

Do Not Turn This Page Until
Directed to do so.

Name: _____

You have 75 Minutes minutes from the start of class to complete
this exam.

You **are** allowed:

One 8.5x11 inch Page of Notes (Front and Back)

You **are not** allowed to use:

Your Book

Your Course Notes

Your Friends

Anyone Else

Your phone or any other devices

If you fail to turn in your test before leaving the room, check your
cell phone, or you use any resource other than your card of notes,
your test will be taken and you will be **given a grade of 0**.

Your note card must be submitted with the test.

**Read the instructions carefully and answer each
question as completely as you can.**

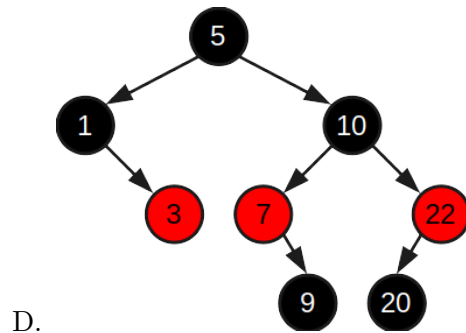
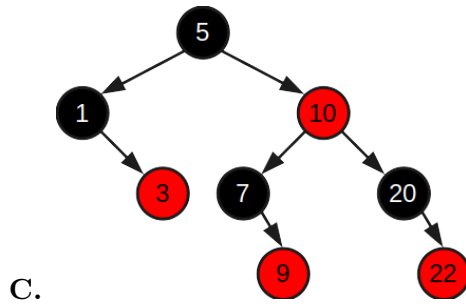
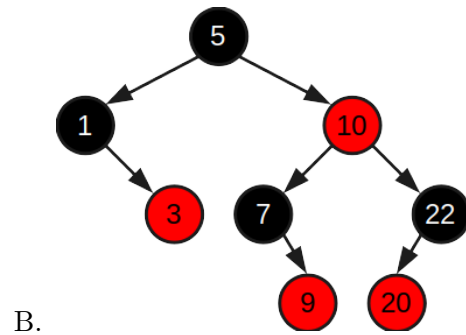
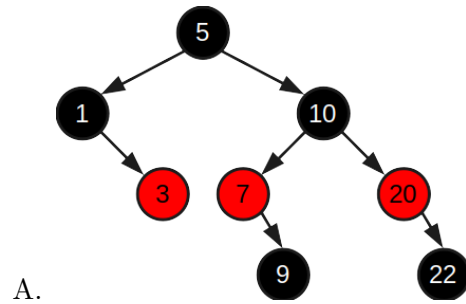
**There will be no talking during the test. This also
means that I will not answer questions. Do what
makes sense for each question.**

Question	Points	Score
1	4	
2	4	
3	4	
4	4	
5	4	
6	4	
7	4	
8	4	
9	4	
10	4	
11	4	
12	4	
13	4	

Question	Points	Score
14	4	
15	4	
16	4	
17	4	
18	4	
19	4	
20	4	
21	4	
22	4	
23	4	
24	4	
25	4	
Total:	100	

This page is left blank to use as scratch paper.

1. (4 points) Starting from an empty tree what will be the resulting Red-Black Tree after adding the following items: [1, 7, 5, 10, 3, 20, 9, 22]



2. (4 points) For any given deletion from a red-black tree how many recolorings at most would be required?

- A. $O(1)$
- B. $O(n)$
- C. $O(\log n)$
- D. $O(n^2)$

3. (4 points) For any given Red-Black tree how many restructurings would be expected (at most) for a insertion operation?
 - A. 0
 - B. n
 - C. 1**
 - D. 2

4. (4 points) Which of the following is not true about Splay trees?
 - A. Does not have a logarithmically bounded height
 - B. Can be degenerate
 - C. Only adjusts itself after insert and delete operations, but not after lookup**
 - D. Keeps most commonly used data at the top.

