Print Name: _____ **ID:** _____

EC 504 Sample Final

Instructions: Put your name and BU ID on this page. I stapler will allow you to stable together at the end of the exam.. This exam is closed book and no notes – except for 2 pages of personal notes to be passed in the end. No use of laptops or internet. You have one hour and fifty minutes. Do the easy ones first.

1. Answer True or False to each of the questions below. Each question is worth 2 points. Answer true only if it is true as stated, with no additional assumptions. **Note if you give a good short explanation of your choice you will receive partial credit even if the answer is wrong!**

   (a) Inserting numbers $1, \ldots, n$ into a binary min-heap in that order will take $\Theta(n)$ time.

   (b) Suppose you have an unsorted array of $2n + 1$ elements. One can determine the n smallest elements of the array in $O(n)$ time.

   (c) The second smallest element in a binary min-heap with all elements with distinct values will always be a child of the root.

   (d) Any sorting algorithm on a set of 32 bit positive integers must have a run time complexity $f(N) \in \Omega(N \log(N))$.

   (e) There are instances of the integer knapsack problem which can be solved in polynomial time by a greedy algorithm.

   (f) Consider a binary search tree with n keys. Finding whether a key value is already in the tree can be done in $O(\log(n))$.

   (g) Consider a binary search tree that was constructed by branching on the median value at each level, storing the keys at the leaves of the tree. Assume there are n keys in the tree, where each key is an ordered pair of values. The worst case complexity to find a key is $O(\log(n))$.

   (h) For the master equation $T(n) = aT(n/b) + n^k$ with $\gamma = \log(a)/\log(b)$ the solution is always a sum of terms powers n^γ and n^k.

   (i) The adjacency matrix for a graphs $G(N, A)$ is always a symmetric matrix, i.e the same above and below the diagonal.

   (j) Given a sorted set of numbers the dictionary search will always perform have complexity $T(n) \in o(log(n))$.

   (k) If you have a minimum distance between (i,j) in an undirected graph $d(i, j) = d(i, k) + d(k, j)$ then either $d(i, k)$ or $d(k, j)$ is a minimum distance but not both.

(l) The best union-find algorithm can perform a sequence random sequence of n unions and m finds in $Tn, m) \in O(n + m)$.

(m) There are exactly 5 distinct binary search trees that include the numbers 1,2 anad 3.

(n) The Johnson All to All distance algorithm for $G(N, A)$ introduces an augmented graph with an extra node and potential function to remove all the negative cycles so that Dijkstras one to all can be used.

2. Consider a list of n elements with distinct values in the range $\{1, \ldots, n^2\}$. Identify which of the following sorting algorithms will produce a sorted list in worst case time $O(n \log n)$ (note I said $O()$, not $\Theta()$).

   (a) Merge sort

   (b) Insertion sort

   (c) Quicksort

   (d) Radix Sort

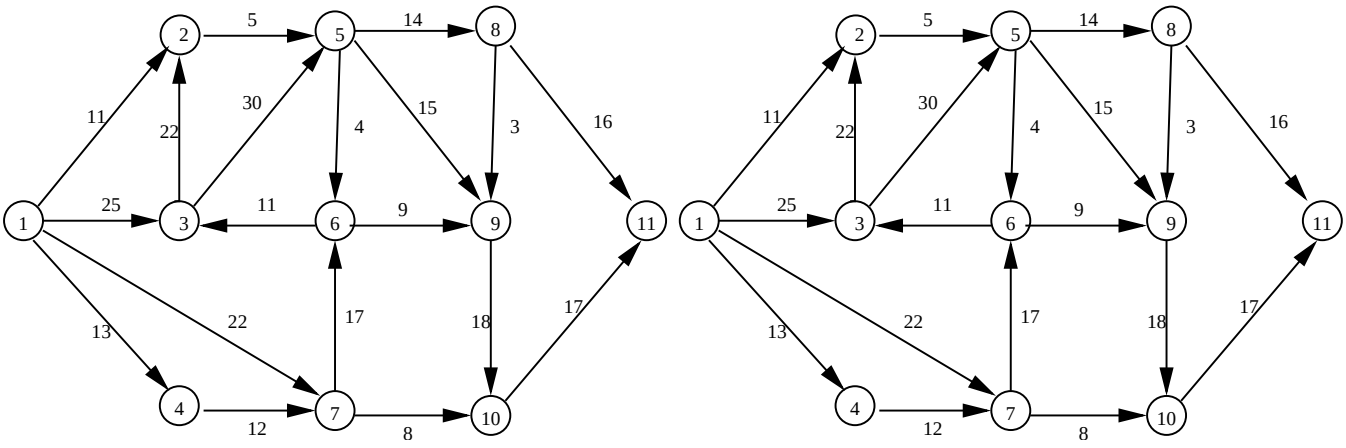
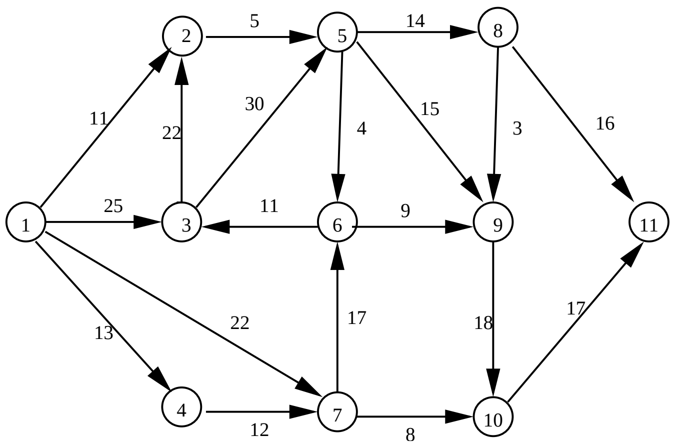


