6.842 Randomness and Computation February 23, 2022

Homework 1

Yuchong Pan *MIT ID: 911346847*

1. *Collaborators and sources:* none.

Proof. We construct an approximation scheme $\mathcal{B}$ for f as follows: On input (x, ε, δ), run $\mathcal{A}(x, \varepsilon)$ independently for $k := \lceil 12 \log(1/\delta) \rceil$ times with outputs $y_1, \ldots, y_k$, and output a median of $y_1, \ldots, y_k$.

Let $t_{\mathcal{A}}(x, \varepsilon)$ be the running time of $\mathcal{A}$ on input (x, ε). Then $\mathcal{B}$ runs in $O(kt_{\mathcal{A}}(x, \varepsilon)) = O(\log(1/\delta) t_{\mathcal{A}}(x, \varepsilon))$. Since $\mathcal{A}$ runs in time polynomial in $1/\varepsilon$ and $|x|$, then $\mathcal{B}$ runs in time polynomial in $1/\varepsilon$, $|x|$ and $\log(1/\delta)$.

By the definition of medians, if more than half of $y_1, \ldots, y_k$ fall in $[f(x)/(1+\varepsilon), f(x)(1+\varepsilon)]$, then $\mathcal{B}(x, \varepsilon, \delta) \in [f(x)/(1+\varepsilon), f(x)(1+\varepsilon)]$. Let $X_1, \ldots, X_k \in \{0, 1\}$ be random variables so that $X_i = 1$ with probability $p := \mathbb{P}[\mathcal{A}(x, \varepsilon) \notin [f(x)/(1+\varepsilon), f(x)(1+\varepsilon)]] \leq 1 - 3/4 = 1/4$. Then $\sum_{i=1}^k \mathbb{E}[X_i] = kp \leq k/4$. Therefore,

$$
\begin{aligned}
&\mathbb{P}\left[\mathcal{B}(x, \varepsilon, \delta) \notin \left[\frac{f(x)}{1+\varepsilon}, f(x)(1+\varepsilon)\right]\right] \\
&\leq \mathbb{P}\left[\text{at least half of } y_1, \ldots, y_k \text{ do not fall in } \left[\frac{f(x)}{1+\varepsilon}, f(x)(1+\varepsilon)\right]\right] \\
&= \mathbb{P}\left[\sum_{i=1}^k X_i \geq \frac{k}{2}\right] \\
&= \mathbb{P}\left[\sum_{i=1}^k X_i \geq (1+1) \cdot \frac{k}{4}\right] \\
&\leq e^{-\frac{k/4}{3}} && \text{(Chernoff bound)} \\
&= e^{-\frac{\lceil 12 \log \frac{1}{\delta} \rceil}{12}} \leq e^{-\frac{12 \log \frac{1}{\delta}}{12}} = \delta.
\end{aligned}
$$

Therefore,

$$
\mathbb{P}\left[\mathcal{B}(x, \varepsilon, \delta) \in \left[\frac{f(x)}{1+\varepsilon}, f(x)(1+\varepsilon)\right]\right] = 1 - \mathbb{P}\left[\mathcal{B}(x, \varepsilon, \delta) \notin \left[\frac{f(x)}{1+\varepsilon}, f(x)(1+\varepsilon)\right]\right] \geq 1 - \delta.
$$

This completes the proof. □

2. *Collaborators and sources:* none.

Proof. Suppose $\binom{m}{t}2^{1-\binom{t}{2}} < 1$. To prove $R(t) > m$, it suffices to show that there exists a 2-edge-coloring of K_m such that for all $S \subset V(K_m)$ of size t, $E(K_m[S])$ is not monochromatic. We randomly color the edges of K_m red or blue, independently and equiprobably. For each $S \subset V(K_m)$ of size t, there are exactly $2^{\binom{t}{2}}$ two-colorings of $E(K_m[S])$, amongst which two are monochromatic colorings (all red and all blue), so

$$\mathbb{P}\left[E\left(K_m[S]\right) \text{ is monochromatic}\right] = \frac{2}{2^{\binom{t}{2}}} = 2^{1-\binom{t}{2}}.$$

By the union bound,

$$\begin{aligned}
&\mathbb{P}\left[\exists S \subset V\left(K_m\right), |S| = t, E\left(K_m[S]\right) \text{ is monochromatic}\right] \\
&\leq \sum_{\substack{S \subset V(K_m) \\ |S| = t}} \mathbb{P}\left[E\left(K_m[S]\right) \text{ is monochromatic}\right] \\
&= \binom{m}{t} 2^{1-\binom{t}{2}} \\
&< 1.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\mathbb{P}\left[\forall S \subset V\left(K_m\right) \text{of size } t, E\left(K_m[S]\right) \text{ is not monochromatic}\right] \\
&= 1 - \mathbb{P}\left[\exists S \subset V\left(K_m\right), |S| = t, E\left(K_m[S]\right) \text{ is monochromatic}\right] \\
&> 1 - 1 = 0.
\end{aligned}$$

This proves that there exists a 2-edge-coloring of K_m such that for all $S \subset V(K_m)$ of size t, $E(K_m[S])$ is not monochromatic. The proof is complete. $\square$

3. *Collaborators and sources:* none.

