Indiana University Bloomington

Fall-2023
CSCI-B505 / INFO-I500
APPLIED ALGORITHMS
Examination – I

November 07, 2023, Tuesday, 6:30 p.m. – 7:30 p.m.

Name & Surname	
University ID	
Signature	

Rules:

1. There are 25 questions in this examination, each question carries 4 marks.
2. Duration of the exam is 60 minutes.
3. Write your name and surname on every page at the designated positions.
4. Put your ID card on your desk so that the proctors can check your identity.
5. The use of lecture notes, books, and any other resources, calculators, computers, mobile phones, and any digital equipment is prohibited.
6. Every student taking this examination is subject to the university discipline code. Any act or attempt of cheating, including helping others, will be considered a violation of the code.

Name & Surname:

1. What is the stopping condition of up-heap bubbling during the insertion of a value x into a max-heap when maintaining the max-heap property?

   (a) When all of the child values are greater than x.

   (b) When all of the child values are less than x.

   (c) When the parent value is greater than x. ✓

   (d) When the parent value is less than x.

2. Consider a scenario where a given stream of numbers [4, 1, 3, 2, 16, 9, 10] is being inserted one by one into an initially empty **max-heap**. Starting with the insertion of 4, followed by 1, and so forth. After every insertion, the heapify process is performed to ensure that the max-heap property is satisfied. Once all the elements are inserted into the max-heap, how many swaps are performed in total?

   (a) 5 ✓

   (b) 6

   (c) 7

   (d) 8

3. Using the array of values given below, perform bottom-up heap construction to create a **min-heap**.

   $$[8,5,9,1,3,6,7]$$

   Which of the following is the resulting array after heap construction?

   (a) $[1,3,6,5,8,9,7]$ ✓

   (b) $[1,3,6,5,7,9,8]$

   (c) $[1,3,5,6,7,8,9]$

   (d) $[1,3,5,6,8,9,7]$

4. In a streaming scenario, where continuous data points are received, which of the following options is the most efficient data structure to keep track of the top k largest elements?

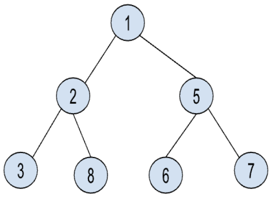
   (a) A sorted array that is updated as new elements arrive.

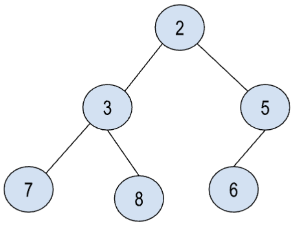
   (b) A min-heap with a fixed capacity of k elements. ✓

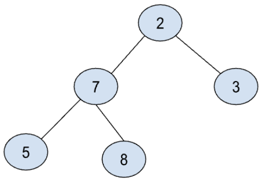
   (c) A self-balancing binary search tree storing all elements.

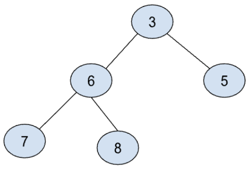
   (d) A max-heap with a fixed capacity of k elements.

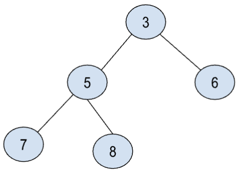
5. Which of the following represents a min-heap after popping two minimum elements from the below given min-heap?



(a) 

(b) 

(c)  ✓

(d) 

6. Suppose we have two max-heaps each of size n. What will be the time complexity to make a new max-heap of size $2n$ from the elements of the two max-heaps?

(a) $O(nlog(n))$

(b) $O(nlog^2(n))$

(c) $O(n)$ ✓

(d) $O(n^2)$

7. Consider the following table representing the weights and values of several items, each with a unit quantity:

Item Number	Weight (Kgs)	Value
1	6	3
2	5	2
3	3	1
4	4	4

The problem at hand is to select a subset of these items such that the total weight does not exceed 10 Kg, while maximizing the overall value. Additionally, the constraint specifies that no item can be divided. The total value of the items chosen by an optimal algorithm is denoted as V_{opt}. On the other hand, a greedy algorithm arranges the items based on their weight-to-value ratios in descending order and packs them in a greedy manner, starting from the first item in the ordered list. The total value of the items selected by the greedy algorithm is denoted as V_{greedy}. What is the value of $V_{opt} - V_{greedy}$?

