CSE 5311 Name _____
Spring 2003
Test 1
Closed Book Questions
Multiple Choice - WRITE your answer to the LEFT of each problem. 4 points each.

1. The reason for marking nodes in a Fibonacci heap is:
   A. to allow computing the value of the potential function.
   B. to assure that the structure is a Fibonacci heap rather than a binomial heap.
   C. to improve the performance of CONSOLIDATE.
   D. to indicate nodes that have lost a child since becoming a child themselves.
2. Which of the following statements is true?
   A. A binary search tree may be assigned legal AVL tree balance factors if and only if it can be legally colored as a red-black tree.
   B. If a binary search tree may be assigned legal AVL tree balance factors, then it can be legally colored as a red-black tree.
   C. If a binary search tree can be legally colored as a red-black tree, then it may be assigned legal AVL tree balance factors.
   D. No binary search tree can be both assigned legal AVL tree balance factors and be legally colored as a red-black tree.
3. Which property does not hold for binomial heaps?
   A. DECREASE-KEY takes O(1) time.
   B. MINIMUM takes O(lg n) time.
   C. Performing n INSERT operations into an empty heap will take O(n) time.
   D. The number of trees is based on the binary representation of the number of stored items.
4. Which algorithm uses Union-Find trees?
   A. Dijkstra's
   B. Kruskal's
   C. Prim's
   D. The MST technique based on Warshall's algorithm.
5. Assuming that $|E| > |V|$, the time for Prim's algorithm using a Fibonacci heap is:
   A. $O(|E|)$
   B. $O(|E| \log |E|)$
   C. $O(|E| \log |V|)$
   D. $O(|V| \log |V| + |E|)$
6. Brent's rehash technique is used to improve the successful search performance of
   A. (separate) chaining
   B. double hashing
   C. linear probing
   D. optimal hashing
7. The perfect hashing method discussed in class may be viewed as improving which technique?
   A. (separate) chaining
   B. double hashing
   C. linear probing
   D. optimal hashing
8. Which of the following properties would not hold for the potential function for a data structure?
   A. Changes in potential are used in computing the amortized cost of an operation.
   B. The potential can become indefinitely large, even though the number of items stored is fixed.
   C. The potential is never negative.
   D. The potential is not stored by an implementation.
9. Which of the following is not true regarding self-adjusting lists?
   A. The count method records the number of requests for each item and approximates the optimal fixed ordering.
   B. The move-to-front method always brings the requested item to the beginning of the list.
   C. The OPT method is efficient and is a good substitute for the move-to-front method.
   D. The transpose method exchanges the requested item with the item before it unless the requested item is already at the beginning of the list.
10. The master method is most closely related to this analysis technique.
    A. Divide-and-conquer
    B. Hiring problem
    C. Recursion-tree
    D. Substitution

==========

11. Suppose a gambling game involves a sequence of rolls from a standard six-sided die. A player wins $1 when the value rolled is the same as the previous roll. If a sequence has 601 rolls, what is the expected amount paid out? (5 points)
12. Suppose a gambling game involves a sequence of rolls from a standard six-sided die. A player wins $1 when the value rolled is larger than the previous roll. If a sequence has 601 rolls, what is the expected amount paid out? (5 points)

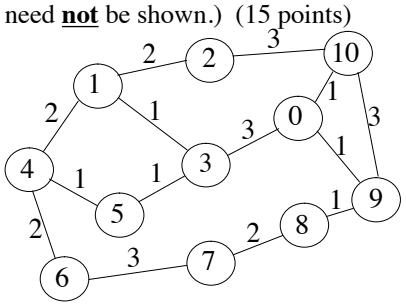
Test 1

Open Book Questions

1. Use the master method to derive an asymptotic bound on $T(n) = T(n/2) + n^2$. (10 points)
2. a. Suppose that the root of a tree in a binomial heap has 4 children. What is the minimum number of nodes that may appear in the tree? (3 points)

   b. Suppose that the rightmost path of a leftist heap has 4 nodes. What is the minimum number of nodes that may appear in the heap? (5 points)

   c. Suppose that the root of a tree in a Fibonacci heap has 4 children. What is the minimum number of nodes that may appear in the tree? (7 points)
3. Does the following graph have a unique minimum spanning tree? Trace the contents of the disjoint set structure during the execution of the modified Kruskal's algorithm that checks for uniqueness. (Galler-Fischer representation, i.e. parent pointers, need **not** be shown.) (15 points)



4. Consider the following information for constructing an optimal binary search tree.

```
n=6;
q[0]=0.01;
key[1]=10;
p[1]=0.2;
q[1]=0.02;
key[2]=20;
p[2]=0.1;
q[2]=0.03;
key[3]=30;
p[3]=0.2;
q[3]=0.04;
key[4]=40;
p[4]=0.2;
q[4]=0.0;
key[5]=50;
p[5]=0.06;
q[5]=0.04;
key[6]=60;
p[6]=0.07;
q[6]=0.03;
w[0][0]=0.010000
w[0][1]=0.230000
w[0][2]=0.360000
w[0][3]=0.600000
w[0][4]=0.800000
w[0][5]=0.900000
w[0][6]=1.000000
w[1][1]=0.020000
w[1][2]=0.150000
w[1][3]=0.390000
w[1][4]=0.590000
w[1][5]=0.690000
w[1][6]=0.790000
w[2][2]=0.030000
w[2][3]=0.270000
w[2][4]=0.470000
w[2][5]=0.570000
w[2][6]=0.670000
w[3][3]=0.040000
w[3][4]=0.240000
w[3][5]=0.340000
w[3][6]=0.440000
w[4][4]=0.000000
w[4][5]=0.100000
w[4][6]=0.200000
w[5][5]=0.040000
w[5][6]=0.140000
w[6][6]=0.030000
```

```
Building c(0,2) using roots 1 thru 2
Building c(1,3) using roots 2 thru 3
Building c(2,4) using roots 3 thru 4
Building c(3,5) using roots 4 thru 5
Building c(4,6) using roots 5 thru 6
Building c(0,3) using roots 1 thru 3
Building c(1,4) using roots 3 thru 3
Building c(2,5) using roots 3 thru 4
Building c(3,6) using roots 4 thru 6
Building c(0,4) using roots 2 thru 3
Building c(1,5) using roots 3 thru 4
Building c(2,6) using roots 4 thru 4
Building c(0,5) using roots 3 thru 3
Building c(1,6) using roots 3 thru 4
Building c(0,6) using roots ? thru ?
Counts - root trick 29 without root trick 50
Average probe length is ?
trees in parenthesized prefix
c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
c(3,3) cost 0.000000
c(4,4) cost 0.000000
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(0,1) cost 0.230000 10
c(1,2) cost 0.150000 20
c(2,3) cost 0.270000 30
c(3,4) cost 0.240000 40
c(4,5) cost 0.100000 50
c(5,6) cost 0.140000 60
c(0,2) cost 0.510000 10(,20)
c(1,3) cost 0.540000 30(20,)
c(2,4) cost 0.710000 30(,40)
c(3,5) cost 0.440000 40(,50)
c(4,6) cost 0.300000 60(50,)
c(0,3) cost 1.100000 20(10,30)
c(1,4) cost 0.980000 30(20,40)
c(2,5) cost 0.940000 40(30,50)
c(3,6) cost 0.740000 40(,60(50,))
c(0,4) cost 1.550000 30(10(,20),40)
c(1,5) cost 1.280000 30(20,40(,50))
c(2,6) cost 1.240000 40(30,60(50,))
c(0,5) cost 1.850000 30(10(,20),40(,50))
c(1,6) cost 1.630000 40(30(20,),60(50,))
c(0,6) cost ???????? ??????????????????
```

Construct the final optimal binary search tree and give its cost. SHOW YOUR WORK. (10 points)

CSE 5311 Name _____
Spring 2003
Test 2
Closed Book Questions
Multiple Choice - WRITE your answer to the LEFT of each problem. 4 points each.

1. The length of the TSP tour found by the triangle inequality technique achieves what minimization ratio?
   A. 0.5
   B. $1 + \in$
   C. 1.5
   D. 2
2. When performing bin packing using the First-Fit Decreasing technique, the total number of items placed in the bins past the optimal bins (1 .. OPT) is bounded by:
   A. $1 + \in$
   B. 2
   C. OPT - 1
   D. OPT
3. Which of the following is NOT required when showing that problem B is NP-complete by a reduction from problem A.
   A. The reduction takes polynomial time.
   B. The reduction has an inverse that takes each instance of problem B to an instance of problem A.
   C. Problem A is NP-complete
   D. The reduction must be consistent for the decision results for each instance of problem A and the corresponding instance of problem B.
4. When combining the left and right parts for the 2-d closest pair algorithm along the vertical dividing line, each left-side point ''near'' the line has its distance computed to no more than this number of right-side points.
   A. 6
   B. 8
   C. lg n
   D. n/2
5. Radix sort is useful for which of the following:
   A. Constructing a longest common subsequence using O(m+n) space.
   B. Constructing a suffix array.
   C. Constructing a suffix tree.
   D. Constructing the polynomial for Karp-Rabin string search
6. Which of the following is not required before relabeling (''lifting'') a vertex to a new height?
   A. Any eligible edges for the present height have been saturated.
   B. Both breadth-first searches have been done.
   C. The vertex is not the source or sink.
   D. The vertex is overflowing.
7. Which of the following is not true for the stable marriages technique that yields the man-optimal matching?
   A. Men propose starting with the beginning of their preference lists.
   B. The order of proposals is irrelevant and the same matching will always be obtained.
   C. When a man proposes to a woman and the woman accepts, they must be paired in the final matching.
   D. Women improve their situation by accepting proposals toward the beginning of their preference lists.
8. Among the problems listed, this problem achieves the best approximation ratio:
   A. Edge coloring
   B. Set cover
   C. TSP with the triangle inequality
   D. Vertex cover
9. Articulation points are found by:
   A. 2-d closest pairs
   B. Convex hull
   C. Depth-first search
   D. Strongly connected components
10. On an augmenting path, a critical edge is:
    A. an edge from the source
    B. an edge into the sink
    C. an edge that used to be saturated
    D. an edge with the minimum residual capacity

==========

11. Fill in the Z table. 10 points

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a	c	a	b	a	c	a	c	a	b	a	c	a	b	a	c

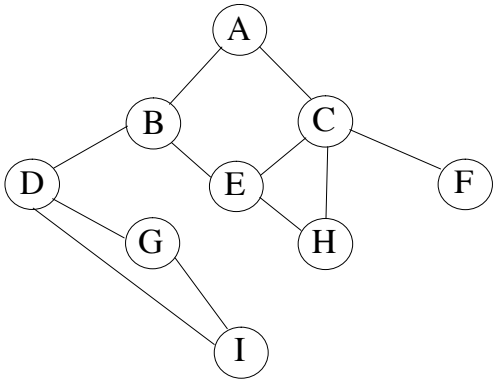
CSE 5311 Name _____
Spring 2003
Test 2
Open Book Questions

1. Give the suffix array for this string. 15 points

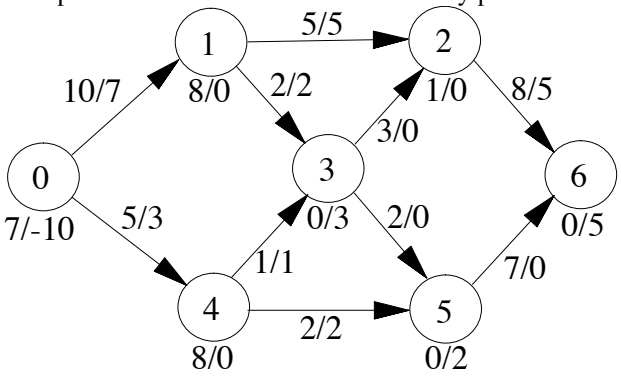
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a	c	a	b	a	c	a	c	a	b	a	c	a	b	a	c

2. Suppose an application requires finding small vertex covers for many undirected graphs. Even though you are initially concerned about this issue, you proceed with great confidence after learning that every graph will be bipartite. Why? 10 points.
3. Trace the execution of the biconnected components algorithm on the following graph by giving the discovery time and ''back'' value for each vertex, along with listing the vertices in each biconnected component. 15 points

Vertex: A B C D E F G H I
Discovery:
Back:



4. List the remaining operations to complete this instance of network flows by push-relabel. 10 points



CSE 5311 Name _____
Summer 2003
Test 1

Closed Book Questions

Multiple Choice - WRITE your answer to the LEFT of each problem. 4 points each.

1. Which of the follow is not true regarding CASCADING-CUT?

   A. It calls CONSOLIDATE.
   B. It may set or clear the marks on some nodes.
   C. It is used by DECREASE-KEY.
   D. It is used by DELETE-KEY.

2. Which of the following statements is true?

   A. A particular AVL tree may have balance factors assigned in only one way.
   B. A particular red-black tree may only be colored in one way.
   C. A rotation takes $\Omega(\lg n)$ time.
   D. A treap can be colored as a red-black tree.

3. Which property holds for binomial heaps?

   A. DECREASE-KEY takes O(1) time.
   B. MINIMUM takes O(1) time.
   C. Performing n INSERT operations into an empty heap will take O(lg n) time.
   D. The number of trees is based on the binary representation of the number of stored items.

4. What is the main contribution of leftist heaps?

   A. The minimum is found in O(1) time.
   B. The amortized complexity of DECREASE-KEY is O(1)
   C. The height of the tree is O(lg n).
   D. They were the first heap to support UNION in O(lg n) time.

5. Assuming that $|E| > |V|$, the time for Dijkstra's algorithm using a Fibonacci heap is:

   A. $O(|E|)$
   B. $O(|E| \log |E|)$
   C. $O(|E| \log |V|)$
   D. $O(|V| \log |V| + |E|)$

6. An optimal binary search tree may be found in time:

   A. O(lg n)
   B. O(n)
   C. $O(n^2)$
   D. $O(n^3)$

7. The perfect hashing method discussed in class depends on which fact?

   A. The sum $1 + 1/x + 1/x^2 + 1/x^3 + 1/x^4 + \ldots < 1/(1-x)$ when $x < 1$
   B. $\ln n < H_n < \ln n + 1$
   C. The expected number of probes for a successful search in Brent's method is less than 2.5
   D. The probability of collisions among n keys stored in a hash table of size n^2 is less than 1/2

8. What is the basis for the potential function for comparing the move-to-front strategy to the optimal off-line strategy?

   A. Dynamic programming
   B. Inversions
   C. Ranks
   D. The expected fixed optimal cost

9. Performing selection in worst-case linear time is based on separating the n input values into how many groups?

   A. 5
   B. lg n
   C. n/lg n
   D. n/5

10. The expected number of random coupons needed to obtain at least one each of n coupon types is.

    A. n
    B. H_n
    C. n ln n
    D. n^2

==========

11. Suppose a gambling game involves a sequence of rolls from a standard six-sided die. A player wins $1 when the value rolled is different from the previous roll. If a sequence has 601 rolls, what is the expected amount paid out? (5 points)

12. Suppose a gambling game involves a sequence of rolls from a standard six-sided die. A player wins k dollars when the value k rolled is smaller than the previous roll. If a sequence has 601 rolls, what is the expected amount paid out? (5 points)

Summer 2003
Test 1

Open Book Questions

1. Use the master method to derive an asymptotic bound on $L(n) = 2L(n/2) + \log n$. (10 points)

2. a. Suppose that the root of a tree in a binomial heap has 5 children. What is the minimum number of nodes that may appear in the tree? (3 points)

   b. What is the minimum number of nodes that may appear in a leftist heap whose rightmost path has 5 nodes? (5 points)

   c. Suppose that the root of a tree in a Fibonacci heap has 5 children. What is the minimum number of nodes that may appear in the tree? (7 points)

3. Give the treap (min-heap ordering) for the following keys and priorities (15 points)

key	priority
10	33
20	12
30	1
40	31
50	17
60	100
70	5
80	22

4. Consider the following information for constructing an optimal binary search tree.

```
n=6;
q[0]=0.01;
key[1]=10;
p[1]=0.29;
q[1]=0.02;
key[2]=20;
p[2]=0.1;
q[2]=0.03;
key[3]=30;
p[3]=0.2;
q[3]=0.04;
key[4]=40;
p[4]=0.2;
q[4]=0.0;
key[5]=50;
p[5]=0.02;
q[5]=0.04;
key[6]=60;
p[6]=0.02;
q[6]=0.03;
w[0][0]=0.010000
w[0][1]=0.320000
w[0][2]=0.450000
w[0][3]=0.690000
w[0][4]=0.890000
w[0][5]=0.950000
w[0][6]=1.000000
w[1][1]=0.020000
w[1][2]=0.150000
w[1][3]=0.390000
w[1][4]=0.590000
w[1][5]=0.650000
w[1][6]=0.700000
w[2][2]=0.030000
w[2][3]=0.270000
w[2][4]=0.470000
w[2][5]=0.530000
w[2][6]=0.580000
w[3][3]=0.040000
w[3][4]=0.240000
w[3][5]=0.300000
w[3][6]=0.350000
w[4][4]=0.000000
w[4][5]=0.060000
w[4][6]=0.110000
w[5][5]=0.040000
w[5][6]=0.090000
w[6][6]=0.030000
```

```
Building c(0,2) using roots 1 thru 2
Building c(1,3) using roots 2 thru 3
Building c(2,4) using roots 3 thru 4
Building c(3,5) using roots 4 thru 5
Building c(4,6) using roots 5 thru 6
Building c(0,3) using roots 1 thru 3
Building c(1,4) using roots 3 thru 3
Building c(2,5) using roots 3 thru 4
Building c(3,6) using roots 4 thru 6
Building c(0,4) using roots 1 thru 3
Building c(1,5) using roots 3 thru 4
Building c(2,6) using roots 4 thru 4
Building c(0,5) using roots 3 thru 3
Building c(1,6) using roots 3 thru 4
Building c(0,6) using roots ? thru ?
Counts - root trick 29 without root trick 50
Average probe length is ????????
trees in parenthesized prefix
c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
c(3,3) cost 0.000000
c(4,4) cost 0.000000
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(0,1) cost 0.320000 10
c(1,2) cost 0.150000 20
c(2,3) cost 0.270000 30
c(3,4) cost 0.240000 40
c(4,5) cost 0.060000 50
c(5,6) cost 0.090000 60
c(0,2) cost 0.600000 10(,20)
c(1,3) cost 0.540000 30(20,)
c(2,4) cost 0.710000 30(,40)
c(3,5) cost 0.360000 40(,50)
c(4,6) cost 0.170000 60(50,)
c(0,3) cost 1.230000 10(,30(20,))
c(1,4) cost 0.980000 30(20,40)
c(2,5) cost 0.860000 40(30,50)
c(3,6) cost 0.520000 40(,60(50,))
c(0,4) cost 1.730000 30(10(,20),40)
c(1,5) cost 1.160000 30(20,40(,50))
c(2,6) cost 1.020000 40(30,60(50,))
c(0,5) cost 1.910000 30(10(,20),40(,50))
c(1,6) cost 1.370000 30(20,40(,60(50,)))
c(0,6) cost ???????? ????????????????????
```

Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

CSE 5311 Name _____
Summer 2003
Test 2

Closed Book Questions

Multiple Choice - WRITE your answer to the LEFT of each problem. 4 points each.

1. When performing bin packing using the ''next fit'' technique, the total number of items placed in the bins past the optimal bins (1 .. OPT) is bounded by:

   A. $1 + \in$
   B. 2
   C. OPT - 1
   D. OPT

2. Each successive time an edge (x, y) becomes critical in the Edmonds-Karp method, the distances for x and y along the augmenting path increase by at least this much:

   A. 1

B. 2
C. 3
D. capacity - flow on the edge where this value is minimized.

3. The length of the TSP tour found by the triangle inequality technique achieves what minimization ratio?

A. 0.5
B. $1 + \in$
C. 1.5
D. 2

4. Which of the following is NOT required when showing that problem B is NP-complete by a reduction from problem A.

A. The reduction takes polynomial time.
B. The reduction has an inverse that takes each instance of problem B to an instance of problem A.
C. Problem A is NP-complete
D. The reduction must be consistent for the decision results for each instance of problem A and the corresponding instance of problem B.

5. Which of the following uses the most angular comparisons in the worst case to find a 2-d convex hull?

A. Ford-Fulkerson
B. Graham scan
C. Jarvis march
D. Karp-Rabin

6. An entry in the Z array contains what value?

A. A pointer or subscript for some suffix of a string.
B. The length of the longest prefix of a string that matches the string starting at this position.
C. The link for continued attempts at matching within this pattern.
D. The number of occurences of the first suffix symbol within the corresponding suffix.

7. The method for constructing a suffix array using radix sort takes worst-case time:

A. $\theta(\lg n)$
B. $\theta(n)$
C. $\theta(n \lg n)$
D. $\theta(n^2)$

8. Among the problems listed, this problem achieves the best approximation ratio:

A. Edge coloring
B. Set cover
C. TSP with the triangle inequality
D. Vertex cover

9. Articulation points are found by:

A. 2-d closest pairs
B. Convex hull
C. Depth-first search
D. Strongly connected components

10. Edge connectivity can be computed using this many instances of network flows:

A. V - 1
B. V
C. E - 1
D. E

==========

11. Explain how the relabel (‘‘lift’’) operation is applied. 10 points

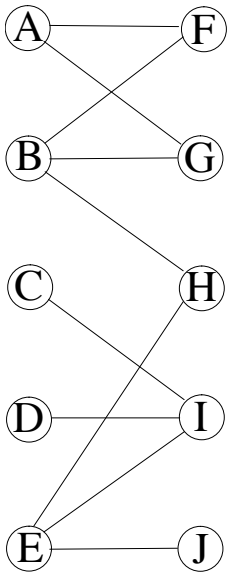
CSE 5311 Name _____
Summer 2003
Test 2

Open Book Questions

1. Fill in the Z table. 10 points

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a	b	a	b	a	c	a	b	a	b	a	c	a	b	a	b

2. List the vertices in a minimum vertex cover for this graph. 10 points.



3. In the k-subset-sum problem, it must be decided whether there is a subset with *exactly* k values chosen from a given set (that may include repeated values and zeroes) that sum to a given value. Show that this problem is NP-complete. 15 points.

4. Solve the following instance of Longest Common Subsequences using the method based on the Longest Strictly Increasing Subsequence problem. 15 points.

0	1	2	3	4	5	6	7	8	9	10	11
a	b	c	d	a	b	c	d	a	b	c	d
d	c	b	a	d	c	b	a	d	c	b	a

CSE 5311 Name _____
Test 1 - Closed Book
Spring 2004 Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 4 points each

1. To reduce the probability of having any collisions to < 0.5 when hashing n keys, the table should have at least this number of elements.

   A. n
   B. n ln n
   C. n^2
   D. n^3

2. Path compression is used in which algorithm?

   A. DECREASE-KEY (e.g. CASCADING-CUT) for Fibonacci heaps
   B. FIND for disjoint sets
   C. UNION for disjoint sets
   D. UNION for Fibonacci heaps

3. When performing selection in worst-cast linear time, roughly what fraction of the set of n keys is kept (in the worst case) for the next round?

