
Volume 50, No. 8, September 2009
ISSN 0888-613X
Special Section on Interval/Probabilistic Uncertainty
Edited by: Van-Nam Huynh and Vladik Kreinovich
ELSEVIER
INTERNATIONAL JOURNAL OF
APPROXIMATE
REASONING
Imprecision and Uncertainty in Intelligent Systems
Editor-in-Chief
Thierry Denoeux
Available online at www.sciencedirect.com
ScienceDirect

International Journal of Approximate Reasoning 50 (2009) 1250–1258

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

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

Optimal triangular decompositions of matrices with entries from residuated lattices

Radim Belohlavek *

SUNY Binghamton, Binghamton, NY 13902, USA
Palacký University, Tomkova 40, 779 00 Olomouc, Czech Republic

ARTICLE INFO

Article history:
Received 26 January 2009
Received in revised form 13 May 2009
Accepted 22 May 2009
Available online 2 June 2009

Keywords:
Matrix decomposition
Residuated lattice
Isotone Galois connection
Fixpoint
Fuzzy logic

ABSTRACT

We describe optimal decompositions of an $n \times m$ matrix I into a triangular product $I = A \triangleleft B$ of an $n \times k$ matrix A and a $k \times m$ matrix B. We assume that the matrix entries are elements of a residuated lattice, which leaves binary matrices or matrices which contain numbers from the unit interval $[0,1]$ as special cases. The entries of I, A, and B represent grades to which objects have attributes, factors apply to objects, and attributes are particular manifestations of factors, respectively. This way, the decomposition provides a model for factor analysis of graded data. We prove that fixpoints of particular operators associated with I, which are studied in formal concept analysis, are optimal factors for decomposition of I in that they provide us with decompositions $I = A \triangleleft B$ with the smallest number k of factors possible. Moreover, we describe transformations between the m-dimensional space of original attributes and the k-dimensional space of factors. We provide illustrative examples and remarks on the problem of computing the optimal decompositions. Even though we present the results for matrices, i.e. for relations between finite sets in terms of relations, the arguments behind are valid for relations between infinite sets as well.

1. Introduction

1.1. Problem setting

The problem discussed in this paper can be described as follows. Let I be an $n \times m$ object–attribute matrix whose entries I_{ij} are elements from a residuated lattice $\mathbf{L} = \langle L, \otimes, \rightarrow, \wedge, \vee, 0, 1\rangle$ (see Section 1.4 for preliminaries), i.e. $I_{ij} \in L$. We look for a decomposition

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

of I into a product $A \triangleleft B$ 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. The composition operator $\triangleleft$ is defined by

$$(A \triangleleft B)_{ij} = \bigwedge_{l=1}^{k} A_{il} \rightarrow B_{lj} \tag{2}$$

with $\bigwedge$ denoting the infimum in $\mathbf{L}$. The operator $\triangleleft$ is known in fuzzy set theory. Namely, $A \triangleleft B$ is called a triangular product, or the inf-$\rightarrow$ product, or Bandler–Kohout product [20].

* Address: SUNY Binghamton, Binghamton, NY 13902, USA. Fax: +1 607 777 4094.
E-mail address: rbelohla@binghamton.edu

0888-613X/$ - see front matter
doi:10.1016/j.ijar.2009.05.006

Note that I, A, and B can be looked at as representing fuzzy relations R_I, R_A, and R_B, i.e. $R_I(i,j) = I_{ij}$, $R_A(i,l) = A_{il}$, and $R_B(l,j) = B_{lj}$, in which context $\triangleleft$ usually appears in fuzzy set theory [14,19]. Two concrete well-known examples are: $L = [0,1]$ and $\rightarrow$ is a residuum of a t-norm $\otimes$; $L = \{0,1\}$ and $\rightarrow$ is a (truth function of) classic implication (i.e. $1 \rightarrow 0 = 0$, $1 \rightarrow 1 = 0 \rightarrow 0 = 0 \rightarrow 1 = 1$). Note that in terms of relations, we present our results for relations between finite sets (namely, for matrices which represent such relations). However, the arguments are valid even for relations between infinite sets (i.e. for "infinite matrices").

1.2. Motivation and factor analysis interpretation

Residuated lattices can be thought of as partially ordered scales of degrees, such as $L = \{0,1\}$ representing a yes-or-no scale; $L = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ representing a scale consisting of "very bad", "bad", "neutral", "good", "very good"; or $L = [0,1]$. An entry $I_{ij} \in L$ of I can be interpreted as a degree to which attribute j applies to object i.

Looking for a decomposition of I can be interpreted as looking for hidden factors in the data represented by I. For the purpose of illustration, consider $L = \{0,1\}$ (binary matrices). Let the $n \times m$ matrix I describe a relationship between objects and attributes. A decomposition $I = A \triangleleft B$ corresponds to a discovery of k factors. Namely, due to (2), the original object–attribute relationship represented by I is described via an object–factor relationship represented by A and a factor–attribute relationship represented by B the following way:

> Object i has attribute j (i.e., $I_{ij} = 1$) if and only if for every factor $l = 1, \ldots, k$: if l applies to i (i.e., $A_{il} = 1$) then j is a particular manifestation of l (i.e., $B_{lj} = 1$).

