Optimal decompositions of matrices with entries from residuated lattices

Radim Belohlavek

Palacký University, 17. listopadu 12, 771 46 Olomouc, Czech Republic
e-mail: radim.belohlavek@upol.cz, phone: +420 585 634 705, fax: +420 585 634 716

Abstract

We describe optimal decompositions of matrices whose entries are elements of a residuated lattice L, such as $L = [0, 1]$. Such matrices represent relationships between objects and attributes with the entries representing degrees to which attributes represented by columns apply to objects represented by rows. Given such an $n \times m$ object-attribute matrix I, we look for a decomposition of I into a product $A \circ B$ of an $n \times k$ object-factor matrix A and a $k \times m$ factor-attribute matrix B with entries from L with the number k of factors as small as possible. We show that formal concepts of I, which play a central role in the Port-Royal approach to logic and which are the fixpoints of particular Galois connections associated to I, are optimal factors for decomposition of I in that they provide us with decompositions with the smallest number of factors. Moreover, we describe transformations between the space of original attributes and the space of factors induced by a decomposition $I = A \circ B$. The paper contains illustrative examples demonstrating the significance of the presented results for factor analysis of relational data. In addition, we present a general framework for a calculus of matrices with entries from residuated lattices in which both the matrix products and decompositions discussed in this paper as well as triangular products and decompositions discussed elsewhere can be regarded as two particular cases of a general type of product and decomposition. We present the results for matrices, i.e. for relations between finite sets in terms of relations, but the arguments behind are valid for relations between infinite sets as well.

Key words: matrix decomposition, residuated lattice, fixpoint, Galois connection, fuzzy logic
MSC codes: 15A33, 15A23, 15A04, 06F05

1. Introduction

1.1. Problem Setting and Paper's Content

Let I be an $n \times m$ object-attribute matrix whose entries are elements from a complete residuated lattice $\mathbf{L} = \langle L, \otimes, \rightarrow, \wedge, \vee, 0, 1 \rangle$ (see Section 1.3). That is, every matrix entry I_{ij} is an element from L. We look for a decomposition

$$I = A \circ B \tag{1}$$

of I into a product $A \circ B$ of an $n \times k$ object-factor matrix A and a $k \times m$ factor-attribute matrix B, with $A_{il}, B_{lj} \in L$, such that the number k of factors is the smallest possible. Note that in Boolean matrix theory, the number k is called the Schein rank [23]. The composition operator $\circ$ is the sup-$\otimes$ composition defined by

$$I_{ij} = \bigvee_{l=1}^{k} A_{il} \otimes B_{lj} \tag{2}$$

with $\bigvee$ denoting the supremum in $\mathbf{L}$. Note that if $L = \{0,1\}$ then $a \otimes b = \min(a,b)$, $\bigvee$ is maximum, and $A \circ B$ is the max-min product of Boolean matrices. If $L = [0,1]$ and $\otimes$ is a t-norm then $A \circ B$ is the max-t-norm product of matrices (fuzzy relations) known from fuzzy set theory [19].

A related type of decomposition, namely

$$I = A \triangleleft B \tag{3}$$

with $A_{il}, B_{lj} \in L$ the same as above and $\triangleleft$ defined by

$$I_{ij} = \bigwedge_{l=1}^{k} A_{il} \to B_{lj} \tag{4}$$

was studied in [8]. In (4), $\bigwedge$ denotes the infimum in $\mathbf{L}$ and the composition operator is called a triangular composition or an inf-$\to$ composition.

In this paper, we are primarily interested in decompositions based on operator $\circ$ because this operator is well known and, as explained in Section 1.2 and demonstrated in Section 4, the decompositions have a natural, easy-to-understand interpretation in terms of factor analysis. In Section 2 and 3, we present selected results on optimal decompositions based on $\circ$ and the corresponding transformations between the spaces of attributes and factors. In Section 5, we present a general framework for a calculus of matrices with entries from residuated lattices in which both the matrix product based on $\circ$ and the one based on $\triangleleft$ are particular cases of a more general product. We show that the results presented in Section 2 and 3 of this paper as well as the results from [8] are particular cases of the results that we work out in the general framework. Hence, proofs are omitted in Section 2 and 3. Instead, we provide references in these sections to the general results and explanatory remarks of Section 5. In Section 6, we conclude the paper and provide an outline of topics for future research.

1.2. Motivation

Residuated lattices can be thought of as partially ordered scales of degrees. An entry $I_{ij} \in L$ of I can be interpreted as a degree to which attribute j (such as "good performance" or "dizziness") applies to object i (such as "product" or "patient"). For $L = \{0,1\}$, in which case I is a Boolean matrix, decompositions $I = A \circ B$ are sought in Boolean factor analysis, see e.g. [13, 15, 25, 29], and data mining, see e.g. [26]. In general, a decomposition $I = A \circ B$ represents a factor model according to which the relationship between n objects (rows of I) and m attributes (columns of I), which is represented by I, is explained by a relationship between the n objects and k new factors, which is represented by A, and a relationship between the k factors and the m original attributes, which is represented by B. Namely, A_{il} can be interpreted as a degree to which factor l applies to object i, and B_{lj} can be interpreted as a degree to which attribute j is a manifestation of factor l. In fuzzy logic, $\bigvee$ and $\otimes$ correspond to the existential quantifier and conjunction, respectively [19, 21]. As a result, (2) implies that according to the factor model given by $I = A \circ B$, the degree to which object i has attribute j can be interpreted as the degree to which there exists a factor l such that l applies to i and such that attribute j is a particular manifestation of factor l. Therefore, a decomposition $I = A \circ B$ provides us with a factor model for data with graded (gradual, fuzzy) attributes. We include an illustrative example in Section 4.

1.3. Preliminaries from residuated lattices

A residuated lattice [19, 22, 28] is an algebra $\mathbf{L} = \langle L, \wedge, \vee, \otimes, \rightarrow, 0, 1\rangle$ such that $\langle L, \wedge, \vee, 0, 1\rangle$ is a lattice with 0 and 1 being the least and greatest element of L, respectively; $\langle L, \otimes, 1\rangle$ is a commutative monoid (i.e. $\otimes$ is commutative, associative, and $a \otimes 1 = 1 \otimes a = a$ for each $a \in L$); $\otimes$ and $\rightarrow$ satisfy adjointness:

$$a \otimes b \leq c \quad \text{iff} \quad a \leq b \rightarrow c \tag{5}$$

for each $a, b, c \in L$. $\mathbf{L}$ is called complete if $\langle L, \wedge, \vee, 0, 1\rangle$ is a complete lattice.

Residuated lattices appear in various areas of mathematics and play a fundamental role in many-valued logics, particularly in fuzzy logic and fuzzy set theory [18, 19, 20, 21]. In fuzzy logic, elements a of L are called truth degrees (or grades). $\otimes$ and $\rightarrow$ are (truth functions of) many-valued conjunction and implication. Examples of residuated lattices include those with the support set $L = [0, 1]$ (real unit interval), $\wedge$ and $\vee$ being minimum and maximum, $\otimes$ being a left-continuous t-norm with the corresponding residuum $\rightarrow$ [19]. Three most important pairs of adjoint operations on $[0, 1]$ are:

$$\text{Łukasiewicz:} \qquad \begin{aligned} a \otimes b &= \max(a + b - 1, 0), \\ a \rightarrow b &= \min(1 - a + b, 1), \end{aligned} \tag{6}$$

$$\text{Gödel:} \qquad \begin{aligned} a \otimes b &= \min(a, b), \\ a \rightarrow b &= \begin{cases} 1 & \text{if } a \leq b, \\ b & \text{otherwise,} \end{cases} \end{aligned} \tag{7}$$

$$\text{Goguen:} \qquad \begin{aligned} a \otimes b &= a \cdot b, \\ a \rightarrow b &= \begin{cases} 1 & \text{if } a \leq b, \\ \frac{b}{a} & \text{otherwise.} \end{cases} \end{aligned} \tag{8}$$

Another commonly used example is a finite linearly ordered $\mathbf{L}$. For instance, one can put $L = \{a_0 = 0, a_1, \ldots, a_n = 1\} \subseteq [0, 1]$ ($a_0 < \cdots < a_n$) with $\otimes$ given by $a_k \otimes a_l = a_{\max(k+l-n,0)}$ and the corresponding $\rightarrow$ given by $a_k \rightarrow a_l = a_{\min(n-k+l,n)}$. Such an $\mathbf{L}$ is called a finite Łukasiewicz chain. Another possibility is a finite Gödel chain which consists of L and restrictions of Gödel operations on $[0, 1]$ to L. A special case of a residuated lattice is the two-element Boolean algebra $\langle\{0,1\},\wedge,\vee,\otimes,\rightarrow,0,1\rangle$, denoted by $\mathbf{2}$, which is the structure of truth degrees of classical logic. That is, the operations $\wedge,\vee,\otimes,\rightarrow$ of $\mathbf{2}$ are the truth functions of the corresponding connectives of classical logic.

Given a residuated lattice $\mathbf{L}$, we define the usual notions [18, 19]: an L-set (fuzzy set, graded set) A in a universe U is a mapping $A\colon U \rightarrow L$, $A(u)$ being interpreted as "the degree to which u belongs to A". L^U (or $\mathbf{L}^U$ if it is desirabe to make the structure of $\mathbf{L}$ explicit) denotes the collection of all L-sets in U. The operations with L-sets are defined componentwise. For instance, the intersection of L-sets $A, B \in L^U$ is an L-set $A \cap B$ in U such that $(A \cap B)(u) = A(u) \wedge B(u)$ for each $u \in U$, etc. $\mathbf{2}$-sets and operations with $\mathbf{2}$-sets can be identified with ordinary sets and operations with ordinary sets, respectively. Binary L-relations (binary fuzzy relations) between X and Y can be thought of as L-sets in the universe $X \times Y$. For L-sets A and B in universe U, we put

$$A \subseteq B \text{ if and only if } A(u) \leq B(u) \text{ for each } u \in U. \tag{9}$$

In this case, we say that A is included in B.

We use well-known properties of residuated lattices and fuzzy sets over residuated lattices which can be found, e.g., in [19, 21, 22].

2. Optimal Decompositions

2.1. Matrix composition as a $\bigvee$-superposition of rectangular matrices

Observe first that $I = A \circ B$ for $n \times k$ and $k \times m$ matrices A and B, in fact, means that I is a $\bigvee$-superposition of particular matrices we call rectangular.

Definition 1. *An $n \times m$ matrix J is called rectangular iff there exist L-sets C in $\{1,\ldots,n\}$ and D in $\{1,\ldots,m\}$ such that $J = C \otimes D$, i.e.*

$$J_{ij} = C(i) \otimes D(j) \tag{10}$$

for $1 \leq i \leq n$, $1 \leq j \leq m$.

