DD2445 COMPLEXITY THEORY: LECTURE 18

Last time

Focus on **AC^0-circuits**

- polynomial size
- constant depth
- unbounded fan-in AND- and OR-gates

PARITY $= \{x \mid x \in \{0,1\}^*, \sum_i x_i \equiv 1 \pmod 2\}$

Will also write PARITY$(x) = \begin{cases} 1 & \text{if } x \in \text{PARITY} \\ 0 & \text{if } x \notin \text{PARITY} \end{cases}$

THM PARITY $\notin AC^0$

Proof idea (roughly):

- Hit circuit with a random restriction
- Simplify $\Rightarrow$ circuit collapses to constant function
- Impossible — still variables left than can flip parity

Key technical tool

HÅSTAD'S SWITCHING LEMMA

Suppose $f: \{0,1\}^n \to \{0,1\}$ can be written as k-DNF. Let ρ uniformly random restriction to t variables. Then for all $s \geq 2$

$$\Pr_\rho[f|_\rho \text{ not } s\text{-CNF}] \leq \left(\frac{(n-t)k^{10}}{n}\right)^{s/2} \quad (*)$$

We proved HSW $\Rightarrow$ PARITY $\notin AC^0$

MAXTERM partial assignment ρ fixing $f|_\rho \equiv 0$

Every clause in k-CNF (negation of) maxterm of size $\leq k$

CLAIM B If all minimal maxterms of f have size $\leq s$, then f can be written as s-CNF

$R_t := \{$all restrictions to t (out of n) variables$\}$

$|R_t| = \binom{n}{t} 2^t$

$B := \{$restrictions $\rho \in R_t$ for which all minimal maxterms of $f|_\rho$ have size $> s\}$

$$\Pr[f|_\rho \text{ not } s\text{-CNF}] \leq \Pr[\rho \in B] = \frac{|B|}{|R_t|}$$

Prove $|B| \ll |R_t|$ by constructing one-to-one mapping

$$m : B \to R_{t+s} \times \{0,1\}^\ell$$

for $\ell = O(s \log k)$

Given $\rho \in B$ construct (τ, c) $\tau \in R_{t+s}$, $c \in \{0,1\}^\ell$ such that c helps to decode ρ hidden in τ

Notation $Vars(\rho) = \{$variables assigned by restriction $\rho\}$

Consider $f \wedge \rho$ for $\rho \in B$

Union of two restrictions ρ & π

- No term in f fixed to 1 by ρ
  (then $f \wedge \rho \equiv 1$ and has no maxterms)
- Not all terms fixed to 0
  (for $f \wedge \rho \equiv 0$ has min maxterm of size 0)
- $f \wedge \rho$ has some minimal maxterm π of size $> s$; $f \wedge \rho\pi \equiv 0$ (by definition)

Fix representation of f as k-DNF
- order terms
- order literals in terms

$(x \wedge y) \vee (\bar{x} \wedge z \wedge u \wedge \bar{w})$
can refer to u as "3rd literal in 2nd term"

Construction of $\tau = \rho\sigma \in R_{t+1,s}$

Look at first term $t_1 \in f$ s.t. $t_1 \wedge \rho \neq 0$

Let $Y_1 := Vars(t_1) \cap Vars(\pi) \neq \emptyset$

π_1 := assignment to Y_1 in π

σ_1 := unique assignment to Y_1 s.t. ~~[illegible]~~ literals over $Vars(Y_1)$ one

c_1 := string with
- $s_1 = |Y_1|$
- positions of these variables in t_1
- values assigned by π_1

Suppose we have constructed

$Y_j, \pi_j, \sigma_j, c_j$ for $j = 1, 2, \ldots, i$

$\rho\pi_1 \cdots \pi_i \subsetneq \rho\pi$

HSWIV

Find first term t_{i+1} not falsified by $\rho\pi_1 \cdots \pi_i$

$Y_{i+1} := Vars(t_1) \setminus Vars(\rho\pi_1 \cdots \pi_i)$

$\pi_{i+1} :=$ assignment to Y_{i+1} in π

$\sigma_{i+1} :=$ unique assignment to Y_{i+1} s.t. [redacted] literals in t_1 over $Vars(Y_{i+1})$ done

$c_{i+1} :=$ string with
- $s_{i+1} = |Y_{i+1}|$
- positions of these variables in t_{i+1}
- values assigned by π_{i+1}

