Previous lecture and this lecture

We want to prove this theorem

**Theorem 1** (Razborov). *Every circuit of depth \( c \) (constant) that computes the *majority* function using AND, OR, NOT, and PARITY gates with *unbounded fan-in*, must have exponential size (in the number of bits).*

The structure of the proof is something like this:

- Circuits (constant depth) \( \approx \) Polynomials (small degree) \( \neq \) Majority
- same on “many” inputs
- different on “many” inputs
We count on how many inputs the two functions differ:

\[ \text{diff}(f, g) \triangleq |\{ y \in \{0, 1\}^n \mid f(y) \neq g(y)\}| \]

In the previous lecture we have seen this:

**Lemma 2** (polynomials vs \( k \)-threshold). For \( n/2 \leq k \leq n \), every polynomial of degree \( \delta \leq 2k - n - 1 \) must differ from the threshold function on at least \( \binom{n}{k} \) inputs.

In this lecture we shall prove the following:

**Lemma 3** (circuits vs polynomials). For every depth-\( d \) circuit with AND, OR, NOT and PARITY gates (of unbounded fan-in) there exists a polynomial \( p \) of degree \( \delta \leq r^d \) which differs from \( C \) on at most \( |C|2^n/2^r \) inputs.

We can already see how these two results imply Theorem 1:

- Take \( k = n/2 + \sqrt{n} \) so that \( \binom{n}{k} \geq 2^n/\sqrt{n} \).
- Set \( r \approx n^{1/2d} \) so that \( \delta \leq 2k - n - 1 \).

Then

\[ \text{diff}(p, f_k) \leq \text{diff}(p, C) + \text{diff}(C, f_k) \]

If \( C \) computes the threshold function \( f_k \) then \( \text{diff}(C, f_k) = 0 \) and the two lemmas tell us

\[ \frac{2^n}{\sqrt{n}} \leq \text{diff}(p, f_k) \leq \text{diff}(p, C) \leq |C|\frac{2^n}{2^r} \]

and thus

\[ |C| \geq \frac{2^r}{\sqrt{n}} \approx \frac{2^{n^{1/2d}}}{\sqrt{n}} \]

Since any circuit computing **majority** can be used to compute \( f_k \) by adding some “dummy inputs 0”:

**Depth-d circuits** (using AND, OR, NOT and PARITY gates of unbounded fan-in) that compute the majority function must have size \( 2^{\Omega(n^{1/2d})} \).
1 Proof of Lemma 3 – probabilistic argument

We shall approximate every bounded depth circuit by some low-degree polynomial. Suppose the circuit we want to approximate is the AND of \( n \) variables
\[
p(x) = x_1 \cdots x_n
\]
and for this purpose we use this (“strange”) polynomial over \( \mathbb{F}_2 \):
\[
\hat{p}(x) = 1 + (1 + x_1) + \cdots + (1 + x_n)
\]
The answer is correct for the input \( x = 1 \)
\[
p(1) = 1 = \hat{p}(1)
\]
Now pick a random subset \( R \) of the variables by including the \( i^{th} \) variable with probability \( 1/2 \) independently from the other variables and look at
\[
\hat{p}_R \triangleq 1 + \sum_{i \in R} (1 + x_i)
\]

Exercise 1. Show that, for every input \( a \neq 1 \)
\[
\Pr_R[\hat{p}_R(a) = 1] = 1/2
\]

Hint: We are in the case \( a_i = 0 \) for some \( i \).

Repeat \( r \) times

Pick \( r \) subsets \( R = \{R_1, \ldots, R_r\} \) at random independently as above and look at
\[
\hat{p}_R(x) \triangleq \hat{p}_{R_1}(x)\hat{p}_{R_2}(x) \cdots \hat{p}_{R_r}(x)
\]
Then, for every input \( a \)
\[
\Pr_R[\hat{p}_R(a) \neq p(a)] \leq 1/2^r \tag{1}
\]
and the degree of \( \hat{p}_R \) is at most \( r \).

Exercise 2. Prove (1).
For every fixed input many polynomials are good
↓
One polynomial is good for many inputs

Our probability space is the set $\Omega$ of all $r$-tuples $R$ of subsets $R_1, \ldots, R_r$ of variables:
$$\Omega = \{0,1\}^n \times \{0,1\}^n \times \cdots \times \{0,1\}^n$$
r times

For every input $a$, we define the random variable $X_a : \Omega \to \mathbb{R}$
$$X_a(R) = \begin{cases} 1 & \text{if } \hat{p}_R(a) \neq p(a) \\ 0 & \text{otherwise} \end{cases}$$

Consider the sum over all possible inputs $a$
$$X \triangleq \sum_{a \in \{0,1\}^n} X_a$$

By linearity of expectation
$$\mathbb{E}[X] = \sum_a \mathbb{E}[X_a] = \sum_a \mathbb{P}[X_a = 1] \leq 2^n/2^r$$

A random variable cannot be always strictly larger than its expectation, that is, there is one $\omega \in \Omega$ such that
$$X(\omega) \leq \mathbb{E}[X]$$

In our case $\omega = R^*$ meaning that

For every positive integer $r$, there exists a polynomial $\hat{p}_{R^*}$ of degree at most $r$ which differs from the AND of $n$ variables on at most $2^n/2^r$ inputs.

