The solution space of sorting by block interchanges and sorting by reversals

Research proposal, 2018 Fall

August 18, 2018

Problem description

A *block interchange* swaps two, not necessarily consecutive blocks in a permutation of the numbers between 1 and n. For example, if we swap $2, 3, 7$ and $4, 1$ in the permutation

$$5, 2, 3, 7, 6, 4, 1$$

we get the permutation

$$5, 4, 1, 6, 2, 3, 7.$$

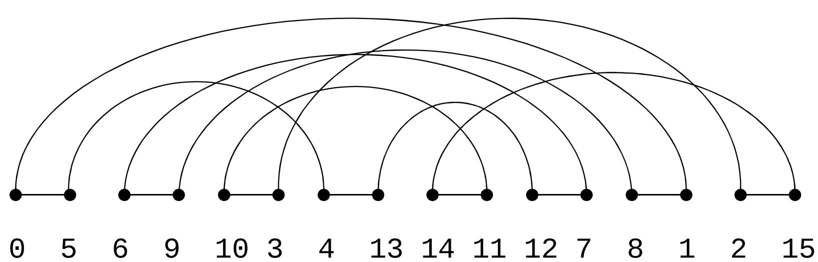
The *Sorting by block interchanges* problem asks for the minimum number of block interchange operations to transform a permutation into the identity permutation. To obtain this number, we have to consider the following graph, called the *graph of desire and reality*. This graph is a drawn multigraph, that is, the drawing is important and also, two vertices might be connected with multiple (at most 2) edges. The construction is the following: replace any number k with $2k-1, 2k$ in the permutation π. Furthermore, frame these numbers between 0 and $2n+1$ where n is the length of π. For example, if π is

$$3, 5, 2, 7, 6, 4, 1$$

then we get

$$0, 5, 6, 9, 10, 3, 4, 13, 14, 11, 12, 7, 8, 1, 2, 15.$$

These numbers will be the vertices of the graph. These vertices are drawn along a line. We connect every second vertex with a straight line, that is, 0 is connected to 5, 6 to 9, etc. Also every even number is connected to the next odd number with an arc. That is, 0 with 1, 2 with 3, etc. In our example, we will get:



It is known that the minimum number of block interchanges necessary to sort the permutation is

$$\frac{n+1-c(\pi)}{2}$$

where n is the length of the permutation and $c(\pi)$ is the number of cycles in the graph of desire and reality. In our example, $n = 7$, and the graph of desire and reality contains 2 cycles. Therefore, 3 block interchange operqations are sufficient to transform the permutation into the identity. Indeed, first swap $3, 5, 2, 7$ and 1 to get

$$1, 6, 4, 3, 5, 2, 7.$$

Then swap 2 and 6 to get

$$1, 2, 4, 3, 5, 6, 7.$$

Then finally, swapping 4 and 3 sorts the permutation. This series of block interchange operartions is called a *sorting scenario*. There might be multiple solutions, that is, there might be many sorting scenarios for a single permutation. The set of sorting scenarios are called the *solution space*. We have the following conjecture:

Conjecture 1. *Let π be an arbitrary permutation. Consider the following graph $G(\pi) = (V, E)$. V is the solution space of π. For any v and $w \in V$, $(v, w) \in E$ if and only if v and w (as sorting scenarios) differ in two consecutive steps. The conjecture is that $G(\pi)$ is connected.*

With other words: the solution space can be explored by small perturbations on the current sorting scenarios. The aim of the proposed research is to prove or disprove this conjecture. If the problem turns to be too easy, we will also work on a similar conjecture on the solution space of sorting by reversals.

Assignment for the first week

Read Chapter 10 from this electronic note:
`https://users.renyi.hu/~miklosi/AlgorithmsOfBioinformatics.pdf`.

This chapter gives the detailed description of the theorem of sorting by block interchanges. Please, also solve the following exercises:

1. The block interchange distance is the minimum number of necessary block interchange operations to sort a permutation. What is the largest possible block interchange distance of a permutation of size n?

2. Prove that the solution space might grow faster than any exponential function of the length of the permutation.

3. Let s be a sorting scenario of π such that 1 is moved to the begining of the permutation in the second block interchange in s. Prove that there is a sorting scenario s' such that 1 is moved to the beginning of the permutation in the first block interchange in s', and s and s' differ only in the first two block interchanges.

4. Let s be a sorting scenario of π. Prove that there is a finite series of sorting scenarios $s = s_0, s_1, \ldots s_k$ such that for each $i = 0, 1, \ldots k-1$, s_i and s_{i+1} differ only in two consecutive steps, and in s_k, 1 is moved to the first position of the permutation in the first block interchange.