6.841 Advanced Complexity Theory Feb 21, 2007

Lecture 5

Lecturer: Madhu Sudan *Scribe: Salman Abolfathe*

1 Overview

- Bounded depth circuits $= AC^0$.
- PARITY$\notin AC^0$.
- Switching lemma $DNF \to CNF$

2 Introduction

AC^i is the class of languages recognized by a circuit of polynomial size and depth $O(\log^i n)$, that can use gates

$$\{\infty - AND, \infty - OR, NOT\}.$$

So for AC^0 circuits should have polynomial size and $O(1)$ depth. Loosely speaking, depth of a circuit shows the amount of parallel time, and size is the amount of work needed for computing.

Through this lecture we assume that the circuits are organized into alternating levels of AND and OR gates. In fact, all NOT gates can be pushed to the first level, and since here we use $\infty - AND$ and $\infty - OR$ gates instead of usual AND and OR gates, we can combine consecutive AND and OR levels. So each circuit by a constant blowup can be organized such that the first level is NOT gate, and the others are alternating AND and OR. In a such circuit, we don't care about the first level (NOT gates) and say the depth is number of AND and OR levels.

As an example the following language, for a constant k, is in AC^0

$$T_{k,n}(x_1, x_2, \dots x_n) = \begin{cases} 1 \text{ if } \sum_i x_i \geq k \\ 0 \text{ otherwise} \end{cases}$$

In fact, $T_{k,n}(x_1, x_2, \dots x_n) = \bigvee_A \bigwedge_{x_i \in A} x_i$, where A ranges over all subsets of $\{x_1, x_2, \dots x_n\}$ of size k.

The goal of this lecture is to show PARITY is not in AC^0. By PARITY we mean

$$PARITY(x_1, x_2, \dots x_n) = \sum_i x_i \ (mod 2).$$

3 Random Restriction

In the last session we studied the idea of restriction of some variables to get a bound on the size of a circuit that can compute a certain language. Here we use another version of this idea, called *random restriction*. Namely, if $x_1, x_2, \dots x_n$ are variables, for each i, with probability q we don't restrict x_i, and restrict it as $x_i = 0$ or $x_i = 1$, each of which with probability $\frac{1-q}{2}$.

$$x_i = \begin{cases} x_i \text{ with prob. } q \\ 0 \text{ with prob. } \frac{1-q}{2} \\ 1 \text{ with prob. } \frac{1-q}{2} \end{cases}$$

and we repeat it for each i independently. Therefore, if we have a function f that can be computed by circuit C, after this restriction, say $\rho(q)$, we get to the function $f|_\rho$ that can be computed by $C|_\rho$, which is hopefully is a simpler circuit.

In our case, the PARITY problem, after restricting some variables we get to the same problem, i.e. PARITY over remaining variables.

$$\text{PARITY over } n \text{ variables} \longrightarrow \text{PARITY over } qn \text{ variables .}$$

4 Switching Lemma

Consider a depth 2 circuit. Then it is either a CNF or DNF formula. By random restriction idea we want to convert a DNF formula to a CNF formula. In general, if we want to write a DNF as a CNF formula we may get to an exponential one (it is because SAT is hard). But after random restriction we may get to a polynomial size one. Indeed, if $q = 0$ then we get the trivial formula, and if $q = 1$ we get the same function. So our goal is to find the largest possible q such that after restriction $\rho(q)$, we get to a polynomial size CNF.

The following lemma, called Switching lemma, is first proved by Furst, Saxe and Sipser in 1981.

Lemma 1 (**Switching Lemma**) *Let $q = n^{-\frac{2}{3}}$, then for any DNF of polynomial size $P(n)$, and $\delta = \frac{1}{poly(n)}$, after random restriction we get to a CNF of size C with probability $(1-\delta)$, where C is constant.*

Before proving Switching lemma, let's use it to prove PARITY$\notin AC^0$.

Theorem 2 *PARITY $\notin AC^0$.*

Proof Suppose there is formula of size $s = poly(n)$ and depth d that computes PARITY$_n$. By induction we can assume there is no formula of size $poly(n)$ and depth $d-1$ that computes PARITY$_n$. Now consider the DNF formulas in the first two levels of this formula, and apply the Switching lemma on them. Since number of these formulas is at most s, with probability $1 - s\delta$ each of DNF's will change to a CNF. Now by the same idea as before, we have two consecutive AND levels and can combine them. Therefore we get to a circuit that compute PARITY on unrestricted variables, and has size $\leq O(s)$ and depth $d-1$, which is a contradiction. For the base of induction it is not hard to see that, if a circuit of depth 2 computes PARITY then its size is $O(2^n)$.

