

Rearrangement Problems on PermutationsReview

Let α and β be two permutations of S_n . Then, the composition of α and β , denoted by $\alpha \circ \beta$, maps i into $(\alpha \circ \beta)(i)$.

$(\alpha \circ \beta)(i) = \alpha(\beta(i))$. For example, $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$, $\beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$,

(or $\alpha = (3142)$, $\beta = (1342)$ in short, then

$$\alpha \circ \beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix} = (3421).$$

(*) If $\alpha \circ \beta = \gamma$, then $\beta = \alpha^{-1} \circ \gamma$. (Notice that S_n is a

symmetric group with "operation", composition.) (*) $(\alpha \circ \beta)^{-1} = \beta^{-1} \circ \alpha^{-1}$.

Problem

Given any two permutations π and σ in G (say S_n) and a set S of generators of G , find a minimum number permutations in S ,

$p_1, p_2, \dots, p_k$ such that $\pi \circ p_1 \circ p_2 \circ \dots \circ p_k = \sigma$, equivalently,

$$\alpha^{-1} \pi = \underbrace{p_k^{-1} \circ p_{k-1}^{-1} \circ \dots \circ p_1^{-1}}_{\text{Minimum length factorization.}}$$

(*) Finding a minimum length factorization of permutations is an NP-hard problem. [Even and Goldreich, J. Algorithm 2 (1981), 311-313.]

Definition (Invariant distance)

Given a set of generators S of a permutation group, A distance between two permutations π and σ (w.r.t. S) can be defined as the minimum number of generators $p_1, p_2, \dots, p_k$ (in S) such that $\pi \circ p_1 \circ p_2 \circ \dots \circ p_k = \sigma$, denoted by $d_S(\pi, \sigma) = k$.

(*) By the definition of "metric", d_S is indeed a metric.

Definition (Left-invariant)

A distance d on G is left-invariant if for all π, σ, τ in G , $d(\pi, \sigma) = d(\tau \circ \pi, \tau \circ \sigma)$.

$$(**) \quad \pi \circ p_1 \circ p_2 \circ \dots \circ p_k = \sigma \Rightarrow (\sigma^{-1} \circ \pi) \circ p_1 \circ p_2 \circ \dots \circ p_k = i \text{ (identity)}.$$

$$d(\pi, \sigma) = d(\sigma^{-1} \circ \pi, i)$$

$$d(\pi, i) = d(\pi^{-1}, i) \swarrow \searrow = d(i, \pi)$$

Definition (Sorting Sequence)

Given a permutation π , a sequence of rearrangements $p_1, p_2, \dots, p_k$ is called a sorting sequence for π if $\pi \circ p_1 \circ p_2 \circ \dots \circ p_k = i$.

How to find the distance?

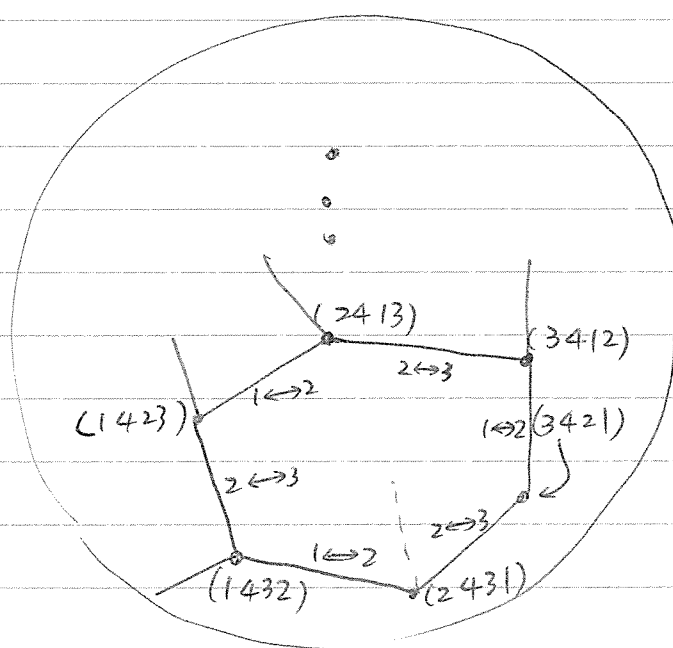
Definition (Cayley graph)

Given a set S of generators of a permutation group G the Cayley graph associated with (S, G) is the graph whose vertices are the elements of G and two vertices are adjacent if the corresponding permutations can be transformed into one another using an element of S .

Example

$G = S_4$, $S = \{\text{exchange of adjacent elements}\}$
 $1 \leftrightarrow 2, 2 \leftrightarrow 3, 3 \leftrightarrow 4$

$G \stackrel{\text{def}}{=} \text{Cayley}(S_4, S)$



(Bubble sort)

3-regular graph of order 24.

