Stanford University	**Spring 2018**
Math 233A: Non-constructive methods in combinatorics	**Instructor: Jan Vondrák**
Lecture date: January 10, 2018	**Original scribe: Apoorva Khare**

Lecture 2. The Lovász Local Lemma

2.1 Introduction and motivation

We start with the *Lovász Local Lemma*, a fundamental tool of the "probabilistic method" and a prototypical non-constructive argument in combinatorics — proving that a certain object exists without showing what it looks like. Often in applying the probabilistic method, one is trying to show that it is possible to avoid "bad events" $\mathcal{E}_1, \ldots, \mathcal{E}_n$ with positive probability, or in other words,

$$\mathbb{P}\left[\bigcap_{i=1}^{n} \overline{\mathcal{E}}_i\right] > 0.$$

Here $\mathcal{E}_i$ are subsets of a probability space Ω (typically a finite set), and $\overline{\mathcal{E}}_i = \Omega \setminus \mathcal{E}_i$ denotes the complementary event for each i.

If $\sum_i \mathbb{P}[\mathcal{E}_i] < 1$, then the above inequality clearly follows, by the "union bound". However, this is often not a strong enough tool, since the sum $\sum_i \mathbb{P}[\mathcal{E}_i]$ may be much larger than 1 even if the events can be avoided.

A weaker constraint on the individual probabilities $\mathbb{P}[\mathcal{E}_i]$ is sufficient if the events $\mathcal{E}_i$ are also independent. In that case if $\mathbb{P}[\mathcal{E}_i] < 1$ for all i, then $\mathbb{P}[\cap_i \overline{\mathcal{E}}_i]$ is clearly positive. The Lovász Local Lemma is an effective refinement of this phenomenon, for events that do not have "too much (inter)dependency" – a notion that will be made precise presently. An additional attractive feature of the Local Lemma is that it does not place any restriction on the (finite) number of events $\mathcal{E}_i$.

2.2 Symmetric Local Lemma and application to hypergraph colorability

Before we state and prove the Local Lemma, we first present a prototypical application of the result, which serves to motivate it.

Example 2.1 (Hypergraph 2-coloring) *Given an integer $k \geq 2$, a k-*uniform hypergraph* $G = (V(G), E(G))$ consists of a finite set of nodes $V(G)$ and a collection of subsets $e_1, \ldots, e_n \subset V(G)$, each of size k, which are termed "edges" (or "hyper-edges"). We want to color each node in $V(G)$ either red or blue. Under what conditions can we guarantee that there is a coloring with no monochromatic edge, i.e., every edge contains both red and blue nodes? Such hypergraphs are said to be* 2-colorable.

Notice, if we color each node red or blue uniformly at random, then the event $\mathcal{E}_i$ that the ith edge is monochromatic has probability 2^{1-k}. Thus if the hypergraph G has less than 2^{k-1} edges, then by the union bound, the probability that there is at least one monochromatic edge is $< 2^{k-1} \cdot 2^{1-k} = 1$. It follows that G is 2-colorable.

However, this argument fails when G has $\geq 2^{k-1}$ edges. In this case, under what assumptions can we prove 2-colorability? One such assumption is that every edge intersects at most d other edges, for some d. Under such an assumption, we will show 2-colorability using the Lovász Local Lemma. (Interestingly, d will be comparable to 2^{k-1}.)

We now define the following notion of mutual independence.

Definition 2.2 *For all integers $n > 0$, define $[n] := \{1, \dots, n\}$. Given events $\mathcal{E}_1, \dots, \mathcal{E}_n \subset \Omega$ and a subset $J \subset [n]$, the event $\mathcal{E}_i$ is said to be* mutually independent of $\{\mathcal{E}_j : j \in J\}$ *if for all choices of disjoint subsets $J_1, J_2 \subset J$,*

$$\mathbb{P}\left[\mathcal{E}_i \cap \bigcap_{j_1 \in J_1} \mathcal{E}_{j_1} \cap \bigcap_{j_2 \in J_2} \overline{\mathcal{E}}_{j_2}\right] = \mathbb{P}[\mathcal{E}_i] \cdot \mathbb{P}\left[\bigcap_{j_1 \in J_1} \mathcal{E}_{j_1} \cap \bigcap_{j_2 \in J_2} \overline{\mathcal{E}}_{j_2}\right].$$

Equipped with this notion, we can state the first form of the Lovász Local Lemma, which will help answer the above question of 2-colorability for k-uniform hypergraphs.

