Order **16:** 231–244, 1999.

On Edge Decompositions of Posets

OLEG PIKHURKO★
Department of Pure Mathematics and Mathematical Statistics, Cambridge University, Cambridge CB2 1SB, U.K. E-mail: o.pikhurko@dpmms.cam.ac.uk

(Received: 19 February 1998; accepted: 6 March 2000)

Abstract. An *edge decomposition* of a poset $\mathcal{P}$ is a collection of chains such that every pair of elements of which one covers the other belongs to exactly one chain. We consider this and the related notion of the *line poset* $L(\mathcal{P})$ which consists of pairs of adjacent elements of $\mathcal{P}$ so that $(x \lessdot y) <_{L(\mathcal{P})} (x' \lessdot y')$ iff $y \leq_{\mathcal{P}} x'$. We present some min-max type results on path-cycle partitions of digraphs which are applicable to poset decompositions. Providing an explicit construction we show that the lattice of the subsets of an n-set admits an edge decomposition into symmetric chains. We demonstrate a few applications of this decomposition. Also, a characterisation of line posets is given.

Mathematics Subject Classification (1991): 06A06.

Key words: chain decomposition, Dilworth's theorem, symmetric chain order.

1. Introduction

There are many important results about decompositions of posets into disjoint chains, when we require that every element of the poset belongs to exactly one chain. We will refer to these as *vertex decompositions*. Typical questions are the following. What is the minimal number of chains of such a partition? Do there exist partitions with some extra properties? Are there any applications of these decompositions?

In this paper we investigate the notion of an *edge decomposition* which is a collection of chains such that every pair of adjacent elements (one covers the other) belongs to exactly one chain and we try to answer the questions above.

Such considerations may arise, for example, when in a computer program we want to operate with posets, and so we wish to represent them efficiently in the memory. If keeping the relational binary $n \times n$-table is impossible or undesirable, we can try to maintain a list of chains completely determining the poset, and a natural question to ask is, for example, how small such a list can be. The related notion of *line poset* (defined later) also arises naturally.

★ Supported by an External Research Studentship, Trinity College, Cambridge.

In Section 3 we present formulas for the minimal number of paths in a cycle-path vertex/edge partition of a digraph. Their corollaries on poset decompositions can be viewed as an analogue of Dilworth's theorem [7].

In Section 4 we provide an explicit edge decomposition of the all-important lattice of subsets of a finite set into symmetric chains. Although the existence of such a partition can be deduced from the results of Anderson [1] and Griggs [10], a constructive proof is a new result. The partition discovered has some extra properties, as is asserted by Theorem 5, and applications: an upper bound on the number of antichains in $L(\mathcal{B}_n)$, a construction of a pair of orthogonal edge partitions of $\mathcal{B}_n$ for odd n, an application to a storage and retrieval problem and to a numerical problem. For details we refer the reader to Section 5.

Finally, we characterise line posets in terms of forbidden configurations and point out which information determines and can be reconstructed from the line poset.

2. Definitions

Let $\mathcal{P} = (X, >)$ be a poset. We say y *covers* x (denoted by $y \gtrdot x$ or $x \lessdot y$) if $y > x$ and no $z \in X$ satisfies $x < z < y$ (such x, y will be also called *adjacent elements*). With every poset $\mathcal{P}$ we associate its *Hasse diagram* $D = D(\mathcal{P})$ which is the digraph with X as the vertex set and $(x, y) \in E(D)$ iff y covers x. Sometimes we switch between the poset and digraph terminology.

A chain in $\mathcal{P}$ is called *skipless* if every element covers its predecessor; skipless chains correspond to *oriented paths* in its Hasse diagram. The *width* $w(\mathcal{P})$ is the maximal size of an antichain in $\mathcal{P}$.

The *line poset* $L(\mathcal{P})$ of a poset $\mathcal{P}$ has as vertices the pairs (x, y) of elements in $\mathcal{P}$ with y covering x and we agree that $(x \lessdot y)$ is less than $(x' \lessdot y')$ in $L(\mathcal{P})$ iff $y \leq x'$. (This operation somewhat resembles taking the line graph, hence the name.)

Every skipless chain in $\mathcal{P}$ corresponds to a skipless chain in $L(\mathcal{P})$ of size smaller by 1. We usually identify these chains.

One can ask which important poset properties are preserved by the operator L. In fact, L preserves very few properties (e.g., self-duality, regularity). As in almost every case it is trivial to find a counterexample/proof we do not dwell on this topic.

A *vertex* (or *chain*) *partition* (or *decomposition*) of $\mathcal{P}$ is a collection of chains such that every $x \in X$ belongs to exactly one chain. An *edge partition* (or *decomposition*) is a family of skipless chains such that every pair $x, y \in X$ with x being covered by y, belongs to exactly one chain. Note that the chains in an edge decomposition are required to be skipless. One can see that edge partitions of $\mathcal{P}$ correspond to skipless chain partitions of $L(\mathcal{P})$.

The subsets of the set $[n] = \{1, \dots, n\}$ partially ordered via the inclusion relation, form the ranked poset $\mathcal{B}_n = (2^{[n]}, \subset)$. The corresponding Hasse diagram is the *oriented n-cube* Q_n. For $\mathcal{B}_n$ the relation 'B covers A' is denoted by $A \sqsubset B$.

We find it useful to identify $A \in \mathcal{B}_n$ with its ()-representation which is the n-sequence of left and right parentheses corresponding to the elements of $\overline{A} = [n] \setminus A$ and A respectively. Likewise, the $(*)$-representation of an element $(A \sqsubset B) \in L(\mathcal{B}_n)$ contains '(' for the elements in $\overline{B}$, ')' for the elements in A and '$*$' for the element in $B \setminus A$.

Generally, let F be a sequence containing left and right parentheses. Consecutively and as long as possible remove matched pairs of adjacent brackets, i.e., substrings '( )'. Clearly, the order of operations does not matter. The elements which would be removed by the above *matching* are called *fixed* or *paired* elements and the remaining ones are called *free*. In particular, the free parentheses always form the following, possibly empty, subsequence:)) . . .)) ((. . . ((.

