CS 710: Complexity Theory 2/16/2009

Lecture 9: Polynomial Approximations

Instructor: Dieter van Melkebeek Scribe: Phil Rydzewski & Piramanayagam Arumuga Nainar

Last time, we proved that no constant depth circuit can evaluate the parity function. We used random restrictions to obtain a bound on the complexity of a circuit evaluating the parity of n inputs. In this lecture, we give an alternative proof of a slightly weaker bound.

Using the random restriction method, we showed that $C_d(\oplus_n) \geq 2^{\Omega(n^{\frac{1}{d-1}})}$. This is a tight bound and uses the property that the parity function is sensitive to every bit of its input. We can also derive a similar bound for the mod_m function defined as follows:

$$\text{mod}_m(x) = \begin{cases} 0 & \text{if } |x| \equiv 0 \pmod{m} \\ 1 & \text{otherwise} \end{cases} \tag{1}$$

where $|x|$ is the number of non-zero bits in the input. Parity is a special case of mod_m at $m = 2$. Another function that we can prove is not in AC^0 is the majority function, which returns 1 when more than half of the input bits are 1, and 0 otherwise. This can be proved by using the parity function as a black box and is left as an exercise.

In this lecture, we use low degree polynomial approximations to show that $C_d(\oplus_n) \geq 2^{\Omega(n^{\frac{1}{2d}})}$. Even though this is a weaker bound, the technique itself is interesting. Moreover, this result applies even if we allow mod_3 gates in the circuit. Finally, prove that constant-depth circuits are unable to approximately evaluate parity.

1 Polynomial approximation method

Theorem 1. $\oplus_n \notin \text{AC}^0$. *Specifically,* $C_d(\oplus_n) \geq 2^{\Omega(n^{\frac{1}{2d}})}$

Proof. We prove this in two steps. First, we show that any constant-depth circuit can be approximated using a low degree multivariate polynomial over the field $\mathbb{Z}_3$. Using $\mathbb{Z}_3$ gives, for free, the ability to mimic mod_3 gates. In general, if we use $\mathbb{Z}_p$ for a prime number p, we can handle mod_p gates. Second, we show that the parity function cannot be approximated using a multivariate polynomial of sufficiently low degree over the field $\mathbb{Z}_3$.

Step 1: Consider a circuit C made of AND, OR, NOT and MOD_3 gates. It can always be represented as a multivariate polynomial of degree n where n is the size of the input. Our goal is to represent it using a polynomial of lower degree, allowing errors if required. A literal x that is directly passed as input to a gate can be represented using the polynomial x. This is the base case of our construction. Now, we can assume that a polynomial P_i can be associated with the i^{th} input of other gate types (AND, OR, NOT, MOD_3). The goal is to construct a polynomial P' that represents the output of the gate. Note that the number of inputs to any gate is at most $|C|$. Construct each of the gates as follows:

NOT: If P' is the polynomial representing the input of a NOT gate, then $1 - P'$ represents the output. Notice that this representation doesn't increase the degree of the polynomial or introduce any additional error.

MOD$_3$: The output of the gate is zero when $\sum_{i=1}^{m} P_i \equiv 0 \pmod 3$. If the summation is 1 or 2, the output is 1. Note that in the field $\mathbb{Z}_3$, $2 \cdot 2 = 1$. So, we can model the gate using the polynomial $P' = (\sum_{i=1}^{m} P_i)^2$. This polynomial accurately models the gate and its degree is at most twice the degree of any of its inputs.

OR: The output of the OR gate is 0 if $(\forall i)\, P_i = 0$. Or, in other words, $(\forall i)\,(1 - P_i = 1)$. Otherwise its output is one. This can be represented as follows:

$$\alpha : P' = 1 - \prod_{i=1}^{m}(1 - P_i) \tag{2}$$

This representation is accurate but the degree of P' may be up to m times the degree of the P_i with the largest degree. This can be much higher than the trivial bound n if there are many gates and many levels in the circuit. To tackle this, we model P' as a random linear combination of P_i for $1 \le i \le n$. Let $r_i \in \mathbb{Z}_3$ be the coefficient associated with P_i, chosen uniformly random. As with MOD$_m$, we square the linear combination to keep the value of P' boolean. This leaves us with:

