LECTURE 7

- Boolean circuits:
  - Directed acyclic graphs (DAGs)
  - nonsource nodes labelled by AND $\wedge$ (fan-in 2), OR $\vee$ (fan-in 2), NOT $\neg$ (fan-in 1)
  - source nodes labelled by variables
  - output: value computed by circuit
- Size = # nodes
  Depth = length of longest path
- Language decided by family of circuits $\{C_n\}_{n \in \mathbb{N}^+}$
- $P_{/poly}$ polynomial-size circuits
- $P \subsetneq P_{/poly}$
- $NP \not\subseteq P_{/poly}$ unless PH collapses
- Almost all functions $f: \{0,1\}^n \to \{0,1\}$ are hard (require circuit size $\Omega(2^n/n)$)

Soooo... What about lower bounds for explicit functions?

Topic of today's lecture

Study subclasses of $P_{/poly}$

Motivation: massively parallel computers

Rough model:

- n processors ($n \to \infty$)
- fast interconnection network ($O(\log n)$ steps for interprocessor communication possible)
- synchronization via global clock

A problem has an efficient parallel algorithm if it can be solved for inputs of size n using $n^{O(1)}$ processors in $\log^{O(1)} n$ time

polylog

DEF 1 (NC) (N = Nick)

$L \in NC^d$ if decided by $\{C_n\}_{n \in \mathbb{N}}$ in poly size and depth $O(\log^d n)$

$$NC = \bigcup_{i \in \mathbb{N}} NC^i$$

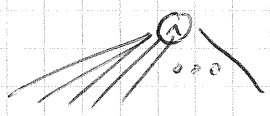
Can also define uniform version by requiring $\exists$ TM M that generates $\{C_n\}$ (in poly time or log space).

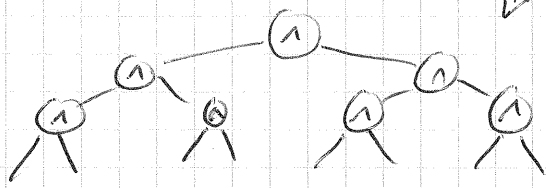
DEF 2 (AC) (A = alternations)

Circuits with unbounded fan-in

$L \in AC^d$ if decided by $\{C_n\}_{n \in \mathbb{N}}$ in poly size and depth $O(\log^d n)$

poly(n) fan-in AND/OR-gate $\Rightarrow$ fan-in 2 tree of depth $O(\log n)$





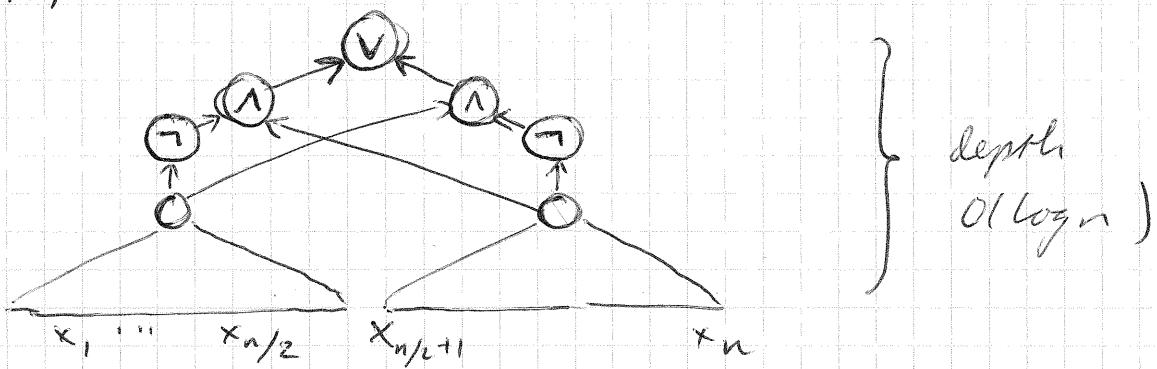
Hence $NC^d \subseteq AC^d \subseteq NC^{d+1}$