3. Edge Decompositions of Minimum Size

Here we present a few Dilworth-type theorems for digraphs and posets.

Let D be any digraph. (Here we allow loops and opposite edges.) Consider partitions of $V(D)$ into vertex-disjoint directed cycles and directed paths. (We consider a single vertex as a path of length zero; loops and pairs of opposite edges are considered as cycles.) Let $m(D)$ be the minimum number of directed paths in a such partition.

On the other hand, let $M(D)$ be the maximum of $|A| - |B|$ taken over all pairs of disjoint sets $A, B \subset V(D)$ such that any directed path connecting two distinct vertices from A contains a vertex of B and any cycle intersecting A intersects B. (In particular, if $(i, i) \in E(D)$ then $i \notin A$.) Clearly, for any such pair (A, B) we have $|P \cap A| \leq |P \cap B| + \varepsilon$, where $\varepsilon = 1$ if P is a directed path and $\varepsilon = 0$ if P is a directed cycle. This implies that $m(D) \geq M(D)$.

We will show that we have in fact equality for any D. This was originally proved by the author by using the linear programming method of Dantzig and Hoffman [6]. Here we present a much simpler argument suggested by Graham Brightwell.

THEOREM 1. *For any digraph D we have $m(D) = M(D)$.*

Proof (Brightwell). Consider the bipartite graph G on two copies of $V(D)$, say $X = \{v^\vee : v \in V(D)\}$ and $Y = \{v^\wedge : v \in V(D)\}$, where we connect $u^\vee$ to $v^\wedge$ if and only if $(u, v) \in E(D)$. It is easy to check that the number of edges missing in a maximum matching in G equals $m(D)$. By the defect form of Hall's theorem, this number equals the maximum of $|Z| - |\Gamma(Z)|$ over $Z \subset X$, where $\Gamma(Z)$ denotes the set of neighbours of Z. Choose any extremal set Z. Let

$$\begin{aligned} A &= \{v \in V(D) : v^\vee \in Z,\ v^\wedge \notin \Gamma(Z)\},\\ B &= \{v \in V(D) : v^\vee \notin Z,\ v^\wedge \in \Gamma(Z)\}. \end{aligned}$$

Let $P = \{v_1, \ldots, v_l\}$ be a directed path in D with $v_1, v_l \in A$, $l \geq 2$. As $v_l^\wedge \notin \Gamma(Z)$, we conclude that $v_{l-1}^\vee \notin Z$. As $v_1^\vee \in Z$, there must be $i \in [1, l-2]$ such that $v_i^\vee \in Z$

but $v_{i+1}^{\vee} \notin Z$. As $v_{i+1}^{\wedge} \in \Gamma(Z)$, we have $v_{i+1} \in B$. Similarly, any cycle intersecting A intersects B. Hence, $m(D) = |A| - |B| \leq M(D)$. □

Remark. For a cycle-free digraph D, the conclusion of Theorem 1 can be deduced from the result of Saks [18, Theorem 5.3] pointed to the author by one of the referees. Saks' result is more general, but it applies only to cycle-free digraphs and its proof is complicated.

The minimum number n of paths in a cycle-path decomposition of $E(D)$ can be computed by applying Theorem 1 to the appropriately defined *directed line graph* of D. However, we present a direct proof which gives a straightforward algorithm for constructing a minimum partition: remove one by one all cycles and then – maximal paths. It turns out that it is enough to consider only pairs $A, B \subset E(D)$ of the following rather special form: take a partition $X \cup Y = V(D)$, let $A = \{(x, y) \in E(D) : x \in X, y \in Y\}$, $B = \{(y, x) \in E(D) : x \in X, y \in Y\}$ and $N(X, Y) = |A| - |B|$.

THEOREM 2. *For any digraph D, the minimum number $n(D)$ of paths in a partition of $E(D)$ into directed paths and cycles is equal to*

$$N(D) = \max\{N(X, Y) : X \cup Y = V(D),\ X \cap Y = \emptyset\}.$$

Proof. It is immediate that $n(D) \geq N(D)$.

As the removal of a cycle does not affect $N(D)$, it is enough to prove the reverse inequality for a cycle-free digraph D. To apply induction on $|E(D)|$ we have to show that $N(D') < N(D)$, where D' is obtained from D by removing the edges of a maximal path $P = (x_1, \ldots, x_k)$.

To see this take a partition $X \cup Y = V(D')$ with $N(D') = N(X, Y)$. Since P is maximal and D is acyclic there is no $y \in V(D)$ with $(y, x_1) \in E(D)$. Therefore, if $x_1 \in Y$, we can move x_1 to X without decreasing $N(X, Y)$. Likewise we may assume $x_k \in Y$. But if we add back the edges of P, we will increase $N(X, Y)$ by 1: if moving along P we change side from Y to X i times, then we go from X to Y $i + 1$ times. This shows that $N(D') < N(D)$ as required. □

The following corollary is obtained by applying Theorem 1 or 2 to the Hasse diagram of a poset $\mathcal{P}$.

COROLLARY 3. *The minimum size of a skipless chain decomposition of a poset $\mathcal{P}$ equals the maximum of $|A| - |B|$ over all disjoint sets $A, B \subset \mathcal{P}$ such that any skipless chain containing two elements from A intersects B.*

The minimum size of an edge decomposition of $\mathcal{P}$ equals the maximum of $e(X, Y) - e(Y, X)$ over all partitions $\mathcal{P} = X \cup Y$, where $e(X, Y)$ denotes the number of elements $(x \lessdot y) \in L(\mathcal{P})$ with $x \in X$ and $y \in Y$.

