ITCS:CCT09 : Computational Complexity Theory	April 15&20, 2009

Lecture 11&12

Lecturer: Jayalal Sarma M.N. *Scribe: Kai Ren*

1 Introduction

In this lecture, we continue to present several results on circuit lowerbound. The first one is to prove the switching lemma and show the circuit lowerbound of $PARITY$ function. Then we show the circuit lowerbound of $PARITY$ implies that $PARITY \notin AC_0$. Before presenting these theorems, we introduce several definitions that are necessary in the proofs:

Definition 1 *A* restriction *on a domain of n variables is a map $\rho : I \to \{0, 1, *\}$ such that $I = \{x_i | 1 \leq i \leq n\}$. Suppose that f is a boolean function with n variables $x_1, \ldots, x_n$. Then f under restriction ρ is defined as $f|_\rho$, which is the result of substituting $\rho(x_i)$ for every variable x_i in f such that $\rho(x_i) \neq *$. We say that all variables x_i such that $\rho(x_i) \neq *$ are* free*, since they are not assigned to any value.*

Definition 2 *Define $\mathcal{R}_n^l$ to be the set of all restrictions ρ on a domain of n variables that leave exactly l variables free , that is, other $n - l$ variables are assigned to either 0 or 1.*

Definition 3 *Consider a DNF formula $F = C_1 \vee \ldots C_k$, and its terms are ordered lexicographically. The* decision tree *for F, $T(F)$ is defined inductively as the following:*

1. If F is the constant function 0 or 1, then $T(F)$ is just a single leaf node with corresponding value 0 or 1.

2. If the first term C_1 of F is not empty, then let F' be the remainder of F so that $F = C_1 \vee F'$. Let K be the set of variables appearing in C_1. The tree $T(F)$ starts with a complete binary tree for K such that at the i'th level we query the i'th variable of K, and proceed left if it is 1 and right if it is 1. Each leaf v_ρ in the tree is associated with a restriction ρ which sets the variables of K according to the path from the root to v_ρ. For each ρ we replace the leaf node, v_ρ, by the subtree $T(F|_\rho)$. (Note that for the unique ρ which satisfies C_1 the leaf v_ρ will remain a leaf and be labeled 1. For all other choices of ρ, the tree that replaces v_ρ is $T(F|_\rho) = T(F'|_\rho)$.

2 The circuit lowerbound of PARITY

Theorem 4 *Any boolean circuits of depth d computing PARITY must have size* $S \geq 2^{\frac{n^{1/(d-1)}}{14}}$.

The proof of this theorem is based on the following lemma, which will be proved in the next section.

Switching Lemma: Let $F = C_1 \vee C_2 \vee \cdots \vee C_k$ be a DNF with terms of size $\leq r$. Let $l = \epsilon n$, for $0 < \epsilon \leq \frac{1}{7}$. Pick $\rho \in R_n^l$ at random, then $Pr[F|_\rho$ does not have a decision tree of height $\leq h] < (7\epsilon r)^h$.

Claim 5 *Let C be an AND/OR circuit of depth d and size S. Let h be given and define $n_d = \frac{n}{14(14h)^{d-1}}$. Choose $\rho \in R_n^{n_d}$ at random, then with probability $1 - S2^{-d}$ every function computed at every gate of C has a decision tree of depth at most h after using ρ.*

Proof First, ρ is chosen at random in an alternative way. Define $n_{i+1} = \frac{n}{14(14h)^i}$ for $0 \leq i \leq n-1$ and $n_0 = n$. Then choose ρ by choosing $\rho_1\rho_2\ldots\rho_d$, where $\rho_i \in R_{n_{i-1}}^{n_i}$ for $1 \leq i \leq n-1$.

We show that for each gate the probability that the corresponding decision tree has depth greater than h, given that its input gates have decision trees of depth at most h, is less than 2^{-h}, and the statement then follows by summing over all gates.

