CS 810: Introduction to Complexity Theory 03/27/2003

Lecture 18: Circuit Lower Bounds for Parity.

Instructor: Jin-Yi Cai Scribe: Matthew Lee, Yingchao Liu, Uchechukwu Okpara

In this lecture we improve the bounds in switching lemma. Using the lemma, we establish lower bounds on size of circuits computing or approximating parity.

1 The Switching Lemma

In the last lecture we proved the following lemma.

Lemma 1. *Let G be a t-And-Or formula $G_1 \wedge G_2 \wedge \cdots \wedge G_w$. For any β, $0 < \beta < t$, let ρ be a random p-restriction, where $p = \frac{\beta}{t-\beta}$, and let $\alpha = \beta / ln \left[\frac{1+\sqrt{1+4e^{\beta}}}{2} \right]$. Then for all $\Delta \geq 0$, we have*

$$\Pr[DC(G|\rho) \geq \Delta] \leq \alpha^{\Delta}.$$

In the above lemma, α is minimized when $\beta = \beta_0 \approx 0.227537$. In which case, $\alpha = \alpha_0 \approx 0.4164447$. Let $\gamma_0 = \beta_0/2 \approx 0.1137685$.

Using Lemma 1, we will prove the Lemma 2, a stronger version of switching lemma. The key is the the following composite property of random restrictions. Observe that a p_1-restriction followed by a p_2-restriction has the same effect with a single p_1p_2-restriction. This property holds because the boolean variables are independently assigned at each step.

Lemma 2. *Let G be a t-And-Or formula $G_1 \wedge G_2 \wedge \cdots \wedge G_w$, and let ρ be a random γ_0/t-restriction. Then for all $\Delta \geq 0$, we claim*

$$\Pr[DC(G|_{\rho}) \geq \Delta] \leq {\alpha_0}^{\Delta}$$

Proof. Let $q = \beta_0/t$ and $p = \frac{q}{2-q}$. Then $q = \frac{2p}{1+p} = \frac{p}{p+\frac{1-p}{2}}$ is the probability a variable is assigned a * in a random p-selection under the condition that it is assigned * or 0.

We have shown that

$$\Pr[DC(G|_{\rho\prime}) \geq \Delta] \leq {\alpha_0}^{\Delta},$$

where $\rho\prime$ is a random p-restriction.

Since $p = \frac{q}{2-q} > \frac{q}{2} = \frac{\gamma_0}{t}$, we know that $\gamma_0/(pt)$ is still a number less than 1. Because of the composite property of random restriction, a random γ_0/t-restriction ρ can be realized by first applying a random p-restriction $\rho\prime$, followed by a $\frac{\gamma_0}{pt}$-restriction. Note that $DC(G|\rho\prime) < \Delta$ means there exists a decision tree with depth less than Δ, which can compute G. Therefore we would also be able to compute G within less than Δ depth with the more stringent restriction ρ. That is to say,

$$\Pr[DC(G|_{\rho}) \geq \Delta] \leq \Pr[DC(G|_{\rho\prime}) \geq \Delta] \leq {\alpha_0}^{\Delta}.$$

□

2 Circuit Lower Bounds

Consider general constant depth circuits. Denote by $C^d(s,t)$ the class of depth d circuits with bfi (the abbreviation of bottom fanin) $\leq t$, and the number of gates above the first level $\leq s$. Denote by $C^d(s)$ the class of depth d circuits without a bfi condition but with total size $\leq s$. It is clear that a circuit in $C^d(s)$ can be considered as a circuit in $C^{d+1}(s,1)$, by adding an extra layer of gates with fan-in 1.

The following lemma is proved using the switching lemma. In Lecture 16, we discussed on how to use the switching lemma to prove circuit lower bounds. We use the same ideas here and hence, we provide only a sketch of the proof.

Lemma 3. *For all $C \in C^d(s, \gamma_0 n^{1/d})$, we have*

$$\Pr[DC(C|_{\rho}) \geq \gamma_0 n^{1/d}] \;\leq\; s \cdot \alpha_0^{\gamma_0 n^{1/d}} \;\approx\; s \cdot 2^{-0.143781 \cdot n^{1/d}},$$

where ρ is a random $1/n^{\frac{d-1}{d}}$-restriction.

