ORIGINAL ARTICLE

Modeling human thinking about similarities by neuromatrices in the perspective of fuzzy logic

Jerzy Grobelny[1] · Rafał Michalski[1] · Gerhard-Wilhelm Weber[2,3]

Abstract

In this work, we propose a new method for modeling human reasoning about objects' similarities. We assume that similarity depends on perceived intensities of objects' attributes expressed by natural language expressions such as low, medium, and high. We show how to find the underlying structure of the matrix with intensities of objects' similarities in the factor-analysis-like manner. The demonstrated approach is based on fuzzy logic and set theory principles, and it uses only maximum and minimum operators. Similarly to classic eigenvector decomposition, we aim at representing the initial linguistic ordinal-scale (LOS) matrix as a max–min product of other LOS matrix and its transpose. We call this reconstructing matrix a neuromatrix because we assume that such a process takes place at the neural level in our brain. We show and discuss on simple, illustrative examples, how the presented way of modeling grasps natural way of reasoning about similarities. The unique characteristics of our approach are treating smaller attribute intensities as less important in making decisions about similarities. This feature is consistent with how the human brain is functioning at a biological level. A neuron fires and passes information further only if input signals are strong enough. The proposal of the heuristic algorithm for finding the decomposition in practice is also introduced and applied to exemplary data from classic psychological studies on perceived similarities between colors and between nations. Finally, we perform a series of simulation experiments showing the effectiveness of the proposed heuristic.

Keywords Similarity perception · Fuzzy logic · Similarity matrix decomposition · Neuromatrices · Linguistic ordinal scales (LOS) · Reconstructing similarity matrix

1 Introduction

In our paper, we present a new approach to modeling and analyzing human perception of similarities that derives both from cognitive psychology and from neurophysiological studies. Though there have been developed a number of models trying to mimic the brain functioning, there is still a variety of issues that are not covered or clear in this regard. In this research, we focus only on modeling of a specific aspect of the human thinking, that is, processing the object's similarities. As a result, our proposal could be treated only as a small subproblem to the general human cognitive architectures available in the literature, such as ACT-R (e.g., [1, 2]), SOAR (e.g., [3, 4]), or EPIC (e.g., [5, 6]). The review of these approaches can be found, for example, in [7–9], whereas interesting discussion of their importance in general artificial intelligence was presented, for example, in [10, 11].

✉ Rafał Michalski
rafal.michalski@pwr.edu.pl
http://www.rafalmichalski.com

Jerzy Grobelny
jerzy.grobelny@pwr.edu.pl
http://www.jerzygrobelny.com

Gerhard-Wilhelm Weber
gerhard.weber@put.poznan.pl

[1] Faculty of Computer Science and Management, Wrocław University of Science and Technology, Wrocław 50-370, Poland

[2] Faculty of Engineering Management, Poznań University of Technology, Poznan 60-965, Poland

[3] Institute of Applied Mathematics, METU, Ankara 06800, Turkey

We assume that similarity depends on perceived intensities of objects' attributes expressed by natural language expressions such as *low*, *medium*, and *high*, that may be provided, for example, by an expert. Ordinal scales are very widely used in psychology, sociology, business or marketing, especially in diverse questionnaire-based studies. In general, one is able to arrange scale items in an order, e.g., from the smallest to the biggest, from the most to the least preferred, important, usable, etc. However, the distances (intervals) between two consecutive items may not be equal throughout the whole scale and are usually unknown. Probably, the most popular scale of this type is the Likert scale [12] and its countless variations (e.g., *1. Strongly disagree*, *2. Disagree*, *3. Neither agree nor disagree*, *4. Agree*, and *5. Strongly agree*). Exact definitions, properties, and applications of different types of scales can be found, for example, in the seminal paper of Stevens [13]. The ordinal scale may be presented in various ways, such as numbers, text, or graphical objects. In this paper, we use scales of ordinal nature with items presented as language expressions and call them linguistic ordinal scales (LOS). They are also sometimes referred to as qualitative ordinal scales (e.g., [14]).

Our main goal is to approximate a matrix containing intensities of objects' similarities evaluated by pairwise comparisons on LOS. For this purpose, we search for such a set of vectors, expressed also on LOS, that, multiplied by their transposes using max and min operators, reconstructs the original matrix as close as possible. The sought set of vectors forms a matrix, which we name the neuromatrix.

Some recent studies involving brain imaging (e.g., [15]) show that tasks differing in cognitive load are related to the existence of brain states that cannot be directly associated with presented stimuli. Thus, modeling and searching for hidden structures of similarities may correspond to the physiology of human brain functioning.

In problem solving approaches, it is more and more common to connect such separate fields as neuroscience and computer science. In this regard, analogies to a physiological human brain functioning are gaining increasing attention. For instance, lately two special issues have been devoted to this subject: *Brain-inspired computing and machine learning* [16] and *Cognitive computing for intelligent application and service* [17]. The researchers, in general, try to mimic the human mental cognition as close as possible. The simple neuron-based models are becoming to a greater extent complex and similar to the real biological mechanisms taking place in human brains (cf. [18]).

Our approach fits well into this trend. We presume that the neuromatrix searched for in our proposal reflects the brain activity occurring at the neural level and represents the objects–factors relations. The matrix of similarities' intensities, in turn, is a kind of a resulting demonstration of those hidden brain processes that can be observed and registered at the behavioral level. Our proposed method may be treated as an attempt to relate neural characteristics of thinking about similarities with their behavioral representations measured on LOS.

1.1 Background and related work

There is a number of various methods developed and extended over the last decades in different fields of science that deal with the data dimensionality reduction problem and finding the data underlying or latent structure. Probably, the most known and popular is the principal component analysis (PCA) that derives directly from the classic linear algebra and involves the eigenvalue and eigenvector decomposition. This approach and its countless modifications have been used by scientists for a variety of purposes in different applications. Currently, it is still common to come across various versions of this method in scientific papers. For example, Zhu et al. [19] combined PCA with linear hashing and manifold learning for similarity search in color images, while Das et al. [20] used PCA to remove irrelevant features in their hybrid neuro-fuzzy reduction model for classification purposes. Other contemporary advancements in regarding PCA may be found, for example, in [21–23]. One of the latest and most comprehensive overviews of the data dimensionality reduction techniques has been provided by Ayesha et al. [24].

Another widely utilized technique used for uncovering hidden structure of surveys' data is factor analysis (FA). Though there is a considerable number of theoretical constraints and practical problems with this approach (e.g., related to applying this method to variables measured on interval scale), it is very popular. The concept of FA is similar to PCA, but due to controlling the within-subjects errors, it is far more computationally challenging. Scientists have also developed more complex and versatile methods that incorporate ideas and computations from FA. These studies led to a common framework called structural equation modeling, which has become, de facto, a standard in studies involving humans answers to various types of surveys. Recommendations on how to apply them in practice are available, for example, in [25], whereas some recent methodological developments in this area can be found, for example, in [26–28].

Most of the available methods assume that the processed data are measured in ratio or interval scales. However, in many cases, especially in the fields of psychology, marketing, education or sociology, the investigators have only ordinal data at their disposal. Unfortunately, it is still quite common for psychologists and researchers from other fields to use FA- or PCA-based methods for the data gathered only on the ordinal scales, which is, strictly

speaking, not fully appropriate. Stevens [13] pays attention to it yet in the 1950s, pointing out that an ordinal scale allows only for determination of greater or less. Thus, the researcher obtains only the rank order of data. In such a case, there is no guarantee that successive intervals on the scale are equal in size.

Though these approaches are useful in many situations, they do not address all the problems associated with their theoretical assumptions (e.g., regarding the nature of the underlying probability distribution). Therefore, researchers have been trying to elaborate new methodologies. The great abundance of misusing the PCA- and FA-based methods is not justified in light of the methodological progress on other than ratio or interval scales. For instance, various equivalents of FA for the ordinal-scale variables were developed, by Jöreskog and Moustaki [29]. These methods, however, still require some assumptions, and they end up with factor loadings which are interpreted as correlations. Despite that, their ideas were further extended, for example, in [30, 31], and recently in [32].

Lately, Revuelta et al. [33] have shown how to apply exploratory and confirmatory analysis to nominal data in Mplus software. Their approach is similar to multinomial logistic regression with unobserved predictors. A version of FA which operates only on binary data (BFA) is described by Belohlavek and Vychodil [34]. Similar proposals were put forward by Boeck and Rosenberg [35] and Widerjans et al. [36]. They developed HICLAS and SIM-CLAS models, respectively. The general idea of these methods consists in reconstructing objects–attributes binary matrix I by two binary matrices of relationships: objects–factors and factors–attributes. Boeck and Rosenberg [35] developed algorithms for performing the reconstruction based on set-theoretical terms, such as association, equivalence, and hierarchical implication. They restrict sets of possible vectors, called as bundles, to those that meet the assumed relationships.

The ordinal HICLAS proposal by Leenen et al. [37] is an extension of the HICLAS model, where the initial binary matrix (objects–attributes) is recreated by bundles containing ordinal-scale variables. They employ specially defined association relationships connecting an AND OR logic to obtain hidden factors and hierarchical relationships between objects and factors.

A similar approach was presented by Ganter and Glodeau [38]. The authors showed the way of searching ordinal factors for an initial binary matrix. This, in turn, is an extension of BFA and aims at a factorization by regular binary factors that are determined individually for each ordinal-scale category. The model corresponds fully to the disjunctive version of ordinal HICLAS proposal [37], though the same result is obtained by a slightly different, more mathematically justified method. Moreover, the

authors analyze properties of their approach in the context of a formal concept analysis. In this respect, they are moving in the direction developed by Belohlavek et al. [39, 40] where BFA is combined with a formal concept analysis. The formal concept is defined as a subset of the attributes' relationships containing these objects that have the same attributes, and the attributes are common for objects from this set. Belohlavek provided many properties of this approach. In particular, he demonstrated that having the binary objects–attributes relationship, one can always find the solution in the form of sets of factors that fully reconstruct the initial matrix. In specific cases, it is also possible to find the solution including fewer vectors than the number of vectors in the initial matrix [34]. Formal concepts allow to restrict the solutions' search space and are generated by contextual knowledge about relationships between objects and attributes. In [41, 42], the binary approach is generalized to fuzzy data and relationships, such that properties and theorems from formal concepts used in BFA apply also if attributes and relations are assessed by the truth value from multimodal logic, i.e., by values from the interval [0, 1]. Recent advances regarding formal concept analysis involving dimensionality reduction may also be found, for example, in [43, 44].

Another original approach to factorization of matrices was put forward by Lin et al. [45], where the matrix with ordinal-scale measures of the objects–attributes relation is explained by vectors representing orthogonal factors.

Apart from the FA group of methods dealing with ordinal-scale data, there are also methods that can be specifically and directly applied for similarities between objects and do not require ratio scale variables. The so-called nonmetric MDS [46–50] approach is one among them. Generally, the MDS class of methods enables one to see the dissimilarities in dimensions which facilitate the identification of possible underlying similarity structures. The final analysis and interpretation of the results, however, might be troublesome if there are more than three dimensions and if an other than Euclidean space is presumed. This trend of research is still active; some of the latest research works include [51–53].

1.2 Contribution

In this research, we put forward to use the fuzzy logic and set theory operators for reconstructing objects' similarities matrix by the searched for (hidden) objects–attributes relation (the neuromatrix). As an input, we take a square, symmetric fuzzy matrix of similarities between objects denoted S (objects–objects, $n \times n$). Based on the structure of these values, we try identify factors that could have possibly shaped the similarity ratings. In technical terms, we want to find the so-called reconstructing vectors

V (objects–factors, $n \times k$) that represent the factors underlying the observed similarities. The idea is similar to classic FA. The number of factors (reconstructing vectors) *k* is either lower or equal to the number of assessed objects. Usually, we would be interested in finding as few factors as possible to reconstruct the initial similarities as well as possible.

The extensive explanations and discussion about the relations, similarities, and differences between the existing methods and our approach is presented in Sect. 2.5. The main contribution of this research can be described under the following three perspectives.

- From a theoretical point of view, we put forward a new methodology that, to the best of our knowledge, has not been earlier developed and utilized. The proposed concept models human cognitive functioning in relation to objects' similarities assessments. The presented approach allows to find hidden objects–attributes' relations based on linguistic expressions and reduce the dimensionality of similarities' matrix. Our approach is theoretically well grounded and soundly logically justified (cf. Sects. 2.1–2.3).
- The technical aspect includes the development of the heuristic algorithm that allows for taking advantage of the theoretical proposal in practical applications. The effectiveness of the put-forward procedure has been proved in performed simulation experiments. The unique characteristics of the method are treating smaller attribute intensities as less important in making decisions about similarities. This feature is consistent with the way the human brain is functioning at a biological level. A neuron fires and passes information further only if input signals are strong enough.
- From the operational and practical application perspective, our proposal extends the arsenal of methods for data dimensionality reduction and finding patterns of an experimental data structure. Additionally, it may be applied for linguistic-based or ordinal data, which is not the case in most other approaches. As it was shown on well-known practical examples, our methodology provides results that can be logically and reasonably interpreted and may allow for better understanding of the examined results. Our approach has a potentially very wide usage in all research concerned with directly assessing objects' similarities.

The rest of this paper is organized in the following way. First, we show and discuss on simple, illustrative examples how the presented model grasps a natural way of reasoning about similarities. In the next section, we discuss relations and differences between our approach and other methods. Then, we provide a description of a heuristic algorithm for finding the underlying structure of the square matrix with

intensities of objects' similarities in the factor-analysis-like manner. Next, we apply our proposal to real experimental data on perceived color similarities provided by Ekman [54] and reanalyzed by Shepard [55, 56] and to an example about subjective nations' similarities described by Kruskal and Wish [57, p. 31]. Section 5 presents experimental simulation results of our algorithm for randomly generated matrices and confront them with a brute force approach. Finally, we sum up the described approach, indicate its possible applications, and broadly discuss possible future studies.

