alpha_{i,\phi(j+1)}]$$

Lenstra [158] pointed out that the BEA is equivalent to the well-known TSP, but, actually, the interpretation of the ME optimization as two TSPs was already shown earlier by the original authors in the publicly available technical report [30] (p. 82). Climer and Zhang [159] have recently presented an approach for converting the matrix reordering problem to one-mode TSP format with additional k dummy cities for cluster boundary detection. The solution provided by the TSP solver is used to rearrange the data matrix. They have reported better results according to the criteria of ME for the examples presented by McCormick *et al.* [29], using the BEA. However, several authors [11,154,160] have revisited the original algorithm, investigated its properties in detail, and found that the BEA provides near-optimal results in different settings and is not trying to fit any specific structural pattern in the data; therefore, sometimes outperforming even dedicated and less universal domain-specific algorithms.

In addition to the BEA, the same group developed another, less known algorithm—the moment ordering algorithm [30,161]. Deutsch and Martin [161] consider the algorithm as a tool 'for analyzing arrays of data whose underlying organization is known but which is hoped that there is a single underlying variable, according to which the rows and columns of the arrays can both be arranged in meaningful one-dimensional orders'. Both of the algorithms provide seriation in the data, but the latter searches for a solution to position all the values along the diagonal.

Similarly to isolated methodological lineage exceptions in other disciplines, several generic methods to reorder objects and matrices were developed by Mullat [31–33] and Vyhandu [34,35,162,163], which were initially mainly used for survey data analysis, but later to most of the scientific disciplines covered in previous sections. Those methods and approach to matrix reordering were mainly influenced by the contributions [29,161] classified under operations research in this review. Mullat formalized a family of meta-heuristics called the monotone systems [31–33] and it was applied to matrix reordering problems by Vyhandu [35]. Vyhandu's matrix reordering techniques can, similarly to McCormick's [29] BEA, reduce the problem into two separate tasks for reordering the rows and columns. Vyhandu [34,35] demonstrated that a specific entity-to-set weight (structural similarity) measure called *conformity* [162,164] is favorable for such task. These methods were enhanced and fine-tuned over three decades resulting a set of matrix reordering tools supporting categorical data, several structural patterns (i.e., the result is not restricted to one specific structural pattern in the output, e.g., block diagonal or checkerboard form), and higher-mode seriation [165–167].

Some recent publications and monographs covering elements from this branch of heuristics and methods include [37,168] (Section 4.2.1), [169] (Section 3.5.4), as well as application examples of text mining [170] and inventory classification [171].

3. CONCLUSIONS

A representative variety of related work on seriation problems was highlighted in the presented historical overview, where independent work in different disciplines has corroborated the advantages of understanding structural patterns in the system by reordering rows and columns in matrices. The real benefit in such an interdisciplinary overview is not about reinventions across the disciplines, but about understanding the differences in order to share methods, approaches, and technical results.

The concept of seriation and matrix reordering can work toward attaching data mining together with the advantages in information visualization and SNA, which emphasize the importance of simultaneous consideration of global and local patterns.

In the future, reordering the matrices should be a ubiquitous and common practice for everybody inspecting any data table. According to Bertin's [7] emphasis, a matrix or data table is never constructed conclusively, but reconstructed until all relations which lie within it can be perceived. However, seriation cannot be considered ubiquitously *usable*, until implemented and shipped as a standard tool in any spreadsheet and internet browser for enabling such analysis. Then one can say that seriation and matrix reordering is usable. That is the main future goal for seriation.

4. ACKNOWLEDGMENTS

This paper is a revised version of the literature review in the author's doctoral dissertation. The author thanks the anonymous reviewers for their insightful comments and suggestions to improve the quality of this paper.

REFERENCES

[1] B. B. Bederson and B. Shneiderman, The Craft of Information Visualization: Readings and Reflections, San Francisco, Morgan Kaufmann, 2003.

[2] B. Shneiderman, Inventing discovery tools: combining information visualization with data mining, Inf Vis 1 (2002), 5–12.

[3] H. Späth, Cluster Analysis Algorithms for Data Reduction and Classification of Objects, Chichester, UK, Ellis Horwood, 1980.