Proof. Suppose that a Boolean function f can be computed by a randomized polynomial-sized circuit family $\mathcal{C} = \{C_n\}_{n\in\mathbb{N}}$. Let $n \in \mathbb{N}$. Let $m \in \mathbb{N}$ be the number of random input bits. Define a $2^n \times 2^m$ matrix M such that

- each row represents a possible combination of inputs $x_1, \ldots, x_n \in \{0, 1\}$;
- each column represents a possible combination of random input bits $r_1, \ldots, r_m \in \{0, 1\}$;
- the entry at the row representing $x_1, \ldots, x_n$ and at the column representing $r_1, \ldots, r_m$ equals the value of C_n on inputs $x_1, \ldots, x_n$ with random input bits $r_1, \ldots, r_m$.

By the definition of polynomial-sized circuit families, each row has at least half of the entries equal to 1, so the total number of 1-entries is at least $2^n \cdot 2^m/2 = 2^{n+m-1}$. Therefore, there exists a column, representing $r_1^*, \ldots, r_m^*$, with at least half of the entries equal to 1; otherwise every column has fewer than half of the entries equal to 1, so the total number of 1-entries is less than $2^m \cdot 2^n/2 = 2^{n+m-1}$, a contradiction.

Construct a deterministic circuit $D_n^{(1)}$ by hard-wiring random input bits $r_1 = r_1^*, \ldots, r_m = r_m^*$. Remove the column representing $r_1^*, \ldots, r_m^*$ and each row representing $x_1, \ldots, x_n$ such that the corresponding entry equals 1, resulting in a new matrix M'. Note that this removes at least half of the rows. We claim that each row of M' has at least half of the entries equal to 1; to see this, we note that the number of columns is decreased by 1, and that the number of 1-entries in each remaining row remains the same. Therefore, we apply the same argument and recurse, until every remaining row is all-zero, obtaining circuits $D_n^{(1)}, \ldots, D_n^{(k)}$. Finally, construct a deterministic circuit D_n by taking the "or" of $D_n^{(1)}, \ldots, D_n^{(k)}$.

Since we remove at least half of the rows in each iteration, then there are at most $O(\log 2^n) = O(n)$ iterations, i.e., $k = O(n)$. Since C_n is polynomial-sized, then the size of C_n is upper bounded by $p(n)$. For each $i \in [k]$, since $D_n^{(i)}$ is obtained by hard-wiring random input bits in C_n, then the size of $D_n^{(i)}$ is upper bounded by $p(n)$. Finally, since D_n is obtained by taking the "or" of $D_n^{(1)}, \ldots, D_n^{(k)}$, then the size of D is upper bounded by $kp(n) + k = O(np(n))$, so D_n is polynomial-sized.

Let $\mathcal{D} = \{D_n\}_{n\in\mathbb{N}}$. We show that $\mathcal{D}$ computes f. Let $x_1, \ldots, x_n \in \{0, 1\}$. If $f(x_1, \ldots, x_n) = 0$, then the row of the original matrix M representing $x_1, \ldots, x_n$ is all-zero and hence never removed, so none of $D_n^{(1)}, D_n^{(2)}, \ldots$ outputs 0 on inputs $x_1, \ldots, x_n$, which implies that D_n outputs 0. If $f(x_1, \ldots, x_n) = 1$, then the row of the original matrix M representing $x_1, \ldots, x_n$ has at least half of the entries equal to 1, and is hence removed in some iteration, say the i^{th} iteration, so $D_n^{(i)}$ outputs 1 on inputs $x_1, \ldots, x_n$, which implies that D_n outputs 1. This shows that D_n outputs $f(x_1, \ldots, x_n)$ for all combinations of inputs $x_1, \ldots, x_n \in \{0, 1\}$. Therefore, $\mathcal{D}$ is a deterministic polynomial-sized circuit family which computes f, completing the proof. □

4. *Collaborators and sources:* none.

Proof. Suppose $e(\Delta\delta+1)(1-1/k)^{\delta} < 1$. WLOG, we assume that each vertex has out-degree δ by removing an outgoing edge from v whenever a vertex v has out-degree greater than δ; this does not increase the in-degree of any vertex, so the maximum in-degree of the resulting graph is at most Δ.