4. Symmetric Edge Partitions of Cubes

The fundamental result of de Brujin, Kruyswijk and Tengbergen [4] (see, e.g., [2, Section 3.1] for a proof) asserts that $\mathcal{B}_n = (2^{[n]}, \subset)$ is a *symmetric chain order*, that is, admits a decomposition into *symmetric chains*. (A chain $x_1 < \cdots < x_k$ in a ranked poset $(\mathcal{P}, r)$ is called *symmetric* if it is skipless and $r(x_1) = r(\mathcal{P}) - r(x_k)$.) This was strengthened by Anderson [1] and Griggs [10], who showed that a LYM poset $\mathcal{P}$ with a unimodal symmetric rank-sequence is a symmetric chain order. (Note that the number of chains is $w(\mathcal{P})$ – minimal possible.)

The latter result is applicable to $L(\mathcal{B}_n)$, which as a regular poset has the LYM property. However, this way we obtain a purely existentional result while one would wish to have an explicit decomposition. Here we provide an explicit construction, which like that of Greene and Kleitman [9] and Leeb (unpublished) on $\mathcal{B}_n$, utilises bracket representations.

THEOREM 4. *For every n, $L(\mathcal{B}_n)$ is a symmetric chain order. In other words, $\mathcal{B}_n$ admits an edge decomposition into symmetric chains.*

Proof. Assume that the numbers $1, \ldots, n$ are placed on a circle clockwise in this order. Let σ denote the *shift permutation* which maps every element to the next position clockwise: $\sigma(k) = k + 1 \pmod n$ and let $\sigma^{(i)}$ be its ith iterate. (These are referred to also as *rotations*.) For clarity of language we use the same symbol σ for the corresponding action on the vertex set and the edge set of Q_n. We will produce a σ-invariant edge partition.

We build, inductively on n, a family $\mathcal{F}_n$ of n-element sequences, starting for the case $n = 1$ with the family $\mathcal{F}_1 = \{ \,(\, \}$. To build $\mathcal{F}_{n+1}$ apply Operations A and B to every sequence $F \in \mathcal{F}_n$ and let $\mathcal{F}_{n+1}$ comprise the resulting sequences. Operation A: add '(' to the right of F. Operation B: add ')' to the right of F and throw away the resulting sequence if it does not contain free elements (i.e., if all its parentheses can be paired).

Proceeding in this way we obtain, for example,

$$\begin{aligned} \mathcal{F}_2 &= \{ \,((\, \}, \\ \mathcal{F}_3 &= \{ \,(((,\ (()\, \} \\ \mathcal{F}_4 &= \{ \,((((,\ ((),\ ()(\, \}. \end{aligned}$$

It is easy to see that $\mathcal{F}_n$ is the set of all n-sequences beginning with '(' which is a free element. (In particular, all right parentheses are paired.)

For any sequence $F \in \mathcal{F}_n$ we build the corresponding chain C_F in $L(\mathcal{B}_n)$ which has length t, where t is the number of free members of F. To obtain the $(*)$-description of the ith element of C_F, $i \in [t]$, we reverse in F the last $i - 1$ free parentheses and replace the ith free element (when counted from the right) by the star $*$. Thus, for example, '(()(()' gives $(()*()$ and $*()))()$ which correspond to the following chain in $L(\mathcal{B}_6)$:

$$\big(\{3, 6\} \sqsubset \{3, 4, 6\}\big) \lessdot \big(\{3, 4, 6\} \sqsubset \{1, 3, 4, 6\}\big).$$

It is easy to see that every C_F is a symmetric chain. We claim that

$$\mathcal{D}_n = \{\sigma^{(j)}(C_F) : F \in \mathcal{F}_n,\ j = 0, \dots, n-1\}$$

is the required edge partition.

We have to prove that for every element $x = (A \sqsubset B)$ in $L(\mathcal{B}_n)$ there are unique $F \in \mathcal{F}_n$ and $j \in [0, n-1]$ such that $x \in \sigma^{(j)}(C_F)$. First we show how to find at least one such pair (F, j).

Step 1. Write x in the $(*)$-representation.

Step 2. Rotate the pattern to bring the star to position 1 and then identify all free parentheses. Clearly, if disregarding the paired elements, our sequence is '$*\,)\dots)\,(\dots($.'

Step 3. Rotate again so that the first free left parenthesis identified in Step 2 (or the star itself if no '(' is free) is moved to position 1. Let j be the number of positions that the star was moved anticlockwise by Steps 2 and 3 combined.

Step 4. Replace the star and all free right parentheses identified in Step 2 by left parentheses. Let F be the resulting sequence.

Obviously, when we pair brackets in F, we obtain the same pairings as in Step 2. This implies that $F \in \mathcal{F}_n$ as it starts with a free '(' and that $x \in \sigma^{(j)}(C_F)$ as required. Here is an illustration for $x = (\{1, 6, 7\} \sqsubset \{1, 4, 6, 7\}) \in L(\mathcal{B}_8)$:

Step 1:	$)\ (\ (\ *\ (\ )\ )\ ($	
Step 2:	$*\ \boxed{(\,)}\ )\ \boxed{(\,)}\ (\ ($	
Step 3:	$(\ (\ *\ \boxed{(\,)}\ )\ \boxed{(\,)}$	(and $j = 1$)
Step 4:	$(\ (\ (\ (\ )\ (\ (\ )$	(this is F)

The uniqueness of (F, j) may be established in different ways. One, which actually gives an alternative definition of $\mathcal{D}_n$, is the following. Given the $(*)$-representation of x, let $g(i) = l_i - r_i$ for $0 \le i \le n-1$, where l_i and $r_i = i - l_i$ are respectively the number of left and right parentheses occuring in the i positions preceding '$*$' clockwise. Now, C_F starts at the element on which g achieves its maximum. (If there are a few such elements, then it is the first one.) Why? Just pair the brackets in the $(*)$-representation of $\sigma^{(-j)}(x) \in C_F$, e.g.,

$$(\,\boxed{(\,(\,)\,(\,)\,)}\,\boxed{(\,)}\,(\,*\,)\,\boxed{(\,)}\,)\,\boxed{(\,)}, \tag{1}$$

and notice that any paired block (boxed regions) contributes 0 to g while any right-hand-sided part of it contributes a strictly negative value. Now, the maximum of g is the number of free left parentheses in (1) and this is achieved for the first time when considering the segment preceding the star, as required.