[4] F. Murtagh, Book review: W. Gaul and M. Schader, Eds., Data, expert knowledge and decisions, Heidelberg: Springer-Verlag, 1988, viii + 380, J Classification 6 (1989), 129–132.

[5] P. Arabie and L. J. Hubert, An overview of combinatorial data analysis, In Clustering and Classification, P. Arabie, L. J. Hubert, and G. De Soete, eds. River Edge, World Scientific, 1996, 5–63.

[6] P. Arabie and L. J. Hubert, Combinatorial data analysis, Annu Rev Psychol 43 (1992), 169–203.

[7] J. Bertin, Graphics and Graphic Information Processing, Berlin, Walter de Gruyter (Translated by W. J. Berg and P. Scott), 1981.

[8] G. Gruvaeus and H. Wainer, Two additions to hierarchical cluster analysis, Br J Math Stat Psychol 25 (1972), 200–206.

[9] Z. Bar-Joseph, D. K. Gifford, and T. S. Jaakkola, Fast optimal leaf ordering for hierarchical clustering, Bioinformatics 17(S1) (2001), S22–S29.

[10] Z. Bar-Joseph, E. D. Demaine, D. K. Gifford, N. Srebro, A. M. Hamel, and T. S. Jaakkola, K-ary clustering with optimal leaf ordering for gene expression data, Bioinformatics 19(9) (2003), 1070–1078.

[11] P. Arabie, L. J. Hubert, and S. Schleutermann, Blockmodels from the bond energy approach, Soc Networks 12 (1990), 99–126.

[12] J. R. King and V. Nakornchai, Machine-component group formation in group technology: review and extension, Int J Prod Res 20(2) (1982), 117–133.

[13] J. F. Goncalves and M. G. C. Resende, An evolutionary algorithm for manufacturing cell formation, Comput Ind Eng 47(2–3) (2004), 247–273.

[14] M. J. O'Brien and L. R. Lyman, Seriation, Stratigraphy, and Index Fossils: The Backbone of Archaeological Dating, New York, Kluwer Academic/Plenum Publishers 1999.

[15] W. H. Marquardt, Advances in archaeological seriation, In Advances in Archaeological Method and Theory, M. B. Schiffer, ed. New York, Academic Press, 1978, 257–314.

[16] L. R. Tucker, The extension of factor analysis to three-dimensional matrices, Contributions to Mathematical Psychology, New York, Holt, Rinehart and Winston, 1964, 110–127.

[17] P. D. Grünwald, The Minimum Description Length Principle, Cambridge, MIT Press, 2007.

[18] J. Rissanen, Modelling by shortest data description, Automatica 14 (1978), 465–471.

[19] C. Faloutsos and V. Megalooikonomou, On data mining, compression, and Kolmogorov, Data Min Knowl Discov 15(1) (2007), 3–20.

[20] I. S. Dhillon, S. Mallela, and D. S. Modha, "*Information-Theoretic Co-Clustering,*" *In Proceedings of The Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD-2003)*, Washington, DC, USA, August 24–27, 2003, 89–98.

[21] I. Liiv, Pattern Discovery Using Seriation and Matrix Reordering: A Unified View, Extensions and an Application to Inventory Management. Ph.D. Thesis, Tallinn University of Technology, 2008.

[22] L. Wilkinson, The Grammar of Graphics, (2nd ed.), New York, Springer-Verlag, 2005.

[23] D. J. Carroll and P. Arabie, Multidimensional scaling, Annu Rev Psychol 31 (1980), 607–649.

[24] E. Forsyth and L. Katz, A matrix approach to the analysis of sociometric data: preliminary report, Sociometry 9(4) (1946), 340–347.

[25] M. Ghoniem, J. Fekete, P. Castagliola, On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis, Inf Vis 4(2) (2005), 114–135.

[26] N. Henry, J. Fekete, and M. J. McGuffin, NodeTrix: a hybrid visualization of social networks, IEEE Trans Vis Comput Graph 13(6) (2007), 1302–1309.

