UNIVERSITY OF MORATUWA

FACULTY OF ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

BSc Engineering Honours Degree

CS2022 & CS2023: DATA STRUCTURES & ALGORITHMS

Time allowed: 2 Hours September 2022

ADDITIONAL MATERIAL: None

INSTRUCTIONS TO CANDIDATES:

1. This paper consists of three (3) parts in 12 pages.
2. Answer all questions in the question paper itself in the space provided. If you make a mistake or need additional space, you may attach additional sheets.
3. Part A contains ten (10) MCQ questions.
4. For each MCQ question in part A, there is only one correct answer, and you are expected to clearly mark only one choice. Each question in part A carries 1 mark.
5. Part B contains seven (7) short answer questions.
6. Part C contains ten (10) essay type questions.
7. The maximum attainable mark for each question in part B and C is given in brackets.
8. The maximum attainable mark for this paper is 100. This examination accounts for 60% of the module assessment.
9. Only calculators approved and labelled by the Faculty of Engineering are permitted.
10. This is a an open for limited content examination. You are authorized to bring 2 (two) A4 page (both side) to the examination and refer to it during the examination if required.
11. Assume reasonable values for any data not given in or with the examination paper. Clearly state such assumptions made on the script.
12. In case of any doubt as to the interpretation of the wording of a question, make suitable assumptions and clearly state them on the script.
13. This paper should be answered only in English.

PART A

1. Assume there is an empty stack. Choose correct output for the following sequence of stack operations.

   push (5), push (8), pop, push (2), push (5), pop, push (1)

   a. 8,5,2,5,1
   b. 8,5,5,21
   c. 2,5,5,1
   d. 5,2,1

2. If $f(n) = (n-2) * (n-4)$, then to which family does the function $f(n)$ belong?
   a. θ (8)
   b. θ (log n)
   c. θ (n)
   d. θ (n^2)

3. Let $G = (V, E)$ be the undirected graph whose edges are $\{(a,b),(b,c),(c,d),(a,d)\}$. If T is the depth-first spanning tree of G rooted at vertex a, it follows that T has __________ branch(es).
   a. 1
   b. 2
   c. 3
   d. 4

4. When does the insertion sort a good choice for sorting an array?
   a. Each component of the array requires a large amount of memory.
   b. Each component of the array requires a small amount of memory.
   c. The array has only a few items out of place.
   d. The processor speed is fast.

5. Merge sort makes two recursive calls. Which statement is true after these recursive calls finish, but before the merge step?
   a. The array elements form a heap.
   b. Elements in each half of the array are sorted amongst themselves.
   c. Elements in the first half of the array are less than or equal to elements in the second half of the array.
   d. None of the above.

6. "Linked list is a method to overcome the problem of array that cannot allow alteration on it once the program is running." Select a correct fact about linked list from the following to justify this statement.
   a. **A linked list** can be used to manage a dynamically growing and shrinking list of data.
   b. A linked list is used to chain a collection of data.
   c. A linked list reallocating the array as needed by adding elements at a time.
   d. A linked list is slow, in efficient and expensive to use.

7. Consider the following statements:
   i. Any comparison-based sorting algorithm requires $\Omega(n \log n)$ swaps.

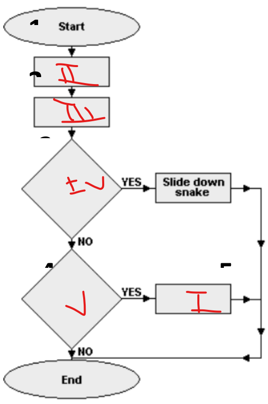
ii. If all edge weights of an undirected connected graph are either 0 or 1, then shortest path between any two nodes can be found in $O(E)$ time, where E is the number of edges in the graph.

Which of the above statement/s is/are TRUE?

a. (i) only
b. (ii) only
c. Both (i) and (ii) are true
d. Both (i) and (ii) are false

8. Choose the correct statement(s) about tree data structure.
   a. Given a set S of n real keys chosen at random from a uniform distribution over $(a: b)$ a binary tree can be structured on S in $O(n)$ expected time.
   b. In the worst case, a red-black tree insertion requires $O(1)$ rotations.
   c. Given a connected, weighted, undirected graph G in which the edge with minimum weight is unique, that edge belongs to every minimum spanning tree of G.
   d. Deleting a node from a binary search tree on n nodes take $O(\log n)$ time in the worst case.

