MTAT.07.003 Cryptology II

Computational Indistinguishability

Sven Laur
University of Tartu

Indistinguishability

A quick recap of hypothesis testing

$s \leftarrow \mathcal{S}_0$

$x \leftarrow f(s)$

Given x and
- description of $\mathcal{S}_0$
- description of $f(\cdot)$

Charlie has to accept or reject
- a hypothesis $\mathcal{H}$

There are several types of hypotheses:

▷ *simple hypotheses* $\mathcal{H} = [s \stackrel{?}{=} s_0]$

▷ *complex hypotheses* $\mathcal{H} = [s \stackrel{?}{=} s_0 \vee s \stackrel{?}{=} s_1 \vee \ldots \vee s \stackrel{?}{=} s_k]$

▷ *trivial hypotheses* that always hold or never hold.

Computational distance

To choose between hypotheses $\mathcal{H}_0 = [s \stackrel{?}{=} s_0]$ and $\mathcal{H}_1 = [s \stackrel{?}{=} s_1]$, we have to distinguish two output distributions $\mathcal{X}_0 = f(s_0)$ and $\mathcal{X}_1 = f(s_1)$.

These distributions are (t, ε)-*indistinguishable* if for all t-time algorithms $\mathcal{A}$:

$$\mathsf{Adv}^{\mathsf{ind}}_{\mathcal{X}_0,\mathcal{X}_1}(\mathcal{A}) = |\Pr[x \leftarrow \mathcal{X}_0 : \mathcal{A}(x) = 0] - \Pr[x \leftarrow \mathcal{X}_1 : \mathcal{A}(x) = 0]| \leq \varepsilon$$

In other terms, the distributions $\mathcal{X}_0$ and $\mathcal{X}_1$ are (t, ε)-indistinguishable if

$$\mathsf{cd}^t_x(\mathcal{H}_0, \mathcal{H}_1) \leq \varepsilon \enspace .$$

Basic properties of computational distance

- **Triangle inequality.** For all simple hypotheses $\mathcal{H}_0$, $\mathcal{H}_1$, $\mathcal{H}_2$:

$$\mathsf{cd}_x^t(\mathcal{H}_0, \mathcal{H}_2) \leq \mathsf{cd}_x^t(\mathcal{H}_0, \mathcal{H}_1) + \mathsf{cd}_x^t(\mathcal{H}_1, \mathcal{H}_2) \enspace .$$

- **Symmetry.** For any two simple hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$:

$$\mathsf{cd}_x^t(\mathcal{H}_0, \mathcal{H}_1) = \mathsf{cd}_x^t(\mathcal{H}_1, \mathcal{H}_0) \enspace .$$

- **Positively definiteness.** For any reasonably large time bound t:

$$\mathsf{cd}_x^t(\mathcal{H}_0, \mathcal{H}_1) = 0 \quad \Leftrightarrow \quad \mathsf{sd}_x(\mathcal{H}_0, \mathcal{H}_1) = 0 \quad \Leftrightarrow \quad \mathcal{H}_0 \equiv \mathcal{H}_1 \enspace .$$

Interactive hypothesis testing

$$\xrightarrow{x_1} \qquad \xleftarrow{y_1} \qquad \cdots \qquad \xrightarrow{x_n}$$

$s \leftarrow \mathcal{S}_0$

Charlie has to accept or reject
– a hypothesis $\mathcal{H}$
based on interaction pattern

$$\star = (x_1, y_1, \ldots, y_{n-1}, x_n).$$

We use analogous notation for computational and statistical distance:

$$\mathsf{cd}^t_\star(\mathcal{H}_0, \mathcal{H}_1) = \max_{\mathcal{A} \text{ is } t\text{-time}} \left|\Pr\left[\mathcal{A}(\star) = 0 | \mathcal{H}_0\right] - \Pr\left[\mathcal{A}(\star) = 0 | \mathcal{H}_1\right]\right| \enspace ,$$

$$\mathsf{sd}_\star(\mathcal{H}_0, \mathcal{H}_1) = \max_{\mathcal{A}} \left|\Pr\left[\mathcal{A}(\star) = 0 | \mathcal{H}_0\right] - \Pr\left[\mathcal{A}(\star) = 0 | \mathcal{H}_1\right]\right| \enspace .$$

These measures also satisfy triangle inequality and other distance axioms.