But now, once that j has been identified, there trivially could not be two suitable F's. $\square$

For the remainder of this work let $\mathcal{D}_n$ denote the edge decomposition of $\mathcal{B}_n$ constructed above. It has the following properties.

THEOREM 5. *Let $C = (A_1 \sqsubset \cdots \sqsubset A_k)$ be one of the chains in $\mathcal{D}_n$. If $A_{i+1} = A_i \cup \{a_i\}$, then the elements $a_{k-1}, \ldots, a_1$ are situated on the circle in this order (clockwise) and between a_{i+1} and a_i (clockwise) there is an even number of places. For each $i \in [k-3]$ there is an element $(B \sqsubset B')$ belonging to a chain of $\mathcal{D}_n$ shorter than C such that*

$$A_i \sqsubset B \sqsubset B' \sqsubset A_{i+3}. \tag{2}$$

Proof. Take the sequence $F \in \mathcal{F}_n$ giving rise to C. (We may assume $j = 0$.) The fact that in F every pair of consecutive free elements contains only paired brackets in between implies the first part of the theorem.

To show the second claim, let F' be the sequence F with the $(i+1)$st free left bracket (if counted from the right) replaced by ')' which is then paired with the $(i+2)$nd free element:

$$\begin{array}{llllllllllllll}
F: & (& \square & (& \square & (& \square & (& \square & (& \square & (& \square & \\
 & & & & & & & & & * & \multicolumn{3}{c}{\ldots\ldots\ldots} & A_i \\
 & & & * & \multicolumn{9}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} & A_{i+3} \\
F': & (& \square & (& \square & \multicolumn{2}{c}{\boxed{(\,\square\,)}} & & \square & (& \square & (& \square & \\
 & & & & & & & & & * & \multicolumn{3}{c}{\ldots\ldots\ldots} & B \\
 & & & * & \multicolumn{9}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} & B'
\end{array}$$

The new sequence corresponds to a chain of length $k-2$ and its ith and $i+1$st elements obviously satisfy the required property. $\square$

5. Applications of the Partition $\mathcal{D}_n$

We would like to include here some applications of the edge partition $\mathcal{D}_n$ built in Theorem 4. Basically, we are inspired by known results where a symmetric vertex decomposition of $\mathcal{B}_n$ is used. We refer the reader to Section 3.4 of Anderson's book [2] for an exposition of a few results of this type. I am grateful to Ian Anderson for drawing my attention to some other applications not surveyed in his book.

5.1. ON THE NUMBER OF ANTICHAINS IN $L(\mathcal{B}_n)$

Let us consider the following question: what is $\varphi(L(\mathcal{B}_n))$, the number of antichains in $L(\mathcal{B}_n)$? The computation of $\varphi(\mathcal{B}_n)$ was an old and difficult problem; a complicated asymptotic formula is established by Korshunov [14].

Here we provide some rough estimates of $\varphi(L(\mathcal{B}_n))$ by applying ideas of Hansel [12] who showed that $2^N \le \varphi(\mathcal{B}_n) \le 3^N$, where $N = w(\mathcal{B}_n) = \binom{n}{\lfloor n/2 \rfloor}$.

Considering all possible subsets of the largest antichain of $L(\mathcal{B}_n)$ we obtain trivially $\varphi(L(\mathcal{B}_n)) \geq 2^m$, where $m = w(L(\mathcal{B}_n)) = \lceil n/2 \rceil \binom{n}{\lfloor n/2 \rfloor}$.

On the other hand, observe that an antichain $A \subset L(\mathcal{B}_n)$ is uniquely determined by the ideal $\Delta(A) = \{x \in L(\mathcal{B}_n) : \exists\, a \in A,\ x \leq a\}$. Consider any $C = (x_1 \lessdot \cdots \lessdot x_l) \in \mathcal{D}_n$. By Theorem 5 we can find y_i in a shorter chain with $x_{i-2} < y_i < x_{i+2}$ for $3 \leq i \leq l-2$. Knowing $\Delta(A) \cap C'$ for every $C' \in \mathcal{D}_n$ shorter than C we know $\Delta(A) \cap \{y_3, \ldots, y_{l-2}\}$. But then it is easy to check that for at most 4 elements of C we cannot deduce whether it is in $\Delta(A)$, and therefore $\Delta(A) \cap C$ can assume at most 5 possible values. Considering consecutively the chains of $\mathcal{D}_n$ in some size-increasing order we conclude that $\varphi(L(\mathcal{B}_n)) \leq 5^m$.

5.2. ORTHOGONAL PARTITIONS OF $L(\mathcal{B}_n)$

Two chains in a poset $\mathcal{P}$ are called *orthogonal* if they have at most one common element. Two vertex partitions $\mathcal{D}$ and $\mathcal{D}'$ are *orthogonal* if any $C \in \mathcal{D}$ is orthogonal to any $C' \in \mathcal{D}'$. This notion is of interest because, as proved by Baumert *et al*. [3], the existence of two orthogonal minimum decompositions implies one, rather strong, Sperner-type property (see [2, Section 3.4.3] for details).

A result of Shearer and Kleitman [13] (see Section 3.4 in [2]) asserts that there exist two orthogonal chain decompositions of $\mathcal{B}_n$ into $\binom{n}{\lfloor n/2 \rfloor}$ chains each. What can be said about $L(\mathcal{B}_n)$?

We define the *complementary chain* $\overline{C}$ of a chain C by replacing every element by its complement, i.e., if $C = (A_1 \subset \cdots \subset A_k)$ then $\overline{C} = (\overline{A_k} \subset \cdots \subset \overline{A_1})$. Let $\overline{\mathcal{D}_n} = \{\overline{C} : C \in \mathcal{D}_n\}$, where $\mathcal{D}_n$ is the decomposition built in Theorem 4.

