International Journal of Foundations of Computer Science

SUCCESSFUL ELEMENTARY GENE ASSEMBLY STRATEGIES

Vladimir Rogojin
Computational Biomodelling Laboratory, Turku Centre for Computer Science,
Department of IT, Åbo Akademi University, Joukahaisenkatu 3-5 A,
Turku, 20520, Finland.
vrogojin@abo.fi
http://combio.abo.fi/∼vrogojin

Received ()
Revised ()
Accepted ()
Communicated by ()

We study elementary gene assembly in ciliates. During sexual reproduction, broken and shuffled gene segments in micronuclei get assembled into contiguous macronuclear genes. We consider here a restricted version of the intramolecular model (called elementary), where at most one gene segment is involved at a time (either inverted, or translocated). Not all gene patterns may be assembled by elementary operations, and not all assembly strategies are successful. For a given gene pattern, we characterize in this paper all successful translocation-only elementary assemblies. We also estimate the number of such assemblies. We solve the problem in terms of graphs and permutations.

Keywords: ciliates; gene assembly; intramolecular model; elementary operations; successful elementary assemblies; sorting permutations

1. Introduction

Ciliates are eukaryotes which exist for over two billion of years. They count thousands of species. One of the unique features for all ciliate organisms is the possession of complexes of 'cilia'. These are hair-like organelles on the surface of a cell used to move the organism in the aqueous space and to capture food from the environment.

The nucleic dualism is another unique feature which unites all ciliate species. Each ciliate organism possess nuclei of two different types called *micronuclei* and *macronuclei*. Almost all RNA-transcriptions happen in macronuclei while micronuclei are transcriptionally silent. However, during matting micronuclei become active and exchange the genetical information in between two matting organisms. Macronuclei from old organisms are destroyed during sexual reproduction, some copies of new micronuclei get transformed into new macronuclei. During this process DNA molecules from the transforming micronuclei are being heavily edited so that macronuclear molecules are formed. This process of DNA transformation (called *gene assembly*) is particularly intricate due to the immense difference in molecular structures of micronuclear and macronuclear versions of genes. A gene is stored in a contiguous form on a macronuclear molecule. On the other hand, the same gene is fragmented when stored on a micronuclear molecule. The gene fragments (called MDSs) are separated by non-codding nucleotide sequences (called IESs). MDSs may be shuffled and some of them may be inverted on the micronuclear molecule. During gene assembly a ciliate has to detect and unscramble all MDSs and to excise all IESs from molecules.

The process of gene assembly is driven by splicing on short nucleotide sequences (called *pointers*). Each pointer occurs on the edges of exactly two different MDSs, where one follows immediately after another one in an assembled gene. One occurrence of the pointer is placed at the end of the preceding MDS and another occurrence is placed at the beginning of the succeeding MDS. Detailed description of the biology of ciliates and of gene assembly process can be found in[7, 8, 15, 24, 25, 28, 29].

There are two molecular models which represent gene assembly in ciliates. Both the *intermolecular model* [17, 18] and the *intramolecular model* [11, 27] define splicing of gene fragments via pointers. In the intermolecular model several molecules participate in the recombination, while the intramolecular model considers folding and recombination within a single molecule. Recent results [3, 23] suggest that there are template molecules aiding the correct alignment of the recombining molecules.

The intramolecular model of [11, 27] has three operations (ld, hi, and dlad) that always apply on a single molecule. In their general form, these operations may manipulate arbitrarily long gene fragments. A restricted version of the intramolecular operations [9, 10, 14] rearranges (translocates or inverts) at most one gene fragment at a time. There are two subclasses of this restricted version of the intramolecular operations called *simple* and *elementary*. We refer to [5, 10, 14, 19, 20] for the definitions and results obtained for simple intramolecular operations. We are concentrating on the elementary operations in this paper.

In contrast to the general type of the intramolecular operations, not all gene patterns may be assembled by simple and elementary intramolecular operations. This fact eventually raises a question: what gene patterns can simple and elementary operations assemble? Since simple gene assembly is deterministic [19] (i.e., either all assembly strategies for a gene pattern are successful, or all of the strategies are unsuccessful) the answer to this question can be obtained when applying a simple assembly strategy for a gene pattern. In other words, a gene pattern can be assembled by simple operations if and only if an arbitrarily chosen simple assembly strategy for the gene pattern is successful. We refer to [21, 22] for the survey on simple and elementary operations.

Unlike simple operations, there may be successful and unsuccessful assembly strategies of elementary operations for a gene pattern [13]. We characterize in the paper all successful elementary assembly strategies for any gene pattern. We estimate the upper bound for the total number of all successful elementary assemblies for any gene pattern. In [13] and in [26] there were characterized those gene patterns which may be assembled by elementary operations. In [26] it was presented an algorithm which decides effectively in less than cubic time whether a given gene pattern may be assembled or not by elementary translocation-only operations. That algorithm was based on the notion of *forbidden* MDSs which are never translocated by elementary operations. It was proved that a gene pattern has an elementary assembly strategy if and only if the forbidden MDSs are sorted in the gene pattern. The main result of this paper gives another way to decide effectively whether a gene pattern may be assembled by elementary operations. One needs to apply an elementary assembly strategy of a special form which is described in this paper. We prove that strategies of this form are applicable if and only if the gene pattern can be assembled by elementary operations.

We study the gene assembly process in terms of graphs and permutations. We represent micronuclear molecules as permutations representing the order of gene fragments. Elementary intramolecular operations are represented as transformations over permutations, assembly strategies as compositions of operations over permutations, and gene assembly process as sorting of permutations. In our solutions we use the notion of graphs to determine the order of intramolecular elementary operations in assembly strategies, and to detect those operations which never may be used in any of assembly strategies applicable to a given gene pattern. To characterize successful elementary assembly strategies we introduced notions of so-called *fixed integers* and of *blocks* for permutations.

In [4], [12], [6], [16] it was presented a different topic on sorting of permutations related to gene transformations. It was studied the problem of sorting signed and unsigned permutations by reversals.

2. Preliminaries

By A^* we denote the set of all finite strings over finite alphabet A. By $\mathsf{dom(u)}$ we define the set of letters occurring in string u. By λ we denote the empty string. We say that string v is a substring of u and denote this by $v \leq u$, if $u = xvy$, where x, y are some strings. We say that string $v = \alpha_1\alpha_2 \dots \alpha_k$, where $\alpha_1, \alpha_2, \dots, \alpha_k \in A$ is a subsequence of string u and denote this fact by $v \leq_s u$, if $u = u_0\alpha_1 u_1\alpha_2 u_2 \dots u_{k-1}\alpha_k u_k$, where $u_0, u_1, u_2, \dots, u_{k-1}, u_k \in A^*$. For a subset $B \subseteq A$ we define the morphism $\Phi_B : A^* \to B^*$ such as $\Phi_B(\alpha) = \alpha$, if $\alpha \in B$, and $\Phi_B(\alpha) = \lambda$, otherwise. For a string u we denote $u|_B = \Phi_B(u)$.

Consider a finite alphabet $A = \{a_1, a_2, \dots, a_n\}$ with order relation $a_1 < a_2 < \dots < a_n$. We define an automorphism $\pi : A \to A$ and say that it is a permutation over A. We identify permutations as strings: $\pi = \pi(a_1)\pi(a_2)\dots\pi(a_n)$. The domain of permutation π is $\mathsf{dom}(\pi) = A$. We say that π is cyclically sorted, if $\pi = a_{k+1}a_{k+2}\dots a_n a_1 a_2 \dots a_k$, for some $1 \leq k \leq n$.

For $B \subseteq A$, we denote by $\pi|_B$ the subsequence of π consisting only of letters from B. Note that $\pi|_B$ is a permutation over B.

We agree about the following notations concerning sets of integers:

- $\Sigma_n = \{1, 2, \ldots, n\}$;
- $N_{p,q} = \{p, p+1, \ldots, q-1, q\}$;
- $N^e_{p,q} = \{i \in N_{p+1,q-1} | i - p \vdots 2\}$;
- $N^o_{p,q} = \{i \in N_{p+1,q-1} | i - p \not\vdots 2\}$;

Note that $N_{p,q} = (N^e_{p,q} \oplus N^o_{p,q}) \cup \{p, q\}$.

By $\lfloor f \rfloor$ for some rational f we denote the biggest integer lower than f. By $\lceil f \rceil$ we denote the smallest integer greater than f.

By $\Gamma = (V, E)$ we define directed graph with the set of vertices V and set of edges $E \in V \times V$.

We make the following conventions about graph $\Gamma = (V, E)$:

- For $p, q \in V$, we write $q \to^+_\Gamma p$, if there is a non-empty path from q to p in Γ. We say that q is a *predecessor* of p, and p is a *successor* of q. If $(q, p) \in E$, then we denote $q \to_\Gamma p$;
- For $p \in V$ we let $\Gamma_p = (T_p, E_p)$, where $T_p = \{r \in V | r \to^+_\Gamma p\} \cup \{p\}$ and $E_p = \{(i, j) \in E | i, j \in T_p\}$. For $p, q \in V$, where $p < q$ we let $\Gamma_{p,q} = (N_{p,q}, E_{p,q})$, where $E_{p,q} = \{(i, j) \in E | i, j \in N_{p,q}\}$. For $p \in V$ we let also $\Gamma^S_p = (S_p, E^S_p)$, where $S_p = \{r \in V | p \to^+_\Gamma r\} \cup \{p\}$ and $E^S_p = \{(i, j) \in E | i, j \in S_p\}$;

For basic notions and results on graph theory we refer to [30].

3. Elementary Gene Assembly

Three intramolecular operations ld, hi and dlad explain the gene assembly process by means of pointers. After application of one of ld, hi and dlad several MDSs get spliced together into one *composite MDS*. For more details see [11, 27, 9]. In this paper we consider only the ld operations and a restricted version of dlad operations called *elementary*, see [10]. We do not consider here the hi operation.

The ld *operation* excises a part of the molecule flanked by a repeating pointer either containing just one IES (*simple case*, Figure 1(i)), or containing all MDSs of the gene (*boundary case*, Figure 1(iv)). As the result, we obtain one circular and one linear molecule. In simple case the circular molecule contains just the IES excised and the linear molecule contains the rest of the nucleotide sequence of the initial molecule (Figure 1(iii)). In boundary case the linear molecule contains the first and the last IESs spliced together and no MDSs. The circular molecule contains all codding sequences from the initial molecule (Figure 1(vi)). For more explanations on simple and boundary ld we refer to Figures 1(i)-(vi).

