Journal of Mathematical Sciences, Vol. 80, No. 2, 1996

MATHEMATICAL ASPECTS OF CONCEPT ANALYSIS

S. O. Kuznetsov

UDC 519.716

1. Introduction

Various applied studies use binary matrices for representing objects from a domain. A row of such a matrix is interpreted as an object and a column is interpreted as a binary attribute. An object possesses an attribute if the corresponding element is 1 and does not possess it if the corresponding matrix element is 0. The more general case, where attributes have more than two values, can be reduced to the binary one. The following problem arises often for these "object - attribute" matrices: given the set of objects B, determine the set of all attributes A that hold for all objects from B and furthermore, determine the set C of all objects that possess the whole set of attributes A. In terms of binary matrices this means that the maximal identity submatrix of the data matrix is sought, i.e., an identity submatrix such that no supermatrix of it is an identity one. This submatrix may be associated with a concept, where the corresponding set of objects is the extent and the set of attributes is the intent of the concept. This model is in accordance with the traditional understanding of the notion of concept, which dates back at least to the *Logique de Port-Royal* of the XVII century.

To introduce more precise definitions we will use Galois connections [8]. In accordance with the "formal concept analysis," we use notations from [80]. To this end, we denote by G the set of objects (from Gegenstand, object (German)) and by M the set of attributes (from Merkmahl, attribute (German)). By I we denote a relation defined on $G \times M$: for $g \in G$, $m \in M$, gIm holds iff the object g possesses the property m (i.e., the element of the corresponding matrix which is in the row g and the column m is a unit one). In accordance with [80], the triple $K = (G, M, I)$ is called a *context*.

Definition 1.1. Let $K = (G, M, I)$ be a context and $A \subseteq G$, $B \subseteq M$ be arbitrary subsets. Then the *Galois connections* $s : G \mapsto M$ and $t : M \mapsto G$ are given in the following way:

$$s(A) = \{m \in M | gIm \text{ for all } g \in A\},$$

$$t(B) = \{g \in G | gIM \text{ for all } m \in B\}.$$

Following [80], we will also write A^I and B^I instead of $s(A)$ and $t(B)$ or just A' and B' when the relation I is fixed. When it does not lead to confusion, we will also use the notation A'' as an abbreviation for both $t(s(A))$ and $(s(t(A))$ (depending on whether $A \subseteq M$ or $A \subseteq G$).

The mappings $A \mapsto A'$ and $B \mapsto B'$, which define Galois connections over the sets $\mathcal{P}(G)$ and $\mathcal{P}(M)$ ($\mathcal{P}(X)$ denotes the power set of X), possess the following properties (see [8], for example):

(1) $A_1 \subseteq A_2$ implies that $A_1' \supseteq A_2'$ for arbitrary $A_1, A_2 \subseteq G$,

(1′) $B_1 \subseteq B_2$ implies that $B_1' \supseteq B_2'$ for arbitrary $B_1, B_2 \subseteq M$,

(2) $A \subseteq A''$ and $A' = A'''$ hold for arbitrary $A \subseteq G$,

(2′) $B \subseteq B''$ and $B' = B'''$ hold for arbitrary $B \subseteq M$.

It is easy to see that the operation $''$ (the double application of the operation $'$, i.e., $s \circ t$ or $t \circ s$) is a closure operation, since the following properties hold for all $X, Y \subseteq G$ or $X, Y \subseteq M$:

- extension: $X \subseteq X''$,
- idempotency: $X'' = X''''$ and
- isotonicity: $X \subseteq Y$ implies $X'' \subseteq Y''$.

Assume that D is a binary matrix that corresponds to the context $K = (G, M, I)$. It is easy to show that an identity submatrix of D, maximal by inclusion, corresponds to a pair (A, B), where $A \subseteq G$, $B \subseteq M$, $A' = B$, $B' = A$. In the sequel, we will speak about pairs of this form in terms of the *formal concept analysis* [80].

Translated from Itogi Nauki i Tekhniki, Seriya Sovremennaya Matematika i Ee Prilozheniya. Tematicheskie Obzory. Vol. 18, Vychislitel'naya Matematika i Kibernetika, 1994

1072-3374/96/8002-1684$15.00

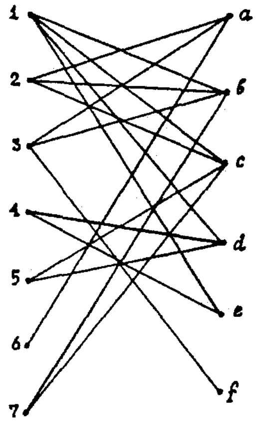


Fig. 1. Bipartite graph corresponding to the context given in Table 1.

Definition 1.2 ([80]). Let $K = (G, M, I)$ be a context. A pair (A, B) is a *concept* of the context K iff $A \subseteq G$, $B \subseteq M$ and $A' = B$, $B' = A$. A and B are called the *extent* and the *intent* of the concept (A, B), respectively.

Contexts and their concepts can also be easily described in graph-theoretic language. Let $K = (G, M, I)$ be a context. Consider a bipartite graph $Z = (V_1 \cup V_2, E)$, $E \subseteq V_1 \times V_2$. The vertices of the first part are in one-to-one correspondence with objects from G and the vertices of the second part are in one-to-one correspondence with attributes from M. For arbitrary vertices $v_i \in V_1$ and $v_j \in V_2$, $(v_i, v_j) \in E$ iff the object from G that corresponds to the vertex v_i possesses the attribute from M that corresponds to the vertex v_j. The concepts of the context K correspond to complete bipartite subgraphs of the graph Z maximal with respect to inclusion, i.e., to graphs of the form $(W_1 \cup W_2, W_1 \times W_2)$, where all vertices of such a subgraph that belong to a common part are adjacent to all vertices of the other part, and no supergraph of this subgraph is a complete bipartite subgraph of the graph Z.

Example. Consider the context (G, M, I) represented by Table 1 taken from [9].

Table 1

$G \backslash M$	1	2	3	4	5	6	7
a	0	1	1	0	0	1	0
b	1	1	1	0	0	0	1
c	1	1	0	0	1	0	1
d	1	0	0	1	1	0	0
e	1	0	0	1	0	0	0
f	0	0	1	0	0	0	0

This context can be represented by the bipartite graph shown in Fig. 1. By way of example, the complete bipartite subgraphs $(\{a, b\} \cup \{2, 3\}, \{a, b\} \times \{2, 3\})$, $(\{d, e\} \cup \{1, 4\}, \{d, e\} \times \{1, 4\})$ of this graph correspond to the concepts $(\{a, b\}, \{2, 3\})$, $(\{d, e\}, \{1, 4\})$ of the context (G, M, I).

In this survey we consider the papers related to mathematical aspects of concept analysis. On the one hand, these are algebraic problems that arise from the lattice nature of the set of all concepts (Sec. 2). Here, the decomposition of concept lattices into smaller ones seems to be one of the most important problems. On the other hand, we will dwell on applications of concepts in data analysis, in particular, on methods of search for dependences between attributes (Sec. 3). Third, we will consider results related to algorithmic problems of concept generation (Sec. 4). And, finally, we will consider problems of different nature related to concepts, including the Zarankiewicz problem, where conditions for a concept of a given size are to be found (Sec. 5).

We assume that the reader is familiar with some elementary notions of the theory of sets, the lattice theory [8, 35], and the theory of computational complexity [1, 34].

2. Algebraic Aspects of Concept Analysis

2.1. The Main Theorem of Formal Concept Analysis

The concepts of a context (G, M, I) are partially ordered in the following way:

$$(A_1, B_1) \leq (A_2, B_2) \rightleftharpoons A_1 \subseteq A_2 \ (B_1 \supseteq B_2).$$

The pair (A_1, B_1) is called a subconcept of the concept (A_2, B_2) and (A_2, B_2) is called a superconcept of the concept (A_1, B_1).

Following [80], by $L(G, M, I)$ we denote the set of all concepts of the context (G, M, I) and by $\mathfrak{B}(G, M, I)$ the partially ordered set $(L(G, M, I), \leq)$ (from Begriff, concept (German)).

Definition 2.1.1. A set D of a lattice L is called *infimum-dense* (*supremum-dense*) if

$$L = \{\bigwedge_{x \in X} x | X \subseteq D\} \ (= \{\bigvee_{x \in X} x | X \subseteq D\}, \text{ respectively}).$$

Haralick [42] found that the set of concepts is closed with respect to certain operations, which define idempotent, commutative, and associative operations (i.e., form semilattices or "idempotent commutative monoids," as Haralick [42] called them). In [80], a more general result was proved.

Theorem 2.1.1 ([80]). *Let (G, M, I) be a context. Then $\mathfrak{B}(G, M, I)$ is a complete lattice called the concept lattice of (G, M, I); its infima and suprema can be described as follows:*

$$\bigwedge_{j \in J}(A_j, B_j) = (\bigcap_{j \in J} A_j, (\bigcap_{j \in J} A_j)''),$$

$$\bigvee_{j \in J}(A_j, B_j) = ((\bigcap_{j \in J} B_j)', \bigcap_{j \in J} B_j).$$

Conversely, if L is a complete lattice, then $L \cong \mathfrak{B}(G, M, I)$ iff there are mappings $\gamma : G \mapsto L$ and $\mu : M \mapsto L$ such that γG is supremum-dense in L, μM is infimum-dense in L, and gIm is equivalent to $\gamma g \leq \mu m$ for all $g \in G$ and $m \in M$, in particular, $L \cong \mathfrak{B}(L, L, \leq)$.

In some papers, for example, in [7], the lattice $\mathfrak{B}(G, M, I)$ is called a "Galois lattice."

Example. The Hasse diagram of the concept lattice, which corresponds to the context (G, M, I) from Sec. 1, is given in Fig. 2. Recall that in the Hasse diagram of a lattice, the vertices correspond to the elements of the lattice, the bottom-up direction corresponds to the order relation of the lattice, and two vertices are joined by an edge if the higher of these two vertices corresponds to an element of the lattice which is immediately superior to the element that corresponds to the other vertex. The diagram illustrates the well-known reciprocal relation between the intent and the extent of concepts (the greater the intent, the less the extent).

Let $J(L)$ be the set of all $\vee$-irreducible elements of the lattice L and $M(L)$ be the set of all $\wedge$-irreducible elements of L, i.e., elements that cannot be represented as suprema (respectively, infima) of some other elements of the lattice. Then, by Theorem 2.1.1., $L \cong \mathfrak{B}(J(L), M(L), \leq)$. The context $(J(L), M(L), \leq)$ is the least context with the property that its concept lattice is isomorphic to L.

The Dedekind–MacNeille completion of a partial order to a lattice (see, e.g., [8, 77]) can be easily described in terms of concept lattices. By way of example, consider a set G of graphs with labeled vertices. The graphs from G are ordered with respect to the subgraph isomorphism relation $\leq$: (for F, $G \in$ G, $F \leq G$ iff F is isomorphic to a subgraph of G). Infimum and supremum operations corresponding to the partial order $\leq$ cannot be defined for graphs. At the same time, these operations can be defined on the Dedekind–MacNeille completion of the partially ordered set (G, $\leq$). The elements of this completion are sets of pairwise incomparable (in the sense of $\leq$) graphs [50]. The mapping of the partially ordered set

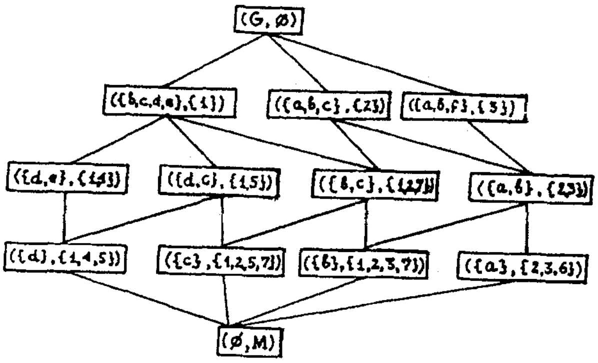


Fig. 2. Hasse diagram for the concept lattice of the context (G, M, I) from Sec. 1.

$(P, \leq)$ into its Dedekind–MacNeille completion $\mathfrak{B}(P, P, \leq)$ is given by the function $\iota: \iota x = ((x], [x))$, where $x \in P$, $(x] = \{y | y \leq x\}, [x) = \{y | y \geq x\}$.

Theorem 2.1.2 ([80]). *The mapping $\iota x = ((x], [x))$ is an order-preserving one-to-one mapping of $(P, \leq)$ into $\mathfrak{B}(P, P, \leq)$. If for $X \subseteq P$ infimum $\wedge X$ and supremum $\vee P$ can be defined in $(P, \leq)$, then $\iota(\wedge X) = \wedge(\iota X)$ and $\iota(\vee X) = \vee(\iota X)$. If λ is a one-to-one order-preserving mapping of $(P, \leq)$ in a complete lattice L, then there exists an order-preserving mapping κ of the lattice $\mathfrak{B}(P, P, \leq)$ into L such that $\lambda = \kappa \circ \iota$.*

2.2. Many-Valued Contexts

The situation of non-binary attributes is frequently encountered in many applied problems of data analysis. Many-valued contexts were proposed in [80] to represent situations of this kind.

Definition 2.2.1 ([80]). A *many-valued context* is a quadruple (G, M, W, I) such that G, M, and W are sets and I is a binary relation between G and $M \times W$ $(I \subseteq G \times M \times W)$, where $gI(m, w_1)$ and $gI(m, w_2)$ imply $w_1 = w_2$ for arbitrary $g \in G$, $m \in M$, $w_1, w_2 \subseteq W$. If $gI(m, w)$ for $g \in G, m \in M, w \in W$, then the object g is said to take the value w for the attribute m. If $|W| = n$, then (G, M, W, I) is called an *n-valued context.*

In the case where the attribute values are understood as nominal data, a many-valued context (G, M, W, I) can be represented by the binary (or "one-valued") context $(G, M \times W, I)$. In this case, the context (G, M, W, I) is called *nominal.* The lattice $\mathfrak{B}(G, M \times W, I)$ is called the *concept lattice* of (G, M, W, I). A context (G, M, W, I) is said to be complete if for any $g \in G, m \in M$ there exists $w \in W$ such that $gI(m, w)$. For a characterization of the concept lattice of a complete n-valued nominal context the following definition is used.

Definition 2.2.2 ([80]). An element d of a complete lattice L has *valence n* if n is the smallest cardinality of a subset D of $L \backslash \{0\}$ containing d, which is maximal with respect to the property that $x \wedge y = 0$ for all $x, y \in D$ such that $x \neq y$. Recall that a lattice is *atomistic* if each lattice element is either 0, or an atom, or a supremum of some atoms.

Theorem 2.2.1 ([80]). *A complete lattice L is isomorphic to a concept lattice of a complete n-valued nominal context iff L is atomistic and has an infimum-dense subset of elements of valence $\leq n$.*

Corollary 2.2.2 ([80]). *A finite lattice L is isomorphic to a concept lattice of a finite complete n-valued nominal context iff L is atomistic and every $\wedge$-irreducible element of L has valence $\leq n$.*

Corollary 2.2.3 ([80]). *A finite lattice L is isomorphic to a concept lattice of a finite complete 2-valued nominal context iff L is atomistic and every $\wedge$-irreducible element of L has a pseudocomplement.*

2.3. Fusion of Contexts

In a series of papers on formal concept analysis, the problems arising from "fusion" of several contexts are studied. From the data analysis standpoint, this problem is interesting in view of data aggregation. In the closely related papers [81, 83] Wille considers the subdirect products studied earlier by him for general complete lattices in [78, 79]. In particular, a relation between the subdirect products of concept lattices and closed relations defined on the sums of the corresponding contexts is established in [84] (see Sec. 2.4). Results from [80] concerning the decomposition of lattices into direct products are found in Secs. 2.5 and 2.6. Before introducing the notion of fusion of contexts, we present some auxiliary definitions and results.

Definition 2.3.1. Let $L_1 = (T_1, \wedge_1, \vee_1)$, $L_2 = (T_2, \wedge_2, \vee_2)$ be lattices on sets T_1, T_2. Let f be a mapping $f: T_1 \mapsto T_2$, a_1, b_1 be any elements of T_1, and $\leq_1$, $\leq_2$ be order relations induced by $\wedge_1$, $\vee_1$, and $\wedge_2$, $\vee_2$. If $a_1 \leq_1 b_1 \to f(a_1) \leq_2 f(b_2)$, then f is called an *order-preserving* mapping from L_1 to L_2. If f is one-to-one, then it is called an *order-preserving homomorphism.* If $f(a_1 \vee_1 b_1) = f(a_1) \vee_1 f(b_1)$, then f is called $\vee$-*homomorphism* from L_1 into L_2. If f is one-to-one, then it is called a $\vee$-*embedding* of L_1 into L_2 (analogously for a $\wedge$-*homomorphism* and a $\wedge$-*embedding*). A mapping from L_1 into L_2 that is a $\wedge$- and $\vee$-homomorphism is called a homomorphism of L_1 into L_2. A mapping from L_1 into L_2 that is a $\vee$- and $\wedge$- embedding is called an *embedding* of L_1 into L_2. An embedding of L_1 into L_2 is an isomorphism if it is a mapping of L_1 onto L_2.

Definition 2.3.2 ([6]). A lattice L is called a *subdirect product* of a family of lattices $(L_s)_{s\in S}$ if there exists an embedding $f: L \mapsto \times_{s\in S} L_s$ such that for each $s \in S$, the mapping $p_s \circ f: L \mapsto L_s$ is onto (where p_s is the projection of $\times_{s\in S} L_s$ onto L_s).

Definition 2.3.3 ([81]). Let $K = (G, M, I)$ be a context. Relation $J \subseteq G \times M$ is called a *closed relation* of the context K if every concept of (G, M, J) is a concept of K.

Let $\mathfrak{S}$ be a complete sublattice of $\mathfrak{B}(G, M, I)$. Then $C(\mathfrak{S}) = \bigcup\limits_{(A,B)\in\mathfrak{S}} A \times B$.

Theorem 2.3.1 ([81]). *C is a bijection from the set of all complete sublattices of $\mathfrak{B}(G, M, I)$ onto the set of all closed relations of (G, M, I), in particular, $C^{-1}(J) = \mathfrak{B}(G, M, J)$ for each closed relation J of (G, M, I).*

The following characterization of closed relations of a context was proposed by B. Ganter.

Theorem 2.3.2 ([81]). *J is a closed relation of a context (G, M, I) iff $J \subseteq I$ and satisfies the following property: $(g, m) \in I \backslash J$ implies that there exists $h \in G$ such that $\{g\}^J \subseteq \{h\}^J$ and $(h, m) \notin I$ and there exist $n \in M$ such that $\{m\}^J \subseteq \{n\}^J$ and $(g, n) \notin J$.*

Definition 2.3.4 *The disjunctive union* of sets $X_s, s \in S$, denoted by $X_1 \dot\cup \ldots \dot\cup X_{|S|}$ is the set $\cup_{s\in S} X_s \times \{s\}$.

Thus, the disjunctive union retains all exemplars of the element $x \in X_1, \ldots, x \in X_t$ by using indices of the sets $X_s, s \in S$.

Definition 2.3.5 ([81]). The *sum of a family of contexts* $(G_t, M_t, I_t)_{t\in T}$ is defined by

$$\sum_{t\in T}(G_t, M_t, I_t) = \left(\dot\cup_{t\in T}\, G_t, \dot\cup_{t\in T}\, M_t, \dot\cup_{t\in T}\, I_t \,\dot\cup \bigcup_{s,t\in T, s\neq t} G_s \times M_t\right).$$

Definition 2.3.6 ([81]). A *bond* from a context (G_s, M_s, I_s) to a context (G_t, M_t, I_t) (denoted by J_{st}) is a set $J \subseteq G_s \times M_t$, for which $\{g\}^J$ is an intent of (G_t, M_t, I_t) and m^J is an extent of (G_s, M_s, I_s) for arbitrary $g \in G_s$ and $m \in M_t$, i.e., the extents of (G_s, M_t, J) are extents of (G_s, M_s, I_s) and the intents of (G_s, M_t, J) are intents of (G_t, M_t, I_t). In the sequel we will write X^t instead of $X^{J_{st}}$ and Y^s instead of $Y^{J_{st}}$.

Lemma 2.3.3 ([81]). *Let J_{rs} be a bond from (G_r, M_r, I_r) to (G_s, M_s, I_s) and let J_{st} be a bond from (G_s, M_s, I_s) to (G_t, M_t, I_t). Then $J_{rs} \circ J_{st} = \{(g, m) \in G_r \times M_t | \{g\}^{ss} \subseteq \{m\}^s\}$ is a bond from (G_r, M_r, I_r) to (G_t, M_t, I_t).*

Corollary. *If J_{rt} is a bond from (G_r, M_r, I_r) to (G_t, M_t, I_t), then the following conditions are equivalent:*
(1)$J_{rt} \subseteq J_{rs} \circ J_{st}$,
(2)$\{g\}^t \subseteq \{g\}^{sst}$ for all $g \in G_r$,
(3)$\{m\}^r \subseteq \{m\}^{ssr}$ for all $m \in M_t$.

Proposition 2.3.4 ([83]). *Let $K_1 = (G_1, M_1, I_1)$, $K_2 = (G_2, M_2, I_2)$ be contexts, and A, B be arbitrary sets such that $A \subseteq G_1$, $B \subseteq M_2$, and*

$$A^2 = \{m \in M_2 | gIm \text{ for all } g \in A\},$$
$$B^2 = \{g \in G_1 | gIm \text{ for all } m \in B\}.$$

Then for $I \subseteq G_1 \times M_2$ the following conditions coincide:

(1) *I is a bond from (G_1, M_1, I_1) to (G_2, M_2, I_2),*

(2) *the mapping $\sigma_I: (A, B) \mapsto (A^{22}, A^2)$ for all $(A, B) \in \mathfrak{B}(K_2)$ is a $\vee$-embedding of the lattice $\mathfrak{B}(K_1)$ into the lattice $\mathfrak{B}(K_2)$ and $\{g\}^{112} = \{g\}^2$ for every $g \in G_1$,*

(3) *the mapping $\sigma^I: (A, B) \mapsto (B^1, B^{11})$ is a $\wedge$-embedding of the lattice $\mathfrak{B}(K_2)$ into $\mathfrak{B}(K_1)$ for $(A, B) \in \mathfrak{B}(K_2)$ and $\{m\}^{221} = \{m\}^1$ for any $m \in M_2$.*

Theorem 2.3.5 ([81]). *Let $(G_t, M_t, I_t)_{t \in T}$ be a family of contexts and ι be an isomorphism from $\times_{t \in T} \mathfrak{B}(G_t, M_t, I_t)$ onto $\mathfrak{B}(\sum_{t \in T}(G_t, M_t, I_t))$ given as $\iota((A_t, B_t)|t \in T) = (\dot{\cup}_{t \in T} A_t, \dot{\cup}_{t \in T} B_t)$. Furthermore, let $J \subseteq \dot{\cup}_{t \in T} G_t \times \dot{\cup}_{t \in T} M_t$ and $J_{st} = J \cap G_s \times M_t$. Then the following conditions are equivalent:*

(1) *$\iota^{-1} C^{-1}(J)$ is a complete subdirect product of the $\mathfrak{B}(G_t, M_t, I_t)_{t \in T}$,*

(2) *J is a closed relation of $\sum_{t \in T}(G_t, M_t, I_t)$ with $J_{tt} = I_t$ for $t \in T$,*

(3) *the J_{st} are bonds from (G_s, M_s, I_s) to (G_t, M_t, I_t) with $J_{tt} = I_t$, and $J_{rt} \subseteq J_{rs} \circ J_{st}$ for $r, s, t \in T$.*

Theorem 2.3.6 ([83]). *Let $(G \times S, M \times S, \bigcup_{r,s \in S} I_{rs})$ be the fusion of contexts $K_s = (G, M, I_s)$, $s \in S$. Then for each $s \in S$, $(g, s) I_{ss} (m, s) \iff g I_s m$ and for every pair $r, s \in S$ the relation I_{rs} is the least (by inclusion) bond from the context $(G \times \{r\}, M \times \{r\}, I_{rr})$ to the context $(G \times \{s\}, M \times \{s\}, I_{ss})$ that satisfies the condition $(g, r) I (m, s)$ iff $g I_r m$ or $g I_s m$.*

Definition 2.3.7 ([83]). Let P be a partially ordered set and α be an isotone mapping from P to the complete lattice T, which preserves suprema and infima of P. If αP is a set of generators of the lattice T, then the pair (T, α) is called a *complete P-lattice.* If $P = \{1, 2, \ldots, n\}$, then the P-lattice is also called an *n-lattice.*

Definition 2.3.8 ([83]). *Let (T_1, α_1) and (T_2, α_2) be complete P-lattices. A homomorphism φ from T_1 into T_2 is a P-morphism if $\varphi\alpha_1 = \alpha_2$. The P-lattices (T_1, α_1) and (T_2, α_2) are called isomorphic if there exists a P-morphism from T_1 into T_2 that is an isomorphism.*

Definition 2.3.9 ([81]). A P-lattice (L, α) is called a *P-product* of complete P-lattices $(L_t \alpha_t)$ if $\alpha p = (\alpha_t p | t \in T) \in \times_{t \in T} L_t$ for $p \in P$ and L is a complete lattice which is a subdirect product of lattices $L_t, t \in T$, generated by αp, $p \in P$.

