Algorithmic Combinatorics WS 2014 / 2015
Tibor Szabó
Shagnik Das

# Solutions to Exercise Sheet 14

**Exercise 1.** Extend the proof of the Local Lemma from class (for two-colouring hypergraphs) to the following more general theorem (which has the optimal constants).

**Theorem 1** (Lovász Local Lemma)**.** *Let $E_1, E_2, \ldots, E_m$ be events in some probability space. Let $d \in \mathbb{N}$ and $p \in [0,1]$ be such that, for every $i \in [m]$, we have*

*(1) $\mathbb{P}(E_i) \leq p$, and*

*(2) there is a set $\Gamma(i) \subseteq [m]\{i\}$ of at most $d$ indices, such that the event $E_i$ is mutually independent of $\{E_j : j \in [m] \setminus (\Gamma(i) \cup \{i\})\}$.*

*If $ep(d+1) \leq 1$, then with positive probability none of the events $E_i$ occur.*

It may help to show that for any $i \in [m]$ and $J \subseteq [m] \setminus \{i\}$, we have $\mathbb{P}\left(E_i | \cap_{j \in J} E_j^c\right) \leq ep$. You may use the estimate $(1 - 1/(d+1))^d \geq e^{-1}$.

<u>Solution:</u> We shall indeed prove $\mathbb{P}\left(E_i | \cap_{j \in J} E_j^c\right) \leq ep$ for any $i \in [m]$ and $J \subseteq [m] \setminus \{i\}$, and shall do so by induction on $|\Gamma(i) \cap J|$.

If $|\Gamma(i) \cap J| = 0$ (including, in particular, the case $J = \emptyset$), then by the independence of $E_i$ and $\{E_j : j \in J\}$, we have

$$\mathbb{P}\left(E_i | \cap_{j \in J} E_j^c\right) = \mathbb{P}(E_i) \leq p < ep,$$

as required.

For the induction step, let $N = \Gamma(i) \cap J$. By the independence of $E_i$ and $\{E_j : j \in J \setminus N\}$, we have

$$\mathbb{P}\left(E_i | \cap_{j \in J} E_j^c\right) = \mathbb{P}\left(E_i | \cap_{j \in N} E_j^c\right) = \frac{\mathbb{P}\left(E_i \cap \left(\cap_{j \in N} E_j^c\right)\right)}{\mathbb{P}\left(\cap_{j \in N} E_j^c\right)} \leq \frac{\mathbb{P}(E_i)}{\mathbb{P}\left(\cap_{j \in N} E_j^c\right)} \leq \frac{p}{\mathbb{P}\left(\cap_{j \in N} E_j^c\right)}.$$

Hence it suffices to show $\mathbb{P}\left(\cap_{j \in N} E_j^c\right) \geq e^{-1}$. If $N = \{j_1, \ldots, j_t\}$, where $t = |N| \leq |\Gamma(i)| \leq d$,

$$\mathbb{P}\left(\cap_{j \in N} E_j^c\right) = \prod_{r=1}^{t} \left(1 - \mathbb{P}\left(E_{j_r} | \cap_{s \leq r-1} E_{j_s}^c\right)\right).$$

$\mathbb{P}\left(E_{j_r} \mid \cap_{s \leq r-1} E_{j_s}^c\right) \leq ep \leq 1/(d+1)$ by the induction hypothesis (which we may apply since $r-1 < |N|$), and so

$$\mathbb{P}\left(\cap_{j \in N} E_j^c\right) = \prod_{r=1}^{t} \left(1 - \mathbb{P}\left(E_{j_r} \mid \cap_{s \leq r-1} E_{j_s}^c\right)\right) \geq \left(1 - \frac{1}{d+1}\right)^d \geq e^{-1},$$

completing the induction step.