(*) The distance of two permutations is equal to the shortest distance in $G \stackrel{\text{def}}{=} \text{Cayley}(S_4, S)$.

Reversals

permutation $\pi = (\pi_1, \pi_2, \dots, \pi_{i-1}, \pi_i, \pi_{i+1}, \dots, \pi_j, \pi_{j+1}, \dots, \pi_n)$

↓

$\rightarrow (\pi_1, \pi_2, \dots, \pi_{i-1}, \pi_j, \pi_{j-1}, \dots, \pi_i, \pi_{j+1})$

Let $\rho = (1, 2, \dots, i-1, j, j-1, \dots, i+1, i, j+1, j+2, \dots, n)$.

Then $\pi \circ \rho =$

Signed permutation

$\rho = (1, 2, \dots, i-1, -j, -j+1, \dots, -i-1, -i, j+1, \dots, n)$

$\pi \circ \rho = (\pi_1, \pi_2, \dots, \pi_{i-1}, -\pi_j, -\pi_{j-1}, \dots, -\pi_{i+1}, -\pi_i, \pi_{j+1}, \dots, \pi_n)$

Problem Use only reversals to find the minimum length factorization.

$\pi \rightarrow \rightarrow \rightarrow \rightarrow \dots \rightarrow i$

$(1, 2, 3, \dots, n)$

Observation $\pi_1 \neq 1$ one reversal is needed!

$\pi_n \neq n$ "

$\pi_{j+1} > \pi_j$ "

⋮

(*) Similarity between permutations

(*) If a group of genes appears consecutively in several species, then they must have been present in the same order in the ancestral species, and were not separated during evolution.

Definition (Linear extension)

The linear extension of (signed or unsigned) permutation π of $\{1, 2, \dots, n\}$ is the permutation of $\{0, 1, 2, \dots, n, n+1\}$ defined by $\pi^l = (0 \pi_1 \pi_2 \dots \pi_n n+1)$. (If signed permutations are considered, then we let "0" be of positive sign.)

Example $(0 \ 4 \ 8 \ 9 \ 7 \ 6 \ 5 \ 1 \ 3 \ 2 \ 10)$ is a linear extension of $(4 \ 8 \ 9 \ 7 \ 6 \ 5 \ 1 \ 3 \ 2)$.

Definition (Break points)

Let π^l be the linear extension of π . A point of π is an ordered pair (π_i^l, π_{i+1}^l) (denoted by $\pi_i^l \circ \pi_{i+1}^l$ in short), $0 \leq i \leq n$.

- If $\pi_{i+1}^l = \pi_i^l + 1$, ^{then} it is called an adjacency (point);
- If $\pi_{i+1}^l = \pi_i^l - 1$, ^{then} it is called a reverse adjacency (point);
- If it is not an adjacency, ^{then} it is called a breakpoint;
- If it is neither an adjacency nor a reverse adjacency, then it is called a strong breakpoint.

e.g.

0 • 4 • 8 • 9 • 7 • 6 • 5 • 1 • 3 • 2 • 10

↑
adjacency

↑
reverse adjacency
(break points)

All the others are strong break points.

$p(\pi)$: The # of points

$bp(\pi)$: The # of breakpoints

$sb(\pi)$: The # of strong breakpoints

(*) Breakpoints yield a first example of an evolutionary distance between genomes: the "breakpoint distance"

between two permutations. (The difference of adjacencies)

$$\downarrow \pi, \sigma$$

$$bp(\pi, \sigma) =_{\text{def}} bp(\sigma^{-1} \circ \pi)$$

(breakpoint distance)

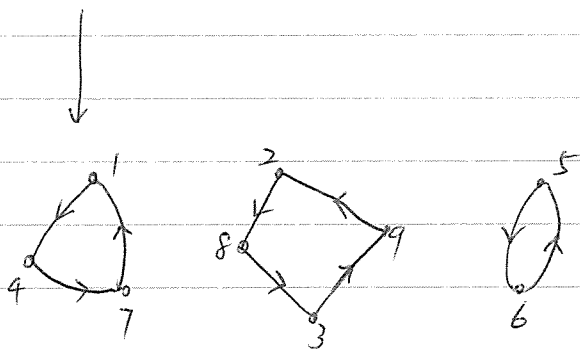
Cycle graph of a permutation

It is known in Algebra that each permutation can be written as a composition of disjoint cycles.

e.g. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 8 & 9 & 7 & 6 & 5 & 1 & 3 & 2 \end{pmatrix}$

$$= (147)(2839)(56)$$

← Cycle decomposition graph



The graph (digraph) which represents the cycles of a permutation is called a cycle decomposition graph of the permutation.

