The Probabilistic Method

Techniques

- Union bound
- Argument from expectation
- **Alterations**
- The second moment method
- The (Lovasz) Local Lemma

And much more

- Alon and Spencer, "The Probabilistic Method"
- Bolobas, "Random Graphs"

Alteration Technique: Main Idea

- A randomly chosen object may not satisfy the property we want
- So, after choosing it we modify the object a little
- In non-elementary situations, the modification itself may be probabilistic
- Or, there might be more than one modification step

Example 1: Independent Set

- $\alpha(G)$ denotes the maximum size of an independent set in G
- Say G has n vertices and m edges
- Intuition: $\alpha(G)$ is proportional to n and inversely proportional to m
- Line of thought: on average a randomly chosen independent set has size μ (proportional to n and inversely proportional to m)
- Problem: random subset of vertices may not be an independent set!!!

A Randomized Algorithm based on Alteration Technique

- Choose a random subset X of vertices where Prob$[v \in X] = p$ (to be determined)
- Remove one end point from each edge in X
- Let Y be the set of edges in X
- Left with at least $|X| - |Y|$ vertices which are independent

$$\mathsf{E}[|X| - |Y|] = np - mp^2 = -m\left(p - \frac{n}{2m}\right)^2 + \frac{n^2}{4m}$$

Thus, choose $p = n/2m$; we get

Theorem

For any graph with n vertices and m edges, there must be an independent set of size at least $n^2/(4m)$.

Example 2: Dominating Set

- Given $G = (V, E)$, $S \subset V$ is a dominating set iff every vertex either is in S or has a neighbor in S
- Intuition: graphs with high vertex degrees should have small dominating set
- Line of thought: a randomly chosen dominating set has mean size μ

A Randomized Algorithm based on Alteration Technique

- Include a vertex in X with probability p
- Let Y = set of vertices in $V - X$ with no neighbor in X
- Output $X \cup Y$

$$\text{Prob}[u \notin X \text{ and no neighbor in } X] = (1-p)^{\deg(u)+1} \leq (1-p)^{\delta+1}$$

where $\deg(u)$ is the degree of u and δ is the minimum degree.

$$\mathsf{E}[|X| + |Y|] \leq n\left(p + (1-p)^{\delta+1}\right) \leq n\left(p + e^{-p(\delta+1)}\right)$$

To minimize the RHS, choose $p = \frac{\ln(\delta+1)}{\delta+1}$

Theorem

There exists a dominating set of size at most $n\frac{1+\ln(\delta+1)}{\delta+1}$

Example 3: 2-coloring of k-uniform Hypergraphs

- $G = (V, E)$ a k-uniform hypergraph.
- Intuition: if $|E|$ is relatively small, G is 2-colorable
- We can shown: $|E| \leq 2^{k-1}$ is sufficient using the argument from expectation, but the bound is too small

Why is the bound too small?

Random coloring disregards the structure of the graph.
Need some modification of the random coloring to improve the bound.

A Randomized Algorithm

1. Order V randomly. For $v \in V$, flip 2 coins:
   - Prob$[C_1(v) = \text{HEAD}] = 1/2$;
   - Prob$[C_2(v) = \text{HEAD}] = p$
2. Color v red if $C_1(v) = \text{HEAD}$, blue otherwise
3. $D = \{v \mid v \text{ lies in some monochromatic } e \in E\}$
4. For each $v \in D$ in the random ordering
   - **If** v is still in some monochromatic e in the first coloring and no vertex in e has changed its color, **then** change v's color if $C_2(v) = \text{HEAD}$
   - **Else** do nothing!

Analysis

$$
\begin{aligned}
\text{Prob}[\text{Coloring is bad}] &\leq \sum_{e \in E} \text{Prob}[e \text{ is monochromatic}] \\
&= 2 \sum_{e \in E} \text{Prob}[e \text{ is red}] \\
&\leq 2 \sum_{e \in E} \Bigg(\text{Prob}[\underbrace{e \text{ was red and remains red}}_{A_e}] \\
&\qquad + \text{Prob}[\underbrace{e \text{ wasn't red but turned red}}_{C_e}] \Bigg)
\end{aligned}
$$

$$
\text{Prob}[A_e] = \frac{1}{2^k}(1-p)^k.
$$

The Event C_e

Let v be the last vertex of e to turn blue $\rightarrow$ red

- $v \in f \in E$ and f was blue (in 1st coloring) when v is considered
- $e \cap f = \{v\}$

For any $e \neq f$ with $|e \cap f| = \{v\}$, let B_{ef} be the event that

- f was blue in first coloring, e is red in the final coloring
- v is the last of e to change color
- when v changes color, f is still blue

$$\text{Prob}[C_e] \leq \sum_{f:|f \cap e|=1} \text{Prob}[B_{ef}]$$

The Event B_{ef}

- The random ordering of V induces a random ordering σ of $e \cup f$
- i_σ = number of vertices in e coming before v in σ
- j_σ = number of vertices in f coming before v in σ

$$\text{Prob}\left[B_{ef} \mid \sigma\right] = \frac{1}{2^k} p \frac{1}{2^{k-1-i_\sigma}} (1-p)^{j_\sigma} \left(\frac{1+p}{2}\right)^{i_\sigma}$$

$$\begin{aligned}
\text{Prob}\left[B_{ef}\right] &= \sum_\sigma \text{Prob}\left[B_{ef} \mid \sigma\right] \text{Prob}[\sigma] \\
&= \frac{p}{2^{2k-1}} \underbrace{\mathsf{E}_\sigma\left[(1-p)^{i_\sigma}(1+p)^{j_\sigma}\right]}_{\leq 1} \\
&\leq \frac{p}{2^{2k-1}}
\end{aligned}$$

Putting it All Together

Let $m = |E|$ and $x = m/2^{k-1}$

$$\begin{aligned}
\text{Prob}[\text{Coloring is bad}] &\leq 2\sum_{e} (\text{Prob}[A_e] + \text{Prob}[C_e]) \\
&< 2m\frac{1}{2^k}(1-p)^k + 2m^2\frac{p}{2^{2k-1}} \\
&= x(1-p)^k + x^2 p \\
&\leq 1
\end{aligned}$$

as long as

$$m = \Omega\left(2^k\sqrt{\frac{k}{\ln k}}\right)$$