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

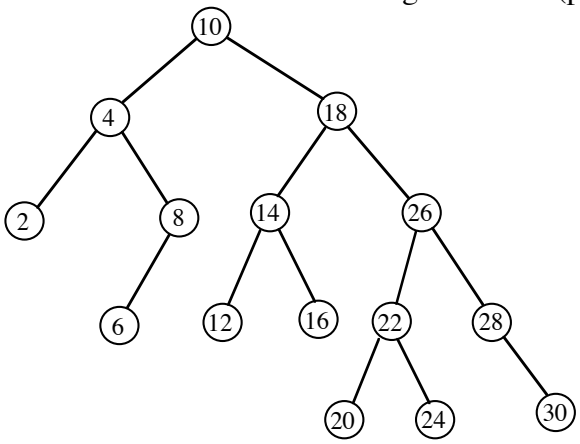
b. $T(n) = T(0.7n) + 1$

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

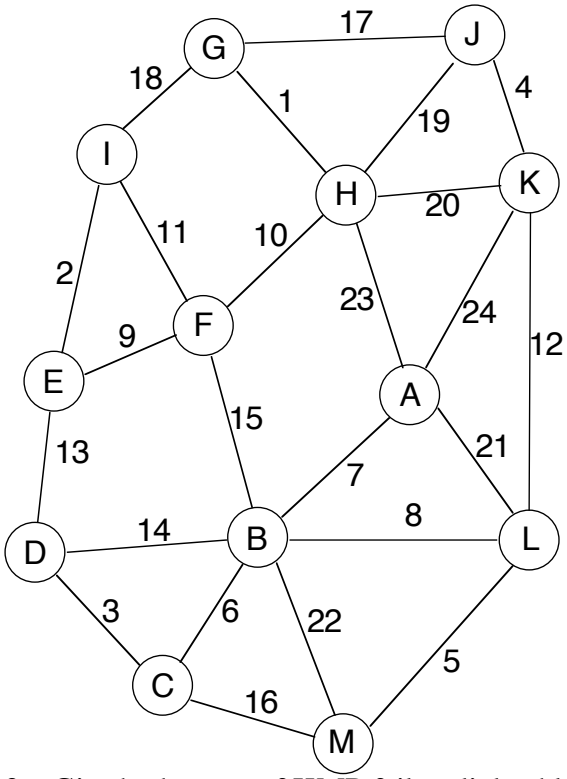
1. When performing selection in worst-cast linear time for n numbers, roughly how many column medians are computed in the first round?
   A. $\frac{n}{5}$
   B. m, the median-of-medians
   C. $.7n$
   D. $W\left(\frac{n}{5}\right)$
2. What is the nature of the signature function for the Karp-Rabin method?
   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 $P \neq NP$)
    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 # _____

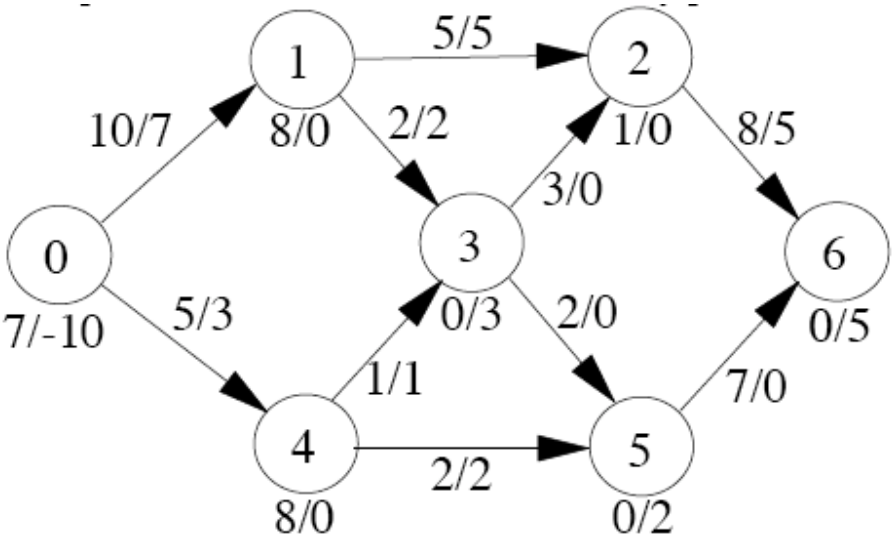
1. Demonstrate Boruvka's algorithm on the following graph. Be sure to indicate the edges that are inserted into the MST in each phase. (15 points)



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