   A. 10%
   B. 20%
   C. 30%
   D. 70%

4. Which Fibonacci heap operation has O(log n) actual cost?

   A. FIB-HEAP-DECREASE-KEY
   B. FIB-HEAP-DELETE
   C. FIB-HEAP-EXTRACT-MIN
   D. FIB-HEAP-UNION

5. The main difference between MTF and OPT for self-organizing linear lists is:

   A. MTF counts inversions
   B. MTF is given the entire request sequence in advance, while OPT receives the requests one-at-a-time
   C. OPT is given the entire request sequence in advance, while MTF receives the requests one-at-a-time
   D. OPT can do transpositions

6. The minimum number of nodes in a tree in a Fibonacci heap where the root has 7 children is:

   A. 34
   B. 61
   C. 121
   D. 128

7. The maximum value of the potential function when comparing MTF and OPT on a list with n elements is:

   A. n
   B. $\frac{n^2-n}{2}$
   C. $\frac{n^2+n}{2}$
   D. n^2-n

8. In the worst case, the number of rotations for inserting a key in a treap with n keys is:

   A. θ(1)
   B. θ(log n)
   C. θ(n)
   D. θ(n log n)

9. Slow convergence toward the optimal fixed ordering is a property of which technique?

   A. Count
   B. Move-ahead-k
   C. Move-to-front
   D. Transpose

10. Sorting the edges is a property of which minimum spanning tree technique?

   A. Boruvka
   B. Kruskal
   C. Prim
   D. Path-based (Warshall)

Long Answer

1. Suppose there are 20 coupon types for the coupon collecting problem and you have already obtained 18 of the coupon types. How many boxes of cereal do you expect (mathematically) to open to get the remaining two coupon types? (5 points)

2. Suppose there are 20 coupon types for the coupon collecting problem. Each of the coupon types is identified by an integer in the range 1 to 20. If you are given two random cereal boxes, what is the probability (ignoring the order in which the coupons were obtained) that the two coupons have consecutive numbers, i.e. some i and i + 1? (5 points)

CSE 5311 Name _____
Test 1 - Open Book
Spring 2004 Student ID # _____

1. Suppose you wish to implement a verifier to test if a given subset (with V - 1 edges) of the set of edges for a graph is a minimum spanning tree. Assuming you already have the code below for Kruskal's algorithm, explain how it should be modified to perform this verification. You may assume that each `struct` in `edgeTab` has a field `subset` that flags whether this edge is in the alleged MST. (15 points)

```
. . .

main()
{
. . .

qsort(edgeTab,numEdges,sizeof(edgeType),weightAscending);
for (i=0;i<numEdges;i++)
{
  root1=find(edgeTab[i].tail);
  root2=find(edgeTab[i].head);
  if (root1==root2)
    printf("%d %d %d discarded\n",edgeTab[i].tail,edgeTab[i].head,
      edgeTab[i].weight);
  else
  {
    printf("%d %d %d included\n",edgeTab[i].tail,edgeTab[i].head,
      edgeTab[i].weight);
    makeEquivalent(root1,root2);
  }
}
if (numTrees!=1)
  printf("MST does not exist\n");
}
```

2. Evaluate the following recurrences using the master method. (15 points)

a. $T(n) = T(0.7n) + n$

b. $T(n) = T(0.7n) + 1$

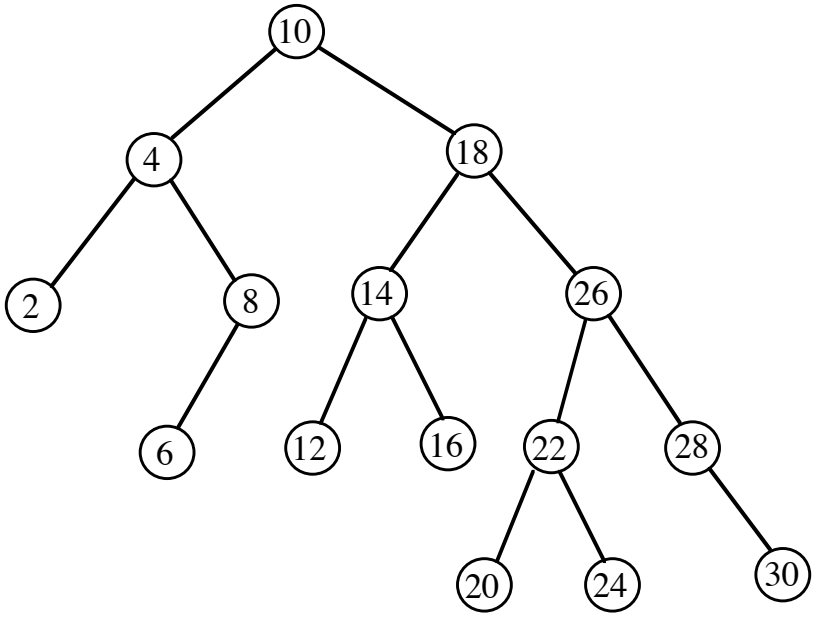
c. $T(n) = 16T\left(\frac{n}{2}\right) + n^3$

3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
n=6;
q[0]=0.01;
key[1]=10;
p[1]=0.19;
q[1]=0.02;
key[2]=20;
p[2]=0.1;
q[2]=0.03;
key[3]=30;
p[3]=0.2;
q[3]=0.04;
key[4]=40;
p[4]=0.2;
q[4]=0.0;
key[5]=50;
p[5]=0.02;
q[5]=0.04;
key[6]=60;
p[6]=0.12;
q[6]=0.03;
w[0][0]=0.010000
```

```
w[0][1]=0.220000
w[0][2]=0.350000
w[0][3]=0.590000
w[0][4]=0.790000
w[0][5]=0.850000
w[0][6]=1.000000
w[1][1]=0.020000
w[1][2]=0.150000
w[1][3]=0.390000
w[1][4]=0.590000
w[1][5]=0.650000
w[1][6]=0.800000
w[2][2]=0.030000
w[2][3]=0.270000
w[2][4]=0.470000
w[2][5]=0.530000
w[2][6]=0.680000
w[3][3]=0.040000
w[3][4]=0.240000
w[3][5]=0.300000
w[3][6]=0.450000
w[4][4]=0.000000
w[4][5]=0.060000
w[4][6]=0.210000
w[5][5]=0.040000
w[5][6]=0.190000
w[6][6]=0.030000
Building c(0,2) using roots 1 thru 2
Building c(1,3) using roots 2 thru 3
Building c(2,4) using roots 3 thru 4
Building c(3,5) using roots 4 thru 5
Building c(4,6) using roots 5 thru 6
Building c(0,3) using roots 1 thru 3
Building c(1,4) using roots 3 thru 3
Building c(2,5) using roots 3 thru 4
Building c(3,6) using roots 4 thru 6
Building c(0,4) using roots 2 thru 3
Building c(1,5) using roots 3 thru 4
Building c(2,6) using roots 4 thru 4
Building c(0,5) using roots 3 thru 3
Building c(1,6) using roots 3 thru 4
Building c(0,6) using roots ? thru ?
Counts - root trick 29 without root
    trick 50
Average probe length is ????
trees in parenthesized prefix
c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
c(3,3) cost 0.000000
c(4,4) cost 0.000000
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(0,1) cost 0.220000 10
c(1,2) cost 0.150000 20
c(2,3) cost 0.270000 30
c(3,4) cost 0.240000 40
c(4,5) cost 0.060000 50
c(5,6) cost 0.190000 60
c(0,2) cost 0.500000 10(,20)
c(1,3) cost 0.540000 30(20,)
c(2,4) cost 0.710000 30(,40)
c(3,5) cost 0.360000 40(,50)
c(4,6) cost 0.270000 60(50,)
c(0,3) cost 1.080000 20(10,30)
c(1,4) cost 0.980000 30(20,40)
c(2,5) cost 0.860000 40(30,50)
c(3,6) cost 0.720000 40(,60(50,))
c(0,4) cost 1.530000 30(10(,20),40)
c(1,5) cost 1.160000 30(20,40(,50))
c(2,6) cost 1.220000 40(30,60(50,))
c(0,5) cost 1.710000 30(10(,20),40(,50))
c(1,6) cost 1.610000 40(30(20,),60(50,))
c(0,6) cost ???????? ????????????
```

4. Insert 23 into the following AVL tree to preserve the AVL properties. (10 points)



CSE 5311
Test 2 - Closed Book
Spring 2004

Name _____

Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. Which of the following problems is not NP-complete? (Assume P ≠ NP)

   A. 3-satisfiability
   B. Testing if a graph is 3-colorable
   C. Testing if a table is in sorted order
   D. Testing if the number of colors needed to edge color a graph is the degree of the graph

2. How many times will -1 occur in the style 1 fail link table for the pattern `abcabd`?

   A. 1
   B. 2
   C. 3
   D. 4

3. How many times will -1 occur in the style 2 fail link table for the pattern `abcabd`?

   A. 1
   B. 2
   C. 3
   D. 4

4. Suppose that you have devised a reduction from a known NP-complete problem A to a problem B in NP. Later you discover that the reduction takes exponential time. What conclusion may be drawn?

   A. P ≠ NP
   B. Problem B is not NP-complete.
   C. You need a reduction that takes linear time to prove that B is NP-complete.
   D. You need a reduction that takes polynomial time to prove that B is NP-complete.

5. Which of the following is not part of the Gale-Shapley algorithm?

   A. Men issue proposals
   B. Finding rotations
   C. Women break engagements to accept proposals from their more-preferred men
   D. Women accept proposals if unengaged

6. Which of the following is not true regarding Strassen's algorithm?

   A. It is not possible to have an asymptotically faster algorithm.
   B. It requires more space than the everyday method.
   C. It uses $\Theta\left(n^{\lg 7}\right)$ scalar additions when multiplying two n x n matrices.
   D. It uses $\Theta\left(n^{\lg 7}\right)$ scalar multiplications when multiplying two n x n matrices.

7. Which algorithm is defined using the notions of left-turn and right-turn?

   A. Closest points in 2-d space

B. Graham scan
C. Jarvis march
D. Suffix array construction

8. Which of the following is not a condition for performing a push operation from u to v?

A. the height at u is one more than the height at v
B. (u, v) has capacity > flow
C. u has excess > 0
D. v has a path to the sink in the residual graph

9. Which of the following does not have an approximation algorithm?

A. Bin packing
B. Edge coloring
C. Traveling salesperson
D. Vertex covering

10. The length of a longest strictly increasing subsequence for `1 4 3 2 3 4 3 6 3 4` is:

A. 4
B. 5
C. 6
D. 7

11. Which longest common subsequence method is potentially the most time-consuming?

A. Compact version of dynamic programming
B. Method based on subsequence indices and longest strictly increasing subsequence
C. Ordinary dynamic programming

12. Consider the technique for determining articulation points using depth-first search. If a vertex has no predecessors (first vertex discovered for a "restart"), it can be an articulation point only if

A. there are no outgoing tree edges
B. there is one outgoing tree edge
C. there is more than one outgoing tree edge
D. there is no back edge returning to this vertex

13. Which of the following problems uses some edges with non-unit capacity when translated to a network flow problem?

A. Bipartite matching
B. Edge connectivity
C. Minimum vertex cover for bipartite graph
D. Vertex connectivity

14. Which string search method is potentially the most time-consuming?

A. Karp-Rabin
B. KMP with fail 1 links
C. KMP with fail 2 links
D. Z table

15. What is the maximum number of times that an edge can become critical in the Edmonds-Karp method?

A. Once
B. Twice
C. (V-2)/2
D. VE

16. Suppose you have a set of points in Euclidean 2-d space. Give an algorithm for finding a ρ-approximation for the minimum traveling salesperson path. Be sure to give the value of ρ. (5 points)

CSE 5311
Test 2 - Open Book
Spring 2004

Name _____

Student ID # _____

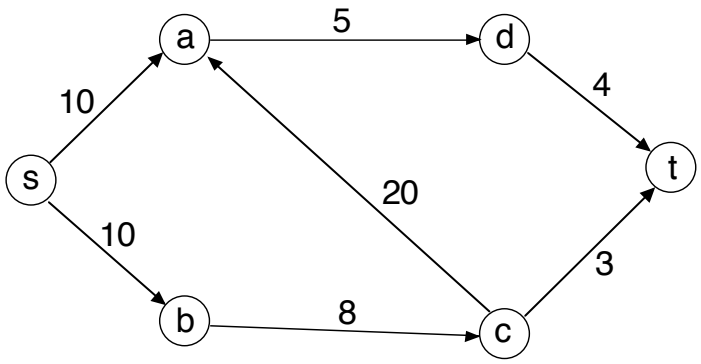
1. Use the Gale-Shapley algorithm to determine the male-optimal solution for the following instance of the stable marriages problem. In addition, show the preference lists at termination. Note that the preference lists are given left-to-right. (15 points)

```
male preference lists are:

1:  1  2  3  4  5

2:  2  3  4  5  1

3:  3  4  2  5  1

4:  1  3  2  5  4

5:  3  5  4  2  1

female preference lists are:

1:  4  3  5  2  1

2:  2  3  5  4  1

3:  1  2  5  4  3

4:  3  2  4  5  1

5:  2  4  3  1  5
```

2. The *hitting set* problem gives a collection C of subsets of a set S and a positive integer k. We would like to know if there is a subset S' of S with $|S'| \leq k$ such that S' contains at least one element from each subset in C.

   Give a proof that *hitting set* is NP-complete by using the fact that vertex cover is NP-complete. (15 points)

3. List the lift and push operations for the preflow-push algorithm on the following network. In addition, give a minimum cut. (20 points)



CSE 5311 Name _____

Test 1 - Closed Book

Summer 2004 Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 4 points each

1. Which of the following random permutation techniques is not uniform?
   A. PERMUTE-BY-SORTING
   B. PERMUTE-WITH-ALL
   C. RANDOMIZE-IN-PLACE
   D. Randomly choose one of the n! ranks and then apply unranking
2. Fibonacci trees are used in the analysis of which technique?
   A. AVL
   B. red-black
   C. splay
   D. treap
3. Which analysis of self-organizing linear search did not involve probabilities?
   A. Count vs. optimal fixed order
   B. Lab Assignment 1
   C. Markov analysis of move-to-front for Zipf's distribution
   D. Move-to-front (online) vs. optimal offline
4. The median of a set of n numbers may be found optimally in time:
   A. $\theta(\log n)$
   B. $\theta(n)$
   C. $\theta(n \log n)$
   D. $\theta(n^2)$
5. To support computing prefix sums of all keys that are no larger than some query key, an augmented binary search tree stores the following at every node:
   A. the sum of all keys in the entire tree
   B. the sum of all keys in the left subtree
   C. the sum of all keys that are no larger than the stored key
   D. the sum of all keys stored in the subtree rooted by this node
6. When using Brent's rehash, the number of previously inserted keys that may move is:
   A. 1
   B. 2
   C. $\frac{1}{\alpha}$
   D. H_m, where m is the number of stored keys
7. Assuming that a random n-permutation is provided, the expected number of hires for the hiring problem is:
   A. 2
   B. H_n
   C. $\frac{1}{n}$
   D. n
8. Which data structure is not used to implement a dictionary?
   A. AVL tree
   B. Red-black tree
   C. Self-organizing list
   D. Union-find

9. Which binary search tree method stores the same information in each node as an unbalanced binary search tree, yet performs retrieval, insertion, and deletion in O(log n) amortized time?
   A. AVL
   B. red-black
   C. splay
   D. treap
10. Which priority queue implementation generalizes binary heaps by increasing the branching?
   A. Binomial heaps
   B. d-heaps
   C. Fibonacci heaps
   D. Leftist heaps
11. Suppose you already have 15 different coupons when there are 20 coupon types. What is the expected number of boxes for obtaining a coupon different from the 15 you already have?
   A. 3
   B. 4
   C. 5
   D. 15
12. When is path compression used?
   A. After an insertion into any type of balanced binary search tree.
   B. After an insertion into a splay tree.
   C. After a FIND operation.
   D. After a UNION operation.

13. How many nodes does a B_5 tree in a binomial heap have? (2 points)

CSE 5311 Name _____
Test 1 - Open Book
Summer 2004 Student ID # _____

1. Give the range of possible heights for an AVL tree with 100 keys. (10 points)
2. Evaluate the following recurrences using the master method. Indicate the case that is used for each. (15 points)

a. $T(n) = 2T\left(\frac{n}{4}\right) + 1$

b. $T(n) = 2T\left(\frac{n}{4}\right) + \sqrt{n}$

c. $T(n) = 2T\left(\frac{n}{4}\right) + n^2$

3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (15 points)

```
n=6;
q[0]=0.01;
key[1]=10;
p[1]=0.09;
q[1]=0.02;
key[2]=20;
p[2]=0.1;
q[2]=0.03;
key[3]=30;
p[3]=0.2;
q[3]=0.04;
key[4]=40;
p[4]=0.2;
q[4]=0.0;
key[5]=50;
p[5]=0.12;
q[5]=0.04;
key[6]=60;
p[6]=0.12;
q[6]=0.03;
w[0][0]=0.010000
w[0][1]=0.120000
w[0][2]=0.250000
w[0][3]=0.490000
w[0][4]=0.690000
w[0][5]=0.850000
w[0][6]=1.000000
w[1][1]=0.020000
w[1][2]=0.150000
w[1][3]=0.390000
w[1][4]=0.590000
w[1][5]=0.750000
w[1][6]=0.900000
```

```
w[2][2]=0.030000
w[2][3]=0.270000
w[2][4]=0.470000
w[2][5]=0.630000
w[2][6]=0.780000
w[3][3]=0.040000
w[3][4]=0.240000
w[3][5]=0.400000
w[3][6]=0.550000
w[4][4]=0.000000
w[4][5]=0.160000
w[4][6]=0.310000
w[5][5]=0.040000
w[5][6]=0.190000
w[6][6]=0.030000
Building c(0,2)
    using roots 1
    thru 2
Building c(1,3)
    using roots 2
    thru 3
Building c(2,4)
    using roots 3
    thru 4
Building c(3,5) using roots 4 thru 5
Building c(4,6) using roots 5 thru 6
Building c(0,3) using roots 2 thru 3
Building c(1,4) using roots 3 thru 3
Building c(2,5) using roots 3 thru 4
Building c(3,6) using roots 4 thru 6
Building c(0,4) using roots 3 thru 3
Building c(1,5) using roots 3 thru 4
Building c(2,6) using roots 4 thru 5
Building c(0,5) using roots 3 thru 4
Building c(1,6) using roots 4 thru 4
Building c(0,6) using roots ? thru ?
Counts - root trick 28 without root
    trick 50
Average probe length is ???
trees in parenthesized prefix
c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
c(3,3) cost 0.000000
c(4,4) cost 0.000000
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(0,1) cost 0.120000 10
c(1,2) cost 0.150000 20
c(2,3) cost 0.270000 30
c(3,4) cost 0.240000 40
c(4,5) cost 0.160000 50
c(5,6) cost 0.190000 60
c(0,2) cost 0.370000 20(10,)
c(1,3) cost 0.540000 30(20,)
c(2,4) cost 0.710000 30(,40)
c(3,5) cost 0.560000 40(,50)
c(4,6) cost 0.470000 60(50,)
c(0,3) cost 0.860000 30(20(10,),)
c(1,4) cost 0.980000 30(20,40)
c(2,5) cost 1.060000 40(30,50)
c(3,6) cost 0.980000 50(40,60)
c(0,4) cost 1.300000 30(20(10,),40)
c(1,5) cost 1.450000 40(30(20,),50)
c(2,6) cost 1.520000 40(30,60(50,))
c(0,5) cost 1.780000 30(20(10,),40(,50))
c(1,6) cost 1.910000 40(30(20,),60(50,))
c(0,6) cost ??? ??????????
```

4. Suppose all 2^k - 1 nodes, along with the sentinel, in a red-black tree are colored black. Explain what will happen if any key is deleted. (10 points)

CSE 5311 Name _____

Test 2 - Closed Book

Summer 2004 Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. Which of the following problems is not NP-complete? (Assume P ≠ NP)
   A. Testing if a graph is 3-colorable
   B. Testing if an undirected graph has a Hamiltonian cycle
   C. Testing if the number of colors needed to edge color a graph is the degree of the graph
   D. 2-satisfiability
2. How many times will -1 occur in the style 1 fail link table for the pattern `abaabaab`?
   A. 1
   B. 2
   C. 3
   D. 4
3. How many times will -1 occur in the style 2 fail link table for the pattern `abaabaab`?
   A. 1
   B. 2
   C. 3
   D. 4
4. Which of the following do Kruskal's and Boruvka's algorithms have in common?
   A. They can take advantage of Union-Find structures
   B. They do not work correctly unless all edge weights are unique
   C. They use min-heaps

D. They use the edges in ascending weight order

5. Under what condition does an instance of stable marriages have only one solution?
   A. Every male preference list is identical to some female preference list
   B. No female appears at the beginning of multiple male preference lists
   C. The male-optimal solution and female-optimal solution are the same
   D. There is only one rotation
6. Which of the following algorithms does not preprocess using a sort?
   A. Closest points in 2-d space
   B. Graham scan
   C. Jarvis march
   D. Kruskal
7. Which algorithm is defined using the notions of left-turn and right-turn?
   A. Closest points in 2-d space
   B. Graham scan
   C. Jarvis march
   D. Longest common subsequence
8. Which of the following is not a condition for performing a lift operation at u?
   A. all exiting saturated edges have heads with height ≥ height of u
   B. all exiting unsaturated edges have heads with height ≥ height of u
   C. u has excess > 0
9. Which of the following does not have a polynomial-time approximation algorithm?
   A. Bin packing
   B. Edge coloring
   C. Traveling salesperson with triangle inequality
   D. Vertex coloring
10. The length of a longest monotone increasing subsequence for `1 4 3 2 3 4 3 6 3 4` is:
   A. 4
   B. 5
   C. 6
   D. 7
11. Which longest common subsequence method is potentially the most space-consuming?
   A. Compact version of dynamic programming
   B. Method based on subsequence indices and longest strictly increasing subsequence
   C. Ordinary dynamic programming
12. Consider the technique for determining articulation points using depth-first search. If a vertex has no predecessors (first vertex discovered for a "restart"), it can be an articulation point only if
   A. there are no incident tree edges
   B. there is one incident tree edge
   C. there is more than one incident tree edge
   D. there is no back edge incident to this vertex
13. Which minimum spanning tree algorithm is the slowest?
   A. Boruvka
   B. Kruskal
   C. Prim
   D. Warshall
14. Which string search method is potentially the most time-consuming?
   A. Karp-Rabin
   B. KMP with fail 1 links
   C. KMP with fail 2 links

D. Z table

15. What is the minimum increase in the tail's distance from the source between the first and second times that an edge becomes critical in the Edmonds-Karp method?
   A. 1
   B. 2
   C. (V-2)/2
   D. VE

16. Explain how the Z Algorithm ("Fundamental String Preprocessing") may be used to find all occurences of string 1 within string 2. (5 points)

CSE 5311 Name _____

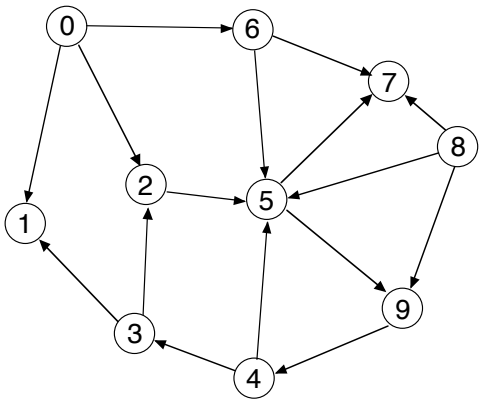
Test 2 - Open Book

Summer 2004 Student ID # _____

1. Use the Gale-Shapley algorithm to determine the male-optimal solution for the following instance of the stable marriages problem. In addition, show the preference lists at termination. Note that the preference lists are given left-to-right. (15 points)

