MIT 6.875/18.425

Foundations of Cryptography
Lecture 4

Course website: ***https://mit6875.github.io/***

Lecture 3 Recap

- **Theorem**: Next-bit Unpredictability = Indistinguishability for PRGs.

  Key Techniques: Hybrid Argument,
  Predicting-to-Distinguishing Reduction.

- **Theorem**: PRG Length Extension

- **New Notion**: Pseudorandom Functions (PRF)

- **Application of PRFs**: Stateless Secret-key Encryption

TODAY

0. Finish up secret-key encryption.

1. **Theorem**: If there are PRGs, then there are PRFs.

   The Goldreich-Goldwasser-Micali (GGM) construction.

2. **More Applications of PRFs**:

   a. Identification Protocols

   b. Authentication

   c. Applications to Learning Theory

   d. (maybe) Natural Proofs

Pseudorandom Functions

Collection of functions $\mathcal{F}_\ell = \{f_k : \{0,1\}^\ell \to \{0,1\}^m\}_{k \in \{0,1\}^n}$

- indexed by a key k
- n: key length, ℓ: input length, m: output length.
- Independent parameters, all poly(sec-param) = poly(n)
- #functions in $\mathcal{F}_\ell \leq 2^n$ (singly exponential in n)

Gen(1^n): Generate a random n-bit key k.

Eval(k, x) is a poly-time algorithm that outputs $f_k(x)$.

Pseudorandom Functions

Collection of functions $\mathcal{F}_\ell = \{f_k: \{0,1\}^\ell \to \{0,1\}^m\}_{k \in \{0,1\}^n}$

- indexed by a key k
- n: key length, ℓ: input length, m: output length.
- Independent parameters, all poly(sec-param) = poly(n)
- #functions in $\mathcal{F}_\ell \leq 2^n$ (singly exponential in n)

$\approx$

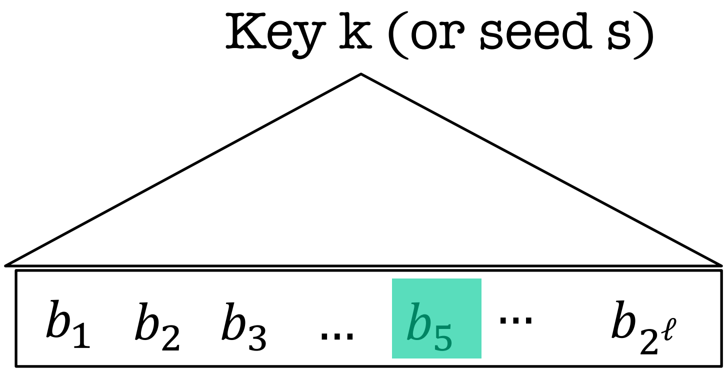
Collection of ALL functions $ALL_\ell = \{f: \{0,1\}^\ell \to \{0,1\}^m\}$

- #functions in $ALL_\ell \leq 2^{m2^\ell}$ (doubly exponential in ℓ)