For a general scale L, such an interpretation of $I = A \triangleleft B$ remains valid but degrees need to be taken into account. In particular, $I = A \triangleleft B$ then means that the degree I_{ij} to which object i has attribute j is the degree to which the following proposition is true: for every factor l, if l applies to i then j is a particular manifestation of l. Concrete example for the binary case: Let objects and attributes be jobs and persons, let $I_{ij} = 1$ mean that person j performs (or is able to perform) job i. Factors in decomposition $I = A \triangleleft B$ can then be interpreted as skills (conditions characterizing the jobs). Namely, with $A_{il} = 1$ being interpreted as "skill l is required for job i" and $B_{lj} = 1$ being interpreted as "person j has skill l", $I = A \triangleleft B$ says that person j is able to perform job i if and only if person j has all skills required for job i.

If k is smaller than m, a decomposition $I = A \triangleleft B$ provides us with a description of objects in terms of a small number of factors which reduce the dimensionality of the original dataset represented by I. If k is the smallest one, the factors can be regarded as a minimal set of descriptive conditions for the n objects with respect to the observed m attributes.

1.3. Related work

Related decompositions, namely $I = A \circ B$ with $\circ$ being the sup-$\otimes$ product defined by $(A \circ B)_{ij} = \bigvee_{l=1}^{k} A_{il} \otimes B_{lj}$, are studied in [5]. Note that for $L = \{0,1\}$, decompositions $I = A \circ B$ are of primary concern in Boolean factor analysis, see e.g. [10,23], and are also studied in data mining, see e.g. [25]. A theoretical analysis of $\circ$-decomposition of binary matrices, its computational complexity, approximation algorithms, and their experimental evaluation are presented in [7]. Note that while technically different, the approach presented in this paper is conceptually similar to that one presented in [5], where instead of isotone Galois connections, used in this paper, we used antitone Galois connections. In the binary case, isotone and antitone Galois connections are mutually definable (one can be obtained from the other by a well-known duality). In the general setting of residuated lattices, they are not [12].

A related problem of decomposition of a binary (or [0,1]-valued) matrix I into $A \triangleleft B$ is known as the problem of inf-$\rightarrow$ (fuzzy) relational equations. This problem has been studied and utilized in various areas for a long time, see e.g. [9,19]. Namely, the problem is, given A and I, find B such that $I = A \triangleleft B$ (or, given B and I, find A such that $I = A \triangleleft B$). This problem is very different from the one discussed in our paper, because in addition to I, one of the other matrices, A or B, is known.

Note also that fuzzy Galois connections, which play an important role in our paper, were studied in several papers including [1,4,12,18,22].

1.4. Preliminaries from residuated lattices

A residuated lattice [14,17,26] 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 the adjointness property

$$a \otimes b \leqslant c \quad \text{iff } a \leqslant b \rightarrow c \tag{3}$$

for every $a, b, c \in L$. A residuated lattice 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 fuzzy logic and fuzzy set theory [3,13,15,16]. 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 $\rightarrow$ [3,14]. Another commonly used example is a finite linearly ordered $\mathbf{L}$, a special case of which is the two-element Boolean algebra

$\langle\{0,1\},\wedge,\vee,\otimes,\to,0,1\rangle$, denoted by **2**, which is the structure of truth degrees of classical logic. That is, the operations $\wedge,\vee,\otimes,\to$ of **2** are the truth functions of the corresponding logical connectives of classical logic.

Given a residuated lattice **L**, an **L**-set (fuzzy set, graded set) A in a universe U is a mapping $A : U \to L$, $A(u)$ being interpreted as "the degree to which u belongs to A". In the following we use well-known properties of residuated lattices and fuzzy sets over residuated lattices which can be found, e.g., in [3,14,16,17].

2. Optimal decompositions

2.1. Matrix composition as a $\bigwedge$-superposition of I-beam matrices

Let us first observe that $I = A \triangleleft B$ for $n \times k$ and $k \times m$ matrices A and B means that I is a $\bigwedge$-superposition of particular matrices.

Definition 1. An $n \times m$ matrix J is called an I-beam matrix (simply I-beam) iff there exist L-sets C in $\{1,\ldots,n\}$ and D in $\{1,\ldots,m\}$ such that

$$J_{ij} = C(i) \to D(j) \tag{4}$$

for $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$. We denote this fact briefly by $J = C \triangleleft D$.

