6.841 Advanced Complexity Theory Feb 26, 2007

Lecture 6

Lecturer: Madhu Sudan *Scribe: David Glasser*

Today's lecture is an alternate proof that Parity is not in AC^0, as shown by Smolensky.

Recall: AC^0 has polynomially-sized circuits of constant depth, with unlimited fan-in ANDs and ORs. Argument is roughly that parity is complex, whereas AC^0 is simple. This is a more algebraic argument than last time.

What measure of complexity can we use to make this argument? The degree of the polynomial which computes the function. What field is this polynomial in? The obvious candidate is $\mathbb{F}_2$. Then we can have a polynomial on n variables. This doesn't work too well: AND ends up being hard ($x_1 x_2 \cdots x_n$ has degree n) and parity is simple ($x_1 + \cdots + x_n$) has degree 1. That's backwards! Instead, we'll use *any* other field, or specifically any field of characteristic not equal to 2.

There is a fundamental trick of complexity theory: instead of writing about 0 and 1, we write about 1 and -1. We can map between this with $1 - 2x$ or $\frac{1-y}{2}$. Or $(-1)^b$. Now, parity is a sum of things in the 0/1 representation, but in the other representation, it's about a product/power, which is more complicated.

We claim that if $p\colon \mathbb{F}^n \to \mathbb{F}$ computes the parity of its input when the inputs are 0 or 1, then a polynomial q defined as $1 - 2p\left(\frac{1-y_1}{2}, \ldots, \frac{1-y_n}{2}\right)$ compute the product of the variables y when each y_i is ± 1. Exercise for the reader: show that p and q have the same degree.

We can use this to prove that the degree of p must be at least n, since we can show (apparently) that a product polynomial must actually contain a $y_1 \cdots y_n$ term. I'm not sure why, but we're going to prove something stronger today I think.

OK, so we've made parity complicated. Can we make AND easy? This is why this is the Razborov-Smolensky theorem: this is Razborov's part. It is the method of approximations. There's some fuzziness involved. The idea is that we have a function that is almost the same as whatever you're actually trying to model.

Lemma: If C is an AC^0 circuit of depth d and size s, then for all ϵ, there exists a $S \subseteq \{0,1\}^n$ and a polynomial $p(x_1, \ldots, x_n)$ of degree D such that for all $x \in S$, $p(x) = C(x)$ and $D \leq \left(\log \frac{s}{\epsilon}\right)^d$ and $|S| \geq (1-\epsilon)2^n$.

Smolensky idea: fix an input, and replace gates probabilistically. Here's a lemma towards this: There exists a distribution of polynomials p of degree $k = O\left(\log \frac{s}{\epsilon}\right)$ such that for all $y_1, y_2, \ldots, y_n$, $\Pr[p(Y) = OR(Y)] \geq 1 - exp(-k)$, where the probability is over the polynomials.

Exact OR is defined via De Morgan as $1 - \prod(1 - y_i)$, but this has degree n: ewww. We'd like low-degree almost-ORs. How does this work? We'll define $Probably - OR(y_1, \ldots, y_m)$ by picking α_1 through α_m at random from $\mathbb{F}$ and look at $\sum \alpha_i y_i$. If the inputs are all 0, then (like OR) the output is zero.

DeMillo-Lipton-Schwartz-Zippel Lemma: If $p\colon \mathbb{F}^n \to \mathbb{F}$ is a nonzero polynomial of degree d, then $\Pr_{x \in S^n}[p(x) = 0] \leq \frac{d}{|S|}$. This is basically just a matter of counting zeros; we can get it by induction on the number of variables.

How do we apply this? Consider the polynomial $p(\alpha_i) = \sum \alpha_i y_i$ (so the α_i are the variables). So when the y_i are not all zero, the probability that the output is nonzero is at least $1 - \frac{1}{\mathbb{F}}$. If we use $\mathbb{F}_3$, this is $\frac{2}{3}$, and it only gets better. We might want to raise this to $|\mathbb{F}| - 1$ to make the answer always 0 or 1; so it ends up having degree 2 (for $\mathbb{F}_3$). So this gives us a weak probabilistic OR.

How about a stronger one? Pick $\alpha^{(1)}$, $\alpha^{(2)}$, $\ldots$, $\alpha^{(k)}$ from F^m independently. We'll let probabilistic OR be the EXACT-OR of the elements $\left(\sum \alpha_i^{(j)} y_i\right)^2$. So this will have degree $2k$ to compute the OR of m things, where k is not a function of m.

Given a circuit, we can replace NOT gates by a degree-1 polynomial; we can DeMorganize AND gates; and we can replace OR gates with this probabilistic OR polynomial. Thus we prove a lemma: given a circuit C of size S and depth d and some $\epsilon > 0$, there exists a distribution of polynomials p of degree at most $\left(\log \frac{s}{\epsilon}\right)^d$ such that for every $x \in \{0,1\}^n$, $\Pr_p[p(x) = C(x)] \geq 1 - \epsilon$. i.e., with a fixed input and picking polynomials randomly, we do well. Each OR gates gives degree k, so by the top of the circuit, it's k^d. For this input, we can choose k large enough to shrink the error, and it gives us the value of k around $\log \frac{s}{\epsilon}$, I suppose.

Now, how do we prove our original lemma, which is about varying inputs? Basically by choosing the set S to be the ones where the lemma we just "proved" holds. So we've shown that AC^0 is simple, basically; now we need to show that parity isn't.

Our new lemma: parity cannot be approximated by low-degree polynomials. So we'll prove that if $|T| \geq (1-\epsilon)2^n$ and $T \subseteq \{-1,1\}^n$, and p is a polynomial of degree D such that p computes the product of its inputs for all $Y \in T$, then $D \geq \Omega\left(\left(\frac{1}{2}-\epsilon\right)\sqrt{n}\right)$. How do we show this?

Consider all functions mapping T into $\mathbb{F}_3$. There are $3^{|T|}$ such functions. All of these are polynomials of degree at most 2 in each variable. Consider $q(y_1,\ldots,y_n) = \sum_d c_d \prod_i y_i^{d_i}$. Because we only care about the behavior when $y_i = \pm 1$, then we can assume that each d_i is either 0 or 1 (since $y_i^2 = 1$).

Consider the product $y_1 y_2 \cdots y_{\frac{3n}{4}}$; it is equal to $y_1 y_2 \cdots y_{\frac{3n}{4}} y_{\frac{3n}{4}+1}^2 \cdots y_n^2 = (\prod y_i)(\prod_{i>\frac{3n}{4}} y_i)$. But this is equal to $p(Y)(\prod_{i>\frac{3n}{4}} y_i)$; so the degree is now $\frac{n}{4}+D$. In fact, for the general case of a function, we can get that $\sum d_i$ is at most $\frac{n}{2}+D$. So if this p exists, then all the functions have polynomials of degree at most $\frac{n}{2}+D$. How may ways can you write these polynomials? $\sum_{i=0}^{\frac{n}{2}+D}\binom{n}{i} \leq \sum_{i=0}^{\frac{n}{2}}\binom{n}{i} + D\binom{n}{\frac{n}{2}} \leq 2^{n-1} + D2^n$. There aren't enough polynomials! So, more or less, this shows that the theorem holds. (In fact, even given gates to calculate parity mod 3, it's hard to calculate mod 2.)