Let $f : V \to \{0, \dots, k-1\}$ be a random coloring of V obtained by choosing for each $v \in V$, $f(v) \in \{0, \dots, k-1\}$ independently according to a uniform distribution. For each $v \in V$, let A_v be the event that there does not exist $u \in V$ such that $(v, u) \in E$ and $f(u) = f(v) + 1 \mod k$. For each $v \in V$, we denote by $N^+(v)$ the set of out-neighbors of v, and denote by $N^-(v)$ the set of in-neighbors of v. Then for each $v \in V$,

$$\begin{aligned}
\mathbb{P}[A_v] &= \mathbb{P}\left[f(u) \neq f(v) + 1 \mod k \;\; \forall u \in N^+(v)\right] \\
&= \sum_{i=0}^{k-1} \mathbb{P}[f(v) = i] \mathbb{P}\left[f(u) \neq f(v) + 1 \mod k \;\; \forall u \in N^+(v) \;\middle|\; f(v) = i\right] \\
&= \sum_{i=0}^{k-1} \mathbb{P}[f(v) = i] \mathbb{P}\left[f(u) \neq i + 1 \mod k \;\; \forall u \in N^+(v)\right] && \text{(independence)} \\
&= \sum_{i=0}^{k-1} \mathbb{P}[f(v) = i] \prod_{u \in N^+(v)} \mathbb{P}\left[f(u) \neq i + 1 \mod k\right] && \text{(independence)} \\
&= k \cdot \frac{1}{k} \cdot \left(1 - \frac{1}{k}\right)^{\delta} = \left(1 - \frac{1}{k}\right)^{\delta}.
\end{aligned}$$

Let $v \in V$. To simplify notations, let $N_*^+(v) = N^+(v) \cup \{v\}$ and $\overline{N_*^+(v)} = V \setminus N_*^+(v)$. Let $I(v) = \{u \in V : u \notin N_*^+(v), N^+(u) \cap N^+(v) = \emptyset\} \subset \overline{N_*^+(v)}$. We show that A_v is independent of all events A_u with $u \in I(v)$. Let $v_1, \dots, v_\ell \in I(v)$. By total independence, for all $\{a_w : w \in V\} \in \{0, \dots, k-1\}^V$,

$$\begin{aligned}
\mathbb{P}\left[f(w) = a_w \; \forall w \in V\right] = \prod_{w \in V} \mathbb{P}\left[f(w) = a_w\right] &= \prod_{w \in N_*^+(v)} \mathbb{P}\left[f(w) = a_w\right] \prod_{w \in \overline{N_*^+(v)}} \mathbb{P}\left[f(w) = a_w\right] \\
&= \mathbb{P}\left[f(w) = a_w \; \forall w \in N_*^+(v)\right] \mathbb{P}\left[f(w) = a_w \; \forall w \in \overline{N_*^+(v)}\right].
\end{aligned}$$

For $i \in [\ell]$, since $v_i \in I(v)$, then A_{v_i} depends upon the values of f on $\overline{N_*^+(v)}$ only. Moreover, A_v depends upon the values of f on $N_*^+(v)$ only. Hence, for all $\{a_w : w \in V\} \in \{0, \dots, k-1\}^V$,

$$\begin{aligned}
&\mathbb{P}\left[A_v \cap \bigcap_{i=1}^{\ell} A_{v_i} \;\middle|\; f(w) = a_w \; \forall w \in V\right] \\
&= \left(\prod_{u \in N^+(v)} \mathbb{1}[a_u \neq a_v + 1 \mod k]\right) \left(\prod_{i=1}^{\ell} \prod_{u \in N^+(v_i)} \mathbb{1}\left[a_u \neq a_{v_i} + 1 \mod k\right]\right) \\
&= \mathbb{P}\left[A_v \mid f(w) = a_w \; \forall w \in V\right] \mathbb{P}\left[\bigcap_{i=1}^{\ell} A_{v_i} \;\middle|\; f(w) = a_w \; \forall w \in V\right]
\end{aligned}$$

$$= \mathbb{P}\left[A_v \;\middle|\; f(w) = a_w \ \forall w \in N_*^+(v)\right] \mathbb{P}\left[\bigcap_{i=1}^{\ell} A_{v_i} \;\middle|\; f(w) = a_w \ \forall w \in \overline{N_*^+(v)}\right].$$