```
male preference lists are:
1: 1 2 3 4 5
2: 3 2 1 5 4
3: 3 4 2 5 1
4: 1 3 2 5 4
5: 2 3 4 5 1
female preference lists are:
1: 4 5 3 2 1
2: 1 2 3 4 5
3: 2 3 4 5 1
4: 3 2 1 5 4
5: 4 5 2 1 3
```

2. Perform depth-first search on the following graph, including start/finish times and edge types (T=tree, B=back, C=cross, F=forward. Assume that the adjacency lists are ***ordered***. Write your answer in the tables below. 15 points



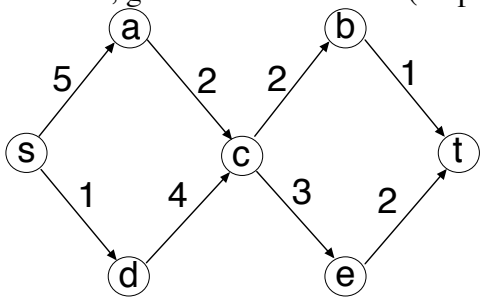
Vertex	Start	Finish
0	1	____
1	____	____
2	____	____
3	____	____
4	____	____
5	____	____
6	____	____
7	____	____

Edge	Type
0 1	____
0 2	____
0 6	____
2 5	____
3 1	____
3 2	____
4 3	____
4 5	____

Edge	Type
6 5	____
6 7	____
8 5	____
8 7	____
8 9	____
9 4	____

8	____	____	5	7	____
9	____	____	5	9	____

3. ***Clearly*** list the lift and push operations for the preflow-push algorithm on the following network. In addition, give a minimum cut. (20 points)



CSE 5311 Name _____
Test 1 - Closed Book
Summer 2005 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 4 points each

1. Assuming that |E| > |V|, the time for Prim's algorithm using a Fibonacci heap is:
   A. O(|E|)
   B. O(|E| log |E|)
   C. O(|E| log |V|)
   D. O(|V| log |V| + |E|)
2. What is the worst-case number of rotations when performing deletion on an AVL tree?
   A. $\Theta(1)$
   B. $\Theta(\log n)$
   C. $\Theta(n)$
   D. No rotations are ever used
3. The maximum value of the potential function when comparing MTF and OPT on a list with *n* elements is:
   A. n
   B. $\frac{n^2-n}{2}$
   C. $\frac{n^2+n}{2}$
   D. n^2-n
4. When performing selection in worst-cast linear time, roughly what fraction of the set of n keys is kept (in the worst case) for the next round?
   A. 10%
   B. 20%
   C. 30%
   D. 70%
5. To support computing prefix sums of all keys that are no larger than some query key, an augmented binary search tree stores the following at every node:
   A. the sum of all keys in the entire tree
   B. the sum of all keys in the left subtree
   C. the sum of all keys that are no larger than the stored key
   D. the sum of all keys stored in the subtree rooted by this node
6. The worst-case placement of keys for an open-address method may be found using:
   A. Binary search on instances of unweighted bipartite matching
   B. Brent's method
   C. Perfect hashing
   D. Weighted bipartite matching by placing weight jP_i at $M_{iS_{ij}}$ for the probe sequence $S_{i1}, S_{i2}, \cdots, S_{im}$
7. Which of the following statements is true?
   A. A binary search tree may be assigned legal AVL balance factors if and only if it may be legally colored as a red-black tree.
   B. If a binary search tree may be assigned legal AVL balance factors, then it may be legally colored as a red-black tree.
   C. If a binary search tree may be legally colored as a red-black tree, then it may be assigned legal AVL balance factors.
   D. No binary search tree may be assigned both legal AVL balance factors and be legally colored as a red-black tree.
8. Path compression is used in which algorithm?
   A. DECREASE-KEY (e.g. CASCADING-CUT) for Fibonacci heaps
   B. FIND for disjoint sets
   C. UNION for disjoint sets
   D. UNION for Fibonacci heaps
9. Which priority queue is defined using the notion of null path length?
   A. Binary heap
   B. Binomial heap
   C. Fibonacci heap
   D. Leftist heap
10. Which situation is true regarding a cascade cut that produces c trees for a Fibonacci heap?
    A. Both the actual and amortized costs are O(1).

B. The actual cost is O(c). The amortized cost is O(1).
C. The actual cost is O(1). The amortized cost is O(c).
D. The potential can become negative.

11. Which of the following is not true regarding the amortized analysis of binary tree traversals?
A. INIT had an amortized cost of 0.
B. SUCC had an actual cost determined by the number of edges followed.
C. SUCC had an amortized cost of 2.
D. The potential was defined without regard for the type of traversal being performed.

12. Which of the following data structures offers similar capabilities and performance characteristics to skip lists?
A. AVL trees
B. Splay trees
C. Treap
D. Union-find with path compression

13. Draw an S_5 tree for a Fibonacci heap. (2 points)

CSE 5311 Name _____

Test 1 - Open Book

Summer 2005 Last 4 Digits of Student ID # _____

1. "During Disney's biggest celebration, find one of 50 character Wobblers inside specially-marked ***Kellogg's***® cereals!".

a. Assuming all Wobblers are equally likely to be the one that occurs in a box, what is the expected number of boxes to obtain all 50 Wobblers ? (3 points. You may leave your answer as an expression.)

b. Under the same assumption as a., what is the expected number of boxes to obtain just 25 different Wobblers (of the available 50)? (7 points. You may leave your answer as an expression.)

2. Evaluate the following recurrences using the master method. Indicate the case that is used for each. (10 points)

a. $T(n) = 8T\left(\frac{n}{2}\right) + n^4$

b. $T(n) = 8T\left(\frac{n}{2}\right) + n^2$

c. $T(n) = 8T\left(\frac{n}{2}\right) + n^3$

3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
n=6;
q[0]=0.12;
key[1]=10;
p[1]=0.12;
q[1]=0.09;
key[2]=20;
p[2]=0.2;
q[2]=0.03;
key[3]=30;
p[3]=0.1;
q[3]=0.04;
key[4]=40;
p[4]=0.2;
q[4]=0.0;
key[5]=50;
p[5]=0.01;
q[5]=0.04;
key[6]=60;
p[6]=0.03;
q[6]=0.02;
w[0][0]=0.120000
w[0][1]=0.330000
w[0][2]=0.560000
w[0][3]=0.700000
w[0][4]=0.900000
w[0][5]=0.950000
w[0][6]=1.000000
w[1][1]=0.090000
w[1][2]=0.320000
w[1][3]=0.460000
w[1][4]=0.660000
w[1][5]=0.710000
w[1][6]=0.760000
w[2][2]=0.030000
w[2][3]=0.170000
w[2][4]=0.370000
w[2][5]=0.420000
w[2][6]=0.470000
w[3][3]=0.040000
w[3][4]=0.240000
w[3][5]=0.290000
w[3][6]=0.340000
w[4][4]=0.000000
w[4][5]=0.050000
w[4][6]=0.100000
w[5][5]=0.040000
w[5][6]=0.090000
w[6][6]=0.020000
Building c(0,2) using roots 1 thru 2
Building c(1,3) using roots 2 thru 3
Building c(2,4) using roots 3 thru 4
Building c(3,5) using roots 4 thru 5
Building c(4,6) using roots 5 thru 6
Building c(0,3) using roots 1 thru 2
Building c(1,4) using roots 2 thru 4
Building c(2,5) using roots 4 thru 4
Building c(3,6) using roots 4 thru 6
Building c(0,4) using roots 2 thru 2
Building c(1,5) using roots 2 thru 4
Building c(2,6) using roots 4 thru 4
Building c(0,5) using roots 2 thru 2
Building c(1,6) using roots 2 thru 4
Building c(0,6) using roots ? thru ?
Counts - root trick 31 without root trick
    50
Average probe length is ???
```

```
trees in
    parenthesized
    prefix

c(0,0) cost
0.000000
c(1,1) cost
0.000000
c(2,2) cost
0.000000
c(3,3) cost
0.000000
c(4,4) cost
0.000000
```

```
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(0,1) cost 0.330000 10
c(1,2) cost 0.320000 20
c(2,3) cost 0.170000 30
c(3,4) cost 0.240000 40
c(4,5) cost 0.050000 50
c(5,6) cost 0.090000 60
c(0,2) cost 0.880000 10(,20)
c(1,3) cost 0.630000 20(,30)
c(2,4) cost 0.540000 40(30,)
c(3,5) cost 0.340000 40(,50)
c(4,6) cost 0.150000 60(50,)
```

```
c(0,3) cost 1.200000 20(10,30)
c(1,4) cost 1.200000 20(,40(30,))
c(2,5) cost 0.640000 40(30,50)
c(3,6) cost 0.490000 40(,60(50,))
c(0,4) cost 1.770000 20(10,40(30,))
c(1,5) cost 1.350000 20(,40(30,50))
c(2,6) cost 0.790000 40(30,60(50,))
c(0,5) cost 1.920000 20(10,40(30,50))
c(1,6) cost 1.540000 40(20(,30),60(50,))
c(0,6) cost ??? ????????????????
```

4. The hash table below was created using double hashing with Brent's rehash. The initial slot ($h_1(key)$) and rehashing increment ($h_2(key)$) are given for each *key*. Show the result from inserting 1300 using Brent's rehash when $h_1(1300) = 5$ and $h_2(1300) = 3$. (10 points)

	key	$h_1(key)$	$h_2(key)$
0			
1	1000	6	2
2			
3	1200	3	3
4	500	3	1
5	12	5	1
6	27	6	3

5. Give an instance of deletion from a red-black tree that will require the maximum number of rotations. Also, explain how the rotations will occur using the four cases. (10 points)

CSE 5311 Name _____
Test 2 - Closed Book
Summer 2005 Last 4 Digits of Student ID # _____
Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. Which of the following problems is not NP-complete? (Assume P ≠ NP)
   A. 3-satisfiability
   B. Testing if a bipartite graph has a vertex cover with no more than k vertices
   C. Testing if a graph is 3-colorable
   D. Testing if the number of colors needed to edge color a graph is the degree of the graph
2. How many times will `-1` occur in the style 1 fail link table for the pattern `acaababc`?
   A. 1
   B. 2
   C. 3
   D. 4
3. How many times will `-1` occur in the style 2 fail link table for the pattern `acaababc`?
   A. 1
   B. 2
   C. 3
   D. 4
4. Radix sort is useful for:

A. Constructing a longest common subsequence using O(m + n) space
B. Constructing a suffix array
C. Constructing the polynomial for Karp-Rabin string search
D. Constructing the Z table for a sequence

5. Which of the following is a deficiency of the maximum capacity path technique?
   A. An augmenting path is blocked if it introduces a cycle of flow.
   B. Augmenting paths will be discovered in descending incremental flow increase order.
   C. Flow decomposition must be applied.
   D. The maximum number of potential augmenting paths depends on the achievable flow, in addition to the number of vertices and edges.
6. Which of the following algorithms does not preprocess using a sort?
   A. Closest points in 2-d space - divide and conquer
   B. Closest points in 2-d space - sweepline
   C. Graham scan
   D. Jarvis march
7. Which algorithm is defined using the notions of left-turn and right-turn?
   A. Closest points in 2-d space
   B. Graham scan
   C. Jarvis march
   D. TSP approximation when vertex distances observe the triangle inequality
8. Which of the following is not a condition for performing a push operation from u to v?
   A. the height at u is one more than the height at v
   B. (u, v) has capacity > flow
   C. u has excess > 0
   D. v has a path to the sink in the residual graph
9. Which of the following does not have a polynomial-time approximation algorithm?
   A. Bin packing
   B. Edge coloring
   C. General traveling salesperson
   D. Vertex cover
10. The length of a longest strictly increasing subsequence for `1 4 3 2 3 4 3 6 3 4` is:
    A. 4
    B. 5
    C. 6
    D. 7
11. Which of the following is not true regarding the reduction from 3-sat to 3-colorability?
    A. The widget appears once for each literal.
    B. The widget will use three colors when corresponding to a satisfied clause.
    C. The widget will use four colors when corresponding to an unsatisfied clause.
    D. There is a triangle of vertices for establishing the colors of TRUE, FALSE, and RED.
12. Consider the technique for determining articulation points using depth-first search. If a vertex has no predecessors (first vertex discovered for a "restart"), it can be an articulation point only if
    A. there are no incident tree edges
    B. there is one incident tree edge
    C. there is more than one incident tree edge
    D. there is no back edge incident to this vertex
13. On an augmenting path, a critical edge is:
    A. an edge from the source
    B. an edge into the sink
    C. an edge that used to be saturated
    D. an edge with the minimum residual capacity
14. Which string search method computes signatures to avoid byte-to-byte comparisons?
    A. Karp-Rabin
    B. KMP with fail 1 links
    C. KMP with fail 2 links
    D. Z table
15. The most general approximation result that can be achieved for an NP-hard problem is:
    A. Approximation Algorithm
    B. Approximation Scheme

C. Fully Polynomial-time Approximation Scheme
D. Polynomial-time Approximation Scheme

16. Explain how the Z Algorithm ("Fundamental String Preprocessing") may be used to determine whether two strings are circularly equal. (5 points)

CSE 5311 Name _____
Test 2 - Open Book
Summer 2005 Last 4 Digits of Student ID # _____

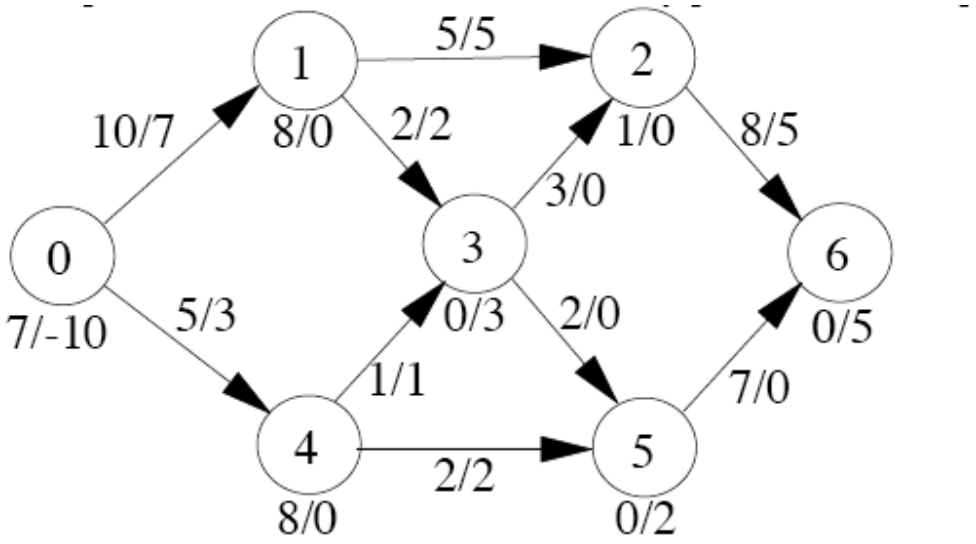
1. Solve the following instance of Longest Common Subsequence using the method based on the Longest Strictly Increasing Subsequence problem. (15 points)

0	1	2	3	4	5	6	7
a	b	d	c	d	e	f	g
b	c	d	a	e	d	g	f

2. Use dynamic programming, either with a table or lists, to determine a subset that sums to 20. DO NOT SOLVE BY INSPECTION! (15 points)

2 3 5 7 11 13 17

3. List the remaining operations to complete this instance of network flows by push-relabel. In addition, give a minimum cut. (20 points)



CSE 5311 Name _____

Test 1 - Closed Book

Summer 2006 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. Sorting the set of input edges is a property of which minimum spanning tree technique?
   A. Boruvka
   B. Kruskal
   C. Path-based (Warshall)
   D. Prim
2. The balance factors in an AVL tree are computed as:
   A. $\text{height}_{\text{right}} - \text{height}_{\text{left}}$
   B. same as the null path length in a leftist heap
   C. the difference of the number of nodes in the left and right subtrees
   D. the distance from a node to the root
3. Suppose you already have 16 different coupons when there are 20 coupon types. What is the expected number of boxes for obtaining a coupon different from the 16 you already have?
   A. 3
   B. 4
   C. 5
   D. 15
4. When performing selection in worst-cast linear time for n numbers, roughly how many column medians are computed in the first round?
   A. $\frac{n}{5}$
   B. m, the median-of-medians
   C. $.7n$
   D. $W\left(\frac{n}{5}\right)$
5. To support computing the number of keys that are smaller than a query key, an augmented binary search tree stores the following at every node:
   A. the count of the number of keys in the entire tree
   B. the sum of all keys in the left subtree
   C. the count of the number of keys stored in the subtree rooted by this node
   D. the sum of all keys stored in the subtree rooted by this node
6. When using Brent's rehash during insertion, the number of previously inserted keys that may move is:
   A. 1
   B. 2
   C. $\frac{1}{\alpha}$
   D. H_m, where m is the number of stored keys
7. Which of the following statements is true?
   A. A binary search tree may be assigned legal AVL balance factors if and only if it may be legally colored as a red-black tree.
   B. If a binary search tree may be legally colored as a red-black tree, then it may be assigned legal AVL balance factors.
   C. If a binary search tree may be assigned legal AVL balance factors, then it may be legally colored as a red-black tree.
   D. No binary search tree may be assigned both legal AVL balance factors and be legally colored as a red-black tree.

8. Which data structure is not used to implement a dictionary?
   A. AVL tree
   B. Red-black tree
   C. Self-organizing list
   D. Union-find
9. The reason for marking nodes in a Fibonacci heap is:
   A. to allow computing the value of the potential function.
   B. to assure that the structure is a Fibonacci heap rather than a binomial heap.
   C. to improve the performance of CONSOLIDATE.
   D. to indicate nodes that have lost a child since becoming a child themselves.
10. Which property does not hold for binomial heaps?
   A. DECREASE-KEY takes O(1) time.
   B. MINIMUM takes O(log n) time.
   C. Performing n INSERT operations into an empty heap will take O(n) time.
   D. The number of trees is based on the binary representation of the number of stored items.
11. When are Fibonacci trees used?
   A. Constructing a priority queue with excellent amortized complexity for DECREASE-KEY.
   B. Defining the potential function for Fibonacci heaps.
   C. Demonstrating worst-case behaviors for AVL trees.
   D. Demonstrating worst-case behaviors for red-black trees.
12. The perfect hashing method discussed in class depends on which fact?
   A. $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \qquad 0 < x < 1$
   B. $\ln n < H_n < \ln n + 1$
   C. The expected number of probes for a successful search in Brent's method is less than 2.5.
   D. The probability of collisions among n keys stored in a hash table of size n^2 is less than 0.5.
13. Dynamic optimality is a concept involving the comparison of
   A. a key-comparison based data structure to hashing.
   B. amortized complexity to actual complexity.
   C. an online data structure to a fixed, unchanging data structure.
   D. an online data structure to an offline data structure.
14. Which of the following minimum spanning tree algorithms is theoretically slowest? Assume that the most efficient data structures are used.
   A. Boruvka
   B. Kruskal
   C. Path-based (Warshall)
   D. Prim
15. The two parts of using a Markov model to analyze a self-adjusting data structure are:
   A. define the potential function, compare to the optimal offline (OPT) method.
   B. determine the probability for each state, then compute the expected number of probes.
   C. simulate the data structure long enough to ensure convergence, the compute the expected number of probes.
   D. use Knuth's root trick to find the optimal subtrees in $O\left(n^2\right)$ time, backtrace to print the tree.
16. Give the red-black (without sentinels) and AVL trees with four levels (maximum distance of a leaf from the root is three edges) with the fewest nodes. (5 points)

