UNIVERSITY OF SOUTHAMPTON COMP1009W1

SEMESTER 2 EXAMINATION 2007/2008

DATA STRUCTURES AND ALGORITHMS

Duration: 120 mins

You must enter your Student ID and your ISS login ID (as a cross-check) on this page. You must not write your name anywhere on the paper.

Student ID:	
ISS ID:	

Question	Marks
1	
2	
3	
4	
Total	

Answer THREE *questions out of* FOUR.

This examination is worth 85%. The tutorials were worth 15%.

University approved calculators MAY be used.

Each answer must be completely contained within the box under the corresponding question. No credit will be given for answers presented elsewhere.

You are advised to write using a soft pencil so that you may readily correct mistakes with an eraser.

You may use a blue book for scratch—it will be discarded without being looked at.

Question 1

(a) Give six container classes that are part of the Java collection and describe briefly what they do and how they are implemented *(12 marks)*

1

2

3

4

5

6

(b) Describe the advantages of using generics rather than building containers for the type `Object` *(4 marks)*

(c) Explain how the iterator pattern works in Java collections. In particular describe the methods used in the iterator interface *(6 marks)*

(d) Explain what the `Comparable<T>` interface is and how it is used to make a class have comparable instances? *(5 marks)*

(e) Explain why it is necessary also to have a `Comparator<T>` interface and how it is used? *(6 marks)*

End of question 1

Q1: (a) $\frac{}{12}$ (b) $\frac{}{4}$ (c) $\frac{}{6}$ (d) $\frac{}{5}$ (e) $\frac{}{6}$ Total $\frac{}{33}$

• Do not write in this space •

Question 2 Merge sort has the form

```
MergeSort(A[1:n]) {
   if (n>1) {
      B = A[1:n/2]
      C = A[n/2+1:n]
      MergeSort(B)
      MergeSort(C)
      Merge(B,C,A)
   }
}
```

The number of comparison operations to merge two arrays of length $n/2$ is n.

(a) Let $C(n)$ be the number of comparison operations. Write down a recurrence relation for $C(n)$ valid if $n = 2^m$ *(4 marks)*

$C(n) =$

(b) Write down the boundary condition $C(1)$ and use the recurrence relation to compute $C(2)$, $C(4)$, and $C(8)$ *(4 marks)*

