# **UC Berkeley - Computer Science**

CS61B: Data Structures

Final Exam, Spring 2017.

This test has 13 questions worth a total of 200 points, and is to be completed in 165 minutes. The exam is closed book, except that you are allowed to use three double sided written cheat sheets (front and back). No calculators or other electronic devices are permitted. Give your answers and show your work in the space provided. Write the statement out below in the blank provided and sign. You may do this before the exam begins.

"I have neither given nor received any assistance in the taking of this exam."

#	Points	#	Points
0	0.5	7	23
1	12	8	0
2	28	9	16
3	14	10	16
4	12	11	12
5	16	12	12
6	12	13	26.5
		TOTAL	200

Name:			
SID:			
Three-letter Login ID:			
Left SID:			
My left neighbor has ID:			
Right SID:			
My right neighbor has ID:			
Exam Room:			

Signature:

#### Tips:

- There may be partial credit for incomplete answers. Write as much of the solution as you can, but bear in mind that we may deduct points if your answers are much more complicated than necessary.
- There are a lot of problems on this exam. Work through the ones with which you are comfortable first. Do not get overly captivated by interesting design issues or complex corner cases you're not sure about.
- Not all information provided in a problem may be useful.
- See the coding reference sheet on the last page for potentially useful data structures.
- Unless otherwise stated, all given code on this exam should compile. All code has been compiled and executed before printing, but in the unlikely event that we do happen to catch any bugs in the exam, we'll announce a fix. Unless we specifically give you the option, the correct answer is not 'does not compile.'
- O indicates that only one circle should be filled in.
- $\square$  indicates that more than one box may be filled in.
- For answers which involve filling in a  $\bigcirc$  or  $\square$ , please fill in the shape completely.

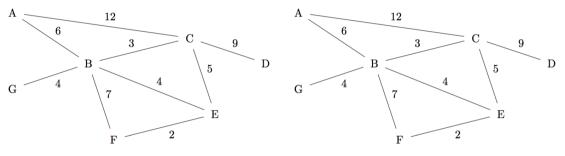
Optional. Mark along the line to show your feelings	Before exam: [☺	©].
on the spectrum between $\ensuremath{ ext{@}}$ and $\ensuremath{ ext{@}}$ .	After exam: [⊗	©].

UC BERKELEY
Login:

**0. So it begins (0.5 points).** Write your name and ID on the front page. Write the exam room. Check the IDs of your neighbors. Write the given statement. Sign. Write your login in the corner of every page. Enjoy your free 0.5 points ©.

## 1. Mystery Spanning Tree 3000 (12 points).

a) (4 pts) For the graph below, list the edges in the order they're added to the MST by Kruskal's and Prim's algorithm. Assume Prim's algorithm starts from vertex A. Assume ties are broken in alphabetical order (i.e. the edge  $\overline{AB}$  would be considered before  $\overline{AC}$ ). Denote each edge with alphabetical overbar notation  $\overline{AB}$ , which represents the edge from A to B. You may not need all blanks. For your convenience, the graph is printed twice (to make running algorithms easier).



Prim's algorithm order:

Kruskal's algorithm order:

b) (2 pts) Is there any vertex for which the shortest paths tree (SPT) is the same as your Prim MST above?

O Yes, and it's \_\_\_\_\_O No

c) (6 pts) For the following propositions, fill in true or false <u>completely</u> and provide a brief explanation. For a proposition that is false, a counter-example suffices. Assume all edge weights are unique.

O True / O False: Adding 1 to the smallest edge across any cut of a graph G must change the total weight of its minimum spanning tree.

O True / O False: The shortest path from vertex A to vertex B in a graph G is the same as the shortest path from A to B using only edges in T, where T is the MST of G.

