# RSA-OAEP IND-CCA2 증명

김동현(wlswudpdlf31@kookmin.ac.kr)

March 18, 2025

## Contents

# 1 논문정보

- 제목: RSA-OAEP is Secure under the RSA Assumption
- 저자: Eiichiro Fujisaki1, Tatsuaki Okamoto, David Pointcheval, and Jacques Stern
- 년도: 2001년
- 초록: 최근 Victor Shoup은 적응적 선택 암호문 공격에 대한 OAEP의 보안성에 관한 널리 받아들여진 결과에 틈이 있음을 지적하였다. 더욱이, 그는 기본 트랩도어 치환의 단방향성만으로는 OAEP의 보안성을 증명할 수 없을 것으로 예상된다는 점을 보였다. 본 논문은 OAEP의 보안성에 대한 또 다른 결과를 제시한다. 즉, 본 논문에서는 무작위 오라클 모델에서, 기본 치환의 부분 영역 단방향성(partial-domain one-wayness) 하에서, OAEP가 적응적 선택 암호문 공격에 대해 의미론적 보안성을 제공함을 증명한다. 따라서, 이는 형식적으로 더 강한 가정을 사용한다. 그럼에도 불구하고, RSA 함수의 부분 영역 단방향성이 (전체 영역) 단방향성과 동치이므로, RSA-OAEP의 보안성은 단순한 RSA 가정만으로도 증명될 수 있음을 알 수 있다. 다만, 그 축소(reduction)는 타이트하지 않다.

# 2 보안 개념

## 2.1 OW trapdoor permutation

트랩도어 치환 체계(Trapdoor permutation scheme) $\Psi = (\mathcal{K}, \mathcal{F}, \mathcal{I})$를 다음과 같이 정의한다.

- $\mathcal{K}(1^\lambda)$: 확률론적 키 생성 알고리즘으로, $1^\lambda$를 입력 받아 $(\mathsf{pk}, \mathsf{sk})$를 생성한다.
- $\mathcal{F}_{\mathsf{pk}}(x)$: 결정론적 알고리즘으로, $\mathsf{pk}$와 $x \in \{0,1\}^\lambda$를 입력 받아 $y \in \{0,1\}^\lambda$를 출력한다.
- $\mathcal{I}_{\mathsf{sk}}(y)$: 결정론적 알고리즘으로, $\mathsf{sk}$와 $y \in \{0,1\}^\lambda$를 입력 받아 $x \in \{0,1\}^\lambda$를 출력한다. $\mathcal{K}(1^\lambda)$로 생성한 모든 $(\mathsf{pk}, \mathsf{sk})$와 모든 $x \in \{0,1\}^\lambda$에 대해, $\mathcal{I}_{\mathsf{sk}}(\mathcal{F}_{\mathsf{pk}}(x)) = x$를 만족한다.

동작시간(Running time) $\tau$를 가지는 공격자 $\mathcal{A}$와 트랩도어 치환 체계 $\Psi$에 대한 일방향성(One-wayness) 실험 $\mathsf{Exp}^{\mathsf{OW}}_{\Psi,\lambda}(\mathcal{A};\tau)$을 다음과 같이 정의한다.

| Challenger $\mathcal{C}$ | $\overset{\mathsf{Exp}^{\mathsf{OW}}_{\Psi,\lambda}}{\Longleftrightarrow}$ | Adversary $\mathcal{A}$ |
|---|---|---|
| $(\mathsf{pk}, \mathsf{sk}) \xleftarrow{\$} \mathcal{K}(1^\lambda)$ | | |
| $x^* \xleftarrow{\$} \{0,1\}^\lambda$<br>$y^* \leftarrow \mathcal{F}_{\mathsf{pk}}(x^*)$ | $\xrightarrow{1^\lambda, \mathsf{pk}, y^*}$ | |
| Return $[x' \overset{?}{=} x^*]$ | $\xleftarrow{x'}$ | $\mathcal{A}$ chooses $x' \in \mathcal{X}$ |

$\mathcal{A}$의 능력치 $\mathsf{Adv}^{\mathsf{OW}}_{\mathcal{A},\Psi}(\lambda,\tau)$를 다음과 같이 정의한다.

