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. $\mathrm{O}(VE)$
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 $P \neq NP$)
    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.

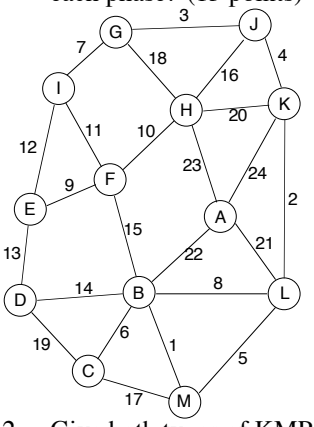
15. The most general approximation result that can be achieved for an NP-hard problem is:
    A. Approximation Algorithm
    B. Approximation Scheme
    C. Polynomial-time Approximation Scheme
    D. Fully Polynomial-time Approximation Scheme
16. What is the Steiner tree problem? (5 points)

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)