COMPSCI 690RA: Problem Set 4

Due: Tuesday, 5/3 by 8pm in Gradescope.

Instructions:

- You are allowed to, and highly encouraged to, work on this problem set in a group of up to three members.
- Each group should **submit a single solution set**: one member should upload a pdf to Gradescope, marking the other members as part of their group in Gradescope.
- You may talk to members of other groups at a high level about the problems but **not work through the solutions in detail together**.
- You must show your work/derive any answers as part of the solutions to receive full credit.

1. Randomized Triangle Coloring (6 points)

A graph is k-colorable if there is an assignment of each node to one of k colors such that no two nodes with the same color are connected by an edge.

1. (2 points) Show that if a graph is 3-colorable then there is a coloring of the graph using just 2 colors such that no triangle in monochromatic. I.e., for any three nodes u, v, w such that $(u, v), (v, w)$, and (u, w) are all edges, we do not have u, v, w all assigned to the same color.

   Start with a valid three coloring of the graph using colors $\{r, g, b\}$. Change all nodes colored b to be colored g. We argue that no triangle is monochromatic in this coloring. Before the swap, all three nodes in each triangle had different colors – since all pairs of the three nodes are connected by an edge. One of the nodes has its color switched from b to g, leaving one node in the triangle colored r and two colored g. Thus the triangle is not monochromatic.

2. (4 points) Consider the following algorithm for coloring a 3-colorable graph with 2 colors so that no triangle is monochromatic. Start with an arbitrary 2-coloring (some edges may be monochromatic, so it's not necessarily a valid coloring). While there are any monochromatic triangles, pick one arbitrarily and change the color of a randomly chosen vertex in that triangle. Give an upper bound on the expected number of steps of this process before a valid 2-coloring with all non-monochromatic triangles is found.

   Hint: Shoot for a polynomial, not an exponential number of steps here. Use the fact that part (1) actually implies the existence of *many* 2-colorings with non-monochromatic triangles.

Let S be some valid 3-coloring of the graph. Let $\mathcal{S}$ be the set of two colorings using the colors r, g that you can generate from S by switching any b vertex to either r or g. Note that by the argument in (1), all these colorings have non-monochromatic triangles. Now, let $\mathbf{X}_i$ be the number of vertices at step i that are colored the same way as the *closest coloring* in S.

Consider recoloring a monochromatic triangle at step i. Assume without loss of generality that all nodes in the triangle are currently colored r. Now, all colorings in $\mathcal{S}$ have one node in this triangle colored r, one colored g, and one that can be colored either g or r (it was originally colored b in S.) If we flip the color of the r node, we move one step further from the closest coloring $\mathcal{S}$. If we flip the g node, we move one step closer. If we flip the g/r node, we don't move any closer or further. Thus, $\mathbf{X}_{i+1}$ is set to $\mathbf{X}_i + 1$, $\mathbf{X}_i - 1$, $\mathbf{X}_i$, each with probability $1/3$.

Now, consider the Markov chain $\mathbf{Y}_0, \mathbf{Y}_1, \ldots,$ which is identical to $\mathbf{X}_0, \mathbf{X}_1, \ldots,$ except that even if there are no monochromatic triangles left, we pick an arbitrary triangle that isn't colored as in any of the colorings in $\mathcal{S}$, and flip a color in this triangle. We continue flipping until $\mathbf{Y}_i = n$. Clearly, this Markov chain can only take more steps to terminate than our original. Additionally, we still have that $\mathbf{Y}_{i+1} = \mathbf{Y}_i$ with probability $1/3$ and $\mathbf{Y}_{i+1} = \mathbf{Y}_i + 1$ with probability *at least* $1/3$ (since in this case it might be that both non-g/r nodes are colored incorrectly in which case $\mathbf{Y}_{i+1} = \mathbf{Y}_i + 1$ with probability $2/3$.)

Finally, let $\mathbf{Z}_0, \mathbf{Z}_1, \ldots$ be identical to $\mathbf{Y}_0, \mathbf{Y}_1, \ldots,$ but with transition probabilities:

$$\begin{aligned}
\Pr[\mathbf{Z}_{i+1} = j | \mathbf{Z}_i = j] &= 1/3 \text{ for } 0 \le j \le n-1 \\
\Pr[\mathbf{Z}_{i+1} = 2 | \mathbf{Z}_i = 1] &= 2/3 \\
\Pr[\mathbf{Z}_{i+1} = j+1 | \mathbf{Z}_i = j] &= 1/3 \text{ for } 2 \le j \le n-1 \\
\Pr[\mathbf{Z}_{i+1} = j-1 | \mathbf{Z}_i = j] &= 1/3 \text{ for } 2 \le j \le n-1 \\
\Pr[\mathbf{Z}_{i+1} = n | \mathbf{Z}_i = n] &= 1.
\end{aligned}$$