$$\mathsf{Adv}^{\mathsf{OW}}_{\mathcal{A},\Psi}(\lambda,\tau) := \Pr[\mathsf{Exp}^{\mathsf{OW}}_{\Psi,\lambda}(\mathcal{A};\tau) = 1].$$

공격자 $\mathcal{A}$의 능력치가 어떤 $\varepsilon$에 대해 다음을 만족할 때, $\Psi$가 $(\tau,\varepsilon)$-일방향성을 가진다고 한다.

$$\mathsf{Adv}^{\mathsf{OW}}_{\mathcal{A},\Psi}(\lambda,\tau) \leq \varepsilon.$$

## 2.2 Partial-domain OW trapdoor permutation

트랩도어 치환 $\Psi = (\mathcal{K}, \mathcal{F}, \mathcal{I})$에서, $\mathcal{F}_{\mathsf{pk}}(x) : \{0,1\}^\lambda \to \{0,1\}^\lambda$를 다음과 같이 표현한다.

$$\mathcal{F}_{\mathsf{pk}} : \{0,1\}^{n+\lambda_1} \times \{0,1\}^{\lambda_0} \to \{0,1\}^{n+\lambda_1} \times \{0,1\}^{\lambda_0}.$$

이때 $\lambda = n + \lambda_0 + \lambda_1$이다.

**메모 1.** 예를 들어, $x = s \parallel t$라고 할 때, $y \leftarrow \mathcal{F}_{\mathsf{pk}}(x)$ 대신 $y \leftarrow \mathcal{F}_{\mathsf{pk}}(s \parallel t)$로 표현할 수 있다.

동작시간 $\tau$를 가지는 공격자 $\mathcal{A}$와 트랩도어 함수 체계 $\Psi$에 대한 부분 일방향성(Partial-domain one-wayness) 실험 $\mathsf{Exp}^{\mathsf{OW\text{-}PD}}_{\Psi,\lambda}(\mathcal{A};\tau)$을 다음과 같이 정의한다.

| Challenger $\mathcal{C}$ | $\overset{\mathsf{Exp}^{\mathsf{OW\text{-}PD}}_{\Psi,\lambda}}{\Longleftrightarrow}$ | Adversary $\mathcal{A}$ |
|---|---|---|
| $(\mathsf{pk}, \mathsf{sk}) \xleftarrow{\$} \mathcal{K}(1^\lambda)$ | | |
| $(s^*, t^*) \xleftarrow{\$} \{0,1\}^{n+\lambda_1} \times \{0,1\}^{\lambda_0}$<br>$y^* \leftarrow \mathcal{F}_{\mathsf{pk}}(s^*, t^*)$ | $\xrightarrow{1^\lambda, \mathsf{pk}, y^*}$ | |
| Return $[s' \overset{?}{=} s^*]$ | $\xleftarrow{s'}$ | $\mathcal{A}$ chooses $s' \in \{0,1\}^{n+\lambda_1}$ |

공격자 $\mathcal{A}$의 능력치 $\mathsf{Adv}^{\mathsf{OW\text{-}PD}}_{\mathcal{A},\Psi}(\lambda,\tau)$를 다음과 같이 정의한다.