[27] J. Czekanowski, Zur differentialdiagnose der Neandertalgruppe, Korrespondenzblatt Deutsch Ges Anthropol Ethnol Urgesch XL(6/7) (1909), 44–47.

[28] J. Bertin, Sémiologie graphique: les diagrammes, les réseaux, les cartes, Paris, Mouton, 1967.

[29] W. T. McCormick, P. J. Schweitzer, and T. W. White, Problem decomposition and data reorganization by a clustering technique, Oper Res 20(5) (1972), 993–1009.

[30] W. T. McCormick, S. B. Deutsch, J. J. Martin, and P. J. Schweitzer, Identification of Data Structures and Relationships by Matrix Reordering Techniques (TR P-512), Arlington, Institute for Defense Analyses, 1969.

[31] J. E. Mullat, Extremal subsystems of monotonic systems I, Automation Remote Control 37 (1976), 758–766.

[32] J. E. Mullat, Extremal subsystems of monotonic systems II, Automation Remote Control 37 (1976), 1286–1294.

[33] J. E. Mullat, Extremal subsystems of monotonic systems III, Automation Remote Control 38 (1977), 89–96.

[34] L. Vyhandu, Fast methods in exploratory data analysis, Trans Tallinn Univ Technol 705 (1989), 3–13.

[35] L. Vyhandu, Some methods to order objects and variables in data systems (in Russian), Trans Tallinn Univ Technol 482 (1980), 43–50.

[36] I. Van Mechelen, H.-H. Bock, and P. De Boeck, Two-mode clustering methods: a structured overview, Stat Methods Med Res 13 (2004), 363–394.

[37] B. G. Mirkin, Mathematical Classification and Clustering (Nonconvex Optimization and Its Applications), Boston-Dordrecht, Kluwer Academic Press, 1996.

[38] F. W. M. Petrie, Sequences in prehistoric remains, J Anthropol Inst Great Britain and England 29(3/4) (1899), 295–301.

[39] L. R. Lyman, S. Wolverton, and M. J. O'Brien, Seriation, superposition, and interdigitation: a history of Americanist graphic depictions of culture change, Am Antiq 63(2) (1998), 239–261.

[40] D. G. Kendall, Seriation from abundance matrices, In Mathematics in the Archaeological and Historical Sciences, F. R. Hodson, D. G. Kendall, and P. Tautu, eds. Chicago, Aldine-Atherton, 1971, 214–252.

[41] F. Hole and M. Shaw, Computer Analysis of Chronological Seriation, Houston, Rice University Studies, 1967.

[42] P. Ihm, *"A contribution to the history of seriation in archaeology,"* in *Proc. Classification—the Ubiquitous Challenge, Proceedings of the 28th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Dortmund, March 911, 2004* 2005, 307–316.

[43] L. R. Lyman, M. J. O'Brien, and R. C. Dunnell, The Rise and Fall of Culture History, New York, Plenum Press, 1997.

[44] G. W. Brainerd, The place of chronological ordering in archaeological analysis, Am Antiq 16(4) (1951), 301–313.

[45] W. S. Robinson, A method for chronologically ordering archaeological deposits, Am Antiq 16(4) (1951), 293–301.

[46] M. Ascher and R. Ascher, Chronological ordering by computer, Am Anthropol 65(5) (1963), 1045–1052.

[47] R. S. Kuzara, G. R. Mead, and K. A. Dixon, Seriation of anthropological data: a computer program for matrix-ordering, Am Anthropol 68(6) (1966), 1442–1455.

[48] J. H. Rowe, Stratigraphy and seriation, Am Antiq 26(3) (1961), 324–330.

[49] R. C. Dunnell, Seriation method and its evaluation, Am Antiq 35(3) (1970), 305–319.

[50] D. J. Lehmer, Robinson's coefficient of agreement. A critique, Am Antiq 17(2) (1951), 151–151.

[51] P. Dempsey and M. Baumhoff, The statistical use of artifact distributions to establish chronological sequence, Am Antiq 28(4) (1963), 496–509.

[52] D. G. Kendall, Incidence matrices, interval graphs and seriation in archaeology, Pacific J Math 28(3) (1969), 565–570.

