**Indiana University Bloomington**

**Spring-2023**
**CSCI-B505 / INFO-I500**
**APPLIED ALGORITHMS**
**Examination – IV**

**April 27, 2023, Thursday, 9:45 a.m. – 10:45 a.m.**

| | |
|---|---|
| Name & Surname | |
| Signature | |

**Rules:**

1. There are 25 questions in this examination.
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 identities.
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 number of edges present in a complete graph having $n$ vertices?

   (a) $(n \cdot (n+1))/2$

   (b) $(n \cdot (n-1))/2$ ✓

   (c) $2n$

   (d) $n \cdot (n+1)$

2. What is the time complexity associated with the getedge(u, v) operation when we use an adjacency matrix for a graph of **n** vertices and **m** edges?

   (a) O(m) (b) O(n) (c) O(1) ✓ (d) O(max(m,n))

3. There is an algorithm **X**. The input for **X** has three types **A**, **B**, **C**. Probability that input happens to be **A** is 0.5, **B** is 0.3 and **C** is 0.2. The cost of running algorithm **X** if input is **A** is 1 if input is **B** is 2, if input is **C** is 3. What is the average-case cost of the algorithm ?

   (a) 1.8 (b) 1.2 (c) 1.5 (d) 1.7 ✓

4. We would like to verify whether A × B = C, where all are n × n matrices. If we run the Freivalds algorithm three times, what is the probability that the Freivalds algorithm says that the A × B = C is true when in reality is A × B = C is false.

   (a) 0.25 (b) 0.5 (c) 1 (d) 0.125 ✓

5. A Hash table uses hashing function h(k) = k mod 10 and linear probing. A few insertions are made to the table, as seen in the image below. What could have been the order in which these insertions been made?

| | |
|---|---|
| 0 | |
| 1 | |
| 2 | 42 |
| 3 | 23 |
| 4 | 34 |
| 5 | 52 |
| 6 | 46 |
| 7 | 33 |
| 8 | |
| 9 | |

   (a) 46, 42, 34, 52, 23, 33

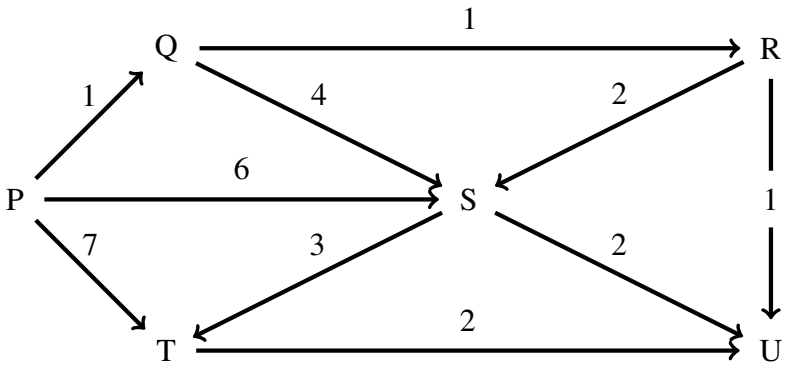
   (b) 34, 42, 23, 52, 33, 46

   (c) 46, 34, 42, 23, 52, 33 ✓

   (d) 42, 46, 33, 23, 34, 52

**Name Surname:**

6. In what order are the vertices visited for the following graph, if we employ Djikstra's algorithm?

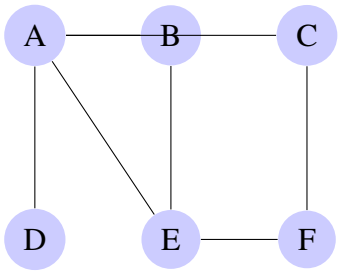



   (a) P Q R S T U
   (b) P Q R U S T ✓
   (c) P Q R U T S
   (d) P Q T R U S

7. $G$ is bipartite graph with $n$ vertices . What is the maximum possible number of edges $G$ can have?