CSE 5311 Name _____

Test 1 - Open Book

Summer 2006 Last 4 Digits of Student ID # _____

1. Suppose that you are given n values. Give a linear-time algorithm to build a binomial heap with the n values. 10 points.
2. Evaluate the following recurrences using the master method. Indicate the case that is used for each. (10 points)

a. $T(n) = 4T\left(\frac{n}{2}\right) + n^3$

b. $T(n) = 4T\left(\frac{n}{2}\right) + n^2$

c. $T(n) = 4T\left(\frac{n}{2}\right) + 1$

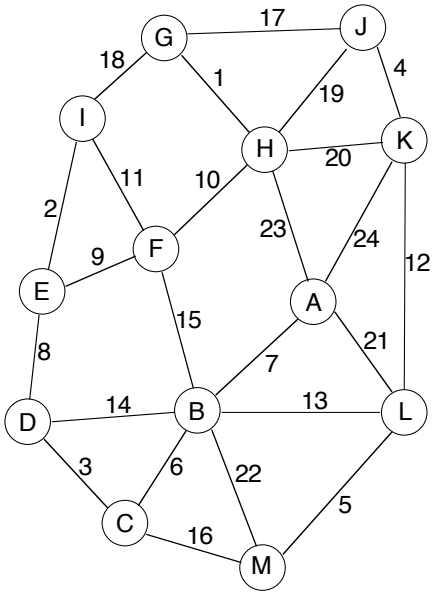
3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
n=7;
q[0]=0.06;
key[1]=1;
p[1]=0.14;
q[1]=0.06;
key[2]=2;
p[2]=0.06;
q[2]=0.06;
key[3]=3;
p[3]=0.08;
q[3]=0.06;
key[4]=4;
p[4]=0.12;
q[4]=0.05;
key[5]=5;
p[5]=0.10;
q[5]=0.05;
key[6]=6;
p[6]=0.02;
q[6]=0.05;
key[7]=7;
p[7]=0.04;
q[7]=0.05;
w[0][0]=0.060000
w[0][1]=0.260000
w[0][2]=0.380000
w[0][3]=0.520000
w[0][4]=0.690000
w[0][5]=0.840000
w[0][6]=0.910000
w[0][7]=1.000000
w[1][1]=0.060000
w[1][2]=0.180000
w[1][3]=0.320000
w[1][4]=0.490000
w[1][5]=0.640000
w[1][6]=0.710000
w[1][7]=0.800000
w[2][2]=0.060000
w[2][3]=0.200000
w[2][4]=0.370000
w[2][5]=0.520000
w[2][6]=0.590000
w[2][7]=0.680000
w[3][3]=0.060000
w[3][4]=0.230000
w[3][5]=0.380000
w[3][6]=0.450000
w[3][7]=0.540000
w[4][4]=0.050000
w[4][5]=0.200000
w[4][6]=0.270000
w[4][7]=0.360000
w[5][5]=0.050000
w[5][6]=0.120000
w[5][7]=0.210000
w[6][6]=0.050000
w[6][7]=0.140000
w[7][7]=0.050000
Average probe length is ???
trees in parenthesized prefix
c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
c(3,3) cost 0.000000
c(4,4) cost 0.000000
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(7,7) cost 0.000000
c(0,1) cost 0.260000 1
c(1,2) cost 0.180000 2
c(2,3) cost 0.200000 3
c(3,4) cost 0.230000 4
c(4,5) cost 0.200000 5
c(5,6) cost 0.120000 6
c(6,7) cost 0.140000 7
c(0,2) cost 0.560000 1(,2)
c(1,3) cost 0.500000 3(2,)
c(2,4) cost 0.570000 4(3,)
c(3,5) cost 0.580000 4(,5)
c(4,6) cost 0.390000 5(,6)
c(5,7) cost 0.330000 7(6,)
c(0,3) cost 0.980000 2(1,3)
c(1,4) cost 0.900000 3(2,4)
c(2,5) cost 0.920000 4(3,5)
c(3,6) cost 0.800000 5(4,6)
c(4,7) cost 0.690000 5(,7(6,))
c(0,4) cost 1.480000 3(1(,2),4)
c(1,5) cost 1.340000 4(3(2,),5)
c(2,6) cost 1.180000 4(3,5(,6))
c(3,7) cost 1.100000 5(4,7(6,))
c(0,5) cost 1.980000 3(1(,2),4(,5))
c(1,6) cost 1.600000 4(3(2,),5(,6))
c(2,7) cost 1.570000 4(3,5(,7(6,)))
c(0,6) cost 2.270000 3(1(,2),5(4,6))
c(1,7) cost 1.990000 4(3(2,),5(,7(6,)))
c(0,7) cost ??? ????????????????
```

4. The hash table below was created using double hashing with Brent's rehash. The initial slot ($h_1(key)$) and rehashing increment ($h_2(key)$) are given for each *key*. Show the result from inserting 9000 using Brent's rehash when $h_1(9000) = 5$ and $h_2(9000) = 4$. (10 points)

	key	$h_1(key)$	$h_2(key)$
0			
1			
2	5000	2	1
3	4000	3	4
4	3000	4	1
5	2000	5	5
6	1000	6	2

5. Demonstrate Boruvka's algorithm on the following graph. Be sure to indicate the edges that are inserted into the MST in each phase. (10 points)



CSE 5311 Name _____

Test 2 - Closed Book

Summer 2006 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. When performing bin packing using the First-Fit Decreasing technique, the total number of items placed in the bins past the optimal bins (1 .. OPT) is no more than:
   A. 1 + ε
   B. 2
   C. OPT - 1
   D. OPT
2. How many times will `-1` occur in the style 1 fail link table for the pattern `abacabac`?
   A. 1
   B. 2
   C. 3
   D. 4
3. How many times will `-1` occur in the style 2 fail link table for the pattern `abacabac`?
   A. 1
   B. 2
   C. 3

D. 4

4. Constructing a suffix array for a sequence with n symbols by using an optimal key-comparison sort has this worst-case time:
   A. $\Theta(n)$
   B. $\Theta(n \log n)$
   C. $\Theta(n^2)$
   D. $\Theta(n^2 \log n)$
5. The algorithm for finding a maximum capacity path for network flows is most similar to which algorithm?
   A. Breadth-first search
   B. Decomposition of a flow into E augmenting paths
   C. Dijkstra
   D. Floyd-Warshall
6. Which of the following algorithms may take time longer than $O(n \log n)$?
   A. Closest points in 2-d space - divide and conquer
   B. Closest points in 2-d space - sweepline
   C. Graham scan
   D. Jarvis march
7. When combining the left and right parts for the 2-d closest pair algorithm along the vertical dividing line, each left-side point "near" the line has its distance computed to no more than this number of right-side points.
   A. 6
   B. 8
   C. lg n
   D. $n/2$
8. Which of the following is not required before relabeling ("lifting") a vertex to a new height?
   A. Any eligible edges for the present height have been saturated.
   B. Both breadth-first searches have been done.
   C. The vertex is not the source or sink.
   D. The vertex is overflowing.
9. Which of the following does not have a polynomial-time approximation algorithm?
   A. Bin packing
   B. Edge coloring
   C. Euclidean traveling salesperson
   D. Vertex coloring
10. The length of a longest increasing subsequence for `1 4 3 2 3 4 3 6 3 4` is:
    A. 4
    B. 5
    C. 6
    D. 7
11. Which of the following is NOT required when showing that problem B is NP-complete by a reduction from problem A?
    A. Problem A is NP-complete.
    B. The reduction has an inverse that takes each instance of problem B to an instance of problem A.
    C. The reduction must be consistent for the decision results for each instance of problem A and and the corresponding instance of problem B.
    D. The reduction takes polynomial time.

12. Consider the technique for determining articulation points using depth-first search. If a vertex has no predecessors (first vertex discovered for a "restart"), it can be an articulation point only if
    A. there are no incident tree edges
    B. there is one incident tree edge
    C. there is more than one incident tree edge
    D. there is no back edge incident to this vertex
13. Which of the following problems uses some edges with non-unit capacity when translated to a network flow problem?
    A. Bipartite matching
    B. Edge connectivity
    C. Minimum vertex cover for bipartite graph
    D. Vertex connectivity
14. Which is not true regarding the Karp-Rabin method?
    A. It always runs in $\Theta(m+n)$ time.
    B. It computes a signature for the pattern and each m contiguous text symbols.
    C. The signature is computed using a polynomial.
    D. When the signature for m contiguous text symbols matches the signature for the pattern, a `strcmp` must be performed.
15. The least general approximation result that can be achieved for an NP-hard problem is:
    A. Approximation Algorithm
    B. Approximation Scheme
    C. Fully Polynomial-time Approximation Scheme
    D. Polynomial-time Approximation Scheme
16. Explain how the Z Algorithm ("Fundamental String Preprocessing") may be used to perform a string search for all occurences (possibly with overlap among occurences) of a given pattern within a given text. Do <u>not</u> explain details of the Z algorithm. (5 points)

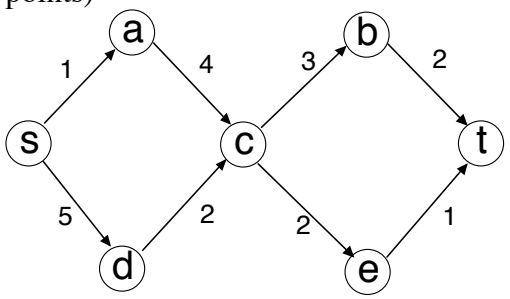
CSE 5311 Name _____

Test 2 - Open Book

Summer 2006 Last 4 Digits of Student ID # _____

1. Explain how to use the Z algorithm to determine a sequence X for the largest possible value k such that for a given sequence W = X^k. Demonstrate your technique for ababab = $(ab)^3$ and abcd = $(abcd)^1$. 15 points
2. Use dynamic programming, either with a table or lists, to determine a subset that sums to 39. DO NOT SOLVE BY INSPECTION! (15 points)

   `2  4  5  11  12  13  17`
3. List the lift and push operations to solve for the maximum flow. In addition, give a minimum cut. (20 points)



CSE 5311 Name _____

Test 1 - Closed Book

Summer 2007 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. During which operation on a leftist heap may subtree swaps be needed?
   A. DECREASE-KEY B. EXTRACT-MIN
   C. UNION D. All of A., B., and C.
2. What is the worst-case number of rotations when performing deletion on an AVL tree?
   A. $\Theta(1)$ B. $\Theta(\log n)$ C. $\Theta(n)$ D. No rotations are ever used
3. Suppose you already have 15 different coupons when there are 20 coupon types. What is the expected number of boxes for obtaining a coupon different from the 15 you already have?
   A. 3 B. 4 C. 5 D. 15
4. When is path compression used?
   A. After an insertion into any type of balanced binary search tree.
   B. After an insertion into a splay tree.
   C. With a FIND operation.
   D. With a UNION operation.
5. To support testing whether an interval intersects some interval in a set of intervals, an augmented binary search tree stores the following at every node:
   A. the count of the number of intervals in the entire tree
   B. the count of the number of intervals stored in the subtree rooted by this node
   C. the maximum value appearing in any interval in the subtree rooted by this node
   D. the sum of the lengths of all intervals in the left subtree
6. Which Fibonacci heap operation has O(log *n*) actual cost?
   A. FIB-HEAP-DECREASE-KEY B. FIB-HEAP-DELETE
   C. FIB-HEAP-EXTRACT-MIN D. FIB-HEAP-UNION
7. If a Fibonacci tree appears as a subtree of an AVL tree, which nodes would be assigned a balance factor of 0?
   A. none of them B. only the leaves
   C. only the root D. the leaves and the root
8. How many inversions are there for the lists 1, 2, 5, 4, 3 and 2, 5, 4, 1, 3?
   A. 2 B. 3 C. 4 D. 5
9. The reason for marking nodes in a Fibonacci heap is:
   A. to allow computing the value of the potential function.
   B. to assure that the structure is a Fibonacci heap rather than a binomial heap.
   C. to improve the performance of CONSOLIDATE.
   D. to indicate nodes that have lost a child since becoming a child themselves.
10. Which property does not hold for binomial heaps?
    A. DECREASE-KEY takes O(log *n*) time.
    B. MINIMUM takes O(1) time.
    C. Performing *n* INSERT operations into an empty heap will take O(*n*) time.
    D. The number of trees is based on the binary representation of the number of stored items.
11. In the worst case, the number of rotations for inserting a key in a treap with *n* keys is:
    A. θ(1) B. θ(log *n*) C. θ(*n*) D. θ(*n* log *n*)
12. Which priority queue is defined using the notion of null path length?
    A. Binary heap B. Binomial heap C. Fibonacci heap D. Leftist heap
13. Which of the following is not a property of splay trees?
    A. Splaying an accessed node to the root can cause the potential to decrease
    B. The amortized cost of a splaying sequence is bounded logarithmically

C. The zig rotation is only applied at the root
D. They are a form of balanced tree

14. The main difference between MTF and OPT for self-organizing linear lists is:
    A. OPT can do transpositions
    B. OPT is given the entire request sequence in advance, while MTF receives the requests one-at-a-time
    C. MTF is given the entire request sequence in advance, while OPT receives the requests one-at-a-time
    D. MTF counts inversions
15. Assuming a random n-permutation is provided, the expected number of hires for the hiring problem is:

    A. 2 B. H_n C. $\frac{1}{n}$ D. n

16. Give a tree of rank 5 that could appear in a Fibonacci heap and has the minimum possible number of nodes. Be sure to include the marks. (5 points)

CSE 5311 Name _____
Test 1 - Open Book
Summer 2007 Last 4 Digits of Student ID # _____

1. Describe the *format* of the system of equations that would be used to perform a Markov analysis on the *transpose* approach to self-adjusting lists for $n = 4$. (10 points)
2. Evaluate the following recurrences using the master method. Indicate the case that is used for each. (10 points)

a. $T(n) = T(0.7n) + n$

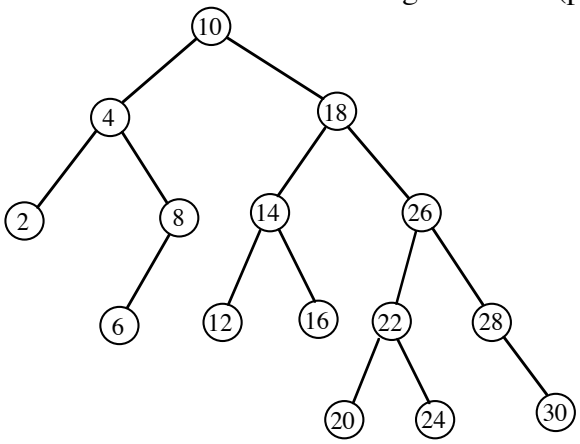
b. $T(n) = T(0.7n) + 1$

c. $T(n) = 16T\left(\frac{n}{2}\right) + n^3$

3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
n=7;
q[0]=0.06;
p[1]=0.15;
q[1]=0.03;
p[2]=0.15;
q[2]=0.05;
p[3]=0.1;
q[3]=0.01;
p[4]=0.05;
q[4]=0.05;
p[5]=0.02;
q[5]=0.05;
p[6]=0.1;
q[6]=0.05;
p[7]=0.08;
q[7]=0.05;
w[0][0]=0.060000
w[0][1]=0.240000
w[0][2]=0.440000
w[0][3]=0.550000
w[0][4]=0.650000
w[0][5]=0.720000
w[0][6]=0.870000
w[0][7]=1.000000
w[1][1]=0.030000
w[1][2]=0.230000
w[1][3]=0.340000
w[1][4]=0.440000
w[1][5]=0.510000
w[1][6]=0.660000
w[1][7]=0.790000
w[2][2]=0.050000
w[2][3]=0.160000
w[2][4]=0.260000
w[2][5]=0.330000
w[2][6]=0.480000
w[2][7]=0.610000
w[3][3]=0.010000
w[3][4]=0.110000
w[3][5]=0.180000
w[3][6]=0.330000
w[3][7]=0.460000
w[4][4]=0.050000
w[4][5]=0.120000
w[4][6]=0.270000
w[4][7]=0.400000
w[5][5]=0.050000
w[5][6]=0.200000
w[5][7]=0.330000
w[6][6]=0.050000
w[6][7]=0.180000
w[7][7]=0.050000
Average probe length is ???
trees in parenthesized prefix
c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
c(3,3) cost 0.000000
c(4,4) cost 0.000000
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(7,7) cost 0.000000
c(0,1) cost 0.240000 1
c(1,2) cost 0.230000 2
c(2,3) cost 0.160000 3
c(3,4) cost 0.110000 4
c(4,5) cost 0.120000 5
c(5,6) cost 0.200000 6
c(6,7) cost 0.180000 7
c(0,2) cost 0.670000 1(,2)
c(1,3) cost 0.500000 2(,3)
c(2,4) cost 0.370000 3(,4)
c(3,5) cost 0.290000 5(4,)
c(4,6) cost 0.390000 6(5,)
c(5,7) cost 0.510000 6(,7)
c(0,3) cost 0.950000 2(1,3)
c(1,4) cost 0.780000 3(2,4)
c(2,5) cost 0.610000 4(3,5)
c(3,6) cost 0.620000 6(5(4,),)
c(4,7) cost 0.700000 6(5,7)
c(0,4) cost 1.260000 2(1,3(,4))
c(1,5) cost 1.030000 3(2,5(4,))
c(2,6) cost 1.030000 4(3,6(5,))
c(3,7) cost 0.930000 6(5(4,),7)
c(0,5) cost 1.570000 2(1,4(3,5))
c(1,6) cost 1.510000 3(2,6(5(4,),))
c(2,7) cost 1.400000 6(4(3,5),7)
c(0,6) cost 2.140000 2(1,4(3,6(5,)))
c(1,7) cost 1.950000 3(2,6(5(4,),7))
c(0,7) cost ??? ????????????????
```

4. Give an example of a splay tree and an access (i.e. a search) such that the potential is increased by splaying the search key. (10 points)
5. Delete 2 from the following AVL tree (preserving AVL properties). (10 points)



Test 2 - Closed Book
Summer 2007 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. When performing selection in worst-cast linear time for n numbers, roughly how many column medians are computed in the first round?
   A. $\frac{n}{5}$
   B. m, the median-of-medians
   C. $.7n$
   D. $W\left(\frac{n}{5}\right)$
2. What is the nature of the signature function for the Karp-Rabin method?
   A. a polynomial of arbitrary precision implemented using a bignum package
   B. it is similar to the KMP failure links
   C. similar to a double hash function for string keys
   D. the remainder by discarding the overflow for a polynomial
3. Which of the following is helpful if you wish to know the farthest pair in a set of points in 2D?
   A. Convex hull
   B. Delaunay triangulation
   C. Euclidean minimum spanning tree
   D. Voronoi diagram
4. While constructing a suffix array for a sequence with n symbols by using radix sorts, sequence symbols are used in which of the radix sorts?
   A. All of them
   B. None of them
   C. The first one
   D. The last one
5. In a maximum flow problem, the number of augmenting paths in a flow decomposition is bounded by:
   A. f
   B. E
   C. V
   D. $O(VE)$
6. Which algorithm is defined using the notions of left-turn and right-turn?
   A. Closest points in 2-d space
   B. Graham scan
   C. Jarvis march
   D. Suffix array construction
7. What data structure is used for the sweep-line status when computing the 2-d closest pair?
   A. BST of points with x-coordinates as the key
   B. BST of points with y-coordinates as the key
   C. Interval tree
   D. Sorted array by ascending x-coordinates
8. The four russians' concept is to:
   A. Implement longest common subsequences using linear space
   B. Pack bits into an efficient storage unit
   C. Trade-off between enumerating situations and referencing these situations
   D. Trade-off between scalar additions and multiplications
9. When coloring the edges of a graph, a dc-path gets inverted because:
   A. All edges in the fan will be colored with d or c.

B. d is a free color for all fan vertices.
C. d is the free color for two fan vertices.
D. We are trying to minimize the number of colors used by the path.

10. Sorting the edges is a property of which minimum spanning tree technique?
A. Boruvka
B. Kruskal
C. Prim
D. Path-based (Warshall)

11. The reduction from 3-sat to 3-colorability will give a graph requiring four colors when:
A. Removing one clause from the 3-sat instance will leave a satisifiable set of clauses.
B. The 3-sat instance is satisfiable.
C. The 3-sat instance is a tautology.
D. The 3-sat instance is unsatisfiable.

12. Which of the following problems is not NP-complete? (Assume P ≠ NP)
A. 3-satisfiability
B. Testing if a graph is 3-colorable
C. Testing if a table is in sorted order
D. Testing if the number of colors needed to edge color a graph is the degree of the graph

13. Which of the following is a deficiency of the maximum capacity path technique?
A. An augmenting path is blocked if it introduces a cycle of flow.
B. Augmenting paths will be discovered in descending incremental flow increase order.
C. Flow decomposition must be applied.
D. The maximum number of potential augmenting paths depends on the achievable flow, in addition to the number of vertices and edges.

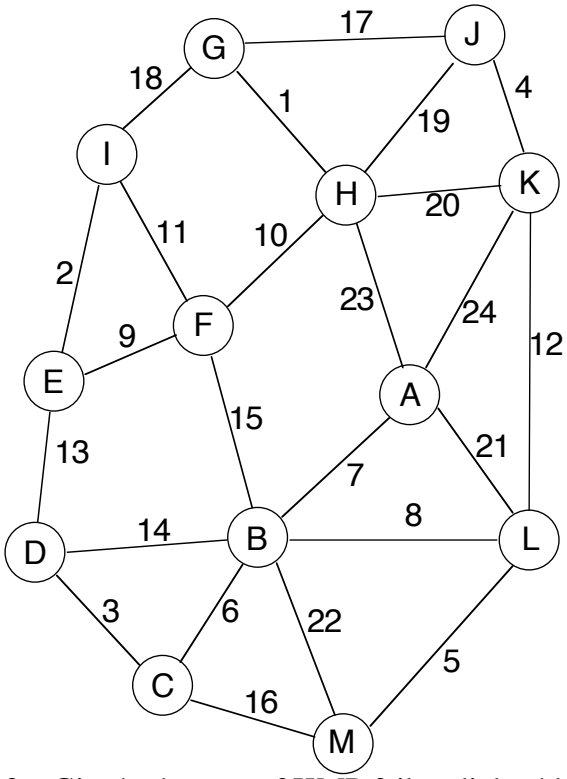
14. Which is not true regarding the first-fit decreasing method for bin packing?
A. Each object is placed by going left-to-right until it fits in a bin or a new bin is allocated.
B. Objects placed in bins beyond the optimal number have sizes no larger than ½.
C. The number of objects placed in bins beyond the optimal number is arbitrary.
D. The number of objects with sizes larger than ½ is a lower bound on the optimal number of bins.

15. The most general approximation result that can be achieved for an NP-hard problem is:
A. Approximation Algorithm
B. Approximation Scheme
C. Fully Polynomial-time Approximation Scheme
D. Polynomial-time Approximation Scheme

16. Suppose you have a set of points in Euclidean 2-d space. Give an algorithm for finding a ρ-approximation for the minimum traveling salesperson path. Be sure to give the value of ρ. (5 points)

CSE 5311 Name _____
Test 2 - Open Book
Summer 2007 Last 4 Digits of Student ID # _____

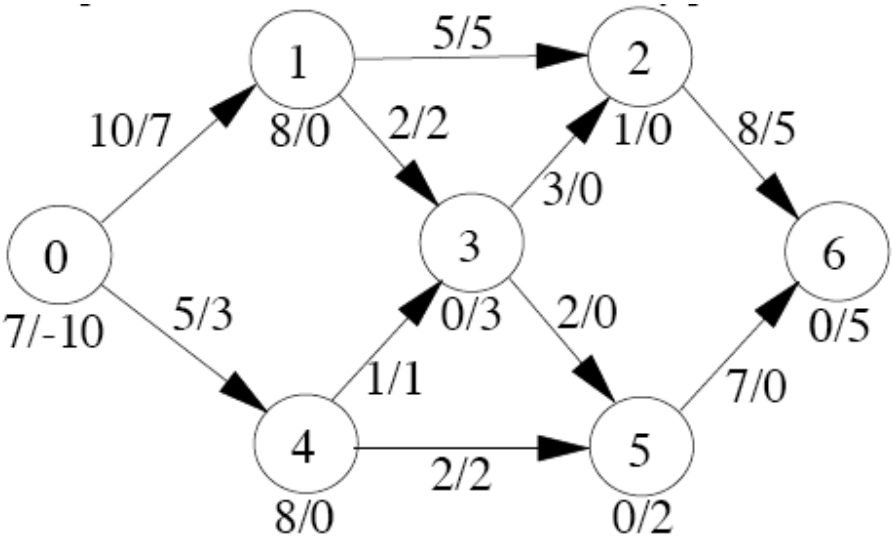
1. Demonstrate Boruvka's algorithm on the following graph. Be sure to indicate the edges that are inserted into the MST in each phase. (15 points)



2. Give both types of KMP failure link tables and the table produced by the Z algorithm for the following string: (15 points)

```
 0  a
 1  c
 2  a
 3  b
 4  a
 5  c
 6  a
 7  c
 8  a
 9  b
10  a
11  c
12  a
13  b
14  a
15  c
16  a
17  c
18  a
19  b
20  a
21  c
22  a
23  c
24  a
```

3. List the remaining operations to complete this instance of network flows by push-relabel. In addition, give a minimum cut. (20 points)


1
2
0
3
6
4
5
5/5
10/7
8/0
2/2
1/0
8/5
3/0
7/-10
5/3
0/3
2/0
0/5
1/1
7/0
2/2
8/0
0/2

CSE 5311 Name _____

Test 1 - Closed Book

Summer 2008 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. Which priority queue is defined using the notion of null path length?
   A. Binary heap B. Binomial heap C. Fibonacci heap D. Leftist heap
2. Dynamic optimality is a concept involving the comparison of
   A. a key-comparison based data structure to hashing.
   B. amortized complexity to actual complexity.
   C. an online data structure to a fixed, unchanging data structure.
   D. an online data structure to an offline data structure.
3. Suppose you already have 16 different coupons when there are 20 coupon types. What is the expected number of boxes for obtaining a coupon different from the 16 you already have?
   A. 3 B. 4 C. 5 D. 15
4. What is required when calling `union(i,j)` for maintaining disjoint subsets?
   A. `i` and `j` are in the same subset
   B. `i` and `j` are leaders for different subsets
   C. `i` and `j` are leaders for the same subset
   D. `i` is the ancestor of `j` in one of the trees
5. To support computing prefix sums of all keys that are no larger than some query key, an augmented binary search tree stores the following at every node:
   A. the sum of all keys in the entire tree
   B. the sum of all keys in the left subtree
   C. the sum of all keys that are no larger than the stored key
   D. the sum of all keys stored in the subtree rooted by this node
6. Which Fibonacci heap operation has $\Theta(\log n)$ worst-case actual cost?
   A. FIB-HEAP-DECREASE-KEY B. FIB-HEAP-DELETE
   C. FIB-HEAP-EXTRACT-MIN D. FIB-HEAP-UNION
7. When are Fibonacci trees used?
   A. Constructing a priority queue with excellent amortized complexity for DECREASE-KEY.
   B. Defining the potential function for Fibonacci heaps.
   C. Demonstrating worst-case behaviors for AVL trees.
   D. Demonstrating worst-case behaviors for red-black trees.
8. To reduce the probability of having any collisions to < 0.5 when hashing n keys, the table should have at least this number of elements.
   A. n B. $n \ln n$ C. n^2 D. n^3
9. The minimum number of nodes in a tree in a Fibonacci heap where the root has 7 children is:
   A. 34 B. 61 C. 121 D. 128
10. Which property does not hold for binomial heaps?
    A. DECREASE-KEY takes $O(1)$ time.
    B. MINIMUM takes $O(\log n)$ time.
    C. Performing n INSERT operations into an empty heap will take $O(n)$ time.
    D. The number of trees is based on the binary representation of the number of stored items.
11. Which of the following is not true regarding the amortized analysis of binary tree traversals?
    A. INIT had an amortized cost of 0.
    B. SUCC had an actual cost determined by the number of edges followed.
    C. SUCC had an amortized cost of 2.
    D. The potential was defined without regard for the type of traversal being performed.