The term "I-beam" comes from a geometric interpretation. For illustration, consider $L = \{0,1\}$. The fact that J is an I-beam matrix means that the entries of J which contain 1s form an area which, up to a permutation, has the form of letter I. For instance, for

$$C = (0\quad 0\quad \mathbf{1}\quad \mathbf{1}\quad \mathbf{1}\quad \mathbf{1}\quad 0\quad 0)^{\mathrm{T}}, \quad \text{and} \quad D = (0\quad 0\quad \mathbf{1}\quad \mathbf{1}\quad \mathbf{1}\quad 0\quad 0),$$

the corresponding I-beam $C \triangleleft D$ is

$$\begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ 0 & 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 & 0 \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \end{pmatrix}.$$

Theorem 1. *For arbitrary $n \times k$ and $k \times m$ matrices A and B, $I = A \triangleleft B$ iff I is a $\bigwedge$-superposition of k I-beam matrices $J_1,\ldots,J_k$, i.e. iff*

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

Proof. Directly from definitions: $I = A \triangleleft B$ means $I_{ij} = (A \triangleleft B)_{ij}$, i.e. $I_{ij} = \bigwedge_{l=1}^{k}(A_{il} \to B_{lj})$. Obviously, this means that I is a $\bigwedge$-superposition of I-beam matrices J_l, $l = 1,\ldots,k$, defined by $(J_l)_{ij} = A_{il} \to B_{lj}$. $\square$

Example 1. For simplicity, consider the following decomposition $I = A \triangleleft B$ of an 4×5 matrix:

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \triangleleft \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

According to Theorem 1, this decomposition can be rewritten as a $\bigwedge$-superposition

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

of I-beams J_1, J_2, J_3, J_4, where J_l results as a $\triangleleft$-product of the lth column of A and the lth row of B. Note that the I-beam shape of J_ls becomes apparent after rearrangement (permutation) of rows and columns. Due to small dimensions, the I-shape is degenerate in case of J_1 and J_4.

2.2. Fixpoints of isotone Galois connection associated with I as optimal factors for decomposition of I

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 fixpoints of certain operators associated with I as factors. The operators form an isotone **L**-Galois connection and were studied in formal concept analysis [12], see also [3,4,21,24] for more information on formal concept analysis of data with graded attributes. In formal concept analysis [11], fixpoints of Galois connections associated with I are called formal concepts. They represent certain biclusters in the data represented by I. In particular, let $X = \{1,\ldots,n\}$ denote the set of objects corresponding to the rows of I, $Y = \{1,\ldots,m\}$ denote the set of attributes corresponding to the columns of I. Formal concepts of I are certain biclusters $\langle C,D\rangle$ with C (called the extent of $\langle C,D\rangle$) and D (called the intent $\langle C,D\rangle$) being L-sets of objects and attributes, respectively. For an object $i \in X$, $C(i)$ represents a degree to which formal concept $\langle C,D\rangle$ applies to i; for an attribute $j \in Y$, $D(j)$ represents a degree to which $\langle C,D\rangle$ applies to j. The fixpoints, i.e. formal concepts which we use as factors in this paper, represent an alternative to the ordinary formal concepts [3,4,24] and were studied in [12].

Isotone **L**-*Galois connections associated with I*. Let $X = \{1,\ldots,n\}$ and $Y = \{1,\ldots,m\}$ be sets (of objects and attributes, respectively), I be an $n \times m$ matrix with entries from a residuated lattice $\mathbf{L} = \langle L,\otimes,\rightarrow,\wedge,\vee,0,1\rangle$. Define operators $^{\cap}: L^X \rightarrow L^Y$ and $^{\cup}: L^Y \rightarrow L^X$, by letting for $C \in L^X$ and $D \in L^Y$,

$$C^{\cap}(j) = \bigvee_{i=1}^{n}(C(i)\otimes I_{ij}), \tag{5}$$

$$D^{\cup}(i) = \bigwedge_{j=1}^{m}(I_{ij}\rightarrow D(j)) \tag{6}$$

for $j \in \{1,\ldots,m\}$ and $i \in \{1,\ldots,n\}$. Furthermore, denote by $\mathcal{B}(X^{\cap},Y^{\cup},I)$ the set of fixpoints of $\langle ^{\cap},^{\cup}\rangle$. That is,

$$\mathcal{B}(X^{\cap},Y^{\cup},I) = \{\langle C,D\rangle \in L^X \times L^Y | C^{\cap} = D, D^{\cup} = C\}.$$

For a fixpoint $\langle C,D\rangle \in \mathcal{B}(X^{\cap},Y^{\cup},I)$, we call C and D the extent and the intent of $\langle C,D\rangle$. Note that every $\langle C,D\rangle \in \mathcal{B}(X^{\cap},Y^{\cup},I)$ is uniquely determined by its extent C as well as by its intent D.

Remark 1. Note that $\mathcal{B}(X^{\cap},Y^{\cup},I)$ equipped with a partial order $\leqslant$ defined by $\langle C_1,D_1\rangle \leqslant \langle C_2,D_2\rangle$ iff $C_1 \subseteq C_2$ (which is equivalent to $D_1 \subseteq D_2$), forms a complete lattice; the compound mapping $^{\cap\cup}: L^X \rightarrow L^X$ is an **L**-closure operator in X; the compound mapping $^{\cup\cap}: L^Y \rightarrow L^Y$ is an **L**-interior operator in X [2,12]. $\mathcal{B}(X^{\cap},Y^{\cup},I)$ is uniquely determined by the set $\mathrm{fix}(^{\cap\cup}) = \{C \in L^X | C = C^{\cap\cup}\}$ of fixpoints of $^{\cap\cup}$, because $\mathcal{B}(X^{\cap},Y^{\cup},I) = \{\langle C,C^{\cap}\rangle | C \in \mathrm{fix}(^{\cap\cup})\}$. As a consequence, $\mathcal{B}(X^{\cap},Y^{\cup},I)$ can be computed by the algorithms for computing sets of fixpoints of **L**-closure operators [6].

