COMP 550

Assignment 3

Siyang Jing

a) The pivot is chosen by line 1 of RANDOMIZED-PARTITION, through randomly selecting a number from p to r. Therefore, each element is chosen with equal probability, i.e. $1/n$.

$$\mathbb{E}[X_i] = \Pr(i\text{th smallest element is chosen as the pivot}) = 1/n$$

b) The events $\{i$th smallest element is chosen as the pivot$\}$ for different i's are disjoint, i.e. if pth smallest element has been chosen as the pivot, then qth smallest element cannot be the pivot, and the probability of their union is 1, i.e. one of them has to happen. When $X_q = 1$, i.e. the qth smallest element has been chosen as the pivot, we have

$$T(n) = T(q-1) + T(n-q) + \Theta(n).$$

Therefore, we have

$$T(n) = \sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n)),$$

and therefore

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

c) $\mathbb{E}[T(n)]$

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

$= \sum_{q=1}^{n} \mathbb{E}[X_q(T(q-1) + T(n-q) + \Theta(n))]$, by linearity of $\mathbb{E}$ operator

we notice that X_q only depends on choice of this pivot, and $(T(q-1) + T(n-q) + \Theta(n))$ does not depend on choice of this pivot, and therefore, they are independent, and therefore we have:

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

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

$= \frac{1}{n}\sum_{q=1}^{n} \mathbb{E}[T(q-1)] + \mathbb{E}[T(n-q)] + \Theta(n)$, by linearity of $\mathbb{E}$

we notice that the term corresponding to $q = k$ and $q = n+1-k$ are identical, just in reverse order, and that if n is odd, then for $q = \frac{n+1}{2}$, $\mathbb{E}[T(q-1)]$ and $\mathbb{E}[T(n-q)]$ are identical, therefore, by a change in the order of the summation, we have

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

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

since $T(1) = T(0) = \Theta(1)$ are ignorable, by a change of variable $q \to q-1$, we have

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

d) $\sum_{k=2}^{n-1} k \lg k$

$= \sum_{k=1}^{n-1} k \lg k$

$=\sum_{k=1}^{\lceil\frac{n}{2}\rceil-1} k\lg k+\sum_{k=\lceil\frac{n}{2}\rceil}^{n-1} k\lg k$

$\leqslant\sum_{k=1}^{\lceil\frac{n}{2}\rceil-1} k\lg\frac{n}{2}+\sum_{k=\lceil\frac{n}{2}\rceil}^{n-1} k\lg n$

$=\sum_{k=1}^{\lceil\frac{n}{2}\rceil-1} k(-1+\lg n)+\sum_{k=\lceil\frac{n}{2}\rceil}^{n-1} k\lg n$

$=\lg n\sum_{k=1}^{n-1} k-\sum_{k=1}^{\lceil\frac{n}{2}\rceil-1} k$

$\leqslant\frac{n^2-n}{2}\lg n-\frac{\frac{n}{2}(\frac{n}{2}-1)}{2}$

$\leqslant\frac{1}{2}n^2\lg n-\frac{1}{8}n^2+\frac{n}{2}\big(\frac{1}{2}-\lg n\big)$

$\leqslant\frac{1}{2}n^2\lg n-\frac{1}{8}n^2$, since $n\geqslant 2\rightarrow\lg n\geqslant 1\rightarrow\frac{1}{2}-\lg n<0\rightarrow\frac{n}{2}\big(\frac{1}{2}-\lg n\big)<0$

e) We first notice that

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

We assume $\Theta(n)$ in the equation above means $c_1 n\leqslant\Theta(n)\leqslant c_2 n$.

We first prove $\mathbb{E}[T(n)]\in\mathrm{O}(n\lg n)$, by finding a, s.t. $\forall n\geqslant 2,\mathbb{E}[T(n)]\leqslant a\,n\lg n$.

