Abhimanyu M A **(11111002)** Sumesh T A **(11111065)**
Department of Computer Science and Engineering
IIT Kanpur

November 14, 2011

ASSIGNMENT 4

Problem 1. **Method of Bounded Difference**

(a) Rumour Spreading
(b) Triangles in a Random Graph

(a) **Rumour Spreading**
We have to look at the first stage where $< \frac{n}{2}$ people know the rumour, let us assume that the people are in some way ordered $P_1, P_2, P_3, \ldots P_n$ let the first i people be the persons knowing rumours, we are now guaranteed that the persons from $P_{n/2+1}, \ldots P_n$ doesn't know the rumour, we now omit everything and check only how many people will know the rumour here. The process for rumour spreading is as follows, in each step a person from $P_1, \ldots P_i$ picks a random person and tells him the rumours, we shall now first say the following thing,

Lemma 1. *On expectation, atleast half the people in* $1, \ldots i$ *will pick people from the set* $P_{n/2+1}, \ldots P_n$

Now consider the random experiment as follows we toss $i/2$ balls into $n/2$ bins , and all we need to find is the expected number of non-empty bins
We shall now define the following random variable
X = Number of bins with atleast 1 ball
We can define

$$X_i = \begin{cases} 1 & \text{if bin } i \text{ has atleast one ball} \\ 0 & \text{otherwise} \end{cases}$$

Let $I = i/2$ and $N = n/2$ We can see that X_i = Probability that the bin is not empty

$$X_i = 1 - \left(1 - \frac{1}{N}\right)^I \tag{1-1}$$
$$X_i = 1 - e^{-I/N} \tag{1-2}$$

Now we can see that

$$E[X] = N - Ne^{-I/N} \tag{1-3}$$
$$E[X] \geq N\left(\frac{I}{N} - \frac{I^2}{2N^2}\right) \tag{1-4}$$
$$E[X] \geq I - \frac{I^2}{2N} \tag{1-5}$$
$$E[X] \geq I\left(1 - \frac{I}{2N}\right) \tag{1-6}$$
$$E[X] \geq \left(\frac{i}{2}\right)\left(1 - \frac{i}{2n}\right) \tag{1-7}$$
$$\tag{1-8}$$

Since $i \leq \frac{n}{2}$

$$E[X] \geq \binom{3i}{8} \tag{1-9}$$

$$\tag{1-10}$$

Note that we already have i knowing nodes, hence the total is $\geq (1+\frac{3}{8})i$ which proves that it increases with a constant factor.
Now we can easily apply MOBD on this since $|E[X_A] - E[X_{A'}]| \leq 1$ this satisfies the Lipschitz Condition and hence we can bound

$$Pr[|X - E[X]| > (11/32)i] \leq exp\left(\frac{-121i^2}{1024n}\right) \tag{1-11}$$

(b) **Triangles in a Random Graph**
Let the grah formed be $G(V,E)$ We shall define X_i as follows

$$X_{ijk} = \begin{cases} 1 & \text{if}(i,j)(j,k)(k,i) \in E \\ 0 & \text{otherwise} \end{cases}$$

i. We can now see that $X = \sum X_{ijk}$ over all distinct triplets $\{i,j,k\}$. We can now by linearity of expectation say that given

$$\begin{aligned} E[X_{ijk}] &= P[X_{ijk} = 1] && (1\text{-}12) \\ &= p^3 && (1\text{-}13) \end{aligned}$$

we have

$$\begin{aligned} E[X] &= \sum E[X_{ijk}] \text{ over all } i,j,k && (1\text{-}14) \\ &= \binom{n}{3} p^3 && (1\text{-}15) \end{aligned}$$

To find $E[X^2]$ we shall examine X.

$$\begin{aligned} X &= (X_1 + X_2 + \ldots X_k) && (1\text{-}16) \\ & \text{Where } k = \binom{n}{3} \\ X^2 &= (X_1 + X_2 + \ldots X_k)^2 && (1\text{-}17) \\ &= \sum X_i \cdot X_j && (1\text{-}18) \\ & \text{Over all } i,j \text{ pairs} \end{aligned}$$