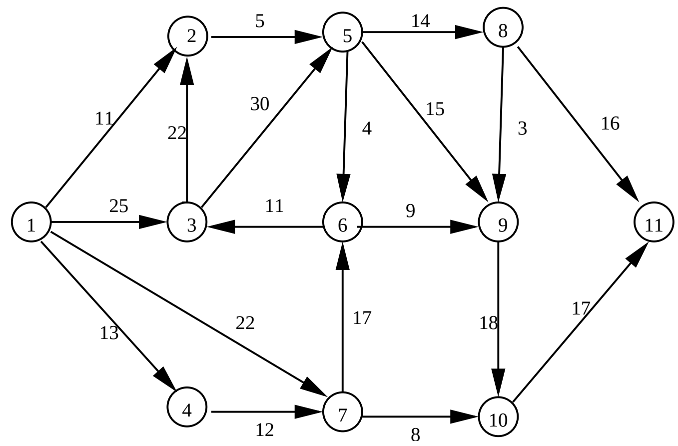
Figure 1:

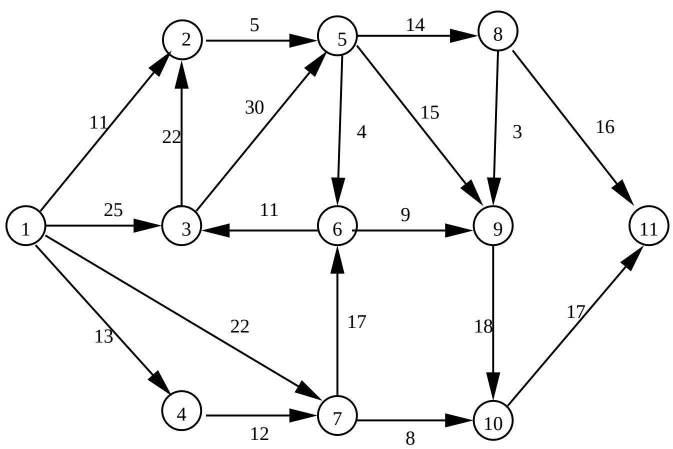
3. (15 pts) Consider Fig. 1 as a directed, capacitated graph, where the numbers on an arc now indicate an arc's capacity to carry flow from node 1 to node 11. In the max-flow algorithm of Ford and Fulkerson, the key step is, once a path has been found, to augment the flow and construct the residual graph for the next iteration.

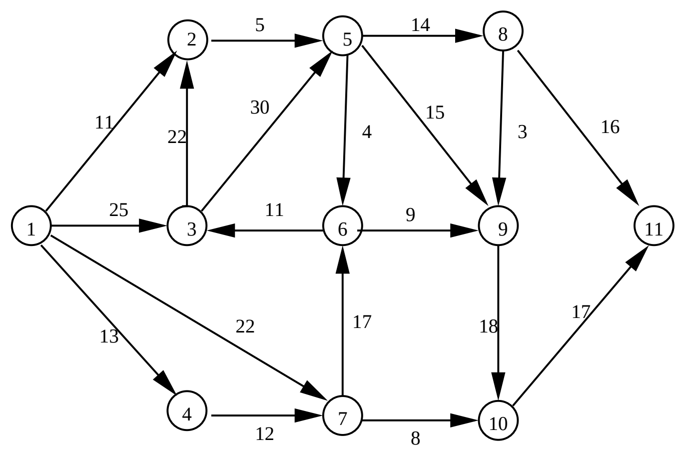
   (a) Suppose your program picks the first augmentation path to be $1 \to 4 \to 7 \to 6 \to 3 \to 2 \to 5 \to 8 \to 11$. Draw the augmented flow path on the left graph above and the residual graph above in the right side. What is the capacity of this path?

