Statistica & Applicazioni - Special Issue, 2011, pp. 61-79

FINDING INCOMPARABLE PAIRS OF SUBSETS BY USING FORMAL CONCEPT ANALYSIS

Hans-Georg Bartel*
Hans-Joachim Mucha**

SUMMARY

In the paper presented here, we use Formal Concept Analysis (FCA) to solve a problem that arises when working with partially ordered sets (posets). In detail, the task here is to look for incomparable subsets which are related to a given poset. A way to solve this problem is to use FCA based on a context which can be derived in some steps from the ζ-matrix of the (simple directed) graph corresponding to the given poset. The requested incomparable subsets result from the set of concepts obtained from this context. For illustrative purposes, small toy data sets are presented. At the end, a real data application to environmental chemistry is given in detail. The data consist of ten chemicals found in the German river Main. As the result a set of twelve incomparable pairs of subsets are figured out.

Keywords: *Bipartite Graph, Adjacency Matrix, Formal Concept Analysis, Partially Ordered Set, Incomparability of Sets.*

1. INTRODUCTION

Rainer Brüggemann set the task of finding a simple method that provides a complete system of incomparable pairs of sets of a given partially ordered set. Brüggemann is owed the essential mathematical foundations of the use of the so-called Hasse diagram technique ((Brüggemann and Halfon, 1995) and many subsequent publications) based on the pioneering paper of Efraim Halfon and Marcello G. Reggiani (Halfon and Reggiani, 1986). The importance of the problem mentioned by Rainer Brüggemann has been underlined with the remark: ''One of the main ideas in partial order theory or in Hasse diagram technique, respectively, is the concept of incomparability''.

The main idea to solve the Brüggemann's task consists (a) of transforming a given partial order $\leq$ (D16) over a set of objects G into the symmetric relation of incomparability $\|$ (D17) over the same set G,

$$(G, \leq) \rightarrow (G, \|), \text{ where } g_i \| g_j \text{ if neither } g_i \leq g_j \text{ nor } g_j \leq g_i \ (g_i, g_j \in G),$$

and (b) of finding all pairs of incomparable sets (D18) $\{G_m, G_n\}$ with $G_m, G_n \subset G$

* Humboldt University Berlin, Department of Chemistry, Brook-Taylor-Straße 2, D-12489 BERLIN, Germany (e-mail: ✉ hg.bartel@yahoo.de)
** Weierstrass Institute of Applied Analysis and Stochastics, Mohrenstraße 39, D-10117 BERLIN, Germany (e-mail: mucha@wias-berlin.de)

and $G_m \| G_n$. For the latter, an algorithm developed by Bernhard Ganter (1984; 2010) is particularly suitable. It allows calculating the set of all concepts in Definition 13 within the scope of Formal Concept Analysis (FCA) (Section 3). In doing so, $(G, \|)$ has to be interpreted as the context in Definition 11 $\mathcal{C} = (G, G_M, \|)$[1].

Brüggemann's task arises in the framework of Hasse diagram technique that uses the terminology of graph theory. Therefore, it is useful to commemorate some basic definitions and concepts of this theory (Sachs, 1970; Chartrand and Zhang, 2005; Bondy and Murty, 2008) at the beginning (Section 2).

2. SOME GRAPH THEORETICAL CONCEPTS AND DEFINITIONS

DEFINITION 1

*A graph $\Gamma = (V, A, \varphi)$ is a triple of two disjoint sets V (vertices) and A (edges, French: **a**rêtes), and a total mapping $\varphi : A \to P_{(2)}(V)$, where $P_{(2)}(V)$ is a set of (either directed or undirected) pairs of vertices. Two types of graphs should be distinguished:*

(*a*) $\Gamma_d = (V, B, \varphi_d)$ *is a* **directed graph**, *if each element b of B is assigned to an ordered pair $(v_i, v_j) \in V \times V$. The elements of B are called directed edges or arcs (German: **B**ögen).*

(*b*) $\Gamma_u = (V, E, \varphi_u)$ *is an* **undirected graph**, *if each element of E is assigned to an unordered pair $\{v_i, v_j\} \in$((V). Here ((V) is the power set of V. The elements of E are called (undirected) edges.*

Two special graphs are to be mentioned:

(*c*) *Both the graph $\Gamma^{(1,1)} = (\{v\},\{l\},\varphi)$ with $\varphi(l) = (v, v)$ or $\varphi(l) = \{v\}$, respectively, and the graph $\Gamma^{(1,0)} = (\{v\},\emptyset,\emptyset)$ are called* **point graph**.

(*d*) *The graph without any vertices (and hence without edges) is called the* **empty graph** *or the* **null graph**. *It is denoted by $\emptyset$.*

DEFINITION 2

Let $e_h \in E$ be an edge in Γ_u and $b_h \in B$ an arc in Γ_d. Then e_h (b_h) is called a **loop** *(directed loop), if $\varphi_u(e_h) = \{v_i\} \in$((V) ($\varphi_d(b_h) = (v_i, v_i) \in V \times V$).*

DEFINITION 3

$\Gamma^s = (V, A, \varphi)$ *is a* **simple graph**, *if all elements $P_h^{(2)}$ of $P_{(2)}(V)$ given by the total mapping $\forall a_j \in A : \varphi(a_j) = P_h^{(2)} \in P_{(2)}(V)$ are the image of exactly one element of A:*

$$\forall(P_h^{(2)} \in P_{(2)}(V),\ a_i, a_j \in A) : (\varphi(a_i) = P_h^{(2)}) \wedge (\varphi(a_j) = P_h^{(2)}) \Rightarrow a_i = a_j,$$

$$\forall(P_g^{(2)}, P_h^{(2)} \in P_{(2)}(V),\ a_i \in A) : (\varphi(a_i) = P_g^{(2)}) \wedge (\varphi(a_i) = P_h^{(2)}) \Rightarrow P_g^{(2)} = P_h^{(2)}.$$

The Definition 3 admits that each of the vertices of a simple graph can be assigned to either none or only one loop.

---

[1] G and G_M have factually the same elements, but the elements of G represent objects and the elements of G_M attributes.

Definition 4
The graph $\Gamma_S = (V_S, A_S, \varphi_S)$ *is a* **subgraph** *of the graph* $\Gamma_H = (V_H, A_H, \varphi_H)$, *if* $V_S \subseteq V_H$, $A_S \subseteq A_H$, *and* $\varphi_S \subseteq \varphi_H$.

Definition 5
A path (German: ***Weg****) in the simple graph* $\Gamma^s = (V, A, \varphi)$ *is a sequence of vertices* $W_{v_\alpha - v_\omega} = (v_\alpha, v_{\alpha+1}, \ldots, v_{\iota-1}, v_\iota, v_{\iota+1}, \ldots, v_{\omega-1}, v_\omega)$ *such that for each of its vertices* v_i $(\alpha \leq i < \omega)$ *exists an edge* a_h *with either* $\varphi_d(a_h) = (v_i, v_{i+1})$, $\varphi_d(a_h) = (v_{i+1}, v_i)$, *or* $\varphi_u(a_h) = \{v_i, v_{i+1}\}$. $W_{v_\alpha - v_\alpha}$ *is called a* **cycle** *if its start vertex* v_α *and its end vertex* v_ω *is the same one. A single vertex* v_α *can be considered as a path* $W_{v_\alpha \bullet} = (v_\alpha, v_\alpha)$.