For brevity, we say just "rectangle" instead of "rectangular matrix". The term comes from a geometric interpretation. For illustration, consider $L = \{0,1\}$. The fact that J is a rectangular binary matrix means that the entries of J which contain 1s form a rectangular area such as

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

or can be brought into a rectangular area by permuting rows and columns. In the above example, $J = C \otimes D$ where C and D are characteristic functions of $\{3,4,5,6\}$ and $\{3,4,5\}$, respectively. The next lemma is a paritular case of Lemma 2 (see Remark 2).

Lemma 1. *$I = A \circ B$ for $n \times k$ and $k \times m$ matrices A and B iff I is a $\bigvee$-superposition of k rectangular matrices $J_1, \dots, J_k$, i.e. iff*

$$I = J_1 \vee J_2 \vee \cdots \vee J_k.$$

Example 1. Consider $L = \{0, 0.1, \dots, 0.9, 1\}$, $a \otimes b = \min(a, b)$, and the following decomposition $I = A \circ B$:

$$\begin{pmatrix} 1.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 0.0 & 0.0 & 1.0 & 0.2 \\ 1.0 & 1.0 & 1.0 & 1.0 & 0.0 & 0.8 \\ 1.0 & 0.2 & 0.0 & 0.0 & 1.0 & 0.2 \end{pmatrix} = \begin{pmatrix} 1.0 & 0.0 & 0.0 & 1.0 \\ 1.0 & 0.0 & 1.0 & 0.7 \\ 0.8 & 1.0 & 0.0 & 0.9 \\ 0.2 & 0.0 & 1.0 & 0.0 \end{pmatrix} \circ \begin{pmatrix} 1.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 1.0 & 1.0 & 0.0 & 0.8 \\ 1.0 & 0.0 & 0.0 & 0.0 & 1.0 & 0.2 \\ 0.8 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{pmatrix}.$$

According to Lemma 1, I is a $\bigvee$-superposition of four matrices, J_1, J_2, J_3, J_4 where J_l is a $\circ$-product of the l-th column of A and the l-th row of B, i.e.

$$\begin{pmatrix} 1.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 0.0 & 0.0 & 1.0 & 0.2 \\ 1.0 & 1.0 & 1.0 & 1.0 & 0.0 & 0.8 \\ 1.0 & 0.2 & 0.0 & 0.0 & 1.0 & 0.2 \end{pmatrix} = \begin{pmatrix} 1.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.8 & 0.8 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.2 & 0.2 & 0.0 & 0.0 & 0.0 & 0.0 \end{pmatrix} \vee$$

$$\begin{pmatrix} 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 1.0 & 1.0 & 0.0 & 0.8 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{pmatrix} \vee \begin{pmatrix} 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 0.0 & 0.0 & 0.0 & 1.0 & 0.2 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 0.0 & 0.0 & 0.0 & 1.0 & 0.2 \end{pmatrix} \vee \begin{pmatrix} 0.8 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.7 & 0.7 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.8 & 0.9 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{pmatrix}.$$

□

2.2. Formal concepts of I as optimal factors for decomposition of I

In this section, we describe decompositions of I which are optimal among all possible decompositions of I in the sense that the number k of factors is the smallest possible. The decompositions use so-called formal concepts of I as factors, which are fixpoints of particular Galois connections associated to I. Moreover, it follows from the proof (see the proof of Theorem 5), any decomposition of I can be extended to at least as good a decomposition which uses formal concepts as factors.

Formal concepts of I. Formal concepts of data tables describing a relationship between objects and attributes are studied in formal concept analysis (FCA) [16]. In the basic setting, FCA deals with data with binary attributes, i.e. with binary matrices I. An extension of FCA which deals with matrices I with entries from residuated lattices has been developed in a series of papers, see e.g. [2, 4, 6, 7]. The basic notions we need are presented below.

Let $X = \{1, \dots, n\}$ and $Y = \{1, \dots, m\}$ be sets (of objects and attributes, respectively), I be an $n \times m$ matrix with entries from a residuated lattice $\mathbf{L} = \langle L, \otimes, \to, \wedge, \vee, 0, 1 \rangle$. Consider the operators ${}^{\uparrow} : L^X \to L^Y$ and ${}^{\downarrow} : L^Y \to L^X$ defined by

$$C^{\uparrow}(y) = \textstyle\bigwedge_{x \in X}(C(x) \to I_{xy}) \quad \text{and} \quad D^{\downarrow}(x) = \textstyle\bigwedge_{y \in Y}(D(y) \to I_{xy}). \tag{11}$$

That is, ${}^{\uparrow}$ assigns an L-set $C^{\uparrow}$ in Y to a given L-set C in X, and ${}^{\downarrow}$ assigns an L-set $D^{\uparrow}$ in Y to a given L-set D in Y. According to the basic principles of first-order fuzzy logic [21], $C^{\uparrow}(y)$ is just the truth degree of the following proposition: "for each object $x \in X$: if x is from C then x has attribute y". Likewise, $D^{\downarrow}(x)$ is the truth degree of "for each attribute $y \in Y$: if y is from D then x has attribute y". The operators $^{\uparrow}$ and $^{\downarrow}$ form a fuzzy Galois connection [2], and the compound operators $^{\uparrow\downarrow}$ and $^{\downarrow\uparrow}$ form particular closure operators in X and Y, respectively [4]. A pair $\langle C, D\rangle$ consisting of an L-set C in X and an L-set D in Y is called a formal concept of I if $C^{\uparrow} = D$ and $D^{\downarrow} = C$, i.e. if $\langle C, D\rangle$ is a fixpoint of $^{\uparrow}$ and $^{\downarrow}$. C and D are called the extent and intent of $\langle C, D\rangle$. For an object x, $C(x)$ is interpreted as a degree to which formal concept $\langle C, D\rangle$ applies to x; for an attribute y, $D(y)$ is interpreted as a degree to which $\langle C, D\rangle$ applies to y. The set of all formal concepts of I is denoted by $\mathcal{B}(X, Y, I)$, i.e.

$$\mathcal{B}(X, Y, I) = \{\langle C, D\rangle \in L^X \times L^Y \mid C^{\uparrow} = D,\ D^{\downarrow} = C\}.$$

A partial order $\leq$ defined by

$$\langle C_1, D_1\rangle \leq \langle C_2, D_2\rangle \text{ iff } C_1 \subseteq C_2 \text{ (iff } D_2 \subseteq D_1)$$

for $\langle C_1, D_1\rangle, \langle C_2, D_2\rangle \in \mathcal{B}(X, Y, I)$, makes $\mathcal{B}(X, Y, I)$ a complete lattice, called the concept lattice of I [7]. Note that $\subseteq$ is defined by (9). For $L = \{0, 1\}$, $\mathcal{B}(X, Y, I)$ coincides with the ordinary concept lattice [16]. Efficient algorithms for computing $\mathcal{B}(X, Y, I)$ exist [10].

Matrices $A_{\mathcal{F}}$ and $B_{\mathcal{F}}$. Let

$$\mathcal{F} = \{\langle C_1, D_1\rangle, \ldots, \langle C_k, D_k\rangle\}$$

be a set of pairs of L-sets C_l and D_l in $\{1, \ldots, n\}$ and $\{1, \ldots, m\}$, respectively, with values from L. In what follows, we always assume that there is a fixed order on the set $\mathcal{F}$ and indicate this order by indices. Thus, we may speak of the first pair in $\mathcal{F}$ which is $\langle C_1, D_1\rangle$, up to the k-th pair which is $\langle C_k, D_k\rangle$. Given $\mathcal{F}$ with such a fixed order, define $n \times k$ and $k \times m$ matrices $A_{\mathcal{F}}$ and $B_{\mathcal{F}}$ by

$$(A_{\mathcal{F}})_{il} = (C_l)(i) \quad \text{and} \quad (B_{\mathcal{F}})_{lj} = (D_l)(i).$$

That is, the l-th column of $A_{\mathcal{F}}$ is the transpose of the vector corresponding to L-set C_l and the l-th row of $B_{\mathcal{F}}$ is the vector corresponding to D_l. (The vectors corresponding to C_l and D_l are $(C_l(1), \ldots, C_l(n))$ and $(D_l(1), \ldots, D_l(m))$.)

Example 2. Let $X = \{1, \ldots, 4\}$, $Y = \{1, \ldots, 6\}$. Let $\mathcal{F} = \{\langle C_1, D_1\rangle, \langle C_2, D_2\rangle\}$ with the vectors corresponding to C_1 and D_1 being (1.0 1.0 0.8 0.2) and (1.0 1.0 0.0 0.0 0.0 0.0), and the vectors corresponding to C_2 and D_2 being (1.0 0.7 0.9 0.0) and (0.8 1.0 0.0 0.0 0.0 0.0). That is, $C_1(1) = 1.0$, $C_1(2) = 1.0$, $C_1(3) = 0.8$, etc. For the matrices $A_{\mathcal{F}}$ and $B_{\mathcal{F}}$ we have

$$A_{\mathcal{F}} = \begin{pmatrix} 1.0 & 1.0 \\ 1.0 & 0.7 \\ 0.8 & 0.9 \\ 0.2 & 0.0 \end{pmatrix} \quad \text{and} \quad B_{\mathcal{F}} = \begin{pmatrix} 1.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.8 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{pmatrix}.$$

□

Next, we show the role of formal concepts of I for decompositions of I. The first theorem says that for every I, there exists a decomposition of I in which formal concepts of I are used as factors. The theorem is a particular instance of Theorem 4 (see Remark 3 (b)).

Theorem 1 (universality of formal concepts as factors)**.** *For every I with entries from a residuated lattice L there exists a finite set $\mathcal{F} \subseteq \mathcal{B}(X, Y, I)$ such that $I = A_{\mathcal{F}} \circ B_{\mathcal{F}}$.*

However, Theorem 1 and its proof (see the proof of Theorem 4) only yield that the number $k = |\mathcal{F}|$ of factors equals the number of attributes (columns) of I, i.e. $k = m$. Obviously, there can be better decompositions, i.e. those with $k < m$. The next theorem shows optimality of using formal concepts as factors for decomposition of I. The theorem is a particular instance of Theorem 5 (see Remark 4 (a)).

Theorem 2 (optimality of formal concepts as factors)**.** *Let $I = A \circ B$ for $n \times k$ and $k \times m$ matrices A and B with entries from L. Then there exists a finite set $\mathcal{F} \subseteq \mathcal{B}(X, Y, I)$ with*

$$|\mathcal{F}| \leq k$$

such that for the $n \times |\mathcal{F}|$ and $|\mathcal{F}| \times m$ matrices $A_{\mathcal{F}}$ and $B_{\mathcal{F}}$ we have

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

From the computational point of view, Theorem 1 and Theorem 2 say that when looking for factors for decompositions of I, we can confine ourselves to formal concepts from $\mathcal{B}(X, Y, I)$.