Theorem 2.3 (Symmetric Lovász Local Lemma) *Suppose $p \in (0, 1)$, $d \geq 1$, and $\mathcal{E}_1, \dots, \mathcal{E}_n$ are events such that $\mathbb{P}[\mathcal{E}_i] \leq p$ for all i. If each $\mathcal{E}_i$ is mutually independent of all but d other events $\mathcal{E}_j$, and $ep(d + 1) \leq 1$, where $e = 2.71828\dots$ is Euler's number, then $\mathbb{P}\left[\cap_{i=1}^n \overline{\mathcal{E}}_i\right] > 0$.*

Remark 2.4 *In the above result, d is sometimes called the "dependence degree". The "local"-ness of the result has to do with the fact that assumptions depend only on d rather than n, the number of events.*

Before we prove the Local Lemma (in a more general form), let us see how it can be used to study 2-colorability for hypergraphs. In the setting of Example 2.1, suppose the hypergraph G has n edges, denoted by $e_1, \dots, e_n$. Let $\mathcal{E}_i$ denote the event that the edge e_i is monochromatic; as computed above, $p = 2^{1-k}$.

We now claim that "$d = d$", that is, the d in Example 2.1 is precisely the d in Theorem 2.3. Indeed, fix an edge e_i; now any conditioning (i.e., node-coloring) on the edges disjoint from e_i is independent of $\mathcal{E}_i$, since the node colors are i.i.d. Bernoulli random variables. Thus, the assumptions of the Symmetric Lovász Local Lemma are indeed satisfied, as long as

$$d + 1 \leq \frac{1}{ep} = \frac{2^{k-1}}{e}.$$

We stress again that this condition is independent of the number of edges in the hypergraph G.

2.3 (Asymmetric) Lovász Local Lemma: statement and proof

We now prove the Symmetric Lovász Local Lemma, i.e., Theorem 2.3. In fact we show a stronger, "asymmetric" version, and use it to prove the symmetric version. This will require the following useful concept.

Definition 2.5 *A (directed) graph $G = (V(G), E(G))$ is a* dependency (di)graph *on events $\mathcal{E}_1, \dots, \mathcal{E}_n$ if $V(G) = [n]$ and each event $\mathcal{E}_i$ is mutually independent of its non-neighbors $\{\mathcal{E}_j : j \neq i,\ (i, j) \notin E(G)\}$.*

Remark 2.6 *Most applications in the literature use the undirected version of the dependency graph; however, there are some applications that use the digraph structure. In such cases, given a directed edge (i, j), i is the source and j the target.*

We can now state the Lovász Local Lemma in its more general form.

Theorem 2.7 ((Asymmetric) Lovász Local Lemma) *Suppose G is a dependency (di)graph for events $\mathcal{E}_1, \ldots, \mathcal{E}_n$, and there exist $x_1, \ldots, x_n \in (0, 1)$ such that*

$$\mathbb{P}[\mathcal{E}_i] \leq x_i \prod_{(i,j)\in E(G)} (1 - x_j), \quad \forall i \in [n]. \tag{2.1}$$

Then,

$$\mathbb{P}\left[\bigcap_{i=1}^{n} \overline{\mathcal{E}}_i\right] \geq \prod_{i=1}^{n}(1 - x_i) > 0. \tag{2.2}$$

Remark 2.8 *Given a set of events $\mathcal{E}_i$, the choice of a dependency digraph G is not unique, nor is the choice of the parameters x_i. Rather, the "user" decides which dependency digraph G and parameters x_i to work with, in a given application. The dependency graph is often clear from the context (e.g. in the hypergraph colorability application above), although the choice of x_i might not be.*