Obviously, the relation
W: $v_h W v_j$ = 'the vertex $v_h \in V$ is connected to the vertex $v_j \in V$ by a path W_{h-j}' (1)

is an equivalence relation that generates a decomposition of the set of vertices V of the graph $\Gamma = (V, A, \varphi)$ in equivalence classes $\{V_1, V_2, \ldots, V_n\} = V$. This partition causes the decomposition of the sets A and φ in the classes $\{A_1, A_2, \ldots, A_n\} = A$ and $\{\varphi_1, \varphi_2, \ldots, \varphi_n\} = \varphi$, respectively, where the elements of A_i and those of φ_i relate only to those of V_i (i = 1, 2, ..., n).

Definition 6
Let $\Gamma = (\{V_1, V_2, \ldots, V_n\}, \{A_1, A_2, \ldots, A_n\}, \{\varphi_1, \varphi_2, \ldots, \varphi_n\})$ *be the result which is obtained when the relation* **W** (1) *was applied to the set* V *of the graph* $\Gamma = (V, A, \varphi)$. *Then each of the subgraphs* $\Gamma_i = (V_i, A_i, \varphi_i)$ (i = 1, 2, ..., n) *is called a* **component** *of* Γ.
If the graph $\Gamma = (V, A, \varphi)$ has only one component (n = 1)*, it is called a* **connected** *one.*

The following partition of the set of vertices V into two classes is of particular importance in this paper:

Definition 7
(*a*) *The graph* $\Gamma_{bp} = (V, A, \varphi) = (\{V_B, V_W\}, A, \varphi)$ *is a* **bipartite graph***, if its set of vertices* V *is partitioned in two classes* V_B and V_W (*where* $(V = V_B \cup V_W,\ V_B \cap V_W = \emptyset)$ *such that every edge of the set* A *connects a vertex in* V_B *with one in* V_W.
(*b*) *The graph* $\Gamma_{cbp} = (\{V_B, V_W\}, A, \varphi)$ *is a* **complete bipartite graph***, if for any two vertices,* $v_b \in V_B$ *and* $v_w \in V_W$, *either* (v_b, v_w), (v_w, v_b) or $\{v_b, v_w\}$ *is an edge of* A.

As examples, graphical representations of some undirected, simple bipartite graphs are shown in Figure 1.

Definition 8
Let $\Gamma_u^s = (V, E, \varphi_u)$ *be an undirected simple graph,* $\varphi_u : E \rightarrow P_{(2)}(V)$, *and* $P_{(2)}^{all}(V)$

the set of all pairs $\{v_i, v_j\}$ *(*$v_i, v_j \in V$*, i = 1, ..., |V|, j = i, ..., |V|). Then the graph* $\bar{\Gamma}^s_u = (V, \bar{E}, \bar{\varphi}_u)$ *is called the* **complement** *or* **inverse** *of the graph* Γ^s_u*, if* $\bar{\varphi}_u$ *is the total mapping* $\bar{\varphi}_u : \bar{E} \to P^{all}_{(2)}(V) \backslash P_{(2)}(V)$*. Therefore, for the set of edges* $\bar{E}$ *in* $\bar{\Gamma}^s_u$ *holds:*

$$\forall \bar{e}_h \in \bar{E} : \bar{e}_h = \{v_i, v_j\} \in \bar{E} \quad \textit{iff} \quad \{v_i, v_j\} \notin E \text{ in } \Gamma^s_u \ (v_i, v_j \in V).$$

Figure 2 shows the graphical representations of three $\Gamma^s_u/\bar{\Gamma}^s_u$-pairs (without consideration of loops).

Of course, the simple (undirected) graph Γ^s_u is the complement $\bar{\bar{\Gamma}}^s_u$ of its complement $\bar{\Gamma}^s_u$: $\bar{\bar{\Gamma}}^s_u = \Gamma^s_u$. Let be $\Gamma^{s,(c)}_u = (V, E^{(c)}, \varphi^{(c)}_u)$ a simple graph in which every pair of distinct vertices is connected by a unique edge and with exactly one loop at each vertex, and $\Gamma^s_u = (V, E, \varphi_u)$a simple graph with the same set of vertices V as that of $\Gamma^{s,(c)}_u$, $E \subseteq E^{(c)}$, and $\varphi_u \subseteq \varphi^{(c)}_u$. Then $\Gamma^s_u \cup \bar{\Gamma}^s_u = \Gamma^{s,(c)}_u$ holds. To illustrate this relationship, an example is given in Figure 3.

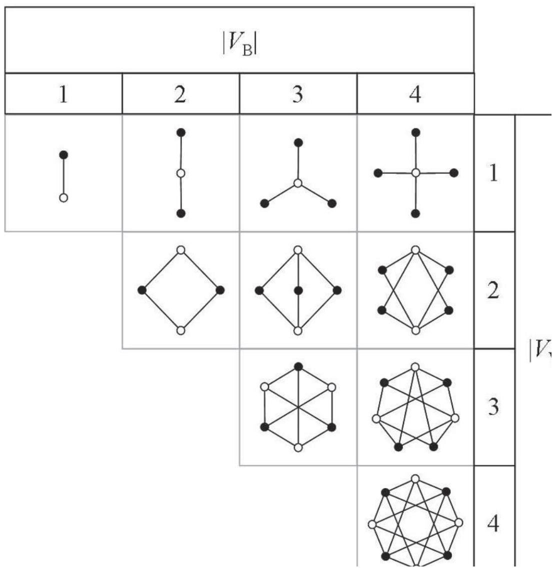


FIGURE 1. - *Graphical representations of undirected, simple complete bipartite* graphs $\Gamma^s_{u,cbp} = (\{V_B, V_W\}, E, \varphi_u)$
(elements of V_B: ● , elements of V_W: ○ , elements of E: –)

At the end of this section, two matrices for representing graphs have to be mentioned.

Definition 9

*Let $\Gamma_d^s = (V, B, \varphi_d)$ be a simple directed graph with $|V| = n$ vertices, which are numbered 1 to n. Then the $n \times n$-*matrix *$\zeta = (\zeta_{ij})$ with*

$$\zeta_{ij} = \begin{cases} 1 & \text{if } \exists b \in B \text{ with } \varphi_d(b) = (v_i, v_j) \in V \times V \\ 0 & \text{else} \end{cases}.$$

is called the **ζ-matrix** *of the graph Γ_d^s.*

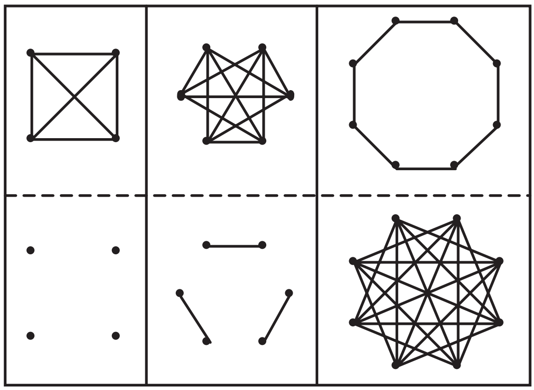


FIGURE 2. - *Three examples of graph/complement pairs (Loops were not drawn)*

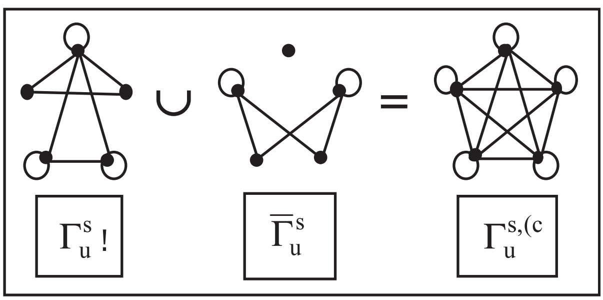


FIGURE 3. - *An example of the graphical illustration of the relationship $\Gamma_u^s \cup \bar{\Gamma}_u^s = \Gamma_u^{s,(c)}$ (• vertex, — edge, ○ loop)*

