International Journal of Approximate Reasoning 173 (2024) 109247

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Semantic explorations in factorizing Boolean data via formal concepts

Radim Belohlavek, Martin Trnecka *

Department of Computer Science, Palacký University Olomouc, Czech Republic

ARTICLE INFO

Keywords:
Boolean matrix factorization
Formal concept
Human concept

ABSTRACT

We use now available psychological data involving human concepts, objects covered by these concepts, and binary attributes describing the objects to explore selected semantic aspects of Boolean matrix factorization. Our basic perspective derives from the intuitive requirement that the factors computed from data should represent natural categories latently present in the data. This idea is examined for factorization algorithms that utilize formal concepts to build factors. We provide several experimental observations which imply that the inspected factorization methods deliver semantically sound factors that resemble significant human concepts of the examined domains.

1. Our goals

Boolean matrix factorization (BMF) has become a widespread topic in modern data analysis. Previous research has focused almost exclusively on developing new approaches and algorithms to factorize Boolean matrices [15]. That is, on finding ways to compute for an input $n \times m$ Boolean matrix I an $n \times k$ Boolean matrix A and a $k \times m$ Boolean matrix B such that the Boolean product $A \circ B$ provides a good approximation of I and yet, the number k of the so-called factors is reasonably small. It has been repeatedly argued that the computed factors, represented by the resulting matrices A and B, are naturally interpretable and easy to understand. This desirable feature of BMF results from the Booleanity of the resulting matrices A and B and the involvement of logical conjunction and disjunction in the Boolean matrix product $\circ$. Yet, even though the developed algorithms have been evaluated not only on synthetic but also on real-world Boolean data, semantic considerations are rather rare in BMF research and restrict to attempts to interpret the computed factors. See, e.g., Miettinen's theses [12] and [13, sec. 4.8], and the works [7,14], in which the authors provide interpretations of extracted factors.

In our paper, we use now available Dutch data—a unique, high-quality psychological data involving human concepts, objects covered by these concepts, and binary attributes describing the objects—to explore various questions inspired by semantic considerations pertaining to BMF. The questions relate to the interpretation and meaningfulness of the computed factors, the ability to discover factors that are hidden in the data, and robustness of factorization. We attempt to provide various viewpoints to consider these questions. In this initial study, we restrict to the BMF algorithms that use formal concepts to build factors, not only due to a limited space but also because of the ability of formal concepts to be interpreted naturally as human concepts. Due to its position among the BMF algorithms utilizing formal concepts, which is explained below, we concentrate on the GreConD algorithm, but also present results for other algorithms. In the next section, we provide the notions and notation essential for BMF. In section 3.1, we

* Corresponding author.
E-mail addresses: radim.belohlavek@acm.org (R. Belohlavek), martin.trnecka@gmail.com (M. Trnecka).

https://doi.org/10.1016/j.ijar.2024.109247
Received 17 March 2024; Received in revised form 12 June 2024; Accepted 3 July 2024

Available online 8 July 2024
0888-613X/

describe the Dutch data, in particular the parts used in our experiments. The questions we explore, our experimental results, and a discussion are presented in section 3.2. Section 4 concludes the paper by providing topics for future research.

2. Notions and notation

To fix terminology and notation, denote by $\{0,1\}^{n\times m}$ the set of all $n\times m$ Boolean matrices. Each entry I_{ij} of a Boolean matrix $I\in\{0,1\}^{n\times m}$ thus equals 0 or 1. In most applications, the rows $i=1,\dots,n$ and the columns $j=1,\dots,m$ represent objects (exemplars) and attributes (properties, features), respectively, and $I_{ij}=1$ indicates that the object i, such as a particular organism or a particular product, has the attribute j, such as "to breathe" or "to contain lithium." The ith row and the jth column of I shall be denoted by $I_{i\_}$ and $I_{\_j}$, respectively.

The basic problem in BMF consists in finding for a given object×attribute matrix $I\in\{0,1\}^{n\times m}$ an object×factor matrix $A\in\{0,1\}^{n\times k}$ and a factor×attribute matrix $B\in\{0,1\}^{k\times m}$ such that k is reasonably small and I approximately equals the well-known Boolean matrix product $A\circ B$. Here, the Boolean matrix product $\circ$ is defined by

$$(A\circ B)_{ij}=\max_{l=1}^{k}\min(A_{il},B_{lj}),$$

and the approximate equality of I and $A\circ B$ is assessed by the metric $E(\cdot,\cdot)$ induced by the matrix L_1-norm, which is defined for matrices $C,D\in\{0,1\}^{n\times m}$ by

$$E(C,D)=\sum_{i,j=1}^{n,m}|C_{ij}-D_{ij}|;$$

see, e.g., [11].

Two particular optimization problems result from the above basic problem [4,14]:

- the approximate factorization problem, in which a threshold ε is prescribed and the smallest number k of factors is sought for which $E(I,A\circ B)\le\varepsilon$, and
- the discrete basis problem in which a number k is prescribed and k factors are sought for which $E(I,A\circ B)$ is as small as possible.

The above notions are presented in the framework of Boolean matrices, as usual in most papers on BMF. Alternatively, they can be presented in the framework of sets and relations as usual, e.g., in formal concept analysis. Namely, as is well known, Boolean matrices $I\in\{0,1\}^{n\times m}$ may be identified with binary relations between sets $\{1,\dots,n\}$ and $\{1,\dots,m\}$, which we denote by a slight abuse of notation by I as well. Similarly, Boolean vectors C and D in $\{0,1\}^n$ and $\{0,1\}^m$ may be identified with the corresponding subsets of $\{1,\dots,n\}$ and $\{1,\dots,m\}$, respectively. That is to say, we identify sets and relations with their characteristic vectors: $I_{ij}=1$, $C_i=1$, and $D_j=1$ correspond to $\langle i,j\rangle\in I$, $i\in C$, and $j\in D$, respectively.

The BMF algorithms to which we restrict utilize the so-called formal concepts associated to the input matrices $I\in\{0,1\}^{n\times m}$, i.e., basic patterns used in formal concept analysis [9,19]. Recall that in formal concept analysis, I_{ij} indicates whether the object i has or does not have the attribute j, which options correspond to $I_{ij}=1$ and $I_{ij}=0$, respectively. A formal concept in I is any pair $\langle C,D\rangle\in\{0,1\}^n\times\{0,1\}^m$, such that C (the so-called extent of the concept) represents the set of all objects sharing all the attributes in D, while D (intent) represents the set of all attributes shared by all the objects in C.

Consider the 4×5 Boolean matrix

$$I=\begin{pmatrix}0&0&0&1&1\\1&1&1&0&0\\1&1&1&1&1\\0&0&1&1&1\end{pmatrix},$$

in which the rows and columns represent the objects $1,\dots,4$ and the attributes $1,\dots,5$, respectively. For instance, the object 2 has the attribute 1, since $I_{21}=1$, but does not have the attribute 4, since $I_{24}=0$. One easily checks that the pair $\langle C_1,D_1\rangle$ representing the sets $C_1=\{2,3\}$ and $D_1=\{1,2,3\}$, or $C_1=(0\ \ 1\ \ 1\ \ 0)$ and $D_1=(1\ \ 1\ \ 1\ \ 0\ \ 0)$ in the vector notation, is a formal concept in I. The pairs $\langle C_2,D_2\rangle=\langle\{3,4\},\{3,4,5\}\rangle$ and $\langle C_3,D_3\rangle=\langle\{1,3,4\},\{4,5\}\rangle$ represent other formal concepts in I. As is apparent from this example, the formal concepts in I correspond—up to a permutation of rows and columns—to maximal rectangular areas in I that are filled with 1s.

Consider now the set $\mathcal{F}=\{\langle C_1,D_1\rangle,\langle C_2,D_2\rangle,\langle C_3,D_3\rangle\}$ of the above formal concepts of the Boolean matrix I. The set $\mathcal{F}$ has the convenient property that it may be used to factorize I in that it induces matrices $A_{\mathcal{F}}\in\{0,1\}^{4\times 3}$ and $B_{\mathcal{F}}\in\{0,1\}^{3\times 5}$, namely

$$A_{\mathcal{F}}=\begin{pmatrix}0&0&1\\1&0&0\\1&1&1\\0&1&1\end{pmatrix}\quad\text{and}\quad B_{\mathcal{F}}=\begin{pmatrix}1&1&1&0&0\\0&0&1&1&1\\0&0&0&1&1\end{pmatrix},$$

for which

$$I=A_{\mathcal{F}}\circ B_{\mathcal{F}}.$$

The general rule to form $A_{\mathcal{F}}$ and $B_{\mathcal{F}}$ from a set $\mathcal{F}$ of formal concepts $\langle C_i, D_i\rangle$ of I is to use (the vectors corresponding to) C_is and D_is as the columns of $A_{\mathcal{F}}$ and the rows of $B_{\mathcal{F}}$, respectively, as in the example above.