[53] D. G. Kendall, Some problems and methods in statistical archaeology, World Archaeol 1(1) (1969), 68–76.

[54] F. R. Hodson, D.G. Kendall, P. Tautu, eds. Mathematics in the Archaeological and Historical Sciences, Chicago, Aldine-Atherton, 1971.

[55] A. C. Spaulding, Some elements of quantitative archaeology, In Mathematics in the Archaeological and Historical Sciences, F. R. Hodson, D. G. Kendall, and P. Tautu, eds. Chicago, Aldine-Atherton, 1971, 3–16.

[56] W. Skrzywan, Metoda grupowania na podstawie tablic prof. Czekanowskiego, Przegl Anthropol 18 (1952), 583–599.

[57] F. A. Szczotka, On a method of ordering and clustering of objects, Zastosow Mat 13 (1972), 23–34.

[58] W. Pluta, Sravnitel'nyi mnogomernyi analiz v ekonomicheskih issledovanijah (in Russian), Moskva, Statistika, 1980.

[59] A. Soltysiak and P. Jaskulski, Czekanowski's diagram. A method of multidimensional clustering, In *Proceedings of the New Techniques for Old Times. CAA 98. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 26th Conference*, Barcelona, Spain, March 25–28, 1998, 1999, 175–184.

[60] R. L. Carneiro, Scale analysis as an instrument for the study of cultural evolution, Southwest J Anthropol 18(2) (1962), 149–169.

[61] V. Elisseeff, Possibilités du scalogramme dans l'étude des bronzes chinois arhaiques, Math Sci Humaines 11 (1965), 1–10.

[62] R. Rao and S. K. Card, *"The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus+Context Visualization for Tabular Information,"* in *Proc. Proceedings of the ACM SIGCHI, Boston, MA, United States, April 24–28, 1994* 1994, 318–322.

[63] M. Friendly and E. Kwan, Effect ordering for data displays, Comput Stat Data Anal 43 (2003), 509–539.

[64] M. Friendly, Corrgrams: exploratory displays for correlation matrices, Am Stat 56 (2002), 316–324.

[65] H. Siirtola, Interaction with the reorderable matrix, In *Proceedings of the International Conference on Information Visualization (IV'99)*, London, July 14–16, 1999, 272–277.

[66] H. Siirtola, Interactive cluster analysis, In *Proceedings of the 8th International Conference on Information Visualisation (IV'04)*, London, UK, July 14–16, 2004, 471–476.

[67] H. Siirtola and E. Mäkinen, Constructing and reconstructing the reorderable matrix, Inf Vis 4(1) (2005), 32–48.

[68] E. Mäkinen and H. Siirtola, Reordering the reorderable matrix as an algorithmic problem, In *Proceedings of The Theory and Application of Diagrams: First International Conference*, Diagrams 2000, Edinburgh, Scotland, UK, September 1–3, 2000, 453–467.

[69] E. Mäkinen and H. Siirtola, The barycenter heuristic and the reorderable matrix, Informatica 29(3) (2005), 357–363.

[70] H. Siirtola, Combining parallel coordinates with the reorderable matrix, In *Proceedings of the conference on Coordinated and Multiple Views In Exploratory Visualization*, London, July 16–18, 2003, 63–74.

[71] C. Mueller, B. Martin, and A. Lumsdaine. Interpreting large visual similarity matrices, In *Proceedings of Asia-Pacific Symposium on Visualization*, Sydney, Australia, 2007, 149–152.

[72] C. Mueller, B. Martin, and A. Lumsdaine. A comparison of vertex ordering algorithms for large graph visualization, In *Proceedings of Asia-Pacific Symposium on Visualization* 2007, Sydney, Australia, 141–148.

[73] R. F. Ling, A computer generated aid for cluster analysis, Commun ACM 16(6) (1973), 355–361.

[74] M. W. Berry, B. Hendrickson and P. Raghavan, Sparse matrix reordering schemes for browsing hypertext, In The Mathematics of Numerical Analysis, Vol. 32, Lectures in Applied Mathematics, J. Renegar, M. Shub, and S. Smale, eds. AMS, Park City, UT, 1996, 99–123.

