6.841: Advanced Complexity Theory Fall 2012

Lecture 4 — September, 18, 2012

Prof. Dana Moshkovitz *Scribe: Ilya Razenshteyn*

1 Overview

In the last lecture we proved that for every k there exists a language in Σ_3 that does not have circuits of size $O(n^k)$.

In this lecture we present a (rather weak) class of circuits AC^0, and show that it can not compute parity. That is, we separate AC^0 and P.

2 AC^0 versus PARITY

Let us consider the following complexity class. Let AC^0 be the set of all languages that can be decided using polynomial-sized constant-depth circuits with gates from the set $\{\wedge, \vee, \neg\}$. In order for this class to be meaningful we allow gates "$\wedge$" and "$\vee$" to have polynomial fan-in (otherwise, any language decidable with such a circuit would depend only on a constant number of variables).

It turns out that for AC^0 we can prove a rather strong lower bound. Let $\mathrm{PARITY}(x_1, x_2, \ldots, x_n) := x_1 \oplus x_2 \oplus \ldots \oplus x_n$.

Theorem 1 (Ajtai [Ajt83], Furst–Saxe–Sipser [FSS84])**.** $\mathrm{PARITY} \notin \mathrm{AC}^0$

2.1 The proof

Our general proof strategy will be as follows. We show that if $f \in \mathrm{AC}^0$, then there is a restriction of all except $n^{\Omega(1)}$ variables such that the restricted version of f has depth-2 polynomial-sized circuit with polynomial fan-ins. Then we show that for PARITY this is not the case.

The key ingredient is the so-called Switching Lemma. To state it we need a notion of *p-restrictions*. Say we have a function $f\colon \{0,1\}^n \to \{0,1\}$. Then we say that a function f' is a p-restriction of f, if it is obtained from f via the following process. Each variable is set to zero with probability $(1-p)/2$, and is set to one with the same probability (so, with probability p a variable is not set).

Theorem 2 (Switching Lemma)**.** *Let c be any constant. Then there exists $b = b(c)$ with the following properties. Suppose a function $f\colon \{0,1\}^n \to \{0,1\}$ has s-size DNF with at most c variables per clause. Let f' be an $n^{-1/2}$-restriction of f. Then, with probability at least $1 - 1/10s$, f' depends on at most b variables, provided that n is sufficiently large.*

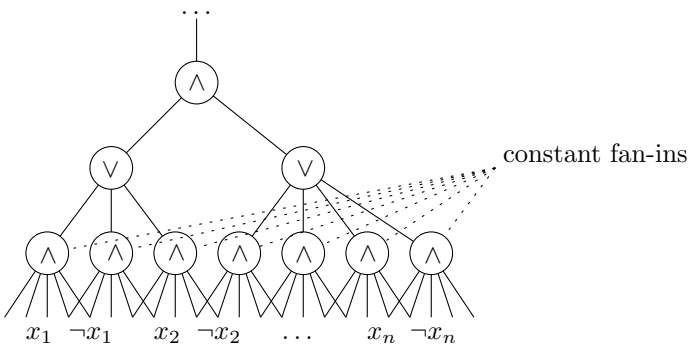
Remark: this theorem is called "Switching Lemma", because it implies that f' can be represented as an $\exp(b)$-size CNF with at most b variables per clause. Thus, we *switch* from DNF to CNF with a small blow-up by restricting variables at random.

Let us first derive the main theorem from Switching Lemma.

Switching Lemma ⇒ the main theorem

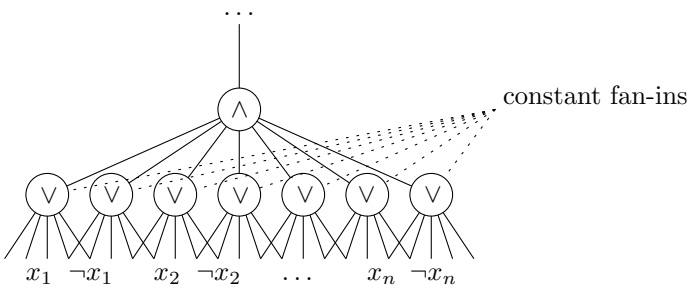
Suppose that PARITY $\in \mathrm{AC}^0$. Let C be the corresponding circuit for n-bit inputs. First, let us transform C as follows (see the picture):

- By duplicating parts of C, let us assume that all gates have fan-out 1;
- By propagating negations, let us assume that negations can only be adjacent to variables: that is, our circuit has $2n$ "inputs": x_1, $\neg x_1$, x_2, $\neg x_2$, …, x_n, $\neg x_n$.
- By merging adjacent similar gates we ensure that ANDs and ORs are grouped into layers: one layer of ANDs, then one layer of ORs, then, again, one layer of ANDs, and so on. The layer adjacent to variables consists of ANDs.
- By introducing dummy gates we maintain the constant fan-in for the first layer.