3. Transformations Between Spaces of Attributes and Factors

In this section, we provide basic results and considerations regarding natural transformations between the m-dimensional space of attributes and the k-dimensional space of factors which are induced by decomposition (1), particularly by matrix B describing a relationship between factors and attributes. Further results are provided within the general framework in Section 5.3.

To facilitate our discussion, we identify the set L^Y of all L-sets in Y with the set L^m of all m-dimensional vectors of grades, i.e. we identify an L-set $P : \{1, \ldots, m\} \to L$ with a vector $(P(1), \ldots, P(m))$. Likewise, we identify an L-set $Q : \{1, \ldots, k\} \to L$ with $(Q(1), \ldots, Q(k))$.

Let thus $I = A \circ B$. In general, we do not assume that $A = A_{\mathcal{F}}$ and $B = B_{\mathcal{F}}$ for some set $\mathcal{F}$ of formal concepts of I. Consider the transformations $g : L^m \to L^k$ and $h : L^k \to L^m$ defined for $P \in L^m$ and $Q \in L^k$ by

$$(g(P))_l = \textstyle\bigwedge_{j=1}^{m} (B_{lj} \to P_j), \qquad (12)$$

$$(h(Q))_j = \textstyle\bigvee_{l=1}^{k} (Q_l \otimes B_{lj}), \qquad (13)$$

for $1 \leq l \leq k$ and $1 \leq j \leq m$. Let us first observe that even on the domains where it makes sense to consider linearity, g and h are non-linear. The following example demonstrates this fact in the case of h.

Example 3. Let $L = [0, 1]$ be equipped with Łukasiewicz t-norm. Let $I = A \circ B$ be

$$\begin{pmatrix} 0.3 & 0.0 & 0.1 \\ 0.3 & 0.7 & 0.5 \\ 0.5 & 0.8 & 0.6 \end{pmatrix} = \begin{pmatrix} 0.2 & 0.8 \\ 0.9 & 0.8 \\ 1.0 & 1.0 \end{pmatrix} \circ \begin{pmatrix} 0.4 & 0.8 & 0.6 \\ 0.5 & 0.2 & 0.3 \end{pmatrix}.$$

Then for $Q_1 = (0.6\ 0.2)$ and $Q_2 = (0.4\ 0.3)$ we have $h(Q_1 + Q_2) = (Q_1 + Q_2) \circ B = (1.0\ 0.5) \circ B = (0.4\ 0.8\ 0.6) \neq (0.0\ 0.6\ 0.2) = (0.0\ 0.4\ 0.2) + (0.0\ 0.2\ 0.0) = Q_1 \circ B + Q_2 \circ B = h(Q_1) + h(Q_2)$. $\square$

$I = A \circ B$ provides a representation of object i by $I_{i\_}$ (i-th row of I) in the space L^m of attributes, and a representation of i by $A_{i\_}$ (i-th row of A) in the space L^k of factors. The next theorem describes basic properties of g and h in case the decomposition of I involves formal concepts of I as factors.

Theorem 3. *Let $I = A_{\mathcal{F}} \circ B_{\mathcal{F}}$ for some $\mathcal{F} \subseteq \mathcal{B}(X, Y, I)$. Then*

$$g(I_{i\_}) = A_{i\_} \quad \textit{and} \quad h(A_{i\_}) = I_{i\_}$$

for every i. Moreover, $B_{\mathcal{F}}$ is the largest of the matrices D for which $I = A_{\mathcal{F}} \circ D$. Likewise, $A_{\mathcal{F}}$ is the largest of the matrices C for which $I = C \circ B_{\mathcal{F}}$.

In addition to the fact that formal concepts as factors are easy to interpret (see Section 4), Theorem 3 shows another reason to look for decompositions that use formal concepts. (a) Such approach guarantees that g and h transform rows of I to rows of A and vice versa. (b) Maximality conditions for $A_{\mathcal{F}}$ and $B_{\mathcal{F}}$ guarantee that the grades to which factors apply to objects and attributes which are implied by the decomposition $I = A_{\mathcal{F}} \circ B_{\mathcal{F}}$ are actually the largest grades for which the factor model still faithfully reconstructs the data represented by I. That is, we avoid models which reconstruct data but provide only lower estimations of grades to which factors apply to objects and attributes. This feature is desirable because otherwise the interpretation of factors might be difficult. As an example, if an attribute is a manifestation of two distinct factors, we may remove it from the first factor (lower the grade to which the factor applies to the attribute) and still have factors which represent the data. Doing so, the first factor becomes "unnatural" (consider factor "speed" from which remove attribute "good performance in 100 m", cf. Section 4). (c) If we subscribe to preference of large As and Bs over smaller ones, as suggested by (b), Theorem 3 can be seen as claiming uniqueness of $B_{\mathcal{F}}$ given $A_{\mathcal{F}}$ and uniqueness of $A_{\mathcal{F}}$ given $B_{\mathcal{F}}$.

The following are the basic properties of g and h (they can be easily checked, (14)–(17) follow from the general results of Section 5.3).

$$
\begin{aligned}
g(\bigwedge_{s\in S} P_s) &= \bigwedge_{s\in S} g(P_s), && (14)\\
h(\bigvee_{t\in T} Q_t) &= \bigvee_{t\in T} h(Q_t), && (15)\\
h(g(P)) &\leq P, && (16)\\
Q &\leq g(h(Q)), && (17)\\
g(a \to P) &= a \to g(P), && (18)\\
h(a \otimes Q) &= a \otimes h(Q), && (19)\\
P \approx P' &\leq g(P) \approx g(P'), && (20)\\
Q \approx Q' &\leq g(Q) \approx g(Q'). && (21)
\end{aligned}
$$

(14) and (15) say that g and h are $\bigwedge$- and $\bigvee$-preserving morphisms between L^m and L^k. An immediate consequence is that

$$
\begin{aligned}
&P \leq P' \text{ implies } g(P) \leq g(P'), && (22)\\
&Q \leq Q' \text{ implies } h(Q) \leq h(Q'), && (23)
\end{aligned}
$$

where $P \leq P'$ means $P_j \leq P'_j$ for all j, and $Q \leq Q'$ means $Q_l \leq Q'_l$ for all l. Properties (22) and (23) are can be regarded as requirements for reasonable transformations between the attribute space and the factor space. Namely, (22) says that the more attributes an object has, the more factors apply, while (23) says that the more factors apply, the more attributes an object has. This is in accordance with the structure of our factor model given by (1). (16) and (17) say that the compositions gh and hg are extensive and intensive, and as is argued in Section 5.3, form a particular closure and interior operator, respectively. (18) and (19) say that g and h are compatible with $\to$-multiplication and $\otimes$-multiplication, respectively. Note that $a \to P$ and $a \otimes Q$, called the $\to$-multiplication of P and $\otimes$-multiplication of Q, are defined by $(a \to P)_j = a \to P_j$ and $(a \otimes Q)_l = a \otimes Q_l$, respectively. To sum up, g is a $\bigwedge$-morphism that preserves $\to$-multiplication and h is a $\bigvee$-morphism that preserves $\otimes$-multiplication. (20) and (21) say that g and h preserve certain natural similarities on L^m and L^k. Namely, for $P, P' \in L^m$ and $Q, Q' \in L^k$, let

$$
P \approx P' = \bigwedge_{j=1}^{m} (P_j \leftrightarrow P'_j), \qquad Q \approx Q' = \bigwedge_{l=1}^{k} (Q_l \leftrightarrow Q'_l),
$$

where $\leftrightarrow$ is the so-called biresiduum defined by $a \leftrightarrow b = (a \to b) \wedge (b \to a)$. For example, for Łukasiewicz t-norm, the corresponding biresiduum is $a \leftrightarrow b = 1 - |a - b|$. The value of $a \leftrightarrow b$ can be interpreted as the grade to which grades a and b are close and, therefore, $P \approx P'$ can be seen as the least grade to which all the coordinates of P and P' are similar, i.e. a degree of similarity between P and P'. It is well-known in fuzzy set theory that $P \approx P' = 1$ iff $P = P'$, $P \approx P' = P' \approx P$, and that $(P \approx P') \otimes (P' \approx P'') \leq (P \approx P'')$, i.e. that $\approx$ is an $\mathbf{L}$-equality, see e.g. [19]. (20) and (21) show that g and h preserve these similarities, i.e. images of two vectors are at least as similar as the two vectors.

Furthermore, Fig. 1 illustrates the following property of g and h. The space L^m of attributes and the space L^k of factors are partitioned into an equal number of convex

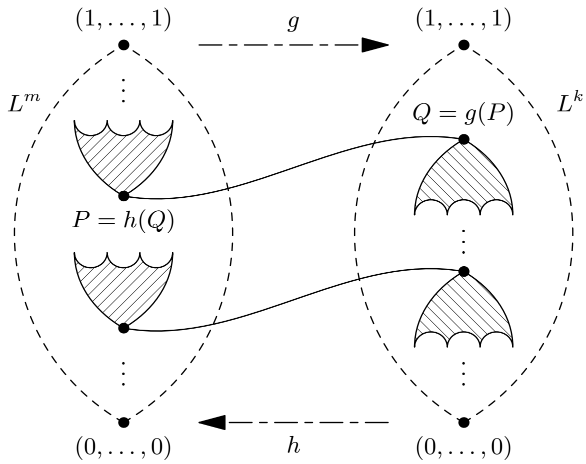
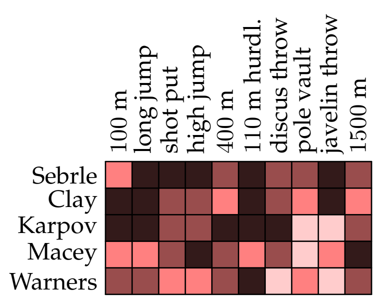


Figure 1: Illustration of property of g and h.

subspaces. The subspaces of the attribute space have least elements, the subspaces of the factor space have greatest elements. One can pair the subspaces in such a way that g maps all vectors of the subspace U of the attribute space to the largest element of the corresponding subspace V of the factor space and conversely, h maps all vectors from V to the least vector from U. This property is proved in the general framework in Section 5.3 (see Theorem 8).

