CS-3334: Advanced Combinatorics **Fall 2022**

Lecture 8: November 1

Lecturer: Kuan Yang *Scribe: Weihao Zhu*

8.1 General Form of the Lovász Local Lemma

Recall that last week we have introduced the definition of the dependency graph: for any event A_i, it is independent from $\{A_j : j \neq i, j \notin N(i)\}$. Also, we have introduced the symmetric form of the Lovász Local Lemma as follows.

Theorem 8.1 (Lovász Local Lemma, symmetric version) *Let $A_1, \ldots, A_n$ be events with $\mathbf{Pr}[A_i] \leq p$. Suppose that each A_i is independent from all other A_j except at most d of them. If $ep(d+1) \leq 1$, then $\mathbf{Pr}[\bigcap \overline{A}_i] > 0$.*

In this section, we will introduce the asymmetric/general form of the Lovász Local Lemma as follows.

Theorem 8.2 (Lovász Local Lemma, asymmetric/general version) *Let $A_1, \ldots, A_n$ be events and A_i is independent from $\{A_j : j \neq i, j \notin N(i)\}$. If there exists $x_1, \ldots, x_n \in [0, 1)$ such that for any $1 \leq i \leq n$,*

$$\mathbf{Pr}[A_i] \leq x_i \cdot \prod_{j \in N(i)} (1 - x_j),$$

then

$$\mathbf{Pr}[\bigcap \overline{A}_i] \geq \prod_{i=1}^{n} (1 - x_i).$$

Proof: We claim that for any $i \notin S \subseteq [n]$, we have

$$\mathbf{Pr}[A_i | \bigcap_{j \in S} \overline{A}_j] \leq x_i.$$

If it holds, then

$$\mathbf{Pr}[\bigcap \overline{A}_i] = \mathbf{Pr}[\overline{A}_i] \cdot \mathbf{Pr}[\overline{A}_2 | \overline{A}_1] \ldots \geq \prod_{i=1}^{n} (1 - x_i),$$

which completes the proof.

Now, let's prove our claim by induction on the size of S. Our claim is trivially true when $|S| = 0$.

We assume that for any set S' of which size is less than S, the claim always holds. Let's consider the set S. For $i \notin S$, let $S_1 = S \bigcap N(i)$ and $S_2 = S \backslash S_1$. Then we have

$$\mathbf{Pr}[A_i | \bigcap_{j \in S} \overline{A}_j] = \frac{\mathbf{Pr}[A_i \bigcap (\bigcap_{j \in S_1} \overline{A}_j) | \bigcap_{j \in S_2} \overline{A}_j]}{\mathbf{Pr}[\bigcap_{j \in S_1} \overline{A}_j | \bigcap_{j \in S_2} \overline{A}_j]} := \frac{\alpha}{\beta}.$$

Note that

$$\alpha \le \mathbf{Pr}[A_i | \bigcap_{j \in S_2} \overline{A}_j] = \mathbf{Pr}[A_i] \le x_i \cdot \prod_{j \in N(i)} (1 - x_j).$$

Also, let $S_1 = \{t_1, \ldots, t_r\}$. We have

$$\begin{aligned}
\beta &= \prod_{k=1}^{r} \mathbf{Pr}[\overline{A}_{t_k} | (\bigcap_{\ell=1}^{k-1} \overline{A}_{t_\ell}) \bigcap (\bigcap_{j \in S_2} \overline{A}_j)] \\
&\ge (1 - x_{t_1}) \ldots (1 - x_{t_r}) \qquad \text{(by induction hypothesis)} \\
&\ge \prod_{j \in N(i)} (1 - x_j).
\end{aligned}$$

Therefore, $\frac{\alpha}{\beta} \le x_i$, which completes the proof. ■

Remark 1. To see the symmetric form, set $x_i = \frac{1}{d+1} < 1$ for all $1 \le i \le n$. Then,

$$x_i \prod_{j \in N(i)} (1 - x_j) \ge \frac{1}{d+1} (1 - \frac{1}{d+1})^d > \frac{1}{e(d+1)} \ge p.$$

Remark 2. In 1985, Shearer proved that the constant e is best possible.