Fixpoints of $\langle ^{\cap},^{\cup}\rangle$ are minimal I-beams covering I. The fixpoints from $\mathcal{B}(X^{\cap},Y^{\cup},I)$ correspond to I-beams which cover I and are minimal w.r.t. a particular partial order $\leqslant_{\mathrm{I}}$ (the subscript I stands for I-beam ordering). We say that an I-beam matrix J corresponds to $\langle C,D\rangle$ iff $J = C \lhd D$, i.e. $J_{ij} = C(i) \rightarrow D(j)$ for all i,j. For L-sets $C_1, C_2 \in L^X$ and $D_1, D_2 \in L^Y$, put

$$\langle C_1,D_1\rangle \leqslant_{\mathrm{I}} \langle C_2,D_2\rangle \quad \text{iff } C_1 \supseteq C_2 \text{ and } D_1 \subseteq D_2,$$

i.e. iff $C_1(i) \geqslant C_2(i)$ for all $i \in \{1,\ldots,n\}$ and $D_1(j) \leqslant D_2(j)$ for all $j \in \{1,\ldots,m\}$. In terms of I-beams, $\langle C_1,D_1\rangle \leqslant_{\mathrm{I}} \langle C_2,D_2\rangle$ means that the I-beam $C_1 \lhd D_1$ corresponding to $\langle C_1,D_1\rangle$ is contained in the I-beam $C_2 \lhd D_2$ corresponding to $\langle C_2,D_2\rangle$, i.e. that $(C_1 \lhd D_1)_{ij} \leqslant (C_2 \lhd D_2)_{ij}$ for every i and j. The following is a crucial property of fixpoints from $\mathcal{B}(X^{\cap},Y^{\cup},I)$.

Theorem 2 [12]. *$\langle C,D\rangle$ is a fixpoint of $\langle ^{\cap},^{\cup}\rangle$ iff the corresponding I-beam is a minimal one which covers I, i.e. iff $\langle C,D\rangle$ is minimal with respect to $\leqslant_{\mathrm{I}}$ such that $I_{ij} \leqslant (C \lhd D)_{ij}$ for all i and j.*

Universality and optimality of fixpoints of $\langle ^{\cap},^{\cup}\rangle$ as factors. 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. In what follows, we always assume that there is a fixed order on the set $\mathcal{F}$ and indicate this order by indexes. Thus, we may speak of the 1st pair in $\mathcal{F}$ which is $\langle C_1,D_1\rangle$, up to the kth 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

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

That is, the lth column of $A_{\mathcal{F}}$ is the transpose of the vector corresponding to L-set C_l and the lth row of $B_{\mathcal{F}}$ is the vector corresponding to D_l. Note that 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. Then

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

The next two theorems contain the main results regarding the triangular decompositions $I = A \triangleleft B$ studied in this paper. The first theorem says that we can always use fixpoints from $\mathcal{B}(X^{\cap}, Y^{\cup}, I)$ as factors for decomposition of I.

Theorem 3 (Universality). *For every I there exists $\mathcal{F} \subseteq \mathcal{B}(X^{\cap}, Y^{\cup}, I)$ such that $I = A_{\mathcal{F}} \triangleleft B_{\mathcal{F}}$. Namely, one can put $\mathcal{F} = \mathcal{B}(X^{\cap}, Y^{\cup}, I)$.*

Proof. Put $\mathcal{F} = \mathcal{B}(X^{\cap}, Y^{\cup}, I)$ and let us denote $\mathcal{B}(X^{\cap}, Y^{\cup}, I)$ simply by $\mathcal{B}$. To show that $I = A_{\mathcal{F}} \triangleleft B_{\mathcal{F}}$, we need to check $I_{ij} = \bigwedge_{\langle C,D\rangle \in \mathcal{B}} C(i) \to D(j)$. On the one hand, $I_{ij} \leqslant \bigwedge_{\langle C,D\rangle \in \mathcal{B}} C(i) \to D(j)$ iff for each $\langle C, D\rangle \in \mathcal{B}$ we have $I_{ij} \leqslant C(i) \to D(j)$ which is equivalent to $I_{ij} \otimes C(i) \leqslant D(j)$ which true because $I_{ij} \otimes C(i) \leqslant C^{\cap}(j) = D(j)$. On the other hand, consider the fixpoint $\langle C_*, D_*\rangle = \langle \{1/i\}^{\cap\cup}, \{1/i\}^{\cap}\rangle \in \mathcal{B}$. We have

$$1 = \{1/i\}(i) \leqslant \{1/i\}^{\cap\cup}(i) = C_*(i),$$

hence $C_*(i) = 1$; and

$$D_*(j) = \{1/i\}^{\cap}(j) = \bigvee_{i' \in X} \{1/i\}(i') \otimes I_{i'j} = I_{ij}.$$

Hence,

$$\bigwedge_{\langle C,D\rangle \in \mathcal{B}} C(i) \to D(j) \leqslant C_*(i) \to D_*(j) = 1 \to I_{ij} = I_{ij},$$

finishing the proof. □

The second theorem says that taking the fixpoints as factors provides us with decompositions with the smallest number k of factors possible.

