A STRONGER ASSOCIATION RULE IN LATTICES, FORMAL CONTEXTS, DATABASES, AND POSETS

GÁBOR CZÉDLI

Abstract. Galois closure operators play an important role in many fields including algebra, formal concept analysis and mining association rules from databases. Given a context $(A^{(0)}, A^{(1)}, \rho)$, the pair of the induced Galois closure operators will be denoted by $\mathcal{G} = \mathcal{G}(A^{(0)}, A^{(1)}, \rho)$. The present paper studies a new pair $\mathcal{C} = \mathcal{C}(A^{(0)}, A^{(1)}, \rho)$ of closure operators, which has been introduced in [3]. After pointing out that $\mathcal{C}$ is of some interest for algebra, decision making and knowledge discovery from databases, we characterize $\mathcal{C}$ as a fixed point of an appropriate contraction map. This easy result leads to a computer program, which is available at the author's home page. Due to this program, there are some statistical results in the paper.

For $\rho \in \{\le, <, \preceq, \prec\}$, the main theorem characterizes finite posets P with $\mathcal{C}(P, P, \rho) = \mathcal{G}(P, P, \rho)$. It is proved that that $\mathcal{C}(J(L), M(L), \le) = \mathcal{G}(J(L), M(L), \le)$ when L is a finite modular lattice.

1. Introduction and motivating examples

Following Wille's terminology, cf. [10] or [6], a triplet

$$(A^{(0)}, A^{(1)}, \rho)$$

is called a *context* if $A^{(0)}$ and $A^{(1)}$ are nonempty sets and $\rho \subseteq A^{(0)} \times A^{(1)}$ is a binary relation. From what follows, we fix a context $(A^{(0)}, A^{(1)}, \rho)$ and let

$$\rho_0 = \rho \text{ and } \rho_1 = \rho^{-1}.$$

From now on, unless otherwise stated, i will be an arbitrary element of $\{0, 1\}$. So whatever we say including i without specification, it will be understood as prefixed by $\forall i$. The set of all subsets of $A^{(i)}$ will be denoted by $P(A^{(i)})$.

It is often, especially in the finite case, convenient to depict our context in the usual form: a binary table with row labels from $A^{(0)}$, column labels from $A^{(1)}$, and a cross in the intersection of the x-th row and the y-th column iff $(x, y) \in \rho$. We will refer to this table as the *context table*. For example, a context is given by Table 1.

Although the concrete meaning of this context about juggling is not relevant for this paper, we make some comments on it, and we refer to Polster [8] and its bibliography for more details. Attribute b_4 means that at least one hand *essentially* leaves the starting position. Attribute b_2 means that the balls are indistinguishable,

Date: Version of September 4, 2006.

Key words and phrases. Association rule, database, data mining, lattice, poset, context, concept lattice, formal concept analysis, Galois connection, functional dependency, juggling.

This research was partially supported by the NFSR of Hungary (OTKA), grant no. T 049433 and K 60148.

i.e, if we do not watch the performance for a while then we cannot tell which ball was thrown first. The rows are well-known three-ball juggling patterns.

		equal hands b_1	equal balls b_2	different heights b_3	moving hand b_4
3 (cascade)	a_1	$\times$	$\times$		
423	a_2	$\times$		$\times$	
Mill's mess	a_3	$\times$	$\times$		$\times$
50505 (snake)	a_4	$\times$			
robot	a_5		$\times$	$\times$	$\times$

Table 1

A mapping $\mathcal{D}^{(i)} : P(A^{(i)}) \to P(A^{(i)})$ is called a *closure operator* if it is *extensive* (i.e., $X \subseteq \mathcal{D}^{(i)}(X)$ for all $X \in P(A^{(i)})$), *monotone* (i.e., $X \subseteq Y$ implies $\mathcal{D}^{(i)}(X) \subseteq \mathcal{D}^{(i)}(Y)$), and *idempotent* (i.e., $\mathcal{D}^{(i)}(\mathcal{D}^{(i)}(X)) = \mathcal{D}^{(i)}(X)$ for all $X \in P(A^{(i)})$). By a *pair of extensive operators* we mean a pair $\mathcal{D} = (\mathcal{D}^{(0)}, \mathcal{D}^{(1)})$ where $\mathcal{D}^{(i)} : P(A^{(i)}) \to P(A^{(i)})$ is an extensive mapping for $i = 0, 1$. If these mappings are closure operators then $\mathcal{D}$ is called a *pair of closure operators*.

If $\mathcal{D} = (\mathcal{D}^{(0)}, \mathcal{D}^{(1)})$ and $\mathcal{E} = (\mathcal{E}^{(0}, \mathcal{E}^{(1)})$ are pairs of extensive operators then $\mathcal{D} \le \mathcal{E}$ means that $\mathcal{D}^{(i)}(X) \subseteq \mathcal{E}^{(i)}(X)$ for all $i \in \{0, 1\}$ and all $X \in P(A^{(i)})$.

Now, associated with $(A^{(0)}, A^{(1)}, \rho)$, we define some pairs of closure operators. The motivation will be given afterwards. For $X \in P(A^{(i)})$ let

$$X\rho_i = \{y \in A^{(1-i)} : \text{ for all } x \in X,\ (x, y) \in \rho_i\},$$

and, again for $X \in P(A^{(i)})$, define

$$\mathcal{G}^{(i)}(X) := (X\rho_i)\rho_{1-i} = \bigcap_{y \in X\rho_i} (\{y\}\rho_{1-i}) \ .$$

Then $\mathcal{G} = (\mathcal{G}^{(0)}, \mathcal{G}^{(1)})$ is the well-known *pair of Galois closure operators*, which plays the main role in formal concept analysis, cf. Wille [10] and Ganter and Wille [6]. The visual meaning of

$$\mathcal{G} = \mathcal{G}(A^{(0)}, A^{(1)}, \rho)$$

is the following. The maximal subsets of ρ of the form $U^{(0)} \times U^{(1)}$ with $U^{(i)} \subseteq A^{(i)}$ are called the (formal) *concepts*, cf. [10] or [6]. Pictorially, they are the maximal full rectangles $U^{(0)} \times U^{(1)}$ of the context table. (Full means that each entry is a cross.) For $X_i \in P(A^{(i)})$ take all maximal full rectangles $U^{(0)} \times U^{(1)}$ such that $X \subseteq U^{(i)}$, then $\mathcal{G}^{(i)}(X)$ is the intersection of all the $U^{(i)}$'s.

Now we define a sequence $\mathcal{C}_i$, $i = 0, 1, 2, \ldots$, of pairs of of closure operators. For $X \in P(A^{(i)})$ let

$$X\psi_i := \{Y \in P(A^{(1-i)}) : \text{ there is a surjection } \varphi : X \to Y \text{ with } \varphi \subseteq \rho_i\}.$$

Let $\mathcal{C}_0 = \mathcal{G}$. If $\mathcal{C}_n$ is already defined then let

$$\mathcal{C}^{(i)}_{n+1}(X) := \mathcal{C}^{(i)}_n(X) \cap \bigcap_{Y \,\in\, X\psi_i} \ \bigcup_{y \,\in\, \mathcal{C}^{(1-i)}_n(Y)} \{y\}\rho_{1-i} \ . \tag{1}$$

This defines the pair $\mathcal{C}_{n+1} = (\mathcal{C}^{(0)}_{n+1}, \mathcal{C}^{(1)}_{n+1})$. Finally, let

$$\mathcal{C} = (\mathcal{C}^{(0)}, \mathcal{C}^{(1)}) := (\bigwedge_{n=0}^{\infty} \mathcal{C}^{(0)}_n, \bigwedge_{n=0}^{\infty} \mathcal{C}^{(1)}_n),$$

which means that, for all $X \in P(A^{(i)})$,

$$\mathcal{C}^{(i)}(X) = \bigcap_{n=0}^{\infty} \mathcal{C}^{(i)}_n(X).$$

Even if the above definitions do not look friendly at the first sight, it was routine to prove in [3] that we have indeed defined pairs of closure operators.