We shall now see that X_i and X_j are not independent, since they could share vertices, we could have them sharing $0, 1, 2$ or 3 vertices.

$$\begin{aligned} E[X^2] &= \sum (E[X_i \cdot X_j]) && (1\text{-}19) \\ &= \binom{n}{3} p^3 \left(\binom{n-3}{3} p^3 + 3\binom{n-3}{2} p^2 + 3(n-3)p + 1\right) && (1\text{-}20) \\ & && (1\text{-}21) \end{aligned}$$

This indendence itself is the reason why we can't apply Chernoff Bound, since they are dependent when the share the vertices.

ii. Chebyshev's Inequality states that given some random variable X with expected value μ and variance σ^2 we have

$$Pr(|X - \mu| > k\sigma) \leq \frac{1}{k^2} \tag{1-22}$$

We also now that

$$\begin{aligned}
\sigma^2 &= E[X^2] - E[X]^2 && \text{(1-23)}\\
&= \binom{n}{3}p^3(\binom{n-3}{3}p^3 + 3\binom{n-3}{2}p^2 + 3(n-3)p + 1 - \binom{n}{3}p^3) && \text{(1-24)}\\
&\leq \binom{n}{3}p^3(\binom{n}{3}p^3 + 3\binom{n}{2}p^2 + 3(n)p + 1 - \binom{n}{3}p^3) && \text{(1-25)}\\
&\leq \binom{n}{3}p^3(\frac{3}{2}n^2p^2 + 3np + 1) && \text{(1-26)}\\
&\leq n^5 && \text{(1-27)}
\end{aligned}$$

In this case we get

$$\sigma \leq n^{\frac{5}{2}} \tag{1-28}$$

For some constant c
Since we have a bound on σ we can apply Chebyshev's Inequality

$$Pr(X > 5/4E[X]) = Pr(|X - E[X]| > 1/4E[X]]) \tag{1-29}$$

From the above equations we can get $E[X] = c' \cdot n^{\frac{1}{2}} \cdot \sigma$ which if substituted we get

$$Pr(|X - E[X]| > c'' \cdot n^{\frac{1}{2}} \cdot \sigma) \leq \frac{1}{c''n} \tag{1-30}$$

This gives us a inverse linear probability ratio.

iii. We shall now use Method of Bounded Difference to prove the bounds. Since $X = X_i$ and we need to bound $X > \alpha E[X]$ we can see that our required function is the sum function, since $X_i = 0, 1$. We have for any two X' and X'' seperated only by a single X_i $|X'' - X'| \leq 1$ this satisfies the Lipschitz Condition and hence we can bound

$$\begin{aligned}
Pr(X > 5/4E[X]) &= Pr(|X - E[X]| > 1/4E[X]]) && \text{(1-31)}\\
Pr(|X - E[X]| > 1/4E[X]]) &\leq exp\left(\frac{-E[X]^2}{32\binom{n}{3}}\right) && \text{(1-32)}\\
&\leq exp\left(\frac{-p^6\binom{n}{3}}{32}\right) && \text{(1-33)}
\end{aligned}$$

We can see that for any value of p there will exists some point where MOBD will give a much better result than Chebyshev's inequality.

$\longrightarrow$ $\mathcal{A}$nswer

Problem 2. **Principle of deffered decision** First we shall use the following basic probability lemmas we will use throughout the question without proof

Lemma 2. *If we select l numbers randomly uniformly distributed in the range $(0, 1)$ then the expected value of the smallest no. (length of the leftmost interval) is $\frac{1}{l+1}$*

Lemma 3. *If there are l random variables uniformly distributed in the range $(\delta, 1)$ then the expected value of the smallest no. is $\delta + \frac{1}{l+1}$*

We define our random variable $X = \sum Xi$ where X_i is the i^{th} edge added to the Kruskal Algorithm.

