

Lecture 11: Polynomial Approximations

Instructor: Dieter van Melkebeek

Scribe: Chetan Rao

In the previous lecture, we discussed Boolean circuits and their lower bounds. In particular, we discussed constant-depth circuits with unbounded fan-in (in-degree) that can compute certain operations such as binary addition in polynomial-size. We also highlighted that binary multiplication and parity (\oplus_n - defined in last lecture) cannot be decided in this framework.

In today's lecture, we will focus on the two different approaches to show that parity, and in turn multiplication, cannot be decided by sub-exponential-sized constant-depth circuits. The following section (Section 1) provides insight on two methods - *polynomial approximations* and *random restrictions*. Sections 2 and 3 conclude with a lower bounds for parity.

1 Lower Bounds for \oplus_n on constant-depth circuits

The literature has concentrated on two methods to establish lower bounds for parity on constant-depth circuits -

- **Random Restrictions method** - This approach uses a p -random restriction (defined in previous lecture) to prove a tight lower bound for circuits - $C_d(\oplus_n) \geq 2^{\Omega(n^{\frac{1}{d-1}})}$. This method also provides similar results for MAJORITY and Modular gates (MOD_k).

Definition 1 (Modular gate). A MOD_k gate is a gate that outputs 0 if the sum of its inputs (mod k) is 0, and 1 otherwise. The precise definition is as follows -

$$\text{MOD}_k(x) = \begin{cases} 0 & \text{if } \sum_i x_i \equiv 0 \pmod{k} \\ 1 & \text{otherwise} \end{cases}$$

where x is the gate input.

- **Polynomial Approximations method** - This method makes use of low-degree polynomial approximations (over finite fields) to show that $C_d(\oplus_n) \geq 2^{\Omega(n^{\frac{1}{2d}})}$ (d is circuit depth). Although this is a weaker bound in comparison with the random restrictions method, the proof technique is very interesting and it allows the use of MOD_3 gates.

2 Polynomial Approximations method

Theorem 1. Given any circuit C_d of depth d that decides \oplus_n , its size is at least $2^{\Omega(n^{\frac{1}{2d}})}$. This result holds even if we allow additional MOD_3 gates.

Proof. We prove the theorem in two steps. In the first step, we approximate the output of constant-depth circuits using a low degree polynomial over the field \mathbb{Z}_3 with a small error. The choice of \mathbb{Z}_3 aids in expressing MOD_3 gates as low degree polynomials. In general, for any prime p , MOD_p

gates have low degree polynomials in the field \mathbb{Z}_p . In the subsequent step, we show that for the parity function such a low-degree polynomial does not exist.

Step 1: Consider a circuit C made of AND, OR, NOT and MOD_3 gates. C can be represented as a multivariate polynomial of degree n where n is the size of the input using a polynomial coefficient for each input. However, our aim is to express C as a lower degree polynomial, allowing reasonably small errors. To achieve this objective, we must represent every literal and gate of the circuit in terms of an approximate low-degree polynomial over \mathbb{Z}_3 i.e. find polynomial $P \in \mathbb{Z}_3$ s.t. on many inputs $X \in \{0, 1\}^n$, $P(X)$ is equal to the value of the gate.

We construct the polynomial inductively from the bottom to top. The base case would be a simple literal:

LITERAL (X): In this case, the polynomial is trivially idempotent - $P_i(X) = X_i$. The degree of the polynomial is 1 and there are no errors induced by this substitution.

In the inductive step, we consider each of the possible gates. We consider the following sequence of gates:

NOT ($-$): If the input of a NOT gate is polynomial $P(X)$, then $P'(X) = 1 - P(X)$ represents the output. This does not increase the degree of the polynomial or introduce any additional error.

MOD₃ : The output $P'(X)$ of the gate is 0 when $\sum_{i=1}^m P_i(X) \equiv 0 \pmod{3}$. The output is 1 if the resultant sum is either 1 or $-1(2)$. Upon squaring the sum of input polynomials, the polynomial exactly behaves like the MOD_3 gate. Hence, the output polynomial is $P'(X) = (\sum_{i=1}^m P_i(X))^2$. However, the degree of the polynomial doubles with this gate and there are no additional errors.