NC^0 – circuits depend only on constant # bits – no fan

AC^0 already nontrivial

EX 3 Let PARITY $= \{x \mid x$ has odd # 1's$\}$

PARITY $\in NC^1$



MAIN GOAL TODAY: PARITY $\notin AC^0$

But before that, back to parallelism

THM 4

A language has an efficient parallel algorithm iff it is in (uniform) NC

Proof sketch:

($\Rightarrow$) N processors, time D
have $O(N)$ nodes in $O(D)$ layers.
ith node in tth layer performs computation of processor i at time t

($\Leftarrow$) Given C_n, assign one node to each processor. Each processor waits for input, computes value, sends to successors.

See textbook for slightly more details. □

Are all poly-time algorithms efficiently parallelizable? I.e., is P = (uniform) NC?
Believe: no. Yet another challenging open problem.

Motivates theory of P-completeness

DEF 5 $\mathcal{L}$ is P-complete if $\mathcal{L} \in P$ and every $\mathcal{L}' \in P$ is logspace-reducible to $\mathcal{L}$.

THM 6 Let $\text{CIRCUITEVAL} = \{\langle C, x\rangle \mid \begin{array}{l} C \text{ circuit with } |x| \text{ inputs} \\ C(x) = 1 \end{array}\}$

Then CIRCUITEVAL is P-complete.

Proof: See textbook (for sketch).

Circuit lower bounds - brief history

Connection to P vs NP in early 70s

Great hopes in 70s & 80s - explicit combinatorial objects easier than TMs

General circuits Best lower bound $5n - o(n)$
($2n$ is trivial)

Restricted models

(a) Monotone circuits (no NOT gates) - exp $\subset$ B

(b) Bounded (constant) depth - exp $\subset$ B

(c) Bounded depth + "counting gates" - frontier

So get smile | for | already | for fairly weak models

Some exciting developments in last few years

But still fairly weak models.

Rest of today's lecture: Focus on (b)

Every Boolean function computable by CNF / DNF formula of exponential size
— depth 2 circuit

For depth 2 also not too hard to prove exponential lower bounds

Techniques don't seem to generalize to depth 3 or larger

THM # [Furst, Saxe, Sipser '81; Ajtai '83]

PARITY $\notin AC^0$

In fact, circuits in bounded depth for parity must have exponential size — proven by Johan Håstad (but we won't try to prove optimal result today)

MAIN TOOL: Random restrictions

Pick / subset / random / of variables

Set to random values true/false (1/0)

Simplify circuit (remove AND-gates with 0-input, OR-gates with 1-input)

HIGH LEVEL IDEA

1. Suppose $\exists$ circuit C_n for PARITY in poly size and depth d
2. Choose random subset of (most) variables — all except n^ε for $\varepsilon > 0$ depending on d — restrict randomly, and simplify C_n
3. Since C_n small and shallow, random restriction will collapse circuit to constant value.
4. But this is impossible — C_n should still compute parity (or negation of parity) of n^ε variables.

Hence no AC^0-circuits for PARITY, Q.E.D.

Notation

k-CNF formula — AND of ORs of size $\le k$

k-DNF formula — OR of ANDs of size $\le k$

f function, ρ partial assignment

$f|_\rho$ — f restricted by ρ

$f|_\rho(\vec{x}) = f(\rho \circ \vec{x})$

$Vars(\rho) = \{$set of variables assigned to by $\rho\}$

LEMMA 8 HÅSTAD'S SWITCHING LEMMA

Suppose $f : \{0,1\}^n \to \{0,1\}$ can be written as a k-DNF formula and let ρ be a uniformly random restriction to t uniformly variables. Then for every $s \ge 2$ it holds that

$$\Pr_\rho\left[f|_\rho \text{ cannot be written as an } s\text{-CNF formula}\right] \le \left(\frac{(n-t)k^{10}}{n}\right)^{s/2} \quad (*)$$

Use switching lemma with

$k = O(1)$
$s = O(1)$
$t \approx n - \sqrt{n}$

$\Rightarrow \Pr\left[f|_\rho \text{ not } s\text{-CNF}\right] \le n^{-c}$ for some constant c

Can apply switching lemma also to $\neg f$, interchanging terms DNF and CNF.