Definition 2.3.10 ([83]). *Let $\mathcal{L}$ be a class of complete lattices and $\mathcal{L}_p$ be a class of complete P-lattices from $\mathcal{L}$. Then $\widetilde{(T, \alpha)}$ is the class of P-lattices from $\mathcal{L}_p$ isomorphic to (T, α), $\widetilde{\mathcal{L}_p}$ is the set of all (T, α) such that $(T, \alpha) \in \mathcal{L}_p$. By definition, $\widetilde{(T_1, \alpha_1)} \geq \widetilde{(T_2, \alpha_2)}$ if there exists a P-embedding from (T_1, α_1) into (T_2, α_2).*

It can be easily seen that $\geq$ is a partial order on $\widetilde{\mathcal{L}_p}$.

Theorem 2.3.7 ([83]). *Let $\mathcal{L}$ be the class of complete lattices closed with respect to the subdirect product of factors with indices from a set I. For (T, α) and $(T_i, \alpha_i)_{i \in I}$ from $\mathcal{L}_p$ the following conditions are equivalent:*

(1) *there exists an isomorphism φ from T onto the direct product of $T_i, i \in I$, such that $\varphi\alpha p = (\alpha_i p)_{i \in I}$ for all $p \in P$,*

(2) *$\widetilde{(T, \alpha)}$ is the supremum of $\widetilde{(T_i, \alpha_i)}, i \in I$, in the partial order $(\mathcal{L}_p, \leq)$.*

This theorem characterizes the subdirect product as the "minimal fusion" of the factors, but does not give an effective way for constructing this fusion.

Definition 2.3.11 ([83]). Let T_1 and T_2 be two complete lattices and τ be a surjective $\wedge$-homomorphism from T_1 into T_2. Then $\underline{\tau}$ is a mapping from T_2 into T_1 defined as $\underline{\tau} = \wedge\tau^{-1}y$.

It is easy to show that $\underline{\tau}$ is an injective $\vee$-homomorphism from T_2 into T_1.

Let T_i, $i \in I$, be a complete lattice and τ_{ij}: $T_i \to T_j$ $(i,j \in I)$ be $\wedge$-homomorphisms satisfying the following conditions:

(1) τ_{ii} is the identity mapping,

(2) $\tau_{ki} \geq \tau_{kj}\tau_{ji}$ for an arbitrary triple $i,j,k \in I$.

Proposition 2.3.8 ([83]). *The set $G(\tau_{ji}|i,j \in I) = \{(\tau_{ji}x)_{j\in I}|x \in T_i,\ i \in I\}\backslash\{0\}$ is supremum-dense in the subdirect product of $S(\tau_{ji}|i,j \in I)$ of the complete lattices T_i, $i \in I$.*

Theorem 2.3.9 ([83]). *Let τ_i: $(T,\alpha) \mapsto (T_i,\alpha_i)$, $i \in I$ be P-morphisms of complete P-lattices. If $\widetilde{(T,\alpha)}$ is the supremum of $\widetilde{(T_i,\alpha_i)}$, then, for $i,j \in I$, $\tau_j\tau_i$ is the greatest $\vee$-homomorphism σ: $T_i \mapsto T_j$ such that $\sigma\alpha_i p \leq \alpha_j p$ for all $p \in P$ $(i,j \in I)$.*

Now, the process of generation of the subdirect product by means of supremum operation can only be represented in the following way [83]: Let $(T_i,\alpha_i)_{i\in I}$ be complete P-lattices.

1. For all $i,j \in I$ define τ_{ij} as the greatest $\wedge$-embeddings: $T_i \mapsto T_j$ such that $\sigma\alpha_i p \leq \alpha_j p$ for all $p \in P$.
2. Construct the set $G(\tau_{ji}|i,j \in I) = \{(\tau_{ji}x)_{j\in I}|x \in T_i,\ i \in I\}$.
3. Generate $T := \{\vee X|X \subseteq G(\tau_{ji}|i,j \in I\}$ and define α: $P \mapsto T$ as $\alpha p = (\alpha_i p)_{i\in I}$ for all $p \in P$.

Then (T,α) is a complete P-lattice and $\widetilde{(T,\alpha)}$ is the supremum of $\widetilde{(T_i,\alpha_i)}$, $i \in I$, i.e., T is the subdirect product of T_i, $i \in I$, generated by αP.

Let $K_s = (G,M,I_s)$, $s \in S$, be contexts. For arbitrary A,B such that $A \subseteq G, B \subseteq M$ define

$$A^s = \{m \in M|gI_s m \text{ for all } g \in A\},$$

$$B^s = \{g \in G|gI_s m \text{ for all } m \in B\}.$$

Let $P = G \cup M$ and $\alpha_s g = (\{g\}^{ss}, \{g\}^s)$, $\alpha_s m = (\{m\}^s, \{m\}^{ss})$ for $g \in G, m \in M, s \in S$. Construct the fusion of the lattices $\mathfrak{B}(K_s),\alpha)$, where $\mathfrak{B}(K_s)_{s\in S}$ are concept lattices for contexts K_s. To this end, Wille [83] uses $\vee$-homomorphisms $\sigma_{sr}: \mathfrak{B}(K_r) \mapsto \mathfrak{B}(K_s)$, $r,s \in S$, where σ_{sr} are the greatest homomorphisms from $\mathfrak{B}(K_r)$ into $\mathfrak{B}(K_s)$ such that $\sigma_{sr}\alpha_r p \leq \alpha_s p$ for all $p \in P$. For each pair $r,s \in S$ define the relation I_{rs} between $G \times \{r\}$ and $M \times \{s\}$ as $(g,r)I_{rs}(m,s) \iff \sigma_{rs}\alpha_r g < \alpha_s m$. If $r = s$, then σ_{ss} is the identity homomorphism and $(gs)I_{ss}(m,s) \iff gI_s m$.

Definition 2.3.12 ([83]). The *fusion* of contexts K_s, $s \in S$, denoted by $\odot$ is a triple $(G \times S, M \times S, \bigcup_{r,s\in S} I_{rs})$, where for all $g \in G$ the object αg is a concept of the context $\odot$ with the extent $\{(g,s)|s \in S\}''$, and for all $m \in M$ the object αm is a concept of the context $\odot$ with the intent $\{(m,s)|s \in S\}''$.

Theorem 2.3.10 ([83]). *$(\mathfrak{B}(\odot),\alpha)$ is a complete P-lattice, which is a representative of the supremum of classes $\widetilde{(\mathfrak{B}(K_s),\alpha_s)}$, $s \in S$; if τ_r is a P-morphism of the lattice $(\mathfrak{B}(\odot),\alpha)$ into $(\mathfrak{B}(K_r),\alpha_r)$ and $\tau_r(A,B) = (C,D)$, then $A \cap (G \times \{r\}) = C \times \{r\}$ and $B \cap (M \times \{r\}) = D \times \{r\}$.*

Wille [83] points out that the use of concept lattices $\mathfrak{B}(K_r)$ and $\mathfrak{B}(K_s)$ in the definition of relation I_{rs} of the fusion of contexts is rather inconvenient. In order to avoid this obstacle, the definition of a bond (Definition 2.3.6) is proposed in [83].

Theorem 2.3.11 ([81]). *Let $(\mathcal{L},\alpha)$ be a P-product of lattices $(\mathfrak{B}(G_t,M_t,I_t),\ \alpha_t)_{t\in T}$ and let ι be an embedding of $\mathfrak{B}(\sum_{t\in T}(G_t,M_t,I_t))$ given as*

$$\iota((A_t,B_t))|_{t\in T}) = (\dot{\cup}_{t\in T} A_t, \dot{\cup}_{t\in T} B_t)).$$

Then $C(\iota\mathcal{L})$ is the relation $J \subseteq (\cup_{t\in T}G_t) \times (\cup_{t\in T}M_t)$, for which $J_{tt} = I_t$ and J_{st} is the least (by inclusion) bond from (G_s,M_s,I_s) to (G_t,M_t,I_t) that includes all pairs from $A_s^p \times B_t^p$ for every pair $p \in P$ such that $\alpha_s p = (A_s^p, B_s^p)$ and $\alpha_t p = (A_t^p, B_t^p)$, where $s,t \in T$, $s \neq t$.

Definition 2.3.13 ([81]). A pair $((G,M,I),\alpha)$ is called a *P-context* if $(\mathfrak{B}(G,M,\ I),\alpha)$ is a complete P-lattice.

Theorems 2.3.6 and 2.3.11 allow one to define the P-fusion of contexts $((G_t,\ M_t,I_t),\alpha)_{t\in T}$ as a P-context $((\dot{\cup}_{t\in T} G_t, \cup_{t\in T}M_t, J)\alpha)$ satisfying the following conditions:

(1) $J_{tt} = I_t$ for $t \in T$,

(2) J_{st} is the least bond from (G_s, M_s, I_s) and (G_t, M_t, I_t) that contains $A_s^p \times B_s^p$ for all $\alpha_s p = (A_s^p, B_s^p)$ and $\alpha_t p = (A_t^p, B_t^p)$, $s, t \in T$, $s \neq t$,

(3) $\alpha p = (\dot\bigcup_{t\in T} A_t^r, \dot\bigcup_{t\in T} B_t^p)$ for all $p \in P$.

2.4. Tensor Products of Concept Lattices

In this section we consider some results concerning the tensor product of concept lattices and the corresponding product of contexts. These results are related mainly to establishing relations between concept lattices and some interesting objects, as well as to the decomposition of concept lattices into "primitives." In this section we will restrict our consideration mainly to the problems related to products of concept lattices. Results of this kind concerning general lattices are found in [84] and in the papers cited therein.

Definition 2.4.1 ([84]). Let L_1 and L_2 be complete lattices. Then the *tensor product* of lattices L_1 and L_2 is $L_1 \otimes L_2 \rightleftharpoons \mathfrak{B}(L_1 \times L_2, L_1 \times L_2, \nabla)$, where $(x_1, x_2)\nabla(y_1, y_2) \rightleftharpoons x_1 \leq y_1$ or $x_2 \leq y_2$ for (x_1, x_2), $(y_1, y_2) \in L_1 \times L_2$.

Definition 2.4.2 ([84]). *The direct product* of contexts $K_1 = (G_1, M_1, I_1)$ and $K_2 = (G_2, M_2, I_2)$ is the context $K_1 \times K_2 = (G_1 \times G_2, M_1 \times M_2, \nabla)$, where $(g_1, g_2)\nabla(m_1, m_2) \rightleftharpoons g_1 I_1 m_1$ or $g_2 I_2 m_2$ for $(g_1, g_2) \in G_1 \times G_2$, $(m_1, m_2) \in M_1 \times M_2$.

Theorem 2.4.1 ([84]). *For arbitrary contexts K_1 and K_2 $\mathfrak{B}(K_1) \otimes \mathfrak{B}(K_2) = \mathfrak{B}(K_1 \times K_2)$.*

A corollary of this theorem asserts the independence of the lattice product of a particular form of contexts, i.e., each of the contexts K_1, K_2 can be replaced by a context isomorphic to the concept lattice, e.g., the context $(J(\mathfrak{B}(K_i)), M(\mathfrak{B}(K_i)), \leq)$, where $J(\mathfrak{B}(K)), M(\mathfrak{B}(K))$ are the sets of $\wedge$-irreducible or $\vee$-irreducible elements of the concept lattice $\mathfrak{B}(K_i)$, $i \in \{1, 2\}$, respectively.

Definition 2.4.3 ([84]). Let $K = (G, M, I)$ be an arbitrary context, $m, n \in M$, $g, h \in G$, and $(k, m) \notin I$. Then $m \searrow h$ if γh is minimal in the set $\{\gamma k | k \in \{g\}''$ and $(k, m) \notin I\}$, $g \nearrow n$, iff μn is maximal in the set $\{\mu p | p \in \{m\}''$ and $(g, p) \notin I\}$.

Definition 2.4.4 ([84]). A context $K = (G, M, I)$ is called *doubly founded* if for arbitrary $m \in M$ and $g \in G$ such that $(g, m) \notin I$ there exist $h \in G$ and $n \in M$: $m \searrow h$ and $g \nearrow n$.

The relations $\searrow$ and $\nearrow$ allow one to study the concept lattice $\mathfrak{B}(G, M, I)$ in terms of the digraph $(G \mathbin{\dot\cup} M, \nearrow \cup \searrow)$. The set of vertices of this graph (denoted by C) is called closed if $g \in G$ and $g \nearrow m$ imply $m \in C$ and $m \searrow g$ and $m \in M$ imply $g \in C$. The closed subsets of the digraph $(G \mathbin{\dot\cup} M, \nearrow \cup \searrow)$ form a complete sublattice of the complete lattice of all subsets of $G \mathbin{\dot\cup} M$. To specify a correspondence between the concept lattice and the introduced digraph, we introduce the following sets for every complete congruence relation θ of the concept lattice $\mathfrak{B}(G, M, I)$:

$$G(\theta) = \{g \in G | \gamma g \text{ is the smallest element of the } \theta\text{-class}\},$$

$$M(\theta) = \{m \in M | \mu m \text{ is the greatest element of the } \theta\text{-class.}\}$$

Theorem 2.4.2 ([84]). *The mapping $\theta \mapsto G(\theta) \mathbin{\dot\cup} M(\theta)$ gives an antiisomorphism of the lattice of all complete congruence relations of the lattice $\mathfrak{B}(G, M, I)$ to the lattice of all closed subsets of the digraph $(G \mathbin{\dot\cup} M, \nearrow \cup \searrow)$.*

This result generalizes the result of [81] that establishes an isomorphism from the lattice of all complete congruence relations onto the lattice of subcontexts of a reduced context (see Section 5.3) (G, M, I) closed with respect to relations similar to $\nearrow$ and $\searrow$. The isomorphism established in Theorem 2.4.2 takes $\wedge$-irreducible elements of the lattice $\mathfrak{B}(G, M, I)$ to $\cup$-irreducible closed subsets of vertices of the graph $(G \mathbin{\dot\cup} M, \nearrow \dot\cup \searrow)$, which are the least closed subsets containing a given element g (these sets are denoted by $\langle g \rangle$). The latter facts imply

Theorem 2.4.3 ([84]). *The mapping $(A, B) \mapsto (A \cap \langle g \rangle, B \cap \langle g \rangle)_{g \in G}$ describes an isomorphism from*

	1	2	3	4
1	1	0	0	0
2	0	1	0	0
3	0	0	1	0
4	0	0	0	1

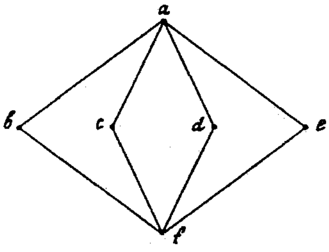


Fig. 3. The context $\mathbb{N}_4 = (\{1,2,3,4\}, \{1,2,3,4\}, =)$ and the Hasse diagram of its concept lattice.

$\mathfrak{B}(G, M, I)$ onto a subdirect product of the completely subdirectly irreducible concept lattices $\mathfrak{B}(\langle g \rangle \cap G, \langle g \rangle \cap M, I \cap \langle g \rangle^2)_{g \in G}$.

This result, as well as the following theorem, allows one to decompose complex concept lattices into products of elementary ones.

Theorem 2.4.4 ([84]). *Let $K_1 = (G_1, M_1, I_1)$, $K_2 = (G_2, M_2, I_2)$ be irreducible doubly founded contexts. Then*

$$\mathfrak{B}(\langle (g_1, g_2) \rangle \cap G_1 \times G_2,\ \langle (g_1, g_2) \rangle \cap M_1 \times M_2, \nabla \cap \langle (g_1, g_2)^2 \rangle \cong$$
$$\mathfrak{B}(\langle g_1 \rangle \cap G_1, \langle g_1 \rangle \cap M_1, I_1 \cap \langle g_1 \rangle^2) \otimes \mathfrak{B}(\langle g_2 \rangle \cap G_2, \langle g_2 \rangle \cap M_2, I_2 \cap \langle g_2 \rangle^2)$$

for any $(g_1, g_2) \in G_1 \times G_2$.

2.5. Scaling of Concept Lattices

In a series of papers [80, 31, 33] on formal concept analysis, the authors considered the idea of the conceptual measurement of concept lattices. As opposed to quantitative measurements, the qualitative ones are based on the notion of order. Therefore, the main notion of conceptual measurement is the notion of an (ordered) scale. A scale is a standard context (see examples below) with a clear concept structure. An original context is interpreted in a scale by means of some measure. In [33] the following four examples of finite scales are considered.

(1) Nominal scales $\mathbb{N}_n = (\{1, 2, \ldots, n\}, \{1, 2, \ldots, n\}, =)$. For $n = 4$, the corresponding contexts and Hasse diagram are given in Fig. 3, where a, b, c, d, e denote the concepts $(\{1,2,3,4\}, \varnothing)$, $(\{1\}, \{1\})$, $(\{2\}, \{2\})$, $(\{3\}, \{3\})$, $(\{4\}, \{4\})$, $(\varnothing, \{1,2,3,4\})$, respectively.

(2) Directed ordinal scales $\mathbb{D}_n = (\{1, 2, \ldots, n\}, \{1, 2, \ldots, n\}, <)$. For $n = 4$, the corresponding context and the Hasse diagram are given in Fig. 4, where a, b, c, d, e denote the concepts $(\{1,2,3,4\}, \varnothing)$, $(\{1,2,3\}, \{4\})$, $(\{1,2\}, \{3,4\})$, $(\{1\}, \{2\})$, $(\varnothing, \{1,2,3,4\})$, respectively.

(3) Undirected ordinal scales $\mathbb{U}_n = (\{1, 2, \ldots, n\}, \{< 2, < 3, \ldots, < n, > 1, > 2, \ldots, > n-1, \} \in)$, where $< k$ denotes the set $\{1, \ldots, k-1\}$, and $> k$ denotes the set $\{k+1, \ldots, n\}$. For $n = 4$, the corresponding context and the Hasse diagram are given in Fig. 5, where a, b, c, d, e, f, g, h, i, j, k denote the concepts $(\{1,2,3,4\}, \varnothing)$, $(\{1,2,3\}, \{< 4\})$, $(\{2,3,4\}, \{> 1\})$, $(\{1,2\}, \{< 3\})$, $(\{2,3\}, \{< 4, > 1\})$, $(\{3,4\}, \{> 2\})$, $(\{1\}, \{< 2\})$, $(\{2\}, \{< 3, > 1\})$, $(\{3\}, \{< 4, > 2\})$, $(\{4\}, \{> 3\})$, $(\varnothing, \{< 2, > 3\})$, respectively.

(4) Boolean scales $\mathbb{B}_n = (\{1, 2, \ldots, n\}, \{1, 2, \ldots, n\}, \neq)$. For $n = 4$, the corresponding context and the Hasse diagram are given in Fig. 6, where a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p denote the concepts $(\{1,2,3,4\}, \varnothing)$, $(\{1,2,3\}, \{4\})$, $(\{1,2,4\}, \{3\})$, $(\{1,3,4\}, \{2\})$, $(\{2,3,4\}, \{1\})$, $(\{1,2\}, \{3,4\})$, $(\{1,3\}, \{2,4\})$, $(\{1,4\}, \{2,3\})$, $(\{2,3\}, \{1,4\})$, $(\{2,4\}, \{1,3\})$, $(\{3,4\}, \{1,2\})$, $(\{1\}, \{2,3,4\})$, $(\{2\}, \{1,3,4\})$, $(\{3\}, \{1,2,4\})$, $(\{4\}, \{1.2,3\})$, respectively.

Other examples of scales are the real ordinal scale $\Sigma_o : (\mathbb{R}, M_O, \in)$, where $M_O = \{(-\infty, r] | r \in \mathbb{R}\}$, the interval real scale $\Sigma_I = (\mathbb{R}, M_I, \in)$, where $M_I = M_O \cup \{\{r - s, r, r + s\} | r \in \mathbb{R}, s \in \mathbb{R}^+$, the real ratio scale $\Sigma_R = (\mathbb{R}, M_R, \in)$, where $M_R = M_\Sigma \cup \{\{r : s, r, r \cdot s\} \{r \in R, s \in \mathbb{R}^+\}$, and their many-dimensional analogs.

	1	2	3	4
1	0	1	1	1
2	0	0	1	1
3	0	0	0	1
4	0	0	0	0

Fig. 4. The context $\mathbb{D}_4 = (\{1,2,3,4\}, \{1,2,3,4\}), <)$ and Hasse diagram of its concept lattice.

	<2	<3	<4	>1	>2	>3
1	1	1	1	0	0	0
2	0	1	1	1	0	0
3	0	0	1	1	1	0
4	0	0	0	1	1	1

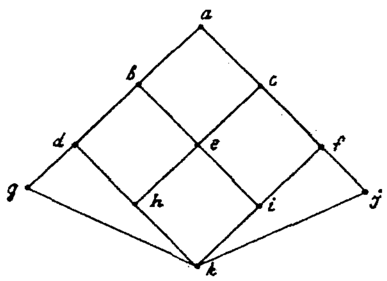


Fig. 5. The context $\mathbb{U}_4 = (\{1,2,3,4\}, \{<2, <3, <4, >1, >2, >3\}, \in\})$ and the Hasse diagram of its concept lattice.

Definition 2.5.1 ([80]). Let $K = (G, M, I)$ be a context and $\mathcal{S} = (G_S, M_S, I_S)$ be a scale. An *(partial) $\mathcal{S}$-measure* of K is a (partial) map σ from G into G_S such that for any extent A of a context $\mathcal{S}$ $\sigma^{-1}A$ is an extent of the context K. A $\mathcal{S}$-measure σ is called full if σ^{-1} defines an isomorphism between concept lattices $\mathfrak{B}(\sigma G, M_S, I_S \cap \sigma \times M_S)$ and $\mathfrak{B}(K)$.

The following two problems are essential for a context K and a scale $\mathcal{S}$: is a given (partial) map a (partial) $\mathcal{S}$-measure of K, and what (partial) $\mathcal{S}$-measures are possible for a given context K and a scale $\mathcal{S}$? We will deal these problems at first for a particular case of scaling, namely for directed ordinal scales.

Definition 2.5.2 ([80]). Let $(\Omega_t)_{t\in T}$ be a family of complete chains, and $\Omega = \times_{t\in T}\Omega_t$. Then $\Omega_{\leq} : (\Omega, \Omega, \leq)$ is called an *ordinal scale of dimension* $|T|$.

Since Ω is a complete lattice, the map $x \mapsto ((x], [x))$ gives an isomorphism from Ω onto $\mathfrak{B}$ by the Main Theorem of concept lattices (Theorem 2.1.1.). The following proposition allows one to study ordinal scaling in terms of lattice theory.

Proposition 2.5.1 ([80]). *For a full $\Omega_{\leq}$-measure μ of a context (G, M, I), let $\overline{\mu}(A, B) \rightleftharpoons \vee\mu A$ for all $(A, B) \in \mathfrak{B}(G, M, I)$. Then the mapping given by $\mu \mapsto \overline{\mu}$ is a bijection from the set of all full $\Omega_{\leq}$-measures of (G, M, I) onto the set of all $\vee$-embeddings of $\mathfrak{B}(G, M, I)$ into Ω; in particular, $\mu g = \overline{\mu}(\{g\}'', \{g\}')$ for all $g \in G$.*

The following theorem allows us to obtain a criterion of existence of an ordinal measure of a context.

