CS 466/666: Algorithm Design and Analysis **University of Waterloo: Fall 2025**

Lecture 20

November 18, 2025

Instructor: Sepehr Assadi

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

Topics of this Lecture

In this lecture, we review probabilistic method (again) and go over one of the strongest tools for it called the Lovász Local Lemma.

1 Motivation

We have used the **probabilistic method** so far in the course multiple times (for instance, in Lecture 14 for showing existence of sparse recovery matrices): to show an object with certain properties exist, we design a randomized process that generates it with a non-zero probability (or in expectation). Let us see another example—and for historical reasons, perhaps the most canonical one—from extremal graph theory.

Ramsey numbers

Let $r, s \geqslant 1$ be integers. Is it the case that *every* large enough graph G, contains either a clique of size r or an independent set of size s? Specifically, define $R(r, s)$ as the smallest integer such that any graph G with $n \geqslant R(r, s)$ vertices either contains an r-clique or an s-independent set. The parameter $R(r, s)$ is referred to as the **Ramsey number** of r and s. Is $R(r, s)$ well-defined for any r and s (as in, it is finite), and if so, what bounds can we prove on its value in terms of r and s?

> **Remark.** Before moving on, let us mention a very common math puzzle in this space: what is the minimum number of people in a party such that either there are 3 people who all know each other, or 3 people who neither know each other? The connection is that the answer is $R(3, 3)$.

Proposition 1. *For any $r, s \geqslant 2$, $R(r, s) \leqslant R(r-1, s) + R(r, s-1)$. As a corollary, $R(r, s) \leqslant 2^{r+s}$.* [1]

[1] We note that the corollary part is loose even given the recurrence and can be improved to getting an upper bound of $\binom{r+s-2}{r-1}$ instead of just 2^{r+s}, but for simplicity, we ignore this stronger bound.

Proof. Let G be an n-vertex graph with no r-clique nor a s-independent set. Consider any vertex v in G and its neighbors $N(v)$. We should have that

$$|N(v)| < R(r-1, s)$$

as otherwise, there is either a s-independent set in $N(v)$, or a $(r-1)$-clique inside it which together with v form an r-clique in G, both contradicting the condition on G. Similarly,

$$|V \setminus (\{v\} \cup N(v))| < R(r, s-1),$$

as otherwise, there is either a r-clique in $V \setminus (\{v\} \cup N(v))$ or a $(s-1)$-independent set which together with v form an s-independent set in G, again, a contradiction. Putting these together, we have

$$|N(v)| + |V \setminus (\{v\} \cup N(v))| = n - 1 < R(r-1, s) + R(r, s-1).$$

This implies that

$$n \leqslant R(r-1, s) + R(r, s-1) - 1.$$

Thus, for any larger graph, we will have either an r-clique or an s-independent set, implying

$$R(r, s) \leqslant R(r-1, s) + R(r, s-1).$$

We can prove the corollary by induction. We have $R(2, 2) = 1$ clearly which satisfies $R(2, 2) \leqslant 2^4$ as our induction base. For the step,

$$R(r, s) \leqslant R(r-1, s) + R(r, s-1) \leqslant 2^{r+s-1} + 2^{r+s-1} = 2^{r+s},$$

concluding the proof. □

An interesting consequence of Proposition 1 is that any graph G either has a clique or an independent set of size $(\log n)/2$ (simply because $R(s, s) \leqslant 2^{2s}$). But how tight is this statement? For instance, could we have proved the same statement by replacing $(\log n)/2$ with, say, $\log^2 n$ or even $\sqrt{n}$? Are there graphs wherein the bound of $\sim \log n$ is asymptotically optimal? How can we find an example? Let us see one of the very first applications of the probabilistic method that shows $\log n$ is asymptotically optimal (this result is due to Erdős and is often credited with popularizing the probabilistic method as a technique).