2 Modeling human thinking about similarities

2.1 Fuzzy-set perspective

The idea of the proposed similarity assessment model for a simple case of a single attribute and two objects can be described in the following way: "the X and Y objects are similar if the intensity rating of this attribute for both objects is high." In the perspective of fuzzy sets and multimodal logic, the intensity level of the attribute *A* may be specified as the membership function value of the object in the set of "objects having the attribute *A* at a high intensity level." Then, $\mu_A(X)$ and $\mu_A(Y)$ denote membership function values of X and Y objects, respectively, belonging to the set of objects having the *A* attribute at a high level of intensity. The relation of similarity can be a fuzzy relation, which is, generally, defined as $R(X, Y) = T(\mu(X), \mu(Y))$, where *T* is any *T*-norm or implication. In the presented approach, the fuzzy relation can be expressed as:

$$\text{SIMILARITY}(X, Y) = \mu_A(X) \text{ AND } \mu_A(Y), \tag{1}$$

where X, Y $\in \{O\}$, and *O* is the set of objects being compared. In terms of fuzzy logic, the described inference model can also be formulated as a logical expression:

$$\begin{aligned} \text{SIMILARITY}(X, Y) \text{ is HIGH IFF} \\ (\text{Truth of } (A(X) \text{ is HIGH}) \\ \text{AND Truth of } (A(Y) \text{ is HIGH})), \end{aligned} \tag{2}$$

where $A(X)$ is an intensity of attribute *A* for object X.

In both cases, we can define and calculate the similarity degree for any two objects. For (1):

$$\text{SIMILARITY}(X, Y) = \min\{\mu_A(X), \mu_A(Y)\}, \tag{3}$$

whereas for (2):

SIMILARITY (X, Y) is HIGH

$$= \min\{\text{Truth of } (A(X) \text{ is HIGH}), \text{ Truth of } (A(Y) \text{ is HIGH})\}. \quad (4)$$

In our approach, we assume that SIMILARITY(X, Y) is determined by a human based on the perceived intensity of attribute A, modeled by membership function values or by the truth of (2). The natural way to define and process attributes' intensities is to use natural language expressions such as *low*, *medium*, *high*, etc. It is reasonable to assume a restricted number of intensity degrees given the psychophysiological resolution, sensitivity of the senses, and the cognitive abilities of the human brain. The intensity granularity may also depend on the context.

The illustrative examples described in the next sections show possible extensions of this way of thinking for a greater number of attributes and objects. In the first example, we took the binary data perspective which is typical to Boolean factor analysis (BFA) [34], whereas in the second one, we extend our considerations to LOS values. In the latter example and the algorithm, we adopted a linguistic model for determining and processing the membership function of the degree of truth for intensity levels of attributes. As membership function values and degrees of truth are usually defined in the range of 0–1, linguistic expressions (*low*, *medium*, and *high*) can be replaced with numerical values from such a range (e.g., 0, 0.5, and 1, respectively). Due to the use of only max–min operators in our approach, such manipulations are not necessary. For clarity, we assign subsequent natural numbers to consecutive levels of attribute intensity.

2.2 Binary similarity data example

Let us say that an expert specifies the similarities between all pairs of six sticks $\{a, b, c, d, e, f\}$. Each stick is characterized by two attributes: its length and diameter. The expert is able to assign each stick to two disjunctive length classes: either *long* or *short*, and two disjunctive diameter classes: *wide* and *narrow*. Table 1 presents an exemplary

Table 1 Sample binary data of objects–attributes relation

Sticks	Length attribute		Width attribute	
	Long (l)	Short (s)	Wide (w)	Narrow (n)
a	1	0	0	1
b	1	0	1	0
c	0	1	1	0
d	0	1	0	1
e	1	0	1	0
f	0	1	1	0

result of such a procedure. Let us consider two extreme approaches for assessing sticks similarities given this objects–attributes relation: liberal and conservative.

2.2.1 Liberal pattern of similarity derivation

In the first, liberal logical pattern of similarity generation (L_LPSG), the expert may regard as similar two sticks that are assigned to the same class of one attribute or to the same categories for both attributes, i.e., "both are *long*" OR "both are *short*" OR "both are *wide*" OR "both are *narrow*."

Let us find objects' relations separately for each attribute by computing the Cartesian product of appropriate vectors taken from Table 1 using logical AND. By applying logical OR to these matrices, we obtain the objects' similarities matrix denoted as $S_{L(Bin)}$ ("L" for "liberal" and "Bin" for "binary"), i.e., s_{ij} is one when in any of these matrices the (i, j)th element is one. The zero value will appear in s_{ij} only if in all matrices the (i, j)th element is equal zero. As a result, we obtain (5):

$$\mathbf{S}_{L(Bin)} = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \end{array} \begin{array}{c} a \; b \; c \; d \; e \; f \\ \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array}. \quad (5)$$

As it was demonstrated by Belohlavek and Vychodil [34], vectors from Table 1 may be treated as a matrix of specifying factors. Thus, the similarity relation matrix can be obtained by (6):

$$\begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \end{array} \begin{array}{c} l \; s \; w \; n \\ \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{array} \circ \begin{array}{c} l \\ s \\ w \\ n \end{array} \begin{array}{c} a \; b \; c \; d \; e \; f \\ \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array} = \mathbf{S}_{L(Bin)}. \quad (6)$$

2.2.2 Conservative pattern of similarity derivation

In the second, conservative logical pattern of similarity generation (C_LPSG), two sticks are similar only when they have both attributes assigned to the same categories, i.e., ("both are *long*" AND "both are *wide*") OR ("both are *long*" AND "both are *narrow*") OR ("both are *short*" AND "both are *wide*") OR ("both are *short*" AND "both are *narrow*").

In this pattern of eliciting similarities, the matrix of the objects–attributes relation consists of column vectors representing joint attributes and may take the following form (7):

$$
\begin{array}{c}
\\ a \\ b \\ c \\ d \\ e \\ f
\end{array}
\begin{array}{cccc}
l\,\mathrm{AND}\,w & l\,\mathrm{AND}\,n & s\,\mathrm{AND}\,w & s\,\mathrm{AND}\,n \\
\left[\begin{array}{cccc}
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0
\end{array}\right]
\end{array}. \quad (7)
$$

By applying the same procedure as for the first way of thinking, we obtain the subsequent similarity matrix denoted as $S_{C(Bin)}$ (8), where "C" refers to conservative and "Bin" to "binary":

$$
\begin{array}{c}
\\ a \\ b \\ c \\ d \\ e \\ f
\end{array}
\begin{array}{cccccc}
a & b & c & d & e & f \\
\left[\begin{array}{cccccc}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1
\end{array}\right]
\end{array} = \mathbf{S}_{C(Bin)}. \quad (8)
$$

It can clearly be observed that matrices $S_{L(Bin)}$ and $S_{C(Bin)}$ differ significantly since they reflect various ways of thinking about similarities.

2.3 Our approach: LOS similarity data example

In our method, we want to factorize matrix S with objects' similarities by finding the matrix V such as $V \circ V^{T} = S$. In contrast to the previous example, here, both matrices S and V include LOS values. It seems that this approach is more realistic than the one employing only binary relations. We modified the binary example presented above, i.e., the objects–attributes relations are given on an ordinal scale by means of natural language expressions, that is, *low*, *medium* (*med*), and *high* represented by numbers 1, 2, and 3,

Table 2 Sample LOS intensities of objects–attributes relation

Sticks	Length attribute		Width attribute	
	Long (*l*)	Short (*s*)	Wide (*w*)	Narrow (*n*)
a	High (3)	Low (1)	Medium (2)	Low (1)
b	Medium (2)	Low (1)	Low (1)	High (3)
c	Low (1)	High (3)	Medium (2)	Low (1)
d	Low (1)	Medium (2)	Medium (2)	High (3)
e	Medium (2)	Low (1)	High (3)	Low (1)
f	Low (1)	High (3)	Medium (2)	Medium (2)

respectively. Table 2 contains possible data under these assumptions.

2.3.1 Liberal pattern of similarity derivation

By applying the schemes of thinking from the binary variables case and using similar logical expressions, one may try to construct similarities matrices for LOS values. Obviously, it is not possible to use the same logical operators as they are only defined for binary variables. The natural extension of the matrix Boolean product is a max–min operation, where OR corresponds to max and AND to min. Within the sets theory, the summation is replaced by the max operator, whereas the multiplication by the min one. It can be noticed that the Boolean matrix product is just specific case of the max–min operation. This type of a construct is used as a method of relations composition, especially in the area of fuzzy sets and fuzzy logic (e.g., [58–60]).

In "Appendix 1," we show that the max–min operation can be used for constructing LOS similarities analogously as in the Boolean data example. The similarities matrix is created by means of a union of objects' similarities relations with respect to individual attributes (simple or complex). Such a procedure is equivalent to performing max–min product of V and V^{T}. The neuromatrix V may contain single vectors of objects–attributes relations or vectors being a logical combination of two or more attributes like in the second scheme of eliciting objects' similarities.

Irrespective of the procedure of determining objects' similarities, we assume that each object is fully similar to itself; therefore, diagonal items have the highest similarity scale value. Additionally, we assume that object i is similar to j with the same extent as object j to i; thus, the similarities matrix is symmetric.

Applying the L_LPSG pattern of eliciting similarities and performing the same max–min operation as in the binary example ($V \circ V^{T}$) on vectors from Table 2, we get (9):

$$
\begin{array}{c}
a \\ b \\ c \\ d \\ e \\ f
\end{array}
\begin{array}{cccc}
l & s & w & n \\
\left[\begin{array}{cccc}
3 & 1 & 2 & 1 \\
2 & 1 & 1 & 3 \\
1 & 3 & 2 & 1 \\
1 & 2 & 2 & 3 \\
2 & 1 & 3 & 1 \\
1 & 3 & 2 & 2
\end{array}\right]
\end{array}
\circ
\begin{array}{c}
l \\ s \\ w \\ n
\end{array}
\begin{array}{cccccc}
a & b & c & d & e & f \\
\left[\begin{array}{cccccc}
3 & 2 & 1 & 1 & 2 & 1 \\
1 & 1 & 3 & 2 & 1 & 3 \\
2 & 1 & 2 & 2 & 3 & 2 \\
1 & 3 & 1 & 3 & 1 & 2
\end{array}\right]
\end{array}
$$

$$
=
\begin{array}{c}
a \\ b \\ c \\ d \\ e \\ f
\end{array}
\begin{array}{cccccc}
a & b & c & d & e & f \\
\left[\begin{array}{cccccc}
3 & 2 & 2 & 1 & 2 & 2 \\
2 & 3 & 1 & 3 & 2 & 1 \\
2 & 1 & 3 & 2 & 2 & 3 \\
1 & 3 & 2 & 3 & 1 & 2 \\
2 & 2 & 2 & 1 & 3 & 2 \\
2 & 1 & 3 & 2 & 2 & 3
\end{array}\right]
\end{array}
= \mathbf{S}_{L(LOS)}. \quad (9)
$$

2.3.2 Conservative pattern of similarity derivation

The second way of eliciting similarities presented in this paper starts with determining combined vectors for objects–attributes relations. For this purpose, instead of binary AND we use the min operator (denoted by ∩). Thus, a vector representing sticks that are *long* and *wide* contains minimal values from columns *long* and *wide* from Table 2. Applying this procedure for all combinations of attributes, we obtain (10):

$$
\begin{array}{c}
\begin{array}{cccc} l\cap w & l\cap n & s\cap w & s\cap n \end{array} \\
\begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \end{array}
\begin{bmatrix}
2 & 1 & 1 & 1 \\
2 & 2 & 1 & 1 \\
1 & 1 & 2 & 1 \\
1 & 1 & 2 & 2 \\
2 & 1 & 1 & 1 \\
1 & 1 & 2 & 2
\end{bmatrix}
\end{array} . \tag{10}
$$

Using the second scheme of thinking (C_LPSG) leads to the following similarities matrix (11):

$$
\begin{array}{c}
\begin{array}{cccccc} a & b & c & d & e & f \end{array} \\
\begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \end{array}
\begin{bmatrix}
2 & 2 & 2 & 2 & 2 & 2 \\
2 & 2 & 1 & 1 & 2 & 1 \\
2 & 1 & 2 & 2 & 1 & 2 \\
2 & 1 & 2 & 2 & 1 & 2 \\
2 & 2 & 1 & 1 & 2 & 1 \\
2 & 1 & 2 & 2 & 1 & 2
\end{bmatrix}
\end{array} = \mathbf{S}_{C(\text{LOS})}. \tag{11}
$$

2.4 Our approach characteristics

In the binary example, we presented two natural ways of obtaining similarities between objects. It can be observed that in this LOS example, other logical patterns of determining objects' similarities may be specified. It stems from the fact that attributes are not necessarily disjunctive. For instance, object *f* in Table 2 was rated by an expert as partly *wide* and partly *narrow* at the same time. Thus, one may deem natural to specify similarity based on a combination of three or more attributes instead of only two of them.

The examples described above show how to obtain a matrix of similarities in a natural and logical way both for the binary and for the LOS variables. There are a number of issues that require discussion and clarifications. First, the patterns of determining similarities between objects based on processing their attributes do not represent all possible ways of doing that. Secondly, some may argue that they are not always consistent with various psychological models of assessing similarities. For example, it is easy to see that the application of L_LPSG does not always preserve the transitivity of relations in a similarity matrix, which is

often assumed or is desirable in psychological studies. It may happen that according to an expert, object *a* is similar to *b* with respect to the length attribute, *b* is similar to *c* with respect to the width attribute; but that does not necessarily mean that *a* should be similar to *c* because it may have different length and width.

In contrast to L_LPSG, the more conservative C_LPSG approach guarantees similarity matrix transitivity for binary data. It results from the assumption about the attributes' disjunction and the similarity construction that requires identical attributes for similar objects. The application of the min operation for combining logically attributes may raise doubts and provoke discussions. Referring to the sticks example, let us consider the situation where an expert evaluated the intensities of attributes using LOS (*low*, *medium*, *high*) as in Table 3. Determining the similarities matrix according to the rule "sticks are similar if they are *long* and *wide*," we obtain (12):

$$
\begin{array}{c}
\begin{array}{cccccc} a & b & c & d & e & f \end{array} \\
\begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \end{array}
\begin{bmatrix}
3 & 1 & 1 & 2 & 1 & 2 \\
1 & 3 & 1 & 1 & 1 & 1 \\
1 & 1 & 3 & 1 & 1 & 1 \\
2 & 1 & 1 & 3 & 1 & 2 \\
1 & 1 & 1 & 1 & 3 & 1 \\
2 & 1 & 1 & 2 & 1 & 3
\end{bmatrix}
\end{array} . \tag{12}
$$

It can be observed that although objects *b* and *c* have identical measures of intensities for the *long* and *wide* attributes (the first and second column of Table 3), their similarity is specified at the lowest level. On the other hand, objects *a* and *f* differ in their attributes' intensities, but are assessed as more similar than objects *b* and *c*, where the attributes' intensities are the same.

