6.841/18.405J: Advanced Complexity Theory March 4, 2002

Lecture 8

Lecturer: Madhu Sudan *Scribe: Minh Nguyen*

1 Introduction

When we introduced the notion of circuits, one of our hopes was to use combinatorial techniques and circuits to prove that $P \neq NP$. We would like to show exponential lower bounds on circuit size for functions in NP, but the best we have been able to show is exponential lower bounds for constant depth circuits.

Today, we introduce the class AC_0 and prove a lower bound for the parity function. We do not introduce the class AC_0 because of its power but because of the powerful techniques used in the proof (algebraic techniques and randomization).

2 Circuit depth

We consider circuits with

- NOT gates (unary function),
- OR & AND gates with unbounded fan-in (the gates have an unlimited number of inputs and we take the OR/AND of all the input bits).

The depth of a circuit is defined as the longest path from input to output. (A circuit is an acyclic graph so "longest path" is well-defined and efficiently computable)

The size of a circuit is defined as the number of wires; if we are not interested in polynomial factors, it is also the number of gates in the circuit. We have seen that the circuit size represents non-uniform time complexity. Circuit depth represents parallel time, i.e. how fast a parallel algorithm can solve a problem. The unbounded fan-in simulates concurrent reading and writing on shared memory cells.

We define AC_0 as the class of constant depth, poly-size circuits with unbounded fan-in OR and AND gates.

3 Parity function

For every n, the parity function is defined as $\bigoplus_n : \{0,1\}^n \to \{0,1\}$, $\bigoplus(x_1 \dots x_n) = \sum x_i \pmod 2$.

Since the OR and AND gates have unbounded fan-in, the OR and AND functions can be computed in constant time. We will show that this is not the case for the parity function.

Theorem 1 *If C is a circuit of depth d computing the parity of n bits, then it must have size at least $2^{n^{\Omega(1/d)}}$.*

Note that we are not proving an impossibility result for constant depth circuits. For instance, there exists an exponential size circuit of depth 2 which computes the parity function by writing $\bigoplus(x_1 \dots x_n)$ in DNF (disjunctive normal form= an OR of ANDs) form, $\bigvee_{S \in \mathcal{F}}(\bigwedge_{i \in S} x_i \cdots)$. There is also a circuit of depth $\log n$ and size n which computes the parity of n bits. Hence we want to show that we cannot have simultaneously small size and small depth: the proof will have to consider these two quantities together.

History of this lower bound:

1. Furst, Saxe, Sipser (83) introduce the method of random restrictions. Their theorem is weaker, the circuit size is superpolynomial in n.

2. Yao (85) strengthens the theorem with an exponential lower bound using the same technique.

3. Hastad (87) obtains a simpler proof and a stronger exponential lower bound on the size of the circuit.

4. Smolensky (87) proves the theorem by algebraic methods, this is the proof we'll see today.

4 Polynomials over $\mathbb{Z}_3$

4.1 $\mathbb{Z}_3$

We consider $\mathbb{Z}_3 = \{-1, 0, 1\}$ with arithmetic mod 3, where we think of 2 as -1 since $-1 \equiv 2 \pmod 3$.

There are two ways to represent the Boolean world in $\mathbb{Z}_3$:

- The obvious one is $\{0, 1\} \subseteq \mathbb{Z}_3$,
- The other one is to use the map $\rho : \{0, 1\} \to \{1, -1\}$, where $\rho(0) = 1, \rho(1) = -1$.

The map ρ is linear: $\forall x \in \{0, 1\}, \rho(x) = 1 - 2x$ and $\forall y \in \{1, -1\}, \rho^{-1}(y) = \frac{1-y}{2}$. We can switch from one representation to the other by a linear transformation over the inputs and think of a function $f : \{0, 1\}^n \to \{0, 1\}$ or $f : \{1, -1\}^n \to \{1, -1\}$ as functions mapping $\mathbb{Z}_3^n \to \mathbb{Z}_3$.

4.2 Polynomials over $\mathbb{Z}_3$

$\mathbb{Z}_3$ is a field: polynomials over $\mathbb{Z}_3$ are well-behaved. In particular, the Schwartz lemma from the previous lecture applies: if $p : \mathbb{Z}_3^n \to \mathbb{Z}_3$ is a non-zero polynomial of degree d, then

$$\Pr_{\bar{a} \overset{R}{\leftarrow} \mathbb{Z}_3^n} [p(\bar{a}) = 0] \leq \frac{d}{3}$$

Note that the Schwartz lemma does not appear to be all that interesting in $\mathbb{Z}_3$: for instance, the polynomial $x_1^3 - x_1$ is zero for any element of the field. However, we will find it useful since we will encounter polynomials of total degree one and for such polynomials the lemma guarantees that the polynomial evaluates to non-zero values on at least 2/3rds of the inputs.