Proof. Let $t = \gamma_0 n^{1/d}$ and $p = 1/n^{\frac{1}{d}}$. Let $C \in C^d(s,t)$. Denote the number of gates on each level as $s_1, s_2, \ldots, s_d$ from the bottom to the top (output level). Clearly, $s_d = 1$ and $\sum_{i=1}^{d-1} s_i = s$. Apply Lemma 2 repeatedly $d-1$ times, each time with a random p-restriction.

After applying the first random restriction, suppose all the s_1 many t-AND-OR circuits at level 1 have DC $\leq t$. Then, we can switch these into t-OR-AND circuits and merge first and second levels eliminating a level. This process also eliminates s_1 gates from the circuit. Nonetheless, for each gate at the bottom level, it is possible that we cannot switch. By Lemma 2, for any one of the s_1 circuits, probability that we fail to to switch is $\leq \alpha_0^t$. So the probability of failure at this level is at most $s_1 \cdot \alpha_0^t$. Accumulating the probability of the failure on each level, and combining with the Lemma 2, we have

$$\Pr[DC(C|_{\rho}) \geq \gamma_0 t] \;\leq\; \sum_i s_i \cdot \alpha_0^t = s \cdot \alpha_0^t = s \cdot \alpha_0^{\gamma_0 n^{1/d}} \;\approx\; s \cdot 2^{-0.143781 \cdot n^{1/d}}$$

Finally, by the composite property of random restrictions, applying $d-1$ random p-restrictions has the same effect as one random p^{d-1}-restriction. The proof is complete. □

We can apply Lemma 3 to $C \in C^d(s)$ by first transforming them into $C \in C^{d+1}(s,1)$. But we can actually do slightly better by a more delicate technique. Here we omit the proof of the following better bound.

Lemma 4. *For all $C \in C^d(s)$, we have*

$$\Pr[DC(C|_\rho) \geq \gamma_0 n^{1/d}] \ < \ s \cdot \alpha_0^{\gamma_0 n^{1/d}} \ \approx \ s \cdot 2^{-0.143781 \cdot n^{1/d}},$$

where ρ is a random $\alpha_0/(2n^{\frac{d-1}{d}})$-restriction.

These results can be used to prove circuit lower bounds for the parity function. Consider any circuit C in $C^d(s, \gamma_0 n^{1/(d-1)})$. Apply $d-2$ rounds of random $1/n^{1/(d-1)}$-restrictions. With probability $> 1 - s \cdot 2^{-0.143781 \cdot n^{1/(d-1)}}$, we get a circuit in $C^2(1, \gamma_0 n^{1/(d-1)})$ after switching and merging. The process is equivalent to applying a single random $n^{(d-2)/(d-1)}$-restriction. Let N be the random variable for number of variables left (i.e., variables assigned *). Then, its expectation $E[N] = n^{1/(d-1)}$. By Chernoff bound we have,

$$\Pr[N \leq \gamma_0 n^{\frac{1}{d-1}}] < e^{-\frac{(1-\gamma_0)^2}{2} \cdot n^{\frac{1}{d-1}}} < e^{-0.3927 n^{\frac{1}{d-1}}}.$$

Hence, if $s < 2^{0.143781 \cdot n^{1/(d-1)}}$, the probability is approaching 1 that both C is reduced to a circuit in $C^2(1, \gamma_0 n^{1/(d-1)})$ and $N > \gamma_0 n^{1/(d-1)}$. Suppose the circuit we started with computes parity on n variables. Then, the circuit obtained after applying the random restriction computes parity on the remaining N variables. Clearly, a $C^2(1,t)$ circuit cannot compute parity on $> t$ variables (see Lecture 16). We have proved the following lemma.

Lemma 5. *For all $C \in C^d(s, \gamma_0 n^{1/(d-1)})$, if C computes the parity function, then its size s must satisfy*

$$s \ \geq \ 2^{0.143781 \cdot n^{1/(d-1)}}.$$

Lemma 5 can be used to obtain lower bounds for general circuits (without bfi). Again, we can simply transform a circuit in $C^d(s)$ into a circuit in $C^{d+1}(s,1)$, then apply Lemma 5. Using a more direct and finer analysis, one can prove the following lemma.

