CS109A Final Examination

Friday, March. 20, 12:25 – 3:15 pm

Directions

The exam is *open book*; any written materials may be used. Answer all 15 questions on the exam paper itself. The total number of points is 300. You may make additional comments at the bottoms of pages or at the end if you feel the need to clarify or justify your answers. Do not forget to **sign the pledge** below.

I acknowledge and accept the honor code.

____

Print your name here: ____

1		5		9		13	
2		6		10		14	
3		7		11		15	
4		8		12			

Problem 1: (20 points) In this question we shall ask about whether an element x is in a set S, by giving conditions on x, S, and other sets. For each of the parts below, answer A, B, or C, with the following meanings.

A: x is surely in the set S.
B: x cannot possibly be in the set S.
C: It cannot be determined from what is given whether x is or is not in S.

Note: The problem is scored 4 points for a correct answer, 0 for no answer, and -2 for an incorrect answer.

a) x is in $S \cup T$ and x is NOT in T. ________

b) x is NOT in $S \cap T$ and x is NOT in T. ________

c) x is NOT in $T - S$ and x is in T. ________

d) x is in $R - (R - T)$ and $S = R \cap T$. ________

e) x is NOT in $R - (R - T)$ and $S = R \cap T$. ________

Problem 2: (20 points) The *symmetric difference* of sets S and T is the set of elements that are in exactly one of S and T. Two expressions for the symmetric difference are $(S - T) \cup (T - S)$ and $(S \cup T) - (S \cap T)$. Prove these two expressions equivalent by showing that an element x in one is also in the other. In writing your proof, you may use as shorthand the following reasons:

(1) Given.
(2) Definition of union.
(3) Definition of intersection.
(4) Definition of set difference.

Suggestion: If you need to consider two cases, divide the page left-to-right and show your reasoning for each case "in parallel."

a) Show that if x is in $(S - T) \cup (T - S)$ then x is in $(S \cup T) - (S \cap T)$.

b) Show that if x is in $(S \cup T) - (S \cap T)$ then x is in $(S - T) \cup (T - S)$.

Problem 3: (15 points) Suppose we use the hash function on integers: $h(i) = i^2 \texttt{ mod } 10$ (i.e., the remainder when i is squared and then divided by 10) and use a hash table with 10 buckets, numbered 0 to 9.

a) If we hash the twenty integers $1, 2, \ldots, 20$ into the table, what is the distribution of integers in the buckets? For each of the buckets, give the number of elements in that bucket: 0: ______ 1: ______ 2: ______ 3: ______ 4: ______ 5: ______ 6: ______ 7: ______ 8: ______ 9: ______

b) If we were to hash the first thousand integers $1, 2, \ldots, 999$ into this hash table, would you expect a more even distribution? ______ Why or why not?

____

____

____

Problem 4: (15 points) The following is an ML declaration suitable for the datatype `ttree`, a "ternary" tree in which each node has "slots" for three children: a left, middle, and right child. As with binary trees, any one or more of these slots may be "null," while the other slots indicate nodes (or in the ML style, indicate the trees rooted at those nodes).

```
datatype 'label ttree =
    null |
    node of
        ('label * ('label ttree) * ('label ttree) * ('label ttree));
```

The following describes a certain `ttree`.

i) The root has label `"a"`, a left child described in (*ii*) below, no middle child, and a right child described in (*iii*) below.

ii) The left child of the root is a leaf with label `"b"`.

iii) The right child of the root has label `"c"`, a left child that is a leaf labeled `"d"`, a middle child that is a leaf labeled `"e"`, and no right child.

Answer the following about this `ttree`.

a) Draw a picture of the tree in the space below. A left child should extend to the left, a middle child be directly below, and a right child extend to the right. Represent each node by its label.

b) Write an ML expression that denotes this `ttree`.

Problem 5: (20 points) Write an ML program `pre` that takes as argument a `ttree` (as defined in Problem 4) and produces a list of this tree's labels, in preorder.

Problem 6: (20 points) Prove by structural induction that in a *nonempty* ternary tree the number of NIL pointers (nulls in the ML datatype) is one more than twice the number of nodes. *Warning*: You will receive no credit if your proof is not a structural induction, e.g., you will receive no credit if it is a complete induction on the height or the number of nodes. To organize your proof, answer the following questions.

a) What is the statement $S(T)$ to be proved?

b) What is a suitable basis case? ____________

c) Prove the basis case.

d) Prove the inductive step. While there are some options regarding the level of detail appearing in the proof, a proof will not be considered correct unless it indicates clearly where and how the inductive hypothesis is used.

Problem 7: (20 points) We would like to find the longest common subsequences of the strings `babcc` and `abacbca`.

a) In the space below, create a table giving the length of the LCS between each pair of prefixes, one from each of the strings. Please use columns to correspond to prefixes of `abacbca` and rows to represent prefixes of `babcc`.

b) What is the length of the LCS of these strings? ______

c) What are all the LCS's of this length? Just give the LCS's; you need not indicate the positions of the two original strings from which they come.