Theorem 2.5.2 ([80]). *Let L be a finite lattice and $\chi \rightleftharpoons \{C_t | t \in T\}$ be a partition of the set $M(L)$ into chains. Then $\hat{\chi}$: $L \mapsto \times_{t\in T}(C_t \cup \{1_L\})$ defined as $\hat{\chi}a \rightleftharpoons (a_t)_{t\in T}$, where $a_t \rightleftharpoons \min\{c \in C_t \cup \{1_L\} | a \leq c\}$, is a $\vee$-embedding. If $\iota: L \mapsto \Omega$ is a $\vee$-embedding into a direct product of complete chains, then there exists a partition $\chi = \{C_t | t \in T\}$ of the set $M(L)$ into chains and a $\vee$-homomorphism κ from Ω into $\times_{t\in T}(C_t \cup \{1_L\})$ such that κ maps $M(\Omega) \cup \{1_\Omega\}$ onto $M(\times_{t\in T}(C_t \cup \{1_L\})) \cup \{1 \times C_t\}$ and $\hat{\chi} = \kappa \circ \iota$.*

	1	2	3	4
1	0	1	1	1
2	1	0	1	1
3	1	1	0	1
4	1	1	1	0

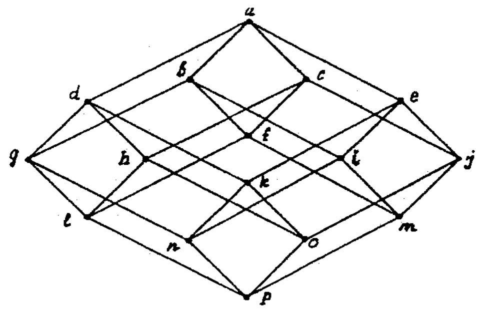


Fig. 6. The context $\mathbb{B}_4 = (\{1,2,3,4\}, \{1,2,3,4\}, \neq)$ and the Hasse diagram of its concept lattice.

Corollary 1 ([80]). *For a finite lattice L, $\vee$-dimension of L equals the width of $M(L)$, and $\vee$-rank of L equals to the cardinality of $M(L)$.*

Corollary 2 ([80]). *For a finite context (G, M, I) there exists a complete $\Omega_{\leq}$-measure: $\Omega = \times_{t\in T}\Omega_T$ iff there exists exactly one partition $\{C_t | t \in T\}$ of the set $M(L)$ into chains such that the length of Ω_t equals the cardinality of C_t for all $t \in T$. In particular, the dimension and the length of the ordinal scale $\Omega_{\leq}$ are equal to the width and the cardinality of $M(\underline{\mathfrak{B}}(G, M, I))$.*

Some properties of the tensor product allowed Wille [84] to prove the following theorem (which is close to Theorem 2.7.2 on dimension from [92]).

Theorem 2.5.3 ([84]). *Let $\mathcal{S} = (C, C, \leq)$ be a scale, where C is a finite chain, and σ be a complete $\mathcal{S}^n$-measure of the context (G, M, I) (where $\mathcal{S}^n$ is a direct product of n contexts $\mathcal{S}$). Then the context (G, M, I) is isomorphic to $(P, P, \not\leq)$ for a finite ordered set P and the lattice $\underline{\mathfrak{B}}(G, M, I)$ is isomorphic to a finite distributive lattice of all order-preserving maps from P into a chain of length 2.*

Corollary ([84]). For a finite partially ordered set P, the minimal number n such that there exists a complete $\mathcal{S}^n$-measure of the context $(P, P, \leq)$ coincides with the order dimension of P.

A criterion for existence of an $\mathcal{S}$-measure for the scale $\mathcal{S}$ of an arbitrary form is given in [47].

Proposition 2.5.4 ([31]). *Let $K = (G, M, I)$ be a context with a scale $\mathcal{S} = (G_\mathcal{S}, M_\mathcal{S}, I_\mathcal{S})$ and σ be a (partial) map from G into $G_\mathcal{S}$, and $K^\sigma = (G, M \dot\cup M_\mathcal{S}, I^\sigma)$, where for $g \in G$, $gI^\sigma m \iff m \in M$ and gIm or $m \in M_\mathcal{S}$ and $\sigma g I_\mathcal{S} m$. The map σ is a (partial) $\mathcal{S}$-measure of the context K iff for every $m \in M_\mathcal{S}$ there exists a set $B \subseteq M$ such that $\{m\}' = \cap_{b\in B}\{b\}'$. Moreover, the $\mathcal{S}$-measure σ of the context K is full iff for every $n \in M$, there exists a set $D \subseteq M_\mathcal{S}$ such that $\{n\}' = \cap_{d\in D}\{d\}'$.*

Using Proposition 2.5.4, we can show that the identity map is a measure of $\mathbb{U}_4$ in $\mathbb{N}_4$ and in $\mathbb{D}_4$, but not in $\mathbb{B}_4$.

Proposition 2.5.5 ([31]). *Let $\mathcal{S}$ be a scale of K such that $\{g\}' \neq \{h\}'$ for all $g, h \in G_\mathcal{S} : g \neq h$. If σ is a (partial) $\mathcal{S}$-measure of K, then a $\vee$-homomorphism $\overline{\sigma}$ of (a principal ideal of) $\underline{\mathfrak{B}}(K)$ into $\underline{\mathfrak{B}}(\mathcal{S})$ is defined by $\overline{\sigma}(A, B) = ((\sigma A)'',)\sigma A)')$. The mapping $\sigma \mapsto \overline{\sigma}$ is a bijection from the set of all (partial) $\mathcal{S}$-measures of K onto the set of all $\vee$-homomorphisms α of (a principal) ideal $\mathfrak{A}$ of) $\underline{\mathfrak{B}}(K)$ into $\underline{\mathfrak{B}}(\mathcal{S})$ with the property that for every $(\{g\}'', \{g\}')$ in $(\mathfrak{A}\cap)\underline{\mathfrak{B}}(K)$ there exists an $h \in G_\mathcal{S}$ with $(\alpha(\{g\}'', \{g\}') = (\{h\}'', \{h\}')$ ($\alpha 0 \neq 0$ is admitted). Moreover, an $\mathcal{S}$-measure σ of K is full iff $\overline{\sigma}$ is injective.* (From here on conditions that correspond to partial maps are given in brackets.)

It is obvious that the proposition dual to Proposition 2.5.5 (where σ^{-1} specifies a $\vee$-homomorphism from $\underline{\mathfrak{B}}(\mathcal{S})$ into $\underline{\mathfrak{B}}(K)$ injective on the image of $\overline{\sigma}$) is also valid.

For finite contexts from the example above, the problem of the existence of corresponding scales can be solved by means of the following:

Proposition 2.5.6 ([31]). *The context K admits a (partial) $\mathcal{S}$-measure*

(1) $\mathcal{S} = \mathbb{N}_n$ *iff the objects of (some extent of) K can be partitioned into n extents.*

(2) $\mathcal{S} = \mathbb{D}_n$ *iff there exists a chain of $n+1$ extents of K including the empty set.*

(3) $\mathcal{S} = \mathbb{U}_n$ *iff there exists a chain of n non-empty extents of K so that their complements (intersected with the largest extent of the chain) are again extents.*

(4) $\mathcal{S} = \mathbb{B}_n$ *iff there exists an independent set of n extents of K (these are extents $A_1, \ldots, A_n$ with $\bigcap_{i=1}^n A_i = \varnothing$ and $A_j \not\subseteq \bigcap_{i\neq j} A_i$ for $j = 1, \ldots, n$), whose union is the set of all objects of (some extent of) K.*

In [31] the scaling of many-valued contexts (see Sec. 2.2) is considered. For a many-valued attribute m the scale $\mathcal{S}_m = (G, M_m, I_m)$, where $M_m \subseteq W$ is the set of values of the attribute m, $gI_m w \iff (g,m)Iw$. Thus, the attribute $m \in M$ corresponds to a (partial) measure $\tilde{m}$ with respect to the context for the scale $\mathcal{S}_m = (G, M_m, I_m)$. The product operator $\times: \mathcal{S} = \times_{t\in T}\mathcal{S}_t = (\times_{t\in T} G_t, N, J)$ is used for composition of scales, where the product can be a direct product or a semiproduct, defined in [33].

Definition 2.5.3 ([78]). Let $K_i = (G_i, M_i, I_i)$, $i \in \{1, \ldots, n\}$, be contexts. Then the context $(G_1 \times \ldots \times G_n,\ M_1 \times \{1\} \cup \ldots \cup M_n \times \{n\}, \nabla)$, where $(g_1, \ldots, g_n)\nabla(m, j) \iff g_j I_j m$, is called a *semiproduct* of the contexts $K_1, \ldots, K_n$ and is denoted by $K_1 \diamond \ldots \diamond K_n$.

The next definition is a trivial generalization of Definition 2.4.2.

Definition 2.5.4. Let $K_i = (G_i, M_i, I_i)$, $i \in \{1, \ldots, n\}$ be contexts. Then the context $(G_1 \times \ldots \times G_n,\ M_1 \times \ldots \times M_n, \nabla)$, where $(g_1, \ldots, g_n)\nabla(m_1, \ldots, m_n)$ iff $g_j I_j m_j$ for some $j \in \{1, \ldots, n\}$, is called a *direct product* of contexts $K_1, \ldots, K_n$ and is denoted by $K_1 \times \ldots \times K_n$.

As shown in [80], the direct product of lattices $\underline{\mathfrak{B}}(G_1, M_1, I_1)$ and $\underline{\mathfrak{B}}(G_2, M_2, I_2)$, where (G_1, M_1, I_1), (G_2, M_2, I_2), are contexts such that $G_1 \cap G_2 = \varnothing$, $M_1 \cap M_2 = \varnothing$, $G_i' = \varnothing$, $M_i' = \varnothing$ for $i \in \{1, 2\}$, is isomorphic to the lattice $\underline{\mathfrak{B}}(G_1 \cup G_2,\ M_1 \cup M_2,\ I_1 \cup I_2 \cup G_1 \times M_2 \cup G_2 \times M_1)$. Definitions of other possible operations over contexts are found, for example, in [33].

A many-valued context with the scale $\mathcal{S} = \prod_{m\in M} \mathcal{S}_m$ is called a scaled context and is denoted by $(K, \prod_{m\in M} \mathcal{S}_m)$. Then its *derived context* (denoted by $\tilde{K} = (G, N, \tilde{J})$) is a context with the set of objects K, the set of attributes that coincides with the set of attributes of $\mathcal{S}$ and relation $\tilde{J}$ defined as $g\tilde{J}n \rightleftharpoons (m(g))_{m\in M} Jn$ for the context $\mathcal{S}$. If the mapping $m(g)$ (we will also write mg) is defined not for all m and g, then $g\tilde{J}n \rightleftharpoons hJn$ for all objects $h = (h_m)_{m\in M}$ from S such that $h_m = m(g)$ if $m(g)$ is defined.

We set, for example, $\tilde{K} = (G, \dot{\bigcup}_{m\in M} M_m, \tilde{I})$, where $g\tilde{I}n$ if $n \in M_m$ and $mgI_m n$ for some $m \in M_m$. For $h \in G$ we denote by $\gamma_m h$ the concept $(\{h\}'', \{h\}')$ of the context $\mathcal{S}_m = (G, M_m, I_m)$.

Proposition 2.5.7 ([31]). *Let (G, N, J) be a context for which $\tilde{m}$ is a partial $\mathcal{S}_m$-measure for all $m \in M$. Then the identity map of G is a $\tilde{K}$-measure of the context (G, N, J).*

The next proposition provides a means for constructing of $\underline{\mathfrak{B}}(\tilde{K})$.

Proposition 2.5.8 ([31]). *There is an isomorphism ι of $\underline{\mathfrak{B}}(\tilde{K})$ onto the $\vee$-semilattice of $\times_{m\in M}\underline{\mathfrak{B}}(\mathcal{S}_m)$ generated by the elements*

$$\iota(\{g\}'', \{g\}') \rightleftharpoons (\gamma_m \tilde{m} g)_{m\in M}$$

with $g \in G$.

Proposition 2.5.9 ([33]). *Let σ be an S-measure of a context $K = (G, M, I)$. Then the map*

$$(A, A') \mapsto (\sigma^{-1}(A), \sigma^{-1}(A)')$$

describes a $\wedge$-homomorphism from the lattice $\underline{\mathfrak{B}}(\mathcal{S})$ into the lattice $\underline{\mathfrak{B}}(K)$. This homomorphism is injective if σ is surjective.

Proposition 2.5.10 ([33]). *Let $\mathcal{S}$ be a scale in which $v \neq w$ implies $\{v\}' \neq \{w\}'$ for all $v, w \in G_{\mathcal{S}}$. Then, for an S-measure σ of a context $K = (G, M, I)$,*

$$(A, A') \mapsto \bar{\sigma}(A, A') \rightleftharpoons (\sigma(A)'', \sigma(A)')$$

describes a $\vee$-preserving map $\bar{\sigma}$ from $\underline{\mathfrak{B}}(K)$ into $\underline{\mathfrak{B}}(\mathcal{S})$; in particular, $\bar{\sigma}(\gamma g) = \gamma_{\mathcal{S}}\sigma(g)$ for all $g \in G$. Conversely, if φ is a $\vee$-homomorphism from $\underline{\mathfrak{B}}(K)$ into $\underline{\mathfrak{B}}(\mathcal{S})$ such that for each $g \in G$ there is a $\tilde{\varphi}(g) \in G_{\mathcal{S}}$ with

$\varphi(\gamma g) = \gamma_S\bar{\varphi}(g)$, then $\bar{\varphi}$ is a S-measure of K. There is a one-to-one correspondence between the S-measures σ (respectively $\bar{\varphi}$) and the specific $\vee$-preserving maps σ (respectively φ). The map σ is full iff $\bar{\sigma}$ is injective.

Proposition 2.5.11 ([33]). *Let $K = (G, M, I)$ be a finite context. If $\iota: \{1, \ldots, n\} \mapsto M$ is a bijection of the set of first n natural numbers onto the set M, then a full $\mathbb{B}_n$-measure of K is given by*

$$\sigma(g) = \{1, \ldots, n\} \setminus \iota(\{g\}')$$

for $g \in G$.

Definition 2.5.5 ([33]). A dichotomic scale is a context of the form $(\{0,1\}, \{0,1\}, =)$. A k-dimensional dichotomic scale is a semiproduct of k dichotomic scales: $D_1 \diamond \ldots \diamond D_k$.

Proposition 2.5.12 ([33]). *A finite context $K = (G, M, I)$ admits a full scale measure into the k-dimensional dichotomic scale iff K is atomistic (i.e., $(\{g\}', \{g\}'')$ is an atom of $\underline{\mathfrak{B}}(K)$ for any $g \in G$). K admits a full scale measure into the k-dimensional dichotomic scale iff K is atomistic and there are at most k pairs of complementary extents, to which all the extents of $\wedge$-irreducible concepts of $\underline{\mathfrak{B}}(K)$ belong.*

2.6. Tolerance and Congruence on Concept Lattices

Another way of representing concept lattices as compositions of smaller lattices is based on the use of a tolerance relation defined on lattice elements.

Definition 2.6.1. A binary relation θ on elements of a complete lattice L is called a *complete tolerance relation* if θ is reflexive, symmetric, and agrees with the lattice operations $\wedge$ and $\vee$, i.e., $x_t\theta y_t$ for all t of a set T, where x_t, $y_t \in L$ for all $t \in T$ implies that $(\wedge_{t\in T}x_t)\theta(\wedge_{t\in T}y_t)$ and $(\vee_{t\in T}x_t)\theta(\vee_{t\in T}y_t)$.

A complete tolerance relation is called *complete congruence relation* if it is transitive (i.e., is an equivalence relation).

Complete congruence relations were studied in detail for general complete lattices (see, for example, [8]). Results concerning complete tolerance relations for general complete lattices are found, e.g., in [84]. In this section, we present only the results on general lattices that will be used in the discussion of tolerances and congruences on concept lattices. Using Lemma 2 from [84], we give a simple definition of a block of tolerance relation.

Definition 2.6.2 ([84]). Let L be a complete lattice. A set S of elements of L is called a *block of tolerance relation* θ if it is maximal by inclusion among subsets of L such that for any pair of elements x, y of this subset $x\theta y$. The set of all blocks of θ is denoted by L/θ.

The blocks of a tolerance relation can also be defined as intervals of the form $[a]_\theta = [a_\theta, (a_\theta)^\theta]$ or $[b]^\theta = [(b^\theta)_\theta, b^\theta]$, where $a_\theta = \wedge\{x \in L | a\theta x\}$ and $a^\theta = \vee\{x \in L | a\theta x\}$ and a, b are arbitrary elements of L. The equivalence of the two definitions is proved, e.g., in [86].

Theorem 2.6.1 ([86]). *The set L/θ of all blocks of θ becomes a complete lattice (called the complete factor lattice of L with respect to θ) by defining*

$$B_1 \leq B_2 \rightleftharpoons \wedge B_1 \leq \wedge B_2 (\rightleftharpoons \vee B_1 \leq \vee B_2) \text{ for } B_1, B_2 \in L/\theta;$$

in particular,

$$\bigwedge_{t\in T} [x_t]^\theta = [\bigwedge_{t\in T} x_t]^\theta$$

and

$$\bigvee_{t\in T} [x_t]_\theta = [\bigvee_{t\in T} x_t]_\theta \text{ for } x_t \leq L.$$

By Theorem 2.6.1, complete tolerance allows one to decompose a lattice L into a set of intervals, which is itself a lattice. Objects of this kind are studied, for example, in [86] as Q-atlases.

Definition 2.6.3 ([86]). Let Q and L_q $(q \in L_q)$ be complete lattices. The family $(L_q)_{q\in Q}$, together with the $\vee$-morphism φ_q^r: $L_q \to L_r$ (in particular, $\varphi_q^r 0_1 = 0_r$) and a $\wedge$-morphism $\psi_q^r : L_r \to L_q$ (in particular, $\psi_q^r 1_r = 1_q$) for each $q \leq r$ in Q is called a *Q-atlas* if the following conditions are satisfied:

(1) $L_q \cap L_r$ is a filter of L_q and an ideal of L_r for all $q \leq r$ in Q.

(2) $\{q \in Q | x \in L_q\}$ is an interval $[x \min, x \max]$ in Q for all $x \in \bigcup_{q \in Q} L_q$.

(3) ϕ^q_q and ψ^q_q are the identities of L_q.

(4) $\varphi^r_q x \leq y$ iff $x \leq \psi^r_q y$.

(5) $\varphi^s_r \varphi^r_q = \varphi^s_q$ and $\psi^r_q \psi^s_r = \psi^s_q$.

(6) $\varphi^r_q x = \varphi^{r \vee s}_{q \vee s} x$ for all $x \in L_q \cap L_{q \vee s}$ and $\psi^t_s y = \psi^{t \wedge r}_{s \wedge r} y$ for all $y \in L_t \cap L_{t \wedge r}$.

The pair $(\bigcup_{q \in Q} L_q \sqsubseteq)$, where $x \sqsubseteq y \rightleftharpoons x\min \leq x\max$ and $\varphi^{ymin}_{xmin} x \leq y$ for all $x, y \in \bigcup_{q \in Q} L_q$, is called a sum of a Q-atlas.

As shown in Theorem 6 from [86], the sum of the Q-atlas is a complete lattice with blocks of complete tolerance isomorphic to the lattices of L_q.

Thus, the construction of a Q-atlas allows one to represent a lattice diagram as the sum of parts with a nonempty intersection in the same way as a geographic map is represented by an atlas.

In [86], for an arbitrary context $K = (G, M, I)$ a relation between complete tolerance of the lattice $\underline{\mathfrak{B}}(K)$ and the so-called block relations of the context K, i.e., the relations $J: I \subseteq J \subseteq G \times M$, for which $\{g\}$ is an intent of the context K, and $\{m\}'$ is an extent of the context K for an arbitrary $g \in G, m \in M$, is established. Thus, every extent and every intent of the context (G, M, J) is, respectively, an extent and an intent of the context (G, M, I). The set of all block relations of a context is a complete lattice with respect to the set-theoretic intersection.

Theorem 2.6.2 ([86]). *For a context $K = (G, M, I)$, there exists an isomorphism β from the lattice of all complete tolerance relations of the lattice $\underline{\mathfrak{B}}(K)$ onto the lattice of all block relations K given by*

$$g\beta(\theta)m \Longleftrightarrow \gamma g\theta(\gamma g \wedge \mu m)(\Longleftrightarrow (\gamma g \vee \mu m)\theta\mu m).$$

Furthermore, $(A, B)\beta^{-1}(J)(C, D) \Longleftrightarrow A \times D \cup B \times C \subseteq J$.

As noted in [86], the blocks of a complete tolerance of a context (G, M, I) correspond to the concepts of the context given by the block relation, i.e., (G, M, J). If the concept lattice of a context (G, M, I) can be represented by a 0–1 matrix, where every concept corresponds to a unit submatrix maximal by inclusion, then for a block relation J, every concept (H, N) of the context (G, M, J) corresponds to the set of all maximal unit submatrices of the matrix $I \cap H \times N$.

A particular case of Q-atlases and Q-sums are the so-called *Q-tied atlases* and *Q-tied sums* [86] encountered in the case where θ is a congruence relation. It is interesting that in this case every congruence is also associated with a context. To describe this result in detail, we present some auxiliary results that allow one to understand the meaning of subcontexts.

Definition 2.6.4 ([29]). Let (G, M, I) be a context and $H \subseteq G$, $N \subseteq M$. Then $(h, M, I \cap H \times N)$ is called a subcontext of (G, M, I).

In the general case, the concept lattice of a subcontext of the context (G, M, I) is not a sublattice of the lattice $\underline{\mathfrak{B}}(G, M, I)$. We can only expect that $\underline{\mathfrak{B}}(G, M, I)$ contains $\underline{\mathfrak{B}}(H, N, I \cap H \times N)$ as a suborder.

Theorem 2.6.3 ([29]). *For arbitrary $N \subseteq M$ the mapping*

$$\underline{\mathfrak{B}}(G, M, I \cap G \times N) \mapsto \underline{\mathfrak{B}}(G, M, I)$$

given by $(A, B) \mapsto (A'', A')$ is an order-embedding and so is the mapping given by $(A, B) \mapsto (B', B'')$.

Theorem 2.6.4 ([29]). *For arbitrary $N \subseteq M$, the mapping*

$$\underline{\mathfrak{B}}(G, N, I \cap G \times N) \mapsto \underline{\mathfrak{B}}(G, M, I)$$

given by

$$(A, B) \mapsto (A, A')$$

is an infimum-preserving order-embedding. Dually, for any $N \subseteq M$ the mapping

$$\underline{\mathfrak{B}}(H, N, I \cap H \times N) \mapsto \underline{\mathfrak{B}}(G, M, I)$$

given by

$$(A, B) \mapsto (B', B)$$

is a supremum-preserving order-embedding.

Definition 2.6.5 ([29]). A subcontext $(H, N, I \cap H \times N)$ of a context (G, M, I) is called *compatible* if $(A \cap H, B \cap N)$ is a concept of $(H, N, I \cap H \times N)$ whenever (A, B) is a concept of (G, M, I).

Proposition 2.6.5 ([80]). *Let (G, M, I) be a context, $\{G_j | j \in J\}$ be a partition of G, and $\{M_k | k \in K\}$ be a partition of M. Then the $\wedge$-embedding of the lattice $\mathfrak{B}(G, M, I)$ into the direct product of lattices of subcontexts $\mathfrak{B}(G_j, M, I \cap G_j \times M)$ is given by the mapping $(A, B) \mapsto (A \cap G_j, (A \cap G_j)')_{j \in J}$, and the $\vee$-embedding of the lattice $\mathfrak{B}(G, M, I)$ into the direct product of subcontexts $\mathfrak{B}(G, M_k, I \cap G \times M_k)$ is given by the mapping $(A, B) \mapsto ((B \cap M_k)', B \cap M_k)_{k \in K}$.*

Theorem 2.6.6 ([29]). *If $(H, N, I \cap H \times N)$ is a compatible subcontext of (G, M, I), then the mapping*

$$\Pi_{H,N} : \mathfrak{B}(G, M, I) \mapsto \mathfrak{B}(H, N, I \cap H \times N)$$

given by

$$(A, B) \mapsto (A \cap H, B \cap N)$$

is a surjective complete lattice homomorphism. If (G, M, I) is finite, then, conversely, for every complete congruence relation θ of $\mathfrak{B}(G, M, I)$ there exists a complete subcontext $(H, N, I \cap H \times N)$ such that θ is the kernel of the homomorphism $\Pi_{H,N}$.

Consider another special type of Q-atlases, where $B_1 \cap B_2 \neq \varnothing$ holds for every covering pair of blocks $B_1 \prec B_2$ of the tolerance θ (i.e., blocks such that $B_1 \leq B_2$ and there is no B': $B_1 \leq B' < B_2$) The corresponding tolerance relation is called a *glued tolerance.* An L/θ-atlas of a glued tolerance is completely determined by the blocks of θ and their intersections. Thus, the definition of the mappings φ and ψ from the definition of a Q-atlas (Definition 2.6.3) is not needed.