Lemma 1. *(cf. [3])* $\mathcal{C} = \mathcal{C}(A^{(0)}, A^{(1)}, \rho)$ *and* $\mathcal{C}_n = \mathcal{C}_n(A^{(0)}, A^{(1)}, \rho)$ $(n = 0, 1, \ldots)$ *are pairs of closure operators. Further,* $\mathcal{G} = \mathcal{C}_0 \geq \mathcal{C}_1 \geq \mathcal{C}_2 \geq \cdots \geq \mathcal{C}$.

It is well-known that, for each context $(A^{(0)}, A^{(1)}, \rho)$, the complete lattices $(\{X \in P(A^{(0)}) : \mathcal{G}^{(0)}(X) = X\}, \subseteq)$ and $(\{X \in P(A^{(1)}) : \mathcal{G}^{(1)}(X) = X\}, \subseteq)$ are dually isomorphic. The analogous statement is far from being true for $\mathcal{C}$; indeed, in case of the context given by Table 1, $|\{X \in P(A^{(0)}) : \mathcal{C}^{(0)}(X) = X\}| = 12$ while $|\{X \in P(A^{(1)}) : \mathcal{C}^{(1)}(X) = X\}| = 10$.

From now on *we always assume* that $(A^{(0)}, A^{(1)}, \rho)$ is *finite*. Then there are only finitely many pairs of operators, whence there is an n with $\mathcal{C} = \mathcal{C}_n = \mathcal{C}_{n+1} = \mathcal{C}_{n+2} = \cdots$.

Now we give our motivations, and this will explain why "association rule" occurs in the title of the paper. Closure operators have been playing an important role in the theory of relational databases and knowledge systems for a long time, cf. e.g., Caspard and Monjardet [2] for a survey. Nowadays most investigations of this kind belong to formal concept analysis, cf. Ganter and Wille [6] for an extensive survey. The theory of mining association rules goes back to Agrawal, Imielinski and Swami [1]; Lakhal and Stumme [7] gives a good account on the present status of this field.

For a data miner, the context is a *huge* binary database, and mining association rules is a popular knowledge discovery technique for warehouse basket analysis. In this case $A^{(0)}$ is the set of costumers' baskets, $A^{(1)}$ is the set of items sold in the warehouse, and the task is to figure out which items are frequently bought together. This information, expressed by so-called "association rules", can help the warehouse in developing appropriate marketing strategies. For example,

$$\{\text{cereal, coffee}\} \to \{\text{milk}\}$$

is an association rule (in many real warehouses), and this association rule says that, with a given probability p, costumers buying cereal and coffee also buy milk. When milk $\in \mathcal{G}^{(1)}\{$cereal, coffee$\}$ then this probability is 1 and we speak about a *strong* association rule.

However, the importance of looking for the hidden regularities and rules is not restricted only to *huge* databases. The success of formal concept analysis, cf. Ganter and Wille [6], or Mendeleyev's classical periodic system of chemical elements show that exploring some rules in *small* databases may also lead to important results. From this aspect, the present paper offers $\mathcal{C}$, a mathematical tool, to formulate some regularities in abstract contexts. Since $\mathcal{C} \leq \mathcal{G}$, the "association rules" corresponding to $\mathcal{C}$ are *stronger* than the previously mentioned ones. It would be nice to find some concrete contexts outside mathematics, say in natural or social sciences, where $\mathcal{C}$ has some real applications. This is much beyond the scope (i.e., lattices and posets) of the present paper but there is a hope, for finding associations is an integral part of any creative activity.

Now we use the context given by Table 1 to develop our ideas further. Let $X = \{a_1, a_2\}$, which is a subset of $A^{(0)} = \{a_1, \dots, a_5\}$. Then $\{a_1, \dots, a_4\} \times \{b_1\}$ is the only relevant maximal full rectangle to compute $\mathcal{G}^{(0)}(X) = \{a_1, \dots, a_4\}$. Notice that $Y = \{b_2, b_3\} \in X\psi_0$ but there is no $y \in \mathcal{G}^{(1)}(Y) = \{b_2, b_3, b_4\}$ with $a_4 \in \{y\}\rho_1$, for $\{a_4\} \times \{b_2, b_3, b_4\}$ is disjoint from ρ. Hence, according to formula (1), $a_4 \notin \mathcal{C}_1^{(0)}(X)$. After the trivial and therefore omitted details we can easily see that $\mathcal{C} = \mathcal{C}_1$ and $\mathcal{C}^{(0)}(X) = \{a_1, a_2, a_3\}$.

Suppose our whole knowledge is decoded in the context and we are asked to associate an element with X. Usually we want an element outside X, and we look for something similar, i.e., we want an element which shares the common attributes of the elements of X. So the first answer is that we should associate some element of $\mathcal{G}^{(0)}(X) \setminus X = \{a_3, a_4\}$. This way we obtain more than one element, but we may want to chose only a single one. For example, if a beginner can perform the juggling patterns a_1 and a_2 then which of a_3 and a_4 should he learn next? Which of a_3 and a_4 should a scientist choose if the context represents something in his research field and choosing both is not permitted? The unique element a_3 of $\mathcal{C}^{(0)}(X) \setminus X$? The other element, a_4? There is no general answer to this question of *decision making* in full generality (and I think that even experienced jugglers would give different answers to the concrete question about Table 1). We just point out that sometimes $\mathcal{C}$ offers a method distinct from coin tossing.

Let us mention that [3] gives a purely universal algebraic theorem which has nothing to do with the notion of $\mathcal{C}$ but $\mathcal{C}$ is heavily used in the proof. This shows that $\mathcal{C}$ is useful in algebra. The rest of the paper is scheduled as follow. First we characterize $\mathcal{C}$ as the largest fixed point of an appropriate contraction map. This theorem was exploited when we wrote a computer program to compute $\mathcal{C}$. The program is available at the author's home page. (Although the source code is in Borland's old Turbo Pascal 7.0 for MS DOS, the executable version runs in today's Windows environment as well.) The next question is how often $\mathcal{C}$ is different from $\mathcal{G}$. There will be experimental results obtained with the help of the program, and there will be mathematical results for some specific contexts obtained from finite lattices or posets.

2. From a fixed point theorem to a program

Given a context $(A^{(0)}, A^{(1)}, \rho)$, let $\mathbf{H} = \mathbf{H}(A^{(0)}, A^{(1)}, \rho)$ be the set of all pairs of extensive operators defined in the previous section. Similarly, the set of all pairs of closure operators will be denoted by $\mathbf{T} = \mathbf{T}(A^{(0)}, A^{(1)}, \rho)$. Then $\mathbf{H} = (\mathbf{H}, \leq)$ and $\mathbf{T} = (\mathbf{T}, \leq)$ are posets, $\mathbf{T}$ is a sub-poset of $\mathbf{H}$, and $\mathcal{G}, \mathcal{C} \in \mathbf{T} \subseteq \mathbf{H}$. Motivated by formula (1), we define a mapping $f : \mathbf{H} \to \mathbf{H}$, $\mathcal{D} = (\mathcal{D}^{(0)}, \mathcal{D}^{(1)}) \mapsto \mathcal{E} = (\mathcal{E}^{(0)}, \mathcal{E}^{(1)})$ by

$$\mathcal{E}^{(i)}(X) := \mathcal{D}^{(i)}(X) \cap \bigcap_{Y \in X\psi_i} \; \bigcup_{y \in \mathcal{D}^{(1-i)}(Y)} \{y\}\rho_{1-i} \; . \tag{2}$$

Clearly, $\mathcal{E} = f(\mathcal{D}) \in \mathbf{H}$, f is a monotone mapping, and $f(\mathcal{C}_{n+1}) = \mathcal{C}_n$ for all n. Since $f(\mathcal{D}) \le \mathcal{D}$ for all $\mathcal{D} \in \mathbf{H}$, we will call f a *contraction map*. The following proposition is mentioned to shed more light on the topic only, and we will not use it in the sequel.