12. Which data structure is not used to implement a dictionary?
    A. AVL tree B. Red-black tree C. Self-organizing list D. Union-find
13. Which of the following is not a property of splay trees?
    A. Splaying an accessed node to the root can cause the potential to decrease
    B. The amortized cost of a splaying sequence is bounded logarithmically
    C. The zig rotation is only applied at the root
    D. They are a form of balanced tree
14. The two parts of using a Markov model to analyze a self-adjusting data structure are:
    A. define the potential function, compare to the optimal offline (OPT) method.
    B. determine the probability for each state, then compute the expected number of probes.
    C. simulate the data structure long enough to ensure convergence, the compute the expected number of probes.
    D. use Knuth's root trick to find the optimal subtrees in $O\left(n^2\right)$ time, backtrace to print the tree.
15. Assuming a random *n*-permutation is provided, the expected number of hires for the hiring problem is:
    A. 2 B. H_n C. $\sqrt{n}$ D. $\ln\ln n$
16. Suppose there are 20 coupon types for the coupon collecting problem. Each of the coupon types is identified by an integer in the range 1 to 20. If you are given two random cereal boxes, what is the probability (ignoring the order in which the coupons were obtained) that the two coupons have consecutive numbers, i.e. some *i* and *i* + 1? (5 points)

CSE 5311 Name _____
Test 1 - Open Book
Summer 2008 Last 4 Digits of Student ID # _____

1. Suppose you are given *n* values. Give a linear-time algorithm to build a binomial heap with the *n* values. 10 points.
2. Evaluate the following recurrences using the master method. Indicate the case that is used for each. (10 points)

a. $T(n) = 4T\left(\frac{n}{2}\right) + n^3$

b. $T(n) = 4T\left(\frac{n}{2}\right) + n^2$

c. $T(n) = 4T\left(\frac{n}{2}\right) + 1$

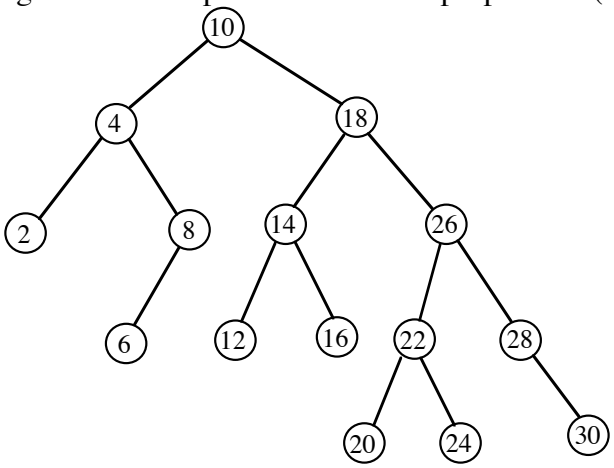
3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)
4. The hash table below was created using double hashing with Brent's rehash. The initial slot ($h_1(key)$) and rehashing increment ($h_2(key)$) are given for each *key*. Show the result from inserting 1300 using Brent's rehash when $h_1(1300) = 5$ and $h_2(1300) = 3$. (10 points)

	key	$h_1(key)$	$h_2(key)$
0			
1	1000	6	2
2			
3	1200	3	3

4	500	3	1
5	12	5	1
6	27	6	3

	key
0	
1	
2	
3	
4	
5	
6	

5. Insert 23 into the following AVL tree to preserve the AVL properties. (10 points)



CSE 5311 Name _____
Test 2 - Closed Book
Summer 2008 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. Which of the following problems uses some edges with non-unit capacity when translated to a network flow problem?
   A. Bipartite matching
   B. Edge connectivity

C. Minimum vertex cover for bipartite graph
D. Vertex connectivity

2. Which is not true regarding the Karp-Rabin method?
   A. It always runs in $\Theta(m+n)$ time.
   B. It computes a signature for the pattern and each m contiguous text symbols.
   C. The signature is computed using a polynomial.
   D. When the signature for m contiguous text symbols matches the signature for the pattern, a `strcmp` must be performed.
3. Which of the following is helpful if you wish to know the farthest pair in a set of points in 2D?
   A. Convex hull
   B. Delaunay triangulation
   C. Euclidean minimum spanning tree
   D. Voronoi diagram
4. Constructing a suffix array for a sequence with n symbols by using an optimal key-comparison sort has this worst-case time:
   A. $\Theta(n)$ B. $\Theta(n \log n)$ C. $\Theta(n^2)$ D. $\Theta(n^2 \log n)$
5. The algorithm for finding a maximum capacity path for network flows is most similar to which algorithm?
   A. Breadth-first search
   B. Decomposition of a flow into E augmenting paths
   C. Dijkstra
   D. Floyd-Warshall
6. Which of the following algorithms may take time longer than $O(n \log n)$?
   A. Closest points in 2-d space - divide and conquer
   B. Closest points in 2-d space - sweepline
   C. Graham scan
   D. Jarvis march
7. What data structure is used for the sweep-line status when computing the 2-d closest pair?
   A. BST of points with x-coordinates as the key
   B. BST of points with y-coordinates as the key
   C. Interval tree
   D. Sorted array by ascending x-coordinates
8. The length of a longest (monotone) increasing subsequence for `1 4 3 2 3 4 3 6 3 4` is:
   A. 4 B. 5 C. 6 D. 7
9. Which of the following is NOT required when showing that problem B is NP-complete by a reduction from problem A?
   A. Problem A is NP-complete.
   B. The reduction has an inverse that takes each instance of problem B to an instance of problem A.
   C. The reduction must be consistent for the decision results for each instance of problem A and and the corresponding instance of problem B.
   D. The reduction takes polynomial time.
10. Which minimum spanning tree algorithm is the slowest?
    A. Boruvka B. Kruskal C. Prim D. Warshall
11. When <u>combining</u> the left and right parts for the 2-d closest pair algorithm along the vertical dividing line, each left-side point "near" the line has its distance computed to no more than this number of right-side points.
    A. 6 B. 8 C. $\lg n$ D. $n/2$
12. Which of the following problems is not NP-complete? (Assume P ≠ NP)

A. Testing if a graph is bipartite
B. Testing if an undirected graph has a Hamiltonian cycle
C. Testing if the number of colors needed to edge color a graph is the degree of the graph
D. 3-satisfiability

13. Which of the following is a deficiency of the maximum capacity path technique?
A. An augmenting path is blocked if it introduces a cycle of flow.
B. Augmenting paths will be discovered in descending incremental flow increase order.
C. Flow decomposition must be applied.
D. The maximum number of potential augmenting paths depends on the achievable flow, in addition to the number of vertices and edges.

14. When performing bin packing using the First-Fit Decreasing technique, the total number of items placed in the bins past the optimal bins (1 .. OPT) is no more than:

A. $1 + \varepsilon$ B. 2 C. OPT - 1 D. OPT

15. Which of the following does not have a polynomial-time approximation algorithm?
A. Bin packing
B. Edge coloring
C. Traveling salesperson with triangle inequality
D. Vertex coloring

16. Give the result of the Z algorithm for the following sequence. (5 points)

```
0 a
1 c
2 a
3 b
4 a
5 c
6 a
7 c
8 a
9 b
10 a
11 c
12 a
13 b
14 a
15 c
16 a
17 c
18 a
19 b
20 a
```

21 c

22 a

23 c

24 a

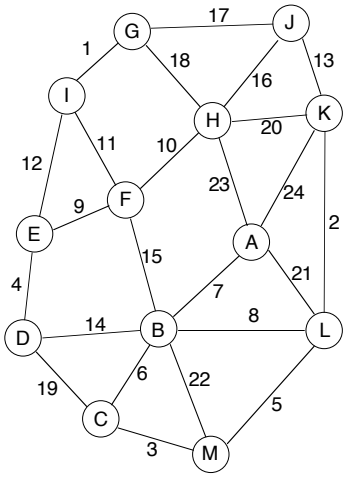
25 b

CSE 5311 Name _____

Test 2 - Open Book

Summer 2008 Last 4 Digits of Student ID # _____

1. Demonstrate Boruvka's algorithm on the following graph. Be sure to indicate the edges that are inserted into the MST in each phase. (10 points)

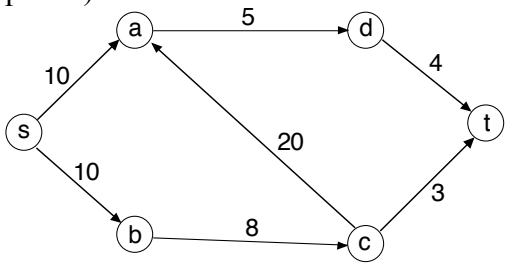


2. Solve the following instance of Longest Common Subsequence using the method based on the Longest Strictly Increasing Subsequence problem. (10 points)

0	1	2	3	4	5	6	7
a	b	d	c	d	e	f	g

b	c	d	f	e	d	g	f

3. List the lift and push operations to solve for the maximum flow. In addition, give a minimum cut. (15 points)



4. Prove that the **SET PACKING** problem is NP-complete (15 points).

Hint 1: There is a straightforward reduction from ***K*-CLIQUE**.

Hint 2: It is often helpful to give an example of the reduction used.

INSTANCE: Collection C of finite sets, positive integer $K \leq |C|$.

QUESTION: Does C contain at least K mutually (i.e. pairwise) disjoint sets?

CSE 5311 Name _____
Test 1 - Closed Book
Summer 2009 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. Which of the following statements is true?

   A. A binary search tree may be assigned legal AVL balance factors if and only if it may be legally colored as a red-black tree.
   B. If a binary search tree may be assigned legal AVL balance factors, then it may be legally colored as a red-black tree.
   C. If a binary search tree may be legally colored as a red-black tree, then it may be assigned legal AVL balance factors.
   D. No binary search tree may be assigned both legal AVL balance factors and be legally colored as a red-black tree.

2. During which operation on a leftist heap may subtree swaps be needed?

   A. DECREASE-KEY B. EXTRACT-MIN
   C. UNION D. All of A., B., and C.

3. What is the worst-case number of rotations when performing deletion on an AVL tree?

   A. $\Theta(1)$ B. $\Theta(\log n)$ C. $\Theta(n)$ D. No rotations are ever used

4. Suppose you already have 15 different coupons when there are 20 coupon types. What is the expected number of boxes for obtaining a coupon different from the 15 you already have?

   A. 3 B. 4 C. 5 D. 15

5. Path compression is used in which algorithm?

   A. DECREASE-KEY (e.g. CASCADING-CUT) for Fibonacci heaps
   B. FIND for disjoint sets
   C. UNION for disjoint sets
   D. UNION for Fibonacci heaps

6. The balance factors in an AVL tree are computed as:

   A. $\text{height}_{\text{right}}$ - $\text{height}_{\text{left}}$
   B. same as the null path length in a leftist heap
   C. the difference of the number of nodes in the left and right subtrees
   D. the distance from a node to the root

7. Which situation is true regarding a cascade cut that produces c trees for a Fibonacci heap?

   A. Both the actual and amortized costs are O(1).
   B. The actual cost is O(c). The amortized cost is O(1).

C. The actual cost is O(1). The amortized cost is O(c).
D. The potential can become negative.

8. Which change to a binary search tree preserves the inorder traversal property?

A. rotation
B. zig-zag
C. zig-zig
D. all of the above

9. Which property does not hold for binomial heaps?

A. DECREASE-KEY takes $O(1)$ time.
B. MINIMUM takes $O(\log n)$ time.
C. Performing n INSERT operations into an empty heap will take O(n) time.
D. The number of trees is based on the binary representation of the number of stored items.

10. Which problem may be addressed using the assignment problem?

A. Optimal binary search tree
B. Optimal hashing
C. Perfect hashing
D. Primary clustering

11. Which of the following data structures offers similar capabilities and performance characteristics to skip lists?

A. AVL trees
B. Splay trees
C. Treap
D. Union-find with path compression

12. Slow convergence toward the optimal fixed ordering is a property of which technique?

A. Count
B. Move-ahead-k
C. Move-to-front
D. Transpose

13. How many inversions are there for the lists $1, 2, 5, 4, 3$ and $2, 5, 4, 1, 3$?

A. 2 B. 3 C. 4 D. 5

14. The number of potential probe sequences when using double hashing with a table with m entries (m is prime) is:

A. $O(\log m)$ B. m C. $m(m-1)$ D. $m!$

15. When using Brent's rehash during insertion, the number of previously inserted keys that may move is:

A. 1
B. 2
C. $\frac{1}{\alpha}$
D. H_m, where m is the number of stored keys

16. Suppose there are 20 coupon types for the coupon collecting problem and you have already obtained 17 of the coupon types. How many boxes of cereal do you expect (mathematically) to open to get the remaining three coupon types? (5 points)

CSE 5311
Test 1 - Open Book
Summer 2009

Name _____

Last 4 Digits of Student ID # _____

1. Give the range of possible heights for an AVL tree with 200 keys. (10 points)

2. Evaluate the following recurrences using the master method. Indicate the case that is used for each. (10 points)

a. $T(n) = T(0.7n) + n$

b. $T(n) = T(0.7n) + 1$

c. $T(n) = 16T\left(\frac{n}{2}\right) + n^3$

3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
n=6;
q[0]=0.12;
key[1]=10;
p[1]=0.12;
q[1]=0.09;
key[2]=20;
p[2]=0.2;
q[2]=0.03;
key[3]=30;
p[3]=0.1;
q[3]=0.04;
key[4]=40;
p[4]=0.2;
q[4]=0.0;
key[5]=50;
p[5]=0.01;
q[5]=0.04;
```

```
key[6]=60;
p[6]=0.03;
q[6]=0.02;
w[0][0]=0.120000
w[0][1]=0.330000
w[0][2]=0.560000
w[0][3]=0.700000
w[0][4]=0.900000
w[0][5]=0.950000
w[0][6]=1.000000
w[1][1]=0.090000
w[1][2]=0.320000
w[1][3]=0.460000
w[1][4]=0.660000
w[1][5]=0.710000
w[1][6]=0.760000
w[2][2]=0.030000
w[2][3]=0.170000
w[2][4]=0.370000
w[2][5]=0.420000
w[2][6]=0.470000
w[3][3]=0.040000
w[3][4]=0.240000
w[3][5]=0.290000
w[3][6]=0.340000
w[4][4]=0.000000
w[4][5]=0.050000
w[4][6]=0.100000
w[5][5]=0.040000
w[5][6]=0.090000
w[6][6]=0.020000
Building c(0,2) using roots 1 thru 2
Building c(1,3) using roots 2 thru 3
Building c(2,4) using roots 3 thru 4
Building c(3,5) using roots 4 thru 5
Building c(4,6) using roots 5 thru 6
Building c(0,3) using roots 1 thru 2
Building c(1,4) using roots 2 thru 4
Building c(2,5) using roots 4 thru 4
Building c(3,6) using roots 4 thru 6
Building c(0,4) using roots 2 thru 2
Building c(1,5) using roots 2 thru 4
Building c(2,6) using roots 4 thru 4
Building c(0,5) using roots 2 thru 2
Building c(1,6) using roots 2 thru 4
Building c(0,6) using roots ? thru ?
Counts - root trick 31 without root trick 50
Average probe length is ???
trees in parenthesized prefix

c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
```

```
c(3,3) cost 0.000000
c(4,4) cost 0.000000
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(0,1) cost 0.330000 10
c(1,2) cost 0.320000 20
c(2,3) cost 0.170000 30
c(3,4) cost 0.240000 40
c(4,5) cost 0.050000 50
c(5,6) cost 0.090000 60
c(0,2) cost 0.880000 10(,20)
c(1,3) cost 0.630000 20(,30)
c(2,4) cost 0.540000 40(30,)
c(3,5) cost 0.340000 40(,50)
c(4,6) cost 0.150000 60(50,)
c(0,3) cost 1.200000 20(10,30)
c(1,4) cost 1.200000 20(,40(30,))
c(2,5) cost 0.640000 40(30,50)
c(3,6) cost 0.490000 40(,60(50,))
c(0,4) cost 1.770000 20(10,40(30,))
c(1,5) cost 1.350000 20(,40(30,50))
c(2,6) cost 0.790000 40(30,60(50,))
c(0,5) cost 1.920000 20(10,40(30,50))
c(1,6) cost 1.540000 40(20(,30),60(50,))
c(0,6) cost ??? ????????????????
```

4. The hash table below was created using double hashing with Brent's rehash. The initial slot ($h_1(key)$) and rehashing increment ($h_2(key)$) are given for each *key*. Show the result from inserting 9000 using Brent's rehash when $h_1(9000) = 5$ and $h_2(9000) = 4$. (10 points)

	key	$h_1(key)$	$h_2(key)$
0			
1			
2	5000	2	1
3	4000	3	4
4	3000	4	1
5	2000	5	5
6	1000	6	2

	key
0	
1	

2

3

4

5

6

5. Suppose all 2^k - 1 nodes, along with the sentinel, in a red-black tree are colored black. Explain what will happen if any key is deleted. (10 points) CSE 5311

Name _____

Test 2 - Closed Book

Summer 2009 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. When performing selection in worst-cast linear time for n numbers, roughly how many column medians are computed in the first round?

   A. $W\left(\frac{n}{5}\right)$ B. $\frac{n}{5}$ C. m, the median-of-medians D. $0.7n$

2. What data structure is used for the sweep-line status when detecting intersections among rectilinear rectangles?

   A. BST of points with x-coordinates as the key
   B. BST of points with y-coordinates as the key
   C. Interval tree
   D. Sorted array by ascending x-coordinates

3. While constructing a suffix array for a sequence with n symbols by using radix sorts, sequence symbols are used in which of the radix sorts?

   A. All of them B. None of them C. The last one D. The first one

4. Constructing a suffix array for a sequence with n symbols by using the original Manber-Myers radix sort construction has this worst-case time:

   A. $\Theta(n)$ B. $\Theta(n \log n)$ C. $\Theta\left(n^2\right)$ D. $\Theta\left(n^2 \log n\right)$

5. In a maximum flow problem, the number of augmenting paths in a flow decomposition is bounded by:

   A. V B. $O(VE)$ C. f D. E

6. Which of the following is not true regarding Strassen's algorithm?

   A. It is not possible to have an asymptotically faster algorithm.
   B. It requires more space than the everyday method.
   C. It uses $\Theta\left(n^{\lg 7}\right)$ scalar additions when multiplying two n x n matrices.
   D. It uses $\Theta\left(n^{\lg 7}\right)$ scalar multiplications when multiplying two n x n matrices.

7. What is the maximum number of times that an edge can become critical in the Edmonds-Karp method?

   A. Once B. Twice C. (V-2)/2 D. VE

8. The worst-case time to determine a longest (monotone) increasing subsequence for a sequence of n numbers is:

   A. $\Theta(m+n)$ B. $\Theta(n)$ C. $\Theta(n \log n)$ D. $\Theta\left(n^2\right)$

9. Which of the following is NOT required when showing that problem B is NP-complete by a reduction from problem A?

   A. The reduction has an inverse that takes each instance of problem B to an instance of problem A.
   B. The reduction takes polynomial time.
   C. The reduction must be consistent for the decision results for each instance of problem A and and the corresponding instance of problem B.
   D. Problem A is NP-complete.

10. On an augmenting path, a critical edge is:

    A. an edge with the minimum residual capacity
    B. an edge from the source
    C. an edge into the sink
    D. an edge that used to be saturated

11. Which algorithm is defined using the notions of left-turn and right-turn?

    A. Closest points in 2-d space B. Graham scan
    C. Jarvis march D. Suffix array construction

12. The reduction from 3-sat to 3-colorability will give a graph requiring four colors when:

    A. The 3-sat instance is satisfiable.
    B. Removing one clause from the 3-sat instance will leave a satisifiable set of clauses.
    C. The 3-sat instance is unsatisfiable.
    D. The 3-sat instance is a tautology.

13. Which of the following is not required before relabeling ("lifting") a vertex to a new height?

A. Both breadth-first searches have been done.
B. The vertex is overflowing.
C. Any eligible edges for the present height have been saturated.
D. The vertex is not the source or sink.

14. When coloring the edges of a graph, a dc-path gets inverted because:

A. d is a free color for all fan vertices.
B. We are trying to minimize the number of colors used by the path.
C. All edges in the fan will be colored with d or c.
D. d is the free color for two fan vertices.

15. The least general approximation result that can be achieved for an NP-hard problem is:

A. Approximation Scheme
B. Approximation Algorithm
C. Fully Polynomial-time Approximation Scheme
D. Polynomial-time Approximation Scheme

16. Give the style 2 KMP failure links for this sequence. (5 points)

