CS 40: Computational Complexity

Sair Shaikh

October 29, 2025

Collaboration Notice: Talked to Henry Scheible '26 to discuss some details.

Problem 1. Suppose that $x \in \{0,1\}^n$ is an unknown n-bit string. A helper reveals to us the bits $\langle x, r^{(i)} \rangle$ (for $1 \leq i \leq n$) where the strings $r^{(1)}, \ldots, r^{(n)} \in_R \{0,1\}^n$ are chosen uniformly at random and independently. Describe a deterministic algorithm that successfully reconstructs x from this information, with probability at least $1/4$.

Solution. The algorithm is as follows: Return $\sum_{i=1}^{n} \langle x, r^{(i)} \rangle r^{(i)}$. We need to show that this returns x with probability at least $\frac{1}{4}$. We break this down into two claims:

1. If $r^{(i)}$ are all linearly independent, then the algorithm returns correctly with certainity.
2. The probability that $r^{(i)}$ are all linearly independent is at least $\frac{1}{4}$.

To show claim (1), notice that if $r^{(i)}$ are linearly independent, they make up a basis as there are $n = \dim_{\mathbb{F}_2}(\{0,1\}^n)$ of them. Then, x can be written as:

$$x = \sum_i \lambda_i r^{(i)}$$

where $\lambda_i = \langle x, r^{(i)} \rangle$ (linear algebra fact). Thus,

$$x = \sum_{i=1}^{n} \langle x, r^{(i)} \rangle r^{(i)}$$

Next, we show claim 2. Let $\overline{E}$ be the event that $r^{(1)}, \cdots, r^{(n-1)}$ are linearly independent. Then, we can write:

$$\mathbf{Pr}[\text{all } n \text{ indep}] = \mathbf{Pr}[r^{(n)} \text{ lin. indep. of first n-1} | \overline{E}] \, \mathbf{Pr}[\overline{E}]$$

Next, we analyze $\overline{E}$ by analyzing E, the event that the first $n-1$ are linearly dependent. Then, there exists $2 \leq i \leq n-1$ such that $r^{(1)}, \cdots, r^{(i-1)}$ are linearly independent but $r^{(i)}$ is

linearly dependent on these. Note that $i = 1$ is not valid as a single vector is always linearly independent. Then, let E_i be the events such that $r^{(i)}$ is dependent on the first $i-1$, given the first $i-1$ are linearly independent. Then, using the union bound:

$$\mathbf{Pr}[E] = \mathbf{Pr}\left[\bigcup_{i=2}^{n} E_i\right] \leq \sum_{i=2}^{n} \mathbf{Pr}[E_i]$$

Thus, we consider $\mathbf{Pr}[E_i]$. The first $i-1$ linearly independent vectors have a span of dimension $i-1$. Thus, the cardinality of their span is 2^{i-1}. Thus, the probability that the next uniform independent random vector, $r^{(i)} \in \text{span}\{r^{(i)}, \cdots, r^{(i-1)}\}$ is $\frac{2^{i-1}}{2^n}$. Thus, $\mathbf{Pr}[E_i] = \frac{2^{i-1}}{2^n}$.

Then, we calculate:

$$\begin{aligned}\mathbf{Pr}[E] &\leq \sum_{i=2}^{n} \mathbf{Pr}[E_i] \\ &= \frac{1}{2^{n-1}} + \frac{1}{2^{n-2}} + \cdots + \frac{1}{4} < \frac{1}{2}\end{aligned}$$

where the inequality follows from the geometric series sum $\frac{1}{2}\sum_i^{\infty} 2^{-i} = \frac{1}{2}$, as a partial sum is smaller than the sum of the series as all terms are positive. Thus, overall, we have:

$$\mathbf{Pr}[\overline{E}] \geq \frac{1}{2}$$

Next, note that by our previous argument for E_i, $\mathbf{Pr}[r^{(n)} \text{ lin. indep. of first n-1} | \overline{E}] = \frac{2^{n-1}}{2^n} = \frac{1}{2}$. Combining these, we have:

$$\mathbf{Pr}[\text{all } n \text{ indep}] = \mathbf{Pr}[r^{(n)} \text{ lin. indep. of first n-1} | \overline{E}]\, \mathbf{Pr}[\overline{E}] \geq \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

That proves claim (2). Together with claim (1) this completes the proof.