Examples

Pseudorandom functions

Let $\mathcal{F}_{\mathrm{all}}$ denote the set of all functions $f : \mathcal{M} \to \mathcal{C}$ and let $\mathcal{F} \subseteq \mathcal{F}_{\mathrm{all}}$ be a function family. Then we can consider the following interactive hypothesis testing scenario. A t-time adversary $\mathcal{A}$ that makes at most q calls to the oracle $\mathcal{O}(\cdot)$ in order to distinguish two worlds (hypotheses):

- ▷ $\mathcal{H}_0$: Oracle chooses $f \xleftarrow[u]{} \mathcal{F}_{\mathrm{all}}$ and for every query x_i replies $y_i \leftarrow f(x_i)$.
- ▷ $\mathcal{H}_1$: Oracle chooses $f \xleftarrow[u]{} \mathcal{F}$ and for every query x_i replies $y_i \leftarrow f(x_i)$.

We say that $\mathcal{F}$ is (t, q, ε)-*pseudorandom function family* if for any t-time adversary $\mathcal{A}$ that makes at most q queries the corresponding advantage

$$\mathsf{Adv}^{\mathsf{ind}}(\mathcal{A}) = \left|\Pr\left[f \xleftarrow[u]{} \mathcal{F}_{\mathrm{all}} : \mathcal{A}^{\mathcal{O}(\cdot)} = 0\right] - \Pr\left[f \xleftarrow[u]{} \mathcal{F} : \mathcal{A}^{\mathcal{O}(\cdot)} = 0\right]\right| \leq \varepsilon \enspace .$$

Pseudorandom permutations

Let $\mathcal{F}_{\mathrm{prm}}$ denote the set of all permutations $f : \mathcal{M} \rightarrow \mathcal{M}$ and let $\mathcal{F} \subseteq \mathcal{F}_{\mathrm{prm}}$ be a permutation family. Then we can consider the following interactive hypothesis testing scenario. A t-time adversary $\mathcal{A}$ that makes at most q calls to the oracle $\mathcal{O}(\cdot)$ in order to distinguish two worlds (hypotheses):

- ▷ $\mathcal{H}_0$: Oracle chooses $f \xleftarrow[u]{} \mathcal{F}_{\mathrm{prm}}$ and for every query x_i replies $y_i \leftarrow f(x_i)$.
- ▷ $\mathcal{H}_1$: Oracle chooses $f \xleftarrow[u]{} \mathcal{F}$ and for every query x_i replies $y_i \leftarrow f(x_i)$.

We say that $\mathcal{F}$ is (t, q, ε)-*pseudorandom permutation family* if for any t-time adversary $\mathcal{A}$ that makes at most q queries the corresponding advantage

$$\mathsf{Adv}^{\mathsf{ind}}(\mathcal{A}) = \left|\Pr\left[f \xleftarrow[u]{} \mathcal{F}_{\mathrm{prm}} : \mathcal{A}^{\mathcal{O}(\cdot)} = 0\right] - \Pr\left[f \xleftarrow[u]{} \mathcal{F} : \mathcal{A}^{\mathcal{O}(\cdot)} = 0\right]\right| \leq \varepsilon \ .$$

Pseudorandom generators

Let f be a function that stretches m-bit seed s to n-bit string. Then we can consider the following classical hypothesis testing scenario. A t-time adversary $\mathcal{A}$ gets x and must distinguish two worlds (hypotheses):

- ▷ $\mathcal{H}_0$: The string x is uniformly chosen over $\{0,1\}^n$.
- ▷ $\mathcal{H}_1$: The string $x \leftarrow f(s)$ for uniformly chosen $s \xleftarrow[u]{} \{0,1\}^m$.

We say that f is (t, ε)-*pseudorandom generator* if for any t-time adversary $\mathcal{A}$ the corresponding advantage is bounded

$$\mathsf{Adv}^{\mathsf{ind}}(\mathcal{A}) = \left|\Pr\left[x \xleftarrow[u]{} \{0,1\}^n : \mathcal{A}(x) = 0\right] - \Pr\left[s \xleftarrow[u]{} \{0,1\}^m : \mathcal{A}(f(s)) = 0\right]\right| \leq \varepsilon \enspace .$$

Practical implementations

- **Pseudorandom functions.** Constructing good pseudorandom functions has never been a an explicit design goal. Cryptographic hash functions $h : \mathcal{M} \times \mathcal{K} \to \mathcal{T}$ with implicit or explicit keys are often treated as pseudorandom functions. However, they are also known to contain much more weaknesses than good block ciphers.

