**CS271 Randomness & Computation** **Spring 2020**

# Lecture 22: April 9

*Instructor: Alistair Sinclair*



In this lecture we return to the discussion of the probabilistic method and introduce a powerful tool known as the *Lovász Local Lemma* (first published by Erdös and Lovász [EL75]). To demonstrate the usefulness of the lemma, we explore its application to random $k$-SAT, packet routing.in networks, and graph coloring.

## 22.1 The Lovász Local Lemma

Recall from earlier lectures that the probabilistic method provides a useful non-constructive strategy for proving the existence (or non-existence) of an object that satisfies some prescribed property. Generally, the argument involves selecting an object randomly from a specific set and demonstrating that it has the desired property with strictly positive probability. This in turn proves the existence of at least one such object. In most of the examples we have seen, the desired property holds not just with positive probability but actually with quite large probability, even with probability approaching 1 as $n \to \infty$. This in turn often leads to an efficient randomized algorithm; we just select an object at random and with high probability it has the desired property.

For some problems, it is natural to describe the selected object in terms of a set of "bad" events $\{A_1, A_2, ..., A_n\}$, whose occurrences render the object undesirable, while the desired property is simply the avoidance of all events in the set. In such scenarios, the existence of a non-trivial lower bound on $\Pr[\bigwedge_{i=1}^n \neg A_i]$ is of particular interest. Clearly, if all "bad" events are independent, and if the probability of each of them satisfies $\Pr[A_i] \leq p$, then the probability that none of the events $\{A_i\}$ occur is simply the product

$$\Pr\left[\bigwedge_{i=1}^n \neg A_i\right] = \prod_{i=1}^n \Pr[\neg A_i] \geq (1-p)^n, \tag{22.1}$$

which is strictly positive (provided only that the trivial condition $p < 1$ holds).

Informally, the Lovász Local Lemma can be viewed as extending the above result to a somewhat more general setting, in which we allow limited dependencies among the events in question. In light of (22.1), the resulting probability that no bad event occurs will typically be exponentially small. Thus the Local Lemma tends to apply in situations where we are looking for a "needle in a haystack," so does not immediately lead to an efficient randomized algorithm. (However, see below and the next lecture for recent developments on constructive versions of the Lemma.)

**Definition 22.1** *An event $A$ is said to be* mutually independent of *a set of events $\{B_i\}$ if for any subset $\beta$ of events or their complements contained in $\{B_i\}$, we have $\Pr[A \mid \beta] = \Pr[A]$.*

**Lemma 22.2 (The Lovász Local Lemma).** *Let $A_1, \ldots, A_n$ be a set of "bad" events with $\Pr[A_i] \leq p < 1$ and each event $A_i$ is mutually independent of all but at most $d$ of the other $A_j$. If $e \cdot p(d+1) \leq 1$ then*

$$\Pr\left[\bigwedge_{i=1}^n \neg A_i\right] > 0.$$

**Note:** Often the Lovász Local Lemma is stated with the condition $e \cdot p(d+1) \leq 1$ replaced by $4pd \leq 1$, which is slightly stronger for $d \leq 2$ but asymptotically weaker. In fact, the constant $e$ above is asymptotically optimal.

We will provide a formal proof of this important lemma shortly, but first we will examine one simple application to the satisfiability properties of Boolean formulas.

### 22.1.1 The existence of a satisfying k-SAT assignment

**Claim 22.3** *Any instance $\varphi$ of k-SAT in which no variable appears in more than $\frac{2^{k-2}}{k}$ clauses is satisfiable.*

As a quick example, the above claim implies that for $k = 10$, any formula in which no variable appears in more than 25 clauses is satisfiable. Note that there is no restriction at all on the total number of clauses!

**Proof:** Suppose we have an arbitrary instance $\varphi$ of k-SAT consisting of $n$ clauses. Let's pick a truth assignment to the variables of $\varphi$ uniformly at random and let $A_i$ denote the event "clause $i$ is not satisfied".

