Lecture 5: October 11

Lecturer: Kuan Yang *Scribe: Weihao Zhu*

5.1 Derandomization

5.1.1 Caro-Wei Inequality

Recall that we have introduced Caro-Wei inequality last time.

Theorem 5.1 (Caro 1979, Wei 1981) *For any graph G,*

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{deg(v)+1}.$$

.

Also, we have introduced its proof via expectation. Here, we will give another algorithmic proof.

Proof: We sort all vertices in degree-non-decreasing order. For each vertex v, we assign $\frac{1}{deg(v)+1}$ to its weight value.

Now, we construct an independent set greedily. At each step, we take the first vertex and remove all its neighbours. Note that the sum of weight value of removed vertices in each step is no larger than 1. As a result, the size of the independent set is at least $\sum_v \frac{1}{deg(v)+1}$, which completes the proof. ■

This proof provides us with a deterministic algorithm that finds an independent set of size at least $\sum_v \frac{1}{deg(v)+1}$. Note that based on the previous proof, we already have a randomized algorithm that finds an independent set of which expected size is exactly $\sum_v \frac{1}{deg(v)+1}$ (generating a random permutation and taking vertices in the way which is used in the probabilistic proof). Now we give a deterministic algorithm that do not use random numbers. We say such deterministic algorithms are *derandomization*.

Now, we give another method for derandomization. Consider an arbitrary order of vertices. For each vertex v, there are two choices: taking v or not taking it. Let $\mathcal{E}$ be the event of taking v, and X be the size of the independent set we obtain (using the randomized algorithm in the probabilistic proof). Note that

$$\mathbf{E}[X] = \mathbf{E}[X \mid \mathcal{E}] \cdot \mathbf{Pr}[\mathcal{E}] + \mathbf{E}[X \mid \overline{\mathcal{E}}] \cdot \mathbf{Pr}[\overline{\mathcal{E}}],$$

which implies that at least one of $\mathbf{E}[X \mid \mathcal{E}]$ and $\mathbf{E}[X \mid \overline{\mathcal{E}}]$ is not less than $\mathbf{E}[X]$. We take vertex v if and only if $\mathbf{E}[X \mid \mathcal{E}]$ is larger. As a result, the size of the independent set we finally obtain is at least $\sum_v \frac{1}{deg(v)+1}$.

5.1.2 Maximum-Cut Problem

Another example we will introduce is the Maximum-Cut Problem, which is as follows.

Problem 5.1 *For a graph $G = (V, E)$, find a maximum cut, which is a partition of the graph's vertices into two complementary sets.*

It is well-known that the Maximum-Cut Problem is NP-hard. However, we have an algorithm that finds a cut with at least $|E|/2$ edges.

For each vertex, we uniformly randomly mark it 0 or 1. For any edge with one endpoint marked 0 and the other one marked 1, we add the edge into the cut. The randomized algorithm finds a cut of which the expected size is $|E|/2$.

Now let's consider the derandomization. List all vertices in an arbitrary order. For each vertex v, there are two choices: marking v as 0 or 1. Then we calculate the expectation of the size of the cut we obtain finally conditioned on marking v as 0 or 1, respectively. If the expectation conditioned on marking it as 0 is larger, then we mark v as 0. Otherwise, we mark it as 1. The final cut we get is at least $|E|/2$. Note that the whole algorithm is deterministic.

Remark To use this method of derandomization, we must have an efficient algorithm to compute the conditioned expectation. For instance, we couldn't do something like this in the proof for Ramsey number, since it is difficult to compute the number of cliques of other subsets.

5.2 Alteration

5.2.1 Lower Bounds of Ramsey Number

Recall that we introduced a lower bound of Ramsey Number in Lecture 3, which is as follows.

Theorem 5.2 (Paul Erdős, 1947) *$R(k, k) > n$ holds if*

$$\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1.$$

In its proof, we color the edges of K_n independently and uniformly at random. Fix a set $S \in \binom{[n]}{k}$. The probability that S induces a monochromatic K_k is $2^{1-\binom{k}{2}}$. We showed in the previous lecture that by the union bound, if $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ there exists a coloring such that no monochromatic K_k exists. However, we could do something clever. If there exists a monochromatic K_k in the graph, we can remove an vertex from the graph to obtain a graph and a coloring without monochromatic K_k.