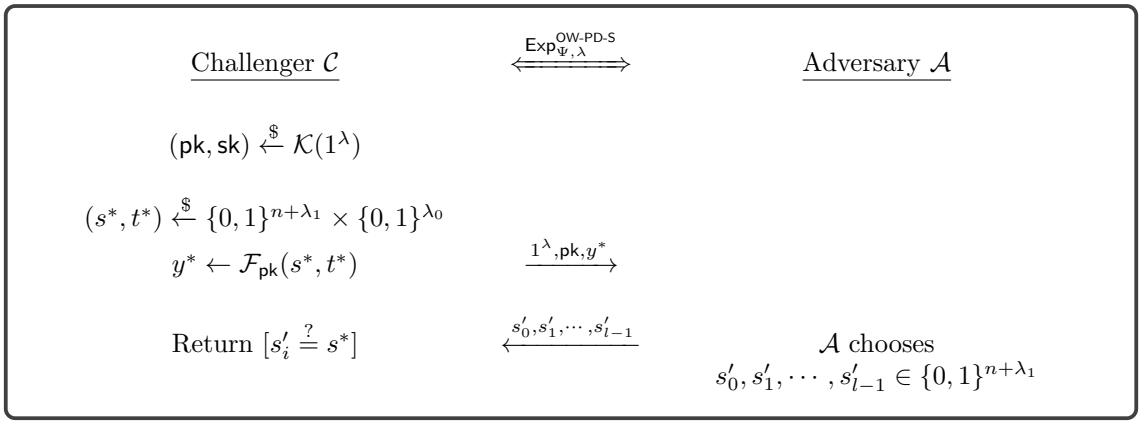
$$\mathsf{Adv}^{\mathsf{OW\text{-}PD}}_{\mathcal{A},\Psi}(\lambda,\tau) := \Pr[\mathsf{Exp}^{\mathsf{OW\text{-}PD}}_{\Psi,\lambda}(\mathcal{A};\rtimes)=1].$$

동작시간 $\tau$를 가지는 공격자 $\mathcal{A}$의 능력치가 어떤 $\varepsilon$에 대해 다음을 만족할 때, $\Psi$가 $(\tau,\varepsilon)$-부분 일방향성을 가진다고 한다.

$$\mathsf{Adv}^{\mathsf{OW\text{-}PD}}_{\mathcal{A},\Psi}(\lambda,\tau) \le \varepsilon.$$

## 2.3 Set partial-domain OW trapdoor permutation

동작시간 $\tau$를 가지고 $l$개의 원소를 출력하는 공격자 $\mathcal{A}$와 트랩도어 함수 체계 $\Psi$에 대한 집합 부분 일방향성(Set partial-domain one-wayness) 실험 $\mathsf{Exp}^{\mathsf{OW\text{-}PD\text{-}S}}_{\Psi,\lambda}(\mathcal{A};\tau,l)$을 다음과 같이 정의한다.




공격자 $\mathcal{A}$의 능력치 $\mathsf{Adv}^{\mathsf{OW\text{-}PD\text{-}S}}_{\mathcal{A},\Psi}(\lambda,\tau,l)$를 다음과 같이 정의한다.

$$\mathsf{Adv}^{\mathsf{OW\text{-}PD\text{-}S}}_{\mathcal{A},\Psi}(\lambda,\tau,l) := \Pr[\mathsf{Exp}^{\mathsf{OW\text{-}PD\text{-}S}}_{\Psi,\lambda}(\mathcal{A};\tau,l)=1].$$

동작시간 $\tau$를 가지는 공격자 $\mathcal{A}$의 능력치가 어떤 $\varepsilon$에 대해 다음을 만족할 때, $\Psi$가 $(l,\tau,\varepsilon)$-집합 부분 일방향성을 가진다고 한다.

