

Claim: $\forall t, \forall c, \forall p(n)$, PARITY cannot be computed using a depth t circuit of size $p(n)$ that has input-level fan-in $\leq c$.

Corollary: $\text{PARITY} \notin \text{AC}^0$.

Corollary follows from Claim because:

- Suppose that PARITY has depth t , polysize circuits, of unbounded input-level fan-in.
- Add one more layer of OR (or AND) gates, that simply copies the values of the inputs.
- Connect the old first layer to this new layer as if those gates were the inputs.
- Then PARITY has depth $(t + 1)$, polysize circuits, of input-level fan-in 1.

Proof of Claim

Case $t = 2$ done.

For $t > 2$, let t be the least such depth, and suppose that S_1, S_2, S_3, \dots is a family of circuits of depth t and polysize computing PARITY. Then we are going to produce S'_1, S'_2, S'_3, \dots , of depth $(t - 1)$ and polysize, thereby getting a contradiction.

To produce S'_n we are going to take S_{4n^2} and set $(4n^2 - n)$ of its variables to 0 and 1 in such a way that we can use the distributive law on levels 1 and 2 to exchange these two levels *without increasing the size of the circuit exponentially*. To do this, it is enough to have a situation where *each gate on level 2 depends on only a constant number of variables*. Then, levels 2 and 3 can be collapsed to a single layer leaving a circuit of depth $(t - 1)$.

How can we come up with the right restriction mapping S_{4n^2} to S'_n ? We use a random restriction and show that the probability of getting an appropriate substitution is positive, so we can conclude that one exists. The random restrictions have the following probabilities:

$$\Pr[x_i \text{ remains }] = \frac{1}{\sqrt{n}}$$

$$\Pr[x_i \text{ is set to } 0] = \frac{1 - \frac{1}{\sqrt{n}}}{2}$$

$$\Pr[x_i \text{ is set to } 1] = \frac{1 - \frac{1}{\sqrt{n}}}{2}$$

Let x_i^r be a random restriction of x_i according to the above probability distribution, so $x_i^r \in \{x_i, 0, 1\}$, and S_n^r is par_m of $\overline{\text{par}}_m$ where $m \leq n$.

Claim: $\Pr[\text{there are fewer than } \frac{\sqrt{n}}{2} \text{ vars in } S_n^r] = O(\frac{1}{\sqrt{n}})$.

Markov Inequality $\Pr[X \geq a] \leq \frac{E(X)}{a}$

Chebyshev Inequality $\Pr[|X - E(X)| \geq a] \leq \frac{V(X)}{a^2}$

Now consider r acting on S_n . The probability of setting a variable to 0 or 1 is $q = 1 - \frac{1}{\sqrt{n}}$, and the probability of leaving it unset is $p = \frac{1}{\sqrt{n}} = 1 - q$.

This is done independently for each variable, and so it is a *binomial distribution* with probability of having exactly k successes (i.e., of leaving exactly k variables unset) at $\binom{n}{k} p^k q^{n-k}$.

Let X be the random var counting the number of successes (i.e., counting the number of unset vars) in n trials.

$$X_i = \begin{cases} 1 & \text{if } i\text{-th trial is a success, i.e., } x_i^r = x_i \\ 0 & \text{otherwise} \end{cases}$$

$$E(X_i) = 1 \cdot p + 0 \cdot q = p, \text{ and}$$

$$E(X) = E(\sum_{i=1}^n X_i) = \sum_{i=1}^n E(X_i) = n \cdot p \text{ and}$$

$$\begin{aligned} \text{Var}(X) &= E(X^2) - E(X)^2 = E\left(\left(\sum_{i=1}^n X_i\right)^2\right) - (n \cdot p)^2 \\ &= E\left(\sum_{i=1}^n \sum_{j=1}^n X_i X_j\right) - (n \cdot p)^2 \\ &= n(n-1)p^2 + \underbrace{n \cdot p}_{(*)} - (n \cdot p)^2 = np - np^2 = np(1-p) \end{aligned}$$

where $(*)$ is for when $i = j$, because $E(X_i X_i) = p$.