Let h_j be the expected number of steps to find a valid two coloring assuming that $\mathbf{X}_0 = j$. We have $h_n = 0$. Also $h_1 = \frac{h_1}{3} + \frac{2h_2}{3} + 1$. Further, $h_j = \frac{h_{j-1}}{3} + \frac{h_{j+1}}{3} + \frac{h_j}{3} + 1$ for all $2 \le j \le n-1$. We claim that $h_j \le h_{j+1} + 3j + 4$. This is shown via induction. We have $h_1 = \frac{3}{2} \cdot \left(\frac{2h_2}{3} + 1\right) = h_2 + 3/2 \le h_2 + 3 + 4$, satisfying the claim. Further, assuming the induction claim for all $j' \le j$,

$$\begin{aligned}
h_j &= \frac{h_{j-1}}{3} + \frac{h_{j+1}}{3} + \frac{h_j}{3} + 1 \implies \\
h_j &\le \frac{3}{2} \cdot \left(\frac{h_j + 3(j-1) + 4}{3} + \frac{h_{j+1}}{3} + 1\right) \\
h_j &\le \frac{h_j}{2} + \frac{3(j-1)}{2} + 2 + \frac{h_{j+1}}{2} + \frac{3}{2} \implies \\
h_j &\le 3(j-1) + 7 + h_{j+1} \\
&= h_{j+1} + 3j + 4.
\end{aligned}$$

Now, in total we have:

$$h_1 \le h_2 + 3 + 4 \le h_3 + 3 \cdot 2 + 4 + \ldots = \sum_{j=1}^{n-1} 3j + 4 = O(n^2).$$

2. Move to Top Shuffling (8 points)

Consider shuffling a deck of n unique cards by randomly picking a card and moving it to the top of the deck. Observe that with probability $1/n$, the top card is picked and so the order does not change from one step to the next.

1. (2 points) Prove that this Markov chain is irreducible and aperiodic.

   Each state has a self loop (pick the top card) to if we show the chain is irreducible then it is also aperiodic. To show that it is irreducible, we just need to argue that there is a path from any state to any other state. Consider some permutation p and consider the series of swaps that selects the cards in reverse order of p. Applying this set of swaps to any permutation transforms it to p. Thus, there is a path in the chain from any permutation to any other, so the state graph is strongly connected, i.e., the chain is irreducible.

2. (2 points) Prove that the chain converges to the the uniform distribution over all $n!$ possible permutations of the cards.

   Since the chain aperiodic and irreducible it has a unique stationary distribution that it converges to. It suffices to check that this stationary distribution is the uniform distribution over permutations. For any permutation p, let $\mathcal{N}(p)$ be the set of permutations that can be transformed to p in one step. These permutations are exactly the permutations where the top card of p is inserted into one of n positions in the deck. For any $p' \in \mathcal{N}(p)$, $P_{p',p} = 1/n$. For any $p' \notin \mathcal{N}(p)$, $P_{p',p} = 0$.

   Thus, letting $\pi \in \mathbb{R}^{n!}$ have $\pi(p) = 1/n!$ for all p (i.e., it is the uniform distribution over permutations) we have for any permutation p:

$$[\pi P](p) = \sum_{p'} \frac{1}{n!} P_{p',p} = \frac{1}{n!} \sum_{\mathcal{N}(p)} \frac{1}{n} = \frac{1}{n!} \cdot n \cdot \frac{1}{n} = \frac{1}{n!}.$$

   Thus $\pi P = \pi$, confirming that the uniform distribution is in fact stationary.

3. (2 points) In class, we argued that after $t = n \log(n/\epsilon)$ steps, the distribution of states q^t in this Markov chain satisfies $\|q^t - \pi\|_{TV} \leq \epsilon$. Say you are a casino, and you offer a game of pure chance where the customer must wager \$1. The game uses the shuffled deck of cards to determine a pay out somewhere between \$0 and \$1000. You have calculated that, when the deck is ordered according to a uniform random permutation (i.e., according to π), your expected winnings per game are \$0.1. How small must you set ϵ to ensure that your expected winnings are at least \$.09?

   In the worst case, q^t places $\|q^t - \pi\|_{TV} = \epsilon$ more mass on permutations leading to the the maximum \$1000 payout, and ϵ less mass on the \$0 payout. Thus, the expected winnings per game under q^t are equal to $0.1 - \epsilon * 1000$. So we must set $\epsilon = 1/100000$ to ensure that the expected winnings remain $\geq .09$.