In fact, let X be the random variable of the number of monochromatic K_ks. Then $\mathbf{E}[X]$ is $\binom{n}{k} \cdot 2^{1-\binom{k}{2}}$. So there exists a coloring such that the number of monochromatic K_k is at most $\mathbf{E}[X]$.

For any monochromatic K_k, we delete a vertex of it from the graph. The remaining size will be at least $n - \mathbf{E}[X]$, which implies the following theorem.

Theorem 5.3 *For any k, n, we have*

$$R(k, k) > n - \binom{n}{k} \cdot 2^{1-\binom{k}{2}}.$$

Remark Theorem 5.2 implies that

$$R(k,k) > (\frac{1}{e\sqrt{2}} + o(1))k \cdot 2^{k/2},$$

while Theorem 5.3 gives us a better lower bound as follows:

$$R(k,k) > (\frac{1}{e} + o(1))k \cdot 2^{k/2}.$$

Note that for general Ramsey Number, we also have a similar lower bound:

Theorem 5.4 *For any s, t, n and $p \in [0,1]$, we have*

$$R(s,t) > n - \binom{n}{s} \cdot p^{\binom{s}{2}} - \binom{n}{t} \cdot (1-p)^{\binom{t}{2}}.$$

This method is called *alteration*. When random structures do not have all the desired properties but may have some "bad parts", we can alter the structure to remove all "bad parts".

5.2.2 Dominating Set

Another classical application of alteration is the problem of minimum dominating set. We first give its definition as follows.

Definition 5.1 *For any graph $G = (V, E)$, a vertex set $U \subseteq V$ is a dominating set if for any vertex v, $N^+(v) \cap U \neq \emptyset$, where $N^+(v)$ is the set of v and all its neighbors.*

Actually, we can always find a dominating set which is not too large.

Theorem 5.5 *For any graph of size n with minimum degree $\delta > 1$, a dominating set of size at most $\frac{\log(\delta+1)+1}{\delta+1} \cdot n$ exists.*

We first consider some naive attempts. If we do something like what we did in the proof of Caro–Wei inequality, we can take a vertex into the dominating set and remove all its neighbors. However, after taking the first vertex, we cannot bound the number of removed vertices at each step, since the minimum degree of vertices in the remaining graph is no longer δ.

Proof: We use the alteration method. Let $p \in [0,1]$ be a fixed parameter to be determined later. We independently pick each vertex into set X with probability p. Consider the vertices that are not dominated by vertices in X. Let $Y = V \setminus N^+(X)$. Clearly, $X \cup Y$ is a dominating set. Now, let's bound the size of $X \cup Y$.

Note that for any vertex v, $\mathbf{Pr}[v \in Y] \leq (1-p)^{1+\delta}$, since neither v nor its neighbors are in X. Thus, we have

$$\begin{aligned}
\mathbf{E}[|X \cup Y|] &= \mathbf{E}[|X|] + \mathbf{E}[|Y|] \\
&\leq pn + (1-p)^{1+\delta} n \\
&\leq (p + e^{-p(1+\delta)})n \qquad \text{(minimize it by setting } p = \frac{\log(\delta+1)}{\delta+1}\text{)} \\
&\leq (\frac{\log(\delta+1)+1}{\delta+1})n,
\end{aligned}$$

which completes the proof. ■

Remark We can also apply the same technique when we are searching for an independent set. We independently pick each vertex into set X with probability p. For each edge, if both of its endpoints are selected, we remove one of them. In this way, the expectation on the size of the independent set is $pn - p^2m$ in expectation. By setting $p = \frac{n}{2m}$, we can show that there exists an independent set of size at least $\frac{n^2}{4m}$. However, this bound is worse than Caro–Wei inequality.

5.3 Chromatic Number

In this section, we will discuss some interesting problems on chromatic number.

5.3.1 Graph with High Chromatic Number

If a graph has a k-clique, then we can say $\chi \geq k$. Conversely, if χ is large, is it always possible to verify it by observing local information? Surprisingly, this is far from being true, even for "locally tree-like" graphs.

Theorem 5.6 (Paul Erdős, 1959) *For any ℓ, k, there exists a graph with girth (the minimal length of cycles) larger than ℓ and chromatic number larger than k.*

Proof: Let $G \sim \mathcal{G}(n,p)$ with $p = (\log n)^2/n$. Let X be the number of cycles of length no larger than ℓ. In K_n, there are exactly $\binom{n}{i} \cdot \frac{(i-1)!}{2}$ cycles of length i. Therefore, we have

$$\mathbf{E}[X] = \sum_{i=3}^{\ell} \binom{n}{i} \cdot \frac{(i-1)!}{2} \cdot p^i \leq \sum_{i=3}^{\ell} n^i \cdot p^i = o(n).$$

According to Markov's inequality, we have $\mathbf{Pr}[X \geq \frac{n}{2}] = o(1) < \frac{1}{2}$.

