Lecture 8 - Derandomization

Wednesday, December 7, 2022 17:03

WE LIKE RANDOMIZED ALGS, ... BUT WE WANT TO MAKE THEM DETERMINISTIC !!

① METHOD OF CONDITIONAL EXPECTATION.

Recall the randomized max cut algorithm that output a cut of size $\geq m/2$

(Just pick a random set S, we know that $E[S] = m/2$)

Pick up a vertex v

We have that (law of conditional expectation):

$$E[\partial(S)] = E[\partial(S) \mid v \in S] \cdot \Pr[v \in S] + E[\partial(S) \mid v \notin S] \cdot \Pr[v \notin S]$$

Can we compute $E[\partial(S) \mid v \in S] \cdot \Pr[v \in S]$?

↳ Indicators!

$$X_e = \begin{cases} 1 & \text{if } e \in \partial(S) \\ 0 & \text{otherwise} \end{cases}$$

For every edge we have that

$$E[X_e \mid v \in S] = \Pr[X_e \mid v \in S] = 1/2$$

(why?)

So $E[\partial(S) \mid v \in S] = E[\partial(S) \mid v \notin S] = m/2$.

So greedily add v to S.

Pick some $u \in V(G)$ s.t. $uv \in E(G)$ [lets just suppose u exists]

$$E[\partial(S) \mid v \in S] = E[\partial(S) \mid v \in S \wedge u \in S] \cdot \Pr[u \in S] + E[\partial(S) \mid v \in S \wedge u \notin S] \cdot \Pr[u \notin S]$$

Indicators: $\Pr[X_e = 1 \mid v \in S \wedge u \in S]$

$= \Pr[X_e = 1 \mid v \in S \wedge u \notin S]$

$= 1/2$ for all $e \neq uv$!!

$$\Pr[X_{uv} = 1 \mid v \in S \wedge u \in S] = 0$$

$$\Pr[X_{uv} = 1 \mid v \in S \wedge u \notin S] = 1$$ ← BETTER!

So $E[\partial(S) \mid v \in S \wedge u \notin S] > m/2$

So decide $u \notin S$.

More generally, let A, B be disjoint subsets of $V(G)$,

$E(A,B)$ be the edges from A to B

$E(A)$ be edges with both ends in A

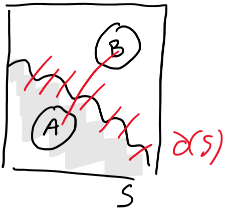
$E(B)$ be edges with both ends in B

THEN $XP(A,B) :=$

$$E[\partial(S) \mid A \subseteq S, B \cap S = \emptyset]$$

$$= \frac{m - |E(A,B)| - |E(A)| - |E(B)|}{2} + |E(A,B)|$$

← EASY TO COMPUTE



Let $v \notin A \cup B$. Conditional expectation yields:

$$
\begin{aligned}
XP(A,B) &= XP(A \cup \{v\}, B) \cdot \Pr[v \in S] \\
&\quad + XP(A, B \cup \{v\}) \cdot \Pr[v \notin S] \\
&= \tfrac{1}{2} XP(A \cup \{v\}, B) + \tfrac{1}{2} XP(A, B \cup \{v\})
\end{aligned}
$$

so $\max\left[XP(A \cup \{v\}, B),\ XP(A, B \cup \{v\})\right]$

$$\geq XP(A,B) \quad !$$

Note that we can compute $XP(A,B)$!

ALG:

```
$A = \emptyset$
$B = \emptyset$
For $i = 1$ to $n$
    if $XP(A \cup \{v_i\}, B) \geq XP(A, B \cup \{v_i\})$
        $A := A \cup \{v_i\}$
    else
        $B := B \cup \{v_i\}$
RETURN $A$
```

Alg always runs in poly time.
always outputs cut (A,B)

