National University of Computer and Emerging Sciences

FAST School of Computing **Spring 2024** **Islamabad Campus**

Assignment # 3

(CS-2009 Design and Analysis of Algorithms – Spring-2024)

Due Date and Time: Thursday, 28th March, 2024 (12:20 pm) Marks: 20

Instructions:

- *Late assignment will not be accepted.*
- *Only handwritten attempt will be graded, i.e., printed attempts will not be graded.*
- *Only the attempts submitted to Mr. Amir (or Mr. Aadil in case of Mr. Amir's unavailability) in the Academic office (till the due date & time) will be considered, i.e., the submissions that will be slided beneath instructors' office doors or submitted elsewhere will not be graded.*
- *There will be no credit if the given requirements are changed.*
- *Your solution will be evaluated in comparison with the best solution (having minimum number of steps).*
- *Please mention the question number and its part (if any) before writing down its solution. Use the same conventions used in the assignment's document.*
- *Plagiarism may result in zero marks in the whole assignments category (all assignments) regardless of the percentage plagiarized.*
- *Whenever a calculation is involved, your solution should show complete steps and a final answer. There will be significant marks for the correct final answer (as far as the assignments are concerned).*
- *In case of unavailability to submit the assignment in-person on the submission day, you must submit it either in-person (in the academic office) before the submission day or (as an emergency measure) email the scanned (using, e.g., CamScanner) copy of the handwritten attempt to your course instructor **before the deadline**.*
- *You must write your roll number, name, and section (Algo. Course section) on your submitted attempt.*
- ***All algorithms must be written in the pseudocode form. For reference, please consult pages # 25, 30, 36, 39, 51, and 83 of the text book (4th Edition).***

Question-1: Rabin Karp [4 Marks]

In a medical imaging research laboratory, scientists are studying brain scans to identify abnormal patterns associated with neurological disorders such as Alzheimer's disease. They observe that certain types of lesions and abnormalities often manifest as diagonal patterns in the brain's imaging data. Write an algorithm that automates the abnormality detection process, by analyzing brain scan data and highlighting regions of the data that exhibit the pattern of a disorder.

You are given brain scan data in the form of a 2D matrix **"T"** of size nxn and a pattern of size mxm **"P"** that represents the disorder where mxm can be 1x1,2x2,3x3 up to nxn.

Identify the regions where the disorder pattern is found in the data matrix.

Example 1:

T=

A B C B A B C T Y

F G H J K L H G V
D V T B Y N G B G
F D G A B C V D E
V J O F G H H F W
S R V D V T W F G
K C D E C CC A B C
X Q R T U O F G H
T Y I O B F D V T

P =

ABC
FGH
DVT

Example 2:

A B A B B A D A
A C D A A B B V
C D G B A H K B
G I K C D V F R
Y B J G D B A V
H K G D W C D B
K C D E C C A B
X Q R T U O F G

P =

BA
CD

Write **"Rabin-Karp"** string matching algorithm, using 2 numbers to find modulo while creating the hash table. Additionally, analyze the time complexity of your algorithm.

Note: You must select those prime numbers that result in the minimum number of spurious hits.

Hint: Traverse through the grid in a diagonal manner to extract the symbols. Instead of going row by row or column by column, you start from the top-left corner and move diagonally, retrieving each symbol one by one.

Question 2: Heaps [4 Marks]

a) Starting with the procedure MAX-HEAPIFY, write pseudocode for the procedure MIN-HEAPIFY (A, i, n), which performs the corresponding manipulation on a min-heap. How does the running time of MIN-HEAPIFY compare with that of MAX-HEAPIFY?

b) Considering the function of Build-Max-Heap, write pseudocode for the procedure BUILD-MIN-HEAP.
c) Comment on whether or not will there be any change in the procedure HEAP-SORT keeping in mind the modifications done in parts (a) and (b). Give valid arguments.
d) Illustrate the operation of BUILD-MAX-HEAP and then HEAP-SORT on the array A = {33, 19, 20, 15, 13, 10, 2, 17, 16, 12}; with root at index 1. Show all steps with elements of array after each operation.

Question 3: String Matching [4 Marks]

a) Suppose that all characters in the pattern P are different. Show how to accelerate NAIVE-STRING-MATCHER to run in O(n) time on an n-character text T.

b) You are supposed to find out how many spurious hits does the Rabin-Karp matcher encounter in the text T = 3141592653589793 when looking for the pattern P =26 working with modulo q calculated by using the given formula.

q = (Last four digits of your roll number) modulus 11

For example, if your roll number is 23i-2159 then q = 2159 modulus 11 = 3.

Question 4: ***String Matching using Finite State Automaton*** **[4 Marks]**

Create a transition table for the pattern P = AAB$#C$#$# by computing transition function for all states:

a) Create Finite Automaton using the transition table, which you have formulated above [2 Marks]
b) Show the trace of using your automaton for the text, T = ##$CAAB$#C$#$#B [2 Marks]

Question 5: ***Knuth-Morris-Pratt Algorithm*** **[4 Marks]**

a) Compute the Prefix function of pattern, P = ABCACBCAB using Knuth-Morris-Pratt algorithm for String matching. [2 Marks]
b) Dry run the Knuth-Morris-Pratt Algorithm on text, T = ABABABCACBCABCAB using the pattern P defined in the part a. Show all the steps. [2 Marks]