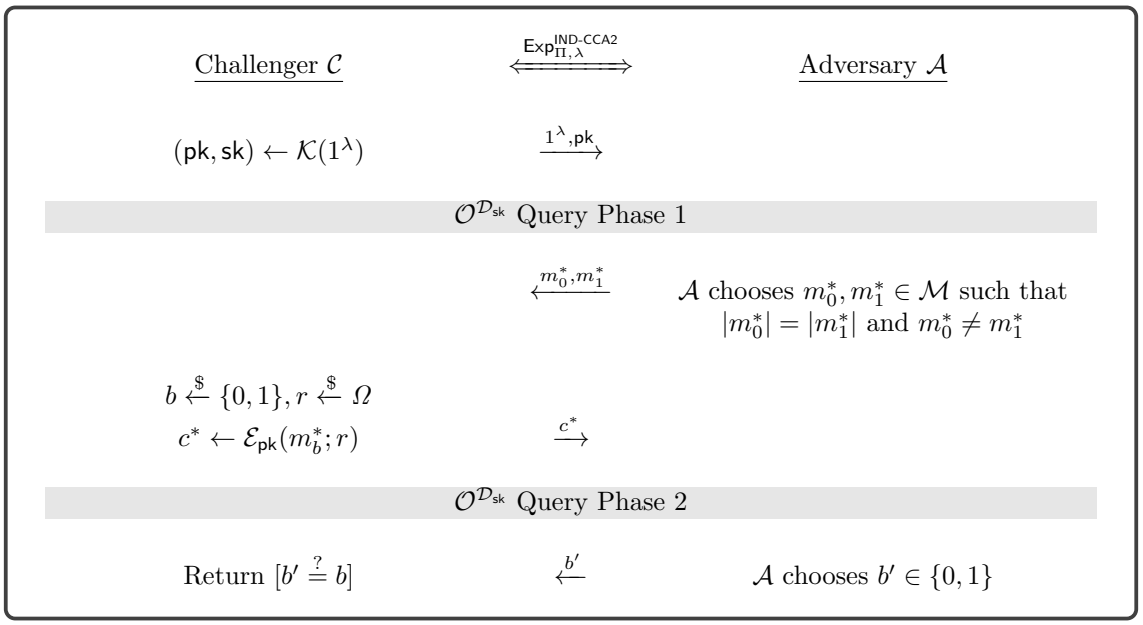
$$\mathsf{Adv}^{\mathsf{OW\text{-}PD\text{-}S}}_{\mathcal{A},\Psi}(\lambda,\tau,l) \le \varepsilon.$$

## 2.4 IND security against CCA2

공개키 암호 체계(Public-key encryption scheme) $\Pi=(\mathcal{K},\mathcal{E},\mathcal{D})$를 다음과 같이 정의한다.

- $\mathcal{K}(1^\lambda)$: 확률론적 키 생성 알고리즘으로, $1^\lambda$를 입력 받아 $(\mathsf{pk},\mathsf{sk})$를 생성한다.
- $\mathcal{E}_{\mathsf{pk}}(m)$: 암호화 알고리즘으로, $\mathsf{pk}$와 $m\in\mathcal{M}$를 입력 받아 $c\in\mathcal{C}$를 출력한다. 확률론적 알고리즘으로, $r\xleftarrow{\$}\Omega$를 추가로 입력 받아 $\mathcal{E}_{\mathsf{pk}}(m;r)$으로 표현할 수도 있다.
- $\mathcal{D}_{\mathsf{sk}}(c)$: 결정론적 복호화 알고리즘으로, $\mathsf{sk}$와 $c\in\mathcal{C}$를 입력 받아 $m\in\mathcal{M}$를 출력한다.

선택 암호문 공격(Adaptive chosen ciphertext attack, 이하 CCA2)을 수행하고 복호화 오라클에 $q$회 질의하는 공격자 $\mathcal{A}$와 공개키 암호 체계 $\Pi$에 대해, 의미론적(Semantic, 혹은 indistinguishability, 이하 IND) 안전성 실험 $\mathsf{Exp}^{\mathsf{IND\text{-}CCA2}}_{\Pi,\lambda}(\mathcal{A};q)$을 다음과 같이 정의한다.

