DD2445 COMPLEXITY THEORY: LECTURE 20

RECAP

Monotone circuits: AND, OR (binary)
no NOT-gates

For $x, y \in \{0,1\}^n$ write $x \leq y$ if $\forall i\ x_i \leq y_i$

MONOTONE FUNCTION $x \leq y \Rightarrow f(x) \leq f(y)$

$CLIQUE_{k,n}$: Input $\binom{n}{2}$ bits – indicators for edges in n-vertex graph
Output: $1 \Leftrightarrow$ graph contains k-clique

THM 4 $\exists \varepsilon > 0\ \forall k \leq n^{1/4}$ no monotone circuit of size $< 2^{\varepsilon\sqrt{k}}$ * computes $CLIQUE_{k,n}$

No implications for general circuits:
Non-monotone circuits can be much more efficient in ~~[illegible]~~ computing monotone functions

However there exist monotone functions for which monotone ~~functions~~ circuits are optimal up to polynomial factors [Berkowitz '82]

f is slice function if $\exists\ k \in \mathbb{N}^+$ s.t.

$$f(x) = \begin{cases} 1 & \text{if } \sum_i x_i > k \\ 0 & \text{if } \sum_i x_i < k \\ \text{something interesting} & \text{if } \sum_i x_i = k \end{cases}$$

And there exist NP-complete slice functions

(*) Not quite the right bound – can get $n^{\Omega(\sqrt{k})}$
Our proof here silently assumes $k \geq n^{\delta}$ or so – can be fixed

For $S \subseteq [n]$, CLIQUE INDICATOR

$$C_S(G) = \begin{cases} 1 & \text{if } S \text{ forms clique in } G \\ 0 & \text{o/w} \end{cases}$$

ORs of too few clique indicators really bad at computing $\mathrm{CLIQUE}_{k,n}$

Create distributions on yes- and no-instances

$\mathcal{Y}$: Choose $K \subseteq [n]$, $|K| = k$, at random
Output graph $G = (V, E)$ with
$V = [n]$, $E = \{(u,v) \mid u \neq v,\ u, v \in K\}$

$\mathcal{N}$: Choose $c : [n] \to [k-1]$ at random
Output graph $G = (V, E)$ with
$V = [n]$, $E = \{(u,v) \mid c(u) \neq c(v)\}$

<u>COROLLARY 6</u>

Suppose $C' = \bigvee_{i=1}^{m} C_{S_i}$
n large enough; $k \leq n^{1/4}$;
$m \leq n^{\sqrt{k}/20}$

Then C' fails on 99% of either $\mathcal{Y}$ or $\mathcal{N}$.

This was where we ended last time

Now we want to show

From small monotone circuit for $\mathrm{CLIQUE}_{k,n}$

$\Downarrow$

Can build OR of somewhat small clique indicators that are decent at distinguishing $\mathcal{Y}$ and $\mathcal{N}$

LEMMA 7 Assume

C monotone circuit of size $s < 2^{\sqrt{k}/2}$

Then $\exists$ collection $S_i \subseteq [n]$ for $i \in [m]$, $m \leq n^{\sqrt{k}/20}$, such that

$$\Pr_{G \sim \mathcal{Y}}\left[\bigvee_{i=1}^{m} C_{S_i}(G) \geq C(G) \right] > 0.9 \qquad (*)$$

$$\Pr_{G \sim \mathcal{N}}\left[\bigvee_{i=1}^{m} C_{S_i}(G) \leq C(G) \right] > 0.9 \qquad (**)$$

From this Thm 4 immediately follows:

- Assume circuit C of size $< 2^{\sqrt{k}/2}$
- Lemma 7 $\Rightarrow$ OR of few clique indicators that do well on both $\mathcal{Y}$ and $\mathcal{N}$
- Contradicts Lemma 6. So no such circuit; QED $\boxtimes$