[75] C.-H. Chen, W. Härdle, and A. Unwin, Handbooks of Computational Statistics: Data Visualization, Heidelberg, Springer-Verlag 2008.

[76] C.-H. Chen, Generalized association plots: Information visualization via iteratively generated correlation matrices, Stat Sinica 12 (2002), 7–29.

[77] C.-H. Chen, H.-G. Hwu, W.-J. Jang, C.-H. Kao, Y.-J. Tien, S. Tzeng, and H.-M. Wu, Matrix visualization and information mining, In *Proceedings in Computational Statistics 2004 (Compstat 2004)*, Prague, Czech Republic, Heidelberg, August 23–27, 2004, 85–100.

[78] L. Wilkinson, M. Friendly, The history of the cluster heat map, Am Stat 63(2) (2009), 179–184.

[79] J. L. Moreno, Who Shall Survive? A New Approach to the Problem of Human Inter-relations, New York, Beacon House, 1934.

[80] J. L. Moreno, Who Shall Survive? Foundations of Sociometry, Group Psychotherapy and Sociodrama, New York, Beacon House, 1953.

[81] L. Katz, On the matric analysis of sociometric data, Sociometry 10(3) (1947), 233–241.

[82] H. Jennings, Structure of leadership-development and sphere of influence, Sociometry 1(1/2) (1937), 99–143.
[83] S. C. Dodd, The interrelation matrix, Sociometry 3(1) (1940), 91–101.
[84] S. Wasserman and K. Faust, Social Network Analysis, Cambridge, Cambridge University Press 1994.
[85] C. O. Beum and E. G. Brundage, A Method for Analyzing the Sociomatrix, Sociometry 13(2) (1950), 141–145.
[86] J. Moody, Peer influence groups: identifying dense clusters in large networks, Soc Networks 23 (2001), 261–283.
[87] J. L. Moreno, Sociogram and Sociomatrix: a note to the paper by Forsyth and Katz, Sociometry 9(4) (1946), 348–349.
[88] E. F. Borgatta and W. Stolz, A note on a computer program for rearrangement of matrices, Sociometry 26(3) (1963), 391–392.
[89] W. de Nooy, A. Mrvar, and V. Batagelj, Exploratory Social Network Analysis with Pajek. Structural Analysis in the Social Sciences, Cambridge University Press, New York, NY, 2005.
[90] F. Murtagh, Multidimensional Clustering Algorithms, Vienna, Physika Verlag, 1985.
[91] H. White, S. A. Boorman, and R. Breiger, Social structure from multiple networks. I. Blockmodels of roles and positions, Am J Econ Sociol 81 (1976), 730–790.
[92] L. J. Hubert, Problems of seriation using a subject by item response matrix, Psychol Bull 81(12) (1974), 976–983.
[93] L. J. Hubert, Seriation using asymmetric proximity measures, Br J Math Stat Psychol 29 (1976), 32–52.
[94] C. H. Coombs, A Theory of Data, New York, John Wiley and Sons, 1964.
[95] L. J. Hubert, P. Arabie, and J. Meulman, Combinatorial Data Analysis: Optimization by Dynamic Programming, Philadelphia, Society for Industrial and Applied Mathematics, 2001.
[96] M. Brusco and S. Stahl, Branch-and-Bound Applications in Combinatorial Data Analysis (Statistics and Computing), New York, Springer, 2005.
[97] S. Kulczynski, Die pflanzenassoziation der pieninen, Bull Int Acad Polon Sci Lett: Class Sci Math Nat Ser., B 2 (1927), 57–203.
[98] P. Legendre and L. Legendre, Numerical ecology, Amsterdam, Elsevier Science, 1998.
[99] Ø. Hammer and D. Harper, Paleontological Data Analysis, Oxford, Blackwell Publishing, 2005.
[100] Ø. Hammer, *PAST—PAlaeontological STatistics*. 2009. http://folk.uio.no/ohammer/past/ (Retrieved October 1, 2009).
[101] K. Clarke and R. Warwick, Change in Marine Communities: An Approach to Statistical Analysis and Interpretation (2nd ed.), Plymouth, UK, Primer-E, 2001.
[102] I. Miklos, I. Somodi, and J. Podani, Rearrangement of ecological data matrices via Markov Chain Monte Carlo simulation, Ecology 86(12) (2005), 3398–3410.