Noting that exactly one of the $2^k$ possible assignments fails to satisfy any particular clause, we have

$$\forall i \in \{1, 2, ..., n\} : \Pr[A_i] = 2^{-k} \equiv p.$$

Furthermore, we observe that each event $A_i$ is independent of all other events $A_j$ *except* those corresponding to clauses $j$ that share at least one variable with clause $i$. Let $d$ denote the largest possible number of such clauses. Clearly, since each variable is assumed to appear in at most $\frac{2^{k-2}}{k}$ clauses, we have

$$d \leq k\frac{2^{k-2}}{k} = 2^{k-2}.$$

The condition $p \leq \frac{1}{4d}$ in the Local Lemma now becomes $\frac{1}{2^k} = p \leq \frac{1}{4d} = \frac{1}{4 \cdot 2^{k-2}}$, which clearly holds. Hence the Lemma implies that

$$\Pr\left[\bigwedge_{i=1}^{n} \neg A_i\right] > 0.$$

Since the probability of picking an assignment that satisfies every clause in $\varphi$ is non-zero, we can invoke the standard argument of the probabilistic method and infer the existence of a satisfying truth assignment. ■

In the above proof, we claimed that each $A_i$ is independent of all $A_j$ for which clauses $i$ and $j$ do not share any variables. This is an instance of the following general principle that is frequently useful in applications of the Local Lemma:

**Proposition 22.4 [Mutual Independence Principle]:** *Suppose that $Z_1, \ldots, Z_m$ is an underlying sequence of independent events, and suppose that each event $A_i$ is completely determined by some subset $S_i \subset \{Z_1, \ldots, Z_m\}$. If $S_i \cap S_j = \emptyset$ for $j = j_1, \ldots, j_k$, then $A_i$ is mutually independent of $\{A_{j_1}, \ldots, A_{j_k}\}$.*

In our above application, the underlying independent events $Z_\ell$ are the assignments to the variables.

### 22.1.2 Proof of the Lovász Local Lemma

The main ingredient in the proof is the following claim.

**Claim 22.5** *For any subset $S \subsetneq \{1,\dots,n\}$, and any $i \in \{1,\dots,n\}$,* $\Pr\left[A_i \;\middle|\; \bigwedge_{j\in S}\overline{A_j}\right] \le \frac{1}{d+1}$.

**Proof:** We proceed by induction on $m := |S|$. The base case of $m = 0$ is true since $\Pr[A_i] \le p \le \frac{1}{e(d+1)} < \frac{1}{d+1}$. For the inductive step ($m > 0$) we first partition $S$ into the two sets $S_1 = S \cap D_i$ and $S_2 = S \backslash S_1$, where $D_i$ is the "dependency set" of $A_i$, i.e., the set of at most $d$ indices $j$ such that $A_i$ is independent of all $A_j$ except for those in this set. Then we may write

$$\Pr\left[A_i \;\middle|\; \bigwedge_{j\in S}\overline{A}_j\right] = \frac{\Pr\left[A_i \wedge \bigwedge_{j\in S_1}\overline{A}_j \;\middle|\; \bigwedge_{k\in S_2}\overline{A}_k\right]}{\Pr\left[\bigwedge_{j\in S_1}\overline{A}_j \;\middle|\; \bigwedge_{k\in S_2}\overline{A}_k\right]} \tag{22.2}$$

We can upper bound the numerator of (22.2) by $\Pr\left[A_i \;\middle|\; \bigwedge_{k\in S_2}\overline{A}_k\right]$, which by mutual independence equals $\Pr[A_i]$.

Denoting $S_1 = \{j_1,\dots,j_r\}$ (and assuming w.l.o.g. that $r > 0$, since otherwise $S_1 = \emptyset$ and the denominator is 1), we can expand the denominator of (22.2) by the chain rule as follows:

$$\begin{aligned}\Pr\left[\bigwedge_{j\in S_1}\overline{A}_j \;\middle|\; \bigwedge_{k\in S_2}\overline{A}_k\right] &= \prod_{l=1}^{r}\left(1-\Pr\left[A_{j_l} \;\middle|\; \left(\bigwedge_{l'<l}\overline{A}_{j_{l'}}\right)\wedge\left(\bigwedge_{k\in S_2}\overline{A}_k\right)\right]\right)\\ &\ge \left(1-\frac{1}{d+1}\right)^d > \frac{1}{e}.\end{aligned}$$

The first inequality here follows by applying the induction hypothesis to each of the factors, noting that the number of events in the conjunction in the conditioning is always less than $m$.