This seemingly paradoxical result may be interpreted in favor of the min operation. It can be treated as a cautious (pessimistic) similarity assessment in a situation when an expert is not fully convinced that the given attribute characterizes the specific object. For instance, sticks *b* and *c* are rated as being *long* to the same, smallest extent, and as medium *wide*. Though both features are measurable, an

Table 3 Sample LOS intensities of objects–attributes relation showing that lower values are less important in our similarity assessment model

Sticks	Long (*l*)	Wide (*w*)	$l \cap w$
a	High (3)	Medium (2)	Medium (2)
b	Low (1)	Medium (2)	Low (1)
c	Low (1)	Medium (2)	Low (1)
d	High (3)	Medium (2)	Medium (2)
e	Low (1)	Low (1)	Low (1)
f	Medium (2)	Medium (2)	Medium (2)

expert could assign the small length value both to the *short* sticks and to the *very short* ones. Likewise, the *medium width* could represent *slightly less than medium width* sticks or somewhat wider than medium *width* sticks. When attributes are hardly to measure or even categorical, such an approach could be even more convincing. If we were, for example, to assess the sticks' colors and specify the attribute as *blue*, then small level of this feature could be attributed both to *navy blue* and to *teal blue*. In such a case, for a sensitive person, these two colors may not be similar. Generally, in the max–min approach, higher levels of a specific attribute intensities (or a combination of many attributes) in both compared objects increase their degree of similarity.

The way of constructing similarities matrices based on the C_LPSG pattern is in concordance with the general idea of the *feature set model of similarity—the contrast model* proposed by Tversky [61]. In this approach, the similarity determination is described as a feature matching process. The model defines the similarity between objects as a linear combination of the measures of their common and distinctive features. In the conservative scheme of constructing objects' similarities based on attributes/factors (C_LPSG) presented here, only common features are taken into account.

Despite all restrictions, the possibility of explaining objects' similarities subjectively expressed on LOS by factors represented on the same type of scale is attractive cognitively and practically. Obviously, the presented above examples are simple and assume full knowledge about objects' attributes. However, our real task being subject to analysis in this paper is the process *reverse* to that presented in the above examples. We try to find unknown neuromatrix V containing factors that reproduce the similarities matrix S which is *known* and has been, for example, acquired from an expert or a group of experts in real contexts.

Based on the extensive theoretical considerations presented in this section, we employ the max–min product of neuromatrix vectors V with their transposed values to produce (reconstruct) the input matrix of similarities. The general idea is presented in Fig. 1 and can also be expressed as (13):

$$V(\text{objects-factors}, n \times k) \circ V^{\mathrm{T}}(\text{factors-objects}, k \times n)$$
$$= S(\text{objects-objects}, n \times n).$$

(13)

Our idea is similar to the PCA concept where orthogonal eigenvectors reconstruct a square symmetric matrix containing either correlations or covariations. Researchers usually take advantage of this approach to reduce the correlation or covariations matrix and represent it by as few

eigenvectors as possible, trying to reproduce the original matrix as close as possible. In our approach, the input is also square and symmetric, but all the similarity values are measured solely on a LOS and they can be represented neither by correlations nor by covariances. Furthermore, we confine solely to max, min, AND, OR operators. However, our main goal is similar to PCA and we want to represent the complex similarity matrix by a simpler structure consisting of reconstructing vectors that would make the data interpretation easier.

Usually in real contexts, attributes' assessments, such as those given in Tables 1 and 2, are not available. What is more, the experts' ways of thinking (L_LPSG, C_LPSG or others) are also not known. When in such circumstances, one finds a decomposition that reconstructs the similarities matrix well, the interpretation of the factors may be a kind of art, which often takes place in the classic factor analysis (FA) or various types of multidimensional scaling (MDS). It seems to be justified to presume that better decompositions signify that the similarity determination mechanism was closer to the max–min composition of relations model. In such a case, one may try to interpret the obtained neuromatrix in terms of the attributes' composition by appropriate LPSG. Likewise in FA, knowledge about the analyzed context facilitates possible factors' explanations also in our approach. A straightforward mathematical background used in our approach is demonstrated in detail in "Appendix 1" using a very simple example.

2.5 Relation to other methods

The method presented in this paper is inspired by the concepts of FA, MDS, HICLASS, Ordinal HICLASS, Ordinal FA, and BFA, which were reviewed in the Introduction section. In particular, the proposal is similar to the
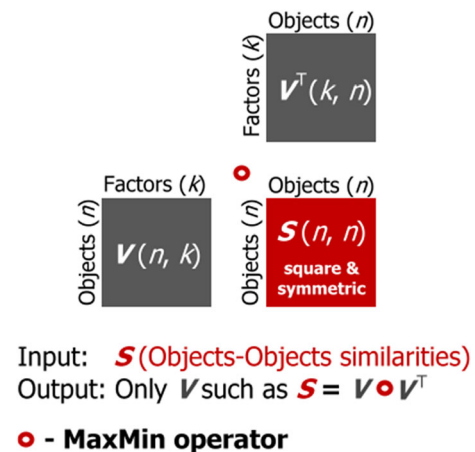

Fig. 1 Schematic illustration of our approach to employ the max–min product of neuromatrix vectors V with their transposed values to produce (reconstruct) the input matrix of similarities ($k \leq n$)

approach originated from fuzzy logic [39], since in both cases the input data represent various intensity levels of similarity relations. In Sect. 2 of our paper, we use ordinal formal contexts for illustrative purposes, but we do not directly refer to the formal concepts analysis. The use of LOS values for describing similarity relationships differentiates our approach form such models as ordinal HICLASS [37] and ordinal FA [38]. In the latter proposal, ordinal scales are employed only for defining searched factors, whereas the matrix being factorized still contains binary values.

Unlike other models, for example, [62–64], in our proposal the formal contexts represented by objects–attributes relationships are not known. In the theoretical setup that we present, we neither generate nor hypothesize about formal contexts. Since the objects–attributes matrix does not exist in our methodology, any type of formal concept analysis is not feasible.

In our approach, we search for unknown factors that explain similarities between objects presented in a square matrix containing pairwise comparisons results. These factors may only, at most, be interpreted as aggregated properties or combinations of objects' attributes which are not known, whereas in formal contexts, objects' attributes are explicitly specified and known before any analysis is conducted.

What is unique in our method is the use of LOS data along with applying only max and min operations in reconstructing the initial matrix. Such a procedure is a generalization of the BFA which is a kind of an extension of the classic FA idea to categorical data. It can also be compared to the association relation in the HICLASS model for ordinal-scale variables, or to the similarity relationship analysis in the fuzzy context and fuzzy concept lattices developed by Belohlavek [39] and extended later by Belohlavek et al. [40–42].

In contrast to the latter approaches, where one analyzes the existing, well-specified fuzzy formal context, we try to find the unknown LOS context, a neuromatrix, understood as an objects–factors relation. Since we take advantage of the simple max–min relations for reconstructing the similarity matrix, there is no need for using multivalued logic formulas, as it is the case in fuzzy concepts approaches.

In general, we search for one objects–factors matrix V (objects–factors, $n \times k$) that reconstructs a square and symmetric matrix S (objects–objects, $n \times n$) of similarities between objects rated on an ordinal scale. Thus, we want V (objects–factors, $n \times k$) $\circ$ V^T (factors-objects, $k \times n$) = S (objects–objects, $n \times n$). In proposals of Belohlavek and colleagues concerned with Boolean factor analysis and factor analyses involving formal concepts, the input data consist of relationships between objects and attributes. The set of objects, attributes, and their

relationships are called a formal context and can be conveniently presented in a form of a matrix: I (objects–attributes, $n \times m$). Based on this input, two additional, distinct matrices are searched for, i.e., A (objects–factors, $n \times k$), and B (factors–attributes, $k \times m$). The relation between these components is (14):

$$I(\text{objects-attributes}, n \times m) = A(\text{objects-factors}, n \times k) \\ \circ B(\text{factors-attributes}, k \times m),$$

(14)

which is totally different than in our proposal.

We provide only one matrix V (objects–factors, $n \times k$) as an output. The only similarity is the correspondence between our matrix V (objects–factors, $n \times k$) and matrix A (objects–factors, $n \times k$) from formal concepts approaches. Matrices I (objects–attributes, $n \times m$) and B (factors-attributes, $k \times m$) from formal concepts decompositions do not appear in our approach, whereas our matrix S (objects–objects, $n \times n$) is not present in papers regarding decompositions of fuzzy contexts both in their fuzzy and Boolean versions and all of their modifications.

Finding factors that try to explain the ordinal-scale similarity matrix is also a main purpose of MDS or nonmetric linear FA concepts [49]. In this trend, all the computations are based on dissimilarities that are represented as distances in a multidimensional space, whereas in our technique we are operating directly on LOS similarity intensities between objects.

3 Algorithm proposal for decomposing the LOS similarity matrix

To apply the demonstrated idea to practically seek data structure and reduce problems' dimensionality, it is necessary to develop a procedure of finding the reconstructing vectors. It would be ideal if such a procedure provides a full decomposition of any square and symmetric LOS data matrix, like it is the case in PCA. However, despite many attempts, we were not able to devise any deterministic algorithm for finding the full decomposition of such an array. Therefore, we present a heuristics that does not guarantee finding the vectors that fully reconstruct the initial matrix.

The general idea of our procedure is summarized in Algorithm 1, which takes as input LOS similarities intensities between all pairs of objects and outputs a neuromatrix with reconstructing vectors and ordered reconstructing vectors. Our procedure does not require defining any parameters.

The key point in the presented process of finding the solution is connected with the observation that within the suggested fuzzy-set theory framework negative values do

not appear. Therefore, if big values are included in the reconstructing vector at the beginning of the procedure, there will be no possibility to decrease the values from the reconstructed matrix by remaining vectors. If one imagines the initial similarity matrix values as cuboids laid down on a plane, where each cuboid consists of the number of unitary cubes corresponding to the degree of similarity between a pair of objects, then the process of reconstructing such a structure is just the superposition (max operation) of quasi-rank-one matrices created by applying the min operation on consecutive reconstructing vectors.

Algorithm 1: Finding reconstructing vectors for LOS similarity matrix

Input: LOS similarity matrix (square, symmetric) to be reconstructed S

Output: LOS neuromatrix with reconstructing vectors V

Procedure:

Step 1 Construct and initialize primary variables. Fill V with zeros and insert columns' maximal values from S into the diagonal of V.

Step 2 Perform the decomposition. Search for reconstructing vectors values V by sequentially processing items from the similarity matrix S and compute auxiliary matrices. Repeat it for all columns in S.

2(a) Compute auxiliary matrices: P denotes a matrix with currently predicted values defined as $V_{\text{current}} \circ V_{\text{current}}^{T}$, R is a matrix with residuals equal $S - P$, and H is a hint matrix defined as $h_{ij} = 0$ if $r_{ij} = 0$, else $h_{ij} = s_{ij}$.

2(b) Select a column in H with the minimal value. If multiple columns satisfy the criterion, take the one with the maximal range. In the case, there is more than one column with the same range, select the one with the maximal range after excluding initial maximal values from those columns. If necessary, repeat the procedure.

2(c) Fill in the selected column by finding the minimal h_{ij} value greater than zero and placing it in the v_{ij} and v_{ji} locations. Next, compute the P and R for both locations.

If there are no negative values in R in only one those locations choose this location.

If in both cases there are no negative values in R, pick the location for which the sum of column values from S is the biggest.

If in both cases in R there are negative values choose the one with the smallest absolute sum of negative residuals.

Step 3 Fine-tune the decomposition. Improvements to the initial solution by making small, local changes to values of V.

V_{inc}: For each v_{ij} repeatedly add one until the sum of absolute values of all residuals is getting smaller or a negative residual appears in R.

V_{dec}: For each v_{ij} repeatedly subtract one until the sum of absolute values of all residuals is getting smaller or a negative residual appears in R.

Step 4 Order the reconstructing vectors. Rearrangement of determined V vectors in a descending order of their importance in reconstructing input matrix S. At first, find this reconstructing vector, for which the rank-one matrix

---

$(v \circ v^{T})$ is the closest to the initial data, i.e., the sum of absolute values of all residuals is the lowest. Then, find the next reconstructing vectors in a decreasing order by checking unions of previously determined rank-one matrices with every candidate from the remaining set of vectors, and selecting this combination which gives the best approximation of S.

Given the problem with additive nature of the max–min operations, the proposed heuristic tries to find such values for reconstructing vectors that, from one hand side, reconstruct as much as possible, but on the other hand they do not produce bigger ratings in other places of the predicted matrix than their equivalents in the initial similarity matrix.

In our method, we use integer values; however, one should bear in mind that these values represent LOS variables just like in the "Appendix 1" example. In a computer program that implements this algorithm, the matrix S contains the initial LOS similarity values; the V neuromatrix includes the searched reconstructing vectors. Additionally, we use three other types of matrices. The predicted (reconstructed) matrix which is computed by the max–min multiplication of the current reconstructing vectors and their transposes $P = V_{\text{current}} \circ V_{\text{current}}^{T}$. The matrix of residuals is calculated as $R = S - P$. If R contains only zeros, then the $V \circ V^{T}$ fully reconstructs S. The hint matrix H that stores those items from S which are not fully reproduced (the residual does not amount to zero) by the current reconstructing vectors. So, if $r_{ij} = 0$, then $h_{ij} = 0$; else, $h_{ij} = s_{ij}$.