Definition 10

*The $n \times n$-*matrix *$\alpha = (\alpha_{ij})$ is the* **adjacency matrix** *of the undirected graph $\Gamma_u = (V, E, \varphi_u)$* with $|V| = n$ vertices, which are numbered 1 to n, if

$$\alpha_{ij} = \begin{cases} |E_q| \text{ where } E_q = \{e_r \in E | \varphi_u(e_r) = \{v_i, v_j\}\} \subseteq E \ (v_i, v_j \in V) & \textit{for } i \neq j \\ 2|E_l| \text{ where } E_l = \{e_m \in E | \varphi_u(e_m) = \{v_i\}\} \subseteq E \ (v_i \in V) & \textit{for } i = j \end{cases}.$$

Adjacency matrices may be defined for directed graphs $\Gamma_d = (V, B, \varphi_d)$, however, this case is of no interest in this paper.

3. SOME REMARKS ON FORMAL CONCEPT ANALYSIS (FCA)

The Formal concept analysis (Ganter and Wille, 1998; Ganter, Stumme and Wille, 2005) has been developed by Rudolf Wille in 1984. It belongs to a field of applied mathematics based on the mathematical formalization of concept and concept hierarchy, whereby it allows representing, analyzing, and constructing conceptual structures in a mathematical way. As examples of the application of FCA to chemistry, see (Bartel and Nofz, 1997; Bartel and Brüggemann, 1998).

Now let us define a context, the basic structure of FCA:

Definition 11
A (single-valued) **context** C *is defined as the triple* $\mathsf{C} = (G, M, H)$, *where G is the set of objects (German:* ***G****egenstände), M is the set of attributes (German:* ***M****erkmale), and* $H \subseteq G \times M$ *is the relation gHm (or* $(g, m) \in H$*) with* $g \in G$ *and* $m \in M$ *which is read: 'The object g has the attribute m.'*

Such a context can be written down as a cross table where the objects of G form the rows and the attributes of M form the columns. A cross **X** means, that the relation gHm is true.

For the definition of a (formal) concept of a given context, first the following definition is needed:

Definition 12
Let $\mathsf{C} = (G, M, H)$ *be a context,* G_I *a subset of G* ($G_I \subseteq G$)*, and* M_J *a subset of M* ($M_J \subseteq M$)*. Then*
(a) G_I' *is the set of the attributes which all objects of* G_I *have:*

$$G_I' = \{m \in M | \forall g_I \in G_I : \ g_I H m\},$$

(b) M_J' *is the set of objects which are valid for all attributes of* M_J*:*

$$M_J' = \{g \in G | \forall m_J \in M_J : \ g H m_J\}.$$

Definition 13
A (formal) **concept** *of the context* $\mathsf{C} = (G, M, H)$ *is defined to be an ordered pair* (G_I, M_J) *with* $G_I \subseteq G$, $M_J \subseteq M$, $G_I' = M_J$ *and* $M_J' = G_I$. G_I *and* M_J *are called the* **extent** *and the* **intent** *of the concept* (G_I, M_J)*, respectively.*

Both, the following definition and the following theorem are fundamental for the FCA.

Definition 14
The set of all concepts $\mathsf{B}(\mathsf{C})$ *of the context* $\mathsf{C} = (G, M, H)$ *can be partially ordered*

by inclusion: if (G_U, M_U), (G_O, M_O)$\in$ $B(C)$ *are concepts a partial order* (see Definition 15) *denoted by* $\leq_{SSR}$, *the* **subconcept-superconcept relation**, *is defined by*

$$(G_U, M_U) \leq_{SSR} (G_O, M_O) \Leftrightarrow G_U \subseteq G_O \ (M_O \subseteq M_U).$$

(G_U, M_U) is called the subconcept of the superconcept (G_O, M_O).

THEOREM 1 (*Fundamental Theorem of FCA*)
The partially ordered set (*D16*) ($B(C)$, $\leq_{SSR}$) *corresponds to the mathematical structure of a complete lattice, the* **concept lattice**[2] *of the context* C. *Here,* $B(C)$ *denotes the set of all concepts of the context* C, *and* $\leq_{SSR}$ *means the subconcept-superconcept relation defined in* (Definition 14).

The concept lattice of a context C reflects the generalized hierarchy and conceptual structure corresponding to C. Thereof will be made use in a later publication.

The determination of the complete set of concepts $B(C)$ of a context C is sufficient to solve the task that Rainer Brüggemann set. To achieve this aim, it is appropriate to establish a link between the graph theoretical terms of Section 2 and those of FCA.

4. CONTEXTS AND CONCEPTS AS BIPARTITE GRAPHS

It may be noted that a context $C = (G, M, H)$ (D11) can formally be considered as a directed bipartite graph $\Gamma_{d,bp}(C) = (\{G, M\}, B, H)$. From $H \subseteq G \times M$ (see Definition 11) and $H \subseteq (G \cup M)^2$ (see Definition 1a) obviously follows $G^2 = M^2 = M \times G = \emptyset$. The elements of B result from the mapping $H : B \rightarrow G \times M$. The ζ-matrix of the graph $\Gamma_{d,bp}(C)$ corresponding to the context C is therefore reduced to a $|G|\times|M|$-Matrix $\zeta_{GM}(\zeta_{ij}^{GM})$ with the elements

$$\zeta_{ij}^{GM} = \begin{cases} 1 & if \ (g_i, m_j) \in H \\ 0 & if \ (g_i, m_j) \notin H \end{cases} \quad (g_i \in G, \ m_j \in M, \ i = 1, \ldots, |G|, \ j = 1, \ldots, |M|) \ .$$

With the substitutions $1 \rightarrow$ **X** and $0 \rightarrow$ 'blank' the ζ_{GM}-matrix is now converted into the cross table of the context $C = (G, M, H)$.

The example of the four elements of Aristotle (384–322 BC) may be useful for a demonstration. This Greek philosopher had associated the elements of Empedocles (c. 490–430 BC) G_{Emp} = {IGNIS, TERRA, AQVA, AER} with the four haptic attributes M_{haptic} = {SICCVS, HVMIDVS, CALDVS, FRIGIDVS} as shown in Figure 4.