Proposition 1. *Given a finite context $(A^{(0)}, A^{(1)}, \rho)$, $\mathbf{T} = (\mathbf{T}, \le)$, the set of pairs of closure operators over $(A^{(0)}, A^{(1)}, \rho)$, is an (upper) semimodular coatomistic meet-semidistributive lattice, and $\mathbf{T}$ is closed with respect to the contraction map f.*

Proof. Let L_i be the poset of closure operators over $A^{(i)}$, $i = 0, 1$. Then, according to Corollaries 30 and 58 in Caspard and Monjardet [2], L_i is a lattice that has some nice properties, including those listed in the proposition. Notice that [2] attributes some of these properties to others, including Demetrovics, Libkin and Muchnik [4], Duquenne [5] and Ore [9]. Since $\mathbf{T}$ is the direct product of L_0 and L_1 and the properties we consider are clearly preserved by finite direct products, the first part of the statement is shown. The statement about the contraction map is included, modulo notational changes, in the proof of Lemma 1 in [3]. $\square$

If $\mathcal{D} \in \mathbf{H}$ and $f(\mathcal{D}) = \mathcal{D}$ then $\mathcal{D}$ is called a fixed point of f. As usual, for $\mathcal{D} \in \mathbf{H}$ the set $\{\mathcal{E} \in \mathbf{H} : \mathcal{E} \le \mathcal{D}\}$ will be denoted by $(\mathcal{D}]$. Since f is a monotone contraction map, $(\mathcal{D}]$ is always closed with respect to f. Remembering that $(A^{(0)}, A^{(1)}, \rho)$ is assumed to be finite, we have the following theorem.

Theorem 1. *$\mathcal{C}$ is the largest fixed point of f in $(\mathcal{G}]$. I.e., $f(\mathcal{C}) = \mathcal{C}$, and for every $\mathcal{D} \in (\mathcal{G}]$, $f(\mathcal{D}) = \mathcal{D}$ implies $\mathcal{D} \le \mathcal{C}$.*

Proof. Since the context is finite, there is a k with $\mathcal{C}_k = \mathcal{C}_{k+1} = \mathcal{C}$. Hence $f(\mathcal{C}) = f(\mathcal{C}_k) = \mathcal{C}_{k+1} = \mathcal{C}$, so $\mathcal{C}$ is a fixed point of f. Clearly, $\mathcal{C} \in (\mathcal{G}]$.

Now let $\mathcal{D} \in (\mathcal{G}]$ be an arbitrary fixed point. Then, using that f is monotone, $\mathcal{D} = f(\mathcal{D}) \le f(\mathcal{G}) = \mathcal{C}_1$. So $\mathcal{D} = f(\mathcal{D}) \le f(\mathcal{C}_1) = \mathcal{C}_2$, etc. Thus $\mathcal{D} \le \mathcal{C}_k = \mathcal{C}$. $\square$

Even if the above theorem is a very simple statement, it is useful from algorithmic point of view. The speed of the obvious algorithm for computing $\mathcal{C}$ depends on how fast the sequence $\mathcal{C}_0 = \mathcal{G}$, $\mathcal{C}_1$, $\mathcal{C}_2, \ldots$ decreases. If we follow what formula (1) says then we obtain $\mathcal{C}_{n+1}$ from $\mathcal{C}_n$ in two steps. In the first step we compute, say, $\mathcal{C}^{(0)}_{n+1}$ from $(\mathcal{C}^{(0)}_n, \mathcal{C}^{(1)}_n)$, and then in the second step we compute $\mathcal{C}^{(1)}_{n+1}$ from $(\mathcal{C}^{(0)}_n, \mathcal{C}^{(1)}_n)$. However, we could obtain a more rapidly decreasing sequence if we performed the second step from $(\mathcal{C}^{(0)}_{n+1}, \mathcal{C}^{(1)}_n)$ instead of $(\mathcal{C}^{(0)}_n, \mathcal{C}^{(1)}_n)$. (This would also mean less memory usage, which would save some additional time, too.) We will refer to this strategy as the *modified algorithm*.

Corollary 1. *The modified algorithm computes $\mathcal{C}$, and it is at least as fast as the straightforward algorithm suggested by formula* (1). *(In fact, it is usually faster.)*

Proof. For $i \in \{0, 1\}$ we define a mapping $f_i : \mathbf{H} \to \mathbf{H}$, $(\mathcal{D}^{(0)}, \mathcal{D}^{(1)}) \mapsto (\mathcal{E}^{(0)}, \mathcal{E}^{(1)})$ such that $\mathcal{E}^{(i)}$ is defined as in formula (2) and $\mathcal{E}^{(1-i)} = \mathcal{D}^{(1-i)}$. (It follows easily from Proposition 1 that $\mathbf{T}$ is closed with respect to the contraction maps f_i, $i \in \{0, 1\}$, but we do not need this fact in the proof.) The modified algorithm produces the sequence

$$f_0(\mathcal{G}),\ f_1(f_0(\mathcal{G})),\ f_0(f_1(f_0(\mathcal{G}))),\ f_1(f_0(f_1(f_0(\mathcal{G})))),\ \ldots\ .$$

Computing two new members of this sequence needs a slightly less computer work than computing one new member of the sequence $\mathcal{C}_0 = \mathcal{G}$, $\mathcal{C}_1$, $\mathcal{C}_2$, ... Hence all we have to show is that, for every n, the $2n$-th member of the first sequence is above $\mathcal{C}$ and below $\mathcal{C}_n$. But this follows via a trivial induction, since $f(\mathcal{D}) \le f_i(\mathcal{D}) \le \mathcal{D}$ and $f(f(\mathcal{D})) \le f_1(f_0(\mathcal{D})) \le f(D)$ hold for all $\mathcal{D} \in \mathbf{H}$. $\square$

It is clear from the proof of Theorem 1 and that of Corollary 1 that instead of the poset $\mathbf{H}$ we could have worked only with the lattice $\mathbf{T}$. However, the advantage of $\mathbf{H}$ is not only to make Theorem 1 stronger. In a practical calculation, like computing $\mathcal{C}^{(i)}(X)$ just for a single X, it gives a better theoretical background: we can reduce the $\mathcal{G}^{(i)}(Y_j)$'s for certain (not necessarily distinct) subsets Y_j of $A^{(0)}$ and $A^{(1)}$ according to (2) in an arbitrary order, and we do not have to care if the actual pair of operators is a pair of closure operators, the process converges to $\mathcal{C}$.

The computer program mentioned before uses the modified algorithm of Corollary 1. When $|A^{(0)}|$ or $|A^{(1)}|$ is large then it is not possible to determine $\mathcal{C}$, at least not with *this* program, for it needs $2^{|A^{(0)}|} + 2^{|A^{(1)}|}$ steps even to store $\mathcal{C}$. However, as it is clear from formula (1), we can determine $\mathcal{C}^{(i)}(X)$ for all X with $|X| \le m$ and all $i \in \{0, 1\}$ without determining $\mathcal{C}$, and this is much faster when m is not too large. The program allows $m \in \{2, \dots, 9\}$ when $|A^{(0)}|, |A^{(1)}| \le 14$, and it allows only $m = 2$ when $|A^{(0)}|, |A^{(1)}| \le 48$. However, the running time even for a single context with $m = 9$ and $|A^{(0)}| = |A^{(1)}| = 14$ is usually too long to wait for. The program can generate and test many random contexts. The experimental results obtained by the program are reported by the following tables.

size	4	5	6	8	10	12	14	20	30	40	48
$\lvert\{\text{tests}\}\rvert$	1000	1000	1000	100	100	100	100	100	100	100	100
$(-,-,-)$	549	757	889	98	100						
$([0,4],-,-)$	549	757	889	98	100	100	100				
$([0,2],-,-)$	535	707	844	97	100	100	100				
$(\{2\},-,-)$	282	517	736	94	100	100	100	100	100	100	100
$(\{2\},\subset,-)$	235	484	721	94	100	100	99	86	18	4	1
$([2,\infty),\subset,\neq)$	0	134	491	96	100						
$([2,4],\subset,\neq)$	0	134	491	96	100	100	99				
$(\{2\},\subset,\neq)$	0	134	470	92	100	100	99	86	18	4	1