Some other procedures could be applied for fine-tuning the decomposition, e.g., instead of adding or subtracting repeatedly ones, a combination of adding and subtracting may be used. We also propose here only one of the possible ways of ordering the reconstructing vectors which, however, does not guarantee finding the best vector arrangement. One may devise different heuristics or, at the expense of computing time, check all combinations of column arrangements.

To give the user idea to what extent the individual vectors reconstruct the input matrix, the relative importance and cumulative relative importance are used. They are computed for a given vector as a percentage value of: (sum of absolute differences from the matrix reconstructed by the given vector minus the sum of absolute differences resulting from reconstructing the initial matrix by all previous vectors) divided by (the maximal possible sum of absolute differences). The consecutive relative importance say by what percentage the initial matrix will be better reconstructed if the given vector is included in the solution,

while all preceding reconstructing vectors are also used. The cumulative relative importance can be interpreted as the percent of the input matrix reconstruction. The full reconstruction occurs when the cumulative relative importance is equal to 100%.

Likewise eigenvalues from the classic approach, the relative importance is used for assessing the usefulness of a given vector in reconstructing S, but their interpretation is different and they should not be confused. Eigenvalues are always associated with their eigenvectors. Here, the relative importance depends on the reconstruction degree of S produced by preceding vectors from the ordered matrix. The sum of consecutive relative importance shows to what extent the initial matrix is reconstructed, but we cannot say that the reconstructing vectors here are orthogonal. They cannot be treated as eigenvectors, and they do not contribute irrespective of other vectors. Changing the order of column vectors in the neuromatrix V would not change the overall reconstruction quality. Residuals will be the same. It results from the fact that the quasi-rank-one matrices produced by individual reconstructing vectors are joined by the union operator which provides the same predicted matrix irrespective of the rank-one matrices order.

In methods like FA, MDS, and our approach, where the initial data matrix is to be represented by a restricted number of dimensions/vectors, there is a problem with determining the final model, i.e., the number of reconstructing vectors in our case. The final model should reproduce S as well as possible, but the number of reconstructing vectors should be as small as possible to provide clear and reasonable interpretations of the underlying data. There have been already different solutions proposed in various methods to tackle this problem. They range from a very simple approach like the Kaiser's heuristic [65] which advices retaining vectors in FA with eigenvalues greater than one, to more complex ones put forward by, for example, Ceulemans and Kiers [66], Preacher et al. [67], Wilderjans et al. [68]. The discussion of strategies that could be applied in such situations was provided, for example, in [36, 69–71].

In our approach, the values in reconstructing vectors should be interpreted as the degree of similarity of a given object with the object's attribute (factor) represented by a given reconstructing vector. We assume that the LOS range is the same for all reconstructing vectors. In classic FA, the factor loadings are just correlations between an object and the hidden factor. Likewise in classic FA, one needs to specify the threshold at which the similarity intensity is high enough and the value for which it should be considered as not meaningful. In the literature regarding FA, various recommendations may be found in this regard. According to some researchers, 0.3 is treated as the minimal value for a factor loading [72, 73], whereas others

classify 0.70 or above as high and 0.5 or lower as low [74]. One of the most popular approaches involves using an absolute value of 0.4 as a cutoff and interpret values of 0.6 as high [72, 75]. Analogically to the recommendations used in FA, we propose to use the similarity scale range median as a threshold in our approach. For example, for nine items LOS with increasing intensities, values bigger than the fifth value will denote a significant degree of similarity between a given object and the specific attribute represented by the reconstructing vector.

As mentioned above, the final model should be as parsimonious as possible and provide reasonable degree of the initial matrix reconstruction. Thus, again some recommendation regarding the acceptable value would be useful. Peterson [76] compared real FA metadata with randomly generated data, and based on the results advocates searching solutions in which the variance explained by the factors exceeds 50%. In our case, instead of variances we use cumulative relative importance that is significantly different but also provides information on how good the approximation is. Given the exemplary data provided in Sect. 4 and experiences from simulation studies presented in Sect. 5, we would rather recommend pursuing solutions with cumulative relative importance higher than 80%. It is also worth noting that our method works for symmetric and asymmetric data. The predicted matrix is always symmetric, and the algorithm tries to find such a symmetric approximation of asymmetric data that the absolute sum of residuals is the lowest.

4 Practical examples

Here, we present two different examples showing how the suggested method may be applied and used for drawing conclusions regarding similarities between objects and reconstructing vectors. These examples are well known in the literature regarding MDS, and the original raw data are easily available. Both of them deal with similarities expressed by humans during pairwise comparisons on ordinal scales.

4.1 Perceived colors' similarities

Ekman, in his work [54], asked participants to assess the degree of similarity between 14 colors. Stimuli were displayed in pairs and subjects rated the *qualitative similarity* on a five-step scale. He rescaled the results to the [0, 1] interval range and applied FA treating these data as correlations. Ekman presented a five-dimensional solution that decently reconstructed the initial quasi-correlation matrix. The obtained eigenvectors were identified as: violet, blue, green, yellow, and red.

Shepard [55, 56] reanalyzed Ekman's data using his nonmetric MDS approach which resulted in two dimensions. He also proved that in this case the relationship between similarities and distances are not linear.

Ekman's and Shepard's original analyses were performed on rational data. Our methodology, by definition, cannot be applied for other than ordinal data. Thus, to make the comparisons more logical we transformed the rational data into ordinal ones by multiplying them by 10 and rounding to whole values. The transformed data are given in "Appendix 2". Then, we reproduced Shepard's result for ordinal data using classic nonmetric MDS with a standard stress value as a goal function, in a MATLAB 7.11.0 (R2010b) version. The outcome, illustrated in Fig. 2, was next compared with the solution provided by our methodology (Table 4).

The application of two versions of our algorithm to these data resulted in finding two different decompositions where one of them was able to fully reconstruct the initial data. The ordering of the obtained vectors (cf. Table 4) shows that the similarity matrix can be reconstructed in 83% by only three vectors, which can be interpreted as red, blue, and green.

The obtained results seem to be qualitatively different from both Ekman's FA which suggested five dimensions and the Shepard's [55, 56] two-dimensional solution. The Ekman's approach was correctly criticized by Shepard for using rescaled similarity ratings as correlations (scalar products). In our approach, we do not use correlations, but

proximity measures and, like Shepard, we do not assume a linear relationship between similarities and distances. Additionally, unlike Shepard, we do not "get something from nothing" which is the case in the Ekman's and Shepard's approaches. Ekman obtains precise points in multidimensional Euclidean space from similarities rated on a five-step scale, whereas Shepard provides metric representation based solely on a nonmetric rank order of those proximity measures.

Both previous approaches applied to this experiment assume that people think in terms of dimensions while performing similarity judgments and/or they are aware that such dimensions exist. This, however, might not be true, all the more that in this specific example the colors were presented only in pairs and subjects could not see the broader picture of the whole experiment. There is another problem with the Shepard's solution: How to interpret the identified dimensions? Although the graphical representation (Fig. 2) resembles to some degree the Newton's color circle, Shepard did not provide substantive and convincing explanation of these two dimensions. The idea of a color circle is to provide a color hue categorization and summarize the additive (subtractive) mixing properties of the so-called primary colors: red, green, and blue (cyan, magenta, yellow), but there were significant problems with identifying color physical properties in two-dimensional Euclidean space. From the perceptual point of view, attempts of representing color hues in a two-dimensional Euclidean space resulted in creating the CIE Lab color system [77] where the equal Euclidean distances between color hues correspond approximately to similar differences in their human subjective perceptions. The Shepard's solution seems to be more similar to this approach than to classic color circles. The CIE Lab space is obtained by linear transformations of human photoreceptors sensitivity to red, green, and blue light components. Since the cones sensitivity functions are not straightforward, the system is far from being perfect. In light of the above, it is quite possible that the process of judging color similarities depends more on physical properties of three different types of retina's cones (red, green, and blue) than the artificially created two-dimensional CIE Lab space for the color hue. The process of estimating the colors' similarities could have been simpler and not based on a dimensional idea. We may try to assign (compare, categorize) a given hue to one of the well-known primary colors. In this sense, our approach and the solution proposed by Ekman would be more appropriate than the Shepard's analysis.

Shepard argues also that his two-dimensional solution accounts for as much as 84% of the overall variance which is much better than the reconstruction obtained by the first two (unrotated) Ekman's dimensions (about 64%). However, according to Kruskal's recommendations [47], only
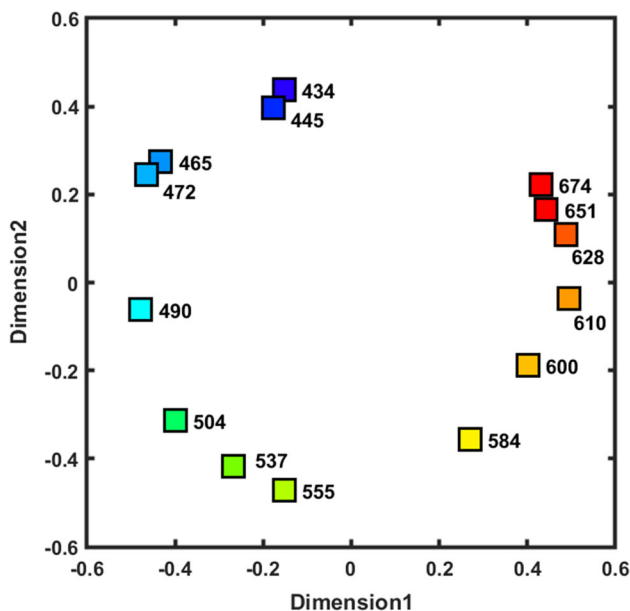

Fig. 2 Illustration of Shepard's solution [55, 56] to Ekman's colors similarities experiment [54]. Colors were generated by Spectra software, which converts the wave lengths to the RGB color system [118] (color figure online)

Table 4 Ordered reconstructing vectors with relative importance for the example with colors [54] (color table online)

| No.	Color	Wave length [nm]	Red v_{14}	Blue v_4	Green v_8	v_2	v_{10}	v_{12}	v_6	v_5	v_{11}	v_9	v_{13}	v_1	v_3	v_7
1.	■	434	1	5	1	10	1	1	2	3	1	1	1	11	5	2
2.	■	445	1	5	2	11	1	1	2	3	1	1	1	1	6	2
3.	■	465	1	9	2	1	1	1	3	6	1	1	1	1	11	2
4.	■	472	1	11	2	1	1	1	4	6	1	1	1	1	1	2
5.	■	490	1	3	4	1	1	1	7	11	1	2	1	1	1	4
6.	■	504	1	2	6	1	2	1	11	2	1	2	1	1	1	7
7.	■	537	1	2	8	1	2	1	2	2	2	3	1	1	1	11
8.	■	555	1	1	11	1	3	1	1	1	1	4	1	1	1	1
9.	■	584	3	1	1	1	7	4	1	1	5	11	3	1	1	1
10.	■	600	4	1	1	1	11	6	1	1	8	1	5	2	1	1
11.	■	610	7	1	1	2	1	9	1	1	11	1	7	2	1	1
12.	■	628	8	1	1	2	1	11	1	1	1	1	10	2	1	1
13.	■	651	9	1	1	2	1	3	1	1	1	1	11	2	2	1
14.	■	674	11	1	1	2	1	3	1	1	1	1	4	3	1	1
First rank-one relative importance [%]			63	59	58	53	51	59	55	57	51	50	61	51	51	55
Relative importances [%]			63	13	7	4	3	2	2	1	1	1	1	1	0.6	0.4
Cumulative relative importances [%]			63	76	83	87	90	92	94	95	96	97	98	99	99.6	100

the stress value around 0.05 corresponds to a good-quality solution (10% fair, 20% poor). Naturally, adding another dimension improves the solution's quality, but then the problem with interpreting dimensions is becoming even bigger since neither the color saturation, nor lightness (brightness) was controlled in this experiment. Our solution based on three attributes exhibits comparable degree of initial values reconstruction as Shepard's two-dimensional or Ekman's five-dimensional solutions and allows for analyzing the experimental data from a different point of view.

4.2 Perceived nations' similarities

The original data from the Kruskal and Wish example [57], p. 31 were collected from a group of 18 students who rated each pair of 12 countries (Brazil, Congo, Cuba, Egypt, France, India, Israel, Japan, China, former times USSR, USA, and former times Yugoslavia) on a scale from 1 "very different" to 9 "very similar." The authors proposed the following interpretation of the three obtained

dimensions: I—Political alignment (noncommunist–communist), II—Economical development (developing–developed), and III—Geography and culture (East–West). A nonmetric MDS with a standard stress value as a goal function performed for this three-dimensional solution in a MATLAB 7.11.0 (R2010b) version provided data illustrated in Fig. 3.

For the purposes of our approach, the original similarity data were rounded and rescaled such that the minimal value equals one, whereas the maximal scale value amounts to seven. They are given in "Appendix 3." Ordered reconstructing vectors provided by our heuristic algorithm are presented in Table 5.

The results of our analysis provide an alternative solution to the one obtained by nonmetric MDS. The first four reconstructing vectors from Table 5 are able to reconstruct the initial nations' similarities in 89.2%, and they may be interpreted as follows: communist or former communist countries (v_3): Cuba, China, Yugoslavia, USSR; countries having nuclear weapons (v_{10}): USSR, France, USA; developing countries (v_6): Congo, Egypt, India, Cuba;

closely cooperating countries in military and economic areas (v_7): USA, Israel, Japan. The military associations seem to be quite justifiable since the study was conducted at the time when the issues regarding the cold war, armament race, or Vietnam War were very popular and constantly present in various media.

The proposed interpretations of dimensions in the nonmetric MDS solution generally seem to be correct; however, if we take closer look on the presented data we would see that the data are sometimes difficult to interpret within the Euclidean space. For instance, China seems to be decidedly more communist than USSR and Cuba; or Japan is significantly less noncommunist than Brazil, which seems to be even more anticommunist than the USA.

The axes rotations do not make the interpretations much easier. It should also be noticed that the standard stress value for this three-dimensional MDS solution amounts to 0.1044; only after adding the fourth dimension, the stress reaches 0.049, which is deemed as a good-quality solution according to [47].