(a) 2

(b) 3

(c) 4 ✓

(d) 5

8. If frequencies for symbols q, w, e, r, and t are 5, 2, 4, 5, and 3 respectively, and 0 is assigned to symbols on the left side of the partition while 1 is assigned to symbols on the right side, then what will the string ”qwert” be encoded as using Fano coding?

(a) 001111011001

(b) 011111000110

(c) 001111001110 ✓

(d) 001111001101

9. Consider the letters A, L, G, O, E, and S with frequencies 4, 3, 7, 2, 6, and 11 respectively. Using these frequencies, perform Huffman coding to generate codewords for each letter.

Now encode the string ’ALGOESE’ using the Huffman codes. Let Y represent the total number of bits in the encoded string.

What is the value of Y?

Y = **19**

10. Which of the following statements about Huffman encoding is true?

(a) Huffman encoding requires knowledge of symbol frequencies.

(b) Huffman encoding always results in the smallest possible encoded message

(c) Huffman coding is a greedy algorithm.

(d) All of the above. ✓

11. In the context of greedy algorithms, which of the following statements is false?

    (a) A greedy algorithm makes the best possible choice at each step.

    (b) A greedy algorithm guarantees an optimal solution for all problems it solves. ✓

    (c) A greedy algorithm never reconsiders previous steps, even if they lead to a suboptimal solution.

    (d) All of the above.

12. Consider a college with a set of classes to be scheduled and multiple classrooms available. The classes have the following start and end times:

    [[7, 10], [9, 12], [11, 16], [13, 16], [15, 17], [10, 12], [14, 15], [6, 8]]

    Where each inner list represents the start and end time of one class.

    If the college uses a greedy algorithm that assigns conflicting classes to different classrooms, what is the minimum number of classrooms needed to schedule all the classes without any time overlap?

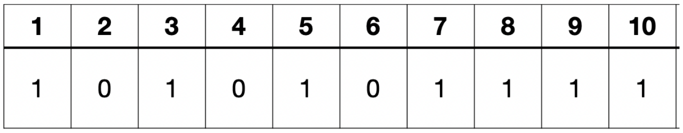
    (a) 2

    (b) 3 ✓

    (c) 4

    (d) 5

13. For the array shown in the figure below, how would we find the predecessor and successor of 6?



1	2	3	4	5	6	7	8	9	10
1	0	1	0	1	0	1	1	1	1

    (a) select(rank(6)), select(rank(6) + 1) ✓

    (b) select(rank(6) - 1), select(rank(6)) + 1

    (c) select(rank(6) + 1), select(rank(6) + 1)

    (d) select(rank(6)), select(rank(6)) + 1

14. We wish to perform radix sort on the numbers [623, 108, 271, 400, 574, 891] to order them in ascending order. After the second pass, what number is second on the list, from the top?

