Lecture 17

Counting is hard for small depth circuits

June 1, 2004
Lecturer: Paul Beame
Notes: Sumit Sanghai

In this lecture we will give bounds on circuit size-depths which compute the function $\oplus_p$. More specifically we will show that a polynomial-sized constant depth $AC^0[q]$ circuit cannot compute $\oplus_p$.

Theorem 17.1 (Razborov,Smolensky). *Let $p \neq q$ be primes. Then $\oplus_p \notin AC^0[q]$.*

We will prove that $S = 2^{n^{\Omega(1/d)}}$ or $d = \Omega(\log n / \log\log S)$. Note that $AC^0[q]$ contains the operations $\wedge$, $\vee$, $\neg$ and $\oplus_q$ where $\oplus_q(x_1, \ldots, x_n) = \begin{cases} 0 & \text{if } \sum_i x_i \equiv 0 \pmod q \\ 1 & \text{otherwise.} \end{cases}$

To prove this theorem we will use the *method of approximation* introduced by Razborov.

Method of Approximation For each gate g in the circuit we will define a family A_g of allowable approximators for g. For the operation Op_g at gate g, we define an approximate version $\widetilde{Op_g}$ such that if $g = Op_g(h_1, \cdots, h_k)$ then $\widetilde{g} = \widetilde{Op_g}(\widetilde{h_1}, \cdots, \widetilde{h_k}) \in A_g$.

We will prove that there are approximators such that $\widetilde{Op}(\widetilde{h_1}, \cdots, \widetilde{h_k})$ and $Op(\widetilde{h_1}, \cdots, \widetilde{h_k})$ differ on only an ϵ-fraction of all inputs implying that the output $\widetilde{f} \in A_f$ differs from f on at most ϵS fraction of all inputs. We will then prove that any function in A_f differs from f on a large fraction of inputs proving that S is large given d.

Proof of Theorem 17.1. We will prove that $\oplus_2 \notin AC^0[q]$ where q is a prime greater than 2. The proof can be extended to replace $\oplus_2$ by any $\oplus_p$ with $p \neq q$.

The Approximators For a gate g of height d' in the circuit, the set of approximators A_g will be polynomials over $\mathbb{F}_q$. of total degree $\leq n^{\frac{d'}{2d}}$.

Gate approximators

- $\neg$ gates: If $g = \neg h$, define $\widetilde{g} = 1 - \widetilde{h}$. This yields no increase in error or degree.
- $\oplus_q$ gates: If $g = \oplus_q(h_1, \ldots, h_k)$, define $\widetilde{g} = (\sum_{i=1}^{k} \widetilde{h_i})^{q-1}$. Since q is a prime, by Fermat's little theorem we see that there is no error in the output. However, the degree increases by a factor of $q - 1$.
- $\vee$ gate:
  Note that without loss of generality we can assume that other gates are $\vee$ gates: We can replace the

$\wedge$ gates by $\neg$ and $\vee$ gates and since the $\neg$ gates do not cause any error or increase in degree we can "ignore" them.
Suppose that $g = \bigvee_{i=1}^{k} h_i$. Choose $\bar{r_1}, \cdots, \bar{r_t} \in_R \{0,1\}^k$. Let $\widetilde{h} = (\widetilde{h_1}, \cdots, \widetilde{h_k})$. Then

$$\Pr[\bar{r_1} \cdot \widetilde{h} \equiv 0 \pmod{q}] = \begin{cases} 1 & \text{if } \bigvee i = 1^k \widetilde{h_i} = 0, \text{ and} \\ \leq 1/2 & \text{otherwise.} \end{cases}$$

(This follows because if $\bigvee_{i=1}^{k} \widetilde{h_i} = 1$ then there exists j such that $\widetilde{h_j} \neq 0$ in which case if we fix the remaining coordinates of $\bar{r_1}$, there is at most one choice for the j^{th} coordinate of $\bar{r_1}$ such that $\bar{r_1} \cdot \widetilde{h} \equiv 0 \pmod{q}$.)
Let $\widetilde{g_j} = (\bar{r_j} \cdot \widetilde{h})^{q-1}$ and define

$$\widetilde{g} = \widetilde{g_1} \vee \cdots \vee \widetilde{g_t} = 1 - \prod_{j=1}^{t} (1 - \widetilde{g_j}).$$

For each fixed vector of inputs $\widetilde{h}$,

$$\Pr[\, \widetilde{g} \neq \bigvee_{i=1}^{k} \widetilde{h_i} \,] \leq (1/2)^t.$$

Therefore, there exists $\bar{r_1}, \cdots, \bar{r_t}$ such that $\widetilde{g}$ and $\bigvee_{i=1}^{k} \widetilde{h_i}$ differ on at most a $(1/2)^t$ fraction of inputs.

Also note that the increase in degree from the $\widehat{h_i}$ to $\widehat{g}$ is $(q-1)t$. We will choose $t = n^{\frac{1}{2d}}/(q-1)$.

Thus we obtain the following lemma:

Lemma 17.2. *Let $q \geq 2$ be prime. Every $\mathsf{AC}[q]$ circuit of size S and depth d has a degree $((q-1)t)^d$ polynomial approximator over $\mathbb{F}_q$ with fractional error at most $2^{-t}S$.*
In particular, setting $t = \frac{n^{1/(2d)}}{q-1}$, there is a degree $\sqrt{n}$ approximator for the output of the circuit having error $\leq 2^{-\frac{n^{1/(2d)}}{q-1}} S$.