By induction: $XP(A,B) \geq m/2$

In the end $A \cup B = V(G)$ so

$$E(G) = E(A) \cup E(B) \cup E(A,B)$$

and so

$$XP(A,B) = E(A,B) = \partial(A)$$

so

$$\partial(A) \geq m/2 \quad !$$

Inspection of $XP(A,B)$ shows that

$$XP(A \cup \{v\}, B) \geq XP(A, B \cup \{v\})$$

if and only if

$$\underbrace{N(v) \cap B}_{\substack{\text{get cut if} \\ v \in A \text{ and not cut} \\ \text{if } v \in B}} \geq \underbrace{N(v) \cap A}_{\substack{\text{get cut if } v \in B \\ \text{and not cut if } v \in A}}$$

So AL6 can be rewritten like this:

```
A = B = ∅
For i = 1 to n
        if |N(v_i) ∩ B| ≥ |N(v_i) ∩ A|
                A = A ∪ {v_i}
        else
                B = B ∪ {v_i}
```

Return A

The fact that this greedy algorithm outputs a cut of size $\geq m/2$ has an easy direct proof.

So why go through analysis of randomized?

A "no creativity needed" way to design alg -

ⓐ Try random solution

ⓑ Analyze expectation

ⓒ Method of conditional expectation to derandomize.

Hm... we got a simple greedy for MAX CUT... what about SET COVER?

In: U, $\mathcal{F} = \{S_1 \ldots\ldots S_m\}$

Task: Find $\mathcal{F}' \subseteq \mathcal{F}$ s.t.

$$C(\mathcal{F}') = \bigcup_{S_i \in \mathcal{F}'} S_i = U$$

and $|\mathcal{F}'|$ is minimized

```
$\mathcal{F}' = \emptyset$
While $C(\mathcal{F}') \neq U$
    Let $S \in \mathcal{F}$ be such that $|C(\mathcal{F}' \cup \{S\})|$ is max
        $\mathcal{F}' = \mathcal{F}' \cup \{S\}$
Return $\mathcal{F}'$
```

* $\mathcal{F}'$ covers U.

★ $|\cup \mathcal{F}'| = ??$

in each round, at least one set $S_i \in \mathcal{F}$ covers a $1/OPT$ fraction of uncovered elements [WHY] ?

$\Downarrow$

If ALG goes on for t rounds then

$$|U \setminus C(F)| \le n \cdot \left(1 - 1/OPT\right)^t \le n \cdot e^{-t/OPT}$$

so for $t = OPT \cdot \ln n$ we get

$$|U \setminus C(F)| \le n \cdot e^{-t/OPT} \le n \cdot e^{-\ln n} \le 1$$

So $|\mathcal{F}'|$ of size $OPT \cdot \ln n + 1$ will cover U.

De-randomizing $(2e)^k poly(n)$ <u>k-path</u>. algorithm?

- → $2^k poly(n)$ deterministic alg to find multi-colored paths
- → random k-coloring multi-colors a path

with probability $\geq 1/e^k$

A perfect k-coloring family of a universe U (of size n) is a family $\mathcal{F}$ of colorings $c : U \to \{1 \dots k\}$ s.t $\forall$ set $S \subseteq U$ of size k

$\exists\, c \in \mathcal{F}$

s.t c multicolors S

Deterministic k-PATH ALGORITHM:

```
Construct perfect k-coloring family $\mathcal{F}$ for $V(G)$
For every col $\in \mathcal{F}$
        If $\exists$ k-path P multi-colored by col
                Return P
Return "NO"
```

Time: Time to construct $\mathcal{F}$ $+ |\mathcal{F}| \cdot 2^k \cdot \text{poly}(n)$

NAIVE CONSTRUCTION OF $\mathcal{F}$:

$\binom{n}{k}$ size & time [HOW?]

Existential Construction:

Thm: $\exists$ perfect k-coloring family of size

$$t = e^k \cdot 2k \cdot \log n$$

Pf: For every $i \leq t$ select $col_i : V(G) \to \{1 \dots k\}$ u.a.r.