    (a) 108 ✓

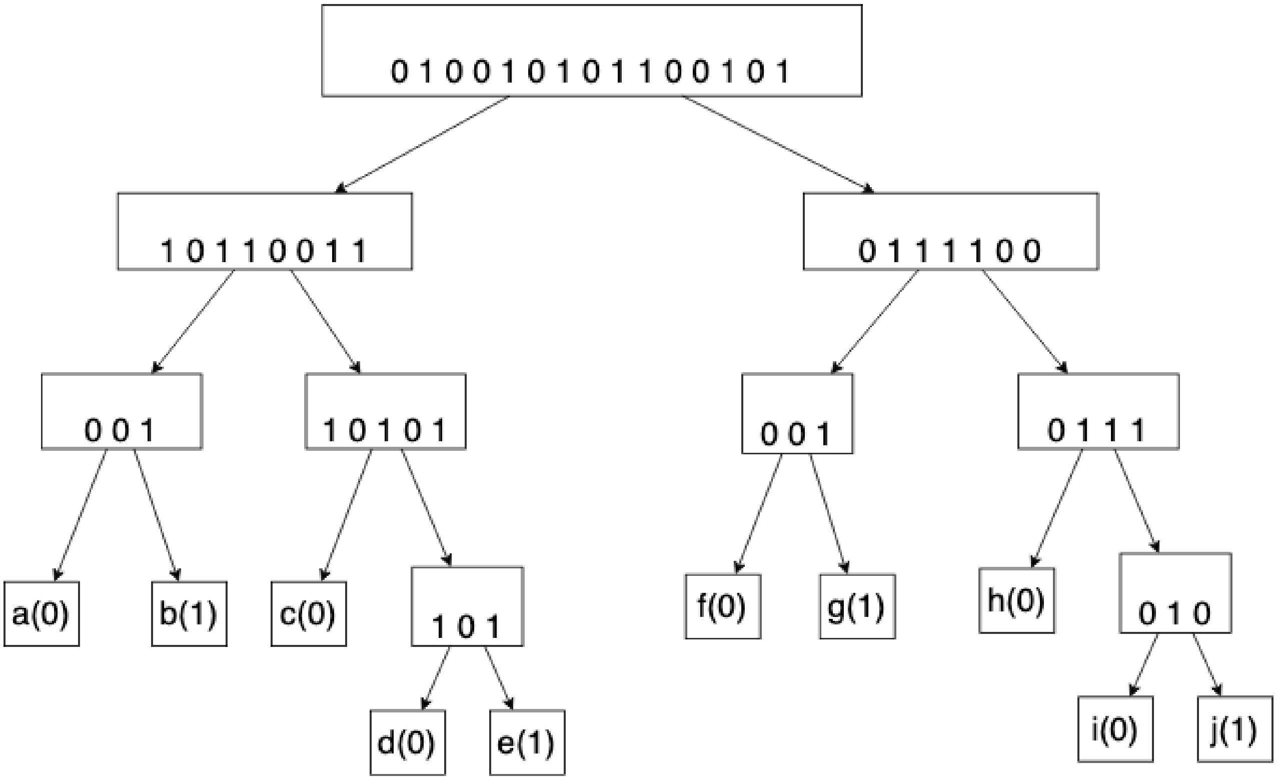
    (b) 271

    (c) 400

    (d) 623

Answer the next 2 questions based on the below information

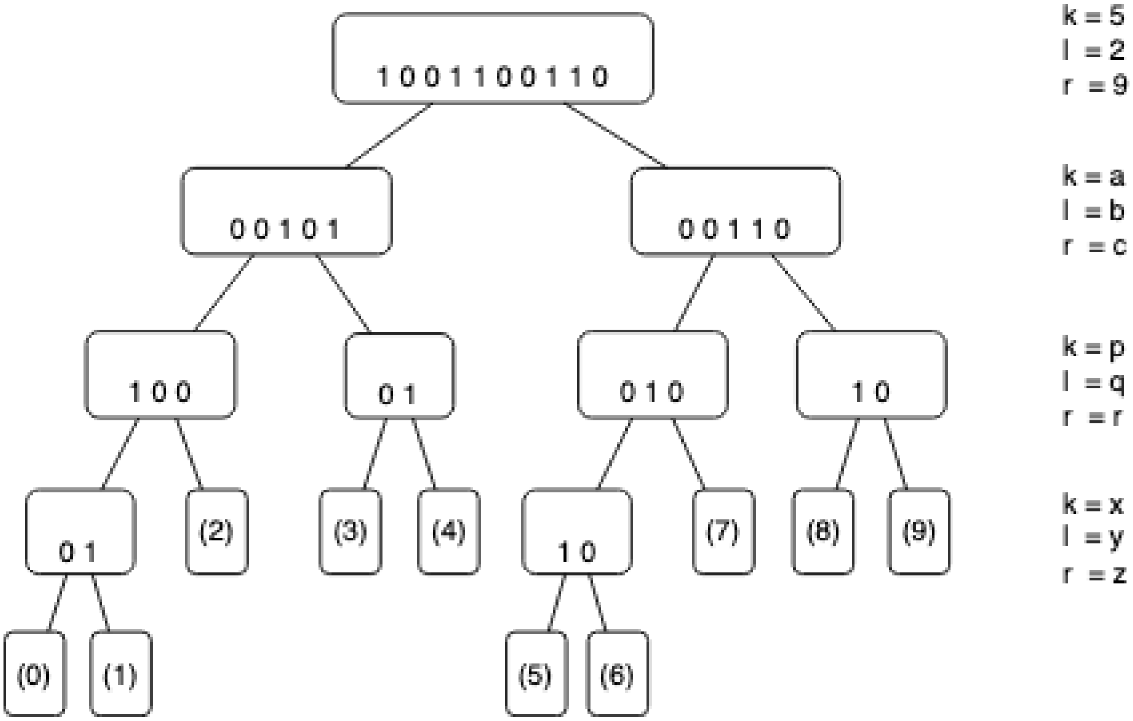
Consider the below wavelet tree constructed for a sequence S[1..15] (i.e **1 based indexing**), where the symbols of S are from the alphabet A = a, b, c, d, e, f, g, h, i, j. At each node, the alphabet is split into two parts following the lexicographical order, e.g., on the root bitmap 0s represent a,b,c,d,e and 1s represent f,g,h,i,j. The left sub-tree is represented by 0's and the right sub-tree is represented by 1s.