The dlad *operation* is applicable on a pair of pointers (let us denote them as p, q) in an alternating repeat. I.e., p and q are placed on the molecule as $\ldots p \ldots q \ldots p \ldots q \ldots$. The *elementary* dlad *operation* is applicable to p, q, if one occurrence of p and q flanks an IES, and another occurrence flanks a micronuclear (non-composite) MDS. As the result, the micronuclear MDS flanked by occurrences of p, q and the IES flanked by another occurrences of p and q switch each others

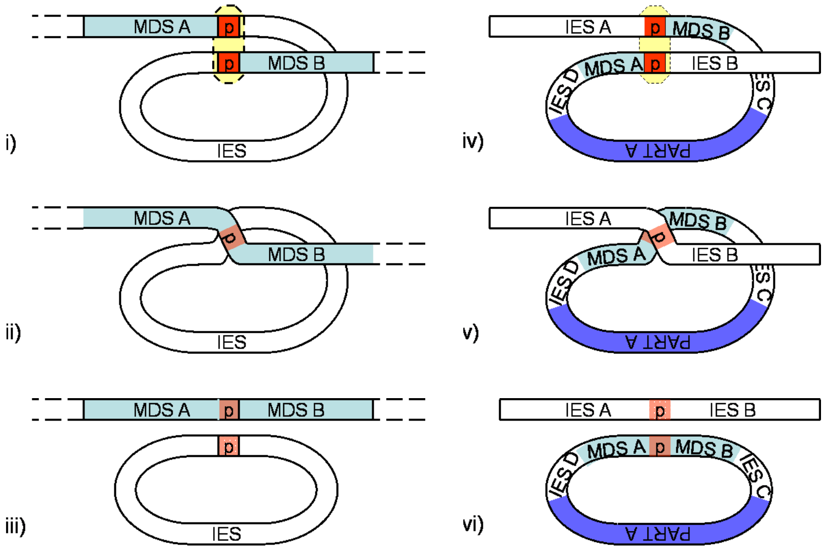


Fig. 1. *Simple* ld, an IES is flanked by p: (i) loop-folding, alignment of occurrences of pointer p; (ii) recombination by pointer p; (iii) result: the IES is excised in the form of circular molecule, *MDS A* and *MDS B* are spliced on common pointer p in the linear molecule. *Boundary* ld, *PART A* contains all MDSs of the gene pattern except of *MDS A* and *MDS B*: (iv) loop-folding, alignment of occurrences of pointer p; (v) recombination by pointer p; (vi) result: *MDS A* and *MDS B* are spliced on common pointer p and are excised together with *PART A* in the form of circular molecule. *IES A* and *IES B* are spliced together and remain in the linear molecule. In both linear and boundary cases one occurrence of p is present in the linear resulting molecule and one occurrence of p is present in the circular resulting molecule, but none of them acts as a pointer.

places. Then a composite MDS is obtained by splicing the MDS ending by p pointer, the MDS starting by p and ending by q pointer, and the MDS starting by q pointer. See Figure 2(i)-(iv) for more clarifications about folding and recombination of the molecule caused by dlad.

As in [13, 26] we formalize gene patterns as permutations by denoting each micronuclear MDS as a letter (an integer number establishing the place for an MDS in the assembled gene). We ignore IESs and place the numbers in the order corresponded MDSs occur in the gene pattern. An assembled gene corresponds then to a sorted permutation.

We formalize the elementary dlad as an operation on permutations. We denote it as ed. For each p, $2 \leq p \leq n-1$, we define ed_p as follows:

$$\mathsf{ed}_p(x\,p\,y\,(p-1)\,(p+1)\,z) = x\,y\,(p-1)\,p\,(p+1)\,z,$$
$$\mathsf{ed}_p(x\,(p-1)\,(p+1)\,y\,p\,z) = x\,(p-1)\,p\,(p+1)\,y\,z,$$

where x, y, z are strings over Σ_n. We denote $\mathsf{Ed} = \{\mathsf{ed}_i \mid 1 \leq i \leq n\}$.

We skip formalization of the ld and IESs. As soon as an ld operation is applicable to the molecule, it can be applied at any latter step during the assembly process.

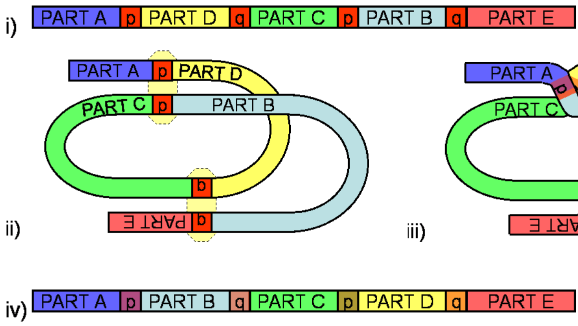


Fig. 2. dlad *operation:* (i) Initial molecule, poinetrs p and q overlap; (ii) Double-loop folding, alignment of occurrences of p and q; (iii) Recombination by pointers p and q; (iv) Resulting molecule, *PART B* and *PART D* changed places of each other. Copies of p and q remain in the molecule, but do not act as pointers.

We may assume that all ld's are applied after all dlad's.

We consider that cyclically sorted permutations correspond to assembled genes. In this way, we formalize the gene assembly process as sorting of a permutation.

We call a *strategy* a composition of ed-operations. Let π be a permutation. We say that a strategy Φ is a *strategy for* π if Φ is applicable to π, and there are no ed operations applicable to $\Phi(\pi)$. We call Φ a sorting strategy for π if $\Phi(\pi)$ is a sorted permutation.

Let $\Phi = \mathsf{ed}_{i_m} \circ \mathsf{ed}_{i_{m-1}} \circ \ldots \circ \mathsf{ed}_{i_1}$ be a strategy. For any two operations ed_{i_k} and ed_{i_l} with $k < l$ we say that ed_{i_k} is used earlier than ed_{i_l} in Φ.

Example 1. *(i) Permutation $\pi_1 = 52413$ has several strategies, but none of them is sorting. Indeed, $\mathsf{ed}_2(\pi_1) = 54123$ and $\mathsf{ed}_3(\pi_1) = 52341$.*

(ii) Permutation $\pi_2 = 1\,3\,5\,2\,4$ has both sorting and non-sorting strategies. Indeed, $\mathsf{ed}_3(\pi_2) = 1\,5\,2\,3\,4$, which is unsortable. However, $\mathsf{ed}_2(\ \mathsf{ed}_4(\ \pi_2)) = 1\,2\,3\,4\,5$ is sorted.

The following example demonstrates that the number of sorting strategies and the ratio of sorting/non-sorting strategies may vary for different sortable permutations.

Example 2. *(i) Permutation $\pi_1 = 1325476\ldots(2n-1)(2n-2)(2n+1)2n$ has only one applicable strategy $\mathsf{ed}_{2n} \circ \mathsf{ed}_{2n-2} \circ \ldots \circ \mathsf{ed}_6 \circ \mathsf{ed}_4 \circ \mathsf{ed}_2$. It is also a sorting strategy for π_1;*

(ii) Permutation $\pi_2 = 13\ldots(2n+1)2n(2n-2)\ldots2$ has exactly $n!$ applicable strategies. All of them sort π_2. Indeed, $\mathsf{ed}_{2n} \circ \mathsf{ed}_{2n-2} \circ\ \ldots\ \circ \mathsf{ed}_2(\pi_2) = 12\ldots(2n+1)$. Any composition of the operations ed_{2i}, $2 \le 2i+1 \le 2n$ also sorts π_2. There are no other operations applicable to π_2;

(iii) Permutation $\pi_3 = 13245768\ldots(4n-3)(4n-1)(4n-2)4n$ has exactly $2^n n!$ applicable strategies. All of these strategies sort π_3. Indeed, we can sort each

substring $(4k+1)(4k+3)(4k+2)(4k+4)$*, where* $0 \le 4k \le 4n-4$*, by applying either* ed_{4k+2} *or* ed_{4k+3}*. The order in which each of the* n $4-$*letter blocks is sorted is arbitrary.*

(iv) Permutation $\pi_4 = 1357246$ *has* $3! = 6$ *sorting strategies and* 6 *unsuccessful strategies. Indeed, one can sort* π_4 *by applying* ed_2*,* ed_4*,* ed_6 *in any order. If one applies in any order either* ed_3 *and* ed_5*, or* ed_2 *and* ed_5*, or* ed_3 *and* ed_6*, then* π_4 *is not sorted;*

(v) Permutation $\pi_5 = 135792468$ *has* $4! = 24$ *sorting strategies and* 36 *unsuccessful strategies. Indeed, by applying operations* ed_2*,* ed_4*,* ed_6*,* ed_8 *in any order one sorts* π_5*. By applying any other subset of operations* $\{\mathsf{ed}_1, \mathsf{ed}_2, \ldots, \mathsf{ed}_8\}$ *in any order we do not get a sorted permutation from* π_5*.*

4. Dependency Graph

We recall now the notion of a dependency graph introduced in [13, 26]. Based on it we characterized in [13, 26] the ed-sortable permutations. We presented a method to decide effectively the ed-sortability.

A dependency graph suggests the order of ed operations to be used in a sorting strategy. Intuitively, if there is a path from vertex p to vertex q, then in any strategy where ed_q is used, operation ed_p is used earlier than ed_q.

Definition 3 ([13, 26]) *We define the* dependency graph *for a permutation* π *over* Σ_n *as a directed graph* $\Gamma_\pi = (\Sigma_n, E_\pi)$*, where for* $1 \le i \le n$ *and* $2 \le j \le n-1$ *we have the following edges:*

- $(i,j) \in E_\pi$*, if* $(j-1)i(j+1) \le_s \pi$*,*
- $(j,j) \in E_\pi$*, if* $(j+1)\,(j-1) \le_s \pi$*, and*
- $(1,1), (n,n) \in E_\pi$*;*

We also denote $\Gamma_\pi = (V_\pi, E_\pi)$ for permutation π, where $V_\pi = \Sigma_\pi$.

For $p, q \in \mathsf{dom}(\pi)$ we define by $\Gamma_{\pi,p}$, $\Gamma_{\pi,p,q}$ and $\Gamma^S_{\pi,p}$ subgraphs of Γ_π as stated in Section 2. I.e., $\Gamma_{\pi,p} = (T_{\pi,p}, E_{\pi,p})$, where $T_{\pi,p} = \{r \in V_\pi | r \to^+_{\Gamma_\pi} p\} \cup \{p\}$ and $E_{\pi,p} = \{(i,j) \in E_\pi | i,j \in T_{\pi,p}\}$. For $p,q \in V_\pi$, where $p < q$ we let $\Gamma_{\pi,p,q} = (N_{p,q}, E_{\pi,p,q})$, where $E_{\pi,p,q} = \{(i,j) \in E_\pi | i,j \in N_{p,q}\}$. $\Gamma^S_{\pi,p} = (S_{\pi,p}, E^S_{\pi,p})$, where $S_{\pi,p} = \{r \in V_\pi | p \to^+_{\Gamma_\pi} r\} \cup \{p\}$ and $E^S_{\pi,p} = \{(i,j) \in E_\pi | i,j \in S_{\pi,p}\}$.

Example 4. *Let* π *be a permutation* $\pi = 1\,3\,9\,5\,2\,4\,7\,6\,8\,10$*. We construct its dependency graph* $\Gamma_\pi = (V_\pi, E_\pi)$ *as follows (see Figure 3).* $V_\pi = \{1,2,3,4,5,6,7,8,9,10\}$*. Integers* 1 *and* 10 *are in self-loops, i.e.,* $(1,1), (10,10) \in E_\pi$*. Substrings* $1\,3$*,* $2\,4$*,* $6\,8$*,* $8\,10$ *are present in permutation* π*. In this way, integers* $2,3,7,9$ *have no incoming edges in* Γ_π*. Subsequence* $9\,7$ *is present in* π*. In this way, integer* 8 *is in self-loop in* Γ_π*. We have substrings* $3\,9\,5$*,* $4\,7\,6$*,* $5\,2\,4\,7$ *in* π*. In this way, we have edges* $(9,4), (7,5), (2,6), (4,6) \in E_\pi$*.*

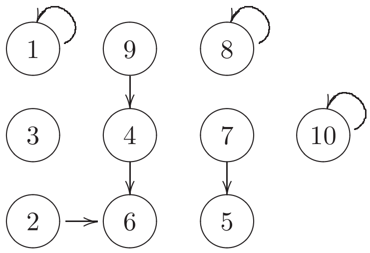


Fig. 3. The dependency graph associated to permutation $\pi = 1\,3\,9\,5\,2\,4\,7\,6\,8\,10$.

