CSCI 305, Homework # 6

Caleb Ouellette

Due date: Midnight, May 14

1. Alternative quicksort analysis (problem 7-3 in the text).

   An alternative analysis of the running time of randomized quicksort focuses on the expected running time of each individial recursive call to RANDOMIZED-QUICKSORT, rather than on the number of comparisons performed.

   (a) Argue that, given an array of size n, the probability that any particular element is chosen as the pivot is $1/n$. Use this to define indicator random variables

   $$X_i = I\{i\text{th smallest element is chosen as the pivot}\}$$

   What is $E[X_i]$? $E[X_i]$ is the expected value of an element being chosen. The probability of choosing any element is $1/n$ so, $E[x_i]$ is $1/n$.

   (b) Let $T(n)$ be a random variable denoting the running time of quicksort on an array of size n. Argue that

   $$E[T(n)] = E\left[\sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n))\right] \tag{1}$$

   The equation come from the code for quicksort. The $\Theta(n)$ is the partition. The $T(q-1)$ is the recursive call for the left side of matrix being quicksorted. The $T(n-q)$ is the rest of the matrix. The indicator variable X_q is the pivot, and has a $\frac{1}{n}$ chance of occuring. Each pivot is equally likely and this equation weights each accordingly.

(c) Show that we can rewrite equation 1 as

$$E[T(n)] = \frac{2}{n}\sum_{q=2}^{n-1} E[T(q)] + \Theta(n) \tag{2}$$

$$E[T(n)] = E\left[\sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n))\right]$$

$$E[T(n)] = \sum_{q=1}^{n} E\left[X_q(T(q-1) + T(n-q) + \Theta(n))\right]$$

$$E[T(n)] = \sum_{q=1}^{n} \frac{1}{n} E\left[T(q-1) + T(n-q) + \Theta(n))\right]$$

$$E[T(n)] = \frac{1}{n}\sum_{q=1}^{n} E\left[T(q-1) + T(n-q) + \Theta(n))\right]$$

$$E[T(n)] = \frac{1}{n}\sum_{q=2}^{n-1} E\left[T(q+1-1) + T(n-1-(q+1)) + \Theta(n))\right]$$

$$E[T(n)] = \frac{1}{n}\sum_{q=2}^{n-1} E\left[T(q) + T(n-q) + \Theta(n))\right]$$

$$E[T(n)] = \frac{1}{n}\sum_{q=2}^{n-1} E\left[2T(q) + \Theta(n))\right]$$

$$E[T(n)] = \frac{2}{n}\sum_{q=2}^{n-1} E\left[T(q)\right] + \Theta(n))$$

(d) Show that

$$\sum_{k=2}^{n-1} k \lg k \leq \frac{1}{2}n^2 \lg n - \frac{1}{8}n^2 \tag{3}$$

(Hint: Split the summation into two parts, one for $k = 2, 3, ..., \lceil n/2 \rceil - 1$ and one for $k = \lceil n/2 \rceil, ..., n-1$

(e) Using the bound from equation 3, show that the recurrence in equation 2 has the solution $E[T(n)] = \Theta(n \lg n)$. (Hint: Show, by substitution, that $E[T(n) \leq an \lg n$ for sufficiently large n and for some positive constant a.)

$$
\begin{aligned}
E[T(n)] &= \frac{2}{n}\sum_{q=2}^{n-1} E\left[T(q)\right] + \Theta(n)) \\
E[T(n)] &\leq \frac{2}{n}\sum_{q=2}^{n-1} an\lg n + \Theta(n)) \\
E[T(n)] &\leq \frac{2a}{n}(\frac{1}{2}n^2\lg n - \frac{1}{8}n^2 + n\Theta(n)) \\
E[T(n)] &\leq 2a(\frac{1}{2}n\lg n - \frac{1}{8}n + \Theta(n)) \\
E[T(n)] &\leq an\lg n - \frac{a}{4n} + 2an \\
E[T(n)] &\leq an\lg n + \frac{an7}{4} \\
an\lg n &\leq an\lg n + \frac{an7}{4}
\end{aligned}
$$

let a be $a \geq 0$ and $n \geq 1$.