OR (\vee): The output $P'(X)$ of the OR gate is 0 iff $(\forall i) P_i(X) = 0$, or equivalently, $(\forall i) (1 - P_i(X) = 1)$. This can be represented as follows:

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

The output polynomial is precise but its degree is dependent on the fan-in m and could be unbounded. In the presence of multiple hierarchies of OR gates the polynomial will have a comparable degree to that of the trivial bound n . To obviate this issue, we pick a randomized linear combination of inputs. This process is repeated t times for a better approximation of the OR gate.

Let $r_i \in \mathbb{Z}_3$ be the coefficient associated with $P_i(X)$ and let the output function be defined as:

$$\beta : P'(X) = \left(\sum_{i=1}^m r_i \cdot P_i(X) \right)^2 \quad (2)$$

Note that we square the linear combination to keep the output $P'(X)$ as Boolean. If the actual output of the OR gate is 0, the approximation ($P'(X)$) is always 0. On the other hand, if the output of OR gate is 1, the approximation errs (outputs 0 when some $P_j(X) = 1$) with a probability $1/3$:

$$\Pr(P'(X) = 0 \mid \vee_{i=1}^m P_i(X) = 1) = \frac{1}{3}$$

To verify that this holds, assume that we pick r_j at the end. Since $P_j(X) = 1$, there is exactly one value of r_j (among $\{0, 1, 2\} = \mathbb{Z}_3$) that makes the polynomial to output 0.

A few key observations of this approximation is that we have constructed a polynomial whose degree doubles instead of fan-in (m) of the OR gate. The error introduced by the approximation is atmost $1/3$. As with any randomized algorithm, we can repeat the above procedure for t independent trials and see if the output of at least one of the them is one.

Combining the formulations 1 and 2, the final approximate polynomial is:

$$P'(X) = P'_\alpha(P'_{\beta_1}(X), \dots, P'_{\beta_t}(X)), \quad (3)$$

where P'_α and P'_{β_i} are the applications of α -formulation and β -formulation (for i th trial) respectively. This resultant polynomial has a degree of $2t$ times the input as each application of equation β -formulation doubles the degree; and has an error of atmost $\mu = \frac{1}{3^t}$ as each of the individual t trials must result in an error.

AND (\wedge): We can simulate an AND gate using the NOT and OR gates. The resulting approximation $P'(X)$ has characteristics of the OR gate approximation ($2t$ degree increment and an error of μ).

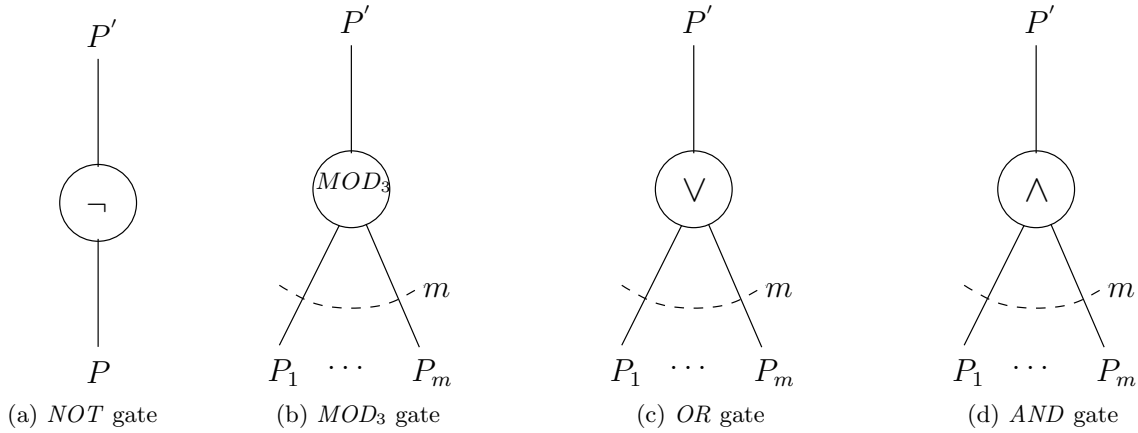


Figure 1: Approximation of Boolean gates

If the depth of the circuit is d , the degree of the polynomial $P(X)$ representing the entire circuit will be at most $(2t)^d$. $P(X)$ gives the wrong value only if the output of at least one of the gates in C is wrong. This happens with probability at most $\mu \cdot |C|$ (tighter bounds would depend on the number of AND/OR gates in C). If we consider all possible 2^n inputs (input size n), the expected number of inputs for which $P(X)$ will output the wrong value is atmost $(\mu \cdot |C|) \cdot 2^n$. This shows the existence of random \mathbb{Z}_3 coefficients for which $P(X)$ is wrong in no more than the expected number. More formally,