$C(1) =$

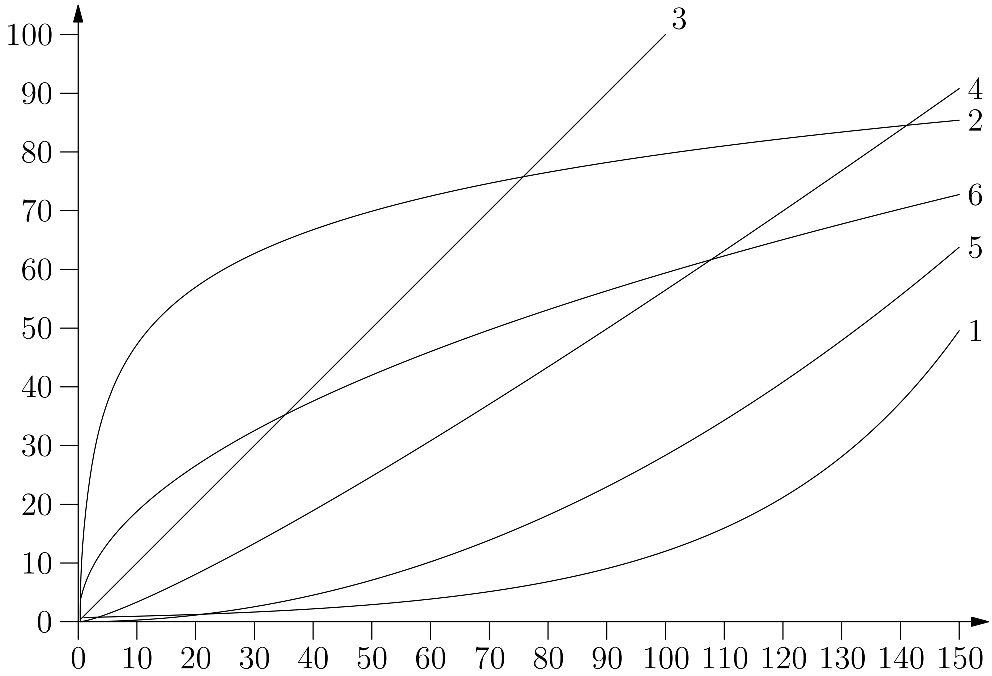
$C(2) =$

$C(4) =$

$C(8) =$

(c) Demonstrate, for $n = 2^m$, that $f(n) = n\log_2(n)$ satisfies the recurrence relation in part (a) *(6 marks)*

(d) The graph below shows the time complexity for the following algorithms (a) $\Theta(\mathrm{e}^{a\,n})$, (b) $\Theta(n^2)$, (c) $\Theta(n\,\log(n))$, (d) $\Theta(n)$, (e) $\Theta(\sqrt{n})$, and (f) $\Theta(\log(n))$. Match the time complexity classes with the curves on the graph.



(6 marks)

1.	**2.**
3.	**4.**
5.	**6.**

• Do not write in this space •

(e) Which of the following statements are true? Give reasons why (marks will only be awarded if correct reasons are given) *(8 marks)*

(i) All $\Theta(n\ \log(n))$ algorithms are faster than all $\Theta(n^2)$ algorithms

(ii) All $O(n^2)$ algorithms run slower than all $O(n)$ asymptotically

(iii) A $\Theta(n!)$ algorithm runs slower than any exponential algorithm in the limit of large n

(iv) An $O(n/\log(n))$ algorithm will run faster than a $\Omega(n\ \log(n))$ algorithm for sufficiently large n

• Do not write in this space •

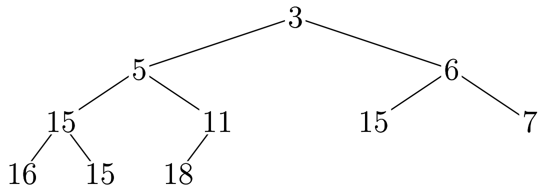
(f) Why is it widely believed that $NP \neq P$? *(5 marks)*

End of question 2

Q2:	(a) $\frac{}{4}$	(b) $\frac{}{4}$	(c) $\frac{}{6}$	(d) $\frac{}{6}$	(e) $\frac{}{8}$	(f) $\frac{}{5}$	Total $\frac{}{33}$

• Do not write in this space •

Question 3 Consider the **heap** represented as a binary tree

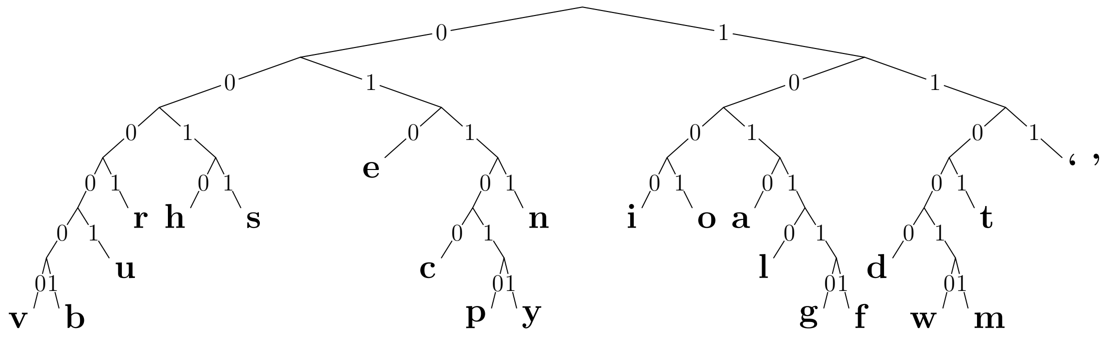


(a) Show how the heap would be stored in the compute memory *(3 marks)*

(b) Draw a binary tree representing the heap after we add 4 to it *(5 marks)*

(c) Draw a binary tree representing the heap above when we run `removeMin()` *(5 marks)*