PRG vs. PRF

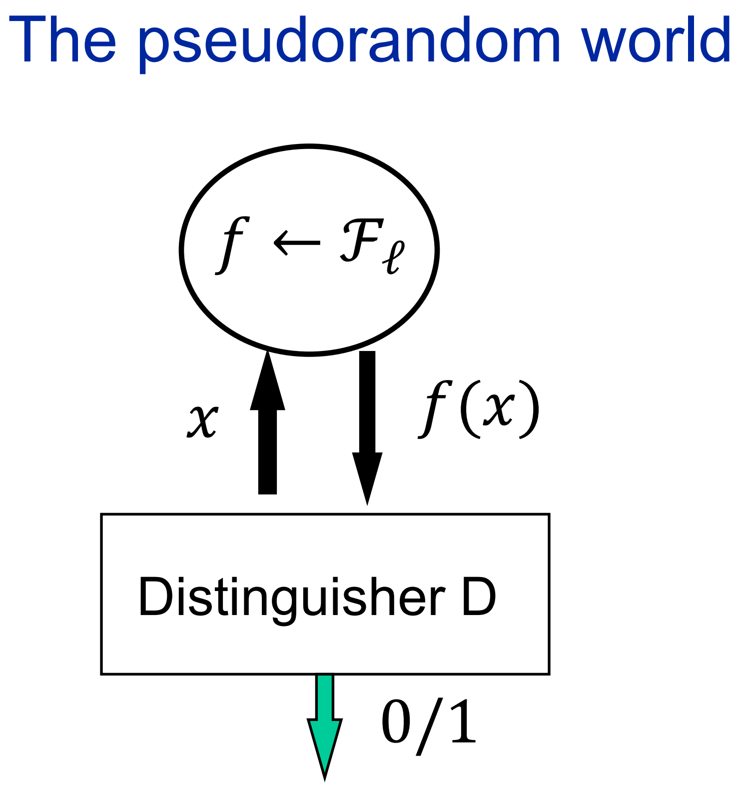
PRG $G(k)$

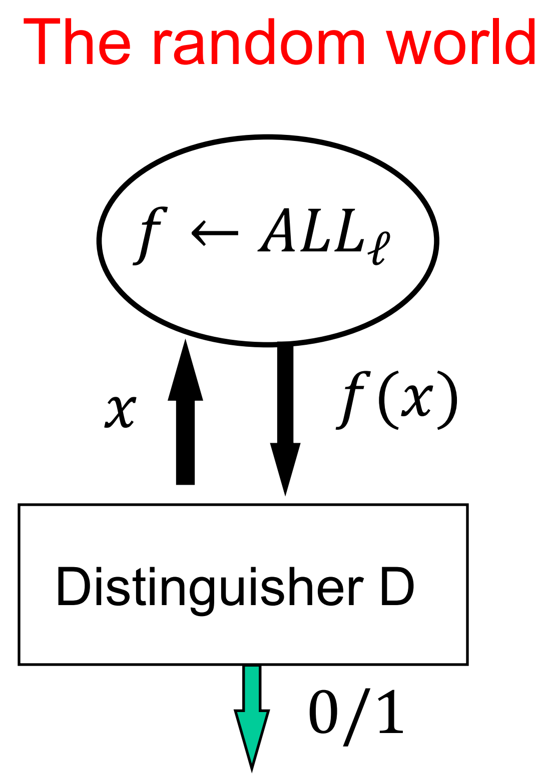
PRF $\mathbf{F}(k,x)$



- Both expand a few random bits into many pseudorandom bits
- With a PRG, accessing the 2^ℓ-th bit takes time 2^ℓ. With a PRF, this can be done in time ℓ.
- **So, a PRF = locally accessible (or random-access) PRG.**

Pseudorandom Functions should be “indistinguishable” from random





For all ppt D, there is a negligible function μ s.t.

$$\left| \Pr[f \leftarrow \mathcal{F}_\ell : D^f(1^n) = 1] - \Pr[f \leftarrow ALL_\ell : D^f(1^n) = 1] \right| \le \mu(n)$$