Remark 2.9 *Theorem 2.7 is sharp when the $\mathcal{E}_i$ are independent, G is empty, and $x_i = \mathbb{P}[\mathcal{E}_i]$ $\forall i$.*

Before we show the Asymmetric Local Lemma, let us quickly see why it implies the Symmetric version. Indeed, if the hypotheses of Theorem 2.3 hold, set $x_i = \frac{1}{d+1}$ $\forall i$. Now the hypotheses imply that there is an undirected dependency graph G in which each node has degree at most d. Therefore,

$$\begin{aligned} x_i \prod_{(i,j)\in E(G)} (1 - x_j) = \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^{\deg(i)} &\geq \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^{d} \\ &\geq \frac{1}{d+1} \cdot \frac{1}{e} \geq \mathbb{P}[\mathcal{E}_i]. \end{aligned}$$

It follows by the Asymmetric Lovász Local Lemma that $\mathbb{P}[\cap_i \overline{\mathcal{E}}_i] > 0$.

Finally, we prove the Asymmetric Lovász Local Lemma.

Proof of Theorem 2.7. Given $S \subset [n]$, define

$$\overline{P}_S := \mathbb{P}\left[\bigcap_{i\in S} \overline{\mathcal{E}}_i\right], \qquad \overline{P}_\emptyset := 1.$$

The result follows once we show, by induction on $|S|$, that for all $S \subset [n]$ and $a \in S$,

$$\frac{\overline{P}_S}{\overline{P}_{S\setminus\{a\}}} \geq 1 - x_a. \tag{2.3}$$

More precisely, we will show by induction on $|S|$ that

$$\overline{P}_S \geq (1 - x_a)\overline{P}_{S \setminus \{a\}} > 0.$$

Indeed, this yields the result, because applying the inequality to $S = [n]$, then $[n-1]$, and so on, yields:

$$\mathbb{P}\left[\bigcap_{i=1}^{n} \overline{\mathcal{E}}_i\right] = \overline{P}_{[n]} \geq (1 - x_n)\overline{P}_{[n-1]} = (1 - x_n)(1 - x_{n-1})\overline{P}_{[n-2]} \geq \cdots \geq \prod_{i=1}^{n}(1 - x_i) > 0,$$

as desired.

Thus it remains to prove (2.3). The base case is when $S = \{a\}$ is a singleton. In this case,

$$\frac{\overline{P}_{\{a\}}}{\overline{P}_{\emptyset}} = \mathbb{P}[\overline{\mathcal{E}}_a] \geq 1 - x_a \prod_{(a,j) \in E(G)} (1 - x_j) \geq 1 - x_a,$$

proving the assertion. Now suppose (2.3) holds for all subsets $S' \subset [n]$ with size at most k, and say $S \subset [n]$ has size $k+1$. To proceed further, let us define the neighborhood of $a \in S$, as well as its "closure", via:

$$\Gamma(a) := \{j \in V(G) : (a,j) \in E(G)\}, \qquad \Gamma^+(a) := \{a\} \cup \Gamma(a). \tag{2.4}$$

Now fix $a \in S$, and compute:

$$\begin{aligned}\overline{P}_S = \mathbb{P}\left[\bigcap_{i \in S} \overline{\mathcal{E}}_i\right] = \mathbb{P}\left[\bigcap_{i \in S \setminus \{a\}} \overline{\mathcal{E}}_i\right] - \mathbb{P}\left[\mathcal{E}_a \cap \bigcap_{i \in S \setminus \{a\}} \overline{\mathcal{E}}_i\right] &\geq \mathbb{P}\left[\bigcap_{i \in S \setminus \{a\}} \overline{\mathcal{E}}_i\right] - \mathbb{P}\left[\mathcal{E}_a \cap \bigcap_{i \in S \setminus \Gamma^+(a)} \overline{\mathcal{E}}_i\right] \\ &= \overline{P}_{S \setminus \{a\}} - \mathbb{P}[\mathcal{E}_a]\overline{P}_{S \setminus \Gamma^+(a)},\end{aligned}$$

