

INSTRUCTIONS:

Write or type your answers on paper. Hand in your answers in class on the due date shown above. Attempt all problems. The maximum grade on this assignment is 20.

Problem 1: [6 pts] (Exercises 7 and 3, Section 7.2) Suppose you are searching in a binary string consisting of 1000 zeros for each of the following patterns:

- (i) 00001
 - (ii) 10000
 - (iii) 01010
- a) How many character comparisons will be made by Horspool's algorithm in searching for each of the three patterns.
 - b) How many character comparisons will be made by Boyer-Moore's algorithm in searching for each of the three patterns.

Problem 2: [8 pts] (inspired from Exercise 2, Section 7.2) Consider the problem of searching for patterns in DNA sequences. A DNA sequence is represented by a text on the alphabet $\{A, C, G, T\}$.

- a) Construct the shift table for Horspool's algorithm for the pattern "TCCTATTCTT".
- b) Construct the good suffix table for Boyer-Moore's algorithm for the same pattern as in Question 2a.
- c) Apply Horspool's algorithm to locate the pattern in Question 2a to the following text:
TTATAGATCTCGTATTCTTTTATAGATCTCCTATTCTT
- d) Apply Boyer-Moore's algorithm to searching for the pattern in the text of Question 2c.

Problem 3: [6 pts] (Exercise 2, Section 7.3) Consider an empty dictionary in which the following keys are inserted : 30, 20, 56, 75, 31, 19, 35, 81. Let the hash function be $H(k) = K \bmod 11$.

- a) Construct the closed hash table.
- b) Find the largest number of key comparisons in a successful search in this table.
- c) Find the average number of key comparisons in a successful search in this table.