CS5160: Topics in Computing
Problem Set 3

Taha Adeel Mohammed CS20BTECH11052

(Crediting the course)

1. Given a monotone Boolean function $f : \{0,1\}^n \to \{0,1\}$ and an input $x \in \{0,1\}^n$, say that the i-th bit x_i of x is "correct" for f if $f(x) = x_i$. Let $c(f)$ denote the expected number of "correct" bits in a uniformly random string x. Show that $c(f) = (n + \mathrm{Inf}(f))/2$. (10 points)

For any $x \in \{0,1\}^n$, we have two cases $\forall\ i \in [n]$:

Case 1: x_i is a sensitive bit
Claim: $f(x) = x_i$, i.e., x_i is a "correct" bit for f.
Proof by contradiction: Assume that $f(x) \neq x_i$.

1. If $x_i = 0$, then $f(x) = 1$ and $f(x^{(i)}) = 0$. But $f(x) > f(x^{(i)})$ while $x \leq x^{(i)}$
2. If $x_i = 1$, then $f(x) = 0$ and $f(x^{(i)}) = 1$. But $f(x) < f(x^{(i)})$ while $x \geq x^{(i)}$

In both cases, we get a contradiction, since f is monotone. Hence our claim is true.

Case 2: x_i is not a sensitive bit
Claim: Either $f(x) = x_i$ or $f(x^{(i)}) = x_i^{(i)}$, i.e., i-th bit is correct for one of x or $x^{(i)}$.
Proof: Since x_i is not a sensitive bit, we have $f(x) = f(x^{(i)})$.

1. If $f(x) = f(x^{(i)}) = x_i$, then i is a correct bit for x and not for $x^{(i)}$.
2. If $f(x) = f(x^{(i)}) = x_i^{(i)}$, then i is a correct bit for $x^{(i)}$ and not for x.

Hence our claim is true. i.e. $\mathbb{E}\left[f(x) = x_i | x_i \text{ is not sensitive}\right] = \frac{1}{2}$.

Now we can write the expected number of correct bits, $c(f)$, as

$$
\begin{aligned}
c(f) &= \mathop{\mathbb{E}}_{x \sim \{0,1\}^n}\left[\sum_{i=1}^{n} \mathbb{I}[f(x) = x_i]\right] && (1)\\
&= \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \left(s(f,x) + (n - s(f,x)) \cdot \mathbb{E}[f(x_i) = x_i | x_i \text{ is not sensitive}]\right) && (2)\\
&= \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \left(s(f,x) + \frac{(n - s(f,x))}{2}\right) && (3)\\
&= \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \frac{n + s(f,x)}{2} && (4)\\
\implies c(f) &= \frac{n + as(f)}{2} = \frac{n + \mathrm{Inf}(f)}{2} && (5)
\end{aligned}
$$

□

2. Let $f : \{-1,1\}^n \to \{-1,1\}$. Give a Fourier formula for the expression

$$\mathbb{E}_{x,y,z\sim\{-1,1\}^n}[f(x)f(y)f(z)f(w)],$$

where x, y, z are chosen uniformly at random from $\{-1,1\}^n$ and $w = x \oplus y \oplus z$, i.e., $w_i = x_i y_i z_i$ for all $i \in [n]$. **(10 points)**