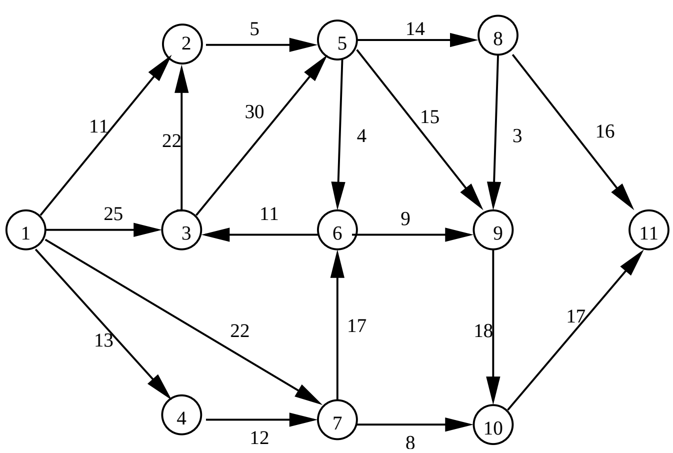
   (b) Now be smarter, starting with shorter paths, beginning with a min hop $1 \to 2 \to 5 \to 8 \to 11$ for the first path and $1 \to 4 \to 7 \to 10 \to 11$ for the second path. Proceeding to usng short augmentation paths to bring you to maximun flow. Draw **all** flow graphs and **all** the residual and after **each** augmentation. What is the maximum flow? Draw the minimum cut to show this right.

