Midterm 2

Name:

SID:

Exam Room:

SID of student to your left:

SID of student to your right:

Do not turn this page until your instructor tells you to do so.

- After the exam starts, please *write your SID on the front of every page*. We may deduct points if your name is missing from any page. You will not be allowed to fill in your name after time is called.
- For short question, your answers must be written clearly inside the box region. Any answer outside the box will not be graded. For longer question, if you run out of space, you must clearly mention in the space provided for the question if part of your answers are elsewhere.
- Try to answer all questions. Not all parts of a problem are weighted equally. Before you answer any question, read the problem carefully. Be precise and concise in your answers.
- You may use the blank page at the back for scratch work, but we will not look at it for grading.
- You may consult only *two sheet of notes*. Apart from that, you may not look at books, notes, etc. Calculators, phones, computers, and other electronic devices are NOT permitted.
- There are **14** pages (7 double sided sheets) on the exam. Notify a proctor immediately if a page is missing.
- **Any algorithm covered in the lecture can be used as a blackbox.**
- **You have 110 minutes: there are 8 questions on this exam worth a total of 110 points.**
- Indicate the discussion you attend on page 2 **after** the exam starts. Fill in the square.
- Good luck!

Select the following by filling in the square. If you go to different sections approximately the same number of times during the semester, feel free to select multiple section times.

TA and section time:

☐ I don't go to discussion section

☐ 9am, Raymond, Etcheverry 3109

☐ 10am, Simin, Wheeler 220

☐ 10am, Zhen, Wheeler 220

☐ 10am, Alex, Soda 310

☐ 11am, Jerry, Hearst Field Annex B5

☐ 11am, Aditya Baradwaj, Etcheverry 3119

☐ 11am, Ching, Soda 310

☐ 12pm, Aditya Mishra, Hearst Field Annex B5

☐ 12pn, Aditya Baradwaj, Cory 289

☐ 12pm, Kevin, Soda 310

☐ 1pm, Raymond, Hearst Field Annex B5

☐ 1pm, Jingcheng, Etcheverry 3113

☐ 2pm, Edmund, Etcheverry 3111

☐ 2pm, Ching, Etcheverry 3113

☐ 3pm, Gary, Leconte 385

☐ 3pm, Mudit, Etcheverry 3113

☐ 4pm, Nathan, Dwinelle 215

☐ 4pm, Kevin, Wheeler 24

☐ 4pm, Alex, Dwinelle 242

☐ 5pm, Pasin, Etcheverry 3119

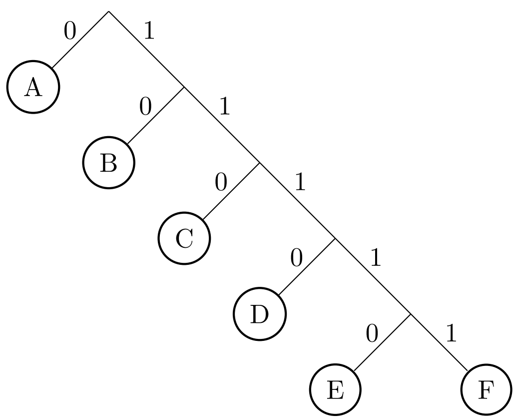
☐ 5pm, Peihan, Dwinelle 79

Name:

1. Huffman Encoding (6 points)

A document consists of letters $\{A, B, C, D, E, F\}$ occurring with various frequencies. Suppose the following is the tree corresponding to an optimal Huffman encoding of the letters $\{A, B, C, D, E, F\}$ (the optimal tree need not be unique).