So let us prove Lemma 7

Set

$$\ell = \sqrt{k}/10$$
$$p = 10\sqrt{k}\log n$$
$$m = (p-1)^{\ell}\,\ell!$$

Observe $m \ll n^{\sqrt{k}/20}$

$$m = (p-1)^{\ell} \cdot \ell! < p^{\ell} \cdot e^{\ell}$$
$$= (k \cdot \log n)^{\sqrt{k}/10} < k^{\sqrt{k}/8}$$
$$\leq n^{\sqrt{k}/32} \ll n^{\sqrt{k}/20}$$

for n large enough (and $k \geq n^{\delta}$ for some $\delta > 0$).

MCA IV

Sort gates of circuit in topological order

Get functions $f_i \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ for $i = 1, \dots, s$ where

(a) f_i = input $x_{u,v}$, or

(b) $f_i = f_j \vee f_k$ $j,k < i$, or

(c) $f_i = f_j \wedge f_k$ $j,k < i$

Function computed by $C = f_s$

Construct sequence of functions $\tilde{f}_1, \dots, \tilde{f}_s$ s.t.

(1) $\tilde{f}_i = \bigvee_{j=1}^{m'} C_{S_j}$ for $|S_j| \leq \ell$, $m' \leq m$

Call this an (m, ℓ)-FUNCTION

(2) $\tilde{f}_i$ approximates f_i well on $\mathcal{Y}$ and $\mathcal{N}$

Construction by induction $\tilde{C} := \tilde{f}_s$

(a) f_i = input $\Rightarrow$ $\tilde{f}_i = f_i$

(b) Define APPROXIMATE OR $\sqcup$ and set $\tilde{f}_i = \tilde{f}_j \sqcup \tilde{f}_k$

(c) Define APPROXIMATE AND $\sqcap$ and set $\tilde{f}_i = \tilde{f}_j \sqcap \tilde{f}_k$

By construction, $\sqcap$ and $\sqcup$ will yield (m, ℓ)-functions

Want to prove four properties

Suppose that $h = f \circ g$ for $\circ \in \{\vee, \wedge\}$ in what follows

(i) $\Pr_{G \sim \mathcal{Y}} \left[\tilde{f} \sqcup \tilde{g}(G) < \tilde{f} \vee \tilde{g}(G) \right] < \frac{1}{10s}$

(ii) $\Pr_{G \sim \mathcal{N}} \left[\tilde{f} \sqcup \tilde{g}(G) > \tilde{f} \vee \tilde{g}(G) \right] < \frac{1}{10s}$

(iii) $\Pr_{G \sim \mathcal{Y}} \left[\tilde{f} \sqcap \tilde{g}(G) < \tilde{f} \wedge \tilde{g}(G) \right] < \frac{1}{10s}$

(iv) $\Pr_{G \sim \mathcal{N}} \left[\tilde{f} \sqcap \tilde{g}(G) > \tilde{f} \wedge \tilde{g}(G) \right] < \frac{1}{10s}$

Assume (i) - (iv) for now. Then

$\Pr_{G \sim \mathcal{Y}}$ [$\tilde{C}$ makes mistake on G; i.e. answers 0 instead of 1] $\leq$

$\sum_{i \in [s]} \Pr$ [mistake on G in (approximation by $\tilde{f}_i$ of) gate f_i] $\leq$

$$s \cdot \frac{1}{10s} = \frac{1}{10}$$

and completely analogously for $G \sim \mathcal{N}$

So if we can construct $\sqcup$ and $\sqcap$ that yield (m, ℓ) - functions that satisfy (i) - (iv), then we are done

OR-APPROXIMATOR II

MCA VI

Given $\tilde{f} = \bigvee_{i=1}^{m_1} C_{S_i}$ $\quad \tilde{g} = \bigvee_{j=1}^{m_2} C_{S'_j}$