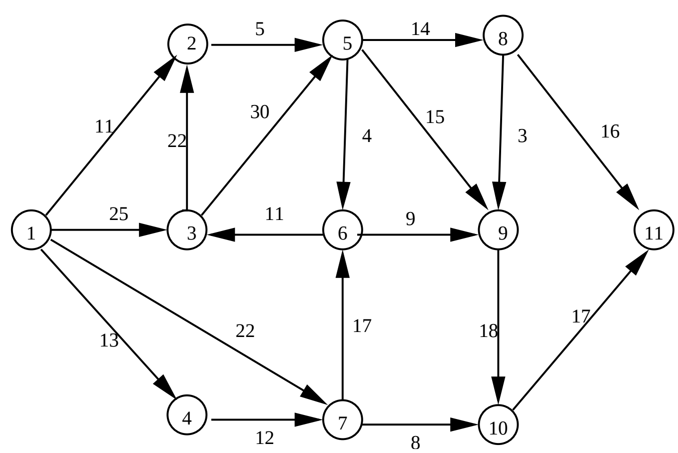

2
5
5
14
8
11
22
30
4
15
3
16
1
25
3
11
6
9
9
11
13
22
17
18
17
4
12
7
8
10

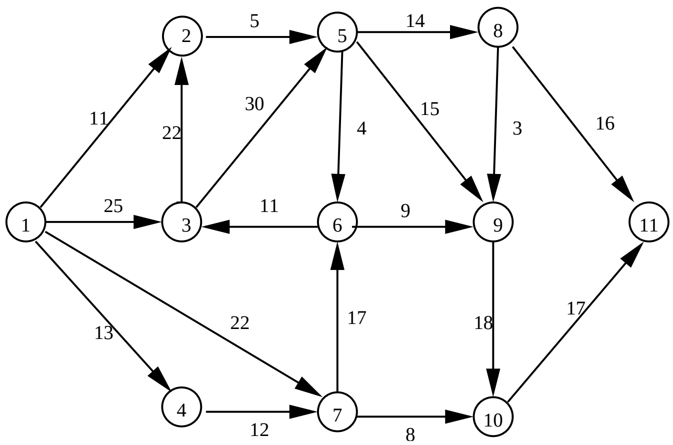

2
5
5
14
8
11
22
30
4
15
3
16
1
25
3
11
6
9
9
11
13
22
17
18
17
4
12
7
8
10

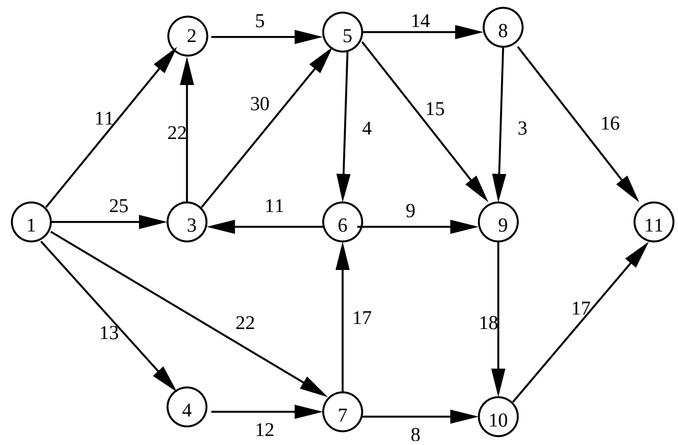

2
5
5
14
8
11
22
30
4
15
3
16
1
25
3
11
6
9
9
11
13
22
17
18
17
4
12
7
8
10

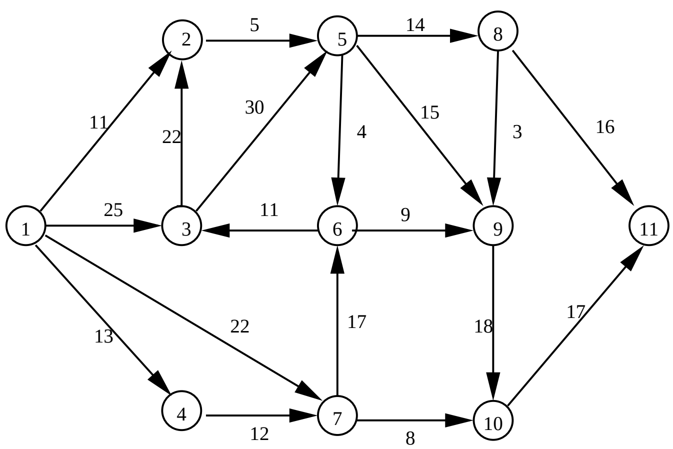

2
5
5
14
8
11
22
30
4
15
3
16
1
25
3
11
6
9
9
11
13
22
17
18
17
4
12
7
8
10

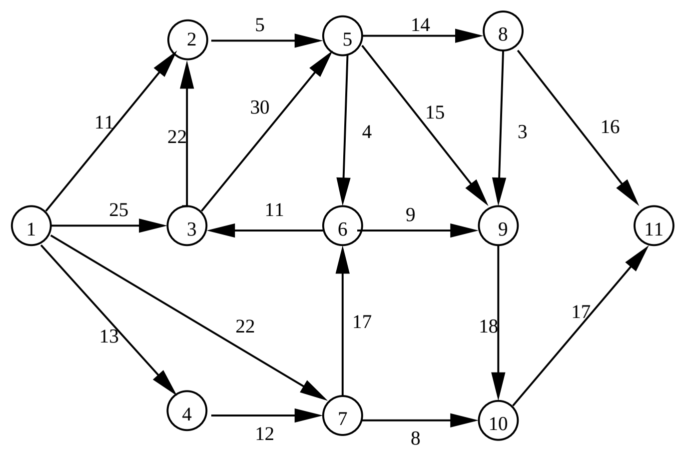

2
5
5
14
8
11
22
30
4
15
3
16
1
25
3
11
6
9
9
11
13
22
17
18
17
4
12
7
8
10

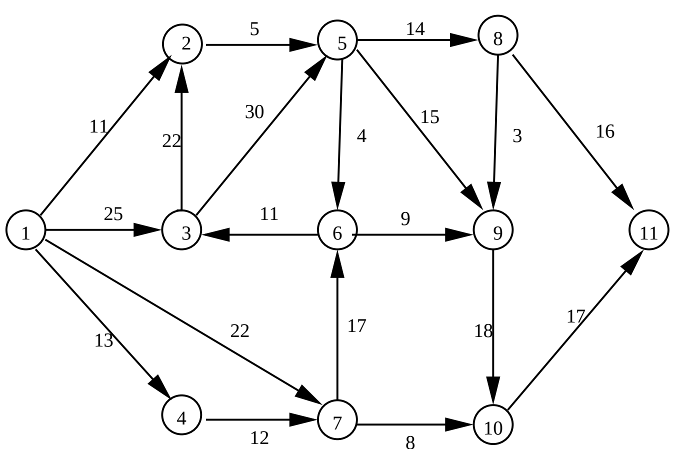

2
5
5
14
8
11
22
30
4
15
3
16
1
25
3
11
6
9
9
11
13
22
17
18
17
4
12
7
8
10

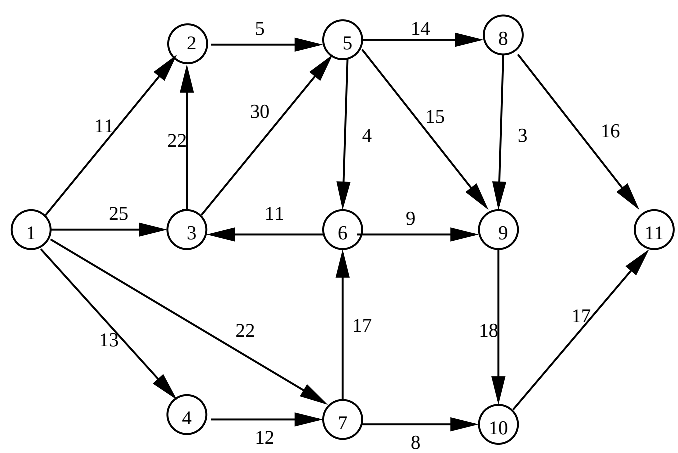

1
2
3
4
5
6
7
8
9
10
11
5
14
11
30
22
15
16
4
3
25
11
9
22
13
17
18
17
12
8


1
2
3
4
5
6
7
8
9
10
11
5
14
11
30
22
15
16
4
3
25
11
9
22
13
17
18
17
12
8


1
2
3
4
5
6
7
8
9
10
11
5
14
11
30
22
15
16
4
3
25
11
9
22
13
17
18
17
12
8


1
2
3
4
5
6
7
8
9
10
11
5
14
11
30
22
15
16
4
3
25
11
9
22
13
17
18
17
12
8


1
2
3
4
5
6
7
8
9
10
11
5
14
11
30
22
15
16
4
3
25
11
9
22
13
17
18
17
12
8


1
2
3
4
5
6
7
8
9
10
11
5
14
11
30
22
15
16
4
3
25
11
9
22
13
17
18
17
12
8

4. You are interested in compression. Given a file with characters, you want to find the binary code which satisfies the prefix property (no conflicts) and which minimizes the number of bits required. As an example, consider an alphabet with 8 symbols, with relative frequency of appearance in a text file give below [1]

alphabet:	D	E	M	O	C	R	A	T
frequency:	15	40	12	7	15	18	35	8

   (a) Determine the Huffman code by constructing a tree with **minimum external path length**. (Please use the "smaller weight to the left" convention.)

   (b) With 0 bit to the left and 1 bit to right, identity the code for each letter and list the number of bits for each letter. (Note: as you descend the tree the bits are code for a letter goes right to left!) **You may want draw first on extra paper and then copy it below to get it this standare the form.**