4. (2 points) Argue that our mixing time bound is essentially tight. In particular, show that if we run the Markov chain for $t \leq cn \log n$ steps for small enough constant c, then $\|q^t - \pi\|_{TV} \geq 99/100$. I.e., we are very far from a uniformly random permutation.

   Hint: Start by arguing that if $t \leq cn \log n$ for small enough c, with high probability there are $\sqrt{n}$ cards which are never swapped in the shuffle. Use the coupon collector analysis from Lecture 2. Then consider the probability that we have $\sqrt{n}$ consecutive cards in order after a uniform random shuffle, vs. after this shuffle starting from an ordered deck.

   Let $\mathbf{T}$ be the expected time to collect $n - \sqrt{n}$ coupons. But the coupon collector analysis, $\mathbb{E}[\mathbf{T}] = n(H_n - H_{\sqrt{n}}) = \Theta(n[\log n - \log(\sqrt{n})]) = \Theta(n \log n)$ since $\log(\sqrt{n}) = 1/2 \cdot \log n$. Additionally, $\text{Var}[\mathbf{T}] = O(n^2)$. So by Chebyshev's inequality, with probability at least 999/1000, $\mathbf{T} \geq \mathbb{E}[\mathbf{T}] - \sqrt{1000}\sqrt{\text{Var}(\mathbf{T})} \geq cn \log n$ for a small enough constant c. Thus, if we make

$< cn \log n$ swaps, with probability at least 999/1000 there will be $\sqrt{n}$ cards that are never swapped.

If we start from an ordered deck there are thus $\sqrt{n}$ consecutive cards which are in order after our shuffle with probability 999/1000. Consider instead a deck that is ordered instead according to a random permutation. Any consecutive run of $\sqrt{n}$ cards is in order with probability $1/\sqrt{n}!$. Additionally, via a union bound over $n - \sqrt{n}$ such sets, the probability of any of them being in order is at most $\frac{n}{\sqrt{n}!} \ll 1/1000$ for large n.

Thus, letting A be the event that there is a run of $\sqrt{n}$ consecutive cards in order, we have:

$$\|q^t - \pi\|_{TV} \geq |q^t(A) - \pi(A)| \geq 999/1000 - 1/1000 > 99/100.$$

This completes the proof.

3. Random Walks and Leverage Scores (8 points)

Consider a random walk on a connected, undirected graph. Let $h_{u,v}$ be the expected number of steps required to reach node v when starting from node u. Define $h_{u,u} = 0$ for all u. We will prove that the effective resistance of edge u, v, $\tau_{u,v} = b_{u,v}^T L^+ b_{u,v}$ is exactly $\tau_{u,v} = \frac{h_{u,v}+h_{v,u}}{2m}$, where m is the number of edges in the graph.

1. (2 points) Let $\mathcal{N}(u)$ be the neighborhood of node u in the graph, and $d_u = |\mathcal{N}(u)|$ be the degree. Fix some node v and argue that for any $u \neq v$,

$$h_{u,v} = 1 + \frac{1}{d_u} \cdot \sum_{w \in \mathcal{N}(u)} h_{w,v}.$$

This gives $n-1$ linear equations (one for each $u \neq v$) that the $h_{u,v}$ values must satisfy. Argue that the values of $h_{u,v}$ are the unique solutions to this set of linear equations.

The formula holds since the random walk transitions from u to some $w \in \mathcal{N}(u)$ in its first step. Thus, the expected time to reach v is equal to the one step from u to w plus the expected time to reach v from each w times the probability of stepping to that w in the first step. I.e., $h_{u,v} = 1 + \sum_{w \in \mathcal{N}(u)} \frac{1}{d_u} \cdot h_{w,v}$, as required.

This system of equations has $n-1$ unknowns ($h_{u,v}$ for all $u \neq v$) and $n-1$ equations: $h_{u,v} - \frac{1}{d_u} \cdot \sum_{w \in \mathcal{N}(u)} h_{w,v} = 1$ for all $u \neq v$. Thus, we just need to verify that these equations are linearly independent, which will imply that there is a unique solution. Let $A \in \mathbb{R}^{n-1 \times n-1}$ be the matrix corresponding to this system of equations. We need to show that A is full rank. A is 1 on the diagonal and $-1/d_u$ for all off diagonal entries in the u^{th} row corresponding to neighbors of u. Scaling rows does not change the rank, so A has the same rank as the matrix A' which is d_u on the diagonal and has all -1s off the diagonal – in the position of edges in the graph. Observe that A' is the $n-1 \times n-1$ submatrix of the $n \times n$ graph Laplacian L with the row and column corresponding to v removed.