Observe that using only $\mathcal{G} = \{\langle C_1, D_1\rangle, \langle C_2, D_2\rangle\}$ as the factors, we have

$$I \neq A_{\mathcal{G}} \circ B_{\mathcal{G}} \quad \text{since} \quad A_{\mathcal{G}} \circ B_{\mathcal{G}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix},$$

Yet, $A_{\mathcal{G}} \circ B_{\mathcal{G}}$ provides a good approximation of I as

$$E(I, A_{\mathcal{G}} \circ B_{\mathcal{G}}) = 2$$

because only two 1-entries of I, namely I_{14} and I_{15}, remain uncovered by the two factors in $\mathcal{G}$.

To observe the approximation capability of a set $\mathcal{G}$ of factors, one utilizes the complement to 1 of a normalized error $\frac{E(I, A_{\mathcal{G}} \circ B_{\mathcal{G}})}{||I||}$, i.e., the function

$$c = 1 - E(I, A_{\mathcal{G}} \circ B_{\mathcal{G}})/||I||, \tag{1}$$

where $||I||$ denotes the number of 1s in I, i.e., $||I|| = \sum_{i,j=1}^{n,m} I_{ij}$. Clearly, c measures coverage of 1s in I by the factors in $\mathcal{G}$ and ranges from 0 to 1. This function shall be used in our experimental evaluation below. For more details on the above concepts, we refer, e.g., to [4].

3. Experiments, results, and discussion

3.1. Data

The Dutch data has been gathered within a psychological study involving hundreds of respondents at the KU Leuven [8], as well as within the preceding study [17]. It represents the largest data regarding human concepts and related phenomena studied within the psychology of concepts and is much larger and more elaborate compared to other available data; see also [16]. For our purpose, it is essential that the data involves selected common human concepts (called categories by the authors), objects covered by these concepts, relevant binary attributes, and a variety of Boolean matrices describing the objects using the attributes. In what follows, we describe the data along with related aspects.

The data involves 10 natural kind concepts and 6 artifact concepts, each represented by a number of objects obtained from over five hundred participants. The natural kind concepts are: “fruit” (with 30 exemplars); “vegetable” (30); “profession” (30); “sport” (30); and 6 animal categories, namely, “amphibian”(5), “bird” (30), “fish” (23), “insect” (26), “mammal” (30), and “reptile” (22).[1] The artifact concepts are: “clothing” (29), “kitchen utensil” (33), “musical instrument” (27), “tool” (30), “vehicle” (30), and “weapon” (20). The concepts have a considerable coverage, especially of two domains which shall be particularly explored in our experiments, namely the animal domain including 129 objects (136 minus the 5 amphibians and 2 reptiles; cf. note 1) and the artifact domain including 166 objects (169 minus three objects, each of which is covered by two distinct concepts). The concepts have a notable coverage of the domains [8].

In addition to the objects representing the included concepts, the Dutch data involves a large number of binary (i.e., yes-no) attributes. These were obtained from over a thousand respondents in two ways. The respondents were first asked to write down attributes relevant to a presented concept name; these are called the category attributes in the Dutch data. Secondly, they were required to write down attributes for a name of an individual object; these are called the exemplar attributes.

For the purpose of creating object-attribute Boolean matrices, the following collections of attributes are considered in the Dutch data: (a) for each concept (i.e., “fruit”, “vegetable”, …, “weapon”), the collection of all category attributes listed for this concept; (b) for each concept, the collection of all exemplar attributes obtained for all the objects in the concept; (c) for each domain (i.e., animal and artifact), the collection of all category attributes listed for the concepts in this domain; (d) for each domain, the collection of all exemplar attributes listed for the objects in this domain. Using these collections of attributes, various Boolean matrices were obtained, in which the entries indicate whether the corresponding object (matrix row) has the corresponding attribute (matrix column).

This way, several matrix types (dimensions) arise, which are described in Table 1 and Table 2. For instance, the first row in Table 1 refers to two types of Boolean matrices: A 30×28 matrix describing which of the 30 objects of the concept “bird” have which of the 28 category attributes listed by the respondents for this concept, and a 30×225 matrix telling which of the 30 objects of “bird” have which of the 225 exemplar attributes of “bird”, i.e., the attributes listed as exemplar attributes for some object of “bird”. Analogously, the 166×301 Boolean matrix in the second row of Table 2 describes which of the 166 objects in the artifact domain have which of the corresponding 301 category attributes.[2]

[1] As the 5 objects of “amphibian” are included in “reptile”, we omit the concept “amphibian” in the considerations below; see [8] for reasons of this inclusion. Note also that the Boolean matrices involving the reptiles include only 20 objects due to lack of familiarity of the respondents with two objects, namely iguanodon and komodo.

[2] That is, the 166 objects are all the objects of “clothing”, “kitchen utensil”, “musical instrument”, “tool”, “vehicle”, and “weapon”; and the 301 attributes are all attributes listed as category attributes for these six concepts.

Table 1
Dimensions of the category-based Boolean matrices used in our experiments.

category	objects	category attributes	exemplar attributes
bird	30	28	225
clothing	29	38	258
fruit	30	32	233
fish	23	32	156
insect	26	37	214
kitchen utensil	33	39	328
mammal	30	34	288
musical instrument	27	39	218
profession	30	21	370
reptile	20	35	179
sport	30	33	382
tool	30	37	285
vegetable	30	30	291
vehicle	30	34	322
weapon	20	32	181

Table 2
Dimensions of the domain-based Boolean matrices used in our experiments.

domain	objects	category attributes	exemplar attributes
animal	129	225	764
artifact	166	301	1,295

The actual Boolean matrices of the types in Tables 1 and 2, were obtained from the respondents (77 respondents were involved in total), 4 respondents per matrix type resulting in 4 Boolean matrices for each type. In addition to these Boolean matrices, the data contains aggregated matrices, one matrix for each type, in which the entries contain values 0 to 4 and equal the number of respondents who claimed that the respective object has the respective attribute. To obtain Boolean matrices from these aggregated matrices, one naturally thresholds the matrix entries: For each aggregated matrix, one obtains 4 Boolean matrices corresponding to the thresholds ≥ 1, ≥ 2, ≥ 3, and $= 4$. We call these thresholded matrices the consensus matrices. For instance, an entry in an ≥ 2 matrix equals 1 if at least 2 respondents agreed that the corresponding object has the corresponding attribute.

All the matrices we use come from [1], which contains adjusted Dutch data along with a convenient Python wrapper. The adjustment aimed at removing minor semantic and technical faults and inconveniences of the original Dutch data [8].

3.2. Main questions considered

In our exploratory study, we consider several questions related to the semantic aspects of BMF. These should be considered as topics of an initial investigation that require further attention. Our explorations are described below.

An obvious question, generally considered in BMF, is whether the computed factors represent good factorizations of the Dutch data. That is, the aim is to assess the computed factors in terms of coverage, particularly coverage by the first few computed factors in terms of the precision of the overall factorization.