We are given a complete graph K_n with random edge weights put uniformly, randomly , independently in the range $(0, 1)$. The algorithm is as follows, we sort the edges, and keep on selecting the lowest weight edge that does not cause a cycle. If we define X_i to be the weight of the i^{th} edge that is selected then we can see that

$$X = \sum_{i=1}^{n-1} (X_i) \tag{2-1}$$

We can apply linearity of expectation on this and obtain

$$E[X] = \sum_{i=1}^{n-1} E[(X_i)] \tag{2-2}$$

Clearly $X_1 = \frac{1}{\binom{n}{2}+1}$ We will uncover the weight of the vertices only when we need to consider that we are going to insert the $(i+1)^{th}$ vertex we can see that the Kruskal forest till now will be a set of isolated vertices and few connected trees. Let $\beta_1, \beta_2, \ldots \beta_{n-i}$ and also that $1 \leq \beta_i \leq i$ be the number of vertices in each of the componenets then we can see that the total number of edges from which the connecting edge can be selected is

$$l_{i+1} = \binom{n}{2} - \sum_{i=1}^{n-i} \binom{\beta_i}{2} \tag{2-3}$$

Since this many points remain the expected value of

$$E[X_{i+1}] = E[X_i] + \frac{1}{l_{i+1} + 1} \tag{2-4}$$

Since we need to get a $E[X_{i+1}] \leq f(i+1)$ we need to find the minimum possible value of l_{i+1} which is obtained when $\beta_1 = i$ and rest all $\beta_i = 1$ that means all the edges form a single connected component and the rest as single vertices, hence

$$l_{i+1} \geq \binom{n}{2} - \binom{i}{2} \tag{2-5}$$

$$E[X_{i+1}] \leq E[X_i] + \frac{2}{n(n-1) - i(i-1)} \tag{2-6}$$

$$E[X_{i+1}] \leq E[X_i] + \frac{2}{(n-1)^2 - (i-1)^2} \tag{2-7}$$

If we let $N = n - 1$ and $I = i - 1$ we get

$$E[X_{i+1}] \leq E[X_i] + \frac{2}{N^2 - I^2} \tag{2-8}$$

$$E[X_{i+1}] \leq E[X_i] + \frac{1}{N}\left(\frac{1}{N-I} + \frac{1}{N+I}\right) \tag{2-9}$$

expanding we get

$$E[X_{i+1}] \leq \frac{1}{N} \sum_{I=0}^{i} \left(\frac{1}{N-I} + \frac{1}{N+I}\right) \tag{2-10}$$

$$E[X_{i+1}] \leq \frac{1}{n} log\left(\frac{n}{n-i}\right) \tag{2-11}$$

Now we move on to analysing $\boldsymbol{E[X]}$

$$\boldsymbol{E[X] \leq \sum_{i=1}^{n-i} \frac{1}{n} log\left(\frac{n}{n-i}\right)} \tag{2-12}$$

$$\boldsymbol{E[X] \leq \frac{1}{n} \sum_{i=1}^{n-i} log\left(\frac{n}{n-i}\right)} \tag{2-13}$$

$$\boldsymbol{E[X] \leq \frac{1}{n} log\left(\prod_{i=1}^{n} \frac{n}{n-i}\right)} \tag{2-14}$$

$$\boldsymbol{E[X] \leq \frac{1}{n} log\left(\frac{n^{n-1}}{(n-1)!}\right)} \tag{2-15}$$

When can now approximate using Sterling's Approximation $\boldsymbol{n! = c\sqrt{n}\left(\frac{n}{e}\right)^n}$

$$\boldsymbol{E[X] \leq \frac{1}{n} log\left(\frac{e^n}{c\sqrt{n}}\right)} \tag{2-16}$$

$$\boldsymbol{E[X] \leq \frac{1}{n}\left(n - log(c\sqrt{(}n))\right)} \tag{2-17}$$

Since $\boldsymbol{c\sqrt{(}n) > 1}$

$$\boldsymbol{E[X] \leq \frac{1}{n}(n)} \tag{2-18}$$

$$\boldsymbol{E[X] = O(1)} \tag{2-19}$$

$\longrightarrow$ $\mathcal{A}$nswer

Problem 3. **Delay Sequences**

(a) Reaching $logn$ in $O(logn)$ steps with $d = 2$ $b = 1$
(b) Find the time to reach n with d neighbours and b option of relaxation

(a) We can first construct a delay sequence the model the problem, if any of the counter took time $\boldsymbol{m}$ to reach the value of $\boldsymbol{log(n)}$ then we can see taht we would get a delay sequence as a "witness" to the "delay" of the counter, As exlained in class we shall now construct a possible delay sequence.