PRF $\Longrightarrow$ Stateless Secret-key Encryption

$Gen(1^n)$: Generate a random n-bit key k that defines

$$f_k: \{0,1\}^\ell \to \{0,1\}^m$$

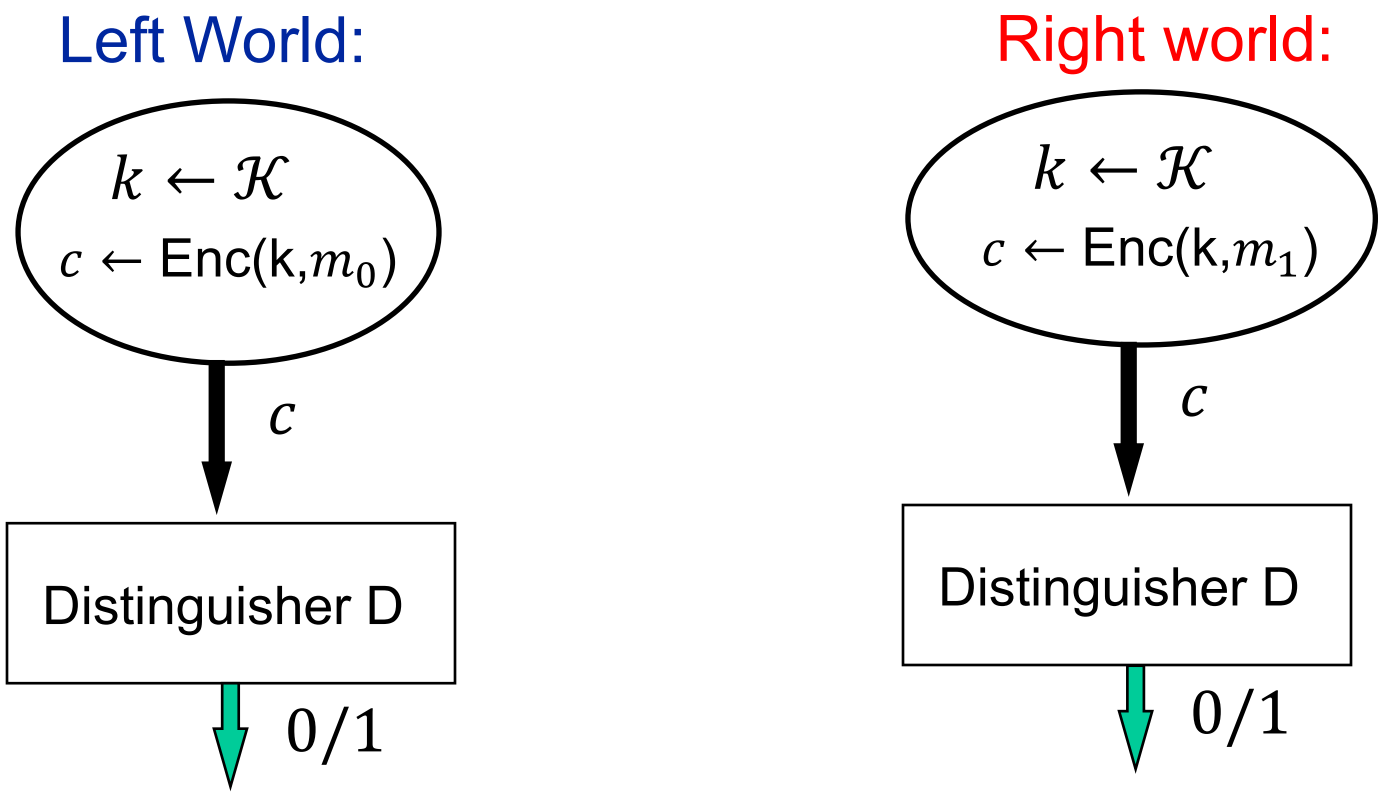
(the domain size, 2^ℓ, had better be super-polynomially large in n)

$Enc(k,m)$: Pick a random x and
let the ciphertext c be the pair $(x, y = f_k(x) \oplus m)$.

$Dec(k, c = (x,y))$: Output $f_k(x) \oplus y$.