Lemma 1. *There exists a choice of r_i 's such that the number of input possibilities for which the approximate polynomial $P(X)$ is not equivalent to the behavior of the circuit is atmost $2^n \cdot \frac{|C|}{3^t}$.*

Proof. For every fixed input X and circuit C , the error probability at any OR/AND gate would be $\mu = \frac{1}{3^t}$. Hence the upper bound of the error in the whole circuit is $\mu \cdot |C|$ -

$$Pr(P(X) \neq C(X)) \leq \mu \cdot |C| = \frac{|C|}{3^t}$$

The expected number of errors given all 2^n input possibilities is -

$$E[\#X \text{ s.t. } (P(X) \neq C(X))] \leq 2^n \cdot \frac{|C|}{3^t}$$

Thus, for some choice of r_i 's we have that -

$$\#X \text{ s.t. } (P(X) \neq C(X)) \leq 2^n \cdot \frac{|C|}{3^t}$$

□

This construction could be generalized in any field \mathbb{Z}_p , for prime p , using the MOD_p gates. From Fermat's Little Theorem, we can use the fact that $a^{p-1} \equiv 1 \pmod{p} \forall a \in \mathbb{Z}_p \setminus \{0\}$ to ensure Boolean values for AND/OR gate approximations. This would result in an approximate polynomial of degree of at most $((p-1) \cdot t)^d$ and maximum error $\frac{1}{p^t}$.

Step 2: In this step, given a polynomial $P(X)$ of some degree that approximates \oplus_n on a subset G of inputs, we establish an upper bound for the degree of the approximate polynomial $P(X)$ 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 .

Initially, we transform the Boolean inputs to a convenient domain. Using a simple linear transformation from Boolean values $\{0, 1\} \mapsto \{1, -1\}$, we reduce \oplus_n to a simple product of literals.

We assume that this polynomial $P(X)$ has degree at most Δ and computes correct values on more than $(1 - \mu)$ fraction of inputs. Let us represent this set of good inputs as $G \subseteq \{-1, 1\}^n$ and hence $\frac{|G|}{2^n} \geq 1 - \mu$.

Lemma 2. *Let P be a polynomial of degree at most Δ that represents $\prod_{i=1}^n x_i$ in a set $G \subseteq \{-1, 1\}^n$. Then each function $f : G \rightarrow \mathbb{Z}_3$ 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 on n inputs has a multivariate polynomial of degree at most n . This is easy to see as we can represent 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 $\{1, 2, \dots, n\}$. Since we represent multiplication with ± 1 inputs, we can rewrite the monomial as:

$$\begin{aligned} \prod_{i \in I} x_i &= \left(\prod_{i \notin I} x_i^2 \right) \left(\prod_{i \in I} x_i \right) \\ &= \left(\prod_{i \notin I} x_i \right) \left(\prod_{i=1}^n x_i \right) \end{aligned} \tag{4}$$

$$\implies \prod_{i \in I} x_i = \left(\prod_{i \notin I} x_i \right) P(X) \tag{5}$$

Equation 4 holds for any input X of n bits as multiplication with squares of each ± 1 input does not change the value of the expression. Equation 5 is an approximation of multiplication and hence, holds only for the set of good inputs (G). The LHS of (5) has degree $|I|$ and the RHS has a degree at most $\Delta + |\bar{I}| = \Delta + n - |I|$. Choosing the minimum degree expression among them helps us express any monomial in Q' with a maximum degree of $\frac{n+\Delta}{2}$ (average of the degrees). \square

Further, we combine the result of Lemmas 1 and 2 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 $f : G \rightarrow \mathbb{Z}_3$ for some good input set (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 , and hence

$$M = \sum_{i=0}^{\frac{n+\Delta}{2}} \binom{n}{i}$$

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

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

Thus,

$$\begin{aligned} M &= 2^{n-1} + 2^n \cdot \Theta\left(\frac{\Delta}{\sqrt{n}}\right) \\ &= 2^n \left(\frac{1}{2} + \Theta\left(\frac{\Delta}{\sqrt{n}}\right)\right) \end{aligned}$$

The number of functions of the form $G \rightarrow \mathbb{Z}_3$ is $3^{|G|}$ as each element of G is assigned to one of 3 possible values of \mathbb{Z}_3 . As 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 .