9. Consider the following flow chart on the left to show some steps for making a move in a "Snake and Ladders" game. Choose which one is the correct missing process/decision steps from the right to complete the flow chart in order (1) to (5).



I	Climbing the ladder
II	Throw the dice
III	Move the counter
IV	Landed on snake head?
V	Landed on bottom of ladder?

a. II, III, IV, I, V
b. II, III, IV, V, I
c. III, IV, V, I, II
d. IV, III, V, I, II

10. Which of the following statements are *correct* regarding the graph representations?
   I. Adjacency matrix representation allows fast checking the presence of an edge.
   II. If the graph is a sparse graph with large number of vertices, then adjacency list representation is more space efficient.
   III. If the adjacency matrix of a graph is symmetric, it must be an undirected graph.

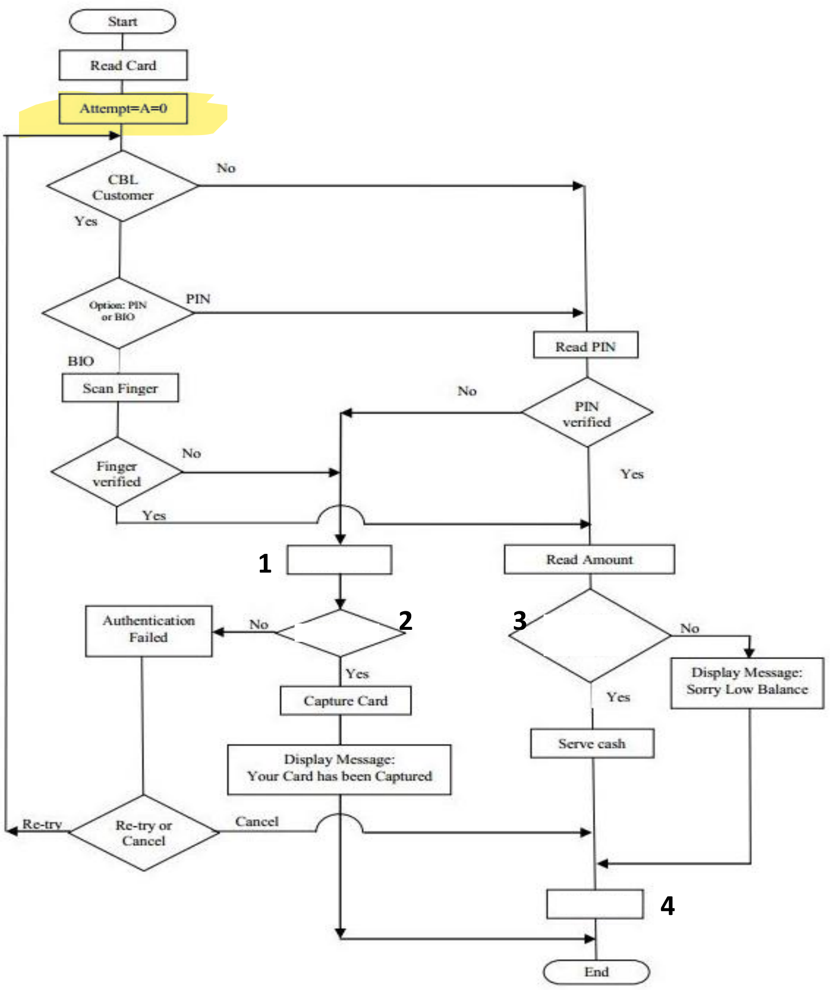
10

a. I Only
b. I and II only
c. I and III only
d. All three

PART B

1. The ATM machines of **State Bank** contain biometric security features besides conventional PIN control. The customers of State Bank have a choice either to use **Card + PIN** or **Card + Biometric** option. Customers of other banks can also withdraw cash from bank's ATM; however, they can only use **Card + PIN** option. A customer's card is captured by the machine after **three** consecutive unsuccessful attempts. The card will be ejected after the successful attempt.

   The following is the incomplete flowchart showing the process of cash withdrawal from State Bank ATM. Assume that transactions other than cash withdrawals are not allowed. Complete the flow chart by adding the missing decision/process(es) 1 to 4.



1	A=A+1
2	If A=3
3	Check balance > amount
4	Eject card

(04 Marks)

2. For each data structure operation given below, choose the correct answer from the following list (not listed in any particular order). Each of which could be the answer for more than one item.