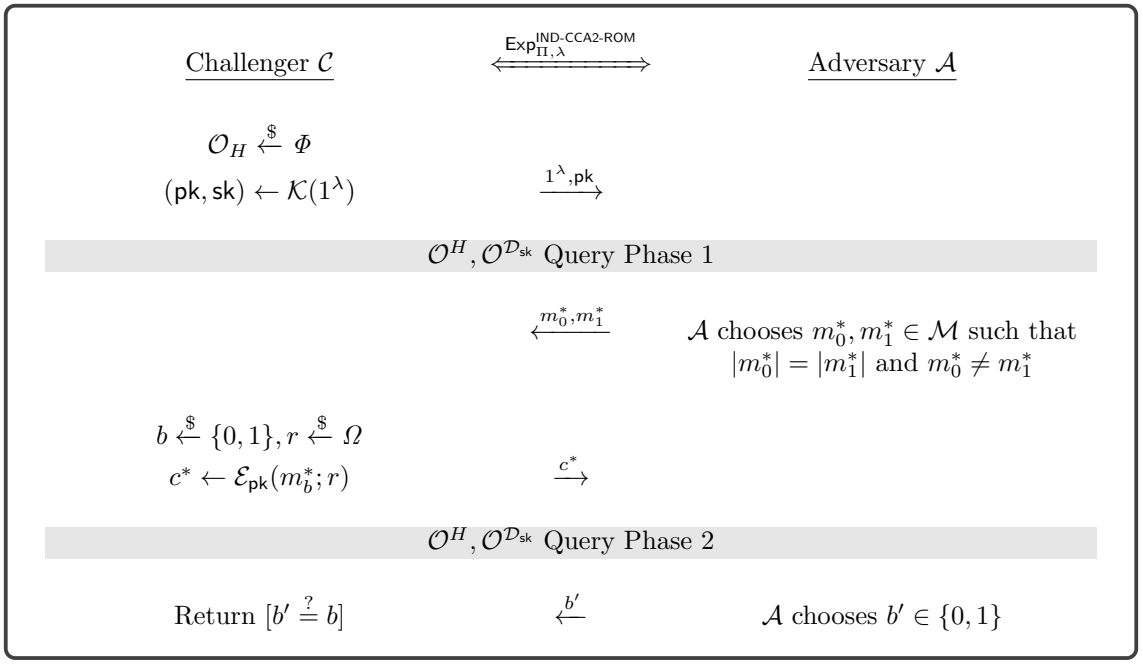



공격자 $\mathcal{A}$의 능력치 $\mathsf{Adv}^{\mathsf{IND\text{-}CCA2}}_{\mathcal{A},\Pi}(\lambda, q)$를 다음과 같이 정의한다.

$$\mathsf{Adv}^{\mathsf{IND\text{-}CCA2}}_{\mathcal{A},\Pi}(\lambda, q) = 2 \cdot \Pr[\mathsf{Exp}^{\mathsf{IND\text{-}CCA2}}_{\Pi,\lambda}(\mathcal{A}; q) = 1] - 1.$$

## 2.5 IND security against CCA2 in ROM

복호화 오라클에 $q_{\mathcal{D}}$회, 랜덤 오라클에 $q_H$회 질의하는 CCA2 공격자 $\mathcal{A}$와 공개키 암호 체계 $\Pi$에 대해, 랜덤 오라클 모델(Random oracle model)에서의 IND 실험 $\mathsf{Exp}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}_{\Pi,\lambda}(\mathcal{A}; q_{\mathcal{D}}, q_H)$을 다음과 같이 정의한다.




공격자 $\mathcal{A}$의 능력치 $\mathsf{Adv}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}_{\mathcal{A},\Pi}(\lambda, q_{\mathcal{D}}, q_H)$를 다음과 같이 정의한다.

$$\mathsf{Adv}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}_{\mathcal{A},\Pi}(\lambda, q_{\mathcal{D}}, q_H) = 2 \cdot \Pr[\mathsf{Exp}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}_{\Pi,\lambda}(\mathcal{A}; q_{\mathcal{D}}, q_H) = 1] - 1.$$

# 3 RSA-OAEP

다음과 같은 두 해시 함수 $H, G$를 준비한다.

$$H : \{0,1\}^{\lambda_0} \to \{0,1\}^{\lambda-\lambda_0} \quad G : \{0,1\}^{\lambda-\lambda_0} \to \{0,1\}^{\lambda_0}.$$

트랩토어 치환 체계 $\Psi = (\mathcal{K}, \mathcal{F}, \mathcal{I})$를 포함하는 OAEP 변환 $(\mathcal{K}, \mathcal{E}, \mathcal{D})$는 다음과 같이 동작한다.