Finally, combining the bounds on the numerator and denominator of (22.2), we get

$$\Pr\left[A_i \;\middle|\; \bigwedge_{j\in S}\overline{A}_j\right] \le \frac{\Pr[A_i]}{1/e} \le e\cdot p \le \frac{1}{d+1}.$$

This completes the induction proof. ∎

We now apply the claim to prove the Lovász Local Lemma.

**Proof of Lemma 22.2:** Expand $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right]$ by the chain rule and then apply Claim 22.5 to each factor of the resulting product:

$$\begin{aligned}\Pr\left[\bigwedge_{i=1}^{n}\overline{A}_i\right] &= \prod_{i=1}^{n}\left(1-\Pr\left[A_i \;\middle|\; \bigwedge_{j<i}\overline{A}_j\right]\right)\\ &\ge \left(1-\frac{1}{d+1}\right)^n > 0.\end{aligned}$$

∎

**Note:** As noted earlier, the Lovász Local Lemma (as evidenced by the above proof) is in general non-constructive. However, important work initiated by Beck [B91], and leading up to a dramatic breakthrough by Moser and Tardos [Mos09,MT10] provides *algorithmic* versions of the Local Lemma in quite general circumstances. Indeed, there are now randomized search algorithms inspired by the Lemma that require even weaker conditions than the Lemma itself! We will discuss some of these developments in the next lecture.

## 22.2 Packet Routing in Networks

As a rather sophisticated application of the Lovász Local Lemma, we look at packet routing in networks.

Consider an undirected graph $G$ and a set of packets $i$, each of which is given a path $P_i$ from some source $s_i$ to some destination $t_i$. Our goal is to establish a schedule that minimizes the time for all packets to travel along their paths from their sources to their destinations. We assume a synchronous model, in which each edge can carry at most one packet (in each direction) in one time step. We will also assume that the paths are *edge-simple*, i.e., no path repeats the same edge.

Since the paths are already specified, the only freedom we have in the schedule is the queueing mechanism: i.e., when multiple packets wish to move along an edge, we can choose which one goes first. Contrast the problem we studied earlier (Lecture 13), where we needed to choose the paths themselves. In the present scenario, we assume nothing about the structure of the network.

It is natural to measure the quality of the schedule in terms of two key parameters. Define the *congestion* $c$ to be the maximum number of paths $P_i$ that go through any one (directed) edge, and the *dilation* $d$ to be the maximum length of any path $P_i$. It is obvious that we can route all the packets in $c \cdot d$ time because each packet traverses at most $d$ edges and can be held up for at most $c-1$ time steps at each edge. Also, a trivial *lower* bound on the time required is $\max\{c, d\}$.

The following remarkable theorem says that this lower bound is achievable (up to constant factors).

**Theorem 22.6** [LMR94] *Assuming that all paths are edge-simple, there is a schedule that achieves time* $O(c+d)$ *with constant size queues.*

**Note:** We will actually prove the slightly weaker time bound of $O\left((c+d)2^{O(\log^\star(c+d))}\right)$, with queue sizes $(\log d)2^{O(\log^\star(c+d))}$. This brings out all the main ideas and avoids some technical complications. Also, $\log^\star n$ is essentially a constant for all practical purposes (e.g., $\log^\star n \leq 5$ for all $n \leq 2^{65536}$).

**Proof:** We assume w.l.o.g. that $c = d$. Consider first a (presumably infeasible) schedule in which each packet $i$ waits time $Z_i$ at its source, and then proceeds directly along $P_i$ without ever waiting. Here the initial delays $Z_i$ are chosen independently and uniformly in $\{1, 2, \ldots, \alpha d\}$, where $\alpha > 1$ is a constant to be chosen later. Clearly, the total length of this (infeasible) schedule is at most $(1+\alpha)d$.

