

CSCI 451/551 Fall 2019 Homework 4

Due: Tue, Nov 5, 11:59pm

Submission instructions:

- Step 1: Combine your answers into a single PDF and submit that using **gradescope** (**Group submissions: create a group and just submit once**)
- Step 2 (only for programming assignments): One person from each group should also upload all source code files to **ecat**. (Group submissions: please make sure all group members are listed)

Problem 1 (10 pts) Exercise 2 (textbook section 6.9, pg 152)

Problem 2 (5 pts) Exercise 3 (textbook section 6.9, pg 152)

There are a few online CLUSTALW servers, one is: <https://www.genome.jp/tools-bin/clustalw>

Problem 3 (10 pts) Exercise 4 (textbook section 6.9, pg 152)

Problem 4 (20 pts) Programming project: *Center Sequence Finder*

Implement the first two steps of the *Center Star Algorithm*. Your program should take as input a FASTA file specifying a set of sequences, as well as two cost function parameters, alpha and beta, where $\delta(x,y) = \alpha > 0$ and $\delta(x,-) = \beta > 0$ (Note: $\delta(x,x) = 0$).

The program should output the center sequence, based on the specified distance function.

Demonstrate the algorithm works on several examples, including the set of sequences from Problem 1.

Bonus (5 pts) Implement the entire Center Star Algorithm so that your program also finds and prints out the multiple sequence alignment.