4. Illustrative example

In order to demonstrate usefulness of the decompositions dealt with in this paper, we now shortly present an illustrative example which is presented in full in [12]. Consider the following 5×10 matrix I:

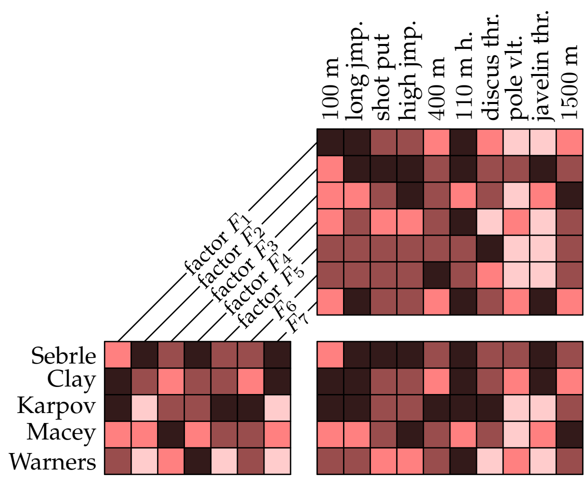


The matrix represents results of top five athletes in 2004 Olympic Games decathlon. The rows and columns correspond to the athletes and decathlon disciplines, respectively. The entries are colored boxes which represent degrees from a five-element Łukasiewicz chain with $L = \{0, 0.25, 0.5, 0.75, 1\}$ (the darker the color, the larger the degree), see Section 1.3. The matrix was obtained by a straightforward transformation to the degrees from L of the actual scores assigned to the athletes according to the IAAF Scoring Tables. This means that the above matrix represents the athletes' scores only approximately (due to rounding,

F_i	*Extent C_i*	*Intent D_i*
F_1	$\{{}^{.5}/\text{Sebrle}, \text{Clay}, \text{Karpov}, {}^{.5}/\text{Macey}, {}^{.75}/\text{Warners}\}$	$\{10, \text{lj}, {}^{.75}/\text{sp}, {}^{.75}/\text{hj}, {}^{.5}/40, 11, {}^{.5}/\text{di}, {}^{.25}/\text{pv}, {}^{.25}/\text{ja}, {}^{.5}/15\}$
F_2	$\{\text{Sebrle}, {}^{.75}/\text{Clay}, {}^{.25}/\text{Karpov}, {}^{.5}/\text{Macey}, {}^{.25}/\text{Warners}\}$	$\{{}^{.5}/10, \text{lj}, \text{sp}, \text{hj}, {}^{.75}/40, 11, {}^{.75}/\text{di}, {}^{.75}/\text{pv}, \text{ja}, {}^{.75}/15\}$
F_3	$\{{}^{.75}/\text{Sebrle}, {}^{.5}/\text{Clay}, {}^{.75}/\text{Karpov}, \text{Macey}, {}^{.5}/\text{Warners}\}$	$\{{}^{.5}/10, {}^{.5}/\text{lj}, {}^{.75}/\text{sp}, \text{hj}, {}^{.75}/40, {}^{.5}/11, {}^{.75}/\text{di}, {}^{.25}/\text{pv}, {}^{.5}/\text{ja}, 15\}$
F_4	$\{\text{Sebrle}, {}^{.75}/\text{Clay}, {}^{.75}/\text{Karpov}, {}^{.5}/\text{Macey}, \text{Warners}\}$	$\{{}^{.5}/10, {}^{.75}/\text{lj}, {}^{.5}/\text{sp}, {}^{.5}/\text{hj}, {}^{.75}/40, 11, {}^{.25}/\text{di}, {}^{.5}/\text{pv}, {}^{.25}/\text{ja}, {}^{.75}/15\}$
F_5	$\{{}^{.75}/\text{Sebrle}, {}^{.75}/\text{Clay}, \text{Karpov}, {}^{.75}/\text{Macey}, {}^{.25}/\text{Warners}\}$	$\{{}^{.75}/10, {}^{.75}/\text{lj}, {}^{.75}/\text{sp}, {}^{.75}/\text{hj}, {}^{.75}/40, {}^{.75}/11, \text{di}, {}^{.25}/\text{pv}, {}^{.25}/\text{ja}, {}^{.75}/15\}$
F_6	$\{{}^{.75}/\text{Sebrle}, {}^{.5}/\text{Clay}, \text{Karpov}, {}^{.75}/\text{Macey}, {}^{.75}/\text{Warners}\}$	$\{{}^{.75}/10, {}^{.75}/\text{lj}, {}^{.75}/\text{sp}, {}^{.75}/\text{hj}, 40, {}^{.75}/11, {}^{.5}/\text{di}, {}^{.25}/\text{pv}, {}^{.25}/\text{ja}, {}^{.75}/15\}$
F_7	$\{\text{Sebrle}, \text{Clay}, {}^{.25}/\text{Karpov}, {}^{.5}/\text{Macey}, {}^{.25}/\text{Warners}\}$	$\{{}^{.5}/10, \text{lj}, {}^{.75}/\text{sp}, {}^{.75}/\text{hj}, {}^{.5}/40, 11, {}^{.75}/\text{di}, {}^{.5}/\text{pv}, \text{ja}, {}^{.5}/15\}\rangle$

Table 1: Factor Formal Concepts for Decathlon Data. Legend: 10—100 m; *lj*—long jump; *sp*—shot put; *hj*—high jump; 40—400 m; 11—110 m hurdles; *di*—discus throw; *pv*—pole vault; *ja*—javelin throw; 15—1500 m.

two close scores are assigned the same grade) but such an approximation is sufficient for the purpose of illustration. Matrix I can decomposed into a product $I = A_{\mathcal{F}} \circ B_{\mathcal{F}}$ of a 5×7 matrix $A_{\mathcal{F}}$ and a 7×10 matrix $B_{\mathcal{F}}$, in which $\mathcal{F} = \{\langle C_l, D_l \rangle \mid l = 1, \ldots, 7\}$ is a set of formal concepts of I. In [12], an efficient approximation algorithm for computing optimal $\circ$-decompositions is provided. The decomposition presented below is computed by this algorithm:



Matrix $A_{\mathcal{F}}$ is the bottom-left matrix with athletes' names labeling the rows; matrix $B_{\mathcal{F}}$ is the top matrix with disciplines' names labeling the columns. As described in Section 2, the l-th column of $A_{\mathcal{F}}$ and the l-th row of $B_{\mathcal{F}}$ are the vectors corresponding to the extent C_l and the intent D_l, respectively, of the l-th factor $F_l = \langle C_l, D_l \rangle$ $(l = 1, \ldots, k)$. The formal concepts (factors) from $\mathcal{F}$ are depicted with a detialed description in Table 1.

The first line of Fig. 2 shows the rectangular matrices (see Definition 1) corresponding to formal concepts from $\mathcal{F} = \{F_1, \ldots, F_7\}$. The second line of Fig. 2 shows the $\bigvee$-superpositions $F_1 \vee \cdots \vee F_l$ of the first l factors $(l = 1, \ldots, 7)$. Note that since $I = A_{\mathcal{F}} \circ B_{\mathcal{F}}$, I equals $F_1 \vee \cdots \vee F_7$. However, as we can see from the visual inspection of the matrices, already the first two or three factors explain the data reasonably well, i.e. both $F_1 \vee F_2$ and $F_1 \vee F_2 \vee F_3$ are reasonable approximations of I.

Let us now consider the interpretation of the first three factors. For this purpose, it is convenient to inspect the rectangular matrices from the first line of Fig. 2 and, to see

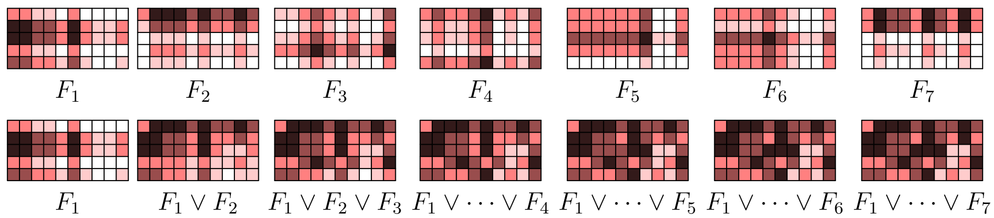


Figure 2: First line: Rectangular matrices representing the factors for decathlon data. Second line: $\bigvee$-superposition of factors.

more details, Table 1. Factor F_1: Manifestations of this factor with grade 1 are 100 m, long jump, and 110 m hurdles. This factor can be interpreted as the ability to run fast for short distances and can thus be termed speed. Note that this factor applies particularly to Clay and Karpov which is well known in the world of decathlon. Factor F_2: Manifestations of this factor with grade 1 are long jump, shot put, high jump, 110 m hurdles, and javelin. F_2 can be interpreted as the ability to apply very high force in a very short time and can thus be termed speed explosiveness. F_2 applies particularly to Sebrle, and to a lesser degree to Clay, who are known for this ability. Factor F_3: Manifestations with grade 1 are high jump and 1500 m. This factor is typical for light, not very muscular athletes (too much muscles prevent jumping high and running long distances). Macey, who is evidently that type among decathletes (196 cm and 98 kg) is the athlete to whom the factor applies to degree 1. These are the most important factors behind data matrix I.

This example demonstrates that the matrix decompositions discussed in this paper can be conveniently used for factor analysis of data matrices with elements from residuated lattices which are interpreted as degrees to which attributes (columns) apply to objects (rows).

5. General framework for matrix (relational) operations involving residuation

This section presents a framework which enables us to consider the $\circ$-decompositions studied in this paper and the $\triangleleft$-decompositions studied in [8] as two particular cases of a general type of decompositon. We restrict to the results directly related to the decomposition problems. More information about the general framework, its role in fuzzy set theory and fuzzy logic, and related work is be presented in a forthcoming paper [9].

5.1. *The framework*

Let for $i = 1, 2, 3$, $\mathbf{L}_i = \langle L_i, \leq_i \rangle$ be a complete lattice. The operations in $\mathbf{L}_i$ are denoted as usual, adding subscript i. That is, the infima, suprema, the least, and the greatest element in $\mathbf{L}_2$ are denoted by $\bigwedge_2$, $\bigvee_2$, 0_2, and 1_2, respectively; the same for $\mathbf{L}_1$ and $\mathbf{L}_3$.

Consider now an operation $\square : L_1 \times L_2 \to L_3$ that commutes with suprema in both arguments. That is, for any $a, a_j \in L_1$ $(j \in J)$, $b, b_{j'} \in L_2$ $(j' \in J')$,

$$(\textstyle\bigvee_{1 j \in J} a_j) \,\square\, b = \bigvee_{3 j \in J} (a_j \,\square\, b) \qquad \text{and} \qquad a \,\square\, (\bigvee_{2 j' \in J'} b_{j'}) = \bigvee_{3 j' \in J'} (a \,\square\, b_{j'}). \tag{24}$$

We call a quadruple $\langle \mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3, \Box\rangle$ satisfying (24) a supremum preserving aggregation structure (aggregation structure for short).

In our setting, $\langle \mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3, \Box\rangle$ plays a role analogous to the role of residuated lattices in case of $\circ$- and $\triangleleft$-decompositions.

Consider a matrix (relation) composition operator $\boxdot$ defined by

$$(A \boxdot B)_{ij} = \bigvee\nolimits_{3l=1}^{k} A_{il} \Box B_{lj}$$

for every $n \times k$ matrix A and $k \times m$ matrix B with $A_{il} \in L_1$, $B_{lj} \in L_2$. The decompositon problem may now be defined as follows. Given an $n \times m$ object-attribute matrix I with $I_{ij} \in L_3$, we look for a decomposition

$$I = A \boxdot B$$

of I into a product $A \boxdot B$ of an $n \times k$ object-factor matrix A and a $k \times m$ factor-attribute matrix B with $A_{il} \in L_1$, $B_{lj} \in L_2$, such that the number k of factors is the smallest possible. As the next example shows, both the $\circ$-decomposition and the $\triangleleft$-decomposition problems are particular cases of the $\boxdot$-decomposition problem.