**Claim 22.7** *If we divide this schedule into "frames" of length* $\ln d$*, we get a decomposition into subproblems whose objectives are to get each packet from its initial position in that frame to its final position in that frame. With positive probability, every edge in every subproblem has congestion at most* $\ln c$*.*

We will prove the Claim shortly using the Local Lemma. For now, assuming the Claim we can divide the problem into $(1+\alpha)d/\ln d$ subproblems, each with congestion and dilation $\ln c$ and $\ln d$ respectively, and solve them recursively. We can then stitch the subschedules together to get an overall schedule. Eventually, we will reach subproblems with constant congestion and dilation, for which we can clearly construct a *feasible* schedule. Since this base schedule is feasible, the final schedule we end up with after stitching back up through the levels of the recursion will also be feasible.

The number of levels of recursion is at most $O(\log^\star(c+d))$ and at each level the schedule length is increased by a constant factor $1+\alpha$. Thus the total length of the schedule is $d2^{O(\log^\star(c+d))}$, as claimed. By combining the facts:

- if a packet is queued in the beginning of a frame, it has to move before that frame ends;

- it takes at most $(\log d)2^{O(\log^\star(c+d))}$ time to clear a frame; and
- for any $k$, if every packet in a queue moves in $k$ time steps, then the queue cannot contain more than $k$ packets,

we see that the queue size is at most $(\log d)2^{O(\log^\star(c+d))}$. ■

It remains to go back and prove the key Claim above.

**Proof of Claim 22.7:** For each edge $e$ define the "bad" event $A_e$ = "$e$ has congestion greater than $\ln c$ in some frame". Notice that $A_e$ can depend only on those $A_{e'}$ for which edges $e, e'$ have a common packet passing through them, and since at most $c$ packets pass through $e$, and each of these packets passes through at most $d$ edges in total, the dependency set of $A_e$ has size at most $cd = d^2$.

Computing an upper bound on the probabilities $\Pr[A_e]$ is a little trickier. Each edge $e$ is visited by at most $c$ packets. However, the length of a frame is $\ln d$, so since each packet suffers a random delay in the range $[1, \dots, \alpha d]$, the probability of any given packet visiting $e$ during a specific frame is only $\frac{\ln d}{\alpha d}$. Since the delays on different packets are independent, the congestion of an edge is distributed as $\mathrm{Bin}(c, \frac{\ln d}{\alpha d})$. Thus taking a union bound over frames (of which there are less than $(1+\alpha)d$), we get

$$\begin{aligned}\Pr[A_e] &\leq (1+\alpha)d \times \Pr[\text{edge } e \text{ has congestion greater than } \ln c \text{ in a fixed frame}] \\ &\leq (1+\alpha)d \times \Pr\left[\mathrm{Bin}\left(c, \frac{\ln d}{\alpha d}\right) > \ln c\right].\end{aligned}$$

Applying the Chernoff bound in the form $\Pr[X \geq (1+\beta)\mu] \leq \exp(-\mu((1+\beta)\ln(1+\beta)-\beta))$ with $\mu = (\ln d)/\alpha$ and $1+\beta = (\ln c)/\mu = \alpha$, and recalling that $c = d$, we get

$$\Pr[A_e] \leq (1+\alpha)d \times \exp\left(-\frac{\ln d}{\alpha}(\alpha \ln \alpha - (\alpha - 1))\right) \leq (1+\alpha)d^{2-\ln\alpha}.$$

The condition of the basic Lovász Local Lemma (Lemma 22.2) requires that this probability be less than $1/e(d^2+1)$ (recall from above that the size of the dependency sets here is at most $d^2$). But we can ensure this (for large enough $d$) by taking $\alpha$ large enough.

This completes the proof of the Claim. ■

**Note:** As is typical in applications of the Local Lemma, the probability of producing a good schedule via the above randomized construction is tiny. However, a (not necessarily practical) algorithmic version of the above result was proved subsequently by Leighton, Maggs and Richá [LMR99].

## 22.3 The General Lovász Local Lemma

In some settings it is useful to have a more flexible version of the Local Lemma, which allows large differences in the probabilities of the "bad" events. We state this next.