Substituting the Fourier representation of f in the given expression, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x,y,z\sim\{-1,1\}^n}[f(x)f(y)f(z)f(w)] \\
&= \mathop{\mathbb{E}}_{x,y,z}\left[\left(\sum_{S\subseteq[n]}\hat{f}(S)\chi_S(x)\right)\left(\sum_{T\subseteq[n]}\hat{f}(T)\chi_T(y)\right)\left(\sum_{U\subseteq[n]}\hat{f}(U)\chi_U(z)\right)\left(\sum_{V\subseteq[n]}\hat{f}(V)\chi_V(w)\right)\right] \\
&= \mathop{\mathbb{E}}_{x,y,z}\left[\sum_{S,T,U,V}\hat{f}(S)\hat{f}(T)\hat{f}(U)\hat{f}(V)\cdot\prod_{i\in S}x_i\cdot\prod_{i\in T}y_i\cdot\prod_{i\in U}z_i\cdot\prod_{i\in V}(x_iy_iz_i)\right] \\
&= \mathop{\mathbb{E}}_{x,y,z}\left[\sum_{S,T,U,V}\hat{f}(S)\hat{f}(T)\hat{f}(U)\hat{f}(V)\cdot\prod_{i\in S\Delta V}x_i\cdot\prod_{i\in T\Delta V}y_i\cdot\prod_{i\in U\Delta V}z_i\right] \text{(wlog, if } i\in S\cap V, x_i^2=1) \\
&= \sum_{S,T,U,V}\hat{f}(S)\hat{f}(T)\hat{f}(U)\hat{f}(V)\cdot\mathop{\mathbb{E}}_{x,y,z}\left[\prod_{i\in S\Delta V}x_i\cdot\prod_{i\in T\Delta V}y_i\cdot\prod_{i\in U\Delta V}z_i\right] \text{(linearity of expectation)} \\
&= \sum_{S,T,U,V}\hat{f}(S)\hat{f}(T)\hat{f}(U)\hat{f}(V)\prod_{i\in S\Delta V}\mathop{\mathbb{E}}_x[x_i]\prod_{i\in T\Delta V}\mathop{\mathbb{E}}_y[y_i]\prod_{i\in U\Delta V}\mathop{\mathbb{E}}_z[z_i] \quad \text{(since } x_i,y_i,z_i \text{ independent)} \\
&= \sum_S \hat{f}(S)^4 \qquad\qquad \text{(since } \mathop{\mathbb{E}}_x[x_i]=0, \text{then if } S\Delta V, T\Delta V, U\Delta V \neq \emptyset)
\end{aligned}$$

Therefore we have

$$\boxed{\mathop{\mathbb{E}}_{x,y,z\sim\{-1,1\}^n}[f(x)f(y)f(z)f(w)] = \sum_S \hat{f}(S)^4}$$

$\square$

3. Let $\rho \in [-1,1]$ and $x \in \{-1,1\}^n$. Recall we say $y \sim N_\rho(x)$ to denote that the random string y is sampled as follows: $y_i = x_i$ with probability $(1+\rho)/2$ and $y_i = -x_i$ with probability $(1-\rho)/2$. For a Boolean function $f : \{-1,1\}^n \to \{-1,1\}$, we define noise stability of f at ρ as follows:

$$\mathrm{Stab}_\rho(f) = \mathbb{E}_{x\sim\{-1,1\}^n, y\sim N_\rho(x)}[f(x)f(y)].$$

Give a Fourier formula for $\mathrm{Stab}_\rho(f)$. **(10 points)**

Simplifying the given expression for $\mathrm{Stab}_\rho(f)$, we have

$$\begin{aligned}
\mathrm{Stab}_\rho(f) &= \mathop{\mathbb{E}}_{x\sim\{-1,1\}^n, y\sim N_\rho(x)}[f(x)f(y)] \\
&= \mathop{\mathbb{E}}_{x}\left[f(x)\cdot \mathop{\mathbb{E}}_{y}[f(y)]\right] && \text{(linearity of expectation)} \\
&= \mathop{\mathbb{E}}_{x}\left[f(x)\cdot \mathop{\mathbb{E}}_{y}\left[\sum_{S\subseteq[n]} \hat{f}(S)\cdot \prod_{i\in S} y_i\right]\right] && \text{(Fourier representation of } f) \\
&= \mathop{\mathbb{E}}_{x}\left[f(x)\cdot \sum_{S\subseteq[n]} \hat{f}(S)\cdot \mathop{\mathbb{E}}_{y}\left[\prod_{i\in S} y_i\right]\right] && \text{(linearity of expectation)}
\end{aligned}$$

In class, we have seen that

$$\mathop{\mathbb{E}}_{y\sim N_\rho(x)}\left[\prod_{i\in S} y_i\right] = \rho^{|S|}\cdot \chi_s(y)$$