- $\mathcal{K}(1^\lambda)$: $(\mathsf{pk}, \mathsf{sk})$를 생성한다. $\mathsf{pk}$는 이후 트랩도어 치환 $\mathcal{F}$에서 사용하며, $\mathsf{sk}$는 $\mathcal{I}$에서 사용한다.
- $\mathcal{E}_{\mathsf{pk}}(m; r)$: $m \in \{0,1\}^n$과 $r \xleftarrow{\$} \{0,1\}^{\lambda_0}$가 주어졌을 때, $s, t$를 다음과 같이 계산한다.

  $$s = (m \parallel 0^{\lambda_1}) \oplus G(r), \quad t = r \oplus H(s).$$

  $s, t$를 계산하는 과정을 도식화하면 그림 1와 같다. 이후 암호문 $c = \mathcal{F}_{\mathsf{pk}}(s, t)$를 출력한다.
- $\mathcal{D}_{\mathsf{sk}}(c)$: $(s, t) = \mathcal{I}_{\mathsf{sk}}(c)$을 계산한 후, $r, M$을 다음과 같이 계산한다.

  $$r = t \oplus H(s) \quad M = s \oplus G(r).$$

  만약 $[M]_{\lambda_1} = 0^{\lambda_1}$이면 $[M]^n$을 출력하고, 아니라면 "Reject"를 출력한다.

