

# The Luby-Rackoff Theorem

Douglas Wikström, `dog@csc.kth.se`  
DD2448 Foundations of Cryptography

February 4, 2010

## Abstract

We give an asymptotic version of Luby-Rackoff, stripped of all concrete parameters. The idea is to illustrate the technique of hybrid arguments by applying it repeatedly in each step of the proof.

## The Luby-Rackoff Theorem

**Definition 1** (Negligible Function). A function  $\epsilon(n)$  is *negligible* if for every constant  $c > 0$ , there exists a constant  $n_0$ , such that  $\epsilon(n) < \frac{1}{n^c}$  for all  $n \geq n_0$ .

**Definition 2** (Pseudo-Random Function). A family of functions  $F : \{0, 1\}^k \times \{0, 1\}^n \rightarrow \{0, 1\}^n$  is *pseudo-random* if for all polynomial time oracle adversaries  $A$

$$\left| \Pr_K \left[ A^{F_K(\cdot)} = 1 \right] - \Pr_{R: \{0,1\}^n \rightarrow \{0,1\}^n} \left[ A^{R(\cdot)} = 1 \right] \right|$$

is negligible.

**Definition 3** (Pseudo-Random Permutation). A family of permutations  $P : \{0, 1\}^k \times \{0, 1\}^n \rightarrow \{0, 1\}^n$  is *pseudo-random* if for all polynomial time oracle adversaries  $A$

$$\left| \Pr_K \left[ A^{P_K(\cdot), P_K^{-1}(\cdot)} = 1 \right] - \Pr_{\Pi \in \mathcal{S}_{2^n}} \left[ A^{\Pi(\cdot), \Pi^{-1}(\cdot)} = 1 \right] \right|$$

is negligible, where  $\mathcal{S}_{2^n}$  is the set of permutations of  $\{0, 1\}^n$ .

**Definition 4** (Feistel Round). A Feistel-round using a function  $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$  is given by  $H_f : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}^n \times \{0, 1\}^n$  where

$$H_f(l, r) = (r, l \oplus f(r)) .$$

**Theorem 5** (Luby and Rackoff). *If  $F : \{0, 1\}^k \times \{0, 1\}^n \rightarrow \{0, 1\}^n$  is a pseudo-random family of functions, then*

$$\begin{aligned} H &: \{0, 1\}^{4k} \times \{0, 1\}^{2n} \rightarrow \{0, 1\}^{2n} \\ H &: ((k_1, k_2, k_3, k_4), x) \mapsto H_{F_{k_4}}(H_{F_{k_3}}(H_{F_{k_2}}(H_{F_{k_1}}(x)))) \end{aligned}$$

*is a pseudo-random family of permutations.*

## Proof of Theorem 5

### Notation

- $F_i$  denotes  $F_{K_i}$  for a random choice of  $K_i$ .
- $R_i$  denotes a randomly chosen function  $\{0, 1\}^n \rightarrow \{0, 1\}^n$ .
- $R$  and  $R'$  denote randomly chosen functions  $\{0, 1\}^{2n} \rightarrow \{0, 1\}^{2n}$ .
- $\Pi$  denotes a randomly chosen permutation of  $\{0, 1\}^{2n}$ .

Given an adversary  $A$ , we also define:

$$\begin{aligned} p_F^A &= \Pr_F \left[ A^{F(\cdot)} = 1 \right] \\ p_{F_1 F_2 F_3 F_4}^A &= \Pr_{F_1, F_2, F_3, F_4} \left[ A^{H_{F_1, F_2, F_3, F_4}(\cdot), H_{F_1, F_2, F_3, F_4}^{-1}(\cdot)} = 1 \right] \\ p_{H, H'}^A &= \Pr_{H, H'} \left[ A^{H(\cdot), H'(\cdot)} = 1 \right] \end{aligned}$$