Lemma 5. [13] *Let π be a permutation over Σ_n and $\Gamma_\pi = (\Sigma_n, E)$ be its dependency graph. If there is a path from p to q in Γ_π, then in any strategy for π where ed_q is used, ed_p is used before ed_q.*

Example 6. *Consider permutation π and its dependency graph Γ_π from Example 4. By Lemma 5 in any strategy applicable to π if ed_6 is used, then operations ed_2, ed_4 and ed_9 are used earlier. Moreover, ed_9 is used earlier than ed_4. In any strategy for π where operation ed_5 is used, operation ed_7 is used earlier.*

The following lemma shows how the dependency graph changes along the transformation of the corresponded permutation by an ed-operation.

Lemma 7. [26] *Let π be a permutation and p an integer such that ed_p is applicable to π. If $\Gamma = (V, E)$ is the dependency graph of π and $\Gamma' = (V, E')$ is the dependency graph of $\mathsf{ed}_p(\pi)$, then*

(i) for any $(i, j) \in (E \setminus E') \cup (E' \setminus E)$, $i = p$, or $j = p - 1$, or $j = p + 1$. Moreover, $(p, j) \in E' \setminus E$ if and only if $(p - 1, j) \in E$ and $(p + 1, j) \in E$;
(ii) $(p - 1, p - 1), (p, p), (p + 1, p + 1) \in E'$.

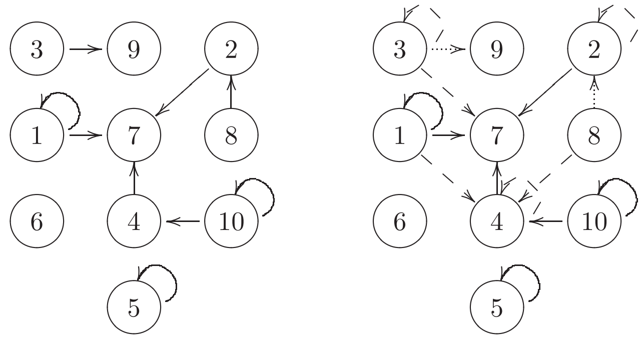
Example 8. *Consider permutations $\pi = 6\,2\,4\,1\,8\,3\,10\,5\,7\,9$ and $\pi' = \mathsf{ed}_3(\pi)$, i.e., $\pi' = 6\,2\,3\,4\,1\,8\,10\,5\,7\,9$. The corresponding graphs are presented on Figure 4. One can see, that after application of ed_3 in the corresponded dependency graph edges $(3, 9)$ and $(8, 2)$ disappear and edges $(2, 2)$, $(3, 3)$, $(4, 4)$, $(1, 4)$, $(3, 7)$, $(8, 4)$ appear.*

5. Fixed integers and blocks

For a permutation π over Σ_n we denote set of integers $Fix(\pi) = \{p|(p-1)p(p+1) \leq_s \pi$ or $(p+1)(p-1) \leq_s \pi\} \cup \{1, n\}$ and call them *fixed integers* of π. Note that the set of fixed integers $Fix(\pi)$ is equivalent to the set of nodes having self-loops in the dependency graph Γ_π.

We agree about the following notations related to the set of fixed integers.

- $\mathbb{B}_\pi = \{N_{p,q} \in 2^{\mathsf{dom}(\pi)} | N_{p,q} \cap Fix(\pi) = \{p, q\}\}$;



(a) (b)

Fig. 4. (a) Dependency graph associated to $\pi = 6\,2\,4\,1\,8\,3\,10\,5\,7\,9$; (b) Dependency graph associated to $\pi' = \mathsf{ed}_3(\pi) = 6\,2\,3\,4\,1\,8\,10\,5\,7\,9$.

- $\mathbb{B}^e_\pi = \{N_{p,q} \in \mathbb{B}_\pi | q - p \vdots 2\}$;
- $\mathbb{B}^o_\pi = \{N_{p,q} \in \mathbb{B}_\pi | q - p \not\vdots 2\}$;
- $B^e_\pi = \bigcup_{N_{p,q} \in \mathbb{B}^e_\pi} N_{p,q}$;
- $B^o_\pi = \bigcup_{N_{p,q} \in \mathbb{B}^o_\pi} N_{p,q}$;
- $N^e_\pi = \bigcup_{N_{p,q} \in \mathbb{B}_\pi} N^e_{p,q}$;
- $N^o_\pi = \bigcup_{N_{p,q} \in \mathbb{B}_\pi} N^o_{p,q}$.

Note that $\mathbb{B}_\pi = \mathbb{B}^e_\pi \oplus \mathbb{B}^o_\pi$, $B^e_\pi \cup B^o_\pi = \mathsf{dom}(\pi)$, $\mathsf{dom}(\pi) \setminus (N^e_\pi \cup N^o_\pi) = Fix(\pi)$.

Example 9. *Let* $\pi = 1\,3\,5\,11\,7\,9\,6\,2\,4\,8\,10\,12$. *Then*

- $\mathbb{B}_\pi = [\{1,2,3,4,5\}, \{5,6,7,8,9,10\}, \{10,11,12\}]$*;*
- $\mathbb{B}^e_\pi = [\{1,2,3,4,5\}, \{10,11,12\}]$*;*
- $\mathbb{B}^o_\pi = [\{5,6,7,8,9,10\}]$*;*
- $B^e_\pi = \{1,2,3,4,5,10,11,12\}$*;*
- $B^o_\pi = \{5,6,7,8,9,10\}$*;*
- $N^e_\pi = \{3,7,9\}$*;*
- $N^o_\pi = \{2,4,6,8,11\}$.

The notion of a *block* plays the key role in the main results of the paper. We say that an element $N_{p,q} \in \mathbb{B}^\pi$ is a *block*. The role of blocks and fixed integers we discuss later.

The following lemma presents the form of subsequences corresponding to blocks. Its proof is straightforward.

Lemma 10. *In any permutation* π *for any block* $N_{p,q} \in \mathbb{B}^e_\pi$ *we have subsequences* $p(p+2)(p+4)\dots(q-2)q \le_s \pi$, $(p+1)(p+3)\dots(q-1) \le_s \pi$, *and for any block* $N_{p,q} \in \mathbb{B}^o_\pi$ *we have subsequences* $p(p+2)(p+4)\dots(q-3)(q-1) \le_s \pi$,

$(p+1)(p+3)\ldots(q-2)q \leq_s \pi$.

We need later the following technical lemma.

Lemma 11. *Let π be a permutation and $N_{p,q} \in \mathbb{B}_\pi$.*

(a) *There are no such integers $i \in N^e_{p,q}$ and $j \in N^o_{p,q}$ where either $(i,j) \in E_\pi$ or $(j,i) \in E_\pi$.*

(b) *Either for all $i \in N^e_{p,q}$ we have subsequences $(i-1)(i+1)i \leq_s \pi$ and for all $j \in N^o_{p,q}$ we have subsequences $j(j-1)(j+1) \leq_s \pi$ or for all $i \in N^e_{p,q}$ we have subsequences $i(i-1)(i+1) \leq_s \pi$ and for all $j \in N^o_{p,q}$ we have subsequences $(j-1)(j+1)j \leq_s \pi$.*

(c) *Let $t, i \in N_{p+1,q-1}$ and $(t,i) \in E_\pi$. If $i < t$, then for any $j \in N_{p+1,q-1}$ where $t \rightarrow^+_{\Gamma_{\pi,p,q}} j$ we have $j < t$. If $i > t$, then for any $j \in N_{p+1,q-1}$ where $t \rightarrow^+_{\Gamma_{\pi,p,q}} j$ we have $j > t$.*

(d) *Graph $\Gamma_{\pi,p,q}$ does not contain non-loop cycles.*

Proof.

(a) Consider an integer $i \in N^e_{p,q}$. Note, that $i-1, i+1 \in N^o_{p,q} \cup \{p,q\}$. By Lemma 10 we have subsequence $(i-1)(i+1) \leq_s \pi$ and we do not have subsequence $(i-1)j(i+1) \leq_s \pi$ for any $j \in N^o_{p,q}$. Similarly, for any integer $i \in N^o_{p,q}$ there is no $j \in N^e_{p,q}$ where $(i-1)j(i+1) \leq_s \pi$.

(b) Let $t \in N^o_{p,q}$ where $t(t-1)(t+1) \leq_s \pi$. If $t-1 > p$, then $(t-2)t(t-1) \leq_s \pi$, and $(t-2)(t-3)(t-1) \leq_s \pi$. By continuing to reason in the same manner we conclude that for all $i \in N^o_{p,t+1}$ we have subsequence $i(i-1)(i+1) \leq_s \pi$ and for all $j \in N^e_{p,t}$ we have subsequence $(j-1)(j+1)j \leq_s \pi$. Similarly, if $t+1 < q$, then $t(t+2)(t+1) \leq_s \pi$. If $t+2 < q$, then $(t+2)(t+1)(t+3) \leq_s \pi$. We continue to reason in the same way. Then for all $i \in N^o_{p,t+1}$ we have subsequence $i(i-1)(i+1) \leq_s \pi$ and for all $j \in N^e_{p,t}$ we have subsequence $(j-1)(j+1)j \leq_s \pi$. I.e., for all $i \in N^o_{p,q}$ we have subsequence $i(i-1)(i+1) \leq_s \pi$ and for all $j \in N^e_{p,q}$ we have subsequence $(j-1)(j+1)j \leq_s \pi$. The symmetrical case where for all $i \in N^o_{p,q}$ we have subsequence $(i-1)(i+1)i \leq_s \pi$ and for all $j \in N^e_{p,q}$ we have subsequence $j(j-1)(j+1) \leq_s \pi$ is proved in the analogous way.

