

CS570 Fall 2021: Analysis of Algorithms Exam I

	Points		Points
Problem 1	20	Problem 5	12
Problem 2	15	Problem 6	10
Problem 3	15	Problem 7	16
Problem 4	12		
	Total	100	

Instructions:

1. This is a 2-hr exam. Open book and notes. No internet access.
2. If a description to an algorithm or a proof is required, please limit your description or proof to within 150 words, preferably not exceeding the space allotted for that question.
3. No space other than the pages in the exam booklet will be scanned for grading.
4. If you require an additional page for a question, you can use the extra page provided within this booklet. However please indicate clearly that you are continuing the solution on the additional page.
5. Do not detach any sheets from the booklet. Detached sheets will not be scanned.
6. If using a pencil to write the answers, make sure you apply enough pressure, so your answers are readable in the scanned copy of your exam.
7. Do not write your answers in cursive scripts.
8. This exam is printed double sided. Check and use the back of each page.

1) 20 pts

Mark the following statements as **TRUE** or **FALSE**. No need to provide any justification.

[**TRUE/FALSE**]

A graph G has a unique MST if and only if all of G 's edges have different weights.

[**TRUE/FALSE**]

The runtime complexity of merge sort can be improved asymptotically by recursively splitting an array into three parts (rather than into two parts)

[**TRUE/FALSE**]

In the interval scheduling problem, if all intervals are of equal size, a greedy algorithm based on the earliest start time will always select the maximum number of compatible intervals.

[**TRUE/FALSE**]

Amortized analysis is used to determine the average runtime complexity of an algorithm.

[**TRUE/FALSE**]

Function $f(n) = 5n^24^n + 6n^43^n$ is $O(n^24^n)$.

[**TRUE/FALSE**]

If an operation takes $O(1)$ time in the worst case, it is possible for it to take $O(\log n)$ amortized time.

[**TRUE/FALSE**]

If you consider every man and woman to be a node and the matchings between them to be represented by edges, the Gale-Shapley algorithm returns a connected graph

[**TRUE/FALSE**]

A Fibonacci Heap extract-min operation has a worst case run time of $O(\log n)$.

[**TRUE/FALSE**]

Consider a binary heap A whose elements have positive or negative key values. If we randomly square the key values for some elements, the heap property can be restored in A in linear time.

[**TRUE/FALSE**]

The order in which nodes in the set $V-S$ are added to the set S can be the same for Prim's and Dijkstra's algorithms when running these two algorithms from the same starting point on a given graph.

2) 15 pts

Consider a full binary tree representing k generations of the Vjestica family tree (i.e. family tree with k levels). It so happens that the Vjestica family is generally very tall, and they have kept records of all its member's heights which has now sparked a researcher's interest in this data. The researcher wants to find out if there is anyone in this family tree who is taller than its parent (if the parent exists in the tree) AND his/her two children (if the children exist in the tree). In other words, the member we are looking for should be taller than anyone they are immediately related to in the family tree.

- a) How many Vjestica's are there in this family tree? (2 pts)
- b) Show that if heights are unique, we can always find such a family member. (5 pts)
- c) Assuming that the heights are unique, design a divide and conquer solution to find such a member in no more than $O(k)$ time. (5 pts)
- d) Show your complexity analysis using the Master Method. (3 pts)

3) 15 pts

There are n distinct jobs, labeled J_1, J_2, \dots, J_n , which can be performed completely independently of one another. Each job consists of two stages: first it needs to be preprocessed on the supercomputer, and then it needs to be finished on one of the PCs. Let's say that job J_i needs p_i seconds of time on the supercomputer, followed by f_i seconds of time on a PC. Since there are at least n PCs available on the premises, the finishing of the jobs can be performed on PCs at the same time. However, the supercomputer can only work on a single job at a time without any interruption. For every job, as soon as the preprocessing is done on the supercomputer, it can be handed off to a PC for finishing.

Let's say that a schedule is an ordering of the jobs for the supercomputer, and the completion time of the schedule is the earliest time at which all jobs have finished processing on the PCs.

- a) Design an algorithm that finds a schedule with as small a completion time as possible. Your algorithm should not take more than $O(n^2)$ time in the worst case. (6 pts)

See part b and C on the next page.