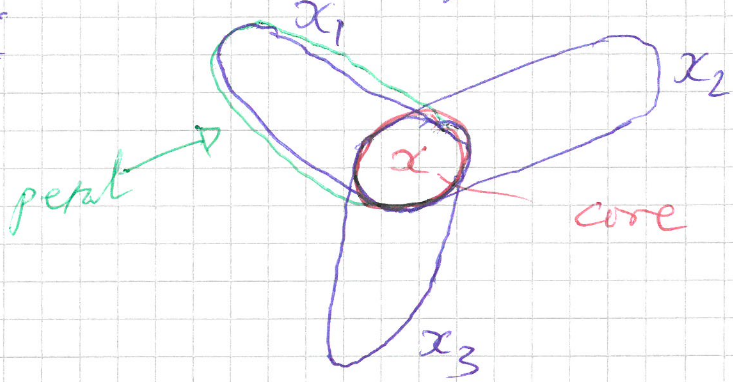
Let $\mathcal{Z} = \{ S_i \mid i \in [m_1] \} \cup \{ S'_j \mid j \in [m_2] \}$
$= \{ Z_1, \dots, Z_{m_1+m_2} \}$

First idea: set $\tilde{h} = \bigvee_{i=1}^{m_1+m_2} C_{Z_i}$

Problem: What if $m_1 + m_2 > m$?

Solution: Identify sets $Z_1, \dots, Z_p$ with common, unique pairwise intersection. Replace $\bigvee_{i=1}^{p} C_{Z_i}$ by C_Z and hope nothing much changes

DEF Sets $X_1, \dots, X_p$ form a SUNFLOWER if $\exists$ centre/core X s.t. for all $1 \le i < j \le p$ $X_i \cap X_j = X$



SUNFLOWER LEMMA [Erdős Rado '60]

$\mathcal{Z}$ collection of distinct sets Z_i
$\forall i \ |Z_i| \le \ell$ If $|\mathcal{Z}| > (p-1)^{\ell} \ell!$ then exist p sets $Z_1, \dots, Z_p \in \mathcal{Z}$ and a set Z such that for $1 \le i < j \le p$ $Z_i \cap Z_j = Z$

Defer proof. Note $Z = \emptyset$ is OK.

If $m_1 + m_2 = |\mathcal{Z}| > m$, apply Sunflower lemma (do "plucking") and replace p clique indicators by new clique indicator for centre Z.
Since $m = (p-1)^\ell \ell!$, can do this until get (m, ℓ)-function.
At most m pluckings.

AND-APPROXIMATOR $\sqcap$

Given $\tilde{f} = \bigvee_{i=1}^{m_1} C_{S_i}$ $\tilde{g} = \bigvee_{j=1}^{m_2} C_{T_j}$
Three steps

1) Consider $h' = \tilde{f} \wedge \tilde{g}$

$$= \bigvee_i \bigvee_j C_{S_i \cup T_j}$$

2) Omit any $S_i \cup T_j$ with $|S_i \cup T_j| > \ell$
3) Reduce remaining clique indicators to at most m by using Sunflower lemma

At most m^2 pluckings

This defines our approximators for gates in circuit. Clearly no errors at input gates. Need to prove (i)-(iv) for $\sqcup$ and $\sqcap$ operations

MCA VIII

(i) $\Pr_{G \sim Y}\left[\tilde{f} \sqcup \tilde{g}(G) < \tilde{f} \vee \tilde{g}(G)\right] < \frac{1}{10s}$

$\Pr\left[\; \downarrow = 0 \quad \text{but} \quad \downarrow = 1 \right]$

Sunflower lemma replaces larger clique indicators by smaller clique indicators

If $C_{Z_i}(G) = 1$ for petal Z_i, then clearly $C_Z(G) = 1$ for core/centre Z.