Theorem 2.6.7 ([86]). *Let L be a lattice of finite length and let $\Sigma(L)$ be the smallest tolerance relation containing all covering pairs of elements in L. Then $\Sigma(L)$ is the smallest glued tolerance relation of L.*

Theorem 2.6.8 ([86]). *Let (G, M, I) be a context such that $\mathfrak{B}(G, M, I)$ has finite length. Then $J = \beta(\Sigma(L)$ (where $\Sigma(L)$ is defined in the preceding theorem) is the smallest block relation of (G, M, I) containing all pairs (g, m) such that $\{g\}'$ is maximal in $\{\{h\}' | h \in G$ and $(h, m) \notin I\}$ or $\{m\}'$ is maximal in $\{\{n\}' | n \in M$ and $(g, n) \notin I\}$; especially, an isomorphism from $\mathfrak{B}(G, M, I)$ onto L/Σ is given by $(H, N) \mapsto \{(A, B) \in L | A \subseteq H$ and $B \subseteq N\}$.*

Definition 2.6.6 ([93]). Let $L(\Sigma)$ be the least glued complete tolerance relation of a lattice L. Then the complete lattice $S(L) = L/\Sigma(L)$ is the *skeleton* of L. This construction may be iterated as follows: $S_0(L) = L$ and $S_r(L) = S(S_{r-1}(L))$ for $R = 1, 2, 3, \ldots$; $S_r(L)$ is called the rth skeleton of L.

In [92], the notion of skeleton is used in the study of free complete distributive lattices.

Definition 2.6.7 ([6]). A lattice L is (α, β) $\wedge$-*distributive* if it satisfies the following condition D:

If $(x_{st})_{s \in S, t \in T}$ is a family of elements in L satisfying the conditions

(1) $0 < |S| \leq \alpha$, $0 < |T| \leq \beta$,

(2) $\bigvee_{t \in T} x_{st}$ exists for each $s \in S$,

(3) $\bigwedge_{s \in S} \bigvee_{t \in T} x_{st}$ exists,

(4) $\bigwedge_{s \in S} x_{s\varphi(s)}$ exists for each function $\varphi \in T^S$ (the set of functions from S into T),

then $\bigvee_{\varphi \in T^S} \bigwedge_{s \in S} x_{s\varphi(s)}$ exists and $\bigwedge_{s \in S} \bigvee_{t \in T} x_{st} = \bigvee_{\varphi \in T^S} \bigwedge_{s \in S} x_{s\varphi(s)}$.

Definition 2.6.8 ([6]). A lattice L is (α, β) $\vee$-*distributive* if it satisfies the condition dual to D (i.e., the condition where the sums and products are interchanged).

Definition 2.6.9 ([6]). A lattice L is *completely* $\wedge$-*distributive* if it is (α, β) $\wedge$-distributive for all α and β (the same for *complete* $\vee$-*distributivity*).

Definition 2.6.10 ([6]). A lattice L is *completely distributive* if it is completely $\wedge$- and $\vee$-distributive.

Thus, the join and meet operations of a completely distributive lattice (in contrast to complete distributive lattices) are defined for countable subsets only.

Theorem 2.6.9 ([84]). *A concept lattice $\mathfrak{B}(G, M, I)$ is completely distributive iff for any $g \in G$ and $m \in M$ such that $(g, m) \notin I$ there exists $h \in G$ and $n \in M$ such that $(g, n) \notin I$, $(h, m) \notin I$, and $h \in \{k\}''$ for all $k \in G \backslash \{n\}'$.*

A *free completely distributive lattice* FCD(S) generated by a set S is specified up to an isomorphism by the fact that every mapping φ from S into a completely distributive complete lattice L can be extended up to the complete homomorphism from FCD(S) into L.

Definition 2.6.11 ([93]). Let $\mathcal{P}(S)$ denote the power set of S and $X, Y \subseteq S$. Then

$$X \Delta Y \rightleftharpoons X \cap Y \neq \varnothing$$

and for $r \in \mathbb{N}$

$$X \sum\nolimits_r^s Y \rightleftharpoons |S \backslash (X \cup Y)| \leq r - 1,$$

where $\sum_0^s \rightleftharpoons \Delta$.

Theorem 2.6.10 ([93]). *For an arbitrary* FCD(S)

$$\mathrm{FCD}(S) \cong \mathfrak{B}(\mathcal{P}(S), \mathcal{P}(S), \Delta),$$

$$S_r(\mathrm{FCD}(S) \cong \mathfrak{B}(\mathcal{P}(S), \mathcal{P}(S), \Delta \cup \sum_r^s).$$

Definition 2.6.12 ([93]). Let S be a finite set; then

$$X \overline{\sum\nolimits_r^s} Y \rightleftharpoons |X| + |Y| \geq |s| + 1 - r.$$

Definition 2.6.13 ([93]). Let X, Y, T be sets; then

$$X \Delta_T Y \rightleftharpoons X \cap Y \cap T \neq \varnothing.$$

Theorem 2.6.11 ([93]). *For a finite set S and $T \subseteq S$ the lattice $\mathfrak{B}(\mathcal{P}(S),\ \mathcal{P}(S),\ \Delta_T \cup \overline{\sum_r^s})$ is a complete sublattice of the lattice $\mathfrak{B}(\mathcal{P}(S),\ \mathcal{P}(S),\ \Delta \cup \overline{\sum_r^s})$.*

Theorem 2.6.12 ([93]). *For a finite set S and any $T, U \subseteq S$ the following equation holds:*

$$\mathfrak{B}(\mathcal{P}(S), \mathcal{P}(S), \Delta_T \cup \overline{\sum\nolimits_r^s}) \cap \mathfrak{B}(\mathcal{P}(S), \mathcal{P}(S), \Delta_U \cup \overline{\sum\nolimits_r^s})$$

$$\cong \mathfrak{B}(\mathcal{P}(S), \mathcal{P}(S), \Delta_{T \cap U} \cup \overline{\sum\nolimits_r^s}).$$

Definition 2.6.14 ([93]). For a finite set S a *bicover of degree* r with bound K is a pair $(\mathcal{X}, \mathcal{Y})$ with $\mathcal{X}, \mathcal{Y} \in \mathcal{P}(S)$ such that, for every $R \subseteq S$ with $|R| = r$, there are $X_R \in \mathcal{X}$ and $Y_R \in \mathcal{Y}$ with $X_R \cap Y_R \subseteq R$ and $|X_R| + |Y_R| \leq k$ and for $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, $X \cap Y \neq \varnothing$ or $|X| + |Y| > k$. Let $\mathrm{bic}_r(S)$ be the smallest number k for which S admits a bicover of degree r with bound k.

Theorem 2.6.13 ([93]). *For a non-empty finite set S and for $r = 1, 2, 3, \ldots, |S|$, $\mathfrak{B}(\mathcal{P}(S),\ \mathcal{P}(S),\ \Delta \cup \overline{\sum_r^s}$ with $|T| = |S| - r$ iff $|S| - r < \mathrm{bic}_r(S)$.*

A tolerance relation defined only on the set of objects G of a context (G, M, I) is studied in [37-39]. Thus, the lattice nature of the set of all concepts is not taken into account. Therefore, we consider results from [37-39] in a separate section (Sec. 5.4.)

2.7. Decomposition of Concept Lattices and Automatic Drawing of Their Diagrams

Various techniques based on the results of formal concept analysis were proposed in the paper [92] as a means for the automatic drawing of Hasse diagrams of lattices.

By finding chains in a concept lattice, we can determine a minimal grid, where the lattice can be embedded: if there are n independent chains, the diagram can be embedded in an n-dimensional grid that is the product of the chains. To determine the minimal number of chains that allows embedding of a concept lattice in their product, Wille [92] proposed the use of the notion of Ferrers dimension.

Definition 2.7.1. Let G and M be finite sets. Then a relation $F \subseteq G \times M$ is called a Ferrers relation if $\underline{\mathfrak{B}}(G, M, F)$ is a chain. The *Ferrers dimension* of a context (G, M, I) is the smallest number of Ferrers relations on $G \times M$ whose intersection is I (or the minimal number of Ferrers relations on $G \times M$ whose union is $G \times M \backslash I$).

Theorem 2.7.1 (Dimension Theorem [92]). *The Ferrers dimension of a context (G, M, I) is equal to the order dimension of $\underline{\mathfrak{B}}(G, M, I)$, i.e., the smallest number of chains which admits an order-embedding of $\underline{\mathfrak{B}}(G, M, I)$ into their direct product.*

Since the order dimension of a partially ordered set $(P, \leq)$ is equal to the order dimension of its Dedekind–MacNeille closure (i.e., $\underline{\mathfrak{B}}(P, P, \leq)$, see Sec. 2.1) the order dimension of an ordered set $(P, \leq)$ is equal to the Ferrers dimension of $(P, P, \leq)$.

Another way to embed a lattice $\underline{\mathfrak{B}}(G, M, I)$ into a product of chains is to use the $\vee$-dimension of the lattice, i.e., the smallest number of chains whose product admits the embedding of the lattice $\underline{\mathfrak{B}}(G, M, I)$.

Theorem 2.7.2 (on $\vee$-dimension [80]). *The $\vee$-dimension of a finite lattice L is equal to the width of the set of all $\wedge$-irreducible elements of L.*

A computer program based on the theorem on $\vee$-dimension is reported in [92] to construct concept lattices. First, the program finds all attributes $m \in M$ such that $\{m\}'$ are not intersections of other extents $\{n\}'$, $n \in M$. Then the attributes are ordered in the following way: $m_1 \leq m_2 \iff \{m_1\}' \subseteq \{m_2\}'$. Thus, a partially ordered set isomorphic to $M(\underline{\mathfrak{B}}(G, M, I))$ is constructed. The partially ordered set obtained inthis way is decomposed into the smallest number of chains by a specialization of the Ford–Fulkerson algorithm, and thus the dimension of the grid is established. Then the operator can choose a basis of the grid. Thereafter the program locates the elements of the lattice in the nodes of the grid and joins them by line segments. Then the projection of the grid that ensures the best appearance of the diagram is sought interactively [55]. A diagram of a lattice constructed by a computer program is usually quite acceptable if the lattice does not differ much from a distributive one ([92]).

A standard way of automatic drawing of lattices in formal concept analysis is based on the decomposition of lattices into products of simpler lattices. Products of this kind include the tensor product [86] (see Sec. 2.4) as well as the substitution product, which corresponds to the substitution sum of contexts (see [56, 84]).

Definition 2.7.2 ([56]). Let $K_1 = (G_1, M_1, I_1)$, $K_2 = (G_2, M_2, I_2)$ be contexts. Let for any $X \subseteq G_j$ and $Y \subseteq M_j$ $(j \in \{1,2\})$, $X^j = \{m \in M_j | gI_jm$ for every $g \in X\}$, $Y^j = \{g \in G_j | gI_jm$ for all $m \in Y\}$. Let $(g, m) \in G_1 \times M_1 \backslash I_1$. Then $G_1(g)G_2 = (G_1 \backslash \{g\}) \cup G_2$, $M_1(m)M_2 = (M_1 \backslash \{m\}) \cup M_2$, and $I_1(g,m)I_2 = \{(h, n) \in I_1 | h \neq g$ or $n \neq m\} \cup G_2 \times \{g\}^1 \cup \{m\}^1 \times M_2 \cup I_2$. The context $(G_1(g)G_2, M_1(m)M_2, I_1(g,m)I_2)$ is called the *substitution sum* of K_2 with K_1 over (g, m) and is denoted by $K_1(g, m)K_2$.

This construction can be understood as substituting K_2 into K_1 at the spot (g, m). For lattices, the counterpart of the substitution sum is the substitution product.

Definition 2.7.3 ([56]). Let L be a lattice and M be a bounded lattice (with 0 and 1) and let $a \not\geq b$ in L. Then $(a] * [b) = \{(u, v) \in (a] \times [b) | u = a \wedge v$ and $u \vee b = v\}$ is an order-isomorphism between $a \wedge [b)$ and $(a] \vee b$. An element (u, v, y) from $((a] * [b))$ $(M \backslash \{0, 1\})$ is denoted by $u[y]v$. Furthermore, $u[0]v = u$ and $u[1]v = v$ for $(u, v) \in (a] * [b)$, where 0 and 1 are the bounds of M (always assume $0 < 1$). On $(a] * M * [b) = \{u[y]v | (u, v) \in (a] * [b)$ and $y \in M\}$ we define the relation of partial order by means of the relation $u[y]v \leq w[z]x = u \leq w$, $y \leq z$, $v \leq x$. It is obvious that the partially ordered set $(a] * M * [b)$ is isomorphic to $(a \wedge [b)) \times M$ and to $M \times ((a] \vee b)$. $L \cup (a] * M * [b)$ together with the transitive closure of the order relation on L and on $(a] * M * [b)$ is a lattice called the *substitution product* of L with M over (a, b) and is denoted by $L(a, b)M$.

(Substitution)**Theorem 2.7.3** ([56]). *Let $\{h\}^{\prime} \neq M_2$ and $\{n\}^{\prime} \neq G_2$ for all $h \in G_2$ and $n \in M_2$. Then $\mathfrak{B}(K_1(g,m)K_2) \not\cong \mathfrak{B}(K_1)\ (\mu_1 m, \gamma_1 g)\mathfrak{B}(K_2)$.*

The program [55] that uses the substitution theorem decomposes sequentially the lattice $\mathfrak{B}(G, M, I)$ into *substitution sums* of indecomposable lattices with diagrams from some library. When the program construct the diagrams of all indecomposable factor lattices, it constructs the diagram of their product. The following theorem establishes the independence of the result of substitution decomposition of the decomposition order.

(Decomposition) **Theorem 2.7.4** ([56]). *Two substitution decompositions of a complete lattice into substitutionally indecomposable factors have the same length and pairwise isomorphic factors.*

Diagrams of large lattices often become illegible when various vertices and edges fuse. The idea of using a nested line diagram [81] consists in distinguishing a part of the diagram and substituting it by a vertex, where the vertex in the new diagram is connected to all vertices from the neighborhood of the removed part of the diagram. The removed parts of the diagram correspond to the vertices of the resulting diagram. Formally, the decomposition of a lattice diagram corresponds to the decomposition of a concept lattice into direct products according to theorems from [80] (Sec. 2.3).

Finally, a representation of lattice diagrams by means of smaller diagrams can be carried out by using the properties of congruence and tolerance relations, as well as their atlases (see Sec. 2.7).

In various studies, different conditions for embedding diagrams in grids are used (see [91]). These conditions, for example, can take into account the rank function (an example of the insufficiency of this condition is presented in [91]) or the minimality of the number of intersections of the resulting diagram. In the studies on automatic drawing of lattice diagrams based on formal concept analysis, it is required (see [91]) that the lines of the diagram which correspond to the ("covering") relation $\prec$ be straight and as steep as possible. In [71], a method of drawing diagrams is studied where every 4-tuple of elements (a, b, c, d) such that $a \prec b$, $a \prec c$, $b \prec d$, $c \prec d$ is depicted as a parallelogram. It was shown that a local-distributive lattice can be represented by such a diagram (a lattice L is called locally distributive if, for an arbitrary $a \in L$, the distributive law holds for the interval $[a, \vee\{b \in L | a \prec b\}]$).

3. Concepts and Dependences of Attributes

3.1. Main Definition

The search for dependences in data is a frequently encountered problem of computer science. This problem was studied within the framework of formal concept analysis, starting from [85, 15], where the definition of dependence in data was given. A similar definition of dependence was introduced earlier in the JSM-method [19, 20] and in [25, 26], but the definitions were given there in other terms. For the sake of uniformity, in this section we will also speak in terms of formal concept analysis. The following definition of dependence was given in [31, 16] (in French and German papers the term implication (Implikation) is used, but we prefer to use the term dependence in line with the English and Russian terminology).

Definition 3.1. Let $K = (G, M, I)$ be a context and $A \subseteq M$, $B \subseteq M$ be arbitrary subsets of attributes. The set of attributes B depends on the set of attributes A (denoted by $A \to B$) if all objects from G that possess the set of attributes A also possess the set of attributes B, i.e., $A' \subseteq B'$ (or $B'' \supseteq A''$).

Thus, the dependence of the set of attributes B on the set of attributes A corresponds to the fact that in the Hasse diagram of the lattice $\mathfrak{B}(G, M, I)$ the concept (A'', A') lies below the concept (B'', B').

3.2. Search for Dependences in the JSM-Method

The first version of the JSM-method of automatic hypothesis generation (named so after John Stuart Mill) was proposed in [19]. In this method, hypotheses concerning the causes of properties are sought among the concepts of the context determined by a set of objects and a set of structural and functional properties (or attributes) of these objects (see the recent papers [22-24] and reviews [48, 50], where a complete list of published papers about the JSM-method is found, including papers about applied studies in pharmacology and technical diagnostics).

Assume that W is a property of objects from a domain under study. Then the input data for the JSM-method can be represented by the sets of positive, negative, and undefined examples. Positive examples are objects that are known to possess the property W, and negative examples are objects that are known not to possess this property. Undefined examples are those that are neither known to possess the property nor known not to possess the property.

In terms of formal concept analysis, this means that three contexts are considered: the positive context $K_+ = (G_+, M_+, I_+)$, the negative context $K_- = (G_-, M_-, I_-)$, and the undefined one $K_\tau = (G_\tau, M_\tau, I_\tau)$. Here G_+, G_-, and G_τ are the sets of positive, negative, and undefined examples, respectively; M is the set of "structural" attributes (that does not contain the "functional" property W); $I_j \subseteq G_j \times M$, $j \in \{+, -, \tau\}$, are relations that specify the structural attributes of positive, negative, and undefined examples. In what follows, we will use X' instead of X^{I_+}, X^{I_-}, X^{I_τ} when it does not cause ambiguity.

Now, the JSM-hypotheses can be defined in the following way. If the intent i_+ of a positive concept (e_+, i_+), i.e., of a concept of a positive context, does not coincide with the intent of a negative concept, then the concept (e_-, i_-) is called a positive hypothesis with respect to the property W. Negative hypotheses are defined dually.

Further requirements (called conditions or empirical dependences) on the form of the hypotheses are proposed in the JSM-method. These requirements are lattice-ordered with respect to their logical strength. The stronger the condition satisfied, the more plausible the hypothesis. For example, a positive "counterexample forbidding" condition requires for a positive hypothesis (e_+, i_+) that no negative example possess all properties from i_+, i.e., $\forall g \in G_-$, $i_+ \not\subseteq \{g\}'$ (analogously for negative hypotheses). The "counterexample forbidding" condition is quite natural and is used in various systems of machine learning and pattern recognition (see, for example, [60]). This condition can be formulated as the requirement that "generalization of positive examples should not cover any negative example." Note that if (e_+, i_+) is a positive JSM-hypothesis that satisfies the "counterexample forbidding" condition with respect to the property W, then $i_+ \to W$ is the dependence in the sense of Definition 3.1 for the context $K_{+-} = (G_+ \cup G_-, M \cup \{W\}, I_+ \cup I_- \cup G_+ \times \{W\})$.

Whereas a graph-theoretic interpretation of a concept for the context (G, M, I) is the maximal-by-inclusion complete bipartite subgraph of a bipartite graph, the interpretation of a hypothesis requires a tripartite graph. The vertices of the first and third parts of this graph (we denote it by T) correspond to the positive and negative examples from the sets G_+ and G_-, and the vertices of the second part correspond to the attributes of the set $M \cup \{W\}$ [50, 51]. The vertex that corresponds to the object $g_+ \in G_+$ is connected with all vertices that correspond to the attributes from $\{g_+\}'$ and only with them. For negative examples the converse holds: the vertex that corresponds to the object $g_- \in G_-$ is adjacent only to the vertices that correspond to the attributes not included in $\{g_-\}'$. Then, a hypothesis that satisfies the "counterexample forbidding" condition corresponds to the tripartite subgraph of T, where the vertices of the third and second parts constitute a maximal bipartite subgraph (say, the subgraph D), and the vertices of the second part of this subgraph dominate the vertices of the first part (i.e., every vertex of the third part is adjacent to a vertex of the second part of D).

Example. Consider the graph T in Fig. 7, where the vertices of the second (middle) part are labeled A, B, C, D, E, F, G. Then the sets of positive and negative examples of the corresponding problem on hypotheses will be $G_+ = \{X_1, X_2, X_3, X_4, X_5\}$, $G_- = \{Y_1, Y_2, Y_3, Y_4\}$, where $X_1' = \{A, B, C\}$, $X_2' = \{A, B, D\}$, $X_3' = \{A, E, F\}$, $X_4' = \{A, C, G\}$, $X_5' = \{A, C, G\}$; $Y_1' = \{A, F, G\}$, $Y_2' = \{A, D, F\}$, $Y_3' = \{B, E, F, G\}$, $Y_4' = \{B, D, F\}$.

Positive concepts (i.e., concepts of the positive context (G_+, M, I_+)) are $(\{X_1, X_2, X_3, X_4, X_5\}, \{A\})$, $(\{X_1, X_2\}, \{A, B\})$, $(\{X_1, X_4, X_5\}, \{A, C\})$, $(\{X_4, X_5\}, \{A, C, G\})$. The second, third, and fourth pairs correspond to the hypotheses satisfying the "counterexample forbidding" condition, i.e., conditions of the form $X \to W$ for the context $K_+ = (G_+ \cup G_-, M \cup \{W\}, I_+ \cup I_- \cup G_+ \times \{W\})$. In the case of the first pair, the vertex of the middle part that has the label A is not adjacent to the first and second vertices of the right (third) part. Therefore, $\{A\} \to W$ is not a positive hypothesis that satisfies the "counterexample forbidding

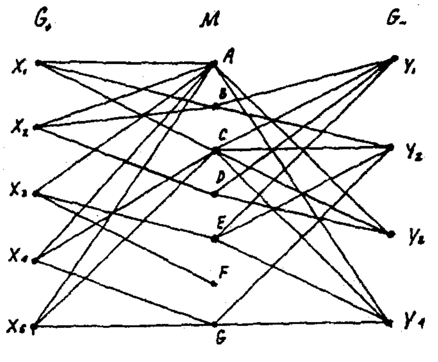


Fig. 7. Tripartite graph corresponding to the problem of hypothesis generation.

condition."

We can propose another description of hypotheses that satisfy the "counterexample forbidding" condition. This condition is especially convenient when we study not a single property W, but a subset of properties of a set of properties M_F. Consider two contexts: $K_s = (G, M_s, I_s)$ and $K_f = (G, M_f, I_f)$, where M_s and M_f are interpreted as the sets of "structural" and "functional" properties of objects from G, respectively (for example, the structural properties of a molecule and the biological properties of the corresponding chemical compound). Then a hypothesis that satisfies the "counterexample forbidding" condition and is about the dependence of the set of functional properties $F \subseteq M_f$ on the set of structural properties $S \subseteq M_s$ (i.e., the dependence $S \to F$) corresponds to the triple $(S^{I_s}, (S^{I_s})^{I_s}, (S^{I_s})^{I_s}$, where $(S^{I_s})^{I_f} = F$.

Hypotheses can be used for classification of undefined examples from G_τ (i.e., for establishing whether they possess the property W or not). If an undefined example $g_\tau \in G_\tau$ possesses all attributes from the intent i_+ of a positive hypothesis (e_+, i_+) (i.e., $\{g_\tau\}' \supseteq i_+$) and does not possess all attributes from the intent of any negative hypothesis, then we can infer that $\{g_\tau\}$ is likely to possess the property W. The pair (e_+, i_+) is called a hypothesis in favor of a positive forecast for g_τ. If an undefined example $g_\tau \in G_\tau$ possesses all attributes from the intent i_- of a negative hypothesis (e_-, i_-) (i.e., $\{g_\tau\}' \supseteq i_-$) and does not possess all attributes from the intent of any positive hypothesis, then we infer that $\{g_\tau\}$ is likely not to possess the property W. The pair (e_-, i_-) is called a hypothesis in favor of a negative forecast for g_τ. If $\{g_\tau\}$ does not include an intent of any negative or positive hypothesis, or includes intents of hypotheses of different signs, then no forecast is produced.

The following graph-theoretic interpretation of a forecast was proposed in [51].