$O(N^2), O\left(N^{1/2}\right), O(log^2 N), O(N \log N), O(N), O(N^2 \log N), O(N^5), O(2^N), O(N^3), O(\log N), O(1)$

1. DeleteMin from a Priority Queue implemented with a binary min heap. <u>O(logN)</u>
2. FindMin in a Priority Queue implemented with a binary search tree. <u>O(N)</u>
3. Pop from a stack implemented with linked list nodes. <u>O(1)</u>

(3*1=03 Marks)

3. Consider the following three functions, $f(x), g(x)$, and $h(x)$ in answering the questions (a) to (c).

$$f(x) = 2x^3 + 4x + 8; \qquad g(x) = x^2 + 2xlogx; \qquad h(x) = x^2 + 2x + 4$$

(a) Is the statement "$f(x) \in \omega(g(x))$" correct? Explain your answer.

(b) Is the statement "Both functions $f(x)$ and $g(x)$ are in the set $\omega(h(x))$" correct? Explain your answer.

([1+1] *2=04 Marks)

4. Illustrate the operation of bubble sort by completing the table below. In successive rows of the table, show the array contents after each pass of the algorithm for the given current state of the array.

9	8	6	7	5	0
8	9	6	7	5	0
8	6	9	7	5	0
8	6	7	9	5	0
8	6	7	5	9	0
8	6	7	5	0	9

(1 *1mark = 06 Marks)

5 What is optimal substructure property? Briefly explain using Knapsack problem.

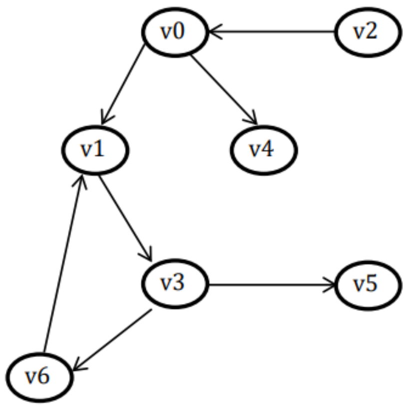
Consider the knapsack problem: Given a set of n items with their weights and values and a knapsack of capacity C, put in the items that maximize the value such that the sum of the weights does not exceed C. The maximum value can be obtained by taking the max of the following two values:

The maximum value obtained by adding n–1 items (excluding the n^{th} item). Value of the n^{th} item plus maximum value obtained by adding n–1 items. (This can only be done if adding the n^{th} item does not exceed the capacity of the bag).

Notice that the maximum value obtained by adding n–1 items is the optimal solution to a subproblem. The optimal solution of the problem can be constructed using the optimal solutions of the sub-problems.

(04 Marks)

6 Consider the directed graph and answer the following questions (a) and (b).



(a) In what order are the vertices visited for a depth-first search that starts at v0?

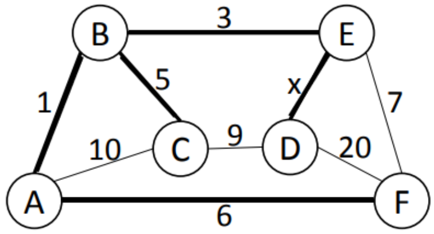
v0 v4 v1 v3 v6 v5 .... **(3 marks)**

(b) In what order the vertices visited for a breadth-first search that starts at v0?

v0 v1 v4 v3 v5 v6 .... **(3 marks)**

(06 Marks)

7. For the following graph, the bold edges form a Minimum Spanning Tree. What can you tell about the range of values for x? Justify your answer.



X ≤ 9 (1 mark) **Justification (2 mark)**

(03 Marks)

30

PART C

1. For the following recurrence equation, solve to get bounds on the runtime.

$$T(2^k) = 3T(2^{k-1}) + 1, \qquad T(1) = 1$$

$= 3^2\ T\ (2^{k-2}) + 1 + 3$
$= 3^3\ T\ (2^{k-3}) + 1 + 3 + 9$ ----------- (1 mark)
… (k steps of recursion)
$= 3^k\ T\ (2^{k-k}) + (1 + 3 + 9 + 27 + \ldots + 3^{k-1})$
$= 3^k + (\ (3^k - 1) / 2\ )$ -----------(1 mark)
$= (\ (2*3^k) + 3^k - 1\ ) / 2$
$= (\ 3 * 3^k\ ) - 1 / 2$
$= 3^{k+1} - 1 / 2$ ----------(1 mark)

(03 Marks)