Round	Counter	Outcome of Toss	Value at Beginning	Value at End
m	C	H	(log n)-1	log n
m-1	C	T	(log n)-1	(log n)-1
m-2	C	H	(log n)-1	(log n)-1
m-3	C‘	T	(log n)-2	(log n)-2
m-4	C‘	H	(log n)-3	(log n)-2
m-5	C‘	H	(log n)-3	(log n)-3
m-6	C“	H	(log n)-5	(log n)-4

We shall now look at the probability of a existence of a delay sequence of this length. the points to be noted are

i. If we get a Head(H) either we reach a previous state of lower value, or move to a counter with lower value hence the number of Heads(H) cannot be more than $\boldsymbol{logn}$.

ii. Each coin toss is random and independent.

We can now see that

$$\begin{aligned}\Pr[\text{The Counter takes } \boldsymbol{m} \text{ steps}] &= \Pr\ [\text{Existence of a } \boldsymbol{m} \text{ long delay sequence}]\\ \Pr\ [\text{Existence of a } \boldsymbol{m} \text{ long delay sequence}] &= \Pr[\text{Single } \boldsymbol{m} \text{ long delay sequence}] \times \Pr[\text{Number of } \boldsymbol{m} \text{ long de}\end{aligned}$$

We shall first see that each delay sequence is obtained by a random coin toss, and the probability of a delay sequence

$$\Pr[\text{Single } \boldsymbol{m} \text{ long delay sequence}] = \boldsymbol{2^{-m}} \tag{3-3}$$

We know that we have exactly $\boldsymbol{logn}$ Heads in the whole delay sequence and in total $\boldsymbol{m}$, so we can first pick the points at which we get a head, which is $= \binom{m}{logn}$. Now at each point we get a head we have three choices either to shift to any of it's neighbours or to shift one down, hence total options $= \boldsymbol{3^{logn}}$. We are now free to decide the starting point and the vertex from which we start each of which gives a $\boldsymbol{logn}$ option.

$$\Pr[\text{Number of } \boldsymbol{m} \text{ long delay sequence}] = \boldsymbol{(logn)^2 \cdot 3^{logn} \cdot \binom{m}{logn}} \tag{3-4}$$

So

$$\Pr[\text{The Counter takes } \boldsymbol{m} \text{ steps}] \quad = \quad \boldsymbol{(logn)^2 \cdot 3^{logn} \cdot \binom{m}{logn} \cdot 2^{-m}} \tag{3-5}$$

To prove that $\boldsymbol{m}$ is $\boldsymbol{O(logn)}$ we will substitute that value for $\boldsymbol{m}$ and see

$$\begin{aligned}\Pr[\text{Algo takes } \boldsymbol{clogn} \text{ steps}] &= (logn)^2 \cdot 3^{logn} \cdot \binom{clogn}{logn} \cdot 2^{-clogn} &\text{(3-6)}\\ &\leq (logn)^2 \cdot 3^{logn} \cdot (ce)^{logn} \cdot 2^{-clogn} &\text{(3-7)}\\ &\leq (logn)^2 \cdot n^{log3} \cdot n^{1+logc} \cdot n^{-c \cdot log2} &\text{(3-8)}\\ &\leq (logn)^2 \cdot n^{log\frac{3ec}{2c}} &\text{(3-9)}\\ & &\text{(3-10)}\end{aligned}$$