O True / O False: Given any cut, the maximum-weight crossing edge is in the maximum spanning tree.

is executed. Which sorts from the list below are monotically improving? Assume that all sorts are as presented during lecture on arrays. Assume insertion sort and selection sort are in-place. Assume heapsort is in-place and that the array acts as a max heap. Assume that Quicksort is non-randomized, uses the leftmost item as pivot, and uses the Hoare partitioning strategy (i.e. using "smaller than" and "bigger than" pointers) from lecture.

<sup>&</sup>lt;sup>1</sup> A BigInteger is an "immutable arbitrary precision integer." It can represent any integer, not just those that fit into 32 bits.

<i>UC BERKELEY</i>
Login:

- **3. Traversals** (14 points). Suppose we have an NAryIntTree, defined as shown below. Any node may have any number of children. If a node is a leaf, children is null. Assume that children[i] is never null for any i.
- a) (6 pts) Fill in the printTreePostOrder method below, which prints the values of the tree in postorder, with one val per line. Your solution **must be recursive** and take linear time in the number of nodes.

<pre>public class NAryIntTree {     private Node root;     public class Node {         public Node[] children;         public int val;     }     public void printTreePostOrder() {</pre>	
}	
public	
/* */ }	

b) (8 pts) Fill in the code below which prints out the values of the tree in level order with one val per line. Your solution must be iterative and take linear time in the number of nodes for any tree.

private void printTreeLevelOrder() { // is a method of NAryIntTree


}

	gin:
4.	Algorithms and Data Structures (12 points).
a)	(4 pts) In class we primarily considered two graph representations: the adjacency list and the adjacency matrix. Antares suggests that we can improve the performance of Dijkstra's algorithm with a third graph representation he calls an "adjacency heap". For each vertex v, v's adjacency heap stores all of v's neighbors in a heap ordered by edge weight, so that the smallest edge adjacent to v is at the root of its heap. Naturally, Antares stores these heaps as arrays. Antares reasons that by considering small edges first, Dijkstra's will be able to complete faster.
	Will using an adjacency heap result in better, equivalent, or worse <u>asymptotic</u> runtime performance for Dijkstra's algorithm than using a regular adjacency list? Assume that we only care about worst case asymptotic performance. Briefly justify your answer.
	OAdjacency heap is better O Performance is the same O Adjacency heap is worse  Justification:
b)	(4 pts) Suppose Antares has conjured up the <b>Gulgate Priority Queue</b> ( <b>GPQ</b> ). Given a GPQ containing N elements, the worst-case running time for insertion, deletion, and change-priority are given as follows: Insertion: $\Theta(N)$ , Deletion: $\Theta(N)$ , Change-Priority: $\Theta(1)$ .
	Suppose we run the implementation of Dijkstra's algorithm provided in class (where every vertex is initially inserted into the PQ with infinite priority) using a GPQ on a graph with V vertices and E edges. What is the worst case runtime of Dijkstra's? <b>Give your answer in big O notation in terms of <math>\underline{V}</math> and <math>\underline{E}</math></b> . Assume that $E >> V$ (this means E is much greater than V).
	Runtime: O()
c)	(4 pts) Suppose Antares has also created a <b>Xelha Quick Union</b> ( <b>XQU</b> ) to check if two vertices are connected while running Kruskal's. Given that there are N items in an XQU, the running time for XQU operations is as follows: Constructor: $\Theta(N)$ , Union: $\Theta(N \log N)$ , Is-Connected: $\Theta(\log N)$
	Suppose we run the implementation of Kruskal's algorithm as presented in class using a XQU and a heap-based priority queue. Recall that in our version of Kruskal's from class, all edges are initially inserted into a regular heap-based priority queue and removed one by one, and added to the MST so long as there are no cycles. What is the worst case runtime of Kruskal's algorithm? Give your answer in big O notation in terms of $\underline{V}$ and $\underline{E}$ . Assume that $E \gg V$ .
	Runtime: O()

<i>UC BERKELEY</i>
Login:

### 5. Potpourri (16 points).

a) (6 pts) We learned in lecture and in lab that we can use an array to compactly store a min/max heap, with formulas to calculate the parent, left child, and right child given a node. Now suppose we want to store a ternary heap, where every node has 0, 1, 2, or 3 children. Would the compact array representation work? If your answer is <u>ves</u>, give formulas on the <u>left</u> to calculate the parent, left child, middle child, and right child. If your answer is <u>no</u>, explain why on the <u>right</u>.