Recall: Definition of Secret-Key Encryption

(for one message)



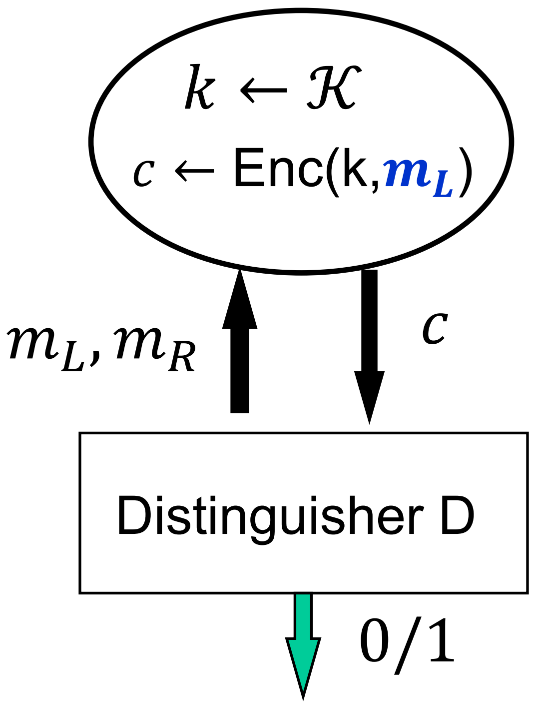
For all m_0, m_1, and all ppt D, there is a negligible function μ s.t.

$$\left| \Pr[k \leftarrow \mathcal{K}: D(Enc(k, m_0)) = 1] - \Pr[k \leftarrow \mathcal{K}: D(Enc(k, m_1)) = 1] \right| \leq \mu(n)$$

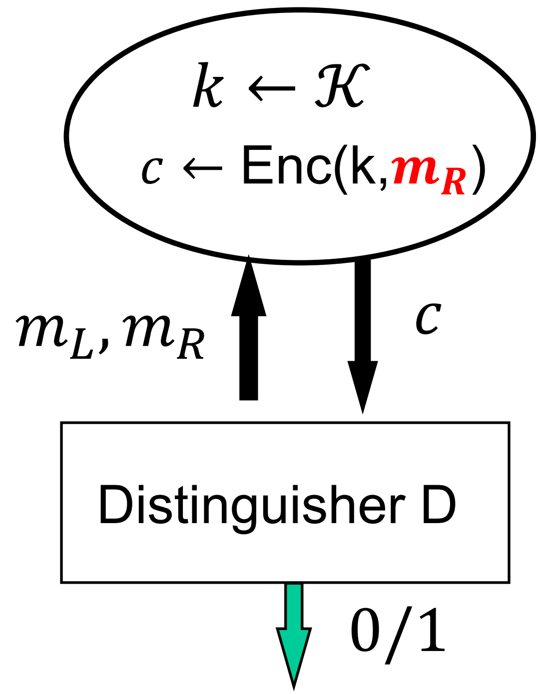
Definition of Secret-Key Encryption

(for many messages)

Left Oracle $Left(\cdot,\cdot)$



Right Oracle $Right(\cdot,\cdot)$



For all ppt D, there is a negligible function μ s.t.

$$\left| \Pr[k \leftarrow \mathcal{K}: D^{Left(\cdot,\cdot)}(1^n) = 1] - \Pr[k \leftarrow \mathcal{K}: D^{Right(\cdot,\cdot)}(1^n) = 1] \right| \leq \mu(n)$$

Proof

Hybrid 0: D gets access to the Left oracle.

$$c = (x, y = f_k(x) \oplus m_L)$$

$\approx$ by PRF security

Hybrid 1: Replace f_k by a random function.

$$c = (x, y = r_x \oplus m_L)$$

$\approx$ by birthday paradox

(w.h.p. all x's distinct)

Hybrid 2: Replace f_k by a random function.

$$c = (x, y = r_x)$$

$\approx$ by birthday paradox

Hybrid 3: Replace f_k by a random function (like H1)

$$c = (x, y = r_x \oplus m_L)$$

$\approx$ by PRF security

Hybrid 4: D gets access to the Right oracle (like H0)

$$c = (x, y = f_k(x) \oplus m_R)$$

TODAY

0. Finish up secret-key encryption.

1. **Theorem**: If there are PRGs, then there are PRFs.

   The Goldreich-Goldwasser-Micali (GGM) construction.