Given the above, it is not clear whether participants in this experiment judged similarities using the dimensional approach. As it can be seen from our analysis, the underlying nations' similarities structure could equally come from the way of thinking resembling the max–min operations performed on the identified attributes.
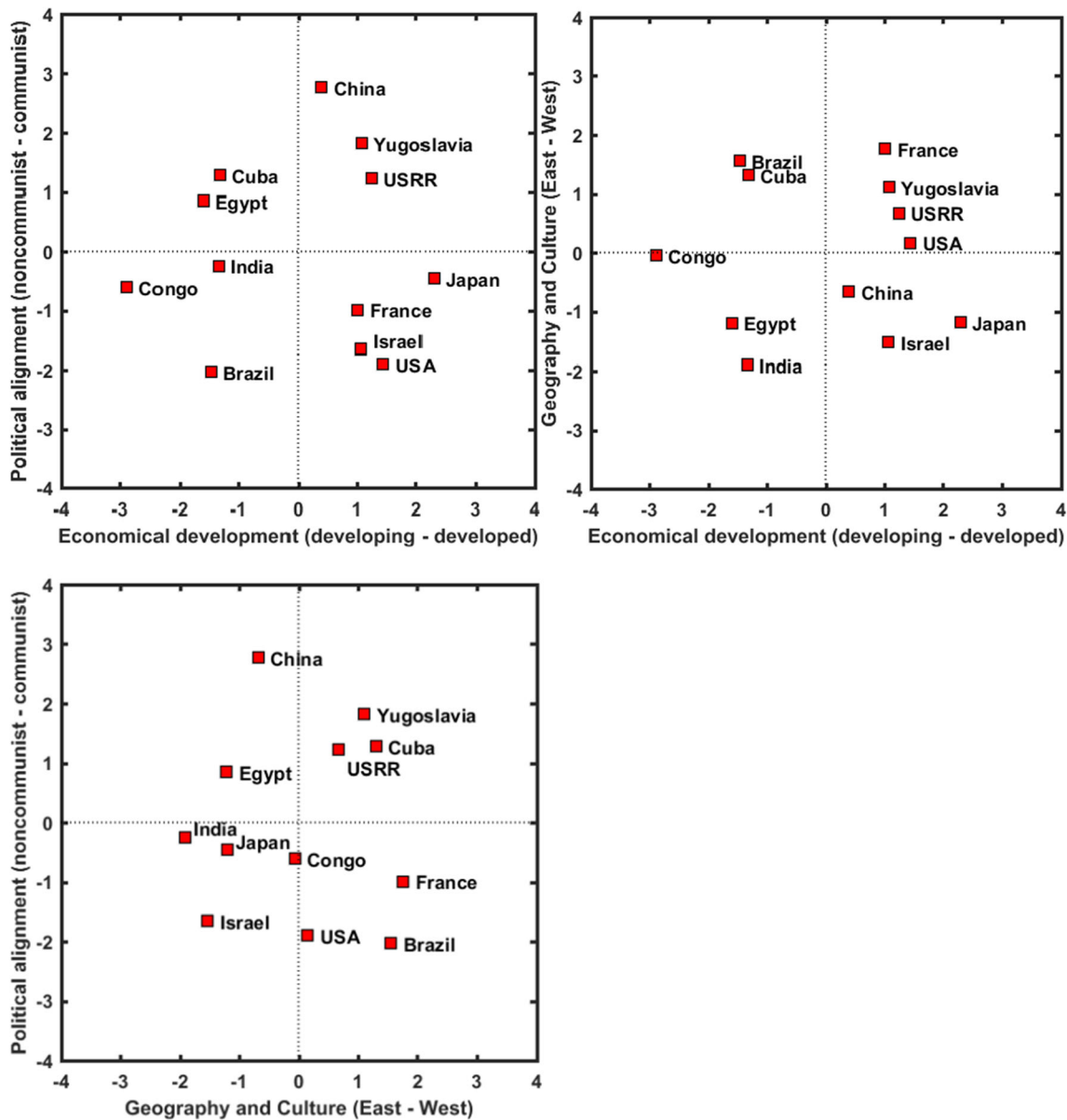

Fig. 3 Nonmetric MDS three-dimensional solution for a nations' similarities example of [57, p. 31]

Table 5 Ordered reconstructing vectors (neuromatrix) with relative importance for the example with nations' similarities [57, p. 31]

No.	Nation	v_3	v_{10}	v_6	v_7	v_9	v_1	v_5	v_{12}	v_2	v_4	v_8	v_1
2.	Congo	2	1	4	1	1	4	1	2	7	1	1	1
4.	Egypt	3	2	5	2	3	2	2	3	4	7	2	2
6.	India	2	2	7	3	4	4	2	2	2	2	2	2
3.	Cuba	7	2	4	2	2	4	2	3	4	2	2	2
9.	China	5	1	1	1	7	1	1	1	3	1	1	1
12.	Yugoslavia	4	3	3	2	3	2	2	7	3	3	2	2
10.	USSR	4	7	2	2	5	2	2	6	2	3	2	2
5.	France	3	4	2	2	2	2	7	4	3	4	3	5
11.	USA	1	4	1	5	2	1	5	1	1	2	5	7
7.	Israel	2	3	3	7	2	3	3	3	2	4	4	3
8.	Japan	2	3	2	4	4	2	3	2	2	3	7	3
1.	Brazil	1	2	2	3	1	7	4	2	1	2	3	4
Rank-one relative importance [%]		67.3	65.8	65.8	64.2	65.8	63.5	64.2	66.5	65.8	65	64.2	64.2
Relative importances [%]		67.3	12.3	5.8	3.8	2.7	1.9	1.5	1.2	0.8	0.4	0.4	0
Cumulative relative importances [%]		67.3	79.6	85.4	89.2	91.9	93.8	95.3	96.5	97.3	97.7	98.1	98.1

5 Simulation studies of the proposed algorithm

5.1 Brute force simulations

We applied the brute force simulations to extend our knowledge about the nature of the problem, its sophistication degree, and to provide a basis for comparison with our heuristics. We assumed that the number of reconstructing vectors is equal to the number of objects; thus, the reconstructing matrix has the same dimensions as the initial data array.

The number of all possible solutions (n_{sol}) depends on the number of analyzed objects (n_{obj}) and the number of LOS items employed (n_{LOS}) and can be calculated in the subsequent way:

$$n_{sol} = n_{LOS}^{n_{obj}^2 - n_{obj}}.$$

Table 6 contains the number of variations for various matrix dimensions and number of LOS items. The conditions in italics presented in Table 6 were analyzed by a brute force algorithm. For those cases, we generated 1000 random squared and symmetric similarity matrices with LOS data and verified all possible variations of max–min

products for each of them. Additionally, we examined the possibility of reconstructing matrices bigger than 5×5 with various n_{LOS} by taking advantage of less number of vectors than the initial matrix dimensions. We confined only to such combinations of n_{LOS}, matrix size, and number of reconstructing vectors when the number of variations was lower than 16 million. Other conditions were not analyzed since they required too much computation time. The results of our simulations are put together in Table 7 and show percentages of fully decomposed matrices (PDF) for each combination of a matrix size, n_{LOS}, and the number of reconstructing vectors applied.

For the following matrices: $n_{LOS} = 2$, $n_{obj} = 11$–20; $n_{LOS} = 3$, $n_{obj} = 8$–13; $n_{LOS} = 4$, $n_{obj} = 7$–11; $n_{LOS} = 5$, $n_{obj} = 6$–10; $n_{LOS} = 6$, $n_{obj} = 5$–9; $n_{LOS} = 8$, $n_{obj} = 5$–8, the decomposition by one reconstructing vector was not found. For these matrices, bigger numbers of vectors were not analyzed due to a great number of variations and unacceptable time of simulations.

The presented brute force results show that the 3×3 matrix size and the full reconstruction were possible for all randomly generated matrices and analyzed n_{LOS} (from 2 up to 10) by 3 or 2 reconstructing vectors. Matrices 4×4 were fully reconstructed at all times by 4 vectors for

Table 6 Number of variations when the number of reconstructing vectors is the same as the number of objects

n_{LOS}	Matrix size		
	3×3	4×4	5×5
2	*64*	*4096*	*1,048,576*
3	*729*	*531,441*	3.5×10^{09}
4	*4096*	*16,777,216*	1.1×10^{12}
5	*15,625*	2.4×10^{08}	9.5×10^{13}
6	*46,656*	2.2×10^{09}	3.7×10^{15}
7	*117,649*	1.4×10^{10}	8×10^{16}
8	*262,144*	6.9×10^{10}	1.2×10^{18}
9	*531,441*	2.8×10^{11}	1.2×10^{19}
10	*1,000,000*	10^{12}	10^{20}

Conditions in italics were analyzed by a brute force algorithm

n_{LOS} – number of LOS items

n_{LOS} = 2–4. It was not possible to check whether this result is also true for larger n_{LOS} since the n_{sol} was too big. The most interesting result was obtained for the matrix 5×5 consisting of values represented on a binary scale (n_{LOS} = 2, cf. Table 7). It occurred that even 5 vectors are not enough to fully reconstruct the initial matrix in all cases.

5.2 Performance of the proposed heuristic algorithm

The effectiveness of our heuristic procedure was examined by performing a simulation experiment. We applied our approach to randomly generated matrices with various combinations of their dimensions and scales. We examined two factors: the matrix dimension and the rating scale type. Matrix sizes ranged from 3×3 up to 13×13 (11 levels), whereas n_{LOS} varied from 2 up to 10 (9 levels). The applied within-subjects design produced 99 different experimental conditions. For every combination of the matrix size and n_{LOS}, we generated 1000 symmetric matrices. We tried to

decompose each of the total 99,000 random matrices by our heuristic procedure. We recorded the highest percentage of the initial matrix reconstruction obtained by applying both versions of the tuning-up procedure. Matrices 3×3 and 4×4 with n_{LOS} from 2 to 10 were fully decomposed for all randomly generated matrices. The remaining results of our simulations presenting PDF and mean percentages of initial matrix reconstructions (MPR) are put together in Table 8.

The simulations show that our heuristic procedure was able to decompose all randomly generated small matrices. This result is the same as applying brute force algorithm. For more complex data, the percentages of full reconstructions gradually decrease. For matrices bigger than 6×6 and n_{LOS} bigger than 3, the possibility of the full decomposition drops radically far below 50%. If the matrix is bigger than 8×8 and n_{LOS} is bigger than 2, then there is almost no chances of getting the full decomposition by our approach. We additionally performed simulations for 14×14 and 15×15 matrices including data with n_{LOS} = 2. In these cases, none of randomly generated similarities' matrices was fully decomposed by our algorithm. However, it is not known whether such full decompositions exist at all for those similarity data.

One should notice that the mean percentage of reconstruction is quite high. Even for the most difficult experimental condition, the value is higher than 90%. The mean percentage of reconstruction for 14×14 and 15×15 matrices with n_{LOS} = 2 amounted to 92.6% (2.33) and 92.3% (2.33), respectively.

6 Applications

The presented method may be widely used in these fields of science in which human perception is involved to judge directly about similarities. It could be especially suited and practically applied in situations where humans are judging objects' similarities in pairwise comparisons, using ordinal

Table 7 Percentages of fully decomposed 1000 random similarity matrices by a brute force algorithm

Matrix size	No of reconstructing vectors																																	
	n_{LOS} = 2					n_{LOS} = 3				n_{LOS} = 4				n_{LOS} = 5			n_{LOS} = 6			n_{LOS} = 7			n_{LOS} = 8			n_{LOS} = 9			n_{LOS} = 10		
	1	2	3	4	5	1	2	3	4	1	2	3	4	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3			
3×3	49	100	100	×	×	32	100	100	×	25	100	100	×	21	100	100	15	100	100	14	100	100	13	100	100	12	100	100	9.2	100	100			
4×4	10	51	94	100	×	3.3	37	92	100	2.1	32	92	100	0.9	25	89	0.4	17	×	0.3	17	×	0.4	13	×	0.4	13	×	0.2	12	×			
5×5	1.4	14	55	91	99	0.2	4.6	36	×	0	2.5	×	×	0	1.2	×																		
6×6	0	1.5	15	50	×	0	0.2	×	×	0	0	×	×																					
7×7	0	0	2.7	×	×	0	0	×	×																									
8×8	0	0	×	×	×																													
9×9	0	0	×	×	×																													
10×10	0	0	×	×	×																													

Table 8 Results of the decompositions obtained by our algorithm for 1000 random similarity matrices

n_{LOS}	Matrix size																	
	5 × 5		6 × 6		7 × 7		8 × 8		9 × 9		10 × 10		11 × 11		12 × 12		13 × 13	
	PFD	MPR	PFD	MPR	PFD	MPR	PFD	MPR	PFD	MPR	PFD	MPR	PFD	MPR	PFD	MPR	PFD	MPR
2	95.5	99.8 (1.03)	83.2	99.2 (1.90)	68.7	98.7 (2.19)	43.6	97.5 (2.75)	26.8	96.8 (2.70)	9.0	95.4 (2.64)	2.5	94.5 (2.60)	0.6	93.8 (2.45)	0.1	93.1 (2.46)
3	93.9	99.7 (1.28)	75.1	99.0 (1.98)	45.4	97.8 (2.59)	20.9	96.4 (2.87)	5.5	95.0 (3.96)	0.9	93.9 (2.69)	0.2	92.7 (2.73)	0	92.2 (2.66)	0	91.5 (2.54)
4	91.9	99.7 (1.30)	66.4	98.7 (2.26)	35.2	97.4 (2.65)	12.0	95.8 (2.95)	2.9	94.4 (2.87)	0.1	93.3 (2.61)	0	92.5 (2.66)	0	91.7 (2.52)	0	91.2 (2.45)
5	91.0	99.7 (1.23)	66.4	98.7 (2.23)	28.1	97.2 (2.55)	7.6	95.5 (2.82)	1.0	94.1 (2.78)	0	93.0 (2.80)	0	92.0 (2.63)	0	91.4 (2.39)	0	90.8 (2.29)
6	89.7	99.6 (1.34)	59.7	98.6 (2.14)	26.1	97.1 (2.66)	6.1	95.3 (2.89)	1.1	94.0 (2.76)	0	92.8 (2.64)	0	92.0 (2.52)	0	91.3 (2.39)	0	90.6 (2.31)
7	88.1	99.6 (1.26)	57.4	98.4 (2.42)	22.2	96.8 (2.74)	5.1	95.4 (2.87)	0.6	93.9 (2.83)	0.2	92.9 (2.64)	0	91.9 (2.58)	0	91.1 (2.54)	0	90.4 (2.32)
8	87.7	99.6 (1.44)	51.9	98.2 (2.41)	18.7	96.9 (2.75)	4.4	95.3 (2.86)	0.4	93.7 (2.72)	0	92.7 (2.65)	0	91.7 (2.63)	0	91.1 (2.38)	0	90.3 (2.34)
9	86.0	99.5 (1.46)	53.6	98.5 (2.24)	18.9	96.9 (2.70)	2.6	95.1 (2.87)	0.2	93.8 (2.85)	0	92.7 (2.73)	0	91.8 (2.59)	0	91.0 (2.37)	0	90.0 (2.37)
10	86.7	99.5 (1.48)	55.0	98.5 (2.27)	16.5	96.7 (2.71)	2.3	95.2 (2.78)	0.3	93.8 (2.73)	0	92.5 (2.68)	0	91.8 (2.43)	0	90.9 (2.33)	0	90.1 (2.25)

scales, such as the Likert's or linguistic ones. They may be, for instance, concerned with benchmark studies in management and marketing like in the HATCO company comparison with nine competitors [72], psychological studies regarding cognition or memory [78–80], or psychological analyses of human perception like in the research on differences between adults and children in recognizing body parts similarities [81].