Stop when $\pi_1 \cdots \pi_m$ assigns $\geq s$ variables

Trim down to size exactly s

$m(\rho) = (\rho\sigma_1 \cdots \sigma_m, c_1 \cdots c_m)$

$= (\rho\sigma, c)$

$\rho\sigma \in R_{t+s}$ by construction

$c \in \{0,1\}^{O(s \log k)}$ argued last time

Remains to show that m is one-to-one

Given $m(\rho) = (\tau, c)$ want to recover $\rho \subseteq \tau$

$\phi|_\tau$ k-DNF formula with lots of terms satisfied (potentially)

But ρ didn't satisfy any terms...

So some assignments in there form σ

HSW V

Find first term t_1 not falsified by τ (either satisfied or not fixed)

This must be t_1 in our construction (first term t_1 s.t. $t_1 \wedge \rho \neq 0$, and σ didn't falsify t_1)

Use $c_1 \in C$ to find
- variables assigned by π_1 (Y_1)
- assignments in π_1

Let $\tau_1 = \tau$ with assignments to Y_1 modified according to π

Now

$$\tau = \rho\, \sigma_1\, \sigma_2 \cdots \sigma_m$$
$$\tau_1 = \rho\, \pi_1\, \sigma_2 \cdots \sigma_m$$

Suppose we have reconstructed

$$\pi_i = \rho\, \pi_1 \cdots \pi_i\, \sigma_{i+1} \cdots \sigma_m$$

Find first term t_{i+1} not falsified by τ_i
Must be same t_{i+1} as in construction
- $\rho\, \pi_1 \cdots \pi_i$ didn't falsify it
- σ_{i+1} designed to satisfy t or leave unfalsified ~~(if $i+1 = m$ & trivially)~~

Use $c_{i+1} \in C$ to find Y_{i+1} and assignments in π_{i+1}. Flip τ_i on π_{i+1} to get

$$\tau_{i+1} = \rho\, \pi_1 \cdots \pi_i\, \pi_{i+1}\, \sigma_{i+2} \cdots \sigma_m$$

When this process ends, have recovered

$\rho\, \pi_1, \cdots \pi_m$ and $\pi_1, \cdots \pi_m$

Read off ρ. So mapping

$$m : B \longrightarrow R_{t+s} \times \{0,1\}^{\ell}$$

one-to-one as claimed.
Håstad's switching lemma follows ☑

(Again, this is a simplified Arora-Barak version — original switching lemma is stronger and yields exponential lower bounds on size for bounded-depth circuits computing parity.)

Question: Why prove lower bound for PARITY? Very simple function...

Choose something harder ⇒ stronger lower bound?

Paradox: In order to prove lower bound for function, need to understand it well. So the function cannot be too hard (or else it becomes too hard to prove lower bound, even if this lower bound is probably true)

Now know PARITY $\notin AC^0$

Strengthen to more general classes of circuits?

- Keep depth $O(1)$
- But allow more kinds of gates

"Counting gates"

$$MOD_m(x) = \begin{cases} 0 & \text{if } \sum x_i \equiv 0 \pmod m \\ 1 & \text{otherwise} \end{cases}$$

Just a single MOD_2-gate can compute parity

DEF ACC^0

For integers $m_1, \dots, m_k > 1$, say that $L \in ACC^0(m_1, \dots, m_k)$ if $\exists$ circuit family $\{C_n\}_{n=1}^{\infty}$ deciding L where C_n has

- gates $\wedge, \vee, \neg, MOD_{m_1}, \dots, MOD_{m_k}$
- constant depth
- polynomial size

ACC^0 = union of languages in $ACC^0(m_1, \dots, m_k)$ for any $k \in \mathbb{N}$ and $m_1, \dots, m_k > 1$

THEOREM 1 [Razborov '87, Smolensky '87]
For primes p, q, $p \neq q$, $MOD_p \notin ACC^0(q)$

We will give the proof for a specific case

THEOREM 2 $PARITY \notin ACC^0(3)$

Proof idea (*Method of Approximations*)

(1) Show circuits in $ACC^0(3)$ are well approximated by "low-degree" polynomial over $GF(3)$

