

CISC-471 WINTER 2016

HOMEWORK 8

Please work on these problems and be prepared to share your solutions with classmates in class on Tuesday March 22. Assignments will **not** be collected for grading.

PROBLEMS

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

Problem 9.8: Design an efficient algorithm for finding the longest exact repeat within a text. (Suggestion: Use a suffix tree to solve this problem.)

Problem 9.10: Design an efficient algorithm for finding the longest exact repeat with at most one mismatch in a text. (Suggestion: Use a suffix tree to solve this problem.)

Problem 9.11: Design an efficient algorithm for finding the longest string shared by two given texts. (This is the longest common substring problem. Suggestion: Use a suffix tree to solve this problem.)

Not from the text : Consider a tree with n leaves. Show that the number of internal nodes can be arbitrarily big. On the other hand, suppose a tree has n leaves and every internal node in the tree has two or more children. Show (using mathematical induction) that in this case the number of internal nodes is at most $n - 1$.