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

From Lemma 1, $\mu \leq \frac{1}{3^t}$ when $\Delta = (2t)^d$. Thus,

$$1 - \frac{|C|}{3^t} \leq 1 - \mu \cdot |C| \leq \frac{1}{2} + \Theta\left(\frac{(2t)^d}{\sqrt{n}}\right) \quad (6)$$

$$\implies |C| \geq 3^t \left[\frac{1}{2} - \Theta\left(\frac{(2t)^d}{\sqrt{n}}\right) \right] \quad (7)$$

For equation (6) to be meaningful, the RHS should be less than 1. Thus, setting $(2t)^d = O(\sqrt{n})$ gives an appropriate value for the RHS in the equation (7). Thus, $t = \Theta(n^{\frac{1}{2d}})$ and this gives $|C| \geq 2^{\Omega(n^{\frac{1}{2d}})}$. \square

The only part of the above analysis that changes when working over \mathbb{Z}_p rather than \mathbb{Z}_3 is that $\Delta = ((p-1) \cdot t)^d$ 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 for distinct primes p and 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.

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

Corollary 1. *A depth d unbounded fan-in circuit that agrees with parity (\oplus_n) 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. Let C be a circuit that is correct on at least $(\frac{1}{2} + \rho)$ of the inputs for $\rho > 0$. 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,

$$\frac{1}{2} + \rho - \frac{|C|}{3^t} \leq \frac{1}{2} + \Theta\left(\frac{(2t)^d}{\sqrt{n}}\right) \implies \rho - \frac{|C|}{3^t} \leq \Theta\left(\frac{\Delta}{\sqrt{n}}\right) \quad (8)$$

$$\implies |C| \geq 3^t \left[\rho - \Theta\left(\frac{(2t)^d}{\sqrt{n}}\right) \right] \quad (9)$$

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 equation (9). Hence, $t = \Theta(n^{\epsilon/(2d)})$, and we get that $|C| \geq 2^{\Omega(n^{\epsilon/(2d)})}$. \square

The above corollary proves the inapproximability of the parity function using constant-depth circuits. There is a stronger result that can be proved using random restrictions.

3 Random Restrictions method

In this section, we provide an alternate proof to Theorem 1. The bound for the circuit size is tighter than what was achieved by the polynomial approximations method in Section 2. However, the proof does not allow the circuit to contain additional MOD_3 gates.

A *restriction* fixes some inputs of a circuit to constant values, and leaves other inputs free. More formally, we define random restrictions as the following:

Definition 2 (Random Restrictions). *A p -random restriction on n variables is a random function $\rho : \{x_1, \dots, x_n\} \rightarrow \{*, 0, 1\}$, such that for each i , independently, $\Pr[\rho(x_i) = *] = p$ and $\Pr[\rho(x_i) = 0] = \Pr[\rho(x_i) = 1] = \frac{1-p}{2}$. If $\rho(x_i) = *$, then we leave x_i as a variable. Otherwise we set it to the result of $\rho(x_i)$.*

Note that if we apply a random restriction to a parity function, we get a parity function or its complement depending on the inputs corresponding to $*$.

Theorem 2. *Given any circuit C_d of depth d that computes \oplus_n , its size is at least $2^{\Omega(n^{\frac{1}{d-1}})}$*

Proof. To prove this theorem, let us assume a neat structure for the circuit C_d that we consider -

- NOT (\neg) gates don't count towards the size or depth of the circuit.
- The circuit gates are organized in alternating layers of AND and OR gates.

We further show that these assumptions don't affect the bounds.

