Lecture 31

Lower Bounds for Constant Depth Circuits

In this lecture we present the details of the result of Furst, Saxe, and Sipser [46] that PARITY is not in AC^0.

Recall that a formula or circuit is in *t-conjunctive normal form* (*t-CNF*) if it is a conjunction of clauses, each clause a disjunction of at most t literals, each literal a variable or its negation. Dually, a formula or circuit is in *t-disjunctive normal form* (*t-DNF*) if it is a disjunction of terms, each term a conjunction of at most t literals. We can assume without loss of generality that no term or clause contains a pair of complementary literals. By convention, the empty conjunction is equivalent to 1 and the empty disjunction is equivalent to 0.

Given a parity circuit of constant depth d, we can assume without loss of generality that the gates are arranged in levels with variables and their negations at level 0 and other levels alternating between disjunctions and conjunctions. We can also assume that the gates at level 1 are disjunctions; if not, consider the dual circuit instead, which is also a parity circuit.

The *length* of a term or clause M, denoted $|M|$, is the number of literals in M. We often omit the symbol $\wedge$ in terms and write $\overline{x}$ for $\neg x$. Thus $\overline{x}y\overline{z}$ means the same as $\neg x \wedge y \wedge \neg z$.

A *partial valuation* of a set X of Boolean variables is an assignment of constants to some of the variables of X, possibly leaving some variables unassigned. Formally, a *partial valuation* on X is a map $\rho : X \to X \cup \{0, 1\}$

such that for each $x \in X$, $\rho(x) \in \{x, 0, 1\}$. We say that x is *unassigned* under ρ if $\rho(x) = x$.

Any partial valuation ρ on X extends to a function on Boolean formulas over X in a natural way, replacing each variable x with $\rho(x)$ and then simplifying wherever possible using the Boolean algebra axioms $0 \vee x = x$, $0 \wedge x = 0$, $1 \vee x = 1$, $1 \wedge x = x$. For example, if $\rho(x) = 1$, $\rho(y) = 0$, and $\rho(z) = z$, then

$$\rho((x \vee y) \wedge (\overline{x} \vee z)) \quad = \quad z.$$

For the remainder of this lecture, we consider randomly chosen partial valuations in which each variable is independently assigned 0 or 1, each with probability $(1 - 1/\sqrt{n})/2$, or left unassigned with probability $1/\sqrt{n}$.

The central idea of the proof is that all the CNF subcircuits at level 2 will very likely become equivalent to DNF circuits after applying a finite number of partial valuations chosen randomly according to this distribution. Thus the chances are good that we will be able to replace all the CNF gates at level 2 with DNF gates, then absorb the disjunctions at level 2 into the disjunctions at level 3, thereby reducing the depth by one level. Continuing in this fashion, we will be able to get rid of all levels except two.

Lemma 31.1 *After a random partial valuation, the probability that there are fewer than $\sqrt{n}/2$ unassigned variables is at most $(2/e)^{\sqrt{n}/2}$.*

Proof. Each variable remains unassigned with probability $1/\sqrt{n}$. There are n variables in all, so the mean number of unassigned variables is $n/\sqrt{n} = \sqrt{n}$. The result now follows immediately from the Chernoff bound (I.7) with $\mu = \sqrt{n}$ and $\delta = 1/2$. □

Lemma 31.1 is important, because it says that after the application of a random partial valuation, it is very likely that there are still enough unassigned variables left that the size of the circuit is still polynomial in the number of inputs.

The following are two technical lemmas that are used in our main development.

Lemma 31.2 *Let c be a constant, and let A be any set of variables of size at least c but $o(n^{1/c})$. The probability that a random partial valuation leaves more than c variables of A unassigned is bounded above by $n^{1-c/2}$ for sufficiently large n.*

Proof. Let X be a random variable representing the number of variables of A left unassigned by the random partial valuation. We wish to estimate $\Pr(X \geq c)$. If $s = |A|$, then the expected number of unassigned variables

is $\mu = sn^{-1/2}$. Plugging this into the Chernoff bound (I.6), for sufficiently large n,

$$\begin{aligned}\Pr(X \geq c) &< e^{c-\mu}(\mu/c)^c \leq (e\mu/c)^c \\ &= (esn^{-1/2}/c)^c = n^{-c/2}(es/c)^c \leq n^{1-c/2},\end{aligned}$$

the last inequality by the assumption that s is $o(n^{1/c})$. □

Lemma 31.3 *Let a and b be constants. Let S be any set of pairwise disjoint sets of variables such that S has size at least $b \log n$ and all elements of S have size less than a. For $A \in S$, let $E(A)$ be the event that a random partial valuation assigns 0 to all variables in A. For sufficiently large n, the probability that $E(A)$ does not occur for any $A \in S$ is bounded above by $n^{b \log(1-2^{-a})}$.*