where the first equality and the inequality are straightforward, and the final equality follows from the mutual independence of $\mathcal{E}_a$ and $\{\mathcal{E}_i : i \notin \Gamma^+(a)\}$. From this computation it follows that

$$\frac{\overline{P}_S}{\overline{P}_{S \setminus \{a\}}} \geq 1 - \mathbb{P}[\mathcal{E}_a] \frac{\overline{P}_{S \setminus \Gamma^+(a)}}{\overline{P}_{S \setminus \{a\}}},$$

where $\overline{P}_{S \setminus \{a\}} > 0$ by the induction hypothesis. Now say $\Gamma(a) \cap S = \{b_1, \ldots, b_d\}$ for some $d \geq 0$, and write the fraction on the right-hand side as a telescoping product:

$$\frac{\overline{P}_{S \setminus \Gamma^+(a)}}{\overline{P}_{S \setminus \{a\}}} = \frac{\overline{P}_{S \setminus \{a, b_1\}}}{\overline{P}_{S \setminus \{a\}}} \frac{\overline{P}_{S \setminus \{a, b_1, b_2\}}}{\overline{P}_{S \setminus \{a, b_1\}}} \cdots \frac{\overline{P}_{S \setminus \{a, b_1, \ldots, b_d\}}}{\overline{P}_{S \setminus \{a, b_1, \ldots, b_{d-1}\}}},$$

where all terms on the right-hand side are strictly positive by the induction hypothesis. By the same hypothesis, each ratio on the right-hand side is bounded above by $\dfrac{1}{1 - x_{b_i}}$. Therefore,

$$\frac{\overline{P}_{S \setminus \Gamma^+(a)}}{\overline{P}_{S \setminus \{a\}}} \leq \frac{1}{1 - x_{b_1}} \cdots \frac{1}{1 - x_{b_d}}.$$

Recalling that by assumption $\mathbb{P}[\mathcal{E}_a] \leq x_a \prod_{b \in \Gamma(a)}(1 - x_b)$, it follows that

$$\frac{\overline{P}_S}{\overline{P}_{S \setminus \{a\}}} \geq 1 - x_a \prod_{b \in \Gamma(a)} (1 - x_b) \prod_{c \in \Gamma(a) \cap S} \frac{1}{1 - x_c} \geq 1 - x_a > 0.$$

This shows (2.3), and with it, the Lovász Local Lemma. $\square$

3 Application to Ramsey numbers

As an application of the asymmetric LLL (Theorem 2.8) we will prove a lower bound for Ramsey numbers $R(3, l)$.

For integers $k, l \geq 2$ the Ramsey number $R(k, l)$ is the minimal positive integer N, such that for every edge-coloring of the complete graph K_N there exists a red clique of size k or a blue clique of size l. By induction on k and l, one can easily prove an upper bound

$$R(k, l) \leq \binom{k+l-2}{k-1}.$$

We will now consider the case $k = 3$ and $l \geq 3$. Then the above inequality reads $R(3, l) \leq \frac{1}{2}l(l+1)$. We would like obtain a lower bound for $R(3, l)$ from the LLL.

In order to do so, we consider a complete graph on n vertices and a random coloring of its edges with the colors red and blue: We color each edge red with probability p and blue with probability $1-p$; independently for all edges. Here the number of vertices n and the probability p will be specified later (depending on the value of l). Our goal is to obtain with positive probability a coloring without a red triangle and without a blue l-clique, since this would establish the lower bound $R(3, l) > n$.