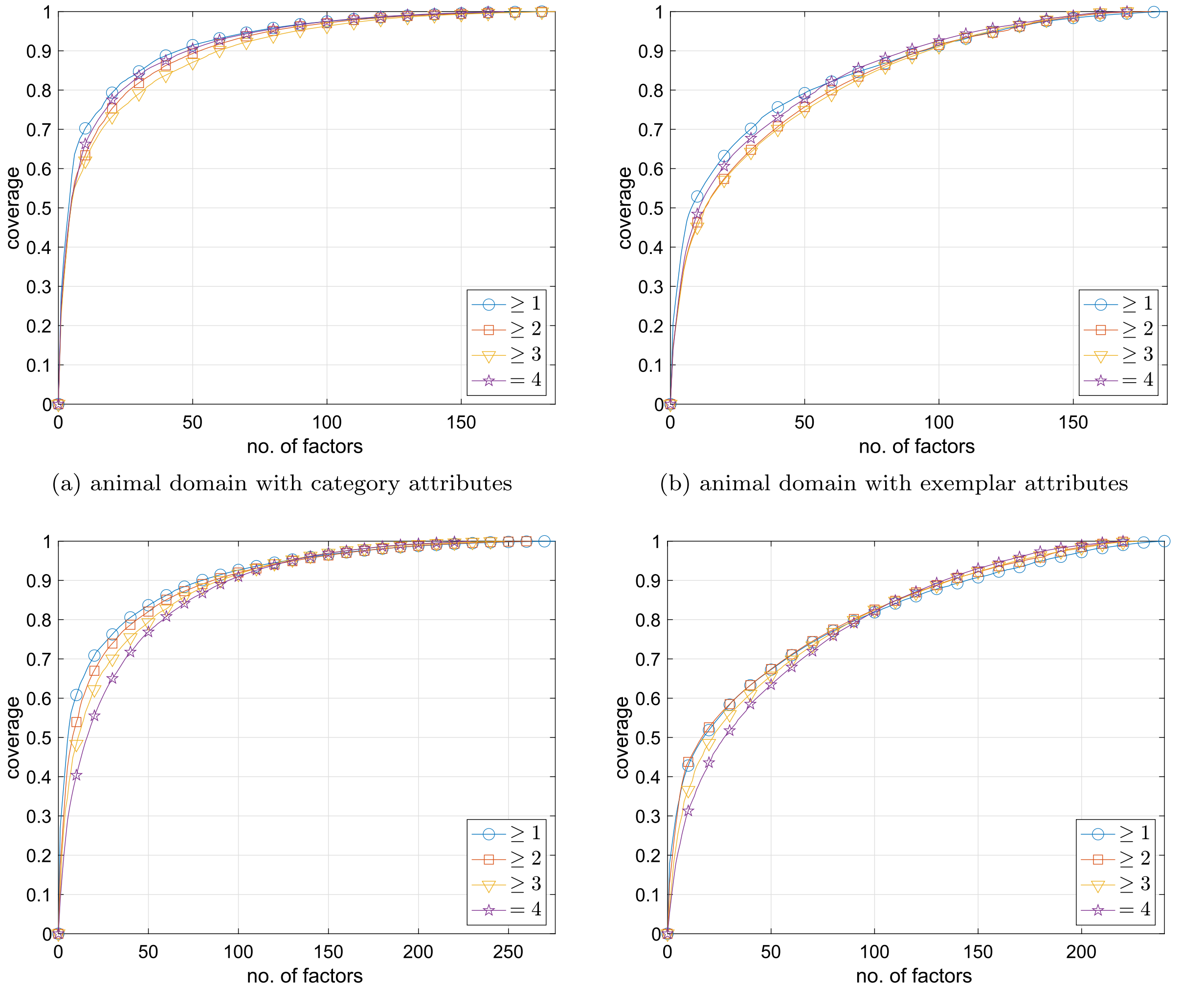
The second question, which already addresses the main topic of our paper, is that of interpretability of factors. This question explores the meanings of the computed factors. We are particularly interested if the factors represent natural concepts, i.e., if they can be easily understood by people.

Related to the second question is another question, namely, whether the human categories underlying Dutch data appear among the computed factors. Since the Dutch data is built around certain human concepts, one is naturally interested in whether these concepts are retrieved by factorization, or are at least similar to the factors computed, and whether they represent important factors in the sense that they are among the first few of the computed factors.

We also look into whether the computed factorizations are robust. Namely, as explained in section 3.1, for each matrix type, the Dutch data contains four Boolean matrices that correspond to four respondents. These four matrices may be regarded as representing four expert views of the corresponding “part of the world” described by the matrix. It is reasonable to assume that these matrices are mutually similar to some considerable degree, and this indeed appears to be the case. We then ask whether the four factorizations of these four Boolean matrices may be regarded as being mutually similar in that they contain similar factors. In a sense, we ask whether a particular factorization method is robust with respect to presumably not substantial differences in the expert knowledge codified by the four matrices.

3.3. Results for the GRECOND algorithm

As mentioned above, we examine in detail the results obtained by the GRECOND algorithm [7, Algorithm 2] for its particular role among the BMF algorithms. Most importantly, GRECOND produces good factorizations from the viewpoint of both the approximate factorization problem and the discrete basis problem. In addition, it is among the fastest BMF algorithms.



(a) animal domain with category attributes
(b) animal domain with exemplar attributes
(c) artifact domain with category attributes
(d) artifact domain with exemplar attributes

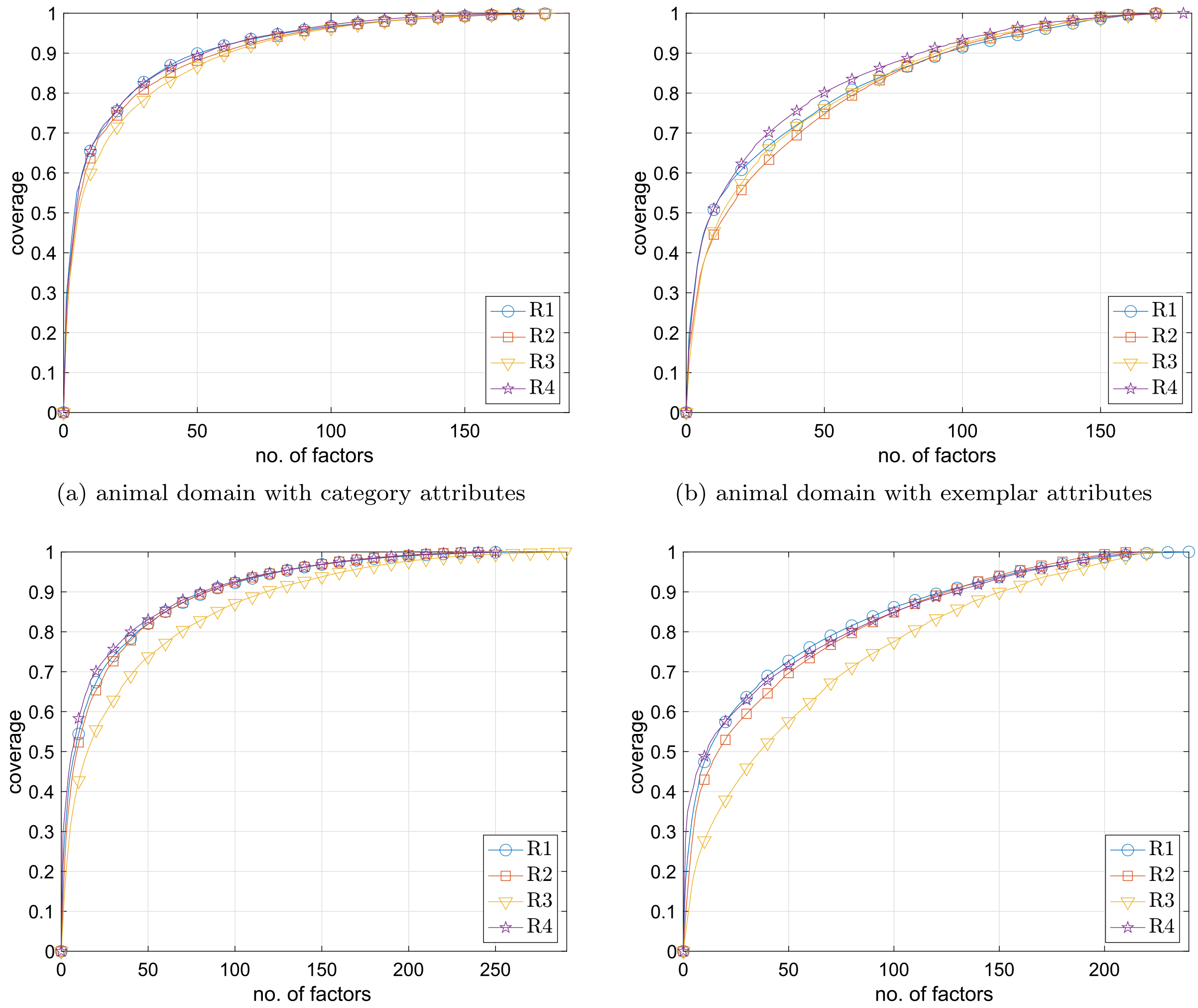
Fig. 1. Factorization of consensus domain-based matrices obtained by GRECOND.