LEMMA 6. *Two elements $x_1 = (A_1 \sqsubset B_1)$ and $x_2 = (A_2 \sqsubset B_2)$ of $L(\mathcal{B}_n)$ can belong simultaneously to $\mathcal{D}_n$ and $\overline{\mathcal{D}_n}$ only if $n = 2k$ is even and $\{|B_1|, |B_2|\} = \{k, k+1\}$.*

Proof. Let $i_h \in [n]$ be the element of B_h not in A_h, $h = 1, 2$, and let pairs (F, j) and (F', j') give rise to chains $C, C' \in \mathcal{D}_n$ such that C' contains x_1 and x_2 while C contains $\overline{x}_1$ and $\overline{x}_2$ respectively. Assume that $j' = 0$ and $x_1 < x_2$, i.e., $B_1 \subset A_2$.

In F' i_2 precedes i_1 and we claim that F' does not contain a free element between them. Indeed, if it be in the position $y \in [n]$, then $y \in \overline{B_1}$ and $y \notin \overline{A_2}$, that is, $\sigma^{(-j)}(y)$ must be a free element in F. But then $\sigma^{(-j)}(y)$ must lie between $\sigma^{(-j)}(i_1) < \sigma^{(-j)}(i_2)$. (In C the element $\overline{x_2}$ comes before $\overline{x_1}$.) This contradiction (on one hand the elements i_2, y, i_1 go clockwise, on the other – anticlockwise) proves the claim.

Thus all the elements between i_2 and i_1 are paired in F'; therefore $B_1 = A_2$ and there must be the same number of left and right parentheses in this interval. Considering $\overline{x_2}, \overline{x_1} \in C$ we show the analogeous statement about the elements between i_1 and i_2 (if going clockwise), which clearly implies the claim. □

COROLLARY 7. *For odd n there is a pair of orthogonal symmetric chain decompositions of $L(\mathcal{B}_n)$.*

Remark. Unfortunately, we do not know if the corresponding claim is true for even n.

5.3. A STORAGE AND RETRIEVAL PROBLEM

Suppose we maintain a database with n records which we number from 1 to n and we wish to organize an efficient searching. We assume that we have queries $Q_1, \dots, Q_M$ each of which we identify with the set of records satisfying it, that is, $Q_i \subset [n]$ and these subsets are not necessarily distinct. One idea, see Ghosh [8], is to find a sequence X of elements of $[n]$ such that every Q_i occurs in X as a subsequence of consecutive terms so that every Q_i can be defined by a starting position in X and the size of Q_i.

In connection with this Lipski [15] considered the following problem. Find the shortest sequence of elements of $X = [n]$ such that X contains every $A \subset [n]$ as a subsequence of $|A|$ consecutive terms. He showed that s_n, the length of an optimal sequence, satisfies

$$\left(\frac{2}{\pi n}\right)^{1/2} 2^n \le (1 + \mathrm{o}(1)) s_n \le \left(\frac{2}{\pi}\right) 2^n. \tag{3}$$

As far as I know, these seems to be the best known bounds to date.

Here we ask what is the value of t_n, the shortest length of a sequence X such that for every $A \sqsubset B \subset [n]$ the sequence X contains A as a subsequence of $|A|$ consecutive terms preceded by x, where $\{x\} = B \setminus A$. Such a situation can happen if every query is a set with a selected point. For example, we search in a dictionary, the allowed queries are of the form "Find *word*" and the answer should give the entry where *word* is defined plus all relevant entries. Applying the ideas of [15] we prove the following upper and lower bounds.

THEOREM 8.

$$\left(\frac{n}{2\pi}\right)^{1/2} 2^n \le (1 + \mathrm{o}(1)) t_n \le \left(\frac{n}{\pi}\right) 2^n. \tag{4}$$

Proof. Clearly, t_n exceeds the number of different pairs $A \sqsubset A \cup \{x\}$ with $|A| = \lfloor n/2 \rfloor$, which implies the lower bound in (4) by Stirling's formula.

On the other hand, associate with every chain $C = (A_1 \sqsubset \cdots \sqsubset A_q)$ in $\mathcal{D}_n$ the sequence of elements of $[n]$ which contains first the elements of A_1 in any order which then are followed by $a_2, \dots, a_q$, where $\{a_i\} = A_i \setminus A_{i-1}$, $i = 2, \dots, q$.

Let $[n] = S \cup T$ be a partition of $[n]$ into 2 parts of (nearly) equal sizes. Let $\phi_1, \dots, \phi_k$ be the sequences corresponding to a symmetric vertex decomposition of 2^S. Also, let $\psi_1, \dots, \psi_l$ be the sequences corresponding to a symmetric edge decomposition of 2^T, each sequence being reversed.

Clearly, for every $A \subset S$ there exists ϕ_i containing A as the first consecutive $|A|$ terms and for every $A \sqsubset A \cup \{x\} \subset T$ there exists ψ_j containing, at the end, A preceded by x.

Now consider the sequence

$$X_1 = \psi_1\phi_1\psi_1\phi_2 \ldots \psi_1\phi_k\psi_2\phi_1\psi_2\phi_2 \ldots \psi_l\phi_k.$$

Take any $A \subset [n]$ and $x \in T \setminus A$. There is ψ_i containing, at the end, x followed by $A \cap T$ and ϕ_j containing $A \cap S$ as an initial subsequence. Therefore, X_1 contains x followed by A. Interchanging S and T, we write a sequence X_2 containing every pair $A \sqsubset A \cup \{x\}$ with $x \in S$. The sequence X_1X_2 is the required (and explicitly constructed) sequence. It is easy to see that the average size of a sequence corresponding to a chain of a symmetric vertex or edge decomposition of $\mathcal{B}_n$ is $(\frac{1}{2} + \mathrm{o}(1))n$. Therefore, $t_n \leq 2\left(\frac{1}{2} + \mathrm{o}(1)\right) nkl$ which gives the desired upper bound by Stirling's formula. □

5.4. ONE NUMERICAL PROBLEM

There exists a so called Audit Expert Mechanism which can be used to protect small statistical databases, see Chin and Ozsoyogly [5]. To find an optimal mechanism the following problem has to be solved. Suppose we operate with n-tuples of nonzero reals $a_1, \ldots, a_n$ and we want to find what is the maximum possible number of subsets $I \subset [n]$ such that $S(I)$ is equal to either 0 or 1. (Here and later we denote $S(I) = \sum_{i\in I} a_i$.) The best possible bound of $\binom{n+1}{\lfloor (n+1)/2 \rfloor}$ was found by M. Miller, Roberts and Simpson [17] and all extremal sequences were characterised by K. Miller and Sarvate [16] (for integers) and by Griggs [11] (for reals). The papers [17, 16] make use of the existence of a symmetric chain decomposition of $\mathcal{B}_n$.