Assuming root is at index:	Impossible, because:
<pre>public int parent(int k) { return;}</pre>	-
<pre>public int left(int k) { return;}</pre>	
<pre>public int middle(int k) { return;}</pre>	
<pre>public int right(int k) { return;}</pre>	

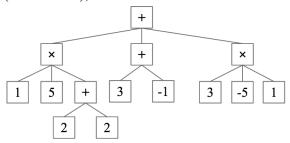
- b) (2 pts) In class, we said that anytime you override equals, you must also override hashCode. Suppose the Yarg class overrides the equals method, but does not override hashCode. Suppose that yargSet is a HashSet<Yarg>. What are the potential direct consequences of not overriding hashCode?
  - O True / O False: yargSet.contains() may return an incorrect result.
    O True / O False: yargSet.contains() runs a much higher risk of taking linear time.
- c) (6 pts) If we wanted to build a generic TrieSet that could hold many different types, we'd need to require all such types to implement some interface, much like items in a TreeSet must implement the Comparable interface (shown below). Give a declaration of an appropriate interface and describe any methods with comments. Provide useful names for your methods and interface (not silly ones, sorry). You may not need all blanks.

d) (2 pts) Suppose the creator of a new DogPicture class is deciding whether or not to implement interface X, where X is the interface from part c. What is the primary consideration of the creator? "Will somebody ever want to build a Trie of DogPictures" is not enough of an answer.

Login:	
<pre>Stock s = nameToStock.get(t.name); s.price = t.price; }</pre>	he code below. Assume the MaxPQ is heap here insertion and deletion operations take que. Assume STOCK_LIST is a list of Stock hay be traded multiple times in one day. ocks(int k) { (); w HashMap<>(); //assume nice hashcode spread //uses bottom up heapification //market closes at 5 PM //waits if no trade available //assume key always in map //may be higher or lower
<pre>ArrayList<stock> returnStocks = new Arra for (int i = 0; i &lt; k; i += 1) {     returnStocks.add(rankedStocks.delM } return returnStocks; }</stock></pre>	//assume k <= s
Where the compareTo method of Stock is defined as publication this.price - s.price; } where price is an	
a) Which of the correctness or compilation issues listed beloapply. If you believe there are compilation errors, consider errors are fixed. Assume k is smaller than the number of stock	the other boxes assuming the compilation
□Compilation: The method is supposed to return a List, bu □Compilation: The HashMap cannot point to items inside of variables are private. □Correctness: The algorithm actually returns the k cheapest □Correctness: The algorithm may return a list with duplicate □Other (explain):	The MaxPQ because the MaxPQ's instance stocks.  es if the same stock is traded multiple times.
b) What is the worst case runtime and space complexity of compilation errors you identified in part a? Give your answeright as possible with no unnecessary lower order terms or conformation of S, T, and k, where S is the number of stocks, T is the number of stocks.	er in O notation. Your bounds should be as constant factors. Give your answers in terms
Runtime complexity: O(	)

CS61B FINAL, SPRING 2017

**7. Arithmetic Tree (23 pts).** An arithmetic tree is a tree that stores an arithmetic expression. For this problem, assume all nodes are either multiplication (represented with  $\times$ ), addition (represented with +), or a number (represented as a written integer). For example, the following tree represents  $(1 \times 5 \times (2 + 2)) + (3 + -1) + (3 \times -5 \times 1)$ , which would evaluate to 7.



Your job is to fill out the code below such that evaluateTree(Node tree) evaluates the arithmetic tree rooted at tree to its correct value. For example, if evaluateTree were applied to the X node at the top left of the figure above, it would return 20. Multiplication and addition operator nodes can have any number of children. You do not need to check for bad inputs (e.g. null children). You may find it easier to work your way from the end of the problem back to the front. You may not need all blanks.