Proof. Let $A \in S$ and let $s = |A|$. For sufficiently large n, the probability of $E(A)$ is $2^{-s}(1-n^{-1/2})^s \geq 2^{-s}/2 \geq 2^{-a}$. The probability that $E(A)$ does not occur is thus bounded by $1-2^{-a}$. Because the elements of S are pairwise disjoint, the events $E(A)$ are independent, therefore the probability that $E(A)$ fails for all $A \in S$ is the product of the probabilities that it fails for each of them, which is bounded by $(1-2^{-a})^{|S|}$. But

$$(1-2^{-a})^{|S|} \leq (1-2^{-a})^{b \log n} = n^{b \log(1-2^{-a})}.$$

□

Now we are ready to give the main part of the argument, which we have split into three lemmas. Let t be a constant. Call a circuit a *t-circuit* if every level-1 gate is of degree at most t; that is, if every level-2 gate is a t-CNF circuit.

Lemma 31.4 *If* Parity *has polynomial-size circuits of depth d, then* Parity *has polynomial-size t-circuits of depth d for some constant t.*

Lemma 31.5 *If* Parity *has polynomial-size t-circuits of depth $d \geq 3$ and $t \geq 1$, then* Parity *has polynomial-size $(t-1)$-circuits of depth d.*

Lemma 31.6 *If* Parity *has polynomial-size 1-circuits of depth $d \geq 1$, then* Parity *has polynomial-size circuits of depth $d-1$.*

Proof of Lemma 31.4. Suppose Parity has circuits of depth d and size n^k. Consider a random partial valuation ρ on the variables of the nth circuit. Let $t = 2k+4$ and $b = (k+1)/\log(3/2)$. For some level-1 clause C, consider the event $|\rho(C)| > t$; that is, more than t literals of C remain unassigned. We show that for sufficiently large n, this event occurs with probability at most $n^{-(k+1)}$.

There are three cases, depending on the size of C.

Case 1 If $|C| \leq t$, then the probability is already 0, so we are done.

Case 2 If $t \leq |C| \leq b \log n$, then by Lemma 31.2, for sufficiently large n, the probability that a random partial valuation leaves more than t literals in C unassigned is bounded above by $n^{1-t/2} = n^{-(k+1)}$.

Case 3 The last case is $|C| \geq b \log n$. If any literal of C is assigned 1, then $\rho(C) = 1$, so $|\rho(C)| = 0$. Thus the probability that there are at least t literals remaining unassigned is bounded by the probability that no literal in C is assigned 1. To calculate this probability, note that for sufficiently large n, the probability that any particular literal is not assigned 1 is

$$1 - \frac{1}{2}(1 - \frac{1}{\sqrt{n}}) \quad = \quad \frac{1}{2}(1 + \frac{1}{\sqrt{n}}) \quad \leq \quad \frac{2}{3}.$$

These events are independent, thus the probability that no literal in C is assigned 1 is at most

$$\left(\frac{2}{3}\right)^{|C|} \quad \leq \quad \left(\frac{2}{3}\right)^{b \log n} \quad \leq \quad n^{b \log(2/3)} \quad = \quad n^{-(k+1)}.$$

We have shown that for sufficiently large n, for each level-1 clause C, the probability that $|\rho(C)| > t$ is at most $n^{-(k+1)}$. By the law of sum, the probability that there exists a level-1 clause with more than t literals is bounded by the sum of these probabilities. Because there are at most n^k level-1 clauses in all,

$$\Pr(\exists C \ |\rho(C)| > t) \quad \leq \quad \sum_C \Pr(|\rho(C)| > t) \quad \leq \quad n^k n^{-(k+1)} \quad = \quad n^{-1},$$

which is vanishingly small.

Now by Lemma 31.1, the probability that there are fewer than $\sqrt{n}/2$ unassigned variables is also vanishingly small. Again by the law of sum, the probability that either event occurs in a single random partial valuation is vanishingly small. That is, with probability tending to 1, the random partial valuation leaves at least $\sqrt{n}/2$ variables unassigned and knocks all level-1 gates down to degree at most t. As the probability of this event is nonzero, there must be a partial valuation that realizes it. By making this partial valuation (and by assigning a few other inputs if necessary), we obtain a circuit for PARITY on $\sqrt{n}/2$ variables and all level-1 gates of degree at most t. These circuits are still polynomial-size, because polynomial in n is still polynomial in $\sqrt{n}/2$. □

Proof of Lemma 31.5. Suppose PARITY has t-circuits of depth $d \geq 3$, $t \geq 1$, and size n^k. Let S be the set of clauses in some level-2 conjunction.