With the relation $g_E I m_h$ ('The element $g_E \in G_{Emp}$ is $m_h \in M_{haptic}$.'), the context C_{Arist} = (G_{Emp}, M_{haptic}, I) with the following cross table will be obtained:

[2] A **lattice** (Birkhoff, 1967) is a poset (see Definition 16) in which any two elements have a supremum (least upper bound) and an infimum (greatest lower bound).

		M_{haptic}			
		SICCVS (s)	FRIGIDVS (f)	HVMIDVS (h)	CALDVS (c)
G_E	IGNIS (IG)	**X**			**X**
	TERRA (TE)	**X**	**X**		
	AQVA (AQ)		**X**	**X**	
	AER (AE)			**X**	**X**

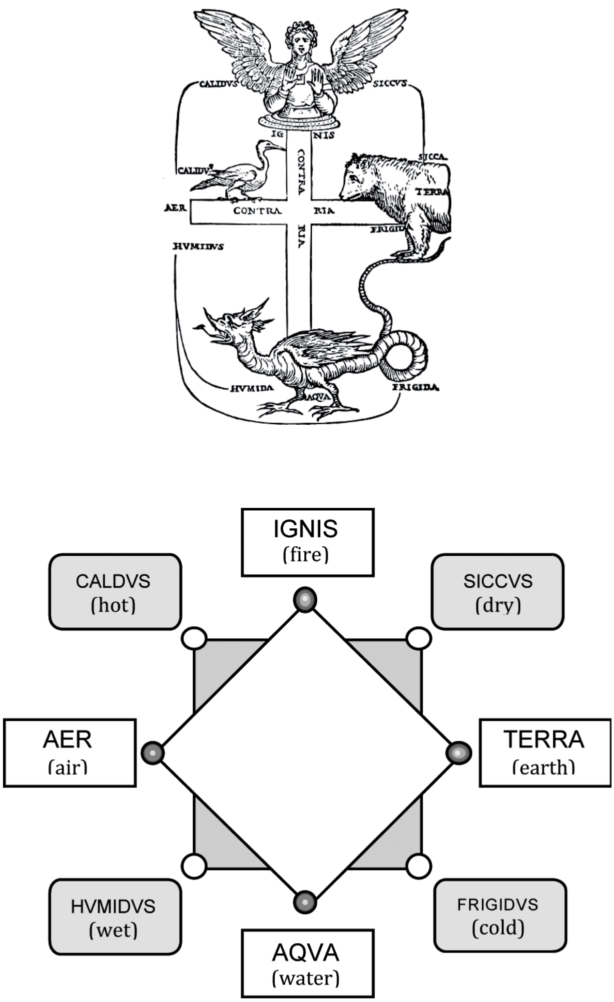


FIGURE 4. - *The Four Elements of Aristotle shown in an alchemical and in a modern representation*

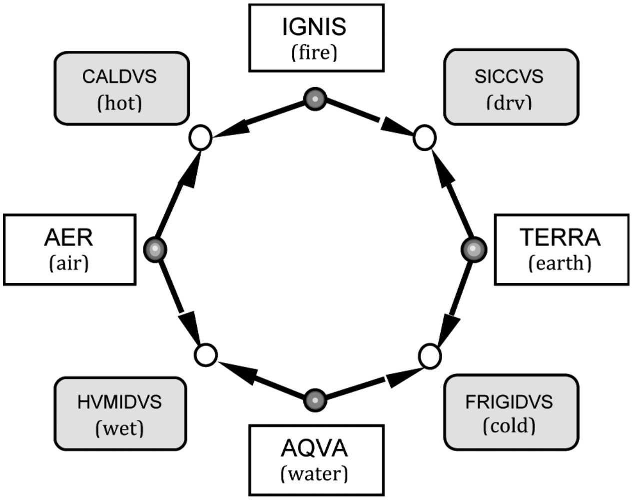


FIGURE 5. - *The bipartite graph* $\Gamma_{\mathrm{d,bp}}(C_{\mathrm{Arist}})$ (◉ $\in G_{Emp}$, ○ $\in M_{haptic}$, *arcs:* ⟶)

The ζ-matrix corresponding with the context C_{Arist} is given by

$$\zeta(C_{\mathrm{Arist}}) = \begin{array}{c} \begin{array}{cccc} \mathrm{s} & \mathrm{f} & \mathrm{h} & \mathrm{c} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array} \begin{array}{c} \mathrm{IG} \\ \mathrm{TE} \\ \mathrm{AQ} \\ \mathrm{AE} \end{array} .$$

A graphical representation of $\zeta(C_{\mathrm{Arist}})$ by a bipartite graph $\Gamma_{\mathrm{d,bp}}(C_{\mathrm{Arist}})$ is shown in Figure 5.

The concepts $B(C)$ associated with a context $C = (G, M, H)$ can also be represented as bipartite graphs. If $(G_I, M_J) \in B(C)$ is a concept (D13) then the corresponding graph $\Gamma^{IJ}_{d,cbp} = (\{G_I, M_J\}, B_{IJ}, H_{IJ})$is a complete bipartite subgraph of the graph $\Gamma_{\mathrm{d,bp}}(C)$ that corresponds to the context C. $B(C)$ can be regarded as the total set of these subgraphs $\Gamma^{IJ}_{d,cbp}$. In the cross table the concepts are maximum rectangles of crosses (after eventually doing some re-sorting).

With regard to these statements, a broadening of the definition of complete bipartite graphs is necessary because in the set $B(C)$ the concepts $(G, \varnothing)$ (meaning: All objects do not have an attribute in common.) and $(\varnothing, M)$ (meaning: No object has all the attributes.) can occur (as 1- or 0-element of the concept lattice). The associated graphs $\Gamma^{Y\varnothing}_{cbp} = (\{Y, \varnothing\}, \varnothing, \varnothing)$ ($Y = G, M$) consist on the one hand of the null graph $\varnothing$ and the other hand of a graph with $|Y|$ components which are the point graphs $\Gamma^{(1,0)}_i = (\{y_i\}, \varnothing, \varnothing)$ ($y_i \in Y$, $i = 1, \ldots, |Y|$). Of course, these graphs have no edges.

In the above example of the context $C_{\mathrm{Arist}} = (G_{\mathrm{Emp}}, M_{\mathrm{haptic}}, I)$ ten concepts can be identified that are listed in the table below. These concepts are associated with each a type of complete bipartite graphs illustrated in Figure 6.

