CSCE 658: Randomized Algorithms

Lecture 14

Samson Zhou

Probabilistic Method

- Suppose we want to argue the existence of a certain desirable object
- Existential argument, non-constructive

- If there is an algorithm that can find it, it must exist!

Ramsey Numbers

- What is the smallest number $n = R(a, b)$ such that in any set of n people, there must be either:
  - a mutual acquaintances
  - b mutual strangers
- $R(a, b)$ are the Ramsey numbers

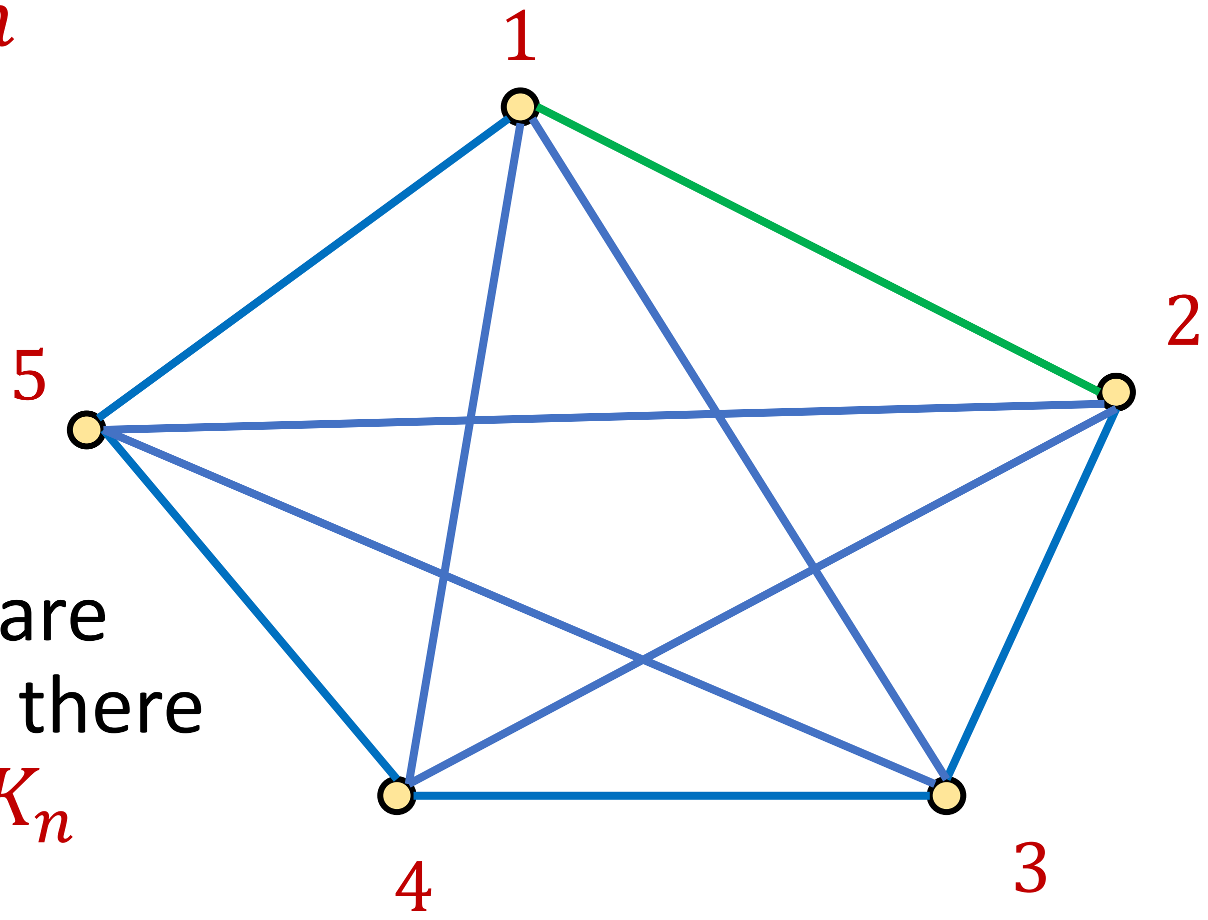
Ramsey Numbers

- We can model a set of n people with a complete graph by coloring an edge (i, j) BLUE if i and j are acquaintances and GREEN if i and j are strangers