In particular, GRECOND uses the formal concepts of the input Boolean matrix I as factors. It computes these factors from I one-by-one using a simple greedy strategy without the need to compute all the formal concepts of I. When computing a new factor, GRECOND attempts to maximize the number of entries in I containing 1 that have not been covered by the previously computed factors. In doing so, it starts with the attribute concepts, i.e., the formal concepts generated by single attributes, and attempts to extend them by further attributes until such an extension produces a factor with a better coverage. The algorithm is guaranteed to produce an exact factorization of I, but may be stopped earlier when a factorization with a desired accuracy has been obtained.

3.3.1. Factorizations obtained by GRECOND

The factorizations obtained by GRECOND for the large, domain-based matrices of Table 2 are displayed in Fig. 1 and Fig. 2. In particular, Fig. 1 provides the graphs of coverage (1) for the consensus based matrices. The four parts, (a)–(d), correspond to the four variants of the domain-based matrices (i.e., the animal or the artifact domain described by the category or the exemplar attributes). For instance, part (b) in Fig. 1 contains the graphs of coverage of the four consensus matrices describing the animal domain with the exemplar attributes. That is, the four Boolean 129×764 matrices corresponding to the thresholds ≥ 1, ≥ 2, ≥ 3, and $= 4$, among the four respondents who had filled the 129×764 matrix; cf. section 3.1. The layout in Fig. 2 is similar, except that the four graphs in each of the parts (a)–(d) correspond to the four domain-based matrices filled by the four respondents denoted R1, R2, R3, and R4 in the legends to the parts.

The particular graphs display the coverage in the usual manner: The horizontal axis represents the numbers $l = 1, 2, \ldots$ of the consecutively computed factors and the vertical one represents the corresponding values of coverage by the first l factors. That is, for a particular matrix I, the value in the graph corresponding to the number l represents the coverage c defined by (1) in which $\mathcal{G}$ denoted the set consisting of the first l factors computed from the matrix I.



(a) animal domain with category attributes (b) animal domain with exemplar attributes

(c) artifact domain with category attributes (d) artifact domain with exemplar attributes

Fig. 2. Factorization of individual respondents' domain-based matrices obtained by GRECOND.

One may observe in Fig. 1 as well as in Fig. 2 that all the graphs have a desired shape and represent good factorizations; see, e.g., [3] for considerations on the quality of factorization and its assessment. In particular, the coverage grows rapidly for the first few factors, presumably the most important ones, and then grows slowly until it reaches the value of 1 corresponding to an exact factorization of the input matrix. From this quantitative viewpoint, all the involved matrices describing the domain data as well as the ability of GRECOND to obtain reasonable factorizations may be regarded as normal, i.e., not displaying unusual characteristics.

While we consider the above factorizations of the large, domain-based matrices to be of primary importance for our purpose of exploring semantic questions, we also include selected factorizations of the smaller, category-based matrices. The factorizations of selected category-based matrices with exemplar attributes obtained from the individual respondents are displayed in Fig. 3. One observes that the involved coverage graphs appear to represent good factorizations as well.

3.3.2. Interpretation of factors and retrieval of human categories by GRECOND

While the results in the previous section indicate that the factorizations computed by GRECOND are of a reasonable quality, they only concern the quantitative aspects of factorizations. As such, they only provide the context for our following considerations. In this section, we examine whether the computed factors can be interpreted in a natural way and whether the human categories underlying the Dutch data are reflected in the factors in a reasonable manner; see section 3.4.

Prior to the presenting the results, let us make a few comments. First, since we employ formal concepts, each factor in principle allows for a logically sound interpretation, which is given by the conjunction "y_1 and … and y_p" of the binary attributes y_j in the concept's intent. This, however, may become difficult with large matrices because a conjunction of, say, tens of attributes may not have an intuitive meaning. One hence needs to look beyond this straightforward option to interpret factors. Second, while our related study [2] implies that most of the human categories of the Dutch data indeed turn out to be extents of formal concepts of the domain-

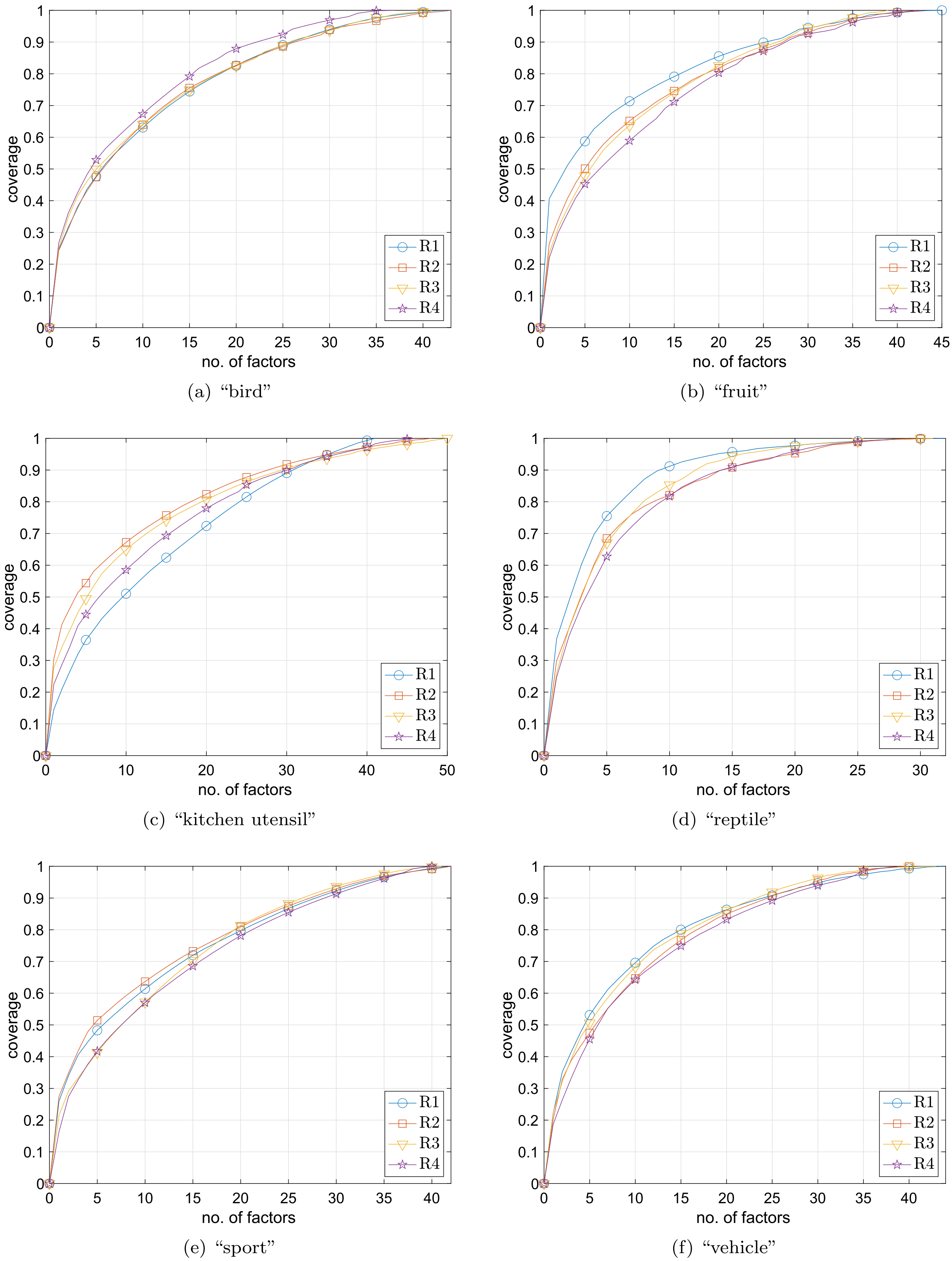


Fig. 3. Factorization of selected individual respondents’ category-based matrices by GreConD.

Table 3
Factors of animal domain with category attributes, consensus matrix for threshold ≥ 2.