Example 4. Let $\langle L, \wedge, \vee, \otimes, \rightarrow, 0, 1\rangle$ be a complete residuated lattice with a partial order $\leq$. The following two particular cases, in which $L_i = L$ and $\leq_i$ is either $\leq$ or the dual of $\leq$ (i.e. $\leq_i = \leq$ or $\leq_i = \leq^{-1}$) are important for our purpose.

(1) Let $\mathbf{L}_1 = \langle L, \leq\rangle$, $\mathbf{L}_2 = \langle L, \leq\rangle$, and $\mathbf{L}_3 = \langle L, \leq\rangle$, let $\Box$ be $\otimes$. Then, as is well known from the properties of residuated lattices [28, 18], $\Box$ commutes with suprema in both arguments. Clearly, the $\boxdot$-composition coincides with the $\circ$-composition (2) and the $\boxdot$-decomposition problem coincides with the $\circ$-decomposition problem.

(2) Let $\mathbf{L}_1 = \langle L, \leq\rangle$, $\mathbf{L}_2 = \langle L, \leq^{-1}\rangle$, and $\mathbf{L}_3 = \langle L, \leq^{-1}\rangle$, let $\Box$ be $\rightarrow$. Then, $\Box$ commutes with suprema in both arguments. Namely, the conditions (24) for commuting with suprema in this case become

$$(\bigvee\nolimits_{j\in J} a_j) \rightarrow b = \bigwedge\nolimits_{j\in J}(a_j \rightarrow b) \qquad \text{and} \qquad a \rightarrow (\bigwedge\nolimits_{j\in J} b_j) = \bigwedge\nolimits_{j\in J}(a \rightarrow b_j)$$

which are well-known properties of residuated lattices. One may thus easily see that in this case, the $\boxdot$-composition coincides with the $\triangleleft$-composition (4) and the $\boxdot$-decomposition problem coincides with the $\triangleleft$-decomposition problem studied in [8].

Remark 1. Note that there are two predecessors to the structure $\langle \mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3, \Box\rangle$. First, it is discussed in [5] that a residuated lattice may alternatively be defined as a bounded lattice with operation $\rightarrow$ that satisfies certain properties for which there exists operation $\otimes$ satisfying adjointness w.r.t. $\rightarrow$, as opposed to the usual definition according to which a residuated lattice is a bounded lattice with a monoidal operation $\otimes$ for which there exists operation $\rightarrow$ satisfying adjointness w.r.t. $\otimes$. The fact that the formal properties of $\rightarrow$ and $\otimes$ used in such two kinds of definition are alsmost the same suggests a "duality" between $\otimes$ and $\rightarrow$ (cf. Example 4 (1) and (2)). The second one is studied in [24] and [11] where a three-sorted residuated structure, slightly more general than the one used in this paper, is investigated. Note that the possibility to obtain two particular kinds of concepts lattices as particular instances, which is proposed in this paper (see Example 6), is not mentioned in [11, 24]. Namely, the motivation in [11, 24] consists in developing formal concepts with extents and intents using different truth degrees and the possibility mentioned in the previous sentence is not realized in those papers. Another important work is [17] in which the authors provide a framework consisting of five lattices and two basic opertions that allows one to consider both antitone and isotone fuzzy Galois connections as a single type of a fuzzy Galois connection. A more comprehensive information about related work is to be presented in [9].

Define operations $\circ_{\Box} : L_1 \times L_3 \to L_2$ and ${}_{\Box}\!\circ : L_3 \times L_2 \to L_1$ (adjoints to $\Box$) by

$$
\begin{aligned}
a_1 \circ_{\Box} a_3 &= \textstyle\bigvee_2 \{a_2 \mid a_1 \Box a_2 \leq_3 a_3\}, && (25)\\
a_3 \,{}_{\Box}\!\circ\, a_2 &= \textstyle\bigvee_1 \{a_1 \mid a_1 \Box a_2 \leq_3 a_3\}. && (26)
\end{aligned}
$$

We put indices in a_1 and the like for mnemonic reasons. For example, a_1 indicates that a_1 is taken from L_1. One may prove several properties of $\Box$, $\circ_{\Box}$, and ${}_{\Box}\!\circ$ which are counterparts to well-known properties of residuated lattices. The following properties are needed in what follows.

$$
\begin{aligned}
& a_1 \Box a_2 \leq_3 a_3 \text{ iff } a_2 \leq_2 a_1 \circ_{\Box} a_3 \text{ iff } a_1 \leq_1 a_3 \,{}_{\Box}\!\circ\, a_2, && (27)\\
& a_1 \Box (a_1 \circ_{\Box} a_3) \leq_3 a_3, && (28)\\
& (a_3 \,{}_{\Box}\!\circ\, a_2) \Box a_2 \leq_3 a_3, && (29)\\
& a \circ_{\Box} (\textstyle\bigwedge_{3 j \in J} c_j) = \textstyle\bigwedge_{2 j \in J} (a \circ_{\Box} c_j), && (30)\\
& (\textstyle\bigvee_{1 j \in J} a_j) \circ_{\Box} c = \textstyle\bigwedge_{2 j \in J} (a_j \circ_{\Box} c), && (31)\\
& c \,{}_{\Box}\!\circ (\textstyle\bigvee_{2 j \in J} b_j) = \textstyle\bigwedge_{1 j \in J} (c \,{}_{\Box}\!\circ\, b_j), && (32)\\
& (\textstyle\bigwedge_{3 j \in J} c_j) \,{}_{\Box}\!\circ\, b = \textstyle\bigwedge_{1 j \in J} (c_j \,{}_{\Box}\!\circ\, b). && (33)
\end{aligned}
$$

One gets various monotony conditions as a corollary of (30)–(33). For example, (30) and (31) imply that $\circ_{\Box}$ is isotone in the second and antitone in the first argument, respectively.

Example 5. (1) If $\mathbf{L}_i$s and $\Box$ are as in Example 4 (1), i.e. in case of $\circ$-composition, we have

$$a_1 \circ_{\Box} a_3 = \bigvee \{a_2 \mid a_1 \otimes a_2 \leq a_3\} = a_1 \to a_3$$

and, similarly, $a_3 \,{}_{\Box}\!\circ\, a_2 = a_2 \to a_3$. Then, for instance, (27) says that $a_1 \otimes a_2 \leq a_3$ iff $a_2 \leq a_1 \to a_3$ iff $a_1 \leq a_2 \to a_3$; (28) says that $a_1 \otimes (a_1 \to a_3) \leq a_3$; and (30) says that $a \to (\bigwedge_{j \in J} c_j) = \bigwedge_{j \in J} (a \to c_j)$.

(2) If $\mathbf{L}_i$s and $\Box$ are as in Example 4 (2), i.e. in case of $\triangleleft$-composition, we have

$$a_1 \circ_{\Box} a_3 = \bigwedge \{a_2 \mid a_1 \to a_2 \geq a_3\} = a_1 \otimes a_3$$

and

$$a_3 \,{}_{\Box}\!\circ\, a_2 = \bigvee \{a_1 \,|\, a_1 \to a_2 \geq a_3\} = a_3 \to a_2.$$

In this case, (27) says that $a_1 \to a_2 \geq a_3$ iff $a_2 \geq a_1 \otimes a_3$ iff $a_1 \leq a_3 \to a_2$; (28) says that $a_1 \to (a_1 \otimes a_3) \geq a_3$; and (30) says that $a \otimes (\bigvee_{j \in J} c_j) = \bigvee_{j \in J}(a \otimes c_j)$.

5.2. Optimal decompositions

Call an $n \times m$ matrix J $\boxdot$-*Cartesian* if there exists an L_1-set C in $\{1, \ldots, n\}$ (that is, C is a mapping of $\{1, \ldots, n\}$ to L_1) and an L_2-set D in $\{1, \ldots, m\}$ such that $J_{ij} = C(i) \Box D(j)$ for $1 \leq i \leq n$, $1 \leq j \leq m$. In this case, we write $J = C \Box D$.

Lemma 2. *$I = A \boxdot B$ for $n \times k$ and $k \times m$ matrices A and B iff I is a $\bigvee_3$-superposition of k $\boxdot$-Cartesian matrices $J_1, \ldots, J_k$, i.e. iff*

$$I = J_1 \vee_3 J_2 \vee_3 \cdots \vee_3 J_k.$$

Proof. Directly from definition: $I = A \boxdot B$ means $I_{ij} = \bigvee_{3\,l=1}^{\;k}(A_{il} \Box B_{lj})$. Obviously, this means that I is a $\bigvee_3$-superposition of $\boxdot$-Cartesian matrices J_l defined by $(J_l)_{ij} = A_{il} \Box B_{lj}$. $\Box$

Remark 2. Clearly, for the settings of Example 4 (1) and (2), $\boxdot$-Cartesian matrices become rectangular matrices introduced in this paper and I-beam matrices introduced in [8]. Hence, Lemma 2 generalizes Lemma 1 as well as Theorem 1 from [8].

Given sets $X = \{1, \ldots, n\}$, $Y = \{1, \ldots, m\}$, and an $n \times m$ matrix I with entries from L_3, let the operators ${}^{\uparrow} : L_1^X \to L_2^Y$ and ${}^{\downarrow} : L_2^Y \to L_1^X$ be defined by

$$C^{\uparrow}(y) = \bigwedge\nolimits_{2\,x \in X}(C(x) \circ_{\Box} I_{xy}) \quad \text{and} \quad D^{\downarrow}(x) = \bigwedge\nolimits_{1\,y \in Y}(I_{xy} \,{}_{\Box}\!\circ\, D(y)). \tag{34}$$

A formal concept of I is then a pair $\langle C, D\rangle$ consisting of an L_1-set C in X and an L_2-set D in Y for which $C^{\uparrow} = D$ and $D^{\downarrow} = C$. $\mathcal{B}(X, Y, I)$ denotes the set of all formal concepts of I, i.e.

$$\mathcal{B}(X, Y, I) = \{\langle C, D\rangle \in L_1^X \times L_2^Y \,|\, C^{\uparrow} = D,\ D^{\downarrow} = C\}.$$

Example 6. (1) If $\mathbf{L}_i$s and $\Box$ are as in Example 4 (1), i.e. in case of $\circ$-composition, then (34) become the concept-forming operators (11) and $\mathcal{B}(X, Y, I)$ is the concept lattice introduced in Section 2.2.