Theorem 4 (Optimality). *Let $I = A \triangleleft B$ for $n \times k$ and $k \times m$ matrices A and B. Then there exists a set $\mathcal{F} \subseteq \mathcal{B}(X^{\cap}, Y^{\cup}, I)$ of fixpoints with*

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

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

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

Proof. Let $I = A \triangleleft B$. Due to Theorem 1, I is an intersection of I-beams $J_1, \ldots, J_k$ which correspond to the columns and rows of A and B, respectively, and cover I. Every J_l contains some I-beam $J'_l \geqslant I$, which is minimal w.r.t. $\leqslant_{\mathrm{I}}$, i.e. $J_l \geqslant J'_l \geqslant I$. Denote by C_l and D_l the L-sets in X and Y for which $J'_l = C_l \triangleleft D_l$. By Theorem 2, $\langle C_l, D_l\rangle$s are fixpoints, i.e. $\langle C_l, D_l\rangle \in \mathcal{B}(X^{\cap}, Y^{\cup}, I)$. Put $\mathcal{F} = \{\langle C_l, D_l\rangle | 1 \leqslant l \leqslant k\}$. Clearly, $|\mathcal{F}| \leqslant k$. Using the assumption, Theorem 1, and the fact that I is the intersection of the collection of all I-beams corresponding to fixpoints (cf. proof of Theorem 3), we get

$$I = \bigwedge_{l=1}^{k} J_l \geqslant \bigwedge_{l=1}^{k} C_l^{\mathrm{T}} \triangleleft D_l = A_{\mathcal{F}} \triangleleft B_{\mathcal{F}} \geqslant \bigwedge_{\langle C,D\rangle \in \mathcal{B}(X^{\cap}, Y^{\cup}, I)} C^{\mathrm{T}} \triangleleft D = I.$$

Therefore, $A_{\mathcal{F}} \triangleleft B_{\mathcal{F}} = I$. □

Example 3. For the purpose of illustration again, let $L = \{0, 1\}$ (binary case). Consider again the decomposition

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

and the corresponding I-beams $J^1, \ldots, J^4$, which are

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Furthermore, consider the fixpoints $\langle C_1, D_1\rangle = \langle \{1,2,3\}, \{3,4,5\}\rangle$, $\langle C_2, D_2\rangle = \langle \{3\}, \{5\}\rangle$, $\langle C_3, D_3\rangle = \langle \{2,4\}, \{2,3,4\}\rangle$, from $\mathcal{B}(X^{\cap}, Y^{\cup}, I)$. Note that for the sake of brevity, we write $C_1 = \{1,2,3\}$ instead of $C_1 = \{1/1, 1/2, 1/3\}$, etc. One can check that each of the I-beams J_l $(l = 1, \ldots, 4)$ contains some of the minimal I-beams corresponding to $\langle C_1, D_1\rangle$, $\langle C_2, D_2\rangle$, or $\langle C_3, D_3\rangle$. Putting now $\mathcal{F} = \{\langle C_1, D_1\rangle, \langle C_2, D_2\rangle, \langle C_3, D_3\rangle\}$, we have $I = A_{\mathcal{F}} \triangleleft B_{\mathcal{F}}$. Denoting by $(A_{\mathcal{F}})_{\_l}$ and $(B_{\mathcal{F}})_{l\_}$ the lth column of $A_{\mathcal{F}}$ and the lth row of $B_{\mathscr{F}}$, $I = A_{\mathscr{F}} \triangleleft B_{\mathscr{F}}$ can further be rewritten as $I = (A_{\mathscr{F}})_{\_1} \triangleleft (B_{\mathscr{F}})_{1\_} \wedge (A_{\mathscr{F}})_{\_2} \triangleleft (B_{\mathscr{F}})_{2\_} \wedge (A_{\mathscr{F}})_{\_3} \triangleleft (B_{\mathscr{F}})_{3\_}$, which shows a $\bigwedge$-decomposition of I into minimal I-beams covering I. In particular, we have

$$\begin{pmatrix} 0&0&1&1&1\\ 0&0&1&1&0\\ 0&0&0&0&1\\ 0&1&1&1&0 \end{pmatrix} = \begin{pmatrix} 0&0&1&1&1\\ 0&0&1&1&1\\ 0&0&1&1&1\\ 1&1&1&1&1 \end{pmatrix} \wedge \begin{pmatrix} 1&1&1&1&1\\ 1&1&1&1&1\\ 0&0&0&0&1\\ 1&1&1&1&1 \end{pmatrix} \wedge \begin{pmatrix} 1&1&1&1&1\\ 0&1&1&1&0\\ 1&1&1&1&1\\ 0&1&1&1&0 \end{pmatrix}.$$

Theorems 3 and 4 say that when looking for factors for decompositions of I, we can confine ourselves to fixpoints from $\mathscr{B}(X^{\cap}, Y^{\cup}, I)$, i.e. to fixpoints of the isotone Galois connection associated with I.

