CSC 263 Homework 2

Due June 22, 2004

1. (20 points) Consider a binary tree T. Let $|T|$ be the number of nodes in T. Let x be a node in T, let L_x be the left subtree of x and let R_x be the right subtree of x. We say that x has the "approximately balanced property," $ABP(x)$, if

$$|R_x| \leq 2|L_x| \qquad \text{and} \qquad |L_x| \leq 2|R_x|.$$

   (a) What is the maximum height of a binary tree T on n nodes where $ABP(root)$ holds?

   (b) We call T an ABP-tree if $ABP(x)$ holds for every node x in T. Prove that if T is an ABP-tree, then the height of T is $O(\log n)$. More precisely, show that

$$height(T) \leq \log_2 n / \log_2 \frac{3}{2}.$$

2. (15 points) In class, we analyzed the worst-case running times for the dictionary operations on BST trees. Now we will consider one possible notion of average case. Assume we have an empty tree T and we do the following operations

$$INSERT(T, 1), INSERT(T, 2), \ldots, INSERT(T, n)$$

   in some order. Assume all orders are equally likely. Since all operations on BST trees depend on the height of the tree, we are interested in $H_{avg}(n)$—the expected height of T after the n inserts.

   (a) Define the probability space that we are dealing with.

   (b) Let A_i be the event that we insert i first. Write an expression for $t(A_i)$, the expected height of the final tree given that we insert i first, in terms of H_{avg}.

   (c) Write a recurrence relation for $H_{avg}(n)$.

   (d) (Extra Credit) Solve the recurrence relation from part (c). Hint: it should be $\Theta(\log n)$.

3. (15 points) We want to augment Red-Black Trees to support the following query, `Average(x)`, which returns the average key-value in the subtree routed at node x (including x itself). The query should work in worst-case time $\Theta(1)$.

   (a) What extra information needs to be stored at each node?

   (b) Describe how to modify `INSERT` to maintain this information. What is its worst-case running time now?

(c) Describe how to modify `DELETE` to maintain this information. What is its worst-case running time now?

4. (25 points) **Programming Question:** Coming soon!