2. **More Applications of PRFs**:

   a. Identification Protocols

   b. Authentication

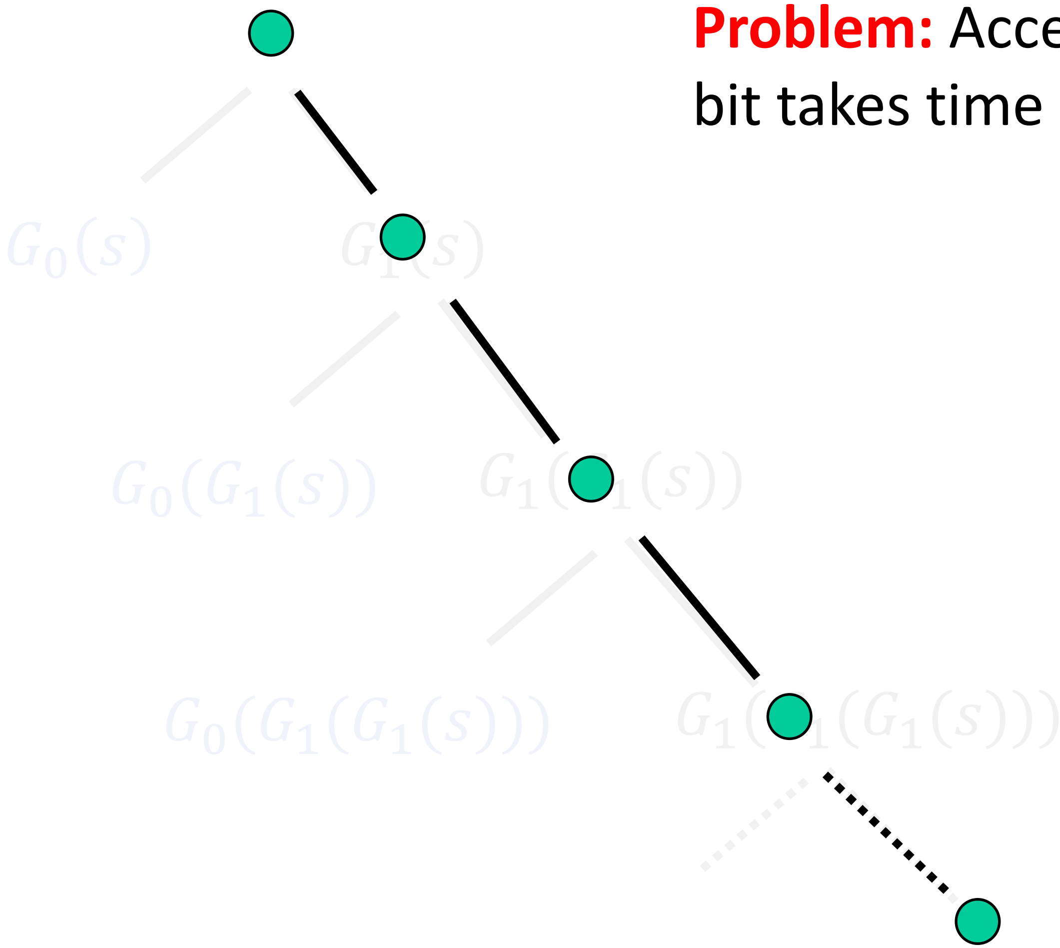
   c. Applications to Learning Theory

Let’s Look Back at Length Extension…

Theorem: Let G: $\{0,1\}^n \rightarrow \{0,1\}^{n+1}$ be a PRG. Then, for every polynomial m(n), there is a PRG G’: $\{0,1\}^n \rightarrow \{0,1\}^{m(n)}$.

Let’s Look Back at Length Extension…

Construction: Let G(s) = $G_0(s)||G_1(s)$ where $G_0(s)$ is 1 bit and $G_1(s)$ is n bits .



Problem: Accessing the i^{th} output bit takes time $\approx i$.