Our proposal assumes that people apply max–min approach on attributes. We demonstrated that in some experimental data applying this method may provide different explanations than dimensional approaches based on Euclidean spaces which are used in classic nonmetric MDS or other FA-like methods.

The suggested, nonstatistic, and nonmetric approach can be applied to any ordinal variables where consecutive values correspond to the perceived level of similarity intensity. A granularity resolution, i.e., scale range, can be freely chosen as it neither affects the theoretical basis nor the described heuristic algorithm. The number of expressions describing similarity levels may, in practice, refer to experts' sensitivity to tell similarities between objects in a specific context.

Reconstructing vectors from our proposal might be interpreted as counterparts of the fuzzy formal contexts. One can apply formal concept analysis to these results for further detailed examination of relations between factors and objects. However, even not using formal analysis, our reconstructing vectors allow for retracting knowledge about factors shaping similarity assessments. It seems particularly important to look for such hidden relations in various practical areas. For example, factors affecting the perception and cognition of visual stimuli are a particular focus of marketing and neuromarketing (see [82–87]). Specifying reliable objects–attributes relations that underlie similarity perception could facilitate the design of more effective graphic marketing messages used in product packages that will stand out from their competitors or advertisements.

7 Future studies

There are some research questions regarding our proposal that should be answered in future studies. They concern detailed properties, practical usefulness, and methodological correctness.

7.1 Algorithmic extensions

Some prospective studies can enclose simulation research on markedly more complex examples, with numerous attributes contributing to peoples' perceptions. For such

significantly more sophisticated problems with a large number of possible factors, one could try to combine our algorithm with the fuzzy formal concept analysis. From the point of view of the described model per se, investigations regarding the influence of its properties on the generated results are of special interest. For instance, how the granularity of a similarity intensity scale impacts the results quality for various tasks.

Possible future research may also focus on applying other than our approaches for determining reconstructing vectors such as deep learning. Some initial works in this regard have already been performed for the Boolean matrix factorization. Frolov et al. [88] used Hopfield-like neural networks for this purpose. They further extended their ideas in a series of subsequent publications [89–91]. Input and output matrices of these approaches are completely different conceptually than our proposal. Thus, direct comparisons are not possible. However, creating a deep learning algorithm similar to the solutions provided by Frolov and colleagues for similarity matrices containing ordinal-scale values would be very interesting.

Another direction of extending or improving our algorithm could utilize some latest developments in optimization research. Though some analyses emphasize the complexity of problem of similar types (e.g., [92]), one may try, for instance, to take advantage of interesting extensions of classical desirability function optimizations, especially the approach involving max- and min-type functions (cf. [93–95]).

7.2 Theoretical advancements

From the theoretical perspective, answering the question under what circumstances the initial objects similarity matrix can be fully decomposed is certainly worth pursuing. A possible conjecture is that such a full decomposition is feasible for matrices where the classic triangle inequality between all similarity ratings is conserved.

A number of future research directions and projects could include the development of a methodology that would deal with multiple similarity matrices. This, probably shall involve some clustering techniques. For the presented approach, one should try to determine in which situations either liberal, conservative models shall be applied. Moreover, an extension of our proposal may include other types of theoretical logical relations or their combinations that have not been demonstrated in this paper.

There could be a potential trade-off concerned with the specific feature of our approach that decreases the significance of smaller attribute intensities. This might increase the danger of excluding some factors from the final reconstructing matrix. The problem may constitute a limitation of our approach, and some further studies may focus on determining whether such a phenomenon exists, and if so, what the scale is.

In recent decades, fuzzy sets have been intensively used as a tool for uncertainty modeling. It results, among others, from the nature of phenomena in many areas of interest in science. It is especially the case in mathematical modeling of both social and technical issues. The framework presented in this research allows for advanced analytical constructions that take into account uncertainty. For example, linguistic expressions of membership function values or degrees of truth can be represented as fuzzy sets. Then, the analysis of similarity relationships would be based on fuzzy sets of second type. Intensive work on extending classic approaches to develop models involving uncertainty can be also seen in construction of algorithms for solving differential equations. For example, in the works [96–99], authors propose algorithms that operate on fuzzy numbers and use the theory of kernel reproduction. Processing and modeling of imprecision involving similar approaches seems to be an interesting approach in analyses related to human behavior. Hence, the cited works suggest a possible direction of more advanced studies in this field.

Another possible extension could involve definitions of the acceptation, rejection, and uncertainty levels of similarity while considering given factors or decompositions. Such modifications are conceptually interesting, despite the danger of decreasing the precision of the similarity matrix reconstruction (cf. [40]).

7.3 Validation studies

A significant problem concerned with validation of the presented methodology is that it is not clear what *true* human being's assessments of similarities are, and how to identify them. The question is what should be the reference? Should this be the results of nonmetric MDS, the structure of neurons activations in the brain? How should these data be aggregated?—Simple average, Choquet integral, etc.

Moreover, in a variety of factor analysis types, one can come up with a number of qualitatively different solutions having similar goodness-of-fit parameters. In a classic factor analysis, it can be done by simple algebraic rotations of eigenvectors. At the end of the day, it is up to the investigator to select the most logical, theoretically and practically justifiable solution, from among those possible. Nevertheless, even without knowing the real true structure of psychological concepts and relations between them, these approaches are useful. They are employed very intensively as they provide, at least, some insight into psychological constructs.

Despite the above-mentioned problems, there are some ways of increasing the belief that the proposal is valid. We showed on two simple, but well-known examples that the idea works, that there is a theoretical background related to how people may think about similarities. The theory corresponds logically to the problem we are trying to solve.

From the technical point of view, the put-forward algorithm for seeking the LOS neuromatrices provided satisfactory solutions both for real experimental examples and in decomposing randomly generated data. This may be considered as a part of an internal validation. The procedure provides reasonable solutions that can be interpreted logically, but qualitatively differently from the case of classic approaches. These outcomes can be treated as a basic external validation.

Naturally, some additional steps may be taken to further validate our proposal. Given the paper length restrictions, other, more detailed and comprehensive attempts can be performed in subsequent studies. In the perspective of neuroscience, future research could be directed to empirical investigations at the biological level focused on checking if the proposed assumptions are reflected in the brain structures' activities. Finding patterns of brain functioning could verify the suggested approach at a physiological level. These investigations could take advantage of brain imaging methods such as PET or fMRI, that occurred to be very helpful in understanding of various neural-based processes (cf., e.g., [100, 101]).

It seems that making decisions about similarities based on hidden factors (neuromatrices) may also influence visual processing strategies. Using considerably technologically improved and significantly less expensive eye tracking techniques to study similarity issue is reasonable. The human visual behavior can be characterized by spatiotemporal oculometric parameters such as saccades and fixations. Their dynamics may be modeled by hidden Markov models (cf. [102–108]), their fuzzy equivalents (see [109–111]), or even Markov switching models (cf. [112, 113]).

It can be attractive from the cognitive and behavioral point of view to conduct experiments focused on determining relations between visual strategies and mechanisms of making decisions about similarities modeled by neuromatrices introduced in this paper. Analogous studies may also be conducted by other tools used in neuroscience such as electroencephalography [114, 115], magnetoencephalography, and facial recognition. Reviews and discussion on the usefulness of various neuroscience methods can be found, for example, in [116, 117].

Some more classic methods may also be applied for extended validation purposes. They may include studies aimed at finding correlations and relations with psychological constructs with known values of similarities between objects or concepts of various types. One may even conduct simulations where similarities would be generated automatically based on objects' physical properties, such as calculated stick lengths. Another new idea is to include retrospective thinking in the experimental procedure. In such a case, after assessing similarities, participants explain how they evaluated them, what factors were the most important, and in what way objects' or concepts' features influenced subjects' perception.

8 Conclusion

In this paper, we propose a new approach for modeling human thinking about objects' similarities by searching the neuromatrix. The neuromatrix consists of vectors that reconstruct the LOS similarity data matrix. We search for internal data structure using fuzzy-set theory simple operators. We have shown that given such assumptions, it is possible to represent the initial matrix as the union of quasi-rank-one matrices.

Conceptually, we presented a model that reflects cognitive processes taking place in the brain. The process is demonstrated in the form of linguistic expressions—provided by a human—about similarities between objects. Uncovering hidden objects–attributes relations in the form of neuromatrices is an attempt of specifying factors shaping decisions about those similarities. From the neuroscientific point of view, we assume that our approach discloses components of the cognitive perception mechanism occurring at the neural level.

Technically, our scientific proposal is very closely related to the idea that underpins PCA, FA, metric and nonmetric MDS and correspondence analysis, namely the eigendecomposition of a matrix of real values. Analogously to classic approaches, we have attempted to represent the initial LOS data matrix as a max–min product of some other LOS matrix, called a neuromatrix, and its transpose. In our opinion, such an approach is appropriate for linguistic based or ordinal data. In the max–min multiplication, the product of two values is replaced by the intersection (min) operator, whereas the summation takes the form of union (max values).

Our methodology is theoretically well grounded and soundly logically justified. The idea is very simple and is based solely on max–min operators. We think that such an approach is closer to real human thinking about similarities than other methods. The max–min operations are truly natural as they are in the inner core of humans' negotiations, games, optimization, orientation and gaining of perspective. Moreover, our methodology has a tendency to relatively diminish the significance of smaller attributes' intensities. This, in turn, fits well to human brain biological

functioning where neurons are activated and pass electrical impulse further only if superposition of input signals exceeds a specific, minimal level. This property of our methodology seems to be better suited to physiological structures of the neurons than other approaches. Such a way of handling similarities is in concordance with intuition. Usually, weights of factors deciding about similarities are nonlinear—some of them are more important than the others. Those less important ones, probably, have less impact on the overall perception.

Below, we present a brief list characterizing our proposal:

- The theoretical foundations are in concordance with the human thinking about similarities and basic neural activity.
- The similarity matrix includes only values measured on an ordinal scale.
- It is possible to decompose the initial objects' similarities matrices.
- The decomposition uses only max and min operators and provides reconstructing vectors that can be interpreted akin to classic approaches.
- The proposal provides logical results; however, their interpretation might differ from classic approaches, which allows for better understanding of the examined phenomenon.
- Brute force simulation results show that starting from the 5×5 matrices it is not possible to fully decompose random objects' similarity matrices.
- Simulation results proved that the proposed heuristic procedure is able to reconstruct randomly generated similarity matrices to a very high extent, even for the most difficult examined cases.

We hope that our proposal would be interesting to other researchers that could apply it in a variety of contexts and extend in various directions including those that are indicated in this paper.

Compliance with ethical standards

Appendix 1: Constructing LOS similarities by the max–min operations

In our approach, the standard formula for matrix multiplication for every i and j takes the following shape:

$$c_{ij} = \max\big\{\min\{a_{i1}, b_{1j}\}, \min\{a_{i2}, b_{2j}\}, \ldots, \min\{a_{in}, b_{nj}\}\big\}. \tag{15}$$

Replacing the max with $\cup$—, and the min with—$\cap$ symbols, the formula may be rewritten as

$$c_{ij} = \big(a_{i1} \cap b_{1j}\big) \cup \big(a_{i2} \cap b_{2j}\big) \cup \cdots \cup \big(a_{in} \cap b_{nj}\big)$$
$$= \bigcup_{r=1}^{n} a_{ir} \cap b_{rj}. \tag{16}$$

For shorthand, we denote the set theory multiplication by the max–min operator of the two LOS matrices as:

$$\mathbf{C}_{\text{LOS}} = \mathbf{A}_{\text{LOS}} \circ \mathbf{B}_{\text{LOS}}. \tag{17}$$

The transpose of the LOS data array is defined in the same way as for regular matrices, that means

$$\mathbf{B}_{\text{LOS}} = \mathbf{A}_{\text{LOS}}^{\text{T}} \text{ if } b_{ij} = a_{ji}, \text{ for each } i \text{ and } j. \tag{18}$$

To illustrate the introduced symbols, let us consider a simple 3×3 matrix that contains assessments of relative similarities between three objects X, Y, Z expressed on LOS {low, med, big}. We additionally assume that each object is very similar to itself and that if X is similar to Y to some degree, then the similarity between Y and X is exactly the same. Thus, we obtain the symmetric matrix with the maximal similarity values at the diagonal, e.g.,

$$\mathbf{E}_{\text{LOS}} = \begin{bmatrix} big & low & big \\ low & big & med \\ big & med & big \end{bmatrix}. \tag{19}$$

The above matrix may be represented by the max–min product of the following square (but not symmetric) matrix and its transpose:

$$\mathbf{V}_{\text{LOS}} = \begin{bmatrix} big & low & low \\ low & big & med \\ big & low & med \end{bmatrix}. \tag{20}$$

In this paper, we call this idea LOS matrix decomposition, and the column vectors of the LOS matrix—the reconstructing vectors (neuromatrix)—the equivalent of

eigenvectors of classical algebra. In this example, we can check that:

$$\mathbf{V}_{\text{LOS}} \circ \mathbf{V}_{\text{LOS}}^{\text{T}} = \begin{bmatrix} big & low & low \\ low & big & med \\ big & low & med \end{bmatrix}$$

$$\circ \begin{bmatrix} big & low & big \\ low & big & low \\ low & med & med \end{bmatrix}$$

$$= \begin{bmatrix} big & low & big \\ low & big & med \\ big & med & big \end{bmatrix} = \mathbf{E}_{\text{LOS}};$$

for example:

$$e_{\text{LOS}(31)} = \max\{\min\{big, big\}, \min\{low, low\}, \min\{med, low\}\}$$
$$= \max\{big, low, low) = big, \qquad (21)$$

$$e_{\text{LOS}(23)} = \max\{\min\{low, big\}, \min\{big, low\}, \min\{med, med\}\}$$
$$= \max\{low, low, med\} = med, \qquad (22)$$

and so on, for every i and j.

