The Prob Method, Turan's Theorem, and Finding Max in Parallel

A Nice Problem

Given n try to find a set $A \subseteq \{1, \ldots, n\}$ such that ALL of the differences of elements of A are DISTINCT.

Try to make A as big as possible.

Students Break into small groups and try to do this.

An Apporach

Let a be a number to be determined.

Pick a RANDOM set of $\{1, \ldots, n\}$ of size a.

What is the probability that all of the diffs in A are distinct?

We hope the prob is strictly greater than 0.

KEY: If the prob is strictly greater than 0 then there must be SOME set of a elements where all of the diffs are distinct.

Determining the Prob

If you pick a RANDOM $A\{1,\ldots,n\}$ of size a what is the probability that all of the diffs in A are distinct?

Determining the Prob

If you pick a RANDOM $A\{1, \ldots, n\}$ of size a what is the probability that all of the diffs in A are distinct?

WRONG QUESTION!

Determining the Prob

If you pick a RANDOM $A\{1, \ldots, n\}$ of size a what is the probability that all of the diffs in A are distinct?

WRONG QUESTION!

If you pick a RANDOM $A \subseteq \{1, \ldots, n\}$ of size a what is the probability that all of the diffs in A are NOT distinct?

We hope that it is NOT 1.

Determining the Prob

If you pick a RANDOM $A \subseteq \{1, \ldots, n\}$ of size a what is the probability that all of the diffs in A are NOT distinct?

Determining the Prob

If you pick a RANDOM $A \subseteq \{1, \ldots, n\}$ of size a what is the probability that all of the diffs in A are NOT distinct?
WRONG QUESTION!

Determining the Prob

If you pick a RANDOM $A \subseteq \{1, \ldots, n\}$ of size a what is the probability that all of the diffs in A are NOT distinct?
WRONG QUESTION!

We only need to show that the prob is LESS THAN 1.

Determining the Prob I

If a RAND $A \subseteq \{1, \ldots, n\}$, size a, want bound on prob all of the diffs in A are NOT distinct. Numb of ways to choose a elements out of $\{1, \ldots, n\}$ is $\binom{n}{a}$.
Two ways to create a set with a diff repeated:
Way One:

- Pick $x < y$. There are $\binom{n}{2} \le n^2$ ways to do that.
- Pick diff d such that $x + d \ne y$, $x + d \le n$, $y + d \le n$. Can do $\le n$ ways. Put $x, y, x + d, y + d$ into A.
- Pick $a - 4$ more elements out of the $n - 4$ left.

Number of ways to do this is $\le n^3 \times \binom{n-4}{a-4}$.
Way Two: Pick $x < y$. Let $d = y - x$ (so we do NOT pick d). Put $x, y = x + d, y + d$ into A. Pick $a - 3$ more elements out of the $n - 3$ left.
Number of ways to do this is $\le n^2 \times \binom{n-3}{a-3}$.

Determining the Prob II

If you pick a RANDOM $A \subseteq \{1, \ldots, n\}$ of size a then a bound on the probability that all of the diffs in A are NOT distinct is

$$\frac{n^3 \times \binom{n-4}{a-4} + n^2 \times \binom{n-3}{a-3}}{\binom{n}{a}} = \frac{n^3 \times \binom{n-4}{a-4}}{\binom{n}{a}} + \frac{n^2 \times \binom{n-3}{a-3}}{\binom{n}{a}}$$

$$= \frac{n^3 a(a-1)(a-2)(a-3)}{n(n-1)(n-2)(n-3)} + \frac{n^2 a(a-1)(a-2)}{n(n-1)(n-2)}$$

(We use $n \geq 5$.)

$$\leq \frac{n^3 a^4}{(n/2)^4} + \frac{n^2 a^3}{(n/2)^3} \leq \frac{16a^4}{n} + \frac{8a^3}{n^2} \leq \frac{32a^4}{n}.$$

ANSWER

RECAP: If pick a RANDOM $A \subseteq \{1, \ldots, m\}$ then the prob that there IS a repeated difference is $\le \frac{32a^4}{n}$.
So WANT

$$\frac{32a^4}{n} < 1$$

Take

$$a = \frac{n^{1/4}}{33}.$$

UPSHOT: For all $n \ge 5$ there exists a set of size $\frac{n^{1/4}}{33}$ of $\{1, \ldots, n\}$ where all of the differences are distinct.

GENERAL UPSHOT

We proved an object existed by showing that the Prob that it exists is NONZERO!
We use this technique in a more sophisticated way.