**Lemma 22.8 (General Lovász Local Lemma).** *Let $A_i, \dots, A_n$ be a set of "bad" events, and let $D_i \subseteq \{A_1, \dots, A_n\}$ denote the "dependency set" of $A_i$ (i.e., $A_i$ is mutually independent of all events not in $D_i$). If there exists a set of real numbers $x_1, \dots, x_n \in [0,1)$ such that $\Pr[A_i] \leq x_i \prod_{j \in D_i}(1-x_j)$ for all $i$, then $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] \geq \prod_{i=1}^n (1-x_i) > 0$.*

**Exercise:** It is left as a straightforward (and strongly recommended) exercise to prove this general version by mimicking the proof of Lemma 22.2. Also, you should check that applying Lemma 22.8 with $x_i = 1/(d+1)$ yields Lemma 22.2 as a special case.

**Corollary 22.9 (Asymmetric Lovász Local Lemma).** *In the same scenario as in Lemma 22.8, if* $\sum_{j \in D_i} \Pr[A_j] \le 1/4$ *for all* $i$ *then* $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] \ge \prod_{i=1}^n (1 - 2\Pr[A_i]) > 0$.

**Proof:** The result follows easily by applying Lemma 22.8 with $x_i = 2\Pr[A_i]$. **Exercise**: check this! ■

### 22.3.1 An Application: Frugal Graph Coloring

We give an application in which the extra flexibility of the Asymmetric Local Lemma, in which the event probabilities $\Pr[A_i]$ can differ a lot, is crucial

**Definition 22.10** *A proper coloring of a graph* $G$ *is called* $\beta$-frugal *if no color appears more than* $\beta$ *times in the neighborhood of any vertex of* $G$.

**Theorem 22.11** [HMR97] *If graph* $G$ *has maximum degree* $\Delta \ge \beta^\beta$ *then* $G$ *has a* $\beta$*-frugal coloring with* $16\Delta^{1+1/\beta}$ *colors.*

**Proof:** A 1-frugal coloring of $G$ is equivalent to a proper coloring of $G^2$, which has maximum degree $\Delta^2$. By Brooks' Theorem [B41], $G^2$ can be colored with $\Delta^2 + 1$ colors and so the theorem holds for $\beta = 1$.

For $\beta \ge 2$, pick a random (not necessarily proper) coloring of $G$ with $Q = 16\Delta^{1+1/\beta}$ colors. We will use the Local Lemma to show that this coloring is proper and $\beta$-frugal with positive probability.

We distinguish events which could prevent our coloring from being proper (Type-1 events) and $\beta$-frugal (Type-2 events):

Type-1 events: for each $\{u, v\}$ in the edge-set of $G$, $A_{uv}$ = "$u, v$ have the same color".

Type-2 events: for each set of $\beta + 1$ neighbors $u_1, \dots, u_{\beta+1}$ of some vertex, $B_{u_1,\dots,u_{\beta+1}}$ = "$u_1, \dots, u_{\beta+1}$ have the same color."

For any $k$, the probability of $k$ vertices having the same color is clearly $1/Q^{k-1}$. Thus

$$\begin{aligned} \Pr[A_{uv}] &= 1/Q; \\ \Pr\left[B_{u_1,\dots,u_{\beta+1}}\right] &= 1/Q^\beta. \end{aligned}$$

Using the Mutual Independence Principle, we can see that each Type-1 event depends on at most $2\Delta$ Type-1 events and $2\Delta\binom{\Delta}{\beta}$ Type-2 events, while each Type-2 event depends on at most $(\beta+1)\Delta$ Type-1 events and $(\beta+1)\Delta\binom{\Delta}{\beta}$ Type-2 events (see Figure 22.1). Note that, since $\beta \ge 2$, the size of the dependency set of a Type-2 event (dependencies of both Type-1 and Type-2) dominates that of a Type-1 event, so it will suffice to sum the probabilities over the dependency set of a Type-2 event.