We first cover the base cases, i.e. select a large enough s.t. $\mathbb{E}[T(2)]\leqslant a\,2\lg 2$.

We then assume $\mathbb{E}[T(k)]\leqslant a\,k\lg k$, then, for all $2\leqslant k<n$.

$$\frac{2}{n}\sum_{q=2}^{n-1}\mathbb{E}[T(q)]\leqslant\frac{2}{n}\sum_{q=2}^{n-1}a\,q\lg q\leqslant\frac{2a}{n}\left(\frac{1}{2}n^2\lg n-\frac{1}{8}n^2\right)\leqslant a\,n\lg n-\frac{a}{4}n$$

$$\mathbb{E}[T(n)]=\Theta(n)+\frac{2}{n}\sum_{q=2}^{n-1}\mathbb{E}[T(q)]\leqslant\Theta(n)+a\,n\lg n-\frac{a}{4}n\leqslant a\,n\lg n+c_2 n-\frac{a}{4}n\quad(*)$$

We let $a>4c_2$, then we have

$$a>4c_2\rightarrow c_2 n-\frac{a}{4}n<0\rightarrow(*)<a\,n\lg n$$

In conclusion, we have successfully found $a=\max\left(\frac{\mathbb{E}[T(2)]}{2\lg 2},4c_2\right)+1$.

We then prove $\mathbb{E}[T(n)]\in\Omega(n\lg n)$, by finding b, s.t. $\forall n\geqslant 2,\mathbb{E}[T(n)]\geqslant b\,n\lg n$.

We first find a lower bound for $\sum_{k=2}^{n-1} k\lg k$.

Since $k\lg k$ is monotonically increasing in k, we have

$$\sum_{k=2}^{n-1} k\lg k\geqslant\int_1^{n-1} x\lg x\,dx=\frac{1}{2}x^2\lg x-\frac{1}{4\log 2}x^2\Big|_{x=1}^{x=n-1}>\frac{1}{2}(n-1)^2\lg(n-1)-\frac{1}{4\log 2}(n-1)^2\quad(**)$$

We notice that $\lg(n)-\lg(n-1)=\lg\left(\frac{n}{n-1}\right)=\lg\left(1+\frac{1}{n-1}\right)\leqslant 1$, so we have:

$(**)\geqslant\frac{1}{2}(n-1)^2(\lg(n)-1)-\frac{(n-1)^2}{4\log 2}>\frac{1}{2}n^2\lg n-n\lg n-\frac{1}{2}(n-1)^2-\frac{(n-1)^2}{2}>\frac{1}{2}n^2\lg n-2n^2$

We proceed our proof of $\mathbb{E}[T(n)] \geqslant b\,n\lg n$, by induction on n.

We first cover the base cases, i.e. select b large enough s.t. $\mathbb{E}[T(2)] \geqslant b\,2\lg 2$.

We then assume $\mathbb{E}[T(k)] \geqslant b\,k\lg k$, then, for all $2 \leqslant k < n$.

$$\frac{2}{n}\sum_{q=2}^{n-1} \mathbb{E}[T(q)] \geqslant \frac{2}{n}\sum_{q=2}^{n-1} b\,q\lg q \geqslant \frac{2b}{n}\left(\frac{1}{2}n^2\lg n - 2n^2\right) \geqslant b\,n\lg n - 4bn$$

$$\mathbb{E}[T(n)] = \Theta(n) + \frac{2}{n}\sum_{q=2}^{n-1} \mathbb{E}[T(q)] \geqslant \Theta(n) + b\,n\lg n - 4bn \geqslant b\,n\lg n + (c_1 - 4b)n \quad (***)$$

We let $b < \frac{c_1}{4}$, then we have

$$(c_1 - 4b)n < 0 \rightarrow (***) > b\,n\lg n$$

In conclusion, we have successfully found $a = \min\left(\frac{\mathbb{E}[T(2)]}{2\lg 2}, \frac{c_1}{4}\right)$