We are now ready to prove the claim.

$$\begin{aligned}
& \Pr[\text{fewer than } \frac{\sqrt{n}}{2} \text{ vars remain in } S_n^r] \\
& \leq \Pr[X \leq \frac{\sqrt{n}}{2}] = \Pr[\sqrt{n} - X \geq \sqrt{n} - \frac{\sqrt{n}}{2}] = \Pr[E(X) - X \geq \frac{\sqrt{n}}{2}] \\
& \leq \Pr[|X - E(X)| \geq \frac{\sqrt{n}}{2}]
\end{aligned}$$

Using Chebyshev inequality, the last line can be bounded by

$$\leq \frac{\text{Var}(X)}{\left(\frac{\sqrt{n}}{2}\right)^2} = \frac{\sqrt{n} - 1}{\binom{n}{4}} \leq 4 \cdot \frac{\sqrt{n}}{n} = 4 \cdot \frac{1}{\sqrt{n}}$$

which finishes the proof of the claim.

We are now going to show that after a random restriction r , the AND gates on the second level depend on a constant number of variables with high probability.

Assume k is strictly bigger than the degree of the poly bounding $S = \langle S_i \rangle$.

We prove by induction on c that the AND gates depend on “too many variables” with probability $O(\frac{1}{n^k})$.

In the first case (large fan-in), we have:

$$\Pr[\text{AND-gate is } \textit{not} 0] = O\left(\frac{1}{n^k}\right)$$

This is because it is very likely that a random r would have set one of the AND's inputs to 0.

So with high probability, the AND-gate depends on no variables.

Detailed proof:

$$\begin{aligned} & \Pr[\text{AND-gate not } 0] \\ & \leq \Pr[\text{all its inputs are non } 0] \\ & \leq \Pr[\text{a fixed input not } 0]^{4k \ln(n)} \\ & \leq \left(\frac{3}{4}\right)^{4k \ln(n)} \end{aligned}$$

since for $n \geq 4$, $1 - \frac{1 - \frac{1}{\sqrt{n}}}{2} \leq \frac{1 + \frac{1}{\sqrt{n}}}{2} \leq \frac{3}{4}$

$$= n^{4k \ln(\frac{3}{4})} \leq n^{-k}$$

In the second case (small fan-in), we have:

$$\Pr[\text{AND-gate depends on } \textit{more than } 18k \textit{ inputs}] = O\left(\frac{1}{n^k}\right)$$

Detailed proof:

We need the following approximation for the binomial distribution:

$$\Pr[X \geq a] = \sum_{i=a}^n \binom{n}{i} p^i (1-p)^{n-i} \leq (np)^a$$

$\Pr[\text{AND-gate depends on more than } a \text{ variables}]$

$$\leq \sum_{i=a}^{4k \ln(n)} \binom{4k \ln(n)}{i} (1/\sqrt{n})^i (1 - 1/\sqrt{n})^{n-i}$$

and using the above approximation,

$$\leq (1/\sqrt{n})^a 2^{4k \ln(n)} = n^{-a/2} n^{8k} = n^{8k-a/2}$$

Now letting $a = 18k$ completes the proof.

Induction Step: Assume the result holds for $(c - 1)$.

Again, there are two cases:

1. Before the random restriction, the AND-gate has many OR-gates below it with *disjoint input variables*
 $(d \cdot \ln(n), d = k \cdot 4^c)$
2. Before the random restriction, the AND-gate has few OR-gates below it with *disjoint input variables*

In the first case (many OR-gates below AND), it is very likely that after the random restriction one of the OR-gates will have all of its inputs set to 0, and so the AND-gate is 0, and so it depends on no variables.