Table 2

In Table 2, size denotes $|A^{(0)}| = |A^{(1)}|$, i.e., only "square" contexts have been tested. The number of contexts tested with the given size is denoted by $|\{\text{tests}\}|$. For a given context, $\mathcal{C} \neq \mathcal{G}$ can be due to some more or less trivial reason like $\mathcal{C}^{(i)}(\emptyset) \neq \mathcal{G}^{(i)}(\emptyset)$. Therefore the program counted those contexts for which there is an $i \in \{0, 1\}$ and an $X \in P(A^{(i)})$ such that $\mathcal{C}^{(i)}(X) \neq \mathcal{G}^{(i)}(X)$ and X satisfies some further conditions. These further conditions are denoted by a vector (α, β, γ). Here α is missing or it is a set of integers, like $[2, 4] = \{2, 3, 4\}$. If α is a set of integers then $|X|$ has to belong to α. If β not missing then it is the "$\subset$" sign and X has to satisfy $X \subset \mathcal{C}^{(i)}(X)$. If γ is not missing then it is the "$\neq$" sign and X has to satisfy $\mathcal{G}^{(i)}(X) \neq A^{(i)}$. For example, the row $(-,-,-)$ gives the number of contexts with $\mathcal{C} \neq \mathcal{G}$, and the entry 484 means that among 1000 random contexts there are 484 contexts containing a 2-element subset X with $\mathcal{C}^{(i)}(X) \neq \mathcal{G}^{(i)}(X)$ and $X \subset \mathcal{C}^{(i)}(X)$.

It is important to emphasize that, for each column, the program produced the given number of random contexts first, and counted those context that have the desired property only afterwards. In other words, different entries in the same column refer to the same set of random contexts. Due to the limited power of the program some entries are missing, but some obvious relations among the numbers in the same column give lower bounds for the missing entries.

We may also ask the question that if we take a random context table of size $n \times n$ and choose an $i \in \{0, 1\}$ and a subset X of $A^{(i)}$ randomly then what is the chance that $(-, -, -)$: $\mathcal{C}(X) \neq \mathcal{G}(X)$, or $([2, \infty), \subset, \neq)$: $2 \leq |X|$ and $X \subset \mathcal{C}^{(i)}(X) \neq \mathcal{G}^{(i)}(X) \neq A^{(i)}$. The experimental results for some values of n are reported in Table 3.

size	3	4	5	6	7	8	9
$\lvert\{\text{tests}\}\rvert$	1000	1000	1000	1000	1000	1000	1000
$(-, -, -)$	17	39	82	185	306	402	571
$(\{2\}, \subset, \neq)$	0	0	0	12	35	54	86

Table 3

Table 2 gives the strong belief[1] that a "medium sized" square context gives an "essentially new" $\mathcal{C}$ with high probability. Here "essentially new" means that the condition $(\{2\}, \subset, \neq)$ holds for some X. However, this probability decreases rapidly when the size of the context grows.

Let us call a random context a *p-random context*, $0 < p < 1$, if we put a cross to each entry with probability p, independently from other entries. So far we have considered 0.5-random contexts. However, we may get different results with other values of p. For example, we tested 100 p-random 40×40-sized contexts with different values of p, and counted the essentially new contexts among them (in the sense of $(\{2\}, \subset, \neq)$). The result is given by Table 4.

$100 \cdot p$	10	20	30	40	50	60	70	80	90
$(\{2\}, \subset, \neq)$	100	68	14	5	2	3	23	64	77

Table 4

Finally, we tested some contexts from the real life: essentially all those contexts from Ganter and Wille [6] which are given by simple context tables (with $\times$ being the only entry) and whose size fits into the program. (Sometimes the context was given by a multi-valued table and we had to reduce it.)

	1.1	1.5a	1.5b	1.13	1.16	1.21	1.23	1.24	2.4	2.13	2.15
$\lvert A^{(0)}\rvert$	8	8	5	5	14	6	6	8	7	12	14
$\lvert A^{(1)}\rvert$	9	5	4	25	16	12	8	8	7	9	9
$(\{2\}, \subset, \neq)$	no	no	no	yes	no	no	yes	no	yes	no	yes
$([0, 6], -, -)$	yes	no	no	yes		yes	yes	no	yes	no	yes

Table 5

[1] Theoretically there could be some unknown hidden connection between $\mathcal{C}$ and the built-in random number generator and this could mislead us, but the chance of this is minimal.

3. Lattices and posets

There are many frequently used relations, and therefore contexts, in lattices and posets (=partially ordered sets). It is probably not always possible to describe those with $\mathcal{C} \neq \mathcal{G}$ in an elegant way. However, there are some particular relations where something interesting can be stated. As usual, for a finite lattice L the set of nonzero join irreducible elements is denoted by $J = J(L)$ while $M = M(L)$ is the set of meet irreducible elements distinct from 1. The context $(J, M, \leq)$ is famous since Wille [10] has pointed out that its concept lattice is isomorphic to L, so this context is a very economic way to describe L up to isomorphism.

Theorem 2. *Let L be a finite lattice. If L is modular then*

$$\mathcal{C}(J(L), M(L), \leq) = \mathcal{G}(J(L), M(L), \leq).$$

Note that the converse is unfortunately not true. There are nonmodular lattices, like the five element ones: N_5 and M_3, for which $\mathcal{C} = \mathcal{G}$. But there are a plenty of nonmodular lattices for which $\mathcal{C} \neq \mathcal{G}$. The simplest such example is perhaps an n-crown with additional 0 and 1 for $n \geq 4$, i.e. the $(2n+2)$-element lattice $(\{0, a_0, \ldots, a_{n-1}, b_0, \ldots, b_{n-1}, 1\}, \leq)$ where the $\{a_0, \ldots, a_{n-1}\}$ is the set of atoms, $\{b_0, \ldots, b_{n-1}\}$ is the set of coatoms, and $a_j < b_k$ iff $k \in \{j, j+1\}$ where $j+1$ is understood modulo n.

Given a context $(A^{(0)}, A^{(1)}, \rho)$, by the *dual context* we mean

$$(A^{(1)}, A^{(0)}, \rho^{-1}).$$

Clearly, if $L_d = (L_d, \leq_d)$ denotes the dual of L, then $(J(L_d), M(L_d), \leq_d)$ is the dual of the context $(J(L), M(L), \leq)$. Similar observations are valid for posets occurring in the next theorem. Hence the lattice duality principle extends to our case and can be used in our proofs.

Proof. Let L be a finite modular lattice. Its ordering relation will also be denoted by $\rho = \rho_0$. Since modularity is a selfdual lattice property, by the duality principle it suffices to show that $\mathcal{C}^{(0)} = \mathcal{G}^{(0)}$. Let $X = \{a_1, \ldots, a_n\} \subseteq J$ with $|X| = n \geq 1$ and let

$$x \in \mathcal{G}^{(0)}(X) = (X\rho_0)\rho_1 = ([a_1 \vee \cdots \vee a_n) \cap M)\rho_1 = (a_1 \vee \cdots \vee a_n] \cap J$$

be an arbitrary element. Let $Y = \{b_1, \ldots, b_n\} \in X\psi_0$. This means that $a_j \leq b_j \in M$ for $j = 1, \ldots, n$ (but the b_j are not necessarily distinct). Then, dually to the above displayed formula, $\mathcal{G}^{(1)}(Y) = [b_1 \wedge \cdots \wedge b_n) \cap M$. Let $b = b_1 \wedge \cdots \wedge b_n$. According to formula (1), we have to show that

$$\text{there exists a } y \in [b) \cap M \text{ such that } x \leq y.$$

This is evident when $x \vee b \neq 1$, for $[x \vee b) \cap M$ is not empty in this case. So we assume that $x \vee b = 1$. Then

$$\begin{gathered} 1 = x \vee b = (b_1 \wedge \cdots \wedge b_n) \vee x \leq b_1 \vee \cdots \vee b_n \vee x = \\ (a_1 \vee b_1) \vee \cdots (a_n \vee b_n) \vee x = \\ (b_1 \vee \cdots \vee b_n) \vee (a_1 \vee \cdots \vee a_n \vee x) = \\ = (b_1 \vee \cdots \vee b_n) \vee (a_1 \vee \cdots \vee a_n) = \\ (a_1 \vee b_1) \vee \cdots (a_n \vee b_n) = b_1 \vee \cdots \vee b_n. \end{gathered}$$

