Parity vs. majority (lecture notes)

Course: Circuit Complexity, Autumn 2024, University of Chicago
Instructor: William Hoza (`williamhoza@uchicago.edu`)

---

1 Using majority gates to compute the parity function

In a previous class, we showed that $\mathsf{PARITY} \notin \mathsf{AC}^0$. In this section, we will do a *reduction* from the parity function to the majority function. This will imply that $\mathsf{MAJORITY} \notin \mathsf{AC}^0$. More precisely, when we say a "reduction from parity to majority," we mean a circuit that computes the parity function using majority gates. This motivates the following definition.

Definition 1 (The complexity class TC^0)**.** A function $f\colon \{0,1\}^* \to \{0,1\}^*$ is in TC^0 if it can be computed by a constant-depth polynomial-size circuit in which every gate is a MAJ gate of unbounded fan-in, and there are literals and constants at the bottom.

The class TC^0 roughly corresponds to *neural networks* in machine learning. We have $\mathsf{AC}^0 \subseteq \mathsf{TC}^0 \subseteq \mathsf{NC}^1$. Our goal is to show that $\mathsf{PARITY} \in \mathsf{TC}^0$. More generally, we will prove that every "symmetric function" is in TC^0.

Definition 2 (Symmetric Boolean function)**.** A function $f\colon \{0,1\}^n \to \{0,1\}$ (or $f\colon \{0,1\}^* \to \{0,1\}$) is *symmetric* if $f(x)$ depends only on the length of x and the Hamming weight of x. The class of symmetric functions is denoted SYM.

For example, PARITY, MAJ, AND, and OR are all symmetric. The following "threshold functions" are also symmetric:

$$T_n^{\leq k}(x) = 1 \iff \sum_{i=1}^n x_i \leq k$$
$$T_n^{\geq k}(x) = 1 \iff \sum_{i=1}^n x_i \geq k.$$

Lemma 1 (Shifting thresholds)**.** *$T_n^{\geq k}$ and $T_n^{\leq k}$ are both in TC^0.*[1]

Proof sketch.

$$T_n^{\geq k}(x_1, \dots, x_n) = \mathsf{MAJ}_{2n+2k}(x_1, \dots, x_n, \underbrace{0, \dots, 0}_{2k \text{ zeroes}}, \underbrace{1, \dots, 1}_{n \text{ ones}}),$$

and similarly

$$T_n^{\leq k}(x_1, \dots, x_n) = \mathsf{MAJ}_{2n+2k}(\neg x_1, \dots, \neg x_n, \underbrace{0, \dots, 0}_{n \text{ zeroes}}, \underbrace{1, \dots, 1}_{2k \text{ ones}}).$$

□

Theorem 1 ($\mathsf{SYM} \subseteq \mathsf{TC}^0$)**.** *If $f\colon \{0,1\}^* \to \{0,1\}$ is symmetric, then $f \in \mathsf{TC}^0$.*

Proof. Let $n \in \mathbb{N}$. There is some set $S \subseteq [n]$ such that for every $x \in \{0,1\}^n$, $f(x) = 1$ if and only if $|x| \in S$, where $|x|$ denotes Hamming weight. Then

$$f(x) = \left(\sum_{k \in S} (T_n^{\leq k}(x) + T_n^{\geq k}(x)) \right) - |S|,$$

which is a threshold of thresholds. □

[1] The function $T_n^{\geq k}$ is parameterized by two values (n and k), so it is not necessarily clear what it means to say that $T_n^{\geq k} \in \mathsf{TC}^0$. The meaning is that $T_n^{\geq k(n)} \in \mathsf{TC}^0$ for every function $k(n)$. Equivalently, for every n and k, there is a constant-depth polynomial-size majority circuit that computes $T_n^{\geq k}$, i.e., the depth and the exponent of the size do not depend on n or k. Similarly with $T_n^{\leq k}$.

Corollary 1 ($\mathsf{MAJ} \notin \mathsf{AC}^0$)**.** *There exists a constant $\alpha > 0$ such that the following holds. Let C be a size-S AC^0_d circuit computing MAJ_n, where $d \leq \frac{\alpha \log n}{\log \log n}$. Then $S = 2^{n^{\Omega(1/d)}}$.*

Proof. Since $\mathsf{PARITY} \in \mathsf{TC}^0$, there is a constant-depth polynomial-size circuit computing $\mathsf{PARITY}_{n'}$ using MAJ_n gates, where $n' = n^{\Omega(1)}$.[2] By replacing each MAJ_n gate with a copy of C, we get an $\mathsf{AC}^0_{O(d)}$ circuit computing $\mathsf{PARITY}_{n'}$ of size $S' = S \cdot \operatorname{poly}(n)$. We proved in a previous class that $S' = 2^{(n')^{\Omega(1/d)}} = 2^{n^{\Omega(1/d)}}$. Therefore, $S \geq 2^{n^{\Omega(1/d)}} / \operatorname{poly}(n)$, which is $2^{n^{\Omega(1/d)}}$ if we pick α small enough. ∎