(2) Show that PARITY cannot be approximated in this way

$GF(3) = \mathbb{F}_3$ = Computing mod 3
mod 3 we have $2 = -1$ so will write
$\mathbb{F}_3 = \{-1, 0, 1\}$

FACT 3 Every $\mathbb{F}_3$-polynomial that computes parity of n variables exactly must have degree n

So if we could get exact representation in (1) we would be done. But we can't (as we will see soon)

Detour Useful ways of thinking about Boolean functions

MA III

Normally $f: \{0,1\}^n \to \{0,1\}$ 0 false, 1 true

Sometimes useful $\tilde{f}: \{-1,+1\}^n \to \{-1,+1\}$ 1 false, −1 true

PARITY computed by $\prod_{i=1}^{n} y_i$

$= -1$ iff odd # $y_i = -1$

Do affine transformation $x \mapsto (1+x)$

$$\prod_{i=1}^{n} (1+x_i) - 1$$

computes PARITY in standard 0/1-setting

Further observations

(i) Going from f to $\tilde{f}$ using $x \mapsto (1+x)$ doesn't change degree

(ii) W.l.o.g. can represent functions $f: \{0,1\} \to \mathbb{F}_2$ and $\tilde{f}: \{-1,1\} \to \mathbb{F}_2$ as MULTILINEAR POLYNOMIALS – degrees of individual variables ≤ 1

$x = 0/1 \Rightarrow x^2 = x$

$x = -1/+1 \Rightarrow x^2 = 1$

FACT 4 Every $f: \{0,1\}^n \to \mathbb{F}_2$ has unique representation as multilinear $\mathbb{F}_2$-polynomial

Proof $f: \{0,1\}^n \to \mathbb{F}_3$ 2^n-dimensional vector space

Multilinear monomials:

- 2^n of them (include ~~of~~ don't include every variable x_i)
- Span space of all functions

So is a basis (can also prove linear independence). Means that representation as linear combination of monomials is unique ☑

Easy to see that $\prod_{i=1}^{n}(1+x_i) - 1$ has degree n (monomial $x_1 x_2 \cdots x_n$ will appear)
Establishes FACT 3

END OF DETOUR

Let us implement step (1) in the proof

Follow not Arora-Barak but Ryan Williams (though differences not huge)
Convenient tool:

DEFINITION 5 A **PROBABILISTIC POLYNOMIAL** for $f: \{0,1\}^n \to \{0,1\}$ with degree d and error ε is a distribution $\mathcal{D}$ over polynomials of degree $\leq d$ such that

$$\forall x \in \{0,1\}^n \quad \Pr_{p \sim \mathcal{D}}\left[p(x) \neq f(x)\right] < \varepsilon$$

Probabilistic polynomial for circuit C: for function computed by C

LEMMA 6 For all circuits C over $\{\vee, \wedge, \neg, MOD_3\}$ of size s and depth d it holds that for all $k \in \mathbb{N}^+$ there is a probabilistic polynomial for C over $\mathbb{F}_3$ with degree $\leq (2k)^d$ and error $\leq s / 3^k$

Remarks

- In general case with MOD_q-gates would get degree $((q-1)k)^d$ and error s/q^k
- Construction of probabilistic polynomial from circuit will be *efficient* (although we don't need it here) – doable in time

$$poly(s)\binom{n}{(2k)^d}$$

Proof of Lem 6

Replace each gate in C by probabilistic polynomial of very low degree (in inputs) and very low error

$$Pr[\text{error at gate}] \leq 1/3^k$$

$$Pr[\text{error in circuit}] \leq \sum_{\text{gate}} Pr[\text{error at gate}] \leq s/3^k$$

Approximate each gate by poly of degree $\leq 2k$
Degrees multiply, but depth $\leq d$
$\Rightarrow$ total degree $\leq (2k)^d$

How to approximate gates g by polys p_g

(i) $g = \neg h \rightsquigarrow p_g = 1 - p_h$

No new errors; degree-1 poly (in # inputs)

(ii) $g = MOD_3(h_1, \ldots, h_t)$

$p_g = \left(\sum_{i=1}^{t} p_{h_i}\right)^2$

Sample p_g by first sampling p_h's and then doing construction

Note $\left(\sum_{i=1}^{t} p_{h_i}\right)^2 \equiv 1 \pmod 3$ iff $3 \nmid \sum_i k_i p_{h_i}$

by Fermat's Little Theorem

No new errors introduced; degree-2 poly (in inputs)

(iii) OR-gate: now we have problems

$\bigvee_{i=1}^{t} h_i$ represented by

$1 - \prod_{i=1}^{t}(1 - p_{h_i})$ — far too high degree!