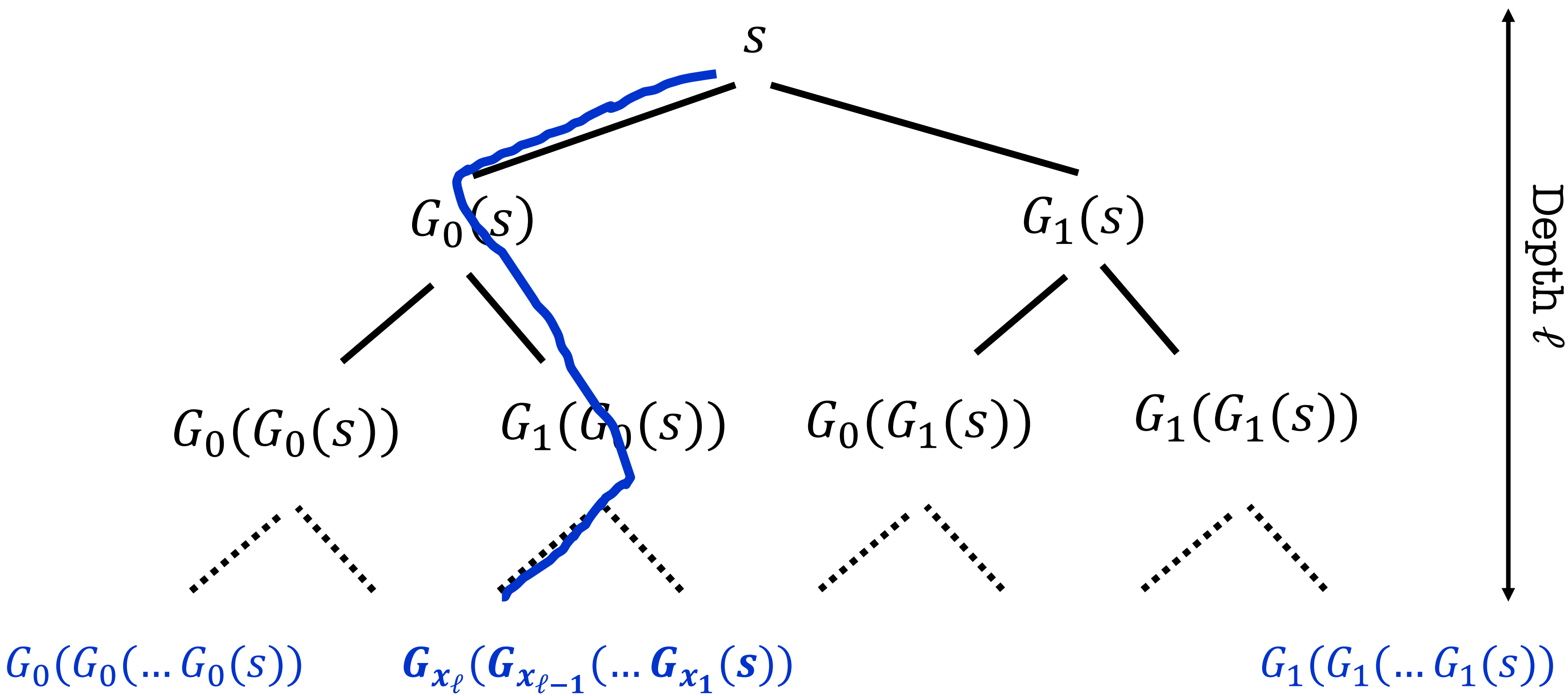
Goldreich-Goldwasser-Micali PRF

Theorem: Let G be a PRG. Then, for every polynomials $\ell = \ell(n)$, $m = m(n)$, there exists a PRF family $\mathcal{F}_\ell = \{f_s: \{0,1\}^\ell \to \{0,1\}^m\}_{s \in \{0,1\}^n}$.

Note: We will focus on $m = \ell$.
The output length could be made smaller (by truncation) or larger (by expansion with a PRG).