Note that $\chi(G) \geq \frac{n}{\alpha(G)}$. It suffices to show $\alpha(G) \leq \frac{n}{k}$. By setting $t = \frac{3 \log n}{p}$, we have

$$\mathbf{Pr}[\alpha(G) \geq t] \leq \binom{n}{t} \cdot (1-p)^{\binom{t}{2}} < n^t \cdot e^{-p \cdot \frac{(t-1)t}{2}} = o(1) < \frac{1}{2}.$$

Let n be sufficiently large such that $\mathbf{Pr}[X < \frac{n}{2} \wedge \alpha(G) < t] > 0$. For each cycle no larger than ℓ, we remove a vertex from it. Suppose the graph we get is G'. Then $|V(G')| \geq \frac{n}{2}$. Also, its girth is larger than ℓ, and $\alpha(G') \leq \alpha(G) \leq \frac{3 \log n}{p}$, which implies that $\chi(G') \geq \frac{np}{6 \log n} > k$. This completes the proof. ■

Remark. Actually, constructing such a graph is not that easy. Here we introduce how to construct a triangle-free graph with large chromatic number. Let G_2 be graph K_2. Given $G_{n-1} = (V, E)$ for $n \geq 3$, construct $G_n = (V \cup V' \cup \{w\}, E \cup E')$ as follows:

- V' is a copy of V;
- For any $(u, v) \in E$, add (u', v) in E' where u' is the copy of u in V';
- For any $v' \in V'$, add (v', w) in E'.

It's easy to check that G_n is triangle-free and its chromatic number equals to n.

Actually, even if we check the chromatic number of a sub-graph induced by $\varepsilon \cdot n$ vertices, we still know nothing about the chromatic number of the whole graph. Paul Erdős proved the following theorem.

Theorem 5.7 (Paul Erdős, 1962) *For any $k > 0$, there exists a positive ε such that for any sufficiently large n, there exists a graph G of size n with $\chi(G) \geq k$, while $\chi(G[S]) \leq 3$ for all $|S| \leq \varepsilon \cdot n$ (where $G[S]$ is the subgraph of G induced by the vertex set S).*

Proof: For a fixed k, let c, ε satisfy $c > 2\ln 2 \cdot k^2 H(1/k)$ and $\varepsilon < 3^3 e^{-5} \cdot c^{-3}$, where

$$H(x) = -x \log_2 x - (1-x)\log_2(1-x)$$

is the entropy function.

Set $p = c/n$ and let $G \sim \mathcal{G}(n,p)$. Now, let's prove that G satisfies all requirements almost surely.

Let's first discuss its chromatic number. If $\chi(G) \leq k$, then $\alpha(G) \geq \frac{n}{k}$. Let X be the number of independent sets of size $\frac{n}{k}$. We have

$$\mathbf{E}[X] = \binom{n}{n/k}(1-p)^{\binom{n/k}{2}} < 2^{n(H(1/k)+o(1))} \cdot e^{-cn/2k^2 \cdot (1+o(1))},$$

which is $o(1)$ by our condition on c. This implies that $\chi(G) > k$ almost surely.

Now, let's consider the other constraint. If there exists a set of size no larger than $\varepsilon \cdot n$ such that the chromatic number of its induced sub-graph is larger than 3, let S be the minimal set such that $\chi(G[S]) = 4$. Thus, for any vertex $v \in S$, $\chi(G[S \setminus \{v\}]) = 3$ which implies that v has at least 3 neighbors in S. Therefore, there are at least $\frac{3|S|}{2}$ in $G[S]$. Let $t = |S|$. We have

$$\begin{aligned}\mathbf{Pr}[\exists \text{a size-}t \text{ induced sub-graph with at least } \frac{3t}{2} \text{ edges}] &\leq \binom{n}{t} \cdot \binom{\binom{t}{2}}{3t/2} \cdot (\frac{c}{n})^{3t/2} \\ &\leq (\frac{ne}{t} \cdot (\frac{te}{3})^{3/2} (\frac{c}{n})^{3/2})^t \\ &\leq (e^{5/2} \cdot 3^{-3/2} \cdot c^{3/2} \cdot \sqrt{t/n})^t.\end{aligned}$$