In contrast we have the following property of approximators for $\oplus_2$.

Lemma 17.3. *For $q > 2$ prime and $n \geq 100$, any $\sqrt{n}$ degree polynomial approximator for $\oplus_2$ over $\mathbb{F}_q$ has error at least $1/5$.*

Proof. Let $U = \{0,1\}^n$ be the set of all inputs. Let $G \subseteq U$ be the set of "good" inputs, those on which a degree $\sqrt{n}$ polynomial a agrees with $\oplus_2$.

Instead of viewing $\oplus_2$ as $\{0,1\}^n \to \{0,1\}$ we consider $\oplus_2' : \{-1,1\}^n \to \{-1,1\}$ where we interpret -1 as representing 1 and 1 as representing 0. In particular, $\oplus_2'(y_1, \cdots, y_n) = \prod_i y_i$. where $y_i = (-1)^{x_i}$. We get that $\oplus_2(x_1, \cdots, x_n) = 1$ if and only if $\oplus_2'(y_1, \cdots, y_n) = -1$.

We can see that the $x_i \to y_i$ map can be expressed using a linear map m as follows
$m(x_i) = 2x_i - 1$ and since q is odd, m has an inverse map $m^{-1}(y_i) = (y_i + 1)/2$

Thus, given a of $\sqrt{n}$-degree polynomial that approximates $\oplus_2$, we can get an approximator a' of $\sqrt{n}$ degree that approximates $\oplus_2'$ by defining

$$a'(y_1, \cdots, y_n) = m(a(m^{-1}(y_1), \cdots, m^{-1}(y_n))).$$

It is easy to see that a' and $\oplus_2'$ agree on the image $m(G)$ of G.

Let $\mathcal{F}_G$ be the set of all functions $f : m(G) \to \mathbb{F}_q$. It is immediate that

$$|\mathcal{F}_G| = q^{|G|}. \tag{17.1}$$

Given any $f \in \mathcal{F}_G$ we can extend f to a polynomial $p_f : \{1,-1\}^n \to F_q$ such that f and p_f agree everywhere on $m(G)$. Since $y_i^2 = 1$, we see that p_f is multilinear. We will convert p_f to a $(n+\sqrt{n})/2$-degree polynomial.

Each monomial $\prod_{i\in T} y_i$ of p_f is converted as follows:

- if $|T| \le (n+\sqrt{n})/2$, leave the monomial unchanged.
- if $|T| > (n+\sqrt{n})/2$, replace $\prod_{i\in T} y_i$ by $a' \prod_{i\in\bar{T}} y_i$ where $\bar{T} = \{1,\dots,n\} - T$. Since $y_i^2 = 1$ we have that $\prod_{i\in T} y_i \prod_{i\in T'} y_i = \prod_{i\in T\Delta T'} y_i$. Since on $m(G)$, $a'(y_1,\dots,y_n) = \prod_{i=1}^n y_i$, we get that $\prod_{i\in T} y_i = a' \prod_{i\in\bar{T}} y_i$ on $m(G)$. The degree of the new polynomial is $|\bar{T}|+\sqrt{n} \le (n-\sqrt{n})/2+\sqrt{n} = (n+\sqrt{n})/2$.

Thus $|\mathcal{F}_G|$ is at most the number of polynomials over $\mathbb{F}_q$ of degree $\le (n+\sqrt{n})/2$. Since each such polynomial has a coefficient over $\mathbb{F}_q$ for each monomial of degree at most $(n+\sqrt{n})/2$,

$$|\mathcal{F}_G| \le q^M \tag{17.2}$$

where

$$M = \sum_{i=0}^{(n+\sqrt{n})/2} \tbinom{n}{i} \le \frac{4}{5}2^n \tag{17.3}$$

for $n \ge 100$. This latter bound follows from the fact that this sum consists of the binomial coefficients up to one standard deviation above the mean. In the limit as $n \to \infty$ this would approach the normal distribution and consist of roughly 68% of all weight. By n around 100 this yields at most 80% of all weight.

From equations 17.1,17.2 and 17.3 we get $|G| \le |M| \le \frac{4}{5}2^n$. Hence the error $\ge 1/5$. ☐

Corollary 17.4. *For $q > 2$ prime, any $\mathsf{AC}^0[q]$ circuit of size S and depth d computing $\oplus_2$ requires $S \ge \frac{1}{5}2^{\frac{n^{\frac{1}{2d}}}{q-1}}$*

Proof. Follows from Lemmas 17.2 and 17.3. ☐

This yields the proof of Theorem 17.1. ☐

From Corollary 17.4, we can see that for polynomial-size $AC[q]$ circuits computing $\oplus_2$, the depth $d = \Omega(\frac{\log n}{\log\log n})$. By the lemma from the last lecture that $\mathsf{NC}^1 \subseteq \mathsf{AC{-}SIZEDEPTH}(n^{O(1)}, O(\frac{\log n}{\log\log n}))$ any asymptotically larger depth lower bound for any function would be prove that it is not in NC^1.

Our inability to extend the results above to the case that q is not a prime is made evident by the fact that following absurd possibility cannot be ruled out.

Open Problem 17.1. Is $NP \subseteq AC^0[6]$?

The strongest kind of separation result we know for any of the NC classes is the following result which only holds for the uniform version of ACC^0. It uses diagonalization.

Theorem 17.5 (Allender-Gore). $\mathrm{PERM} \notin \mathsf{UniformACC}^0$.