Let's introduce a simple application of the Lovász Local Lemma. Consider a k-SAT formula:

$$\varphi = c_1 \wedge c_2 \wedge \ldots \wedge c_m$$

of which each clause has exactly k literals. Suppose that each variable appears in at most d clauses, then based on the Lovász Local Lemma, we can claim that there exists a satisfying assignment when $e \cdot kd \cdot 2^{-k} \le 1$.

However, the Lovász Local Lemma only tells us the existence of such assignment. Can we find such a satisfying assignment in polynomial time?

8.2 Algorithmic Lovász Local Lemma

In this section, we will discuss the algorithmic Lovász Local Lemma, which was awarded 2020 Gödel Prize.

Let's start from a computationally hard example. Let $q = 2^k$ and $f : [q] \to [q]$ be a bijection. Let $y \in [q]$ be a fixed element. We sample $x \in [q]$ uniformly at random. Define A_i as the bad event that $f(x)$ and y disagree at the i-th bit. All A_i's are mutually independent, so the Lovász Local Lemma applies. This means that there exists x such that $f(x) = y$. However, this conclusion is meaningless as we have already known that f is a bijection. Also, finding such an x may be extremely hard. (For instance, consider the problem of discrete logarithm: $f : \mathbf{F}_q \to \mathbf{F}_q = g^x$.)

The example above shows that it's sometimes hard for us to find an assignment such that no "bad events" occur if we add no constraints to events. For simplicity, we only talk about random variable models, where each event only depends on some variables.

Robin Moser and Gábor Tardos gave the following algorithm:

- Step 1: Initialize each variable a random value independently.

- Step 2: While some bad event A_i occurs (if several bad events occur simultaneously, pick A_i arbitrarily), re-sample all variables that A_i depends on. Denote by $vbl(A_i)$ the set of these variables.

In 2010, they proved the following theorem.

Theorem 8.3 (Robin Moser & Gábor Tardos, 2010) *If the condition of Lovász Local Lemma holds, then Moser-Tardos algorithm returns an assignment that no bad event occurs in expected linear time. In particular, the expected rounds of re-sampling is no more than*

$$E = \sum_{i=1}^{n} \frac{x_i}{1 - x_i}.$$

Proof: Let the excution log L be the sequence of A_i's that are picked in step 2. $|L|$ may be infinite, but we claim that $\mathbf{E}[|L|] \leq E$.

Construct witness trees as follows for each time $t \leq |L|$. Let $L = (A_{l_1}, A_{l_2}, \ldots, A_{l_t}, \ldots)$. Read prefix $A_{l_t}, \ldots, A_{l_1}$.

- Let the root of the witness tree $T(t)$ be a vertex labelled with l_t.
- For $t' = t - 1, \ldots, 1$:
  - If none of the events corresponding to vertices in T shares variables with $A_{l_{t'}}$, continue.
  - Otherwise, find a deepest node v such that $vbl(A_{[v]}) \cap vbl(A_{l_{t'}}) \neq \emptyset$ and add a vertex labelled with $l_{t'}$ as v's child.

The following picture demonstrates a valid witness tree for better understanding.

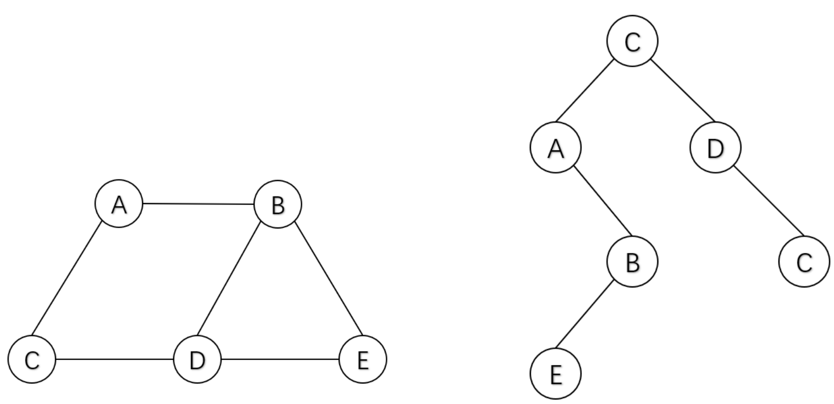

Figure 8.1: The left picture is the dependency graph of events, while the right one is a valid witness tree when $L = (C, E, B, D, A, B, B, E, C)$.