Hence $b_1 \vee \cdots \vee b_n = 1$ and this happens in the interval $[b, 1] = [b, b \vee x]$. Since L is modular, this interval is isomorphic to the interval $[b \wedge x, x]$. But $x \in J$, so x is join irreducible also in the interval $[b \wedge x, x]$, whence 1 is join irreducible in $[b, 1]$, and we conclude that $b_j = 1$ for some j. But this is a contradiction, for $b_j \in M$ and $1 \notin M$. Thus we have shown that $\mathcal{C}_1^{(0)}(X) = \mathcal{G}^{(0)}(X)$ when $X \neq \emptyset$. When X is the empty set then $\emptyset\rho_0 = M$ and therefore $\mathcal{C}_1^{(0)}(\emptyset) = \mathcal{G}^{(0)}(\emptyset) = \emptyset$. This proves $\mathcal{C}_1^{(0)} = \mathcal{G}^{(0)}$, whence $\mathcal{C}^{(0)} = \mathcal{G}^{(0)}$. $\square$

In order to formulate the main theorem, we need some definitions. Let $Q = (Q, \leq)$ be a finite poset. Let $\max(Q)$ resp. $\min(Q)$ denote the set of maximal resp. minimal elements of Q. Notice that Q is an antichain iff $\max(Q) = \min(Q) = Q$. If Q is a chain then the length of Q, denoted by $\mathrm{length}(Q)$, is $|Q| - 1$. In the general case, $\mathrm{length}(Q)$ is the maximum of the set $\{\mathrm{length}(C) : C \subseteq Q \text{ and } C \text{ is a chain}\}$. If X is a subset of Q then $L(X)$ denotes the set of lower bounds of X:

$$L(X) = \{y \in Q : y \leq x \text{ for all } x \in X\},$$

and, dually, $U(X)$ denotes the set of upper bounds of X. In particular, $U(\emptyset) = L(\emptyset) = Q$. For $X = \{x_1, \ldots, x_n\}$ we will write $U(x_1, \ldots, x_n)$ instead of $U(\{x_1, \ldots, x_n\})$, and the same convention applies for L. The disjoint union (or cardinal sum) of the posets $(Q_1, \leq_1)$ and $(Q_2, \leq_2)$ is $(Q_1 \cup Q_2, \leq_1 \cup \leq_2)$ where Q_1 is assumed to be distinct from Q_2. For example, an n-element antichain is the disjoint union of n chains of length 0.

Finally, we have to define three kinds of posets, cf. also Figure 1. For $1 \leq m, n$ we define an $(m + n + 1)$-element poset $T_{mn} = \{a_1, \ldots, a_m,\ b,\ d_1, \ldots, d_n\}$ such that $\min(T_{mn}) = \{a_1, \ldots, a_m\}$, $\max(T_{mn}) = \{d_1, \ldots, d_n\}$ and $a_j < b < d_k$ for all $(j, k) \in \{1, \ldots, m\} \times \{1, \ldots, n\}$. In particular, T_{11} is the three element chain. For $2 \leq m, n$ we define two $(m + n)$-element posets, $G_{mn} = \{a_1, \ldots, a_m, b_1, \ldots, b_n\}$ and $H_{mn} = \{a_1, \ldots, a_m, b_1, \ldots, b_n\}$ such that $\min(G_{mn}) = \min(H_{mn}) = \{a_1, \ldots, a_m\}$, $\max(G_{mn}) = \max(H_{mn}) = \{b_1, \ldots, b_n\}$ and we have $a_j < b_k$ for all $(j, k) \in \{1, \ldots, m\} \times \{1, \ldots, n\}$ in G_{mn} while $a_j < b_k$ iff $1 \in \{k, j\}$ in H_{mn}.

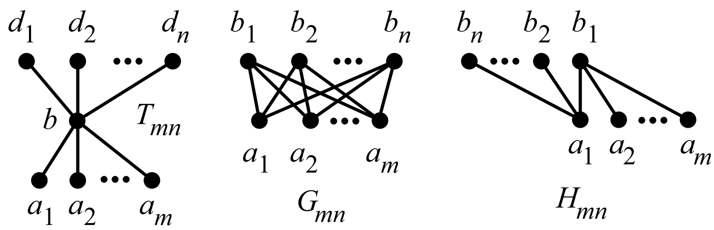

Figure 1

Theorem 3. *Let $Q = (Q, \leq)$ be a finite poset, and let us consider the contexts $(Q, Q, <)$, $(Q, Q, \prec)$, $(Q, Q, \leq)$ and $(Q, Q, \preceq)$. Then we have*

(A) $\mathcal{C}(Q, Q, <) = \mathcal{G}(Q, Q, <)$ if and only if $U(Q \setminus \max(Q)) \neq \emptyset$ and $L(Q \setminus \min(Q)) \neq \emptyset$.

(B) $\mathcal{C}(Q, Q, \prec) = \mathcal{G}(Q, Q, \prec)$ if and only if $\mathrm{length}(Q) \leq 1$, $U(Q \setminus \max(Q)) \neq \emptyset$ and $L(Q \setminus \min(Q)) \neq \emptyset$.

(C) $\mathcal{C}(Q,Q,\leq)=\mathcal{G}(Q,Q,\leq)$ *if and only if either* $|\max(Q)|=|\min(Q)|=1$*, or* $|\max(Q)|\geq 2$*,* $|\min(Q)|\geq 2$ *and*

$$(\forall x,y,z,t\in\max(Q))\ (x\neq y \text{ and } z\neq t \text{ imply } L(x,y)=L(z,t)) \text{ and}$$
$$(\forall x,y,z,t\in\min(Q))\ (x\neq y \text{ and } z\neq t \text{ imply } U(x,y)=U(z,t)).$$

(D) $(Q,Q,\preceq)$*. Then* $\mathcal{C}(Q,Q,\preceq)=\mathcal{G}(Q,Q,\preceq)$ *if and only if either* Q *is (isomorphic to)* T_{mn} *for some* $m,n\geq 1$*, or* Q *is* H_{mn} *or* G_{mn} *for some* $m,n\geq 2$*, or* $\text{length}(Q)\leq 1$ *and* Q *is a disjoint union of chains.*

Proof. Let $\rho=\rho_0$ denote the relation of the context in question, and remember that ρ_1 stands for ρ^{-1}. Since the conditions in the theorem are selfdual, by the duality principle it will suffice to deal with $\mathcal{C}^{(0)}$ and $\mathcal{G}^{(0)}$. Formula (1) will be used often without referring to it. Notice that $\mathcal{C}=\mathcal{G}$ iff $\mathcal{C}_1=\mathcal{G}$, and this fact will be used implicitly either (so we usually drop the subscript of $\mathcal{C}_1$).

(A) Suppose that $\mathcal{C}=\mathcal{C}(Q,Q,<)$ coincides with $\mathcal{G}=\mathcal{G}(Q,Q,<)$. Let $A=Q\setminus\max(Q)$ and $B=Q\setminus\min(Q)$. By way of contradiction, suppose that $U(A)$ or $L(B)$ is empty. By the duality principle, it suffices to consider the case when $U(A)$ is empty. Then $A\neq\emptyset$, $A\rho_0=\emptyset$ and $\mathcal{G}^{(0)}(A)=(A\rho_0)\rho_1=Q$. Let $x\in\max(Q)\subseteq\mathcal{G}^{(0)}(A)=\mathcal{C}^{(0)}(A)$. Clearly, $A\psi_0$ is not empty, so we can choose a $Y\in A\psi_0$. However, since x is a maximal element, $x\in\{y\}\rho_1$, i.e. $x<y$, holds for no $y\in\mathcal{G}^{(1)}(Y)$. Hence $x\notin\mathcal{C}^{(0)}(A)$, a contradiction.

