Chapter 23

Probabilistic Method and Lovász Local Lemma

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

23.1 Motivation

We have used the **probabilistic method** so far in the course multiple times: to show an object with certain properties exists, 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.

23.1.1 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 n-vertex graph G with $n \geqslant R(r, s)$ 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 $R(r, s)$ even finite for every r and s, 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 no two of whom know each other (see **??**)?[a]
>
> [a]The connection is that the answer is $R(3, 3)$.

There is a somewhat easy upper bound for Ramsey numbers.

Proposition 23.1. *For every $r, s \geqslant 2$,*

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

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 give 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 an s-independent set. Consider any vertex v in G and its neighbors $N(v)$. We should have that

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

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

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

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

$$|N(v)| + |V \setminus (\{v\} \cup N(v))| = n - 1 \leqslant R(r-1, s) + R(r, s-1) - 2.$$

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 $R(2, 2) = 2$ 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 23.1 is that any graph G has either 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 23.2. *There are n-vertex graphs that contain neither a clique nor an independent set of size more than $2 \log n + c$ for some absolute constant $c > 0$.*

Proof. Let G be a random graph by picking 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,

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

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,

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

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. $\square$

This concludes another application of probabilistic method. An important remark is in order: in all the applications of the probabilistic method we have seen so far, we in fact managed to show the object in question exists 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 in which we are really searching for a "needle in a haystack"? This is going to be the topic of this chapter and the next one. Let us use a couple of simple examples to showcase such a scenario.

23.1.2 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,

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

thus, there is always an independent set with the desired size. $\square$

Statement 2: Suppose we have a graph $G = (V, E)$ with maximum degree Δ with an arbitrary partitioning of vertices $V = V_1 \sqcup V_2 \sqcup \ldots \sqcup V_r$ into 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 $(u, v) \in E$, the probability that both u and v are also sampled is at most

$$\Pr\left(\text{both } u \text{ and } v \text{ sampled in } S\right) \leqslant \frac{1}{(6\Delta)^2},$$

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 anyway.

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 23.1.

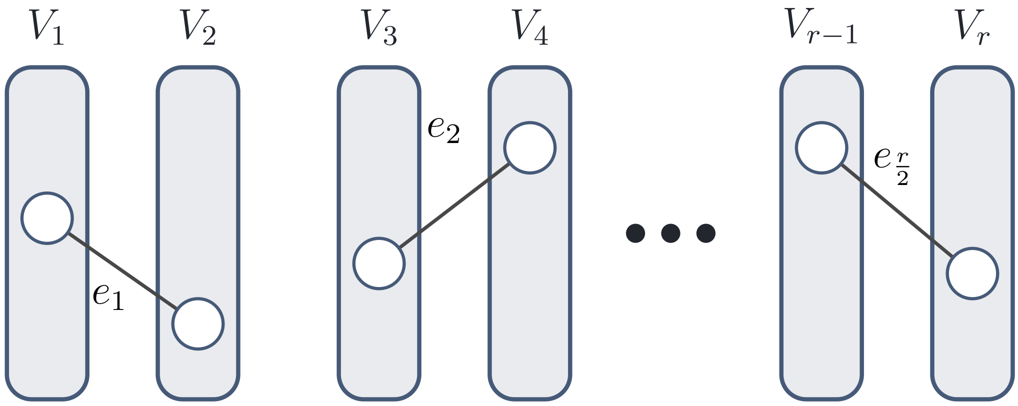

Figure 23.1: Consider the $r/2$ edges above that do not share any layers with each other. The probability that each of them is "bad", i.e., has both its endpoints sampled in S is $1/(6\Delta)^2$ as we calculated earlier. If we do a union bound, the probability that none of them is bad can only be bounded 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.

23.1.3 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 $(u, v, w) \in E$, the probability that *all* endpoints are same-colored 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"—is not monochromatic—with a "good enough" probability, i.e., $1-\Theta(1/\Delta)$. However, once again, this is not good enough to *union bound* over all hyperedges which are possibly $\Theta(n\Delta)$ many. So, this approach does not seem to work. $\square$

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

23.2 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 $i \in [n]$,

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

in words, the bad event B_i is mutually independent of any subset of other events (and their $2^{|J|}$ ways of them happening or not happening) 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 (see Figure 23.2).