   (c) Compute the average number of bits per symbol in this code? Is it less than 3? (You can leave the answer as a fraction since getting the decimal value is difficult without a calculator.) Give an example of frequencies for these 8 symbols that would saturate 3 bits per letter?

[1] From the ancient Greek prefex: demotic =*common people* + suffex: -kratia = *strength* . In late 18th century (originally denoting an opponent of the *aristocrats* in the French Revolution of 1790): from French *democrate*, on the pattern of aristocrate *aristocrat*.

5. Consider the following knapsack problem. We have six objects with different values and Vi and sizes wi. The object values and sizes are listed in the table below:

Object number	1	2	3	4	5	6
Value	10	11	12	13	14	15
Size	1	1	2	3	4	3

(a) Suppose we are given a knapsack with total size 11, and you are allowed to use fractional assignments. What is the maximum value that fits in the knapsack for the above tasks?

(b) What is the maximum value of the tasks that fit entirely in the knapsack with size 11 — i.e. the integer knapsack.

6. (15 pts) Consider searching in the "text" $T(1:N)$ of length $N = 25$

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
b	a	c	b	b	a	c	b	a	b	a	c	b	a	b	a	c	b	a	c	b	a	b	a	c

for the following string of length $M = 8$ as an array $P(1:M)$ (i.e. $P(i), i = 1, 2, \cdots M$)

b	a	c	b	a	b	a	c

using the KMP algorithm.

(a) Give the prefix function for above string: That is the value of $\pi(i)$ for $i = 1, \cdots 8$. [2] It is useful to copy the pattern in $P(1:M)$ to slide it against $P(1:M)$ After a failure at $q+1$ you can safely shift by $\pi(q)$ before starting to search again. Note overlapping finds are ok!)

[2] **Recall that the prefex function looks at the first q values of $P(1:q)$ and ask how far you can shift to match to have largest prefi match the suffix in these q terms. $\pi(q) = MAX\{k < q$ such that$P(1:k) = P(1+q-k:q)\}$.**

(b) Find all the instances of this pattern in the text. That is give the start index in the text for the aligned pattern. (You can do this even if you have not got the right prefex function. Slide $P(1:M)$ against $T(1:N)$.)

(c) Specify all the non-trivial shift (i.e grater than 1 unit) that occurred in the scan using the KMP algorithm.