[103] H. Mannila, Finding total and partial orders from data for seriation, Discovery Sci 5255 (2008), 16–25.
[104] G. Garriga, E. Junttila, and H. Mannila, Banded structure in binary matrices, In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2008)*, Las Vegas, Nevada, August 24–27, 2008, 292–300.
[105] R. R. Sokal and P. H. A. Sneath, Principles of Numerical Taxonomy, San Francisco, W. H. Freeman, 1963.
[106] J. A. Hartigan, Direct Clustering of a Data Matrix, J Am Stat Assoc 67(337) (1972), 123–129.
[107] J. A. Hartigan, Clustering Algorithms, New York, John Wiley and Sons 1975.
[108] Y. Cheng and G. M. Church, *“Biclustering of expression data,”* in *Proc. Proceedings of the 8th International Conference on Intelligent Systems for Molecular Biology, La Jolla / San Diego, CA, USA, August 19–23, 2000,* 2000, 93–103.
[109] B. G. Mirkin and S. N. Rodin, Graphs and Genes, Berlin, Springer-Verlag 1984.
[110] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein, Cluster analysis and display of genome-wide expression patterns, Proc Natl Acad Sci USA 95(25) (1998), 14863–14868.
[111] S. C. Madeira and A. L. Oliveira, Biclustering algorithms for biological data analysis: a survey, IEEE Trans Comput Biol Bioinform 1(1) (2004), 24–45.
[112] A. Tanay, R. Sharan, and R. Shamir, Biclustering algorithms: a survey, Handbook of Computation Molecular Biology, Boca Raton, CRC Press, 2006, 26.1–26.17.
[113] A. Prelic, S. Bleuler, P. Zimmermann, A. Wille, P. Bühlmann, W. Gruissem, L. Hennig, L. Thiele, and E. Zitzler, A systematic comparison and evaluation of biclustering methods for gene expression data, Bioinformatics 22(9) (2006), 1122–1129.
[114] R. M. Haralick, The diclique representation and decomposition of binary relations, J ACM 21(3) (1974), 356–366.
[115] R. Wille, Concept lattices and conceptual knowledge systems, Comput Math Appl 23 (1992), 493–515.
[116] B. Ganter and R. Wille, Formal Concept Analysis. Mathematical Foundations, Berlin, Springer, 1999.
[117] N. Gale and W. C. Halperin, Unclassed matrix shading and optimal ordering in hierarchical cluster analysis, J Classification 1 (1984), 75–92.
[118] D. L. Vandev and Y. G. Tsvetanova, *Ordering of Hierarchical Classifications* 1997.
[119] D. L. Vandev and Y. G. Tsvetanova, Perfect chains and single linkage clustering algorithm, In *Proceedings of Statistical Data Analysis (SDA-95)*, Varna, Bulgaria, September 23–28, 1995, 99–107.
[120] G. Caraux and S. Pinloche, PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order, Bioinformatics 21(7) (2005), 1280–1281.
[121] G. Caraux, Reorganisation et représentation visuelle d’une matrice de donnée numériques: un algorithme itératif, Rev Stat Appl 32(4) (1984), 5–23.
[122] J. L. Burbidge, The new approach to production, Prod Eng 40(12) (1961), 769–784.
[123] J. L. Burbidge, Production flow analysis, Prod Eng 42(12) (1963), 742–752.
[124] J. McAuley, Machine grouping for efficient production, Prod Eng 51(2) (1972), 53–57.
[125] J. L. Burbidge, Production flow analysis, Prod Eng 50(4) (1971), 139–152.
[126] S. P. Mitrofanov, Nauchnye osnovy gruppovoj tehnologii (in Russian), Leningrad, Lenizdat, 1959.
[127] S. P. Mitrofanov, The Scientific Principles of Group Technology, London, National Lending Library for Science and Technology, 1966.
[128] J. L. Burbidge, A manual method for production flow analysis, Prod Eng 56 (1977), 34–38.
[129] J. R. King, Machine-component group formation in production flow analysis: an approach using a rank order clustering algorithm, Int J Prod Res 18(2) (1980), 213–232.

