6.5420 Lecture 2

The probabilistic method

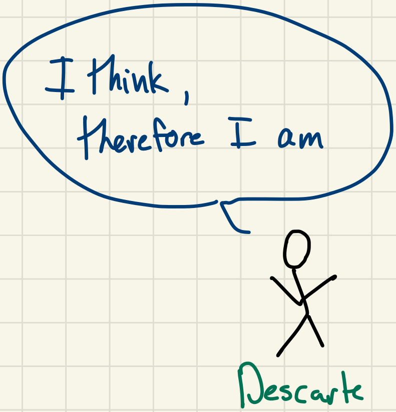
- hypergraph coloring
- dominating set

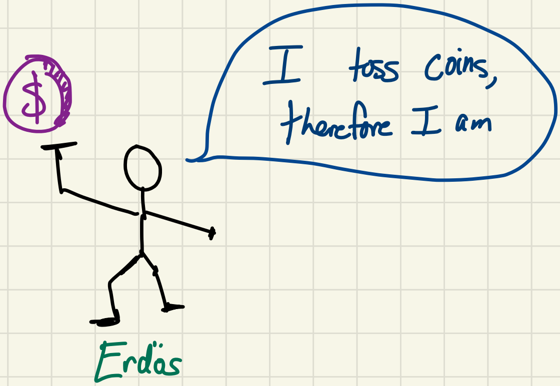
The Probabilistic Method

(+ excuse for a probability review)

Plan: Show object exists by showing that

$\underbrace{\text{probability it exists}}_{\text{can only be 0 or 1 since it either exists or it doesn't}}$ is $\underset{\text{so must be 1}}{>0}$





Example X is a set of elements

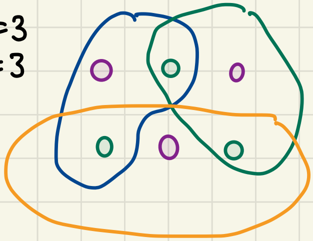
Input Given $\underbrace{S_1, S_2, \ldots, S_m}_{\text{each of size } \ell} \subseteq X$

Output Can we 2-color objects in X s.t. each set S_i not monochromatic? ← NP-hard problem!

Important special case: $m < 2^{\ell-1}$ not too many sets

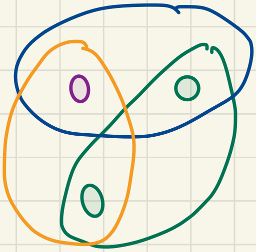
Thm if $m < 2^{\ell-1}$, $\exists$ proper 2-coloring

$m=3$
$\ell=3$



$3 < 2^{\ell-1}$

vs.



$m=3$
$\ell=2$

$3 > 2^{\ell-1}$

(note that no other coloring works either)

Proof

- randomly color elts of X red/blue
  (independently, prob ½)

want to "delete" all colorings that make any set monochromatic & show that there is a leftover good coloring

- $\forall i, \quad \Pr[S_i \text{ monochromatic}] = \underbrace{\frac{1}{2^\ell}}_{\text{all red}} + \underbrace{\frac{1}{2^\ell}}_{\text{all blue}} = \frac{1}{2^{\ell-1}}$

- $\Pr[\exists\, i \text{ s.t. } S_i \text{ monochromatic}]$

$$\le \sum_i \Pr[S_i \text{ monochromatic}]$$

union bnd
recall: $\Pr[A \cup B] \le \Pr[A] + \Pr[B]$

$$\le m \cdot \frac{1}{2^{\ell-1}}$$

$$< \frac{2^{\ell-1}}{2^{\ell-1}} = 1$$

assumption on m

$\therefore \Pr[\text{all } S_i \text{ 2-colored}] > 0$

$\Longrightarrow \exists$ setting of colors which gives legal 2-coloring ▨

i.e. there are lots of colorings, but if rule out monochromatic ones, still have some left over. We don't know how many.