(2) If $\mathbf{L}_i$s and $\Box$ are as in Example 4 (2), , i.e. in case of $\triangleleft$-composition, (34) become

$$C^{\uparrow}(y) = \bigvee\nolimits_{x \in X}(C(x) \otimes I_{xy}) \quad \text{and} \quad D^{\downarrow}(x) = \bigwedge\nolimits_{y \in Y}(I_{xy} \to D(y)),$$

i.e. the operators denoted by ${}^{\cap}$ and ${}^{\cup}$ in [8], and $\mathcal{B}(X, Y, I)$ coincides with the set of the fixpoints of these operators, denoted by $\mathcal{B}(X^{\cap}, Y^{\cup}, I)$ in [8].

Denote for $\mathcal{F} = \{\langle C_1, D_1\rangle, \ldots, \langle C_k, D_k\rangle\} \subseteq \mathcal{B}(X, Y, I)$ by $A_{\mathcal{F}}$ and $B_{\mathcal{F}}$ the $n \times k$ and $k \times m$ matrices given by $(A_{\mathcal{F}})_{il} = (C_l)(i)$ and $(B_{\mathcal{F}})_{lj} = (D_l)(j)$. Again, we assume a fixed order on $\mathcal{F}$ given by the indices $1, \ldots, k$. We are ready to present the theorems on $\boxdot$-decompositons which generalize the results from Section 2 and from [8]. Denote by $\subseteq_i$ the inclusion relation induced by $\leq_i$, cf. (9). That is, if C, C' are L_i-sets in universe U, we put $C \subseteq_i C'$ if and only if for each $u \in U$, $C(u) \leq_i C'(u)$, for $i = 1, 2, 3$.

Theorem 4 (universality)**.** *Let $(a_3 \,{}_{\square}\!\circ\, 1_2) \square 1_2 = a_3$ for every $a_3 \in L_3$ or $1_1 \square (1_1 \circ_{\square} a_3) = a_3$ for every $a_3 \in L_3$. For every matrix I there exists a finite set $\mathcal{F} \subseteq \mathcal{B}(X, Y, I)$ such that $I = A_{\mathcal{F}} \boxdot B_{\mathcal{F}}$.*

Proof. Let $(a_3 \,{}_{\square}\!\circ\, 1_2) \square 1_2 = a_3$ for every $a_3 \in L_3$. Denote for $l \in \{1, \ldots, m\}$, $\langle C_l, D_l\rangle = \langle \{1_2/l\}^{\downarrow}, \{1_2/l\}^{\downarrow\uparrow}\rangle$. Here, $\{1_2/l\}$ is a singleton in $\{1, \ldots, m\}$, i.e. an L_2-set defined by $\{1_2/l\}(l) = 1_2$ and $\{1_2/l\}(j) = 0_2$ for $j \neq l$. $\langle C_l, D_l\rangle$ are particular formal concepts from $\mathcal{B}(X, Y, I)$ because $\langle D^{\downarrow}, D^{\downarrow\uparrow}\rangle$ is a formal concept for every L_2-set D in Y.

The latter claim follows from the following argument: (a) ${}^{\uparrow}$ and ${}^{\downarrow}$ form a Galois connection [27] between partially ordered sets $\langle L_1^X, \subseteq_1\rangle$ and $\langle L_2^Y, \subseteq_2\rangle$; (b) $D^{\downarrow} = D^{\downarrow\uparrow\downarrow}$ is one of the basic properties of Galois connections [27]; (c) $\langle D^{\downarrow}, D^{\downarrow\uparrow}\rangle$ thus satisfies the definition of a formal concept. To verify (a) is a matter of routinely checking the definition of a Galois connection conditions, i.e. that ${}^{\uparrow}$ and ${}^{\downarrow}$ are antitone, $C \subseteq_1 C^{\uparrow\downarrow}$, and $D \subseteq_2 D^{\downarrow\uparrow}$. Namely, the antitony of ${}^{\uparrow}$ follows from the antitony of $\circ_{\square}$ in the first argument which follows from (31); $C \subseteq_1 C^{\uparrow\downarrow}$ iff $C(x) \leq_1 I_{xy} \,{}_{\square}\!\circ\, C^{\uparrow}(y)$ for each $x \in X$ iff (due to (27)) $C(x) \square C^{\uparrow}(y) \leq_3 I_{xy}$. Now, since $\square$ preserves suprema, it is isotone in both arguments and hence $C(x) \square C^{\uparrow}(y) \leq_3 C(x) \square (C(x) \circ_{\square} I_{xy}) \leq_3 I_{xy}$, the last inequality being true due to (28); the antitony of ${}^{\downarrow}$ and $D \subseteq_2 D^{\downarrow\uparrow}$ can be checked analogously using (32), (27), and (29).

To verify $I = A_{\mathcal{F}} \boxdot B_{\mathcal{F}}$, i.e.

$$I_{xy} = \textstyle\bigvee_{3l=1}^{n} C_l(x) \square D_l(y),$$

observe that the "$\geq$"-part follows from $I_{xy} \geq_3 C_l(x) \square D_l(y)$ which is equivalent to $C_l(x) \leq_1 I_{xy} \,{}_{\square}\!\circ\, D_l(y)$, the last inequality being true due to $C_l(x) = \bigwedge_{1y\in Y}(I_{xy} \,{}_{\square}\!\circ\, D_l(y))$. The "$\leq$"-part follows from the fact that $\bigvee_{3l=1}^{n}(C_l(x) \square D_l(y)) \geq_3 C_y(x) \square D_y(y) = \{1_2/y\}^{\downarrow}(x) \square \{1_2/y\}^{\downarrow\uparrow}(y) \geq_3 (I_{xy} \,{}_{\square}\!\circ\, 1_2) \square 1_2 = I_{xy}$, the last equality being true due to the assumption $(a_3 \,{}_{\square}\!\circ\, 1_2) \square 1_2 = a_3$. Putting thus $\mathcal{F} = \{\langle C_l, D_l\rangle \mid l = 1, \ldots, m\}$, we get $I = A_{\mathcal{F}} \circ B_{\mathcal{F}}$.

If $1_1 \square (1_1 \circ_{\square} a_3) = a_3$ for every $a_3 \in L_3$, one proceeds analogously with $\langle C_l, D_l\rangle = \langle \{1_1/l\}^{\uparrow\downarrow}, \{1_1/l\}^{\uparrow}\rangle$. $\square$

Remark 3. (a) Note that due to (27), $(a_3 \,{}_{\square}\!\circ\, 1_2) \square 1_2 \leq_3 a_3$ and $1_1 \square (1_1 \circ_{\square} a_3) \leq_3 a_3$ for every $a_3 \in L_3$. However, the converse inequalities, assumed by Theorem 4, may not be satisfied, as the following example shows.

Let $L_1 = \{0, 1\}$, $L_2 = [0, 1]$, $L_3 = [0, 1]$, let $\leq_1$, $\leq_2$, $\leq_3$ be the usual total orders on L_1, L_2, and L_3, respectively. Let $\square$ be defined by $a_1 \square a_2 = \min(a_1, a_2)$. Then $\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{L}_3$, and $\square$ satisfy (24). However, $(a_3 \,{}_{\square}\!\circ\, 1_2) \square 1_2 \geq_3 a_3$ is violated for $0 < a_3 < 1$. Indeed, for $0 < a_3 < 1$, $(a_3 \,{}_{\Box}\!\circ\, 1_2) = \bigvee_1 \{a_1 \mid a_1 \Box 1_2 \leq a_3\} = 0_1$, hence $(a_3 \,{}_{\Box}\!\circ\, 1_2) \Box 1_2 = 0_1 \Box 1_2 = \min(0,1) = 0 < a_3$.

(b) For the setting of Example 4 (1), $(a_3 \,{}_{\Box}\!\circ\, 1_2) \Box 1_2 = a_3$ and $1_1 \Box (1_1 \circ_{\Box} a_3) = a_3$ become $(1 \to a_3) \otimes 1 = a_3$ and $1 \otimes (1 \to a_3) = a_3$, respectively, and hence both are satisfied. As a consequence, Theorem 1 is a consequence of Theorem 4.

(c) For the setting of Example 4 (2), $(a_3 \,{}_{\Box}\!\circ\, 1_2) \Box 1_2 = a_3$ becomes $(a_3 \to 0) \to 0 = 0$ which is not satisfied in general. However, $1_1 \Box (1_1 \circ_{\Box} a_3) = a_3$ becomes $1 \to (1 \otimes a_3) = a_3$ which is always true. As a consequence, Theorem 3 from [8] is a consequence of Theorem 4.

(d) Let $L_1 = L_3$ be supports of a residuated lattice with a partial order $\leq$, $L_2 = \{0,1\}$ with its natural order, let $a_1 \otimes a_2 = \min(a_1, a_2)$ for $a_1 \in L_1$, $a_2 \in L_2$. Decompositions corresponding to the aggregation structure consisting of L_1, L_2, and L_3 are studied in [1]. It may be checked that the results from [1] may easily be obtained from those presented in this section.

Theorem 5 (optimality)**.** *Let $I = A \boxdot B$ for $n \times k$ and $k \times m$ matrices A and B with entries from L. Then there exists a finite set $\mathcal{F} \subseteq \mathcal{B}(X, Y, I)$ of formal concepts of I with*

$$|\mathcal{F}| \leq k$$

such that for the $n \times |\mathcal{F}|$ and $|\mathcal{F}| \times m$ matrices $A_{\mathcal{F}}$ and $B_{\mathcal{F}}$ we have

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

Proof. Let $I = A \boxdot B$. According to Lemma 2 and its proof, I is a $\bigvee_3$-superposition of $\boxdot$-Cartesian matrices $J_1, \ldots, J_k$ defined as follows. Denote the L-sets in $\{1, \ldots, n\}$ and $\{1, \ldots, m\}$ corresponding to the l-th column of A and the l-th row of B by G_l and H_l, respectively, and put $J_l = G_l \Box H_l$, i.e. $(J_l)_{ij} = G_l(i) \Box H_l(j)$. Then,

$$G_l \Box H_l \subseteq_3 I. \tag{35}$$

We need the following claim which generalizes Theorem 4 of [3].

Claim. $\langle C, D\rangle \in \mathcal{B}(X, Y, I)$ iff $\langle C, D\rangle$ is a maximal pair that is contained in I. Maximality is considered w.r.t. a partial order $\sqsubseteq$ defined by $\langle C_1, D_2\rangle \sqsubseteq \langle C_2, D_2\rangle$ iff $C_1 \subseteq_1 C_2$ and $D_1 \subseteq_2 D_2$; a pair $\langle C, D\rangle$ is said to be contained in I if $C \Box D \subseteq_3 I$, i.e. $C(x) \Box D(y) \leq_3 I_{xy}$ for every $x \in X$, $y \in Y$.