15. What is $rank(i, 8)$?
    - (a) 1 ✓
    - (b) 2
    - (c) 3
    - (d) 4

16. What is $access(6)$?
    - (a) i
    - (b) c
    - (c) d ✓
    - (d) a

17. When using the Range Quantile Query (RQQ) method to find the 5th smallest element within the range S[2..9] in the given wavelet tree, assuming a 1-based indexing for all indices and the left sub-tree is represented by 0's and the right sub-tree is represented by 1s,



What are the values of p, q, and r in the above figure?

(a) 5, 2, 9

(b) 1, 2, 2

(c) 1, 2, 3 ✓

(d) 1, 2, 5

18. What does the stable sorting refer to?

(a) Time complexity remains the same irrespective of whether the input is sorted or not

(b) Time complexity remains constant up to a certain input size and beyond that input size the time complexity goes exponential

(c) The order of the equal-valued elements on the input is maintained after the sorting. ✓

(d) It is proven in theoretical computer science that the upper bound for some algorithms does not exist. These algorithms are called stable algorithms

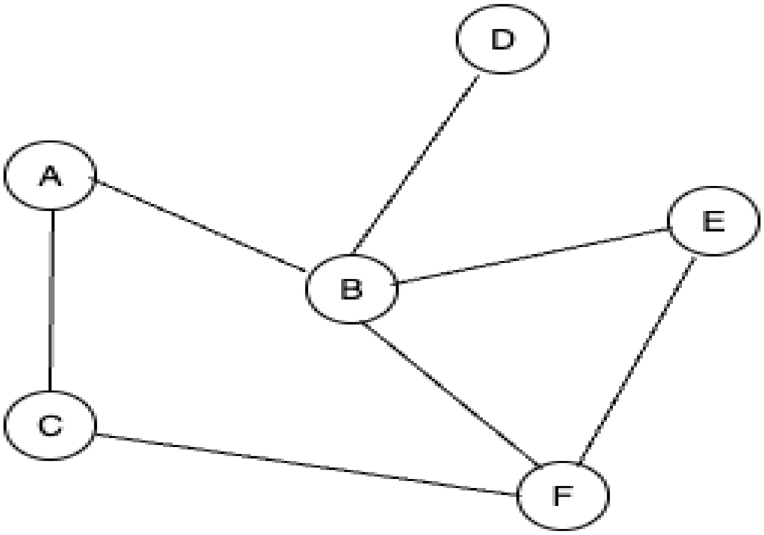
19. Which of the following is an advantage of using an adjacency list instead of an adjacency matrix for graph representation?

    (a) Edge insertion is faster
    (b) Edge removal is faster
    (c) Time taken to check for the presence of an edge is lesser
    (d) Auxiliary space complexity is lesser ✓

20. What is the maximum number of edges that can exist in a complete graph with n vertices?

    (a) (n*(n+1))/2
    (b) (n*(n-1))/2 ✓
    (c) N
    (d) None of the Above

21. Based on the provided graph, please choose the correct sequences for both Breadth-First Search (BFS) and Depth-First Search (DFS) traversals. In each option, the first mentioned traversal is BFS, and the second is DFS. Assume that the initial node specified in your option serves as the starting point for both BFS and DFS.