Goldreich-Goldwasser-Micali PRF

Construction: Let G(s) = $G_0(s)||G_1(s)$ where $G_0(s)$ and $G_1(s)$ are both n bits each.



Each path/leaf labeled by $x \in \{0,1\}^\ell$ corresponds to $f_s(x)$.

Goldreich-Goldwasser-Micali PRF

Construction: Let G(s) = $G_0(s)||G_1(s)$ where $G_0(s)$ and $G_1(s)$ are both n bits each.

The pseudorandom function family $\mathcal{F}_\ell$ is defined by a collection of functions f_s where:

$$f_s(\underbrace{x_1 x_2 \ldots x_\ell}_{\ell\text{-bit input}}) = \boldsymbol{G_{x_\ell}(G_{x_{\ell-1}}(\ldots G_{x_1}(s))}$$

- f_s defines 2^ℓ pseudorandom bits.
- The x^{th} bit can be computed using ℓ evaluations of the PRG G (as opposed to $x \approx 2^\ell$ evaluations as before.)

PRG Repetition Lemma

Lemma: Let G be a PRG. Then, for every polynomial L=L(n), the following two distributions are computationally indistinguishable:

$$\left(\boldsymbol{G}(\boldsymbol{s_1}), \boldsymbol{G}(\boldsymbol{s_2}), \dots, \boldsymbol{G}(\boldsymbol{s_L})\right) \approx \left(\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_L}\right)$$

Proof: **By Hybrid Argument.**

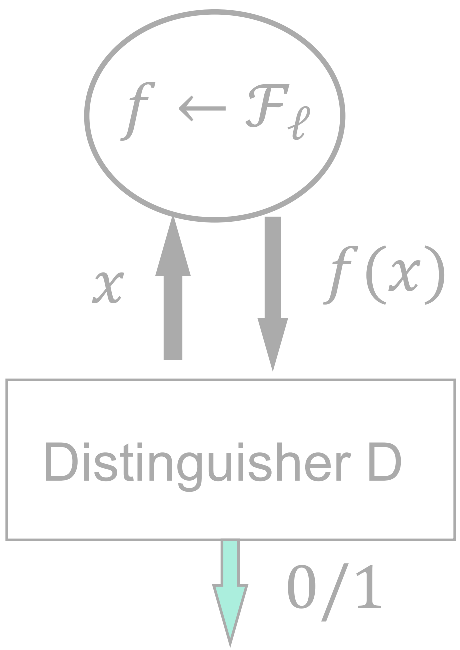
If there is a ppt distinguisher between the two distributions with distinguishing advantage ε, then there is a ppt distinguisher for G with advantage $\geq \varepsilon/L$.

GGM PRF: Proof of Security

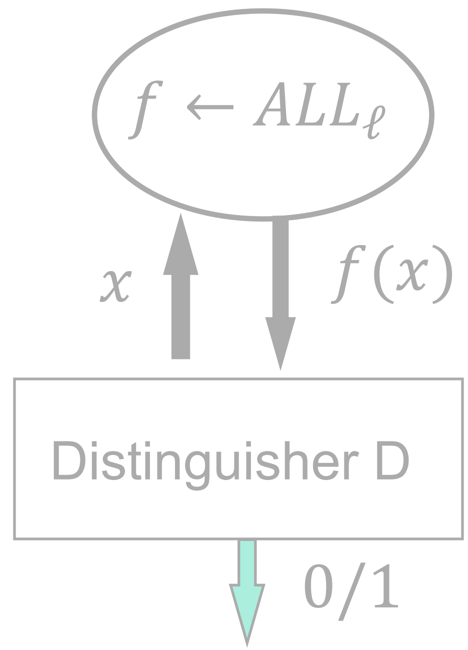
By contradiction. Assume there is a ppt D and a poly function p s.t.

$$\left| \Pr[f \leftarrow \mathcal{F}_\ell : D^f(1^n) = 1] - \Pr[f \leftarrow ALL_\ell : D^f(1^n) = 1] \right| \geq 1/p(n)$$

