198:538 Lecture Given on April 14th, 1998

Lecturer: Professor Eric Allender
Scribe: Sunny Daniels

November 25, 1998

1 The result to be proved

In this lecture, we will prove the following result that we made use of in the previous lecture:

Fact 1 *Let k and a be any two positive integers. Then there exists a positive integer b, whose value depends only on the value of a, and a polynomial-time algorithm $\mathcal{A}$ that takes the string 1^n as input, for any $n \in \mathbb{N}$ with $n \geq 2$, and outputs a Boolean matrix[1] $N = (\nu_{ij})$. $a < b$, and the matrix N satisfies the following conditions:*

1. *N has n^k rows.*
2. *N has $b \ln n$ columns.*[2]
3. *Each row has exactly $a \ln n$ entries equal to 1.*
4. *For any two rows i and l, if we let*

$$\begin{aligned} C_i &= \{j : \nu_{ij} = 1\} \\ C_l &= \{j : \nu_{lj} = 1\} \end{aligned}$$

then:

$$|C_i \cap C_l| \leq \ln n$$

2 Picking the value of b:

Now, we will define a real number η by setting:

$$\eta = e^{2k+15+2a}$$

[1] i.e. a matrix of which all of the elements are either 1 or 0.

[2] Since $b \ln n$ is not, in general, an integer, we really mean "$\lfloor b \ln n \rfloor$ columns". For neatness, we will omit the "$\lfloor$" and "$\rfloor$" whenever we are using real numbers to specify quantities that are obviously integers, and tacitly assume that these real numbers are truncated to integers.

We will then define the value of b by setting:

$$b = 1 + \eta a^2$$

3 The Algorithm $\mathcal{A}$:

The algorithm $\mathcal{A}$ for building the matrix N is actually a very simple "greedy" algorithm; the algorithm is as follows:

1. Let $\mathcal{L}$ be a list of all of the possible Boolean row vectors of length $b \ln n$ of which exactly $a \ln n$ entries have the value 1. The order in which the row vectors are listed in $\mathcal{L}$ is irrelevant to this algorithm. Since the *total* number of row vectors of length $b \ln n$ is $2^{b \ln n}$, which is clearly polynomial in n, both the length of the list $\mathcal{L}$ and the amount of time required to construct it are polynomial in n.

2. The matrix N is constructed a row at a time, starting at the top row and finishing at the bottom row.

3. This algorithm has an integer variable ρ which indicates the number of the next row of N to be constructed. Initially, ρ equals 1.

4. The algorithm then looks for the first element of the list $\mathcal{L}$ for which copying this element of $\mathcal{L}$ to row ρ of N would give:

   $$|C_i \cap C_\rho| \leq \ln n, \text{for every } i \in \mathbb{Z} \text{ with } i < \rho \tag{1}$$

   where C_i and C_ρ are defined by:

   $$\begin{aligned} C_i &= \{j : \nu_{ij} = 1\} \\ C_\rho &= \{j : \nu_{\rho j} = 1\} \end{aligned}$$

   The first such element of $\mathcal{L}$ is then copied to row ρ of the matrix N. The value of ρ is then incremented.

   The question that immediately arises is: What if no elements of the list $\mathcal{L}$ satisfy condition 1? We will resolve this question later by showing that this undesirable situation is impossible.

5. If there are still rows of the matrix N remaining to be constructed (i.e. $\rho < n^k$), then we branch back to step 4 above. Otherwise, the algorithm has finished constructing the matrix, so it terminates.

4 Proof that the algorithm $\mathcal{A}$ works

Now, assume temporarily that the undesirable situation mentioned in step 4 of algorithm $\mathcal{A}$ is impossible. It is then obvious that the algorithm terminates, since each iteration of the main loop adds a row to the matrix until the matrix is fully constructed. It should also be obvious, from the specification of the algorithm, that the matrix constructed by the algorithm satisfies the required conditions given in Fact 1.

Hence, all that remains to be proved (in order to prove that Fact 1 is true) is that the undesirable situation mentioned in step 4 of algorithm $\mathcal{A}$ never occurs. The remainder of this lecture will be devoted to proving that this undesirable situation never occurs.

5 Another Chernoff Bound

From Alon and Spencer [**?**], we have the following result. This result is one of the class of results known as *Chernoff Bounds*.

