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. $height_{right} - height_{left}$
   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 $O(n^2)$ 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. $T(n) = 4T\left(\frac{n}{2}\right) + 1$

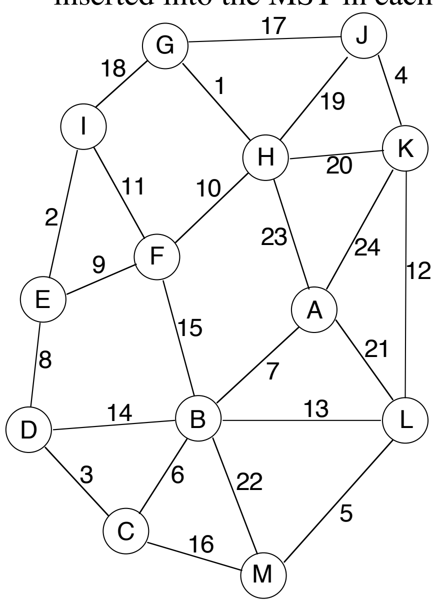
3. Construct the final optimal binary search tree (using Knuth's root trick) and give its cost. SHOW YOUR WORK. (10 points)

```
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)