(c) Consider integers $t, t' \in N_{p+1,q-1}$ where $(t,t') \in E_\pi$. Assume $t \in N^o_{p,q}$ (case $N^e_{p,q}$ is symmetric). Assume that $t' > t$. If there is $t'' \in N_{p+1,q-1} \setminus \{t'\}$ such that $(t', t'') \in E_\pi$, then we show that $t'' > t$. Indeed, $(t'-1)t(t'+1) \leq_s \pi$, and $(t''-1)t'(t''+1) \leq_s \pi$. Since $t \in N^o_{p,q}$, then by Property (a) $t' \in N^o_{p,q}$ as well. Since $t' > t$, then according to Lemma 10 $tt' \leq_s \pi$. Then $(t'-1)(t'+1)t' \leq_s \pi$. Then by Property (b) for any $i \in N^o_{p,q}$ we have subsequence $(i-1)(i+1)i \leq_s \pi$. By Property (a) integer $t'' \in N^o_{p,q}$. Then $(t''-1)t'(t''+1)t'' \leq_s \pi$. Then according to Lemma 10 $t'' > t' > t$. By continuing to reason in the same manner we proof that for any $s \in N_{p+1,q-1}$ such that $t' \rightarrow^+ s$ we have that $s > t'$. If there is an integer

$t''' \in N_{p+1,q-1}$ such that $(t, t''') \in E_\pi$, then $t''' > t$. Indeed, by Property (a) $t''' \in N^o_{p,q}$. Then $(t'''-1)(t'''+1)t''' \leq_s \pi$. Since $(t, t''') \in E_\pi$, we have $(t'''-1)t(t'''+1)t''' \leq_s \pi$. By Lemma 10 $t''' > t$. We repeat reasoning as for the case $t \to^+ t' \to^+ t''$. Then we show that for any integer $s \in N_{p+1,q-1}$ such that $t''' \to^+ s$ we have that $s > t''' > t$. I.e., for any $s \in N_{p+1,q-1}$, $t \to^+ s$ we have $s > t$. Case $t' < t$ where $(t, t') \in E_\pi$ is symmetric;

(d) Assume that there is a non-loop cycle containing integers only from one block. Consider an integer i from the cycle. Then we have $i \to^+ i$. According to Property (c) either $i > i$ or $i < i$. This is a contradiction.

□

6. Forbidden Integers

Consider the set of integers $F(\pi) = \{p \in \mathsf{dom}(\pi) |$ there is no strategy applicable to π where ed_p is used$\}$. We say that the set $F(\pi)$ is the set of *forbidden integers* in π.

The set $F(\pi)$ was characterized in [26].

Theorem 12. [26] *For a permutation π over Σ_n and $p \in \Sigma_n$, $p \in F(\pi)$ if and only if the subgraph $\Gamma_{\pi,p} = (T_{\pi,p}, E_{\pi,p})$ is cyclic or $q-1, q \in T_{\pi,p}$ for some q.*

Note that $Fix(\pi) \subseteq F(\pi)$. Contrary to the rest of forbidden integers, fixed integers can be "detected" without constructing and analyzing the dependency graph. For each integer p one needs to check whether either subsequence $(p-1)p(p+1)$ or subsequence $(p+1)(p-1)$ is present in π. Also, integers 1 and n are always fixed. We remind that in the dependency graph Γ_π integers from $Fix(\pi)$ are represented by vertices with self-loops.

Example 13. *Consider permutation $\pi = 8\,1\,10\,3\,12\,5\,2\,7\,4\,6\,9\,11\,13$. Its dependency graph is presented in Figure 5. Integers 1, 7, 13 are fixed. Integers 3 and 11 are forbidden, since there is a path to them in Γ_π from fixed integer 7. Integer 9 is forbidden since there is edge $(1, 9)$ in Γ_π and integer 1 is fixed. In this way, $F(\pi) = 1, 3, 7, 9, 11, 13$.*

7. Sortable permutations

We recall the following theorem from [26] describing an effective procedure to check the ed-sortability of a permutation.

Theorem 14. *Permutation π is sortable if and only if $\pi|_{F(\pi)}$ is sorted.*

Example 15. *Let us consider permutation π from Example 13. Set of forbidden integers is sorted in π: $\pi|_{F(\pi)} = 1\,3\,7\,9\,11\,13$. In this way, permutation π is ed-sortable. Indeed, $(\mathsf{ed}_8 \circ \mathsf{ed}_4 \circ \mathsf{ed}_{12} \circ \mathsf{ed}_6 \circ \mathsf{ed}_2 \circ \mathsf{ed}_{10})(\pi) = 1\,2\,3\,4\,5\,6\,7\,8\,9\,10\,11\,12\,13$.*

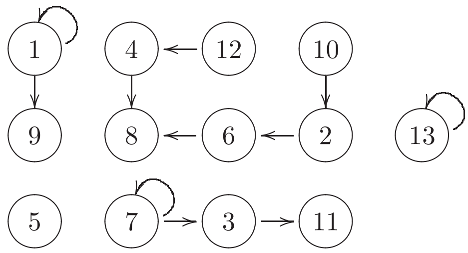


Fig. 5. Dependency graph associated to $\pi = 8\ 1\ 10\ 3\ 12\ 5\ 2\ 7\ 4\ 6\ 9\ 11\ 13$.

By using Lemma 10 we obtain the following lemma and corollary which represents the form of subsequences of integers of a block $N_{p,q} \in \mathbb{B}^o_\pi$, where π is a sortable permutation.

Lemma 16. *Let π be a permutation. If for a block $N_{p,q} \in \mathbb{B}^o_\pi$ and $k \in N_{p+1,q-1}$ we have either $kpq \leq_s \pi$, or $pqk \leq_s \pi$, then π cannot be sorted.*

Proof.
Consider an unsigned permutation π and a block $N_{p,q} \in \mathbb{B}_\pi$ such that $kpq \leq_s \pi$ where $k \in N_{p+1,q-1}$. Then $k \in N^o_{p,q}$. It follows by Lemma 10 that there is such integer $k' \in N^o_{p,q}$ that $k'p(k'+2) \leq_s \pi$. Then by Theorem 12 $k'+1 \in F(\pi)$. It follows by Lemma 10 that $k'(k'+2)(k'+1) \leq_s \pi$. Then either $(k'+1)q \leq_s \pi$ or $q(k'+1) \leq_s \pi$. In case $q(k'+1) \leq_s \pi$ permutation π cannot be sorted. If $(k'+1)q \leq_s \pi$, then exists integer $k'' \in N^o_{p,q}$, $k'' > k'$ such that $k''(k'+1)(k''+2) \leq_s \pi$. Then $(k''+1) \in F(\pi)$ as well. Moreover, either $(k''+1)q \leq_s \pi$ or $q(k''+1) \leq_s \pi$. In this way, we can find integer $t \in N^e_{p,q}$, where $k'+1 \in T_{\pi,t}$ such that $qt \leq_s \pi$. Since $t \in F(\pi)$, permutation π cannot be sorted.

In the similar way we show that if $pqk' \leq_s \pi$ for some $k' \in N_{p+1,q-1}$, then π cannot be sorted.

□

Corollary 17. *Let π be a sortable permutation and $N_{p,q} \in \mathbb{B}^o_\pi$, then*

(i) for any integer $t \in N^o_{p,q}$, if $pq \leq_s \pi$, then $(t-1)(t+1)t \leq_s \pi$. Otherwise, $t(t-1)(t+1) \leq_s \pi$;
(ii) for any integer $t \in N^e_{p,q}$, if $pq \leq_s \pi$, then $t(t-1)(t+1) \leq_s \pi$. Otherwise, $(t-1)(t+1)t \leq_s \pi$;

Example 18. *Consider the following permutations:*

- *Consider permutation $\pi_1 = 2\,4\,1\,3\,5\,7\,9\,6\,8\,10$. We have $\mathbb{B}^o_{\pi_1} = \{N_{5,10}\}$ and $\mathbb{B}^e_{\pi_1} = \{N_{1,5}\}$. For block $N_{1,5}$ we have subsequences $135 \leq_s \pi_1$ and $24 \leq_s \pi$. For block $N_{5,10}$ we have subsequences $579 \leq_s \pi_1$ and $6\,8\,10 \leq_s \pi_1$. Permutation π_1 is sortable. For integers 6, 8 from $N^o_{5,10}$ we have subsequences*

$576 \leq_s \pi_1$ *and* $798 \leq_s \pi_1$. *For integers* 7 *and* 9 *from* $N^e_{5,10}$ *we have subsequences* $768 \leq_s \pi_1$ *and* $9\,8\,10 \leq_s \pi_1$*;*

- *Let* $\pi_2 = 241365$. *Then* $\mathbb{B}_\pi = \mathbb{B}^o_\pi = \{N_{1,6}\}$. *We have subsequence* $216 \leq_s \pi_2$. *We observe that* π_2 *cannot be sorted by* ed.

8. Applicable strategies

The following lemma characterizes strategies applicable to π.

Lemma 19. *Let* π *be a permutation. A strategy* Φ *can be applied to* π *if and only if*

(i) a subgraph induced by $\mathsf{dom}(\Phi)$ *from the dependency graph* Γ_π *is acyclic, and*
(ii) for any $i \in \mathsf{dom}(\Phi)$ *we have* $T_{\pi,i} \subseteq \mathsf{dom}(\Phi)$*, and*
(iii) for any $i \in \mathsf{dom}(\Phi)$ *if* $j \in T_{\pi,i}$*, then* ed_j *is used before* ed_i *in* Φ*, and*
(iv) there is no integer $t \in \mathsf{dom}(\Phi)$ *such that* $t+1 \in \mathsf{dom}(\Phi)$.

Proof.

The direct implication follows by Lemma 5, Theorem 12 and the following observation [13]: for any integer t ed-operations ed_t and ed_{t+1} can never be used in the same strategy for any permutation.

To prove the inverse implication we show that if conditions (i)–(iv) are satisfied for a permutation π and a composition $\Phi = \Phi' \circ \mathsf{ed}_i$, then ed_i is applicable to π. Indeed, assume ed_i is not applicable to π. Then, either $i \in F(\pi)$ or $i \notin F(\pi)$, but $|T_{\pi,i}| > 1$. If $i \in F(\pi)$, then by Theorem 12 either $\Gamma_{\pi,i} = (T_{\pi,i}, E_{\pi,i})$ contain cycle or there is $j \in T_{\pi,i}$ such that $j+1 \in T_{\pi,i}$. Then by condition (ii) we have either cycle in the graph induced by $\mathsf{dom}(\Phi)$ or $j, j+1 \in \mathsf{dom}(\Phi)$. But this contradicts to conditions (i) and (iv). If $i \notin F(\pi)$ and $|T_{\pi,i}| > 1$, then by condition (iii) for any $j \in T_{\pi,i}$ operation ed_j is used in Φ before ed_i, what contradicts to our assumption about the structure of Φ. In this way, if conditions (i)–(iv) are satisfied for a permutation π and a composition $\Phi = \Phi' \circ \mathsf{ed}_i$, then $T_{\pi,i} = \{i\}$ and ed_i is applicable to π.

The inverse implication we prove by induction. Let $\Phi = \mathsf{ed}_{i_m} \circ \mathsf{ed}_{i_{m-1}} \circ \ldots \circ \mathsf{ed}_{i_2} \circ \mathsf{ed}_{i_1}$. The graph induced by $\mathsf{dom}(\Phi)$ from the dependency graph Γ_π we denote as $\Gamma_{\pi,\Phi} = (\mathsf{dom}(\Phi), E_{\pi,\Phi})$, where $E_{\pi,\Phi} = \{(i,j) | i,j \in \mathsf{dom}(\Phi)\}$. We denote π_s and Φ_{s+1} for all s, $0 \leq s \leq m$ as follows:

$$\pi_s = \begin{cases} \pi, & \text{if } s = 0 \\ \mathsf{ed}_{i_s}(\pi_{s-1}), & \text{if } s > 0 \end{cases} \qquad \Phi_{s+1} = \begin{cases} \mathsf{ed}_{i_m} \circ \mathsf{ed}_{i_{m-1}} \circ \ldots \circ \mathsf{ed}_{i_{s+1}}, & \text{if } s < m \\ id, & \text{if } s = m \end{cases}$$

We show that for any $0 \leq s \leq m$ properties (i)–(iv) hold for π_s and Φ_{s+1}.

