

CISC-471 WINTER 2016

HOMEWORK 4

Please work on these problems and be prepared to share your solutions with classmates in class on Monday Feb. 8. Assignments will **not** be collected for grading.

PROGRAMMING

Write a program in the language of your choosing (I recommend Python) and verify that it works on the sample data (using the on-line Rosalind platform). For each problem be prepared to tell us why you think your algorithm is correct (whether your program worked on the sample data or not). Also provide an estimate of the time and space complexity of your algorithm.

Greedy Motif Finding:

<http://rosalind.info/problems/3d/>

A greedy heuristic for finding motifs, GREEDYMOTIFSEARCH, is described in section 5.5 of the text. ~~Implement this algorithm and try it on the Rosalind test data. Also see if your implementation is any help in solving the following problem.~~

The page linked to above is no longer available so use the following for DNA strings.

```
AAATTGACGCAT
GACGACCACGTT
CGTCAGCGCCTG
GCTGAGCACCGG
AGTACGGGACAG
```

and find the best 3-mer motif.

Problem 5.18: Design an input for GREEDYMOTIFSEARCH algorithm that causes the algorithm to output an incorrect result. That is, create a sample that has a strong pattern that is missed because of the greedy nature of the algorithm. If `optimalScore` is the score of the strongest motif in the sample and `greedyScore` is the score returned by GREEDYMOTIFSEARCH, how large can `optimalScore/greedyScore` be?

PROBLEMS

These questions come from *An Introduction to Bioinformatics Algorithms* by Neil C. Jones and Pavel A. Pevzner.

Problem 5.4: Perform the BREAKPOINTREVERSALSORT algorithm with $\pi = 3$ 4 6 5 8 1 7 2 and show all intermediate permutations (break ties arbitrarily). Since

BREAKPOINTREVERSALSORT is an approximation algorithm, there may be a sequence of reversals that is shorter than the one found by BREAKPOINTREVERSALSORT. Could you find such a sequence of reversals? Do you know if it is the shortest possible sequence of reversals?

Problem 5.5: Find a permutation with no decreasing strips for which there exists a reversal that reduces the number of breakpoints.