Now, consider properties of the witness trees. For convenience, denote by $[v]$ the label assigned to vertex v.

- $T(t_1) \neq T(t_2)$ for different times $t_1 \neq t_2$.

  If $A_{l_{t_1}} \neq A_{l_{t_2}}$, then the roots of $T(t_1)$ and $T(t_2)$ have different labels. If $A_{l_{t_1}} = A_{l_{t_2}} = A_r$, then label r appears different times in $T(t_1)$ and $T(t_2)$, which implies that $T(t_1) \neq T(t_2)$.

- For any $T = T(t)$ and $u, v \in T$ of the same depth, $vbl(A_{[u]}) \cap vbl(A_{[v]}) = \emptyset$.

The first property implies that

$$\mathbf{E}[|L|] = \sum_T \mathbf{E}[X_T] = \sum_T \mathbf{Pr}[T \text{ is a witness tree}].$$

We claim that

$$\mathbf{Pr}[T \text{ appears as a witness tree for some time } t] \leq \prod_{v \in T} \mathbf{Pr}[A_{[v]}].$$

In order to illustrate the above inequality more clearly, we give two simple examples. Consider T is a tree with one single vertex A as its root. If T is a valid witness tree for some time t, then A happens at the beginning, which implies that $\mathbf{Pr}[T$ appears as a witness tree for some time $t] \leq \mathbf{Pr}[A]$. If T is a tree with two vertices and A is its root while B is a child of A, then clearly B happens at the beginning. After re-sampling $vbl(B)$, event A occurs. Therefore, the probability that T is a valid witness tree is no larger than $\mathbf{Pr}[B] \cdot \mathbf{Pr}[A]$.

Now, we start to prove our claim strictly. In general, consider the reverse BFS order of T: $v_1, v_2, \ldots$. Assume for each variable, we have an infinite list of values, of which each is independently sampled and then fixed. When simulating the Moser-Tardos algorithm or checking $A_{[v_1]}, A_{[v_2]}, \ldots$ independently, we look up the value table of each variable instead of sampling. We prove our claim by induction on the depth from bottom to top.

For each $v \in T$ and any $u \in T$ with $vbl(A_{[u]}) \cap vbl(A_{[v]}) \neq \emptyset$, u is deeper than v if and only if $A_{[u]}$ appears before $A_{[v]}$ in the execution log. For any $z \in vbl(A_{[v]})$, let $n_{z,v}$ be the number of u's before v such that $z \in vbl(A_{[u]})$. In the simulation of the Moser-Tardos algorithm, when checking whether A_v occurs, look up the $(n_{z,v} + 1)$-th value of variable z. When checking the reverse BFS order sequence $A_{[v_1]}, A_{[v_2]}, \ldots$, we also look up the $(n_{z,v} + 1)$-th value of variable z at the time checking $A_{[v]}$. So the event that T is valid has the same distribution as the sequence occur. Namely,

$$\mathbf{Pr}[T \text{ is valid for some time } t] = \prod_{v \in T} \mathbf{Pr}[A_{[v]}].$$

Certainly,

$$\mathbf{Pr}[T \text{ is a witness tree } T(t)] \leq \prod_{v \in T} \mathbf{Pr}[A_{[v]}],$$

which proves our claim.

Let W be the set of all possible witness trees.

$$\mathbf{E}[|L|] = \sum_{T \in W} \mathbf{Pr}[T = T(t) \text{ for some } t] \leq \sum_{T \in W} \prod_{v \in T} \mathbf{Pr}[A_{[v]}].$$

If $T \in W$, then T has the following properties:

- T is finite;
- For any $u \to v$ in T, $A_{[u]}$ and $A_{[v]}$ overlap;
- For any $u, v \in T$ have the same depth, $A_{[u]}$ and $A_{[v]}$ are disjoint.