■

5 Proof of Switching Lemma

Let $f = T_1 \vee T_2 \vee \ldots T_m$ be a formula that solves PARITY. Call each T_j a term. Suppose we restrict variables in two stages

Stage 1

Restrict variable with probability $\sqrt{q}$, i.e.

$$x_i = \begin{cases} x_i \text{ with prob. } \sqrt{q} \\ 0 \text{ with prob. } \frac{1-\sqrt{q}}{2} \\ 1 \text{ with prob. } \frac{1-\sqrt{q}}{2} \end{cases}$$

$$f \longrightarrow f|_{\rho_1}.$$

Stage 2

Restrict variables in $f|_{\rho_1}$ with probability $\sqrt{q}$, i.e.

$$x_i = \begin{cases} x_i \text{ with prob. } \sqrt{q} \\ 0 \text{ with prob. } \frac{1-\sqrt{q}}{2} \\ 1 \text{ with prob. } \frac{1-\sqrt{q}}{2} \end{cases}$$

$$f|_{\rho_1} \longrightarrow f|_{\rho_1 \cup \rho_2}.$$

Claim 1 There is a constant c such that all terms of $f|_{\rho_1}$ are of size $\leq c$ (with high probability).

Consider two cases

- T_i has a *large* size

  Suppose T_i consists of $k = \Omega(\log n)$ variables. Then after restriction T_i is non-zero iff non of the variables restricted to 0. Therefore

$$Pr(T_i \neq 0) \leq (\frac{1-\sqrt{q}}{2})^k \leq \frac{1}{poly}.$$

- T_i has*small* size

  Suppose T_i contains $k' = O(\log n)$ variables. Then the probability that at least c of them remain unrestricted is

$$Pr(c \text{ variables remain unrestriced}) \leq k'^c(\sqrt{q})^c \leq (\frac{k'}{\sqrt{q}})^c \leq \frac{1}{poly}$$

Therefore **Claim 1** is true.

Claim 2 Each term of $f|_{\rho_1 \cup \rho_2}$ depends on b_c variables, again with high probability. Here b_c is a constant depends on c.

We prove it by induction on c. Consider two cases.

- There are $l = \Omega(\log n)$ terms $T_1, T_2, \ldots T_l$ in $f|_{\rho_1}$ such that they have disjoint variables.

  In this case, the probability that $f|_{\rho_1 \cup \rho_2} = 1$ is

$$Pr(f|_{\rho_1 \cup \rho_2} = 1) \geq \log n(\frac{1-\sqrt{q}}{2})^c \to 1$$

  then with high probability $f|_{\rho_1 \cup \rho_2} = 1$.

- $T_1, T_2, \ldots T_l$, where $l = O(\log n)$, are maximal disjoint terms in $f|_{\rho_1}$, i.e. any other term has at least one variable in $T_1, T_2, \ldots T_l$.

  Let H be set of variables in $T_1, T_2, \ldots T_l$, and Y be the remained variables, and assume after restriction ρ_2, H changes to H' and Y to Y'. Since each T_i has most c variables (Claim 1) then $\#H \leq cl$. Therefore after the second restriction, with high probability $\#H' \leq c'$, for some constant c'.

  Now set all variables in H, 0 or 1, then $f|_{\rho_1 \cup \rho_2 \cup H'}$ depends just on b_{c-1} variables. It is because we assigned either 0 or 1 to each variable in $T_1, T_2, \ldots T_l$, and by the maximality assumption in $T_1, T_2, \ldots T_l$, each of the remained terms has at most $c-1$ variables. So by the induction assumption $f|_{\rho_1 \cup \rho_2 \cup H'}$ has at most b_{c-1} variables.

  Set $b_c = c' + 2^{c'} b_{c-1}$. Then $f|_{\rho_1 \cup \rho_2}$ depends on at most b_c variables. Because there are at most c' variables in H', and for each of $2^{c'}$ assignments of variables in H', $f|_{\rho_1 \cup \rho_2 \cup H'}$ depends on b_{c-1} variables. Note that all of these arguments are true with high probability.

So after two restrictions ρ_1 and ρ_2, with high probability, we get to a formula that has constant number of variables. Now use the distributive law to switch the order of AND and OR gates. Since number of variables in constant, number of gates after that is also constant.