Here, applying a symmetric chain decomposition of $L(\mathcal{B}_n)$, we can find K, the maximum possible number of elements $(I \sqsubset J) \in L(\mathcal{B}_n)$ such that $\{S(I), S(J)\} = \{0, 1\}$ over all real sequences $a_1, \ldots, a_n$. Actually, we can allow zero entries for, as we will see later, this does not affect K. Apparently, this problem does not have such an application like that of the original problem, but it might be of some interest especially as an unexpected application of a symmetric chain decomposition of $L(\mathcal{B}_n)$.

The expression $(a)^i$ is a shorthand for a repeated i times. Also we assume that all n-tuples have their entries ordered nondecreasingly.

THEOREM 9. *For $n \geq 2$ we have*

$$K = \lceil n/2 \rceil \binom{n}{\lfloor n/2 \rfloor}, \tag{5}$$

and this value is achieved for and only for the following sequences: $((-1)^k, (+1)^k)$, $((-1)^{k-1}, (+1)^{k+1})$ *and* $((-1)^{k-1}, 0, (+1)^k)$ – *for* $n = 2k$ *and* $((-1)^k, (+1)^{k+1})$ – *for* $n = 2k + 1$.

Proof. Let m be the largest index for which $a_m < 0$. (If $a_1 \geq 0$, let $m = 0$.) Define $f : 2^{[n]} \to 2^{[n]}$ by the formula

$$f(I) = I \triangle [m] = (I \setminus [m]) \cup ([m] \setminus I), \quad I \subset [n].$$

One can easily check that $I \subset J \subset [n]$ implies $S(f(I)) \leq S(f(J))$.

$\mathcal{D}_n$ can be viewed as a collection of symmetric chains in $2^{[n]}$. Let $X_r \sqsubset \cdots \sqsubset X_{n-r}$ be one such chain. The sequence

$$S(f(X_r)), \ldots, S(f(X_{n-r}))$$

is nondecreasing and therefore 0 and 1 can occur side by side there at most once. As every $A \sqsubset B$ is present in exactly one chain and f is a bijection preserving or reversing the $\sqsubset$-relation, K does not exceed the total number of chains, which gives the required upper bound.

A moment's thought reveals that a necessary and sufficient condition for an n-tuple to be optimal is the following. If $n = 2k + 1$ then for every $A \sqsubset B \subset X$, $|A| = k$, we have $S(f(A)) = 0$ and $S(f(B)) = 1$. If $n = 2k$ then for every $A \sqsubset B \sqsubset C \subset X$, $|A| = k - 1$, among the numbers

$$S(f(A)) \leq S(f(B)) \leq S(f(C)) \tag{6}$$

there is a 0 adjacent a 1.

This condition is fulfilled for the sequences mentioned in the statement. Indeed, let us consider $((-1)^k, (+1)^{k+1})$, for example. Here $m = k$ and for any $A \sqsubset B$ with $|A| = k$ we have

$$S(f(A)) = S(A \triangle [k]) = (-1)(k - s) + (k - s) = 0, \tag{7}$$

where $s = |A \cap [k]|$. Similarly, $S(f(B)) = 1$ so the sequence is optimal.

We claim that these are essentially the cases of the equality. Let us do the harder case $n = 2k$. If for some $i \neq j$ we have $a_i \neq \pm 1$ and $a_j \neq \pm 1$, then $A \sqsubset A \cup \{i\} \sqsubset A \cup \{i, j\}$ with any $A \in X^{(k-1)}$, $A \not\ni i, j$, obviously violates the condition. If for exactly one i we have $a_i \neq \pm 1$, then considering $A \sqsubset A \cup \{i\} \sqsubset C$ we conclude that $S(f(A \cup \{i\})) = 0$ for any $A \in (X \setminus \{i\})^{(k-1)}$. Suppose $a_i \geq 0$, for example. Then $S(f(A \cup \{i\})) = k - j - 1 + a_i = 0$, where j is the total number of elements equal to -1 (so $2k - 1 - j$ elements equal $+1$). If $a_i = 0$, then we have the third example mentioned in the theorem. If $a_i \geq 2$, then $j \geq k + 1$ and any sequence (6) with $C \not\ni i$ violates the condition. Finally, if $|a_i| = 1$ for every i, then arguing as in (7) we deduce that we can have either k or $k + 1$ positive entries. □

6. Characterisation of Line Posets

Here we ask ourselves when a given poset $\mathcal{L}$ is the line poset of some $\mathcal{P}$ and what information about $\mathcal{P}$ can be reconstructed from $L(\mathcal{P})$. (Of course, it is implicitly understood that we operate with isomorphism classes of posets.)

Now, $L(\mathcal{P})$ cannot contain elements w, x, y, z such that $w \lessdot y$, $x \lessdot y$, $w \lessdot z$ but $x \not\lessdot z$; call this configuration N. Indeed, if y and z cover w they must be of the form $(a \lessdot b)$, $(a \lessdot c)$, where $w = (d \lessdot a)$, some $a, b, c, d \in \mathcal{P}$. Then the relation $x \lessdot y$ implies that $x = (e \lessdot a)$ which implies that $x \lessdot z$.

