Lecture 9

Proposition 10. $R(\mathrm{Simon}_n) = \Omega(\sqrt{n})$.

We will need the following lemma

Lemma 6. *Let $f, T\colon D \coloneqq D_0 \dot\cup D_1 \subseteq \Sigma^n \to \{0,1\}$. Let $f(D_0) = \{0\}$ and $f(D_1) = \{1\}$. Suppose μ_0 is a distribution on D_0 and μ_1 is a distribution on D_1. Let μ denote the distribution on D such that $x \leftarrow \mu$ is defined by $b \leftarrow \{0,1\}$ and $x \leftarrow \mu_b$. Let $P_1 \subseteq D_1$. Suppose that for all $b \in \{0,1\}$,*

$$\Pr[T(x) = b \mid x \leftarrow \mu_0] = \Pr[T(x) = b \mid x \in P_1, x \leftarrow \mu_1]. \tag{94}$$

Then

$$\Pr[T(x) = f(x) \mid x \leftarrow \mu] \le \frac{1}{2} + \frac{1}{2}\Pr[x \notin P_1 \mid x \leftarrow \mu_1]. \tag{95}$$

Proof.

$$\begin{aligned}
&\Pr[T(x) = f(x) \mid x \leftarrow \mu] \\
=&\frac{1}{2}\Pr[T(x) = 0 \mid x \leftarrow \mu_0] + \frac{1}{2}\Pr[T(x) = 1 \mid x \leftarrow \mu_1] && \text{definition of } \mu \\
=&\frac{1}{2}\Pr[T(x) = 0 \mid x \leftarrow \mu_0] + \frac{1}{2}(\Pr[T(x) = 1 \mid x \in P_1, x \leftarrow \mu_1]\Pr[x \in P_1 \mid x \leftarrow \mu_1] \\
&\quad + \frac{1}{2}\Pr[T(x) = 1 \mid x \notin P_1, x \leftarrow \mu_1]\Pr[x \notin P_1 \mid x \leftarrow \mu_1]) && \text{law of total probability} \\
\le&\frac{1}{2}\Pr[T(x) = 0 \mid x \leftarrow \mu_0] + \frac{1}{2}\Pr[T(x) = 1 \mid x \leftarrow \mu_0] + \frac{1}{2}\Pr[x \notin P_1 \mid x \leftarrow \mu_1] && \text{by lemma condition} \\
=&\frac{1}{2} + \frac{1}{2}\Pr[x \notin P_1 \mid x \leftarrow \mu_1],
\end{aligned}$$

as required. □

Comment: Apply this lemma to $f = \mathrm{Simon}_n$ and T the (function induced by the) decision tree.

Proof of proposition 10. (A more rigorous version of de Wolf's exposition.) By the averaging argument/easy direction of Yao's principle (i.e., the arguments we used at the beginning of the randomized lower bound proof for OR_n), it suffices to show the following. Suppose T is a DDT and μ is a distribution over D, and

$$\Pr[T(x) = \mathrm{Simon}_n(x) \mid x \leftarrow \mu] \ge 2/3, \tag{96}$$

then the depth d of T is at least $\Omega(\sqrt{n})$. We also assume wlog that T never queries the x at the same index twice, i.e., in all paths from root to leaf, the labels of the nodes are distinct; because a DDT that does this can be simulated by another DDT of no greater depth that doesn't do this. We may also assume T is balanced, i.e., every root to leaf path is length d.

To define μ, we first define two distributions μ_0 and μ_1 on D_0 and D_1 respectively by the following sampling procedures. Then we define $x \leftarrow \mu$ by $b \leftarrow \{0,1\}$ and $x \leftarrow \mu_b$.

1. Definition of $x \leftarrow \mu_0$. For each $s \in \{0,1\}^k$, pick a distinct value in $\{0,1,\ldots,n-1\}$ for $x(s)$ uniformly at random. (So x is a uniformly random permutation of $\{0,1,\ldots,n-1\}$.)

2. Definition of $x \leftarrow \mu_1$. Pick $a \leftarrow \{0,1\}^k - \{0^k\}$, then for each set $\{s, s \oplus a\}$, where $s \in \{0,1\}^k$, pick a distinct value in $\{0,1,\ldots,n-1\}$ for $x(s) = x(s \oplus a)$ uniformly at random. Comment: the distribution defined is independent of how the "for each" loop is ordered.

If $x \leftarrow \mu_0$, the sequence of d responses to the d queries T makes is a uniformly random sequence of d distinct elements in $\{0,1,\ldots,n-1\}$.