4.3 Examples

The function $AND(x_1 \ldots x_n) : \{0, 1\}^n \to \{0, 1\}$ can be computed by the polynomial $x_1 \cdot x_2 \ldots \cdot x_n$. Similarly, $OR(x_1 \ldots x_n) = 1 - \prod_{i=1}^n (1 - x_i)$. In each case, the function is computed by a polynomial over $\mathbb{Z}_3$ of degree 1 in each variable. This comes from the following fact:

Fact 2 *For every $f : \{0, 1\}^n \to \{0, 1\}$, we can find a polynomial $q : \mathbb{Z}_3^n \to \mathbb{Z}_3$ such that q has degree 1 in each variable and agrees with f on $\{0, 1\}^n$.*

(proof using AND, OR functions or using interpolation)

We have a similar fact for $g : \{1, -1\}^n \to \{1, -1\}$: g is computed by some polynomial of degree 1 in each variable. Suppose that $g(\rho(x)) = \rho(f(x))$ and the boolean function f is represented by the polynomial p. Then $g(y)$ is represented by the polynomial $1 - 2[p\left(\frac{1-y_1}{2}, \frac{1-y_2}{2}, \cdots, \frac{1-y_n}{2}\right)]$.

The parity function has a nice formulation in the $\{1, -1\}$ representation:

$$\bigoplus_n : \{1, -1\}^n \to \{1, -1\}, \bigoplus(x_1 \ldots x_n) = \prod_{i=1}^n x_i$$

5 Proof of the theorem

Proof Idea The main intuition behind the proof are the following insights.

- The degree of a function is a measure of its complexity.
- Parity has high degree since $\bigoplus(x_1 \dots x_n) = \prod_{i=1}^{n} x_i$.
- Circuits in AC_0 compute low degree functions.

There is a caveat in this reasoning: the functions AND and OR have also degree n and belong to AC_0! But if we delete part of the input to the function, we can represent the rest with a small depth circuit and a small degree polynomial. We will show that AC_0 "essentially" computes only small degree polynomials. These ideas are formalized in the next three lemmas. ■

Lemma 3 *If $f : \{0,1\}^n \to \{0,1\}$ is computed by a depth d circuit of size $s = 2^{n^{o(1/d)}}$, then there exist a set $S \subseteq \{0,1\}^n$ of size $|S| > \frac{3}{4} \cdot 2^n$ and a polynomial $p : \mathbb{Z}_3^n \to \mathbb{Z}_3$ of total degree $\leq (\log s)^{O(d)}$ such that $p(x) = f(x), \forall x \in S$.*

Lemma 4 *If there exists a degree D polynomial $p : \mathbb{Z}_3^n \to \mathbb{Z}_3$ such that $p(x) = \bigoplus(x)$ for all $x \in S \subseteq \{0,1\}^n$, then every Boolean function $f : S \to \{0,1\}$ is represented by a polynomial of total degree $\frac{n}{2} + D$ and of degree 1 in each variable.*

Lemma 5 *If $h_1, \cdots, h_N : S \to \{0,1\}$ are such that for every $f : S \to \{0,1\}, \exists \alpha_1, \cdots, \alpha_N$ such that $f = \sum_{i=1}^{N} \alpha_i h_i$, then $N \geq |S|$.*

Proof of Theorem 1: Proof by contradiction. Suppose there exists a circuit C of depth d and size $s = 2^{n^{o(1/d)}}$ computing the parity function.

- By Lemma 3, the parity function is computed on S of size $|S| \geq \frac{3}{4} \cdot 2^n$ by a polynomial of degree $D \leq (\log s)^{O(d)}$.
- By Lemma 4, every Boolean function on S is computed by a polynomial of degree $\frac{n}{2} + D \leq \frac{n}{2} + (\log s)^{O(d)}$.
- Let $h_1, \cdots, h_N$ be all the monomials $x_1^{i_1} \cdots x_n^{i_n}$ where $i_j \in \{0,1\}$ and $\sum i_j \leq \frac{n}{2} + D$. Note that $h_1, \cdots, h_N$ generate all Boolean functions on S, hence by Lemma 5, we must have that $N \geq |S| \geq \frac{3}{4} \cdot 2^n$.