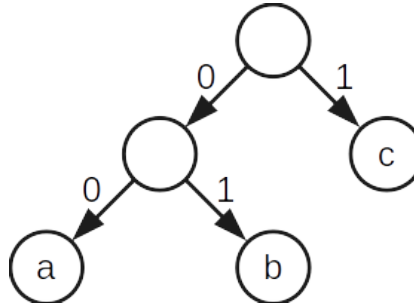
5. (4 points) What is the key difference between an (a, b)-Tree and a B-Tree
 - A. a B-Tree has data only external nodes
 - B. a B-Tree is more general than an (a, b)-tree
 - C. a B-Tree is an (a, b)-Tree where $a = \text{ceil}(b / 2)$**

6. (4 points) You work for a company that is interested in developing a language and programming system to compete with Java and C#. You have been tasked with the development of a garbage collection system for recently de-allocated objects. Which of the following data structures would you use to implement such a structure (optimally).
 - A. Red-Black Tree
 - B. Map
 - C. Splay Tree**
 - D. Priority Queue (Heap-Based)

7. (4 points) Select the one which makes the following statement false. Compared to a LinkedList a Skip List -----?
 - A. does not have $O(1)$ insertion and deletion time
 - B. has $O(\log n)$ search time
 - C. increases the difficulty to access items in the middle**

8. (4 points) The following narrative describes an operation on one of the data structures we studied. Please select which data structure and operation is describe, as follows:
 Given an item i, search the structure until either i is found, or null is reached. If null is reached, then we can add i to the structure. We then set the value for the level to be i and we randomly flip a coin, until “tails” is seen, each time heads is seen we add i into “level” and then increment level.
 - A. Insertion into a Red-Black Tree
 - B. Insertion into a Perfect SkipList
 - C. Insertion into a Random SkipList**
 - D. Insertion into a Splay Tree

9. (4 points) Using the following Huffman Coding Tree, which of the following is a correct encoding of ACAB:



- A. 010011
B. 01010100
C. 0010001
D. 10101011
10. (4 points) We use variable length codes in Huffman coding to optimize the compression of the encoded data. What property of these prefix free codes makes it so that Huffman coding actually works?
- A. We don't actually use prefix free codes in Huffman Coding
B. They are uniquely decipherable since each code is of the same length.
C. They are uniquely decipherable since no code is a prefix of another.
D. Huffman coding doesn't produce optimal encoding of the data.
11. (4 points) You have just been hired into a company developing a new text editor marketed towards software engineers. For your first project you have been tasked with the implementation of a feature which indicates when programmer's have misspelled a word in their code. Of the following, which data structure would be the best choice to build your implementation on top of?
- A. Deque
B. 2,4-Tree
C. Trie
D. SkipList
12. (4 points) If I am determining whether to use a Heap or an LinkedList as the basis of a PriorityQueue, we know that the difference lies in the fact that for a Heap both `offer` and `poll` are $O(\log n)$ but in an LinkedList `offer` is $O(1)$, while `poll` is $O(n)$. What does this mean.
- A. If I am offering seldom, and polling often, I should choose the heap.
B. The Linked List is much faster when offering data, so just use it always.
C. If I will be offering seldom, but polling often, I should choose to use a Heap.
D. The Heap is best for small amounts of data.

13. (4 points) Lets say that we have a priority Queue which we are using to sort people by their age prior to gaining access to a prize drawing. We have the following people:

Larry, 22

Sara, 19

Bob, 5

Qing, 98

Lowell, 18

Sunshine, .5

Assuming that we use a Max-Heap as the basis of the Queue (and their ages are a value in years). After offering each individual to the Priority Queue, and we perform the following operations:

1 peek, 3 polls, 1 peek, 1 poll, then offer the following individuals:

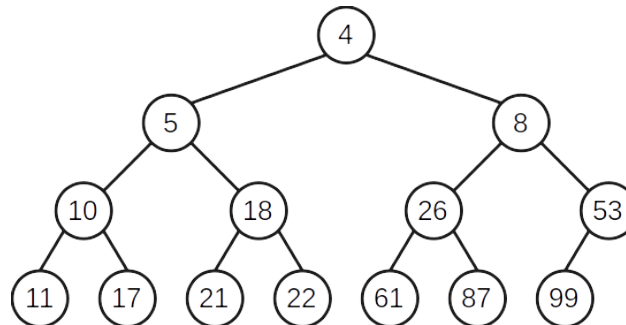
Ginny, 13

Lora, 8

Then perform an additional peek, who is the person on top of the Queue?