```
 0 a
 1 c
 2 a
 3 b
 4 a
 5 c
 6 a
 7 c
 8 a
 9 b
10 a
11 c
12 a
13 b
14 a
15 c
16 a
17 c
18 a
19 b
```

```
20 a
21 c
22 a
23 c
24 a
25 b
```

CSE 5311 Name _____
Test 2 - Open Book
Summer 2009 Last 4 Digits of Student ID # _____

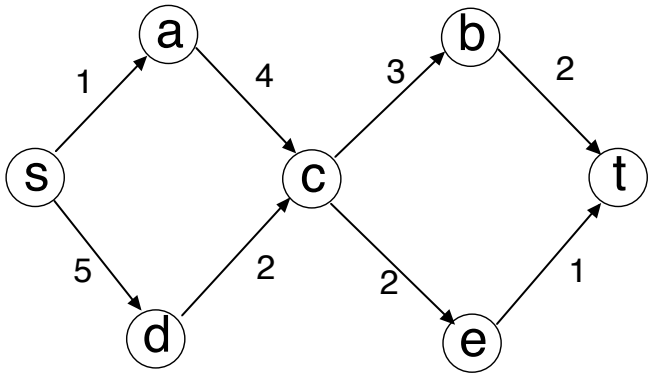
1. Explain how to use the Z algorithm to determine the maximum overlap for a suffix of string X and a prefix of string Y. For example, if X="dogcatcat" and Y="catcatdog", then the maximum overlap would be "catcat". Do not explain details of the Z algorithm. 10 points

2. Solve the following instance of Longest Common Subsequence using the method based on the Longest Strictly Increasing Subsequence problem. (10 points)

0	1	2	3	4	5	6	7
a	b	d	c	d	e	b	a

b	c	d	c	e	d	b	a

3. List the lift and push operations to solve for the maximum flow. In addition, give a minimum cut. (15 points)



4. Show that deciding whether an undirected graph is 5-colorable is NP-complete by a simple reduction from the 3-colorability problem. In addition to your proof, give an example of your reduction on a 3-colorable graph. (15 points)

CSE 5311 Name _____

Test 1 - Closed Book

Summer 2010 Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. During which operation on a leftist heap may subtree swaps be needed?
   A. DECREASE-KEY B. EXTRACT-MIN
   C. UNION D. All of A., B., and C.
2. What is the worst-case number of rotations when performing deletion on an AVL tree?
   A. $\Theta(1)$ B. $\Theta(\log n)$ C. $\Theta(n)$ D. No rotations are ever used
3. Suppose you already have 15 different coupons when there are 20 coupon types. What is the expected number of boxes for obtaining a coupon different from the 15 you already have?
   A. 3 B. 4 C. 5 D. 15
4. When is path compression used?
   A. After an insertion into a splay tree.
   B. With a FIND operation.
   C. After an insertion into any type of balanced binary search tree.
   D. With a UNION operation.
5. To support computing prefix sums of all keys that are no larger than some query key, an augmented binary search tree stores the following at every node:
   A. the sum of all keys in the entire tree
   B. the sum of all keys in the left subtree
   C. the sum of all keys that are no larger than the stored key
   D. the sum of all keys stored in the subtree rooted by this node
6. To reduce the probability of having any collisions to < 0.5 when hashing n keys, the table should have at least this number of elements.
   A. n B. $n \ln n$ C. n^2 D. n^3
7. If a Fibonacci tree appears as a subtree of an AVL tree, which nodes would be assigned a balance factor of 0?
   A. none of them B. only the leaves
   C. only the root D. the leaves and the root
8. How many inversions are there for the lists 1, 2, 5, 4, 3 and 2, 5, 4, 1, 3?
   A. 2 B. 3 C. 4 D. 5
9. The reason for marking nodes in a Fibonacci heap is:
   A. to assure that the structure is a Fibonacci heap rather than a binomial heap.
   B. to allow computing the value of the potential function.
   C. to indicate nodes that have lost a child since becoming a child themselves.
   D. to improve the performance of CONSOLIDATE.
10. Which property does not hold for binomial heaps?
    A. Performing n INSERT operations into an empty heap will take O(n) time.
    B. The number of trees is based on the binary representation of the number of stored items.
    C. DECREASE-KEY takes O(log n) time.
    D. MINIMUM takes O(1) time.
11. In the worst case, the number of rotations for inserting a key in a treap with n keys is:
    A. θ(1) B. θ(log n) C. θ(n) D. θ(n log n)
12. Which priority queue is defined using the notion of null path length?
    A. Leftist heap B. Binomial heap C. Fibonacci heap D. Binary heap
13. When performing selection in worst-cast linear time for n numbers, roughly how many column medians are computed in the first round?
    A. $\frac{n}{5}$ B. m, the median-of-medians C. $0.7n$ D. $W\left(\frac{n}{5}\right)$
14. The main difference between MTF and OPT for self-organizing linear lists is:
    A. OPT can do transpositions
    B. OPT is given the entire request sequence in advance, while MTF receives the requests one-at-a-time
    C. MTF is given the entire request sequence in advance, while OPT receives the requests one-at-a-time
    D. MTF counts inversions
15. Assuming a random n-permutation is provided, the expected number of hires for the hiring problem is:
    A. H_n B. 2 C. $\sqrt{n}$ D. $\ln\ln n$

16. Suppose there are 20 coupon types for the coupon collecting problem. Each of the coupon types is identified by an integer in the range 1 to 20. If you are given two random cereal boxes, what is the probability (ignoring the order in which the coupons were obtained) that the two coupons have consecutive numbers, i.e. some i and $i + 1$? (5 points)

CSE 5311 Name _____

Test 1 - Open Book

Summer 2010 Student ID # _____

1. The hash table below was created using double hashing with Brent's rehash. The initial slot ($h_1(key)$) and rehashing increment ($h_2(key)$) are given for each *key*. Show the result from inserting 1300 using Brent's rehash when $h_1(1300) = 5$ and $h_2(1300) = 3$. (10 points)

	key	$h_1(key)$	$h_2(key)$
0			
1	1000	6	2
2			
3	1200	3	3
4	500	3	1
5	12	5	1
6	27	6	3

	key
0	
1	
2	
3	
4	
5	
6	

2. Evaluate the following recurrences using the master method. Indicate the case that is used for each. (10 points)

a. $T(n) = 4T\left(\frac{n}{2}\right) + n^3$

b. $T(n) = 4T\left(\frac{n}{2}\right) + n^2$

c. $T(n) = 4T\left(\frac{n}{2}\right) + 1$

3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
n=7;
q[0]=0.06;
p[1]=0.15;
q[1]=0.03;
p[2]=0.15;
q[2]=0.05;
p[3]=0.1;
q[3]=0.01;
p[4]=0.05;
q[4]=0.05;
p[5]=0.02;
q[5]=0.05;
p[6]=0.1;
q[6]=0.05;
p[7]=0.08;
q[7]=0.05;
w[0][0]=0.060000
w[0][1]=0.240000
w[0][2]=0.440000
w[0][3]=0.550000
w[0][4]=0.650000
w[0][5]=0.720000
w[0][6]=0.870000
w[0][7]=1.000000
w[1][1]=0.030000
w[1][2]=0.230000
w[1][3]=0.340000
w[1][4]=0.440000
w[1][5]=0.510000
w[1][6]=0.660000
w[1][7]=0.790000
w[2][2]=0.050000
w[2][3]=0.160000
w[2][4]=0.260000
w[2][5]=0.330000
w[2][6]=0.480000
w[2][7]=0.610000
w[3][3]=0.010000
w[3][4]=0.110000
w[3][5]=0.180000
w[3][6]=0.330000
w[3][7]=0.460000
w[4][4]=0.050000
w[4][5]=0.120000
w[4][6]=0.270000
w[4][7]=0.400000
w[5][5]=0.050000
w[5][6]=0.200000
w[5][7]=0.330000
w[6][6]=0.050000
w[6][7]=0.180000
w[7][7]=0.050000
Average probe length is ???
trees in parenthesized prefix
c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
```

```
c(3,3) cost
0.000000
c(4,4) cost
0.000000
c(5,5) cost
0.000000
c(6,6) cost
0.000000
c(7,7) cost
0.000000
c(0,1) cost
0.240000 1
c(1,2) cost
0.230000 2
c(2,3) cost 0.160000 3
c(3,4) cost 0.110000 4
c(4,5) cost 0.120000 5
c(5,6) cost 0.200000 6
c(6,7) cost 0.180000 7
c(0,2) cost 0.670000 1(,2)
c(1,3) cost 0.500000 2(,3)
c(2,4) cost 0.370000 3(,4)
c(3,5) cost 0.290000 5(4,)
c(4,6) cost 0.390000 6(5,)
c(5,7) cost 0.510000 6(,7)
c(0,3) cost 0.950000 2(1,3)
c(1,4) cost 0.780000 3(2,4)
c(2,5) cost 0.610000 4(3,5)
c(3,6) cost 0.620000 6(5(4,),)
c(4,7) cost 0.700000 6(5,7)
c(0,4) cost 1.260000 2(1,3(,4))
c(1,5) cost 1.030000 3(2,5(4,))
c(2,6) cost 1.030000 4(3,6(5,))
c(3,7) cost 0.930000 6(5(4,),7)
c(0,5) cost 1.570000 2(1,4(3,5))
c(1,6) cost 1.510000 3(2,6(5(4,),))
c(2,7) cost 1.400000 6(4(3,5),7)
c(0,6) cost 2.140000 2(1,4(3,6(5,)))
c(1,7) cost 1.950000 3(2,6(5(4,),7))
c(0,7) cost ??? ????????????????
```

4. Fill in the `min` and `max` blanks for the following instance of a van Emde Boas tree for the set {0, 1, 8, 10, 11, 12, 13}. You should give these as values in the local universe (`0..u-1`). Instead of using the symbol "/" for NIL, use the symbol "∅". (10 points)

```
root (base 0) u 16 min _______ max _______
    summary (base 0) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 0) u 2 min _______ max _______
        cluster[1] (base 2) u 2 min _______ max _______
    cluster[0] (base 0) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 0) u 2 min _______ max _______
        cluster[1] (base 2) u 2 min _______ max _______
    cluster[1] (base 4) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 4) u 2 min _______ max _______
        cluster[1] (base 6) u 2 min _______ max _______
    cluster[2] (base 8) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 8) u 2 min _______ max _______
        cluster[1] (base 10) u 2 min _______ max _______
    cluster[3] (base 12) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 12) u 2 min _______ max _______
        cluster[1] (base 14) u 2 min _______ max _______
```

5. Give the range of possible heights for an AVL tree with 500 keys. Your answer should be two natural numbers giving the minimum and maximum heights. (A tree with one node has height 0.) Show your work! (10 points)

CSE 5311 Name _____

Test 2 - Closed Book

Summer 2010 Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. Constructing a suffix array for a sequence with n symbols by using the original Manber-Myers radix sort construction has this worst-case time:

   A. $\Theta(n)$ B. $\Theta(n \log n)$ C. $\Theta(n^2)$ D. $\Theta(n^2 \log n)$

2. What is the nature of the signature function for the Karp-Rabin method?
   A. a polynomial of arbitrary precision implemented using a bignum package
   B. similar to a double hash function for string keys
   C. it is similar to the KMP failure links
   D. the remainder by discarding the overflow for a polynomial
3. Which of the following is helpful if you wish to know the farthest pair in a set of points in 2D?
   A. Euclidean minimum spanning tree

B. Voronoi diagram
C. Convex hull
D. Delaunay triangulation

4. While constructing a suffix array for a sequence with n symbols by using radix sorts, sequence symbols are used in which of the radix sorts?
   A. The last one
   B. The first one
   C. None of them
   D. All of them
5. In a maximum flow problem, the number of augmenting paths in a flow decomposition is bounded by:
   A. f
   B. E
   C. V
   D. $O(VE)$
6. Which algorithm is defined using the notions of left-turn and right-turn?
   A. Graham scan
   B. Closest points in 2-d space
   C. Jarvis march
   D. Suffix array construction
7. What data structure is used for the sweep-line status when computing the 2-d closest pair?
   A. BST of points with x-coordinates as the key
   B. Sorted array by ascending x-coordinates
   C. BST of points with y-coordinates as the key
   D. Interval tree
8. The four russians' concept is to:
   A. Trade-off between enumerating situations and referencing these situations
   B. Trade-off between scalar additions and multiplications
   C. Implement longest common subsequences using linear space
   D. Pack bits into an efficient storage unit
9. When coloring the edges of a graph, a dc-path gets inverted because:
   A. All edges in the fan will be colored with d or c.
   B. d is a free color for all fan vertices.
   C. d is the free color for two fan vertices.
   D. We are trying to minimize the number of colors used by the path.
10. Which longest common subsequence method is potentially the most time-consuming?
    A. Compact version of dynamic programming based on divide-and-conquer
    B. Method based on subsequence indices and longest strictly increasing subsequence
    C. Ordinary dynamic programming using full matrix
    D. Four russians' implementation of C.
11. The reduction from 3-sat to 3-colorability will give a graph requiring four colors when:
    A. The 3-sat instance is satisfiable.
    B. The 3-sat instance is a tautology.
    C. The 3-sat instance is unsatisfiable.
    D. Removing one clause from the 3-sat instance will leave a satisifiable set of clauses.
12. Which of the following problems is not NP-complete? (Assume P ≠ NP)
    A. Testing if a table is in sorted order
    B. 3-satisfiability
    C. Testing if a graph is 3-colorable
    D. Testing if the number of colors needed to edge color a graph is the degree of the graph
13. Which of the following is a deficiency of the maximum capacity path technique?
    A. An augmenting path is blocked if it introduces a cycle of flow.
    B. The maximum number of potential augmenting paths depends on the achievable flow, in addition to the number of vertices and edges.
    C. Augmenting paths will be discovered in descending incremental flow increase order.
    D. Flow decomposition must be applied.
14. Which is not true regarding the first-fit decreasing method for bin packing?
    A. Each object is placed by going left-to-right until it fits in a bin or a new bin is allocated.
    B. The number of objects placed in bins beyond the optimal number is arbitrary.
    C. The number of objects with sizes larger than ½ is a lower bound on the optimal number of bins.

D. Objects placed in bins beyond the optimal number have sizes no larger than ½.

15. The most general approximation result that can be achieved for an NP-hard problem is:
   A. Approximation Algorithm
   B. Approximation Scheme
   C. Polynomial-time Approximation Scheme
   D. Fully Polynomial-time Approximation Scheme

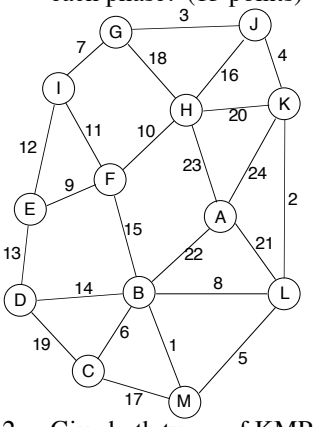
16. What is the Steiner tree problem? (5 points)

CSE 5311 Name _____

Test 2 - Open Book

Summer 2010 Student ID # _____

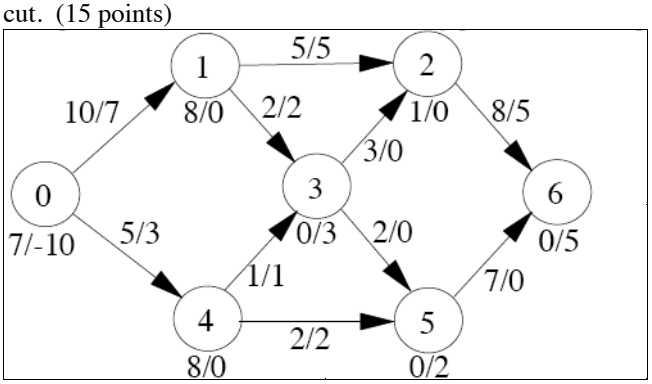
1. Demonstrate Boruvka's algorithm on the following graph. Be sure to indicate the edges that are inserted into the MST in each phase. (15 points)



2. Give both types of KMP failure link tables and the table produced by the Z algorithm for the following string: (20 points)

```
 0  a
 1  c
 2  a
 3  b
 4  a
 5  c
 6  a
 7  c
 8  a
 9  b
10  a
11  c
12  a
13  b
14  a
15  c
16  a
17  c
18  a
19  b
20  a
21  c
22  a
23  c
24  a
```

3. List the remaining operations to complete this instance of network flows by push-relabel. In addition, give a minimum cut. (15 points)



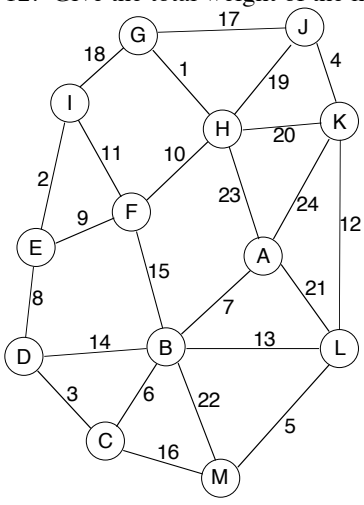
CSE 5311 Name _____

Test 1 - Closed Book

Summer 2011 Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. The number of potential probe sequences when using double hashing with a table with m entries (m is prime) is:
   A. $O(\log m)$ B. m C. $m(m-1)$ D. $m!$
2. What is the worst-case number of rotations when performing insertion on an AVL tree?
   A. $\Theta(1)$ B. $\Theta(\log n)$ C. $\Theta(n)$ D. No rotations are ever used
3. Suppose you already have 10 different coupons when there are 20 coupon types. What is the expected number of boxes for obtaining a coupon different from the 10 you already have?
   A. 2 B. 3 C. 4 D. 5
4. Which of the following is not true regarding Bloom filters?
   A. An indication that a candidate element is not in the set is always correct.
   B. They are an especially compact dictionary representation.
   C. Several hash functions are used.
   D. The optimal bit array size depends on the number of items and the false positive probability.
5. If the universe size at the root of a van Emde Boas tree is u, then the number of children is:
   A. 2 B. $\log u$ C. $\log \log u$ D. $\sqrt{u}$
6. The summary structures are critical for implementing which operation efficiently for van Emde Boas trees?
   A. Member B. Minimum C. Successor D. all of these
7. How many inversions are there for the lists 1, 2, 5, 4, 3 and 2, 5, 4, 3, 1?
   A. 2 B. 3 C. 4 D. 5
8. Suppose there are n entries in a Fibonacci heap. The maximum number of trees is:
   A. $\theta(\log \log n)$ B. $\theta(\log n)$ C. $\theta(n)$ D. $\theta(n \log n)$
9. Which property does not hold for binomial heaps?
   A. DECREASE-KEY takes $O(1)$ time.
   B. MINIMUM takes $O(\log n)$ time.
   C. Performing n INSERT operations into an empty heap will take $O(n)$ time.
   D. The number of trees is based on the binary representation of the number of stored items.
10. What is the main contribution of leftist heaps?
    A. The minimum is found in $O(1)$ time.
    B. The amortized complexity of DECREASE-KEY is $O(1)$.
    C. The height of the tree is $O(\log n)$.
    D. The UNION is computed in $O(\log n)$ time.
11. Dynamic optimality is a concept involving the comparison of
    A. a key-comparison based data structure to hashing.
    B. amortized complexity to actual complexity.
    C. an online data structure to a fixed, unchanging data structure.
    D. an online data structure to an offline data structure.
12. Give the total weight of the minimum spanning tree for this graph. (5 points)



13. Give an example of a binary search tree that can be colored as a legal red-black tree, but does not satisfy the AVL balance conditions. (6 points)
14. Give the binomial (min) heap that results from inserting 1, 2, 3, 4, 5, 6, 7, 8 (in that order) into an empty heap. (6 points)

CSE 5311 Name _____

Test 1 - Open Book

Summer 2011 Student ID # _____

1. The hash table below was created using double hashing with Brent's rehash. The initial slot ($h_1(key)$) and rehashing increment ($h_2(key)$) are given for each *key*. Show the result from inserting 7095 using Brent's rehash when $h_1(7095) = 4$ and $h_2(7095) = 4$. (10 points)

	key	$h_1(key)$	$h_2(key)$
0	7000	0	5
1	7081	4	2
2	7002	2	1
3			
4	7004	4	3
5			
6	7006	6	5

	key
0	
1	
2	
3	
4	
5	
6	

2. Evaluate the following recurrences using the master method. (10 points)

a. $T(n) = T(0.7n) + n$

b. $T(n) = T(0.7n) + 1$

c. $T(n) = 16T\left(\frac{n}{2}\right) + n^3$

3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
n=6;
q[0]=0.01;
key[1]=10;
p[1]=0.19;
q[1]=0.02;
key[2]=20;
p[2]=0.1;
q[2]=0.03;
key[3]=30;
p[3]=0.2;
q[3]=0.04;
key[4]=40;
p[4]=0.2;
q[4]=0.0;
key[5]=50;
p[5]=0.02;
q[5]=0.04;
key[6]=60;
p[6]=0.12;
q[6]=0.03;
```

```
w[0][0]=0.010000
w[0][1]=0.220000
w[0][2]=0.350000
w[0][3]=0.590000
w[0][4]=0.790000
w[0][5]=0.850000
w[0][6]=1.000000
w[1][1]=0.020000
w[1][2]=0.150000
w[1][3]=0.390000
w[1][4]=0.590000
w[1][5]=0.650000
w[1][6]=0.800000
w[2][2]=0.030000
w[2][3]=0.270000
w[2][4]=0.470000
w[2][5]=0.530000
w[2][6]=0.680000
w[3][3]=0.040000
w[3][4]=0.240000
```

```
w[3][5]=0.300000
w[3][6]=0.450000
w[4][4]=0.000000
w[4][5]=0.060000
w[4][6]=0.210000
w[5][5]=0.040000
w[5][6]=0.190000
w[6][6]=0.030000
Building c(0,2) using roots 1
   thru 2
Building c(1,3) using roots 2
   thru 3
Building c(2,4) using roots 3
   thru 4
Building c(3,5) using roots 4
   thru 5
Building c(4,6) using roots 5
   thru 6
Building c(0,3) using roots 1
   thru 3
```