- **Pseudorandom permutations**. Block ciphers are specifically designed to be pseudorandom permutations. This is the most thoroughly studied branch of practical primitive design and we have many good candidates.

- **Pseudorandom generators.** Stream ciphers are designed to be fast pseudorandom generators. However, we know much more about block ciphers than about stream ciphers. In fact, there is no widely adopted stream cipher standard. There are also more secure constructions based on number theoretical constructions but they are much slower.

Guessing Games

Simplest guessing game

Consider the simplest attack scenario:

1. $\mathcal{S}_0$ is a uniform distribution over two states s_0 and s_1.
2. $\mathcal{H}_0$ and $\mathcal{H}_1$ denote simple hypotheses $[s \stackrel{?}{=} s_0]$ and $[s \stackrel{?}{=} s_1]$.
3. Given $x \leftarrow f(s)$, Charlie must choose between hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$.

The probability of an incorrect guess

$$\begin{aligned}\Pr[\text{Failure}] &= \Pr[\mathcal{H}_0] \cdot \Pr[\mathcal{A}(x) = 1|\mathcal{H}_0] + \Pr[\mathcal{H}_1] \cdot \Pr[\mathcal{A}(x) = 0|\mathcal{H}_1] \\ &= \frac{1}{2} \cdot \Bigl(\underbrace{\Pr[\mathcal{A}(x) = 1|\mathcal{H}_0]}_{\text{False negatives}} + \underbrace{\Pr[\mathcal{A}(x) = 0|\mathcal{H}_1]}_{\text{False positives}}\Bigr) \\ &= \frac{1}{2} + \frac{1}{2} \cdot \underbrace{\Bigl(\Pr[\mathcal{A}(x) = 0|\mathcal{H}_1] - \Pr[\mathcal{A}(x) = 0|\mathcal{H}_0]\Bigr)}_{\pm \mathsf{cd}^t_x(\mathcal{H}_0,\mathcal{H}_1)} \quad .\end{aligned}$$

Guessing game with a biased coin

Let $\mathcal{D}$ be a distribution over $\{0,1\}$ such that $\Pr\left[i \leftarrow \mathcal{D} : i = 0\right] \leq \frac{1}{2}$ and consider a guessing game $\mathcal{G}$ between a challenger and an adversary $\mathcal{A}$:

$$\mathcal{G}^{\mathcal{A}}$$

$$\left[\begin{array}{l} i \leftarrow \mathcal{D} \\ b \leftarrow \mathcal{A}(f(s_i)) \\ \textbf{return } [b \stackrel{?}{=} i] \end{array}\right.$$

For this game, the adversary succeeds with probability

$$\begin{aligned} \Pr\left[\mathsf{Success}\right] &= \Pr\left[\mathcal{H}_0\right] \cdot \Pr\left[\mathcal{A} = 0 | \mathcal{H}_0\right] + \Pr\left[\mathcal{H}_1\right] \cdot \Pr\left[\mathcal{A} = 1 | \mathcal{H}_1\right] \\ &\leq \Pr\left[\mathcal{H}_1\right] \cdot \left(1 + \Pr\left[\mathcal{A} = 0 | \mathcal{H}_0\right] - \Pr\left[\mathcal{A} = 0 | \mathcal{H}_1\right]\right) \\ &\leq \Pr\left[\mathcal{H}_1\right] + \mathsf{cd}_x^t(\mathcal{H}_0, \mathcal{H}_1) \enspace . \end{aligned}$$

Choosing between many values

Now consider a game

$$\begin{array}{l} \mathcal{G}^{\mathcal{A}} \\ \left[\begin{array}{l} s \leftarrow \mathcal{S}_0 \\ s' \leftarrow \mathcal{A}(f(s)) \\ \textbf{return } [s \stackrel{?}{=} s'] \end{array} \right. \end{array}$$

