

ICCS481: Midterm Exam

Kriangsak Thuiprakhon
kriangsak.thi@student.mahidol.edu

1: UNSTARING

In order for any sequences of N operation (ADDVERTEX, ADDEDGE,UNSTAR) to deliver in linear time, it means that we need to show that it takes amortized constant time per operation.

Lemma 0.1. Any sequence of sequences of N operation (ADDVERTEX, ADDEDGE,UNSTAR) take $O(n)$

Proof. We will prove this lemma using the potential method. First, let us define the potential function:

$$\phi^{(i)} = \sum_u [\deg(u) - 3] \text{ For all } u \text{ where } \deg(u) \geq 3$$

It is obvious to see that the above function will always be positive Now we will use this function to show that the amortized cost per one of the above iterations takes constant time. recall the equation of the amortized cost(\hat{c}) where c denotes the actual cost per iteration.

$$\hat{c} = c + \Delta\phi$$

- For ADDVERTEX

$$\begin{aligned}\hat{c} &= c + \Delta\phi \\ &= 1 + \phi^{(i)} - \phi^{(i-1)} \\ &= 1 + \phi^{(i)} - \phi^{(i)} \\ &= 1 \\ &= O(1)\end{aligned}$$

- For ADDEDGE. There are two cases to consider:

- the case when adding an edge but every node in graph G still has degree less than 3. In this case, it is the same as what is being done in ADDVERTEX Operation since there is no change in the potential function:(i.e., the potential function does not change because there is no node that will have a degree greater than 3).
- The case when adding an edge and it turns that node to be a node of degree greater than 3, (i.e., 4). then,

$$\begin{aligned}\hat{c} &= c + \Delta\phi \\ &= 1 + \phi^{(i)} - \phi^{(i-1)} \\ &= 1 + \phi^{(i)} - (\phi^{(i)} - 1) \\ &= 2 \\ &= O(1)\end{aligned}$$

- For UNSTAR) with a node of degree d.

$$\begin{aligned}\hat{c} &= c + \Delta\phi \\ &= d + \phi^{(i)} - \phi^{(i-1)} \\ &= d + d(3 - 3) - (d - 3) \\ &= 3 \\ &= O(1)\end{aligned}$$

Since every function takes amortized constant time, we conclude that the sequence of N operation takes amortized $O(N)$. \square

2: Parallel Sublinear-Work Approximate Median

(i)

- Show that work bound is $O(n)$

Proof. From Data Structure, we did in the end analyze the running time of some flavor of ε -approximate median. We have proved the work bound:

Algorithm 1: `findSmallestK(k, S)` — find the smallest k numbers in S

```

if  $|S| \leq k$  then
   $\perp$  return  $S$ 
else
  Pick  $p \in S$  uniformly at random //  $p = \text{random.choice}(S)$ 
  Compute  $L = [\mathbf{x}$  for  $\mathbf{x}$  in  $S$  if  $\mathbf{x} \leq p]$  and  $R = [\mathbf{x}$  for  $\mathbf{x}$  in  $S$  if  $\mathbf{x} > p]$ 
  if  $|L| \geq k$  then
     $\perp$  return findSmallestK(k, L)
  else
     $\perp$  return  $L + \text{findSmallestK}(k - |L|, R)$ 

```

Note that the picture is taken from DS lecture note by the same professor teaching this very class. As we can see that the running time for this algorithm depends on the size of the two arrays, L and R . Let's write the recurrence of work as:

$$W(n) = W(\max(|L|, |R|)) + O(n)$$

From this on, we can see that our running time will be probabilistic, that is, we need to find the probability that the length of them will fall in a good range, between $n/4$ to $3n/4$, which amounts to the $\Pr[X_n \leq \frac{3n}{4}]$ of $\frac{1}{2}$, denote $X_n = \max(|L|, |R|)$. Now, if we denote the expected work of the algorithm to be \bar{W} , then:

$$\begin{aligned}
 \bar{W} &\leq \mathbb{P}[X_n = i] \cdot \bar{W}(i) + cn \\
 &\leq \mathbb{P}[X_n \leq 3n/4] \cdot \bar{W}(3n/4) + \mathbb{P}[X_n > 3n/4] \cdot \bar{W}(i) + cn \\
 \Rightarrow \bar{W} &\leq \bar{W}(3n/4) + 2cn
 \end{aligned}$$