To use the approach for dimensionality reduction and searching for the data structure, it is necessary to check whether it is possible to represent the initial matrix as a union (the equivalent of summation in set theory) of outer products of the LOS vectors. In other words, to perform a quasi-rank-one decomposition using only the set theory operators plus the transpose. Let the union (max) operator for two LOS matrices be defined as follows:

$$\mathbf{C}_{\text{LOS}} = \mathbf{A}_{\text{LOS}} \cup \mathbf{B}_{\text{LOS}}, \qquad (23)$$

where $c_{ij} = \max\{a_{ij}, b_{ij}\}$ for every i and j.

It can be shown that if for the given LOS matrix $\mathbf{C}_{\text{LOS}}$ there exists such a LOS matrix $\mathbf{V}_{\text{LOS}}$ that

$$\mathbf{V}_{\text{LOS}} \circ \mathbf{V}_{\text{LOS}}^{\text{T}} = \mathbf{C}_{\text{LOS}}, \qquad (24)$$

then the initial matrix $\mathbf{C}_{\text{LOS}}$ may also be represented by a union of the matrices constructed in a similar way as the outer product in the conventional algebra, i.e.,

$$\mathbf{C}_{\text{LOS}} = \mathbf{v}_1 \circ \mathbf{v}_1^{\text{T}} \cup \mathbf{v}_2 \circ \mathbf{v}_2^{\text{T}} \cup \ldots \cup \mathbf{v}_n \circ \mathbf{v}_n^{\text{T}} = \bigcup_{k=1}^{n} \mathbf{v}_k \circ \mathbf{v}_k^{\text{T}}, \qquad (25)$$

where

$$\mathbf{v}_1 = \begin{bmatrix} w_{11} \\ w_{21} \\ \vdots \\ w_{n1} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} w_{12} \\ w_{22} \\ \vdots \\ w_{n2} \end{bmatrix}, \ldots, \mathbf{v}_k = \begin{bmatrix} w_{1k} \\ w_{2k} \\ \vdots \\ w_{nk} \end{bmatrix}, \ldots, \mathbf{v}_n$$

$$= \begin{bmatrix} w_{1n} \\ w_{2n} \\ \vdots \\ w_{nn} \end{bmatrix}, \qquad (26)$$

are the column vectors of $\mathbf{V}_{\text{LOS}}$ of size n.

Proof of (25) First, according to the product formula from (16) for a LOS data matrix, each component of

$$\mathbf{C}_{\text{LOS}} = \mathbf{V}_{\text{LOS}} \circ \mathbf{V}_{\text{LOS}}^{\text{T}} \qquad (27)$$

may be calculated as:

$$c_{ij} = \bigcup_{k=1}^{n} w_{ik} \cap w_{kj}^{\text{T}}, \qquad (28)$$

where w_{ik} is an item of the $\mathbf{V}_{\text{LOS}}$ array. Because of (18), we obtain:

$$c_{ij} = \bigcup_{k=1}^{n} w_{ik} \cap w_{jk}. \qquad (29)$$

On the other hand, we state that $\mathbf{C}_{\text{LOS}}$ can be represented by a union of outer products of LOS reconstructing vectors (cf. (25)). Let us denote the individual outer product array as

$$\mathbf{C}_k = \mathbf{v}_k \circ \mathbf{v}_k^{\text{T}}. \qquad (30)$$

Since $\mathbf{v}_k$ is a column vector and $\mathbf{v}_k^{\text{T}}$ is a row vector, there is only one component from (16) involved in obtaining every outer product item for the kth array, namely:

$$w_{ik} \cap w_{kj}^{\text{T}}. \qquad (31)$$

Due to (18), (31) is equivalent to

$$w_{ik} \cap w_{jk}. \qquad (32)$$

Now, making a union of all of the outer products, and taking advantage of (23) we arrive at the following formula, identical to (29):

$$c_{ij} = \left(w_{i1} \cap w_{j1}\right) \cup \left(w_{i2} \cap w_{j2}\right) \cup \cdots \cup \left(w_{in} \cap w_{jn}\right)$$
$$= \bigcup_{k=1}^{n} w_{ik} \cap w_{jk},$$

which was to be demonstrated.

For the example presented in (20), the outer product decomposition takes the following form:

$$
\left(\begin{bmatrix} big \\ low \\ big \end{bmatrix} \circ \begin{bmatrix} big & low & big \end{bmatrix} \right)
$$

$$
\cup \left(\begin{bmatrix} low \\ big \\ low \end{bmatrix} \circ \begin{bmatrix} low & big & low \end{bmatrix} \right)
$$

$$
\cup \left(\begin{bmatrix} low \\ med \\ med \end{bmatrix} \circ \begin{bmatrix} low & med & med \end{bmatrix} \right)
$$

$$
= \begin{bmatrix} big & low & big \\ low & low & low \\ big & low & big \end{bmatrix} \cup \begin{bmatrix} low & low & low \\ low & big & low \\ low & low & low \end{bmatrix}
$$

$$
\cup \begin{bmatrix} low & low & low \\ low & med & med \\ low & med & med \end{bmatrix} = \begin{bmatrix} big & low & big \\ low & big & med \\ big & med & big \end{bmatrix} = \mathbf{E}_{\mathrm{LOS}}.
$$

$$(33)$$

Appendix 2: The similarity S array for the example with colors

The original data from [54] were rounded and rescaled such that the minimal value equals one, whereas the maximal amounts to eleven. Colors in the second column were generated by Spectra software, which converts the wavelengths to the RGB color system [118].

Appendix 3: The similarity array S for the example with nations

The original data from [57, p. 31] were rounded and rescaled such that the minimal value equals one, while the maximal scale value amounts to seven.

No	Nation	1.	2.	3	4.	5.	6.	7.	8.	9.	10.	11.	12.
1.	Brazil	7	4	4	2	4	4	3	3	1	2	4	2
2.	Congo	4	7	4	4	3	4	2	2	3	2	1	3
3.	Cuba	4	4	7	4	3	3	3	2	5	4	2	4
4.	Egypt	2	4	4	7	4	5	4	3	3	3	2	3
5.	France	4	3	3	4	7	2	3	3	3	4	5	4
6.	India	4	4	3	5	2	7	3	4	3	4	3	3
7.	Israel	3	2	3	4	3	3	7	4	2	3	5	3
8.	Japan	3	2	2	3	3	4	4	7	3	4	5	3
9.	China	1	3	5	3	3	3	2	3	7	5	2	4
10.	USSR	2	2	4	3	4	4	3	4	5	7	4	6
11.	USA	4	1	2	2	5	3	5	5	2	4	7	3
12.	Yugoslavia	2	3	4	3	4	3	3	3	4	6	3	7

No	Color	Wave length [nm]	1.	2.	3	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
1.		434	11	10	5	5	3	2	2	1	1	2	2	2	2	3
2.		445	10	11	6	5	3	2	2	2	1	1	2	2	2	2
3.		465	5	6	11	9	6	3	2	2	1	1	1	1	2	1
4.		472	5	5	9	11	6	4	2	2	1	1	1	1	1	1
5.		490	3	3	6	6	11	7	4	4	2	1	1	1	1	1
6.		504	2	2	3	4	7	11	7	6	2	2	1	1	1	1
7.		537	2	2	2	2	4	7	11	8	3	2	2	1	1	1
8.		555	1	2	2	2	4	6	8	11	4	3	1	1	1	1
9.		584	1	1	1	1	2	2	3	4	11	7	5	4	3	3
10.		600	2	1	1	1	1	2	2	3	7	11	8	6	5	4
11.		610	2	2	1	1	1	1	2	1	5	8	11	9	7	7
12.		628	2	2	1	1	1	1	1	1	4	6	9	11	10	8
13.		651	2	2	2	1	1	1	1	1	3	5	7	10	11	9
14.		674	3	2	1	1	1	1	1	1	3	4	7	8	9	11

References

1. Anderson JR (1996) ACT: a simple theory of complex cognition. Am Psychol 51(4):355–365
2. Ritter FE, Tehranchi F, Oury JD (2019) ACT-R: a cognitive architecture for modeling cognition. WIREs Cogn Sci 10(3):e1488
3. Milnes BG, Pelton G, Doorenbos R, Laird MH, Rosenbloom P, Newell A (1992) A specification of the soar cognitive architecture in Z. Carnegie Mellon University, USA, Technical Report
4. Laird JE (2012) The soar cognitive architecture. MIT Press, London
5. Kieras DE, Meyer DE (1997) An overview of the EPIC architecture for cognition and performance with application to human–computer interaction. Hum Comput Interact 12(4):391–438
6. Meyer DE, Glass JM, Mueller ST, Seymour TL, Kieras DE (2001) Executive-process interactive control: a unified computational theory for answering 20 questions (and more) about cognitive ageing. Eur J Cogn Psychol 13(1–2):123–164
7. Langley P, Laird JE, Rogers S (2009) Cognitive architectures: research issues and challenges. Cogn Syst Res 10(2):141–160
8. Thórisson K, Helgasson H (2012) Cognitive architectures and autonomy: a comparative review. J Artif Gen Intell 3(2):1–30
9. Kotseruba I, Tsotsos JK (2020) 40 years of cognitive architectures: core cognitive abilities and practical applications. Artif Intell Rev 53(1):17–94
10. Laird JE, Lebiere C, Rosenbloom PS (2017) A standard model of the mind: toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. AI Mag 38, no. 4, Art. no. 4
11. Lieto A, Bhatt M, Oltramari A, Vernon D (2018) The role of cognitive architectures in general artificial intelligence. Cogn Syst Res 48:1–3
12. Likert R (1932) A technique for the measurement of attitudes. Arch Psychol 22(140):55
13. Stevens SS (1946) On the theory of scales of measurement. Science 103:677–680
14. Franceschini PF, Galetto M, Varetto M (2004) Qualitative ordinal scales: the concept of ordinal range. Quality Engineering 16(4):515–524. https://doi.org/10.1081/QEN-120038013
15. Taghia J et al (2018) Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition. Nat Commun 9(1):2505
16. Iliadis LS, Kurkova V, Hammer B (2020) Brain-inspired computing and machine learning. Neural Comput Appl 32(11):6641–6643
17. Zhang Y, Abbas H, Li Y (2020) Cognitive computing for intelligent application and service. Neural Comput Appl 32(9):4315–4316
18. Filippov VA, Bobylev AN, Busygin AN, Pisarev AD, Udovichenko SYu (2020) A biomorphic neuron model and principles of designing a neural network with memristor synapses for a biomorphic neuroprocessor. Neural Comput Appl 32(7):2471–2485
19. Zhu X, Li X, Zhang S, Xu Z, Yu L, Wang C (2017) Graph PCA hashing for similarity search. IEEE Trans Multimed 19(9):2033–2044
20. Das H, Naik B, Behera HS (2020) A hybrid neuro-fuzzy and feature reduction model for classification. Adv Fuzzy Syst 2020:1–15
21. Gou J, Xue Y, Ma H, Liu Y, Zhan Y, Ke J (2020) Double graphs-based discriminant projections for dimensionality reduction. Neural Comput Appl
22. Peng C, Chen Y, Kang Z, Chen C, Cheng Q (2020) Robust principal component analysis: a factorization-based approach with linear complexity. Inf Sci 513:581–599
23. Sahoo TK, Banka H, Negi A (2020) Novel approaches to one-directional two-dimensional principal component analysis in hybrid pattern framework. Neural Comput Appl 32(9):4897–4918
24. Ayesha S, Hanif MK, Talib R (2020) Overview and comparative study of dimensionality reduction techniques for high dimensional data. Inf Fusion 59:44–58
25. Schreiber JB, Nora A, Stage FK, Barlow EA, King J (2006) Reporting structural equation modeling and confirmatory factor analysis results: a review. J Educ Res 99(6):323–338
26. Guo J, Marsh HW, Parker PD, Dicke T, Lüdtke O, Diallo TMO (2019) A systematic evaluation and comparison between exploratory structural equation modeling and bayesian structural equation modeling. Struct Equ Model Multidiscipl J 26(4):529–556
27. Marsh HW, Guo J, Dicke T, Parker PD, Craven RG (2020) Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and Set-ESEM: optimal balance between goodness of fit and parsimony. Multivar Behav Res 55(1):102–119
28. Liang X, Yang Y, Cao C (2020) The performance of ESEM and BSEM in structural equation models with ordinal indicators. Struct Equ Model Multidiscipl J 1–14
29. Jöreskog KG, Moustaki I (2001) Factor analysis of ordinal variables: a comparison of three approaches. Multivar Behav Res 36(3):347–387
30. Katsikatsou M, Moustaki I, Yang-Wallentin F, Jöreskog KG (2012) Pairwise likelihood estimation for factor analysis models with ordinal data. Comput Stat Data Anal 56(12):4243–4258
31. Katsikatsou M, Moustaki I (2016) Pairwise likelihood ratio tests and model selection criteria for structural equation models with ordinal variables. Psychometrika 81(4):1046–1068
32. Papageorgiou I, Moustaki I (2019) Sampling of pairs in pairwise likelihood estimation for latent variable models with categorical observed variables. Stat Comput 29(2):351–365
33. Revuelta J, Maydeu-Olivares A, Ximénez C (2019) Factor analysis for nominal (first choice) data. Struct Equ Model Multidiscipl J 1–17
34. Belohlavek R, Vychodil V (2010) Discovery of optimal factors in binary data via a novel method of matrix decomposition. J Comput Syst Sci 76(1):3–20
35. Boeck PD, Rosenberg S (1988) Hierarchical classes: model and data analysis. Psychometrika 53(3):361–381
36. Wilderjans TF, Ceulemans E, Mechelen IV (2012) The SIM-CLAS model: simultaneous analysis of coupled binary data matrices with noise heterogeneity between and within data blocks. Psychometrika 77(4):724–740
37. Leenen I, Mechelen IV, Boeck PD (2001) Models for ordinal hierarchical classes analysis. Psychometrika 66(3):389–403
38. Ganter B, Glodeanu CV (2012) Ordinal factor analysis. In: Domenach F, Ignatov DI, Poelmans J (eds) Formal concept analysis. Springer, Berlin, pp 128–139
39. Belohlavek R (2000) Similarity relations in concept lattices. J Logic Comput 10(6):823–845
40. Belohlavek R, Dvořák J, Outrata J (2007) Fast factorization by similarity in formal concept analysis of data with fuzzy attributes. J Comput Syst Sci 73(6):1012–1022
41. Belohlavek R, Krmelova M (2014) Factor analysis of ordinal data via decomposition of matrices with grades. Ann Math Artif Intell 72(1):23–44
42. Belohlavek R, Trneckova M (2019) Factorization of matrices with grades via essential entries. Fuzzy Sets Syst 360:97–116