#	∼	extent
1	typical animal	alligator, anchovy, ant, bat, beaver, bee, beetle, bison, blackbird, blindworm, boa, bumblebee, butterfly, caiman, canary, carp, cat, caterpillar, centipede, chameleon, chickadee, chicken, cobra, cockchafer, cockroach, cod, cow, cricket, crocodile, crow, cuckoo, deer, dog, dolphin, donkey, dove, dragonfly, dromedary, duck, eagle, earwig, eel, elephant, falcon, flatfish, flee, fly, fox, frog, fruit fly, gecko, giraffe, goldfish, grasshopper, hamster, hedgehog, heron, herring, hippopotamus, horse, horsefly, iguana, kangaroo, ladybug, leech, lion, lizard, llama, louse, magpie, monitor lizard, mosquito, moth, mouse, orca, ostrich, owl, parakeet, parrot, peacock, pelican, penguin, pheasant, pig, pike, piranha, plaice, python, rabbit, ray, robin, rooster, salamander, salmon, sardine, seagull, shark, sheep, snake, sole, sparrow, spider, squid, squirrel, stickleback, stork, swallow, swan, swordfish, toad, tortoise, trout, turkey, turtle, viper, vulture, wasp, wolf, wood louse, woodpecker, worm, zebra
2	bird	blackbird, canary, chickadee, chicken, crow, cuckoo, dove, duck, eagle, falcon, heron, magpie, ostrich, owl, parakeet, parrot, peacock, pelican, pheasant, robin, rooster, seagull, sparrow, stork, swallow, swan, turkey, woodpecker
3	fish	anchovy, carp, cod, eel, flatfish, herring, pike, plaice, ray, salmon, sardine, sole, stickleback, swordfish, trout
4	mammal	beaver, bison, cat, cow, deer, dog, donkey, dromedary, elephant, fox, hamster, hedgehog, hippopotamus, horse, kangaroo, lion, llama, monkey, mouse, pig, polar bear, rabbit, rhinoceros, sheep, squirrel, tiger, wolf, zebra
5	insect	ant, bee, beetle, bumblebee, butterfly, centipede, cockchafer, cockroach, cricket, dragonfly, earwig, flee, fly, fruit fly, grasshopper, horsefly, ladybug, louse, mosquito, moth, spider, wasp, wood louse
6	reptile	alligator, blindworm, boa, caiman, chameleon, cobra, crocodile, gecko, iguana, lizard, monitor lizard, python, salamander, snake, viper

based Boolean matrices, it would be deceptive to assume that the domain-based matrices split in a disjoint-like manner into parts corresponding to the underlying human categories, hence to assume that the domain-based matrices are amenable to disjoint-like, trivial factorizations.[3]

We now examine in detail the first couple of factors computed by GreConD from selected domain-based matrices.[4] The first six factors computed from the animal domain matrix with category attributes representing the consensus of ≥ 2 (see section 3.1) are displayed in Table 3. The rows correspond to the factors in the order they were computed while the columns contain the number of the particular factor (#), the possible interpretation of the factor (∼), and the extent of the formal concept representing the factor.

The first factor is represented by a general concept which in fact is characterized by 18 of the 225 category attributes of the animal domain and contains about the same portion of objects of each human category of the animal domain. Since its characteristic attributes contain "belongsToNature," "breathes," "canBecomeIll," "doesNotResembleAHumanBeing," "hasAMouth," "hasEyes," and "isEatenByOtherAnimals" and similar attributes by and large applicable to animals, it may be interpreted as "typical animal." Such a factor, albeit considerably general, is rather natural and easy to understand. The factor has a large coverage and appears as a result of the greedy approach of GreConD. While it takes some consideration to find the meaning of the first factor, each of the next five factors has a very straightforward meaning which is provided in the table. Most importantly, these five factors correspond exactly to the five human categories underlying the animal domain of the Dutch data. Moreover, even though the extents of these factors do not coincide with the collections of objects representing the human categories, they extents match the collections significantly. All of the 28 objects of extent of the first factor all belong to the collection of 30 objects representing the human category "bird" in the Dutch data.[5] Hence, none of these objects fails to be a bird and the factor almost exhausts the whole category "bird" of Dutch data. The situation with the other factors vs. the corresponding human categories of the Dutch data is similar. One may thus conclude that not only the factors computed by GreConD have natural meanings but also contain, among the first six ones, all the five human categories that one would desire to retrieve as factors.

Similar observations can be made for the other matrices representing the animal domain; see, e.g., Tables 4 and 5 for the category attributes with threshold ≥ 3 and the exemplar attributes with threshold ≥ 2, respectively.[6]

Even though the artifact domain appears more complicated from the psychological viewpoint [8], as the involved categories tend to be less salient compared to those of the animal domain, our results imply similar conclusions. As an example, consider

[3] This is apparent from some of the results presented below and also from an analysis of the domain-based matrices, which we do not present. On the other hand, the underlying human categories may naturally be turned into factors which then have a good coverage and may be completed to reasonable overall factorizations by computing further factors.

[4] The results, however, are representative for all the domain-based matrices.

[5] One needs to consult the Dutch data to see this and some other observations we present in this section.

[6] We focus on the thresholds ≥ 2 and ≥ 3 as these seem to be of particular interest in that thresholds ≥ 1 and = 4 seem to be too low and too high, respectively.

Table 4
Factors of animal domain with category attributes, consensus matrix for threshold ≥ 3.

#	∼	extent
1	typical animal	alligator, anchovy, bat, beaver, bison, blackbird, boa, caiman, canary, carp, cat, chameleon, chickadee, chicken, cobra, cod, cow, crocodile, crow, cuckoo, deer, dog, dolphin, donkey, dove, dromedary, duck, eagle, eel, elephant, falcon, flatfish, fox, frog, gecko, giraffe, goldfish, hamster, hedgehog, heron, herring, hippopotamus, horse, iguana, kangaroo, lion, lizard, llama, magpie, monitor lizard, mouse, orca, ostrich, owl, parakeet, parrot, peacock, pelican, penguin, pheasant, pig, pike, piranha, plaice, polar bear, python, rabbit, ray, rhinoceros, robin, rooster, salamander, salmon, sardine, seagull, shark, sheep, snake, sole, sparrow, sperm whale, squirrel, stickleback, stork, swallow, swan, swordfish, tiger, toad, tortoise, trout, turkey, turtle, viper, vulture, whale, wolf, woodpecker, zebra
2	insect	ant, bee, beetle, bumblebee, butterfly, caterpillar, centipede, cockchafer, cockroach, cricket, dragonfly, earwig, flee, fly, fruit fly, grasshopper, horsefly, ladybug, leech, louse, mosquito, moth, spider, wasp, wood louse, worm
3	bird	blackbird, canary, chickadee, chicken, crow, cuckoo, dove, duck, eagle, falcon, heron, magpie, owl, parakeet, parrot, peacock, pelican, pheasant, robin, rooster, seagull, sparrow, stork, swallow, swan, turkey, vulture, woodpecker
4	mammal	beaver, bison, cat, cow, deer, dog, donkey, dromedary, fox, hamster, horse, kangaroo, lion, llama, monkey, mouse, polar bear, rabbit, sheep, squirrel, tiger, wolf, zebra
5	fish	anchovy, carp, cod, eel, herring, pike, plaice, ray, salmon, sardine, sole, stickleback, trout
6	reptile	alligator, anchovy, blindworm, boa, caiman, carp, caterpillar, centipede, chameleon, cobra, cod, crocodile, dinosaur, earwig, eel, flatfish, flee, frog, gecko, goldfish, grasshopper, herring, iguana, lizard, louse, monitor lizard, pike, piranha, plaice, python, ray, salamander, salmon, shark, snake, sole, squid, stickleback, swordfish, toad, tortoise, trout, turtle, viper, wood louse

Table 5
Factors of animal domain with exemplar attributes, consensus matrix for threshold ≥ 2.