(d) Given the Hoffman tree



decode 0110101011000001011101110010 *(6 marks)*

(e) Give a high level description (ignoring implementation details) of how a Huffman tree is constructed for the set of English letters. *(8 marks)*

(f) Describe how a heap is used to construct the Huffman tree *(4 marks)*

(g) Describe how a heap is used in Heap Sort *(2 marks)*

End of question 3

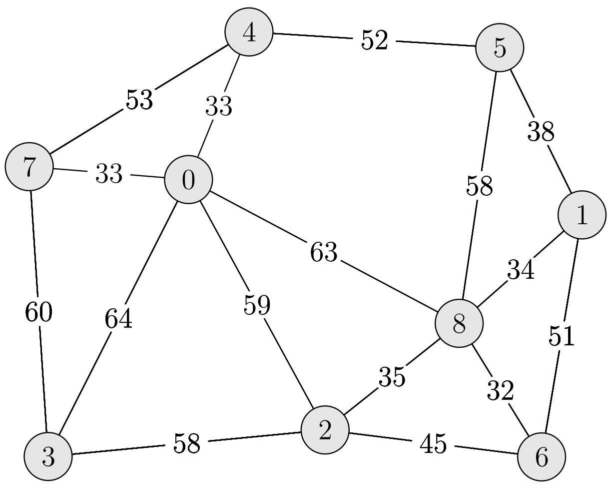
Q3: (a) $\frac{\quad}{3}$ (b) $\frac{\quad}{5}$ (c) $\frac{\quad}{5}$ (d) $\frac{\quad}{6}$ (e) $\frac{\quad}{8}$ (f) $\frac{\quad}{4}$ (g) $\frac{\quad}{2}$ Total $\frac{\quad}{33}$

• Do not write in this space •

Question 4

(a) Describe Kruskal's algorithm for finding minimal spanning trees *(10 marks)*

(b) For the following graph, show in what order Kruskal's algorithm would add edges to the minimum spanning tree



(8 marks)

1.	**2.**
3.	**4.**
5.	**6.**
7.	**8.**

The disjoint set class is described by the following program

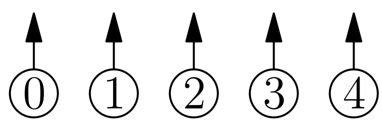
```
public class DisjSets
{
    private int[] s;

    public DisjSets(int numElements) {
        s = new int[numElements];
        for(int i=0; i<s.length; i++)
            s[i] = -1;
    }

    public void union(int root1, int root2) {
        if (s[root2]<s[root1]) {
            s[root1] = root2;
        } else {
            if (s[root1]==s[root2])
                s[root1]--;
            s[root2] = root1;
        }
    }

    public int find(int x) {
        if (s[x]<0)
            return x;
        else
            return s[x] = find(s[x]);
    }
}
```

We assume that we have created an instance of the disjoint sets class

```
DisjSets disjset = new DisjSets(5);
```

Below we show the initial settings of the array s and a graphical representation of the forest (set of trees) representing the array.

0	1	2	3	4
-1	-1	-1	-1	-1



(c) Show the state of array and the forest after performing the following operation

```
disjset.union(disjset.find(1), disjset.find(3));
```

(3 marks)

0	1	2	3	4

(d) Show the state of array and the forest after performing the following operation

```
disjset.union(disjset.find(4), disjset.find(3));
```

(3 marks)

0	1	2	3	4

(e) Show the state of array and the forest after performing the following operation

```
disjset.union(disjset.find(0), disjset.find(2));
```

(3 marks)

0	1	2	3	4

(f) Show the state of array and the forest after performing the following operation

```
disjset.union(disjset.find(4), disjset.find(2));
```

(3 marks)

0	1	2	3	4

(g) Show the state of array and the forest after performing the following operation

```
disjset.find(2);
```

(3 marks)

0	1	2	3	4

End of question 4

Q4: (a) $\frac{}{10}$ (b) $\frac{}{8}$ (c) $\frac{}{3}$ (d) $\frac{}{3}$ (e) $\frac{}{3}$ (f) $\frac{}{3}$ (g) $\frac{}{3}$ Total $\frac{}{33}$

• Do not write in this space •

END OF PAPER