In the asymptotic case we need not bother about the $\boldsymbol{(logn)^2}$ factor since $\boldsymbol{n^2}$ will be much higher and hence $\frac{(logn)^2}{n^2}$ will be very less. So now to say this with high probability all we need is the value for which

$$\begin{aligned} log\left(\frac{3e}{c}\right) &\leq -2 &\text{(3-11)}\\ \frac{3e}{c} &\leq e^{-2} &\text{(3-12)}\\ \frac{1}{c} &\leq \frac{e^{-3}}{3} &\text{(3-13)}\\ c &\geq 3e^3 &\text{(3-14)}\end{aligned}$$

Hence we get the probability that the counter set will take more than $\boldsymbol{3e^3}$ steps to reach $\boldsymbol{logn}$ is $\leq \frac{1}{n^2}$ hence since $\boldsymbol{3e^3}$ is a constant we can say that it will reach $\boldsymbol{logn}$ in $\boldsymbol{O(logn)}$ steps with w.h.p

(b) We can see that the same delay sequence with certain modification can apply here,

Round	Counter	Outcome of Toss	Value at Beginning	Value at End
m	C	H	n-1	log n
m-1	C	T	n-1	n-1
m-2	C	H	n-1	n-1
m-3	C‘	T	n-1-b	n-1-b
m-4	C‘	H	n-2-b	n-1-b
m-5	C‘	H	n-2-b	n-2-2b
m-6	C“	H	n-3-3b	n-2-3b

Here we just carry forward all the arguments we made for the other, and we can see that the probability for existence of a tree is also the same, but what has changed is the number of ways it can branch.
Option of branching into $d+1$ counters but is not available at every counter, that is possible at only $\frac{n}{b}$ places since after each of such branching the value reduces by b . We are not really sure about the $+1$ but since we are overestimating it is accetable. Now it can also be seen that the maximum number of heads is n and in which we can choose branchings.

$$\text{Number of ways to branch} = \binom{n}{n/b}(1+d)^{n/b} \tag{3-15}$$

So putting all the rest of the terms that have not changed we get.

$$\begin{aligned}
\Pr[\text{The Counter takes } m \text{ steps}] &= n^2 \cdot \binom{n}{n/b} \cdot (1+d)^{n/b} \cdot \binom{m}{n} \cdot 2^{-m} && (3\text{-}16)\\
\Pr[\text{The Counter takes } m \text{ steps}] &= n^2 \cdot (be)^{n/b} \cdot (1+d)^{n/b} \cdot \left(\frac{me}{n}\right)^n \cdot 2^{-m} && (3\text{-}17)\\
& && (3\text{-}18)
\end{aligned}$$

We now look at $m = Cn/b = cn$

$$\begin{aligned}
Pr &= n^2 \cdot (be)^{n/b} \cdot (1+d)^{n/b} \cdot \left(\frac{Ce}{b}\right)^{Cn/b} \cdot 2^{-Cnb} && (3\text{-}19)\\
Pr &= n^2 \cdot (beD(Ceb)^C)^{n/b} \cdot 2^{-Cnb} && (3\text{-}20)\\
Pr &= n^2 \cdot 2^{((n/b)(log(beD)+Clog(Ceb)))-Cnb} && (3\text{-}21)
\end{aligned}$$

To get a good bound we will set $Pr = n^2 \cdot 2^{-n}$ we will not bother about n^2 since asymptotically this value will decrease rapidly. Hence

$$\begin{aligned}
(n/b)(log(beD) + Clog(Ceb)) - Cnb &\leq -n && (3\text{-}22)\\
log(beD) + Clog(Ceb) &\leq (Cb-1)b && (3\text{-}23)\\
log(beD) + Clog(Ceb) &\leq Cb^2 && (3\text{-}24)\\
C(b^2 - log(Ceb)) &\geq log(beD) && (3\text{-}25)\\
Cb^2 &\geq logb + 1 + logD && (3\text{-}26)\\
C &\geq 2 + logD && (3\text{-}27)
\end{aligned}$$

Hence if $C \geq 3 + \frac{logd}{b}$ satisfies the equation, hence for a given graph it is a constant.

$\longrightarrow$ $\mathcal{A}$nswer

*Submitted by Abhimanyu M A **(11111002)** Sumesh T A **(11111065)** on November 14, 2011.*