Graphs and Ind Sets

Notation:

1. A Graph is (V, E) where V is the set of vertices and E is a set of pairs of vertices. Easy to draw!
2. An Ind Set in a graph (V, E) is a set $V' \subseteq V$ such that there are NO edges between elements of V'.
3. If (V, E) is a graph and $v \in V$ then the degree of v, denoted d_v, is the number of edges coming out of it.

DO EXAMPLES ON BOARD.

Turan's Theorem

Theorem

If $G = (V, E)$ is a graph, $|V| = n$, and $|E| = e$, then G has an ind set of size at least

$$\frac{n}{\frac{2e}{n} + 1}.$$

We proof this using Probability, but first need a lemma.

Lemma

Lemma

If $G = (V, E)$ *is a graph. Then*

$$\sum_{v \in V} deg(v) = 2e.$$

Proof in class using counting.

Proof of Turan's Theorem

Theorem

If $G = (V, E)$ is a graph, $|V| = n$, and $|E| = e$, then G has an ind set of size

$$\geq \frac{n}{\frac{2e}{n} + 1}.$$

Proof: Take the graph and RANDOMLY permute the vertices. (DO EXAMPLE ON BOARD.) The set of vertices that have NO edges coming out on the right form an Ind Set. Call this set I.

How Big is I?

How big is I

How Big is I?

How big is I
WRONG QUESTION!

How Big is I?

How big is I
WRONG QUESTION!
What is the EXPECTED VALUE of the size of I.
(NOTE- we permuted the vertices RANDOMLY)

What is Prob $v \in I$

Let $v \in V$. What is prob that $v \in I$ DRAW PICTURE OF v. v has degree d_v. How many ways can v and its vertices be laid out: $(d_v + 1)!$. In how many of them is v on the right? $d_v!$.

$$\Pr(v \in I) = \frac{d_v!}{(d_v + 1)!} = \frac{1}{d_v + 1}.$$

Hence

$$E(|I|) = \sum_{v \in V} \frac{1}{d_v + 1}.$$

How Big is this Sum?

Need to find lower bound on

$$\sum_{v \in V} \frac{1}{d_v + 1}.$$

Rephrase

NEW PROBLEM:
Minimize

$$\sum_{v \in V} \frac{1}{x_v + 1}$$

relative to the constraint:

$$\sum_{v \in V} x_v = 2e.$$

KNOWN: This sum is minimized when all of the x are $\frac{2e}{|V|} = \frac{2e}{n}$. So the min the sum can be is

$$\sum_{v \in V} \frac{1}{\frac{2e}{n} + 1} = \frac{n}{\frac{2e}{n} + 1}.$$

DONE!

$$\sum_{v \in V} \frac{1}{x_v + 1} \geq \sum_{v \in V} \frac{1}{\frac{2e}{n} + 1} = \frac{n}{\frac{2e}{n} + 1}.$$

Finding Max in Parallel!

PROBLEM: We want to find MAX really really fast in parallel. (Connection to Turan's theorem will be later.)

1. Want to find MAX of n elements and only have ONE round. How many comparisons do we need to do in that round?

Finding Max in Parallel!

PROBLEM: We want to find MAX really really fast in parallel.
(Connection to Turan's theorem will be later.)

1. Want to find MAX of n elements and only have ONE round. How many comparisons do we need to do in that round?

$$\binom{n}{2}$$

Finding Max in Parallel!

PROBLEM: We want to find MAX really really fast in parallel. (Connection to Turan's theorem will be later.)

1. Want to find MAX of n elements and only have ONE round. How many comparisons do we need to do in that round?

$$\binom{n}{2}$$

2. What if you have two rounds? Can you do this problem with subquadratic number of comparisons? HAVE STUDENTS DO IN GROUPS.

Finding Max in 2 Rounds in Parallel!

a and b to be picked later such that $a + b = 1$.

1. First Round: Divide the n elements into n^a groups of n^b each. Find max of each group. Takes $n^a \binom{n^b}{2} \sim n^{a+2b}$ comparisons.
2. Second Round: Find max of the winners. $\binom{n^a}{2} \sim n^{2a}$ comparisons.

Want to MINIMIZE $\max\{n^{a+2b}, n^{2a}\}$.
Technique: Make them equal $a + 2b = 2a$ and $a + b = 1$ imply $a = 2/3$, $b = 1/3$. Each round uses n^{2a} which is also n^{a+2b} which is also $n^{4/3}$.
UPSHOT: Can find max of n elements in 2 rounds using $\sim n^{4/3}$ comparisons-per-round.

Can we do Better?

Can find max of n elements in 2 rounds using $\sim n^{4/3}$ comparisons-per-round.

Finding Max in 3 Rounds in Parallel!

a and b to be picked later such that $a + b = 1$.