```
Building c(1,4)   Building c(1,6) using roots 3     c(1,3) cost 0.540000 30(20,)
   using roots       thru 4                         c(2,4) cost 0.710000 30(,40)
   3 thru 3       Building c(0,6) using roots ?     c(3,5) cost 0.360000 40(,50)
Building c(2,5)      thru ?                         c(4,6) cost 0.270000 60(50,)
   using roots    Counts - root trick 29 without    c(0,3) cost 1.080000 20(10,30)
   3 thru 4          root trick 50                  c(1,4) cost 0.980000 30(20,40)
Building c(3,6)   Average probe length is ????     c(2,5) cost 0.860000 40(30,50)
   using roots    trees in parenthesized prefix     c(3,6) cost 0.720000 40(,60(50,))
   4 thru 6       c(0,0) cost 0.000000              c(0,4) cost 1.530000
Building c(0,4)   c(1,1) cost 0.000000                 30(10(,20),40)
   using roots    c(2,2) cost 0.000000              c(1,5) cost 1.160000
   2 thru 3       c(3,3) cost 0.000000                 30(20,40(,50))
Building c(1,5)   c(4,4) cost 0.000000              c(2,6) cost 1.220000
   using roots    c(5,5) cost 0.000000                 40(30,60(50,))
   3 thru 4       c(6,6) cost 0.000000              c(0,5) cost 1.710000
Building c(2,6)   c(0,1) cost 0.220000 10              30(10(,20),40(,50))
   using roots    c(1,2) cost 0.150000 20           c(1,6) cost 1.610000
   4 thru 4       c(2,3) cost 0.270000 30              40(30(20,),60(50,))
Building c(0,5)   c(3,4) cost 0.240000 40           c(0,6) cost ????????
   using roots    c(4,5) cost 0.060000 50              ????????????????
   3 thru 3       c(5,6) cost 0.190000 60
                  c(0,2) cost 0.500000 10(,20)
```

4. Fill in the `min` and `max` blanks for the following instance of a van Emde Boas tree for the set {1, 8, 9, 10, 12, 13, 14}. You should give these as values in the local universe (`0..u-1`). Instead of using the symbol "/" for NIL, use the symbol "∅". (10 points)

```
root (base 0) u 16 min _______ max _______
    summary (base 0) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 0) u 2 min _______ max _______
        cluster[1] (base 2) u 2 min _______ max _______
    cluster[0] (base 0) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 0) u 2 min _______ max _______
        cluster[1] (base 2) u 2 min _______ max _______
    cluster[1] (base 4) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 4) u 2 min _______ max _______
        cluster[1] (base 6) u 2 min _______ max _______
    cluster[2] (base 8) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 8) u 2 min _______ max _______
        cluster[1] (base 10) u 2 min _______ max _______
    cluster[3] (base 12) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 12) u 2 min _______ max _______
        cluster[1] (base 14) u 2 min _______ max _______
```

Give the range of possible heights for a red-black tree with 200 keys. Your answer should be two natural numbers giving the minimum and maximum heights. (A tree with one node has height 0.) Show your work! (10 points)

CSE 5311 Name _____

Test 2 - Closed Book

Summer 2011 Last 4 Digits of Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. When performing selection in worst-cast linear time for n numbers, roughly how many column medians are computed in the first round?

   A. $W\left(\frac{n}{5}\right)$

   B. m, the median-of-medians

   C. $.7n$

D. $\frac{n}{5}$

2. What is the nature of the signature function for the Karp-Rabin method?
   A. the remainder by discarding the overflow for a polynomial
   B. it is similar to the KMP failure links
   C. similar to a double hash function for string keys
   D. a polynomial of arbitrary precision implemented using a bignum package
3. Which of the following is helpful if you wish to know the farthest pair in a set of points in 2D?
   A. Euclidean minimum spanning tree
   B. Delaunay triangulation
   C. Convex hull
   D. Voronoi diagram
4. Which of the following is not a goal for suffix array construction methods?
   A. Fast in practice
   B. Based on radix sort
   C. Lightweight
   D. Linear-time in worse case
5. In a maximum flow problem, the number of augmenting paths in a flow decomposition is bounded by:
   A. V
   B. $\mathrm{O}(VE)$
   C. f
   D. E
6. Which algorithm is defined using the notions of left-turn and right-turn?
   A. Graham scan
   B. Closest points in 2-d space
   C. Suffix array construction
   D. Jarvis march
7. What data structure is used for the sweep-line status when computing the 2-d closest pair?
   A. BST of points with y-coordinates as the key
   B. Interval tree
   C. BST of points with x-coordinates as the key
   D. Sorted array by ascending x-coordinates
8. The four russians' concept is to:
   A. Pack bits into an efficient storage unit
   B. Trade-off between enumerating situations and referencing these situations
   C. Implement longest common subsequences using linear space
   D. Trade-off between scalar additions and multiplications
9. When coloring the edges of a graph, a dc-path gets inverted because:
   A. All edges in the fan will be colored with d or c.
   B. d is a free color for all fan vertices.
   C. d is the free color for two fan vertices.
   D. We are trying to minimize the number of colors used by the path.
10. How many times will `-1` occur in the style 2 fail link table for the pattern `abacabac`?
    A. 1
    B. 2
    C. 3
    D. 4
11. How many times will `-1` occur in the style 1 fail link table for the pattern `abacabac`?
    A. 1
    B. 2
    C. 3
    D. 4
12. Which of the following problems is NP-complete? (Assume P ≠ NP)
    A. 3-satisfiability
    B. Testing if a graph is 2-colorable
    C. Testing if a table is in sorted order
    D. Verifying a solution to traveling salesperson
13. Which of the following is a deficiency of the maximum capacity path technique?
    A. Augmenting paths will be discovered in descending incremental flow increase order.

B. Flow decomposition must be applied.
C. An augmenting path is blocked if it introduces a cycle of flow.
D. The maximum number of potential augmenting paths depends on the achievable flow, in addition to the number of vertices and edges.

14. Constructing a suffix array for a sequence with n symbols by using an optimal key-comparison sort has this worst-case time:

A. $\Theta(n)$ B. $\Theta(n \log n)$ C. $\Theta(n^2)$ D. $\Theta(n^2 \log n)$

15. Which of the following does not have a polynomial-time approximation algorithm?
A. Bin packing
B. Edge coloring
C. Traveling salesperson with triangle inequality
D. Vertex coloring

16. Determine a monotone longest increasing subsequence for the sequence below. (5 points)

```
1  4  3  2  3  4  3  6  3  4
```

CSE 5311 Name _____

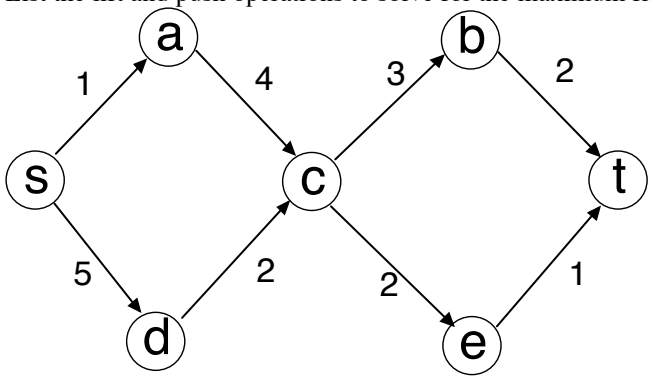
Test 2 - Open Book

Summer 2011 Last 4 Digits of Student ID # _____

1. Fill in the blanks in the following instance of a suffix array with lcp values and ranks. As usual, s[21] is NULL ('\0'). (15 points)

i	sa	suffix	lcp	s	rank	lcp[rank]
0	21		-1	0	__	__
1	20	0	0	1	__	__
2	7	00100101001010	1	0	__	__
3	15	001010	4	0	__	__
4	2	0010100100101001010	6	1	20	4
5	10	00101001010	9	0	7	3
6	18	010	1	1	15	2
7	5	0100100101001010	3	0	2	__
8	13	01001010	6	0	10	__
9	0	0100101001001010	8	1	18	__
10	___	0100101001010	11	0	5	9
11	16	01010	3	0	13	8
12	3	010100100101001010	5	1	21	7
13	11	__101001010	8	0	8	6
14	19	__0	0	1	16	5
15	6	100100101001010	2	0	3	4
16	14	1001010	5	0	11	3
17	1	10010100100101001010	7	1	19	2
18	9	100101001010	10	0	6	1
19	___	1010	2	1	14	0
20	4	10100100101001010	4	0	1	0
21	12	101001010	7		0	-1

2. List the lift and push operations to solve for the maximum flow. In addition, give a minimum cut. (15 points)



3. Give the result of the Z algorithm and both KMP methods for the following sequence. (20 points)

```
0 a
```

1 c

2 a

3 b

4 a

5 c

6 a

7 c

8 a

9 b

10 a

11 c

12 a

13 b

14 a

15 c

16 a

17 c

18 a

19 b

20 a

21 c

22 a

23 c

24 a

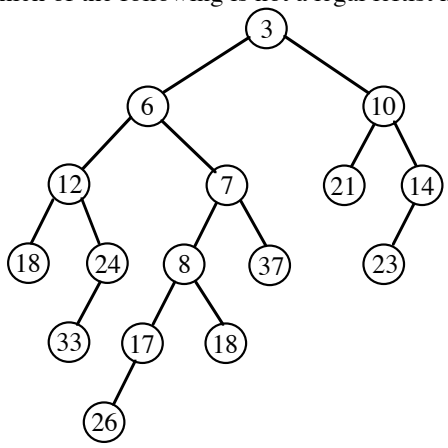
25 b

CSE 5311 Name _____
Test 1 - Closed Book
Summer 2012 Last 4 Digits of Student ID # _____

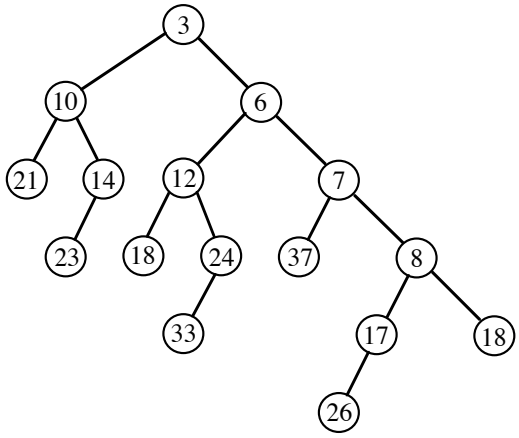
Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. When are Fibonacci trees used?
   A. Defining the potential function for Fibonacci heaps.
   B. Constructing a priority queue with excellent amortized complexity for DECREASE-KEY.
   C. Demonstrating worst-case behaviors for red-black trees.
   D. Demonstrating worst-case behaviors for AVL trees.
2. Which of the following is not a legal leftist heap?

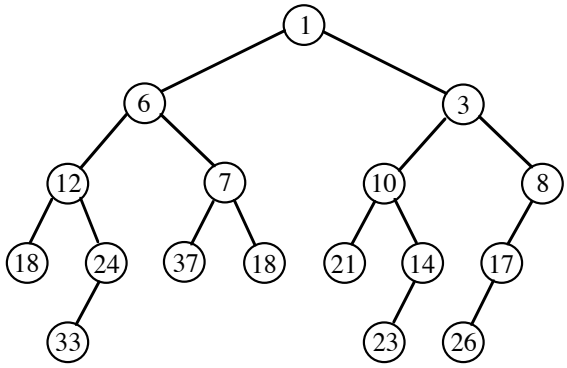
A.



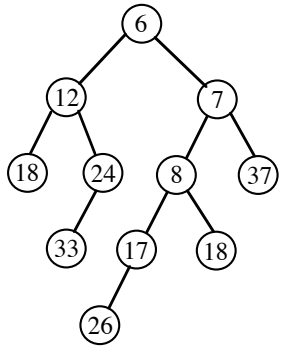
B.



C.



D.



3. Dynamic optimality is a concept involving the comparison of
   A. a key-comparison based data structure to hashing.
   B. amortized complexity to actual complexity.
   C. an online data structure to an offline data structure.
   D. an online data structure to a fixed, unchanging data structure.
4. Suppose you already have 16 different coupons when there are 20 coupon types. What is the expected number of boxes for obtaining a coupon different from the 16 you already have?
   A. 4 B. 5 C. 6 D. 10
5. If the universe size at the root of a van Emde Boas tree is u, then the number of children is:
   A. $\lg u$ B. $\log u$ C. $\log \log u$ D. $\sqrt{u}$
6. To support computing prefix sums of all keys that are no larger than some query key, an augmented binary search tree stores the following at every node:
   A. the sum of all keys in the entire tree
   B. the sum of all keys in the left subtree
   C. the sum of all keys that are no larger than the stored key
   D. the sum of all keys stored in the subtree rooted by this node
7. Which Fibonacci heap operation has $\Theta(\log n)$ worst-case actual cost?
   A. FIB-HEAP-EXTRACT-MIN B. FIB-HEAP-UNION
   C. FIB-HEAP-DECREASE-KEY D. FIB-HEAP-DELETE
8. What is the main contribution of leftist heaps?
   A. The minimum is found in $O(1)$ time.
   B. The amortized complexity of DECREASE-KEY is $O(1)$.
   C. The height of the tree is $O(\log n)$.
   D. The UNION is computed in $O(\log n)$ time.
9. The number of potential probe sequences when using linear probing with a table with m entries (m is prime) is:

A. $O(\log m)$ B. m C. $m(m-1)$ D. $m!$

10. Which of the following is not true regarding the amortized analysis of binary tree traversals?
    A. INIT had an amortized cost of 1.
    B. SUCC had an actual cost determined by the number of edges followed.
    C. SUCC had an amortized cost of 2.
    D. The potential was defined with regard to the type of traversal being performed.
11. Which data structure is not used to implement a dictionary?
    A. vEB tree B. Red-black tree C. Self-organizing list D. Splay tree
12. The reason for marking nodes in a Fibonacci heap is:
    A. to indicate nodes that have lost a child since becoming a child themselves.
    B. to allow computing the value of the potential function.
    C. to assure that the structure is a Fibonacci heap rather than a binomial heap.
    D. to improve the performance of CONSOLIDATE.
13. If a Fibonacci tree appears as a subtree of an AVL tree, which nodes would be assigned a balance factor of 0?
    A. none of them B. only the root
    C. only the leaves D. the leaves and the root
14. The main difference between MTF and OPT for self-organizing linear lists is:
    A. MTF is given the entire request sequence in advance, while OPT receives the requests one-at-a-time
    B. OPT counts inversions
    C. OPT is given the entire request sequence in advance, while MTF receives the requests one-at-a-time
    D. MTF can do transpositions
15. To reduce the probability of having any collisions to < 0.5 when hashing n keys, the table should have at least this number of elements.
    A. n B. $n \ln n$ C. n^2 D. n^3
16. Suppose you roll three standard six-sided dice. What is the probability that the sum of the three values rolled does not exceed 11? Show your work. (5 points)

CSE 5311 Name _____
Test 1 - Open Book
Summer 2012 Last 4 Digits of Student ID # _____

1. Fill in the `min` and `max` blanks for the following instance of a van Emde Boas tree for the set {1, 2, 3, 8, 9, 10, 14, 15}. You should give these as values in the local universe (`0..u-1`). Instead of using the symbol "/" for NIL, use the symbol "∅". (10 points)

```
root (base 0) u 16 min _______ max _______
    summary (base 0) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 0) u 2 min _______ max _______
        cluster[1] (base 2) u 2 min _______ max _______
    cluster[0] (base 0) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 0) u 2 min _______ max _______
        cluster[1] (base 2) u 2 min _______ max _______
    cluster[1] (base 4) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 4) u 2 min _______ max _______
        cluster[1] (base 6) u 2 min _______ max _______
    cluster[2] (base 8) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 8) u 2 min _______ max _______
        cluster[1] (base 10) u 2 min _______ max _______
    cluster[3] (base 12) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 12) u 2 min _______ max _______
        cluster[1] (base 14) u 2 min _______ max _______
```

2. Evaluate the following recurrences using the master method. Indicate the case that is used for each. (10 points)

a. $T(n) = 4T\left(\frac{n}{2}\right) + n^3$

b. $T(n) = 4T\left(\frac{n}{2}\right) + n^2$

c. $T(n) = 4T\left(\frac{n}{2}\right) + 1$

3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
n=6;
q[0]=0.01;
key[1]=10;
p[1]=0.19;
q[1]=0.02;
key[2]=20;
p[2]=0.1;
q[2]=0.03;
key[3]=30;
p[3]=0.2;
q[3]=0.04;
key[4]=40;
p[4]=0.2;
q[4]=0.0;
key[5]=50;
p[5]=0.02;
q[5]=0.04;
key[6]=60;
p[6]=0.12;
q[6]=0.03;
w[0][0]=0.01000
   0
w[0][1]=0.22000
   0
w[0][2]=0.35000
   0
w[0][3]=0.59000
   0
w[0][4]=0.79000
   0
w[0][5]=0.85000
   0
w[0][6]=1.00000
   0
w[1][1]=0.02000
   0
w[1][2]=0.15000
   0
w[1][3]=0.39000
   0
w[1][4]=0.59000
   0
w[1][5]=0.650000
w[1][6]=0.800000
w[2][2]=0.030000
w[2][3]=0.270000
w[2][4]=0.470000
w[2][5]=0.530000
w[2][6]=0.680000
w[3][3]=0.040000
w[3][4]=0.240000
w[3][5]=0.300000
w[3][6]=0.450000
w[4][4]=0.000000
w[4][5]=0.060000
w[4][6]=0.210000
w[5][5]=0.040000
w[5][6]=0.190000
w[6][6]=0.030000
Building c(0,2) using roots 1
   thru 2
Building c(1,3) using roots 2
   thru 3
Building c(2,4) using roots 3
   thru 4
Building c(3,5) using roots 4
   thru 5
Building c(4,6) using roots 5
   thru 6
Building c(0,3) using roots 1
   thru 3
Building c(1,4) using roots 3
   thru 3
Building c(2,5) using roots 3
   thru 4
Building c(3,6) using roots 4
   thru 6
Building c(0,4) using roots 2
   thru 3
Building c(1,5) using roots 3
   thru 4
Building c(2,6) using roots 4
   thru 4
Building c(0,5) using roots 3
   thru 3
Building c(1,6) using roots 3
   thru 4
Building c(0,6) using roots ?
   thru ?
Counts - root trick 29 without
   root trick 50
Average probe length is ????
trees in parenthesized prefix
c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
c(3,3) cost 0.000000
c(4,4) cost 0.000000
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(0,1) cost 0.220000 10
c(1,2) cost 0.150000 20
c(2,3) cost 0.270000 30
c(3,4) cost 0.240000 40
c(4,5) cost 0.060000 50
c(5,6) cost 0.190000 60
c(0,2) cost 0.500000 10(,20)
c(1,3) cost 0.540000 30(20,)
c(2,4) cost 0.710000 30(,40)
c(3,5) cost 0.360000 40(,50)
c(4,6) cost 0.270000 60(50,)
c(0,3) cost 1.080000 20(10,30)
c(1,4) cost 0.980000 30(20,40)
c(2,5) cost 0.860000 40(30,50)
c(3,6) cost 0.720000 40(,60(50,))
c(0,4) cost 1.530000
   30(10(,20),40)
c(1,5) cost 1.160000
   30(20,40(,50))
c(2,6) cost 1.220000
   40(30,60(50,))
c(0,5) cost 1.710000
   30(10(,20),40(,50))
c(1,6) cost 1.610000
   40(30(20,),60(50,))
c(0,6) cost ????????
   ?????????????
```

4. The hash table below was created using double hashing with Brent's rehash. The initial slot ($h_1(key)$) and rehashing increment ($h_2(key)$) are given for each *key*. Show the result from inserting 5559 using Brent's rehash when $h_1(5559) = 1$ and $h_2(5559) = 4$. (10 points)

	key	$h_1(key)$	$h_2(key)$
0			
1	2199	1	4
2	3299	2	6
3	5500	5	5
4	4400	4	3
5	1111	5	2
6			

	key
0	
1	
2	
3	
4	
5	
6	

5. Give the precise range of possible heights for an AVL tree with 300 keys. (10 points)

CSE 5311 Name _____
Test 2 - Closed Book
Summer 2012 Student ID # _____

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. Constructing a suffix array for a sequence with *n* symbols by using the original Manber-Myers radix sort construction has this worst-case time:

   A. $\Theta(n)$ B. $\Theta(n \log n)$ C. $\Theta(n^2)$ D. $\Theta(n^2 \log n)$

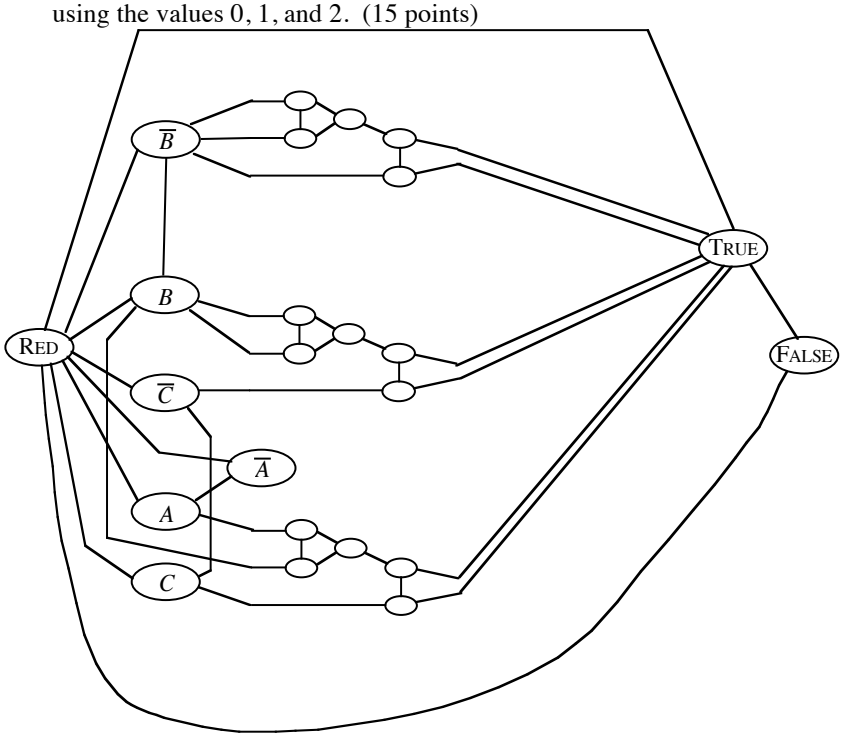
2. What is the nature of the signature function for the Karp-Rabin method?
   A. it is similar to the KMP failure links
   B. the remainder by discarding the overflow for a polynomial
   C. a polynomial of arbitrary precision implemented using a bignum package
   D. similar to a double hash function for string keys
3. Which of the following is helpful if you wish to know the farthest pair in a set of points in 2D?
   A. Voronoi diagram
   B. Convex hull
   C. Delaunay triangulation
   D. Euclidean minimum spanning tree
4. While constructing a suffix array for a sequence with *n* symbols by using radix sorts, sequence symbols are used in which of the radix sorts?
   A. The last one
   B. The first one
   C. None of them
   D. All of them
5. In a maximum flow problem, the number of augmenting paths in a flow decomposition is bounded by:
   A. the number of edges in the graph
   B. the number of saturated edges
   C. the number of unsaturated edges
   D. the number of vertices
6. Which algorithm is defined using the notions of left-turn and right-turn?
   A. Suffix array construction
   B. Graham scan
   C. Closest points in 2-d space
   D. Jarvis march
7. What data structure is used for the sweep-line status when computing the 2-d closest pair?
   A. BST of points with y-coordinates as the key
   B. Sorted array by ascending x-coordinates
   C. BST of points with x-coordinates as the key
   D. Interval tree
8. The four russians' concept is to:
   A. Trade-off between generating situations and referencing these situations
   B. Trade-off between scalar additions and multiplications
   C. Implement longest common subsequences using linear space