Lemma 3. *If C_d is a circuit of size $|C_d|$ and depth d , then there is a circuit C'_d of size at most $2 \cdot |C_d|$ and depth d that computes the same function and such that all the NOT gates are applied at the input level.*

Proof. We prove the stronger statement that, for every gate g of C_d , there is a gate g' in C'_d whose output is the complement of g . Then we just let the output of C'_d be the complement of the output gate of C_d . Let us order the gates of C_d as $g_1, \dots, g_{|C_d|}$ in such a way that if the gate g_i uses the output of gate g_j as an input then $j < i$. We describe an inductive construction. In the base case, if g_1 is an AND gate (OR gate), then g'_1 is an OR gate (AND gate), whose inputs are the complements of the inputs of g_1 . It follows from De Morgan's law that the output of g'_1 is the complement of the output of g_1 . In the inductive step, if we have constructed gates g'_1, \dots, g'_i whose outputs are the complement of g_1, \dots, g_i , then g'_{i+1} has similar complemented inputs as g_{i+1} and from De Morgan's law, it follows that the output of g'_{i+1} is the complement of the output of g_{i+1} . \square

Lemma 4. *If C_d is a circuit of size $|C_d|$ and depth d , then there is a circuit C'_d of size at most $d \cdot |C_d|$ and depth d that computes the same function and such that the gates are arranged in d sequential layers which have alternating AND and OR gates.*

Proof. The associativity of AND and OR ensures that every input-output path in the circuit has an alternation between AND gates and OR gates. If there is an AND gate (OR gate) g in the circuit one of whose inputs is coming from another AND gate (OR gate) g' : then we can connect the inputs of g' directly to g . This does not alter the size or depth of the circuit. Repeatedly merging the similar gates, we eventually get an alternating circuit C'_d which has the same size and depth of C_d .

Finally, we arrange the gates in d layers so that each layer is entirely AND or OR gates, and wires go only from lower-numbered layers to higher-numbered layers. Finally, we replace all wires that jump many layers by a path of alternating fan-in 1 AND gates and fan-in 1 OR gates. This step increases the size of the circuit by atmost a factor of d . \square

Further, we use the following lemma (Switching Lemma) to establish a contradiction for a polynomial-sized circuit C_d that decides PARITY.

Lemma 5 (Switching Lemma). *Let φ be a k -CNF formula. We then apply a p -random restriction that leaves a fraction p of variables unassigned. For the remaining fraction $(1-p)$ of the variables, independently hardwire 0 or 1 based on the random restriction.*

Then, for every k' ,

$$\Pr[\varphi|_p \text{ cannot be expressed as a } k'\text{-DNF}] \leq (5pk)^{k'}$$

Note that the statement is trivial if the restriction can set all variables ($p = 1$). For practical purposes, the values of p , k and k' should concur with the inequality - $|C| \cdot (5pk)^{k'} < 1$.

Proof Idea. Consider an AND of ORs, where each OR has size at most k . Notice that a random restriction is not very likely to set an OR to 0 since all literals involved need to be set to 0 for that to happen. But if k is small, there is a nontrivial probability that this happens. There are two cases:

1. There are a large number of pairwise disjoint ORs. In that case, there are many independent events that can set the AND gate to 0, namely each of those pairwise disjoint ORs being set to 0. Since each of those events happens with a nontrivial probability, the odds are that the random restriction will set the AND gate to 0, in which case it can trivially be written as a DNF with small bottom fan-in.
2. There is not a large number of pairwise disjoint ORs. Let V be a minimal set of variables such that each OR queries at least one variable from V . Since there is a lot of overlap among the ORs, V is small. If we query all the variables in V , then we have essentially reduced our problem to a simpler one of the same type, namely the transformation of a CNF with bottom fan-in at most $k - 1$. This is because each of the ORs contains at least one literal in V . We then repeat the case distinction to that simpler problem, depending on the setting of the variables in V .

Along every branch of this process, we will eventually end up in case 1. Since there are at most k steps and each step involves querying a small number of variables, we end up with a decision tree of small depth that represents the given CNF under a random restriction with high probability. A decision tree of small depth can be turned into a DNF with small bottom fan-in by writing down an OR over all paths in the decision tree that lead to acceptance of the AND of all the conditions that define the path. \square