$$\beta : P' = (\sum_{i=1}^{m} r_i \cdot P_i)^2 \tag{3}$$

This makes the degree of P' at most twice that of the degree of its inputs. But it is definitely not an accurate description of an OR gate. Evaluate the probability of P_i being different from the boolean expression $\vee_{i=1}^{n} P_i$. If $P_i = 0$ for all i, then irrespective of the values picked for the coefficients, the output is correct. If $P_i = 1$ for at least one i, then $\displaystyle\sum_{i=1}^{m} r_i \cdot P_i$ is $\displaystyle\sum_{i|P_i=1} r_i$. This is the wrong value, 0, in one out of three cases for a random assignment of the coefficients. Thus, P' can introduce errors in the representation with a probability at most $\frac{1}{3}$. As with any randomized algorithm, we can repeat the above calculation for, say, t independent trials and see if the output of at least one of the trials is one. (Note: An output of one will always be correct but an output of zero may be wrong). This leads us to the third, and final, formulation of P'.

$$P' = P'_{\alpha}(P'_{\beta_1}(\hat{P}), \ldots P'_{\beta_t}(\hat{P})) \tag{4}$$

Here, P'_{α} is the application of P' as described in eqn. 2 on t inputs. P'_{β_k} is the k^{th} trial using the formulation of P' in eqn. 3. $\hat{P}$ is a shorthand for $P_1, P_2, \ldots P_m$. The above formulation produces a wrong output if all the trials produce the wrong output, i.e. with probability at most $\frac{1}{3^t}$. The degree of P' increases by a factor of $2t$: a factor t for the α-formulation and a factor of 2 for the β-formulation.

AND: We can handle an AND gate in a similar way, resulting in an approximation P' with at most a factor of $2t$ blow-up in the degree, and giving an imprecise value with probability at most $\frac{1}{3^t}$.

If the depth of the circuit is d, the degree of the polynomial P representing the entire circuit will be at most $(2t)^d$. P gives the wrong value only if the output of at least one of the gates in C was wrong. This happens with probability at most $\frac{|C|}{3^t}$. Note, this is not a very tight upper bound but it is enough for this proof. A tighter bound would depend on the number of OR gates in C. By averaging, the expected number of inputs for which P will give the wrong value is at most $\frac{|C|}{3^t}2^n$ since there are 2^n possible inputs of length n. There exists a choice for the random coefficients for which P is wrong in no more than the expected number, derived above. More formally,

Lemma 1. *There exists a choice of r_i's such that there exists a set $G \subseteq \{0,1\}^n$ of relative size $\mu(G) \geq 1 - \frac{|C|}{3^t}$ such that $(\forall x \in G)\, P(x) = C(x)$, where P is a polynomial of degree at most $(2t)^d$ constructed as described above.*

Here, $\mu(G)$ is the relative size of G with respect to the set of all possible inputs to C and is equal to $\frac{|G|}{2^n}$. This construction can be generalized to work over any field $\mathbb{Z}_p$ for prime p, thus allowing mod_p gates. The property of $\mathbb{Z}_3$ we used is that $a^2 \equiv 1 \pmod 3$ for all $a \not\equiv 0 \pmod 3$. Thus, squaring a polynomial ensures boolean values. To work over $\mathbb{Z}_p$, we would instead raise polynomials to the power $p-1$ as $a^{p-1} \equiv 1 \pmod p$ for all $a \not\equiv 0 \pmod p$. The degree of the resulting polynomial is at most $(p \cdot t)^d$ rather than $(2t)^d$.

Step 2: In this step, given a polynomial P of some degree that approximates $\oplus_n$ on a subset G of inputs, we establish an upper bound below which every function of n inputs has a corresponding polynomial approximating it over G. By equating the number of such functions to the number of polynomials with degrees not greater than the established upper bound, we derive the lower bound on the depth of circuit C.

As a first step, we transform the inputs to a slightly more convenient domain: $\{-1,1\}$ instead of $\{0,1\}$.

Proposition 1. *Suppose there exists a polynomial P of degree at most Δ that computes $\oplus_n$ on a set $G \subseteq \{0,1\}^n$. Then there exists a polynomial P' of degree at most Δ and a set $G' \subseteq \{-1,1\}^n$ such that $\mu(G') = \mu(G)$ and $(\forall x \in G') \prod_{i=1}^{n} x_i = P'(x)$.*

The reason is that parity on boolean inputs is equivalent to multiplication over $\{-1,1\}$.