  How many such monomials are there?

$$
\begin{aligned}
N &\leq \sum_{i=0}^{\frac{n}{2}+D} \binom{n}{i} && \text{choose how to distribute up to } n/2 + D \text{ ones in a vector of size } n \\
N &\leq \sum_{i=0}^{\frac{n}{2}} \binom{n}{i} + \sum_{i=\frac{n}{2}+1}^{\frac{n}{2}+D} \binom{n}{i} \\
N &\leq 2^{n-1} + D \cdot \left(\frac{2^n}{\sqrt{n}}\right) && \text{each term } \binom{n}{i} \text{ for } n/2 < i \leq n/2 + D \text{ is smaller than } \left(\frac{2^n}{\sqrt{n}}\right)
\end{aligned}
$$

  If we assume that $s < 2^{n^{o(1/d)}}$, then $\frac{D}{\sqrt{n}} \leq \frac{1}{4}$ and $N < \frac{3}{4} \cdot 2^n$. This contradicts Lemma 5, hence there is no such circuit C.

■

6 Proofs of lemmas

Proof of Lemma 5:

Note that the set of functions $f : S \to \{0,1\}$ are members of the vector space $\mathbb{Z}_3^{|S|}$ that contain the "unit" functions $\{\delta_x\}_{x \in S}$ where $\delta_x(y) = 1$ if $y = x$ and 0 otherwise. Furthermore the δ_x functions are linearly independent of each other. Thus any collection of functions $h_1, \ldots, h_N$ that generate all the functions f must have size $N \geq |S|$. The following paragraph elaborates further on this proof.

Every function $f : S \to \{0,1\}$ can be viewed as a vector of size $|S|$ over $\{0,1\}$. We will ignore most of these functions and just focus on the δ_x functions defined above. Consider the $|S| \times |S|$ matrix F whose xth row is the vector corresponding to the function δ_x. Note this matrix is simply the identity matrix. Now suppose there exists functions $h_1, \cdots, h_N : S \to \mathbb{Z}_3$ such that for every function $f : S \to \{0,1\}$, there exists $\alpha_1, \cdots, \alpha_N \in \mathbb{Z}_3$. such that $f = \sum_{i=1}^{N} \alpha_i f_i$. In particular, let $\alpha_{x,1}, \ldots, \alpha_{x,N}$ be the multipliers needed to get the function δ_x. Let us represent the functions $h_1, \ldots, h_N$ as $|S|$ dimensional vectors over $\mathbb{Z}_3$ as well, and consider the N by $|S|$ matrix H whose rows are the vectors $h_1, \ldots, h_N$. Now let A be the $|S|$ by N matrix whose rows are indexed by $x \in S$ and columns by index $i \in \{1, \ldots, N\}$ and where the entry $A_{x,i} = \alpha_{x,i}$. By construction, $A \cdot H = F$! Now we get $|S| = \text{rank}(F) = \text{rank}(A \cdot H) \leq \min\{\text{rank}(A), \text{rank}(H)\} \leq N$. ■

Proof of Lemma 4: Let S be a subset of $\{0,1\}^n$. Assume that $\bigoplus : \{0,1\}^n \to \{0,1\}$ is represented by a polynomial $p : \mathbb{Z}_3^n \to \mathbb{Z}_3$ of degree D. Let $T \subseteq \{1,-1\}^n$ be the associated set, $T = \{\rho(x) | x \in S\}$. Then the map $\rho_{\bigoplus} : T \to \{1,-1\}, \rho_{\bigoplus}(y_1 \ldots y_n) = \prod_{i=1}^{n} y_i$ agrees with the polynomial $r(\mathbf{y}) = 1 - 2p\left(\frac{1-\mathbf{y}}{2}\right)$ on the set T. r is a polynomial of total degree D: the hypothesis about the parity function holds in the $\{1,-1\}$ representation as well.

Consider a Boolean function $f : S \to \{0,1\}$. Let $g : T \to \{1,-1\}$ be the associated function which is represented by a polynomial, i.e. by a summation of monomials. Let $\{A_i\}$ be monomials of total degree less or equal to $n/2$ and $\{B_i\}$ be monomials of total degree more than $n/2$.

$$Q_g = \sum_i \alpha_i A_i + \sum_i \beta_i B_i$$

Let $C_j = \frac{\prod_{i=1}^{n} x_i}{B_j}$ (complement of the variables appearing in B_j). Then $B_j \cdot C_j = \rho_{\bigoplus}(x_1 \ldots x_n)$ and in the $\{1,-1\}$ representation, $B_j = C_j \cdot \rho_{\bigoplus}(x_1 \ldots x_n)$.

$$Q_g = \sum_i \alpha_i A_i + \rho_{\bigoplus} \sum_i \beta_i C_i$$

g is represented on T by the polynomial Q_g of total degree at most $\frac{n}{2} + D$ and of degree at most 1 in each variable (by substituting $x_i^2 = 1$). Switching back to the $\{0,1\}$ representation, we have that the function f is represented by the polynomial $Q_f(x) = \frac{1 - Q_g(1-2x)}{2}$ of degree $\frac{n}{2} + D$ on S.

■