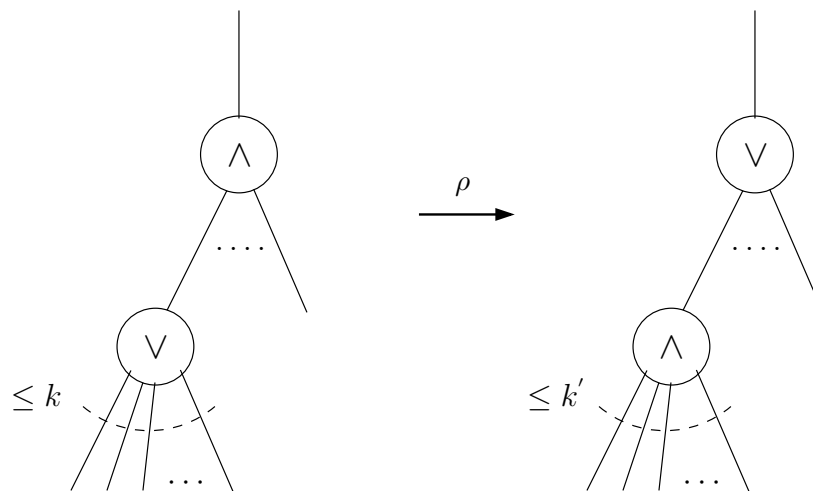


Figure 2: k -CNF to k' -DNF

Given a circuit C_d that decides PARITY, we can rewrite it in terms of alternate levels of AND or ORs without much change in size (Lemmas 3 and 4). Note that we can also switch from a DNF to a CNF by considering the negation of the circuit.