Proposition 2. *There are n-vertex graphs that do not contain neither a clique nor an independent set of size more than $2 \log n + \Theta(1)$.*

Proof. Let G be a random graph, meaning that we pick each edge independently with probability $1/2$. Let $k > 1$ be a parameter to be chosen later. For any set S of vertices with size k,

$$\Pr(S \text{ is a } k\text{-clique or a } k\text{-independent set}) = 2 \cdot 2^{-\binom{k}{2}},$$

because the two events are disjoint and either happens if all edges are picked or none are picked. By union bounding over all choices of S, we have,

$$\begin{aligned}
\Pr(G \text{ has a } k\text{-clique or a } k\text{-independent set}) &\leqslant \sum_{S \subseteq V : |S| = k} \Pr(S \text{ is a } k\text{-clique or a } k\text{-independent set}) \\
&= \binom{n}{k} \cdot 2 \cdot 2^{-\binom{k}{2}} \\
&< n^k \cdot 2^{-(\binom{k}{2} - 1)}.
\end{aligned}$$

We now pick k such that the RHS above is (much) less than one, which requires

$$n^k \leqslant 2^{\binom{k}{2} - 1} \implies k \cdot \log n \leqslant \frac{k \cdot (k-1)}{2} - 1.$$

This implies that for some constant $C > 0$, setting $k = 2 \log n + C$ ensures that the RHS of the main inequality above is at most 1. Thus, with non-zero probability, we have a graph G without a k-clique or k-independent set, concluding the proof. □

This concludes another application of probabilistic method. An important remark is in order: in all the applications we have used so far, we in fact managed to show the object exist even with high (constant) probability: this somehow means that we were searching for a "hay in a haystack"! But, what if we are in a scenario that we are really searching for a "needle in a haystack"? This is going to be the topic of our lecture (and the next one). Let us use a couple of simple examples to showcase such a scenario.

1.1 Large Independent Sets in Bounded-Degree Graphs

Let us consider two different statements about independent sets in graphs.

Statement 1: Suppose we have a graph $G = (V, E)$ with maximum degree Δ. Prove that there is always an independent set of size at least $n/(\Delta + 1)$ in G.

Proof. Pick π to be a random permutation of vertices and let S be the following independent set: iterate over vertices in the order of π, pick the first available vertex in S and remove all its neighbors; continue like this until all vertices are either picked in S or are removed. Clearly, S is an independent set. Moreover, by linearity of expectation,

$$\mathbb{E}\,|S| = \sum_{v \in V} \Pr\left(v \in S\right) = \sum_{v \in V} \Pr\left(\pi(v) \text{ is the minimum } \pi\text{-value for } \{v\} \cup N(v)\right) = n \cdot \frac{1}{\Delta + 1};$$

thus, there is always an independent set with the desired size. □

Statement 2: Suppose we have a graph $G = (V, E)$ with maximum degree Δ with an arbitrary partitioning of vertices $V = V_1 \sqcup V_2 \sqcup \ldots \sqcup V_r$ in to r sets of size 6Δ. Prove that there is always an independent set in G that picks exactly one vertex from each V_i.

Notice that this statement is a qualitative generalization of the first one since it puts more structure on the independent set, although quantitatively it is a bit weaker because the independent set size we get here is only $n/6\Delta$. Let us see a proof attempt.

Proof Attempt? Suppose we pick one vertex uniformly at random from each V_i to obtain a set S. What is the probability that S is an independent set?

For a *single* edge $(u, v) \in V$, the probability that both u and v are also sampled is at most

$$\Pr\left(\text{both } u \text{ and } v \text{ sampled in } S\right) \leqslant \frac{1}{(36\Delta)^2},$$

since each vertex is marginally sampled with probability at most $1/6\Delta$ and if vertices are in two different V_i's they are independent, and if they are inside the same one, the probability is zero any way.

So, a single edge is "good"—i.e., does not violate independent set property—with a "good enough" probability, i.e., $1 - \Theta(1/\Delta^2)$. However, this is not good enough to do a *union bound* over all edges which are possibly $\Theta(n\Delta)$ many. So, this approach does not seem to be able to prove existence of the type of the independent set we want. □

While our proof attempt for Statement 2 failed, we can also see that something went wrong with our analysis attempt and not necessarily with the randomized process we had. See the example in Figure 1.

