Homework 5

Dushan Terzikj

12 Mar, 2018

1 Problem 1

a) Please check 'quicksort.cpp' and follow the comments.

b) **Worst case:**

Array sorted (descending or ascending):
Picking the first two elements would mean that you will pick the two smallest (or largest) elements in the whole array. The partitioning part is $\Theta(n)$, therefore we get:

$$T(n) = T(n-2) + T(0) + T(0) + \Theta(n)$$

Obviously, this will be repeated $\frac{n}{2}$ times, each of which is $\Theta(n) \Rightarrow T(n) = \Theta(n^2)$ In this case, two of the three partitions would be of length 0.

Best case:
The array is divided into three parts which are close to have the same length.

$$T(n) = T(n-a-b-2) + T(a) + T(b) + \Theta(n)$$

where a is the length of the middle partition and b is the length of the right partition. Ideally we would have $a = b = \frac{n}{3}$

$$\Rightarrow T(a) = T(\frac{n}{3}) \tag{1}$$
$$T(b) = T(\frac{n}{3}) \tag{2}$$
$$T(n-a-b-2) = T(\frac{n}{3} - 2) \tag{3}$$

The -2 term in (3) can be omitted for large cases. Finally we have something like this:

$$T(n) = 3T(\frac{n}{3}) + \Theta(n)$$

Using <u>master method's case 2</u> we have: $T(n) = \Theta(nlgn)$

c) Please check 'randomized_quicksort.cpp' and follow the comments.

2 Problem 2

a) Proof by induction:

$$\sum_{k=2}^{n-1} klgk \leq \frac{1}{2}n^2lgn - \frac{1}{8}n^2 \tag{4}$$

Base case, $n = 3$

$$\sum_{k=2}^{2} klgk \leq \frac{1}{2}3^2lg3 - \frac{1}{8}3^2$$

$$\sum_{k=2}^{2} klgk = 2lg2 = 2$$

$$\frac{1}{2}3^2lg3 - \frac{1}{8}3^2 = \frac{9}{2}lg3 - \frac{9}{8} = 6.01$$

$$\Rightarrow 2 \leq 6.01$$

Base case holds and it is proven.

Induction step:
Assume that the equation (4) holds $\forall m \leq n$. Let us prove equation (4) for n+1:

$$\begin{aligned}\sum_{k=2}^{n} klgk &\leq \frac{1}{2}(n+1)^2lg(n+1) - \frac{1}{8}(n+1)^2 \\ &= \frac{1}{2}(n^2+2n+1)lg(n+1) - \frac{1}{8}(n^2+2n+1) \\ &= \frac{1}{2}n^2lg(n+1) + nlg(n+1) + \frac{1}{2}lg(n+1) - \frac{1}{8}n^2 - \frac{1}{4}n - \frac{1}{8} \\ &= \left(\frac{1}{2}n^2lg(n+1) - \frac{1}{8}n^2\right) + nlg(n+1) + \frac{1}{2}lg(n+1) - \frac{1}{4}n - \frac{1}{8}\end{aligned}$$

The sum in the above equation can also be written as:

$$\sum_{k=2}^{n} klgk = \sum_{k=2}^{n-1} klgk + nlgn \tag{5}$$

$$\sum_{k=2}^{n-1} klgk + nlgn \leq \left(\frac{1}{2}n^2lg(n+1) - \frac{1}{8}n^2\right) + nlg(n+1) + \frac{1}{2}lg(n+1) - \frac{1}{4}n - \frac{1}{8}$$

Since $lgn \approx lg(n+1)$, using equation (4) we already know $\sum_{k=2}^{n-1} klgk \leq \frac{1}{2}n^2lgn - \frac{1}{8}n^2$ we can cancel these two terms in the upper equation.

$$\begin{aligned}
&nlgn \leq nlgn - \frac{1}{4}n - \frac{1}{8} + \frac{1}{2}lgn \\
\Rightarrow &\frac{1}{4}n + \frac{1}{8} \leq \frac{1}{2}lgn \\
\Rightarrow &\frac{1}{2}n + \frac{1}{4} \leq lgn \\
\Rightarrow &2^{\frac{1}{2}n+\frac{1}{4}} \geq n
\end{aligned}$$