2 Majority is hard, even if we are allowed to use parity gates

Definition 3 (The complexity class $\mathsf{AC}^0[m]$)**.** An $\mathsf{AC}^0_d[m]$ circuit is a depth-d circuit in which we can use AND gates, OR gates, and MOD_m gates of unbounded fan-in. A MOD_m gate computes the function

$$\mathsf{MOD}_m(x) = \begin{cases} 0 & \text{if } \sum_i x_i \equiv 0 \pmod m \\ 1 & \text{if } \sum_i x_i \not\equiv 0 \pmod m. \end{cases}$$

At the bottom, there are literals and constants. (Negations do not count toward the size or depth of the circuit.) A function $f \colon \{0,1\}^* \to \{0,1\}$ is in $\mathsf{AC}^0[m]$ if it can be computed by constant-depth polynomial-size $\mathsf{AC}^0[m]$ circuits.

For example, a MOD_2 gate computes the parity function. For this reason, $\mathsf{AC}^0[2]$ is also often denoted $\mathsf{AC}^0[\oplus]$. We have

$$\mathsf{AC}^0 \subseteq \mathsf{AC}^0[\oplus] \subseteq \mathsf{TC}^0 \subseteq \mathsf{NC}^1.$$

The first containment is strict ($\mathsf{AC}^0 \neq \mathsf{AC}^0[\oplus]$), because $\mathsf{PARITY} \notin \mathsf{AC}^0$. In this section, we will prove that the second containment is also strict ($\mathsf{AC}^0[\oplus] \neq \mathsf{TC}^0$). This is equivalent to proving that $\mathsf{MAJ} \notin \mathsf{AC}^0[\oplus]$. The first step of the proof is to generalize our probabilistic polynomial construction to $\mathsf{AC}^0[\oplus]$ circuits.

Theorem 2 (Probabilistic polynomials for $\mathsf{AC}^0[\oplus]$)**.** *Let $f \colon \{0,1\}^n \to \{0,1\}$ be a $\mathsf{AC}^0_d[\oplus]$ circuit of size $S \geq n$, and let $\varepsilon > 0$. Then f can be computed with error ε by a probabilistic polynomial over $\mathbb{F}_2$ of degree $(\log(S/\varepsilon))^{O(d)}$.*

Proof sketch. Repeat the proof that AC^0 can be simulated by probabilistic polynomials, and use the fact that the parity function can be computed exactly by a degree-1 polynomial over $\mathbb{F}_2$. ∎

The second step, which is not so easy, is to show that low-degree polynomials over $\mathbb{F}_2$ cannot approximate the majority function (just like we showed that low-degree polynomials over $\mathbb{F}_3$ cannot approximate the parity function). Specifically, we will bound the success probability when the input is chosen uniformly at random:

Theorem 3 (Low-degree polynomials over $\mathbb{F}_2$ have low correlation with the majority function)**.** *If p is an n-variate degree-D polynomial over $\mathbb{F}_2$, then*

$$\Pr_{x \in \{0,1\}^n}[p(x) = \mathsf{MAJ}(x)] \leq \frac{1}{2} + O(D/\sqrt{n}).$$

The first step in the proof of Theorem 3 is to show that if q is a nonzero low-degree polynomial, then every point in $\mathbb{F}_2^n$ is close to a point that q accepts. Let $\Delta(\cdot,\cdot)$ denote Hamming distance.

Lemma 2 (If q has low degree, then every point is close to $q^{-1}(1)$)**.** *Suppose $q \colon \mathbb{F}_2^n \to \mathbb{F}_2$ is a nonzero polynomial. Then for every $x \in \mathbb{F}_2^n$, we have $\Delta(x, q^{-1}(1)) \leq \deg(p)$.*

[2] Here we are using the fact that MAJ_a reduces to MAJ_b whenever $a \leq b$. This is because $\mathsf{MAJ}_a(x) = \mathsf{MAJ}_{a+1}(x1)$ if a is even, and $\mathsf{MAJ}_a(x) = \mathsf{MAJ}_{a+1}(x0)$ if a is odd.