43. Sumangali K, Kumar CA (2019) Knowledge reduction in formal contexts through CUR matrix decomposition. Cybern Syst 50(5):465–496

44. Singh PK (2020) Bipolar δ-equal complex fuzzy concept lattice with its application. Neural Comput Appl 32(7):2405–2422

45. Lin MM, Dong B, Chu MT (2005) Integer matrix factorization and its application [Online]. Available http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.157.7359. Accessed: 05 Jun 2020

46. Shepard RN (1962) The analysis of proximities: multidimensional scaling with an unknown distance function. I. Psychometrika 27(2):125–140

47. Kruskal JB (1964) Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. Psychometrika 29:1–27

48. Kruskal JB (1964) Nonmetric multidimensional scaling: a numerical method. Psychometrika 29:115–129

49. Kruskal JB, Shepard RN (1974) A nonmetric variety of linear factor analysis. Psychometrika 39(2):123–157

50. Guttman L (1968) A general nonmetric technique for finding the smallest coordinate space for a configuration of points. Psychometrika 33(4):469–506

51. Vera JF, Mair P (2019) SEMDS: an R package for structural equation multidimensional scaling. Struct Equ Model Multidiscipl J 26(5):803–818

52. Borg I (2020) Data fit (stress) vs. model fit (recovery) in multidimensional scaling. Austrian J Stat 49(2):43–52

53. Lu S-T, Zhang M, Li Q-N (2020) Feasibility and a fast algorithm for Euclidean distance matrix optimization with ordinal constraints. Comput Optim Appl 76(2):535–569

54. Ekman G (1954) Dimensions of color vision. J Psychol 38(2):467–474

55. Shepard RN (1962) The analysis of proximities: multidimensional scaling with an unknown distance function. II. Psychometrika 27(3):219–246

56. Shepard RN (1980) Multidimensional scaling, tree-fitting, and clustering. Science 210(4468):390–398

57. Kruskal J, Wish M (1978) Multidimensional scaling. Sage, Thousand Oaks

58. Lee-Kwang H, Song Y-S, Lee K-M (1994) Similarity measure between fuzzy sets and between elements. Fuzzy Sets Syst 62(3):291–293

59. Wang W-J (1997) New similarity measures on fuzzy sets and on elements. Fuzzy Sets Syst 85(3):305–309

60. Vassilevska V, Williams R, Yuster R (2009) All pairs bottleneck paths and max–min matrix products in truly subcubic time. Theory Comput 5:173–189

61. Tversky A (1977) Features of similarity. Psychol Rev 84(4):327–352

62. Singh PK, Gani A (2015) Fuzzy concept lattice reduction using Shannon entropy and Huffman coding. J Appl Non-Class Log 25(2):101–119

63. Singh PK, Cherukuri AK, Li J (2017) Concepts reduction in formal concept analysis with fuzzy setting using Shannon entropy. Int J Mach Learn Cybern 8(1):179–189

64. Singh PK (2018) Cloud data processing using granular based weighted concept lattice and Hamming distance. Computing 100(10):1109–1132

65. Kaiser HF (1960) The application of electronic computers to factor analysis. Educ Psychol Meas 20(1):141–151

66. Ceulemans E, Kiers HAL (2006) Selecting among three-mode principal component models of different types and complexities: a numerical convex hull based method. Br J Math Stat Psychol 59(1):133–150

67. Preacher KJ, Zhang G, Kim C, Mels G (2013) Choosing the optimal number of factors in exploratory factor analysis: a model selection perspective. Multivar Behav Res 48(1):28–56

68. Wilderjans TF, Ceulemans E, Meers K (2013) CHull: a generic convex-hull-based model selection method. Behav Res 45(1):1–15

69. Shepard RN (1974) Representation of structure in similarity data: problems and prospects. Psychometrika 39(4):373–421

70. Wilderjans TF, Ceulemans E, Kiers HAL, Meers K (2009) The LMPCA program: a graphical user interface for fitting the linked-mode PARAFAC-PCA model to coupled real-valued data. Behav Res Methods 41(4):1073–1082

71. Wilderjans T, Ceulemans E, Mechelen IV (2008) The CHIC model: a global model for coupled binary data. Psychometrika 73(4):729–751

72. Hair JF, Tatham RL, Anderson RE, Black W (1998) Multivariate data analysis, 5th edn. Prentice Hall, Upper Saddle River

73. Tabachnick BG, Fidell LS (2000) Using multivariate statistics, 4th edn. Allyn and Bacon, Boston

74. Kaufman AS (1994) Intelligent testing with the WISC-III. Wiley, New York

75. Stevens J (2002) Applied multivariate statistics for the social sciences. Lawrence Erlbaum, Mahwah

76. Peterson RA (2000) A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. Mark Lett 11(3):261–275

77. Robertson AR (1977) The CIE 1976 color-difference formulae. Color Res Appl 2(1):7–11

78. Brady TF, Konkle T, Alvarez GA, Oliva A (2008) Visual long-term memory has a massive storage capacity for object details. PNAS 105(38):14325–14329

79. Konkle T, Brady TF, Alvarez GA, Oliva A (2010) Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. J Exp Psychol Gen 139(3):558–578

80. Hout MC, Goldinger SD, Brady KJ (2014) MM-MDS: a multidimensional scaling database with similarity ratings for 240 object categories from the massive memory picture database. PLoS ONE 9(11):e112644

81. Jacobowitz D (1973) Development of semantic structures. Unpublished Ph.D. dissertation, University of North Carolina, Chapel Hill

82. Koenigs M, Tranel D (2008) Prefrontal cortex damage abolishes brand-cued changes in cola preference. Soc Cogn Affect Neurosci 3(1):1–6

83. Astolfi L et al. (2009) The track of brain activity during the observation of TV commercials with the high-resolution EEG technology. Comput Intell Neurosci, vol 2009, no. Article ID 652078

84. Khushaba RN, Greenacre L, Kodagoda S, Louviere J, Burke S, Dissanayake G (2012) Choice modeling and the brain: a study on the Electroencephalogram (EEG) of preferences. Expert Syst Appl 39(16):12378–12388

85. Khushaba RN, Wise C, Kodagoda S, Louviere J, Kahn BE, Townsend C (2013) Consumer neuroscience: assessing the brain response to marketing stimuli using electroencephalogram (EEG) and eye tracking. Expert Syst Appl 40(9):3803–3812

86. Lahmiri S (2018) Neuromarketing perspective of consumer choice. [Online]. Available https://www.igi-global.com/gateway/chapter/199641. Accessed 20 Jul 2018

87. Grigaliunaite V, Pileliene L (2016) Emotional or rational? The determination of the influence of advertising appeal on advertising effectiveness. Sci Ann Econ Bus 63(3):391–414

88. Frolov AA, Sirota AM, Húsek D, Muraviev IP, Polyakov PA (2004) Binary factorization in hopfield-like neural networks: single-step approximation and computer simulations. Neural Netw World 14(2):139–152

89. Frolov AA, Husek D, Muraviev IP, Polyakov PY (2007) Boolean factor analysis by attractor neural network. IEEE Trans Neural Netw 18(3):698–707

90. Frolov AA, Husek D, Polyakov PY (2009) Recurrent-neural-network-based boolean factor analysis and its application to word clustering. IEEE Trans Neural Netw 20(7):1073–1086

91. Frolov AA, Húsek D, Polyakov PY (2016) Comparison of seven methods for boolean factor analysis and their evaluation by information gain. IEEE Trans Neural Netw Learn Syst 27(3):538–550

92. Shitov Y (2020) Factoring a band matrix over a semiring. Fuzzy Sets Syst 382:165–171

93. Akteke-Özturk B, Köksal G, Weber GW (2018) Nonconvex optimization of desirability functions. Quality Eng 30(2):293–310

94. Akteke-Öztürk B, Weber G-W, Köksal G (2017) Optimization of generalized desirability functions under model uncertainty. Optimization 66(12):2157–2169

95. Akteke-Öztürk B, Weber G-W, Köksal G (2020) Generalized desirability functions: a structural and topological analysis of desirability functions. Optimization 69(1):115–130

96. Arqub OA, AL-Smadi M, Momani S, Hayat T (2016) Numerical solutions of fuzzy differential equations using reproducing kernel Hilbert space method. Soft Comput 20(8):3283–3302

97. Arqub OA (2017) Adaptation of reproducing kernel algorithm for solving fuzzy Fredholm–Volterra integrodifferential equations. Neural Comput Appl 28(7):1591–1610

98. Arqub OA, Al-Smadi M, Momani S, Hayat T (2017) Application of reproducing kernel algorithm for solving second-order, two-point fuzzy boundary value problems. Soft Comput 21(23):7191–7206

99. Arqub OA, Al-Smadi M (2020) Fuzzy conformable fractional differential equations: novel extended approach and new numerical solutions. Soft Comput 24(16):12501–12522

100. Kropat E, Tikidji-Hamburyan RA, Weber G-W (2017) Operations research in neuroscience. Ann Oper Res 258(1):1–4

101. Çevik A, Weber G-W, Eyüboğlu BM, Oğuz KK, Initiative TADN (2017) Voxel-MARS: a method for early detection of Alzheimer's disease by classification of structural brain MRI. Ann Oper Res 258(1):31–57

102. Zucchini W, MacDonald IL, Langrock R (2016) Hidden Markov models for time series: an introduction using R, 2nd edn. Chapman and Hall/CRC, Boca Raton

103. Grobelny J, Michalski R (2017) Applying hidden Markov models to visual activity analysis for simple digital control panel operations. In: Świątek J, Wilimowska Z, Borzemski L, Grzech A (eds) Advances in intelligent systems and computing. Springer, Berlin, pp 3–14

104. Grobelny J, Michalski R (2017) Zastosowanie modeli Markowa z ukrytymi stanami do analizy aktywności wzrokowej w procesie oceny wirtualnych opakowań techniką porównywania parami. Zeszyty Naukowe Politechniki Poznańskiej Organizacja i Zarządzanie 73:111–125

105. Coutrot A, Hsiao JH, Chan AB (2018) Scanpath modeling and classification with hidden Markov models. Behav Res Methods 50(1):362–379

106. Ulutas BH, Özkan NF, Michalski R (2020) Application of hidden Markov models to eye tracking data analysis of visual quality inspection operations. Cent Eur J Oper Res 28(2):761–777

107. Grobelny J, Michalski R (2020) Investigating human visual behavior by hidden markov models in the design of marketing information. In: Cassenti DN (ed) Advances in human factors and simulation. Springer, Berlin, pp 234–245

108. Grobelny J, Michalski R (2012) Hidden Markov models for visual processing of marketing leaflets. In: Cassenti DN, Scataglini S, Rajulu SL, Wright JL (eds) Advances in simulation and digital human modeling. Springer, Berlin, pp 72–79

109. Mohamed MA, Gader P (2000) Generalized hidden Markov models. I. Theoretical frameworks. IEEE Trans Fuzzy Syst 8(1):67–81

110. Verma NK, Hanmandlu M (2010) Additive and nonadditive fuzzy hidden Markov models. IEEE Trans Fuzzy Syst 18(1):40–56

111. Winkley J, Jiang P (2014) Adaptive probability scheme for behaviour monitoring of the elderly using a specialised ambient device. Int J Mach Learn Cybern 5(2):293–307

112. Temoçin BZ, Weber G-W (2014) Optimal control of stochastic hybrid system with jumps: a numerical approximation. J Comput Appl Math 259:443–451

113. Savku E, Weber G-W (2018) A stochastic maximum principle for a markov regime-switching jump-diffusion model with delay and an application to finance. J Optim Theory Appl 179:696–721

114. Choi Y, Park J, Shin D (2017) A semi-supervised inattention detection method using biological signal. Ann Oper Res 258(1):59–78

115. Rasoulzadeh V, Erkus EC, Yogurt TA, Ulusoy I, Zergeroğlu SA (2017) A comparative stationarity analysis of EEG signals. Ann Oper Res 258(1):133–157

116. Harris JM, Ciorciari J, Gountas J Consumer neuroscience for marketing researchers. J Consum Behav 17(3): 239–252

117. Zhao Y, Siau K (2016) Cognitive neuroscience in information systems research. J Datab Manag (JDM) 27(1):58–73

118. Glynn EF II (2001) Spectra software. EFG Computer Lab, Overland Park