Also, $L(\mathcal{P})$ cannot contain the configuration C_n, $n \geq 3$, made of elements y and $x_1, \ldots, x_n$ such that $x_1 \lessdot y \lessdot x_n$ and $x_i \lessdot x_{i+1}$, for $i \in [n-1]$. Indeed, suppose the contrary. Clearly, $\mathcal{P}$ contains elements $z_0 \lessdot z_1 \lessdot \cdots \lessdot z_n$ such that $x_i = (z_{i-1} \lessdot z_i)$. But y covers the same element as x_2 and is covered by the same element as x_{n-1}, so $y = (z_1 \lessdot z_{n-1})$ and $n = 3$; but then $y = x_2$, which is a contradiction.

For a poset $\mathcal{P}$ let $T(\mathcal{P}) = (\mathcal{C}, k, l, u)$ be the quadruple with $\mathcal{C}$ being a subposet of $\mathcal{P}$ spanned by the nonextremal elements, that is by $\{a \in \mathcal{P} : \exists\, b, c \in \mathcal{P},\ b < a < c\}$ and k is the number of pairs $(a \lessdot b)$ with $a, b \in \mathcal{P} \setminus \mathcal{C}$ while the functions $l, u : \mathcal{C} \to \mathbf{N_0}$ are given by

$$\begin{aligned} l(a) &= |\{x \in \mathcal{P} \setminus \mathcal{C} : x \lessdot a\}|, \\ u(a) &= |\{x \in \mathcal{P} \setminus \mathcal{C} : x \gtrdot a\}|, \quad a \in \mathcal{C}. \end{aligned}$$

It is easy to see that $T(\mathcal{P})$ determines $L(\mathcal{P})$.

The following theorem states that the above examples provide a complete answer to our two questions.

THEOREM 10. *A poset $\mathcal{L}$ is isomorphic to $L(\mathcal{P})$ for some $\mathcal{P}$ if and only if $\mathcal{L}$ contains neither configuration N nor any of C_n, $n \geq 3$. Furthermore, $T(\mathcal{P})$ determines $L(\mathcal{P})$ and can be reconstructed from it.*

Proof. Given a poset $\mathcal{L}$ without N or C_n let X be two disjoint copies of its vertex set, namely $X = \{x^\wedge, x^\vee : x \in \mathcal{L}\}$. Let $x^\wedge \sim y^\vee$ if $x \lessdot y$; let $x^\wedge \sim y^\wedge$ if for some $s \in \mathcal{L}$ we have $s \gtrdot x$ and $s \gtrdot y$; let $x^\vee \sim y^\vee$ if for some $s \in \mathcal{L}$, $s \lessdot x$ and $s \lessdot y$.

We claim that $\sim$ is an equivalence relation. Indeed, if $x^\wedge \sim y^\wedge$ and $y^\wedge \sim z^\wedge$ then there are $s, t \in \mathcal{L}$ such that $x, y \lessdot s$ and $y, z \lessdot t$. But then t must cover x for otherwise x, y, s, t would span a forbidden configuration. So $x, z \lessdot t$ and $x^\wedge \sim z^\wedge$. The remaining cases are equally easy.

Let $\overline{x}$ denote the equivalence class of $x \in X$. Define the poset $\mathcal{P}$ (also denoted by $L^{-1}(\mathcal{L})$) on $V = X/{\sim} = \{\overline{x} : x \in X\}$ by $A < B$, $A, B \in V$, iff in $\mathcal{L}$ there exist $y \leq z$ with $y^\vee \in A$ and $z^\wedge \in B$. One can check that this is indeed an ordering. For example, to check its transitivity, let $A < B$ and $B < C$, choose $w \leq x$ and $y \leq z$ in $\mathcal{L}$ with $w^\vee \in A$, $x^\wedge, y^\vee \in B$ and $z^\wedge \in C$; then $x^\wedge \sim y^\vee$ implies that $w \leq x \lessdot y \leq z$ and $A < C$.

Let us show that $\overline{x^\wedge}$ covers $\overline{x^\vee}$. Assuming the contrary we find $z \geq y$ and $w \geq v$ in $\mathcal{L}$ with $z^\wedge \sim x^\wedge$, $y^\vee \sim w^\wedge$ and $v^\vee \sim x^\vee$. By the definition of $\sim$, some $t \in \mathcal{L}$ covers both x and z, some $s \in \mathcal{L}$ is covered by both x and v and $v \leq w \lessdot y \leq z$ – which implies that $\mathcal{L}$ contains some C_n, which is forbidden.

We claim that $\mathcal{L} \cong L(\mathcal{P})$ via the map F which sends $x \in \mathcal{L}$ to $(\overline{x^\vee} \lessdot \overline{x^\wedge})$. First note that F is an order preserving map: if $x \gtrdot y$ in $\mathcal{L}$, then $x^\vee \sim y^\wedge$, which implies $F(x) \gtrdot F(y)$ as desired. Next, F is injective for if $F(x) = F(y)$ then $x^\wedge \sim y^\wedge$ and

$x^\vee \sim y^\vee$, which implies that for some w and z we have $w \lessdot x \lessdot z$ and $w \lessdot y \lessdot z$; but as $\mathcal{L}$ does not contain configuration C_3 we conclude that $x = y$. To show that F is surjective take any $(A \lessdot B) \in L(\mathcal{P})$. As $A < B$, for some $\mathcal{L}$-elements $x \leq y$ we have $A = \overline{x^\vee}$, $B = \overline{y^\wedge}$. But it is easy to see that $\overline{x^\wedge} \leq \overline{y^\wedge}$, which implies that $(A \lessdot B) = (\overline{x^\vee} \lessdot \overline{x^\wedge}) = F(x)$. Finally, if $F(x) \lessdot F(y)$ then $x^\wedge \sim y^\vee$ and $x \lessdot y$. This proves completely that $\mathcal{L} \cong L(\mathcal{P})$.