If for all possible states $s_i, s_j \in \mathrm{supp}(\mathcal{S}_0)$ distributions $f(s_i)$ and $f(s_j)$ are (t, ε)-indistinguishable, then for all t-time algorithms

$$\min_s \Pr[s] - \varepsilon \leq \Pr[\text{Success}] \leq \max_s \Pr[s] + \varepsilon \enspace .$$

The corresponding proof

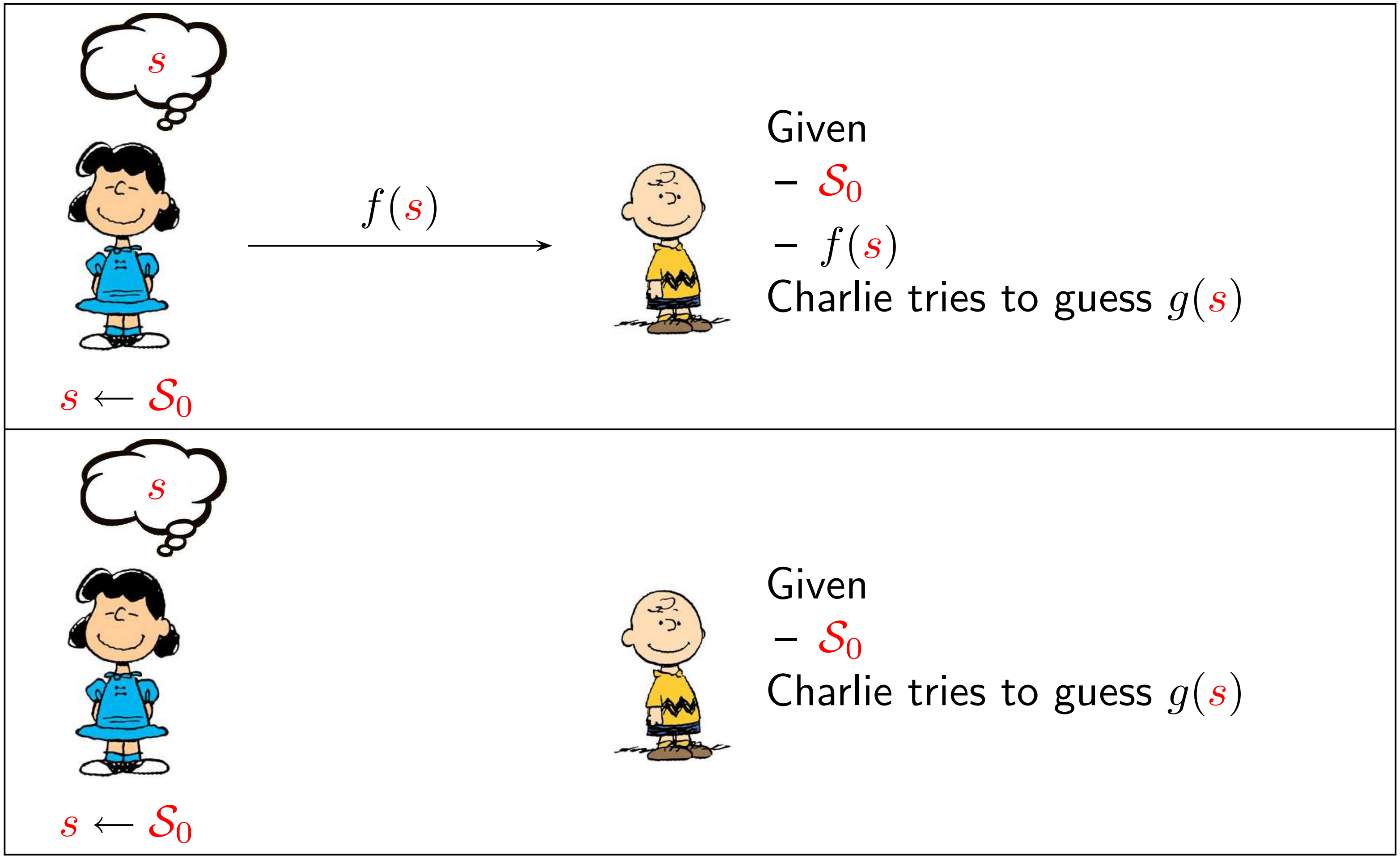
Let s_* the element with the maximal probability over $\mathcal{S}_0$. Then