For a given gate, we proof it by the induction on the depth of the gate. As the base case, consider an OR gate at level 1. This can be viewed as a DNF with terms of size 1, meaning that we can apply the switching lemma. Thus, when picking a restriction $\rho_1 \in R_{n_0}^{n_1}$ at random, we get that:

$$Pr[F_{\rho_1} \text{ does not have a decision tree of depth at most } h] < (7 \cdot \frac{1}{14} \cdot 1)^h = 2^{-h}$$

In the case of AND gate at level 1, the similar result can be got from the decision tree of the negation.

For the induction step, all gates at levels 1 to i have decision trees of depth $\leq h$ after using $\rho_1 \ldots \rho_i$.

Consider an OR gate at level $i+1$. Its inputs have decision trees of depth $\leq h$, which can be rewritten to DNF's with terms of size $\leq h$. Since each root-to-leaf path in the decision trees can be expressed as a term of DNF with at most h variables. Now the OR gate at level $i+1$ has only OR gates as inputs. If all OR gate are collapsed into one OR gate, then the circuit turns into a DNF with terms of size at most h (see Figure 1).

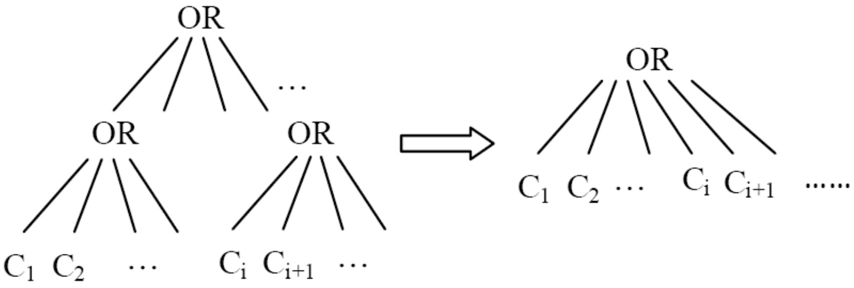

Figure 1: Collapse all OR gates into one.

If picking $\rho_{i+1} \in R_{n_i}^{n_{i+1}}$ at random, by switching lemma, we have: $\epsilon = \frac{n_{i+1}}{n_i} = \frac{1}{14h}$

Thus,

$$Pr[F_{\rho_1 \cdots \rho_{i+1}} \text{ does not have a decision tree of depth at most } h] < (7 \cdot \frac{1}{14h} \cdot h)^h = 2^{-h}$$

Similar results can be achieve for the AND gate by negating the expression. ■

Proof (Proof for Theorem 4) Given a circuit C of depth d and size S computing PARITY. Let $h = \log S$. Assume that the topmost gate is an OR gate. According to the proof of Claim 5, there exists a $\rho \in R_n^{n_{h-1}}$ such that the input gates of the topmost OR gate have decision trees of depth at most h after applying ρ. Then the circuit after applying ρ can be expressed as a DNF formula F with terms of size at most h.

However, for a PARITY function of n_{d-1} variables, its DNF formula F requires terms of size n_{d-1}. Since if one of the term has less than n_{d-1} variables, then the variable can be set to either 0 or 1 when finding a restriction that satisfies this term, which cannot be the case. Therefore, it should have

$$\begin{aligned}
& h \geq n_{d-1} = \frac{n}{14(14h)^{d-2}} \\
\Rightarrow \quad & (14h)^{d-1} \geq n \\
\Rightarrow \quad & h \geq \frac{1}{14} n^{\frac{1}{d-1}} \\
\Rightarrow \quad & S \geq 2^{\frac{n^{\frac{1}{d-1}}}{14}}
\end{aligned}$$

■

Claim 6 *PARITY* $\notin AC^0$

Proof By the proof of Theorem 4, a circuit computing PARITY for n input variables and of constant height d requires size $S \geq 2^{\frac{n^{\frac{1}{d-1}}}{14}} = 2^{\Omega(n)}$, and is not polynomial in size. Therefore, PARITY $\notin AC^0$. ■

3 Switching Lemma

Definition 7 *Define* $Stars_k(r,h)$ *to be the set of* k *sequences* $(\beta_1, \ldots, \beta_k)$ *such that for every* j*,* $\beta_j \in \{*, -\}^r \backslash \{-\}^r$ *and the total number of *'s over all the* β_j*'s is* h*.*