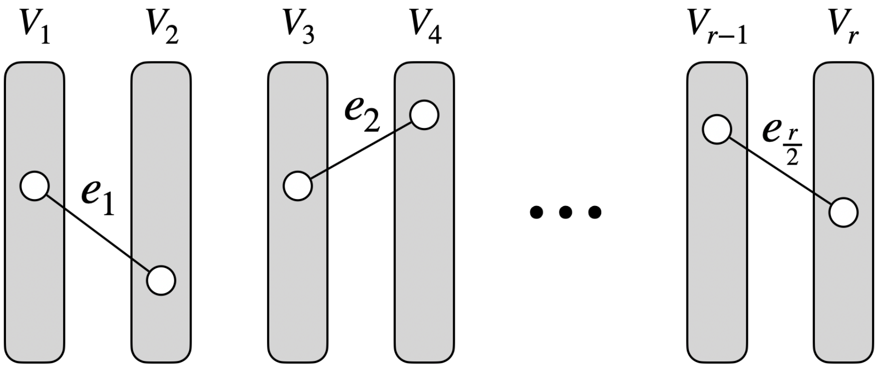

Figure 1: Consider the $r/2$ edges above that do not share any layers with each other. The probability that each of them is "bad", i.e., has both its endpoints sampled in S is $1/(6\Delta)^2$ as we calculated earlier. If we do a union bound, the probability that none of them is bad can only be bounded by $r/2 \cdot 1/(6\Delta^2)$ which is much larger than one even! (when $\Delta \ll n/\Delta$). However, at least in this case, the probability that all of these edges are good is at least $(1 - 1/(6\Delta^2))^{r/2} > 0$ since these edges are all independent.

As the above example suggests, we need a tool that can take into account the *independence* between the "bad" events better than just applying a union bound to them.

1.2 Coloring 3-Uniform Hypergraphs

Let us consider another example of the same flavor. A *hypergraph* $H = (V, E)$ over vertices V is similar to a graph except that it has *hyperedges* that can connect more than 2 vertices together. I.e., each hyperedge is a set of vertices $\{u_1, u_2, \ldots, u_k\} \subseteq V$. The degree of a vertex in a hypergraph is the number of hyperedges this vertex belongs to.

For any integer $k \geqslant 1$, a *k-uniform hypergraph* is a hypergraph wherein every hyperedge connects exactly k distinct vertices together. So, a 'normal' graph is a 2-uniform hypergraph. We can extend the definition of *vertex coloring* to hypergraphs by defining a coloring to be *proper* as before if it does not create any monochromatic hyperedge: this means that in this coloring, not *all* vertices in the same hyperedge should receive the same color (but it is okay if more than one vertex receives the same color as long as at least two distinct colors appear on vertices of the hyperedge). Notice that this means coloring hypergraphs is "easier" than graphs in the sense that we may need fewer colors even.

Recall that every graph with maximum degree Δ can be colored with $\Delta+1$ colors. One can try to extend this to other k-uniform hypergraphs for $k > 2$ as well. The following is an example:

Statement 3: Suppose we have a 3-uniform hypergraph $H = (V, E)$ with maximum degree Δ. Prove that H can always be properly vertex colored with $O(\sqrt{\Delta})$ colors.

Proof Attempt? Suppose we color the vertices of H randomly from a set of $4\sqrt{\Delta}$ colors. What is the probability that this is a proper coloring?

For a *single* hyperedge $(u, v, w) \in V$, the probability that *all* endpoints are colored the same is

$$\Pr(u, v, \text{ and } w \text{ are colored the same}) \leqslant \frac{1}{(4\sqrt{\Delta})^2} = \frac{1}{16\Delta}.$$

So, again, a single hyperedge is "good"—i.e., is not monochromatic—with a "good enough" probability, i.e., $1 - \Theta(1/\Delta)$. However, once again, this is not good enough to do a *union bound* over all hyperedge which are possibly $\Theta(n\Delta)$ many. So, this approach does not seem to work for us. □

Yet again, the problem is in the union bound step even though clearly two hyperedge which are "very far" from each other in the graph—specifically, do not share any common vertices—have nothing to do with each other.

2 Lovász Local Lemma

With the above examples in mind, we can now present the statement of the *Lovász Local Lemma (LLL)*, one of the strongest tools in the probabilistic methods literature, that addresses the above scenarios.