Properties (i)–(iv) hold for π_0 by the condition of the inverse implication.

Assume that for some k, $1 \leq k < m$ properties (i)–(iv) hold for all π_j, where $0 \leq j \leq k$. Then, as it was shown, $T_{\pi_k, i_{k+1}} = \{i_{k+1}\}$ and $\mathsf{ed}_{i_{k+1}}$ can be applied to

π_k. By Lemma 7 it follows that $E_{\pi_{k+1},\Phi} \subseteq E_{\pi_k,\Phi}$. Indeed,

$$E_{\pi_{k+1}} \setminus E_{\pi_k} = \{(i_{k+1}, j) | \{(i_{k+1}-1, j), (i_{k+1}+1, j)\} \cap E_{\pi_k} \neq \emptyset\} \cup \\ \cup \{(i_{k+1}-1, i_{k+1}-1), (i_{k+1}, i_{k+1}), (i_{k+1}+1, i_{k+1}+1)\}.$$

By property (iv) $i_{k+1}-1, i_{k+1}+1 \notin \mathsf{dom}(\Phi_\mathsf{k})$ since $i_{k+1} \in \mathsf{dom}(\Phi)$. Then there is no $j \in \mathsf{dom}(\Phi)$ such that $(i_{k+1}, j) \in E_{\pi_{k+1}} \setminus E_{\pi_k}$. Since for any $(i_{k+1}, j) \in E_{\pi_k}$ where $j \in \mathsf{dom}(\Phi)$ we have $(i_{k+1}, j) \notin E_{\pi_{k+1}}$, then $(E_{\pi_{k+1}} \setminus \{(i_{k+1}-1, i_{k+1}-1), (i_{k+1}, i_{k+1}), (i_{k+1}+1, i_{k+1}+1)\}) \subseteq E_{\pi_k}$. I.e., $E_{\pi_{k+1},\Phi} \subseteq E_{\pi_k,\Phi}$. Moreover, for any integer $i \in \mathsf{dom}(\Phi_\mathsf{k})$ such that $i_{k+1} \in T_{\pi_k,i}$ we have $T_{\pi_k,i} \setminus T_{\pi_{k+1},i} = \{i_{k+1}\}$. For all $i \in \mathsf{dom}(\Phi_\mathsf{k})$ such that $i_{k+1} \notin T_{\pi_k,i}$ we have $T_{\pi_k,i} = T_{\pi_{k+1},i}$. We have $\mathsf{dom}(\Phi_\mathsf{k}) \setminus \mathsf{dom}(\Phi_{\mathsf{k+1}}) = \{i_k\}$. In this way, for any $i \in \mathsf{dom}(\Phi_{\mathsf{k+1}})$ we have $T_{p_{i_{k+1}},i} \subseteq T_{p_{i_k},i}$. Properties (i), (ii), (iii) and (iv) hold for π_{k+1}.

□

The following corollary follows immediately.

Corollary 20. *Let π be a permutation and Γ_π be its dependency graph. Let $T_{\pi,i} = \{i\}$. Then ed_i is applicable to π.*

9. Sorting Strategies

Let π be a permutation and let $\mathbb{B}^o_\pi = \{N_{p_1,q_1}, N_{p_2,q_2}, \ldots, N_{p_k,q_k}\}$. We define a set of subsets of $\mathsf{dom}(\pi)$ $\mathbb{S}_\pi = \{S \in 2^{\mathsf{dom}(\pi)} | S = (B^e_\pi \cap N^o_\pi) \cup (\bigcup_{1 \le i \le k} (N^o_{p_i,t_i} \cup N^e_{t_i,q_i}))$, where $t_i \in N^e_{p_i-2,q_i}$ for all $i, 1 \le i \le k\}$.

Example 21. *Let us consider permutation $\pi = 1\,3\,5\,11\,7\,9\,6\,2\,4\,8\,10\,12$ from Example 9. Here $B^e_\pi \cap N^o_\pi = \{2, 4, 11\}$, $\mathbb{B}^o_\pi = [\{5, 6, 7, 8, 9, 10\}]$. Then $\mathbb{S}_\pi = [\{2, 4, 7, 9, 11\}, \{2, 4, 6, 9, 11\}, \{2, 4, 6, 8\}]$.*

We prove in this section that any set from $\mathbb{S}_\pi$ is the domain of a sorting strategy for permutation π, and viceversa, any strategy with the domain from $\mathbb{S}_\pi$ is a sorting strategy for π.

In the following two lemmas we analyze pathes in the dependency graph between integers from different blocks. We show that integers from set $B^e_\pi \cap N^o_\pi$ may have as predecessors only integers from the same set $B^e_\pi \cap N^o_\pi$. Integers from a block $N_{p,q} \in \mathbb{B}^o_\pi$ may have as predecessors only integers from the same block $N_{p,q}$ and integers from set $B^e_\pi \cap N^o_\pi$.

Lemma 22. *Let π be a sortable permutation. Then for any integer $i \in (B^e_\pi \cap N^o_\pi)$ we have $T_{\pi,i} \subseteq (B^e_\pi \cap N^o_\pi)$, and $\Gamma_{\pi,i}$ is an acyclic graph.*

Proof.

Let $i \in \mathsf{dom}(\Phi) \cap (B^e_\pi \cap N^o_\pi)$ and $(j, i) \in E_\pi$. I.e., $(i-1)j(i+1) \le_s \pi$, where $p(i-1)q \le_s \pi$, if $i-1 > p$ and $p(i+1)q \le_s \pi$, if $i+1 < q$. In this way, $pjq \le_s \pi$. If $j \in N_{p+1,q-1}$, then it follows by Lemma 11 that $j \in N^o_{p,q}$. Graph $\Gamma_{\pi,p,q}$ does not have cycles by Lemma 11.

If $j \notin N_{p+1,q-1}$, then by Theorem 14 $j \notin F(\pi)$, and so, $j \notin Fix(\pi)$. Moreover, since $j \notin F(\pi)$, then by Theorem 12 graph $\Gamma_{\pi,j}$ does not have a cycle.

If $j \in N_{p',q'}$ such that $N_{p',q'} \in \mathbb{B}^e_\pi$ and $p' \neq p$, $q' \neq q$, then $j \in N^o_\pi$. Indeed, if $j \in N^e_\pi$, by Lemma 10 we would have either $p'pq'q \leq_s \pi$, or $p'pqq' \leq_s \pi$, or $pp'qq' \leq_s \pi$, or $pp'q'q \leq_s \pi$. Since $p, p', q, q' \in Fix(\pi)$ and $p < q$, $p' < q'$, and either $q' < p$ or $q < p'$, then by Theorem 14 permutation π cannot be sorted, contradiction.

Assume that $j \in N_{p',q'}$ where $N_{p',q'} \in \mathbb{B}^o_\pi$ and $p' \neq p$, $q' \neq q$. Then by Lemma 16 since π is sortable we have either $p'jq' \leq_s \pi$, or $q'p'j \leq_s \pi$, or $jq'p' \leq_s \pi$. But, then either

$$\begin{array}{lll} p'pq'q \leq_s \pi, \text{ or} & p'pqq' \leq_s \pi, \text{ or} & pp'qq' \leq_s \pi, \text{ or} \\ pp'q'q \leq_s \pi, \text{ or} & q'p'q \leq_s \pi, \text{ or} & pq'p' \leq_s \pi. \end{array}$$

Then π is non-sortable, contradiction.

In this way, for any j where $(j,i) \in E_\pi$ we have $j \in (B^e_\pi \cap N^o_\pi)$. Then $T_{\pi,i} \subseteq (B^e_\pi \cap N^o_\pi)$. Moreover, $j \notin F(\pi)$, then by Theorem 12 $\Gamma_{\pi,j}$ does not have cycles. Then graph $\Gamma_{\pi,i}$ does not have cycles either.

□

Lemma 23. *Let π be a sortable permutation. Then for any integer $i \in B^o_\pi$ we have that*

(i) $T_{\pi,i} \subseteq (B^e_\pi \cap N^o_\pi) \cup N_{p+1,q-1}$;
(ii) If $i \in N^e_{p,q}$, then $T_{\pi,i} \cap N_{p+1,q-1} \subseteq N^e_{p,q}$. Otherwise, $T_{\pi,i} \cap N_{p+1,q-1} \subseteq N^o_{p,q}$;
(iii) Set $\Gamma_{\pi,i}$ is an acyclic graph.

Proof.
Assume that $i \in N_{p+1,q-1}$ and $N_{p,q} \in \mathbb{B}^o_\pi$. By Lemma 11 it follows that there are no cycles in $\Gamma_{\pi,p,q}$. Moreover, if $i \in N^o_{p,q}$ ($i \in N^e_{p,q}$), then $T_{\pi,i} \cap N_{p+1,q-1} \subseteq N^o_{p,q}$ ($T_{\pi,i} \cap N_{p+1,q-1} \subseteq N^e_{p,q}$).

Assume that $(j,i) \in E_\pi$. I.e., $(i-1)j(i+1) \leq_s \pi$.

Let $j \in N_{p+1,q-1}$. If $i \in N^e_{p,q}$, then by Lemma 11 $j \in N^e_{p,q}$ as well. Otherwise, $j \in N^o_{p,q}$.

Let us consider the case where $j \notin N_{p+1,q-1}$. Since π is sortable, by Lemma 16 we have either $pjq \leq_s \pi$, or $qpj(i+1) \leq_s \pi$, or $(i-1)jqp \leq_s \pi$. It follows by Theorem 14 that $j \notin F(\pi)$. In particular we have $j \notin Fix(\pi)$. Moreover, since $j \notin F(\pi)$, then by Theorem 12 set $\Gamma_{\pi,j}$ does not contain cycles.

Let $j \in N_{p',q'}$ where $p' \neq p$ and $q' \neq q$. Let either $N_{p',q'} \in \mathbb{B}^e_\pi$ or $N_{p',q'} \in \mathbb{B}^o_\pi$.

- If $N_{p',q'} \in \mathbb{B}^e_\pi$, then $j \in N^o_{p',q'}$. Indeed, if we would have $j \in N^e_{p',q'}$, then

$$\begin{array}{lll} pp'qq' \leq_s \pi, \text{ or} & p'pqq' \leq_s \pi, \text{ or} & pp'q'q \leq_s \pi, \text{ or} \\ p'pq'q \leq_s \pi, \text{ or} & p'q'qp \leq_s \pi, \text{ or} & p'qq'p \leq_s \pi, \text{ or} \\ p'qpq' \leq_s \pi. & & \end{array}$$

In all these cases $\pi|_{Fix(\pi)}$ is not sorted. Then π is non-sortable, contradiction.