Lemma 1 *Let Y be the sum of mutually independent indicator random variables*[3]*, $\mu = E[Y]$. [From the statement of this result (without proof) in Alon and Spencer [**?**], it is unclear as to whether or not it is necessary for the indicator random variables to be identically distributed. In our application of this result, the indicator random variables* will *be identically distributed.] For every real number $\epsilon > 0$, we have:*

$$Pr[|Y - \mu| > \epsilon\mu] < 2e^{-c_\epsilon \mu}$$

where c_ϵ (which is greater than zero) is defined by:

$$c_\epsilon = \min\left[-\ln(e^\epsilon(1+\epsilon)^{-(1+\epsilon)}), \frac{\epsilon^2}{2}\right]$$

6 An Elementary Result about Probabilities

Lemma 2 *Let $E_1, E_2, \ldots, E_N$ be any finite collection of events over the same sample space S. Then:*

$$Pr[E_1 \cup E_2 \cup \ldots \cup E_N] \leq Pr[E_1] + Pr[E_2] + \ldots + Pr[E_N]$$

Proof: For any two events E_i and E_j over the same sample space S, a standard elementary result in probability theory tells us that:

$$Pr[E_i \cup E_j] = Pr[E_1] + Pr[E_2] - Pr[E_1 \cap E_2]$$

Now, another elementary result in probability theory tells us that all probabilities are non-negative, so:

$$Pr[E_1 \cap E_2] \geq 0$$

[3] *Indicator random variables* are random variables whose value is always either zero or one

and hence:

$$Pr[E_i \cup E_j] \leq Pr[E_1] + Pr[E_2]$$

From this result, lemma 2 follows immediately by induction.

7 Conclusion of Proof

Assume for now that:

Claim 1 *If:*

1. *B is any fixed subset of size $a \ln n$ of the set:*

$$U = \{1, 2, \ldots, b \ln n\}$$

2. *A is any randomly-chosen subset of size $a \ln n$ of the same set U.*

then:

$$Pr_{\mathcal{E}_\mathcal{L}}(|A \cap B| > \ln n) < \frac{1}{n^{2k}} \tag{2}$$

Now, consider the situation in which, for an arbitrary iteration of the loop in algorithm $\mathcal{A}$, step 4 is executed. Now, pretend momentarily that we were to *randomly* pick an element of $\mathcal{L}$, and copy it to row ρ of N. We will use the symbol "$\mathcal{E}_\mathcal{L}$" to denote this experiment of randomly picking an element of $\mathcal{L}$, and the symbol "$Pr_{\mathcal{E}_\mathcal{L}}$" to denote the corresponding probability function.

The probability that the new row of N would *fail to* satisfy property 1 in step 4 would then be the probability that one of $|C_1 \cap C_\rho|$, $|C_2 \cap C_\rho|$, ..., $|C_{\rho-1} \cap C_\rho|$ was greater than $\ln n$. By lemma 2, this probability would be at most:

$$\sum_{i=1}^{\rho-1} Pr_{\mathcal{E}_\mathcal{L}}[|C_i \cap C_\rho| \geq \ln n]$$

which, by equation 2, is *strictly* less than:

$$\frac{\sum_{i=1}^{\rho-1} i}{n^{2k}} \leq \frac{\rho^2}{n^{2k}}$$

Since the termination condition on the loop ensures that $\rho \leq n^k$, this quantity is at most $\frac{n^{2k}}{n^{2k}} = 1$. Hence, the probability that the hypothetical randomly-chosen row would fail to satisfy property 1 is strictly less than 1. Since this row of the matrix is randomly chosen from the list $\mathcal{L}$, this implies that the list $\mathcal{L}$ must contain at least one element that *does* satisfy property 1. This proves that step 4 in algorithm $\mathcal{A}$ can never fail.

