COM S 311 SPRING 2021
EXAM 2
Due: April 22 7:59 p.m.
You must turn in a single pdf with your typed answers by 7:59 p.m.

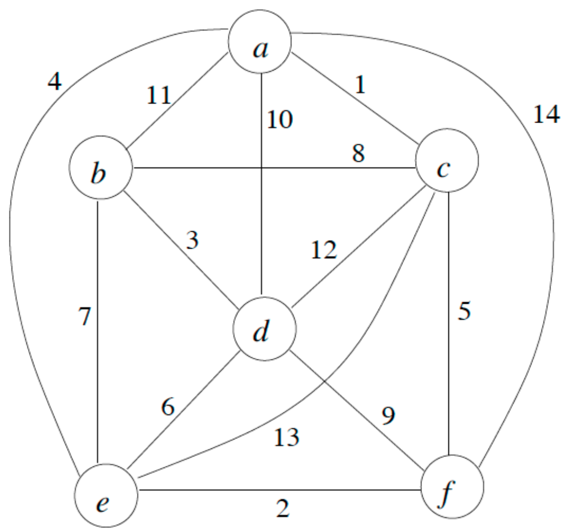
Guidelines

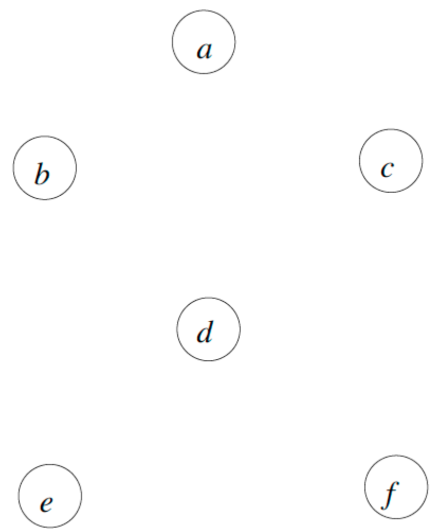
- For each problem, if you write the statement "I do not know how to solve this problem" (and nothing else), you will receive 20% credit for that problem. If you do write a solution, then your grade could be anywhere between 0% to 100%. To receive this 20% credit, you must explicitly state that you do not know how to solve the problem.
- You are **not** allowed to discuss the problems with anyone. You are allowed to use the textbook and notes. Do **not** copy solutions from the internet. Your writing should demonstrate that you understand the proofs completely.
- When proofs are required, you should make them both clear and rigorous. Do not handwaive.
- Please submit your assignment on the given Canvas exam.
  - You **must** type your solutions, except manual tree drawing. Please submit a PDF version.
  - Please make sure that the file you submit is not corrupted and that its size is reasonable (e.g., roughly at most 10-11 MB).

    If we cannot open your file, your exam will not be graded.
- Any concerns about grading should be expressed within one week of returning the exam.

Problem 1: (Minimum Spanning Trees) (20 Points)

(a) Construct a minimum spanning tree for the graph below on the left. Draw the tree by adding edges to the graph on the right.





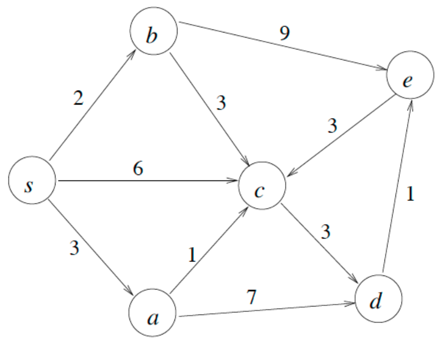
(b) Fill out the table below with the edges of the above minimum spanning tree according to their orders of selection by Kruskal's and Prim's algorithms, respectively.

	1	2	3	4	5
Kruskal's	(,)	(,)	(,)	(,)	(,)
Prim's	(,)	(,)	(,)	(,)	(,)

(c) Suppose all edges in a graph G have different edge weights. Show that G has a unique minimum spanning tree.

Problem 2: (Shortest Paths) (20 Points)

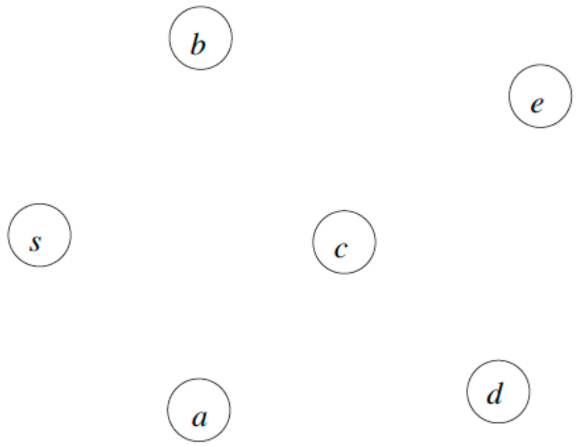
(a) Consider the graph G below.



Show the execution of Dijkstra's algorithm on G by filling out the table below. Assume the source is vertex s. You must show the changes in the d-array at the end of each iteration. Iteration 0 refers to the situation just before the first iteration of the **while** loop. Also, fill in the vertex that is selected (i.e., extracted from the priority queue) at each iteration.