#	∼	extent
1	typical animal	alligator, anchovy, ant, bat, beaver, beetle, bison, blackbird, blindworm, boa, bumblebee, butterfly, caiman, carp, caterpillar, centipede, chameleon, chickadee, cockchafer, cockroach, cod, cricket, crocodile, crow, cuckoo, deer, dolphin, dove, dromedary, duck, eagle, earwig, eel, elephant, falcon, flatfish, flee, fly, fox, frog, fruit fly, gecko, giraffe, grasshopper, hedgehog, heron, herring, hippopotamus, horse, iguana, kangaroo, ladybug, leech, lion, lizard, llama, louse, magpie, monitor lizard, moth, mouse, orca, ostrich, owl, parrot, peacock, pelican, penguin, pheasant, pike, piranha, plaice, polar bear, python, rabbit, ray, rhinoceros, robin, salamander, salmon, sardine, seagull, shark, snake, sole, sparrow, sperm whale, spider, squid, squirrel, stickleback, swan, swordfish, tiger, toad, tortoise, trout, turkey, turtle, viper, vulture, whale, wolf, wood louse, woodpecker, worm, zebra
2	fish	anchovy, carp, cod, eel, flatfish, herring, pike, plaice, ray, salmon, sardine, sole, stickleback, swordfish, trout
3	bird	blackbird, canary, chickadee, chicken, crow, cuckoo, dove, falcon, magpie, owl, parakeet, parrot, pheasant, robin, rooster, seagull, sparrow, swallow, turkey, woodpecker
4	insect	ant, bee, beetle, bumblebee, butterfly, centipede, cockchafer, cockroach, cricket, dragonfly, earwig, flee, fly, fruit fly, grasshopper, horsefly, ladybug, louse, mosquito, moth, spider, wasp, wood louse
5	mammal	beaver, bison, cat, cow, deer, dog, donkey, dromedary, elephant, fox, giraffe, hamster, hedgehog, hippopotamus, horse, kangaroo, lion, llama, monkey, mouse, pig, polar bear, rabbit, rhinoceros, sheep, squirrel, tiger, wolf, zebra
6	reptile	boa, cobra, python, snake

Table 6 providing the factors computed from the artifact domain matrix with category attributes representing the consensus of ≥ 2. The first factor again represents a "typical artifact," as its characteristic attributes include "canBeMadeOfDifferentMaterials," "canBeUsedSeveralTimes," "canBreak," "existsInDifferentColors," and other typical features. The next six factors then correspond to

Table 6
Factors of artifact domain with category attributes, consensus matrix for threshold ≥ 2.

#	~	extent
1	typical artifact	(hot air) balloon, adjustable spanner, airplane, anvil, apron, axe, bass guitar, bassoon, bathing suit, beanie, belt, bicycle, blouse, boat, boots, bowl, bra, bus, cap, car, carriage, cart, chisel, clamp, clarinet, club, coat, colander, crowbar (koevoet), cymbals, dress, drill, drum, drum set, dungarees, electric kettle, file, flute, fork, fridge, glass, gocart, grater, grinding disc, guitar, hammer, harpsichord, hat, helicopter, hovercraft, jeans, jeep, kettle, kick scooter, knife, lawn mower, level, microwave oven, mittens, mixer, motorbike (brommer), motorbike (moto), mug, nail, oven, paint brush, pan, pan flute, panties, pants, percolator, pickaxe, plane, plate, plough, pot, pullover, pyjamas, recorder, rope, saw, saxophone, scales, scarf, scissors, scooter, screwdriver, shield, shirt, shoes, shorts, shovel, sieve, skateboard, skirt, sled, slingshot, socks, spatula, spear, spoon, stove, submarine, subway train, suit, sweater, synthesizer, t-shirt, tambourine, taxi, tie, toaster, tongs, towel, tracksuit, tractor, trailer, train, tram, trombone, truck (camion), truck (vrachtwagen), trumpet, vacuum cleaner, van, whip, whisk, wire brush, wrench, zeppelin
2	musical instr.	accordion, banjo, bass guitar, bassoon, cello, clarinet, cymbals, double bass, drum, drum set, flute, guitar, harmonica, harp, organ, pan flute, piano, recorder, saxophone, synthesizer, tambourine, triangle, trombone, trumpet, violin
3	clothing	bathing suit, beanie, blouse, bra, coat, dress, dungarees, hat, jeans, mittens, panties, pants, pullover, pyjamas, scarf, shirt, shorts, skirt, socks, suit, sweater, t-shirt, top, tracksuit
4	kitchen ut./tool	adjustable spanner, axe, bottle, bowl, can opener, chisel, clamp, colander, crowbar, drill, electric kettle, file, filling knife, fork, fridge, glass, hammer, kettle, knife, lawn mower, level, microwave oven, mixer, mug, nail, nutcracker, oil can, oven, paint brush, pan, percolator, pickaxe, place mat, plate, pot, rope, saw, scales, scissors, screwdriver, shovel, sieve, spatula, spoon, stove, tongs, towel, vacuum cleaner, wheelbarrow, whisk, wire brush, wok, wrench
5	vehicle	airplane, boat, bus, car, helicopter, hovercraft, jeep, submarine, subway train, taxi, tractor, train, tram, truck (camion), truck (vrachtwagen)
6	weapon	bazooka, canon, club, dagger, double-barreled shotgun, grenade, knuckle dusters, machine gun, pistol, rifle, slingshot, spear, sword, tank, whip
7	tool	adjustable spanner, chisel, clamp, crowbar, file, filling knife, hammer, level, nail, oil can, plane, saw, screwdriver, tongs, wire brush, wrench

the six human categories underlying the artifact domain in the Dutch data. The correspondence is rather straightforward with the exception of factor 4, which contains object of "kitchen utensil" but also quite a few objects of "tool."

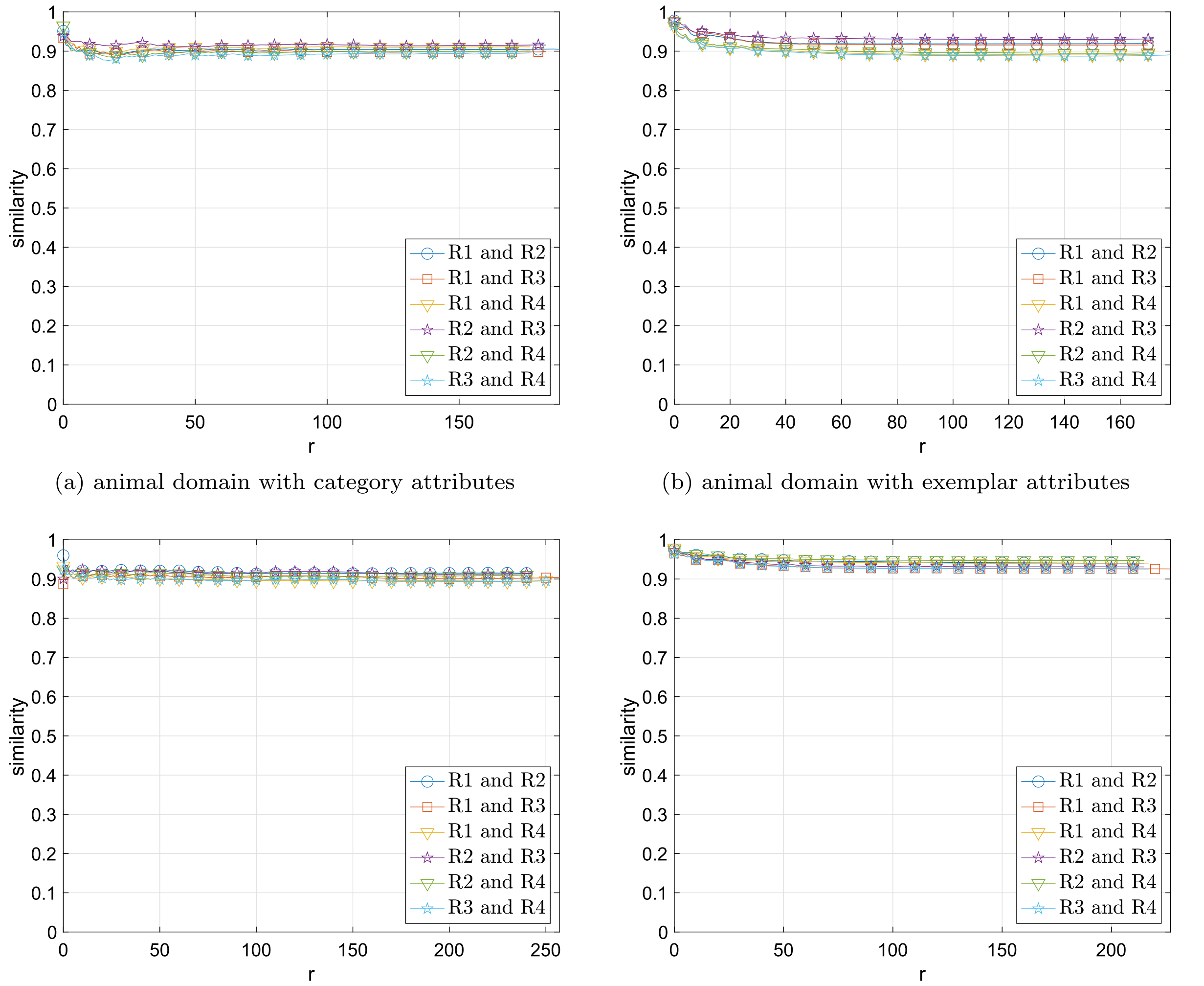
Even though the domain-based matrices are of our primary interest, let us note that we also observed the factors obtained for the smaller, category-based matrices with the exemplar attributes.[7] In many cases, particularly for the non-artifact categories, the first computed factors appear to have natural interpretations as well in that they represent naturally grouped collections of objects. For instance, we observed factors in the matrix describing "bird" that may be interpreted as "pet bird," "predatory bird," and "migratory bird."