Let T be a maximal subset of S such that no two clauses of T contain the same variable, either positively or negatively. By *maximal*, we mean that there is no proper superset of T satisfying this property. Such a set T can be constructed by considering all the clauses in S in some order, taking the next clause into T if it does not have a variable in common with any of the clauses previously taken.

Let $a = 2k + 4$ and $b = -(k + 1)/\log(1 - 2^{-a})$. Consider the effect of a random partial valuation on the variables of T. We consider two cases.

Case 1 If T contains at least $b \log n$ elements, then we have at least $b \log n$ clauses, no two of which share a variable. By Lemma 31.3, for sufficiently large n, some clause in T receives all 0 with probability at least $1 - n^{b \log(1-2^{-a})} = 1 - n^{-(k+1)}$, in which case the entire conjunction at level 2 disappears.

Case 2 If T contains at most $b \log n$ elements, then $\bigcup T$ contains at most $bt \log n$ elements and shares a variable with all clauses in S (otherwise T was not maximal). By Lemma 31.2, for sufficiently large n, the probability that a random partial valuation leaves more than a variables of $\bigcup T$ unassigned is bounded above by $n^{1-a/2} = n^{-(k+1)}$. Thus with very high probability, there are at most $2a$ literals $\ell_1, \ldots, \ell_{2a}$ of $\bigcup T$ unassigned, and every clause of $\rho(S)$ that is still of size t must contain one of these literals. Let φ_0 be the conjunction of all clauses of $\rho(S)$ of size at most $t-1$. Of the remaining clauses of $\rho(S)$, let φ_j be the conjunction of those containing the literal ℓ_j and no literal ℓ_i for $i < j$, and let φ'_j be φ_j with all occurrences of ℓ_j deleted. Then φ_j is equivalent to $\ell_j \vee \varphi'_j$, and the original conjunction after the partial evaluation is equivalent to

$$\bigwedge_{j=0}^{2a} \varphi_j \quad \equiv \quad \varphi_0 \wedge \bigwedge_{j=1}^{2a} (\ell_j \vee \varphi'_j).$$

Using the distributive laws of Boolean algebra, this can be expressed as a disjunction of at most 2^{2a} $(t-1)$-CNF circuits of the form $\varphi_0 \wedge \psi_1 \wedge \cdots \wedge \psi_{2a}$, where each ψ_i is either ℓ_i or φ'_i.

We have shown that for sufficiently large n, with probability at least $1 - n^{-(k+1)}$, any t-CNF gate at level 2 becomes equivalent to a disjunction of a constant number of $(t-1)$-CNF gates under a random partial valuation. This disjunction can be merged with the disjunctions at level 3 to give a $(t-1)$-circuit.

The remainder of the proof proceeds as in Lemma 31.4. Briefly, the probabilities of each of these events and the event that there remain at least $\sqrt{n}/2$ unassigned variables tend to 1 sufficiently fast that they all occur simultaneously with nonzero probability. □

Proof of Lemma 31.6. This is the easy one. A polynomial-size 1-circuit of depth $d \geq 1$ is equivalent to a circuit of depth $d-1$ simply by bypassing the singleton gates at level 1. (In fact, this is just a special case of Lemma 31.5 for $t = 1$.) □

Lemma 31.7 *There is no $(n-1)$-CNF or $(n-1)$-DNF parity circuit on n inputs.*

Proof. An $(n-1)$-CNF circuit on n inputs is a conjunction of clauses with at most $n-1$ literals per clause. Any partial valuation of a parity circuit is a parity circuit on the remaining variables. But by setting at most $n-1$ variables, we can make all the literals in some clause 0, thus the whole circuit has constant value 0 regardless of the values of the remaining variables, so it cannot be a parity circuit. The argument for $(n-1)$-DNF circuits is similar. □

Combining Lemmas 31.4–31.7, we have

Theorem 31.8 (Furst, Saxe, and Sipser [46]) $\textsc{Parity} \notin AC^0$.

Proof. By repeating the constructions of Lemmas 31.4–31.6, we could start with any family of circuits for PARITY of constant depth and polynomial size and reduce them to a family of circuits for PARITY of depth 2, polynomial size, and constant indegree at level 1. These circuits would be t-DNF or t-CNF circuits for some constant t. But this is impossible by Lemma 31.7. □