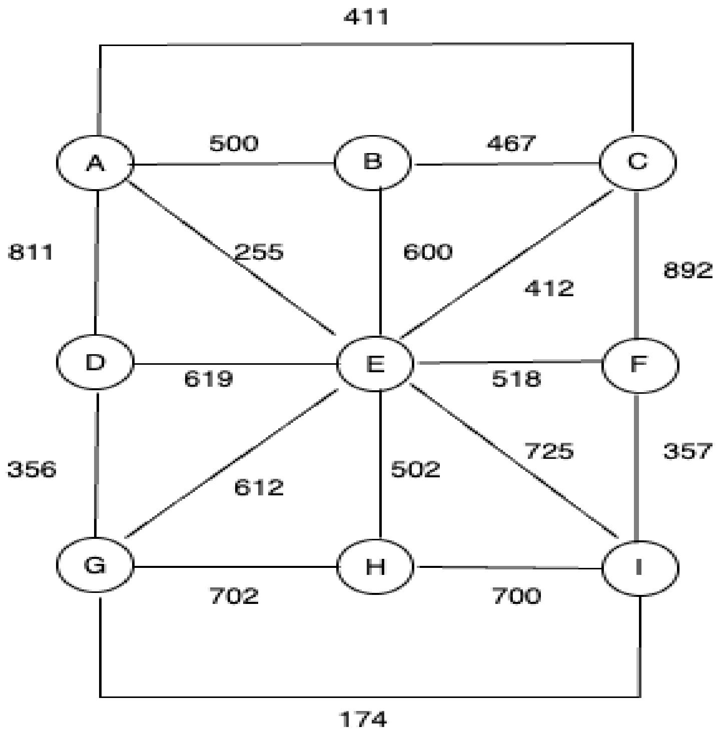
    (a) A - B - C - D - E - F
        A - B - F - D - C - E

    (b) D - B - E - F - C - A
        A - B - F - D - C - E

    (c) B - D - A - E - C - F
        A - B - F - D - C - E

    (d) A - B - C - D - E - F
        A - C - F - B - D - E ✓

Answer the next two questions based on the below graph



22. Based on the given weighted graph and the Kruskal Algorithm, determine the correct order of edges that are added when constructing the Minimum Spanning Tree (MST).

    (a) GI - AE - GD - IF - AC - CE - BA - EH

    (b) GI - AE - GD - IF - AC - CE - CB - BA

    (c) GI - AE - GD - IF - AC - CB - EH - EF ✓

    (d) GI - GD - IF - FE - EA - CE - CA - AD

23. Consider the given weighted graph. If we start Prim's algorithm at vertex A, which edge will be the third to be added to the spanning tree?

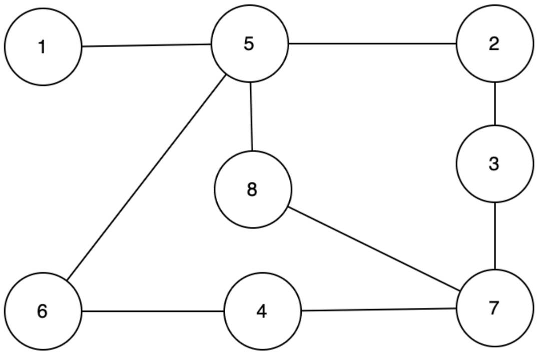
    (a) The edge BC ✓

    (b) The edge EH

    (c) The edge IF

    (d) The edge AE

24. Consider the below undirected graph. While using the Greedy graph coloring heuristic, How many number of colors do you need if you visit the nodes in the order 1,2,3,4,5,6,7,8 such that no two adjacent vertices share the same color?



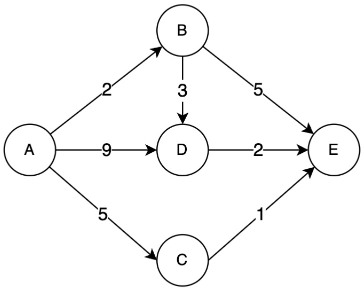
(a) 2

(b) 3

(c) 4 ✓

(d) 5

25. Consider the below-weighted graph with vertices A, B, C, D, and E. Fill in the D[] values after each iteration when applying Dijkstra's algorithm, starting from vertex A.



	A	B	C	D	E
Iteration 0	0	∞	∞	∞	∞
Iteration 1	0	2	5	9	∞
Iteration 2	0	2	5	5	7
Iteration 3	0	2	5	5	6
Iteration 4	0	2	5	5	6