6.841 Advanced Complexity Theory February 16, 2005

Lecture 5

Lecturer: Madhu Sudan *Scribe: Victor Chen*

1 Administration

Reminder to all that problem set 1 is due today. And please sign up for scribing and get on the class mailing list. Note that we meet on next Tuesday instead of Monday.

2 Overview

Today we show that PARITY is not in AC_0. AC_0 is a family of circuits with constant depth, polynomial size, and unbounded fan-in for the AND and OR gates. We establish this result through an application of the Switching Lemma. This result is the first use of randomization in its full power in complexity.

Circuits were defined in previous lectures. In this lecture, we always assume that the circuits are organized into alternating levels of AND and OR gates. We can make such an assumption since we can convert circuits into this convenient form with only a constant factor of blowup.

The Switching Lemma is first proved by Furst, Saxe, and Sipser in FOCS 81, and readers can find the paper in the Journal of Mathematical Systems Theory 1984. We will highlight their work by using the Lemma though the version we prove today will not be as strong as we claimed in the last lecture. Johan Håstad, in 1986, proved a more general and powerful form of the Lemma, and interested readers can find it in his PhD thesis at MIT. There is also a survey on the Lemma written by Paul Beame.

3 The Switching Lemma

Let us review some terminology. A *term* is an AND of literals, and a *clause* is an OR of literals. We say a formula is a k-DNF if it is an OR of terms with each term having at most k literals, and we say a formula is a ℓ-CNF if it is an AND of clauses with each clause having at most ℓ literals. And we define the size of a formula to be the number of literals in it. Now if a formula has variables $x_1, \ldots, x_n$, then a p-restriction is a function that sets each x_i independently to be 0 with probability $\frac{1-p}{2}$, 1 with probability $\frac{1-p}{2}$, and $*$ with probability p. A restriction fixes some variables, and those assigned a $*$ remain variables.

Here is a rough version of what the Switching Lemma says.

Lemma 1 *Given a k-DNF formula of size s on variables $x_1, \ldots, x_n$, a p-restriction yields a ℓ-CNF formula with size s' with probability $1 - \delta$, provided k, ℓ, s, s', p, and δ are appropriately chosen.*

Håstad's work shows that we can have s arbitrary and s' uncontrolled. If k is small, say $2^{(\log n)^{\epsilon}}$, then ℓ is also small, say $2^{(\log n)^{2\epsilon}}$. We can pick $p \approx O(1/k)$ which gives failure probability when $\delta = (7kp)^{\ell}$. However, this is not the result we will prove. Instead, in Furst, Saxe, and Sipser's version, $s = \mathrm{poly}(n), p = n^{-2/3}, k = \infty, \ell = O(1)$, and $\delta = 1/\mathrm{poly}(n)$. We quantify the parameters by asserting that for all s and δ, there exists ℓ and s' such that the Lemma holds.

3.1 PARITY $\notin \mathrm{AC}_0$

So why does the Switching Lemma imply that PARITY is not in AC_0? Say there exists a circuit of size $s = \mathrm{poly}(n)$ computing parity of n bits, and it has minimal depth d. Hit the circuit with a p-restriction, where $p = n^{-2/3}$. Say the first level of the circuit consists of OR gates. The probability that the circuit defined by the first OR gate does not have a ℓ-CNF formula of size s' is at most δ. Let $\delta = \frac{1}{10s}$. Then by the Union bound, the probability that there exists an OR gate at level 2 that does not get switched is at most $s\delta = 1/10$. This implies that with high probability, there exists a circuit whose first level is now AND and the second level is now OR. The ordering is switched now. By merging the second level with the third, we have circuits of size at most $ss' = \mathrm{poly}(n)$ and depth at most $d-1$. If a function computes parity, then its restriction is still a parity function. So this contradicts the minimality of d.

Does this technique also work for some other functions, say MAJORITY? Yes, a similar argument with induction also work. Alternately, one can also show this by reducing PARITY to MAJORITY with constant depth, polynomial size circuits.

3.2 Proof Sketch of the Lemma

The proof proceeds in three stages. In the first stage, we show that there exists a restriction such that each of the AND gates at level 1 have constant fan-in. Then in the second stage, we apply a second restriction so that the OR gates at the second level depend on only a constant number of input bits. Then in the third stage, we argue that we can switch the ordering of the two levels.

Stage 1:
Hit the circuit with a $\sqrt{p}$ restriction, i.e., x_i is set to $*$ with probability $\sqrt{p}$, 1 with probability $(1-\sqrt{p})/2$ and 0 with probability $(1-\sqrt{p})/2$.