If you're stuck on this problem, come back later!

```
CS61B FINAL, SPRING 2017
Login: _____
public abstract class OperatorNode extends Node {
   public int numArgs() { /* Returns number of args for this node. */ }
   public abstract int apply(int arg1, int arg2);
   public void processNode(Stack<Integer> stk) {
       int res = stk.pop();
       for( ) {
       stk.push(res);
   }
}
public class ArgNode extends Node {
   public int value;
   @Override
   public void processNode(Stack<Integer> stk) {
   }
}
/* Don't overthink this! */
public class MultiplicationNode extends OperatorNode {
   @Override
   public int apply(int arg1, int arg2) {
       return _____;
   }
}
public class AdditionNode extends OperatorNode {
   public int apply(int arg1, int arg2) {
```

**8. PNH (0 points).** This United States President won office with the smallest fraction of the popular vote in the history of United States presidential elections.

}

}

return \_\_\_\_\_;

9. Asymptotics (16 points). For each of the code snippets below, give the <u>best</u> and *worst case* runtimes in terms of N. Give the <u>best runtimes</u> in the column to the <u>left</u>, and the *worst* in the column to *the right*.

Best: Worst: public static void f1(int N) {  $\Theta$ \_\_\_\_\_ if (N == 0) { return; }

```
f1(N / 2);
                           f1(N / 2);
                           g(N); // runs in \Theta(N^2) time
                        }
                       public static int f2(String[] x, int i) {
_Θ____
          _Θ____
                           int N = x.length;
                           int total = 0;
                           try {
                               while (i < N) {
                                   total += x[i].length();
                                   i += 1;
                               }
                           } catch(NullPointerException e) {
                               x[i] = "null";
                               total += f2(x, i);
                           return total;
                       }
                       Assume t is a binary IntTree with N nodes:
          \Theta____
                       public static void f3(IntTree t) {
                           t.value = t.value * 2;
                           if (t.left != null) { f3(t.left);
                           if (t.right != null) { f3(t.right);
                       }
_Θ____
                       Assume t is a binary IntTree with N nodes:
                       public static void f4(IntTree t) {
                           t.value = t.value * 2;
                           if (t.left != null) { f4(t.left);
                                                                  }
                           t.right = t.left;
                           if (t.right != null) { f4(t.right);
                           t.left = t.right;
                       }
```

<i>CS61B</i>	FINAL,	<b>SPRING</b>	2017
Login:			

10. Return of the XelhaTree (16 points). Write a function validXelhaTree which takes an IntTree and a List and returns true if the IntTree is a XelhaTree for the list. You may not need all lines. A XelhaTree is valid if it obeys the min heap property, and if an in-order traversal of the XelhaTree yields the list of items passed to createXelhaTree (in the same order). One line if statements with {} on the same line are fine. You may not need all the blanks. Assume there are no duplicates.

lic class XelhaTreeTest {	
<pre>public static class IntTree {</pre>	
<pre>public int item;</pre>	
<pre>public IntTree left, right;</pre>	
}	
<pre>public static IntTree createXelhaTree(List<integer> x) { } /** If x is null, returns largest possible integer 2147483647 private static int getItem(IntTree x) {    if (x == null) { return Integer.MAX_VALUE; }    return x.item;</integer></pre>	
1	
<pre>public static boolean isAHeap(IntTree xt) {</pre>	
	}
<pre>public static void getTreeValues(IntTree xt, List<integer> tree</integer></pre>	eValues){
	,
<pre>public static boolean validXelhaTree(IntTree xt, List<integer></integer></pre>	<pre>}  vals) {</pre>
<pre>List<integer> treeValues = new ArrayList<integer>();</integer></integer></pre>	
<pre>/* getTreeValues adds all values in xt to treeValues */</pre>	
<pre>getTreeValues(xt, treeValues);</pre>	
getireevalues(xt, theevalues);	
return isAHeap(xt) &&	ì
I E LUI II I DAIIEAP (AL) KK	ſ