No.	IG	TE	AG	AE	s	f	h	c
1	X	X	X	X				
2	X			X				X
3			X	X			X	
4		X	X			X		
5	X	X			X			
6				X			X	X
7			X			X	X	
8		X			X	X		
9	X				X			X
10					X	X	X	X

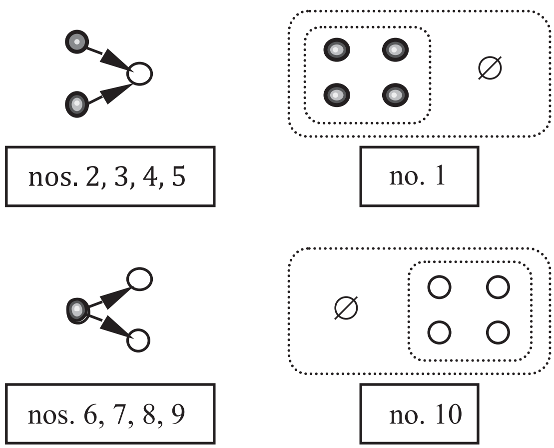


FIGURE 6. - *The ten concepts of the context* C_{Arist} *and the types of graphs assigned to them* (◉ $\in G_{Emp}$, ○ $\in M_{haptic}$, *arcs:* ⟶, *null graph:* ∅)

5. PARTIALLY ORDERED SETS AND THEIR GRAPHS

Now we need to recall three definitions that are important to solve Brüggemann's task:

DEFINITION 15
A **partial order** *over a set S is a binary relation denoted by* $\leq$ *which is reflexive, antisymmetric, and transitive, i.e.*

- (*reflexivity*) $\forall s \in S : s \leq s$
- (*antisymmetry*) $\forall s, t \in S : (s \leq t) \wedge (t \leq s) \Rightarrow s = t$
- (*transitivity*) $\forall r, s, t \in S : (r \leq s) \wedge (s \leq t) \Rightarrow r \leq t$.

DEFINITION 16
Let S be a set, in which a partial order $\leq$ is declared, then $(S, \leq)$ *is called a* **partially ordered set** *or* **poset**.

DEFINITION 17
If $(S, \leq)$ *is a poset then the elements* $s, t \in S$ *are called* **comparable** (*denoted by st*) *if* $s \leq t$ *or* $t \leq s$. *Otherwise the elements* $s, t \in S$ *are* **incomparable** (*denoted by* $s \| t$).

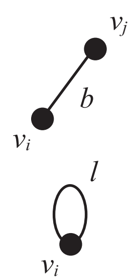
Both, comparability and incomparability are symmetric relations. Of course, only the first one is reflexive.

Like any set in which a relation is declared, a poset $(S, \leq)$can be assigned a graph called the relation graph $\Gamma_d^s(S, \leq) = (S, B_{\leq}, \varphi_{\leq})$ where $b_{ij} \in B_{\leq}$ with $\varphi_{\leq}(b_{ij}) = (s_i, s_j) \in S \times S$ if $s_i \leq s_j$ holds in $(S, \leq)$. This relation graph $\Gamma_d^s(S, \leq) = (S, B_{\leq}, \varphi_{\leq})$ has the following properties:

- Because of the antisymmetry of $\leq$, $\Gamma_d^s(S, \leq)$ is directed and simple.
- Since $\leq$ is reflexive, each vertex $s \in S$ has a loop $b_s \in B_{\leq}$ with $\varphi_{\leq}(b_s) = (s, s) \in S \times S$.
- The transitivity of $\leq$ causes that the arc $b_{hj} \in B_{\leq}$ with $\varphi_{\leq}(b_{hj}) = (s_h, s_j)$ exists if the arcs $b_{hi}, b_{ij} \in B_{\leq}$ ($\varphi_{\leq}(b_{hi}) = (s_h, s_i)$,$\varphi_{\leq}(b_{ij}) = (s_i, s_j)$) are present. Therefore, there is no cycle $W_{s_\alpha - s_\alpha}$ (apart from loops) of $\Gamma_d^s(S, \leq)$ in which all the arcs have the same direction.

Before drawing the graphical representation of graphs $\Gamma_d^s(S, \leq)$, the following convention should be agreed:

CONVENTION 1
If in the directed graph $\Gamma_d = (V, B, \varphi_d)$ *the mapping of the arc* $b \in B$ *is* $\varphi_d(b) = (v_i, v_j) \in V \times V$ *then the graphical representation of b can be a line (instead of an arrow) where in the paper plane, the graphical representation of the vertex* v_i *must be located below the graphical representation of the vertex* v_j.
A loop $l \in B$*at the vertex* $v_i \in V$ *is drawn on the usual way, but without direction indicator.*



For demonstration we choose the poset $(N, |)$ as an example where N is a set of ten integers: N = {2,3,4,6,8,15,18,45,72,90}, and | is the partial order $m|n$ = 'The integer m is a divisor of the integer n.'. $(N, |)$ is assigned the graph $\Gamma_d^s(N, |) = (N, B_|, \varphi_|)$. Here $b_{ij} \in B_|$ with $\varphi_|(b_{ij}) = (n_i, n_j) \in N \times N$ if the relation $n_i | n_j$ is true in $(N, |)$.

Taking into account the Convention 1, the graphical representation of the graph $\Gamma_d^s(N, |) = (N, B_|, \varphi_|)$ shown in Figure 7 is obtained.

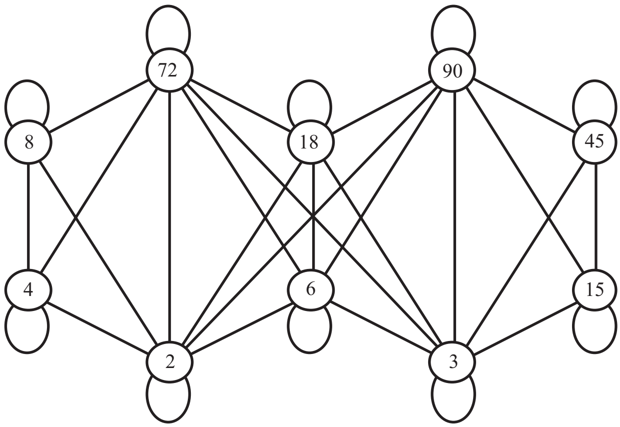


FIGURE 7. - (a) *Convention 1 is taken into account: Graphical representation of the graph* $\Gamma_d^s(N, |)$ *(N =* {2,3,4,6,8,15,18,45,72,90}*: set of vertices)* or (b) *Convention 1 is not taken into account: Graphical representation of the graph* $\Gamma_u^s(N, -)$*, respectively*

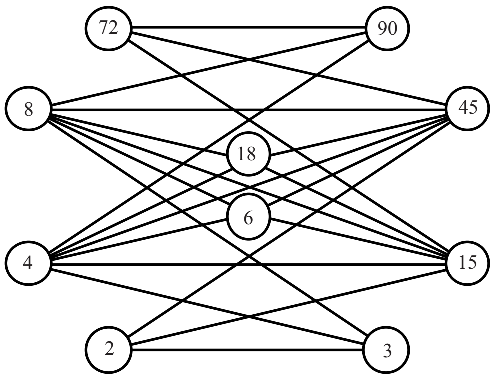


FIGURE 8. - *Graphical representation of the complement of the graph* $\Gamma_u^s(N, -)$ (see Figure 7b): $\bar{\Gamma}_u^s(N, -) = \Gamma_u^s(N, \|)$)

Converting of the directed relation graph $\Gamma_d^s(S, \leq) = (S, B_\leq, \varphi_\leq)$ into an undirected one $\Gamma_u^s(S, -) = (S, E_-, \varphi_-)$ using the unique assignments

$$\forall b_{ij} \in B_\leq : (b_{ij} \rightarrow e_{ij} \in E_-) \wedge (\varphi_\leq(b_{ij}) = (s_i, s_j) \in S \times S) \rightarrow \varphi_-(e_{ij}) = \{s_i, s_j\} \in \Pi(S)$$

then an edge e_{ij} and the mapping $\varphi_-(e_{ij}) = \{s_i, s_j\}$, respectively, show that the elements s_i and s_j of S are comparable s_i — s_j:

$$\forall e_{ij} \in B_- \textit{ with } \varphi_-(e_{ij}) = \{s_i, s_j\} : e_{ij} \in B_- \Leftrightarrow s_i \;\; — s_j.$$

If the Convention 1 is not observed in Figure 7 then the picture shown there represents the undirected graph $\Gamma_u^s(N,-) = (N, E_-, \varphi_-)$.