ROADMAP

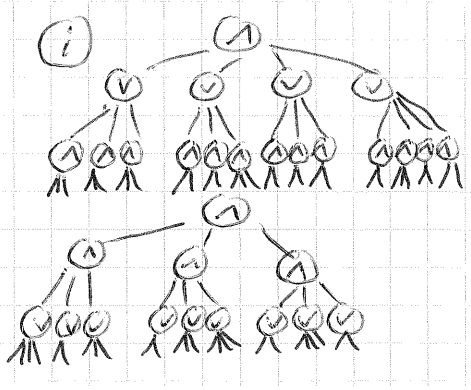
PARITY $\notin AC^0$

(1) Use switching lemma to prove ~~$AC^0 \notin$~~

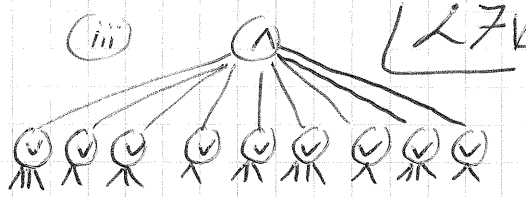
(2) Prove switching lemma

Outline (a) Suppose circuit of depth d

(i) Hit with restriction on bottom layers - turns DNFs into CNFs with high probability



(iii) But if so, then can collapse two layers
⇒ depth decreased to $d-1$ with high probability



(iv) Repeat for $d-2$ steps
⇒ get k-CNF or k-DNF formula

(v) Just fix k more variables to make disjunction false / conjunction true ⇒ fixed value

(vi) But still variables left that can flip value of fixed function. Contradiction

Formal proof of Thm 7

Start with poly-size AC^0 circuit of constant depth
Do initial preprocessing to get circuit s.t.
(a) all fan-outs are 1 (circuit is a tree) (DAG)
(b) all NOT gates are at input level (only negate variables)
(c) AND and OR gates alternate (at each level only $\vee$ or $\wedge$)
(d) bottom level has AND gates with fanin 1.

Claim A This preprocessing can be done without increasing depth and blowing up size only polynomially

Proof of claim: Exercise.

Suppose resulting circuit has depth d and $\leq n^b$ gates. Let $n_0 = n =$ # variables

For $i = 1, 2, \ldots, d-2$ steps $\quad n_i =$
- Restrict $n_{i-1} - \sqrt{n_{i-1}}$ variables, leaving $\sqrt{n_{i-1}}$ unset
- Collapse circuit one layer
- keep bottom layer at constant fan-in

$$n_i = \sqrt{n_{i-1}} = \dots = \sqrt[2^i]{n}$$

Let $k_i = 10 \cdot 2^i$ $(= O(1))$.

Suppose bottom layer contains $\wedge$ gates, level above $\vee$ gates
(opposite case entirely symmetric)

Show that whp after ith restriction have
depth $-(d-i)$ circuit with $\le k_i$ fan-in at bottom level

By assumption each such $\vee$ gate computes k_i-DNF. In $(i+1)$st step,

Use Håstad's switching lemma with

$$n = n_{i \text{ in}} = n^{1/2^i}$$
$$t = \text{min } n_i - n_{i+1} = n^{1/2^i} - n^{1/2^{i+1}}$$
$$k = k_i$$
$$s = k_{i+1}$$

For a fixed $\vee$ gate, k_i-DNF turns into k_{i+1}-CNF
except with probability

$$\left(\frac{n^{1/2^{i+1}} \cdot k_i^{10}}{n^{1/2^i}} \right)^{k_{i+1}/2} \le K \cdot \left(n^{-\frac{1}{2^{i+1}}} \right)^{k_{i+1}/2} \le n^{-\frac{5}{2}\sigma} \le 1/(10 n^\sigma)$$

(K: some constant $\le k_d^{10kd}$)

for n large enough.