Now it is simple to derive the conclusion of the Local Lemma. Indeed, the probability that none of the events $E_i$ occurs is

$$\mathbb{P}\left(\cap_{i \in [m]} E_i\right) = \prod_{i \in [m]} \mathbb{P}\left(E_i^c \mid \cap_{j<i} E_j^c\right) = \prod_{i \in [m]} \left(1 - \mathbb{P}\left(E_i \mid \cap_{j<i} E_j^c\right)\right) \geq (1-ep)^m > 0.$$

(Observe that $ep < ep(d+1) \leq 1$.)

**Exercise 2.** In class we showed that, for a $k$-uniform hypergraph $\mathcal{F}$ with $\Delta(L(\mathcal{F})) \leq 2^{k-4}$, the expected number of recolourings in the algorithmic Local Lemma is $O(m \log m)$. Show that, with more careful analysis, this bound can be greatly improved to $O\left(\frac{n}{k} \log m\right)$.

Solution: Recall how we obtained the $O(m \log m)$ bound. We constructed a rooted ordered tree to keep track of the recolourings of the randomised colouring algorithm. By showing that, apart from the top-level calls, each random recolouring of the $k$ elements of a set could be efficiently encoded in $k-1$ bits, we deduced that it is very unlikely that there should be many lower-level recolourings, and hence in expectation each top-level call leads to $O(\log m)$ lower-level recolourings.

We also showed that at the end of a top-level call for a set $F$, any set intersecting $F$ is properly coloured, and hence will not subsequently appear as a top-level call. In particular, each set can be invoke a top-level call at most once, and hence there are at most $m$ top-level calls, giving the $O(m \log m)$ bound.

However, the previous fact shows that the sets involved in the top-level calls must actually form a matching. As a matching of $k$-sets in $[n]$ can have size at most $\frac{n}{k}$, we get the improved $O\left(\frac{n}{k} \log m\right)$ bound on the number of recolourings.

**Exercise 3.** Recall that the Ramsey number $R(k,k)$ is the smallest $n$ such that any two-colouring of the edges of $K_n$ must contain a monochromatic copy of $K_k$.

(i) By colouring edges randomly, show that if $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$, then $R(k,k) > n$. Deduce that $R(k,k) \geq \frac{1}{e\sqrt{2}}(1+o(1))k2^{k/2}$. [This is from Discrete Math I.]

(ii) Obtain a $\sqrt{2}$-factor improvement of the result in (i) by 'correcting' a random colouring by removing monochromatic cliques: show that for any integer $n$, $R(k,k) > n - \binom{n}{k} 2^{1-\binom{k}{2}}$. Deduce that $R(k,k) \geq \frac{1}{e}\left(1+o(1)\right) k2^{k/2}$.

(iii) Improve the bound by yet another $\sqrt{2}$-factor with the Local Lemma: show that if $e\binom{k}{2}\binom{n-2}{k-2}2^{1-\binom{k}{2}} \leq 1$, then $R(k,k) > n$. Deduce the bound $R(k,k) \geq \frac{\sqrt{2}}{e}(1+o(1))k2^{k/2}$.

<u>Solution:</u>

(i) Given $n$, and consider colouring the edges of $K_n$ independently, uniformly at random. For a given set of $k$ vertices, the probability they induce a monochromatic clique is $2^{1-\binom{k}{2}}$, since there are two possible colours, and each of the $\binom{k}{2}$ edges will be given that colour with probability $1/2$. As there are $\binom{n}{k}$ sets of $k$ vertices, the expected number of monochromatic cliques of size $k$ is $\binom{n}{k}2^{1-\binom{k}{2}}$. By assumption, this is strictly less than 1, which is only possible if there is some edge-colouring of $K_n$ without any monochromatic $k$-clique. Hence we must have $R(k,k) > n$.