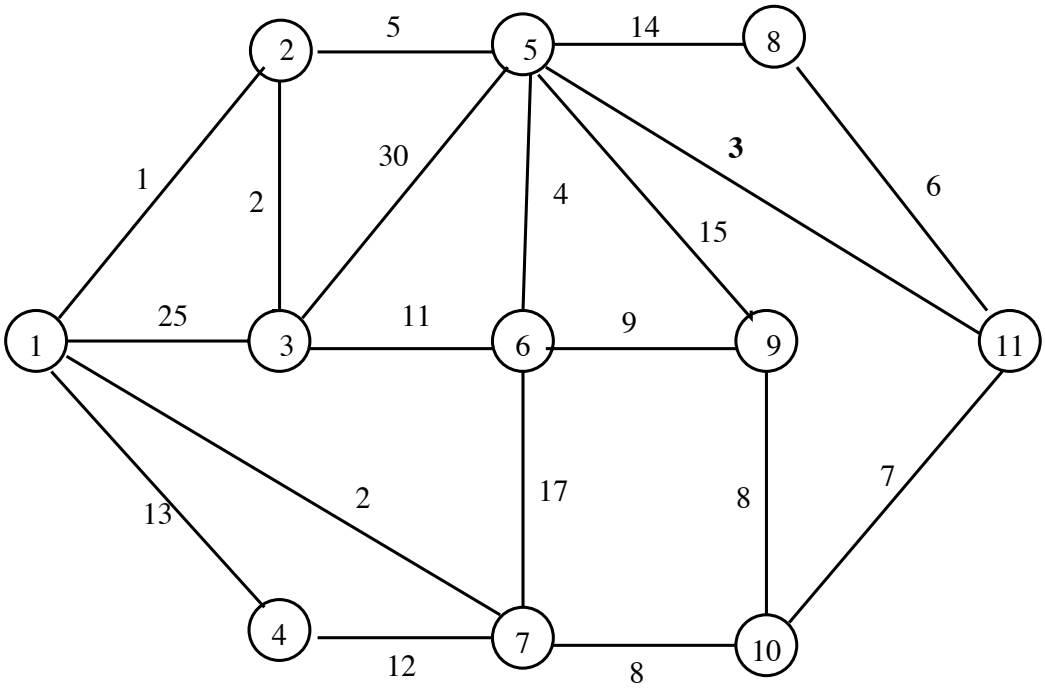


Figure 2:

7. Consider the undirected weighted graph in Figure 2. (There are copies of this graph at the end of the exam.) [3]

   (a) Use Dijkstra's algorithm to find the shortest paths from node 1 to all of the nodes. (Dijkstra's, like Prim's minimum spaning tree algorithm, adds one node at a time.)

   (i) Make a series of tables showing the values of the distance array $d(i)$ and the predecessor array $\pi(i)$ after each update

   (ii.) Draw the final tree on the figure with the final values of $d(i)$ and $p(i)$ at each node.

   (b) Repeat the exercise above except now use using Belllman-Ford this time (Bellman-Ford, like Kruskal's minimum spaning tree algorithm, adds one arc at a time.)

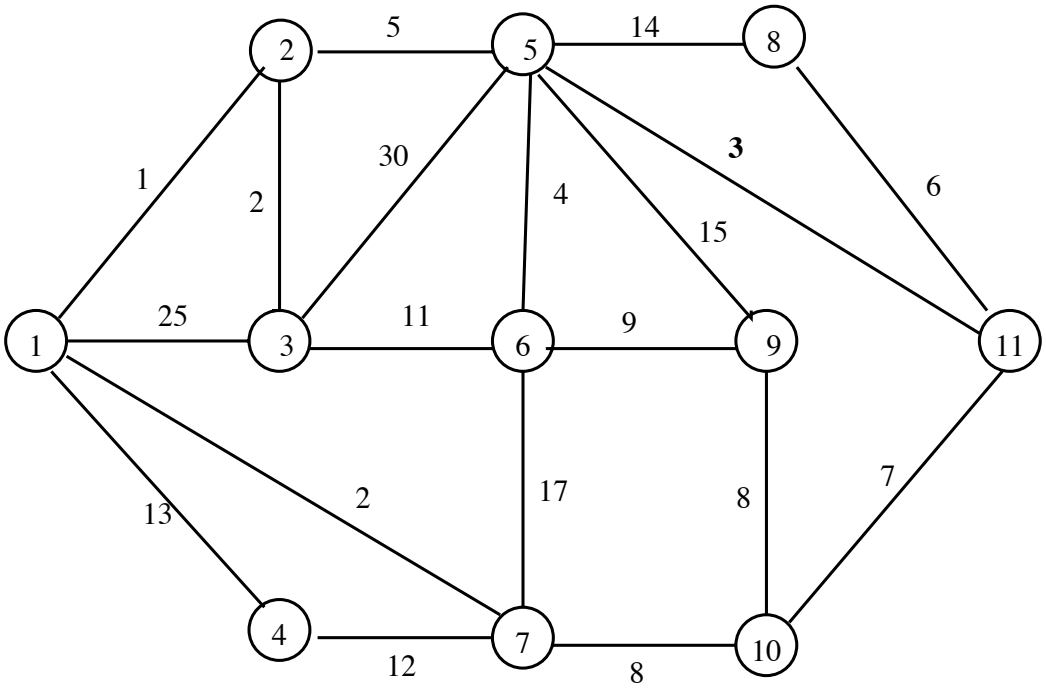
   (i.) Make a series of tables showing the values of the distance array $d(i)$ and the predecessor array $\pi(i)$ after each update.