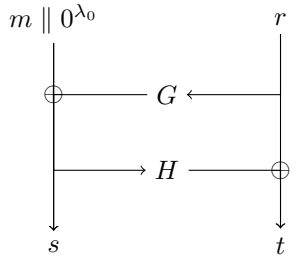


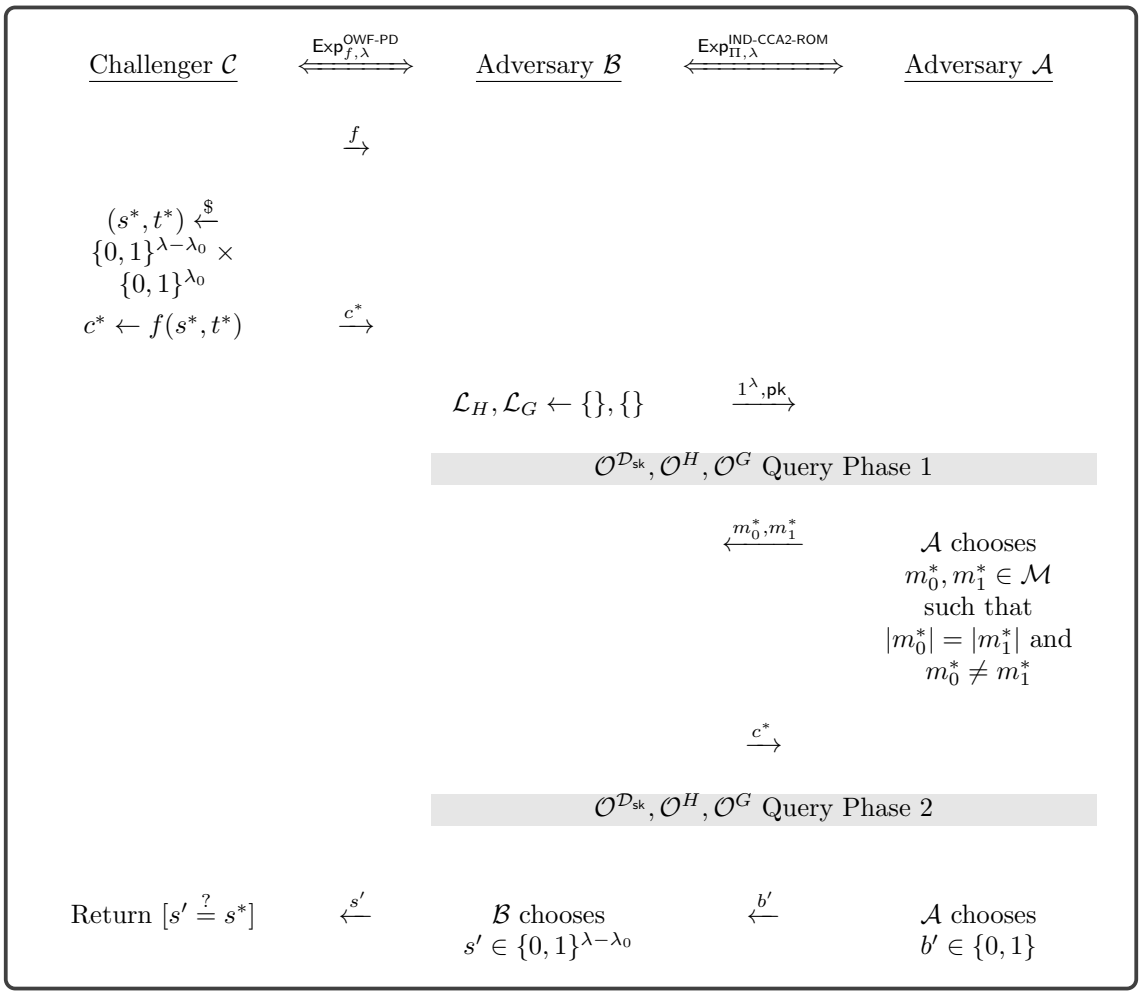

Figure 1: $\mathcal{E}_{\mathsf{pk}}(m; r)$에서 $s, t$를 계산하는 과정

# 4 증명

**보조정리 1.** *공격자 $\mathcal{A}$를 OAEP 변환 $(\mathcal{K}, \mathcal{E}, \mathcal{D})$에 대해 $\tau$를 가지고, 복호화 오라클 $\mathcal{O}^{\mathcal{D}}$와 랜덤 오라클 $\mathcal{O}^H, \mathcal{O}^G$에 각각 $q_{\mathcal{D}}, q_H, q_G$회 질의하는 IND-CCA2 공격자라 하자. 어떤 OW-PD-S 공격자 $\mathcal{B}$에 대해, 다음을 만족한다.*

$$\mathsf{Adv}^{\mathsf{OW\text{-}PD\text{-}S}}_{\mathcal{B},\Psi}(\lambda, \tau', q_H) \geq \frac{\mathsf{Adv}^{\mathsf{IND\text{-}CCA2}}_{\mathcal{A},\Pi}(\lambda, \tau, q_{\mathcal{D}}, q_H, q_G)}{2} - \frac{2q_{\mathcal{D}}q_G + q_{\mathcal{D}} + q_G}{2^{\lambda_0}} - \frac{2q_{\mathcal{D}}}{2^{\lambda_1}}.$$

*이 때, $\tau' \leq \tau \cdot q_H \cdot q_G \cdot (T_{\mathcal{F}} + O(1))$이고, $T_{\mathcal{F}}$는 트랩도어 치환 $\mathcal{F}$의 시간 복잡도를 의미한다.*




**How $\mathcal{B}$ simulate $\mathcal{O}^G$?**

- If $\gamma \in \mathcal{L}_G$, then response $G_\gamma$ and $\mathcal{L}_G \leftarrow \mathcal{L}_G \cap (\gamma, G_\gamma)$.
- Otherwise, do following:
  - For some $\delta \in \mathcal{L}_H$, if $c^* = f(\delta, \gamma \oplus H_\delta)$, then $G_\gamma \leftarrow \delta \oplus (m_b \parallel 0^{\lambda_1})$.
  - For all $\delta \in \mathcal{L}_H$, if $c^* \neq f(\delta, \gamma \oplus H_\delta)$, then $G_\gamma \xleftarrow{\$} \{0,1\}^\lambda$.
  - response $G_\gamma$ and $\mathcal{L}_G \leftarrow \mathcal{L}_G \cap (\gamma, G_\gamma)$.

**How $\mathcal{B}$ simulate $\mathcal{O}^H$?**

- If $\delta \in \mathcal{L}_H$, then response $H_\delta$.
- Otherwise, response $H_\delta \xleftarrow{\$} \{0,1\}^\lambda$ and $\mathcal{L}_H \leftarrow \mathcal{L}_H \cap (\delta, H_\delta)$.

**How $\mathcal{B}$ simulate $\mathcal{O}^{\mathcal{D}_{\mathsf{sk}}}$?**

- If $c = f(\delta, H_\delta \oplus \gamma)$ and $[G_\gamma \oplus \delta]_{\lambda_1} = 0^{\lambda_1}$, then response $[G_\gamma \oplus \delta]^n$.
- Otherwise, response $\mathsf{reject}$.