3.3.3. Robustness of factorizations of respondents' matrices by GRECOND

In the preceding section, we observed that the collections $\mathcal{F}^{\geq 2}$ and $\mathcal{F}^{\geq 3}$ of the first seven factors computed from both the consensus animal domain matrices with the category attributes corresponding to thresholds ≥ 2 and ≥ 3, respectively, are considerably similar. The general question behind this observation is that of robustness of a factorization algorithm in the sense that the sets $\mathcal{F}_1$ and $\mathcal{F}_2$ of factors computed by the algorithm from two Boolean matrices, $I_1, I_2 \in \{0,1\}^{n\times m}$, representing possibly different but similar views of a certain domain of interest, respectively, are similar.

We shall use the measure $S(\mathcal{F},\mathcal{G})$ of similarity of two collections of factors, $\mathcal{F}$ and $\mathcal{G}$, designed in [6]. The values of $S(\mathcal{F},\mathcal{G})$ range from 0 to 1 with 1 indicating $\mathcal{F}=\mathcal{G}$. Put briefly, $S(\mathcal{F},\mathcal{G})$ expresses the extent to which for each factor in $\mathcal{F}$ there is a similar factor in $\mathcal{G}$ and vice versa, with the similarity of factors assessed by the well-known Jaccard similarity. We shall consider the domain-based matrices with both the category and the exemplar attributes filled by the four respondents, R1–R4, and inquire whether the

[7] For the category-based matrices, category attributes seem to be much less distinctive. This probably results in a less clear meaning of the factors obtained for the small matrices with category attributes.



(a) animal domain with category attributes (b) animal domain with exemplar attributes

(c) artifact domain with category attributes (d) artifact domain with exemplar attributes

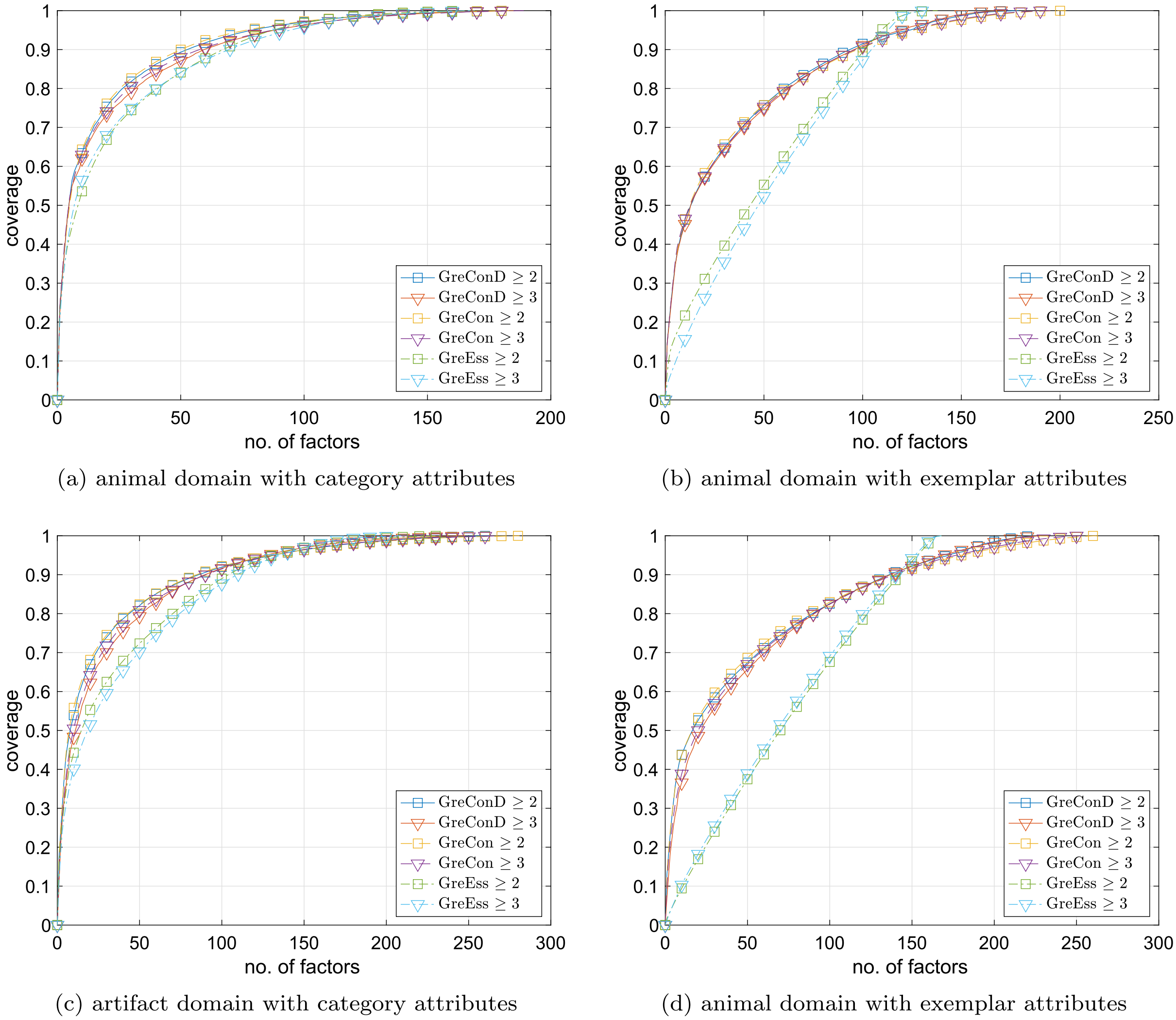
Fig. 4. Similarity of factorizations computed from respondents' domain-based matrices.

factorizations computed from the four matrices of each of the four types described in Table 2 are mutually similar. Since for each type, e.g., the animal domain with category attributes, the four matrices provided by R1–R4 may be regarded as representing four personal views of the same part of the world, the collections of factors computed from these matrices by a robust factorization method indeed should be mutually similar. Fig. 4 supports this intuition: The graphs of the mutual similarities $S(\mathcal{F}_i(r), \mathcal{F}_j(r))$ of the collections $\mathcal{F}_i(r)$ and $\mathcal{F}_j(r)$ of the first r factors extracted from the matrices filled by respondents Ri and Rj, respectively, indeed reveal a considerable similarity of the sets of r most important factors for almost every number r. Analogous observations can be made for the smaller, category-based matrices, which we do not presented in detail. Note also that these observations are related to the problem of handling noise in BMF and are congruent with previous observations regarding GRECOND [5].

3.4. Results for other algorithms utilizing formal concepts

In this section, we briefly discuss factorizations by other significant algorithms that utilize formal concepts as factors. For one, we intend to demonstrate that these algorithms also produce reasonably good factorizations with naturally interpretable factors. On the other hand, our experiments appear to point out what has not been observed before, namely considerable semantical differences between BMF algorithms.