Hence no errors. No "false negatives" introduced

(ii) $\Pr_{G \sim N}\left[\tilde{f} \sqcup \tilde{g}(G) > \tilde{f} \vee \tilde{g}(G)\right] < \frac{1}{10s}$

$\Pr\left[\; \downarrow = 1 \text{ and } \downarrow = 0 \right]$

Replacing $Z_1, \ldots, Z_p$ by Z can introduce error if

$\forall i \quad C_{Z_i}(G) = 0 \qquad A_i$

but $\quad C_Z(G) = 1 \qquad B$

$G \sim N$ constructed from $c : [n] \to [k-1]$

Error if

A_i : c not one-to-one on Z_i

B_i : c one-to-one on Z

Want to show $\Pr\left[\bigwedge_i A_i \wedge B\right] < 2^{-p}$

$\leq 1/(10 m^2 s) \qquad (\dagger)$

(by choice of parameters)

And we make $\leq m$ pluckings, so if we can show $(\dagger)$, then we are done

MCA IX

$$\Pr\left[\bigcap_i A_i \cap B\right] =$$

$$= \Pr[B] \cdot \Pr\left[\bigcap_i A_i \mid B\right]$$

Conditioned on B all events A_i independent, because petals don't intersect outside of centre (and edges in disjoint subsets of vertices in graph are independent). So:

$$\Pr\left[\bigcap_i A_i \mid B\right] = \prod_i \Pr[A_i \mid B]$$

And conditioning on no collisions for c in centre Z only makes it less likely that c has collisions in Z_i.

Formally

$$\Pr[A_i] = \Pr[A_i \mid B] \cdot \Pr[B] + \Pr[A_i \mid \overline{B}] \cdot \Pr[\overline{B}]$$

$$= \Pr[A_i \mid B] \cdot \Pr[B] + 1 \cdot \Pr[\overline{B}]$$

$$\geq \Pr[A_i \mid B] \cdot \Pr[B]$$

$$\geq \Pr[A_i \mid B]$$

But $|Z_i| = \ell = \sqrt{k}/10$, meaning that c ~~very~~ likely to be one-to-one ~~from~~ Z_i to $[k-1]$ by the Birthday bound (see last lecture)

$$\Pr[A_i] \leq \frac{1}{2}$$

Summing up

$$\Pr[\wedge_i A_i \wedge B] = \Pr[B] \cdot \Pi_i \Pr[A_i \| B]$$
$$\leq \Pi_i \Pr[A_i] < 2^{-p}$$

which shows (†)

(iii) $\Pr_{G\sim y}[\tilde{f} \wedge \tilde{g}(G) < f \wedge \tilde{g}(G)] < \frac{1}{10s}$

$\Pr[\ \downarrow = 0$ but $\downarrow = 1\ ]$

$\tilde{f} \wedge \tilde{g}(G) = V_i V_j\ C_{S_i \cup T_j}$ so

first step introduces no errors

$\tilde{f} \wedge \tilde{g}(G) = 1$ if choose clique K s.t.
$S_i \cup T_j \subseteq K$ for some i,j

$C_{S_i \cup T_j}$ discarded if $|S_i \cup T_j| > \ell$ — introduces error in step 2

But this is quite a large clique indicator — unlikely to be 1 anyway

Proved last lecture:

$$|Z| = \ell \Rightarrow \Pr_{G\sim y}[C_Z(G) = 1] < n^{-\sqrt{k}/20} < \frac{1}{10sm^2}$$

And we ignore at most m^2 $S_i \cup T_j$'s,
so $\Pr[\text{error}] < \frac{1}{10s}$ by union bound.

Step 3 introduces no error (as in (i)).