If this happens for all $\vee$ gates in next-to-bottom layer, can collapse $\wedge$ gates with $\wedge$ gate above

$\Rightarrow$ depth reduces by 1; fan-in $k_{i+1} = O(1)$.
k_{i+1}-CNFs

In next step, reduce k_{i+1}-CNFs to k_{i+2}-DNFs in the same way

Now that we only process each node/gate once

i.e. at most n^6 times.

We can fail if $\exists$ circuit/gate where collapse does not happen.

Look at random restriction ρ constructed as above

$$\Pr[\rho \text{ fails to collapse } C] \leq \Pr[\exists \text{ gate } v \text{ s.t. } \rho \text{ fails to collapse } v]$$

$$\leq \sum_{v \text{ gate}} \Pr[\rho \text{ doesn't collapse } v] \leq$$ (union bound)

$$\leq n^6 \cdot \frac{1}{10 n^6} = 1/10$$

So whole process works with $> 90\%$ prob.

In particular, must $\exists$ one good ρ that after $d-2$ steps collapses C to k_{d-2}-CNF or k_{d-2}-DNF

Set $k_{d-2} = O(1)$ additional variables to falsify disjunction or satisfy conjunction.

But still $n^{1/2^{d-1}} - k_{d-2}$ variables left, so function is not constant

Hence the circuit cannot have been computing parity

QED

Proof of Hastad's switching lemma

Original proof technically fairly complicated
Will do more intuitive proof by Razborov

MINTERM partial assignment to variables fixing $f = 1$
MAXTERM — " — $f = 0$

Every conjunction / term in k-DNF is minterm of size $\leq k$
Every disjunction in k-CNF is (negation of) maxterm of size $\leq k$

Assume we pick minterms and maxterms minimal

Claim B If all (minimal) maxterms of Boolean function f are of size $\leq s$ then f can be written as an s-CNF formula.

Proof Exercise

Hence if $f|_\rho$ is not an s-CNF then it has a maxterm of size $\geq s+1$.

R_t = all restrictions on t variables

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

B = bad restrictions ρ for which $f|_\rho$ not s-CNF

Random restriction succeeds with prob $1 - \frac{|B|}{|R_t|}$

Prove $|B|$ small compared to $|R_t|$

Construct [mapping] [one-to-one (injective)]

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

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

$$\frac{|B|}{|R_t|} \leq \frac{|R_{t+s} \times \{0,1\}^\ell|}{|R_t|} = \frac{\binom{n}{t+s} 2^{t+s} 2^{O(s \log)}}{\binom{n}{t} 2^t} =$$

$$\frac{\binom{n}{t+s} k^{O(s)}}{\binom{n}{t}} \leq \left[\begin{array}{l} k, s \text{ constants} \\ t \text{ close to } n \end{array}\right]$$

$$\frac{\binom{n}{t} / n^s}{\binom{n}{t}} \cdot k^{O(s)} = n^{-\Omega(s)}.$$

Claim C The above handwaving can be worked out to prove (*). ~~by e stat~~

Proof Show that for $t > n/2$

$$\binom{n}{t+s} \leq \binom{n}{t} \left(\frac{e(n-t)}{n}\right)^s$$

by first proving

$$\binom{n}{t+s} = \binom{n}{t}\binom{n-t}{s} / \binom{t+s}{s}$$

using

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k$$

Details left as exercise.

