

CMPE 59H: Bioinformatics, Fall 2017
Assignment 1 - Multiple Sequence Alignment
Due: 17/11/2017, 23:00

Implement a multiple sequence alignment algorithm to align three sequences. In order to obtain the multiple sequence alignment of three sequences, you should first align all pairs of sequences using the Needleman-Wunsch dynamic programming *global sequence alignment* algorithm and determine the pair with the highest pairwise alignment score. Then, you should align the profile of this pair with the third sequence to obtain the final multiple sequence alignment. As the scoring function, use match reward of 1 and indel/mismatch penalty of 0.

Your program should take three DNA sequences as input and should output (i) the pairwise alignments of the sequences and their scores, (ii) the alignment of the profile with the third sequence and the score of the alignment, (iii) the maximum alignment score of the final multiple sequence alignment, and (iv) the final multiple alignment achieving this maximum score. If there are multiple alignments with highest score, you may print one of them.

Test your program by aligning the sequences in the test-seq1.txt and test-seq2.txt files and include the produced output in your report. Include *screenshots* from running your programs on the test sequences and the results produced.

Submission: Please submit your assignment using Moodle. Upload a single zip file named as YourNameSurname.zip. Your zip file should include your report, your source code, and the corresponding readme file. You can use any programming language of your choice. But, your readme file should clearly explain how to run your program.

Late Submission: You are allowed a total of 3 late days on homeworks with no late penalties applied. You can use these 3 days as you wish. For example, you can submit the first homework 2 days late, then the second homework 1 day late. After using these 3 extra days, 10 points will be deducted for each late day.