

Homework Assignment #3

Due: 2020/04/29 12:00

Assessment policy:

- Give full points when correct, $1/n$ for solving each n subproblems. 0 for totally wrong or none, -1 for each errors.
- There may be partial points for proofs if the direction is correct.

1. Apply Longest Common Subsequence (9 pts)

Determine an LCS of the following sequences. (3 pts each)

(a) DNA strands

$$\begin{aligned} X &= \text{ACCGGTCGAGTGC GCGGAAGCCGGCCGAA} \\ Y &= \text{GTCGTTCGGAATGCCGTTGCTCTGTAAA} \end{aligned}$$

As it's too difficult to solve this by hand, we decide to only partially contribute to science of genomics. Find LCS for X_{11} and Y_{10} .

(b) binary sequences: $\langle 1, 0, 1, 0, 0, 1, 0, 1 \rangle$ and $\langle 0, 1, 0, 1, 1, 0, 1, 1, 1, 0 \rangle$.

(c) For two sequences $X = \langle A, B, C, B, D, A, B \rangle$ and $Y = \langle B, D, C, A, B, A \rangle$, count how many longest common sequences exist. You may omit duplicate subsequences, for instance two identical subsequences $\langle A, A \rangle$ and $\langle A, A \rangle$ are counted as one.

2. Memoized Longest Common Subsequence (10 pts)

Give a pseudocode for a top-down memoized version of the LCS-LENGTH procedure TOP-DOWN-LCS-LENGTH. Your code should run in $O(mn)$ time.

(Do not write in actual programming language, like Java or Python)