To get a good lower bound on $R(k,k)$, we need to choose $n$ as large as possible while $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ holds. We can bound the left-hand side by

$$\binom{n}{k}2^{1-\binom{k}{2}} \leq \left(\frac{ne}{k}\right)^k 2^{1-\binom{k}{2}} = 2\left(\frac{ne\sqrt{2}}{k2^{k/2}}\right)^k.$$

Thus if $n = \frac{2^{-1/k}}{e\sqrt{2}}k2^{k/2} = \frac{1}{e\sqrt{2}}(1+o(1))k2^{k/2}$, the above expression is equal to 1, and so we deduce $R(k,k) > \frac{1}{e\sqrt{2}}(1+o(1))k2^{k/2}$.

(ii) For larger $n$, the expected number of monochromatic cliques will be large, so we cannot hope to find a monochromatic-$k$-clique-free $K_n$ by taking a random edge colouring. However, if the number of monochromatic cliques is not *too* large, we can remove a vertex from each such clique to be left with a good colouring on a smaller (but not too small) number of vertices.

Indeed, if we start with $n$ vertices, and remove one vertex from every monochromatic clique, we would on average be left with at least $n - \binom{n}{k}2^{1-\binom{k}{2}}$ vertices, showing there exists a monochromatic-$k$-clique-free graph of at least that size, and hence $R(k,k) \geq n - \binom{n}{k}2^{1-\binom{k}{2}} \geq n - 2\left(\frac{ne\sqrt{2}}{k2^{k/2}}\right)^k$.

As before, to get a good bound on $R(k,k)$, we should choose $n$ to maximise this lower bound. Differentiating with respect to $n$ and setting the derivate equal to 0, we find

$$2^{(3-k)/2}e\left(\frac{ne\sqrt{2}}{k2^{k/2}}\right)^{k-1} = 1,$$

or

$$n = \frac{2^{(k-3)/(2k-2)}e^{-1/(k-1)}}{e\sqrt{2}}k2^{k/2} \sim \frac{1}{e}(1+o(1))k2^{k/2},$$

giving the required bound on $R(k,k)$.

(iii) For the final[1] improvement, we make use of the Lovász Local Lemma. Once again, we shall colour the edges of $K_n$ independently and uniformly at random. For each set $K \subset V(K_n)$ of $k$ vertices, let $E_K$ be the event that $k$-clique on $K$ is monochromatic. As before, we have $\mathbb{P}(E_K) = 2^{1-\binom{k}{2}}$.

The event $E_K$ is determined by the edges supported on $K$, and hence is mutually independent of $\{E_{K'} : |K \cap K'| \leq 1\}$, since these events do not depend on any of the edges of $K$. Thus the number of events that $E_K$ is not mutually independent of (including $E_K$ itself) is at most $\binom{k}{2}\binom{n-2}{k-2}$, since we must choose some edge of $K$ that they have in common, and then can choose the remaining $k-2$ vertices freely.

Taking these values as $p$ and $d+1$ respectively, it follows from the Local Lemma that if

$$e\binom{k}{2}\binom{n-2}{k-2}2^{1-\binom{k}{2}} \leq 1,$$

then with positive probability our random colouring of $K_n$ will not have any monochromatic $k$-cliques, and thus $R(k,k) > n$.

Asymptotically, the left-hand side is

$$ek^2\left(\frac{ne}{k-2}\right)^{k-2}2^{-k(k-1)/2} = \frac{ek^2}{2}\left(\frac{ne}{(k-2)2^{(k+1)/2}}\right)^{k-2}.$$

This will be smaller than 1 when $n = \frac{\sqrt{2}}{e}(1+o(1))k2^{k/2}$, as claimed.

**Exercise 4.** The city of London is surrounded by the M25 motorway, a circular road that directs traffic around the city without congested its inner roads. It is approximately 110 miles long and, as per UK traffic regulations, has 30 streetlights per mile, and thus a total of 3300 lampposts.

To comply with recent environmental guidelines, the Mayor of London wants to illuminate the M25 with environmentally-friendly lightbulbs that will consume less power while maintaining adequate light coverage. To find the best lightbulb for the job, he commissions London's 300 different lighting firms to submit prototypes for evaluation.