For every set S,

$$\Pr[col_i \text{ multicolors } S] = \frac{k!}{k^k} \geq \frac{1}{e^k}$$

$$\Pr[\nexists i : col_i \text{ multicolors } S]$$

$$\leq \left(1 - \frac{1}{e^k}\right)^t$$

$$= \left(1 - \frac{1}{e^k}\right)^{e^k \cdot 2k \cdot \log n}$$

$$\leq e^{-2k \log n}$$

$$= 1/n^{2k}$$

By UNION BOUND over all $S \in \binom{V(G)}{k}$

$$\Pr[\exists S \subseteq V(G), |S| = k, \nexists i : col_i \text{ multicolors } S]$$

$$\leq 1/n^{2k} \cdot n^k \leq 1/n^k < 1$$

So a perfect k-coloring family of size t exists! ∎

(But how to construct it deterministically ??)

Next: A k-universal family of size $k^{O(k)} \cdot n^2$ (that we can construct in $k^{O(k)} \cdot n^2$ time)

Pairwise Independent Maps

Assume n is a prime (just add vertices until it is)

Suppose vertices are numbered $0, \ldots, n-1$

For $a, b \in \{0 \ldots n-1\}$ define

$$f_{a,b} : \{0 \ldots n-1\} \to \{0 \ldots n-1\} \quad \text{as}$$

$$f_{a,b}(x) = ax + b \pmod{n}$$

Claim: (i) $\forall x, y \quad \Pr_{a,b}\left[f_{a,b}(x) = y\right] = 1/n$

(ii) $\forall x_1, y_1, \overset{\neq x_1}{x_2}, y_2 : \quad \Pr\left[f_{a,b}(x_1) = y_1 \wedge f_{a,b}(x_2) = y_2\right] = 1/n^2$

Pf:

(ii) implies (i) [WHY?]

so only prove (ii)

the system

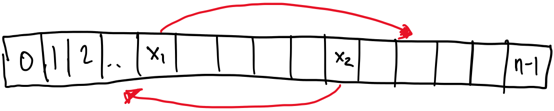
$$\begin{aligned} a x_1 + b &= y_1 \\ a x_2 + b &= y_2 \end{aligned}$$

solves to

$$\begin{aligned} a(x_1 - x_2) &= y_1 - y_2 \\ a &= (y_1 - y_2)(x_1 - x_2)^{-1} \leftarrow \text{OK SINCE } x_1 \neq x_2 \end{aligned}$$

$$b = y_1 - ax_1$$

↳ has unique solution.



Knowing the location of x_1 gives no information about location of x_2

Let $g : \{0 \ldots n-1\} \to \{0 \ldots k^2\}$ be a function s.t

$$\forall i, \ |g^{-1}(i)| \le \frac{n}{k^2}$$

(e.g. $g(i) = i \bmod k^2+1$)

Then $\forall\, x_1, x_2 \in \{0 \ldots n-1\}$ s.t $x_1 \neq x_2$ we have that:

$$\Pr_{a,b}\left[g(f_{a,b}(x_1)) = g(f_{a,b}(x_2))\right]$$

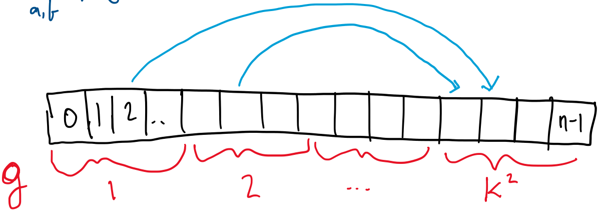
$$= \sum_{y=0}^{n-1} \Pr_{a,b}[f_{a,b}(x_1) = y] \cdot \Pr\left[f_{a,b}(x_2) \in g^{-1}(y) \mid f_{a,b}(x_1) = y\right]$$

$$\le \sum_{y=0}^{n-1} (1/n) \cdot |g^{-1}(y)| \cdot 1/n$$

$$\le n/k^2 \cdot 1/n = 1/k^2$$

SUMMARIZING

$$\Pr_{a,b}\left[g\left(f_{a,b}(x_1)\right) = g\left(f_{a,b}(x_2)\right)\right] \le 1/k^2$$



Lemma: $\forall$ set $S \subseteq \{1 \ldots n-1\}$ s.t $|S| = k$
$\exists\, a,b$ s.t $g(f_{a,b}())$ is
injective on S.