Proof of Claim. If $\langle C, D\rangle \in \mathcal{B}(X, Y, I)$, we have $D = C^{\uparrow}$, hence $D(y) = \bigwedge_{2 x \in X} (C(x) \circ_{\Box} I_{xy}) \leq_2 C(x) \circ_{\Box} I_{xy}$ for every x, y. According to (27), $C(x) \Box D(y) \leq_3 I_{xy}$, i.e. $C \Box D \subseteq_3 I$, showing that $\langle C, D\rangle$ is contained in I. If $\langle C_1, D_1\rangle$ is contained in I and $\langle C, D\rangle \sqsubseteq \langle C_1, D_1\rangle$ then it follows from $C_1 \Box D_1 \subseteq_3 I$, (27), and the antitony of $\circ_{\Box}$ in the first argument that

$$D_1(y) \leq_2 C_1(x) \circ_{\Box} I_{xy} \leq_2 C(x) \circ_{\Box} I_{xy}$$

for every x, y, and thus $D_2 \subseteq_2 C^{\uparrow} = D$, showing $D = D_1$. Similarly, $C = C_1$, proving that $\langle C, D\rangle$ is maximal w.r.t. $\sqsubseteq$ among the pairs contained in I.

Conversely, let $\langle C, D\rangle$ be maximal pair contained in I. $C \boxempty D \subseteq_3 I$ implies $D(y) \leq_2 \bigwedge_2(C(x) \circ_\boxempty I_{xy}) = C^\uparrow(y)$ for each y, i.e. $D \subseteq_2 C^\uparrow$. Note that $\langle C, C^\uparrow\rangle$ is contained in I since

$$C(x) \boxempty \bigwedge_{2x\in X}(C(x) \circ_\boxempty I_{xy}) \leq_3 C(x) \boxempty (C(x) \circ_\boxempty I_{xy}) \leq_3 I_{xy}.$$

Therefore, if $D \neq C^\uparrow$ were the case, one would have $D \subset_2 C^\uparrow$ and $\langle C, C^\uparrow\rangle$ would be contained in I and would be larger w.r.t. $\sqsubseteq$ than $\langle C, D\rangle$, contradicting the assumption. QED (Claim)

Due to (35), Claim implies that there exists a formal concept $\langle C_l, D_l\rangle \in \mathcal{B}(X, Y, I)$ such that $G_l \subseteq_1 C_l$ and $H_l \subseteq_2 D_l$. As shown above, since $\langle C_l, D_l\rangle$ is a formal concept, we have $C_l \boxempty D_l \subseteq_3 I$. Finally, for $\mathcal{F} = \{\langle C_1, D_1\rangle, \ldots, \langle C_k, D_k\rangle\}$ we get $|\mathcal{F}| \leq k$ and conclude

$$I = A \boxdot B = \bigvee_{3\,l=1}^{k} G_l \boxempty H_l \subseteq_3 \bigvee_{3\,l=1}^{k} C_l \boxempty D_l = A_\mathcal{F} \boxdot B_\mathcal{F} \subseteq_3 I,$$

i.e. $A_\mathcal{F} \boxdot B_\mathcal{F} = I$, finishing the proof. $\square$

Remark 4. (a) Theorem 5 generalizes both Theorem 2 and Theorem 4 from [8]. Both are particular cases of Theorem 5 for the settings of Example 4 (1) and (2), respectively.

(b) In Theorem 4 and its proof, we identified particular formal concepts that can be used to factorize I provided the condition in the first sentence of the theorem is satisfied. The following statement, similar to Theorem 4, follows directly from Theorem 5: If $I = A \boxdot B$ for some A and B, then $I = A_\mathcal{F} \boxdot B_\mathcal{F}$ for some finite $\mathcal{F} \subseteq \mathcal{B}(X, Y, I)$ (i.e., if I is decomposable at all, then I may be decomposed using formal concepts as factors). Namely, if $I = A \boxdot B$, I is a $\bigvee_3$-superposition of k $\boxdot$-Cartesian matrices. Then, a reasoning similar to that used in the proof of Theorem 5 yields a set $\mathcal{F}$ of at most k formal concepts of I for which $I = A_\mathcal{F} \boxdot B_\mathcal{F}$.

(c) Not every I is decomposable. For example, for $L_1 = L_2 = \{0, 1\}$, $L_3 = [0, 1]$, and $a \boxempty b = \min(a, b)$, if I contains and entry diferent from both 0 and 1, I is clearly not decomposable.

5.3. Transformations between spaces of attributes and factors

In this section, we present results on transformations between the m-dimensional space L_3^m of attributes and the k-dimensional space L_1^k of factors which are induced by decomposition (25). The results generalize most results from Section 3 and include some further ones. Again, we identify L_3-sets $P : \{1, \ldots, m\} \to L_3$ with vectors $(P(1), \ldots, P(m))$ of elements from L_3, and L_1-sets $Q : \{1, \ldots, k\} \to L_1$ with $(Q(1), \ldots, Q(k))$.

Let thus $I = A \boxdot B$ and consider $g : L_3^m \to L_1^k$ and $h : L_1^k \to L_3^m$ defined for $P \in L_3^m$ and $Q \in L_1^k$ by

$$(g(P))_l = \bigwedge_{1\,j=1}^{m} (P_j \,{}_\boxempty\!\circ\, B_{lj}), \quad (36)$$

$$(h(Q))_j = \bigvee_{3\,l=1}^{k} (Q_l \boxempty B_{lj}), \quad (37)$$

for $1 \leq l \leq k$ and $1 \leq j \leq m$.

Remark 5. For the setting of Example 4 (1), (36) and (37) become (12) and (13), respectively. For the setting of Example 4 (2), (36) and (37) become

$$
\begin{aligned}
(g(P))_l &= \textstyle\bigwedge_{j=1}^{m}(P_j \to B_{lj}), && (38)\\
(h(Q))_j &= \textstyle\bigwedge_{l=1}^{k}(Q_l \to B_{lj}), && (39)
\end{aligned}
$$

which are the transformations described in [8] (equations (7) and (8)).

As is shown in Example 3, mappings g and h are non-linear in general. We now explore the properties of g and h. Denoting by $I_{i\_}$ and $A_{i\_}$ the i-th row of I and A, respectively, $I = A \circ B$ and (13) immediately yields

$$h(A_{i\_}) = I_{i\_} \tag{40}$$

for $i = 1, \ldots, n$. The next lemma describes properties of g. Particularly, it shows that if the columns of A are extents of formal concepts of I which correspond to the rows of B (the rows of B need not be intents) then we also have

$$g(I_{i\_}) = A_{i\_}. \tag{41}$$

Lemma 3. *If $I = A \circ B$ then $(g(I_{i\_}))_l \geq_1 A_{il}$ for every i and l. If, moreover, every column of A is the extent induced by the corresponding row of B, i.e. $A_{\_l} = B_{l\_}^{\downarrow}$, then $g(I_{i\_}) = A_{i\_}$.*

Proof. Because $I = A \boxdot B$, we have $(g(I_{i\_}))_l = \bigwedge_{1\,j=1}^{m}(I_{ij} \,{}_{\Box}\!\circ\, B_{lj}) = \bigwedge_{1\,j=1}^{m}((\bigvee_{3\,l=1}^{k} A_{il} \Box B_{lj}) \,{}_{\Box}\!\circ\, B_{lj})$. Thus, in order to check $(g(I_{i\_}))_l \geq_1 A_{il}$, we need to verify

$$\bigwedge_{1\,j=1}^{m}((\bigvee_{3\,l'=1}^{k} A_{il'} \Box B_{l'j}) \,{}_{\Box}\!\circ\, B_{lj}) \geq_1 A_{il}. \tag{42}$$

Clearly, (42) holds true iff for each j we have $(\bigvee_{3\,l'=1}^{k} A_{il'} \Box B_{l'j}) \,{}_{\Box}\!\circ\, B_{lj} \geq_1 A_{il}$. This inequality is equivalent to $A_{il} \Box B_{lj} \leq_3 \bigvee_{3\,l'=1}^{k} A_{il'} \Box B_{l'j}$ which is evidently true. If, in additon, $A_{\_l} = B_{l\_}{}^{\downarrow}$, then $(g(I_{i\_}))_l = \bigwedge_{1\,j=1}^{m}(I_{ij} \,{}_{\Box}\!\circ\, B_{lj}) = B_{l\_}{}^{\downarrow}(l) = A_{il}$, finishing the proof. $\Box$

The next lemma shows what happens if the rows of B are the intents corresponding to the columns of A (columns of A need not be extents).

Lemma 4. *Let $I = A \boxdot B$. If every row of B is the intent induced by the corresponding column of A, i.e. $B_{l\_} = A_{\_l}{}^{\uparrow}$, then B is the largest matrix for which $I = A \boxdot B$. That is, if $I = A \boxdot B'$ then $B'_{lj} \leq_2 B_{lj}$ for every l and j.*

Proof. The assertion follows from the fact that for every decomposition $I = A \boxdot B$, matrix D defined by $D_{lj} = \bigwedge_{2\,i=1}^{m}(A_{il} \circ_{\Box} I_{ij})$ is the largest one for which $I = A \boxdot D$. To check this, one can check that adjointness property implies that if $I = A \boxdot B$ then $B_{lj} \leq_2 \bigwedge_{2\,i=1}^{m}(A_{il} \circ_{\Box} I_{ij}) = D_{lj}$. Furthermore, as one can verify, using $A_{il} \Box (A_{il} \circ_{\Box} I_{ij}) \leq_3 I_{ij}$, we get $I_{ij} = (A \boxdot B)_{ij} \leq_3 (A \boxdot D)_{ij} \leq_3 I_{ij}$. Now, if every row of B is the intent induced by the corresponding column of A then, by definition of D, $B = D$, i.e. B itself is the largest matrix for which $I = A \boxdot B$. $\Box$

As a consequence, we get the following theorem:

Theorem 6. *Let $I = A_{\mathcal{F}} \boxdot B_{\mathcal{F}}$ for a finite set $\mathcal{F} \subseteq \mathcal{B}(X, Y, I)$. Then*

$$g(I_{i\_}) = A_{i\_} \quad \textit{and} \quad h(A_{i\_}) = I_{i\_}$$

for every i. Moreover, $B_{\mathcal{F}}$ is the largest of the matrices D for which $I = A_{\mathcal{F}} \circ D$. Likewise, $A_{\mathcal{F}}$ is the largest of the matrices C for which $I = C \circ B_{\mathcal{F}}$.

Proof. The first part follows directly from Lemma 3 and 4. The fact that $A_{\mathcal{F}}$ is the largest one follows from symmetry under transposition of the matrices in $I = A_{\mathcal{F}} \boxdot B_{\mathcal{F}}$ (or can be proved directly the same way as the maximality of B in Lemma 4). □

The orderings $\leq_3$ and $\leq_1$ on L_3 and L_1 induce coordinate-wise orderings of vectors in the spaces L_3^m of attributes and L_1^k of factors, defined by $P \leq_3 P'$ if $P_j \leq_3 P'_j$ for all j, and $Q \leq_1 Q'$ if $Q_l \leq_1 Q'_l$ for all l. The following theorem shows basic properties of g and h w.r.t. these orderings.