Definition 3.2.1. The following problem is called the problem on "domination by the parts of complete bipartite graphs" (DPCBG):

INSTANCE. Tetrapartite graph $G = (V_1 \cup V_2 \cup V_3 \cup V_4, E)$, $E \subseteq (V_1 \times V_2) \cup (V_2 \times V_3) \cup (V_3 \times V_4)$. The graphs B_1, B_2, B_3 are the subgraphs of the graph G induced by the sets of vertices $(V_1 \cup V_2)$, $(V_2 \cup V_3)$, $(V_3 \cup V_4)$, respectively.

QUESTION. Does there exist a complete bipartite subgraph $C = (W_2 \cup W_3, W_2 \times W_3)$ of the graph B_2 such that it is maximal by inclusion, $W_2 \subseteq V_2$, $W_3 \subseteq V_3$, the set of vertices W_2 dominates V_1, and the set of vertices W_3 dominates V_4?

Definition 3.2.2. The following problem is called the "problem on a hypothesis in favor of a positive forecast" (HFPF):

INSTANCE. Input data: contexts $K_{+-} = (G_+ \cup G_-, M \cup \{W\}, I_+ \cup I_- \cup G_+ \times \{W\})$, $K_\tau = (G_\tau, M, I_\tau)$, and the objects $g_\tau \in G_\tau$, where $M = V_1 \cup V_3$, $G_+ = V_2$, $I_+ = \bigcup_{v_i^2 \in V_2} \{v_i^2\} \times \{v_1, \dots, v_r\}$, and $\{v_1, \dots, v_r\}$ is the union of the set of all vertices of V_3 adjacent to the vertex $v_i^2 \in V_2$ and the set of all vertices of V_1 not adjacent

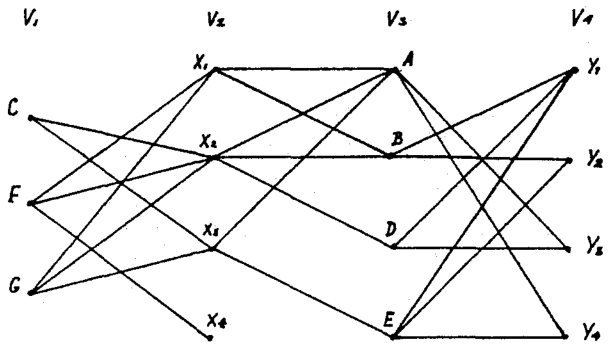


Fig. 8. Tetrapartite graph corresponding to the problem of a forecast from the example.

to the vertex $v_i^2 \in V_2$, $G_- = V_4$, $I_- = \bigcup_{v_k^4 \in V_4} \{v_k^4\} \times \{w_1, \ldots, w_q\}$, where $v_k^4 \in V_4$ and $\{w_1, \ldots, w_q\}$ is the set of all vertices of V_3 not adjacent to the vertex $v_k \in V_4$.

QUESTION. Does there exist a hypothesis (e_+, i_+) with respect to the property W that satisfies the "counterexample forbidding" condition and is a hypothesis in favor of a positive forecast for g_+, i.e., $i_+ \to W$ is the dependence for the context $K_{+-} = (G_+ \cup G_-, M \cup \{W\}, I_+ \cup I_- \cup G_+ \times \{W\})$ and $i_+ \subseteq \{g_\tau\}'$?

Lemma 3.2.1 ([51]). *The problem* DPCBG *for the tetrapartite graph G given in Definition 3.1 has a solution iff the corresponding problem of HFPF has a solution.*

Example. Consider the graph in Fig. 8, where the vertices of the first part are labeled C, F, G, and the vertices of the third part are labeled A, B, D, E. Then the sets of positive and negative examples in the corresponding HFPF problem for the undefined example g_τ: $\{g_\tau\}' = \{A, B, D, E\}$ will be $G_+ = \{X_1, X_2, X_3, X_4\}$, $G_- = \{Y_1, Y_2, Y_3, Y_4\}$, where $X_1' = \{A, B, C\}$, $X_2' = \{A, B, D\}$, $X_3' = \{A, E, F\}$, $X_4' = \{A, C, G\}$; $Y_1' = \{A\}$, $Y_2' = \{A, D\}$, $Y_3' = \{B, E\}$, $Y_4' = \{B, D\}$. The pairs $(\{X_1, X_2, X_3, X_4\}, \{A\})$, $(\{X_1, X_2\}, \{A, B\})$, $(\{X_1, X_4\}, \{A, C\})$ are the concepts for the positive context $K_+ = (G_+, M, I_+)$. Only the second pair is in favor of a positive forecast for g_τ, since in the case of the first pair the first and second vertices of the fourth part are not dominated and in the case of the third pair the vertex with label C does not lie in the third part.

Finn [22] proposed the so-called "generalized JSM-method." Each "generalized" hypothesis concerning the property W is a triple of the form $\langle (e_+, i_+), \mathcal{B}, W \rangle$, where (e_+, i_+) is the concept of the positive context $K_+ = (G_+, M, I_+)$, W is the property under study, and $\mathcal{B}$ is the family of all "specific obstacles" (different from (-)-hypotheses): the sets of form H', where $H \subseteq G_-$ and H' is minimal by inclusion among all sets of the form $\{Y' | Y \subseteq G_-, i_+ \subset Y'\}$, i.e., among intersections of the attributes of negative examples that possess properties from i_+). The generalized hypothesis $\langle (e_+, i_+)$, B, $W\rangle$ means that "W depends on i_+ in the absence of the sets of attributes from the family B." Contrary to the case of hypotheses that satisfy the "counterexample forbidding" condition, i_+ can coincide with an intent i_- of the negative context $K_- = (G_-, M, I_-)$. Many applications can justify the introduction of generalized hypotheses, for example, the pharmacological one, where the "structural causes" of a biological activity of a chemical compound, i.e., some parts of the corresponding molecule, can be suppressed by specific parts of the same molecule.

The JSM-method was defined above in terms of formal concept analysis. It can also be formulated for a more general case, namely, for the case where a semi-lattice operation $\sqcap$ is used instead of the set-theoretic intersection $\cap$ in the definition of $'$ and $''$ operations (and therefore, for the generation of concepts and hypotheses). In [50] a semi-lattice operation was defined for sets of graphs and numerical intervals and in [97] for data with numerical parameters. The description of a special logical language used for formulation of the JSM-theory can be found in [2-5]. Section 5.4 contains some results concerning the relation of concepts to tolerance relations. In Sec. 3.6 we present results concerning the stability of hypotheses, and in Sec. 4.6

we consider algorithmic complexity issues of hypothesis generation.

3.3. Bases of Dependences

The problem of generation of a minimal set of dependences that can reproduce the whole set of dependences is considered in [85, 16] (dependences are called *implications* there).

Definition 3.3.1 ([16]). A dependence $A \to B$ is called *informative dependence* (ID) if $B \not\subset A$.

Definition 3.3.2 ([16]). A dependence $A \to B$ is called *a maximal informative dependence* (MID) if for all X and Y such that $X, Y \subseteq G$, the conditions $(X \subseteq A), (B \subseteq Y), (X \to Y)$ imply $(X = A)$ and $(Y = B)$.

Proposition 3.3.1 ([16]). *Let $A \subseteq M$. The dependence $A \to A''$ is a maximal informative one iff $X'' \subseteq A$ holds for an arbitrary $X \subseteq A$.*

Definition 3.3.3 ([16]). A set $A \subseteq M$ is called a *gap* (*lacune*) if $A \neq A''$, i.e., the set A is not closed. The gap A is called *irreducible* if the dependence $A \to A''$ is maximal.

Note that if A is an irreducible gap, then the whole interval (A, A'') consists of reducible gaps. Proposition 3.3.1 allows one to establish a one-to-one correspondence between irreducible gaps and maximal dependences.

As noted in [15, 16], the following relations hold for dependences (as well as for ID from [16]).

(1) If $A \to B$ and $C \to D$, then $A \cup C \to B \cup D$,

(2) If $A \to B$ and $B \to C$, then $A \to C$,

(3) If $A \to B$, then $A \cup C \to B$, where $A \cup C \to B$ can be not an ID.

In [16], relations (1)-(3) are called inference rules. The set $\mathcal{J}'$ of ID is said to be deduced from $\mathcal{J}$ (denoted by $\mathcal{J} \vdash \mathcal{J}'$) if $\mathcal{J}'$ can be obtained from $\mathcal{J}$ by a sequence of applications of rules (1)-(3).

Definition 3.3.4 ([16]). Let $\mathcal{J}$ be a set of ID. A dependence $i \in \mathcal{J}$ is called *redundant* if $\mathcal{J} \backslash \{i\} \vdash \{i\}$.

The following notion was introduced in [16] as a means of distinguishing minimal non-redundant subsets of ID that can generate the whole set of ID.

Definition 3.3.5 ([16]). Let $A \subseteq M$; then the set $\overline{A} = A \cup \{B'' | B \subset A, B \in \mathcal{J}, B'' \neq A''\}$ is called the *presaturation* (*pré-saturé*) of A with respect to the set of gaps $\mathcal{J}$.

It can be easily seen that the ID $A \to \overline{A}$ is redundant with respect to the set of ID $\{B \to B'' | B \subset A, B \in \mathcal{J}\}$.

Duquenne and Guigues [16] showed that the presaturation operation (taking A to $\overline{A}$) has the following properties:

(4) the inclusions $A \subseteq \overline{A} \subseteq A''$ hold for arbitrary $A \subseteq G$,

(5) if I is the set of irreducible gaps, then $\overline{A} = A \cup \{B'' | B \subset A, B \in I, B'' \neq A''\}$ holds for an arbitrary $A \subseteq G$,

(6) the operator $A \mapsto \overline{A}$ is a monotonic one, i.e., $B \subset C$ implies $\overline{(A \cup B)} \subseteq \overline{(A \cup C)}$,

(7) if $\overline{A} = A''$, then $A \to A''$ is a $\mathcal{J}$-redundant dependence, where $\mathcal{J}$ is the set of all MID. The converse does not hold, in general,

(8) if $A \in I$ and $A = \overline{A}$, then the dependence $A \to A''$ is not $\mathcal{J}$-redundant (where $\mathcal{J}$ is the set of all MID).

Definition 3.3.6 ([16]). Let $A \subset G$. We define a sequence of the form $A_0 = A,\ A_{k+1} = \overline{A_k}, k \in \mathbb{N}$. Then $\tilde{A} = \cup\{A_k | k \in \mathbb{N}\}$ is called *$\mathcal{J}$-saturation* of A.

It is easy to see that the tilde ($\tilde{\ }$) denotes a closure operation and that property (9) (see [16]) holds:

(9) If $\tilde{A} = A$, then the dependence $A \to A''$ is redundant with respect to the set of all MID.

Definition 3.3.7 ([16]). A set $A \subset M$ is called a *node of non-redundancy* (NR) if $A = \tilde{A}$ and $A \neq A''$.

Proposition 3.3.2 ([16]). *A set $A \subset M$ is an* NR *iff $A = \overline{A}$ and $A \neq A''$.*

Definition 3.3.8 ([16]). A set $A \subseteq M$ is called a *minimal* NR (MNR) if $(C'' = A'')$ and $(C = \tilde{C})$ imply $C = A$ for an arbitrary $C \subseteq A$.

Note that if A is a MNR and $B \subseteq A \subseteq B''$, then $A = \tilde{B}$. Denote by N the set of all NR and by N_0 the set of all MNR. For sets X, Y by $[X, Y]$ denote the set $\{Z | X \subseteq Z \subseteq Y\}$.

Proposition 3.3.3 ([16]). *Let B be a gap and $\mathcal{J}$ be a family of dependences; then*

(1) *the dependence $B \to B''$ is $\mathcal{J}$-redundant if $[B, B''] \cap N = \varnothing$;*

(2) *the dependence $B \to B''$ is $\mathcal{J} \backslash \mathcal{J}_A$-irredundant if $B \in \mathcal{J}_A$, where $A \in N_0$ and $\mathcal{J}_A = \{B \subseteq M | B \subseteq A \subseteq B''\} = \{B \subseteq M | \tilde{B} = A\}$.*

Define the equivalence relation $\sim$ for objects from M: $A, B \subseteq M, A \sim B \rightleftharpoons \tilde{A} = \tilde{B}$. Proposition 3.3.3 implies the following:

Theorem 3.3.4 ([16]). *Let $K = (G, M, I)$ be a context and $\mathcal{J}$ be a family of ID of this context. Then an arbitrary minimal subfamily $\mathcal{J}^m$, such that one can deduce $\mathcal{J}$ from it, has the form $\mathcal{J}^m = \{A \to A'' | A \in R\}$, where R is a representation system of equivalence classes of* MNR.

Wille [85] proposed a different means of deriving a minimal subset of the set of dependences.

Definition 3.3.9 ([85]). Let $K = (G, M, I)$ be a context. A set of attributes $A \subseteq M$ is called a *proper premise* (*echte Prämisse*) if $A \neq A'' \neq \cup\{(A \backslash \{n\}'' | n \in A\}$. A dependence $A \to B$ is called proper if A is a proper premise and $B = A'' \backslash \cup \{(A \backslash \{n\})'' | n \in A\}$.

If we have a list L of all proper dependences of the context $K = (G, M, I)$, then the set of all concepts of $\mathcal{B}(K)$ can be obtained by the use of the fact that $A'' = A \cup \{Y | (X \to Y) \in L, X \subseteq A\}$.

The following recursive definition of a dependence base was proposed in [77].

Definition 3.3.10. Let $K = (G, M, I)$ be a context. A set $P \subseteq M$ is called a *pseudointent* if $P'' = P$ and $Q'' \subseteq P$ for all pseudointents Q such that $Q \subset P$.

Then the set of all base dependences is $\{P \to (P'' \backslash P) | P$ is a pseudointent$\}$.

3.4. Dependences in Many-Valued contexts

In [31, 80, 94] various types of dependences in many-valued contexts are considered. It can be easily shown that certain 0-1 contexts can be introduced such that dependences therein are naturally related to dependences in initial many-valued contexts.

Definition 3.4.1. Let $K = (G, M, W, I)$ be a many-valued context (see Sec. 2.2). For $Y, Z \subseteq M$, Z is called *functionally dependent* on Y if, for all $g, h \in G$, $y(g) = y(h)$ for all $y \in Y$, implies $z(g) = z(h)$ for all $z \in Z$, i.e., there is a function $f\colon W^Y \to W^Z$ such that $f(y(g))_{y \in Y} = (z(g))_{z \in Z}$ for all $g \in G$.

Definition 3.4.2. Let $K = (G, M, W, I)$ be a many-valued context, then $K_f \rightleftharpoons (G \times G, M, I_f)$ is a context such that $(g, h) I_f m \Leftrightarrow m(g) = m(h)$ for $g, h \in G$, $m \in M$.

Proposiition 3.4.1 ([94]). *For a many-valued context $K = (G, M, W, I)$ $Y, Z \subseteq M$, Z is functionally dependent on Y in K iff $Y \to Z$ is a dependence in K_f, i.e., $Y' \subseteq Z'$ for K_f.*

Definition 3.4.3. Let $K = (G, M, W, I)$ be a many-valued context, $Y, Z \subseteq M$, $\delta \in R$, $\delta \geq 0$. Z is called *δ-dependent* on Y if for all $g, h \in G$, $|y(g) - y(H)| \leq \delta$, for all $y \in Y$, implies $|z(g) - z(h)| \leq \delta$ for all $z \in Z$.

Definition 3.4.4. Let $K = (G, M, W, I)$ be a many-valued context, then $K_\delta \rightleftharpoons (G \times G, M, I_\delta)$, where for $g, h \in G$, $m \in M$ $(g, h) I_\delta m$ iff $|m(g) - m(h)| \leq \delta$.

Proposition 3.4.2 ([94]). *For a many-valued context $K = (G, M, W, I)$, $Y, Z \subseteq M$, Z is δ-dependent on Y iff $Y \to Z$ is a dependence in K_f, i.e., $Y' \subseteq Z'$ in K_δ.*

Definition 3.4.5. Let $K = (G, M, W, I)$ be a many-valued context and $\leq$ be a relation of partial order on W. For $Y, Z \subseteq M$, Z is called ordinally dependent on Y if, for all $g, h \in G$, $y(g) \leq y(h)$ implies $z(g) \leq z(h)$ for all $z \in Z$.

Definition 3.4.6 Let $K = (G, M, W, I)$ be a many-valued context. Then $K_O \rightleftharpoons (G \times G, M, I_O)$, where, for $g, h \in G$, $m \in M$, $(g, h) I_O m \Longleftrightarrow m(g) \leq m(h)$.

Proposition 3.4.3 ([94]). *Z functionally depends on Y in $K = (G, M, W, I)$ for $Y, Z \subseteq M$ iff $Y \to Z$ is a dependence in K_f, i.e., $Y' \subseteq Z'$ in K_O.*

3.5. Dependences and Scaling

In papers [88, 33] the relationships between dependences in contexts and dependences in their scales are studied.

Definition 3.5.1 ([88]). Let $(K, \prod_{m\in M} S_m)$ be a scaled context, where $K = (G, M, I)$ and $R \subset M$ be a subset of attributes. Then the context $(G, N_R, \bar{J}_R)$ is called the *derived context* for the set R, where N_R is the set of all attributes $n \in N$ such that $(g_m)_{m\in M} J_n \iff (h_m)_{m\in M} Jn$ for all elements of $\times_{m\in M}$ such that $g_r = h_r$ for all $r \in R$ (thus, N_R is the set of attributes determined by components from R), $\bar{J}_R = J \cap G \times N_R$.

Several definitions of dependences in scaled contexts were considered in [88].

Definition 3.5.2 ([88]). Let $(K, \prod_{m\in M} S_m)$ be a scaled context. A set of attributes $Y \subset M$ *depends* on a set of attributes $X \subset M$, if every intent of the derived context for $(K, \prod_{m\in X\cup Y} S_m)$ is an intent of the derived context for $(K, \prod_{m\in X} S_m)$.

Definition 3.5.3 ([88]). A set of attributes of a scaled context $(K, \prod_{m\in M} S_m)$ *weakly depends* on the set attributes X if every intent of the context $(K, \prod_{m\in Y} S_m)$ is an intent of the context $(K, \prod_{m\in X} S_m)$.

Proposition 3.5.1 ([88]). *Let $(K, \prod_{m\in M} B_m)$ be a Boolean-scaled context of a complete many-valued context K, where K is a direct product of scales. Then, for every $X, Y \subseteq M$, the following conditions are equivalent:*

(1) *Y depends on X in $(K, \prod_{m\in M} B_m)$,*

(2) *Y weakly depends on X in $(K, \prod_{m\in M} B_m)$,*

(3) *Y functionally depends on X in $(K, \prod_{m\in M} B_m)$.*

Definition 3.5.4 ([88]). Let K be a field. Then the relation $\perp$ of the context $(K, K, \perp)$ is defined as $h \perp k \rightleftharpoons hk = 0$. The fusion of such contexts is defined as $\prod_{m\in M} S_m \rightleftharpoons (K^{|M|}, K^{|M|}, \perp)$, where $\vec{a} \perp \vec{b} \rightleftharpoons ab = 0$. This context is called a *linearly scaled context.*

Proposition 3.5.2 ([88]). *Let $(G, M, \prod)$ be a linearly scaled context with respect to the field K of a complete many-valued context. For arbitrary $X, Y \subseteq M$ the set Y depends on X in $(G, M, \prod)$ iff every tuple $(y(g))_{g\in G}$, where $y \in Y$, linearly depends on $\{(x(g)_{g\in G} | x \in X\}$ in $K^{|G|}$.*

Definition 3.5.5 ([33]). Let P be a partially ordered set with the order relation $\geq$; then the context $(P, P, \geq)$ is called an *ordinal scale* and is denoted by O_P.

Definition 3.5.6 ([33]). Let $K = (G, M, W, I)$ be a complete many-valued context with partial order $\leq$ on $W \times W$. Then $[x, y, z]$ denotes that $x \leq y \leq z$ or $x \geq y \geq z$, i.e., y lies between x and z. A set $Y \subseteq M$ *weakly depends* on $X \subseteq M$ if, for all $g, h \in G$ the fact $[x(g), x(h), x(k)]$ for all $x \in X$ implies $[y(g), y(h), y(k)]$ for all $y \in Y$.

Proposition 3.5.3 ([33]). *Let $K = (G, M, W, I)$ be a complete many-valued context scaled by the direct product $\times_{m\in M} O_P$ so that the values of each attribute $m \in K$ comprise the partially ordered set P_m. Then for $X, Y \subseteq M$ the following conditions are equivalent:*

(1) *Y ordinally depends on X in K,*

(2) *Y depends on X in $(K; \times_{m\in M} O_{P_m})$,*

(3) *Y weakly depends on X in $(K; \times_{m\in M} O_{P_m})$.*

Proposition 3.5.4 ([33]). *Let $K = (G, M, I)$ be a complete many-valued context with partially ordered set of attribute values M. Let $K_o = (G \times G, M, I_o)$ be a context such that $(g, h) I_o m \rightleftharpoons m(g) \leq m(h)$ and let $K_{io} = (G \times G \times G, M, I_{io})$ be a context such that $(g, h, k) I_{io} m \iff [m(g), m(h), m(k)]$. Then for $X, Y \subseteq M$:*

(1) *Y ordinally depends on X iff $X' \subseteq Y'$ for the context K_o;*

(2) *Y interordinally depends on X iff $X' \subseteq Y'$ for the context K_{io}.*

Definition 3.5.7 ([33]). Let K_1 and K_2 be contexts: $K_1 = (G, M_1, I_1)$, $K_2 = (G, M_2, I_2)$. The context $(G, M_1 \dot{\cup} M_2, I_1 \dot{\cup} I_2)$ is denoted by $K_1 | K_2$.

Proposition 3.5.5 ([33]). *Let K be a complete many-valued context (G, M, W, I) scaled by the apposition*

$$\times_{m\in M} O_{P_m} | \times_{m\in M} O_{P_m}$$

so that the values of each attribute $m \in K$ constitute the ordered set P_m. The following conditions are *equivalent for $X, Y \subseteq M$:*

(1) *Y interordinally depends on X in K,*

(2) *Y depends on X in $(K; \times_{m\in M} O_{P_m} | \times_{m\in M} O_{P_m})$,*

(3) *Y weakly depends on X in $(K; \times_{m\in M} O_{P_m} | \times_{m\in M} O_{P_m})$.*

3.6. Stability of Dependences

Definition 3.1 of the dependence on attributes of the form $X \to W$ assumes that the property W is caused by common properties of objects from X' that have the property W. All attributes that do not hold for these objects are considered implicitly inessential for the dependence $X \to W$. It is reasonable to consider that the greater the set of examples X', the more plausible the dependence $X \to W$. However, in a case of, say, the dependence $X_1 \to W_1$, the examples can be "too similar" as obtained, for instance, in a single series of experiments. In another case, say of the dependence $X_2 \to W_2$, the examples can be "more independent" and, therefore, can vary from each other considerably, except for the attributes from X. This means, in particular, that the second dependence can be obtained starting from a lesser number of examples, i.e., for a subcontext $(\underline{G}, M, \underline{I})$ of the initial context (G, M, I), $\underline{G} \subseteq G$, $\underline{I} = I \cap \underline{G} \times M$, since the independence of data allows one to separate faster and more reliably the essential data from the inessential data. This means, in turn, that the second dependence holds for a greater number of subcontexts of the context (G, M, I) induced by subsets of the set of objects G, i.e., this dependence is more "stable" with respect to the randomness of selecting data that correspond to the context (G, M, I). The fewer the number of hypotheses, the greater the average stability of the hypotheses.

The idea of stability was used for analysis of the plausibility of dependences of different nature, for example, in methods of nonparametric statistics, namely, in that of the jackknife and bootstrap methods [17]. In [49], the notion of the stability of JSM-hypotheses was introduced. Here we present the main results from [49] in terms of formal concept analysis.

Let $K = (G, M, I)$ be a context and $H : \langle X \to W\rangle$ be a dependence over the attributes of this context. We introduce the following notation:

$$\langle H\rangle_j = \{Y \subseteq G | Y \subset X', |Y| = k, Y' = X\},$$

$$\langle H\rangle_\Sigma = \bigcup_{j=2}^{n-1} \langle H\rangle_j,$$

$$\gamma(j, H) = |\langle H\rangle_j|, \ \gamma(\Sigma, H) = |\langle H\rangle_\Sigma|, \ n = |X'|.$$

In the cases where a fixed dependence is treated, we will omit the arguments j and H of the function γ, i.e., we will simply write γ_j or γ_Σ.