- A. Lowell
- B. Bob
- C. Ginny**
- D. Sara

14. (4 points) Which type of heap does the following represent?



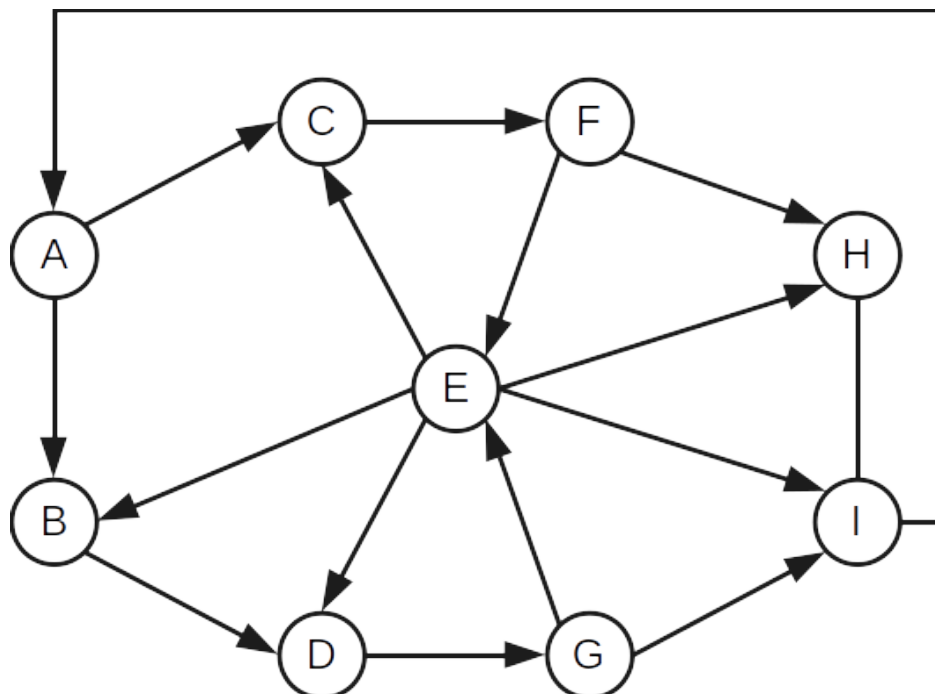
- A. Max Heap
- B. Neither
- C. Min Heap**

15. (4 points) If I have a heap with 33 nodes, how many levels does the heap have?