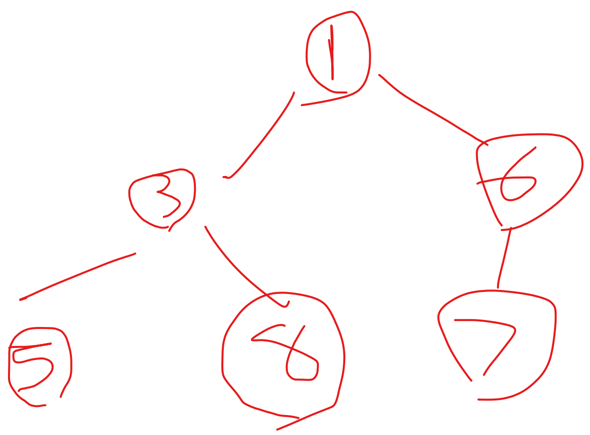
2. Suppose that we first insert an element x into a **binary search tree** that does not already contain x. Suppose that we then immediately delete x from the tree. Will the new tree be identical to the original one? If *yes* give the reason.

Yes. (2 marks)
x will be inserted as a leaf, with no further changes to the tree. A leaf has no children. When we remove a node without children, the rest of the tree remains unchanged. (3 marks)

(05 Marks)

3. Draw the **binary min heap** that results from inserting the integer: 7, 5, 6, 3, 8, 1 in that order into an initially empty binary min heap.

1 mark for each insertion of integers into the binary min heap.



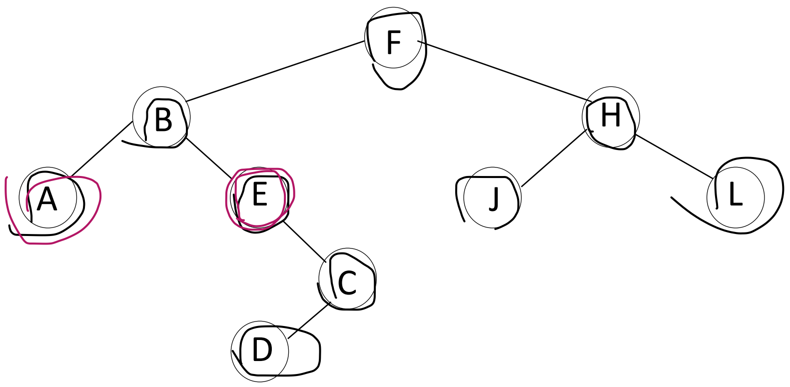
(06 Marks)

4. You are given an array of integers $\boldsymbol{A}$ of length $\boldsymbol{n}$ and another integer $\boldsymbol{q}$. Write an $\boldsymbol{O}(\boldsymbol{n}\,\boldsymbol{log}\,\boldsymbol{n})$ algorithm to find a pair of elements of A such that their sum is q. If no such pair exists, report it briefly.

> **Sort the array. This takes O (n log n)**
>
> **Then for each x in array A, use binary search to look for T – x. This will take O(nlogn).**
>
> **So, overall search is O(n log n)**

(06 Marks)

5. Consider the **binary tree** shown below. For each of the traversals listed, give the order in which the nodes are visited.



Preorder	F	B	A	E	C	D	H	J	L
Inorder	A	B	E	D	C	F	J	H	L
Postorder	A	D	C	E	B	J	L	H	F

(3*2=06 Marks)

6. Write the final result after inserting keys 5, 19, 28, 15, 20, 17, 10, 33 into a **hash table** with collisions resolved by (a) chaining, (b) linear probing. Let the table have eight slots with addresses starting at 0 and let the hash function be $\boldsymbol{h}(\boldsymbol{k}) = \boldsymbol{k}\ \boldsymbol{mod}\ \mathbf{8}$.

(a)

0				
1	33	17		17 33
2	10			10
3	19			19
4	20	28		28 20
5	5			5
6				
7	15			15

(b)