- What is the smallest number $n = R(a, b)$ such there must be either:
  - BLUE induced complete subgraph K_a
  - GREEN induced complete subgraph K_b

Ramsey Numbers

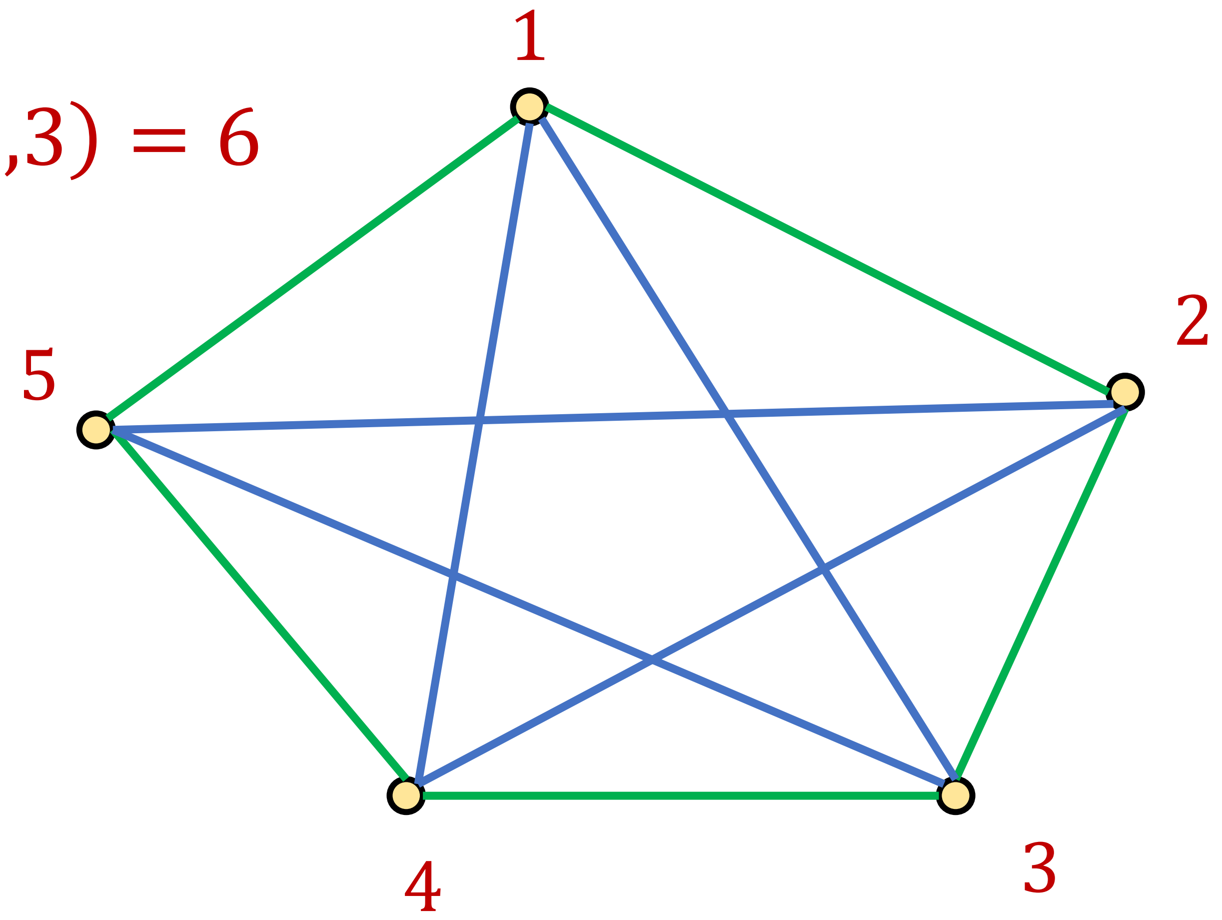
- $R(2,n) = n$



- If no edges are green, then there exists blue K_n

Ramsey Numbers

- $R(3,3) > 5$
- In fact, $R(3,3) = 6$



Ramsey Numbers

- Finding the precise value of $R(a, b)$ is quite difficult
- $R(3,3) = 6$
- $R(4,4) = 18$
- $43 \leq R(5,5) \leq 48$
- $102 \leq R(6,6) \leq 161$
- $205 \leq R(7,7) \leq 497$