8 Proof of Claim 1

Take any fixed subset B of size $a \ln n$ of the set:

$$U = \{1, 2, \ldots, b \ln n\}$$

Let A be any randomly-chosen subset of size $a \ln n$ of the same set U. Define:

$$D = \Pr_{\mathcal{E}_\mathcal{L}}(|A \cap B| > \ln n)$$

So far, we have defined all of our probabilities with respect to the experiment $\mathcal{E}_\mathcal{L}$. However, in order to derive our desired upper bound on D, we will have to consider a different experiment $\mathcal{E}_\mathcal{B}$. This new experiment $\mathcal{E}_\mathcal{B}$ will consist of $b \ln n$ independent Bernoulli trials. Each Bernoulli trial will have a probability of success of $\frac{a}{b}$. For each $i \in \{1, 2, \ldots, b \ln n\}$, we will define:

1. $i \in A$ if the outcome of the ith trial is success.
2. $i \notin A$ if the outcome of the ith trial is failure.

Clearly then, the distribution of $|A|$ is binomial. Then the most likely value of $|A|$ is $a \ln n$ (e.g. see Feller [**?**]). Therefore:

$$\frac{D}{a \ln n} \leq D \cdot \Pr_{\mathcal{E}_\mathcal{L}}(|A| = a \ln n) \tag{3}$$

But then, by the properties of conditional probabilities and the definitions of $\mathcal{E}_\mathcal{L}$ and $\mathcal{E}_\mathcal{B}$, we have:

$$D = \Pr_{\mathcal{E}_\mathcal{B}}(|A \cap B| > \ln n || A| = a \ln n)$$

Then, since all probabilities are non-negative:

$$D \cdot \Pr_{\mathcal{E}_\mathcal{L}}(|A| = a \ln n) < \sum_{i=1}^{b \ln n} \Pr_{\mathcal{E}_\mathcal{B}}(|A \cap B| > \ln n || A| = i) \Pr_{\mathcal{E}_\mathcal{B}}(|A| = i) \tag{4}$$

But then, by the partition theorem:

$$\sum_{i=1}^{b \ln n} \Pr_{\mathcal{E}_\mathcal{B}}(|A \cap B| > \ln n || A| = i) \Pr_{\mathcal{E}_\mathcal{B}}(|A| = i) = \Pr_{\mathcal{E}_\mathcal{B}}(|A \cap B| > \ln n) \tag{5}$$

Now, the distribution of $|A \cap B|$ (w.r.t. the experiment $\mathcal{E}_\mathcal{B}$) is also binomial, since whether or not each $i \in B$ is in $A \cap B$ is effectively determined by an independent Bernoulli trial whose probability of success is $\frac{a}{b}$. Hence, by the properties of the binomial distribution, the expected value of $|A \cap B|$ w.r.t. the experiment $\mathcal{E}_\mathcal{B}$ is $\frac{a^2 \ln n}{b}$. Let's call this expected value μ. Then, by our choice of b in section 2, we have:

$$\mu = \frac{a^2 \ln n}{1 + \eta a^2}$$

Now, since a^2 is positive, and $\ln n$ and η are non-negative, we have:

$$\frac{a^2 \ln n}{2\eta a^2} \leq \frac{a^2 \ln n}{1 + \eta a^2} \leq \frac{\ln n}{\eta} \tag{6}$$

Now, let's define:

$$\widehat{\mu} = \frac{a^2 \ln n}{2\eta a^2} = \frac{\ln n}{2\eta}$$

so that:

$$\widehat{\mu} \leq \frac{a^2 \ln n}{1 + \eta a^2} = \mu$$

Our choice of μ, together with inequality 6, also implies that:

$$\eta\mu \leq \ln n$$

Therefore "$|A \cap B| > \ln n$" is a stronger condition than "$|A \cap B| > \eta\mu$"; hence:

$$\Pr_{\mathcal{E}_\mathcal{B}}[|A \cap B| > \ln n] \leq \Pr_{\mathcal{E}_\mathcal{B}}[|A \cap B| > \eta\mu]$$

Putting this together with equations 3, 4 and 5 then gives:

$$\frac{D}{a \ln n} < \Pr_{\mathcal{E}_\mathcal{B}}[|A \cap B| > \eta\mu] \tag{7}$$

Now, we define a real number ϵ by:

$$\epsilon = \eta - 1$$

Since obviously $\eta > 1$, we have $\epsilon > 0$. Then, by lemma 1, we have:

$$\Pr_{\mathcal{E}_\mathcal{B}}[||A \cap B| - \mu| > \epsilon\mu] < 2e^{-c_\epsilon \mu} \tag{8}$$

where c_ϵ is defined by:

$$c_\epsilon = \min\left[-\ln\left(e^\epsilon (1+\epsilon)^{-(1+\epsilon)}\right), \frac{\epsilon^2}{2}\right]$$