Now, we can show that

$$\begin{aligned}&\sum_{t \leq \varepsilon \cdot n} \mathbf{Pr}[\exists \text{a size-}t \text{ induced sub-graph with at least } \frac{3t}{2} \text{ edges}] \\ \leq &\sum_{t \leq \varepsilon \cdot n} (e^{5/2} \cdot 3^{-3/2} \cdot c^{3/2} \cdot \sqrt{t/n})^t \\ = &o(1),\end{aligned}$$

which implies that $\forall |S| \leq \varepsilon \cdot n$, $\chi(G[S]) \leq 3$ holds with high probability. This completes the proof. ■

5.3.2 2-colorable Hypergraphs Revisit

In Lecture 3, we have introduced the definition of $m(k)$, which is the minimum number of edges in a k-uniform hypergraph that is not 2-colorable. We have also proved that $2^{k-1} \leq m(k) \leq k^2 \cdot 2^k$. Today, we will give a better bound of $m(k)$.

Theorem 5.8 (Radhakrishnan & Srinivasan, 2000) $m(k) = \Omega(2^k \cdot \sqrt{\frac{k}{\log k}})$.

Proof: (by Cherkashin & Kozik, 2015) Suppose $H = (V, E)$ with $|E| = m$. We randomly select a weight function $w : V \to [0, 1]$. Also, we color all vertices according to their weights from small to large. For each vertex v, color it 0 if it does not form an all-0 edge. Otherwise, we color it 1. In this way, we construct a coloring which leads to no all-0 edge. Now, let's bound the probability that an all-1 edge exists.

We can see that for any color-1 vertex v, there exists an edge $e \in E$ such that v has the largest weight in e. This implies that for an all-1 edge f, there exists an edge e such that the vertex with the smallest weight in f has the largest weight in e. In this case, we say the edge pair (e, f) is conflicting. We try to bound the probability that a conflicting edge pair exists.

Let $L = [0, \frac{1-p}{2})$, $M = [\frac{1-p}{2}, \frac{1+p}{2}]$ and $R = (\frac{1+p}{2}, 1]$ where p is determined later. For any edge $e \in E$, $\mathbf{Pr}[\forall v \in e, w(v)$ lies in $L] = \mathbf{Pr}[\forall v \in e, w(v)$ lies in $R] = (\frac{1-p}{2})^k$. Therefore, $\mathbf{Pr}[\exists$ all L or all R edge$] \leq 2m \cdot (\frac{1-p}{2})^k$.

Now, let's assume that no all-L or all-R edge exists. If (e, f) is conflicting, then $e \cap f$ is a single vertex of which the weight value lies in M. Assume that $v = e \cap f$. Then, we have

$$\begin{aligned}\mathbf{Pr}[v \text{ has largest weight value in } e \text{ and smallest weight value in } f] &= \int_{(1-p)/2}^{(1+p)/2} x^{k-1}(1-x)^{k-1}dx \\ &\leq p/4^{k-1}.\end{aligned}$$

Therefore, we have

$$\begin{aligned}\mathbf{Pr}[\exists \text{ conflicting } (e,f)] &\leq 2m\Big(\frac{1-p}{2}\Big)^k + m^2 \cdot \frac{p}{4^{k-1}} \\ &< 2^{1-k}me^{-pk} + (2^{1-k}m)^2 p \\ &< 1.\end{aligned}$$

The last inequality is true when $m \leq c \cdot 2^k \sqrt{\frac{k}{\log k}}$ for some sufficiently small constant $c > 0$. This completes the proof. ∎

5.4 Chebyshev's Inequality and Second Moment Method

Markov's inequality is an important tool when bounding probability. It states that $\mathbf{Pr}[X \geq a] \leq \frac{\mathbf{E}[X]}{a}$ for $a > 0$. However, can we do better?

Theorem 5.9 (Chebyshev's Inequality) $\mathbf{Pr}[|X - \mathbf{E}[X]| \geq t] \leq \frac{\mathbf{Var}[x]}{t^2}$.

Remark Variance $\mathbf{Var}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2] = \mathbf{E}[X^2] - \mathbf{E}[X]^2$, which is usually denoted by σ^2. $\mathbf{E}[X]$ is usually denoted by μ.