$$\begin{aligned}
\Pr[\mathsf{Success}] &= \sum_{s \neq s_*} \Pr[s] \cdot \Pr[\mathcal{A}(f(s)) = s] \\
&\quad + \Pr[s_*] - \sum_{s \neq s_*} \Pr[s_*] \cdot \Pr[\mathcal{A}(f(s_*) = s)] \\
&\leq \Pr[s_*] + \sum_{s \neq s_*} \Pr[s] \cdot \underbrace{|\Pr[\mathcal{A}(f(s)) = s] - \Pr[\mathcal{A}(f(s_*)) = s]|}_{\leq \varepsilon} \\
&\leq \Pr[s_*] + \varepsilon \enspace .
\end{aligned}$$

The proof of the lower bound is analogous.

Semantic Security

Semantic security



Formal definition

Consider the following games:

$$\mathcal{G}_0^{\mathcal{A}} \quad \left[\begin{array}{l} s \leftarrow \mathcal{S}_0 \\ g' \leftarrow \mathcal{A}(f(s)) \\ \textbf{return } [g' \stackrel{?}{=} g(s)] \end{array}\right. \qquad \mathcal{G}_1^{\mathcal{A}} \quad \left[\begin{array}{l} s \leftarrow \mathcal{S}_0 \\ g' \leftarrow \mathrm{argmax}_{g'} \Pr\left[g(s) = g'\right] \\ \textbf{return } [g' \stackrel{?}{=} g(s)] \end{array}\right.$$

Then we can define a true guessing advantage

$$\begin{aligned} \mathsf{Adv}^{\mathsf{sem}}_{f,g}(\mathcal{A}) &= \Pr\left[\mathcal{G}_0^{\mathcal{A}} = 1\right] - \Pr\left[\mathcal{G}_1^{\mathcal{A}} = 1\right] \\ &= \Pr\left[s \leftarrow \mathcal{S}_0 : \mathcal{A}(f(s)) = g(s)\right] - \max_{g'} \Pr\left[g(s) = g'\right] \quad . \end{aligned}$$

IND $\Longrightarrow$ SEM

Theorem. If for all $s_i, s_j \in \mathrm{supp}(\mathcal{S}_0)$ distributions $f(s_i)$ and $f(s_j)$ are $(2t, \varepsilon)$-indistinguishable, then for all t-time adversaries $\mathcal{A}$:

$$\mathsf{Adv}^{\mathsf{sem}}_{f,g}(\mathcal{A}) \leq \varepsilon \enspace .$$

Note that

- ▷ function g might be randomised,
- ▷ function $g : \mathcal{S}_0 \rightarrow \{0,1\}^*$ may extremely difficult to compute,
- ▷ it might be even infeasible to get samples from the distribution $\mathcal{S}_0$.

Proof Sketch

Coin fixing

If $g : \mathcal{S}_0 \times \Omega \to \mathcal{Y}$ is a randomised function, then by definition

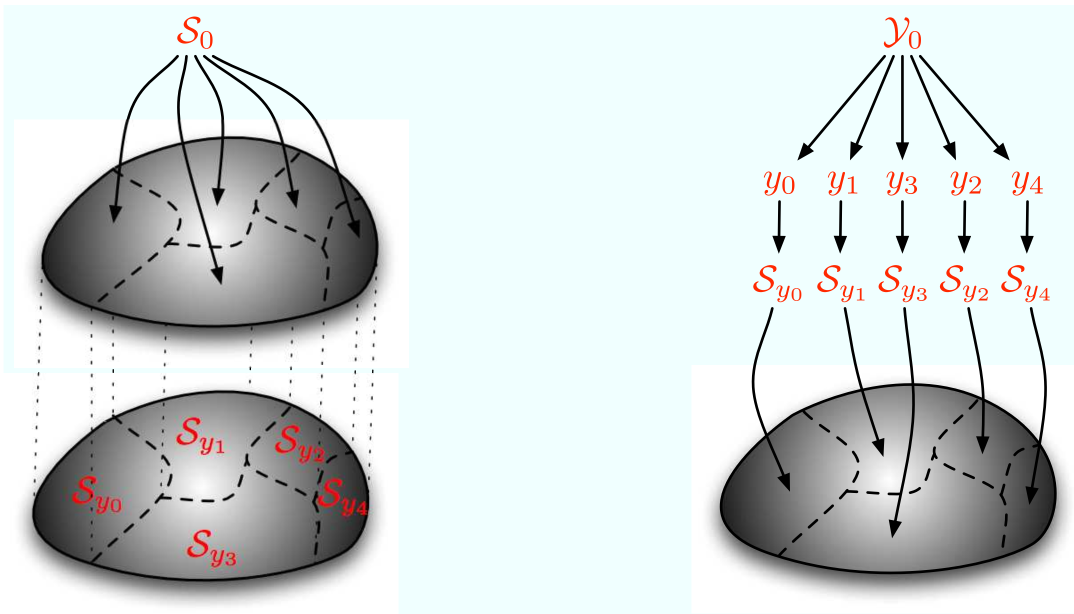
$$\mathsf{Adv}^{\mathsf{sem}}_{f,g}(\mathcal{A}) = \sum_{\omega \in \Omega} \Pr\left[\omega\right] \cdot \mathsf{Adv}^{\mathsf{sem}}_{f,g_\omega}(\mathcal{A})$$