Lemma 8 $|Stars_k(r,h)| < (\frac{r}{ln2})^h$

Proof Define α by $(1+\frac{1}{\alpha})^r = 2$. Then $\ln(1+1/\alpha) = \frac{\ln 2}{r}$. By using $1+x < e^x$ for $x \neq 0$, we get

$$\frac{\ln 2}{r} = \ln(1+1/\alpha) < \ln(e^{1/\alpha}) = \frac{1}{\alpha}$$

Thus,

$$\alpha < \frac{r}{\ln 2}$$

We use induction on h to prove that $|Stars_k(r,h)| < \alpha^h$. In the base case, it is trivial that $|Stars_k(r,0)| < \alpha^0$. For the induction part, assume that for all $h < k$, the inequality holds. Consider that β_1 has i *'s. Then the number of possible values for β_1 is $\binom{i}{r}$. We then get:

$$\begin{aligned}
|Stars_k(r,h)| &= \sum_{i=1}^{\min(r,h)} \binom{i}{r} Stars_{k-1}(r,h-i) \\
&\leq \sum_{i=1}^{\min(r,h)} \binom{i}{r} Stars_k(r,h-i) \\
&< \sum_{i=1}^{r} \binom{i}{r} \alpha^{h-i} \\
&= \alpha^h \sum_{i=1}^{r} \binom{i}{r} (1/\alpha)^i \\
&= \alpha^h[(1+1/\alpha)^r - 1] \\
&= \alpha^h
\end{aligned}$$

■

Lemma 9 *(Switching Lemma) Let $F = C_1 \vee C_2 \vee \cdots \vee C_k$ be a DNF with terms of size $\leq r$. Let $l = \epsilon n$, for $0 < \epsilon \leq \frac{1}{7}$. Pick $\rho \in R_n^l$ at random, then $Pr[F|_\rho$ does not have a decision tree of height $\leq h] < (7\epsilon r)^h$.*

Proof Let S be the set of restriction in R_n^l such that for $\rho \in S$, $F|_\rho$ doesn't have a decision tree of height h. Since The probability we want to bound equals to $|S|/|R_n^l|$, we first obtain a bound on $|S|$ by defining a 1-1 map from S to a small set.

We will define a 1-1 map $S \to H$, where $H = R_n^{l-h} \times Stars_k(r,h) \times \{0,1\}^h$. Given some $\rho \in S$, and let π be the restriction corresponding to the first h variables of lexicographically first path in $T(F|_\rho)$ that has length $\geq h$. We use the formula F and π to determine the image of ρ.

Let C_{v_1} be the first term of F that is not set to 0 by ρ, that is the first term of $F|_\rho$. And let π_1 be the part of π in C_{v_1}. Also, let σ_1 be the unique restriction satisfying C_{v_1} on the variables of π_1. For $i > 1$ let C_{v_i} be the first term of $F|_{\rho\pi_1\ldots\pi_{i-1}}$, and let π_i be the part of π in C_{v_i}. Also, let σ_i be the unique restriction satisfying C_{v_i} on the variables of π_i. Note that π_i may not restrict every variable of C_{v_i}, since π has only restricted h variables and the height of $T(F|_\rho)$ may be higher than h. Thus, we have $\pi_1\pi_2\ldots\pi_k = \pi$. The relation between these notions and $T(F|_\rho)$ is shown in Figure 2.

Before defining the 1-1 map $S \to H$, some notations should be defined. For every $i = 1,\ldots,k$, let the j'th component of β_i be $*$ if and only if the j'th variable in C_{v_i} is set by σ_i. Also, define $\delta \in \{0,1\}^h$ to be the bit-string for which the i'th bit is 1 if and only if π and $\sigma_1\ldots\sigma_k$ agree on the i'th variable. By the above notions, we get the 1-1 map such that for every $\rho \in S$, $\rho \mapsto (\rho\sigma_1\ldots\sigma_k, (\beta_1,\ldots,\beta_k), \delta)$. Note that $\rho\sigma_1\ldots\sigma_k \in R_n^{l-h}$.