(iv) $\Pr_{G \sim N}\left[\tilde{f} \sqcap \tilde{g}(G) > \tilde{f} \wedge \tilde{g}(G) \right] < \frac{1}{10s}$

$\Pr[\ \downarrow = 1 \text{ but } \downarrow = 0\ ]$

Step 1 just rewrites $\tilde{f} \wedge \tilde{g}$ as

$\bigvee_i \bigvee_j C_{S_i \cup T_j}$ — no error

In step 2 we throw away terms — can't make function go from 0 to 1

In step 3 we do plucking — can happen for $Z_1, \ldots, Z_p$ with centre Z that

$\forall i \ C_{Z_i}(G) = 0$ but $C_Z(G) = 1$

By analysis in (ii), probability that this happens is $< \frac{1}{10 m^2 s}$

At most m^2 pluckings — do union bound — done. Lemma 7 follows $\boxed{/\!/}$

It remains to prove the Sunflower Lemma.

Proof of Sunflower lemma

Have collection $\mathcal{Z}$ of $>(p-1)^{\ell}\,\ell!$ sets
of cardinality ℓ disjoint

Want to find sunflower of size $p>1$ (with p petals).

Induction over ℓ

Base case ($\ell = 1$): $|\mathcal{Z}| \geq p$; all $|Z_i| = 1$

Pick all sets; centre $Z = \emptyset$.

Induction step

Try again to find sunflower with empty centre.
maximal
Let $\mathcal{M} \subseteq \mathcal{Z}$ collection of pairwise disjoint sets ($Z_i \cap Z_j = \emptyset$ for $Z_i, Z_j \in \mathcal{M}$, $Z_i \neq Z_j$). If $|\mathcal{M}| = p$, then done.

Otherwise $\forall\, Z^* \in \mathcal{Z}\ \exists\, x \in Z^*$ s.t.
$x \in \bigcup_{Z_i \in \mathcal{M}} Z_i$ (by maximality of $\mathcal{M}$)

$$\left| \bigcup_{Z_i \in \mathcal{M}} Z_i \right| \leq (p-1)\,\ell \text{ , so}$$

some $x^* \in \bigcup_{Z_i \in \mathcal{M}} Z_i$ appears in fraction $\frac{1}{(p-1)\ell}$ of all sets in $\mathcal{Z}$, or in

$$> \frac{(p-1)^{\ell}\,\ell!}{(p-1)\,\ell} = (p-1)^{\ell-1}(\ell-1)! \text{ sets}$$

Fix $\mathcal{Z}^* = \{ Z \in \mathcal{Z}^* \mid x^* \in Z \}$ and look
at $\mathcal{Z}' = \{ Z \setminus \{x^*\} \mid Z \in \mathcal{Z}^* \}$

$\mathcal{Z}'$ contains $\geq (p-1)^{\ell-1}(\ell-1)!$ sets of size $\leq \ell-1$. Apply induction hypothesis to find sunflower in $\mathcal{Z}'$

$Z_1', \ldots, Z_p'$

these sets are not in $\mathcal{Z}$.

But set $Z_i = Z_i' \cup \{x\}$ for $i = 1, \ldots, p$

then get sunflower in $\mathcal{Z}$

The lemma follows by the induction principle $\boxed{/\!/}$

Frontiers in circuit complexity

THM [Williams '10] $NEXP \not\subseteq ACC^0$

uses that if $f \in ACC^0$, then f has depth-2 circuit with symmetric top gate (output depends on # input wires 1 only, not which wires) with AND-gates feeding in

SYM-gates quasipolynomial fan-in $2^{\log^k n}$

AND-gates polylogarithmic fan-in $\log^k n$

Plus lots & lots of other stuff

Other problems / approaches

- Prove lower bounds for circuits of polynomial size and logarithmic depth
  Or even $O(n)$ size.
- Branching programs (won't have time to talk about it now)
- Communication complexity
  Deep and fascinating connections
  Many great open problems
- Natural proofs barrier by Razborov & Rudich — argues why current techniques are unlikely to work [Could be great, but hard, paper to read up on and present.]
- Lots of work also on algebraic circuits — very active area
  (And here circuits in constant depth 4 can do anything that poly-size circuits can do!?)