<i>UC BERKELEY</i>
Login:

11. MaxPQ (12 points). Complete the implementation of MaxPQ using data structures from the reference sheet on the last page of the exam. You may not need all blanks. Write at most one statement per line.

oublic MaxF 	PQ() {		 	
}			 	
oublic clas	ss		 	{
}				
oublic Iter 	n delMax() {			
}				
oublic void	l insert(Iter	m x) {		

Welcome to the Chill Out Zone.

CS61B FINAL, SPRING 2017 Login:
12. Danger and Optimization (12 points).
a) (5 pts) Suppose you provide a computing service where users can upload lists of integers and receive back the numbers in sorted order. Which sorts below would be appropriate to choose for this task assuming you want to prevent users from submitting inputs that either result in terrible <sup>2</sup> runtime or cause an exception? <b>Assume you are using Java</b> .
O Appropriate / O Inappropriate : Merge Sort O Appropriate / O Inappropriate : Insertion Sort O Appropriate / O Inappropriate : Quicksort using Hoare partitioning and that starts by shuffling O Appropriate / O Inappropriate : LSD O Appropriate / O Inappropriate : Recursive MSD
b) (5 pts) Suppose you provide a service where users can upload a list of names (stored as Java Strings) and it will return the list of all unique Strings. Which set implementations below would be appropriate to choose for this task, assuming you want to prevent users from submitting inputs that either result in terrible runtime or cause an exception?
O Appropriate / O Inappropriate : 2-3 Tree based Set O Appropriate / O Inappropriate : LLRB based Set O Appropriate / O Inappropriate : Hash based Set O Appropriate / O Inappropriate : Trie based Set O Appropriate / O Inappropriate : TST based Set
c) (2 pts) Suppose you provide a service where users can upload their own custom graphs and a start vertex, and you will find the list of all vertices reachable from the start. Which graph search algorithms would be appropriate to choose for this task, assuming you want to prevent users from submitting inputs that either result in terrible runtime or cause an exception?
O Appropriate / O Inappropriate : Recursive DFS O Appropriate / O Inappropriate : BFS
d) (0 pts) What should Josh name his future kid (assume female if you want a gender specific name)?

<sup>&</sup>lt;sup>2</sup> By terrible we mean: Imagine you are demoing your website and want to impress someone with its speed. Does it run so slow that you are embarrassed? If so, that is terrible.

<i>UC BERKELEY</i>
Login:

13. Reduc	tions	(26.5 po	ints). C	Often in cor	npu	ıter	science, j	problen	ns are	just c	ther p	roblems in	disg	uise.
Complete	each	problem	below	according	to	the	direction	ns give	n. Ma	ny of	these	problems	are	very
challengin	g.													

a)	(3 pts) Describe an algorithm to find a <b>maximum</b> spanning tree. Your algorithm must use Kruskal's
	as a "black box," that is, without any modifications. Your answer should be brief.

- b) (5 pts) Suppose you want to find the SPT of a graph, but where you redefine the total cost of a path as follows. Let cost(List<Edge>) be the sum of the weights of the edges, plus the number of edges. In other words, we want to run Dijkstra's taking into account not just the weights of the edges, but also the number of edges. Describe an algorithm to find this shortest paths tree. Your algorithm must use Dijkstra's as a "black box". Your answer should be brief.
- c) (5 pts) Dijkstra's algorithm sometimes fails on graphs with negative edges. Suppose we have a graph G with a single negative edge with weight -Q, and we want to find the shortest path. Suppose we construct a new graph G' where every edge has Q added to its weight. If we run Dijkstra's on G', is the resulting shortest paths tree always a correct shortest paths tree for G? If yes, explain why. If no, provide a counter-example.