where $g_\omega(s) \doteq g(s;\omega)$ is a deterministic function.

Hence, the advantage is maximised by a deterministic function, since

$$\sum_{\omega \in \Omega} \Pr\left[\omega\right] \cdot \mathsf{Adv}^{\mathsf{sem}}_{f,g_\omega}(\mathcal{A}) \leq \max_{\omega \in \Omega} \mathsf{Adv}^{\mathsf{sem}}_{f,g_\omega}(\mathcal{A}) \enspace .$$

Sampling idiom



Let $\mathcal{S}_{y_i}$ be the conditional distribution over the set $\{s \in \mathcal{S}_0 : g(s) = y_i\}$ and $\mathcal{Y}_0$ distribution of final outcomes $g(s)$. Then we get the distribution $\mathcal{S}_0$ if we first draw y from $\mathcal{Y}_0$ and then choose s according to $\mathcal{S}_y$.

Choosing between many values

As we can transform the security game into a new game

$$
\begin{aligned}
&\mathcal{G}_0^{\mathcal{A}}\\
&\left[\begin{array}{l}
y \leftarrow \mathcal{Y}_0\\
s \leftarrow \mathcal{S}_y\\
g' \leftarrow \mathcal{A}(f(s))\\
\textbf{return } [g' \stackrel{?}{=} g(s)]
\end{array}\right.
\end{aligned}
$$

where the adversary $\mathcal{A}$ must choose between hypotheses $\mathcal{H}_{y_0} = [y \stackrel{?}{=} y_0]$ for all possible outcomes $y_0 \in \mathcal{Y}_0$, we can establish

$$\Pr\left[\mathcal{G}_0^{\mathcal{A}} = 1\right] \leq \max_{y_0, y_1 \in \mathcal{Y}} \mathsf{cd}^{2t}_{f(s)}(\mathcal{H}_{y_0}, \mathcal{H}_{y_1}) + \max_{y_1} \Pr\left[y \leftarrow \mathcal{Y}_0 : y = y_1\right] \enspace .$$

Indistinguishability of conditional distributions

Fix $y_0, y_1 \in \mathcal{Y}$ and let $\mathcal{S}_{y_0}$ and $\mathcal{S}_{y_1}$ be the corresponding distributions. Then for any $2t$-time $\mathcal{B}$ the acceptance probabilities are

$$p_i = \sum_{s_0,s_1} \Pr\left[s \leftarrow \mathcal{S}_{y_0} : s = s_0\right] \Pr\left[s \leftarrow \mathcal{S}_{y_1} : s = s_1\right] \Pr\left[\mathcal{B}(f(s_i)) = 1\right] \quad .$$

Now the difference of acceptance probabilities can be bounded

$$\begin{aligned} |p_0 - p_1| &\leq \sum_{s_0,s_1} \Pr\left[s_0\right] \Pr\left[s_1\right] \left|\Pr\left[\mathcal{B}(f(s_0)) = 1\right] - \Pr\left[\mathcal{B}(f(s_1)) = 1\right]\right| \\ &\leq \max_{s_0,s_1} \left|\Pr\left[\mathcal{B}(f(s_0)) = 1\right] - \Pr\left[\mathcal{B}(f(s_1)) = 1\right]\right| \leq \varepsilon \end{aligned}$$

since all states in $\mathcal{S}_0$ are $(2t, \varepsilon)$-indistinguishable.