[130] M. P. Chandrasekharan and R. Rajagopalan, Groupability: an analysis of the properties of binary data for group technology, Int J Prod Res 27 (1989), 1035–1052.

[131] K. R. Kumar and M. P. Chandrasekharan, Group efficacy: a quantitative criterion for goodness of block diagonal forms of binary matrices in group technology, Int J Prod Res 28(2) (1990), 233–243.

[132] R. G. Askin and K. S. Chiu, A graph partitioning procedure for machine assignment and cell formation in group technology, Int J Prod Res 28(8) (1990), 1555–1572.

[133] R. G. Askin, S. H. Cresswell, and J. B. Goldberg, A Hamiltonian path approach to reordering the part-machine matrix for cellular manufacturing, Int J Prod Res 29(6) (1991), 1081–1100.

[134] R. Rajagopalan and J. L. Batra, Design of cellular production systems—a graph theoretic approach, Int J Prod Res 13(6) (1975), 567–579.

[135] R. K. Gunasingh and R. S. Lashkari, Machine grouping problem in cellular manufacturing systems an integer programming approach, Int J Prod Res 27(9) (1989), 1465–1473.

[136] C. H. Chu and J. C. Hayya, A fuzzy clustering approach to manufacturing cell formation, Int J Prod Res 29(7) (1991), 1475–1487.

[137] H. Hwang and J. U. Sun, A genetic-algorithm-based heuristic for the GT cell formation problem, Comput Ind Eng 30(4) (1996), 941–955.

[138] G. C. Onwubolo and M. Mutingi, A genetic algorithm approach to cellular manufacturing systems, Comput Ind Eng 39 (2001), 125–144.

[139] J. Balakrishnan and P. D. Jog, Manufacturing cell formation using similarity coefficients and a parallel genetic TSP algorithm formulation and comparison, Math Comput Model 21(12) (1995), 61–73.

[140] A. Kusiak and Y. K. Chung, GT/ART: using neural network to form machine cells, Manuf Rev 4(4) (1991), 293–301.

[141] S. Kaparthi and N. C. Suresh, Machine-component cell formation in group technology: a neural network approach, Int J Prod Res 30(6) (1992).

[142] H. A. Rao and P. Gu, Design of cellular manufacturing systems: a neural network approach, Int J Control Autom Syst: Res Eng 2(4) (1993), 407–424.

[143] C. H. Chu, An improved neural network for manufacturing cell formation, Decision Support Syst 20 (1997), 279–295.

[144] A. Kusiak and C. H. Cheng, A branch-and-bound algorithm for solving the group technology problem, Ann Oper Res 26 (1990), 415–431.

[145] A. Kusiak, Branching algorithms for solving the group technology problem, J Manuf Syst 10(4) (1991), 332–343.

[146] A. Kusiak, J. W. Boe, and C. H. Cheng, Designing cellular manufacturing systems: branch-and-bound and A* approaches, IEE Trans 25(4) (1993), 46–56.

[147] V. Venugopal and T. T. Narendran, Cell formation in manufacturing systems through simulated annealing: an experimental evaluation, Eur J Oper Res 63 (1992), 409–422.

[148] Y.-T. Park and U. Wemmerlöv, A shop structure generator for cell formation research, Int J Prod Res 32(10) (1994), 2345–2360.

[149] F. Marcotorchino, Block seriation problems: A unified approach. Reply to the problem of H. Garcia and J. M. Proth (Applied Stochastic Models and Data Analysis, 1(1) (1985), 25–34)', Appl Stoch Models Data Anal 3(2) (1987), 73–91.

[150] F. Marcotorchino, Seriation problems: an overview, Appl Stoch Models Data Anal 7(2) (1991), 139–151.