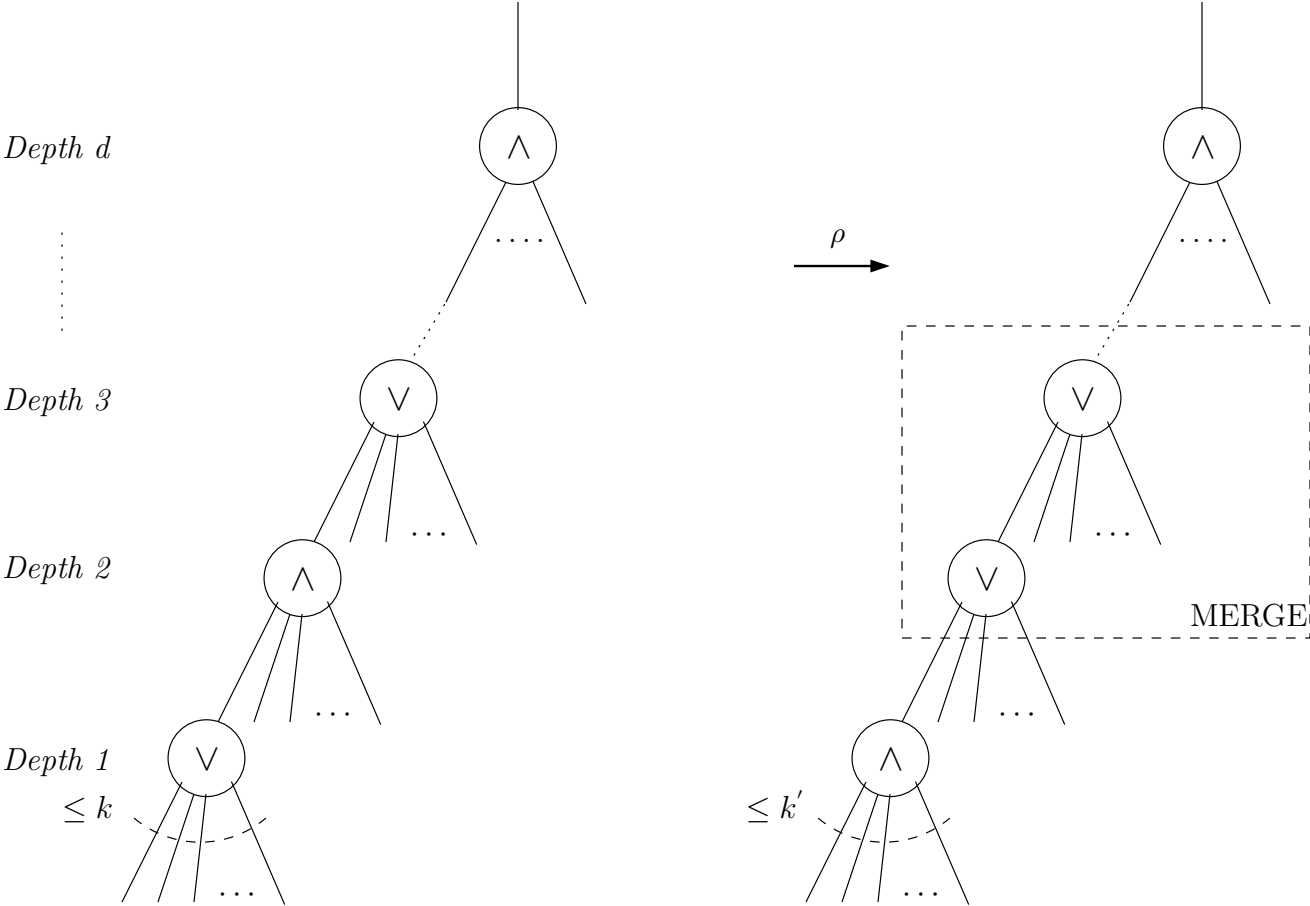


Figure 3: Decreasing the depth by merging adjacent levels while maintaining a small bottom fan-in

Next, we use repeated invocations of Switching Lemma to reduce the circuit depth by 1 in each iteration (as shown in Figure 3) until we are left with a circuit of depth 2. Suppose that the bottom gates are ANDs (ORs). To apply the switching lemma, we need to ensure the gates at the bottom of the circuit have small fan-in. To ensure this, we insert dummy OR (AND) gates below the AND (OR) gates. Namely, for each input x to the AND (OR) gate, we replace that with x OR x . Now, we apply the switching lemma to the AND of ORs (OR of ANDs) we have created. With high probability, each application is successful in creating an OR of ANDs (AND of ORs) with small bottom fan-in and without setting too many variables. Now the second bottom-most and third bottom-most levels are both ORs (ANDs) and can be merged. This reduces the depth of the circuit by 1 (back down to d since we added a level initially) without adding extra gates.

Now the circuit still has small bottom fan-in, so we can apply the switching lemma again. We repeat this process until we get a circuit C' of depth 2. At this point, if C computed \oplus_n , then C' computes \oplus_m on some m -subset of the variables (those that were unset by the random restrictions; $E(m) = n \cdot p^{d-1}$).

This number of unset variables m is larger than the small bottom fan-in of the resulting k -CNF (or k -DNF) circuit C' and hence, PARITY cannot be decided by C' . Therefore, we arrive at a contradiction. \square

Corollary 2. *The size of circuit C_d of constant depth d required to compute \oplus_n correctly on at least $(\frac{1}{2} + \frac{1}{2^{n^{1/d}}})$ fraction of the inputs, is at least $2^{\Omega(n^{\frac{1}{d}})}$.*

$$C_d(\oplus_n) \geq 2^{\Omega(n^{\frac{1}{d}})}$$

Proof. The proof follows similar ideas as that of Corollary 1, but is more involved. \square

4 Next Lecture

In the next lecture, we will focus on parallelism. We will use circuits to model parallel computations.

5 References

M. Furst, J.B. Saxe, M. Sipser. *Parity, circuits, and the polynomial-time hierarchy*. Mathematical Systems Theory, 17(1), p.13-27, 1984..

J. Hastad. *Computational limitations for small depth circuits*. MIT Press, 1986. Ph.D. thesis.

N. Linial, Y. Mansour, N. Nisan. *Constant depth circuits, Fourier transform, and learnability*. In Proceedings of the 30th Annual Symposium on Foundations of Computer Science (FOCS 89). p.574-579.

A.A. Razborov. *Lower bounds on the size of bounded depth networks over a complete basis with logical addition*. Matematicheskie Zametki, 41, p.598-607, 1987.

Roman Smolensky. *Algebraic methods in the theory of lower bounds for boolean circuit complexity*. In Proceedings of the 19th ACM Symposium on Theory of Computing, pages 77-82, 1987.

Acknowledgements

In writing the notes for this lecture, I perused the notes by Seeun William Umboh and Piramanayagam Arumuga Nainar for Lecture 08 and Lecture 09 from the Spring 2007 offering of CS 810, and the notes by Chris Hopman and Phil Ryzdewski for Lecture 08 and Lecture 09 respectively from the Spring 2010 offering of CS 710.