Definition 3.6.1 ([49]). Let $K = (G, M, I)$ be a context. Then for a dependence $X \to W$ of the context K *stability indices* are defined as follows:

(1) the stability index of the jth level $(2 \le j \le n-1)$ has the form

$$J_j = \frac{\gamma_j}{\binom{n}{j}},$$

(2) the integral stability index has the form

$$J_\Sigma = \frac{\gamma_\Sigma}{2^n - n - 2},$$

(3) the averaged stability index has the form

$$J_a = \frac{1}{n-2}\left(\sum_{j=2}^{n-1} J_j\right).$$

The stability indices of the dependence $X \to W$ are related to X in the same way as the sample variance (computed by the jackknife method, see [17]) is related to the sample mean in statistics.

The following property of stability indices follows from a simple property of monotone Boolean functions: the relative amount of units of a monotone Boolean function in the $(j+1)$st layer of a Boolean hypercube is greater than that in the jth layer.

Lemma 3.6.1 ([49]). *Let $K = (G, M, I)$ be a context. Then for a dependence $X \to W$ of the context K, the following inequalities hold: $J_2 \leq \ldots \leq J_{n-1}$.*

We will see now how the stability indices change as the set of objects G is updated by new objects.

After the arrival of k new objects the stability indices of dependences after will be supplied by the superscript k, e.g., J_j^k. We also set, for the sake of convenience,

$$J_n = 1,$$
$$J_j = 0 \text{ for } j \in Z \setminus \{2, \ldots, n\}.$$

Theorem 3.6.2 ([49]). *Let $K = (G, M, I)$ be a context. Let the set G be updated with k new objects, then the stability indices are within the following bounds:*

$$\underline{J}_j^k \leq J_j^k \leq \overline{J}_j^k,$$
$$\underline{J}_\Sigma^k \leq J_\Sigma^k \leq \overline{J}_\Sigma^k,$$

where

$$\underline{J}_j^k = \frac{1}{\binom{n+k}{j}} \left(\gamma_j + \binom{k}{1} \gamma_{j-1} + \ldots + \binom{k}{k-1} \gamma_{j-k+1} + \gamma_{j-k} \right),$$

$$\overline{J}_j^k = \frac{1}{\binom{n+k}{j}} \left(\gamma_j + \binom{n}{j-1} + \ldots + \binom{n+k-1}{j-1} \right),$$

$$\underline{J}_\Sigma^k = \frac{2^k \cdot \gamma_\Sigma + 2^k - 1}{2^{n+k} - (n+k+2)},$$

$$\overline{J}_\Sigma^k = \frac{\gamma_\Sigma + 2^n (n^k - 1) - k}{2^{n+k} - (n+k+2)}$$

Consider now the limits of the upper and lower bounds of stability indices as $k \to \infty$. The behavior of the lower bounds of the layered indices differs, namely, the indices of the higher levels tend to 1, while the indices of the lower levels tend to 0. In fact, by Theorem 3.6.2, we have

$$J_{n+k-1}^k \geq \frac{1}{\binom{n+k}{n+k-1}} \left(\gamma_{n-1} + \binom{k}{k-1} \right) \quad \text{and} \quad \lim_{k \to \infty} J_{n+k-1}^k = 1.$$

On the other hand,

$$\underline{J}_2^k = \frac{1}{\binom{n+k}{2}} \cdot \gamma_2 \text{ and } \lim_{k \to \infty} \underline{J}_2^k = 0.$$

The question concerning behavior of the lower bounds of the middle layer indices and of the averaged index remains open. The limit of the lower bound of the integral stability index is strictly positive and less than 1:

$$\lim_{k \to \infty} \underline{J}_\Sigma^k(k) = \frac{\gamma_\Sigma + 1}{2^n} \geq 0.$$

The upper bounds of the stability indices behave uniformly: they increase monotonically and tend to 1 as $k \to \infty$.

The analysis of the asymptotic behavior allows one to advance conjectures only about the integral index J_Σ: most likely, it will grow since it has small "decreasing reserve." Now, if we consider a set of objects X' that "support" the dependence $X \to W$ as having arisen from updating an initial set of size $r < |X'|$, then we can conclude that dependences with greater $|x'|$ are likely to have greater J_Σ than dependences with lesser $|X'|$. A "soft" dependence of J_Σ on the number of examples allows one to prefer the integral stability index as the most informative one. On the one hand, it reflects explicitly the stability of a dependence and, on the other hand, it reflects implicitly the number of objects that "support" the dependence, i.e., $|X'|$.

The idea of stability can be realized in different stability indices. In certain situations some of them are preferable to others. Consider the index J_F based on the stability of forecast. Let $K_+ = (G_+, M, I_+)$, $K_- = (G_-, M, I_-)$, $K_\tau = (K_\tau, M, I_\tau)$ be a positive, a negative, and an undefined context, relative to the property $W \not\subseteq M$, respectively (see Sec. 3.2). Let F_+, F_- be sets of all positive and negative forecasts obtained on the basis of dependences for K_+ and K_-. Then J_F is defined as the fraction of all subsets of the set $G_+ \cup G_-$ for which the sets of all forecasts coincide with the forecasts generated from the whole set of objects $G_+ \cup G_-$. Examples of other definitions can be found in [49].

3.7. Partial Dependences on Attributes

The notion of partial dependence was introduced in [59] (under the name of *implication partielle*). On the one hand, this notion generalizes the notion of dependence; on the other hand, it "inverts" the latter, since a dependence $A \to B$ corresponds to the partial dependence $B \xrightarrow{p} A$, where p is a measure of partial dependence.

Let $K = (G, M, I)$ be a context, $B \subseteq M$, and $|B|$ denote the number of objects that have the set of attributes B. Then the "probability" of B is defined as $P(B) = \frac{|B|}{|G|}$ and the "conditional probability" of the set of properties B_2 with respect to the set of attributes B_1 is defined as

$$P(B_2|B_1) = \begin{cases} \frac{|B_1 \cap B_2|}{|B_1|} & \text{if } |B_1| \neq 0, \\ 1 & \text{otherwise.} \end{cases}$$

Definition 3.7.1 ([59]). A *partial dependence* (PD) $B_1 \xrightarrow{p} B_2$ is a triple of the form $\langle B_1, B_2, p\rangle$, where $B_1, B_2 \subseteq M$, and $p = P(B_2|B_1)$.

Definition 3.7.2 ([59]). The set $\mathcal{J}(\mathfrak{B}(K)) = \{A \xrightarrow{p} B | A, B \subseteq M, p = P(B|A)\}$ is the set of PD for the lattice generated by the context $K = (G, M, I)$.

Theorem 3.7.1 ([59]). *Let M_1, M_2, M_3, $M_4 \subseteq M$ be some intents, where $M_1 \subseteq M_2 \subseteq M_4$, $M_1 \subseteq M_3 \subseteq M_4$. Then*

$$P(M_2|M_1)P(M_4|M_2) = P(M_3|M_1)P(M_4|M_3) = P(M_4|M_1).$$

The transitivity of a partial dependence (i.e., $A \xrightarrow{p} B,\ B \xrightarrow{p} C \Rightarrow A \xrightarrow{pq} C$) does not hold for intents that are not contained in each other as in Theorem 3.7.1.

Definition 3.7.3 ([59]). Let $\mathcal{J} \subseteq \{A \xrightarrow{p} B | A \subset B \subseteq M,\ p \in Q \cap [0,1]\}$ (where Q is the set of rational numbers) be a set of partial dependences. Then the context $K = (G, M, I)$ such that $\mathcal{J} \subseteq \mathcal{J}(\mathfrak{B}(K))$ (i.e., $(A \xrightarrow{p} B) \in \mathcal{J}$ implies $P(B|A) = p$ for the context K) is called a *realization* of $\mathcal{J}$.

Definition 3.7.4 ([59]). Two realizations (G, M, I) and (H, M, J) of the set of partial dependences are called *equivalent*

$$\frac{|\{g \in G | gIb \text{ for all } b \in B\}|}{|\{g \in H | gJb \text{ for all } b \in B\}|} = \text{const}$$

if for an arbitrary $B \subseteq M$.

The following theorem from [59] establishes a criterion of realization of sets of PD.

Theorem 3.7.2. *A set of* PD $\mathcal{J} \subseteq \{A \xrightarrow{p} B | A \subset B \subseteq M,\ p \in Q \cap [0,1]\}$ *can be realized iff the following system is solvable for variables $p_B^A \in Q \cap [0,1]$, $A \subset B \subseteq M$:*

$$\sum_{\{N \subseteq M | N \supseteq Q\}} p_N^S (-1)^{|N \setminus S|} \geq 0 \text{ for all } S \subseteq M, \tag{1}$$

$$p^A_B = 0 \;\Rightarrow\; p^D_N = 1 \text{ for all } A \subseteq B \subseteq N \subseteq M, \tag{2}$$

$$p^S_B = p^S_A \cdot p^A_B \text{ for all } S \subseteq A \subseteq B \subseteq M, \tag{3}$$

$$p^A_B = p \text{ if } (A \xrightarrow{p} B) \in \mathcal{J}, \tag{4}$$

where p is the variable that corresponds to $P(B|K)$.

The problem of the search for a minimal set of partial dependences that allows one to reconstruct the whole set of PD (in a way similar to that in the case of dependences, see [16, 77, 85] and Sec. 3.3) is stated in [59].

Definition 3.7.5 ([59]). A partial dependence $A \xrightarrow{p} B$ is *derived* from the set of PD $\mathcal{J}$ (denoted by $\mathcal{J} \vdash A \xrightarrow{p} B$) iff the system of equations from Theorem 3.7.2 for the set $\mathcal{J}$ has a unique solution such that $p^A_B = p$.

Definition 3.7.6 ([59]). The set $\mathcal{J}^{<1}\mathfrak{B}(K) = \{A \xrightarrow{p} B | A \subset B\}$, where A is a set of intents of the concepts from $\mathfrak{B}(K)$, is called a set of *proper* PD (PPD).

Definition 3.7.7 ([59]). A set of PPD $\mathcal{J} \subseteq \mathcal{J}^{<1}(\mathfrak{B}(K))$ is called a set of *generating elements* for $\mathcal{J}^{<1}(\mathfrak{B}(K))$ iff $\mathcal{J} \vdash \mathcal{J}^{<1}(\mathfrak{B}(K))$. A minimal-by-inclusion set of generating elements is called a *base*.

The following theorem on the size of a base holds.

Theorem 3.7.3 ([59]). *If $\mathcal{J} \subseteq \mathcal{J}^{<1}(\mathfrak{B}(K))$ is a base of $\mathcal{J}(\mathfrak{B}(K))$, then $|\mathcal{J}| \leq |\mathfrak{B}(K)| - 1$.*

The following theorem establishes the type of lattices, where the upper bound from Theorem 3.7.3 is sharp. Therefore, the construction of a minimal base does not always lead to the result desired, i.e., to a considerable reduction of the size of PPD.

Theorem 3.7.4 ([59]). *Let $\mathcal{L}$ be a finite lattice with the set $J(\mathcal{L})$ of $\wedge$-irreducible elements such that $|\mathcal{L}| - |J(\mathcal{L})|| \leq 2$. Then the relation $|\mathcal{J}| = |\mathcal{L}| - 1$ holds for an arbitrary base*

$$\mathcal{J} \subseteq \mathcal{J}^{<1}(\mathfrak{B}(K))$$

and an arbitrary context K such that $\mathfrak{B}(K) \cong \mathcal{L}$.

4. Algorithmic Problems of Concept Generation

4.1. Crucial Problems. Main Algorithms for Concept Generation

In this section we dwell upon the problems of the algorithmic generation of concepts. These problems are essential for applications, where algorithmic efficiency is one of the main issues.

It is easy to show that for some contexts the number of all concepts can be exponential with respect to the size of the context. Consider, e.g., the context $K = (G, G, \neq)$, where $|G| = n$ (i.e., the Boolean scale of dimension n, see Sec. 2.5.). The number of all concepts is $2^n - 1$. An additional difficulty is the intractability of the problem of determining the number of all concepts (Theorem 4.1.1.). Knowledge of this number could be helpful in effective resources allocation. In [74] the following definitions were introduce to capture the notion of a "hard" enumeration problem.

Definition 4.1.1 ([74]). Let a counting Turing machine (CTM) be a nondeterministic Turing machine (TM) (for the definition of TM see [34]) with an auxiliary output device that prints the number of accepting computations induced by the input in binary notation on a special tape. A CTM has polynomial-time complexity if the longest accepting computation induced by the set of all inputs of size n takes pol(n) steps, where pol(n) is some polynomial of n. #P is the class of functions that can be computed by counting Turing machines of polynomial-time complexity. A problem Π_1 from #P is #P-complete if an arbitrary problem Π_2 from #P can be polynomially reduced by Turing to Π_1 (i.e., if there is a polynomial-time algorithm that solves Π_2 using an oracle that outputs solutions for instances of Π_1 in unit time).

In the cases known so far, the counting problems that correspond to NP-complete problems are #P-complete. Some examples of #P-complete problems that correspond to polynomial-time decision problems are found in [75].

Theorem 4.1.1 ([47]). *The following problem "number of all concepts" is #P-complete:*

INPUT. Context $K = (G, M, I)$.

OUTPUT. The number of all concepts from $\mathfrak{B}(K)$.

The theorem makes us either use the most effective algorithms for the generation of concepts or generate the sets of the "most interesting" concepts. The second way is considered in more detail in Secs. 4.3 and 4.4. In this section we consider some algorithms for concept generation.

An estimate of the time and space complexities of algorithms that generate the set of all concepts must, of course, be a function of $|G| = n$ and $|M| = n$. Moreover, since the size of $\mathfrak{B}(K)$ can be exponential in $|G|$ and $|M|$, and the problem of determining $|\mathfrak{B}(K)|$ is #P-complete, $|\mathfrak{B}(K)| = H$ should be taken as a parameter of an algorithmic complexity estimate.

A review of several algorithms for the generation of the set of all concept lattices for a given context can be found in [36]. A comparative study of four algorithms presented there shows that the upper complexity bound of all algorithms is quadratic in the number of concepts generated, except for the algorithm of B. Ganter [27]. For the algorithm of Ganter this bound is linear. In [36] the results of an experimental study are presented (unfortunately, the author of [36] did not give a description of the input data used therein). The algorithm of Norris [62] proved to be the fastest in the experiments of the author of [36]; the algorithm of Ganter was a bit slower. The algorithm of Bordat [33] was two times slower than that of Norris, and the algorithm of Chein [12] was two times slower than that of Bordat. Below, we present two of the algorithms from [50] and the CbO algorithm from [52] similar to that of Ganter.

The Algorithm of Chein ([12]).

Suppose we are given a context $K = (G, M, I)$. The input of the algorithm is the set of all pairs of the form $(\{g\}, \{g\}')$, $g \in G$. This set is called the set of objects of the first level or L_1. Let the set $L_k = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ of the kth level be generated; then the set L_{k+1} can be generated in the following way. For all $i < j$ the set $Y_{ij} = Y_i \cap Y_j$ is computed.

If Y_{ij} is not contained in a pair (X, Y_{ij}) from L_{k+1}, then $(X_i \cap Y_j, Y_{ij})$ is put in L_{k+1}.

If $Y_{ij} = Y_i$, then (X_i, Y_i) is removed from L_k.

If $Y_{ij} = Y_j$, then (X_i, X_j) is removed from L_k.

The process is iterated until $|L_K| < 2$ at a step K. The union of L_k for all k: $1 \le k \le K$ comprises the set of all concepts.

The Algorithm of Norris ([62]).

Let the objects from G be numbered, i.e., $G = \{g_1, \ldots, g_m\}$. Suppose that $L_1 = \{(\{g_1\}, \{g_1\}')\}$ and the sets $L_1, \ldots, L_k$, $k < m$ are already constructed. Then the set L_{k+1} is constructed in the following way. Consider a pair $(\{g_{k+1}\}, \{g'_{k+1}\})$ and all objects from L_{k+1} of the form (X_i, Y_i). If $Y_i \subseteq \{g_{k+1}\}'$, then (X_i, Y_i) is replaced by $(X_i \cup \{g_{k+1}\}', Y_i)$. If $Y_i \not\subseteq \{g_{k+1}\}'$, but $Y_i \cap \{g_{k+1}\}' \neq \varnothing$, then the object $(X_i \cup \{g_{k+1}\}, Y_i \cap \{g_{k+1}\}')$ is added to L_k. Now, if $\{g_k\}' \subseteq Y_i$, then the pair $(\{g_{k+1}\}, \{g_{k+1}\}')$ is also added to L_k. The resulting set comprises L_{k+1}. The set $L(K)$ is obtained as L_m for $m = |G|$.

Close-by-One Algorithm ([52])

The algorithm called "Close-by-One" (CbO), which seems to be quite close to the algorithm of Ganter, was proposed in [52]. An analog of the CbO algorithm, which did not make use of lexicographical ordering, was proposed in [96] for the top-down strategy.

We assume that all objects from G are numbered, and therefore a set $X \subseteq G$ can be represented by a respectively ordered tuple. The numbering of objects from G induces lexicographical ordering of sets from $\mathcal{P}(G)$. For the sake of convenience, we can represent the process of constructing intersections as a top-down one, which generates some tree whose vertices correspond to concepts. During this process, the objects from G can be labeled or remain unlabeled in each vertex independently. The following procedure is based upon the "depth-first" strategy, though other strategies are possible as well. Y denotes the extent of a current concept.

Step 0. There is only one root vertex where all objects are unlabeled, $Y := \varnothing$.

Step 1. The current vertex corresponds to the concept with the extent Y. The first unlabeled element of G, say X_i, is taken, $(Y \cup \{X_i\})'$ and $(Y \cup \{X_i\})''$ are computed. (In doing so we say that Y is closed by X_i.) Thus, a new vertex that corresponds to $(Y \cup \{X_i\})''$ is generated and connected to the vertex associated with Y.

Step 2. If $(Y \cup \{X_i\})'$ contains objects with numbers less than those of the objects from Y or the number of X_i, then the concept with the extent $(Y \cup \{X_i\})''$ has already been generated. All elements of G are labeled at the vertex $(Y \cup \{X_i\})''$ (thus, the branch will not be extended). If $(Y \cup \{X_i\})''$ has not yet been generated, we label additionally the element X_i at the vertex Y and all elements of $(Y \cup \{X_i\})''$ at the vertex $(Y \cup \{X_i\})''$.

Step 3. If all elements of G are labeled at $(Y \cup \{X_i\})''$, we go to Step 4. Otherwise, $Y := (Y \cup \{X_i\})''$, and we return to Step 1.

Step 4. We backtrack the tree upwards to the nearest vertex with unlabeled elements of G. If such a vertex exists and corresponds, say, to the object Z, then $Y := Z$ and we have to go to step 1. If such a vertex does not exist, then this means that all concepts have been generated and the algorithm halts.

Consider the top-down strategy in the case where $M = \{a_1, \ldots, a_n\}$, $G = \{X_1, \ldots, X_n, X_{n+1}\}$, and $X_i' = M \backslash \{a_i\}, 1 < i < n$, $X_{n+1}' = \{a_n\}$ (recall that the top-down and the bottom-up directions are specified with respect to the order on concepts). In the case of the top-down strategy, the process begins with the generation of the concept $\langle h, \mathrm{Par}(h) \rangle$ with the greatest (by inclusion) intent and proceeds then to concepts with smaller intents. Then the intent $\langle \{a_n\}, \{X_1, \ldots, X_{n-1}, X_{n+1}\} \rangle$ can be generated from an arbitrarily greater concept $\langle h, \mathrm{Par}(h) \rangle$ as the closure $(h \cup \{X_{n+1}\})'$. Since there are $2^n - 1$ concepts of this kind, the number of ways in which $\langle \{a_n\}, \{X_1, \ldots, X_n\} \rangle$ can be generated is the same.

Notwithstanding the pessimistic implications of the consideration above, we can state that the number of concepts generated exponentially many times and the number of all concept generations are not really great. This is substantiated by the following theorem, which is a particular case of the theorem from [52] on the algorithm generating all elements of a semilattice from a given set of generators.

Theorem 4.1.2. *Let $K = (G, M, I)$ be a context. Then the set of all elements of the lattice $\mathfrak{B}(K)$ can be generated with the use of $O(mnH)$ of space and $O((m+n)nH)$ of time.*

The MI algorithm from [96] can be obtained from the CbO algorithm by changing the provisions of Step 1, so that at each vertex only those X_i are chosen for which $(Y \cup \{X_i\})''$ are maximal by inclusion among all $(Y \cup \{X_j\})''$. These are the only vertices, where the tree is generated further. The MI algorithm was proposed by O. M. Anshakov and K. P. Khazanovskii in [96] for the bottom-up strategy as the "Minimal Intersection" algorithm.

The upper bound of the memory needed for the MI algorithm is also linear with respect to the number of concepts. In fact, the space required is less than that required by the CbO algorithm because only minimal (maximal) intersections are stored. The time complexity of the MI algorithm is worse because testing the uniqueness of a concept generation (i.e., verifying whether it was not generated before) cannot be accomplished in the way it was done in the CbO algorithm (Step 2). The uniqueness of a concept generation is tested with the use of the MI algorithm by comparing the concept generated to all concepts generated earlier. This leads to $O(K^2)$ time complexity of the MI algorithm.

The algorithm of Bordat [9] is similar to the MI algorithm in computing at each step for a concept constructed the least concept that majorizes it (in the sense of partial order on concepts). The advantage of the algorithm of Bordat consists in the quick test for uniqueness of a current intersection (i.e., in veryifying whether it corresponds to a concept already generated). This test is accomplished by means of the "trie" data structure, which allows a logarithmic-time uniqueness test and/or insertion of a new concept in the data structure. Thus, the time complexity of the algorithm of Bordat is linear with respect to the size of the concept lattice generated, as in the case of the Ganter algorithm and that of CbO. Moreover, the "trie" structure allows the algorithm from [9] to obtain the Hasse diagram of a concept lattice as a by-product. However, the creation of the "trie" structure is essentially a serial one (since it demands the solution of hard sequencing problems in realizing the evolution of the "trie" for parallel processors), unlike the algorithms from [52, 531] and [27], where tests for uniqueness can be accomplished locally for every new intersection generated.

In [96, 25], other algorithms can be found that construct the set of all concepts of a given context. However, unlike the algorithms from [52, 53, 9, 27], these algorithms have complexity estimates worse than linear (with respect to the size of the set of concepts generated).

4.2. Complexity of Some Problems of Enumeration and Decision Concerning Concepts

In the case where dependences are sought for sets of attributes, the pessimistic result of Theorem 4.1.1 warrants a statement of the problem of generating a subset of the most interesting dependences.

Assume that we study dependences of the form $Y \to W$ with respect to a fixed property W, which does not contain the set of attributes M, and all objects from G are partitioned into sets of positive G_+, negative G_- and undefined G_τ examples with respect to the property W (see Sec. 3.2.) Thus, dependences of the type $Y \to W$, $Y \subseteq M$ are sought for the context $(G_+ \cup G_-, M \cup \{W\}, I_+ \cup I_- \cup G \times \{W\})$. Then dependences of the form $X \to W$ such that X is minimal by inclusion among all Y such that $Y \to W$ (these dependences are called minimal) can be considered the most interesting ones, since they are "supported" by a greater number of examples than dependences with greater antecedents (X' is maximal among those Y': $Y' \to W$). At the same time, minimal dependences are more "decisive" or more informative, since they can lead to more forecasts (see Sec. 3.2).

In the case where there are no negative examples, there can be no more than $|M|$ minimal dependences, and they can be generated in time which is no more than cubic in $|M|$ (see [95, 96]). However, in the general case, where $G_+ \neq \varnothing, G_- \neq \varnothing$, Theorem 4.2.1 leaves us no hope for fast generation of all minimal dependences (unless P = #P).

Theorem 4.2.1. *The following problem is #P-complete.*

INPUT. Contexts $K_+ = (G_+, M, I_+)$, $K_- = (G_-, M, I_-)$, $K_{+-} = (G_+ \cup G_-, M \cup \{W\}, I_+ \cup I_- \cup G_+ \times \{W\}$.

OUTPUT. The number of minimal dependences of the context K_{+-}.

The #P-completeness of the problem of generation of minimal dependences in the particular case of the context K_{+-} implies the #P-completeness of the problem concerning the number of minimal dependences with W: $|W| > 1$.

When W is fixed, the following functionals, depending on the sizes of the intent and extent of a dependence $X \to W$, can be proposed as measures of the "quality" of a hypothesis:

(1) $|X|$,

(2) $|X'|$,

(3) $|X| + |X'|$,

(4) $q|X| + |X'|$, $0 < q < 1$,

(5) $|X| + q|X'|$, $0 < q < 1$,

(6) $|X| \cdot |X'|$.