Probabilistic Method for Ramsey Numbers

- If $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, then $R(k,k) > n$ (Erdös)

- Consider a random coloring of K_n, so that each edge is colored BLUE with probability $\frac{1}{2}$ and GREEN with probability $\frac{1}{2}$
- For any fixed set S of k vertices, the probability S is monochromatic is $\frac{1}{2^{\binom{k}{2}}} + \frac{1}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}$

Probabilistic Method for Ramsey Numbers

- By a union bound, the probability that there exists a set of k vertices is monochromatic is $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$.
- Then with nonzero probability, algorithm finds a coloring with no monochromatic K_k
- Thus, there exists a graph coloring with no monochromatic K_k
- $R(k,k) > n$

Probabilistic Method

- Suppose we want to argue the existence of a certain desirable object
- Existential argument, non-constructive

- If there is an algorithm that can find it, it must exist!

Probabilistic Method

- Suppose we want to argue the existence of a certain desirable object
- Existential argument, non-constructive

- A random variable cannot always be less than its expected value
- A random variable cannot always be more than its expected value

Probabilistic Method for Graph Cuts

- Any undirected graph G with m edges has a cut of at least $\frac{m}{2}$ edges

- Consider a random cut of G formed by putting each vertex into A with probability $\frac{1}{2}$ and into B with probability $\frac{1}{2}$
- Let the edges be $e_1, \ldots, e_m$ and let X_i denote whether e_i crosses the cut

Probabilistic Method for Graph Cuts

- The probability that e_i crosses the cut (A, B) is $\frac{1}{2}$
- $E[X_i] = \frac{1}{2}$
- Let $|C(A, B)|$ denote the size of the cut (A, B)
- $E[|C(A, B)|] = E\left[\sum_{i \in [m]} X_i\right] = E[X_1] + \cdots + E[X_m] = \frac{m}{2}$
- Thus, there exists a cut of size $\frac{m}{2}$

k-SAT

- In the k-SAT problem, we are given a conjunctive normal form (CNF) formula, i.e., an AND of OR's, $f(x_1, \dots, x_n)$ with m clauses $C_1, \dots., C_m$ and k distinct variables per clause

- Example for $k = 4$:

$$(x_2 \vee \neg x_4 \vee x_5 \vee x_7) \wedge (x_1 \vee \neg x_3 \vee x_6 \vee x_8)$$

Probabilistic Method for k-SAT

- Suppose $m < 2^k$, we claim f must be satisfiable!

Probabilistic Method for k-SAT

- Suppose $m < 2^k$, we claim f must be satisfiable!
- Suppose we assign each variable x_i a separate random TRUE/FALSE value
- For each $i \in [m]$, we have $\Pr[C_i \text{ is FALSE}] \le 1/2^k$
- By a union bound
- $\Pr[f(x_1, \dots, x_n) = \text{FALSE}] \le \sum_{i \in [m]} \Pr[C_i \text{ is FALSE}]$

$$\le \frac{m}{2^k} < 1$$

Probabilistic Method for k-SAT

- In the k-SAT problem, we are given a CNF formula $f(x_1, \dots, x_n)$ with m clauses $C_1, \dots, C_m$ and k distinct variables per clause

- If $m < 2^k$, then f is satisfiable
- What about $m \geq 2^k$?