   (a) $n-1$ (b) $n/2$ (c) $n^2$ (d) $n^2/4$ ✓

8. The Breadth First Search algorithm has been implemented using the queue data structure. One possible order of visiting the nodes of the following graph is:




   (a) EABDFC
   (b) ABCDEF
   (c) BEAFCD
   (d) EABFDC ✓

9. Which of the following algorithms can be used to most efficiently determine the presence of a cycle in a given graph?

   (a) Prim's algorithm
   (b) Kruskal's algorithm
   (c) Depth first traversal✓
   (d) Misra-Gries Algorithm

**Name Surname:**

10. What is the effect of increasing the number of hash functions used in a Bloom Filter?

    (a) The probability of false positives increases.

    (b) The probability of false positives decreases. ✓

    (c) The time complexity of search decreases.

    (d) The search time of the filter decreases.

11. In a Bloom filter, how many hash functions should be used to keep the false alarm probability at 1/64 ?

    (a) 6 ✓ (b) 9 (c) 5 (d) 7

12. There is a hash table with size 5. What is the probability of no collision after inserting 5 elements?

    (a) $\frac{1}{5^5}$ (b) $\frac{5}{5^5}$ (c) $\frac{5!}{5^5}$ ✓ (d) $\frac{1}{5^6}$

13. Which of the following is an advantage of adjacency list representation over adjacency matrix representation of a graph?

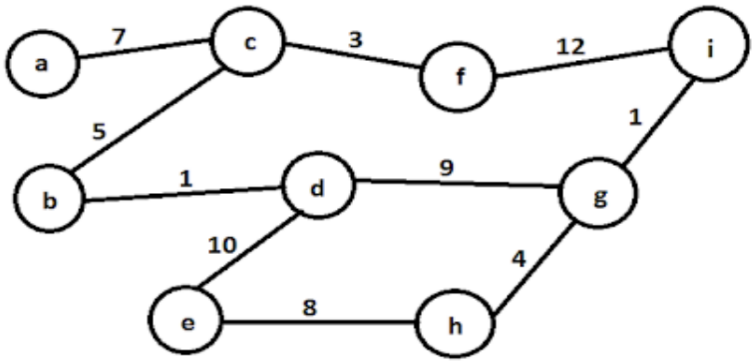
    (a) In adjacency list representation, space is saved for sparse graphs.

    (b) DFS and BFS can be done in $O(V+E)$ time for adjacency list representation. These operations take $O(V^2)$ time in adjacency matrix representation. Here V and E are the number of vertices and edges respectively.

    (c) Adding a vertex in adjacency list representation is easier than adjacency matrix representation.

    (d) All of the above✓

14. If Prim's algorithm were to be applied to the graph below, which will be the 5th vertex to be added to the spanning tree being formed by starting from the vertex 'c'? Note: vertex 'c' is being considered as the first node.




    (a) g (b) a ✓ (c) i (d) e

**Name Surname:**

15. In Bloom Filters, how many hash functions are required if the required false alarm probability is (1/32), and how many bits per element are required respectively?

(a) 5, 7.2 ✓ (b) 6, 10 (c) 5, 10 (d) 6, 7.2

16. Select one element RANDOMLY from the streaming sequence. Use the values which are random values in the range [1..100] for the probability generation such that an event with probability $p$ is assumed to have happened if $r/100 \leq 100p$, for the corresponding random number $r$ given in the table, where sampling is performed when that condition is met. What will be the final element after the streaming sequence is passed?

| Index | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Sequence | 47 | 23 | 12 | 49 | 7 | 3 | 22 | 31 | 41 | 1 |
| Random Number | - | 65 | 45 | 21 | 15 | 87 | 95 | 20 | 10 | 12 |

(a) 1 (b) 41 ✓ (c) 31 (d) 49

17. Given the streaming data S = 1,1,2,3,4,5,1,1,1,5,3,3,1,1,2,5,3,2,3,3. Use the Misra-Gries algorithm to find the final state of the counters and the 0.45 heavy hitters given the $\varepsilon = 0.25$.
Note: {Cx: 1, Cy:2} represents that there are 2 counters Cx and Cy, where the element x has count 1 and the element y has count 2.