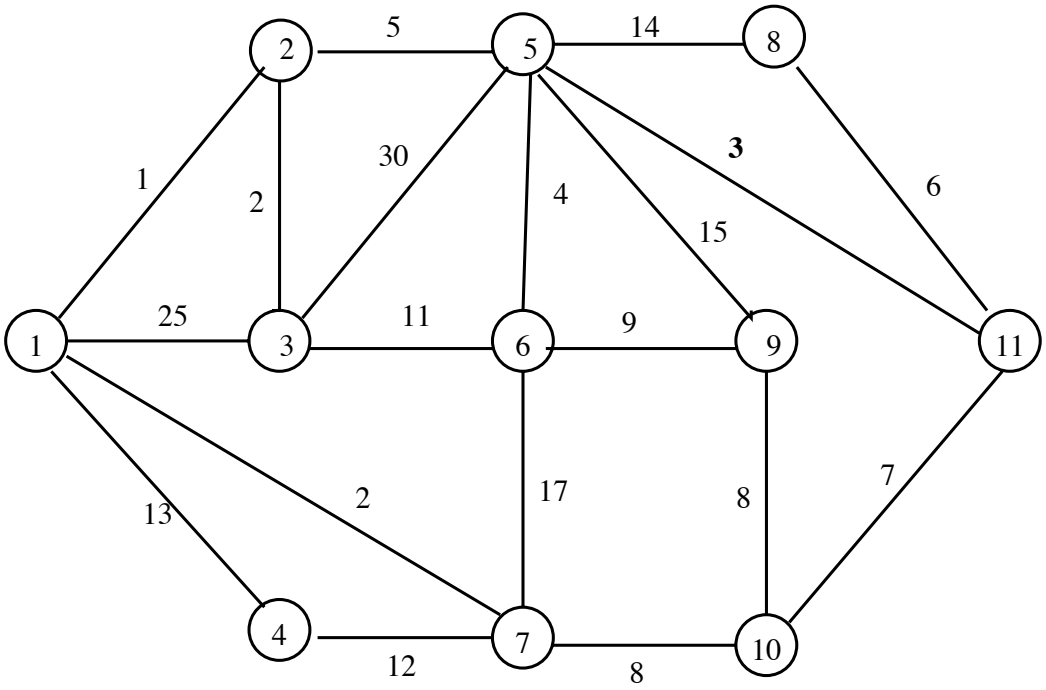
   (ii.) Draw the final tree on the figure with the final values of $d(i)$ and $p(i)$ at each node.