Problem 2. A function $f : \{0,1\}^* \to \{0,1\}^*$ is said to be *weakly one-way* if there exists a constant $d > 0$ such that for every randomized polynomial-time algorithm A that attempts to invert f, for large enough n,

$$\mathbf{Pr}_{x \in_R \{0,1\}^n}[A(f(x), 1^n) \notin f^{-1}(f(x))] > n^{-d}.$$

Suppose we construct a new function family F by setting

$$F(x_1, \ldots, x_m) = (f(x_1), \ldots, f(x_m)),$$

where each $x_i \in \{0,1\}^n$ and $m = 2n^{1+d}$. Prove that F is a (strongly) one-way function, i.e., what we've defined in class as a "one-way function."

[Omitted part of the question]

Fix an $i \in [m]$ and, for this i, consider a single iteration of the inner loop, when INVERT is called with input $f(x)$. Intuitively, this iteration has some non-negligible probability of successfully returning a value for an average x. Let's consider the good set G_i of all x's for which such success happens with probability at least $1/(2mn^c)$. Prove the following two claims.

- **Claim 1.** There exists $i \in [m]$ such that

  $$\mathbf{Pr}_{x \in_R \{0,1\}^n}[x \in G_i] \geq 1 - \frac{1}{2n^d}.$$

  You'll have to use the fact that $m = 2n^{1+d}$ somewhere.

- **Claim 2.** For the particular i guaranteed by the previous claim, the probability that INVERT successfully outputs a value in the ith iteration of the outer loop is at least $1 - n^{-d}$. This probability is taken over both $x \in \{0,1\}^n$ (where $y = f(x)$ is the input to INVERT) and the internal randomness of INVERT.

  Hint: Split the failure event into two parts: one where $x \in G_i$ and one where $x \notin G_i$.

Finally, using the two claims, obtain the contradiction that f isn't weakly one-way after all.

Solution. Claim 1. We will prove this by contrapositive. Assume that:

$$\forall i \in [m] : \mathbf{Pr}_{x \in_R \{0,1\}^n}[x \in G_i] < 1 - \frac{1}{2n^d}$$

We will show that this implies that B cannot successfully invert with a high enough probability.

Let A be the event that all (randomly chosen) $x_i \in G_i$ for all $i \in [m]$ and $\overline{A}$ be the event that $x_i \notin G_i$ for at least one i. Using law of total probability, we can write:

$$\mathbf{Pr}_{\forall i: x_i \in_R \{0,1\}^n}[B \text{ inverts}] = \mathbf{Pr}[B \text{ inverts}|A]\,\mathbf{Pr}[A] + \mathbf{Pr}[B \text{ inverts}|\overline{A}]\,\mathbf{Pr}[\overline{A}]$$

where B inverts implies inverting on $x_1, \cdots, x_m$.

First, we bound $\mathbf{Pr}[A]$, noting that the x_i are chosen independently, as follows:

$$\begin{aligned}
\mathbf{Pr}[A] &= \mathbf{Pr}\left[\bigwedge_{i=0}^{m} x_i \in G_i\right] \\
&= \prod_{i=1}^{m} \mathbf{Pr}[x_i \in G_i] \\
&< (1 - \frac{1}{2n^d})^m \\
&\leq e^{-\frac{2n^{1+d}}{2n^d}} = e^{-n}
\end{aligned}$$

where we used $(1 - \alpha) \leq e^{-\alpha}$ in the last step. Next, we will upper-bound $\mathbf{Pr}[B \text{ inverts}|\overline{A}]$. Define E_i to be the event that B sucessfully inverts given $x_i \notin G_i$. Since $\overline{A}$ is the event that at least one $x_i \notin G_i$, we can write:

$$\mathbf{Pr}[B \text{ inverts}|\overline{A}] \leq \mathbf{Pr}\left[\bigcup_{i=1}^{m} E_i\right]$$