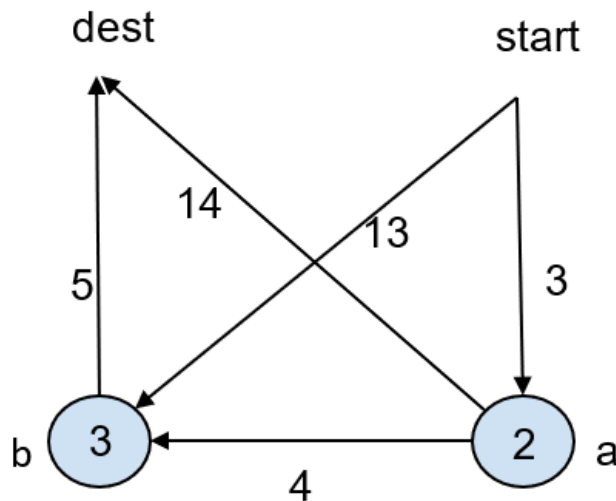
b) Prove that your algorithm produces an optimal solution. (7 pts)

c) Analyze the complexity of your solution (2 pts)

4) 12 pts

You are in the city called *start*, and you want to drive your new electric car to the city called *dest*. There are many possible paths to take for this trip, but the good news is that if you take any of these paths, the recharging stations along the way are designed such that if you stop at every recharging station on your way, you will always have enough charge to reach *dest*. You can consider the map as a **directed weighted** graph with recharging stations at some of its nodes. The edge weights C_e represent the time it takes to drive through edge e . You are currently in the city *start* with a full charge and you want to optimize your trip to get to *dest* as fast as possible. What you also need to consider is that some charging stations are faster than others. In fact, for each charging station i , we know that we need to wait a fixed amount of time C_i to recharge. Give an efficient algorithm to find out the fastest path from city *start* to city *dest* (minimum total time) assuming that you must stop and recharge at every charging station on that path. No proof necessary.

Example: Assume charging in cities a and b takes 2 and 3 units of time respectively.



The optimal path is start -> a -> b -> dest. The total time is 17:

- 3 hrs driving between start to a
- 2 hours charging at a
- 4 hours driving from a to b
- 3 hours charging at b
- 5 hours driving from b to dest

More space provided on the next page

Space for problem 4

5) 12 pts

The United States Commission of Southern California Universities (USCSCU) is researching the impact of class rank on student performance. For this research, they want to find a list of students ordered by GPA containing every student in California. However, each school only has an ordered list of its own students by GPA and the commission needs an algorithm to combine all the lists. There are a few hundred colleges of interest, but each college can have thousands of students, and the USCSCU is terribly underfunded so the only computer they have on hand is an old vacuum tube computer that can do about a thousand operations per second. They are also on a deadline to produce a report so every second counts. Find the fastest algorithm for yielding the combined list and give its runtime in terms of the total number of students (m) and the number of colleges (n).

6) 10 pts

Consider the Gale-Shapley algorithm operating on n men and n women, with women proposing.

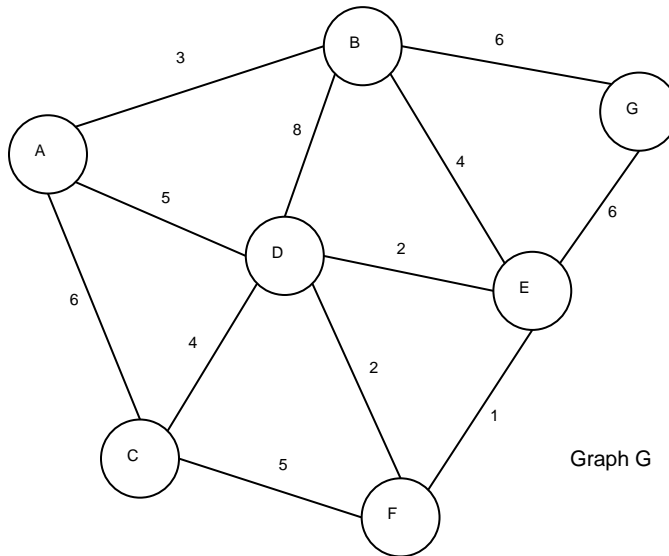
a) What is the maximum number of times a woman may be rejected (with respect to the problem size n)? Give an example where this can happen. (5 pts)

b) Now, consider the following modification to the G-S algorithm: At each iteration, we always pick the free woman with the highest average preference among men, i.e. the most “popular” remaining woman (when taking an average across all men’s preference lists).

Prove or disprove: this will help reduce the number of rejections for some women. (5 pts)

7) 16 pts

For the next 3 multiple choice questions consider the following graph G.



1. In graph G, if we use Kruskal's Algorithm to find the MST, what is the third edge added to the solution? Select all correct answers (4 pts)
 - a. E-F
 - b. D-E
 - c. A-B
 - d. C-F
 - e. D-F
2. In graph G, if we use Prim's Algorithm to find MST starting at A, what is the second edge added to the solution? (4 pts)
 - a. B-G
 - b. B-E
 - c. D-E
 - d. A-D
 - e. E-F
3. What is the cost of the MST in the Graph? (4 pts)
 - a. 18
 - b. 19
 - c. 20
 - d. 21
 - e. 22

See next page for multiple choice problem 4

4. Which of the following Asymptotic notation is $O(n^2)$? Select all apply. (4 pts)

- a. $O(7 \log n^3)$
- b. $O(n \log n^n)$
- c. $O(3^n)$
- d. $O(\log 10^n)$
- e. $O(5n^2 + n \log n)$
- f. $O(2000)$

Additional space

Additional Space

Additional Space