Using this, we can write

$$\begin{aligned}
\mathrm{Stab}_\rho(f) &= \mathop{\mathbb{E}}_{x}\left[f(x)\cdot \sum_{S\subseteq[n]} \hat{f}(S)\cdot \rho^{|S|}\cdot \chi_s(x)\right] \\
&= \sum_{S\subseteq[n]} \hat{f}(S)\cdot \rho^{|S|}\cdot \mathop{\mathbb{E}}_{x}[f(x)\cdot \chi_s(x)] && \text{(linearity of expectation)} \\
&= \sum_{S\subseteq[n]} \hat{f}(S)\cdot \rho^{|S|}\cdot \hat{f}(S) && (\mathbb{E}[\chi_s\chi_t] = 0 \text{ if } S\neq T \text{ and } 1 \text{ otherwise}) \\
&= \sum_{S\subseteq[n]} \hat{f}(S)^2\cdot \rho^{|S|}
\end{aligned}$$

Therefore we have

$$\boxed{\mathrm{Stab}_\rho(f) = \sum_{S\subseteq[n]} \hat{f}(S)^2\cdot \rho^{|S|}}$$

$\square$

4. **Let $\epsilon > 0$. Prove that for every Boolean function $f : \{-1,1\}^n \to \{-1,1\}$, there exists a Boolean function $g : \{-1,1\}^n \to \{-1,1\}$ depending on at most $2^{O(\mathrm{as}(f)/\epsilon)}$ variables such that g differs from f on at most an ϵ fraction of inputs. Recall $\mathrm{as}(f)$ denotes the average sensitivity of f. (15 points)**

5. **A tournament is a directed graph obtained by assigning a direction to each edge in an undirected complete graph. (See Figure 1.) We say that a tournament is acyclic if it contains no directed cycles. Note that a tournament can be represented by a string in $\{0,1\}^{\binom{n}{2}}$, where every edge is represented by a bit and its value represents the orientation of the edge. Thus, we can define the following Boolean function $T_{\text{acyclic}} : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ such that $T_{\text{acyclic}}(x) = 1$ if and only if x defines an acyclic tournament.**
   Prove that $D(T_{\text{acyclic}}) \geq \binom{n}{2} - \frac{n}{2}$. Recall $D(f)$ is the deterministic decision tree complexity of f. (15 points)

Consider an input $x \in \{0,1\}^{\binom{n}{2}}$ such that $T_{\text{acyclic}}(x) = 1$. Since x defines an acyclic tournament, there exists a topological ordering for our given vertices. Let $v_1, v_2, \ldots, v_n$ be the vertices in the topological ordering.

Now consider the following adversial argument for any deterministic decision tree T that computes T_{acyclic}. We iteratively tell that while we have not queried all edges, if the edge (v_i, v_j) is queried, then we set $x_{(v_i,v_j)} = 1$ if $i < j$ and $x_{(v_i,v_j)} = 0$ otherwise.

We state that after querying for $i < \binom{n}{2}$ edges, there exist inputs $x, y \in \{0,1\}^{\binom{n}{2}}$ such that $T_{\text{acyclic}}(x) \neq T_{\text{acyclic}}(y)$ and T has queried for the same set of edges for both x and y.

1. For an unqueried edge (v_i, v_j), if $i < j$, and we set $x_{(v_i,v_j)} = 1$, then $T_{\text{acyclic}}(x) = 0$ since we have a cycle (v_j, v_i, v_j).
2. Or else if $x_{(v_i,v_j)} = 0$, then $T_{\text{acyclic}}(x) = 1$ since we have a topological ordering.

Therefore we need to query all the edges to get the correct answer. Hence we have

$$\boxed{D(T_{\text{acyclic}}) = \binom{n}{2}}$$

□

6. **Let T be a tournament and v be a vertex of T. We say that v is a *source* if all edges incident on v are directed *away* from it. Not every tournament has a source. Therefore we can consider the following Boolean function SRC : $\{0,1\}^{\binom{n}{2}} \to \{0,1\}$ defined as SRC$(x) = 1$ if and only if the tournament given by x has a source. Show that $D(\text{SRC}) = O(n)$.** **(15 points)**

We propose the following algorithm for querying the edges of the tournament T to find a source:

1. We maintain two sets *Possible* and *Rejected* containing the vertices that are possible sources and rejected sources respectively. Initially, $Possible = \{1, 2, \ldots, n\}$ and $Rejected = \emptyset$.
2. While $|Possible| > 1$: Take two vertices $u, v \in Possible$ and query the edge (u, v).
3. (Correctness proof) Whatever the direction of the edge (u, v), we can reject one of u or v as a possible source. Hence we remove one of them from *Possible* and add it to *Rejected*.
4. Therefore, w.l.o.g. if (u, v) is directed from u to v, update $Possible = Possible \setminus \{v\}$ and $Rejected = Rejected \cup \{v\}$.
5. Once $|Possible| = 1$, we query all the edges incident on the vertex in *Possible* and check if all of them are directed away from it. If so, we have found a source and $\text{SRC}(T) = 1$. Otherwise, $\text{SRC}(T) = 0$.