### Overview of Proof

Our goal is to prove that

$$\left| p_{F_1 F_2 F_3 F_4}^A - p_{\Pi, \Pi^{-1}}^A \right|$$

is negligible. We proceed by a *hybrid argument*. We expand the absolute value and apply the triangle inequality and the lemmas below

$$\begin{aligned} & \left| p_{F_1 F_2 F_3 F_4}^A - p_{R_1 R_2 R_3 R_4}^A + p_{R_1 R_2 R_3 R_4}^A - p_{R, R'}^A + p_{R, R'}^A - p_{\Pi, \Pi^{-1}}^A \right| \\ & \leq \underbrace{\left| p_{F_1 F_2 F_3 F_4}^A - p_{R_1 R_2 R_3 R_4}^A \right|}_{\leq \epsilon_1 \text{ by Lemma 6}} + \underbrace{\left| p_{R_1 R_2 R_3 R_4}^A - p_{R, R'}^A \right|}_{\leq \epsilon_2 \text{ by Lemma 8}} + \underbrace{\left| p_{R, R'}^A - p_{\Pi, \Pi^{-1}}^A \right|}_{\leq \epsilon_3 \text{ by Lemma 7}} \\ & \leq \epsilon_1 + \epsilon_2 + \epsilon_3 \end{aligned}$$

which is then negligible. Throughout we assume without loss that:

1. **Remember Previous Queries.** The adversary  $A$  never asks the same query twice, since it can store all previous queries in table at a polynomial cost.
2. **No Trick-Queries.** The adversary  $A$  with oracles  $O_1$  and  $O_2$  never asks for both  $y = O_1(x)$  and  $O_2(y)$  or  $x = O_2(y)$  and  $O_1(x)$ , since the relations  $O_2(O_1(x)) = x$  and  $O_1(O_2(y)) = y$  are guaranteed to hold for both  $(O_1, O_2) = (H_{F_1, F_2, F_3, F_4}, H_{F_1, F_2, F_3, F_4}^{-1})$  and  $(O_1, O_2) = (\Pi, \Pi^{-1})$ .

### Technical Lemmas

#### Pseudo-Random Functions Look Like Random Functions, Even Within the Feistel Network

**Lemma 6.** For each polynomial time oracle adversary  $A$

$$\left| p_{F_1 F_2 F_3 F_4}^A - p_{R_1 R_2 R_3 R_4}^A \right| < \epsilon_1$$

for some negligible function  $\epsilon_1$ .

*Proof.* This follows from the triangle inequality and the pseudo-randomness of  $F$ . To see this, first note that

$$\begin{aligned} & \left| p_{F_1 F_2 F_3 F_4}^A - p_{R_1 R_2 R_3 R_4}^A \right| \\ &= \left| p_{F_1 F_2 F_3 F_4}^A - p_{F_1 F_2 F_3 R_4}^A + p_{F_1 F_2 F_3 R_4}^A - \dots - p_{F_1 R_2 R_3 R_4}^A + p_{F_1 R_2 R_3 R_4}^A - p_{R_1 R_2 R_3 R_4}^A \right| \\ &\leq \left| p_{F_1 F_2 F_3 F_4}^A - p_{F_1 F_2 F_3 R_4}^A \right| + \dots + \left| p_{F_1 R_2 R_3 R_4}^A - p_{R_1 R_2 R_3 R_4}^A \right| . \end{aligned}$$

If the lemma is false, then at least one of the absolute values in the sum is not negligible, say the first (the other cases follows similarly). This means that there exists a constant  $c > 0$  and an infinite set  $\mathcal{N}$  such that

$$\left| p_{F_1 F_2 F_3 F_4}^A - p_{F_1 F_2 F_3 R_4}^A \right| \geq \frac{1}{n^c}$$

for  $n \in \mathcal{N}$ . We could then construct an adversary  $A'$  expecting a single oracle  $T(\cdot)$  that executes  $A$  and simulates  $F_1, F_2,$  and  $F_3$  to  $A$ . Any query to  $F_4$  (or  $R_4$ ), is forwarded to  $T(\cdot)$  and the result handed by to  $A$ . It follows that

$$\left| p_{F_4}^{A'} - p_{R_4}^{A'} \right| = \left| p_{F_1 F_2 F_3 F_4}^A - p_{F_1 F_2 F_3 R_4}^A \right| \geq \frac{1}{n^c} ,$$

which contradicts the pseudo-randomness of  $F$ . Thus, the lemma must be true.  $\square$

### Without Trick Queries Random Permutations Look Like Random Functions

**Lemma 7.** *For each polynomial time oracle adversary  $A$*

$$\left| p_{R, R'}^A - p_{\Pi, \Pi^{-1}}^A \right| < \epsilon_3$$

for some negligible function  $\epsilon_3$ .