[151] U. Wemmerlöv and N. L. Hyer, Cellular manufacturing in the U.S. industry: a survey of users, Int J Prod Res 27(9) (1989), 1511–1530.

[152] N. Singh, Design of cellular manufacturing systems: an invited review, Eur J Oper Res 69 (1993), 284–291.

[153] C. T. Mosier, J. Yelle, and G. Walker, Survey of similarity coefficient based methods as applied to the group technology configuration problem, Omega 25(1) (1997), 65–79.

[154] C. H. Chu and M. Tsai, A comparison of three array-based clustering techniques for manufacturing cell formation, Int J Prod Res 28(8) (1990), 1417–1433.

[155] J. Miltenburg and W. Zhang, A comparative evaluation of nine well-known algorithms for solving the cell formation problem in group technology, J Oper Manage 10(1) (1991), 44–72.

[156] L. Kandiller, A comparative study of cell formation in cellular manufacturing systems, Int J Prod Res 32(10) (1994), 2395–2429.

[157] U. Wemmerlöv and N. L. Hyer, Research issues in cellular manufacturing, Int J Prod Res 25(3) (1987), 413–431.

[158] J. K. Lenstra, Clustering a data array and the traveling-salesman problem, Oper Res 22(2) (1974), 413–414.

[159] S. Climer and W. Zhang, Rearrangement clustering: pitfalls, remedies, and applications, J Mach Learn Res 7 (2006), 919–943.

[160] P. Arabie and L. J. Hubert, The bond energy algorithm revisited, IEEE Trans Syst Man Cybern 20(1) (1990), 268–274.

[161] S. B. Deutsch and J. J. Martin, An ordering algorithm for analysis of data arrays, Oper Res 19(6) (1971), 1350–1362.

[162] L. Vyhandu, Fast methods for data processing (in Russian), Computer Systems: Computer Methods for Revealing Regularities, Novosibirsk, Institute of Mathematics Press, 1981, 20–29.

[163] L. Vyhandu, Rapid data analysis methods (in Russian), Trans Tallinn Univ Technol 464 (1979), 21–39.

[164] I. Liiv, R. Kuusik, and L. Vyhandu, Conformity analysis with structured query language, In *Proceedings of the 6th International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases*, Corfu Island, Greece, February 16–19, 2007, 187–189.

[165] A. Saar and R. Kuusik, Metod vydeleniya yader v trjohmernom predstavlenii sociologicheskih dannyh (in Russian), Opyt Izucheniya Teleradio-Zhurnalistiki i Obshhestvennogo Mneniya, Tallinn, Periodika, 1987, 175–189.

[166] R. Kuusik, Metod determinacionnogo analiza v sluchae trjohmernogo predstavleniya sovokupnosti dannyh (in Russian), Opyt Izucheniya Teleradio-Zhurnalistiki i Obshhestvennogo Mneniya, Tallinn, Periodika, 1987, 208–221.

[167] R. Kuusik and A. Saar, Ob opyte primeneniya metoda vydeleniya yader v trjohmernom predstavljenii sociologicheskih dannyh (in Russian), Opyt Izucheniya Teleradio-Zhurnalistiki i Obshhestvennogo Mneniya, Tallinn, Periodika, 1987, 191–207.

[168] B. G. Mirkin and I. Muchnik, Clustering and multidimensional scaling in Russia (1960–1990): a review, In Clustering and Classification, P. Arabie, L. J. Hubert, and G. De Soete, eds. River Edge, World Scientific 1996, 295–339.

[169] E. Tyugu, Algorithms and Architectures of Artificial Intelligence, Amsterdam, IOS Press 2007.

[170] I. Liiv, A. Vedeshin, and E. Täks, Visualization and structure analysis of legislative acts: a case study on the law of obligations, In *Proceedings of the 11th ACM International Conference on Artificial Intelligence and Law (ICAIL), Proceedings of the Conference*, Stanford University, June 4–8, 2007, 189–190.

[171] I. Liiv, Visualization and data mining method for inventory classification, In *Proceedings of The 2007 IEEE International Conference on Service Operations and Logistics, and Informatics*, Philadelphia, August 27–29, 2007, 472–477.