Given a collection of "bad" events $B_1, \ldots, B_n$, the **dependency graph** of these events is defined as follows: there is a vertex $i \in [n]$ for each event B_i. The neighbors of $i \in [n]$ are denoted by $N(i) \subseteq [n]$. The neighbors are defined such that for any $i \in [n]$,

$$\Pr\left(B_i \mid \wedge_{j \in J} B_j\right) = \Pr\left(B_i\right) \qquad \forall J \subseteq [n] \setminus (\{i\} \cup N(i))\,;$$

in words, the bad event B_i is mutually independent of any subset of other events except for its neighbors. In other words, regardless of what is happening to non-neighbors of i, the probability of B_i happening is not going to change. The simplest form of LLL is the following:

Theorem 3 (Symmetric LLL). *Suppose $B_1, \ldots, B_n$ are a collection of events. If:*

1. $\Pr\left(B_i\right) \leqslant p$ *for every* $i \in [n]$, *for some* $p \in (0, 1)$;
2. *and, the events admit a dependency graph with maximum degree* $d \geqslant 1$,

Then, as long as

$$e \cdot p \cdot (d+1) \leqslant 1, \qquad\qquad (e \sim 2.73\cdots \text{ is the natural number})$$

*the probability that **none of** $B_1, \ldots, B_n$ happens is strictly more than zero.*

Roughly speaking, Theorem 3 states that if the dependencies of our bad events to each other are small, say, d, then we can get away with bounding the probability of each bad event with roughly (but not exactly) $1/d$ (which is enough for union bound over d neighbors of a single event but not the entire event-sets) instead of $1/n$ which is needed for union bound.

Let us see how we can apply LLL to the scenarios of the previous section.

2.1 Application 1: Independent Sets in Bounded-Degree Graphs

Proposition 4. *Suppose we have a graph $G = (V, E)$ with maximum degree Δ with an arbitrary partitioning of vertices $V = V_1 \sqcup V_2 \sqcup \ldots \sqcup V_r$ in to r sets of size 6Δ. Then, there is always an independent set in G that picks exactly one vertex from each V_i.*

Proof. We do the same exact strategy as our 'proof attempt' for statement 2 in the previous section. Suppose we pick one vertex uniformly at random from each V_i to obtain a set S. For any edge $e = (u, v)$, define the bad event B_e to be the event that both u and v are sampled. As calculated earlier,

$$\Pr(B_e) \leqslant \frac{1}{36\Delta^2} := p.$$

What is the degree of the dependency graph? Well, the edge $e = (u, v)$ can depend on the choice of all other edges that go out of the same level as u or v (because they share the 'same source of randomness' as the one that determines the status of u and v being in S or not). Any other edge however is entirely independent. The total number of such edges is

$$2 \cdot (6\Delta \cdot \Delta) = 12\Delta^2.$$

Thus, the maximum degree d of the dependency graph on $\{B_e\}_{e\in E}$ is at most $12\Delta^2$, and hence

$$p \cdot (d+1) = \frac{1}{36\Delta^2} \cdot \left(12\Delta^2 + 1\right) < \frac{1}{e}.$$

We can now apply LLL (Theorem 3) and obtain that the probability that none of $\{B_e\}_{e\in E}$ happens is more than zero. This means, at least one choice of S leads to an independent set, which proves the original statement. □

2.2 Application 2: Coloring 3-Uniform Hypergraphs

Proposition 5. *Suppose we have a 3-uniform hypergraph $H = (V, E)$ with maximum degree Δ. Then, H can always be properly vertex colored with $\sqrt{3e \cdot \Delta}$ colors*

Proof. We do the same exact strategy as our 'proof attempt' for statement 3 in the previous section. Suppose we pick color each vertex uniformly at random from $\left\{1, \ldots, \sqrt{3e \cdot \Delta}\right\}$. For any hyperedge $e = (u, v, w)$, define the bad event B_e to be the event that all of u, v, w are colored the same. As calculated earlier,

$$\Pr(B_e) = \frac{1}{(\sqrt{3e \cdot \Delta})^2} = \frac{1}{3e \cdot \Delta} := p.$$

What is the degree of the dependency graph? The hyperedge $e = (u, v, w)$ can depend on the choice of all other hyperedges that share at least one vertex with e which are at most $3 \cdot (\Delta - 1)$ many (because they share the 'same source of randomness' as the one that determines the color of u, v, w). Any other hyperedge however is entirely independent. Thus, the maximum degree d of the dependency graph on $\{B_e\}_{e\in E}$ is at most 3Δ, and hence

$$p \cdot (d+1) = \frac{1}{3e\Delta} \cdot (e\Delta) \leqslant \frac{1}{e}.$$

We can now apply LLL (Theorem 3) and obtain that the probability that none of $\{B_e\}_{e\in E}$ happens is more than zero. This means, at least one choice of the coloring leads to a proper vertex coloring, which proves the original statement. □