In the second part it is enough to show that for any poset $\mathcal{R}$ we have $T(\mathcal{R}) \cong T(\mathcal{P})$, where $\mathcal{P} = L^{-1}(\mathcal{L})$, $\mathcal{L} = L(\mathcal{R})$. To build a natural isomorphism $H : \mathcal{C}(\mathcal{R}) \to \mathcal{C}(\mathcal{P})$ take, for any element $a \in \mathcal{C}(\mathcal{R})$, some $b \lessdot a$ which exists as a is a non-extremal element of $\mathcal{R}$. Now let $H(a) = \overline{x^\wedge}$, where $x = (b \lessdot a) \in \mathcal{L}$ and $\sim$ is as above. To show that H is well defined let b' be another choice of b and denote $y = (b' \lessdot a)$. Let c be an element covering a. Then $(a \lessdot c)$ covers in $\mathcal{L}$ both x and y, so by the definition of $\mathcal{P}$ we have $x^\wedge \sim y^\wedge$. Also, $H(a) \in \mathcal{P}$ is not extremal as

$$\overline{(b \lessdot a)^\vee} < H(a) < \overline{(a \lessdot c)^\wedge}.$$

Next, H is an order-preserving bijection. Indeed, let $a \gtrdot b$ in $\mathcal{C}(\mathcal{R})$. Choose $c \lessdot b$. Then $H(a) = \overline{(b \lessdot a)^\wedge}$ and $H(b) = \overline{(c \lessdot b)^\wedge}$. But $(c \lessdot b)^\wedge \sim (b \lessdot a)^\vee$ and we have $H(a) > H(b)$ by the definition of the order on $\mathcal{P}$. To show that H is injective choose any $a, a' \in \mathcal{C}(\mathcal{R})$. Then $H(a) = H(a')$ implies that $y = (c \lessdot a)^\wedge \sim y' = (c' \lessdot a')^\wedge$, some $c, c' \in \mathcal{R}$. Therefore there is $x \in \mathcal{L}$ covering both y and y' which implies $a = a'$ in $\mathcal{R}$ as required. To establish the surjectivity of H consider $x = \overline{(a \lessdot b)^\vee} \in \mathcal{C}(\mathcal{P})$, for example. Observe first that $a \in \mathcal{R}$ is not extremal. Indeed, take any $y \in \mathcal{P}$ covered by x; as we have already shown any pair $y \lessdot x$ is of the form $\overline{(c \lessdot d)^\vee} \lessdot \overline{(c \lessdot d)^\wedge}$ which implies $d = a$ and $c \lessdot a$. Now $H(a) = \overline{(c \lessdot a)^\wedge} = x$ as required. Again, any two adjacent elements of $\mathcal{C}(\mathcal{P})$ can be represented as $\overline{(a \lessdot b)^\vee} \lessdot \overline{(a \lessdot b)^\wedge}$ and then they are the images of two adjacent elements, $a \lessdot b$ of $\mathcal{C}(\mathcal{R})$, which implies that $\mathcal{C}(\mathcal{P}) \cong \mathcal{C}(\mathcal{R})$.

Clearly, H preserves k, l and u, which completes the proof. □

Acknowledgements

The author is extremely grateful to Andrew Thomason for helpful discussions and to the anonymous referees for pointing out many mistakes in the first version of the paper.

References

1. Anderson, I. (1967) Some problems in combinatorial number theory, Ph.D. Thesis, University of Nottingham.
2. Anderson, I. (1987) *Combinatorics of Finite Sets*, Oxford University Press.
3. Baumert, L. D., McEliece, R. J., Rodermich, R. J. and Rumsey, H. (1980) A probabilistic version of Sperner's theorem, *Ars Combin.* **9**, 91–100.

4. Bruijn, N., Tengbergen, C. and Kruyswijk, D. (1951) On the set of divisors of a number, *Nieuw Arch. Wisk.* **23**, 191–193.
5. Chin, F. and Ozsoyoglu, G. (1982) Auditing and inference control in statistical databases, *IEEE Trans. Software Eng.* **SE-8**, 574–582.
6. Dantzig, G. B. and Hoffman, A. J. (1956) Dilworth's theorem on partially ordered sets, in H. W. Kuhn and A. W. Tucker (eds.), *Linear Inequalities and Related Systems*, Ann. Math. Stud. 38, pp. 207–214.
7. Dilworth, R. P. (1950) A decomposition theorem for partially ordered sets, *Ann. of Math.* **51**(2), 161–166.
8. Ghosh, S. P. (1975) Consecutive storage of relevant records with redundancy, *Comm. ACM* **18**, 464–471.
9. Greene, C. and Kleitman, D. J. (1976) Strong versions of Sperner's theorem, *J. Combin. Theory Ser. A* **20**, 80–88.
10. Griggs, J. R. (1977) Sufficient conditions for a symmetirc chain order, *SIAM J. Appl. Math.* **32**, 807–809.
11. Griggs, J. R. (1999) Database security and the distribution of subset sums in $\mathbf{R}^m$, in *Graph Theory and Combinatorial Biology (Balatonlelle, 1996)*, János Bolyai Math. Soc. 7, pp. 223–252.
12. Hansel, G. (1966) Sur de nombre des fonctions booléennes monotones de n variables, *C.R. Acad. Sci. Paris Sér. A-B* **262**, 1088–1090.
13. Kleitman, D. J. and Shearer, J. B. (1979) Probabilities of independent choices being ordered, *Stud. Appl. Math.* **60**, 271–276.
14. Korshunov, A. D. (1981) The number of monotone Boolean functions, *Probl. Kibernet.* **38**, 5–108.
15. Lipski, W. (1978) On strings containing all subsets as substrings, *Discrete Math.* **21**, 253–259.
16. Miller, K. C. and Sarvate, D. G. (1995) An application of symmetric chains to a statistical database compromise prevention problem, *Bull. ICA* **13**, 57–64.
17. Miller, M., Roberts, I. and Simpson, J. (1991) Application of symmetric chains to an optimization problem in the security of statistical databases, *Bull. ICA* **2**, 47–58.
18. Saks, M. (1986) Some sequences associated with combinatorial structures, *Discrete Math.* **59**, 135–166.