For each 3-element subset T of the vertex set let $\mathcal{A}_T$ be the event that the three vertices in T form a red triangle. Note that for each such T we have $\mathbb{P}(\mathcal{A}_T) = p^3$ and the number of these events $\mathcal{A}_T$ is $\binom{n}{3}$. Furthermore, for each l-element subset S of the vertex set let $\mathcal{B}_S$ be the event that the l vertices in S form a blue clique. Note that for each such S we have

$$\mathbb{P}(\mathcal{B}_S) = (1-p)^{\binom{l}{2}}$$

and the number of these events $\mathcal{B}_S$ is $\binom{n}{l}$.

Let us now define a dependency graph for these events. We join two events of the form $\mathcal{A}_T$ or $\mathcal{B}_S$, if the corresponding sets S or T intersect in at least two vertices (i.e. if they share an edge). It is clear that an event $\mathcal{A}_T$ or $\mathcal{B}_S$ is mutually independent of all events with which it does not share an edge. Hence the graph we just defined is indeed a dependency graph.

Let us now bound the degrees in this graph. First consider a vertex $\mathcal{A}_T$. It is connected to at most $3n$ events $\mathcal{A}_{T'}$. (In order to be connected, the triangles T and T' need to share an edge. T has 3 edges and for each of them there are at most n choices of a third vertex to form a triangle T'). Also, trivially $\mathcal{A}_T$ is connected to at most $\binom{n}{l}$ events $\mathcal{B}_S$. (We use a trivial bound here since the actual number of dependencies is not much less.) Now, let us consider a vertex $\mathcal{B}_S$. It is connected to at most $\binom{l}{2}n$ events $\mathcal{A}_T$ (In order to have a connection, the sets S and T need to intersect in at least 2 elements. There are $\binom{l}{2}$ choices of two elements in S and for each of them at most n choices for the third vertex in T). Also, trivially $\mathcal{B}_S$ is connected to at most $\binom{n}{l}$ events $\mathcal{B}_{S'}$.

So in order to apply the LLL we need to find positive real numbers $x, y \in (0, 1)$ with

$$p^3 = \mathbb{P}(\mathcal{A}_T) \leq x(1-x)^{3n}(1-y)^{\binom{n}{l}}$$

and

$$(1-p)^{\binom{l}{2}} = \mathbb{P}(\mathcal{B}_S) \leq y(1-x)^{\binom{l}{2}n}(1-y)^{\binom{n}{l}}.$$

Here x and y shall be the values x_i in Theorem 2.8. Note that we choose the same value x for all events $\mathcal{A}_T$ (since the situation is symmetric for all choices of T) and the same value y for all events $\mathcal{B}_S$ (since the situation is symmetric for all choices of S).

Let us assume that for some value of n we can find positive real numbers $p, x, y \in (0, 1)$ such that both of the above inequalities are fulfilled. Then the LLL yields that with positive probability none of the events $\mathcal{A}_T$ and $\mathcal{B}_S$ happen. This means that there is a coloring of the edges of a complete graph on n vertices such that there is no red triangle and no blue l-clique. Therefore we have $R(3, l) > n$, if we can find $p, x, y \in (0, 1)$ as above.

Let us now try to find such $p, x, y \in (0, 1)$ for n as large as we can. We guess that $y = \frac{1}{\binom{n}{l}}$ would be a smart choice, because then the term $(1-y)^{\binom{n}{l}}$ is roughly constant (around e^{-1}). Furthermore, we observe that p and x need to fulfill the following inequalities:

$$p^3 \leq x(1-x)^{3n}(1-y)^{\binom{n}{l}} \leq x$$

and

$$e^{-p\binom{l}{2}} \approx (1-p)^{\binom{l}{2}} \leq y(1-x)^{\binom{l}{2}n}(1-y)^{\binom{n}{l}} \leq (1-x)^{\binom{l}{2}n} \approx e^{-xn\binom{l}{2}}.$$

Hence we need $p \geq xn \geq p^3 n$ and therefore $p \leq \frac{1}{\sqrt{n}}$. Finally, to guess the dependence of n on l we note

$$e^{-p\binom{l}{2}} \approx (1-p)^{\binom{l}{2}} \leq y(1-x)^{\binom{l}{2}n}(1-y)^{\binom{n}{l}} \leq y = \frac{1}{\binom{n}{l}} \approx e^{-l\log n},$$

hence $pl^2 \geq p\binom{l}{2} \geq l \log n$ and therefore $l \geq \frac{1}{p}\log n \geq \sqrt{n}\log n$.