1. First Two Rounds: Divide the n elements into n^a groups of n^b each. Find max of each group using the two round algorithm. Takes

   Numb of groups $\times$ Numb of comps needed for each group

   which is
   $$n^a \times (n^b)^{4/3} = n^{a+(4b/3)}$$

2. ThirdRound: Find max of the winners. n^{2a} comparisons.

Want to MINIMIZE $\max\{n^{a+\frac{4b}{3}}, n^{2a}\}$.
Technique: Make them equal $a + \frac{4b}{3} = 2a$ and $a + b = 1$ imply $a = 4/7$, $b = 3/7$. Each round uses n^{2a} which is also $n^{a+4b/3}$ which is also $n^{8/7}$.
UPSHOT: Can find max of n elements in 3 rounds using $\sim n^{8/7}$ comparisons-per-round.

Lower Bounds

1. We can find MAX in 1 round with $O(n^2)$ comps-per-round. Can we do better?
2. We can find MAX in 2 round with $O(n^{4/3})$ comps-per-round. Can we do better?
3. We can find MAX in 3 round with $O(n^{8/7})$ comps-per-round. Can we do better?

One Round Lower Bounds

Assume that we can find the MAX of n elements in 2 rounds with n^x comparisons-per-round. Assume, by way of contradiction, that $x < 2$.
Adversary Argument: We run the algorithm and WE answer the comparisons in a consisten way that will show that $x \geq 2$.
KEY: this is NOT an algorithm. This is OUR giving the algorithm answers to make it take a long time.

1. Algorithm says what its comparisons are for the first round. KEY- This is a Graph on n vertices and n^x edges.
2. Since $x < 2$ and there are $\binom{n}{2}$ pairs of elements there are two elements that were NOT compared to each other. Call them x and y.
3. WE answer the comparisons so that x and y beat all of their oponents. We set the other comparisons also (arbitrarily). Since x and y never played each other this is possible.

After the one round, DO NOT KNOW what the max is since it could be x or y. This contradicts this being an algorithm that can

Two Round Lower Bounds

Can we do Better than $n^{4/3}$ for 2 Rounds?- VOTE!

Two Round Lower Bounds

Can we do Better than $n^{4/3}$ for 2 Rounds?- VOTE! NO.
Assume, by way of contradiction, that we can find the MAX of n elements in 2 rounds with n^x comparisons-per-round where $x < \frac{4}{3}$.
Adversary Argument: We run the algorithm and WE answer the comparisons in a consisten way that will show that $x \ge \frac{4}{3}$.
KEY: this is NOT an algorithm. This is OUR giving the algorithm answers to make it take a long time.

1. Algorithm says what its comparisons are for the FIRST ROUND. KEY- This is a Graph on n vertices and n^x edges.
2. We find the Independent set of size $\frac{n}{\frac{2n^x}{n}+1} \sim\ge n^{2-x}$
3. WE answer the comparisons so that all elements of that ind set WIN. So now there are n^{2-x} candidates for MAX.
4. In Round 2 they have to find MAX of n^{2-x} elements. So will take at least n^{4-2x} comparisons.

If first round uses n^x comps then second round uses n^{4-2x} comps.
KEY: one can show that if $x < 4/3$ then $4 - 2x > 4/3$ (Next slide)

Algebra

$$\begin{aligned} x &< \tfrac{4}{3} \\ -2x &> -\tfrac{8}{3} \\ 4-2x &> 4-\tfrac{8}{3} = \tfrac{4}{3} \end{aligned}$$

UPSHOT: If first round uses $< n^{4/3}$ comps then second round has to use $\geq n^{4/3}$ comps.

Three Round Lower Bounds

Can we do Better than $n^{8/7}$ for 3 Rounds?- VOTE!

Three Round Lower Bounds

Can we do Better than $n^{8/7}$ for 3 Rounds?- VOTE! NO.
Assume, by way of contradiction, that we can find the MAX of n elements in 3 rounds with n^x comparisons-per-round where $x < \frac{8}{7}$.
Adversary Argument: We run the algorithm and WE answer the comparisons in a consisten way that will show that $x \ge \frac{8}{7}$.
KEY: this is NOT an algorithm. This is OUR giving the algorithm answers to make it take a long time.

1. Algorithm says what its comparisons are for the FIRST ROUND. KEY- This is a Graph on n vertices and n^x edges.
2. We find the Independent set of size $\frac{n}{\frac{2n^x}{n}+1} \sim\ge n^{2-x}$
3. WE answer the comparisons so that all elements of that ind set WIN. So now there are n^{2-x} candidates for MAX.
4. In Rounds 2 and 3 they have to find MAX of n^{2-x} comparisons. We already know this REQUIRES $n^{4(2-x)/3}$ by the LOWER BOUND on 2-round Max.

If first round uses n^x comps then second and third round uses