To prove the converse, suppose that A has an upper bound a and B has a lower bound b. We can assume that $a\in\max(Q)$ and $b\in\min(Q)$. If A or B is empty then Q is an antichain, and $\mathcal{C}=\mathcal{G}$ follows from the fact that $X\psi_i$ is empty when X is nonempty. Hence we assume that neither A nor B is empty. Clearly, $x<a$ for all $x\in A$ and $b<y$ for all $y\in b$. Moreover, $a\in B$ and $b\in A$, and therefore $b<a$. Notice that, for any $\emptyset\neq U\subseteq Q$, $U\rho_0\subseteq B$ and $U\rho_1\subseteq A$.

Let X be a subset of Q. If $X=\emptyset$ then $X\rho_0=Q$ gives $\mathcal{G}^{(0)}(X)=Q\rho_1=\emptyset$, so $\mathcal{C}^{(0)}(\emptyset)=\mathcal{G}^{(0)}(\emptyset)$. If $X\not\subseteq A$ then $X\psi_0=\emptyset$ yields $\mathcal{C}^{(0)}(X)=\mathcal{G}^{(0)}(X)$.

Now let $X\subseteq A$. Then $X\rho_0\supseteq\{a\}$ yields

$$\mathcal{G}^{(0)}(X)=(X\rho_0)\rho_1\subseteq\{a\}\rho_1=L(a)\setminus\{a\}\subseteq A.$$

Now let $x\in\mathcal{G}^{(0)}(X)$ and $Y\in X\psi_0$ be arbitrary. Then $Y\rho_1\subseteq A$ gives $\mathcal{G}^{(1)}(Y)=(Y\rho_1)\rho_0\supseteq A\rho_0\ni a$. Since $x\in A$, $x\in\{a\}\rho_1$. Hence a can play the role of y in formula (1) and we obtain $x\in\mathcal{C}_1^{(0)}(X)$. This shows that $\mathcal{C}_1=\mathcal{G}$, thus $\mathcal{C}=\mathcal{G}$, proving part (A) of the theorem.

(B) Suppose $\text{length}(Q)\geq 2$. Then we can choose $a,b,c\in Q$ such that $a\prec b$, $b\prec c$ and $c\in\max(Q)$. Let $X=\{a,b\}$. Then $\mathcal{G}^{(0)}(X)=(\{a,b\})\rho_0)\rho_1=\emptyset\rho_1=Q$. If $\mathcal{C}=\mathcal{G}$ then $c\in Q=\mathcal{G}^{(0)}(X)=\mathcal{C}^{(0)}(X)$ and $Y=\{b,c\}\in X\psi_0$ implies that $c\in\{y\}\rho_1$, i.e. $c\prec y$, for some $y\in\mathcal{G}^{(1)}(Y)$, which contradicts $c\in\max(Q)$. Hence $\mathcal{C}=\mathcal{G}$ implies $\text{length}(Q)\leq 1$. Then $(Q,Q,\prec)$ is exactly the same context as $(Q,Q,<)$, and the rest of part (B) follows from part (A).

(C) Consider the context $(Q,Q,\leq)$ and suppose that $\mathcal{C}=\mathcal{G}$. Suppose first that $|\min(Q)|=1$, i.e., Q has a unique least element 0. Let $X=\emptyset$. Then $\mathcal{C}^{(0)}(X)=\mathcal{G}^{(0)}(X)=Q\rho_1=\{0\}$ and $Y=\emptyset\in X\psi_0$ yields that there is a $y\in\mathcal{G}^{(1)}(Y)$ with $0=x\in\{y\}\rho_1$. Thus $\mathcal{G}^{(1)}(Y)=\mathcal{G}^{(1)}(\emptyset)=(\emptyset\rho_1)\rho_0=Q\rho_0=\{z\in Q: t\leq z \text{ for all } t\in Q\}$ is nonempty. Therefore Q has a greatest element and $|\max(Q)|=1$. Now the duality principle gives that $|\min(Q)| = 1$ iff $|\max(Q)| = 1$, and the condition of (C) holds.

Now suppose that $|\min(Q)| > 1$, then $|\max(Q)| > 1$ either. By way of contradiction let us assume that $L(u, v)$ (where $u \neq v$) is not a constant on $\max(Q)$. Then we can choose a three element subset $\{a, b, c\}$ of $\max(Q)$ such that $L(a, b) \nsubseteq L(a, c)$. Then there is an element $x \in L(a, b) \setminus L(a, c)$. Let $X = \{a, b\}$. We obtain $\mathcal{G}^{(0)}(X) = (X\rho_0)\rho_1 = L(U(a, b)) = L(\emptyset) = Q$, so $c \in \mathcal{G}^{(0)}(X)$. Let $Y = X$. Then $Y \in X\psi_0$ and $\mathcal{C} = \mathcal{G}$ imply that there is an element $y \in \mathcal{G}^{(1)}(Y)$ with $c \in \{y\}\rho_1$, i.e, $c \leq y$. Since $c \in \max(Q)$, $c = y \in \mathcal{G}^{(1)}(Y) = (Y\rho_1)\rho_0 = U(L(a, b))$. This and $x \in L(a, b)$ yield $x \leq c$, contradicting $x \in L(a, b)\setminus L(a, c)$. This shows that L is constant on $\{(u, v) : u, v \in \max(Q)$ and $u \neq v\}$. It follows from the duality principle that U is constant on $\{(u, v) : u, v \in \min(Q)$ and $u \neq v\}$.

Now, to prove the converse, suppose first that $0, 1 \in Q$, i.e., $|\max(Q)| = |\min(Q)| = 1$. Then $1 \in U(Q) = U(L(\emptyset)) = \mathcal{G}^{(1)}(\emptyset)$. Since $\mathcal{G}^{(1)}$ is monotone, $1 \in \mathcal{G}^{(1)}(Y)$ for any $Y \subseteq Q$. Moreover, $\{1\}\rho_1 = Q$. Hence 1 can always serve as y in formula (1), and we conclude that $\mathcal{C} = \mathcal{G}$.

From now on we suppose that $|\max(Q)| = |\min(Q)| \geq 2$, L is constant on $\{(u, v) : u, v \in \max(Q)$ and $u \neq v\}$ and U is constant on $\{(u, v) : u, v \in \min(Q)$ and $u \neq v\}$. Then $\mathcal{G}^{(0)}(\emptyset) = L(U(\emptyset)) = L(Q) = \emptyset$ gives $\mathcal{C}^{(0)}(\emptyset) = \mathcal{G}^{(0)}(\emptyset)$.

Now let us consider a nonempty subset X of Q, and an arbitrary $Y \in X\psi_0$. Then Y is nonempty either. We distinguish two cases according to $U(Y)$.

First suppose that $U(Y)$ is nonempty, and let us fix an element $z \in U(Y)$. Since $Y \in X\psi_0$, $U(X) \supseteq U(Y)$, so $U(X) \supseteq \{z\}$, whence $\mathcal{G}^{(0)}(X) = L(U(X)) \subseteq L(\{z\}) = \{z\}\rho_1$. On the other hand, the transitivity of the ordering gives $U(Y) \subseteq U(L(Y)) = \mathcal{G}^{(1)}(Y)$, whence $z \in \mathcal{G}^{(1)}(Y)$. Now it is clear from formula (1) that $\mathcal{C}^{(0)}(X) = \mathcal{G}^{(0)}(X)$.

Secondly, we suppose that $U(Y)$ is empty. Then there are $y_1, y_2 \in Y$ and $z_1, z_2 \in \max(Q)$ such that $y_1 \leq z_1$, $y_2 \leq z_2$ and $z_1 \neq z_2$. Since $\mathcal{G}^{(1)}(Y) = U(L(Y))$ is an order filter including Y, $\{z_1, z_2\} \subseteq \mathcal{G}^{(1)}(Y)$. Now let x be an arbitrary element of $\mathcal{G}^{(0)}(X)$, and choose an element $\tilde{x} \in \max(Q)$ such that $x \leq \tilde{x}$. If $\tilde{x} = z_j$ for some $j \in \{1, 2\}$ then we can chose $y = \tilde{x} = z_j$ in formula (1). Hence we can assume that $|\{\tilde{x}, z_1, z_2\}| = 3$. Using the assumption that L is constant for distinct maximal elements we obtain

$$\tilde{x} \in \mathcal{G}^{(1)}(\{\tilde{x}, z_1\}) = U(L(\tilde{x}, z_1)) = U(L(z_1, z_2)) = \\ \mathcal{G}^{(1)}(\{z_1, z_2\}) \subseteq \mathcal{G}^{(1)}(\mathcal{G}^{(1)}(Y)) = \mathcal{G}^{(1)}(Y),$$

and therefore the choice $y = \tilde{x}$ for formula (1) works again. This shows that $\mathcal{C}_1^{(0)}(X) = \mathcal{G}^{(0)}(X)$ for any $X \in P(A^{(0)})$. So $\mathcal{C} = \mathcal{G}$, proving part (C).