Lemma 2. *Suppose there exists a polynomial P' of degree at most Δ that represents multiplication in a set $G' \subseteq \{-1,1\}^n$. Then each function $f : G' \to \mathbb{Z}_3$ has a multivariate polynomial Q over $\mathbb{Z}_3$ of degree at most $\frac{n+\Delta}{2}$ such that it represents f, i.e. $(\forall x \in G') f(x) = Q(x)$.*

Proof. Every function f has a multivariate polynomial of degree at most n. This is trivial because we can hardwire every possible input using monomials of degree n. Let us start from one such polynomial Q' (such that $f = Q'$ on G'). Consider a monomial in Q' of the form $\prod_{i \in I} x_i$ where I is a subset of the input bits. Because we are only concerned with ± 1 inputs, we can rewrite it as:

$$\begin{aligned}
\prod_{i \in I} x_i &= \Big(\prod_{i \notin I} x_i^2\Big)\Big(\prod_{i \in I} x_i\Big) \\
&= \Big(\prod_{i \notin I} x_i\Big)\Big(\prod_{i=1}^{n} x_i\Big) && (5) \\
\implies \prod_{i \in I} x_i &= \Big(\prod_{i \notin I} x_i\Big) P'(x) && (6)
\end{aligned}$$

Eqn. 5 holds for any input x of n bits but eqn. 6 holds only for the inputs in the set G'. The LHS of 6 has degree $|I|$. The RHS has a degree at most $\Delta + |\bar{I}| = \Delta + n - |I|$. Averaging these gives a minimum degree $\leq \frac{n+\Delta}{2}$. Thus, we can make the degree of every monomial in Q' to not exceed $\frac{n+\Delta}{2}$. □

Given the lemmas, we are now ready to prove the theorem. Suppose there exists a circuit C of depth d computing $\oplus_n$. From Lemma 1, there exists a polynomial P' of degree at most $\Delta = (2t)^d$ that computes parity on a set G of relative size at least $1 - \frac{|C|}{3^t}$. Consequently, from Lemma 2, all functions $f : G' \to \mathbb{Z}_3$ for some G' such that $|G| = |G'|$ can be represented using a multivariate polynomial of degree at most $\frac{n+\Delta}{2}$. The total number of such polynomials must be at least the number of functions f from G' to $\mathbb{Z}_3$.

The number of multivariate polynomials with degree at most $\frac{n+\Delta}{2}$ is exactly 3^M where M is the number of monomials of degree at most $\frac{n+\Delta}{2}$. There are $\binom{n}{i}$ monomials of degree i, so

$$M = \sum_{i}^{\frac{n+\Delta}{2}} \binom{n}{i}$$

The number of monomials of degree $\leq \frac{n}{2}$ will be 2^{n-1} - half of the 2^n possible monomials. The remaining $\frac{\Delta}{2} = \Theta(\Delta)$ terms in the summation will be lower than $\binom{n}{\frac{n}{2}}$ - the maximum possible number for any degree. Using Stirling's approximation, we can show that:

$$\binom{n}{\frac{n}{2}} = \Theta\left(\frac{2^n}{\sqrt{n}}\right)$$

Thus, $M = 2^{n-1} + 2^n \cdot \Theta\left(\frac{\Delta}{\sqrt{n}}\right) = 2^n\left(\frac{1}{2} + \Theta\left(\frac{\Delta}{\sqrt{n}}\right)\right)$.

The number of functions of the form $G' \to \mathbb{Z}_3$ is $3^{|G'|}$ as one of 3 possible values can be assigned to each element of G'. Because the number of functions of this form must be at most the number of polynomials of degree at most $(n+\Delta)/2$, $3^{|G'|} \leq 3^M$ or, in other words, $|G'| \leq M$. This gives us the following bound on the size of G'.

$$\mu(G') = \frac{|G'|}{2^n} \leq \frac{M}{2^n} \leq \frac{1}{2} + \Theta\left(\frac{\Delta}{\sqrt{n}}\right)$$