By definition of G_i, we know that $\mathbf{Pr}[E_i] < \frac{1}{2mn^c}$ for each i. Thus, using the union-bound, we get:

$$\mathbf{Pr}[B \text{ inverts}|\overline{A}] \leq \sum_{i=1}^{m} \mathbf{Pr}[E_i] \leq m \cdot \frac{1}{2mn^c} = \frac{1}{2n^c}$$

Using these two bounds, we conclude,

$$\begin{aligned}
\mathbf{Pr}[B \text{ inverts}] &= \mathbf{Pr}[B \text{ inverts}|A]\,\mathbf{Pr}[A] + \mathbf{Pr}[B \text{ inverts}|\overline{A}]\,\mathbf{Pr}[\overline{A}] \\
&\leq \mathbf{Pr}[A] + \mathbf{Pr}[B \text{ inverts}|\overline{A}] \\
&< e^{-n} + \frac{1}{2n^c} \\
&= \frac{1}{n^c}
\end{aligned}$$

where the last inequality is because $e^{-n} < \frac{1}{2n^c}$ for sufficiently large n. Thus, by contrapositive, as B succeeds with probability $> n^{-c}$, we note that there is at least one such i such that $\mathbf{Pr}[x \in G_i] \geq 1 - \frac{1}{2n^d}$.

Solution. Claim 2. Use the i from the previous part to fix an iteration i of the outer loop. Let E be the event that INVERT does not invert $f(x)$ (in the ith outer loop). We want to show that:

$$\mathbf{Pr}[E] < \frac{1}{n^d}$$

where the probability is over the randomness of INVERT and x. Using the law of total probability, we write this:

$$\mathbf{Pr}[E] = \mathbf{Pr}[E|x \in G_i]\,\mathbf{Pr}[x \in G_i] + \mathbf{Pr}[E|x \notin G_i]\,\mathbf{Pr}[x \notin G_i]$$

From the guarantee provided by the last part, we note that:

$$\mathbf{Pr}[x \notin G_i] \leq \frac{1}{2n^d}$$

Next, we look to bound $\mathbf{Pr}[E|x \in G_i]$. Let E_j be the event that INVERT does not invert $f(x)$ on the jth iteration of the inner loop (still in the outer loop i). Then, from the definition of G_i, we note that:

$$\mathbf{Pr}[E_j|x \in G_i] < 1 - \frac{1}{2mn^c}$$

Moreover, we have:

$$\mathbf{Pr}[E|x \in G_i] = \mathbf{Pr}\left[\bigcap_{i=1}^{2mn^{1+c}} E_j|x \in G_i\right]$$

Since the inner loops all share the same ith entry, i.e. $f(x)$, they only depend on the randomness of INVERT. In particular, as each inner loop iteration uses independent randomness for each of the other entries, the events E_j are independent given any x. Thus, we can bound:

$$\begin{aligned}
\mathbf{Pr}[E|x \in G_i] &= \prod_{i=1}^{2mn^{1+c}} \mathbf{Pr}[E_j|x \in G_i] \\
&\leq \left(1 - \frac{1}{2mn^c}\right)^{2mn^{1+c}} \\
&\leq e^{-\frac{2mn^{1+c}}{2mn^c}} \\
&= e^{-n} < \frac{1}{n^d}
\end{aligned}$$

where the 2nd last inequality again uses $(1 - \alpha) \leq e^{-\alpha}$, and the last holds for large enough n. Thus, for large enough n, we have:

$$\mathbf{Pr}[E] < \frac{1}{n^d}$$

Thus, the probability that INVERT successfully inverts $f(x)$ in the ith iteration is at least $1 - n^{-d}$. Since INVERT checks before returning an inversion, it only fails if it fails to return

in any iteration of the outer loop. Since the probability INVERT successfully inverts in any outer loop iteration is at least as much as the probability it successfully inverts in the ith iteration, we conclude that INVERT is a randomized algorithm that can invert $f(x)$ with probability $\geq 1 - n^{-d}$ (where the probability is over x and the randomness of INVERT). Thus, f is not weakly one-way.