3. Transformations between spaces of attributes and factors

We now describe mappings 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. 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 $(Q(1), \ldots, Q(k))$. Note that we are dealing with spaces L^m and L^k which are, however, not linear spaces (vector spaces). Namely, the vector components are elements of a residuated lattice rather than a field and we use operations of residuated lattices rather than fields.

Let $I = A \triangleleft B$ (we do not assume that $A = A_{\mathscr{F}}$ and $B = B_{\mathscr{F}}$ for some $\mathscr{F} \subseteq \mathscr{B}(X^{\cap}, Y^{\cup}, I)$). 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

$$(g(P))_l = \bigwedge_{j=1}^{m} (P_j \to B_{lj}), \tag{7}$$

$$(h(Q))_j = \bigwedge_{l=1}^{k} (Q_l \to B_{lj}) \tag{8}$$

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

$I = A \triangleleft B$ provides us with a representation of object i by the ith row $I_{i\_}$ of I in the space L^m of attributes, and a representation of i by the ith row $A_{i\_}$ of A in the space L^k of factors. Obviously, $I = A \triangleleft B$ and (8) immediately yield

$$h(A_{i\_}) = I_{i\_} \tag{9}$$

for $i = 1, \ldots, n$. The next lemma describes properties of g. Particularly, it shows that if the columns of A are extents of the fixpoints of $\langle {}^{\cap}, {}^{\cup} \rangle$ which correspond to the rows of B (the rows of B need not be intents) then we also have

$$g(I_{i\_}) = A_{i\_}. \tag{10}$$

Lemma 1. *If $I = A \triangleleft B$ then $(g(I_{i\_}))_l \geqslant A_{il}$ for every i and l. If, moreover, every column of A is the extent induced by the corresponding row of B, i.e. $A_{\_l} = B_{l\_}^{\cup}$, then $g(I_{i\_}) = A_{i\_}$.*

Proof. Since $I = A \triangleleft B$, we have

$$(g(I_{i\_}))_l = \bigwedge_{j=1}^{m} (I_{ij} \to B_{lj}) = \bigwedge_{j=1}^{m} \left(\left(\bigwedge_{l'=1}^{k} A_{il'} \to B_{l'j} \right) \to B_{lj} \right).$$

Thus, in order to check $(g(I_{i\_}))_l \geqslant A_{il}$, we need to verify

$$A_{il} \leqslant \bigwedge_{j=1}^{m} \left(\left(\bigwedge_{l'=1}^{k} A_{il'} \to B_{l'j} \right) \to B_{lj} \right), \tag{11}$$

which holds true iff for every j,

$$A_{il} \otimes \left(\bigwedge_{l'=1}^{k} A_{il'} \to B_{l'j} \right) \leqslant B_{lj}.$$

The last inequality is true because

$$A_{il} \otimes \left(\bigwedge_{l'=1}^{k} A_{il'} \to B_{l'j} \right) \leqslant A_{il} \otimes (A_{il} \to B_{lj}) \leqslant B_{lj}.$$

If every column of A is the extent induced by B, then $g(I_{i\_}) = A_{i\_}$ by definition of g. □

The next lemma describes the situation in which the rows of B are the intents corresponding to the columns of A (columns of A need not be extents).

Lemma 2. *Let $I = A \triangleleft B$. If every row of B is the intent induced by the corresponding column of A, i.e. $B_{l\_} = A_{\_l}^{\cap}$ then B is the smallest matrix for which $I = A \triangleleft B$. That is, if $I = A \triangleleft B'$ then $B_{lj} \leqslant B'_{lj}$ for every l and j.*

Proof. If every row of B is the intent induced by the corresponding column of A then, by definition, $B_{lj} = \bigvee_{i=1}^{m}(A_{il} \otimes I_{ij})$. If $I = A \triangleleft B'$, then using adjointness one can easily verify that $\bigvee_{i=1}^{m}(A_{il} \otimes I_{ij}) \leqslant B'_{lj}$, i.e. $B_{lj} \leqslant B'_{lj}$, verifying the claim. □

As a consequence, we get the following theorem:

Theorem 5. *Let $I = A_{\mathcal{F}} \triangleleft B_{\mathcal{F}}$ for a set $\mathcal{F} \subseteq \mathcal{B}(X^{\cap}, Y^{\cup}, I)$ of fixpoints. Then*

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

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

Proof. The first part follows directly from Lemmas 1 and 2. The fact that $A_{\mathcal{F}}$ is the largest one can be proved the same way we proved that B is the smallest one in Lemma 2. □

Theorem 5 shows another reason to look for decompositions of I in the form $I = A_{\mathcal{F}} \triangleleft B_{\mathcal{F}}$, i.e. reason to take fixpoints of $\langle {}^{\cap}, {}^{\cup} \rangle$ for factors. Namely, such an approach guarantees that g and h transform rows of I to rows of A and *vice versa.*