Now, clearly the condition "$||A \cap B| - \mu| > \epsilon\mu$" is weaker than the condition "$|A \cap B| - \mu > \epsilon\mu$"; hence:

$$\Pr_{\mathcal{E}_\mathcal{B}}[|A \cap B| - \mu > \epsilon\mu] < \Pr_{\mathcal{E}_\mathcal{B}}[||A \cap B| - \mu| > \epsilon\mu]$$

Then, by equation 8, together with our definition of ϵ, we have:

$$\Pr_{\mathcal{E}_\mathcal{B}}[|A \cap B| > \eta\mu] = \Pr_{\mathcal{E}_\mathcal{B}}[|A \cap B| - \mu > \epsilon\mu] \leq 2e^{-c_\epsilon \mu}$$

and therefore, by equation 7:

$$\frac{D}{a \ln n} < 2e^{-c_\epsilon \mu} \tag{9}$$

Now, if we invert the sign of both sides of the equation defining c_ϵ, we get:

$$-c_\epsilon = \max\left[\ln\left(e^\epsilon(1+\epsilon)^{-(1+\epsilon)}\right), -\frac{\epsilon^2}{2}\right] \tag{10}$$

Then, since the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 2a(\ln n)e^{x\mu}$ is obviously monotonic (since a, $\ln n$ and μ are nonnegative), the above equation implies:

$$2a(\ln n)e^{-c_\epsilon \mu} = \max\left[\widehat{A}, \widehat{B}\right]$$

where:

$$\widehat{A} = 2a(\ln n)e^{\mu \ln(e^\epsilon(1+\epsilon)^{-(1+\epsilon)})}$$

and:

$$\widehat{B} = 2a(\ln n)e^{-\mu\frac{\epsilon^2}{2}}$$

8.1 Proof that $\widehat{A} \leq \frac{1}{n^{2k}}$

Now, evaluating $\widehat{A}$ for our chosen values of μ and ϵ gives us:

$$\begin{aligned}
\widehat{A} &= (2a\ln n)e^{\mu \ln(e^\epsilon(1+\epsilon)^{-(1+\epsilon)})} \\
&= (2a\ln n)e^{\mu(\epsilon-(1+\epsilon)\ln(1+\epsilon))} \\
&= (2a\ln n)e^{\mu(\eta-1-\eta\ln\eta)}
\end{aligned}$$

Now, since k and a are both greater than or equal to 0, $2k+15+2a \geq 1$. Since we define η to be $e^{2k+15+2a}$, this implies that $\ln\eta \geq 1$. Therefore $\eta - 1 - \eta\ln\eta \leq 0$. Since we define $\widehat{\mu}$ in such a way that $\widehat{\mu} < \mu$, and the exponential function is monotonic, we have:

$$e^{\widehat{\mu}(\eta-1-\eta\ln\eta)} \geq e^{\mu(\eta-1-\eta\ln\eta)}$$

Then, since $2a$ and $\ln n$ are both non-negative:

$$(2a\ln n)e^{\widehat{\mu}(\eta-1-\eta\ln\eta)} \geq (2a\ln n)e^{\mu(\eta-1-\eta\ln\eta)}$$

And therefore:

$$\begin{aligned}
\widehat{A} &\leq (2a\ln n)e^{\mu(\eta-1-\eta\ln\eta)} \\
&= (2a\ln n)e^{(\ln n - \widehat{\mu} - (\ln\eta)(\ln n))} \\
&= (2a\ln n)ne^{-\widehat{\mu}}e^{-(\ln\eta)(\ln n)} \\
&= \frac{(2a\ln n)n}{n^{2k+15+2a}n^{\frac{1}{2\eta}}}
\end{aligned}$$

Then, since $n \geq 2$ and $\eta \geq 1$, we have $n^{\frac{1}{2\eta}} > 1$. Therefore:

$$\widehat{A} \leq \frac{(2a\ln n)n}{n^{2k+15+2a}}$$

Now, since $k \geq 0$, $a \geq 0$ and $n \geq 2$, we have:

- $(2a \ln n)n \geq 0$
- $2a \leq 2^{2a}$
- $\ln n \leq n$

and therefore:

$$\widehat{A} \leq \frac{1}{n^{2k}}$$

8.2 Proof that $\widehat{B} \leq \frac{1}{n^{2k}}$

To prove this, we will prove an equivalent assertion, namely that:

$$\widehat{B}n^{2k} \leq 1$$

Now:

$$\begin{aligned} \widehat{B}n^{2k} &= 2n^{2k}a(\ln n)e^{-\mu\frac{(\eta-1)^2}{2}} \\ &= \frac{2n^{2k}a(\ln a)}{e^{\mu\frac{(\eta-1)^2}{2}}} \end{aligned}$$

Now, since $(\eta - 1)^2 > 0$ and $\widehat{\mu} \leq \mu$ (where $\widehat{\mu}$ is as defined in subsection 8.1 above), we can increase the quantity in the last equation above by replacing μ in the denominator by $\widehat{\mu}$. Hence:

$$\begin{aligned} \widehat{B}n^{2k} &\leq \frac{2n^{2k}a(\ln n)}{e^{\widehat{\mu}\frac{(\eta-1)^2}{2}}} \\ &= \frac{2n^{2k}a(\ln n)}{e^{\frac{a^2\ln n}{2\eta a^2}\cdot\frac{(\eta a)^2}{2}}} \\ &= \frac{2n^{2k}a(\ln n)}{e^{\frac{\ln n}{2\eta}\cdot\frac{(\eta-1)^2}{2}}} \\ &= \frac{2n^{2k}a(\ln n)}{n^{\frac{(\eta-1)^2}{4\eta}}} \end{aligned}$$

Now, since obviously $\eta \geq 2$, $(\eta - 1)^2 \geq (\frac{\eta}{2})^2$. Therefore we can increase the quantity in the last equation above by replacing "$(\eta - 1)^2$" in the exponent in the denominator by "$(\frac{\eta}{2})^2$". Hence:

$$\begin{aligned} \widehat{B}n^{2k} &\leq \frac{2n^{2k}a(\ln n)}{n^{\frac{(\frac{\eta}{2})^2}{4\eta}}} \\ &= \frac{2an^{2k}(\ln n)}{n^{\frac{\eta}{16}}} \end{aligned}$$

Then, since $n \geq 2$ and $a \geq 1$, we have $2a \leq 2^{2a}$ and $\ln n \leq n$. Therefore:

$$\begin{aligned}
\widehat{B}n^{2k} &\leq \frac{n^{2a} \cdot n^{2k} \cdot n}{n^{\frac{\eta}{16}}} && (11)\\
&= \frac{n^{2a+2k+1}}{n^{\frac{\eta}{16}}} && (12)\\
&= \frac{n^{2a+2k+1}}{(n^{\eta})^{\frac{1}{16}}} && (13)\\
&= \frac{n^{2a+2k+1}}{(n^{e^{2k+15+2a}})^{\frac{1}{16}}} && (14)
\end{aligned}$$

Now, there is a standard result of calculus that says that, for any real number α, $1 + \alpha \leq e^{\alpha}$. Therefore:

$$\begin{aligned}
e^{2k+15+2a} &= e^1 5 e^{2k+2a}\\
&\geq (1+15)(1+2k+2a)\\
&= 16(2a+2k+1)
\end{aligned}$$

Therefore, by equation 14, we have:

$$\begin{aligned}
\widehat{B}n^{2k} &\leq \frac{n^{2a+2k+1}}{(n^{16(2a+2k+1)})^{\frac{1}{16}}}\\
&= \frac{n^{2a+2k+1}}{n^{2a+2k+1}}\\
&= 1
\end{aligned}$$

as required.

8.3 Conclusion of Proof

Since we have now proved that both $\widehat{A}\frac{1}{n^{2k}}$ and $\widehat{B}\frac{1}{n^{2k}}$, equation 10 implies that:

$$2a(\ln n)e^{-c_\epsilon \mu} \leq \frac{1}{n^{2k}}$$

then, since $a \ln n$ is positive, we can multiply equation 9 through by $a \ln n$ to get:

$$D < 2a(\ln n)e^{-c_\epsilon \mu}$$

And therefore:

$$D < \frac{1}{n^{2k}}$$

this completes the proof of claim 1. It therefore follows that the algorithm A works correctly.