Claim: With probability $1 - \frac{1}{\mathrm{poly}(n)}$, the DNF formula is a c-DNF.
Fix an AND gate. Consider two cases.
Case 1: The fan-in is at least $10c\log s$.
Since the gate has a lot of inputs, we should expect that at least one of them is assigned 0 by the restriction. So if after the restriction, the gate still has fan-in at least c, then it has no input of 0. The probability that the gate is not zero is at most

$$(\frac{1-\sqrt{p}}{2})^{10c\log s} \leq \frac{1}{100\mathrm{poly}(s)}.$$

Case 2: The fan-in is at most $10c\log s$.
If the fan-in is at least c after restriction, then at least c of inputs are assigned $*$. But since the original fan-in is small, by our choice of restriction, this should happen with small probability. More specifically, the probability that i variables out of $10c\log s$ are unset by the restriction is at most

$$\binom{10c\log s}{i}(\sqrt{p})^i \leq \frac{1}{\mathrm{poly}(s)}.$$

We also need to make sure that the number of unset variables is not too small, since otherwise the circuit would be constant. But this occurs with small probability by Chernoff's bound.

Hence, for a sufficiently large circuit, there exists some restriction such that the induced circuit computes PARITY, is polynomial sized, and every gate in the first level has constant fan-in.

Stage 2:
Now we assume that the circuit is polynomial size, has depth d, and the gates in the first level have fan-in at most c. We want to show that there exists a random restriction such that in the induced circuit, the gates in the second level depend on at most b_c inputs, where b_c is a constant. Again hit the circuit with a $\sqrt{p}$-restriction. We argue by induction that for every c, there exists b_c such that the probability some second level gate depends on more than b_c inputs is small.

Fix a second level gate and call it A. When $c = 1$, A is simply the OR of variables. A dual argument analogous to the one in Step 1 will show that it has small fan-in with high probability. For the induction, we consider two cases, one in which A is wide, i.e., has $O(\log s)$ terms that are disjoint. In other words, it has at least $O(\log s)$ first level gates connecting to A such that they all have disjoint inputs. And the other case is when A is narrow.

Case 1: A is wide.
Call the disjoint terms $T_1, \ldots, T_m$. If after the restriction A depends on a lot of inputs, then it is not set to 1. Since A is an OR gate, this implies all terms are not set to 1. But this probability is small. More specifically, $\Pr(T_i = 1 \text{ after restriction}) \geq (1/3)^c$. So the probability that there exists a $j \in \{1, \ldots, m\}$ such that $T_j = 1$ is at least

$$1 - (1 - (1/3)^c)^m = 1 - \frac{1}{\text{poly}(s)}.$$

Case 2: A is narrow.
A has at most $O(\log s)$ disjoint terms. Let H be the set of inputs appearing in a maximal collection of disjoint terms of A. So $|H| = O(c \log s)$, and H intersects with each term of A. Let H' be the set of variables that are left unset (assigned a $*$ by the restriction) in H. With an argument similar to the one in Step 1 Case 2, one can show that with probability at least $1 - 1/\text{poly}(s)$, $|H'|$ is at most a constant c'. Now, there are $2^{c'}$ possible assignments to the variables in H'. Consider an arbitrary assignment ρ_i. Since H intersects with each term of A, ρ_i will induce each term of A to have size at most $c - 1$. By the inductive hypothesis, the probability that A induced by ρ_i depends on more than b_{c-1} inputs is small. Set $b_c = c' + 2^{c'} b_{c-1}$. So if A depends on more than b_c inputs, then either $|H|' > c'$, or at least one assignment to the variables in H' causes the induced A to depend on more than b_{c-1} inputs. Both of these events are small. This completes the proof of induction.

Stage 3:
We have shown that the gates in the second level depend on only a constant number of inputs. Now simply switch the order of AND and OR in these two levels using the distributive law. The blowup maybe be exponential in the number of dependent inputs, but this is still a constant.

4 Conclusion

The general form of the Switching Lemma due to Håstad is not proved here. The argument is more complicated; when a circuit is hit with a restriction, the observation of some terms being canceled introduces bias. But Håstad showed that conditioning on these events only help the outcome. Interested readers are referred to Håstad's PhD thesis. Next time, we will cover another proof that shows PARITY $\in \mathrm{AC}_0$. Instead of using the probabilistic method, the proof uses algebraic technique.

References

[1] Merrick Furst, James B. Saxe, and Michael Sipser. Parity, Circuits, and the Polynomial-Time Hierarchy. Mathematical Systems Theory **17**, 13-27 (1984).