Problem 8: (20 points) Suppose strings of lengths 1 through 10 are equally likely, i.e., 10% probability each. Also, suppose that we wish to store strings in linked lists of cells consisting of room for k bytes and a pointer to the next cell. Assume that the number of bytes taken by a cell with room for k characters is $k+2$ bytes (i.e., pointers take 2 bytes, and there is no requirement that fields begin at a position that is a multiple of 4).

a) Describe briefly how you would calculate the average number of bytes needed to store strings if $k = 5$.

---

b) Give the average number of bytes required for the following values of k. $k = 1$: ________ $k = 2$: ________ $k = 3$: ________ $k = 4$: ________

Problem 9: (20 points) Let S be the set of pairs (c, j), where c is one of the colors red, green, or blue (abbreviate r, g, and b, respectively), and j is a nonnegative integer. We wish to show that S is a countable set.

a) Define a one-to-one correspondence f from the nonnegative integers to S.

---

b) Explain briefly why f is a one-to-one correspondence.

Problem 10: (30 points) Let the domain D be the integers from 1 to 10, and let R be the relation on D defined by aRb if and only if a and b have a common divisor greater than 1. For example, $4R4$ and $4R6$, but the following are false: $1R5$, $2R5$, and $3R5$. Tell whether R has each of the following properties and either give a brief proof if so or a counterexample if not.

a) R is reflexive? ________

b) R is symmetric? ________

c) R is antisymmetric? ________

d) R is transitive? ________

e) R is a partial order? ________

f) R is an equivalence relation? ________

Problem 11: (20 points) Prove by induction on n that for all $n \geq 1$, the sum

$$\frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \frac{1}{3 \times 4} + \cdots + \frac{1}{n \times (n+1)} \qquad \left(\text{that is, } \sum_{i=1}^{n} \frac{1}{i(i+1)}\right)$$

equals $\frac{n}{n+1}$. To organize your proof, answer the following questions.

a) What is the statement $S(n)$ to be proved?

____

b) What is a suitable basis case? ________________

c) Prove the basis case.

____

____

____

d) When we prove the inductive step, we assume $S(n)$. What must we prove, using this assumption?

____

e) Prove the inductive step. While there are some options regarding the level of detail appearing in the proof, a proof will not be considered correct unless it indicates clearly where and how the inductive hypothesis is used.

____

____

____

____

____

____

Problem 12: (20 points) We wish to represent sets S and T that consist of $n/2$ integers chosen from the integers 1 through n, and we shall consider three data structures to do so:

A: A linked list (unsorted, no duplicates).
B: A hash table with n buckets.
C: A Boolean array indexed by integers `1..n`, used as a characteristic vector.

For each of the operations suggested below, give the big-oh running time of the operation, assuming the most efficient possible implementation. For the hash table, you should give the average running time, and in the other two cases give the worst-case running time.

a) Insertion of an element x into a set S.

A: ____________________ B: ____________________ C: ____________________

b) Computing the union of the two sets S and T.

A: ____________________ B: ____________________ C: ____________________

Problem 13: (20 points) Consider the following fragment of Pascal code. Note that `ODD` is a built-in Pascal function, not a procedure call, and can be assumed to take $O(1)$ time.

```
     begin
(1)      sum := 0;
(2)      for i := 1 to n do
(3)          if ODD(i) then
(4)              for j := 1 to i do
(5)                  sum := sum + A[j]
(6)          else
(7)              for j := i to n do
(8)                  sum := sum - A[j]
     end;
```

In the space below, draw a tree showing how the lines are grouped into progressively larger statements. For each statement, indicate the lines it covers and give a smooth and tight big-oh bound on its running time (as a function of the value of n).

Problem 14: (20 points) For each of the following, tell whether the stated big-oh relationship holds. If not, you can just say "false." If the relationship does hold, then you must give witnesses n_0 and c that are minimal in the sense that no smaller n_0 for the same c works, and no smaller c works for the same n_0. Note that there may be more than one minimal pair of witnesses for a given problem, so there is not always a unique correct answer.

a) $7n + 50$ is $O(n^2)$.

b) 10^n is $O(n^n)$.

c) Let $f(n) = n^3$ for $n = 0, 1, \ldots, 1000$, and $f(n) = n^2$ for $n > 1000$. $f(n)$ is $O(n^2)$.

d) Let $f(n)$ be as in (c). Let $g(n) = n^3$ for even n and $g(n) = 10n$ for odd n. $f(n)$ is $O\big(g(n)\big)$.

Problem 15: (20 points) The following is a binary search tree formed by inserting into an empty tree the integers 1, 2, 3, and 4, in some order.

```
    3
   / \
  1   4
   \
    2
```

a) What are all the orders in which the four integers could have been inserted to form this particular binary search tree?

b) There are many other binary search trees that can be formed from the integers 1 through 4 by insertion in some other order. In the space on the next page, show all the possible BST's that have 2 at the root.

Answer for part (b):

c) In the space below, show all the BST's with integers 1 through 4 that have 1 at the root.

d) How many BST's with the integers 1 through 4 have 4 at the root? ________