Which solves to $O(n)$

\square

- Show that span bound is $O(\log^2 n)$

Proof. In class, we can use FILTER to do the partition in L and R . and we proved that FILTER has span of $O(\log n)$. In the above algorithm, we can see that we can also observe that we can partition up to at most $\log n$ w.h.p. therefore, we have the span of this algorithm to be $O(\log^2 n)$

\square

(ii) we can do so by doing pfor and since randomly picking a number and accessing an array take $O(1)$, then we have work of $O(T)$ and span of $O(1)$

(iii) We want to promise that what the algorithm returns has the rank between the given range. That is:

$$Pr[Y_i \in_R X \mid n(1/2+\varepsilon) \leq Rank_X(Y_i) \leq n(1/2+\varepsilon)] = 1 - 2 \cdot Pr[Rank_X(Y_i) < n(1/2+\varepsilon)] - (*)$$

Using Union Bounds and symmetry, Let Y denote an indicator random variable

$$Y_i = \begin{cases} 1 & \text{if } Rank_X(Y_i) < n(1/2 - \varepsilon) \\ 0 & \text{otherwise} \end{cases}$$

From this, what is the expectation of Y ? To find the expectation of Y , we first need to find the probability that an element drawn is bad, by bad i mean, is not in the range of approximate median. This is simply:

$$Pr[Rank_X(Y_i) < n(1/2 - \varepsilon)] = \frac{n(1/2 - \varepsilon)}{n} = \frac{1}{2} - \varepsilon$$

By definition of expectation,

$$\mathbb{E}[Y] = T(\frac{1}{2} - \varepsilon)$$

From now we can apply Chernoff-Hoffding bounds. We want something like $Pr[Y < T/2]$ and the expectation is $\mu = T(\frac{1}{2} - \varepsilon)$. We need to massage it a little bit.

$$Pr[Y < \mu + T\varepsilon] \leq \exp\{-2\varepsilon^2 T\}$$

This means in * we have that

$$Pr[Y_i \in_R X \mid n(1/2 + \varepsilon) \leq Rank_X(Y_i) \leq n(1/2 + \varepsilon)] \leq 1 - 2\exp\{-2\varepsilon^2 T\}$$

now we set:

$$\delta = 2 \cdot e^{-2\varepsilon^2 T}$$

Therefore,

$$T = \frac{\ln(\delta/2)}{-2\varepsilon^2}$$

3: JL

1. For $0 < t < 1/2$ if $R \in_R -1, 1$, then $\mathbb{E}[e^{tR}] \leq e^{t^2/2}$

Proof. By Definition of expectation,

$$\begin{aligned} \mathbb{E}[e^{tR}] &= \frac{1}{2}(e^t + e^{-t}) \\ &= \frac{1}{2} \left(\sum_{k: \text{even}} \frac{t^k}{k!} \right) \leq \sum_{k \geq 0} \frac{(t^2/2)^k}{k!} = e^{t^2/2} \end{aligned}$$

□

2. show that $\mathbb{E}[e^{tV_i}] \leq e^{\frac{1}{2}t^2 \cdot \|x\|_2^2}$

Proof.

$$\begin{aligned}\mathbb{E}[e^{tV_i}] &= \mathbb{E}[e^{t \sum_{j=1}^D R_j x_j}] \\ &= \prod_{j=1}^D [\mathbb{E}[e^{(tx_j)R_j}]]\end{aligned}$$

from 1

$$\begin{aligned}&\leq \prod_{j=1}^D [e^{x^2 t^2 / 2}] \\ &\leq e^{\frac{t^2 \|x\|_2^2}{2}}\end{aligned}$$

□

3. Let $G \sim \mathcal{N}(0,1)$. Show that

$$\mathbb{E}[e^{tG}] = e^{t^2/2}$$

Proof. By definition of expectation of continuous random variable:

$$\mathbb{E}[e^{tG}] = \int_{-\infty}^{\infty} e^{tG} \cdot \frac{e^{-G^2/2}}{\sqrt{2\pi}} dG = e^{t^2/2}$$

By Wolfram Alpha

□

4. For $\|x\|_2^2 = 1$, show that

$$\mathbb{E}[e^{tV_i^2}] \leq \frac{1}{\sqrt{1-2t}}$$

Proof. from the given convolution trick we have:

$$\begin{aligned}\mathbb{E}[e^{tV_i^2}] &= \mathbb{E}_G[\mathbb{E}_V[e^{\sqrt{2t}GV}]] \\ &= \mathbb{E}_G[e^{tG^2}] \text{ from 2} \\ &\leq \frac{1}{\sqrt{1-2t}} \text{ from the bound in class}\end{aligned}$$

□