D. Pack bits into an efficient storage unit

9. When coloring the edges of a graph, a dc-path gets inverted because:
   A. d is the free color for two fan vertices.
   B. We are trying to minimize the number of colors used by the path.
   C. All edges in the fan will be colored with d or c.
   D. d is a free color for all fan vertices.
10. What is the nature of the linear-space method for the longest common subsequence problem?
   A. Build a suffix array and lcp array for the concatenated input sequences
   B. Recursive divide-and-conquer
   C. Radix sort
   D. Use a polynomial for the signature function
11. The reduction from 3-sat to 3-colorability will give a graph requiring four colors when:
   A. Removing one clause from the 3-sat instance will leave a satisifiable set of clauses.
   B. The 3-sat instance is unsatisfiable.
   C. The 3-sat instance is satisfiable.
   D. The 3-sat instance is a tautology.
12. How many times will -1 occur in the style 1 fail link table for the pattern `abcabd`?
   A. 1
   B. 2
   C. 3
   D. 4
13. How many times will -1 occur in the style 2 fail link table for the pattern `abcabd`?
   A. 1
   B. 2
   C. 3
   D. 4
14. Vizing's theorem is the basis for:
   A. Approximate bin packing.
   B. Traveling salesperson approximation when the triangle inequality applies.
   C. Approximate edge coloring.
   D. Approximate a vertex coloring.
15. The technique for approximating a subset cover proceeds by:
   A. Choosing the subset with the smallest fraction of its elements uncovered
   B. Choosing the subset with the smallest number of uncovered elements
   C. Choosing the subset with the largest fraction of its elements uncovered
   D. Choosing the subset with the largest number of uncovered elements
16. Describe the initialization for push/relabel techniques. You may give an example. (5 points)

CSE 5311 Name _____
Test 2 - Open Book
Summer 2012 Student ID # _____

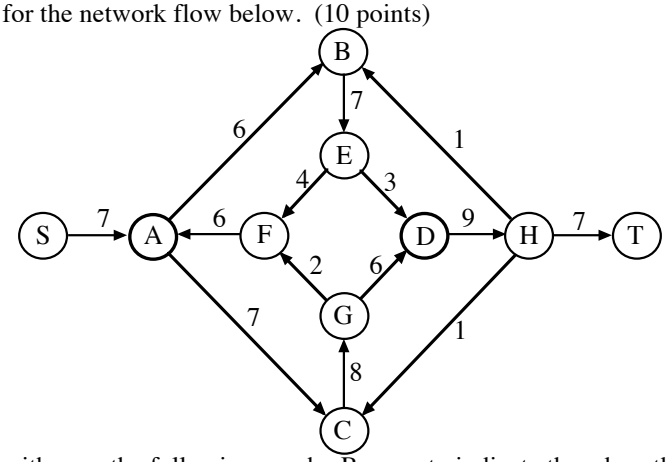
1. Give the set of clauses for the instance below of the reduction from 3-sat to 3-colorability ***and*** an appropriate coloring using the values 0, 1, and 2. (15 points)



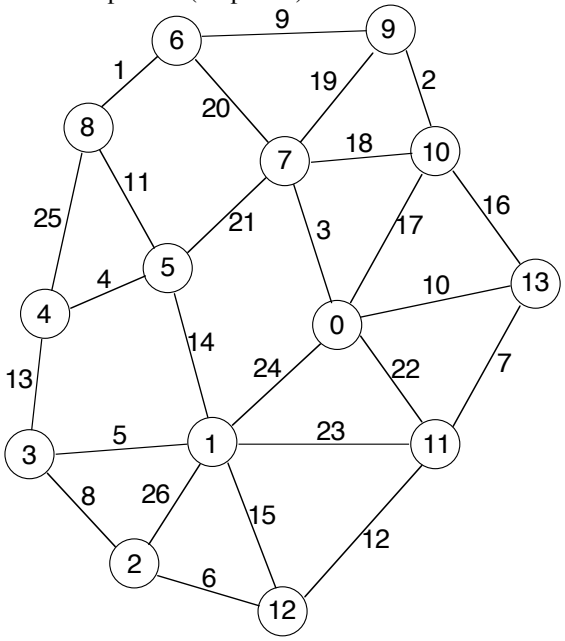
2. Give both types of KMP failure link tables for the following string: (10 points)

0	C
1	A
2	B
3	A
4	B
5	C
6	C
7	A
8	B
9	C
10	A
11	B
12	A
13	B
14	C

3. Give a flow decomposition for the network flow below. (10 points)



4. Demonstrate Boruvka's algorithm on the following graph. Be sure to indicate the edges that are inserted into the MST in each phase. (15 points)



CSE 5311 Name _____

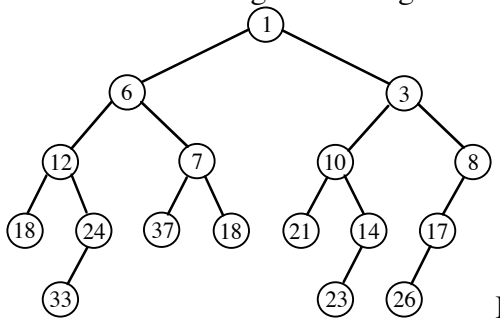
Test 1 - Closed Book

Summer 2014 Student ID # _____

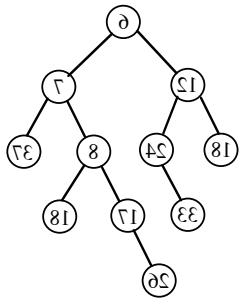
Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. During which operation on a leftist heap may subtree swaps be needed?
   A. DECREASE-KEY B. EXTRACT-MIN
   C. UNION D. All of A., B., and C.
2. Which of the following is not true regarding the amortized analysis of binary tree traversals?
   A. INIT had an amortized cost of 1.
   B. SUCC had an actual cost determined by the number of edges followed.
   C. SUCC had an amortized cost of 2.
   D. The potential was defined with regard to the type of traversal being performed.
3. Suppose you already have 10 different coupons when there are 20 coupon types. What is the expected number of boxes for obtaining a coupon different from the 10 you already have?
   A. 2 B. 3 C. 4 D. 5
4. When is path compression used?
   A. After an insertion into a splay tree.
   B. With a FIND operation.
   C. After an insertion into any type of balanced binary search tree.
   D. With a UNION operation.
5. Which of the following data structures offers similar capabilities and performance characteristics to skip lists?
   A. AVL trees B. Splay trees C. Treap D. Union-find with path compression
6. If a Fibonacci tree appears as a subtree of an AVL tree, which nodes would be assigned a balance factor of 0?
   A. none of them B. only the leaves
   C. only the root D. the leaves and the root
7. How many inversions are there for the lists 1, 2, 5, 4, 3 and 1, 2, 3, 4, 5?
   A. 2 B. 3 C. 4 D. 5
8. Which of the following is not a legal leftist heap?

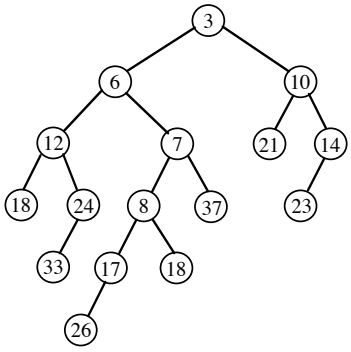
A.



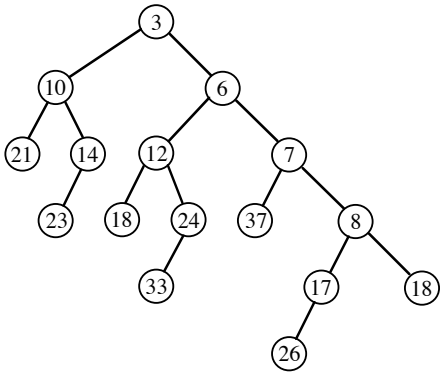
B.



C.



D.



9. Which situation is true regarding a cascading cut that produces c trees for a Fibonacci heap?

A. Both the actual and amortized costs are O(1).
B. The actual cost is O(c). The amortized cost is O(1).
C. The actual cost is O(1). The amortized cost is O(c).
D. The potential can become negative.

10. Which property does not hold for binomial heaps?
A. MINIMUM takes O(1) time.
B. Performing n INSERT operations into an empty heap will take O(n) time.
C. The number of trees is based on the binary representation of the number of stored items.
D. DECREASE-KEY takes O(log n) time.

11. When using Brent's rehash, the number of previously inserted keys that may move is:
A. 1 B. 2 C. $\frac{1}{\alpha}$ D. H_m, where m is the number of stored keys

12. Which priority queue implementation generalizes binary heaps by increasing the branching?
A. Binomial heaps B. d-heaps C. Fibonacci heaps D. Leftist heaps

13. What is minimized in the dynamic programming solution to the subset sum problem?
A. The number of input values used to sum to each $C(i)$ B. S_j
C. m D. The index stored for each $C(i)$

14. What is the worst-case number of rotations when performing deletion on an AVL tree?
A. $\Theta(1)$ B. $\Theta(\log n)$ C. $\Theta(n)$ D. No rotations are ever used

15. If the universe size at the root of a van Emde Boas tree is u, then the number of children is:
A. lg u B. log u C. log log u D. $\sqrt{u}$

16. Give the binomial min-heap that results from inserting 1, 2, 3, 4, 5, 6, 7, 8 (in that order) into an empty heap. (5 points)

CSE 5311 Name _____
Test 1 - Open Book
Summer 2014 Student ID # _____

1. "During Disney's biggest celebration, find one of 50 character Wobblers inside specially-marked ***Kellogg's***® cereals!".

a. Assuming all Wobblers are equally likely to be the one that occurs in a box, what is the expected number of boxes to obtain all 50 Wobblers? (3 points. You may leave your answer as an expression.)
b. Under the same assumption as a., what is the expected number of boxes to obtain just 25 different Wobblers (of the available 50)? (7 points. You may leave your answer as an expression.)

2. Evaluate the following recurrences using the master method. Indicate the case that is used for each. (10 points)

a. $T(n) = 8T\left(\frac{n}{2}\right) + n^4$

b. $T(n) = 8T\left(\frac{n}{2}\right) + n^2$

c. $T(n) = 8T\left(\frac{n}{2}\right) + n^3$

3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
n=6;
q[0]=0.12;
key[1]=10;
p[1]=0.12;
q[1]=0.09;
key[2]=20;
p[2]=0.2;
q[2]=0.03;
key[3]=30;
p[3]=0.1;
q[3]=0.04;
key[4]=40;
p[4]=0.2;
q[4]=0.0;
key[5]=50;
p[5]=0.01;
q[5]=0.04;
key[6]=60;
```

```
p[6]=0.03;
q[6]=0.02;
w[0][0]=0.120000
w[0][1]=0.330000
w[0][2]=0.560000
w[0][3]=0.700000
w[0][4]=0.900000
w[0][5]=0.950000
w[0][6]=1.000000
w[1][1]=0.090000
w[1][2]=0.320000
w[1][3]=0.460000
w[1][4]=0.660000
w[1][5]=0.710000
w[1][6]=0.760000
w[2][2]=0.030000
w[2][3]=0.170000
w[2][4]=0.370000
w[2][5]=0.420000
w[2][6]=0.470000
w[3][3]=0.040000
w[3][4]=0.240000
w[3][5]=0.290000
w[3][6]=0.340000
w[4][4]=0.000000
w[4][5]=0.050000
w[4][6]=0.100000
```

```
w[5][5]=0.040000
w[5][6]=0.090000
w[6][6]=0.020000
Building c(0,2) using roots 1 thru 2
Building c(1,3) using roots 2 thru 3
Building c(2,4) using roots 3 thru 4
Building c(3,5) using roots 4 thru 5
Building c(4,6) using roots 5 thru 6
Building c(0,3) using roots 1 thru 2
Building c(1,4) using roots 2 thru 4
Building c(2,5) using roots 4 thru 4
Building c(3,6) using roots 4 thru 6
Building c(0,4) using roots 2 thru 2
Building c(1,5) using roots 2 thru 4
Building c(2,6) using roots 4 thru 4
Building c(0,5) using roots 2 thru 2
Building c(1,6) using roots 2 thru 4
Building c(0,6) using roots ? thru ?
Counts - root trick 31 without root
    trick 50
Average probe length is ???
trees in parenthesized prefix
c(0,0) cost 0.000000
c(1,1) cost 0.000000
c(2,2) cost 0.000000
c(3,3) cost 0.000000
c(4,4) cost 0.000000
```

```
c(5,5) cost 0.000000
c(6,6) cost 0.000000
c(0,1) cost 0.330000 10
c(1,2) cost 0.320000 20
c(2,3) cost 0.170000 30
c(3,4) cost 0.240000 40
c(4,5) cost 0.050000 50
c(5,6) cost 0.090000 60
c(0,2) cost 0.880000 10(,20)
c(1,3) cost 0.630000 20(,30)
c(2,4) cost 0.540000 40(30,)
c(3,5) cost 0.340000 40(,50)
c(4,6) cost 0.150000 60(50,)
c(0,3) cost 1.200000 20(10,30)
c(1,4) cost 1.200000 20(,40(30,))
c(2,5) cost 0.640000 40(30,50)
c(3,6) cost 0.490000 40(,60(50,))
c(0,4) cost 1.770000 20(10,40(30,))
c(1,5) cost 1.350000 20(,40(30,50))
c(2,6) cost 0.790000 40(30,60(50,))
c(0,5) cost 1.920000 20(10,40(30,50))
c(1,6) cost 1.540000 40(20(,30),60(50,))
c(0,6) cost ??? ????????????????
```

4. Fill in the `min` and `max` blanks for the following instance of a van Emde Boas tree for the set {1, 3, 5, 6, 7, 13, 14, 15}. You should give these as values in the local universe (`0..u-1`). Instead of using the symbol "/" for NIL, use the symbol "∅". (10 points)

```
root (base 0) u 16 min _______ max _______
    summary (base 0) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 0) u 2 min _______ max _______
        cluster[1] (base 2) u 2 min _______ max _______
    cluster[0] (base 0) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 0) u 2 min _______ max _______
        cluster[1] (base 2) u 2 min _______ max _______
    cluster[1] (base 4) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 4) u 2 min _______ max _______
        cluster[1] (base 6) u 2 min _______ max _______
    cluster[2] (base 8) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 8) u 2 min _______ max _______
        cluster[1] (base 10) u 2 min _______ max _______
    cluster[3] (base 12) u 4 min _______ max _______
        summary (base 0) u 2 min _______ max _______
        cluster[0] (base 12) u 2 min _______ max _______
        cluster[1] (base 14) u 2 min _______ max _______
```

5. The hash table below was created using double hashing with Brent's rehash. The initial slot ($h_1(key)$) and rehashing increment ($h_2(key)$) are given for each *key*. Show the result from inserting 1300 using Brent's rehash when $h_1(1300) = 5$ and $h_2(1300) = 3$. (10 points)

	key	$h_1(key)$	$h_2(key)$
0			
1	1000	6	2
2			
3	1200	3	3
4	500	3	1
5	12	5	1
6	27	6	3

	key
0	
1	
2	
3	
4	
5	
6	

CSE 5311 Name _____

Test 2 - Closed Book

Summer 2014

Multiple Choice. Write your answer to the LEFT of each problem. 3 points each

1. What is the nature of the signature function for the Karp-Rabin method?
   A. it is similar to the KMP failure links
   B. the remainder by discarding the overflow for a polynomial
   C. a polynomial of arbitrary precision implemented using a bignum package
   D. similar to a double hash function for string keys
2. When performing selection in worst-cast linear time for n numbers, roughly how many column medians are computed in the first round?
   A. $W\left(\frac{n}{5}\right)$ B. m, the median-of-medians C. $0.7n$ D. $\frac{n}{5}$
3. While constructing a suffix array for a sequence with n symbols by using radix sorts, sequence symbols are used in which of the radix sorts?
   A. All of them B. None of them C. The last one D. The first one
4. Which of the following is a deficiency of the maximum capacity path technique?
   A. Augmenting paths will be discovered in descending incremental flow increase order.
   B. Flow decomposition must be applied.
   C. An augmenting path is blocked if it introduces a cycle of flow.
   D. The maximum number of potential augmenting paths depends on the achievable flow, in addition to the number of vertices and edges.
5. In a maximum flow problem, the number of augmenting paths in a flow decomposition is bounded by:
   A. V B. $O(VE)$ C. f D. E
6. Which of the following is not true regarding Strassen's algorithm?
   A. It is not possible to have an asymptotically faster algorithm.
   B. It requires more space than the everyday method.

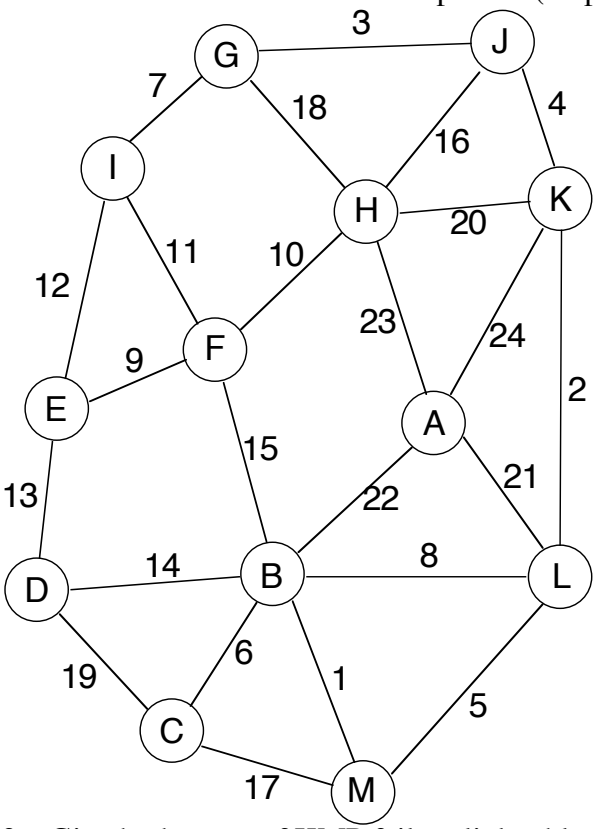
C. It uses $\Theta\left(n^{\lg 7}\right)$ scalar additions when multiplying two n x n matrices.

D. It uses $\Theta\left(n^{\lg 7}\right)$ scalar multiplications when multiplying two n x n matrices.

7. The four russians' concept is to:
   A. Implement longest common subsequences using linear space
   B. Pack bits into an efficient storage unit
   C. Trade-off between enumerating situations and referencing these situations
   D. Trade-off between scalar additions and multiplications
8. Which is not true regarding the first-fit decreasing method for bin packing?
   A. Each object is placed by going left-to-right until it fits in a bin or a new bin is allocated.
   B. Objects placed in bins beyond the optimal number have sizes no larger than ½.
   C. The number of objects placed in bins beyond the optimal number is arbitrary.
   D. The number of objects with sizes larger than ½ is a lower bound on the optimal number of bins.
9. Which of the following is NOT required when showing that problem B is NP-complete by a reduction from problem A?
   A. The reduction has an inverse that takes each instance of problem B to an instance of problem A.
   B. The reduction takes polynomial time.
   C. The reduction must be consistent for the decision results for each instance of problem A and and the corresponding instance of problem B.
   D. Problem A is NP-complete.
10. The algorithm for finding a maximum capacity path for network flows is most similar to which algorithm?
    A. Breadth-first search
    B. Decomposition of a flow into *E* augmenting paths
    C. Dijkstra
    D. Floyd-Warshall
11. Which minimum spanning tree algorithm is the slowest?
    A. Boruvka B. Kruskal C. Prim D. Warshall
12. The reduction from 3-sat to 3-colorability will give a graph requiring four colors when:
    A. Removing one clause from the 3-sat instance will leave a satisifiable set of clauses.
    B. The 3-sat instance is unsatisfiable.
    C. The 3-sat instance is satisfiable.
    D. The 3-sat instance is a tautology.
13. Which of the following is not required before relabeling ("lifting") a vertex to a new height?
    A. Both breadth-first searches have been done.
    B. The vertex is overflowing.
    C. Any eligible edges for the present height have been saturated.
    D. The vertex is not the source or sink.
14. Which of the following does not have a polynomial-time approximation algorithm?
    A. Bin packing
    B. Edge coloring
    C. Traveling salesperson with triangle inequality
    D. Vertex coloring
15. How many times will `-1` occur in the style 2 fail link table for the pattern `acaababc`?
    A. 1 B. 2 C. 3 D. 4
16. Suppose you have a set of points in Euclidean 2-d space. Give an algorithm for finding a ρ-approximation for the minimum traveling salesperson path. Be sure to give the value of ρ. (5 points)

Test 2 - Open Book
Summer 2014

1. Demonstrate Boruvka's algorithm on the following graph. Be sure to indicate the edges that are inserted into the MST in each phase. (10 points)

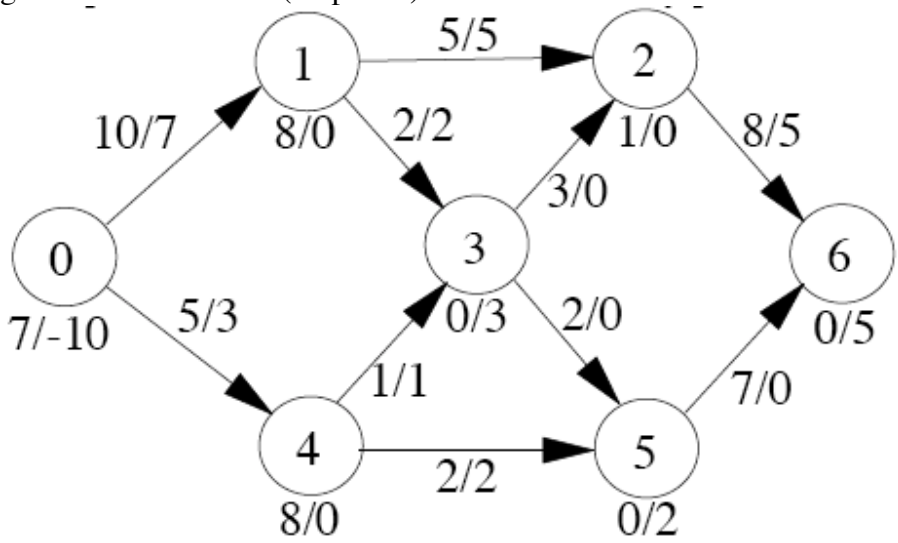


2. Give both types of KMP failure link tables and the table produced by the Z algorithm for the following string: (15 points)

```
 0  a
 1  c
 2  a
 3  b
 4  a
 5  c
 6  a
 7  c
 8  a
 9  b
10  a
11  c
12  a
13  b
14  a
15  c
16  a
17  c
18  a
19  b
20  a
21  c
22  a
```

```
23   c
24   a
```

3. List the remaining operations to complete this instance of network flows by push-relabel. In addition, give a minimum cut. (15 points)



4. Use dynamic programming, either with a table or lists, to determine a subset that sums to 20. DO NOT SOLVE BY INSPECTION! (10 points)

```
2   3   5   7   11   13   17
```