Let W' be the set of trees that only satisfy the second property. Let W'_B be the set of trees in W' and rooted at event B. We generate trees in W'_B by a random process (Galton-Watson process):

- Let B be the root of the tree.
- For any vertex v, we find all its "potential" children $N^+(v) = N([v]) \cup \{[v]\}$ whose variables overlap with $vbl(A_{[v]})$.
- For each "potential" child A_i, add a vertex labelled with i as the child of vertex v in the tree with probability x_i (x_i is the value corresponding to event A_i in the statement of the local lemma) and call it an alive children of v.

Let $D(v)$ be the set of alive children of vertex v. Let P_T be the probability that Galton-Watson process generates T. Thus, we have

$$\begin{aligned}
P_T &= \frac{1}{x_B} \prod_{v \in T} x_{[v]} \prod_{v \in T} \prod_{k \in N^+(v) \setminus D(v)} (1 - x_k) \\
&= \frac{1 - x_B}{x_B} \prod_{v \in T} \frac{x_{[v]}}{1 - x_{[v]}} \prod_{k \in N^+(v)} (1 - x_k) \\
&= \frac{1 - x_B}{x_B} \prod_{v \in T} x_{[v]} \prod_{k \in N(v)} (1 - x_k) \\
&\geq \frac{1 - x_B}{x_B} \prod_{v \in T} \mathbf{Pr}[A_v].
\end{aligned}$$

Clearly, $\sum_{T \in W'_B} P_T = 1$. Therefore,

$$\sum_{T \in W_B} \prod_{v \in T} \mathbf{Pr}[A_{[v]}] \leq \sum_{T \in W'_B} P_T \cdot \frac{x_B}{1 - x_B} = \frac{x_B}{1 - x_B},$$

which implies that

$$\mathbf{E}[|L|] \leq \sum_{i=1}^{n} \frac{x_i}{1 - x_i}.$$

This completes the whole proof. ■

8.3 Several Examples

In this section, we will introduce several classical applications of the Lovász Local Lemma.

8.3.1 Ramsey Number, Revisit

Theorem 8.4 (Spencer, 1977) *If*

$$e\left(\binom{k}{2}\binom{n}{k-2} + 1\right) \cdot 2^{1-\binom{k}{2}} < 1,$$

then $R(k, k) > n$.

Proof: Color K_n randomly. For any set of vertices S of size k, let E_S be the event that S induces a monochromatic K_k. Thus, $\mathbf{Pr}[E_S] = 2^{1-\binom{k}{2}}$.

For any k-vertex sets S, E_S is independent from all E_T where $|S \cap T| < 2$. Therefore, the maximal degree of the dependency graph is at most $\binom{k}{2} \cdot \binom{n}{k-2}$. Then, the Lovász Local Lemma applies. ■

Remark. Optimizating the choice of n, it gives the best bound so far

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

Recall that by the union bound we obtain $R(k,k) > (1/(e\sqrt{2}) + o(1)) \cdot k \cdot 2^{k/2}$, and by the alteration method we obtain $R(k,k) > (1/e + o(1)) \cdot k \cdot 2^{k/2}$. The Lovász Local Lemma does not improve much.

Let $K = \binom{n}{k}$ be the number of all events, then $d = |N(S)| \approx K^{1-O(1/k)}$. There are so many "dependencies", so the Lovász Local Lemma does not work well.

Now, let's consider $R(k,3)$. Let p be a fixed parameter to be determined later. For each vertex, color it 0 with probability p, and 1 with probability $1-p$. Let S, T be two vertex sets where $|S| = 3$ and $|T| = k$. Define A_S as the event that S forms a monochromatic K_3 with color 0 and B_T as the event that T forms a monochromatic K_k with color 1. Clearly,

$$\mathbf{Pr}[A_S] = p^3, \ \mathbf{Pr}[B_T] = (1-p)^{\binom{k}{2}},$$

and two event are adjacent in the dependency graph if the intersection of their corresponding subsets has size at least 2.

For A_S, there exists at most $3(n-3)$ S' such that $A_S \sim A_{S'}$ and at most $\binom{n}{k}$ T' such that $A_S \sim B_{T'}$. For B_T, there exists at most $\binom{k}{2}(n-2) < \frac{k^2 n}{2}$ S' such that $B_T \sim A_{S'}$ and at most $\binom{n}{k}$ T' such that $B_T \sim B_{T'}$.