(D) Consider the context $(Q, Q, \preceq)$ and suppose Q is one of the posets listed in (D). We need to show that $\mathcal{C} = \mathcal{G}$. If $\mathrm{length}(Q) \leq 1$ then $(Q, Q, \preceq)$ coincides with $(Q, Q, \leq)$ and part (C) easily implies that $\mathcal{C} = \mathcal{G}$. So we can assume that Q is T_{mn} for some $m, n \geq 1$. By the duality principle, it suffices to show that $\mathcal{C}^{(0)} = \mathcal{G}^{(0)}$. Let

$$K = \{X \in P(Q) : (\forall Z \in P(Q))\, (Z \subset X \Rightarrow \mathcal{G}^{(0)}(Z) \subset \mathcal{G}^{(0)}(X))\}.$$

If $\mathcal{C}^{(0)}$ and $\mathcal{G}^{(0)}$ would agree on K then for any $X \in P(Q)$ we could take a minimal element Z of $\{X' \in P(Q) : \mathcal{G}^{(0)}(X') = \mathcal{G}^{(0)}(X)\}$, and from $Z \in K$ we could deduce

$$\mathcal{C}^{(0)}(X) \supseteq \mathcal{C}^{(0)}(Z) = \mathcal{G}^{(0)}(Z) = \mathcal{G}^{(0)}(X),$$

implying $\mathcal{C}^{(0)} = \mathcal{G}^{(0)}$.

Hence it suffices to show that for all $X \in K$, $\mathcal{C}^{(0)}(X) = \mathcal{G}^{(0)}(X)$. Moreover, it suffices to consider a small subset K' of K such that for all X in K there is an automorphism of Q which maps X to an element of K'. Let $A = \{a_1, \dots, a_m\}$, $A^+ = A \cup \{b\}$, $D = \{d_1, \dots, d_n\}$, $D^+ = D \cup \{b\}$, and assume that $m \geq 2$ and $n \geq 2$. (The case $m = 1$ or $n = 1$ is simpler and will not be detailed.) Then an appropriate K' is given by the second row in Table 6, where, for brevity, we write x, y instead of $\{x, y\}$:

X	$\emptyset$	a_1	a_1, a_2	b	a_1, b	a_1, d_1	b, d_1	d_1	d_1, d_2
$X \in K'$?	yes	yes	yes	yes	yes	yes	no	yes	yes
$\mathcal{G}^{(0)}(X)$	$\emptyset$	a_1	A^+	b	A^+	Q	b, d_1	b, d_1	Q
$\mathcal{G}^{(1)}(X)$	$\emptyset$	a_1, b	Q	b	a_1, b	Q	D^+	d_1	D^+

Table 6

Now we can easily list all possible Y's from formula (1) (up to isomorphism, again) and check that $\mathcal{C}^{(0)}(X) = \mathcal{G}^{(0)}(X)$ for $X \in K'$; the tedious details will be omitted.

Now, in order to prove the converse direction, assume that $\mathcal{C} = \mathcal{G}$. If $\text{length}(Q) = 0$ then Q is an antichain, which is a disjoint union of chains, and there is nothing to prove.

Now assume that $\text{length}(Q) = 1$ and Q is not a disjoint union of chains. Then part (C) of the theorem applies, so $2 \leq |\max(Q)|$, $2 \leq |\min(Q)|$, L is constant on $\{(u, v) : u, v \in \max(Q),\ u \neq v\}$ and U is constant on $\{(u, v) : u, v \in \min(Q), u \neq v\}$. Since Q is not a disjoint union of chains, there are $a_1, b_1, b_2 \in Q$ such that $a_1 < b_1$ and $a_1 < b_2$, or dually. So we can assume that $a_1 < b_1$ and $a_1 < b_2$. If $c \in \max(Q) \cap \min(Q)$ then $\emptyset = L(b_1, c) \neq L(b_1, b_2) \supseteq \{a_1\}$ would lead to a contradiction. Therefore Q is the disjoint union of $\max(Q)$ and $\min(Q)$. Notice also that the Hasse diagram of Q is connected as a graph, for otherwise we could find an $x \in \max(Q)$ with $L(b_1, x) = \emptyset$. Let

$$B = \{x \in \max(Q) : a_1 \leq x\}, \text{ and remember that } b_1, b_2 \in B.$$

Since a_1 is connected with other elements of $\min(Q)$ in the graph, there is an $a_2 \in \min(Q) \setminus \{a_1\}$ which is less than some element of B. So we can assume that $a_2 < b_1$. Let

$$A := \{x \in \min(Q) : x < b_1\}, \text{ and notice that } a_1, a_2 \in A.$$

If $c \in \max(Q) \setminus B$ then $a_1 \notin L(b_1, c) = L(b_1, b_2) \supseteq \{a_1\}$ would be a contradiction. Hence $B = \max(Q)$, and we obtain $A = \min(Q)$ similarly.

Now Q is the disjoint union of A and B. Let $m = |A|$ and $n = |B|$. If, for $a \in A$ and $b \in B$, $a < b$ holds only when $\{a_1, b_1\} \cap \{a, b\} \neq \emptyset$ then Q is H_{mn}. Otherwise we may suppose that $a_2 < b_2$. Then, for any $b \in B \setminus \{b_1\}$, $a_2 \in L(b_1, b_2) = L(b_1, b)$ yields $a_2 < b$. Hence $U(a_1, a_2) = B$, and for any $a \in A \setminus \{a_1\}$ we have $U(a_1, a) = U(a_1, a_2) = B$. This means that $Q = G_{mn}$, and the case $\text{length}(Q) = 1$ is settled.

Now suppose that $\mathrm{length}(Q) \geq 2$ and introduce the notation

$$\mathrm{mid}(Q) = Q \setminus (\max(Q) \cup \min(Q)).$$

Let us observe that for any $u, v \in Q$,

$$\text{if } u \prec v \text{ then } \mathcal{G}^{(0)}(\{u,v\}) = \{x : x \preceq v\}$$
$$\text{and } \mathcal{G}^{(1)}(\{u,v\}) = \{x : u \preceq x\}. \tag{3}$$

Indeed, $\mathcal{G}^{(0)}(\{u,v\}) = (\{u,v\}\rho_0)\rho_1 = \{v\}\rho_1 = \{x : x \preceq v\}$, and the other equation follows by duality.

First we consider the case when $\mathrm{length}(Q) \geq 3$. Then there are elements $a, b, c \in Q$ and $d \in \max(Q)$ such that $a \prec b \prec c \prec d$. Let $X = \{b, d\}$. Then $\mathcal{C}^{(0)}(X) = \mathcal{G}^{(0)}(X) = (X\rho_0)\rho_1 = \emptyset\rho_1 = Q$. Let $Y = \{c, d\} \in X\psi_0$. Then for any $y \in \mathcal{G}^{(1)}(Y)$ we have $c \leq y$ by (3), so $a \not\succeq y$, whence $a \notin \{y\}\rho_1$, and $a \notin \mathcal{C}^{(0)}(X) = Q$ by formula (1), a contradiction. Hence $\mathrm{length}(Q) \geq 3$ is excluded, and from now on we assume that $\mathrm{length}(Q) = 2$.