Pf: Pick random a,b.
We have that $\forall x_1, x_2$

$$\Pr_{a,b}\left[g\left(f_{a,b}(x_1)\right) = g\left(f_{a,b}(x_2)\right)\right] \le 1/k^2$$

And so by UNION BOUND over all $\binom{k}{2}$
choices of x_1, x_2 in S we have

$$\Pr_{a,b}\left[\exists x_1, x_2 : g\left(f_{a,b}(x_1)\right) = g\left(f_{a,b}(x_2)\right)\right] \le \binom{k}{2}/k^2$$

$$\le 1/2$$

So $\Pr_{a,b}\left[f_{a,b} \text{ is injective on } S\right] \ge 1/2$

So $\exists\, f_{a,b}$ which is injective on S ■

UNIVERSE REDUCTION

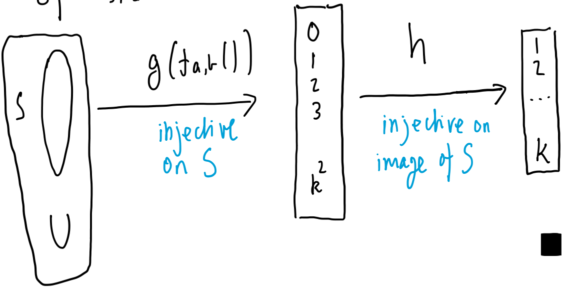
Lemma: Let $\mathcal{H}$ be a k-perfect coloring family of $\{0 \ldots k^2\}$. Then exists a perfect k-coloring family $\mathcal{F}$ of $V(G)$ of size $n^2 \cdot |\mathcal{H}|$. Given $\mathcal{H}$ we can construct $\mathcal{F}$ in time $|\mathcal{H}| \cdot \mathrm{poly}(n)$.

Proof $\mathcal{F}$ contains a function for every $f_{a,b}$ and every $h \in \mathcal{H}$.

It maps $i \in \{0 \ldots n-1\}$ to $h(g(f_{a,b}(i)))$

$|\mathcal{F}| \le |\mathcal{H}| \cdot n^2$ and we can construct it efficiently.

Let S be an arbitrary set in $\{0 \ldots n-1\}$ of size k



■

Corollary: $\exists$ perfect k-coloring family of size $k^{O(k)} n^2$

(constructible in same-ish time)

pf: Lemma + naive $\binom{n}{k}$ - construction of $\mathcal{H}$ with $n = k^2$!

Corollary to Corollary: $K^{O(k)}$ time algorithm for K-path

CAN WE DO BETTER?

SPLITTER LEMMA

$\forall n, k, t,$

a k/t-perfect coloring family $\mathcal{H}$ for universe of size $\le n$

$$\Downarrow$$

a k-perfect coloring family $\mathcal{F}$ for universe of size $\le n$

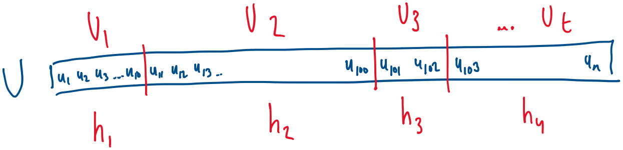
with $\mathcal{F} \le |H|^t \cdot n^t$

proof:

LET $U = u_1 \; u_2 \; u_3 \; \dots\dots \; u_n$

Each function $f \in \mathcal{F}$ is defined by:

- ★ A contiguous k-partition of U into $U_1 \dots U_k$
- ★ A function $h_i : U_i \to \{1 \dots k/t\}$ for each $i \le t$ (with $h_i \in \mathcal{H}$)



We define $f(u_i)$ as follows:

Let $j \le t$ be s.t $u_i \in U_j$:

SET $f(u_i) = k/t \cdot (j-1) + h_j(u_i)$

$$|\mathcal{F}| \le \underbrace{n^{k/t}}_{\text{choices for } U_1 \dots U_{k/t}} \cdot \underbrace{|\mathcal{H}|^t}_{\text{choices for } h_1 \dots h_t}$$