Proof:

$$\begin{aligned}\sigma^2 = \mathbf{E}[(X-\mu)^2] &= \mathbf{Pr}[|X-\mu| \geq t] \cdot \mathbf{E}[(X-\mu)^2 \mid |X-\mu| \geq t] + \mathbf{Pr}[|X-\mu| \leq t] \cdot \mathbf{E}[(X-\mu)^2 \mid |X-\mu| \leq t] \\ &\geq \mathbf{Pr}[|X-\mu| \geq t] \cdot t^2.\end{aligned}$$

∎

The use of Chebyshev's inequality is called the second moment method. Now, we will introduce two applications.

5.4.1 Distinct Sums Problem

Problem 5.2 *Let S be a positive integer set of size k of which all 2^k subset sums are distinct. What is the minimum possible value of the largest element in S?*

A simple argument shows that $\max S \geq 2^k/k$ since all subset sums are at most $k \max S$. However, we can bound $\max S$ in a more clever way, because most subset sums "concentrate" to the mean value by the Chebyshev's inequality.

Theorem 5.10 $\max S \gtrsim \frac{2^k}{\sqrt{k}}$.

Proof: Let $S = \{x_1, \ldots, x_k\}$ and $n = \max S$. For $1 \leq i \leq k$, choose $\varepsilon_i \in \{0,1\}$ independently and uniformly at random. Let $X = \sum \varepsilon_i x_i$. Thus, we have $\mu = \mathbf{E}[X] = \frac{\sum x_i}{2}$. Also, the variance $\sigma^2 = \mathbf{Var}[X] = \frac{\sum x_i^2}{4} \leq \frac{nk^2}{4}$.

By Chebyshev's inequality, $\mathbf{Pr}[|X-\mu| < n\sqrt{k}] \geq \frac{3}{4}$. Since X takes distinct values for distinct $(\varepsilon_1, \ldots, \varepsilon_k) \in \{0,1\}^k$, we have $\mathbf{Pr}[X = r] \leq 2^{-k}$ for all r. Thus, we have $\mathbf{Pr}[|X-\mu| < n\sqrt{k}] \leq 2^{-k} \cdot 2n\sqrt{k}$, which implies that $2^{-k} \cdot 2n\sqrt{k} \leq \frac{3}{4}$. This completes the proof. ∎

Remark In 2020, Dubroff, Fox and Xu proved that $\max S \gtrsim \left(\sqrt{\frac{2}{\pi}} + o(1)\right)\frac{2^k}{\sqrt{k}}$.

5.4.2 Weierstrass Approximation Theorem

Now, we will introduce an application of the second moment method to analysis.

Theorem 5.11 (Weierstrass, 1885) *Suppose $f : [0,1] \to \mathbb{R}$ is continuous. For every $\varepsilon > 0$, there exists a polynomial $p(x)$ such that*

$$\forall x \in [0,1], \quad |p(x) - f(x)| \leq \varepsilon.$$

Proof: (by Bernstein, 1912) Since $[0,1]$ is compact, f is uniformly continuous and bounded. Without loss of generality, assume $|f(x)| \leq 1$. There exists $\delta > 0$ such that $|f(x) - f(y)| \leq \frac{\varepsilon}{2}$ for all $|x - y| \leq \delta$.

Now, we approximate f by

$$P_n(x) = \sum_{i=0}^{n} E_i(x) f(\frac{i}{n}),$$

where

$$E_i(x) = \mathbf{Pr}[\mathrm{Bin}(n,x) = i] = \binom{n}{i} x^i (1-x)^i.$$

Note that $E_i(x)$ peaks at $\frac{i}{n}$ and decays away from $\frac{i}{n}$. Since $\mathrm{Bin}(n,x)$ has expectation nx and variance $nx(1-x) \leq \frac{n}{4}$, with Chebyshev's inequality we have

$$\sum_{i:|i-nx|>n^{2/3}} E_i(x) = \mathbf{Pr}[|\mathrm{Bin}(n,x) - nx| > n^{2/3}] \leq n^{-1/3}.$$

Note that $\sum_{i=0}^{n} E_i(x) = 1$. Taking $n > \max\{64\varepsilon^{-3}, \delta^{-3}\}$, we have

$$|P_n(x) - f(x)| \leq \sum_{i=0}^{n} E_i(x)|f(\frac{i}{n}) - f(x)| \leq \sum_{|i-nx|\leq n^{2/3}} E_i(x) \cdot \frac{\varepsilon}{2} + 2n^{-1/3} < \varepsilon,$$

which completes the proof. ∎