Can we explicitly output a good 2-coloring?

- could try all 2-colorings ← "brute force"
  (exponential time)

- could make stronger assumption:

Old: Thm if $m < 2^{\ell-1}$, ∃ proper 2-coloring

↙ even smaller!

New Thm if $m < 2^{\ell-2}$, ∃ proper 2-coloring
& can find it quickly!

Since $\text{prob}[\exists i \text{ st. } S_i \text{ monochromatic}] \le \frac{m}{2^{\ell-1}} \le \frac{1}{2}$

so a random coloring of X is good
with prob $\ge 1/2$. ← p
(if not good, recolor until you find a good one)

how many times do you expect to need to recolor? 2 ← $1/p$

Recall

prob of "heads"

Given coin with bias p

What is expected number of tosses until see heads?

answer: $1/p$

Note tension between ability to

find good solution

+

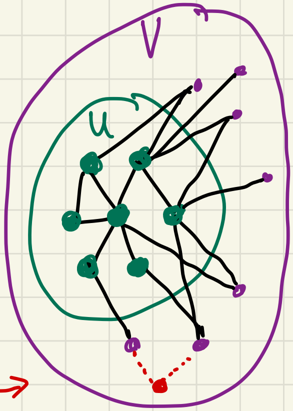
Strength of assumptions

(NP-hard, polytime, linear time, ...)

(in example: m vs. $2^{\ell-c}$)

Another example:

<u>def</u>. given graph $G=(V,E)$
$U \subseteq V$ is a "dominating set"
if every node $v \in V \setminus U$ has
at least one nbr in U.



no nodes that don't have at least one connection to U

<u>Note</u>: Finding min size dominating set is NP-hard.
(in fact one of the 1st known...)

Thm Given G with min degree Δ.
Then G has a dominating set of size $\leq \frac{4n \cdot \ln(4n)}{\Delta+1}$

Pf.

Construct $\hat{U}$: Place each node $v \in V$ into $\hat{U}$ indep. with prob

$$p = \frac{\ln 4n}{\Delta+1}$$

Is $\hat{U}$ a dominating set?

for $w \in V$, $\Pr[w$ has no nbr in $\hat{U}$ & is not in $\hat{U}]$

$$\leq (1-p)^{\Delta+1}$$

← uses independence in constructing $\hat{U}$

$\Pr[\exists\ w \in V$ s.t. w has no nbr in $\hat{U}$ & w not in $\hat{U}]$

$$\leq n \cdot (1-p)^{\Delta+1} \qquad \text{union bnd}$$

$$\leq n\left(1 - \frac{\ln 4n}{\Delta+1}\right)^{\frac{\Delta+1}{\ln 4n} \cdot \ln 4n} \approx n \cdot \underbrace{e^{-\ln 4n}}_{\frac{1}{4n}} = \frac{1}{4}$$

Useful:
$\lim_{x \to \infty} \left(1 - \frac{1}{x}\right)^x \to \frac{1}{e}$

So prob $[\hat{U}$ is not a dominating set$] \leq 1/4$

How big is $\hat{u}$?

$$E[\,|\hat{u}|\,] = n \cdot p \leftarrow$$

why?

let $\delta_w = \begin{cases} 1 & \text{if } w \text{ placed in } \hat{u} \\ 0 & \text{o.w.} \end{cases}$ indicator variables

$E[\delta_w] = p$

$|\hat{u}| = \sum_{w \in V} \delta_w$

linearity of expectations

$E[\,|\hat{u}|\,] = E[\sum \delta_w] = \sum E[\delta_w]$

$= \sum p = n \cdot p$

$$\Pr[\,|\hat{u}| > 4 \cdot np\,] < \frac{1}{4}$$

recall: Markov's ≠

$\Pr[X > c \cdot E[X]] \leq \frac{1}{c}$

So $\Pr[\hat{u}$ is dominating set of size $\leq \frac{4n \ln 4n}{\Delta + 1}]$

$\geq 1 - \frac{1}{4} - \frac{1}{4}$

$\geq \frac{1}{2} > 0$

2 badevents:

- not D.S. prob $\leq 1/4$
- too big prob $\leq 1/4$

so it exists! ∎

A third example: Sum-free subsets

A is subset of positive integers (>0)

Def. A is "sum-free" if $\nexists\ a_1, a_2, a_3 \in A$ s.t. $a_1 + a_2 = a_3$

Thm (Erdős '65)

$\forall\ B = \{b_1 \cdots b_n\}$ $\exists$ sum-free $A \subseteq B$ st. $|A| > \frac{n}{3}$

note: not true if $|A|$ needs to be greater than $\frac{12n}{29}$

Example

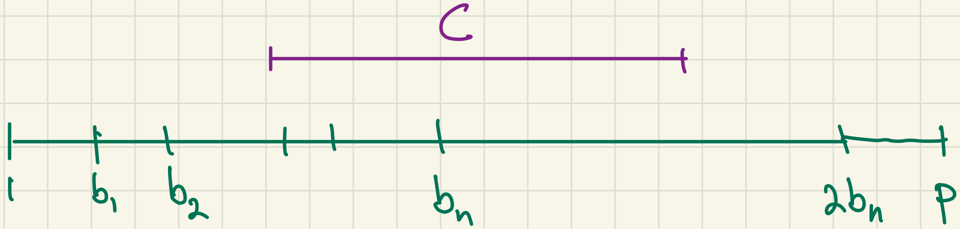
$B = \{1 \ldots n\}$

can take $A = \{\lceil \frac{n}{2} \rceil, \ldots, n\}$

Proof wlog b_n is max

pick prime $p > 2b_n$ s.t. $p \equiv 2 \pmod{3}$

i.e. $p = 3k+2$ for some int k



Let $C = \{k+1, \ldots, 2k+1\}$ "middle third of $[p]$"

$\mathbb{Z}_p = \{0, \ldots, p-1\}$

$\mathbb{Z}_p^* = \{1, \ldots, p-1\}$ group has unique multiplicative inverses mod p

(need p to be prime for this)

e.g. $\mathbb{Z}_3^* = \{1,2\}$

$1 \cdot 1 \equiv 1 \mod 3$

$2 \cdot 2 \equiv 1 \mod 3$

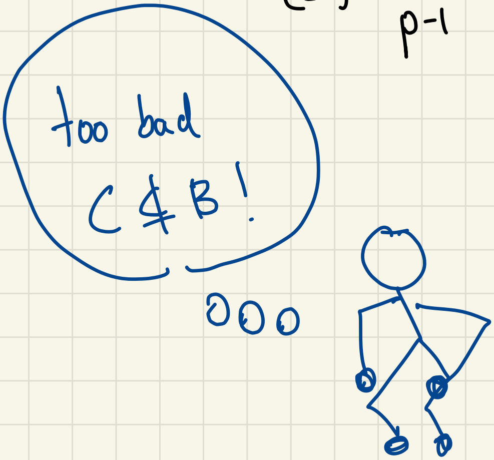
Note: (1) $C \subseteq \mathbb{Z}_p$

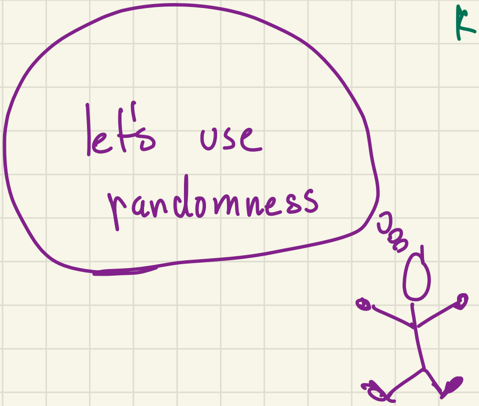
(2) C sum-free, even in $\mathbb{Z}_p$

↖ why?