LEMMA 7 $\forall k \in \mathbb{N}^+ \; \forall n \in \mathbb{N}^+ \; \exists$ probabilistic polynomial for OR_n over $\mathbb{F}_3$ of degree $2k$ and error $\leq 1/3^k$

Note Error probability independent of n!

Proof Pick uniformly random $\bar{r} \in \mathbb{F}_3^n$

Set $p_1(x) = \sum_{i=1}^{n} r_i \cdot x_i$

$OR_n(x) = 0 \Rightarrow p_1(x) = 0$

$OR_n(x) = 1 \Rightarrow p_1(x) \in \{-1, 0, 1\}$

In this case $Pr[p_1(x) = 0] = 1/3$

Fix $x_i = 1$. Compute $\sum_{j \neq i} v_j \cdot x_j = a$

a is some element in $F_3 = \{-1, 0, 1\}$

There is exactly 1 out of 3 choices for v_i s.t.

$$v_i + a = \sum_{j=1}^{n} v_j \cdot x_j = 0$$

Set $p_2(x) = (p_1(x))^2 \in \{0, 1\}$

$OR_n(x) = 1 \Rightarrow Pr[p_2(x) = 1] = 2/3$

Pick $g_1, \ldots, g_k = \left(\sum v_i^k \cdot x_i\right)^2$

in this way

$$OR_n(x) = 1 \Rightarrow Pr\left[\forall j \; g_j(x) = 0\right] = 1/3^k$$

Set $p(x) = 1 - \prod_{j=1}^{k} (1 - g_j(x))$

$OR_n(x) = 0 \Rightarrow p(x) = 0$ with probability 1

$OR_n(x) = 1 \Rightarrow Pr[p(x) = 1]$

$= Pr[\exists j \; g_j(x) = 1]$

$= 1 - 1/3^k$

Degree $2k$; error $\leq 1/3^k$ as claimed ∎

(iv) AND-gate: Use De Morgan's Law

$$AND(x_1, \ldots, x_n) = \neg OR(\neg x_1, \ldots, \neg x_n)$$

So take polynomial

$$1 - p(1-x_1, \ldots, 1-x_n)$$

as constructed before

Final probabilistic polynomial = that of output gate

Degree $\leq (2k)^d$ by construction

Failure probability

$$\Pr[\text{failure in computation in circuit}]$$

$$\leq \sum_{\text{gates}} \Pr[\text{failure at particular gate}]$$

$$\leq s \cdot 1/3^k$$

(very pessimistic bound)

Lemma 6 follows ▨

COROLLARY 8 For any circuit C over $\{\vee, \wedge, \neg, MOD_3\}$ of size s and depth d and $k \in \mathbb{N}^+$ there exists an $\mathbb{F}_3$-polynomial p of degree $\leq (2k)^d$ such that

$$\Pr_{x \sim \{0,1\}^n}[C(x) \neq p(x)] \leq s/3^k$$

Proof For $x \sim \{0,1\}^k$ and $p \sim \mathcal{D}$ define 0/1-valued $x_{p,x}$ by

$$x_{p,x} = 1 \iff C(x) = p(x)$$

Then

$$\sum_{x \in \{0,1\}^n} \mathop{\mathbb{E}}_{p \sim \mathcal{D}} [x_{p,x}] \geq 2^n (1 - 5/3^k)$$

for distribution $\mathcal{D}$ constructed in proof of Lem 6.

By linearity of expectation

$$\mathop{\mathbb{E}}_{p \sim \mathcal{D}} \left[\sum_{x \in \{0,1\}^n} x_{p,x} \right] \geq 2^n (1 - 5/3^k)$$

This means there must exist at least one polynomial p^* with

$$\sum_x x_{p^*x} \geq 2^n (1 - 5/3^k)$$

But this means that

$$\Pr_{x \sim \{0,1\}^n} [p^*(x) \neq C(x)] \leq 5/3^k \qquad \square$$

So small, low-depth $ACC^0(3)$-circuits can be approximated well by low-degree polynomials In step (2) we want to show that this is not true for parity function.