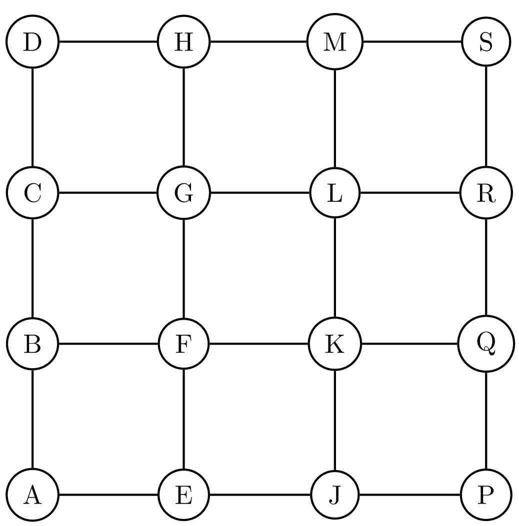
Suppose the frequencies of E and F are 1 and 1 (both occur only once) respectively. Here frequency refers to the number of occurrences in the document. What is the smallest possible frequency for each of the following letters?



Letter	Frequency
A	
B	
C	

2. Max Flow (10 points)

Consider the following undirected graph with all edge capacities equal to 1. An undirected edge of capacity 1 can carry flow of up to 1 unit in either direction.



(a) What is the value of the maximum flow from A to S?

(b) What is the value of the maximum flow from F to L?

(c) Call an edge to be critical, if removing the edge reduces the value of the maximum flow from F to L.

For each of the following edges, exhibit a minimum cut they are part of, thereby proving that they are critical. Specify the minimum cut by marking the vertices on one side of the cut in the table.

(a) Edge FB

Vertices on B side of cut :

A	B	C	D	E	F	G	H	J	K	L	M	P	Q	R	S
	X														

(b) Edge HM

Vertices on H side of cut :

A	B	C	D	E	F	G	H	J	K	L	M	P	Q	R	S
							X								

3. Alternative Recurrence Relations (20 points)

For the longest increasing subsequence and edit distance problems, here are candidate alternate definitions of subproblems. For each of these subproblems, **fill in** the square for either "Recurrence relation below" OR "NO recurrence relation". If there exists a valid recurrence relation which solves the problem, write the recurrence relation in the answer box. For parts (a) - (c), you are trying to solve the longest increasing subsequence problem and for parts (d) and (e), you are trying to solve the edit distance problem.

(To avoid guessing, if you incorrectly say "NO recurrence relation" then you lose twice as many points).

(a) $T[i]$ = length of longest increasing sequence in $A[i, \ldots, n]$ starting at $A[i]$ (i.e. $A[i]$ is part of the longest increasing sequence).

☐ Recurrence relation below

☐ NO recurrence relation

Name:

(b) $T[i]$ = length of longest increasing sequence in $A[1, \ldots, i]$.

☐ Recurrence relation below

☐ NO recurrence relation

(c) $T[i]$ = length of longest increasing sequence in $A[i, \ldots, n]$

☐ Recurrence relation below

☐ NO recurrence relation

(d) $E[i, j]$ = edit distance of $A[i, \ldots, j]$ to $B[i, \ldots j]$.

☐ Recurrence relation below

☐ NO recurrence relation

(e) $E[i, j, k, \ell]$ = edit distance of $A[i, \ldots, j]$ to $B[k, \ldots, \ell]$.

☐ Recurrence relation below

☐ NO recurrence relation

Name:

4. HORN-SAT (6 points)

On running the Horn-SAT algorithm on a horn-SAT formula, the following assignment is obtained:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
T	T	T	F	F	F	F	F	F	T

Suppose we modify the horn-SAT formula in the following ways, indicate whether the new formula is Necessarily Unsatisfiable or Necessarily Satisfiable or Can't Say. **Fill in** one of the choices.

(a) Add new constraint $x_1 \wedge x_2 \longrightarrow x_4$

- ☐ Necessarily Unsatisfiable
- ☐ Necessarily Satisfiable
- ☐ Can't Say

(b) Add new constraint $x_1 \wedge x_2 \longrightarrow \overline{x_3}$

- ☐ Necessarily Unsatisfiable
- ☐ Necessarily Satisfiable
- ☐ Can't Say

5. Duality (12 points)

(a) Write down the dual for the following linear program:

$$\begin{aligned} \text{Maximize} \quad & x_1 + 3x_3 \\ \text{subject to} \quad & \\ & 2x_1 + 4x_2 \leq 120 \\ & 5x_1 - 6x_2 \leq 20 \\ & 7x_1 + 8x_2 + 9x_3 \leq 100 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Use y_i as decision variables for the dual.