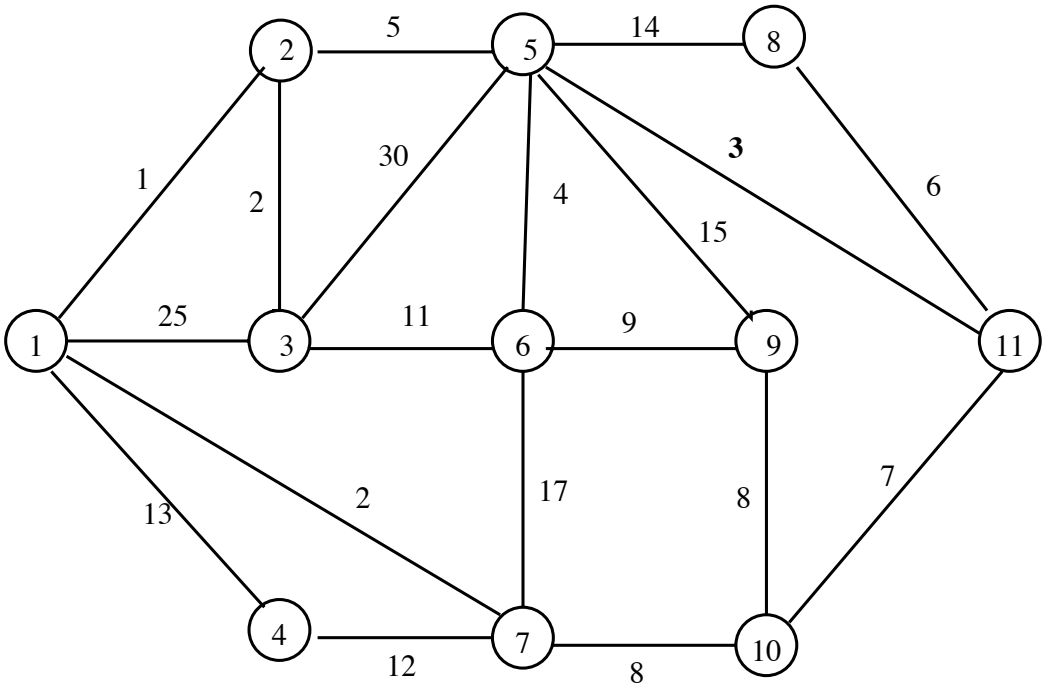
   (c) Computed the minimum spanning tree considering all arcs to be bi-directional (in spite of the figure!) using Prim's algorithm starting form node 1, listing in order the total weight and predecessors as they are modified at each step.

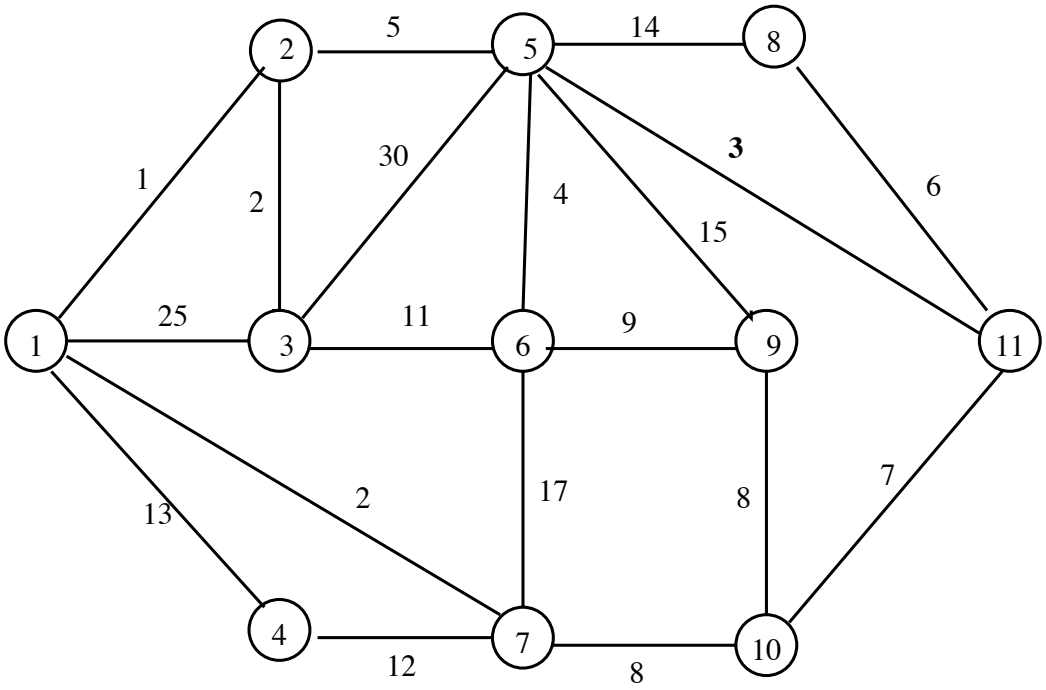
   (d) Are the final trees in all these cases the same? Explain

[3] **Note in part a and b below you start with $d(1) = 0$, $d(i > 1) = \infty$ and $\pi(i) = -1$. The table at each step only needs to show the values of $d(i)$ and $\pi(i)$ that change!**


1
2
3
4
5
6
7
8
9
10
11
5
14
1
2
30
3
4
6
15
25
11
9
13
2
17
8
7
12
8


1
2
3
4
5
6
7
8
9
10
11
5
14
1
2
30
3
4
6
15
25
11
9
13
2
17
8
7
12
8


1
2
3
4
5
6
7
8
9
10
11
5
14
30
3
6
1
2
4
15
25
11
9
13
2
17
8
7
12
8


1
2
3
4
5
6
7
8
9
10
11
5
14
30
3
6
1
2
4
15
25
11
9
13
2
17
8
7
12
8