IITM-CS6840: Advanced Complexity Theory January 29, 2012

Lecture No. 12 : Sabbotovskaya's method

Lecturer: Jayalal Sarma M.N. *Scribe: Sajin Koroth*

Theme: Circuit Complexity-Lower Bounds using Restrictions

Lecture Plan:In today's lecture we will be seeing Sabbotovskaya's formula lowerbound method. This method uses random restrictions to prove the existence of a function whose formula size is $\Omega(n^{\frac{3}{2}})$. We will also show Andrev's method which extends Sabbotovskaya's method to prove the existence of an explicit function whose formula size is $\Omega(n^{\frac{5}{2}})$.

1 Sabbotavskaya's lowerbound

Recall that we defined random restrictions in the following way

Definition 1 (Random Restriction on all but k variables)**.** It is the set of all functions δ described as above with the additional requirement that exactly k indices are unassigned.

$$R_k = \{\delta \mid | \{i \mid \delta(i) = *\} |= k\}$$

where $\delta : [n] \to \{0, 1, *\}$ is interpreted as a restriction as follows :

$$\delta(i) \quad = \quad \begin{cases} 0 & \text{assign 0 to } x_i \\ 1 & \text{assign 1 to } x_i \\ * & x_i \text{ is unassigned} \end{cases}$$

Randomly choosing such a function is equivalent to for each index $i \in [n]$ first tossing a coin to decide to leave it unassigned or to fix it, and then if it is decided to fix it, toss another coin and fix the index according to the outcome of the toss. Recall that $L(f)$ denoted the minimum size of a formula computing f and $\mathbb{E}_\delta \left[L\left(f \mid_\delta\right)\right]$ denoted the expected formula size of a formula when hit with a random restriction in R_k. We also saw how to use Sabbotavskaya's method to obtain a lowerbound for PARITY function.

Theorem 2 (Sabbotavskaya (1961))**.**

$$\mathbb{E}_{\delta \in R_k} \left[L\left(f \mid_\delta\right)\right] \leq \left(\frac{k}{n}\right)^{\frac{3}{2}} L(f)$$

Proof. To prove the theorem we will prove the following lemma □

Lemma 3.

$$\mathbb{E}_{\delta\in R_{n-1}}\left[L\left(f\mid_{\delta}\right)\right]\le\left(1-\frac{3}{2n}\right)L(f)$$

Note that the above lemma would imply Sabbotavskaya's theorem. This is because we can apply the theorem repeatedly due to the following observation,

Fact 4.

$$\mathbb{E}_{\delta\in R_{i}}\left[L\left(f\mid_{\delta}\right)\right]\le\left(1-\frac{3}{2n}\right)\mathbb{E}_{\delta\in R_{i+1}}\left[L\left(f\mid_{\delta}\right)\right]$$

This is because there exists a restriction $\delta^{'}\in R_{i+1}$ which achieves $L(f\mid_{\delta'})=\mathbb{E}_{\delta'\in R_{i+1}}\left[L\left(f\mid_{\delta'}\right)\right]$. Note that $f\mid_{\delta'}$ is a function on $i+1$ variables, hence $\delta\in R_i$ can be thought of as restricting just one variable in $f\mid_{\delta'}$, hence allowing us to use the lemma.

Note that $\left(1-\frac{3}{2n}\right)\le\left(1-\frac{1}{n}\right)^{\frac{3}{2}}$. Now Lemma 3 along with Lemma 4 gives

$$\mathbb{E}_{\delta\in R_{k}}\left[L\left(f\mid_{\delta}\right)\right]\le\left(1-\frac{1}{n}\right)^{\frac{3}{2}}\left(1-\frac{1}{n-1}\right)^{\frac{3}{2}}\cdots\left(1-\frac{1}{k+1}\right)^{\frac{3}{2}}L(f)\le\left(\frac{k}{n}\right)^{\frac{3}{2}}L(f)$$

Hence the theorem.

It remains to prove Lemma 3. Before proceeding to the proof let us introduce some notation and assumptions. We will be working with the basis $\Omega=\{\vee_2,\wedge_2,\neg\}$. Without loss of genearlity we will be assuming that all the negation gates are pushed down to the inputs. This can be done with the use of De Morgan's laws without increasing the size of the formula realizing f. Let n_{x_i} denote the number of occurances of x_i or $x_i^{'}$ and let x_i^j denote the j^{th} occurance of the variable x_i.