Therefore,

$$\begin{aligned}
&\mathbb{P}\left[A_v \cap \bigcap_{i=1}^{\ell} A_{v_i}\right] \\
&= \sum_{\{a_w\}_{w\in V} \in \{0,\ldots,k-1\}^V} \mathbb{P}\left[f(w) = a_w \ \forall w \in V\right] \mathbb{P}\left[A_v \cap \bigcap_{i=1}^{\ell} A_{v_i} \;\middle|\; f(w) = a_w \ \forall w \in V\right] \\
&= \sum_{\{a_w\}_{w\in V} \in \{0,\ldots,k-1\}^V} \mathbb{P}\left[f(w) = a_w \ \forall w \in N_*^+(v)\right] \mathbb{P}\left[f(w) = a_w \ \forall w \in \overline{N_*^+(v)}\right] \\
&\qquad \mathbb{P}\left[A_v \;\middle|\; f(w) = a_w \ \forall w \in N_*^+(v)\right] \mathbb{P}\left[\bigcap_{i=1}^{\ell} A_{v_i} \;\middle|\; f(w) = a_w \ \forall w \in \overline{N_*^+(v)}\right] \\
&= \left(\sum_{\substack{a_w \in \{0,\ldots,k-1\} \\ \forall w \in N_*^+(v)}} \mathbb{P}\left[f(w) = a_w \ \forall w \in N_*^+(v)\right] \mathbb{P}\left[A_v \;\middle|\; f(w) = a_w \ \forall w \in N_*^+(v)\right]\right) \\
&\quad \left(\sum_{\substack{a_w \in \{0,\ldots,k-1\} \\ \forall w \in \overline{N_*^+(v)}}} \mathbb{P}\left[f(w) = a_w \ \forall w \in \overline{N_*^+(v)}\right] \mathbb{P}\left[\bigcap_{i=1}^{\ell} A_{v_i} \;\middle|\; f(w) = a_w \ \forall w \in \overline{N_*^+(v)}\right]\right) \\
&= \mathbb{P}\left[A_v\right] \mathbb{P}\left[\bigcap_{i=1}^{\ell} A_{v_i}\right].
\end{aligned}$$

This proves that A_v is independent of all events A_u with $u \in I(v)$. Since D is simple, then the maximum degree of the dependency digraph of $\{A_v : v \in V\}$ is at most

$$\begin{aligned}
\max_{v\in V} |V \setminus (I(v) \cup \{v\})| &= \left|\left\{u \in V : u \in N^+(v) \ \vee \ \left(u \neq v \ \wedge \ N^+(u) \cap N^+(v) \neq \emptyset\right)\right\}\right| \\
&= \left|N^+(v) \cup \bigcup_{u \in N^+(v)} \left(N^-(u) \setminus \{v\}\right)\right| \\
&\leq \left|N^+(v)\right| + \sum_{u \in N^+(v)} \left(\left|N^-(u)\right| - 1\right) \\
&\leq \delta + \delta(\Delta - 1) = \Delta\delta.
\end{aligned}$$

Since $e(\Delta\delta + 1)(1 - 1/k)^{\delta} < 1$, then by the Lovász local lemma,

$$\mathbb{P}[\forall v \in V, \exists u \in N^+(v) \text{ s.t. } f(u) = f(v) + 1 \mod k] = \mathbb{P}\left[\bigcap_{v \in V} \overline{A_v}\right] > 0.$$

This implies that there exists a coloring $f : V \to \{0, \ldots, k-1\}$ such that for all $v \in V$, there exists $u \in N^+(v)$ such that $f(u) = f(v) + 1 \mod k$.

$\ell \leftarrow 1$
pick $v_1 \in V$ arbitrarily
while *true* **do**
    pick $u \in N^+(v)$ such that $f(u) = f(v) + 1 \mod k$
    if $\exists i \in [\ell]$ *s.t.* $v_i = u$ **then**
        return $(v_i, \ldots, v_\ell, u)$

Algorithm 1: An algorithm which finds a simple directed cycle whose length is a multiple of k, given a digraph D with $e(\Delta\delta + 1)(1 - 1/k)^\delta < 1$ and a coloring $f : V \to \{0, \ldots, k-1\}$ such that for all $v \in V$, there exists $u \in N^+(v)$ such that $f(u) = f(v) + 1 \mod k$, where the out-degree of every vertex in D is δ, and the maximum in-degree of D is Δ.

Consider Algorithm 1. Since the number of vertices in D is finite, the algorithm eventually visits a vertex which has been visited before, so Algorithm 1 terminates. We show that the returned walk $(v_i, \ldots, v_\ell, u)$ is a simple directed cycle whose length is a multiple of k. By definition, since u is the first vertex on this list that has occurred before, then $(v_i, \ldots, v_\ell, u)$ is a simple directed cycle. Since $f(v_j) = f(v_{j-1}) + 1 \mod k$ for each $j \in \{i+1, \ldots, \ell\}$, then an inductive argument implies that

$$f(v_i) = f(u) = f(v_\ell) + 1 \mod k = f(v_i) + \ell - i + 1 \mod k.$$

Therefore, the length of the simple directed cyle $(v_i, \ldots, v_\ell, u)$, which equals $\ell - i + 1$, is a multiple of k. This completes the proof. $\square$