Generation of dependences satisfying conditions with the above functionals is similar to the inductive biases proposed in various settings of machine learning (e.g., in the INDUCE system [60] or in the GUHA-method [66]).

Obviously, Theorem 4.1.1 implies that the problems of determining the numbers of dependences such that $f \leq C$, $f \geq C$ (where f is one of the functionals above and C is a parameter) are #P-complete. At the same time, the results concerning the corresponding decision problems are not always NP-complete, even when sets of positive and negative examples are both not empty.

In Table 2, we present results from [47, 51, 53, 54, 95] on the complexity of decision problems concerning concepts with restrictions on the sizes of the intents and extents.

Table 2

	$\leq$	$=$	$\geq$
$\|i\|$	P	NP	P
$\|e\|$	P	NP	P
$\|e\| + \|i\|$	NP	NP	P

Here, i and e denote the intent and the extent of a concept, respectively; P symbolizes that there exists a polynomial algorithm, and NP denotes NP-completeness of the problem. For instance, the upper left element of the table means that the problem "does there exist a concept such that $|i| \leq C$?" can be solved by a polynomial algorithm. The element in the bottom line and the middle column is indicative of the fact that the problem "does there exist a concept such that $|e| + |i| = C$" is an NP-complete one.

Results from [51] on decision problems concerning dependences of the form $X \rightarrow W$ for a context $K_{+-} = (G_+ \cup G_-, M \cup \{W\}, I_+ \cup I_- \cup G_+ \times \{W\}$ are presented in Table 3.

Table 3

	$\leq$	$=$	$\geq$
$\|X\|$	NP	NP	P
$\|X'\|$	P	NP	NP
$\|X\| + \|X'\|$	NP	NP	NP

For example, the element in the third column and the third row of the table denotes that the decision problem concerning dependence of the form $X \rightarrow W$, where $|X| + |X'| = C$, is NP-complete. It is obvious that the NP-completeness of a problem from Table 2 implies NP-completeness of the corresponding problem from Table 3 (since a context of the form K_{+-} degenerates into a context of the form K_+ when $G_- = \varnothing$).

The lower row of Table 2 may seem paradoxical; in fact, for positive examples only the problem of generating a minimal hypothesis is very unlikely to be polynomial-time solvable, whereas there exists a polynomial algorithm for the search for a maximal hypothesis. Graph interpretation of hypotheses may clarify such a situation: as shown in [47, 54] a hypothesis maximal in an $|X| + |X'|$ functional corresponds to a maximal-size complete bipartite subgraph of a bipartite graph. Such a subgraph can be found in polynomial time, e.g., after reducing it to the maximal matching problem [54]. The NP-completeness of the problem of the minimal (with respect to the functional $|X| + |X'|$) hypothesis was proved [50] by reducing it to the problem of the inclusion-maximal matching of minimal size [34].

4.3. Complexity of the Forecast (Classification) Problems

Here, we present some results concerning computational problems of generating forecasts (classification) for objects from G_τ of a context $K = (G_\tau, M, I_\tau)$ in the sense of the definitions from Sec. 3.2. In the general case, the problem of the existence of a hypothesis in favor of a positive forecast is intractable. To be more exact, the following theorem holds.

Theorem 4.3.1 ([51]). *The following problem is* NP*-complete*:

INSTANCE. Input data: contexts $K_{+-} = (G_+ \cup G_-, M \cup \{W\}, I_+ \cup I_- \cup I_+\{W\})$, $K_\tau = (G_\tau, M, I_\tau)$, and an object $g_\tau \in G_\tau$.

QUESTION. Does there exist a hypothesis (e_+, i_+) concerning the property W and satisfying the "counterexample forbidding" condition in favor of a positive forecast for g_τ, i.e., $i_+ \rightarrow W$ is a dependence for the context $(G_+ \cup G_-, M \cup \{W\}, I_+ \cup I_- \cup G_+\{W\})$ and $i_+ \subseteq \{g_\tau\}'$.

In the case where $G_- = \varnothing$, the problem has a trivial algorithm running in time polynomial with respect to the input size. This algorithm computes $\{m\}^{I_+}$ for every $m \in M$. If there exists $m \in M$ such that $\{m\}^{I_+I_+} \subseteq \{g_\tau\}^{I_\tau}$, then the forecast for g_τ will be positive; otherwise it will be undefined.

A forecast problem can be solved in polynomial time for one of the following cases:

$$M = \{g_\tau\}' \ (V_1 = \varnothing),$$
$$G_+ = \{g_\tau\}' \ (V_2 = V_3),$$
$$G_- = \varnothing \ (V_4 = \varnothing),$$

(where V_1, V_2, V_3, V_4 are sets of vertices of the graph from Definition 3.2.2), i.e., when the tetrapartite graph from Definition 3.2.1 degenerates into a tripartite one. The problem is polynomial-time solvable also in the case where $|\{g_\tau\}'|$ is constant (see [51]).

4.4. Algorithmic Complexity of the Problem of Computing Stability Indices

Theorem 4.4.1 ([49]). *Let $K = (G, M, I)$ be a context. Then the problem of determining the stability index J_Σ of a dependence $X \to W$ of the context K* (see Sec. 3.6) *as well as the problem of determining J_j (where $2 \le j \le |X'|$) is #P-complete.*

Theorem 4.4.2 ([49]). The *stability indices $J_j = \frac{\gamma_j}{\binom{n}{j}}$ for $1 \le k \le n-1$ of a dependence $X \to W$ can be computed in time linear with respect to $\max_{2 \le i \le j} i\gamma_i$. The integral stability index $J_\Sigma = \frac{\gamma_\Sigma}{2^n - n - 2}$ of a dependence $X \to W$ can be computed in time linear with respect to γ_Σ (where $n = |X'|$).*

This result, together with the result concerning the #P-completeness of the determination of stability indices, means that the algorithm presented in [49] for the proof of Theorem 4.6.2 is optimal modulo some polynomial of n (see [75]).

We can propose computation of the upper and lower bounds of the integral J_Σ and average J_a stability indices on the basis of the following:

Theorem 4.4.3 ([49]). *The following inequalities hold for the integral and average stability indices of a dependence $X \to W$ (where $n = |X|$, $2 \le k, r \le n-1$):*

$$\frac{\gamma_1 + \ldots + \gamma_k}{\binom{n}{2} + \ldots + \binom{n}{k}} \le J_\Sigma \le \frac{\gamma_{n-r} + \ldots + \gamma_{n-1}}{\binom{n}{n-r} + \ldots + \binom{n}{n-1}},$$

$$\frac{1}{k}\left(\frac{\gamma_2}{\binom{n}{2}} + \ldots + \frac{\gamma_k}{\binom{n}{k}} \le J_a\right) \le \frac{1}{r}\left(\frac{\gamma_{n-r}}{\binom{n}{n-r}} + \ldots + \frac{\gamma_{n-1}}{\binom{n}{n-1}}\right).$$

Thus, these algorithms for approximate computation of J_Σ and J_a have polynomial time complexity with respect to the numbers of the levels up to which the summation is carried out (i.e., k and r).

5. Miscellaneous

5.1. The Origin of Contexts

In [73] the following model of the context origin was studied. Let there be given a set V, whose elements are called preconcepts. Some objects from V are interpreted as objects (the set of objects is denoted by $\mathcal{G}$), other objects from V are interpreted as attributes (the set of objects is denoted by $\mathcal{M}$). Now, the relation $\Delta \subseteq \mathcal{G} \times \mathcal{M}$ is defined in the following way: for any $X \in \mathcal{G}, Y \in \mathcal{M}$, $X \Delta Y \rightleftharpoons X \cap Y \ne \varnothing$. Then the context is defined as a triple $(\mathcal{G}, \mathcal{M}, \Delta)$.

By way of example, Stahl and Wille [73] describe a study of the use of the conjunction "et" ("and", French) in the book *Le Petit Prince* by A. de Saint-Exupéry. A list consisting of 206 occurrences of "et" was considered as the set V of preconcepts. 38 subsets of V are formed as a result of finding out which occurrences of "et" can be replaced by paraphrases such as "alors," "de même que," "c'est pourquoi," etc.; these subsets are interpreted as objects. Meanings of the conjunctions are grasped by 14 properties such as "additive," "comparative," "temporal," etc., whereby subsets of V come into existence, which represent attributes.

Stahl and Wille [73] explore how a given context can be derived from an appropriate set of preconcepts.

Definition 5.1.1 ([73]). Let there be given a context $K = (G, M, I)$ and a set V. A pair of injective mappings $\alpha: G \mapsto \mathcal{P}(V)$ and $\beta: M \mapsto \mathcal{P}(V)$ is called a *representation* of a context K on the set V if gIm iff $\alpha g \cap \beta m \neq \varnothing$ (i.e., K is isomorphic to $(\alpha G, \beta M, \Delta)$.

Definition 5.1.2 ([73]). A pair (A, B) is called a *preconcept* of a context $K = (G, M, I)$ if $A \subseteq G, B \subseteq M, A \subseteq B', B \subseteq A'$.

Definition 5.1.3 ([73]). A *predomain* of a context $K = (G, M, I)$ is a set $\mathcal{H}$ of preconcepts of K that satisfies the following conditions:

(1) for $g, h \in G$, $g \neq h$, there exists a pair $(A, B) \in \mathcal{H}$ such that

$$|\{g, h\} \cap A = 1,$$

(2) for $m, n \in M, m \neq n$, there exists a pair $(A, B) \in \mathcal{H}$ such that

$$|\{m, n\} \cap B| = 1,$$

(3) for gIm there exists a pair (A, B) such that

$$g \in A \text{ and } m \in B.$$

The following theorem establishes a relationship between the predomains and representations of a context.

Theorem 5.1.1 ([73]). *Let $\mathcal{H}$ be a predomain of a context $K = (G, M, I)$. For $g \in G$ and $m \in M$,*

$$\mathcal{H}_g = \{(A, B) \in \mathcal{H} | g \in A\},$$

$$\mathcal{H}_m = \{(A, B) \in \mathcal{H} | m \in B\}.$$

Then the mappings $g \mapsto \mathcal{H}_g$ and $m \mapsto \mathcal{H}_m$ yield a representation of K on $\mathcal{H}$. Conversely, let (α, β) be any representation of K on the set V. For $v \in V$, $A_v = \{g \in G | v \in \alpha g\}$ and $B_v = \{m \in M | v \in \beta m\}$. Then

$$\mathcal{H} = \{(A_v, B_v) | v \in V\}$$

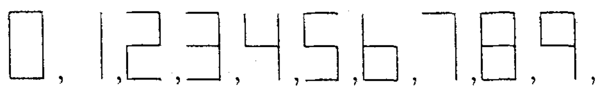
is a predomain of K such that

$$\mathcal{H}_g = \{(A_v, B_v) | v \in \alpha g\},$$
$$\mathcal{H}_m = \{(A_v, B_v) | v \in \beta m\}.$$

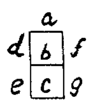
An example of a predomain is the set $\{(\{g\}, \{g\}') | g \in G\} \cup \{(\{m\}, \{m\}') | m \in M\}$ and the set $\{(\{g\}, \{m\}) | gIm\}$, where $\{g\}' \neq \varnothing$ and $\{m\}' = \varnothing$ for all $g \in G$, $m \in M$. It is natural that the set of all concepts of a context is also a predomain.

It is useful to have a means for finding a predomain of the least size for a context. However, there are no hopes for finding a fast algorithm for computing such a predomain, since the problem is NP-complete even in the case where preconcepts are sought among the set of concepts.

The following example of a search for the least predomain is considered in [73]. Let the digit descriptions



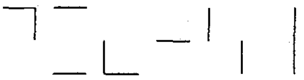
be objects, and the seven line segments of the figure



be attributes (with labels standing at the corresponding segments). The relation I between objects from the set G and attributes from the set M are given in a natural way: for $g \in G$ and $m \in M$, gIm iff the description of the object g has the attribute m. K has a natural representation on the set of seven line segments. The question as to whether K can have representations of less size is answered positively. In fact, consider the predomain which consists of the following six concepts:

$$A = (\{0,2,3,7,8,9\},\{a,f\}),\quad B = (\{0,2,3,5,8\},\{a,c\}),$$
$$C = (\{0,2,6,8\},\{c,e\}),\quad D = (\{2,3,4,5,6,8,9\},\{b\}),$$
$$E = (\{0,4,5,6,8,9\},\{d,g\}),\quad F = (\{0,1,3,4,7,9\},\{f,g\}).$$

The intents of these concepts correspond to the following six figures:



A B C D E F

which can give all ten digits.

5.2. Concept Analysis of Paired Comparisons

In [57] the means of formal concept analysis were used in the study of preference relations. Let A be a set of alternatives and $R \subseteq A \times A$ a preference relation: iRj means that the subject prefers alternative j to alternative i. A pair (A, R) is called a tournament if for any pair $i, j \in A$ either iRj or jRi. Luksch and Wille [59] studied subsets of A, whose elements are in the same preference relation with any alternative not in the set. These subsets can be interpreted as clusters given by the relation R. Here are precise definitions.

Definition 5.2.1 ([57]). A set $S: S \subseteq A$ is called a *superalternative* of a tournament (A, R) if for every $a \in A\backslash S$ either aRs for all $s \in S$ or sRa for all $s \in S$.

Proposition 5.2.1 ([57]). *A nonempty subset S of a tournament (A, R) is a superalternative iff*

$$S = \{a \in A \mid \wedge\,\mu S \wedge \bigvee \gamma S < \gamma a \leq \bigvee \gamma S\} =$$
$$= \{a \in A \mid \bigwedge \mu S \leq \mu a < \bigwedge \mu S \vee \bigvee \gamma S\}$$

and

$$A\backslash S = \{a \in A \mid \gamma a \leq \wedge \mu S \text{ or } \mu a \geq \vee \gamma S\}.$$

Here, $\bigwedge \mu S$ stays for $\bigwedge_{s\in S} \mu S$, analogously for $\bigvee$ and γ. For the definitions of μ and γ see Sec. 2.1.

Proposition 5.2.2 ([57]). *Let $S_1, \ldots, S_n$ be pairwise disjoint alternatives of a tournament (A, R) and let $(\underline{A}, \underline{R})$ be the tournament with $\underline{A} = \{S_1, \ldots, S_n\} \cup \{\{a\} \mid a \in A\backslash(\cup_{i=1}^{n} S_i)$ and for $S, T \in \underline{A}$, $S\underline{R}T$ iff sRt for all $s \in S$ and $t \in T$. Then $\mathfrak{B}(\underline{A}, \underline{A}, \underline{R})$ is isomorphic to the sublattice of $\mathfrak{B}(A, A, R)$ consisting of all concepts (X, Y) with $S_i \subseteq X$ or $S_i \cap X = 0$ for each i.*

Proposition 5.2.3 ([57]). *Let S be a superalternative of a tournament (A, R). Then the mapping $(X, Y) \mapsto (X \cap S, Y \cap S)$ defines an isomorphism from each of the intervals $[\bigwedge \mu S \wedge \bigvee \gamma S, \bigvee \gamma S]$ and $[\bigwedge \mu S, \bigwedge \mu S \vee \bigvee \gamma S]$ of $\mathfrak{B}(A, A, R)$ onto the concept lattice $\mathfrak{B}(S, S, R \cap S \times S)$.*

Proposition 5.2.4 ([57]). *Let S and T be superalternatives of a tournament (A, R) for which $S \cap T$, $S\backslash T$, and $T\backslash S$ are not empty. Then $S \cup T$, $S \cap T$, $S\backslash T$, and $T\backslash S$ are superalternatives of (A, R) and either sRt for all $s \in S$ and $t \in T$, or tRs for all $t \in T$ and $s \in S$.*

In [57], a sketch of an algorithm for computing superalternatives of a given tournament is considered. Suppose that a superalternative S has been found; then other superalternatives are sought for the tournaments $(S, R \cap S \times S)$ and $(A/S, R/S)$, where $A/S = A\backslash S \cup \{S\}$, $R/S = \{R \cap (A\backslash S) \times (A\backslash S)) \cup \{(a, S) \mid aRs$ for $s \in S\} \cup \{(S, a) \mid sRa$ for $s \in S\}$. Since the set A is finite, the algorithm terminates after finitely many steps and yields a sequence $S_1, \ldots, S_n$ of indecomposable superalternatives, which is called a decomposition into indecomposables of the underlying tournament.

Proposition 5.2.5 ([57]). *Two arbitrary decompositions into indecomposable of a tournament have equal lengths and consist of pairwise isomorphic superalternatives.*

A decomposition of a tournament (A, R) can be used for construction of the Hasse diagram of the lattice $\mathfrak{B}(A, A, R)$. The program mentioned in [57] has a library of standard diagrams of indecomposable tournaments of sizes not exceeding a certain value. Having constructed a decomposition of a tournament, the program constructs the diagram of $\mathfrak{B}(A, A, R)$ in the inverse order.

5.3. Characterization of "Good" Contexts

Novotný and Pawlak [63-65] introduced the notion of a "good" context whose definition was based on the notions of a "black box" and a "rough top equality." Since we will not deal with these two notions in our paper, we will present a definition of a good context that was given in [14] on the basis of Theorem 4.6 from [64]. The "goodness" of a context means that each attribute corresponds to a certain set of objects, and, therefore, the representation by the context is adequate.

Definition 5.3.1 ([14]). Let $K = (G, M, I)$ be a context and

$$\theta = \{(x, y) | (x, y) \in M \times M: \{x\}' = \{y\}'\}$$

be an equivalence relation defined on pairs of attributes. Then for $x \in M$, $[x]_\theta \rightleftharpoons \{y | (x, y) \in \theta\}$ and $[X]_\theta \rightleftharpoons \cup_{x \in X} [x]_\theta$ for $X \subseteq M$. The context $K = (G, M, I)$ is called *good* if $X' = Y'$ implies $[X]_\theta = [Y]_\theta$ for any $X, Y \subseteq M$.

It was shown in [14] that the relation θ from Definition 5.3.1 is a congruence on the semilattice $(\mathcal{P}(M), \cup)$.

A context such that $(x, y) \in \theta$ iff $x = y$ is called *reduced* in [14]. Thus, a reduced good context is a context where each attribute is in one-to-one correspondence with a set of objects.

Theorem 5.3.1 ([14]). *Let $K = (G, M, I)$ be a context; then the following conditions are equivalent:*

(1) *K is a reduced good context,*

(2) *the mapping $t: \mathcal{P}(M) \mapsto \mathcal{P}(G)$ taking an arbitrary set $Y \subseteq M$ into Y' is injective,*

(3) *$\{m\}' \not\supseteq (M \backslash \{m\})'$ holds for an arbitrary $m \in M$.*

(4) *A set of attributes $X \subseteq M$ depends on a set of attributes $Y \subseteq M$ (i.e., $Y \to X$) iff $X \subseteq Y$.*

Corollary. *Let $K = (G, M, I)$ be a reduced good context. Then for any sets $H \supseteq G$, $N \subseteq M$, $N \neq \varnothing$, the following assertions hold:*

(1) *the subcontext $(G, M, I \cap G \times N)$ of the context (G, M, I) is a good reduced context;*

(2) *supercontext (H, M, J), where $I \cap G \times M = J$, is a good reduced context.*

Definition 5.3.2 ([14]). Let S be a nonempty set and k a natural number. A family of sets $X_i \subseteq S, i = 1, \ldots, k$, is an irredundant subset system in S whenever $X_i \not\subseteq \cup_{j=1,\ j \neq i}^{k} X_j$ for every $i = 1, \ldots, k$.

It is easy to see that for every irredundant subset system we can introduce a representative set, i.e., a subset $\{x_i, i \leq k\}$ of S such that $x_i \in X_j$ iff $i = j$ for $i, j \in \{1, \ldots, k\}$.

Theorem 5.3.2 ([14]). *Let $K = (G, M, I)$ be a reduced good context, $\{X_i, i \leq n\}$ an irredundant subset system in M, and $\{y_i, i \leq n\}$ an arbitrary n-element set. Then the context $C = (G, \{y_i, i \leq n\}, J)$ defined by the relation $t_C(\{y_i\}) = t_K(X_i)$, $i = 1, \ldots, n$, is a reduced good context.*

Theorem 5.3.3 ([14]). *Let $K = (G, M, I)$ be a reduced good context. Then there exists a uniquely determined irredundant subset system $\{X_m, m \in M\}$ in G such that $G \backslash M = \{m\}'$ for any $m \in M$.*

Corollary. *For a context $K = (G, M, I)$ we have*

(1) *if $|G| < |M|$, then K is not a reduced good context,*

(2) *if $|G| = |M|$, then K is a reduced good context iff $K \cong (G, G, \neq)$.*

Theorem 5.3.4 ([14]). *Let $K = (G, M, I)$ be a context. Then the following assertions are equivalent:*

(1) *K is a reduced good context,*

(2) *the mapping $s: \mathcal{P}(G) \mapsto \mathcal{P}(M)$ taking a set $X \subseteq G$ to X' is surjective,*

(3) *there is a subset $H \subseteq G$ such that $(H, M, I \cap H \times M) = (M, M, \neq)$.*

It is obvious that the dual assertion holds too.

Corollary. *Let $K = (G, M, I)$ be a context. Then the following assertions are equivalent:*

(1) *the mappings s, t are injective,*

(2) *the mappings* s, t *are surjective,*

(3) *the mapping* s *is bijective,*

(4) *the mapping* t *is bijective,*

(5) $|G| = |M|$ *and* $\mathfrak{B}(K) \cong 2^{|G|}$ $(\cong 2^{|M|})$,

(6) $|G| = |M|$ *and* $K \cong (G, G, \neq)$ $(\cong (M, M, \neq)$,

(7) *the mappings* s *and* t *are mutually inverse, i.e.,* $s \circ t = 1_{\mathcal{P}(G)}$ *and* $t \circ s = 1_{\mathcal{P}(M)}$.

We present one more similar result from [14].

Theorem 5.3.5 ([14]). *Let* $K = (G, M, I)$ *be a context and* n *a positive integer. Then the following conditions are equivalent:*

(1) $\mathfrak{B}(K) \cong \mathbf{2}^n$,

(2) *there is a set* $N \subseteq M$ *such that*

(a) $|N| = n$,

(b) $(G, N, I \cap G \times N)$ *is a reduced good context,*

(c) $\{h\}' = P'$ *for any* $h \in M \backslash N$, *where* P *is a uniquely determined subset of* N,

(2′) *there is a set* $H \subseteq G$ *such that*

(a) $|H| = n$,

(b) $(M, H, I \cap M \times H)$ *is a reduced good context,*

(c) $\{f\}' = F'$ *for any* $f \in G \backslash H$, *where* F *is a uniquely determined subset of* G,

(3) *there exist sets* $N \subseteq M$ *and* $H \subseteq G$ *such that*

(a) $|N| = |H| = n$,

(b) $(H, N, I \cap H \times N) = (N, N, \neq)$,

(c) *for any* $h \in M \backslash N$, $f \in G \backslash H$ *there are uniquely determined sets* $M \subseteq N$, $F \subseteq H$ *such that* $\{h\}' = M'$, $\{f\}' = F'$.

5.4. Tolerance on the Set of Objects (G)

The tolerance relation defined on objects from the set G of a context $K = (G, M, I)$ was studied in [37-39]. For $X, Y \subset G$ the tolerance was defined as $X \Theta Y = X' \Theta Y' \neq \varnothing$. In particular, the following problem was studied: what sets M allow one to define tolerance in such a way. Various generalizations of tolerance were studied. Contrary to [86], the tolerance is not extended to the set of concepts of the context K. A condition of coincidence of the sets of all intents with the family of blocks of tolerance on G was obtained. This condition is shown in the following theorem from [39].

Theorem 5.4.1. *Let* $K = (G, M, I)$ *be a context, and* $C = \{(e_1, i_1), \ldots, (e_n, i_n)\}$ *be the set of all concepts for* K. *The set of all intents* $E = \{e_1, \ldots, e_n\}$ *coincides with the set of blocks of* Θ *iff the following two conditions are satisfied:*

(1) $e_i \not\subseteq e_j$ *for all* $e_i, e_j \in E$.

(2) *Let* $\tilde{e} \subset G$ *and* $\tilde{e} \notin E$. *Then the existence of* $e_{i_1}, \ldots, e_{i_k} \in E$ *such that* $\tilde{e} \subset \cup_{l=1}^{k} e_{i_l}$ *implies the existence of* $x_1, \ldots, x_n \in G$ *such that* $\{x_1, \ldots, x_n\} \subseteq \tilde{e}$ *and for all* $e_{i_l} (1 \leq l \leq k)$ *either* $\{x_1, \ldots, x_n\} \not\subseteq e_{i_l}$ *or* $\cap_{l=1}^{k} e_{i_l} \subseteq \tilde{e}$.