	$d[\,]$						selected
iter.	s	a	b	c	d	e	node
0							s
1							
2							
3							
4							
5							
6							

(b) Draw a shortest path tree for the graph in (a) with source vertex s.



(c) Suppose that all edge weights in a given graph $G := (V, E)$ are NOT negative, and that the shortest path distances in G from a source $s \in V$ to each vertex $v \in V$ are unique. Let V_k denote the vertices in V with the k-closest shortest path distances from s.

Example: Let $V = \{s, u, v\}$ and $0, 2, 3$ be the shortest path distances of the vertices s, u, v from s in G respectively. Then, s is the 1-closest vertex, u is the 2-closest vertex, and v is the 3-closest vertex. Therefore we have $V_1 = \{s\}$, $V_2 = \{s, u\}$, $V_3 = V$.

Prove by contradiction that a shortest path from the source vertex $s \in V$ to a k-closest vertex $x \in V$ consists only of vertices in V_k.

Problem 3: (P is Closed under Reverse Complement) (20 Points)

Consider the operation of reverse complement on a language over the alphabet $\{0, 1\}$. The reverse complement of a string $x \in \{0, 1\}^*$ of length n is a string $y \in \{0, 1\}^*$ of the same length such that $y_k = 1 - x_{n-k+1}$ for $k = 1, 2, ..., n$, where y_k is the bit at index k of string y. Let $\overline{x}$ denote the reverse complement of string x. For example, if $x = 011101$, then $\overline{x} = 010001$. Note that $\overline{x}$ is obtained from x by reversing its bit sequence and complementing each bit (changing 1 to 0 and 0 to 1). Let L be a language over the alphabet $\{0, 1\}$. The reverse complement of L is defined as

$$\overline{L} = \{\overline{x} : x \in L\}.$$

Show that if L is in P, then $\overline{L}$ is also in P. Note that the reverse complement of L is not related to the set complement of L. Your algorithm takes as input a binary string in $\{0, 1\}^*$.

Problem 4: (String Transformation is in NP) (20 Points)

A string $x \in \{0,1\}^*$ is transformed into another string y by a sequence of operations of two types. Let x be a string of n bits, indexed from 1 to n, where $x_{i,j}$ denotes a substring of x, consisting of consecutive bits from index i to index j in the same order, with $1 \leq i \leq j \leq n$. Let xyz denote the concatenation of strings x, y, and z. An inversion operation $r(i,j)$ turns string x into $x_{1,i-1}\overline{x_{i,j}}x_{j+1,n}$, with $1 \leq i \leq j \leq n$, where $\overline{x_{i,j}}$ is the reverse complement of $x_{i,j}$ (see problem 3). A deletion operation $d(i,j)$ turns string x into $x_{1,i-1}x_{j+1,n}$. For example, an inversion operation $r(3,6)$ transforms string 11011100 into string 11000100 ($x_{3,6} = 0111$, $\overline{x_{3,6}} = 0001$), which is further transformed into string 100100 by a deletion operation $d(2,3)$. Show that the problem of deciding whether a string can be transformed into another string by a sequence of at most k operations of substring deletions and inversions is in NP. Specifically, the problem is defined as the formal language

ST = $\{< x, y, k >$: there exists a sequence of at most k operations
of deletions and inversions to transform binary strings x into y $\}$.

Show that ST is in NP. Note that your algorithm takes two types of input: one type is an ordinary input including two strings x and y along with an integer k, and the other is a certificate.

Problem 5: (Restricted CNF-SAT is in P) (20 Points)

A boolean formula in conjunctive normal form, or CNF, is expressed as an AND of clauses, each of which is the OR of one or more literals. A restricted CNF formula meets the following requirements. For every clause with two or more literals in the formula, if the clause contain literal x, then no clause can contain $\neg x$, and if it contain $\neg x$, then no clause can contain x. If any clause with only one literal contains x or $\neg x$, then this literal cannot occur in any clause with two or more literals. For example, the boolean formulas,

$(x_1 \vee \neg x_2) \wedge (\neg x_3 \vee x_4 \vee x_5) \wedge x_6 \wedge \neg x_6$, and
$(x_1 \vee \neg x_2) \wedge (\neg x_3 \vee x_4 \vee x_5) \wedge x_6 \wedge \neg x_7$,

meet the requirements, and the following formulas do not,

$(x_1 \vee \neg x_2) \wedge (\neg x_3 \vee x_4 \vee x_5) \wedge x_2$, and
$(x_1 \vee \neg x_2) \wedge (\neg x_1 \vee x_2 \vee x_5)$.

Show that the problem of deciding whether a restricted boolean formula in CNF is satisfiable is in P. Specifically, the problem is defined as the formal language

RES-CNF $= \{< \theta >: \theta$ is a satisfiable restricted boolean formula in CNF$\}$.

Show that RES-CNF is in P. Note that input to your algorithm is a restricted boolean formula in CNF.