- Let $j \in N_{p',q'}$ where $p' \neq p$, $q' \neq q$, and $N_{p',q'} \in \mathbb{B}^o_\pi$. Since π is sortable, we have either $p'jq' \leq_s \pi$, or $q'p'j \leq_s \pi$, or $jq'p' \leq_s \pi$.
  - If $pq \leq_s \pi$, then we have either
  $$\begin{array}{lll} p'pq'q \leq_s \pi, \text{ or} & p'pqq' \leq_s \pi, \text{ or} & pp'q'q \leq_s \pi, \text{ or} \\ pp'qq' \leq_s \pi, \text{ or} & q'p'pq \leq_s \pi, \text{ or} & q'pp'q \leq_s \pi, \text{ or} \\ pq'p'q \leq_s \pi, \text{ or} & pq'qp' \leq_s \pi, \text{ or} & pqq'p' \leq_s \pi. \end{array}$$
    In all these cases $\pi|_{Fix(\pi)}$ is not sorted. Then, π is non-sortable, contradiction;
  - If $qp \leq_s \pi$, then we have either
  $$\begin{array}{lll} p'q'qp \leq_s \pi, \text{ or} & p'qq'p \leq_s \pi, \text{ or} & p'qpq' \leq_s \pi, \text{ or} \\ qp'pq' \leq_s \pi, \text{ or} & qpp'q' \leq_s \pi, \text{ or} & q'p'qp \leq_s \pi, \text{ or} \\ q'qp'p \leq_s \pi, \text{ or} & q'qpp' \leq_s \pi, \text{ or} & qq'p'p \leq_s \pi, \text{ or} \\ qq'pp' \leq_s \pi, \text{ or} & qpq'p' \leq_s \pi. & \end{array}$$
    In all these cases $\pi|_{Fix(\pi)}$ is not sorted. Then, π is non-sortable, contradiction.

In this way, if $(j,i) \in E_\pi$ where $i \in N_{p,q}$ and $N_{p+1,q-1} \in \mathbb{B}^o_{p,q}$, then either $j \in N_{p,q}$ as well, or $j \in N^o_{p',q'}$ where $p' \neq p$, $q' \neq q$ and $N_{p',q'} \in \mathbb{B}^e_\pi$. If $j \in B^e_\pi$, then by Lemma 22 $T_{\pi,i} \cap (B^e_\pi \cap N^o_\pi) \neq \emptyset$. If $j \in N_{p+1,q-1}$ and $i \in N^e_\pi$, then $j \in N^e_\pi$ as well. If $j \in N_{p+1,q-1}$ and $i \in N^o_\pi$, then $j \in N^o_\pi$ as well. In this way, there are no integers $j \in N_{p'+1,q'-1}$ where $(j,i) \in E_\pi$, $p' \neq p$, $q' \neq q$, and $N_{p',q'} \in \mathbb{B}^o_\pi$. Then we have that $T_{\pi,i} \cap (B^e_\pi \cap N^o_\pi) \neq \emptyset$ and $T_{\pi,i} \cap N_{p+1,q-1} \neq \emptyset$. For any j where $(j,i) \in E_\pi$, since $\Gamma_{\pi,j}$ does not have a cycle, then $\Gamma_{\pi,j}$ does not have a cycle either.

□

The following two theorems state the main results of the paper.

Theorem 24. *Let π be a* ed*-sortable permutation. Let Φ be a strategy with* $\mathsf{dom}(\Phi) \in \mathbb{S}_\pi$ *and where for any $i \in \mathsf{dom}(\Phi)$ any operation ed_j with $j \in T_{\pi,i} \cap \mathsf{dom}(\Phi)$ is used earlier than ed_i in Φ. Then strategy Φ is applicable to π.*

Proof.
Let π be a ed-sortable permutation.

Consider a strategy $\Phi = \phi_m \circ \phi_{m-1} \circ \ldots \circ \phi_2 \circ \phi_1 = \mathsf{ed}_{i_m} \circ \mathsf{ed}_{i_{m-1}} \circ \ldots \circ \mathsf{ed}_{i_2} \circ \mathsf{ed}_{i_1}$ with the domain from $\mathbb{S}_\pi$, and where for any $i \in \mathsf{dom}(\Phi)$ any operation ed_j with $j \in T_{\pi,i} \cap \mathsf{dom}(\Phi)$ is used earlier than ed_i. We show that Φ can be applied to π. I.e., according to Lemma 19 we have to show that

(a) a subgraph $\Gamma_{\Phi,\pi}$ induced by $\mathsf{dom}(\Phi)$ from the dependency graph Γ_π is acyclic;
(b) for any $i \in \mathsf{dom}(\Phi)$ we have $T_{\pi,i} \subseteq \mathsf{dom}(\Phi)$;
(c) there is no integer $t \in \mathsf{dom}(\Phi)$ such that $t+1 \in \mathsf{dom}(\Phi)$.

By definition of $\mathbb{S}_\pi$ for any $1 \leq s \leq m$ either $i_s \in (B^e_\pi \cap N^o_\pi)$ or $i_s \in B^o_\pi$. If $i_s \in (B^e_\pi \cap N^o_\pi)$, then by Lemma 22 graph Γ_{π,i_s} has no cycles. Moreover, there is no integer $t \in T_{\pi,i_s}$ such that $t+1 \in T_{\pi,i_s}$. If $i_s \in B^o_\pi$, then by Lemma 23 graph Γ_{π,i_s} has no cycles, and there is no integer $t \in T_{\pi,i_s}$ such that $t+1 \in T_{\pi,i_s}$. I.e., conditions (a) and (c) are satisfied.

We show that for any $i \in \mathsf{dom}(\Phi)$ we have $T_{\pi,i} \subseteq \mathsf{dom}(\Phi)$.

- Let $i \in \mathsf{dom}(\Phi)$ and $i \in (B^e_\pi \cap N^o_\pi)$. Then by Lemma 22 $T_{\pi,i} \subseteq (B^e_\pi \cap N^o_\pi)$. By definition of $\mathbb{S}_\pi$ set $\mathsf{dom}(\Phi)$ contains $(B^e_\pi \cap N^o_\pi)$. I.e., for any $i \in (B^e_\pi \cap N^o_\pi)$ we have $T_{\pi,i} \subseteq \mathsf{dom}(\Phi)$;
- Let $i \in \mathsf{dom}(\Phi)$ and $i \in (B^o_\pi \cap N^o_\pi)$. I.e., let $i \in N^o_{p,q}$ for some $N_{p,q} \in \mathbb{B}^o_\pi$. Then, by Lemma 23 set $T_{\pi,i}$ contains some subset of $(B^e_\pi \cap N^o_\pi)$ and a subset from $N^o_{p,q}$. By Corollary 17 it follows that if there is an integer $j \in N^o_{p,q}$ such that $(j,i) \in E_\pi$, then $j < i$. Then by Lemma 11 for any $j \in T_{\pi,i} \cap N_{p+1,q-1}$ we have $j \leq i$. By the definition of the set $\mathbb{S}_\pi$ all $j \in N^o_{p,i}$ are in $\mathsf{dom}(\Phi)$. I.e., $T_{\pi,i} \cap N_{p+1,q-1} \subseteq \mathsf{dom}(\Phi)$. Set $T_{\pi,i} \setminus N_{p,q} \subseteq (B^e_\pi \cap N^o_\pi)$. As it was mentioned above, $(B^e_\pi \cap N^o_\pi) \subseteq \mathsf{dom}(\Phi)$. I.e., $T_{\pi,i} \setminus N_{p,q} \subseteq \mathsf{dom}(\Phi)$. In this way, for any $i \in \mathsf{dom}(\Phi)$ such that $i \in (B^o_\pi \cap N^o_\pi)$ we have $T_{\pi,i} \subseteq \mathsf{dom}(\Phi)$;
- Equivalently as above, we show that for any $i \in \mathsf{dom}(\Phi)$ such that $i \in (B^o_\pi \cap N^e_\pi)$ we have $T_{\pi,i} \subseteq \mathsf{dom}(\Phi)$.

In this way, condition (b) is satisfied. Then Φ is applicable to π.

□

Theorem 25. *Let π be a* ed*-sortable permutation. Then a strategy Φ sorts π if and only if $\mathsf{dom}(\Phi) \subseteq \mathbb{S}_\pi$ and for any $i \in \mathsf{dom}(\Phi)$ for any $j \in T_{\pi,i} \cap \mathsf{dom}(\Phi)$ operation ed_j is used earlier than ed_i in Φ.*

Proof.

Direct implication: Let Φ be a strategy that sorts permutation π. Then for any $i \in \mathsf{dom}(\Phi)$ for any $j \in T_{\pi,i} \cap \mathsf{dom}(\Phi)$ operation ed_j is used before ed_i in Φ. Assume that $\mathsf{dom}(\Phi) \notin \mathbb{S}_\pi$. That means that one or both of the conditions below are satisfied:

(i) $(B^e_\pi \cap N^o_\pi) \nsubseteq \mathsf{dom}(\Phi)$;
(ii) There is a block $N_{p,q} \in \mathbb{B}^o_\pi$ and integers $t_1, t_2 \in N_{p+1,q-1}$ from that block, where $t_2 \geq t_1$, such that $t_1, t_1-1, t_2, t_2+1 \notin \mathsf{dom}(\Phi)$;

Assume that property (i) holds for Φ. Let $N_{p,q} \in \mathbb{B}^e_\pi$. By Lemma 11 for any integer $i \in N^o_{p,q}$ we have either $i(i-1)(i+1) \leq_s \pi$, or for any integer $i \in N^o_{p,q}$ we have $(i-1)(i+1)i \leq_s \pi$. Without loose from generality we consider for the former case the smallest $t \in N^o_{p,q}$ such that $t \notin \mathsf{dom}(\Phi)$. And for the latter case we consider the biggest integer $t \in N^o_{p,q}$ such that $t \notin \mathsf{dom}(\Phi)$. In the former case, if $t-1 = p$, then we have $tp \leq_s \pi$. Since $p \in Fix(\pi)$, then $p \notin \mathsf{dom}(\Phi)$. Then Φ does not sort π. Then, $t-1 > p$, and $t-1 \in \mathsf{dom}(\Phi)$. Then $t-2 \notin \mathsf{dom}(\Phi)$, what contradicts to our choice of t. In the latter case, if $t+1=q$, then we have $qt \leq_s \pi$. Since $q \in Fix(\pi)$, then $q \notin \mathsf{dom}(\Phi)$. In this way, Φ does not sort π. Then we assume that $t+1<q$. Then $t+1 \in \mathsf{dom}(\Phi)$. Then $t+2 \notin \mathsf{dom}(\Phi)$, what contradicts to our choice of t. In this way, if Φ satisfies (i), it does not sort π.

Assume that Φ satisfies property (ii). Let $N_{p,q} \in \mathbb{B}^o_\pi$. Then without loose from generality we may take the smallest t_1 from $N_{p+1,q-1}$ such that $t_1, t_1-1 \notin \mathsf{dom}(\Phi)$ and the greatest $t_2 \in N_{p+1,q-1}$ such that $t_2, t_2+1 \notin \mathsf{dom}(\Phi)$. We show that $t_1 \in N^o_{p,q}$ and $t_2 \in N^e_{p,q}$. Indeed, if t_1 is the smallest from $N_{p+1,q-1}$ such that $t_1-1, t_1 \notin \mathsf{dom}(\Phi)$, then $t_1 - 1 \geq p$. If $t_1-1=p$, then $t_1 = p+1$. I.e., $t_1 \in N^o_{p,q}$. If $t_1-1>p$, then $p+1, p+3, \ldots, t_1-2 \in \mathsf{dom}(\Phi)$. I.e., $t_1 = p+2k_1+1$ for some $k_1 > 1$. This means that $t_1 \in N^o_{p,q}$. We show in the analogous way that $t_2 = q-2k_2-1$ for some $k_2 \geq 0$, i.e., $t_2 \in N^e_{p,q}$, since $N_{p,q} \in \mathbb{B}_\pi$.