From Lemma 1, $\mu(G') \geq 1 - \frac{|C|}{3^t}$ when $\Delta = (2t)^d$. Thus,

$$\begin{aligned} 1 - \frac{|C|}{3^t} &\leq \mu(G') \leq \frac{1}{2} + \Theta\left(\frac{(2t)^d}{\sqrt{n}}\right) \\ \Longrightarrow |C| &\geq 3^t\left[\frac{1}{2} - \Theta\left(\frac{(2t)^d}{\sqrt{n}}\right)\right] \end{aligned}$$

Setting $(2t)^d = O(\sqrt{n})$ gives a tight value for the RHS in the last equation. Thus, $t = \Theta(n^{\frac{1}{2d}})$. This gives $|C| \geq 2^{\Omega(n^{\frac{1}{2d}})}$. □

The only part of the above analysis that changes when working over $\mathbb{Z}_p$ rather than $\mathbb{Z}_3$ is that $\Delta = (p \cdot t)^d$ rather than $(2t)^d$. Thus the result holds with the same lower bound on $|C|$ for boolean circuits with mod_p gates for any prime p. In fact, the argument in the above proof can be generalized to give a lower bound for circuits with mod_p gates to compute mod_q (recall that parity is the special case of $q = 2$). This is achieved by viewing Step 2 as harmonic analysis over $\mathbb{Z}_2$ and then generalizing that to harmonic analysis over $\mathbb{Z}_q$. As this generalization takes a bit of work to prove, we leave it at that.

Because the lower bound for parity was proved by viewing parity as multiplication, we get a lower bound for multiplication as well.

Corollary 1. *The decision variant of binary multiplication is not in* AC^0.

We further use the proof above to give a lower bound on circuits that even approximate parity.

Corollary 2. *A depth d unbounded fan-in circuit that agrees with parity on a fraction at least $\frac{1}{2} + \frac{1}{n^{(1-\epsilon)/2}}$ of $\{0,1\}^n$ must have size $2^{\Omega(n^{\epsilon/2d})}$.*

Proof. Suppose we have a circuit that is correct on at least $\frac{1}{2} + \rho$ of the inputs. Similar to Theorem 1, we can prove that there exists a polynomial of degree $\Delta = (2t)^d$ that is correct on a set G' that is at least $\frac{1}{2} + \rho - \frac{|C|}{3^t}$ of $\{0,1\}^n$. From *Step 2* of the proof above,

$$\frac{1}{2} + \rho - \frac{|C|}{3^t} \leq \frac{1}{2} + \Theta\left(\frac{(2t)^d}{\sqrt{n}}\right) \implies \rho - \frac{|C|}{3^t} \quad \leq \quad \Theta\left(\frac{\Delta}{\sqrt{n}}\right) \tag{7}$$

$$\implies |C| \quad \geq \quad 3^t \left[\rho - \Theta\left(\frac{(2t)^d}{\sqrt{n}}\right)\right] \tag{8}$$

Note that the $(2t)^d/\sqrt{n}$ term is $\Omega(1/\sqrt{n})$, so ρ must also be $\Omega(1/\sqrt{n})$ to ensure the lower bound we get is even positive. If we let $\rho = 1/n^{(1-\epsilon)/2}$, we set $(2t)^d = \Theta(n^{\epsilon/2})$ to optimize the RHS of 8. So $t = \Theta(n^{\epsilon/(2d)})$, and we get that $|C| \geq 2^{\Omega(n^{\epsilon/(2d)})}$. □

The above corollary proves the inapproximability of the parity function using constant depth circuits. There is another such result that can be proved using random restrictions. It is as follows:

Theorem 2. *A depth d unbounded fan-in circuit that agrees with parity on a fraction at least $\frac{1}{2} + \frac{1}{2^{\Omega(n^{1/d})}}$ of $\{0,1\}^n$ must have size $2^{\Omega(n^{1/d})}$.*

This is interesting because even trivial functions can guess parity correctly on half of the inputs. This is slightly weaker than the $2^{\Omega(n^{\frac{1}{d-1}})}$ bound we derived last lecture but it disproves approximability rather than computability of the parity function. We will see more such results of inapproximability when we discuss pseudo-randomness.

2 Next lecture

Next lecture, we will discuss parallelism where we distribute the computational task among multiple processors to reduce the time complexity.