CS 6810: Theory of Computing **Fall 2023**

Lecture 11: Sep 26, 2023

Lecturer: Eshan Chattopadhyay *Scribe: Sanjit Basker and Owen Oertell*

1 AC^0 lower bounds

Recall: AC^0 is the set of languages decidable by a family of constant-depth circuits with unbounded fan-in: the allowed gates are $\vee, \wedge, \neg$.

Theorem 1.1 (Razborov-Smolensky)**.** *$\mathsf{MAJORITY} \notin \mathsf{AC}^0$.*

We'll use $\mathbb{F}_2^n$ and $\{0,1\}^n$ interchangeably in this proof. We define the Hamming weight $|x|$ as the number of 1s in x.

Plan for proof: We will show that there is a weakness in AC^0 which is not in $\mathsf{MAJORITY}$. Specifically, we shall show low degree approximators exist for all languages computed by AC^0, but that there are no low-degree approximators for MAJORITY.

Let $P_{n,d} = \{n\text{-variate polynomials over } \mathbb{F}_2^n \text{ of degree} \leq d\}$. Note that we have $x^2 = x$ for any $x \in \mathbb{F}_2$. So we will be working with multi-linear polynomials, i.e. any occurrence of a variable has degree at most 1. Some examples are $x_1 + x_2 + \cdots + x_n$ and $x_1x_2 + x_3x_5$. A non-example is $x_1x_2^2$.

Definition 1.2. *Let U_n denote the uniform distribution on bitstrings of length n.*

Definition 1.3. *A function $f : \{0,1\}^n \to \{0,1\}$ ϵ-approximates $g : \{0,1\}^n \to \{0,1\}$ if*

$$\Pr_{x \sim U_n}[f(x) = g(x)] \geq 1 - \epsilon.$$

Definition 1.4 (Probabilistic Polynomial)**.** *A distribution $\mathcal{P}$ on $P_{n,d}$ is a probabilistic polynomial of degree $\leq d$.*

Instead of thinking about a function f that ϵ-approximates g, it is easier to think about the following notion.

Definition 1.5 (Probabilistic pointwise approximation)**.** *A probabilistic polynomial $\mathcal{P}$ approximates $g : \{0,1\}^n \to \{0,1\}$ if for all $x \in \{0,1\}^n$,*

$$\Pr_{\mathcal{P}}[\mathcal{P}(x) = g(x)] \geq 1 - \epsilon$$

Claim 1.6. *Suppose g has an ϵ probabilistic pointwise approximator $\mathcal{P}$ of degree d. Then $\exists f \in P_{n,d}$ that ϵ-approximates g.*

Proof. By definition, we know that

$$\mathop{\mathbb{E}}_{p \sim \mathcal{P}}[1_{P(x)=g(x)}] \geq 1 - \epsilon$$

for all $x \in \{0,1\}^n$. Thus,

$$\mathop{\mathbb{E}}_{x \sim \mathcal{U}_n} \mathop{\mathbb{E}}_{p \sim \mathcal{P}}[1_{P(x)=g(x)}] \geq 1 - \epsilon$$

We can switch the expectation,

$$\mathop{\mathbb{E}}_{p\sim\mathcal{P}}\mathop{\mathbb{E}}_{x\sim\mathcal{U}_n}[1_{P(x)=g(x)}] \geq 1-\epsilon$$

Then, we know that $\exists p \in \operatorname{supp}(\mathcal{P})$ such that $\Pr_x[p(x) = g(x)] \geq 1 - \epsilon$.

$\square$

Fact 1.7. *Any function $f : \{0,1\}^n \to \{0,1\}$ can be completely approximated by $p \in P_{n,n}$.*

You can prove this, for example, by doing casework on the $f_0(x') = f(x'|0)$ and $f_1(x') = f(x'|1)$, or by using polynomial interpolation.

1.1 Probabilistic Polynomials for AC^0

We'll start with approximating AND and OR gates. Note that the output of a NOT gate can be trivially approximated by subtracting the polynomial that approximates its argument from 1.