Since both $t_1, t_2 \notin \mathsf{dom}(\Phi)$, $t_1 < t_2$ and Φ sorts π, we have $pt_1t_2q \leq_s \pi$. Since $t_1 \in N^o_{p,q}$, $pq \leq_s \pi$, and π is sortable, by Corollary 17 we have $(t_1-1)(t_1+1)t_1 \leq_s \pi$. Since $t_2 \in N^e_{p,q}$ and $(t_1-1)t_1t_2 \leq_s \pi$, there is such $t'_1 \in N^o_{p,q}$, $t_2 > t'_1 > t_1$ that $(t'_1-1)t_1(t'_1+1) \leq_s \pi$. Consider set $S_{\pi,t_1} \cap N_{p+1,q-1}$. Since $(t'_1-1)t_1(t'_1+1) \leq_s \pi$, we have $|S_{\pi,t_1} \cap N_{p+1,q-1}| > 1$. Since $t'_1 > t_1$, by Lemma 11 for any $t \in S_{\pi,t_1} \cap N_{p+1,q-1}$ we have $t \geq t_1$. Assume that for all integers $t \in S_{\pi,t_1} \cap N_{p+1,q-1}$ we have $t \leq t_2$. The biggest integer from $S_{\pi,t_1} \cap N_{p+1,q-1}$ we denote by t'. By Lemma 11 $t' \in N^o_{p,q}$, since $t_1 \in N^o_{p,q}$. Then $t' \neq t_2$, i.e., $t' < t_2$. Since Φ is applicable to π we have $t' \notin \mathsf{dom}(\Phi)$ because $t_1 \notin \mathsf{dom}(\Phi)$. Since Φ sorts π, we have $t_1t't_2 \leq_s \pi$. Moreover, $(t_1+1)t't_2 \leq_s \pi$ because $(t_1-1)(t_1+1)t_1 \leq_s \pi$. Since $t_1+1, t_2 \in N^e_{p,q}$, then there is $t'' \in N^o_{p,q}$, where $t_1+1 < t'' < t_2$ such that $(t''-1)t'(t''+1) \leq_s \pi$. By our assumption $t'' < t'$, and so, $t''t' \leq_s \pi$. By Corollary 17 we have $(t''-1)(t''+1)t'' \leq_s \pi$, i.e., we have $t't'' \leq_s \pi$, contradiction. In this way, there are integers from $S_{\pi,t_1} \cap N_{p+1,q-1}$ greater than t_2. In particular, $t'' > t_2$. Then $t''-1 \geq t_2$. Since $(t''-1)t' \leq_s \pi$ and $t't_2 \leq_s \pi$, we have $t'' \neq t_2$, i.e., $t''-1 > t_2$. Since $t''-1, t_2 \in N^e_{p,q}$, we have that $t_2(t''-1) \leq_s \pi$, contradiction. In this way, there are no such t_1, t_2 from $N_{p+1,q-1}$ with the property (ii) such that Φ sorts π.

Inverse implication: Let $\mathsf{dom}(\Phi) \in \mathbb{S}_\pi$ and for any $i \in \mathsf{dom}(\Phi)$ for any $j \in T_{\pi,i} \cap \mathsf{dom}(\Phi)$ operation ed_j is used before ed_i in Φ. According to Theorem 24 strategy Φ is applicable to π. Consider a block $N_{p,q} \in \mathbb{B}^e_\pi$. Let $\pi' = \Phi(\pi)$. Then $p(p+1)(p+2)\ldots(q-2)(q-1)q \leq \pi'$. Consider a block $N_{p',q'} \in \mathbb{B}^o_\pi$. By the definition of $\mathbb{S}_\pi$ we have $t \in N^e_{p'-2,q'}$ such that $t, t+1 \notin \mathsf{dom}(\Phi)$. Then we have substrings $p'(p'+1)(p'+2)\ldots(t-2)(t-1)t \leq \pi'$ and $(t+1)(t+2)\ldots(q'-2)(q'-1)q' \leq \pi'$. Since π is sortable, by Theorem 14 $\pi|_{F(\pi)}$ is sorted. Clearly $\pi|_{F(\pi)} = \pi'|_{F(\pi)}$. I.e., $F(\pi)$ is sorted in π' as well. Since $Fix(\pi) \subseteq F(\pi)$, then $Fix(\pi)$ is sorted in π'. Then one can observe that π' is sorted.

$\square$

Example 26. *Consider the following permutations:*

- $\pi_1 = 2\,4\,1\,3\,5\,7\,9\,6\,8\,10$ *has* $\mathbb{S}_{\pi_1} = [\{2,4,6,8\},\ \{2,4,6,9\},\ \{2,4,7,9\}]$. *All* *strategies with these domains sort π_1. There are no other strategies sorting π_1;*

- *Permutation $\pi_2 = 241365$ has three sets in $\mathbb{S}_{\pi_2}$: $\{2,4\}$, $\{2,5\}$ and $\{3,5\}$. As one can see, none of these strategies is applicable to π_2, and π_2 cannot be sorted.*

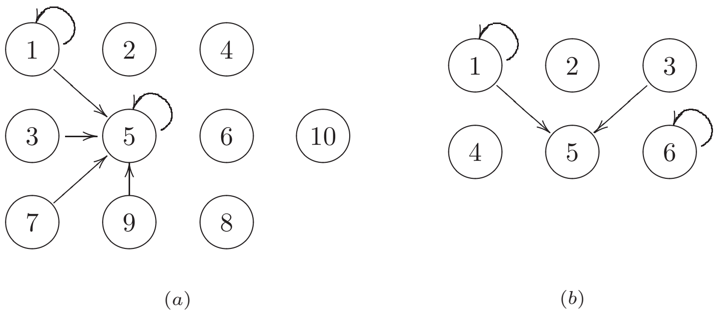


Fig. 6. Dependency graphs associated to (a) $\pi_1 = 2\,4\,1\,3\,5\,7\,9\,6\,8\,10$, and (b) $\pi_2 = 2\,4\,1\,3\,6\,5$.

10. Elementary assemblies

In this section we count the total number of different domains of ed sorting strategies for a permutation. We estimate the upper bound for total number of all sorting strategies for a permutation.

The following theorem characterizes permutations in terms of the number of ed-sorting strategies (modulo domain).

Theorem 27. *Consider a sortable permutation π having subset of blocks $\mathbb{B}^o = [N_{p_1,q_1}, N_{p_2,q_2}, \ldots, N_{p_l,q_l}]$. Then all sorting strategies for π have exactly $(q_1 - p_1 + 1)(q_2 - p_2 + 1)\ldots(q_l - p_l + 1)2^{-l}$ different domains.*

Proof.

Indeed, by Theorem 25 a permutation can be sorted by only strategies with their domains from $\mathbb{S}_\pi$. By definition of $\mathbb{S}_\pi$ all sets from $\mathbb{S}_\pi$ contain set $B^e_\pi \cap N^o_\pi$. For any block $N_{p',q'} \in \mathbb{B}^o_\pi$ there exist exactly $\frac{q'-p'+1}{2}$ different sets $N_{p',q'} \cap S$ for all $S \in \mathbb{S}_\pi$. In this way, if set $\mathbb{B}^o = [N_{p_1,q_1}, N_{p_2,q_2}, \ldots, N_{p_l,q_l}]$, then $|\mathbb{S}_\pi| = (q_1 - p_1 + 1)(q_2 - p_2 + 1)\ldots(q_l - p_l + 1)2^{-l}$.

$\square$

The following corollary follows by Theorems 24, 25 and 27.

Corollary 28. *A permutation π is ed-sortable if and only if there is a strategy Φ with $\mathsf{dom}(\Phi) \in \mathbb{S}_\pi$ applicable to it. Moreover, if π is sortable, then there are exactly*

$(q_1 - p_1 + 1)(q_2 - p_2 + 1)\dots(q_l - p_l + 1)2^{-l}$ different domains of sorting strategies for π.

Example 29. *We characterize permutation $\pi = 9\,11\,5\,7\,1\,3\,2\,4\,6\,8\,10$ in terms of the number of domains of its sorting strategies. We have $Fix(\pi) = 1, 4, 8, 11$, $\mathbb{B}^o_\pi = \{N_{1,4}, N_{8,11}\}$, and $\mathbb{B}^e = \{N_{4,8}\}$. A strategy $\mathsf{ed}_9 \circ \mathsf{ed}_7 \circ \mathsf{ed}_5 \circ \mathsf{ed}_2(\pi) = 11\,1\,2\,3\,4\,5\,6\,7\,8\,9\,10$ sorts π. Then there are sorting strategies with $\frac{4*4}{4} = 4$ different domains. Indeed, strategies*

$$\begin{aligned}
\mathsf{ed}_9 \circ \mathsf{ed}_7 \circ \mathsf{ed}_5 \circ \mathsf{ed}_2(\pi) &= 11\,1\,2\,3\,4\,5\,6\,7\,8\,9\,10\\
\mathsf{ed}_9 \circ \mathsf{ed}_7 \circ \mathsf{ed}_5 \circ \mathsf{ed}_3(\pi) &= 11\,1\,2\,3\,4\,5\,6\,7\,8\,9\,10\\
\mathsf{ed}_{10} \circ \mathsf{ed}_7 \circ \mathsf{ed}_5 \circ \mathsf{ed}_2(\pi) &= 9\,10\,11\,1\,2\,3\,4\,5\,6\,7\,8\\
\mathsf{ed}_{10} \circ \mathsf{ed}_7 \circ \mathsf{ed}_5 \circ \mathsf{ed}_3(\pi) &= 9\,10\,11\,1\,2\,3\,4\,5\,6\,7\,8
\end{aligned}$$

sort π. There are no sorting strategies for π with other domains.

In the following lemma we estimate the maximal number of different sorting strategies with the same domain.

Lemma 30. *Let $D \subseteq \Sigma_n$. Let Π_D be a set of all permutations over Σ_n on which we can apply a strategy with the domain D. Let $\pi \in \Pi_D$ be a permutation where set D induces an isolated subgraph in the dependency graph Γ_π. Then*

(i) π has exactly $|D|!$ different strategies with the domain D;
(ii) π has the maximal number of strategies with the domain D among all permutations from Π_D.

Proof.

(i) Since there are no edges in Γ_π between integers from D, then by Lemma 19 any strategy Φ_D with $\mathsf{dom}(\Phi_D) = D$ is applicable to π. I.e., operations from D may be applied in any order to π. In this way, there are $|D|!$ different strategies with the domain D which are applicable to π.