Now consider the case $x \leftarrow \mu_1$. Let $t \in \{1,\ldots,d\}$. Let $v_1,\ldots,v_{t-1} \in \{0,1,\ldots,n-1\}$ be distinct. Let $s_1,\ldots s_t$ denote the sequence of indices that T queries on x given $x(s_1) = v_1, \ldots, x(s_{t-1}) = v_{t-1}$. (Note $s_1,\ldots,s_t$ are uniquely defined, in particular, s_1 is the label of the root of T.) Say the sequence $x(s_1),\ldots,x(s_t)$ is good if all its values are all distinct. Let μ_1 be the distribution μ conditioned on the second case. Then, writing Pr for probability over $x \leftarrow \mu_1$, we have

$$\begin{aligned}
&\Pr[x(s_1),\ldots,x(s_t) \text{ is good} \mid x(s_1) = v_1,\ldots,x(s_{t-1}) = v_{t-1}] \\
=&\Pr[x(s_t) \notin \{x(s_1) = v_1,\ldots,x(s_{t-1}) = v_{t-1}\} \mid x(s_1) = v_1,\ldots,x(s_{t-1}) = v_{t-1}] \\
=&\Pr[a(x) \notin \{s_1 \oplus s_t,\ldots,s_{t-1} \oplus s_t\} \mid x(s_1) = v_1,\ldots,x(s_{t-1}) = v_{t-1}] && a(x) = \text{the } a \text{ corresp. to } x
\end{aligned}$$

Comment: the point of conditioning like this is to explicitly see that s_t is *fixed* and not a function of x; without such conditioning, the queried indices are generally functions of x and we would need to argue why, e.g., we can't have $s_1 = 0^k$ and $s_t = a(x)$, so that $a(x)$ is always in $\{s_1 \oplus s_t\}$. This is why I have chosen to be more rigorous here than de Wolf's exposition. The set $\{s_1 \oplus s_t, \ldots, s_{t-1} \oplus s_t\}$ in the last equation is the set that contains $t-1$ elements: $s_i \oplus s_t$ where $i \in [t-1]$. In class, I incorrectly thought $\{s_1 \oplus s_t, \ldots, s_{t-1} \oplus s_t\}$ was a set containing $\binom{t-1}{2}$ elements, which led to the confusion later on that got corrected by Victor.

Since the v_is are distinct, conditioning on $x(s_1) = v_1, \ldots, x(s_{k-1}) = v_{k-1}$ implies that $a(x)$ cannot belong to $\{s_i \oplus s_j \mid i, j \in [t-1], i \neq j\} \cup \{0^k\}$ but can take any other value. Since a is initially chosen uniformly from $\{0,1\}^k - \{0^k\}$, $a(x)$ is uniformly distributed over the set of other values, i.e.,

$$\{0,1\}^k - \{0^k\} - \{s_i \oplus s_j \mid i, j \in [t-1], i \neq j\}, \tag{97}$$

which has at least $2^k - 1 - \binom{t-1}{2}$ elements. Therefore, by the union bound

$$\Pr[a(x) \notin \{s_1 \oplus s_t, \ldots, s_{t-1} \oplus s_t\} \mid x(s_1) = v_1, \ldots, x(s_{k-1}) = v_{k-1}] \geq 1 - \frac{t-1}{2^k - 1 - \binom{t-1}{2}}. \tag{98}$$

Write x is t-good if the responses to the first t queries T makes on x are distinct. Then, since the above analysis holds for all distinct $v_1, \ldots, v_{t-1}$, we have

$$\Pr[x \text{ is } k\text{-good} \mid x \text{ is } (k-1)\text{-good}] \geq 1 - \frac{t-1}{2^k - 1 - \binom{t-1}{2}}, \tag{99}$$

using the fact that $\Pr[A \mid \dot\cup_i B_i] \geq \min_i \Pr[A \mid B_i]$.

Therefore, since the last inequality holds for all $t \in \{1, \ldots, d\}$,

$$\begin{aligned}
\Pr[x \text{ is } d\text{-good}] \geq & \prod_{t=1}^{d} \left(1 - \frac{t-1}{2^k - 1 - \binom{t-1}{2}}\right) \\
\geq & 1 - \sum_{t=1}^{d} \frac{t-1}{2^k - 1 - \binom{t-1}{2}} && \forall a, b \in [0,1], (1-a)(1-b) \geq 1 - a - b
\end{aligned}$$

Assume wlog that d is such that $1 + \binom{d-1}{2} \leq 2^k/2$ (else we're done) so

$$\Pr[x \text{ is } d\text{-good}] \geq 1 - \frac{2}{2^k}\frac{1}{2}d(d-1) \geq 1 - \frac{d^2}{2^k}. \tag{100}$$

Conditioned on the event that x is d-good, the sequence of d responses to the d queries T makes is a uniformly random sequence of d distinct elements. Comment: this is intuitive but can also verify this by computing a product of conditional probabilities.

Therefore, letting $P := \{x \in D_1 \mid x \text{ is } d\text{-good}\}$, we can apply lemma 6 to find that

$$\Pr[T(x) = \text{Simon}_n(x) \mid x \leftarrow \mu] \leq \frac{1}{2} + \frac{1}{2}\frac{d^2}{2^k}. \tag{101}$$

Therefore, we must have $d \geq \sqrt{2^k/3} = \Omega(\sqrt{n})$, as required. □

Remark 11. The D_0 of Simon_n is the same as the D_0 of Collision_n (when n is a power of 2). On the other hand, the D_1 of Simon_n is a subset of D_1 of Collision_n. Therefore, any randomized decision tree that computes Collision_n (with bounded-error 1/3) can also be used to compute Simon_n (with bounded-error 1/3). Therefore $R(\text{Collision}_n) \geq R(\text{Simon}_n)$. Therefore $O(\sqrt{n}) \geq R(\text{Collision}_n) \geq R(\text{Simon}_n) \geq \Omega(\sqrt{n})$, where the first inequality is from a few lectures ago and the last inequality is what we just proved. So $R(\text{Simon}), R(\text{Collision}) = O(\sqrt{n})$.