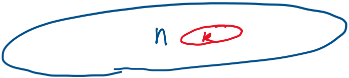
Consider a set $S \subseteq U$ of size k

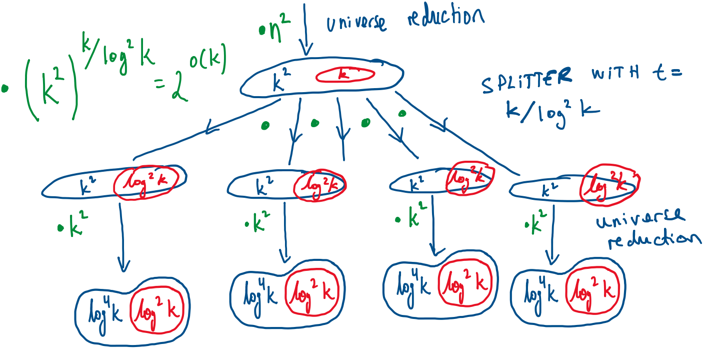
Let $U_1 \dots U_t$ be such that $\forall i \le t : |S \cap U_i| \le k/t$

Let $h_1 \dots h_t$ be functions in $\mathcal{H}$ s.t h_i is injective on $S \cap U_i$

Then $f_{[U_1, U_2 \dots U_t, h_1, h_2, \dots h_t]}$ is injective on S!

SCHEME FOR k-perfect coloring family for universe of size n, of size $e^{k+o(k)} \cdot n^2$





$$\bullet\ \left(k^2\right)^{k/\log^2 k} = 2^{O(k)}$$

Exists: $e^{\log^2 k}\cdot \log(\log^4 k)$

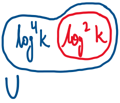
Will Find: $e^{\log^2 k}\cdot \log^5 k$

Total size:

$$\Bigg(\underbrace{e^{\log^2 k}\cdot \log^5 k}_{\text{BASE}}\cdot \underbrace{k^2}_{\text{U.R}}\Bigg)^{k/\log^2 k}\cdot \underset{\text{SPLITTER}}{2^{O(k)}}\cdot \underset{\text{U.R}}{n^2}$$

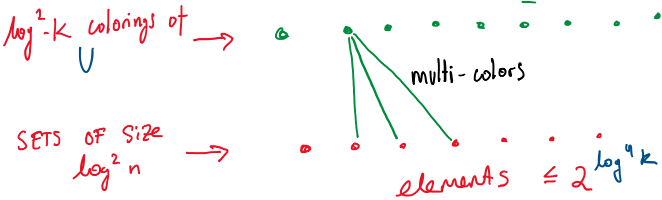
$$= e^{k+o(k)}\cdot n^2$$

Final task: finding a $\log^2 k$- perfect coloring family for universe of size $\log^4 k$



IDEA: REDUCE TO SET COVER!!

SETS $< k^{\log^4 k}$



$\log^2 k$-perfect coloring family ⟷ SET COVER

$\exists$ $\log^2$-perfect coloring family of size $e^{\log^2 k}\cdot\log(\log^4 k)$ ⟶ $\exists$ SET COVER OF SIZE $e^{\log^2 k}\cdot\log(\log^4 k)$

↓ FIND USING APPX ALG

SET COVER OF SIZE $e^{\log^2 k}\cdot\log(\log^4 k)\cdot\underbrace{\log^4 k}_{\text{APPX RATIO}}$

$\le e^{\log^2 k}\cdot\log^5 k$

⟵ $\log^2$-perfect coloring family of size $\le e^{\log^2 k}\cdot\log^5 k$

CONCLUSION:

CAN CONSTRUCT K-PERFECT COLORING FAMILY OF SIZE $e^{k+o(k)}\cdot n^2$

IN TIME $e^{k+o(k)}\, n^2$

$\Downarrow$

DETERMINISTIC k-PATH IN TIME

$(2e)^{k+o(k)}\cdot\text{poly}(n)$