*Proof.* For our class of adversaries that remember previous queries and does not ask trick questions, the only way to distinguish the oracle pair  $(R, R')$  from  $(\Pi, \Pi^{-1})$  is if two queries to  $R$  or  $R'$  respectively, give the same reply. This happens with some negligible probability  $\epsilon_3$ , since  $A$  only asks a polynomial number of queries.  $\square$

### Without Trick Queries A Feistel Network Based On Random Functions Look Like Random Functions

**Lemma 8.** *For each polynomial time oracle adversary  $A$*

$$\left| p_{R_1 R_2 R_3 R_4}^A - p_{R, R'}^A \right| < \epsilon_2$$

for some negligible function  $\epsilon_2$ .

*Proof.* Let  $(O_1^i, O_2^i)$  be a hybrid oracle pair that equals  $(H_{R_1, R_2, R_3, R_4}, H_{R_1, R_2, R_3, R_4}^{-1})$  for the first  $i$  queries and then equals  $(R, R')$  for the remaining  $q - i$  queries, where  $q$  is the total number of queries. Note that  $q$  is polynomial, since  $A$  is polynomial. We clearly have  $p_{O_1^0, O_2^0}^A = p_{R_1 R_2 R_3 R_4}^A$  and  $p_{O_1^q, O_2^q}^A = p_{R, R'}^A$ . Using the triangle inequality we get

$$\left| p_{R_1 R_2 R_3 R_4}^A - p_{R, R'}^A \right| = \left| \sum_{i=1}^q \left( p_{O_1^i, O_2^i}^A - p_{O_1^{i-1}, O_2^{i-1}}^A \right) \right| \leq q \max_i \left\{ \left| p_{O_1^i, O_2^i}^A - p_{O_1^{i-1}, O_2^{i-1}}^A \right| \right\} .$$

Let  $(l_k^j, r_k^j)$  denote the  $j$ th intermediate left-right value pair of the  $k$ th execution of the Feistel network behind  $(O_1^i, O_2^i)$  for  $1 \leq k \leq i$  and  $j = 0, \dots, 4$ . Thus, the  $j$ th query-reply pair may be written  $((l_k^j, r_k^j), (l_k^j, r_k^j))$ .

Let  $E_{\text{good}}$  denote the event that  $r_k^1 \neq r_{k'}^1$  and  $l_k^3 \neq l_{k'}^3$  for every  $k \neq k'$ . It follows that  $\Pr[\overline{E_{\text{good}}}]$  is negligible, since (1) for every fixed  $(l_k^0, r_k^0) \neq (l_{k'}^0, r_{k'}^0)$

$$\Pr [r_k^1 = r_{k'}^1] = \Pr [R_1(r_k^0) \oplus l_k^0 = R_1(r_{k'}^0) \oplus l_{k'}^0]$$

is negligible and there are at most  $q^2$  pairs to consider, and (2) the argument is similar for the case of  $l_k^3 \neq l_{k'}^3$ . We conclude that

$$\left| p_{O_1^i, O_2^i}^A - p_{O_1^{i-1}, O_2^{i-1}}^A \right|$$

only changes by a negligible quantity if we condition on  $E_{\text{good}}$ .

We now consider the case where the  $i$ th query  $x = (l_i^0, r_i^0)$  is to the first oracle of  $A$ , and investigate the distribution of the reply  $y$  conditioned on the event  $E_{\text{good}}$  when the first oracle is  $O_1^{i-1}$  and  $O_1^i$  respectively. The case where the  $i$ th query is to the second oracle is similar.

If the first oracle of  $A$  is  $O_1^{i-1}$ , then clearly  $y$  is uniformly distributed in  $\{0, 1\}^{2n}$ . If on the other hand the first oracle of  $A$  is  $O_1^i$ , then we note that

$$l_i^3 = r_i^2 = l_i^1 \oplus R_2(r_i^1) \quad \text{and} \quad r_i^3 = l_i^2 \oplus R_3(r_i^2) = r_i^1 \oplus R_3(l_i^3) . \quad (1)$$

Thus, when  $r_i^1$  is distinct from all  $r_k^1$  and  $l_i^3$  is distinct from all  $l_k^3$  for each  $k < i$ , then  $R_2(r_i^1)$  and  $R_3(l_i^3)$  are uniformly and independently distributed in  $\{0, 1\}^n$ , which makes  $(l_i^3, r_i^3)$  uniformly and independently distributed in  $\{0, 1\}^{2n}$  by (1). We conclude that the reply  $y$  is also uniformly and independently distributed in  $\{0, 1\}^{2n}$ .  $\square$