0	33	33
1	17	17
2	10	10
3	19	19
4	28	28
5	5	5
6	20	20
7	15	15

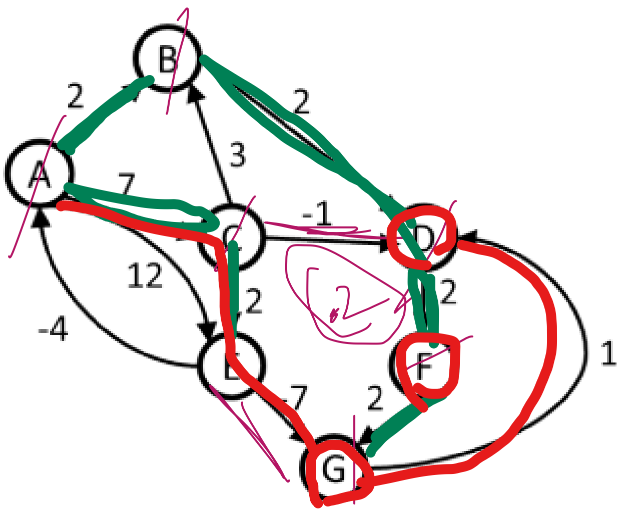
(4+4=08 Marks)

7. Given a stack of integers, write a pseudo code for external method that sorts the stack in ascending order, so that the smallest element appears on top of the stack. You can use the methods ***pop, top, push***, and ***size*** as already implemented methods. Let ***n*** be the size of array and **arrStack** is the final stack output after executing the pseudo code.

```
Sort (a):

    n ← size(a)
    arrStack ←[]
    temp ←[]


    your code here


    for i ← n to 1:                                ......... (1 mark)

        max ← a.pop()

        for j ← 0 to i-1:                          ............. (1 mark)

            k ← a.pop()

            if k>max then:

                temp.push(max)                     ............... (2 marks)

                max ← k

            else: temp.push(k)

        a.push(max)

        while temp is not empty:

            arrStack.push(temp.pop())  ........ (2 marks)
```

(06 Marks)

8. Consider the following directed, weighted graph:



a. Even though the graph has negative weight edges, step through ***Dijkstra's algorithm*** to calculate supposedly shortest paths from A to every other vertex. Show your steps in the table. Cross out old values and write in new ones, from left to right within each cell, as the algorithm proceeds. Also list the vertices in the order which you marked them known.

vertex	known	distance	path
A	Y	0	
B	Y	2	A
C	Y	7	A
D	Y	4	B
E	Y	12, 9	A C
F	Y	6	D
G	Y	8	F

(1+1+2+2=06 Marks)

b. Dijkstra's algorithm found the wrong path to some of the vertices. For just the vertices where the wrong path was computed, indicate both the path that was computed and the correct path.

To G → A, B, D, F, G but => A, C, E, G

To D → A, B, D but => A, C, E, G, D

To F → A, B, D, F but => A, C, E, G, D, F

(03 Marks)

c. What single edge could be removed from the graph such that Dijkstra's algorithm would happen to compute correct answers for all vertices in the remaining graph?

E – G

(03 Marks)

9. You can use a tree to represent a file system, with each node representing a directory or file. Since a directory can contain more than 2 files, a binary tree may seem like a poor choice. But you can do it using a structure where a node N's "left" child is the first child file inside N, and the "right" child is the next sibling file/dir within the same parent directory of N. (A normal file that isn't a directory will always have a null first child.) For example, consider the set of directories and files

shown below at left. The file system contains files such as "/bin/rm" and "/usr/lib/code/java". You can represent this set of files using the binary tree shown at right.

```
bin
    cat
    ls
    rm
    tar
home
    msahami
    stepp
    xuamyj
usr
    include
        h
    local
    lib
        apt
        code
            cpp
            java
            python
        dpkg
    share
    src
var
```

```
                 ""
                /
            bin___________
           /              \
       cat                 home_____________________
          \               /                         \
           ls      msahami                           usr
             \            \                         /   \
              rm           stepp            include     var
                \               \          /       \
                 tar             xuamyj   h         local
                                                         \
                                                          lib
                                                         /   \
                                                      apt     share
                                                         \         \
                                                          code      src
                                                         /    \
                                                      cpp      dpkg
                                                         \
                                                          java
                                                              \
                                                               python
```

For this problem, write a recursive function named **removeFile** that deletes a directory or file from the file system, including all descendent files/dirs inside it. Your function accepts two parameters: a reference to a **FileTreeNode** pointer for the root of the file system binary tree, and a string for the full path of the file or directory to delete, with / slashes between directories, such as "/usr/local/lib". If there is no node in the tree that corresponds to the string path passed in, your function should not modify the tree.

```
removeFile(dir):
    if dir exists then:
        for the_file in dir:
            file_path ← os.path.join(dir, the_file)

            if file_path exists then:
                os.unlink(file_path)
            else:
                removeFile (file_path)
                os.rmdir(file_path)
```

(08 Marks)