Now, any graph Laplacian L has a one-dimensional null space spanned by the all ones vector. Assume for the sake of contradiction that A' has a null space. Then there is some nonzero vector $w' \in \mathbb{R}^{n-1}$ with $A'w' = 0$. Let $w \in \mathbb{R}^n$ be equal to w' but with a 0 appended in its last position. Then we can see that $Lw = [A'w', \langle L_{v,:}, w \rangle] = [0, \langle L_{v,:}, w \rangle]$. I.e., Lw is a vector with possibly just a single non-zero entry in its last position. Such a vector cannot exist: if $\langle L_{v,:}, w \rangle = 0$, it would mean that w is in the null-space of L which is not possible, since L

has a one-dimensional nullspace spanned by the all ones vector, and since w is not spanned by the all ones vector. Alternatively, if $\langle L_{v,:}, w\rangle \neq 0$, then $1^T L w = \langle L_{v,:}, w\rangle \neq 0$, which is impossible since $1^T L = 0$.

2. (2 points) View the graph as a resistor network with unit resistance on each edge. For any vertex v, consider an electrical flow in which d_u units of current are introduced at each vertex u, and all $\sum_{u\in V} d_u = 2m$ units of current are removed at vertex v. I.e., letting $B \in \mathbb{R}^{m\times n}$ be the vertex-edge incidence matrix and $f^e \in \mathbb{R}^m$ be the flow, we have $B^T f^e = \chi_v$ where $\chi_v(v) = d_v - 2m$ and $\chi_v(u) = d_u$ for $u \neq v$. Prove that in this flow, letting $\phi(u)$ be the voltage of vertex u,

$$d_u = d_u \cdot [\phi(u) - \phi(v)] - \sum_{w\in\mathcal{N}(u)} [\phi(w) - \phi(v)].$$

For any $u \neq v$, the total flow out of node u is d_u. Also, since all resistors have unit resistance, the flow along edge (u, w) is equal to $\phi(u) - \phi(w)$. Thus, we have:

$$\begin{aligned} d_u = \sum_{w\in\mathcal{N}(u)} [\phi(u) - \phi(w)] &= \sum_{w\in\mathcal{N}(u)} [\phi(u) - \phi(v)] - \sum_{w\in\mathcal{N}(u)} [\phi(w) - \phi(v)] \\ &= d_u \cdot [\phi(u) - \phi(v)] - \sum_{w\in\mathcal{N}(u)} [\phi(w) - \phi(v)]. \end{aligned}$$

3. (2 points) Use the above to conclude that $h_{u,v} = \phi(u)-\phi(v)$, where $\phi(\cdot)$ is the voltage function as in part (2) for the electrical flow f^e with $B^T f^e = \chi_v$.

   Dividing each side of the equation in (2) by d_u and rearranging terms we have:

$$[\phi(u) - \phi(v)] = 1 + \frac{1}{d_u} \sum_{w\in\mathcal{N}(u)} [\phi(w) - \phi(v)].$$

   Thus, these equations are identical to those in part (1) and since we argued this system of equations has a unique solution, we have $[\phi(u) - \phi(v)] = h_{u,v}$ for all $u \neq v$.

4. (2 points) Complete the proof, showing that $\tau_{u,v} = \frac{h_{u,v}+h_{v,u}}{2m}$. **Hint:** Given two electrical flows $f_1 = B^T x_1$ and $f_2 = B^T x_2$, what is the electrical flow with $f = B^T(x_1 - x_2)$?

   For a vertex demand vector x, the induced voltages are given by L^+x. This, for a vertex demand vector $x_1 - x_2$ the voltages are given by $L^+x_1 - L^+x_2$. I.e., the voltages are linear in the demand vectors.

   Observe that for $x = \chi_v - \chi_u$, $x(u) = 2m$, $x(v) = -2m$ and $x(w) = 0$ for all other w. $\frac{1}{2m}x$ is thus the unit demand flow from u to v. We know that the effective resistance is exactly the voltage drop $\phi(u) - \phi(v)$ in this flow. Thus, letting $\phi_v = L^+\chi_v$, $\phi_u = L^+\chi_u$, and $b_{u,v}$ be the vector with a 1 at position u and a -1 at position v, we have:

$$\tau_{u,v} = b_{u,v}^T[L^+ \cdot \frac{1}{2m}x] = \frac{1}{2m} b_{u,v}^T L^+(\chi_v - \chi_u) = \frac{b_{u,v}^T\phi_v - b_{u,v}^T\phi_u}{2m}.$$

   Now, by part (3), $b_{u,v}^T\phi_v = h_{u,v}$ and $b_{u,v}^T\phi_u = -h_{v,u}$ and , which finally gives

$$\tau_{u,v} = \frac{h_{u,v} + h_{v,u}}{2m}.$$