Welcome to

6.842

Randomness & Computation

Lecturer:

Ronitt Rubinfeld

TA: Shankha Biswas

Today:

Course Overview

The probabilistic method

- hypergraph coloring
- dominating set

What is this course about?

- How can randomness help?
  - algorithm design
    simpler, faster, new problems
  - show existence of combinatorial objects
    good solutions, codes, nice graphs
  - easy to verify proofs
    interactive proofs, PCPs
  - distributed algorithms
  - learning algorithms
  - testing algorithms

- Do we require randomness?
  - can we use less
  - can we do without it?
  - in what settings do we need it?
- Settings in which randomness is inherent:
  - uniform generation - approximate counting
  - learning theory
  - testing
- relation to complexity theory
  - hardness vs. randomness
  - hard core sets
  - ...

- Tools:

  Fourier representation

  random walks / Markov chains

  algebraic techniques

  probabilistic proofs

  Lovasz Local Lemma

  graph expansion, extractors

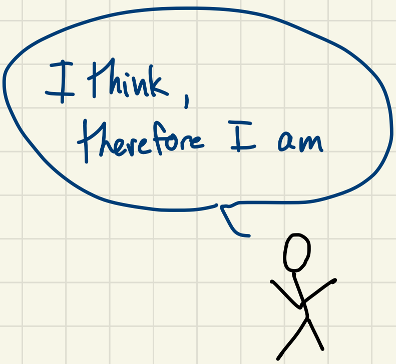
  Szemeredi Regularity Lemma

The Probabilistic Method

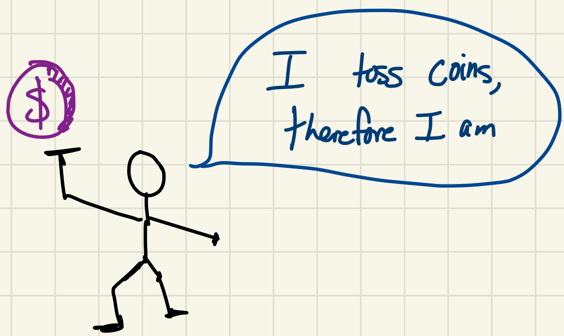
(+ excuse for a probability review)

Plan: Show object exists by showing that

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



Descarte



Erdös

Example X is a set of elements

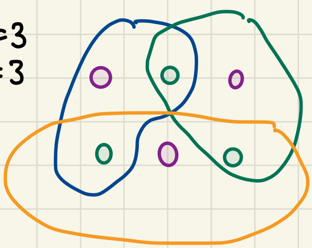
Input Given $\underbrace{S_1, S_2, \ldots, S_m}_{\text{each of size } \ell} \subseteq X$

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

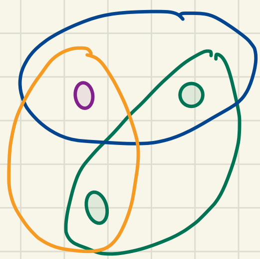
Thm if $m < 2^{\ell-1}$, $\exists$ proper 2-coloring

$m=3$
$\ell=3$



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

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

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

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

$$\leq \sum_i \Pr[S_i \text{ monochromatic}]$$ union bnd

recall: $\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$

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

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

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

$\Rightarrow \exists$ 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?

"brute force"

- could try all 2-colorings (exponential time)

- could make stronger assumption:

Old: $\boxed{\underline{\text{Thm}} \text{ if } m < 2^{\ell-1}, \exists \text{ proper 2-coloring}}$

even smaller!

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

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

so a random coloring of X is good with prob $\geq 1/2$. $\leftarrow p$

(if not good, recolor until you find a good one)

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

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

+

Strength of assumptions

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

Another example:

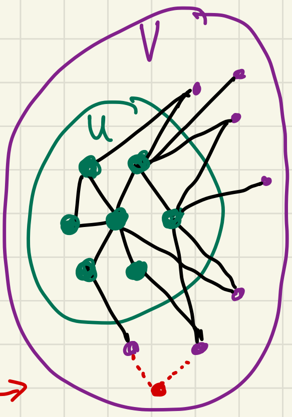
def. given graph $G = (V, E)$

$U \subseteq V$ 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



Note: 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 $\le \frac{4n \cdot \ln(4n)}{\Delta+1}$

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 $\hat{U}]$

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

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

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

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

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

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] $\le 1/4$

How big is $\hat{U}$?