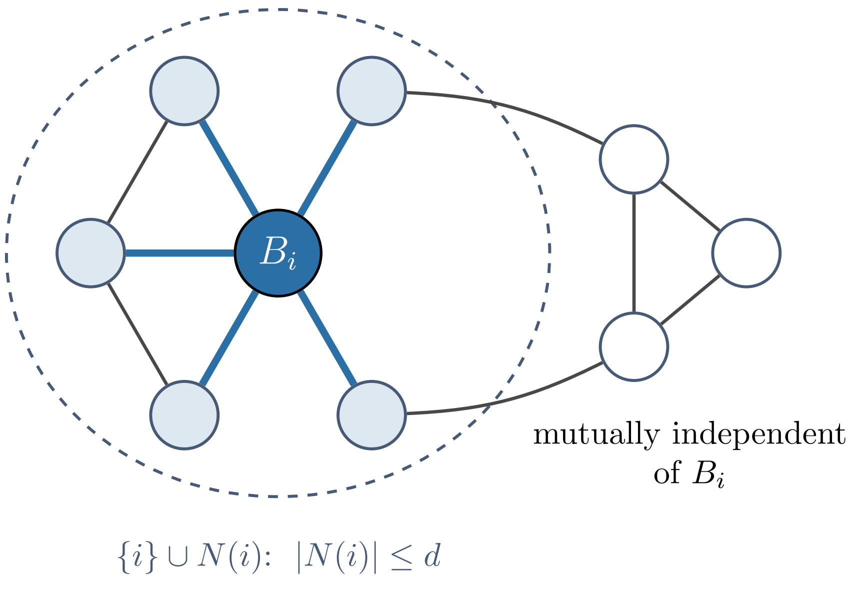

Figure 23.2: The dependency graph of the bad events $B_1, \ldots, B_n$. The event B_i (blue) is joined to the events in its neighborhood $N(i)$ and $|N(i)| \leqslant d$, the maximum degree of the graph. The event B_i is *mutually independent* of any subset of the remaining events, i.e. of every B_j with j outside $\{i\} \cup N(i)$: conditioning on any of them leaves $\Pr(B_i)$ unchanged. Note that a neighbour of B_i may itself depend on events far from B_i.

The simplest form of LLL is the following:

Theorem 23.3 (**Symmetric LLL**)**.** *Suppose $B_1, \ldots, B_n$ are a collection of events. If:*

1. $\Pr(B_i) \leqslant p$ *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

$$e \cdot p \cdot (d+1) \leqslant 1, \qquad (e \approx 2.718\cdots \text{ is the Euler's number})$$

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

Roughly speaking, Theorem 23.3 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.

23.2.1 Application 1: Independent Sets in Bounded-Degree Graphs

Proposition 23.4. *Suppose we have a graph $G = (V, E)$ with maximum degree Δ with an arbitrary partitioning of vertices $V = V_1 \sqcup V_2 \sqcup \ldots \sqcup V_r$ into 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 follow 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 23.3) 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. □

23.2.2 Application 2: Coloring 3-Uniform Hypergraphs

Proposition 23.5. *Suppose we have a 3-uniform hypergraph $H = (V, E)$ with maximum degree Δ. Then, H can always be properly vertex colored with $\sqrt{3e\Delta}$ colors.*

Proof. We follow the same exact strategy as our 'proof attempt' for statement 3 in the previous section. Suppose we color each vertex uniformly at random from $\sqrt{3e\Delta}$ colors. 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 $d \leqslant 3\Delta - 3$, and hence

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

We can now apply LLL (Theorem 23.3) 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. $\square$

Exercises

Exercise 23.1. In this exercise we pin down the smallest diagonal Ramsey number exactly, $R(3,3) = 6$. Improve the bounds in Proposition 23.1 in this special case to prove $R(3,3) \leqslant 6$. Then, show that $R(3,3) > 5$ by exhibiting a graph on 5 vertices that contains neither a triangle nor an independent set of size 3. These two parts together conclude that $R(3,3) = 6$.

Exercise 23.2. Let Φ be a Boolean formula in conjunctive normal form in which every clause is a disjunction of exactly k literals over k *distinct* variables (a k-*CNF* formula). Suppose that every clause shares at least one variable with at most d other clauses. Assign each variable independently and uniformly at random a value in $\{\text{TRUE}, \text{FALSE}\}$, and for each clause C let B_C be the (bad) event that C is left unsatisfied.

1. Show that $\Pr(B_C) = 2^{-k}$ for every clause C, and that $\{B_C\}_C$ admits a dependency graph of maximum degree at most d.
2. Use LLL (Theorem 23.3) to conclude that if $e \cdot (d+1) \leqslant 2^k$, then Φ is satisfiable.

Exercise 23.3 (♦). Proposition 23.2 uses a union bound to produce a graph with no clique or independent set of size k. Here we sharpen it with the *alteration method*, gaining a factor of $\Theta(k)$. As in Proposition 23.2, let G be a random graph on n vertices in which each edge is present independently with probability $1/2$, and let X be the number of subsets $S \subseteq V$ with $|S| = k$ that form either a clique or an independent set.

1. Show that $\mathbb{E}[X] = \binom{n}{k} \cdot 2^{1-\binom{k}{2}}$.
2. Fix an outcome of G with $X \leqslant \mathbb{E}[X]$ and delete one vertex from each such "bad" subset S. Argue that the remaining graph has no clique or independent set of size k, and conclude that
$$R(k,k) > n - \binom{n}{k} \cdot 2^{1-\binom{k}{2}} \qquad \text{for every } n \geqslant k.$$
3. By choosing n appropriately, deduce that $R(k,k) = \Omega\left(k \cdot 2^{k/2}\right)$, an improvement of a factor of $\Theta(k)$ over the bound implied by Proposition 23.2.