Theorem 5.4.1 is valid not only for a binary tolerance, but for n-ary tolerance as well (in this case it is defined as $\Theta(X_1, \ldots, X_n) = X_1' \cap \ldots \cap X_n' \neq \varnothing$). If the set of all intents of the context is represented by an n-ary tolerance (in the sense of Theorem 5.4.1), it can be also represented by an m-ary tolerance defined in the same way on G for an arbitrary m: $\min\limits_{e_i \in E} |e_i| > n \geq m$.

5.5. The Zarankiewicz Problem

The Zarankiewicz problem consists in determination of the number $k(a, b, m, n)$ such that an arbitrary binary $m \times n$ matrix (i.e., a context (G, M, I) such that $|G| = m$, $|M| = n$) with $k(a, b, m, n)$ units (i.e., $|I| = k(a, b, m, n)$) contains an identity submatrix of the size $a \times b$, where $2 \leq a \leq m$, $2 \leq b \leq n$ (Zarankiewicz himself stated the problem for $a = b = 2$ [95]). The problem was extensively studied in the fifties and sixties. We present the main results concerning the estimates of k ($k(a,a)$ denotes $k(a,a,n,n)$).

In [46] it was proved that

$$k(2,2,p^2+p,p^2) = p^3+p^2+1$$

for the case where p is prime.

Hartman et al. [43] proved that

$$C_1 n^{4/3} < k(2,n) < C_2 n^{3/2},$$

where C_1 and C_2 are constants. The following result was obtained in [13] for $1 \le b \le m$ and $n \ge (a-1)\binom{m}{b}$:

$$k(a,b,m,n) = (b-1)n + (a-1)\binom{m}{b} + 1$$

The following upper bound was obtained in [67]:

$$k(2,2,m,n) < \frac{1}{2}(n + n\sqrt{n + 4mn(m-1)}) + 1.$$

When $n = m$, the latter inequality gives

$$k(2,n) \le \frac{1}{2}(n + n\sqrt{4n-3}) + 1.$$

This bound is sharper than the bound

$$k(2,n) \le 1 + 2n + [n^{3/2}]$$

given in [46]. The above inequality from [67] was shown there to turn into the equality in infinitely many cases. Moreover,

$$k(2,2,p^2+p+1,p^2+p+1) = p^3+2p^2+2p+2,$$

where p is a power of a prime.

The following results were proved in [44, 40, 41]:

$$\lim_{n\to\infty} k(2,3,n,n)n^{-3/2} = 2,$$

$$[b/2]^{1/2} < \lim_{n\to\infty}\ \inf k(2,b,n,n)n^{-3/2} < \lim_{n\to\infty}\ \sup k(2,b,n,n)n^{-3/2} \le (k-1)^{1/2} \quad [44];$$

$$k(a,b,m,n) = \left[\frac{(a^2-1)n + (b-1)\binom{m}{a}}{a}\right] + 1$$

for

$$(b-1)\binom{m}{a} + 1 \ge n \ge \ell(m,a,b),$$

where $\ell(m,a,b)$ equals approximately

$$(b-1)\binom{m}{a}/(a+1) \ [40];$$

(exact values of $\ell(m,a,b)$ for small a are given in [40]);

$$k(a,b,m,n) \le 1 + [nu], \text{ for } 3 \le a \le m,\ 3 \le b \le n,\ n \ll m^a$$

and

$$u = v + \frac{1}{2}(a-1),$$

$$v = z + \frac{(a^2-1)}{24z} + \frac{(a^2-1)(a^2-9)}{1920z^3} + \frac{(a^2-1)(a^2-4)(a^2-25)}{41472z^5},$$

$$nz^a = (b-1)m(m-1)\dots(m-a+1) \ [41].$$

The results of [99] were improved in [100] as follows:

$$k(a,n) < [1/2n(a-1) + (a-1)^{1/a}(n-3/8(a-1)^{1-1/a},$$

$$k(a,n) < [n(a-1)/\epsilon + (a-1)^{1/a}n^{2-1/a}],$$

where

$$\varepsilon = (2(n(a-1))^{1/a} - 1)/((n/(a-1))^{1/a} - 1).$$

The second estimate is better than the first only in the case where $a = 2$ or $a = 3$, as well as for considerably small n when $a \geq 4$.

Erdös and Spenser [18] consider the function $B(a,b,m,n,e)$ defined as the least number of concepts of the size $a \times b$ of a context (G, M, I) with $|G| = m$, $|M| = n$, and $|I| = e$. It is obvious that $B(a,b,m,n,e) > 0 \iff k(a,b,m,n) \leq e$. The following estimate of $B(a,b,m,n,e)$ was obtained in [18]:

$$\binom{m}{a}\binom{n}{b}\frac{\binom{mn-ab}{e-ab}}{\binom{mn}{e}} \geq B(a,b,m,n,e) \geq \binom{n}{b}\binom{\left(\frac{m\binom{e/m}{b}}{\binom{n}{b}}\right)}{a}.$$

This estimates of $B(a,b,m,n,e)$ allowed Erdös and Spenser [18] to obtain the following bounds for $k(a,n)$:

$$(a!)^{2/a^2}n^{2-2/a}(1-o(1)) \leq k(a,n) \leq (a-1)^{1/a}n^{2-1/a}(1+o(1)).$$

In [69] the upper bound

$$k(a,b,m,n) \leq \frac{b-1}{\binom{p}{a-1}}\binom{m}{a} + \frac{(p+1)(a-1)}{a}n + 1$$

was proved for all integers $p \geq a-1$.

Let $T_{a+1,m,b-1}$ be the maximal number of subsets of size $a-1$ that can be packed in a set of size m in a way such that no subset of the size a is in more than $b-1$ subsets. In [69] it was proved that

$$k(a,b,m,n) = \left[\frac{b-1}{a}\binom{m}{a} + \frac{a^2-1}{a}n\right] + 1$$

for

$$\max\left[\frac{b-1}{a+1}\binom{m}{a}, (b-1)\binom{m}{a} - T_{a+1,m,b-1}\right] \leq n \leq (b-1)\binom{m}{a}.$$

The paper [61] generalizes the result from [44]: the author proved that

$$\lim_{n\to\infty} k(2,b,n,n)n^{-3/2} = (b-1)^{1/2}$$

for an arbitrary $b \geq 2$.

REFERENCES

1. A. Aho and J. Hopcroft, *The Design and Analysis of Computer Algorithms*, Addison-Wesley (1974).
2. O. M. Anshakov, D. P. Skvortsov, and V. K. Finn, "Logical means of the JSM-method of automatic hypothesis generation," *Nauchn. Tekh. Inf., Ser. 2 (Autom. Document. Math. Linguist.)*, No. 9, 9-17 (1978).
3. O. M. Anshakov, D. P. Skvortsov, V. K. Finn, and V. G. Ivashko, "Logical means of the JSM-method of automatic generation: main notions and the system of inference rules," *Nauchn. Tekh. Inf., Ser. 2 (Autom. Document. Math. Linguist.)*, No. 9, 10-18 (1987).

4. O. M. Anshakov, V. K. Finn, and D. P. Skvortsov, "On logical means of formalization of plausible inferences," In: *VIII International Congress of Logic, Methodol., and Phil. of Sci., Abstracts*, 212-215, Nauka, Moscow, (1987).

5. O. M. Anshakov, V. K. Finn, and D. P. Skvortsov, "On axiomatization of many-valued logics associated with the formalization of plausible reasonings," *Stud. Log.*, **25**, No. 4, 23-47 (1989).

6. R. Balbes and P. Dwinger, *Distributive Lattices*, University of Missouri Press (1974).

7. M. Barbut and B. Monjardet, *Ordre et classification,* **II**, Hachette, Paris (1970).

8. G. D. Birkhoff, *Lattice Theory*, Amer. Math. Soc. (1979).

9. J. P. Bordat, "Calcul pratique du treillis de Galois d'une correspondance," *Math. Sci. Hum.*, No. 96, 31-47 (1986).

10. P. Burmeister, "Merkmalsimplikationen bei unvollständigem Wissen," In: *Arbeitstagung Begriffdanalyse und Künstliche Intelligenz* (W. Lex, ed.), Institut für Informatik, Clausthal, pp.15-46 (1991).

11. T. Bylander, D. Allemang, M. C. Tanner, and J. R. Josephson, "The computational complexity of abduction," *Artif. Intell.,* **49**, No. 1, 25-60 (1991).

12. M. Chein, "Algorithme de recherche des sous-matrices premières d'une matrice," *Bull. Math. R.S. Roumanie*, No. 13, No. 1, 21-25 (1969).

13. K. Čulik, "Teilweise Lösung eines verallgemeinerten Problem von K. Zarankiewicz," *Ann. Polon. Math.*, **3**, No.1, 165-168 (1956).

14. J. Duda, "Boolean concept lattices and good contexts," *Čas. Pest. Mat.*, **114**, No.2, 165-175 (1989).

15. V. Duquenne, "Contextual implications between attributes and some representation properties for finite lattices," In: *Beiträge zur Begriffsanalyse* (B. Ganter, R. Wille, and K. E. Wolff, eds.), B. I. Wissenschaft-Verlag, Mahnnheim, pp.213-239 (1987).

16. V. Duquenne and J.-L. Guigues, "Familles minimales d'implications informatives résultant d'un tableau de données binaires," *Math. Sci. Hum.*, **24**, No. 95, 5-18 (1986).

17. B. Efron, *Nontraditional Methods of Many-Dimensional Statistical Information,* [Russian translation], Moscow (1988).

18. P. Erdös and J. Spenser, *Probability Methods in Combinatorics* [Russian translation], Moscow (1976).

19. V. K. Finn, "Data bases with incomplete information and a new method of automatic hypothesis generation," In: *Dialog and Factographic Information Systems* [in Russian], Moscow, pp.153-156 (1981).

20. V. K. Finn, "On machine-oriented formalization of plausible reasoning in the style of F. Backon–J. S. Mill," *Semiotika Informatica,* No. 20, 35-101 (1983).

21. V. K. Finn, "Plausible inference and plausible reasoning," *Itogi Nauki i Tekhniki, Seriya Teoriya Veroyatnostei, Mat. Statist. Teoret. Kibern.*, **28**, 3-84 (1988).

22. V. K. Finn, "On the generalized JSM-method of automatic hypothesis generation," *Semiotika Informatika*, No. 29 93-123 (1989).

23. V. K. Finn, "JSM-method of automatic generation with an order relation," *Semiotika Informatika,* No. 31, 69-103 (1990).

24. V. K. Finn, "Plausible reasoning in systems of JSM type," *Itogi Nauki i Tekhniki, Seriya Informatika*, **15**, 54-101 (1991).

25. J. G. Ganascia, "CHARADE: A rule system learning system," In: *Proc. of the 10th International Joint Conference on Artificial Intelligence (IJCAI-87)*, Milan, Italy, August 23-28 (1987), 117-124 (1987).

26. J. G. Ganascia, "Improvement and refinement of the learning bias semantic," In: *Proc. of the 8th Europ. Conf. on Artificial Intelligence (ECAI-88)*, pp.81-86, (1988).

27. B. Ganter, "Two basic algorithms in concept analysis," *preprint 831*, Technische Hochshule Darmstadt (1984).

28. B. Ganter, "Algorithmen zur Formalen Begriffsanalyse," In: *Beiträge zur Begriffsanalyse*, (B. Ganter, R. Wille, and K. E. Wolff, eds.), B. I. Wissenschaftsverlag, Mannheim, 241-254, (1987).

29. B. Ganter, "Composition and Decomposition in Formal concept Analysis," In: *Classification and Related Methods of Data Analysis* (H. H. Bock, ed.), North-Holland, Amsterdam, 561-566, (1988).

30. B. Ganter, K. Rindfrey, and M. Skorsky, "Software for formal concept analysis," In: *Classification as a Tool of Research* (W. Gaul and M. Schader, eds.), North-Holland, Amsterdam, 161-167, (1986).

31. B. Ganter, J. Stahl, and R. Wille, "Conceptual measurement and many-valued contexts," In: *Classification as a Tool of Research* (W. Gaul and M. Schader, eds.), North-Holland, Amsterdam, 169-176, (1986).

32. B. Ganter and R. Wille, "Implikationen und Abhängigkeiten zwischen Merkmalen," In: *Die Klassifikation und ihr Umfeld*, (P. O. Degens, H.-J. Hermes, and O. Opitz, eds.), Index-Verlag, Frankfurt, 171-185, (1986).

33. B. Ganter and R. Wille, "Conceptual scaling," In: *Applications of Combinatorics and Graph Theory to the Biological and Social Sciences* (F. Roberts, ed.), Springer, New York, 139-167, (1989).

34. M. Gary and D. Johnson, *Computers and Intractability* [Russian translation], Mir, Moscow (1982).

35. G. Grätzer, *General Lattice Theory*, Birkhäuser Verlag, Basel–Stuttgart (1978).

36. A. Guénoche, "Construction du treillis de Galois d'une relation binaire," *Math. Inf. Sci. Hum.*, No. 109, 41-53 (1990).

37. S. M. Gusakova, "Canonical representation of similarity," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)*, No. 9, 19-22 (1987).

38. S. M. Gusakova and V. K. Finn, "On formalization of local and global similarity," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)*, No. 6, 16-19 (1986).

39. S. M. Gusakova and V. K. Finn, "On new means of formalization of the notion of similarity," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)*, No. 10, 14-22 (1987).

40. R. K. Guy, "A problem of Zarankiewicz," In: *Theory of Graphs*, (G. Katona, ed.) Akad. Kiado, Budapest, 119-150, (1968).

41. R. K. Guy and Š. Znám, "A problem of Zarankiewicz," *Recent Progr. Combinator.*, New York-London, 237-243, (1969).

42. R. M. Haralick, "The diclique representation and decomposition of binary relations," *J. ACM*, **21**, 356-366 (1974).

43. S. Hartman, J. Mycielski, and C. Ryll-Nardzewski, "Système speciaux de points a coordonnees entieres," *Coll. Math.*, **3**, No. 1, 84-85 (1954).

44. C. Hylten-Cavallius, "On a combinatorial problem," *Colloq. Math.*, **6**, 59-65 (1958).

45. U. Kipke and R. Wille, "Begriffsverbände als Ablaufschemata zur Gegenstandsbestimmung," In: *Die Klassifikation und ihr Umfeld* (P. O. Degens, H.-J. Hermes, and O. Opitz, eds.), Indeks-Verlag, Frankfurt, 164-170, (1986).

46. T. Kővari, V. Sós, and P. Turan, "On a problem of K. Zarankiewicz," *Colloq. Math.*, **3**, No. 1, 50-57 (1954).

47. S. O. Kuznetsov, "Interpretation on graphs and complexity characteristics of a search for specific patterns," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)* No. 1, 23-27 (1989).

48. S. O. Kuznetsov, "Introduction to the JSM-method," *Semiotika i Informatika*, No. 31, 5-40 (1990).

49. S. O. Kuznetsov, "Stability as an estimate of the degree of substantiation of hypotheses derived on the basis of operational similarity," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)* No. 12, 21-29 (1990).

50. S. O. Kuznetsov, "JSM-method as a system of machine learning," *Itogi Nauki i Tekhniki, Seriya Informatika*, **15**, 17-54 (1991).

51. S. O. Kuznetsov, "Complexity of learning and classification based on a search for set intersection," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)* No. 5, 1-11 (1991).

52. S. O. Kuznetsov, "Fast algorithm for construction of all intersections of the objects from a finite semilattice," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)* No. 9, 11-21 (1993).

53. S. O. Kuznetsov, Algorithmic complexity of the generation of hypotheses and classification based on the search for set intersection *Doklady Akad. Nauk*, **335**, No. 3, 300-303, (1994) (*Physics-Doklady*, **39**, No. 3, 142-145 (1994)).

54. V. E. Levit, "Algorithm of search for a submatrix of maximal perimeter that consists of units for the case of 0-1 matrix," *Systems of transmitting and processing of information, Collection of papers of IPPI of Soviet Acad. Sci.*, 42-45, (1988).

55. P. Luksch, Skorsky M., and R. Wille, "On drawing concept lattices with a computer," In: *Classification as a Tool of Research* (W. Gaul, and M. Schader, eds.), Elsevier, 269-274, (1986).

56. P. Luksch and R. Wille, "Substitution decomposition of concept lattices," In: *Contributions to General Algebra 5* (J. Czermak, G. Eigenthaler, H. K. Kaiser, and W. B. Müller, W. Nobauer, eds.), Holder-Pichler-Tempsky, Wien, 213-220, (1987).

57. P. Luksch and R. Wille, "Formal concept analysis of paired comparisons," In: *Classification and Related Methods of Data Analysis* (H. H. Bock, ed.), North-Holland, Amsterdam, 567-576, (1988).

58. P. Luksch and R. Wille, "A mathematical model for conceptual knowledge systems," In: *Classification, Data Analysis, and Knowledge Organization* (H. H. Bock and P. Ihm, eds.), Springer, Heidelberg, 156-162, (1990).

59. M. Luxenburger, "Implications partielles dans un contexte," *Math. Inf. Sci. Hum.*, **29**, No.113, 35-55 (1991).

60. R. S. Michalski and R. E. Stepp, "Learning from observation: conceptual clustering," *Machine Learning: an Artificial Intelligence Approach* (R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, eds.), Tioga, Palo Alto, California, 41-81, (1983).

61. M. Mörs, "A new result on the problem of Zarankiewicz," *J. Comb. Theory, Ser. A*, **31**, No. 2, 126-130 (1981).

62. E. M. Norris, "An algorithm for computing the maximal rectangles in a binary relation," *Rev. Roum. de Math. Pures Appliquées*, **23**, No. 2, 243-250 (1978).

63. M. Novotný and Z. Pawlak, "Black box analysis and rough top equality," *Bull. Acad. Sci. Math.*, **33**, Nos. 1-2, 105-113 (1985).

64. M. Novotný and Z. Pawlak, "Concept forming and black boxes," *Bull. Acad. Sci. Math.*, **35**, Nos. 1-2, 133-141 (1987).

65. M. Novotný and Z. Pawlak, "Algebraic theory of independence in information systems," *Fundamenta Informaticae*, **14**, No. 4, 454-476 (1991).

66. P. Pudlak and F. Springsteel, "Complexity in mechanized hypothesis formation," *Theor. Comp. Sci.*, **8**, No. 2, 203-225 (1979).

67. I. Reiman, "Über ein Problem von K. Zarankiewicz," *Acta Math. Acad. Sci. Hungar.*, **9**, 269-273 (1958).

68. K. Reuter and R. Wille, "Complete congruence relations of concept lattices," *Acta Sci. Math.*, **51**, Nos. 3-4, 319-327 (1987).

69. S. Roman, "A problem of Zarankiewicz," *J. Comb. Th. (A)*, **18**, No. 2, 187-198 (1975).

70. J. Schmidt, "Zur Kennzeichnung der Dedekind-MacNeilleschen Hülle einer geordneten Menge," *Arch. Math.*, **7**, No.4, 241-249 (1956).

71. M. Skorsky, "How to draw a concept lattice with parallelograms," In: *Klassifikation und Ordnung* (R. Wille, ed.), INDEKS-Verlag, Frankfurt, 191-196, (1989).

72. F. Springsteel, "Complexity of hypothesis formation problems," *Int. J. Man-Mach. Studies*, **15**, 319-332 (1981).

73. J. Stahl and R. Wille, "Preconcepts and set representation of contexts," In: *Classification as a Tool of Research*, (W. Gaul and M. Schader, eds.), Elsevier, 431-438, (1986).

74. L. G. Valiant, "The complexity of computing the permanent," *Theoretical Computer Science*, **8**, No. 2, 189-201 (1979).

75. L. G. Valiant, "The complexity of enumeration and reliability problems," *SIAM J. Comput.*, **8**, No. 3, 410-421 (1979).

76. F. Vogt, C. Wachter, and R. Wille, "Data analysis based on a conceptual file," In: *Classification, Data Analysis, and Knowledge Organization* (H. H. Bock and P. Ihm, eds.), Springer, Heidelberg, 131-142, (1990).

77. M. Wild, "Implicational bases for finite closure systems," In: *Arbeitstagung, Begriffsanalyse und Künstliche Intelligenz* (W. Lex, ed.), 147-169, (1991).

78. R. Wille, "Subdirekte Produkte und konjunkte Summen," *J. reine angew. Math.*, **239/240**, 333-338 (1970).

79. R. Wille, "Subdirekte Produkte vollständiger Verbände," *J. reine angew. Math.*, **283/284**, 53-70 (1976).

80. R. Wille, "Restructuring lattice theory: an approach based on hierarchies of concepts," In: *Ordered Sets* (I. Rival, ed.), Reidel, Dordrecht–Boston, 445-470, (1982).

81. R. Wille, Subdirect decomposition of concept lattices, *Algebra Universalis*, **17**, No. 3, 275-287 (1983).

82. R. Wille, "Liniendiagramme hierarchischer Begriffssysteme," In: *Anwendungen der Klassifikation: Datenanalyse und Numerische Klassifikation*, (H. H. Bock, ed.), INDEKS-Verlag, Frankfurt, 32-51, (1984).

83. R. Wille, "Sur la fusion des contexts individuels," *Math. Sci. Hum.*, **85**, 57-71 (1984).

84. R. Wille, "Tensorial decomposition of concept lattices," *Order*, **2**, 81-95 (1985).

85. R. Wille, "Finite distributive lattices as concept lattices," *Atti Inc. Logica Matematica (Siena)*, **2**, 635-648 (1985).

86. R. Wille, "Complete tolerance relations of concept lattices," In: *Contributions to General Algebra Proceedings of a Conference*, June 21-24, 1984, (B. G. Teubner, ed.), 397-415, (1985).

87. R. Wille, "Bedeutungen von Begriffsverbänden" In: *Beiträge zur Begriffsanalyse* (B. Ganter, R. Wille, and K. E. Wolff, eds.), B.I.-Wissenschaftsverlag, Mannheim, 161-211, (1987).

88. R. Wille, "Subdirect product construction of concept lattices," *Discr. Math.*, **63**, 305-313 (1987).

89. R. Wille, "Dependences between many-valued attributes," In: *Classification and Related Methods of Data Analysis* (H. H. Bock, ed.), North-Holland, Amsterdam, 581-586, (1988).

90. R. Wille, "Geometric representation of concept lattices," In: *Conceptual and Numerical Analysis of Data* (O. Opitz, ed.), Springer, Heidelberg, 239-255, (1989).

91. R. Wille, "Knowledge acquisition by methods of formal concept analysis," In: *Data Analysis, Learning Symbolic and Numeric Knowledge* (E. Diday, ed.), Nova Sci. Publ., New York–Budapest, 365-380, (1989).

92. R. Wille, "Lattices in data analysis: how to draw them with a computer," In: *Algorithms and Order* (I. Rival, ed.) Kluwer, Dordrecht-Boston, 33-58, (1989).

93. R. Wille, "The skeletons of free distributive lattices," *Discr. Math.*, **88**, Nos. 2-3, 309-320 (1991).

94. R. Wille, "Concept lattices and conceptual knowledge systems," *Comput. Math. Appl.*, **23**, No. 6-9, 493-515 (1992).

95. M. I. Zabezhailo, "On some exhaustive search problems that arise in automatic generation of hypotheses by the JSM-method," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)*, No. 1, 28-31 (1988)

96. M. I. Zabezhailo, V. G. Ivashko, S. O. Kuznetsov, M. A. Mikheenkova, K. P. Khazanovskii, and O. M. Anshakov, "Algorithmic and program means of the JSM-method of automatic hypothesis generation," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)*, No. 10, 1-14 (1987).

97. M. I. Zabezhailo, "To the problem of the extension of the JSM-method of automatic hypothesis generation to data with numeric parameters," *Nauchn. Tekh. Inf., Ser. 2 (Automat. Document. Math. Linguist.)* No. 11, 11-21 (1992).

98. K. Zarankiewicz, "Problem P101," *Colloq. Math.*, **2**, 301 (1951).

99. S. Znám, "On a combinatorial problem of K. Zarankiewicz," *Colloq. Math.*, **11**, No. 1, 81-84 (1963).

100. Š. Znám, "Two improvements of a result concerning a problem of K. Zarankiewicz," *Colloq. Math.*, **13**, No.2, 255-258 (1965).