The first step in the case $\mathrm{length}(Q) = 2$ is to show that for any $b \neq c$

$$\text{if } b, c \in \mathrm{mid}(Q) \text{ and } b \parallel c \text{ then } |L(b,c)| \leq 1 \text{ and } |U(b,c)| \leq 1. \tag{4}$$

Suppose, by way of contradiction, that $d_1, d_2 \in U(b,c)$ and $d_1 \neq d_2$. Let $X = \{d_1, d_2\}$, and choose an element a such that $a \prec b$. Since $X \subseteq \max(Q)$, we obtain that $a \in Q = \emptyset\rho_1 = \mathcal{G}^{(0)}(X) = \mathcal{C}^{(0)}(X)$. Let $Y = X \in X\psi_0$. By formula (1) there is a $y \in \mathcal{G}^{(1)}(Y)$ with $a \preceq y$. But $Y\rho_1 \supseteq \{b, c\}$ implies $y \in \mathcal{G}^{(1)}(Y) = (Y\rho_1)\rho_0 \subseteq \{b,c\}\rho_0$, i.e., $b \preceq y$ and $c \preceq y$. Since $b \parallel c$, we obtain $a \prec b \prec y$, which contradicts $a \preceq y$. This and the duality principle prove (4).

Now, to sharpen the previous assertion, we prove that for any $b \neq c$

$$\text{if } b, c \in \mathrm{mid}(Q) \text{ and } b \parallel c \text{ then } L(b,c) = U(b,c) = \emptyset. \tag{5}$$

Suppose the contrary. By the duality principle, we may assume that $L(b,c)$ is nonempty. Let $L(b,c) = \{a\}$. We can choose $d_1, d_2 \in \max(Q)$ such that $b \prec d_1$ and $c \prec d_2$. If possible, then we choose them equal: $d_1 = d_2$. Let $X = \{b, c\}$. If $U(b,c)$ is nonempty then $d_1 = d_2$, $X\rho_0 = \{d_1\}$ and we have $d_1 \in \mathcal{G}^{(0)}(X) = \mathcal{C}^{(0)}(X)$. If $U(b,c)$ is empty then so is $X\rho_0$ and we have $d_1 \in \mathcal{G}^{(0)}(X) = \mathcal{C}^{(0)}(X)$ again. Let $Y = X = X\psi_0$. Then, by formula (1), $d_1 \preceq y$ for some $y \in \mathcal{G}^{(1)}(Y)$. Since $d_1 \in \max(Q)$, $d_1 = y \in \mathcal{G}^{(1)}(Y) = (Y\rho_1)\rho_0 = \{a\}\rho_0$. This gives $a \preceq d_1$, which contradicts $a \prec b \prec d_1$. This shows (5).

Based on (5) we can prove even more: for any elements of Q we have

$$\text{if } c \in Q,\ b \in \mathrm{mid}(Q) \text{ and } b \parallel c \text{ then } L(b,c) = U(b,c) = \emptyset. \tag{6}$$

Suppose the contrary. By (5), $c \notin \mathrm{mid}(Q)$. By the duality principle we can assume that $c \in \max(Q)$. Then $U(b,c) = \emptyset$. Let $a \in L(b,c)$ and choose an element $d \in \max(Q)$ with $b \prec d$. For $X = \{a, d\}$ from $X\rho_0 = \emptyset$ we obtain $c \in Q = \mathcal{G}^{(0)}(X) = \mathcal{C}^{(0)}(X)$. Let $Y = \{b, d\} \in X\psi_0$. Then $c \preceq y$ for some $y = \mathcal{G}^{(1)}(Y)$ by (1) and $b \preceq y$ by (3). This together with $b \parallel c$ imply $c \prec y$, which contradicts $c \in \max(Q)$. This proves (6)

Now we are in the position to show that

$$\text{if } b \in \mathrm{mid}(Q) \text{ then there is no } c \in Q \text{ with } b \parallel c. \tag{7}$$

Suppose the contrary, and choose $a, d \in Q$ with $a \prec b \prec d$. If $X = \{a, d\}$ and $Y = \{b, d\} \in X\psi_0$ then, exactly the same way as in the previous step, we obtain an element y with $c \preceq y$ and $b \preceq y$, and we conclude that $y \in U(b, c)$, which contradicts (6). This proves (7)

Now, since length$(Q) = 2$, we can choose elements $a_1 \prec b \prec d_1$ in Q. It follows from (7) that for any further element x either $x < b$ or $b < x$. Let $m = |\{x \in Q : x < b\}|$ and $n = |\{x \in Q : b < x\}|$, then clearly Q is T_{mn}. □

We conclude the paper by open problem about $\mathcal{C}$. For motivation and a possible application cf. [3]. Let us say that $(A^{(0)}, A^{(1)}, \rho)$ is a decomposable context if there are nonempty sets $B^{(i)}$ and $C^{(i)}$ with $B^{(i)} \cup C^{(i)} = A^{(i)}$ and $B^{(i)} \cap C^{(i)} = \emptyset$ such that

$$\rho = (\rho \cap (B^{(0)} \times B^{(1)})) \cup (\rho \cap (C^{(0)} \times C^{(1)})).$$

Otherwise $(A^{(0)}, A^{(1)}, \rho)$ is called an *indecomposable context*. We say that it is a *uniform context* if $|\{x\}\rho_i| = |\{y\}\rho_i|$ for all $x, y \in A^{(i)}$. In the terminology of context tables, if any two columns contain the same number of crosses and any two rows contain the same number of crosses. For example, each finite block design (P, B, I) and, in particular, each finite projective space (P, L, I) is a uniform context.

Problem 1. *Is it true that for each indecomposable uniform context* $(A^{(0)}, A^{(1)}, \rho)$ *with* $|A^{(0)}| \geq 3$ *and* $|A^{(1)}| \geq 3$ *there exists an* $i \in \{0, 1\}$ *and there are* $x, y, z \in A^{(i)}$ *such that*

$$\mathcal{C}^{(i)}(\{x, y\}) \cap \mathcal{C}^{(i)}(\{y, z\}) \cap \mathcal{C}^{(i)}(\{z, x\}) = \emptyset?$$

References

[1] R. Agarwal, T. Imielinski and A. Swami: Mining association rules between sets of items in large databases, in *Proceedings of the 1993 ACM SIGMOND international conference on Management of Data (SIGMOND'93)*, pp. 207-216, ACM Press, May 1993.

[2] N. Caspard and B. Monjardet: The lattice of closure systems, closure operators, and implicational systems on a finite set: a survey, *Discrete Applied Mathematics* **127** (203), 241-269.

[3] G. Czédli: 2-uniform congruences in majority algebras and a closure operator, *Algebra Universalis*, to appear

[4] J. Demetrovics, L.0. Libkin and I.B. Muchnik: Functional dependencies and the semilattice of closed classed, *Proceedings of MFDBS 89, Lecture Notes in Computer Sci.,* vol. **364**, Springer, Berlin, 1989, pp. 136–147.

[5] V. Duquenne: The lattice of all 1-meet-subsemilattices of a finite lattice, 1987 (unpublished notes).

[6] B. Ganter and R. Wille: *Formal Concept Analysis: Mathematical foundations*, Springer-Verlag, Berlin — Heidelberg 1999.

[7] L. Lakhal and G. Stumme: Efficient mining of association rules based on formal concept analysis, in *Formal Concept Analysis*, Lecture Notes in Computer Science, vol. 3626/2005 (editors: B. Ganter, G. Stumme and R. Wille), pp. 180–195, Springer-Verlag, Berlin — Heidelberg 2005.

[8] B. Polster: *The mathematics of juggling*, Springer-Verlag, New York 2003.

[9] O. Ore: Combinations of closure relations, *Ann. Math.* **14** (1943), 514–533.

[10] R. Wille: Restructuring lattice theory: an approach based on hierarchies of concepts: *Ordered sets* (Banff, Alta., 1981) *NATO Adv. Study Inst. Ser. C: Math. Phys. Sci.* **83**, Reidel, Dordrecht-Boston, Mass., 1982, pp. 445–470.

University of Szeged, Bolyai Institute, Szeged, Aradi vértanúk tere 1, HUNGARY 6720

E-mail address: `czedli@math.u-szeged.hu`

URL: `http://www.math.u-szeged.hu/~czedli/`