To do the proof we will be exploiting the following property of the tree corresponding to the minimal formula realizing f. Whenever an input variable x_i appears as an input to a gate $g_j\in\{\vee_2,\wedge_2\}$ it cannot appear as a leaf in the sub-tree rooted at the other input to g_j because if it did we could replace it with a constant without affecting the function computed by the formula and reducing the size thus contradicting the minimality of the formula. The safe replacement depends on the type of gate g_j. For example if $g_j=\vee_2$, then we could replace the occurance of x_i in the sub-tree with 0, this is because when $x_i=0$ the assignment is obviously correct and when $x_j=1$ since g_j is an OR gate its output is 1 irrespective of what the sub-tree rooted at sibling of x_i. Similarly for AND gate you can see that setting the second occurance to 1 does not change the function computed by the formula.

The above mentioned property gurantees that in the sub-tree rooted at sibling of x_i there must be another variable $x_j, j \neq i$ as the function computed by the sub-tree is non-trivial as the formula we assumed to be the minimal one. Hence on one fixing of x_i for each occurancce of x_i at least one occurance of another variable $x_j, j \neq i$ gets killed (if $g_j = \vee$ we set $x_i = 1$ thus fixing the OR gate and if $g_j = \wedge$ we set $x_i = 0$ thus fixing the AND gate). Hence we get that

$$(L(f) - L(f|_{x_i=0})) + (L(f) - L(f|_{x_i=1})) \geq 2n_{x_i} + n_{x_i} = 3n_{x_i}$$

Also note that the leaves of the formula are the variables, hence

$$L(f) \geq \sum_{i=1}^{n} n_{x_i}$$

Combining the above two observations we get that

$$\sum_{i=1}^{n} (L(f) - L(f|_{x_i=0})) + (L(f) - L(f|_{x_i=1})) \quad \geq \quad 3\sum_{i=1}^{r} n_{x_i}$$

Note that

$$\mathbb{E}_{\delta \in R_{n-1}}[L(f)] \quad = \quad \frac{1}{2n} \sum_{1 \leq i \leq n} \sum_{b \in \{0,1\}} L(f|_{x_i=b})$$

because choosing $\delta \in R_{n-1}$ is equivalent to choosing an index uniformly at random from $[n]$ and then fixing it to a value chosen uniformly at random from $\{0, 1\}$.

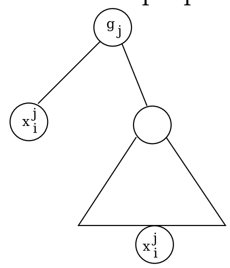
Adding an subtracting $L(f)$ we get,

$$\begin{aligned}
\mathbb{E}_{\delta \in R_{n-1}}[L(f)] &= L(f) - \frac{1}{2n}\frac{1}{2n} \sum_{1 \leq i \leq n} \sum_{b \in \{0,1\}} (L(f) - L(f|_{x_i=b})) \\
&\leq L(f) - \frac{1}{2n} \sum_{i=1}^{n} 3n_{x_i} \\
&\leq L(f) - \frac{3}{2n} L(f) \\
&= \left(1 - \frac{3}{2n}\right) L(f)
\end{aligned}$$

Hence the lemma.

Recall that we have already showed a lower bound for parity which was $\Omega((\frac{n}{4})^{\frac{3}{2}})$. We can use that result to get a better lowerbound for an explicit function

Figure 1: Minimality of formula : A property of the corresponding tree



Theorem 5 (Andreev, 1986)**.** *There is an explicit function that requires* $\Omega\left(n^{\frac{5}{2}}\right)$ *size for any formula*

Proof Sketch:

Consider the function $f : \{0,1\}^{2n} \to \{0,1\}$ which is evaluated as

$$\phi\left(\bigoplus_{i=1}^{\frac{n}{\log n}} x_i, \bigoplus_{i=\frac{n}{\log n}}^{\frac{2n}{\log n}} x_i, \ldots, \bigoplus_{i=\frac{(\log n-1)n}{\log n}}^{n} x_i\right)$$

where ϕ is the function whose truth table is given by the bits $x_{n+1}, \ldots, x_{2n}$. We also know that there exists some function $\phi^{'}$ which requires size at least $\frac{2^n}{n}$. This function need not be explicit but we are guranteed that this functions truth table will come as $x_{n+1}, \ldots, x_{2n}$ for some input setting. Hence on that $\phi^{'}$ on $\log n$ variables we will require a circuit of size $\frac{n}{\log n}$. Also note that to compute the parity inputs for this function we would require, $\log n$ circuits of size $\Omega\left(\left(\frac{n}{4\log n}\right)^{\frac{3}{2}}\right)$. Now we need to argue about random restrictions applied on the formula computing this function and then apply Sabbotavskaya's lowerbound appropriately. The proof needs more technical details. The detailed proof can be found at the reference for this lecture.