Lemma 6. *For all $C \in C^d(s)$, if C computes the parity function, then its size s must satisfy*

$$s \geq 2^{0.143781 \cdot n^{\frac{1}{d-1}}}$$

3 Inapproximability Type Lower Bounds

Now we consider the inapproximability type lower bound. By *inapproximability*, we mean circuits with certain restrictions (like size, depth and bfi) cannot compute parity on significantly more than half of the possible cases. Specifically, for the parity function, we can simply guess 0 and 1 as the function value. So, it is easy to get 50% success. We want to show that, one cannot do significantly better. The decision tree depth lower bound is ideally suited for deriving the inapproximability type lower bound, and the decision tree perspective was introduced precisely for this reason. Our goal is to show that, when the sizes of the circuits are below some lower bound, the circuits will make asymptotically 50% error on all possible inputs.

Let C be a depth d circuit. Note that after some restriction ρ, if C is reduced to a decision tree of depth smaller than the number of variables left, then for exactly half of the 0-1 extensions of ρ, C agrees on the parity. This is because at every leaf of the decision tree, the circuit C is completely determined.

Consider Pr[$C(x_1, \dots, x_n) = \oplus(x_1, \dots, x_n)$], where $\oplus(x_1, \dots, x_n)$ denotes the parity function, and the probability is over all 2^n assignments. This random restriction technique can be realized by first assigning any random restriction, followed by an unbiased 0-1 assignments for all the remaining variables. Let E_1 denote the event that after the random restriction, we end up with a decision tree of depth not more than t, and let E_2 denote the event that the number of variables N assigned to * is more than t. Then let $E = E_1 \wedge E_2$, and let $[\,C = \oplus\,]$ denote $[\,C(x_1, \dots, x_n) = \oplus(x_1, \dots, x_n)\,]$ for convenience. As we already pointed out, $\Pr[\,C = \oplus \,|\, E\,] = 1/2$ due to a property of the parity function.

Expending in terms of conditional probabilities, we have

$$\begin{aligned}
\Pr[C = \oplus] &= \Pr[E] \cdot \Pr[C = \oplus | E] + \Pr[\neg E] \cdot \Pr[C = \oplus | \neg E] \\
&= (1 - \Pr[\neg E]) \cdot \Pr[C = \oplus | E] + \Pr[\neg E] \cdot \Pr[C = \oplus | \neg E] \\
&= \Pr[C = \oplus | E] + \Pr[\neg E](Pr[C = \oplus | \neg E] - \Pr[C = \oplus | E]).
\end{aligned}$$

As we noted, $\Pr[\,C = \oplus \,|\, E\,] = 1/2$, and $\Pr[C = \oplus | \neg E] \leq 1$. Then substitute these 2 observations into the above equation,

$$\left| \Pr[C = \oplus] - \frac{1}{2} \right| \;\leq\; \frac{1}{2} \Pr[\neg E].$$

Since $\Pr[C = \oplus] + \Pr[C \neq \oplus] = 1$, we have

$$\Pr[\,C = \oplus\,] - \Pr[\,C \neq \oplus\,] = 2\left(\Pr[\,C = \oplus\,] - \frac{1}{2}\right),$$

and hence

$$|\Pr[\,C = \oplus\,] - \Pr[\,C \neq \oplus\,]| \;\leq\; \Pr[\neg E].$$

Now we specify the parameters of the random restrictions. Let $m = \gamma_0 n^{1/d}$. First consider any $C \in C^d(s, \gamma_0 m)$. Let $t = \gamma_0 m$ and apply Lemma 3. With a random $1/n^{(d-1)/d}$-restriction, we have

$$\Pr[\neg E_1] \leq s\alpha_0^t \approx s\dot{2}^{0.143781 \cdot m}.$$

Again by using the Chernoff bound, we estimate $\Pr[\neg E_2] = \Pr[N \leq \gamma_0 m]$ as follows.

$$\Pr[\neg E_2] \leq e^{-\frac{(1-\gamma_0)^2}{2} m} < e^{0.3927m}.$$

Thus $\Pr[\neg E_2]$ is dominated by $\Pr[\neg E_1]$. This analysis gives the following bound.

Lemma 7. *For all* $C \in C^d(2^{0.07189n^{1/d}}, \gamma_0 n^{1/d})$*, we have*

$$|\Pr[\,C = \oplus\,] - \Pr[\,C \neq \oplus\,]| \;\leq\; 2^{-0.07189n^{1/d}}.$$

Again straightforward application of the above lemma gives inapproximability results for general circuits (without bfi). A more careful analysis leads to the following lemma.

Lemma 8. *For all circuits* $C \in C^d(2^{0.07189n^{1/d}})$*, we have*

$$|\Pr[C = \oplus] - \Pr[C \neq \oplus]| \leq 2^{0.07189n^{1/d}}.$$