We now appeal to the Asymmetric Local Lemma (Corollary 22.9). With slight abuse of notation we concatenate all of the events $A_{uv}$ and $B_{u_1,\dots,u_{\beta+1}}$ into a sequence of events $\{A_i\}$. Then we have, for any $i$,

$$\begin{aligned} \sum_{j \in D_i} \Pr[A_j] &\le \left[(\beta+1)\Delta \times \frac{1}{Q}\right] + \left[(\beta+1)\Delta\binom{\Delta}{\beta} \times \frac{1}{Q^\beta}\right] \\ &\le \frac{(\beta+1)\Delta}{Q} + \frac{(\beta+1)\Delta^{\beta+1}}{\beta! Q^\beta} && \text{bounding } \binom{\Delta}{\beta} \text{ by } \frac{\Delta^{\beta+1}}{\beta!} \\ &= \frac{(\beta+1)}{16\Delta^{1/\beta}} + \frac{\beta+1}{\beta! 16^\beta} && \text{expanding Q} \\ &\le 1/4 && \text{provided } \Delta > \beta^\beta \text{ and } \beta > 2 \end{aligned}$$

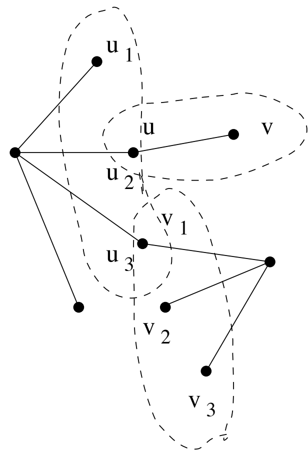


Figure 22.1: Dependency sets of Type-2 events in the proof of Theorem 22.11: the Type-2 event $B_{u_1,\ldots,u_{\beta+1}}$ can depend on another Type-2 event $B_{v_1,\ldots,v_{\beta+1}}$ or a Type-1 event $A_{uv}$ only if $\{u_1,\ldots,u_{\beta+1}\} \cap \{v_1,\ldots,v_{\beta+1}\} \neq \emptyset$ or $\{u_1,\ldots,u_{\beta+1}\} \cap \{u,v\} \neq \emptyset$, respectively. The dashed curves define the the sets of vertices $\{u_1,\ldots,u_{\beta+1}\}$, $\{v_1,\ldots,v_{\beta+1}\}$ and $\{u,v\}$.

Thus $\Pr[\text{the random } Q\text{-coloring is } \beta\text{-frugal}] = \Pr\left[\bigwedge \overline{A}_i\right] > 0$ and so there exists a $\beta$-frugal $Q$-coloring of $G$. ■

**Note:** It is instructive to attempt to prove the above result using the basic form of the Local Lemma (Lemma 22.2). In that case we would have to use the uniform bound $\Pr[A_i] \le p = 1/Q$, together with the dependency set size of at least $d \ge \frac{(\beta+1)\Delta^{\beta+1}}{\beta!}$. But now for $\Delta = \beta^\beta$ we get that $pd \gg 1$, so there is no hope of applying the lemma. The reason the asymmetric version works is that, although some of the events in the dependency sets have large probability, most have small probability.

# References

[B91] J. Beck, "An Algorithmic Approach to the Lovász Local Lemma. I.," *Random Structures and Algorithms* **2** (1991), pp. 343–366.

[EL75] P. Erdös and L. Lovász, "Problems and results on 3-chromatic hypergraphs and some related questions," in *Infinite and Finite Sets*, A. Hajnal et al. eds., North-Holland, 1975, pp. 609–628.

[HMR97] H. Hind, M. Molloy and B. Reed, "Coloring a graph frugally," *Combinatorica* **17** (1997), pp. 469–482.

[LMR94] F.T. Leighton, B.M. Maggs and S.B. Rao, "Packet Routing and Job-Shop Scheduling in O(Congestion + Dilation) Steps," *Combinatorica* **14** (1994), pp. 167–186.

[LMR99] F.T. Leighton, B.M. Maggs and A. Richá, "Fast Algorithms for Finding O(Congestion + Dilation) Packet Routing Schedules," *Combinatorica* **19** (1999), pp. 375–401.

[Mos09] R. Moser, "A constructive proof of the Lovász local lemma," *Proceedings of ACM STOC*, 2009.

[MT10] R. Moser and G. Tardos, "A constructive proof of the general Lovász local lemma," *Journal of the ACM*, 2010.