O Yes, because:	O No, counter-example:			

d) (6 pts) Suppose that we're using a programming language Zulg where instead of comparison we have a *zelch* operation. Suppose that we prove that "puppy, cat, dog" requires  $\Omega(N \log \log N)$  *zelch* operations. Assume that zelch takes constant time. For each of the following statements, determine whether the answer is false, true, or the answer depends on whether P = NP.

O True / O False / O P=NP?: A <i>zelch</i> based sort requires at least $\Omega(N \log \log N)$ <i>zelch</i> operations.
O True / O False / O P=NP?: Sorting an array in Zulg requires $\Theta(N \log \log N)$ time in the worst case
O True / O False / O P=NP?: The optimal sorting algorithm in Zulg requires $O(N \log \log N)$ time in
the worst case.
O True / O False / O P=NP?: All sorting algorithms in Zulg require $O(N \log \log N)$ time in the worst
case.

<sup>&</sup>lt;sup>3</sup> Recall that "puppy, cat, dog" is a game from lecture where we have N boxes, each containing a unique object (e.g. a puppy, a cat, and a dog) of known size, and our job is to determine which box contains which object.

Login: \_\_\_\_\_

e) (6 pts) Suppose we have the abstract data type MinimumPQ, defined as an interface in Java as shown below:

```
public interface MinimumPQ<Item extends<Comparable<Item>> {
   public void add(Item x);
   public Item removeMin();
   public Item min();
}
```

For each statement below, state whether it is true, false, or "depends on whether P = NP". Assume that all implementations are correct.

- O True / O False / O P=NP?: There exists a possible MinPQ implementation for which add requires  $\Theta(\log \log N)$  time in the worst case.
- O True / O False / O P=NP?: There exists a possible MinPQ implementation for which removeMin requires  $\Theta(\log \log N)$  time in the worst case.
- O True / O False / O P=NP?: There exists a possible MinPQ implementation for which add and removeMin require  $\Theta(\log \log N)$  time in the worst case.
- O True / O False / O P=NP?: There exists a possible MinPQ implementation for which add requires  $\Theta(1)$  time in the worst case.
- O True / O False / O P=NP?: There exists a possible MinPQ implementation for which removeMin requires  $\Theta(1)$  time in the worst case.
- O True / O False / O P=NP?: There exists a possible MinPQ implementation for which add and removeMin require  $\theta(1)$  time in the worst case.
- f) (1.5 pts) Consider the LongestPath problem, i.e. given a graph, does there exist a path with total weight k or greater? Suppose that we prove that LongestPath cracks 3SAT (i.e. 3SAT reduces to longest path). For each of the following statements, determine whether the answer is false, true, or the answer depends on whether P = NP. Let N be the number of edges.
- O True / O False / O P=NP?: There exists an algorithm to solve LongestPath in  $O(N^k)$  time.
- O True / O False / O P=NP?: There exists an algorithm to check a supposed solution to LongestPath in  $O(N^k)$  time.
- O True / O False / O P=NP?: There exists an algorithm to calculate the length in bytes of the shortest Java program that solves LongestPath.



... and that's it!

### Nothing written on this page will be graded.

### **Data Structures Reference:**

```
MinPQ<Item extends Comparable<Item>>> {
HashSet<Key> {
  void add(Key k)
                                     MinPQ(Comparator<Item> c)
  boolean contains(Key k)
                                     void insert(Item x)
}
                                     Item min()
                                     Item delMin()
                                  }
HashSet is the same except Key must
                                  Uses natural order unless comparator given during construction.
implement Comparable<Key>
                                  Stack<Item> {
HashMap<Key, Value> {
                                                           Queue<Item> {
  void put(Key k, Value v)
                                    void push(Item x)
                                                             void enqueue(Item x)
  boolean containsKey(Key k)
                                                             Item dequeue()
                                    Item pop()
                                  }
                                                           }
  Value get(k)
TreeMap is the same except Key must
implement Comparable<Key>
Assume all of these classes implement Iterable and have a size() method.
```