Postmortem

We have now formally shown that if $f : \mathcal{M} \times \mathcal{K} \to \mathcal{C}$ is a (t, ε)-pseudorandom function family then it is difficult to approximate a predicate $g(x)$ given only the value $f(x, k)$ and black-box access to the function $f_k(\cdot)$.

However, this general semantic security guarantee has also limitations:

▷ The proof is non-constructive.

▷ The theorem does not hold if $\mathcal{S}_0$ is specified by the adversary.
For example, if adversary can influence which messages are enciphered.

Switching lemma

Motivation

Block ciphers are designed to be pseudorandom permutations. However, it is much more easiser to work with pseudorandom functions. Therefore, all classical security proofs have the following structure:

1. Replace pseudorandom permutation family $\mathcal{F}$ with the family $\mathcal{F}_{\mathrm{prm}}$.
2. Use the PRP/PRF switching lemma to substitute $\mathcal{F}_{\mathrm{prm}}$ with $\mathcal{F}_{\mathrm{all}}$.
3. Solve the resulting combinatorial problem to bound the advantage:
   - ▷ All output values $f(x)$ have uniform distribution.
   - ▷ Each output $f(x)$ is independent of other outputs.

More formally, let $\mathcal{G}_0$ the original security game and $\mathcal{G}_1$ and $\mathcal{G}_2$ be the games obtained after replacement steps. Then

$$\mathsf{Adv}^{\mathsf{win}}_{\mathcal{G}_0}(\mathcal{A}) = \Pr\left[\mathcal{G}_0^{\mathcal{A}} = 1\right] \leq \mathsf{cd}^t_\star(\mathcal{G}_0, \mathcal{G}_1) + \mathsf{sd}_\star(\mathcal{G}_1, \mathcal{G}_2) + \Pr\left[\mathcal{G}_2^{\mathcal{A}} = 1\right] \quad .$$

PRP/PRF switching lemma

Theorem. Let $\mathcal{M}$ be the input and output domain for $\mathcal{F}_{\mathrm{all}}$. Then the permutation family $\mathcal{F}_{\mathrm{prm}}$ is (q, ε)-pseudorandom function family where

$$\varepsilon \leq \frac{q(q-1)}{2\,|\mathcal{M}|} \quad .$$

Theorem. Let $\mathcal{M}$ be the input and output domain for $\mathcal{F}_{\mathrm{all}}$. Then for any $q \leq \sqrt{|\mathcal{M}|}$ there exists a $\mathrm{O}(q \log q)$ distinguisher $\mathcal{A}$ that achieves

$$\mathsf{Adv}^{\mathsf{ind}}_{\mathcal{F}_{\mathrm{all}},\mathcal{F}_{\mathrm{prm}}}(\mathcal{A}) \geq 0.316 \cdot \frac{q(q-1)}{|\mathcal{M}|} \quad .$$

Birthday paradox

Obviously $f \notin \mathcal{F}_{\mathrm{prm}}$ if we find a collision $f(x_i) = f(x_j)$ for $i \neq j$.

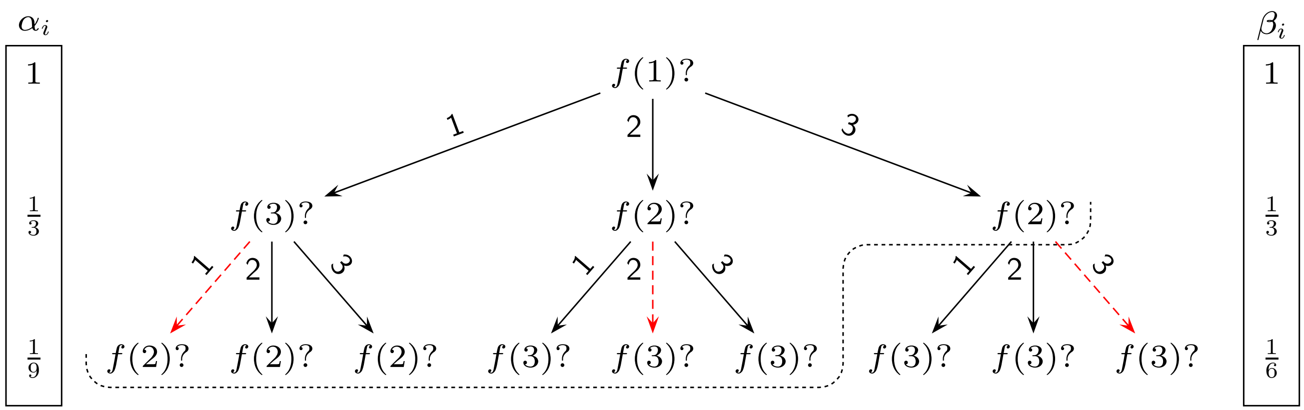
For the proof note that:

- ▷ If $x_1, \ldots, x_q$ are different then the outputs $f(x_1), \ldots, f(x_q)$ have uniform distribution over $\mathcal{M} \times \ldots \times \mathcal{M}$ when $f \xleftarrow[u]{} \mathcal{F}_{\mathrm{all}}$.
- ▷ Hence, the corresponding adversary $\mathcal{A}$ that outputs 0 only in case of collision obtains

$$\begin{aligned}\mathsf{Adv}^{\mathsf{ind}}_{\mathcal{F}_{\mathrm{all}},\mathcal{F}_{\mathrm{prm}}}(\mathcal{A}) &= \Pr\left[\mathsf{Collision}|\mathcal{F}_{\mathrm{all}}\right] - \Pr\left[\mathsf{Collision}|\mathcal{F}_{\mathrm{prm}}\right]\\ &= \Pr\left[\mathsf{Collision}|\mathcal{F}_{\mathrm{all}}\right] \geq 0.316 \cdot \frac{q(q-1)}{|\mathcal{M}|} \enspace .\end{aligned}$$

Distinguishing strategy as decision tree

Let $\mathcal{A}$ be a deterministic distinguisher that makes *up to* q oracle calls.



Then $\Pr\left[\text{Vertex } u|\mathcal{F}_{\mathrm{prm}}\right]$ and $\Pr\left[\text{Vertex } u|\mathcal{F}_{\mathrm{all}} \wedge \neg\text{Collision}\right]$ might differ. However, if $\mathcal{A}$ makes *exactly* q queries then all vertices on decision border are sampled with uniform probability and thus

$$\Pr\left[\mathcal{A}=0|\mathcal{F}_{\mathrm{prm}}\right] = \Pr\left[\mathcal{A}=0|\mathcal{F}_{\mathrm{all}} \wedge \neg\text{Collision}\right] \quad .$$

The corresponding proof

Obviously, the best distinguisher $\mathcal{A}$ is deterministic and makes exactly q oracle calls. Consequently,

$$\begin{aligned}\Pr\left[\mathcal{A}=0|\mathcal{F}_{\mathrm{all}}\right] &= \Pr\left[\mathsf{Collision}|\mathcal{F}_{\mathrm{all}}\right]\cdot\Pr\left[\mathcal{A}=0|\mathcal{F}_{\mathrm{all}}\wedge\mathsf{Collision}\right]\\ &\quad+\Pr\left[\neg\mathsf{Collision}|\mathcal{F}_{\mathrm{all}}\right]\cdot\Pr\left[\mathcal{A}=0|\mathcal{F}_{\mathrm{all}}\wedge\neg\mathsf{Collision}\right]\\ &\leq\Pr\left[\mathsf{Collision}|\mathcal{F}_{\mathrm{all}}\right]+\Pr\left[\mathcal{A}=0|\mathcal{F}_{\mathrm{prm}}\right]\end{aligned}$$

and thus also

$$\mathsf{Adv}^{\mathsf{ind}}_{\mathcal{F}_{\mathrm{all}},\mathcal{F}_{\mathrm{prm}}}(\mathcal{A})\leq\Pr\left[\mathsf{Collision}|\mathcal{F}_{\mathrm{all}}\right]\quad.$$

Now observe

$$\Pr\left[\bigvee_{i\neq j} f(x_i)=f(x_j)\right]\leq\sum_{i\neq j}\Pr\left[f(x_i)=f(x_j)\right]=\frac{q(q-1)}{2}\cdot\frac{1}{|\mathcal{M}|}\quad.$$