Proof. Let $r(y) = q(x+y)$. Then r is another nonzero polynomial with $\deg(r) = \deg(q)$. Let y be the indicator for some minimal nonzero term of r. Then $r(y) = 1$, so $q(x+y) = 1$. ☐

Let $B_w(x)$ be the Hamming ball of radius w centered at $x \in \{0,1\}^n$, i.e., $B_w(x) = \{y : \Delta(x,y) \leq w\}$. Based on Lemma 2, we now show that arbitrary functions on small Hamming balls can be interpolated by low-degree polynomials.

Lemma 3 (Low-degree interpolation on small Hamming balls)**.** *Let $x \in \mathbb{F}_2^n$, let $w \leq n$, and let $f\colon B_w(x) \to \mathbb{F}_2$ be any function. There exists a polynomial $q\colon \mathbb{F}_2^n \to \mathbb{F}_2$ of degree at most w such that $q(y) = f(y)$ for every $y \in B_w(x)$.*

Proof. Let $\mathcal{P}$ be the space of all n-variate polynomials of degree at most w over $\mathbb{F}_2$. Let $\mathcal{F}$ be the space of all functions mapping $B_w(x)$ to $\mathbb{F}_2$. Define $\Psi\colon \mathcal{P} \to \mathcal{F}$ by the rule $\Psi(q) = q|_{B_w(x)}$, i.e., $\Psi(q)$ is q restricted to $B_w(x)$. Our goal is to prove that Ψ is surjective. Observe that $|\mathcal{P}| = |\mathcal{F}| = 2^{|B_w(x)|}$. Therefore, it is equivalent to show that Ψ is injective, which is what we will do next.

Suppose $\Psi(q_1) = \Psi(q_2)$, i.e., q_1 and q_2 are polynomials of degree at most w that agree on $B_w(x)$. Let $q' := q_1 + q_2$. Then q' is a polynomial of degree at most w that is zero on $B_w(x)$. By Lemma 2, this implies that $q' \equiv 0$, and hence $q_1 \equiv q_2$. ☐

Using Lemma 3, we will now show that there is a low-degree reduction from arbitrary functions to the majority function, analogous to our analysis of polynomials approximating PARITY over $\mathbb{F}_3$.

Lemma 4 (Low-degree reduction from arbitrary functions to the majority function)**.** *Every function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ can be written in the form $f(x) = p_0(x) + p_1(x) \cdot \mathsf{Maj}(x)$, where p_0 and p_1 have degree at most $n/2$.*

Proof. Let p_0 agree with f on $B_{\lceil n/2 \rceil - 1}(0^n)$. Let p_1 agree with $f + p_0$ on $B_{\lfloor n/2 \rfloor}(1^n)$. ☐

Proof of Theorem 3. Let $S = \{x : p(x) = \mathsf{Maj}(x)\}$. By Lemma 4, every function $f\colon S \to \mathbb{F}_2$ can be written as $f(x) = p_0(x) + p_1(x) \cdot p(x)$, a polynomial of degree at most $n/2 + D$. The number of functions $f\colon S \to \mathbb{F}_2$ is $2^{|S|}$. On the other hand, the number of polynomials of degree at most $n/2 + D$ is $2^{\sum_{i=0}^{n/2+D} \binom{n}{i}}$. Therefore, $|S| \leq \sum_{i=0}^{n/2+D} \binom{n}{i} \leq 2^n \cdot (1/2 + O(D/\sqrt{n}))$. ☐

Corollary 2 ($\mathsf{MAJ} \notin \mathsf{AC}^0[\oplus]$)**.** *Every $\mathsf{AC}^0_d[\oplus]$ circuit computing MAJ_n has size $2^{n^{\Omega(1/d)}}$.*

Using similar techniques, one can show more generally that $\mathsf{MAJ} \notin \mathsf{AC}^0[m]$ whenever m is a power of a prime. However, when m is composite, these techniques break down. It is an open problem to prove that $\mathsf{MAJ} \notin \mathsf{AC}^0[6]$. In fact, it is an open problem to rule out the ridiculous suggestion that $\mathsf{NP} \subseteq \mathsf{AC}^0[6]$! However, the situation is not completely bleak; there are some known techniques for proving that "somewhat explicit" functions cannot be computed by $\mathsf{AC}^0[6]$ and similar classes. For example, Murray and Williams proved that $\mathsf{NQP} \not\subseteq \mathsf{ACC}$ [MW18], where NQP denotes nondeterministic quasipolynomial time and $\mathsf{ACC} = \bigcup_m \mathsf{AC}^0[m]$.

References

[MW18] Cody Murray and Ryan Williams. "Circuit lower bounds for nondeterministic quasi-polytime: an easy witness lemma for NP and NQP". In: *Proceedings of the 50th Annual Symposium on Theory of Computing (STOC)*. 2018, 890–901. DOI: 10.1145/3188745.3188910.