Let $\Gamma_d^s(S,\leq)$ be associated with the ζ-Matrix $(S,\leq)$ then $\Gamma_u^s(S,-)$ can be derived from the adjacency matrix $(S,-)$. As for the elements ζ_{ij} of $(S,\leq)$

$$\zeta_{ij} = 1 \Rightarrow \zeta_{ji} = 0 \quad (i \neq j) \text{ and } \forall \zeta_{ii} \in (S,\leq)\text{: } \zeta_{ii} = 1$$

holds, the relationship between the two matrices is

$$(S,\leq) + \zeta^{\mathsf{T}(S,\leq)} = (S,-), \tag{2}$$

where $\zeta^{\mathsf{T}(S,\leq)}$ is the transposed matrix of the matrix $(S,\leq)$, i.e.: $\zeta_{ij} \in (S,\leq)$, $\zeta_{ji}^T \in \zeta^{\mathsf{T}(S,\leq)}$: $\zeta_{ij} = \zeta_{ji}^T$.

Obviously the complement $\bar{\Gamma}_u^s(S,-)$ of the graph $\Gamma_u^s(S,-)$ is identical to the graph $\Gamma_u^s(S,\|)$ whose edges connect all pairs of vertices where their associated elements are incomparable. In Figure 8, the complement $\bar{\Gamma}_d^s(N,-) = \Gamma_d^s(N,\|)$ of the graph $\Gamma_d^s(N,-)$ (Figure 7b) is shown.

The adjacency matrix $\alpha(S,\|)$ of the graph $\Gamma_u^s(S,\|)$ is obtained from the one $(S,-)$of $\Gamma_u^s(S,-)$ (eq. (2)) by the following transformation:

$$\alpha^{\|}{}_{ij}(S) \in \alpha(S,\|) = \left\{ \begin{array}{c} 1 \\ 0 \end{array} \right\} \text{ if } \alpha^{-}{}_{ij}(S) \in \alpha(S,-) \left\{ \begin{array}{c} > 0 \\ = 0 \end{array} \right\}. \tag{3}$$

For the following, the further definition is required:

DEFINITION 18

If every element q of a set Q is incomparable with all the elements r of a set R, then Q and R are called **incomparable sets**[3] $Q\|R$:

Q$\|R$ if $\forall (q \in Q, r \in R) : q\|r$. Here, $Q\|R = R\|Q$ holds.

6. DETERMINING THE CARDINALITY OF PAIRS OF INCOMPARABLE SETS OF A POSET

The objective of the task given by Rainer Brüggemann is to find all pairs of incomparable sets for any prescribed partially ordered set $(G, \leq)$ that is, in practice, generally derived from a data matrix

$$X(G,M) = \begin{pmatrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1|M|} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i1} & \cdots & x_{ij} & \cdots & x_{i|M|} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{|G|1} & \cdots & x_{|G|j} & \cdots & x_{|G|,|M|} \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_{|G|} \end{pmatrix}$$

[3] In Definition 18 and in the following we prefer the more appropriate term ''incomparable sets'' instead of ''separated subsets''. The latter is used in connection with Hasse diagram technique.

with real elements x_{ij} of $|G|$ objects G (rows) and $|M|$ attributes M (columns). It is supposed that there are no two $\mathbf{x}_i$, $\mathbf{x}_j$ in $\mathbf{X}(G,M)$ with $\mathbf{x}_i = \mathbf{x}_j$ if $i \neq j$. Therefore, in the set G each object is different to each other. In general, however, it is possible that an object is a representative of a class of equivalent elements.

Here, the following is true for the poset $(G, \leq)$ based on the objects:

- $(g_h, g_k) \in (G, \leq)$ if $x_h(\rightarrow g_h \in G) \leq x_k(\rightarrow g_k \in G)$, i.e., $\forall(x_{hj} \in x_h, x_{kj} \in x_k) : x_{hj} \leq x_{kj}$. In this case the objects g_h and g_k are comparable: g_h — g_k. Therefore, the element $\zeta_{hk}^{(G)}$ of the ζ-matrix $\zeta(G)$ assigned to the relation graph $\Gamma_d^s(G, \leq)$ is $\zeta_{hk}^{(G)} = 1$. (4a)
- $(g_h, g_k) \notin (G, \leq)$ if $\neg\forall(x_{hj} \in x_h, x_{kj} \in x_k) : (x_{hj} \leq x_{kj}) \vee (x_{hj} \geq x_{kj})$. g_h and g_k are now incomparable: $g_h // g_k$, and $\zeta_{hk}^{(G)} = 0$ holds. (4b)

In this way, from the data matrix $\mathbf{X}(G,M)$ the $|G|^2$-ζ-Matrix $\zeta(G)$ corresponding to the relation graph $\Gamma_d^s(G, \leq)$ was determined. According to equation (2) the adjacency matrix $(G, -)$ to the graph $\Gamma_d^s(G, -)$ indicating the comparability of the objects can then be calculated. By means of the transformation (3) this results in the adjacency matrix $\alpha(G,||)$ of the graph $\bar{\Gamma}_d^s(G, -) = \Gamma_d^s(G,||)$, from which the incomparability of pairs of objects can be read off.

The matrix $\alpha(G,||)$ corresponds to the context $C = (G, G_M, ||)$ if the following replacements are made in its tabular form:

$\rightarrow$ if $\alpha^{||}_{hk} \in \alpha(G, ||) = 1 \rightarrow c_{hk} \in C = \mathbf{X}$; (5a)

$\rightarrow$ if $\alpha^{||}_{hk} \in \alpha(G, ||) = 0 \rightarrow c_{hk} \in C =$ 'blank' (or to improve readability: **.**). (5b)

Using the context $C = (G, G_M, ||)$ (the meaning of G_M, see note 1) the set $B(C)$ (see Definition 14) of all (formal) concepts can now be determined with the help of FCA. To do so, the procedure *ConImp* of Peter Burmeister (Burmeister, 2003) was used. This software is based on algorithms developed by Bernhard Ganter (1984; 1989).

Since in the context $C = (G, G_M, ||)$ the relation $||$ is symmetrical and the attribute set is equal to the set of objects, the set C even is symmetrical. Therefore, in the set of concepts C to each of its concepts $(G_h, G_k) \in B(C)$ with $G_h, G_k \subseteq G$ exists just another one (G_k, G_h).

According to Brüggemann's task, the two sets that appear in the concept (G_h, G_k) contain one of the required (maximum) set of incomparable pairs: $G_h||G_k$. Because of the symmetry property $G_h||G_k = G_k||G_h$, only half of the concepts contained in the set $B(C)$ is needed to solve Brüggemann's task.