Apply the Lovász Local Lemma, if there exists p, x, y such that

$$\begin{cases} p^3 \le x(1-x)^{3n}(1-y)^{\binom{n}{k}} \\ (1-p)^{\binom{k}{2}} \le y(1-x)^{k^2 n/2}(1-y)^{\binom{n}{k}} \end{cases},$$

then $R(k,3) > n$.

By setting $p = c_1 \cdot n^{-1/2}$, $k = c_2 \cdot n^{1/2} \log n$, $x = c_3 \cdot n^{-3/2}$ and $y = c_4/\binom{n}{k}$, we have $R(k,3) > c_5 \cdot k^2/\log^2 k$. The best known lower bound is $c_6 \cdot k^2/\log k$. Analogously, $R(k,4) > k^{\frac{5}{2}+o(1)}$, which is better than any known result without the Lovász Local Lemma.

8.3.2 Large Independent Sets from Partition

Previously, we have introduced the Caro-Wei inequality, where we learned how to find an independent set of size at least $\frac{|V|}{\Delta+1}$ when given a graph with maximal degree Δ. Today, we will show that there exists a large independent set from any "good" partition.

Theorem 8.5 *Let $G = (V,E)$ be a graph with maximal degree at most Δ. $V = V_1 \cup \ldots \cup V_r$ is a parition where $|V_i| \ge 2e\Delta$ for any $1 \le i \le r$. Then, there exists an independent set which contains a vertex from each V_i.*

Proof: Let $k = \lceil 2e\Delta \rceil$ and assume that $|V_i| = k$ for all $1 \le i \le r$. Pick $v_i \in V_i$ u.a.r. For any edge $e \in E$, let B_e be the event that both of its endpoints are chosen. Thus, $\mathbf{Pr}[B_e] \le \frac{1}{k^2}$. In the dependency graph,

$B_e \sim B_f$ if there exists V_i that intersects both e and f. Therefore, the maximal degree of the dependency graph $d \leq 2k\Delta$. Then, the Lovász Local Lemma applies. ■

Remark. Some choices of bad events are better than others. If we define $A_{i,j} = \{v_i \sim v_j\}$ for any $1 \leq i < j \leq r$, then $\mathbf{Pr}[A_{i,j}] \leq \frac{\Delta}{k}$. In the dependency graph, $A_{i,j} \sim A_{k,l}$ if $\{i,j\} \cap \{k,l\} \neq \emptyset$. The maximal degree of the dependency graph is $d \leq 2k\Delta$. However, this upper bound is still too large.

8.3.3 Directed Cycles of Length Divisible by k

Theorem 8.6 (Alon & Linial, 1989) *For any directed graph G with minimal out-degree at least δ and maximal in-degree at most Δ contains a cycle of length divisible by k when*

$$k \leq \frac{\delta}{1 + \log(1 + \delta\Delta)}.$$

Proof: Assume that every vertex $v \in V$ has out-degree δ. (Otherwise, we delete some edges from v.) Assign $x_v \in \mathbb{Z}/k\mathbb{Z}$ to v uniformly randomly. Now, we look for cycles that the label increase by 1 at each step.

Let $A_v = \{\text{none out-neighbor of } v \text{ has label } x_v + 1\}$. Thus,

$$\mathbf{Pr}[A_v] = (1 - 1/k)^{\delta} \leq e^{-\delta/k}.$$

Let $N^{out}(v)$ be the set of out-neighbors of vertex v. Naively we may use the dependency graph where $A_u \sim A_v$ if and only if $\{u\} \cup N^{out}(u)$ intersects $\{v\} \cup N^{out}(v)$.

In fact we can construct a directed dependency graph and improve the bound. Note that $\Pr[A_v]$ is $(1-1/k)^{\delta}$ as long as $N^{out}(v)$ are free, even if v is assigned. So A_v is independent from all A_u's where $N^{out}(v)$ does not intersect $\{u\} \cup N^{out}(u)$. Therefore, the maximal degree of the dependency graph $d \leq \Delta\delta$. As

$$e^{1-\delta/k}(1 + \Delta\delta) \leq 1,$$

we are done by the Lovász Local Lemma. ■

Remark. The dependency is not symmetric in this proof.