(a) {C1: 6, C5: 2, C3: 5, C2: 2}, (1, 3) ✓
(b) {C1: 6, C5: 2, C3: 5, C2: 2}, (1, 2, 3, 5)
(c) {C1: 6, C5: 2, C3: 5, C4: 1}, (1, 3)
(d) {C1: 7, C5: 3, C3: 6, C2: 3}, (1, 3)

18. After running the count min sketch in a distributed environment we got the following matrices. Which of the following is a 0.4 heavy hitter? The functions used are,

$$h_1(x) = 2x + 9 \bmod 5$$

$$h_2(x) = 3x + 7 \bmod 5$$

$$h_3(x) = 7x + 3 \bmod 5$$

$$\begin{bmatrix} 2 & 1 & 0 & 1 & 1 \\ 1 & 2 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 3 & 0 \\ 1 & 0 & 0 & 3 & 1 \\ 0 & 1 & 3 & 1 & 0 \end{bmatrix}$$

(a) 7 ✓ (b) 4 (c) 6 (d) 3

19. The maximum degree of any vertex in a simple graph with $n$ vertices is

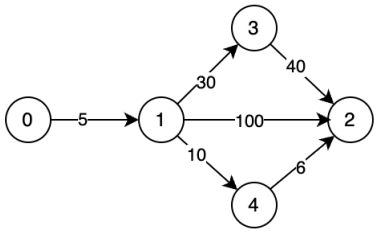
(a) n - 1 ✓ (b) n + 1 (c) 2n - 1 (d) n

20. For perfect hashing using h(k) = k mod 5 as the hash function at level 1. What should be the size of level 2 for the given sequence 15, 2, 1, 5, 20, 31, 12, 13, 34.

(a) 18 (b) 19 ✓ (c) 20 (d) 21

**Name Surname:**

21. Using Dijkstra's single source shortest path for the source node **0** what is the distance array (D) **after two iterations** for all the nodes of the algorithm for the following graph.




(a) $[0, \infty, \infty, \infty, \infty]$

(b) $[0, 5, 21, 35, 15]$

(c) $[0, 5, 105, 35, 15]$ ✓

(d) $[0, 5, 100, 20, 10]$

22. For the given psuedo code what is the probability of not finding the element?

```
Algorithm FindElement(e, A):
    j <- 1
    repeat
        idx <- rand(0, len(A));
        result <- A[idx];
        j <- j + 1
    until (result != e) or (j = k + 1)
    return result
```

(a) $(1/n)^k$ (b) $((n-1)/n)^k$ ✓ (c) 1/n (d) 1 - 1/n

23. While estimating the number of distinct integers in a stream we observed that the ratio of minhash to the hash size as 0.1. What is the estimated number of distinct values

(a) 8 (b) 9 (c) 10 ✓ (d) 11

24. Which of the following statements is TRUE regarding the Misra-Gries algorithm?

(a) Misra-Gries algorithm provides an exact count of distinct elements in a streaming data.

(b) Misra-Gries algorithm maintains a fixed-size sample of elements from the stream.

(c) Misra-Gries algorithm is used for finding the median element in a streaming data.

(d) Misra-Gries algorithm provides an approximate count of frequent elements in a streaming data.✓

25. What is the number of elements in the stream if the following array is obtained during count min sketch

$$\begin{bmatrix} 2 & 2 & 2 & 2 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 3 & 1 & 1 & 1 & 3 \\ 1 & 1 & 3 & 3 & 1 \end{bmatrix}$$

(a) 7 (b) 8 (c) 9 ✓ (d) 10