

CSCI 451/551 Fall 2023 Homework 2

Due: Sept 21, at the start of class

(Group submissions: please make sure all group members are listed)

Problem 1 (5 pts) Exercise 2 (textbook section 2.7, pg 54)

2. Can you propose a dynamic programming solution to solve the longest common subsequence problem?

Problem 2 (10 pts) Exercise 3 (textbook section 2.7, pg 54)

3. Given two sequences S and T (not necessarily the same length), let G , L , and H be the scores of an optimal global alignment, an optimal local alignment, and an optimal global alignment without counting the beginning space of S and end space of T , respectively.

- (a) Give an example of S and T so that all three scores G , L , and H are different.
- (b) Prove or disprove the statement $L \geq H \geq G$.

Problem 3 (15 pts) Implement the *global* alignment algorithm from class. Your program should read in a FASTA file (see HW1). You can assume that the file just contains two sequences, e.g.

```
> seq1
ACTGGGAAA
> seq2
CTGGAACA
```

You can assume a simplified scoring function δ that has the following form:

$\delta(\text{match}) = a$

$\delta(\text{mismatch}) = b$

$\delta(\text{insertion/deletion}) = c$

The a , b , and c values and filename should be supplied to your program via command-line parameters, e.g. `myAlign 2 -1 -1 test.fa`

(or something like this)

Align the first string with the second string. Print out one optimal alignment.

Demonstrate your algorithms on two test cases (use screen shots to show runs). Include your source code.

Bonus question (1 pt): Print out all optimal alignments.