CS 2401 - Introduction to Complexity Theory

Lecture #6: Fall, 2015

Lecturer: Toniann Pitassi

Scribe Notes by: Margarita Castro

1 Error amplification

1.1 Review of last class

Definition A language $L \in$ BPP if there exists a PTM M and a polynomial p such that M is polytime in $|x|$ and:

$$\begin{aligned} \forall x \in L \qquad & Pr_{r,|r|=p(|x|)}(M(x,r)=1) \geq \frac{2}{3} \\ \forall x \notin L \qquad & Pr_{r,|r|=p(|x|)}(M(x,r)=1) \leq \frac{1}{3} \end{aligned}$$

Definition A language $L \in$ RP if there exists a PTM M and a polynomial p such that M is polytime in $|x|$ and:

$$\begin{aligned} \forall x \in L \qquad & Pr_{r,|r|=p(|x|)}(M(x,r)=1) \geq \frac{2}{3} \\ \forall x \notin L \qquad & Pr_{r,|r|=p(|x|)}(M(x,r)=1) = 0 \end{aligned}$$

1.2 Error amplification

Theorem 1 *Let L be a language such that there is a PTM M such that:*

$$\begin{aligned} \forall x \in L \qquad & Pr_{r,|r|=p(|x|)}(M(x,r)=1) \geq n^{-c} \\ \forall x \notin L \qquad & Pr_{r,|r|=p(|x|)}(M(x,r)=1) = 0 \end{aligned}$$

Then for every $d > 0$ there is a polytime PTM M' such that:

$$\begin{aligned} \forall x \in L \qquad & Pr_{r,|r|=p(|x|)}(M'(x,r)=0) \leq \frac{1}{2^{n^d}} \\ \forall x \notin L \qquad & Pr_{r,|r|=p(|x|)}(M'(x,r)=1) = 0 \end{aligned}$$

Proof

The main idea is to construct a new PTM $M'(x,r')$, where $|r'| = k \cdot r$.
We need to divide r into k equal size pieces: $r_1, \ldots, r_k$. Then run M for every r_i: $M(x,r_1), \ldots, M(x,r_k)$.
We have two options:

- If any $M(x,r_i)$ outputs 1, then M' outputs 1
- otherwise M' outputs 0

Therefore:

$$\forall x \in L \qquad Pr_{r,|r|=p(|x|)}(M'(x,r') = 0) \leq \frac{1}{2^k}$$
$$\forall x \notin L \qquad Pr_{r,|r|=p(|x|)}(M'(x,r') = 1) = 0$$

Finally, we just need to pick $k = n^d$ and we are done.

□

Theorem 2 *Let L be a language and suppose that there exists a poly-time PTM M such that:*

$$\forall x \in L \; Pr_{r,|r|=p(|x|)}(M(x,r) = L(x)) \geq \frac{1}{2} + |x|^{-c}$$

Then for every $d > 0$ there is a polytime PTM M' such that:

$$\forall x \in L \; Pr_{r,|r|=p(|x|)}(M'(x,r) = L(x)) \geq 1 + 2^{-n^d}$$

Proof

Main idea: create a PTM M' by running M k times. We accept if the majority of the outputs are 1, otherwise we reject. First, we divide r into k equal size pieces: $r_1, \ldots, r_k$. Then run M for every r_i and name the outputs $y_1, \ldots, y_k$.

Lets define the random variable X_i as:

$$X_i = \begin{cases} 1 & \text{if } y_i = L(x) \\ 0 & o.w \end{cases}$$

Note that $X_1, \ldots, X_k$ are independent boolean random variables and that:

$$E(X_i) = P(X_i = 1) \geq \frac{1}{2} + |x|^{-c}$$

To continue with our proof we are going to used the **Chernoff Bound**:

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with expected value p. Then:

$$Pr\left(\left|\sum_{i=1}^{k} X_i - pk\right| > \delta pk\right) < e^{-\frac{-\delta^2}{4}pk}$$

In our problem we have that $p = \frac{1}{2} + |x - |^{-c}$. We can use $\delta = |x|^{-c}/2$ and $k = 8|x|^{2d+c}$. Therefore, the probability we output the wrong answer is:

$$Pr\left(\frac{1}{|x|}\sum_{i=1}^{k} X_i > \frac{1}{2} + |x - |^{-c}\right) < e^{-\frac{1}{4|x|^{-2c}}\frac{1}{2}8|x|^{2c+d}} \leq 2^{-n^d}$$

□

2 BPP and P\poly

Definition A language L is in P\poly if it can be computed by a family of circuits $C = \{C_1, C_2, \dots\}$, where $|C_i|$ is polynomial in i. C_i accepts exactly the string in L of length i.