$$E[|\hat{U}|] = n \cdot p$$

why?

indicator variables

let $\sigma_w = \begin{cases} 1 & \text{if } w \text{ placed in } \hat{U} \\ 0 & \text{o.w.} \end{cases}$

$E[\sigma_w] = p$

$|\hat{U}| = \sum_{w \in V} \sigma_w$

linearity of expectations

$E[|\hat{U}|] = E[\sum \sigma_w] = \sum E[\sigma_w]$

$= \sum p = n \cdot p$

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

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 bad events:

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

so it exists! $\blacksquare$

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 $A \subseteq B$ s.t. $|A| > \frac{n}{3}$

note: not true if $|A|$ is only greater than $\frac{12n}{29}$

Example

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

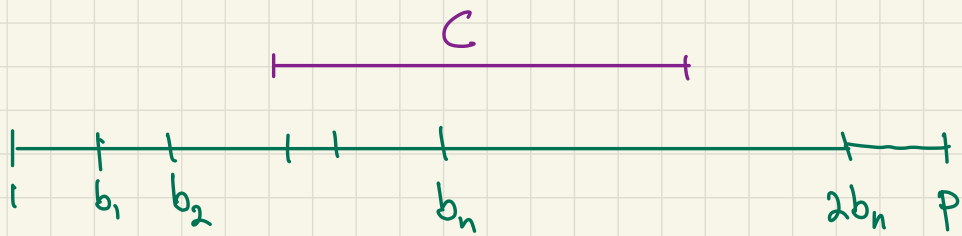
can take $A = \{\lceil \frac{n}{2} \rceil, \ldots, n\}$

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"

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

$\mathbb{Z}_p^* = \{1, \ldots, p-1\}$ group has multiplicative inverses mod p (need p to be prime for this)

unique

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

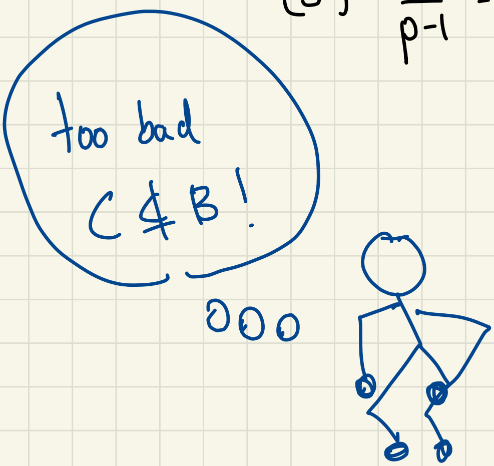
$$1 \cdot 1 \equiv 1 \mod 3$$
$$2 \cdot 2 \equiv 1 \mod 3$$

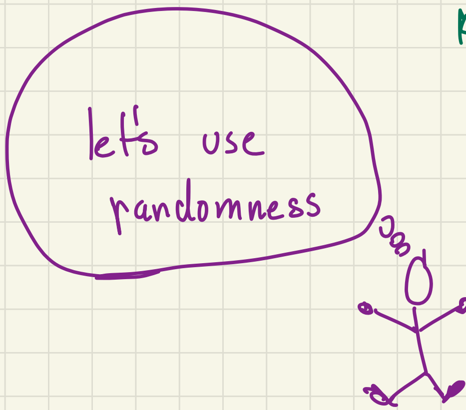
Note: (1) $C \subseteq \mathbb{Z}_p$

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





What if $B \cap C$ is big?

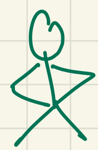
we are done!

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

if $x_1 + x_2 = x_3$

then $a \cdot x_1 + a \cdot x_2 = a \cdot x_3$

So what??



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

we will use it "backwards"!!

Constructing A:

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

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

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

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 $b_i, b_j, b_k \in A_x$ s.t. $b_i + b_j = b_k$

then $x \cdot b_i + x \cdot b_j = x \cdot b_k$ (mod p) ← also not mod p

all in C by construction

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

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

Pf

Fact $\forall\, y \in \mathbb{Z}_p^* \;\&\; \forall\, i$
exactly **one** $x \in \mathbb{Z}_p^*$ satisfies
$$y \equiv x \cdot b_i \pmod p$$

follows from unique inverse property when p is prime

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

$\forall\, i$, fact $\Rightarrow$ $|C|$ choices of x s.t.
$$x \cdot b_i \pmod p \in C$$

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

b_i maps to C under X

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

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

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

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

$\blacksquare$