Assignment 8

SFWRENG 2CO3: Data Structures and Algorithms–Winter 2024

Deadline: March 22, 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. Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be an *undirected tree* (the graph $\mathcal{G}$ is *undirected*, *connected*, and has $|\mathcal{N}| = |\mathcal{E}| + 1$ if we count edges (v, w) and (w, v) as the same edge). Let $m, n \in \mathcal{N}$. We say that the *distance* between m and n, denoted by $\text{dist}(m, n)$, is the length of the shortest path between m and n.

P1.1. Prove that $\text{dist}(m, n) = \text{dist}(n, m)$.

P1.2. Prove that there is a *unique* path without repeated nodes and edges from node m to node n with length $\text{dist}(m, n)$.

P1.3. Prove the triangle inequality $\text{dist}(m, n) \leq \text{dist}(m, x) + \text{dist}(x, n)$.

P1.4. Provide an algorithm that computes the distance $d = \max_{m,n \in \mathcal{N}} \text{dist}(m, n)$ that is the maximum distance between any pair of nodes in $\mathcal{G}$ in $\mathcal{O}(|\mathcal{N}| + \mathcal{E})$.

Problem 2. Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be a graph such that the existence of a path $m \ldots n$ represents that $(m, n) \in R$ for some relationship R. Hence, R is the transitive closure of $\mathcal{E}$. (For example, $\mathcal{E}$ could represent the *parentOf* relationship, in which case R represents *ancestorOf*.)

We want to send the derived relationship R to a remote site with limited internet access. To do so, we want to find a small representation $E \subseteq (\mathcal{N} \times \mathcal{N})$ of $\mathcal{E}$ such that the transitive closure of E is R. In specific, we want to find *minimal tc-subset* E such that $\textsc{TransitiveClosure}(E) = R$ and we cannot remove any edges from E (more formally, $\textsc{TransitiveClosure}(E \setminus \{(m, n)\}) \neq R$ for all edges $(m, n) \in E$).

P2.1. Assume the graph $\mathcal{G}$ is *undirected*. Provide a highly efficient algorithm with a running time of $\mathcal{O}(|\mathcal{N}| + |\mathcal{E}|)$ to compute a minimal tc-subset of $\mathcal{E}$.

P2.2. Assume the graph $\mathcal{G}$ is *directed*. Provide a highly efficient algorithm with a running time of $O(|\mathcal{N}|+|\mathcal{E}|)$ to compute a minimal tc-subset of $\mathcal{E}$.

Hint: To do so, first answer some elementary questions about minimal subsets. For example:

(a) Say a node n has only one outgoing edge $(m, n) \in \mathcal{E}$. Can a minimal subset E exist with $(m, n) \notin E$?

(b) Consider a directed acyclic graph with a node n that has outgoing edges $(m, n_1), (m, n_2) \in \mathcal{E}$. When can we remove (m, n_1) in favor of keeping (m, n_2)?

(c) Say a set of nodes $S \subseteq \mathcal{N}$ are a strongly connected component. What is the minimal number of edges any minimal tc-subset requires between nodes in this strongly connected component S?

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.