any 2 elements sum to $\geq 2k+2$ & at most $4k+2 \equiv k \pmod{3k+2}$

(3) $\frac{|C|}{p-1} = \frac{k+1}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}$



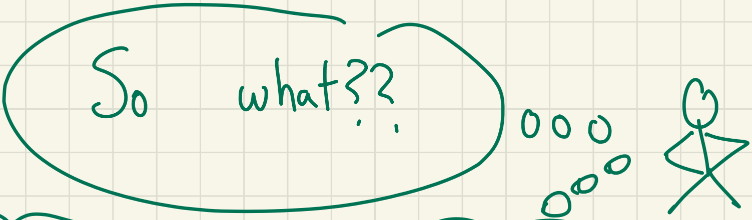


What if $B \cap C$ is big? e.g. if $|B \cap C| = \frac{|B|}{3}$

we are done!

Cool idea: Sum-freeness extends to linear fcfns of elements

$$b_1 + b_2 = b_3$$
iff
$$a \cdot b_1 + a \cdot b_2 = a \cdot b_3$$



So what??

we need sum-freeness of B, not lin fcfns of B!

we will use it "backwards"!!

above we note: $B \cap C$ big $\Rightarrow$ we are done

cool idea $\Rightarrow$
if $\exists x$ st. $\{x \cdot b_i \mid b_i \in B\} \cap C$ big $\Rightarrow$ we are done

but is there such an x?

Constructing A:

pick $x \in_R \overset{\mathbb{Z}_p^* \downarrow}{\{1..p-1\}}$

then use x to define (random) linear map $f_x(a) = x \cdot a \bmod p$

let $A_x \leftarrow \{ b_i \text{ st. } \underbrace{x \cdot b_i \bmod p}_{f_x(b_i)} \in C \}$

so A_x are elts of B in preimage of C under f_x

"x maps these guys to middle third"

Claim 1 A_x is sum-free

Pf. if not, then

let $b_i, b_j, b_k \in A_x$ st. $b_i + b_j = b_k$

then $\underbrace{x \cdot b_i} + \underbrace{x \cdot b_j} = \underbrace{x \cdot b_k}$ (mod p)

all in C by construction

also not mod p

$\Longrightarrow$ C not sumfree in $\mathbb{Z}_p$

$\rightarrow\leftarrow$ ∎

Claim 2 $\exists\, x$ st. $|A_x| > \frac{n}{3}$

Pf

Fact $\forall\, y \in \mathbb{Z}_p^*$ + $\forall\, i$
exactly one $x \in \mathbb{Z}_p^*$ satisfies

$$y \equiv x \cdot b_i \pmod p$$

for each b_i + for each y, exactly one x maps b_i to y

follows from unique inverse property when p is prime

Fact $\Longrightarrow$

$$\forall\, y \in \mathbb{Z}_p^*,\ \forall\, i \quad \Pr_x[\, b_i \text{ mapped to } y \text{ via } x\,] = \frac{1}{p-1}$$

$\forall\, i$, fact $\Rightarrow$ $|C|$ choices of x s.t.

$$x \cdot b_i \pmod p \in C$$

define $\sigma_i^{(x)} \leftarrow \begin{cases} 1 & \text{if } x \cdot b_i \bmod p \in C \\ 0 & \text{o.w.} \end{cases}$ $\leftarrow$ b_i maps to C under X

$$E_x[\sigma_i^{(x)}] = \Pr_x[\sigma_i^{(x)} = 1] = \frac{|C|}{p-1} > \frac{1}{3}$$

$$E_x[|A_x|] = E_x\Big[\sum_i \underbrace{\delta_i^{(x)}}_{\text{\# of b's that map to C under x}}\Big] = \sum_i E_x[\delta_i^{(x)}]$$

$$> \frac{n}{3}$$

$\Rightarrow$ at least one x s.t. $|A_x| > \frac{n}{3}$

■