Furthermore, since the existing concepts $(G, \emptyset)$ and $(\emptyset, G_M)$ in $B(C)$ are irrelevant in the sense of Brüggemann's task, the total number of the requested pairs of incomparable sets of objects $\{G_h, G_k\} \in B((G, \leq))$ is $|B|-1$:

$(G_h, G_k) \in B(C) \rightarrow \{G_h, G_k\} \in B((G, \leq))$ with $G_h||G_k$, $h < k$, $G_h, G_k \subset G$, $G_h \neq \emptyset$, $G_k \neq \emptyset$, (6)

i.e., $|B((G, \leq))| = |B|-1$.

In the case of a given data matrix **X**(G,M), for the determination of the incomparable sets of objects, we are looking for, the following scheme is:

(S_0) *Starting point:* data matrix **X**(G,M);
(S_1) Determination of the poset $(G, \leq)$ and of the $\zeta(G)$-matrix according to (4a,b), respectively;
(S_2) Calculation of the adjacency matrix $\boldsymbol{\alpha}(G, -)$ according to (2);
(S_3) Calculation of the adjacency matrix $\boldsymbol{\alpha}(G, \|)$ according to (3);
(S_4) Convert the adjacency matrix $\boldsymbol{\alpha}(G, \|)$ into the context $C = (G, G_M, \|)$ according to (5a,b);
(S_5) Calculation of the set of concepts $B(C)$ with respect to the context $C = (G, G_M, \|)$ by using FCA;
(S_6) *Result*: altogether $\frac{1}{2}|B|-1$ pairs of incomparable sets of $B((G, \leq))$ according to (6).

The task was to look for incomparable sets of couples in $(G, \leq)$. The solution of Brüggemann's task made more illustrative without a factual information loss if the graph $\Gamma_d^s(G, \leq) = (G, E_\leq, \varphi_\leq)$ is not connected, that is, if the graph decomposes into n $(n \geq 2)$ components (classes):

$$(G, E_\leq, \varphi_\leq) = (\{G_1, \ldots, G_n\}, \{E_{\leq,1}, \ldots, E_{\leq,n}\}, \{\varphi_{\leq,1}, \ldots, \varphi_{\leq,n}\}) = \\ = \bigcup_{i=1}^{n} (G_i, E_{\leq,i}, \varphi_{\leq,i}) = \bigcup_{i=1}^{n} \Gamma_i. \quad (7)$$

Since each vertex of a component Γ_i is not incident to any vertex of a different component Γ_j $(i \neq j)$, no object of the class G_i is comparable to any object of a disjoint class G_j:

$$\forall (g_h \in G_i,\ g_k \in G_j) : g_h \| g_k \text{ with } G_i \cap G_j = \emptyset.$$

In this case the whole problem can be reduced to a separate treatment of each of the components. The matrix $\zeta(G)$ is now a block matrix

$$\zeta(G) = \begin{pmatrix} \zeta_1(G_1) & 0 & 0 & 0 \\ 0 & \zeta_2(G_2) & 0 & 0 \\ & & \ddots & \\ 0 & 0 & & 0 \\ 0 & 0 & 0 & \zeta_n(G_n) \end{pmatrix}, \quad (8)$$

so that in the above scheme, starting from $\mathbf{S_1}$ the steps $\mathbf{S_2}$ to $\mathbf{S_6}$ can be run separately for each i = 1, 2, ..., n, according to matrix $\zeta_i(G_i)$.

If there are point graphs $\Gamma_j^{(1,0)} = (\{g_j\},\emptyset,\emptyset)$ then for the decomposition (7) it is advisable to summarize their associated objects $g_j \in G$ in a particular class $I(G)$ and to continue further recalculating without them. Since for each element $g_i \in I(G)$ holds that it is incomparable with any other object of G $(g_j \| (G \backslash g_j))$, this approach is justified. In this way, without any actual loss of information, the clarity of the result also is improved when working with an otherwise not further divided poset $(G \backslash I(G), \leq)$.

7. AN EXAMPLE FOR DEMONSTRATION

A data table Table 1 in Carlsen and Brüggemann (2011) was used as a small demonstration example that is already covered in detail in this issue. More information on this data matrix can be found in the article by Carlsen and Brüggemann (2011).

(S_0) The chemical class B (polychlorinated biphenyls) were not taken into account in this paper, so now the investigated 10×3 data matrix can be derived from Table 1 below.

TABLE 1. - *Data to be examined for the sample data matrix from Carlsen and Brüggemann (2011)*

ID	Name	Volat	Sedim	Advec
A_ph	Phenanthrene	3	2	4
A_py	Pyrene	3	3	4
A_fl	Fluoranthene	2	3	4
C_ch	Chloroform	4	1	2
C_tt	Tetrachlormethane	4	1	3
C_tr	Trichlorethene	4	2	2
C_pe	Tetrachlorethene ("Per")	3	2	3
D_at	Atrazine	1	2	4
D_nt	Nitrilotriacetic acid	1	1	1
D_ed	EDTA	1	1	3

The set of objects G_{chem} to be examined is:

$$G_{\text{chem}} = \{\text{A\_ph, A\_py, A\_fl, C\_ch, C\_tt, C\_tr, C\_pe, D\_at, D\_nt, D\_ed}\}.$$

(S_1) The ζ-matrix $\zeta(G_{\text{chem}})$:

ζ	A_ph	A_py	A_fl	C_ch	C_tt	C_tr	C_pe	D_at	D_nt	D_ed
A_ph	1	1	0	0	0	0	0	0	0	0
A_py	0	1	0	0	0	0	0	0	0	0
A_fl	0	1	1	0	0	0	0	0	0	0
C_ch	0	0	0	1	1	1	0	0	0	0
C_tt	0	0	0	0	1	0	0	0	0	0
C_tr	0	0	0	0	0	1	0	0	0	0
C_pe	1	1	0	0	0	0	1	0	0	0
D_at	1	1	1	0	0	0	0	1	0	0
D_nt	1	1	1	1	1	1	1	1	1	1
D_ed	1	1	1	0	1	0	1	1	0	1

Obviously, the object D_nt is the only minimal element of the poset (G_{chem}, $\leq$). Therefore, the graph $\Gamma_d^s(G_{chem}, \leq) = (G_{chem}, E_{\leq}, \varphi_{\leq})$ must be connected.