Objective function:

Constraints:

Use below space for your work (we will use it while awarding partial credit):

(b) How does the dual change if we replace the constraint $2x_1 + 4x_2 \leq 120$ by an equality $2x_1 + 4x_2 = 120$? (No need to write out the dual again, just specify the change/changes to be made)

Name:

6. Lecture Scheduling (25 points)

A university wants to renovate lecture halls with the student's tuition. However, it must also schedule all the lectures of the existing courses. As a scheduling officer, you want to determine what is the minimum number of classrooms (lecture halls) that you need to schedule N lectures. Lectures are given to you in start-time finish-time pairs:

$$A[(s_1, f_1), (s_2, f_2), \ldots, (s_N, f_N)]$$

The university would like to schedule these lectures in the available classrooms without any conflicts – A conflict occurs if start time of a lecture is less than the finish time of an already scheduled lecture in a classroom.

(a) Design a greedy algorithm to find the minimum number of classrooms that are needed to schedule all lectures such that there are no conflicts.

i. What does the greedy algorithm do at each step? (Describe the main idea of the algorithm.

Space for scratch work: (we won't grade this)

ii. Suppose your greedy algorithm finds a schedule with c classrooms, why is there no schedule with $c - 1$ classrooms? (briefly argue in two-three sentences)

(b) After running the greedy algorithm, the administration realized that not all lectures can be scheduled in the available c classrooms. Therefore, the university has now decided to cancel some of these lectures.

The administration would like to try out solving a linear program to determine the set of lectures to be cancelled. Suppose t_i is the loss in tuition from cancelling the i^{th} lecture, write a linear program to determine the set of lectures to be cancelled so as to minimize total loss in tuition.

i. What are the variables of the linear program?

ii. What is the objective function?

iii. What are the constraints?

iv. It turns out that the optimal solution to the above linear program has a total tuition loss $= T$, but the university was forced to incur a loss strictly larger than T. When can this happen?

7. Taco Trucks (20 points)

You are running a Taco truck company with 4 trucks. There are n cities that your company operates in. On day 0, all trucks are parked in city numbered 1. For the next m days, there is one street festival each day in one of the cities. Specifically, on day i, there is a street festival in city $c[i]$. In order to cater to a street festival on day i you need at least one truck to go the city $c[i]$.

Let $d[i, j]$ denote the distance between cities i and j.

Given the schedule $c[\cdot]$ of the street festivals and the distances $d[\cdot, \cdot]$ between the cities, design a dynamic programming algorithm that finds the optimal way to move the trucks around, so as to cater to every street festival while minimizing the total distance driven by all the trucks put together.

(a) What are the subproblems?

(b) What are the base cases? Give an expression for the final answer in terms of the subproblems.

(c) What is the recurrence relation?

Name:

8. Floyd-Warshall Paths (11 points)

Suppose $d(\cdot,\cdot,\cdot)$ is the dynamic programming table for an execution of the Floyd-Warshall algorithm on a graph G with n nodes. Recall that the subproblems $d(i,j,k)$ for $i,j,k \in \{1,\ldots,n\}$ is defined as follows:

$d(i,j,k) =$ length of the shortest path from i to j that only uses vertices in $\{1,\ldots,k\}$ as intermediate vertices

The recurrence relation is given by:

$$d(i,j,k) = \min\{d(i,j,k-1), d(i,k,k-1) + d(k,j,k-1)\}$$

Design an algorithm that given two vertices i and j, recovers the shortest path from i to j using the $d(.,.,.)$ table. The run-time of your algorithm to find the shortest path from i to j, must be equal to $O((\#\text{ of hops on shortest path from } i \text{ to } j) \cdot \log n)$.

Write Pseudocode below

DO NOT TEAR ANY PAGES OFF YOUR EXAM.

(Extra page for scratch work.)