Theorem 7. *For $P, P' \in L_3^m$ and $Q, Q' \in L_1^k$:*

$$P \leq_3 P' \textit{ implies } g(P) \leq_1 g(P'), \tag{43}$$
$$Q \leq_1 Q' \textit{ implies } h(Q) \leq_3 h(Q'), \tag{44}$$
$$h(g(P)) \leq_3 P, \tag{45}$$
$$Q \leq_1 g(h(Q)), \tag{46}$$
$$g(P) = ghg(P), \tag{47}$$
$$h(Q) = hgh(Q), \tag{48}$$
$$g(\textstyle\bigwedge_{3 s \in S} P_s) = \textstyle\bigwedge_{1 s \in S} g(P_s), \tag{49}$$
$$h(\textstyle\bigvee_{1 t \in T} Q_t) = \textstyle\bigvee_{3 t \in T} h(Q_t). \tag{50}$$

Proof. (43) and (44) follow from definitions of g and h and the fact that $\Box$ is isotone and ${}_{\Box}\!\circ$ is isotone in the first argument. (45):

$$\begin{aligned} h(g(P))_j &= \bigvee\nolimits_{3 l=1}^{k} (g(P)_l \Box B_{lj}) = \bigvee\nolimits_{3 l=1}^{k} ((\bigwedge\nolimits_{1 j'=1}^{m} (P_{j'} \,{}_{\Box}\!\circ\, B_{lj'})) \Box B_{lj}) \leq_3 \\ &\leq_3 \bigvee\nolimits_{3 l=1}^{k} ((P_j \,{}_{\Box}\!\circ\, B_{lj}) \Box B_{lj}) \leq_3 P_j \end{aligned}$$

because $(a_3 \,{}_{\Box}\!\circ\, a_2) \Box a_2 \leq_3 a_3$. (46) can be shown in a similar way. (43)–(46) mean that g and h form a residuated pair of mappings [14]. (47)–(50) thus follow from the properties of residuated mappings. □

Remark 6. As is easily seen, for the setting of Example 4 (1), Theorem 7 generalizes the corresponding properties from Section 3, namely (14)-(17), and provides further properties. Likewise, for the setting of Example 4 (2), Theorem 7 generalizes the corresponding properties from Section 3 of [8].

The next theorem shows that g and h partition the space of attributes and the space of factors into particular convex subsets. Recall that a subset $S \subseteq L^p$ is called convex if $V \in S$ whenever $U \leq V \leq W$ for some $U, W \in S$. Let for $P \in L_3^m$ and $Q \in L_1^k$ denote by $g^{-1}(Q)$ the set of all vectors mapped to Q by g and by $h^{-1}(P)$ the set of all vectors mapped to P by h, i.e. $g^{-1}(Q) = \{P \in L_3^m \mid g(P) = Q\}$, and $h^{-1}(P) = \{Q \in L_1^k \mid h(Q) = P\}$.

Theorem 8. *(i) If $g^{-1}(Q) \neq \emptyset$ then $g^{-1}(Q)$ is a convex partially ordered subspace of the attribute space and $h(Q)$ is the least element of $g^{-1}(Q)$.*

(ii) If $h^{-1}(P) \neq \emptyset$ then $h^{-1}(P)$ is a convex partially ordered subspace of the attribute space and $g(P)$ is the largest element of $h^{-1}(P)$.

Proof. (i) Let $g^{-1}(Q) \neq \emptyset$. Let P be from $g^{-1}(Q)$, i.e. $g(P) = Q$. Then, in particular, $Q \leq_1 g(P)$. Using (44) and (45), $h(Q) \leq_3 h(g(P)) \leq_3 P$, hence $h(Q) \leq_3 P$. Moreover, using (47) we get $Q = g(P) = ghg(P) = gh(Q)$, hence $h(Q) \in g^{-1}(Q)$. Therefore, $h(Q)$ is the least vector of $g^{-1}(Q)$. Let now $U, W \in g^{-1}(Q)$ and $U \leq_3 V \leq_3 W$. (43) yields $Q = g(U) \leq_1 g(V) \leq_1 g(W) = Q$, hence $g(V) = Q$, proving that $g^{-1}(Q)$ is convex. The proof of (ii) is similar. □

Remark 7. For the setting of Example 4 (1), Theorem 8 provides the result behind Fig. 1 (see the discussion at the end of Section 3). For the setting of Example 4 (2), Theorem 7 generalizes Theorem 6 from [8].

6. Conclusions and future research

We describe optimal decompositions of matrices with entries from residuated lattices. Factors in such decompositions are formal concepts in the sense of Port-Royal logic. In addition, we describe transformations between the space of original attributes and the space of new factors. Main results, comments and an illustrative example are presented for the important case of $\circ$-decompositions. In addition, we present a general framework which enables one to generalize the results regarding $\circ$-decomposition and $\triangleleft$-decomposition and prove results within this framework. The topics for future research include the following ones.

- Algorithms for computing decompositions, computational complexity and approximability of decomposition problems. In [12] we showed that the results presented in this paper are relevant for a design of approximation algorithms. Namely, an efficient greedy algorithm which computes formal concepts as factors is presented and experimentally evaluated in [12] on datasets with thousands of rows and hundreds of attributes. The case of binary data is studied in [13].

- Approximate decompositions, i.e. decompositions in which the input matrix I is to be decomposed into A and B in such a way that with an appropriate definition of approximate equality, I is approximately equal to the product of A and B.

– Concept lattices and further topics in decompositions in the general framework presented in Section 5. This includes further study and possible variants of the general framework which turnes out to be an interesting generalization of residuated lattices. Such a study should make it possible to provide further general results on the structures related to decompositons. As a particular example, we did not provide a generalization of (18)–(21) because there is no direct way to represent the terms involved in the expressions in the general framework. One way toward such a generalization is to consider a particular case of the general framework in which all the three lattices have the same support set and which satisfy further conditions.

– Further applications of the general framework in fuzzy set theory and its applications, e.g. in providing a common generalization of the so-called sup-t-norm and inf-residuum type of fuzzy relational equations [19]. One practical consequence of such generalization is the fact that solution methods for both types of equations may be investigated for the general type at once.

Acknowledgment. This paper is an extended version of a contribution presented at IEEE IS 2008. Support by grant No. 202/10/0262 of the Czech Science Foundation and research plan MSM 6198959214 is gratefully acknowledged. I thank the anonymous reviewer for helpful suggestions.

References

[1] E. Bartl, R. Belohlavek, J. Konecny, Optimal decompositions of matrices with grades into binary and graded matrices, *Annals of Mathematics and Artificial Intelligence* (in press), DOI: 10.1007/s10472-010-9185-y.

[2] R. Belohlavek, Fuzzy Galois connections, *Math. Logic Quarterly* **45** (1999), 497–504.

[3] R. Belohlavek, Lattices of fixed points of fuzzy Galois connections, *Math. Logic Quarterly* **47** (2001), 111–116.

[4] R. Belohlavek, Fuzzy closure operators, *J. Mathematical Analysis and Appl.* **262**(2001) 473–489.

[5] R. Belohlavek, *Fuzzy Relational Systems: Foundations and Principles,* Kluwer, Academic/Plenum Publishers, New York, 2002.

[6] R. Belohlavek, Concept equations, *J. Logic and Computation* Vol. **14** No. 3(2004), 395–403.

[7] R. Belohlavek, Concept lattices and order in fuzzy logic, *Annals of Pure and Applied Logic* **128**(1–3)(2004), 277–298.

[8] R. Belohlavek, Optimal triangular decompositions of matrices with entries from residuated lattices, *Int. J. Approximate Reasoning* **50**(8)(2009), 1250–1258.

[9] R. Belohlavek, Sup-t-norm and inf-residuum are one type of relational product: unifying framework and consequences (submitted).

[10] R. Belohlavek, B. De Baets, J. Outrata, V. Vychodil, Computing the lattice of all fixpoints of a fuzzy closure operator, *IEEE Transactions on Fuzzy Systems* **18**(3)(2010), 546–557.

[11] R. Belohlavek, V. Vychodil, What is a fuzzy concept lattice? Proc. CLA 2005, 3rd Int. Conference on Concept Lattices and Their Applications, pp. 34–45.

[12] R. Belohlavek, V. Vychodil, Factor analysis of incidence data via novel decomposition of matrices, *Lecture Notes in Computer Science* **5548**(2009), 83–97.

[13] R. Belohlavek, V. Vychodil, Discovery of optimal factors in binary data via a novel method of binary matrix decomposition, *Journal of Computer and System Sciences* **76**(1)(2010), 3–20.

[14] T.S. Blyth, *Lattices and Ordered Algebraic Structures*, Springer, 2005.

[15] A. A. Frolov, D. Húsek, I. P. Muraviev, P. A. Polyakov, Boolean factor analysis by Hopfield-like autoassociative memory, *IEEE Transactions on Neural Networks* Vol. **18**, No. 3, May 2007, pp. 698–707.

[16] B. Ganter, R. Wille, *Formal Concept Analysis. Mathematical Foundations,* Springer, Berlin, 1999.

[17] G. Georgescu, A. Popescu, Non-dual fuzzy connections, *Archive for Mathematical Logic* **43**(2004), 1009–1039.

[18] J. A. Goguen, L-fuzzy sets, *J. Math. Anal. Appl.* **18**(1967), 145–174.

[19] S. Gottwald, *A Treatise on Many-Valued Logics,* Research Studies Press, Baldock, Hertfordshire, England, 2001.

[20] S. Gottwald, Mathematical fuzzy logics, *Bulletin of Symbolic Logic* **14**(2)(2008), 210–239.

[21] P. Hájek, *Metamathematics of Fuzzy Logic,* Kluwer, Dordrecht, 1998.

[22] U. Höhle, On the fundamentals of fuzzy set theory, *J. Math. Anal. Appl.* **201**(1996), 786–826.

[23] K. H. Kim, *Boolean Matrix Theory and Applications*, Marcel Dekker, Inc., New York, 1982.

[24] S. Krajči, A generalized concept lattice. *Logic J. of IGPL* **13**(2005), 543–550.

[25] G. Markowsky, A. Wohlgemuth, Mathematical Immunogenetics II: Antibody Incidence Structure, *J. Theor. Biology* **102** (1983), 425–437.

[26] P. Miettinen, T. Mielikäinen, A. Gionis, G. Das, H. Mannila, The discrete basis problem, *Lecture Notes in Artificial Intelligence* **4213**(2006), 335–346.

[27] O. Ore, Galois connexions, *Trans. Amer. Math. Soc.* **55** (1944), 493–513.

[28] M. Ward, R.P. Dilworth, Residuated lattices, *Trans. Amer. Math. Soc.* **45**(1939), 335–354.

[29] A. Wohlgemuth, G. Markowsky, Mathematical Immunogenetics I: Mathematics and Language, *J. Theor. Biology* **102**(1983), 411–424.