We have to argue the mapping from H to S that recovers ρ from $\rho\sigma_1\ldots\sigma_k, (\beta_1,\ldots,\beta_k), \delta$. The reconstruction is iterative. Suppose that we have recovered $\pi,\ldots,\pi_{i-1},\sigma_1,\ldots,\sigma_{i-1}$, and $\rho\pi_1\ldots\pi_{i-1}\sigma_i\ldots\sigma k$. Notice that for $i < k$, $C_{v_i}|\rho\pi_1\ldots\pi_{i-1}\sigma_i = 1$ and $C_j|\rho\pi_1\ldots\pi_{i-1}\sigma_i = 0$ for all $j < v_i$. Thus we can recover v_i as the index of the first term of F that is not set to 0 by $\rho\pi_1\ldots\pi_{i-1}\sigma_i\ldots\sigma k$.

Now, using C_{v_i} and β_i, we know those variables in C_{v_i} that are only set by σ_i. Hence we get σ_i. Then by using δ and σ_i, we can imply π_i. We repeat the whole procedure until we find $\pi_1,\ldots,\pi_k$ and $\sigma_1,\ldots,\sigma_k$. Then we can easily reconstruct ρ by removing the restriction of $\pi_1,\ldots,\pi_k$ from $\rho\pi_1\ldots\pi_k$.

With this mapping, we have shown that $|S| < |H|$, and $|H| \leq |R_n^{l-h}| \cdot |Stars_k(r,h)| \cdot 2^h$. Therefore,

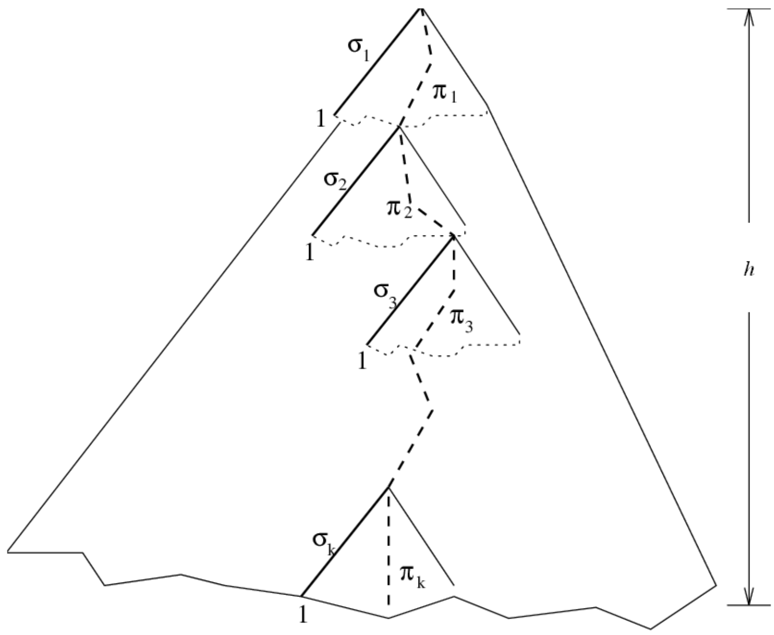

Figure 2: Decision Tree $T(F|_\rho)$

$$
\begin{aligned}
\frac{|S|}{R_n^l} &\leq \frac{|R_n^{l-h}|}{R_n^l} \cdot |Stars_k(r,h)| \cdot 2^h \\
&\leq \frac{|R_n^{l-h}|}{R_n^l} \cdot (\frac{2r}{\ln 2})^h \\
&\leq \frac{\binom{n}{l-h} 2^{n-l+h}}{\binom{n}{l} 2^{n-l}} \cdot (\frac{2r}{\ln 2})^d \\
&\leq \frac{l^h}{(n-l)^h} \cdot (\frac{4r}{\ln 2})^h \\
&= (\frac{4\frac{l}{n}r}{(1-\frac{l}{n})\ln 2})^h \\
&= \frac{4\epsilon r}{(1-\epsilon)\ln 2})^h \leq (\frac{4\epsilon r}{\frac{6}{7}\ln 2})^h < (7\epsilon r)^h
\end{aligned}
$$

■