Algorithm

Final Exam

June 11, 2023

Please answer the following questions with **explanation**. No points will be given if there is no explanation provided.
[Note] **Kindly take note that the use of ChatGPT is permitted. However, it is important to refrain from simply copying and pasting responses from ChatGPT, as this typically results in receiving a lower score.**

1. Given a directed graph $G = (V, E)$, and two vertices u and v in V, we call vertex v is reachable from u, if there exists a directed path from u to v. A vertex s in V is called a source vertex if every vertex in V is reachable from s.

   (a) **(10%)** Given a directed graph $G = (V, E)$, and a specified vertex v in V, design a linear time algorithm (i.e., your algorithm should run in $O(|V| + |E|)$-time) to determine if v is a source vertex. You need to describe the data structure used in your algorithm.

   (b) **(15%)** Given a directed acyclic graph $G = (V, E)$, you are asked to determine if G contains a source vertex. If you apply the algorithm of the subproblem (a) on every vertex of G, you will get algorithm runs in $O(|V|^2 + |V||E|)$-time. It is not desirable. Design a more efficient algorithm for this problem. Analyze the time complexity of your algorithm.

   (c) **(15%)** Given a directed graph $G = (V, E)$, you are asked to determine if G contains a source vertex. Note that the given graph may contain directed cycles. As in subproblem (b) an $O(|V|^2 + |V||E|)$-time algorithm is not acceptable. Design a more efficient algorithm for this problem. Analyze the time complexity of your algorithm.

2. Please answer the following two questions about the maximum flow problem.

   (a) **(10%)** Fig. 1 is a flow network where each edge is labeled with its capacity. Please derive the maximum flow from vertex s to vertex t.

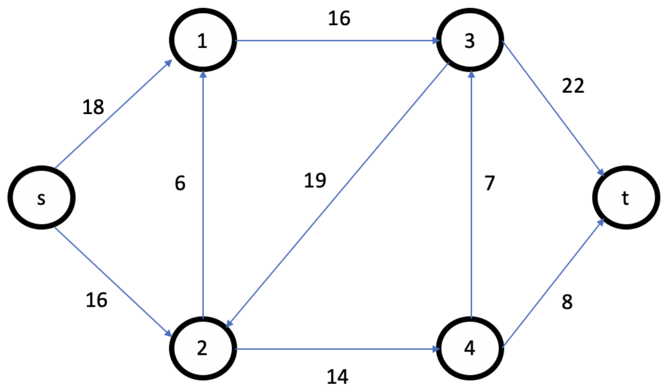

Figure 1: Flow Network

(b) **(10%)** Fig. 2 is a residual network where each eadge is labeled with its flow and its capacity. You are asked to:

i. write down the augmenting paths used by the Ford-Fulkerson algorithm

ii. identify the minimum S-T cut for the network

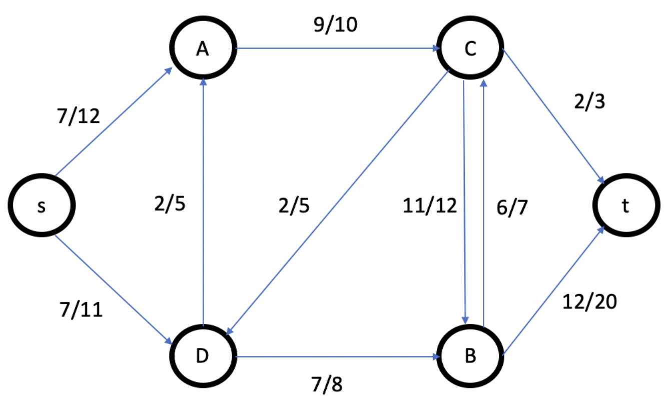

Figure 2: Residual Network

3. **(15%)** Given a sequence $K =< k_1, k_2, ..., k_n >$ of n distinct keys in sorted order such that $k_1 < k_2 < ... < k_n$, and we wish to build a binary search tree from these keys. For each key k_i, we have a probability p_i that a search will be for k_i. Some searches may be for values not in K, and so we

also have $n+1$ "dummy keys" $d_0, d_1, d_2, \ldots, d_n$ representing values not in K. In particular, d_0 represents all values less that k_1, d_n represents all values greater that k_n, and for $i = 1, 2, \ldots, n-1$, the dummy keys d_i represents all values between k_i and k_{i+1}. For each dummy key d_i, we have a probability q_i that a search will correspond to d_i. Determine the cost and structure of an optimal binary search tree in the expected cost of search time for a set of $n = 7$ keys with the following probabilities: $p_1 = 0.04$, $p_2 = 0.06$, $p_3 = 0.08$, $p_4 = 0.02$, $p_5 = 0.10$, $p_6 = 0.12$, $p_7 = 0.14$, $q_0 = 0.06$, $q_1 = 0.06$, $q_2 = 0.06$, $q_3 = 0.06$, $q_4 = 0.05$, $q_5 = 0.05$, $q_6 = 0.05$, and $q_7 = 0.05$.

4. In the Middle Ages, an alchemist pursued the goal of increasing the length of gold using various formulas. However, formulas are not always perfect. There exist n types of formulas, each represented by (a_i, b_i, k_i) for $1 \le i \le n$, where $a_i \ge 1$ and $b_i \ge 1$. If the i-th formula is applied at the start of a round, the gold's length immediately increases by a_i. Subsequently, for the next k_i rounds, including the round of formula usage, the gold's length decreases by b_i at the end of each round. The effects of the formulas can accumulate. Thus, at the conclusion of each round, the decrease in gold length is equal to the sum of b_i for all formulas i that remain in effect.

   In each round, the alchemist can use at most one formula, and each formula can only be used once. The alchemist has the freedom to choose any subset of formulas and use them in any desired order. Initially, the gold available is a mere piece of gold leaf, meaning its length is ZERO. The objective is to maximize the length of the gold.

   (a) **(10%)** Suppose that $n = 4$ and the formulas are $(5, 3, 2)$, $(20, 33, 1)$, and $(30, 115, 1)$. What is the maximum lenght of the gold?

   (b) **(25%)** Define the following notations:

   - $A(i, j)$ represents the maximum length of the gold under the following conditions: (1) Only the first i formulas are taken into account, and (2) j formulas, including the i-th formula, are in effect when the value $A(i, j)$ is achieved.
   - $B(i, j)$ represents the maximum length of the gold under the following conditions: (1) Only the first i formulas are taken into account, (2) j formulas, including the i-th formula, are in effect when the value $B(i, j)$ is achieved, and (3) the i-th formula is used in the first round.

   $A(i, j)$ and $B(i, j)$ are defined to be $-\infty$ if it is not possible. Based on the optimal substructure, write down the recurrence formula of $A(i, j)$ and $B(i, j)$. If needed, you can assume that the formulas are sorted in a particular order.