If we consider the first two layers, we can see that they form a bunch of DNFs. Let us make an $n^{-1/2}$-restriction, then using Switching Lemma we can switch to CNFs with constant clause sizes, and then merge two successive layers of ANDs. By union bound switching succeedes with probability at least 1/10.

Now we are left with a circuit with depth smaller than the original one. We pay for it with at most constant blow-up in size, and we are left with $\Theta(n^{1/2})$ variables with probability exponentialy close to 1 (it follows from Chernoff Bound). Note that the bottom fan-ins remain constant.



Now we again apply the same idea. We have a bunch of CNFs. We switch them to DNFs by taking a $n^{-1/4}$-restriction, and applying Switching Lemma (note that, despite Switching Lemma formally claims to switch DNFs to CNFs, we can apply it in the other direction as well, since small DNFs are negations of small CNFs, and vice versa).

If the depth of the initial circuit is d, then after at most d such iterations we are left with an $n^{-\Omega(2^{-d})}$-restriction of the original function, which can be computed with poly-size depth-2 circuit. But since the initial function was PARITY, we derive that there exists a $n^{O(1)}$-size depth-2 circuit that computes PARITY on $n^{\Omega(1)}$ variables (all restrictions of PARITY are either PARITY or its negations). The following exercise refutes such possibility.

Exercise 3. *Any depth-2 circuit for computing $x_1 \oplus x_2 \oplus \ldots \oplus x_n$ must have size at least $2^{\Omega(n)}$.*

Thus, we obtain a lower bound $\exp(\Omega(n^{2^{-d}}))$ on a size of depth-d circuits for parity.

It is "only" left to prove Switching Lemma.

Proof of Switching Lemma

Suppose we have a c-DNF on n variables that consists of clauses $T_1, T_2, \ldots, T_k$, and we want to find a constant b such that whenever we make an $n^{-1/2}$-restriction we get a function that depends only on b variables with probability at least $1 - 1/n^{O(1)}$.

The main idea is to consider two cases. Wlog we can assume that $\{T_1, T_2, \ldots, T_l\}$ is the maximum set of clauses that are pairwise disjoint. Let l^* be some parameter that we will choose later.

First Case. Suppose that $l \geq l^*$. Then, there is a substantial probability that at least one of the clauses $T_1, T_2, \ldots, T_l$ will be identically equal to one, and, as a result, the whole DNF will shrink to a constant function.

The probability that this collapse will not occur is at most

$$\left(1 - \left(\frac{1 - n^{-1/2}}{2}\right)^c\right)^l.$$

If we choose l^* to be equal to $\alpha 2^c \log n$, where α is a sufficiently large constant, then we can make this probability to be smaller than any inverse polynomial (α depends only on the degree of this polynomial).

Second Case. Suppose that $l < l^*$. Then $T_1, T_2, \ldots, T_l$ consist of at most cl variables. Using Chernoff Bound one can see that the probability that at least βc of those variables survive the restriction is at most $O((n^{-1/2}l)^{\beta c})$.

Again, by taking β to be sufficiently large constant, we can make this error to be smaller than any inverse polynomial (β depends on its degree and on c).

Suppose that at most βc variables survived. Then, the crucial insight is that if we fix these variables to some values, then **we get a function that is representable with $(c-1)$-DNF**! This is because we chose $T_1, T_2, \ldots, T_l$ as a maximum set of disjoint clauses, so, each of the other clauses share at least one variable with $T_1, T_2, \ldots, T_l$. That is, we can fix these variables to all possible sets of values, and invoke Switching Lemma for $c - 1$ with smaller probability of failure (by a constant factor).

To sum up, we have that with probability at most (any) inverse polynomial we get a function that depends on more than $b(c) := \beta c + 2^{\beta c} \cdot b(c-1)$ variables, where $b(c-1)$ is a constant from the Switching Lemma statement. So, by induction, we have $b(c) = \exp(c^2)$.

Remark: the reason why we need to take p polynomially small is to get the polynomially

small bound on the probability that more than some constant number of variables will survive in $T_1, T_2, \ldots, T_l$ during the p-restriction.

Remark: the lower bound $\exp(\Omega(n^{2^{-d}}))$ on a depth-d circuit size for parity we got is far from being tight. The tight lower bound $\exp(\Omega(n^{1/d}))$ is due to Håstad [Hås86]. He came up with a much stronger version of Switching Lemma.

References

[Ajt83] M. Ajtai. Σ_1^1-Formulae on finite structures. *Annals of Pure and Applied Logic*, 24(1):1–48, 1983.

[FSS84] Merrick L. Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17(1):13–27, 1984.

[Hås86] Johan Håstad. Almost optimal lower bounds for small depth circuits. In *STOC*, pages 6–20, 1986.