In particular, we discuss factorizations of selected Dutch data matrices obtained by GRECON and GREESS. The GRECON algorithm was proposed in [7, Algorithm 1] but a very similar approach appeared earlier in [10] as an approach to a so-called tiling of Boolean data. The idea of GRECON is similar to that of GRECOND in that the algorithm selects formal concepts of I one-by-one in a greedy manner to maximize the number of the previously uncovered entries of I that contain 1. Unlike GRECOND, the GRECON algorithm selects the factors from the whole concept lattice associated to I, for which purpose all concepts of the concept lattice must first be computed. This idea may be implemented in a significantly more efficient way [18] compared to the original, direct proposition [7].



(a) animal domain with category attributes (b) animal domain with exemplar attributes

(c) artifact domain with category attributes (d) animal domain with exemplar attributes

Fig. 5. Factorizations of domain data by GreConD, GreCon, and GreEss.

Yet, computing the concept lattice presents a time-consuming step of the algorithm. The purpose of the GreEss algorithm, proposed in [4], is to utilize a deeper insight into the structure of formal concepts and to produce factorizations involving a small number of factors. This is achieved by first reducing the input matrix I to its so-called essential part, factorizing the essential part, and retrieving from this factorization the desired factorization of I.

The factorizations of the domain-based matrices are depicted in Fig. 5. They concern the consensus matrices for thresholds ≥ 2 and ≥ 3. One observes that both GreCon and GreEss display the usual pattern of coverage by the consecutively computed factors, which we also observed with the other domain-based matrices and the category-based matrices. In addition, while the graphs of GreCon's coverage are highly similar to those of GreConD, those of GreEss are slightly different, particularly for matrices with the exemplar attributes: As a result of GreEss's design, the coverage by GreEss does not grow as quickly for the first couple of factors which is compensated by reaching the full coverage, i.e., the exact decomposition of the input matrix, with a smaller number of factors. In addition, inspection of the meaning of the factors computed by both GreCon and GreEss revealed that, as with GreConD, the first couple of factors have an intuitive meaning. These features may be regarded as a confirmation with quality data of trends that have been observed by the previous studies.

However, while providing intuitively interpretable factors, neither GreCon nor GreEss were able to retrieve the human categories underlying the domain data as cleanly as the simple GreConD algorithm. This is particularly true of GreEss, in which case the apparent semantic drawback results from the approach employed by the algorithm, namely, removal of the technically inessential but semantically possibly significant 1s of the input matrix. The discussed property also applies to GreCon for which it may be regarded as perhaps pointing out a semantic advantage of the attribute-based strategy of GreConD over the unrestricted greedy strategy of GreCon. The situation is illustrated symbolically in Table 7, in which the symbol "≈" indicates a possible but questionable

Table 7
Interpretation of the first six factors obtained by GreConD, GreCon, and GreEss for the animal domain with category attributes and consensus threshold ≥ 2.

#	GreConD	GreCon	GreEss
1	typical animal	typical animal	typical animal
2	bird	≈bird	≈mammal
3	fish	fish	fish
4	mammal	≈insect	unclear
5	insect	bird	small bird
6	reptile	≈reptile	snake

interpretation. The possible semantic differences between BMF algorithms, which our observations indicate, are important from the practical viewpoint and hence worth a detailed exploration in future studies.

4. Conclusions and future research

We address the problem of semantic considerations pertaining to Boolean matrix factorization. We consider this a significant topic that the previous studies have not dealt with properly. In this first study, we examine the factorization algorithms that utilize formal concepts as factors, and utilize now available high-quality psychological data for our purpose.

The main conclusion of our inquiry is that in most of the examined cases, the inspected factorization methods produce good-quality factorizations with easily interpretable, natural factors. The GreConD algorithm, in particular, proved to have the ability to retrieve among the computed factors the human categories around which the testing data has been built. That is to say, to retrieve human categories that are inherently present in the data. In addition, the algorithm appears to be robust in that it tends to produce similar factorizations from matrices representing different views of the same domain of interest. Our considerations suggest several future research topics of which we mention the following three:

- Extend the present study to include further significant factorization algorithms. In particular, include algorithms that commit the so-called overcover error, such as the Asso algorithm [14]. A more comprehensive exploration of semantic properties of the algorithms shall help reveal important differences and features of the algorithms that have not been apparent in the previous studies. Our results provide first examples of such differences.
- Our study is enabled by the availability of the Dutch data—a high-quality psychological data involving human concepts, which is considerably more extensive compared to other data of similar purpose. While the Dutch data shall be used to examine other factorization algorithms, additional quality data is needed, but not presently available, to proceed with further semantic considerations as well as to better support the obtained conclusions. Obtaining such data collections represents a challenging important goal.
- Make semantic considerations a significant part in designing factorization algorithms. Semantic considerations may lead to a discovery of useful, psychologically justified views of what makes a factorization a good factorization. For this purpose, benchmark methods of quantitative nature need to be developed. In a broader perspective, pay attention to semantic considerations and to the utilization of psychological knowledge in explorations of Boolean data.

CRediT authorship contribution statement

Radim Belohlavek: Writing – original draft, Investigation, Formal analysis. **Martin Trnecka:** Writing – original draft, Investigation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] R. Belohlavek, T. Mikula, Dutch concepts, GitHub repository, Available at: https://github.com/mikulatomas/dutch-concepts, 2023.
[2] R. Belohlavek, T. Mikula, Are human categories formal concepts? A case study using Dutch data, Int. J. Gen. Syst. (2024), https://doi.org/10.1080/03081079.2024.2311917, in press.
[3] R. Belohlavek, J. Outrata, M. Trnecka, Toward quality assessment of Boolean matrix factorizations, Inf. Sci. 459 (2018) 71–85.

[4] R. Belohlavek, M. Trnecka, From-below approximations in Boolean matrix factorization: geometry and new algorithm, J. Comput. Syst. Sci. 81 (8) (2015) 1678–1697.
[5] R. Belohlavek, M. Trnecka, Handling noise in Boolean matrix factorization, Int. J. Approx. Reason. 96 (2018) 78–94 (preliminary version in Proc. ICJCAI 2017, 1433–1439).
[6] R. Belohlavek, M. Trnecka, Basic level in formal concept analysis: interesting concepts and psychological ramifications, in: Proc. IJCAI, 2013, pp. 1233–1239.
[7] R. Belohlavek, V. Vychodil, Discovery of optimal factors in binary data via a novel method of matrix decomposition, J. Comput. Syst. Sci. 76 (1) (2010) 3–20 (preliminary version in Proc. SCIS & ISCIS 2006).
[8] S. De Deyne, S. Verheyen, E. Ameel, W. Vanpaemel, M.J. Dry, W. Voorspoels, G. Storms, Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts, Behav. Res. Methods 40 (4) (2008) 1030–1048.
[9] B. Ganter, R. Wille, Formal Concept Analysis: Mathematical Foundations, Springer, 1999.
[10] F. Geerts, B. Goethals, T. Mielikäinen, Tiling databases, in: Proceedings of Discovery Science, 2004, pp. 278–289.
[11] K.H. Kim, Boolean Matrix Theory and Applications, Marcel Dekker, New York, 1982.
[12] P. Miettinen, The Discrete Basis Problem, MSc thesis, University of Helsinki, 2005.
[13] P. Miettinen, Matrix decomposition methods for data mining: Computational complexity and algorithms, PhD thesis, University of Helsinki, 2009.
[14] P. Miettinen, T. Mielikäinen, A. Gionis, G. Das, H. Mannila, The discrete basis problem, IEEE Trans. Knowl. Data Eng. 20 (10) (2008) 1348–1362 (preliminary version in Proc. PKDD 2006).
[15] P. Miettinen, S. Neumann, Recent developments in Boolean matrix factorization, in: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 2020, pp. 4922–4928.
[16] G.L. Murphy, The Big Book of Concepts, MIT Press, Cambridge, Massachusetts, 2002.
[17] W. Ruts, S. De Deyne, E. Ameel, W. Vanpaemel, T. Verbeemen, G. Storms, Dutch norm data for 13 semantic categories and 338 exemplars, Behav. Res. Methods Instrum. Comput. 36 (2004) 506–515.
[18] M. Trnecka, R. Vyjidacek, Revisiting the GreCon algorithm for Boolean matrix factorization, Knowl.-Based Syst. 249 (2022) 108895.
[19] R. Wille, Restructuring lattice theory: an approach based on hierarchies of concepts, in: I. Rival (Ed.), Ordered Sets, in: NATO Advanced Study Institutes Series, vol. 83, Kluwer, Dordrecht, 1982, pp. 445–482.