Assignment 11

SFWRENG 2CO3: Data Structures and Algorithms–Winter 2024

Deadline: April 10, 2024*

Department of Computing and Software
McMaster University

Please read the *Course Outline* for the general policies related to assignments.

Plagiarism is a *serious academic offense* and will be handled accordingly.
All suspicions will be reported to the *Office of Academic Integrity*
(in accordance with the Academic Integrity Policy).

This assignment is an *individual* assignment: do not submit work of others. All parts of your submission *must* be your own work and be based on your own ideas and conclusions. Only *discuss or share* any parts of your submissions with your TA or instructor. You are *responsible for protecting* your work: you are strongly advised to password-protect and lock your electronic devices (e.g., laptop) and to not share your logins with partners or friends!

If you *submit* work, then you are certifying that you are aware of the *Plagiarism and Academic Dishonesty* policy of this course outlined in this section, that you are aware of the Academic Integrity Policy, and that you have completed the submitted work entirely yourself. Furthermore, by submitting work, you agree to automated and manual plagiarism checking of all submitted work.

Late submission policy. Late submissions will receive a late penalty of 20% on the score per day late (with a five hour grace period on the first day, e.g., to deal with technical issues) and submissions five days (or more) past the due date are not accepted. In case of technical issues while submitting, contact the instructor *before* the deadline.

Problem 1. Assume we have a list of N real number values that are uniformly distributed between $[0 \ldots 1)$. Consider the following algorithm.

Algorithm UniformSort(L) **:**

```
1: B := an array of N empty linked lists.
2: for each v ∈ L do
3:     Add v to B[⌊N · v⌋].
4: end for
5: for each c ∈ B do
6:     Sort the linked list c using selection sort.
7: end for
8: return the lists B[0], ..., B[N − 1] concatenated.
```

P1.1. Prove that this algorithm is correct.

P1.2. What is the worst-case runtime complexity of this algorithm?

P1.3. Consider two values $v, w \in L$. What is the probability that both values end up in the same linked list at $B[i]$, $0 \leq i < N$?

*This assignment is intended as a *bonus* and as an opportunity to improve ones grades if needed. Remember that the best ten out of twelve assignments will count toward your final grade. If you are happy with your previous assignments, then you are *not* required to make this assignment. The material covered in this assignment will *not* be part of the final exam.

P1.4. Consider a value $v \in L$ that is placed in the linked list at $B[i]$, $0 \leq i < N$. How many other values do you expect to end up in $B[i]$?

P1.5. What is the expected runtime of the above algorithm?

Problem 2. There was a terrible accident with our strings due to which they all got messed up! Hence, our substring search algorithms no longer work.

P2.1. Assume that all our strings got rotated. For example, the string "example" got rotated by 3 characters, and now we have "pleexam". Given two strings A and B, write an algorithm that can determine whether A can be produced by rotating B zero-or-more characters in worst-case $O(|A|)$. Argue why your algorithm is correct and has the stated complexity.

P2.2. Assume that a large collection of documents got corrupted due to which zero-or-more characters have mutated (changed into others, but no insertions or deletions). You want to find those documents that match some keyword K (up-to-c corruptions). For example, if we are looking for the string "keyword" with $c = 2$, then "sayword" and "!esword" match the search. Argue why your algorithm is correct and what the complexity of your algorithm is.

Assignment Details

Write a report in which you solve each of the above problems. Your submission:

1. must start with your name, student number, and MacID;
2. must be a PDF file;
3. must have clearly labeled solutions to each of the stated problems;
4. must be clearly presented;
5. must *not* be hand-written: prepare your report in LaTeX or in a word processor such as Microsoft Word (that can print or export to PDF).

Submissions that do not follow the above requirements will get a grade of zero.

Grading

Each problem counts equally toward the final grade of this assignment.