(ii) Let π' be a permutation from Π. If D induces a cyclic subgraph in the dependency graph $\Gamma_{\pi'}$, then by Lemma 19 there is no strategy Φ_D with $\mathsf{dom}(\Phi_D) = D$ which is applicable to π'. Then we assume that D induces some non-trivial forest F (an acyclic subgraph with at least one edge) in $\Gamma_{\pi'}$. We denote by N_F some number of orderings ρ of integers from D (i.e., ρ is a permutation over D) which we count as follows. For all pairs $p, q \in D$ where there is a path from q to p in the forest F we count the total number N_F of orderings ρ with a subsequence $pq \leq_s \rho$. In other words, N_F is the number of all those strategies for π', where condition (i) of Lemma 5 does not hold. Clearly, N_F is the number of all strategies with the domain D which cannot be applied to π'. Since F contains at least one edge, then

$N_F \geq 1$. Then $|D|! - N_F$ is the number of all applicable strategies to π', and $|D|! - N_F < |D|!$. In this way, there are no permutations from Π having the number of strategies with the domain D greater than $|D|!$.

□

The following theorem estimates the upper bound for the number of ed-sorting strategies.

Theorem 31. *Let π be a permutation over Σ_n with $|\mathbb{B}^o_\pi| = k \geq 0$ blocks $N_{p_i,q_i} \in \mathbb{B}^o_\pi$ for all $0 \leq i \leq k$, and with $|\mathbb{B}^e_\pi| = l \geq 0$ blocks $N_{r_i,s_i} \in \mathbb{B}^e_\pi$ for all $0 \leq i \leq l$. Then π may have at most*

$$\left(\sum_{0 \leq i \leq l} \frac{s_i - r_i}{2} + \sum_{0 \leq i \leq k} \frac{q_i - p_i - 1}{2}\right)! \prod_{0 \leq i \leq k} \frac{q_i - p_i}{2}$$

different sorting strategies.

Proof.
Consider a permutation π over Σ_n with $|\mathbb{B}^o_\pi| = k \geq 0$ blocks $N_{p_i,q_i} \in \mathbb{B}^o_\pi$, $0 \leq i \leq k$, and with $|\mathbb{B}^e_\pi| = l \geq 0$ blocks $N_{r_i,s_i} \in \mathbb{B}^e_\pi$, $0 \leq i \leq l$. Assume that the dependency graph Γ_π has no edges (p,q), where either $q \in (B^e_\pi \cap N^o_\pi)$ or $q \in B^o_\pi$. From the definition of $\mathbb{S}_\pi$ it follows that any strategy with the domain from $\mathbb{S}_\pi$ has exactly $N = \sum_{0 \leq i \leq l} \frac{s_i - r_i}{2} + \sum_{0 \leq i \leq k} \frac{q_i - p_i - 1}{2}$ operations ed. As we assumed, all ed operations from a strategy with the domain from $\mathbb{S}_\pi$ can be applied in any order. Then by Lemma 30 it follows that there are $N!$ strategies with the same domain $S \in \mathbb{S}_\pi$ which are applicable to π. Also, by Lemma 30 there are no other permutations over Σ_n having more than $N!$ different sorting strategies with the same domain.

All sorting strategies are with domains from $\mathbb{S}_\pi$, and all strategies with their domains from $\mathbb{S}_\pi$ are sorting strategies by Theorem 25. By Theorem 27 there are exactly $\prod_{0 \leq i \leq k} \frac{q_i - p_i}{2}$ different domains of sorting strategies.

Then permutation π has $(N!)\left(\prod_{0 \leq i \leq k} \frac{q_i - p_i}{2}\right)$ different sorting strategies.

□

Note, that in order to compute the exact number of domains and strategies for a permutation π over Σ_n, it is necessary to analyze the dependency graph Γ_π and to detect what sets of operations can be applied to π. Also, one have to compute the total number of all strategies for π with the domain S for each applicable set S to π. This should be done by using the induced graph by S in Γ_π. The problem of computation of the number of all strategies for π and their domains remains open.

11. Summary

In this paper we gave a general form for successful assembly strategies by elementary intramolecular operations. We used a formalization based on permutations. We introduced the notions of fixed integers and blocks. Knowing the general form of elementary assembly strategies one can decide whether a given gene pattern may be assembled by elementary operations by applying any of these strategies. We refer to [26] for another decision method. We also estimated the maximal number of all successful strategies for a gene pattern.

The result from this paper may be useful for solving the following open problems:

- counting the exact number of all strategies for a permutation π;
- counting the number of ed-sortable permutations of fixed length n;
- searching in polynomial time optimal parallel ed-sorting strategies for a permutation (note, that for the moment the analogous problem for general intramolecular operations may be solved in time roughly estimated as $O(n^n)$, see [1, 2]);
- effective detection of gene patterns with inverted MDS's which can be assembled to the genes by elementary operations.

Acknowledgments. The work of Vladimir Rogojin is supported by Academy of Finland, project 203667 and by Science and Technology Center in Ukraine, project 4032. Vladimir Rogojin is on leave of absence from Institute of Mathematics and Computer Science of Academy of Sciences of Moldova, Chisinau MD-2028 Moldova. The author is very grateful to Ion Petre and to the reviewers for the valuable comments.

References

[1] Alhazov, A., Li, C., and Petre, I.. Computing the graph-based parallel of gene assembly, *Theoretical Computer Science*, (2008), to appear.

[2] Alhazov, A., Petre, I., and Rogojin, V.. Computing the parallel complexity of signed graphs: an improved algorithm. *Theoretical Computer Science*, (2008), to appear.

[3] Angeleska, A., Jonoska, N., Saito, M., and Landweber, L.F., RNA-Template Guided DNA Assembly. *Journal of Theoretical Biology* **248**, Elsevier (2007), 706–720.

[4] Berman, P., and Hannenhalli, S., Fast sorting by reversals. *Combinatorial Pattern Matching, Lecture Notes in Comput. Sci.*, **1072**, (1996), 168 – 185.

[5] Brijder, R., Langille, M. and Petre, I., A String-based Model for Simple Gene Assembly. In: E. Csuhaj-Varju, Z. Esik (Eds.) *Proceedings of FCT 2007*, Lecture Notes in Computer Science, **4639**, Springer-Verlag Berlin Heidleberg, (2007), 161 - 172.

[6] Caprara, A., Sorting by reversals is difficult. In S. Istrail, P. Pevzner and M. Waterman (eds.) *Proceedings of the 1st Annual International Conference on Computational Molecular Biology*, (1997), 75 – 83.

[7] Chang, W.J., Bryson, P.D., Liang, H., Shin, M.K., Landweber, L., The evolutionary origin of a complex scrambled gene. *Proceedings of the National Academy of Sciences of the US*, **102**(42), (2005), 15149–15154.

[8] Chang, W.J., Kuo, S., Landweber, L., A new scrambled gene in the ciliate *Uroleptus*. *Gene*, **368**, (2006), 72 – 77.

[9] Ehrenfeucht, A., Harju, T., Petre, I., Prescott, D. M., and Rozenberg, G., *Computation in Living Cells: Gene Assembly in Ciliates*, Springer (2003).

[10] Ehrenfeucht, A., Petre, I., Prescott, D. M., and Rozenberg, G., Universal and simple operations for gene assembly in ciliates. In: V. Mitrana and C. Martin-Vide (eds.) *Words, Sequences, Languages: Where Computer Science, Biology and Linguistics Meet*, Kluwer Academic, Dortrecht, (2001), 329–342.

[11] Ehrenfeucht, A., Prescott, D. M., and Rozenberg, G., Computational aspects of gene (un)scrambling in ciliates. In: L. F. Landweber, E. Winfree (eds.) *Evolution as Computation*, Springer, Berlin, Heidelberg, New York, (2001), 216–256.

[12] Hannenhalli, S., and Pevzner, P. A., Transforming cabbage into turnip (Polynomial algorithm for sorting signed permutations by reversals). In: *Proceedings of the 27th Annual ACM Symposium on Theory of Computing*, (1995), 178 – 189.

[13] Harju, T., Petre, I., Rogojin, V. and Rozenberg, G., Patterns of Simple Gene Assembly in Ciliates, *Discrete Applied Mathematics*, **156**(14), Elsevier, (2008), 2581-2597.

[14] Harju, T., Petre, I., and Rozenberg, G., Modelling simple operations for gene assembly. In: J.Chen, N.Jonoska, G.Rozenberg (Eds.) *Nanotechnology: Science and Computation*, (2006), 361–376.

[15] Jahn, C. L., and Klobutcher, L. A., Genome remodeling in ciliated protozoa. *Ann. Rev. Microbiol.* **56**, (2000), 489–520.

[16] Kaplan, H., Shamir, R., and Tarjan, R. E., A faster and simpler algorithm for sorting signed permutations by reversals. *SIAM J. Comput.* **29**, (1999), 880 – 892.

[17] Landweber, L. F., and Kari, L., The evolution of cellular computing: Nature's solution to a computational problem. In: *Proceedings of the 4th DIMACS Meeting on DNA-Based Computers*, Philadelphia, PA, (1998), 3 – 15.

[18] Landweber, L. F., and Kari, L., Universal molecular computation in ciliates. In: L. F. Landweber and E. Winfree (eds.) *Evolution as Computation*, Springer, Berlin Heidelberg New York, (2002).

[19] Langille, M., Petre, I., Simple gene assembly is deterministic. *Fundamenta Informaticae* **72**, IOS Press, (2006), 1–12.

[20] Langille, M., Petre, I., Sequential vs. Parallel Complexity in Simple Gene Assembly.

Theoretical Computer Science, **395**(1), Elsevier B.V., (2008), 24–30.
[21] Langille, M., Petre, I. and Rogojin, V., Three models for gene assembly in ciliates: a comparison, In: *Proceedings of BIONETICS 2008*, (2008).
[22] Langille, M., Petre, I. and Rogojin, V., Three models for gene assembly in ciliates: a comparison, *Technical report 878*, TUCS, 2008.
[23] Nowacki, M., Vijayan, V., Zhou, Y., Schotanus, K., Doak, T.G., Landweber, L.F., RNA-mediated epigenetic programming of a genome-rearrangement pathway. *Nature* **451**, doi:10.1038/nature06452, (2008), 153–158.
[24] Prescott, D. M., The DNA of ciliated protozoa. *Microbiol. Rev.*, **58**(2), (1994), 233 – 267.
[25] Prescott, D.M., DNA manipulations in ciliates. In: W.Brauer, H.Ehrig, J.Karhumäki, A.Salomaa (Eds.) *Formal and Natural Computing*, LNCS 2300, Springer, (2002), 394 - 417.
[26] Petre, I., and Rogojin, V., Decision Problem for Shuffled Genes, *Information and Computation*, **206**(11), Elsevier, (2008), 1346-1352.
[27] Prescott, D. M., Ehrenfeucht, A., and Rozenberg, G., Molecular operations for DNA processing in hypotrichous ciliates. *Europ. J. Protistology*, **37**, (2001), 241–260.
[28] Swanton, M.T., Heumann, J.M., Prescott, D.M., Gene-sized DNA molecules of the macronuclei in three species of hypotrichs: size distribution and absence of nicks. *Chromosoma*, **77**, (1980), 217–227.
[29] Yao, M.C., Fuller, P., Xi, X., Programmed DNA Deletion As an RNA-Guided System of Genome Defense, *Science*, **300**, (2003), 1581 - 1584.
[30] West, D. B., *Introduction to Graph Theory*, Prentice Hall, Upper Saddle River, NJ (1996).