Lemma 4 (low degree polynomials). Let $p(x) \triangleq p_1(x) \cdot p_2(x) \cdots p_m(x)$, where $p_1, \ldots, p_m$ are polynomials of degree at most $d$. For any positive integer $r$, there exists a polynomial $\hat{p}$ such that

1. The degree of $\hat{p}$ is at most $rd$
2. $\hat{p}$ differs from $p$ on at most $2^n/2^r$ inputs.

1.1 Approximate an entire circuit – Lemma 3

Approximating the AND gate is not enough

We first show that one cannot approximate an entire circuit by approximating “locally” every single gate. The output of an AND gate can be also seen as the polynomial \( p(y) = y_1 \cdots y_m \) of its “direct inputs in the circuit”:

\[
\begin{array}{c}
\text{\( x_1 \)} \\
\vdots \\
\text{\( x_n \)} \\
\downarrow \\
\text{\( y'_1 \)} \\
\downarrow \\
\text{\( \Lambda \)} \\
\downarrow \\
\text{\( y'_m \)} \\
\downarrow \\
\text{\( \Lambda \)} \\
\downarrow \\
\text{\( y_1 \)} \\
\downarrow \\
\text{\( \Lambda \)} \\
\downarrow \\
\text{\( y_m \)} \\
\downarrow \\
\text{\( \Lambda \)} \\
\downarrow \\
\text{\( C(x_1, \ldots, x_n) \)}
\end{array}
\]

We know that there is a polynomial \( p_\Lambda \) which differs from \( y_1 \cdots y_m \) on at most \( 2^m/2^r \) values. So we could “locally” replace every AND gate in the circuit by \( p_\Lambda \) on the “direct inputs” of the gate under consideration. In some cases, however, the final polynomial will differ from \( C \) on more than \( |C|2^n/2^r \) inputs (Exercise 1).

Lemma 4 ⇒ Lemma 3

Each gate outputs some “intermediate” function \( g_i(x_1, \ldots, x_n) \) of the input of the circuit, and the output of the circuit is the output of the “last” gate

\[
C(x) = g_s(x)
\]

where \( s = |S| \) is the number of gates. If we replace the \( i^{th} \) function \( g_i(x_1, \ldots, x_n) \) by some other function \( \hat{g}_i(x_1, \ldots, x_n) \), then the “new circuit” \( \hat{C} \) differs from the previous one on at most \( \text{diff}(g_i, \hat{g}_i) \) inputs. This argument can be used to derive Lemma 3 from Lemma 4 (Exercise 3).
Exercise Set 9 – HS12
(Linear Algebra Methods in Combinatorics)

You can submit solutions also by email by the next lecture – 28.11.2012. These exercises are non-graded but you get feedback on your submitted solutions.

The next three exercises are on the proof of Razborov lower bound.

Exercise 1. We know that for every $m$ there is a polynomial $p_{\land}$ which differs from the AND of $m$ bits on at most $2^m / 2^r$ of the possible 0/1-inputs. Consider the following “local procedure” to approximate a given circuit $C$ with a low-degree polynomial. For each AND gate consider its “direct inputs in the circuit” (left figure below):

\[ \land \]
\[ C(x_1, \ldots, x_n) \]

(Each $y_i$ and $y'_i$ is either the output of another gate or one of the variables) Replace each gate by the polynomial $p_{\land}$ on the “direct inputs” of the gate (see right figure above). Show that there is a circuit $C$ for which the resulting polynomial $\hat{C}$ differs from $C$ on more than $|C|2^m / 2^r$ inputs.

Exercise 2. Prove Lemma 4 in the lecture notes by adapting the idea for the special case $p = x_1 \cdots x_m$.

Exercise 3. Derive Lemma 3 from Lemma 4 (see lecture notes).

The next two exercises on depth-2 circuits will help for future exercises.

Exercise 4. Consider this type of depth-2 circuits:
where each of the OR gates is connected to a subset of the input variables (no variable is negated) and the output is the XOR of all these gates. For any \( u \) and \( v \), let \( e_{u,v} \) denote the input in which only variables \( x_u \) and \( x_v \) are set to 1:

\[
e_{u,v} = (0, \ldots, 0, 1, 0, \ldots, 0, 1, 0, \ldots, 0)
\]

Consider the graph over \( n \) nodes whose edges are

\[
E = \{(u, v) : C(e_{u,v}) = 1\}
\]

Show that

\[
(u, v) \in E \iff s - |D_u \cap D_v| = 1 \mod 2
\]

where \( s \) is the number of OR gates and \( D_u \) is the set of OR gates that is not connected to variable \( x_u \).

**Exercise 5.** Also in this exercise we restrict our attention to depth-2 circuits \( C \) of the type described in Exercise 4 ("XOR-of-OR circuits"). Consider the following \( n \times n \) bipartite graph \( G = (V_1 \cup V_2, E) \). The vertices of each side \( V_i \) correspond to all 0/1-vectors of length \( k \) (so \( n = |V_1| = |V_2| = 2^k \)). Two vertices \( u \in V_1 \) and \( v \in V_2 \) are adjacent if and only if \( \langle u, v \rangle = 1 \) over \( \mathbb{F}_2 \).

Show that there is a depth-2 "XOR-of-OR" circuit \( C \) that computes the edges of this graph, that is

\[
E = \{(u, v) : C(e_{u,v}) = 1\}
\]

and that uses \( O(\log n) \) OR gates.