$$\Pr[\text{AND-gate is } \textit{not} = 0] = O\left(\frac{1}{n^k}\right)$$

Detailed proof:

$$\begin{aligned} & \Pr[\text{AND-gate not } 0] \\ & \leq \Pr[\text{none of the OR-gates becomes } 0] \\ & \leq \Pr[\text{a fixed OR-gate doesn't become } 0]^{d \cdot \ln(n)} \end{aligned}$$

and for $n \geq 4$,

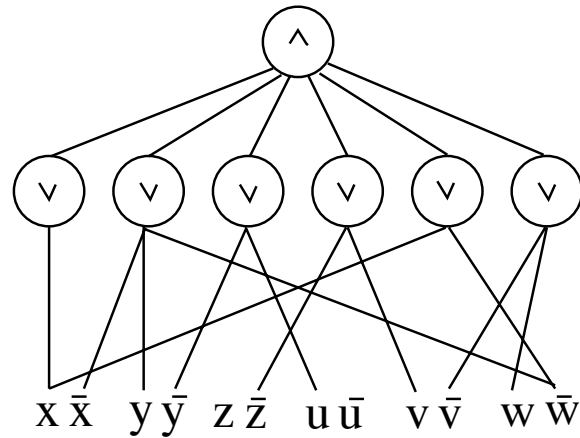
$$\begin{aligned} & \leq (1 - 4^{-c})^{d \cdot \ln(n)} \\ & = n^{d \cdot \ln(1 - 4^{-c})} \end{aligned}$$

and since $\ln(1 - x) \leq -x$,

$$\leq n^{-d \cdot 4^{-c}} = n^{-k}$$

In the second case (few OR-gates below AND), choose a *maximal* set of OR-gates with *disjoint input variables*.

Let H be the set of the variables that occur in these OR gates.



In this example, a maximal set of ORs would consist of gates (numbered from left) $\{1, 3, 4\}$ and $H = \{x, y, u, z, v\}$.

$$|H| \leq c \cdot d \cdot \ln(n).$$

Each of the OR-gates has a variable in H . (*)

Let $\{\tau_1, \tau_2, \dots, \tau_l\}$, $l = 2^{|H|}$, be all the truth assignments to variables in H .

Let $A_i = C^{\tau_i}$, after simplifying, where C represents the whole AND-OR circuit.

By (*), each OR gate in A_i either disappears or “loses” an input.

So the input fan-in in each A_i is at most $c - 1$.

By the induction hypothesis, the probability that A_i^r depends on more than e_{c-1} variables is bounded above by $O(\frac{1}{n^k})$.

If $H = \{x_1, x_2\}$, then

$$C = (\bar{x}_1 \wedge \bar{x}_2 \wedge A_1) \vee (\bar{x}_1 \wedge x_2 \wedge A_2) \vee (x_1 \wedge \bar{x}_2 \wedge A_3) \vee (x_1 \wedge x_2 \wedge A_4)$$

Generalize, set $h = |H|$,

$$C = \bigvee_{\tau \in 2^h} (x_1^\tau \wedge \dots \wedge x_h^\tau) \wedge A_\tau \quad (**)$$

where x_i^τ is x_i if $\tau(x_i) = 1$, and \bar{x}_i otherwise.

Let h be the number of variables in H after a random restriction.

$$\Pr[h > 4cd + 2k] = O\left(\frac{1}{n^k}\right)$$

Detailed proof:

Since $|H| \leq cd \ln(n)$ and using our binomial approximation

$$\begin{aligned} & \Pr[h > a] \\ & \leq 2^{cd \ln(n)} \cdot (1/\sqrt{n})^a \\ & \leq n^{2cd} \cdot n^{-a/2} \\ & = n^{2cd - a/2} \end{aligned}$$

Solving $2cd - a/2 = -k$ for a , we get that $a = 4cd + 2k$.

Thus, with high probability

$$2^h \leq 2^{4cd+2k} =: m$$

Hence, there are at most m terms in (**), so

$$\boxed{e_c = m \cdot e_{c-1} + (4cd + 2k)}$$

and $\Pr[C \text{ depends on more than } e_c \text{ variables }]$

$$\leq \Pr[h > 4cd + 2k] + m \cdot \Pr[A_j \text{ depends on more than } e_{c-1} \text{ variables }]$$

$$\leq O\left(\frac{1}{n^k}\right) + m \cdot O\left(\frac{1}{n^k}\right) = O\left(\frac{1}{n^k}\right)$$