Dependency Graph

- Let $E_1, \dots, E_n$ be events and let G be a graph on the vertices $[n] := \{1, \dots, n\}$
- G is called a dependency graph for the events $E_1, \dots, E_n$ if and only if E_i is mutually independent of all events E_j for which (i, j) is not an edge in E

- G models the dependencies between the events $E_1, \dots, E_n$

Lovász Local Lemma

- Theorem: Let $E_1, \ldots, E_n$ be events and let G be their dependency graph. Suppose for all $i \in [n]$,

$$\Pr[E_i] \leq p, \qquad \deg(i) \leq d, \qquad 4dp \leq 1$$

- Then $\Pr\left[E_1^C \cap E_2^C \cap \cdots \cap E_n^C\right] > 0$, where E_i^C denotes the complement of E_i

Lovász Local Lemma

- To show $\Pr\left[E_1^C \cap E_2^C \cap \cdots \cap E_n^C\right] > 0$, it suffices to show $\Pr\left[E_i \mid E_1^C \cap E_2^C \cap \cdots \cap E_{i-1}^C\right] \leq 2p$ for all $i \in [n]$.

Lovász Local Lemma

- To show $\Pr\left[E_1^C \cap E_2^C \cap \cdots \cap E_n^C\right] > 0$, it suffices to show $\Pr\left[E_i \mid E_1^C \cap E_2^C \cap \cdots \cap E_{i-1}^C\right] \leq 2p$ for all $i \in [n]$.

- Indeed:

$$\Pr\left[E_1^C \cap E_2^C \cap \cdots \cap E_n^C\right] = \Pi_{i=1}^{n} \Pr\left[E_i^C \mid E_1^C \cap E_2^C \cap \cdots \cap E_{i-1}^C\right]$$

$$\geq \Pi_{i=1}^{n}(1-2p) > 0$$

Lovász Local Lemma

- To show $\Pr\left[E_i \mid E_1^C \cap E_2^C \cap \cdots \cap E_{i-1}^C\right] \leq 2p$ for all $i \in [n]$, we instead show $\Pr\left[E_i \mid \cap_{j \in S} E_j^C\right] \leq 2p$ for all $|S| \leq s$
- Use induction on s
- Our assumption is that for all $i \in [n]$:

$$\Pr[E_i] \leq p, \qquad \deg(i) \leq d, \qquad 4dp \leq 1$$

- Base case follows from assumption for $s = 1$

Lovász Local Lemma

- Assume true for $s-1$, show $\Pr\left[E_i \mid \cap_{j\in S} E_j^C\right] \leq 2p$ for all $|S| \leq s$
- Let Λ be the neighbors of i in G
- By joint probability,

$$\Pr\left[E_i \mid \cap_{j\in S} E_j^C\right] = \frac{\Pr\left[E_i \cap \bigcap_{j\in\Lambda} E_j^C \mid \bigcap_{j\in S\setminus\Lambda} E_j^C\right]}{\Pr\left[\bigcap_{j\in\Lambda} E_j^C \mid \bigcap_{j\in S\setminus\Lambda} E_j^C\right]}$$

Lovász Local Lemma

- The numerator is $\Pr\left[E_i \cap \bigcap_{j\in\Lambda} E_j^C \mid \bigcap_{j\in S\setminus\Lambda} E_j^C\right]$
- We have $\Pr\left[E_i \cap \bigcap_{j\in\Lambda} E_j^C \mid \bigcap_{j\in S\setminus\Lambda} E_j^C\right] \le \Pr\left[E_i \mid \bigcap_{j\in S\setminus\Lambda} E_j^C\right]$
- Since E_i is independent of E_j for $j \in S \setminus \Lambda$, then $\Pr\left[E_i \mid \bigcap_{j\in S\setminus\Lambda} E_j^C\right] = \Pr[E_i] \le p$

Lovász Local Lemma

- The denominator is $\Pr\left[\cap_{j\in\Lambda} E_j^C \mid \cap_{j\in S\setminus\Lambda} E_j^C\right]$
- Our assumption is that for all $i \in [n]$:

$$\Pr[E_i] \le p, \qquad \deg(i) \le d, \qquad 4dp \le 1$$

- By a union bound,

$$\Pr\left[\cap_{j\in\Lambda} E_j^C \mid \cap_{j\in S\setminus\Lambda} E_j^C\right] \ge 1 - \sum_{j\in\Lambda} \Pr\left[E_j \mid \cap_{j\in S\setminus\Lambda} E_j^C\right]$$