We now turn our attention to further properties of mappings g and h which are induced by B via (7) and (8). Note first that the pair $\langle g, h\rangle$ forms the (antitone) **L**-Galois connection induced by B, which was studied in [1] to which we refer for the properties of g and h mentioned below. First, g and h satisfy the following properties:

$$S(P, P') \leqslant S(g(P'), g(P)), \tag{12}$$
$$S(Q, Q') \leqslant S(h(Q'), h(Q)), \tag{13}$$
$$P \leqslant h(g(P)), \tag{14}$$
$$Q \leqslant g(h(Q)) \tag{15}$$

for any $P, P' \in L^m$ and $Q, Q' \in L^k$. Here, $S(\cdot, \cdot)$ denotes the subsethood degree defined for $G, H \in L^p$ by $S(G, H) = \bigwedge_{i=1}^{p}(G(i) \to H(i))$. A consequence of (12) and (13) is that $P \leqslant P'$ implies $g(P) \leqslant g(P')$ and $Q \leqslant Q'$ implies $h(Q) \leqslant h(Q')$. From (12), (13) it further follows that

$$g(a \to P) = a \to g(P), \tag{16}$$
$$h(a \to Q) = a \to h(Q), \tag{17}$$
$$g(P) = ghg(P), \tag{18}$$
$$h(Q) = hgh(Q), \tag{19}$$
$$g\left(\bigvee_{s \in S} P_s\right) = \bigwedge_{s \in S} g(P_s), \tag{20}$$
$$h\left(\bigvee_{t \in T} Q_t\right) = \bigwedge_{t \in T} h(Q_t), \tag{21}$$

see [1]. Properties (16) and (17) can be seen as properties which are analogous to homogeneity of linear mappings. Note that for $a \in L$ and $P \in L^m$, $Q \in L^k$, the vectors $a \to P$ and $a \to Q$ are defined by $(a \to P)_j = a \to P_j$ and $(a \to Q)_l = a \to Q_l$. Properties (20) and (21) say that g and h are dual $\bigvee$-morphisms; they can be seen as properties which are analogous to additivity of linear mappings. As a consequence, we get:

$$g\left(\bigvee_{s \in S}(a_s \to P_s)\right) = \bigwedge_{s \in S}(a_s \to g(P_s))$$

and

$$h\left(\bigvee_{t \in T}(a_t \to Q_t)\right) = \bigwedge_{t \in T}(a_t \to h(Q_t)).$$

The next theorem shows that g and h partition the space of attributes and the space of factors into particular convex subsets. A subset $S \subseteq L^p$ is called convex if $V \in S$ whenever $U \leqslant V \leqslant W$ for some $U, W \in S$. Let for $P \in L^m$ and $Q \in L^k$ 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. $g^{-1}(Q) = \{P \in L^m | g(P) = Q\}$, and $h^{-1}(P) = \{Q \in L^k | h(Q) = P\}$. We get:

Theorem 6

(i) $g^{-1}(Q)$ *is a convex partially ordered subspace of the attribute space and* $h(Q)$ *is the largest element of* $g^{-1}(Q)$.
(ii) $h^{-1}(P)$ *is a convex partially ordered subspace of the attribute space and* $g(P)$ *is the largest element of* $h^{-1}(P)$.

Proof. (i) Let $P \in g^{-1}(Q)$. Then, $Q = g(P)$, thus particularly, $Q \leqslant g(P)$. Using (14) and (13), $P \leqslant h(g(P)) \leqslant h(Q)$. Moreover, using (18) we get $Q = g(P) = ghg(P) = gh(Q)$, hence $h(Q) \in g^{-1}(Q)$. $h(Q)$ is thus the largest vector from $g^{-1}(Q)$. Let now $U, W \in g^{-1}(Q)$ and $U \leqslant V \leqslant W$. (12) yields $Q = g(U) \geqslant g(V) \geqslant g(W) = Q$, hence $g(V) = Q$, proving that $g^{-1}(Q)$ is convex. The proof of (ii) is dual. □

Theorem 6 provides us with the following insight to the transformations g and h: The space L^m of attributes and the space L^k of factors are partitioned into an equal number of convex subspaces (i.e. there is a bijective mapping between the subspaces of L^m and L^k) which have largest 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 largest vector from U.

4. Future research

This paper presented theorems regarding optimal triangular decompositions of matrices with degrees from residuated lattices. Most importantly, we proved that optimal decompositions, i.e. those with the smallest number of factors (smallest inner dimension) can be attained by using fixpoints of the isotone Galois connection associated with the input matrix. These fixpoints are known as formal concepts in formal concept analysis and can be computed by existing algorithms. Furthermore, we presented results describing transformations between the space of original attributes and the space of factors.