Analyzing the number of queries made by the algorithm, we have that step 2 does $n - 1$ queries, since in each iteration we remove one vertex from *Possible*. In step 5, we query at most $n - 1$ edges. Hence the total number of queries is $2n - 2 = O(n)$.

Hence, since we have that max cost of the algorithm is $2n - 2$, we have that the Deterministic Decision Tree complexity $\leq 2n - 2 = O(n)$. Therefore we have

$$\boxed{D(\text{SRC}) = O(n)}$$

□

7. **For $1 \leq t \leq n$, let Tht : $\{0,1\}^n \to \{0,1\}$ be the threshold function defined as follows:**

$$\text{Tht}(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} x_i \geq t \\ 0 & \text{otherwise} \end{cases}$$

Prove that deg(Tht) $= n$, i.e., any polynomial representing Tht must have full degree n. **(15 points)**

In class, we have seen that

$$\deg(f) = n \text{ iff } |X^{\text{even}}| \neq |X^{\text{odd}}| \tag{6}$$

where $X^{\text{even}} = \{x | f(x) = 1 \text{ and } |x| \text{ is even}\}$ and $X^{\text{odd}} = \{x | f(x) = 1 \text{ and } |x| \text{ is odd}\}$.

Let Tht_t be the threshold function for $t \in [n]$. We have that

$$\mathrm{Tht}_t(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^n x_i \geq t \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

This implies that

$$X^{\text{even}} = \{x | \sum_{i=1}^{n} x_i \geq t \text{ and } |x| \text{ is even}\}$$
$$X^{\text{odd}} = \{x | \sum_{i=1}^{n} x_i \geq t \text{ and } |x| \text{ is odd}\}$$

By choosing $2k$ or $2k+1$ bits from n bits and setting them to 1, while the rest are set to 0, we can get all the elements of X^{even} and X^{odd} respectively. Hence we have

$$|X^{\text{even}}| = \sum_{k=\lceil t/2 \rceil}^{\lfloor n/2 \rfloor} \binom{n}{2k}$$
$$|X^{\text{odd}}| = \sum_{k=\lfloor t/2 \rfloor}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1}$$

Therefore we have

$$|X^{\text{even}}| - |X^{\text{odd}}| = \sum_{i=t}^{n} (-1)^i \cdot \binom{n}{i}$$
$$\implies |X^{\text{even}}| - |X^{\text{odd}}| = \sum_{i=0}^{n-t} (-1)^i \cdot \binom{n}{i} \qquad \left(\text{using } \binom{n}{i} = \binom{n}{n-i}\right)$$

Hence to prove $X^{\text{even}} \neq X^{\text{odd}}$, we need to show that

$$\sum_{i=0}^{k} (-1)^i \cdot \binom{n}{i} \neq 0, \text{ where } k = n - t, \forall\, t \in [n], \text{ i.e. } 0 \leq k \leq n-1; \tag{8}$$

We can prove this by induction on n.
Base case: For k = 0, it holds trivially.
Induction hypothesis: Assume that it holds for some $k \in [n-2]$. Then for $k+1$, we have

$$\begin{aligned}
\sum_{i=0}^{k+1}(-1)^i\binom{n}{i} &= (-1)^{k+1}\binom{n}{k+1} + \sum_{i=0}^{k}(-1)^i\binom{n}{i} \\
&= (-1)^{k+1}\binom{n}{k+1} + (-1)^k\binom{n-1}{k} \\
&= (-1)^{k+1}\left\{\binom{n}{k+1} - \binom{n-1}{k}\right\} \\
&= (-1)^{k+1}\binom{n-1}{k+1}
\end{aligned}$$

Hence the induction hypothesis holds.

Therefore we have that $|X^{\text{even}}| \neq |X^{\text{odd}}|$ for all $t \in [n]$. Hence deg $(\mathrm{Tht}_t) = n \;\forall\; t \in [n]$.