$$\ge 1 - \sum_{j\in\Lambda} 2p \ge 1 - 2pd \ge \frac{1}{2}$$

Lovász Local Lemma

- Assume true for $s-1$, show $\Pr\left[E_i \mid \cap_{j\in S} E_j^C\right] \leq 2p$ for all $|S| \leq s$
- Let Λ be the neighbors of i in G
- By conditional probability,

$$\Pr\left[E_i \mid \cap_{j\in S} E_j^C\right] = \frac{\Pr\left[E_i \cap \bigcap_{j\in\Lambda} E_j^C \mid \bigcap_{j\in S\setminus\Lambda} E_j^C\right]}{\Pr\left[\bigcap_{j\in\Lambda} E_j^C \mid \bigcap_{j\in S\setminus\Lambda} E_j^C\right]} \leq \frac{p}{(1/2)} = 2p$$

Lovász Local Lemma

- **Theorem**: Let $E_1, \ldots, E_n$ be events and let G be their dependency graph. Suppose for all $i \in [n]$,

$$\Pr[E_i] \leq p, \qquad \deg(i) \leq d, \qquad 4dp \leq 1$$

- Then $\Pr\left[E_1^C \cap E_2^C \cap \cdots \cap E_n^C\right] > 0$, where E_i^C denotes the complement of E_i

Probabilistic Method for k-SAT

- In the k-SAT problem, we are given a CNF formula $f(x_1, \ldots, x_n)$ with m clauses $C_1, \ldots, C_m$ and k distinct variables per clause

- If $m < 2^k$, then f is satisfiable
- What about $m \geq 2^k$?

Resampling Algorithm for k-SAT

- We say clauses C_i and C_j intersect if there exists a variable x_k (or its negation) that appears in both C_i and C_j

- Theorem: If each clause intersects with at most $d \leq \frac{2^k}{4}$ other clauses, then f is satisfiable

Resampling Algorithm for k-SAT

- Suppose we assign each variable x_i a separate random TRUE/FALSE value
- For each $i \in [m]$, we have $\Pr[C_i \text{ is FALSE}] \leq 1/2^k$
- If each clause intersects with at most $d \leq \frac{2^k}{4}$ other clauses, then by the Lovász Local Lemma, the algorithm finds satisfying assignment with nonzero probability
- Thus by the probabilistic method, the assignment must be satisfiable

Resampling Algorithm for k-SAT

- Suppose we assign each variable x_i a separate random TRUE/FALSE value
- As long as there is a clause C_j that is unsatisfied, we resample all the variables in C_j independently and uniformly at random
- Algorithm may never terminate?
- Algorithmic version of the Lovász Local Lemma (we will not cover this)

Edge-Disjoint Paths

- Suppose there are n pairs of users who want to communicate over a network. Find a routing such that no communication paths for each pair share any edges
- Theorem: Let P_i be the set of paths that pair i can use. Suppose:
  - $|P_i| \geq m$ for all $i \in [n]$
  - For all $i \neq j$ and any path $P \in P_i$, there are at most k other paths $P' \in P_j$ that conflict with P
- If $\frac{8nk}{m} \leq 1$, then there exists a routing with no conflicting paths

Edge-Disjoint Paths

- Suppose $|P_i| = m$ and choose a random path from each P_i, independently for each $i \in [n]$
- Let $E_{i,j}$ be the event that the paths chosen from P_i and P_j conflict
- After fixing a path from P_i, there are at most k conflicting paths P_j among m possible paths, so that $\Pr[E_{i,j}] \leq k/m$

- Set $p = k/m$ in the Lovász Local Lemma

Edge-Disjoint Paths

- Since $E_{i,j}$ is independent of $E_{x,y}$ for $x, y \notin \{i, j\}$, then each vertex in the dependency graph has degree less than $2n$
- Set $d < 2n$ in the Lovász Local Lemma
- Then $4pd < 4\left(\frac{k}{m}\right)(2n) = \frac{8nk}{m} \leq 1$
- By the Lovász Local Lemma, the algorithm finds a disjoint routing with nonzero probability
- Thus by the probabilistic method, there exists a disjoint routing