Future research will include a further study of triangular decompositions, including approximate decomposition of matrices, i.e. decompositions in which I is required to be approximately equal to $A \triangleleft B$. Another important problem is the problem of computing the optimal decompositions. As mentioned above, the fixpoints from $\mathscr{B}(X^{\cap}, Y^{\cup}, I)$, which are crucial for the optimal decompositions, can be computed using existing algorithms. Therefore, a similar approach can be followed as the one to the computation of optimal $\circ$-decompositions which is presented in [7,8]. The third problem is to compare the resulting method of factor analysis based on triangular decompositions to classic methods of factor analysis. Belohlavek and Vychodil [7,8] indicate that relational decompositions, of which the $\circ$-decomposition as well as the $\triangleleft$-decomposition are particular examples, have the ability to reveal factors which are not revealed by classic factor analysis. This is not surprising because the mathematics behind relational decompositions is completely different from the mathematics behind classic factor analysis. One important aspect is that factor analysis based on relational decompositions is congruent with the semantics of relational data, such as binary or ordinal data. As a result, the factors delivered by factor analysis based on relational decompositions are easy to interpret and have a natural meaning. Contrary to that, as repeatedly observed in the literature, see e.g. [25], classic factor analysis of relational data delivers results which are difficult to interpret. A typical example, reported in the literature, is negative real-valued coefficients which typically result in classic factor analysis of binary data. A thorough study of these methodological problems also remains for future research. An important issue for further research regarding in particular triangular decompositions of graded matrices which are not binary is applying these decompositions to data from various areas. In [8] we presented an analysis of 2004 Olympic Decathlon data using $\circ$-decomposition mentioned in Section 1.3 which demonstrates that natural factors can be revealed from graded data using such decompositions. We assume that triangular decompositions will be useful in factor analysis of data from psychology due to their ability to explain graded attributes in terms of satisfaction of a small number of conditions.

Acknowledgements

I thank the anonymous reviewers for helpful comments. This paper is an extended version of a contribution presented at ROGICS 2008. Support by research plan MSM 6198959214 and by Grant No. 1ET101370417 of GA AV ČR is gratefully acknowledged.

References

[1] R. Belohlavek, Fuzzy Galois connections, Mathematical Logic Quarterly 45 (1999) 497–504.
[2] R. Belohlavek, Fuzzy closure operators, Journal of Mathematical Analysis and Applications 262 (2001) 473–489.
[3] R. Belohlavek, Fuzzy Relational Systems: Foundations and Principles, Kluwer, Academic/Plenum Publishers, New York, 2002.
[4] R. Belohlavek, Concept lattices and order in fuzzy logic, Annals of Pure and Applied Logic 128 (1–3) (2004) 277–298.
[5] R. Belohlavek, Optimal decompositions of matrices with entries from residuated lattices, submitted for publication.
[6] R. Belohlavek, B. De Baets, J. Outrata, V. Vychodil, Computing all fixpoints of a fuzzy closure operator, submitted for publication.
[7] 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, in press, doi:10.1016/j.jcss.2009.05.002.
[8] R. Belohlavek, V. Vychodil, Factor analysis of incidence data via novel decomposition of matrices, Lecture Notes in Computer Science 5548 (2009) 83–97.

[9] A. Di Nola, E. Sanchez, W. Pedrycz, S. Sessa, Fuzzy Relation Equations and Their Applications to Knowledge Engineering, Kluwer, Dordrecht, 1989.
[10] A.A. Frolov, D. Húsek, I.P. Muraviev, P.A. Polyakov, Boolean factor analysis by Hopfield-like autoassociative memory, IEEE Transactions on Neural Networks 18 (3) (2007) 698–707.
[11] B. Ganter, R. Wille, Formal Concept Analysis. Mathematical Foundations, Springer, Berlin, 1999.
[12] G. Georgescu, A. Popescu, Non-dual fuzzy connections, Archive for Mathematical Logic 43 (2004) 1009–1039.
[13] J.A. Goguen, L-fuzzy sets, Journal of Mathematical Analysis and Applications 18 (1967) 145–174.
[14] S. Gottwald, A Treatise on Many-Valued Logics, Research Studies Press, Baldock, Hertfordshire, England, 2001.
[15] S. Gottwald, Mathematical fuzzy logics, Bulletin of Symbolic Logic 14 (2) (2008) 210–239.
[16] P. Hájek, Metamathematics of Fuzzy Logic, Kluwer, Dordrecht, 1998.
[17] U. Höhle, On the fundamentals of fuzzy set theory, Journal of Mathematical Analysis and Applications 201 (1996) 786–826.
[18] J. Järvinen, M. Kondo, J. Kortelainen, Logics from Galois connections, International Journal of Approximate Reasoning 49 (3) (2008) 595–606.
[19] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic. Theory and Applications, Prentice-Hall, 1995.
[20] L.J. Kohout, W. Bandler, Relational-product architectures for information processing, Information Sciences 37 (1-3) (1985) 25–37.
[21] S. Krajči, A generalized concept lattice, Logic Journal of IGPL 13 (5) (2005) 543–550.
[22] H. Lai, D. Zhang, Concept lattices of fuzzy contexts: formal concept analysis vs. rough set theory, International Journal of Approximate Reasoning 50 (5) (2009) 695–707.
[23] D.S. Nau, G. Markowsky, M.A. Woodbury, D.B. Amos, A mathematical analysis of human leukocyte antigen serology, Mathematical Biosciences 40 (1978) 243–270.
[24] S. Pollandt, Fuzzy Begriffe, Springer-Verlag, Berlin, 1997.
[25] P. Miettinen, T. Mielikäinen, A. Gionis, G. Das, H. Mannila, The discrete basis problem, Lecture Notes in Artificial Intelligence 4213 (2006) 335–346.
[26] M. Ward, R.P. Dilworth, Residuated lattices, Transactions of the American Mathematical Society 45 (1939) 335–354.