Mathematical Foundations of Computer Science
Constructor University
Dr. Jürgen Schönwälder

Module: CH-233
Date: 2024-09-13
Due: 2024-09-20

Problem Sheet #2

Problem 2.1: *boyer moore algorithm* (2+2 = 4 points)

You have designed a simple robot that can turn left (L), turn right (R), move one step forward (F), and pause (P) for short time. The robot is programmed by a sequence of robot instructions. For example, the sequence FFLFLFRFRFFLFRF will direct the robot through the maze shown on the slides discussing maze generation algorithms. Using the Boyer Moore algorithm, we can determine whether a robot program contains certain movement sequences.

Let $\Sigma = \{L, R, F, P\}$ be an alphabet and $t \in \Sigma^*$ be a text of length n describing a program for the robot. Let $p \in \Sigma^*$ be a pattern of length m. We are looking for the first occurrence of p in t.

Consider the text $t = FPLFLFRFRFPLFPLFRF$ and the pattern $p = LFPLF$.

a) Execute the Boyer-Moore string search algorithm with the good suffix rule only. How many alignments are used? How many comparisons are done?

b) Execute the Boyer-Moore string search algorithm with the bad character rule and the good suffix rule. How many alignments are used? How many comparisons are done?

Problem 2.2: *big O notation* (1+1 = 2 points)

a) Sort the functions

$$\begin{aligned}
f_1(n) &= \frac{1}{2} n \log n \\
f_2(n) &= n^2 \\
f_3(n) &= \sqrt{n^3} \\
f_4(n) &= n^n \\
f_5(n) &= 100n^2 + 10n^3 \\
f_6(n) &= 2 \log n \\
f_7(n) &= (n^2)^2 \\
f_8(n) &= \log \log n
\end{aligned}$$

in increasing order concerning their big O membership. (It is sufficient to provide the correct order.)

b) Given the functions $f, g, h : \mathbb{N} \to \mathbb{N}$, show that the following transitivity property holds: If $f \in O(g)$ and $g \in O(h)$, then $f \in O(h)$.

Problem 2.3: *clustering using minimum spanning trees* (1+2+1 = 4 points)

You collected data about the music some of your friends listen to. You obtained the following table, where each number represents the number of songs someone heard of a given genre of music:

Person	Pop	Rock	Classic	Jazz
Alice	7	9	0	0
Bob	5	7	2	0
Carol	3	3	1	7
Dan	9	0	0	1
Erin	8	1	0	4
Frank	4	9	0	1
Grace	2	2	8	4

You want to cluster your friends based on the music they have been listening to.

a) For each pair of friends, you caluclate their music listening distance by summing of the absolute distances in each genre. For example, the musical distance between Alice and Bob would be $(7-5)+(9-7)+(2-0)+(0-0)=6$ (this is also known as a Manhattan distance). The result can be represented by a weighted graph where the vertices are your friends and the weighted edges indicate the calculated distance of their music preferences. Draw the graph and label the edges with the calculated distances.

b) Calculate a minimum spanning tree over the graph using Kruskal's algorithm. You can refer to your friends using the first letter of their name.

c) To identify four clusters, remote the three most expensive edges from the spanning tree. Who are the members of the four clusters?