Motivated by this we assume $l \geq 20\sqrt{n}\log n$ and choose $y = \frac{1}{\binom{n}{l}}$, $x = \frac{1}{9n^{3/2}}$ and $p = \frac{1}{3\sqrt{n}}$ (the constants here are not really important, they are just chosen in such a way that the inequalities work out). Then we have

$$(1-y)^{\binom{n}{l}} = \left(1 - \frac{1}{\binom{n}{l}}\right)^{\binom{n}{l}} \geq e^{-1.01}$$

if n is sufficiently large (since $(1-\frac{1}{m})^m \to e^{-1}$ as $m \to \infty$). Furthermore for sufficiently large n we also have

$$(1-x)^{3n} = \left(1 - \frac{1}{9n^{3/2}}\right)^{3n} \geq 1 - \frac{1}{3\sqrt{n}} \geq e^{-0.01}$$

for sufficiently large n. Thus,

$$p^3 = \frac{1}{27n^{3/2}} \leq \frac{1}{9n^{3/2}} \cdot e^{-1.02} \leq x(1-x)^{3n}(1-y)^{\binom{n}{l}},$$

which establishes the first desired inequality.

For the second inequality note that for sufficiently small $h > 0$ we have $1 - h \geq e^{-2h}$. So for sufficiently large n we get

$$(1-x)^{\binom{l}{2}n} \geq e^{-2xn\binom{l}{2}} = e^{-\frac{2}{9\sqrt{n}}\binom{l}{2}}.$$

Furthermore, using $l \geq 20\sqrt{n}\log n$,

$$y = \frac{1}{\binom{n}{l}} \geq \frac{1}{n^l} = e^{-l\log n} \geq e^{-l^2\frac{1}{20\sqrt{n}}} \geq e^{-l(l-1)\frac{1}{19\sqrt{n}}} \geq e^{-l(l-1)\frac{1}{18\sqrt{n}}+1.01} = e^{-\frac{1}{9\sqrt{n}}\binom{l}{2}+1.01}.$$

Hence

$$(1-p)^{\binom{l}{2}} \le e^{-p\binom{l}{2}} = e^{-\frac{1}{3\sqrt{n}}\binom{l}{2}} = e^{-\frac{1}{9\sqrt{n}}\binom{l}{2}+1.01} e^{-\frac{2}{9\sqrt{n}}\binom{l}{2}} e^{-1.01} \le y(1-x)^{\binom{l}{2}n}(1-y)^{\binom{n}{l}},$$

which verifies the second desired inequality.

Thus, for $l \ge 20\sqrt{n}\log n$ we can find $p, x, y \in (0,1)$ with the two desired inequalities. This implies (by the LLL, as described above) that $R(3,l) > n$ whenever $l \ge 20\sqrt{n}\log n$ and n sufficiently large.

Note that $n \le \frac{l^2}{(40\log l)^2}$ implies

$$20\sqrt{n}\log n \le 20\frac{l}{40\log l}\log(l^2) = l.$$

Therefore we have $R(3,l) \ge \frac{l^2}{(40\log l)^2}$ for all sufficiently large l. This proves the following theorem.

Theorem 3.1 *There is a positive constant $c > 0$, such that $R(3,l) \ge c\frac{l^2}{\log^2 l}$ for all $l \ge 3$.*

So we have proved $c\frac{l^2}{\log^2 l} \le R(3,l) \le \frac{1}{2}l(l+1)$. We conclude this section by remarking that the true answer is $R(3,l) = \Theta\left(\frac{l^2}{\log l}\right)$. (The upper bound was proved by Ajtai, Komlós, Szemerédi and the lower bound by Jeong Han Kim.)