We can naïvely approximate the AND gate by:

$$\wedge(x_1,\ldots,x_n) \Rightarrow P_{\wedge,\text{naïve}}(x_1,\ldots,x_n) = \prod x_i$$

Although this is a complete approximation, it is quite high degree. So let's try to find a better one.

For $S \subseteq [n]$, define $P_{\wedge,S} = 1 - \sum_{i\in S}(1-x_i)$. Sample S according to the following: for each $i \in [n]$, place i in S independently and with probability $\frac{1}{2}$. This then gives a distribution $\mathcal{P}$ over polynomials.

We care the most about: $x = 1^n$, we get that $P_{\wedge,S}(x) = 1 - \sum_{i\in S} 0 = 1 = \wedge(x)$. On the other hand, in the case that $x \neq 1^n$, then we want to find the probability that

$$\Pr[P_{\wedge,S}(x) = \wedge(x)] = \Pr[P_{\wedge,S}(x) = 0].$$

Pick an i such that $x_i = 0$. Now, sample S by deciding on the fate of x_i last, i.e. $S = T \cup X$ where T is a fixed subset of $[n] \setminus \{i\}$, and X is either $\{\}$ or $\{i\}$, with probability $\frac{1}{2}$. Since $x_i = 0$, we know that $1 - x_i = 1$, so

$$\begin{aligned} P_{\wedge,S}(x) &= 1 - \sum_{j\in S}(1-x_j) \\ &= \begin{cases} 1 - \sum_{j\in T}(1-x_j) - 1, & \text{with probability } \frac{1}{2} \\ 1 - \sum_{j\in T}(1-x_j), & \text{with probability } \frac{1}{2} \end{cases} \end{aligned}$$

The branches in the last expression are equal to 0 and 1 in some order, so this shows that for all $x \neq 1^n$, $\Pr_{p\sim\mathcal{P}}[p(x) = \wedge(x)] = \frac{1}{2}$

However, this polynomial isn't a satisfactory approxomation because it only detects the no-cases with probabililty $\frac{1}{2}$. We can "boost" its accuracy by taking k such polynomials and multiplying them together, à la naïve AND:

$$P_\wedge^k = \prod P_{\wedge,S_i} = P_{\wedge,n}(P_{n,s_1},\ldots P_{n,s_k})$$

This gives us a better approximator with accuracy 1 when $x = 1^n$ and accuracy $1 - \left(\frac{1}{2}\right)^k$ when $n \neq 1^n$.

Taking $k = \log(1/\epsilon)$ gives a low-degree ϵ-approximator for AND.

The approximator for OR with similar properties can be found by composing De Morgan's Law with this entire construction.

Let's try to compose these approximators inductively: suppose that our circuit C is an OR of ℓ inputs, which can be approximated by the polynomials P_{c_i} for $i = 1, \ldots, \ell$. We apply the k-approximator for OR to these polynomials to get the approximator $P_C = P_\vee^k(P_{c_1}, \ldots, P_{c_\ell})$. We know that $\deg P_C = k \cdot \max(\deg P_{c_i})$.

We'll use the union bound to evaluate the correctness of the new circuit: the bad events are any of the $\ell + 1$ polynomials (the inputs and the polynomial that evaluates the OR) being wrong, and each bad event occurs with probability ϵ'. So the final probability of being wrong is at most $(\ell + 1)\epsilon'$.

Every time that we reduce the depth of this circuit, we know that we reduce the degree of our polynomial by d. Therefore, the degree of probabilistic polynomial approximating a circuit of t layers has degree k^t. We can use a similar argument on the correctness. In general, we have that for a circuit of size s and depth t, there exists a probabilistic polynomial of degree is $\log^t\left(\frac{1}{\epsilon}\right)$ and with error probability $s \cdot \epsilon$.

Coalescing this all,

Theorem 1.8. *For any circuit $C \in \mathsf{AC}^0$, with size s and depth t. There is a $\log^t(s/\epsilon)$ degree ϵ-probabilistic polynomial approximator.*