(S$_2$) The adjacency matrix $(G_{chem}, -)$= $(G_{chem}, \leq)$ + $\zeta^{\mathsf{T}(G_{chem}, \leq)}$:

	A_ph	A_py	A_fl	C_ch	C_tt	C_tr	C_pe	D_at	D_nt	D_ed
A_ph	2	1	0	0	0	0	1	1	1	1
A_py	1	2	1	0	0	0	1	1	1	1
A_fl	0	1	2	0	0	0	0	1	1	1
C_ch	0	0	0	2	1	1	0	0	1	0
C_tt	0	0	0	1	2	0	0	0	1	1
C_tr	0	0	0	1	0	2	0	0	1	0
C_pe	1	1	0	0	0	0	2	0	1	1
D_at	1	1	1	0	0	0	0	2	1	1
D_nt	1	1	1	1	1	1	1	1	2	1
D_ed	1	1	1	0	1	0	1	1	1	2

(S$_3$), **(S$_4$)** The adjacency matrix α(G_{chem}, ||) or the context C = (G_{chem}, $G_{\mathrm{M,chem}}$, ||), respectively

	A_ph	A_py	A_fl	C_ch	C_tt	C_tr	C_pe	D_at	D_nt	D_ed
A_ph	.	.	**X**	**X**	**X**	**X**	.	.	.	.
A_py	.	.	.	**X**	**X**	**X**	.	.	.	.
A_fl	**X**	.	.	**X**	**X**	**X**	**X**	.	.	.
C_ch	**X**	**X**	**X**	.	.	.	**X**	**X**	.	**X**
C_tt	**X**	**X**	**X**	.	.	**X**	**X**	**X**	.	.
C_tr	**X**	**X**	**X**	.	**X**	.	**X**	**X**	.	**X**
C_pe	.	.	**X**	**X**	**X**	**X**	.	**X**	.	.
D_at	.	.	.	**X**	**X**	**X**	**X**	.	.	.
D_nt	.	.	.	.	.	.	.	.	.	.
D_ed	.	.	.	**X**	.	**X**	.	.	.	.

(For the representation of the matrix α(G_{chem}, ||) change **X** $\rightarrow$ 1 and **.** $\rightarrow$ 0 in the context C = (G_{chem}, $G_{\mathrm{M;chem}}$, ||) shown here.)

A – ph	A – py	A – fl	C – ch	C – tt	C – tr	C – pe	D – at	D – nt	D – ed		A – ph	A – py	A – fl	C – ch	C – tt	C – tr	C – pe	D – at	D – nt	D – ed
X	X	X	X	X	X	X	X	X	X		.	.	.	.	.	.	.	.	.	.
.	.	X	X	X	X	.	X	.	.	ǁ	.	.	.	.	.	.	X	.	.	.
X	X	X	.	X	.	X	X	.	X	ǁ	.	.	.	.	.	X	.	.	.	.
.	.	X	.	X	.	.	X	.	.	ǁ	.	.	.	.	.	X	X	.	.	.
X	X	X	.	.	X	X	X	.	.	ǁ	.	.	.	.	X	.	.	.	.	.
.	.	X	.	.	X	.	X	.	.	ǁ	.	.	.	.	X	.	X	.	.	.
X	X	X	.	.	.	X	X	.	X	ǁ	.	.	.	X	.	X	.	.	.	.
X	X	X	.	.	.	X	X	.	.	ǁ	.	.	.	X	X	X	.	.	.	.
.	.	X	.	.	.	.	X	.	.	ǁ	.	.	.	X	X	X	X	.	.	.
X	.	.	X	X	X	X	.	.	.	ǁ	.	.	X	.	.	.	.	.	.	.
X	.	.	.	X	.	X	.	.	.	ǁ	.	.	X	.	.	X	.	.	.	.
X	.	.	.	.	X	X	.	.	.	ǁ	.	.	X	.	X	.	.	.	.	.
X	.	.	.	.	.	X	.	.	.	ǁ	.	.	X	X	X	X	.	.	.	.
.	.	.	.	.	.	X	.	.	.		.	.	X	X	X	X	.	X	.	.
.	.	.	.	.	X	.	.	.	.		X	X	X	.	X	.	X	X	.	X
.	.	.	.	.	X	X	.	.	.		.	.	X	.	X	.	.	X	.	.
.	.	.	.	X	.	.	.	.	.		X	X	X	.	.	X	X	X	.	.
.	.	.	.	X	.	X	.	.	.		.	.	X	.	.	X	.	X	.	.
.	.	.	X	.	X	.	.	.	.		X	X	X	.	.	.	X	X	.	X
.	.	.	X	X	X	.	.	.	.		X	X	X	.	.	.	X	X	.	.
.	.	.	X	X	X	X	.	.	.		.	.	X	.	.	.	.	X	.	.
.	.	X	.	.	.	.	.	.	.		X	.	.	X	X	X	X	.	.	.
.	.	X	.	.	X	.	.	.	.		X	.	.	.	X	.	X	.	.	.
.	.	X	.	X	.	.	.	.	.		X	.	.	.	.	X	X	.	.	.
.	.	X	X	X	X	.	.	.	.		X	.	.	.	.	.	X	.	.	.
.	.	.	.	.	.	.	.	.	.		X	X	X	X	X	X	X	X	X	X

(S_5), (S_6) The set of concepts $B(C)$ of the context C = (G_{chem}, $G_{\text{M,chem}}$, ǁ) and the set of pairs of incomparable sets $U((G_{\text{chem}}, \leq)) \subset B(C)$.

The number of concepts is $|B| = 26$. The twelve elements of the set $U((G_{\text{chem}}, \leq))$ are edged with a double line in the following table representation of $B(C)$.

8. CONCLUSIONS

The present paper has demonstrated that FCA provides a convenient method for solving Brüggemann's task. In the future, the influence of existing components in the graph of a given poset should be investigated in detail.

REFERENCES

Bartel H.-G., Brüggemann R. (1998). Application of formal concept analysis to structure-activity relationships. *Fresenius Journal of Analytical Chemistry*, **361**, 23-28.

Bartel H.-G., Nofz M. (1997). Exploration of NMR data of glasses by means of formal concept analysis. *Chemometrics and intelligent laboratory systems*, **36**, 53-63.

Birkhoff G. (1967). *Lattice Theory*. Volume XXV of the American Mathematical Society Colloquium Publications. Rhode Island.

Bondy J.A., Murty U.S.R. (2008). *Graph Theory*. Springer, New York.

Brüggemann R., Halfon E. (1995). *Theoretical Base of the Program "Hasse"*. GSF-Bericht 20/95, München-Neuherberg.

Burmeister P. (2003). *Formal concept analysis with ConImp: introduction to the basic features*. Self-publishing, Darmstadt.

Carlsen L., Brüggemann R. (2011). Risk assessment of chemicals in the river Main (Germany): application of selected partial order ranking tools. *Statistica & Applicazioni*, this issue.

Chartrand G., Zhang P. (2005). *Introduction to Graph Theory.* McGraw-Hill Education, Boston.

Ganter B. (1984). *Two basic algorithms in concept analysis.* Technical Report FB4-Preprint No. 831, TH Darmstadt.

Ganter B. (2010). Two Basic Algorithms in Concept Analysis. *Proceedings of the 8^{th} International Conference on Formal Concept Analysis* 2010 (Lecture Notes in Computer Science, Volume 5986/2010). Springer, Berlin, 312-340.

Ganter B., Stumme G., Wille R. (Eds.) (2005). Formal Concept Analysis: Foundations and Applications. (Lecture Notes in Artificial Intelligence, no. 3626). Springer, Berlin.

Ganter B., Wille R. (1998). *Formal Concept Analysis: Mathematical Foundations.* Springer, Berlin.

Halfon E., Reggiani M.G. (1986). On Ranking Chemicals for Environmental Hazard. *Enviromental Science & Technology*, **20**, 1173-1179.

Sachs H. (1970). *Einführung in die Theorie der endlichen Graphen, Teil 1.* Teubner, Leipzig.

Birkhoff G. (1967). *Lattice Theory*. Volume XXV of the American Mathematical Society Colloquium Publications. Rhode Island.