(i) We consider (i) a cycle (loop), $\bigcirc_i$.

(ii) If the cycle decomposition graph of a permutation is connected (strongly);

then the permutation is called a circular permutation,

denoted by $[x_1, x_2, \dots, x_n]$. So, a circular permutation

has n equivalent forms $(x_1, x_2, \dots, x_n), (x_2, x_3, \dots, x_n, x_1), \dots,$

$(x_n, x_1, \dots, x_{n-1})$.

Cycle graph of π

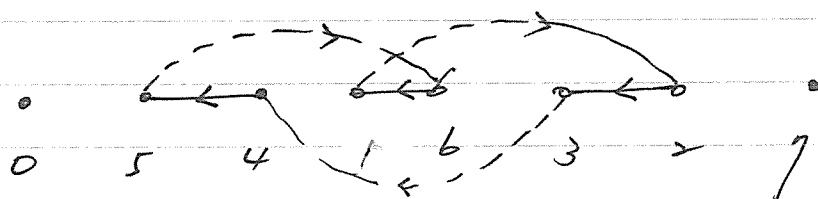
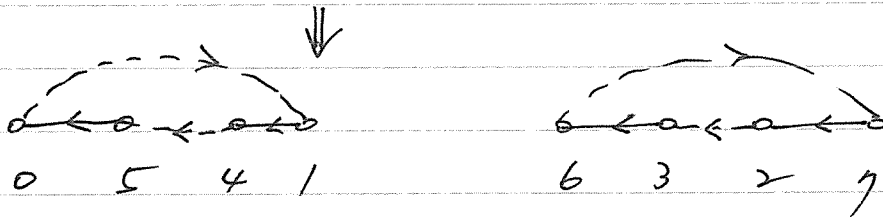
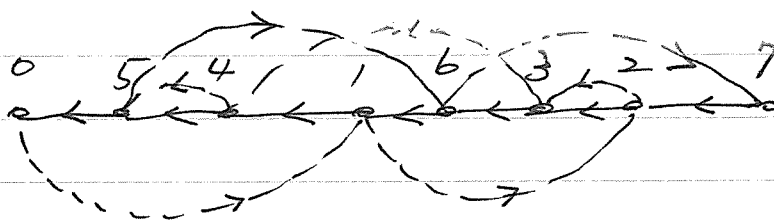
Definition (Cycle graph)

The cycle graph of a permutation π of $\{1, 2, \dots, n\}$ is the directed graph $G(\pi)$ with vertex set $\{0, 1, 2, \dots, n, n+1\}$ and whose arcs are :

- (1) black arcs (π_i^l, π_{i-1}^l) for $1 \leq i \leq n+1$, and
(reality)
- (2) gray arcs $(i, i+1)$ for $0 \leq i \leq n$.
(desired)

Example $\pi = (5 \ 4 \ 1 \ 6 \ 3 \ 2)$, π^l : linear extension of π

$G(\pi)$:



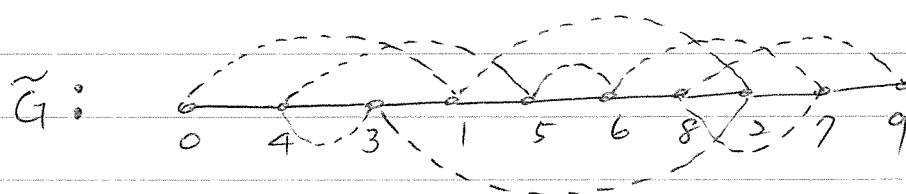
alternating cycles $\leftarrow$ The # of $c(G(\pi))$.

Definition (Breakpoint graph)

The breakpoint graph of a permutation π is the undirected graph $BG(\pi)$, whose vertex set is the vertex set of the cycle graph of π , and whose edges are the arcs of the cycle graph of π taken without their orientation.

e.g. circular permutation (4 3 1 5 6 8 2 7)

extension 0 4 3 1 5 6 8 2 7 9



Incident edges "..."

$\tilde{G}$ is obtained by adding incident edges 01, 12, 23, ..., 89.

$\tilde{G}$ is an eulerian graph.

$\tilde{G}$ can be decomposed into alternating cycles: (circuits)

(0, 4, 5, 1), (5, 6), (4, 3), (6, 8, 9, 7), (3, 1, 2, 8, 7, 2)

Let $c^*(BG(\pi))$ denote the number of cycles in a maximal alternating cycle decomposition of $BG(\pi)$.

The goal of finding minimum length factorization can be accomplished by using both the cycle graph and breakpoint graph of π . Though, not able to find an explicit solution, we are able to estimate the upper bounds, hopefully a tight bound.