So we're done if we can construct one-to-one mapping

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

Look at $f|_\rho$ for bad $\rho \in B$.
No term in k-DNF becomes 1
Some terms might become 0, but not all
$f|_\rho$ has some (minimal) maxterm π of size $> s$.

Let $\rho \sqcup \pi$ denote union of restrictions $\rho : S \to \{0,1\}$, $\pi : T \to \{0,1\}$
for $S \cap T = \emptyset$ disjoint

Look at $\rho \sqcup \pi$
Map ρ to $\rho\sigma \in R_{\sigma+s}$
Add extra info so that $\rho\sqcup\pi$ and π can be recovered from $\rho\sigma \Rightarrow \rho$ recoverable
$\Rightarrow$ mapping one to one.

Order terms of f in some fixed order
In each term, order variables in some fixed order.
$\rho \sqcup \pi$ sets f to 0
ρ does not

Look at first term t_1 not falsified by ρ
Let $Y_1 = Vars(t_1) \cap Vars(\pi) \neq \emptyset$
Let π_1 restriction to Y_1 that coincides with π
Let σ_1 unique restriction to Y_1 that satisfies t_1

In step $i+1$, have defined

$\pi_1, \ldots, \pi_{i+1} \subsetneq \pi$

$\sigma_1, \ldots, \sigma_{i+1}$

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

$Y_{i+1} = (\mathrm{Vars}(t_{i+1}) \cap \mathrm{Vars}(\pi)) \setminus \mathrm{Vars}\,\rho\,\pi_1 \ldots \pi_i$

π_{i+1} restriction to Y_{i+1} coinciding with π

σ_{i+1} unique restriction to Y_{i+1} satisfying t_{i+1}

Stop when $\pi_1 \cdots \pi_m$ assigns $\geq s$ variables, trim down π_m so that size exactly $= s$.

Map ρ to $(\rho\,\sigma_1 \cdots \sigma_m, c)$

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

Given $\rho\,\sigma_1 \ldots \sigma_m$, find t_1: first term not falsified

ρ does not falsify it
σ_1 does not by construction
σ_i, $i > 1$, don't assign to $\mathrm{Vars}(t_1)$

Let $s_1 = |\mathrm{Vars}(\pi_1)|$

Let c contain

- s_1
- indices of these vars in t_1 (remember fixed order)
- values assigned by π_1

At most k vars in term $\Rightarrow$ $O(s_1 \log k)$ bits needed

Given this, can reconstruct π_1

Can flip $\rho\,\sigma_1\sigma_2 \ldots \sigma_m$ to

$\rho\,\pi_1\sigma_2 \ldots \sigma_m$

L7 XIV

Find t_2 – first term not falsified by $\rho \pi_1 \sigma_2 \cdots \sigma_m$

Add to c

- $s_2 = |\text{Vars}(\pi_2)|$
- indices in t_2 of these vars
- their values in π_2

$O(s_2 \log k)$ bits

Reconstruct π_2

Flip $\rho \pi_1 \sigma_2 \sigma_3 \cdots \sigma_m$ to $\rho \pi_1 \pi_2 \sigma_3 \cdots \sigma_m$

Repeat.

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

$\Rightarrow$ read off $\rho \Rightarrow$ one-to-one mapping

Total length of c

$O((s_1 + s_2 + \cdots + s_m) \log k) = O(s \log k)$.

But $\sum_{i=1}^{m} s_i = s$ — The lemma follows.

SUMMING UP

$NC^i \subseteq AC^i \subseteq NC^{i+1}$

- Defined AC^i NC^i

poly(n) size ~~log~~ $O(\log^i n)$ depth unbounded / bounded fan-in

- $NC \approx$ efficiently parallelizable problems
- $P \neq NC$? P-completeness
- General circuit lower bounds only $\Omega(n)$
- But can prove $AC^0 \neq NC^1$

  PARITY $\in NC^1$ PARITY $\notin AC^0$