- A. 1
- B. 5
- C. 6**
- D. 2

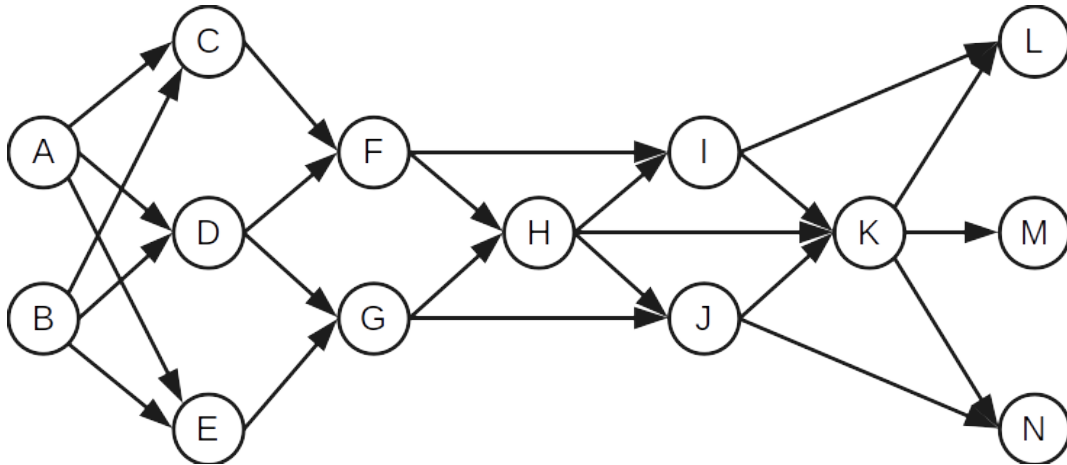
16. (4 points) Which of the following best describes the greedy algorithm design strategy?
- A. The problem is divided into smaller versions of the problem until they can be immediately computed, the solution is then built by aggregating these small solutions into the whole.
 - B. The problem is divided into smaller subproblems whose answer is stored for later use or when the same subproblem is seen. The larger problem is solved by selecting the optimal solution from the aggregation of the subproblems.
 - C. The problem is solved optimally in the global sense by selecting locally optimal choices at each stage of the problem.**
17. (4 points) True or False: Unit testing is capable of detecting all errors in a program.
- A. True
 - B. False**
18. (4 points) A Map is a data structure which?
- A. an ordered collection of unique values
 - B. relates a unique key to one or more data values.
 - C. relates a unique key to a data value (not necessarily unique)**
 - D. is a fancy array
19. (4 points) If the load factor in a hash table grows too large, the operation overhead also increases, leading to poor performance. One strategy we have to reduce loading is to resize the table. Why does this strategy work?
- A. This actually does nothing.
 - B. decreases the number of entries in the tables which in turn reduces the load factor.
 - C. Increases the number of buckets and thus reduces the load factor.**
 - D. Splits out the entries and reduces clusters.
20. (4 points) A MultiSet is a data structure which
- A. Is a fancy name for an array
 - B. Contains an unordered collection of unique items
 - C. Contains an unordered collection of items which may contain duplicates.**
 - D. Associates a set of unique keys to values (not necessarily unique)
21. (4 points) I am interested in building a quick-and-dirty web crawler which simply extracts every word (except stop words such as “a, an, the, by, be, ...”) which occur on each web page visited. I then want to store these words such that for each unique word I want to know which sites that it appears on. Which of the following data structures would be best suited for this task?
- A. Set
 - B. MultiSet
 - C. MultiMap**
 - D. Map

22. (4 points) Linear and Quadratic probing are techniques used to handle collisions when inserting or searching for items in a Map/HashTable. Which of the following is not disadvantage of this technique?
- A. Upon deletion we must use a special marker to note that we have not yet reached the end.
 - B. These approaches tend to cluster items together causing slowdown on search, insert, and delete operations
 - C. Depending on the size of the bucket selected, there could be significant slowdown on search, insert, or delete operations.**
23. (4 points) I am currently designing a simple text-based dungeon game. At any point in the game a player has a choice on the direction they wish to go. Not all directions are available in every room of the dungeon. Furthermore, some rooms, preclude the ability to go back the way you came or even lock the door after you enter. What data structure would you select to optimally represent the order of the path a player has taken through the dungeon?
- A. Tree
 - B. DAG
 - C. Stack**
 - D. Queue
24. (4 points) Which of the following represents a simple path through the following graph:



- A. A B E F H
- B. A C F E C
- C. A C F E H I**
- D. I A C F H I

25. (4 points) Given the graph shown below, which of the following is a valid topological sort of the graph?



- A. A B C D E F G I H J K N M L
- B. A C B D F E G H I J K N M L
- C. A B C D F E G H I J K N M L**
- D. B A C D F E G I H J K N L M

End of Test