Each firm provides a sample of 11 lightbulbs. To ensure that no firm has all its lightbulbs in a favourable stretch of the highway, all 3300 lightbulbs are mixed together and then placed, in some arbitrary order, in the M25's lampposts. The Mayor intends to keep these lightbulbs in place for a month and evaluate their efficiency before making a final decision about which lightbulb to use in the long-term.

[1] Indeed, despite turning 40 this year, this bound (due to Spencer) remains the best-known lower bound. While that may seem a long time without progress, I'd like to point out that its been a good 46 years since the lunar landing, in which time we still haven't got past first base with the moon.

Unfortunately, after a few days, he realises that this experiment is rather expensive, and decides the test has to be scaled down[2]. Thus one of each company's 11 lightbulbs will be switched off. However, in the interests of public safety, no two neighbouring lightbulbs should both be switched off, for fear of creating too long a dark stretch on the motorway.

Show that, regardless of how the lightbulbs were initially distributed, it is always possible to safely turn off one lightbulb from each company.

Solution: As we have not spent much time studying lightbulbs in this course, let us first rephrase the problem in more familiar terms. We shall make an auxiliary graph $G$, where the vertices of $G$ are the lampposts, and edges in $G$ correspond to neighbouring lampposts. Since the M25 is a circular road, it follows that $G$ will be a cycle with 3300 vertices.

Now we have an arbitrary partition of the vertices, $V(G) = \sqcup_{i=1}^{300} V_i$, where $V_i$ is the set of lightbulbs from the $i$th lighting firm. In particular, $|V_i| = 11$ for all $i$. We wish to turn off one lightbulb from each company; that is, some $x_i \in V_i$ for $1 \le i \le 300$. However, to do so safely, we must not turn off two neighbouring lightbulbs, and so $X = \{x_i : 1 \le i \le 300\}$ must be an independent set. In the parlance of graph theory, such a set $X$ is often referred to as an *independent transversal*.

To show that an independent transversal exists, we apply the Lovász Local Lemma. For each $i$, $1 \le i \le 300$, let $x_i$ be chosen independently and uniformly at random from $V_i$. As we want $X = \{x_i : 1 \le i \le 300\}$ to be an independent set, we will define the event $E_e$ for each edge $e \in E(G)$ of both of the endpoints of $e$ being selected in $X$. Since $X$ consists of exactly one vertex from each class $V_i$, if the endpoints of $e$ belong to the same class, the event $E_e$ never occurs. Otherwise the endpoints belong to different classes and are chosen independently, each with probability $1/11$, and so the event $E_e$ holds with probability $1/11^2 = 1/121$. Thus, setting $p = 1/121$, we always have $\mathbb{P}(E_e) \le p$.

We must take a little care in determining the dependencies of the events $E_e$. Note that if the endpoints of $e$ belong to classes $V_i$ and $V_j$, then the event is determined solely by knowing $x_i$ and $x_j$. Hence $E_e$ is mutually independent of all events corresponding to edges spanned by $V \setminus (V_i \cup V_j)$.

As $G$ is a cycle, every vertex is incident to 2 edges. Since $|V_i \cup V_j| \le |V_i| + |V_j| = 22$ (the inequality appears because we might have $i = j$), there are at most $2 \cdot 22 = 44$ edges involving vertices in $V_i \cup V_j$. This includes the edge $e$ itself, so we have $d + 1 \le 44$.

By the Local Lemma, since $ep(d+1) \le 44e/121 < 1$, it follows that with positive probability, none of the events $E_e$ occur. In this case, $X$ is an independent set in $G$, thus showing the existence of an independent transversal.

[2] An alternative would have been to raise taxes to fund the project, but he is a proud patriot, and, after a rather poor showing at the FIFA World Cup 2014, decides his country can ill afford to surrender either of her two advantages over France: lower taxes and finer food.