$2^{\frac{1}{2}n+\frac{1}{4}} \geq n$ is always true, therefore the induction step is proven. QED.

b) Proof by induction:

$$E[T(n)] \geq cnlgn \tag{6}$$

According to the slides:

$$E[T(n)] = \frac{2}{n}\sum_{k=2}^{n-1} E[T(k)] + \Theta(n) \tag{7}$$

Base case: n=3,

$$\begin{aligned}
E[T(3)] &= \frac{2}{3}\sum_{k=2}^{2} E[T(k)] + \Theta(3) = \frac{2}{3}\Theta(1) \\
&\frac{2}{3}\Theta(1) \geq 3clg3 \\
\Rightarrow &\frac{2}{9}\Theta(1) \geq clg3 \\
\Rightarrow &\ 0.14\Theta(1) \geq c
\end{aligned}$$

For $0 < c \leq 0.14$ This equation holds. Therefore, the base case is proven.

Induction step: Assume that equation (6) holds $\forall m \leq n$. Let us show that it also holds for n+1.

$$\begin{aligned}
E[T(n+1)] &= \frac{2}{n}\sum_{k=2}^{n} E[T(k)] + \Theta(n+1) \\
&= \frac{2}{n}\sum_{k=2}^{n-1} E[T(k)] + \frac{2}{n}E[T(n)] + \Theta(n)
\end{aligned}$$

$$\Rightarrow \frac{2}{n}\sum_{k=2}^{n-1} E[T(k)] + \frac{2}{n}E[T(n)] + \Theta(n) \geq c(n+1)lg(n+1)$$
$$= cnlg(n+1) + clgn$$

Multiply everything by $\frac{n}{2}$

$$\sum_{k=2}^{n-1} E[T(k)] + E[T(n)] + \frac{n}{2}\Theta(n) \geq \frac{1}{2}cn^2lg(n+1) + \frac{1}{2}cnlgn$$

Since $\frac{1}{2}c$ can also be in the interval $(0, 0.14]$, using equation (6) and our initial assumption, we can eliminate the elements from equation (6) from both sides.

$$\begin{aligned}
\frac{n}{2}E[T(n)] - \frac{n}{2}\Theta(n) + \frac{1}{2}\Theta(n^2) &\geq \frac{1}{2}cn^2lgn \\
nE[T(n)] - \Theta(n^2) + \Theta(n^2) &\geq cn^2lgn \\
nE[T(n)] &\geq cn^2lgn \\
E[T(n)] &\geq cnlgn
\end{aligned}$$

We got the same form as equation (6). The induction step has been proven. QED.

3 Problem 3

We know that $lgn! = \sum_{k=1}^{n} lgk$. From this we can derive the upper bound of $\Theta(nlgn)$.

$$\begin{aligned}
lg(n!) &= \sum_{k=1}^{n} lg(k) \\
&= lg(1) + lg(2) + ... + lg(n)
\end{aligned}$$

$$\begin{aligned}
\sum_{k=1}^{n} lg(n) &= lg(n) + lg(n) + ... + lg(n) \\
&= nlg(n)
\end{aligned}$$

Obviously, $lg(1) + lg(2) + ... + lg(n) \leq nlg(n) \Rightarrow O(nlgn)$
And we can get the lower bound by doing a similar thing, but throwing away the first half of the sum:

$$lg(1) + lg(2) + ... + lg(\frac{n}{2}) + ... + lg(n) \geq lg(\frac{n}{2}) + ... + lg(n)$$

$$lg(\frac{n}{2}) + ... + lg(n) \geq \sum_{k=1}^{\frac{n}{2}} lg(\frac{n}{2})$$

where

$$\sum_{k=1}^{\frac{n}{2}} lg(\frac{n}{2}) = \frac{n}{2} lg(\frac{n}{2})$$

Since $\Theta = O \cap \Omega$, we have proved that $lg(n!) = \Theta(nlgn)$