Theorem 3 *$L \in BPP \Rightarrow L \in P\backslash poly$*

Proof Let $L \in$ BPP.
By error amplification, $\exists M'(x, r)$, $\forall |x| = n$ and $|r| = m$ $(m > n)$ such that:

$$\forall x, |x| = n \qquad Pr(M'(x, r) \neq L(x)) \leq 2^{-(n+1)}$$

We will say that r is bad for x if $M(x, r) \neq L(x)$.

For every x, the number of bad strings r is less or equal to:

$$\frac{2^m}{2^{n+1}}$$

So there is at most k r that are bad for some x, where:

$$k = 2^n \cdot \frac{2^m}{2^{n+1}} = \frac{2^m}{2}$$

In other words, at least $2^m - \frac{2^m}{2}$ choices of r are good for every x. So lets pick a r that is good for every x of length n, r^*. We can use r^* to create a circuit C for L on inputs of length n that outputs $M(x, r^*)$. Therefore, our circuit C will satisfy $C(x) = L(x)$ for every $x \in \{0, 1\}^n$.

□

Note: We can't have $L \in$ P because we need a different r^* for each x of length n. In other words, we can't find in polytime an r^* that is good for all x.

Theorem 4 *$L \in BPP \Rightarrow L \in \Sigma_2^p$*

Proof Following the previews proof:

$$\begin{aligned} \forall x \in L &\quad \Rightarrow \quad Pr(M'(x, r) = 1) \geq 1 - 2^{-n} \\ \forall x \notin L &\quad \Rightarrow \quad Pr(M'(x, r) = 1) \leq 2^{-n} \end{aligned}$$

Lets fix x. Then M is defining a set S_x where S_x is a set of r such that M accepts (x, r). We have two options:

- $|S_x| \geq (1 - 2^{-n}) \cdot 2^m$ (Huge set)
- $|S_x| \leq 2^{m-n}$ (very small)

We want to distinguish between this two possibilities. To do that we are going to define a *shift*: let $S \subseteq \{0, 1\}^m$ and $u \subseteq \{0, 1\}^m$, then $S + u$ represent the shift of S by u.

Example

$$S = \begin{pmatrix} 1011 \\ 1000 \\ 0110 \\ 1101 \end{pmatrix} \qquad u = [1110] \qquad S + u = \begin{pmatrix} 0101 \\ 0110 \\ 1000 \\ 0011 \end{pmatrix}$$

We are going to use two claims to prove the theorem.

Claim 5 *$\forall S \subseteq \{0,1\}^m$ $|S| \leq 2^{m-n}$ and every k vectors $u_1, \ldots, u_k$ ($k = m/(n+1)$):*

$$\bigcup_{i=1}^{k}(S + u_i) \neq \{0,1\}^m$$

Proof

$$\left|\bigcup_{i=1}^{k}(S + u_i)\right| \leq k \cdot 2^{m-n} = (\frac{m}{n} + 1)2^{m-n}$$

$\square$

Claim 6 *$\forall S \subseteq \{0,1\}^m$ $|S| \geq (1 - 2^{-n})2^m$ $\exists u_1, \ldots, u_k$ ($k = m/(n+1)$) such that:*

$$\bigcup_{i=1}^{k}(S + u_i) = \{0,1\}^m$$

Proof For $r \in \{0,1\}^m$ let B_r be the event that $r \notin \bigcup_{i=1}^{k}(S + u_i)$. We will show that:

$$\forall r\ Pr(B_r) < 2^{-m}$$

Therefore, there exists a vector $u_1, \ldots, u_k$ that is good for all r. Lets write B_r as:

$$B_r = \bigcap_{i=1}^{k} B_r^i$$

where B_r^i is the event that $r \notin S + u_i$, which is equivalent to $r + u_i \notin S$.

For each r we are going to count the total number of $u_1, \ldots, u_k$ that are bad for r.For a random u_i, $r = u_i$ is uniform in $\{0,1\}^m$. So:

$$Pr(r = u_i \in S) \geq 1 - 2^n \qquad \text{because } |S| = (1 - 2^{-n})2^m$$

So $Pr(B_r) \leq (2^(-n))^k$. Summing over all r:

$$\text{number of bad vectors} = 2^{-nk}2^m = 2^{-n(m/n+1)}2^m = 2^n$$

Finally, $\exists u_1, \ldots, u_k$ good for all r. $\square$

Now, with claim 5 and 6 we have that:

$$x \in L \Leftrightarrow \exists u_1, \ldots, u_k \, \forall r \in \{0,1\}^m \left(\bigvee_{i=1}^{k} M(x, r + u_i) \right)$$

$\square$

Summary:

$$L \subseteq NL \subseteq P \subseteq RP \subseteq BPP \subseteq \Sigma_2^p \subseteq PH \subseteq EXP$$