

AAGB (TD4)

Genome Rearrangements

1 Introduction

Let's consider the permutation $\pi = (\pi_1 \pi_2 \dots \pi_n)$

1. What is a reversal, $\rho(i, j)$, of an interval $[i, j]$? In genome rearrangements, what does it represent?
2. What is the *reversal distance problem*? What is the output?
3. Give a biologically more relevant version of this problem.
4. Give the pseudo code of a simple algorithm performing a reversal sort based on the increase of the length of the ordered prefix. What is the maximal number of steps to sort a permutation of length n ?
5. Apply this algorithm to the permutation $\pi = (5\ 4\ 1\ 2\ 3)$. Do you think it is optimal?

2 Break point Reversal Sort

1. Let's define $b(\pi)$ the number of break points in a permutation π . Each time a *good reversal* is done, what is the behaviour of $b(\pi)$? Deduce a lower bound for the reversal sort problem.
2. Give the pseudo code of a reversal sorting algorithm that use the $b(\pi)$ information at each step.
3. Apply this sorting algorithm to the following permutations :
→ (5 4 1 2 3)
→ (8 2 7 6 5 1 4 3)
→ (3 5 8 6 4 7 9 2 1 10 11)
4. How many break points the algorithm described in 2. eliminates at each step?
5. With unsigned data, the problem of finding minimal reversal distances is NP-complete (*Caprara, 1999*). The best approximation ratio reached is 1.375 (*Hannenhalli, 1998*). Let's define $A(\pi)$, the solution of your algorithm, and $OPT(\pi)$, the optimal solution. The approximation ratio is defined as $A(\pi)/OPT(\pi)$. How many times worse is the algorithm given in 2. as compared to the optimal solution?

3 Introduction to the reversal sort with signed data

Let's consider the following permutation $\pi : (1\ -2\ 4\ 5\ -3\ 6)$

1. Give the break point graph representation for this permutation. How many breakpoints ($b(\pi)$) and cycles ($c(\pi)$) are there?

2. Each step of the algorithm tries to reduce the number of cycles by one. How many cycles are there for the Identity permutation? Bafna and Pevzner demonstrated in 1996 that $d(\pi) \geq n + 1 - c(\pi)$. How could you justify it?
3. Apply the reversal sort with signed data algorithm to this permutation.