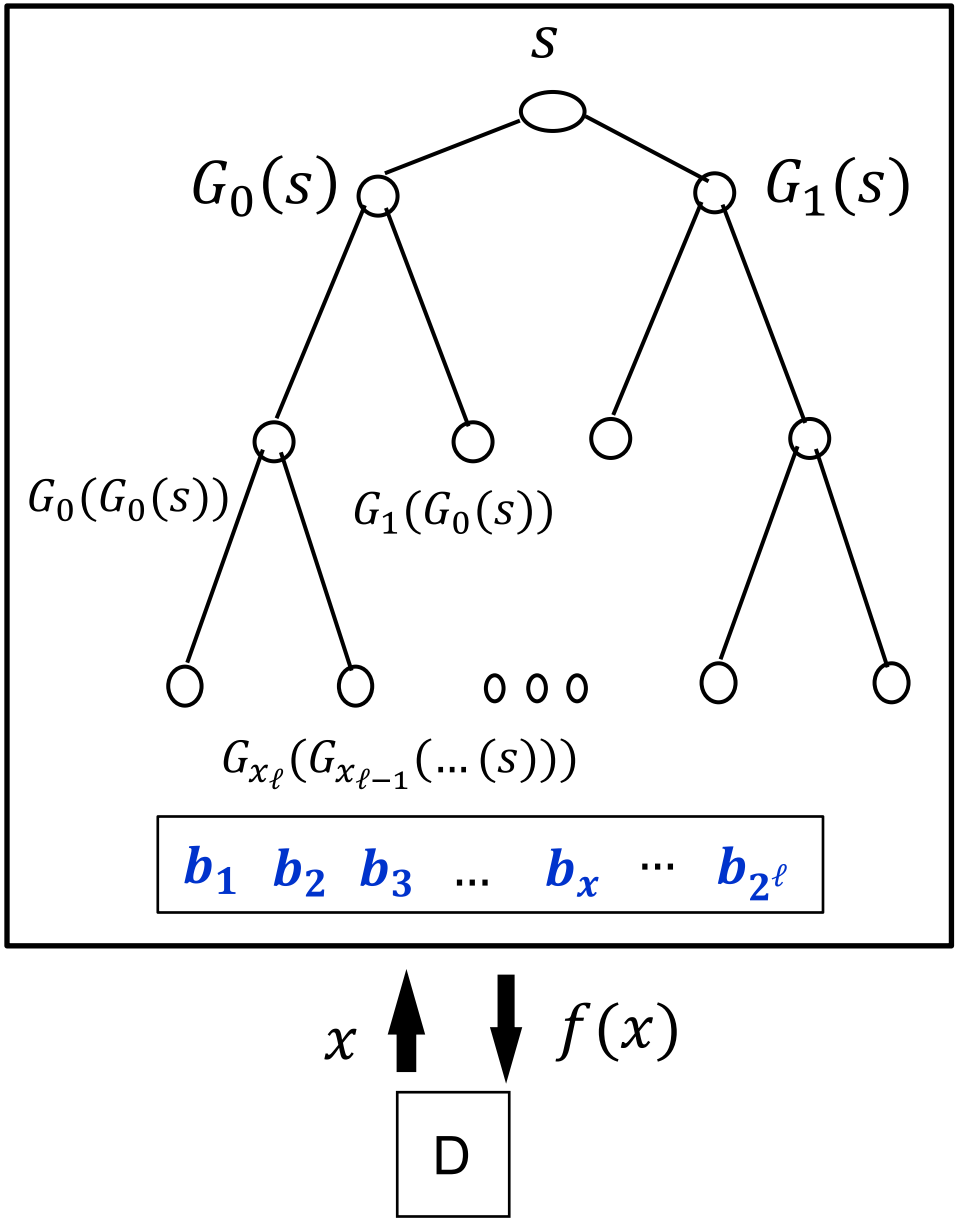
The pseudorandom world



The random world

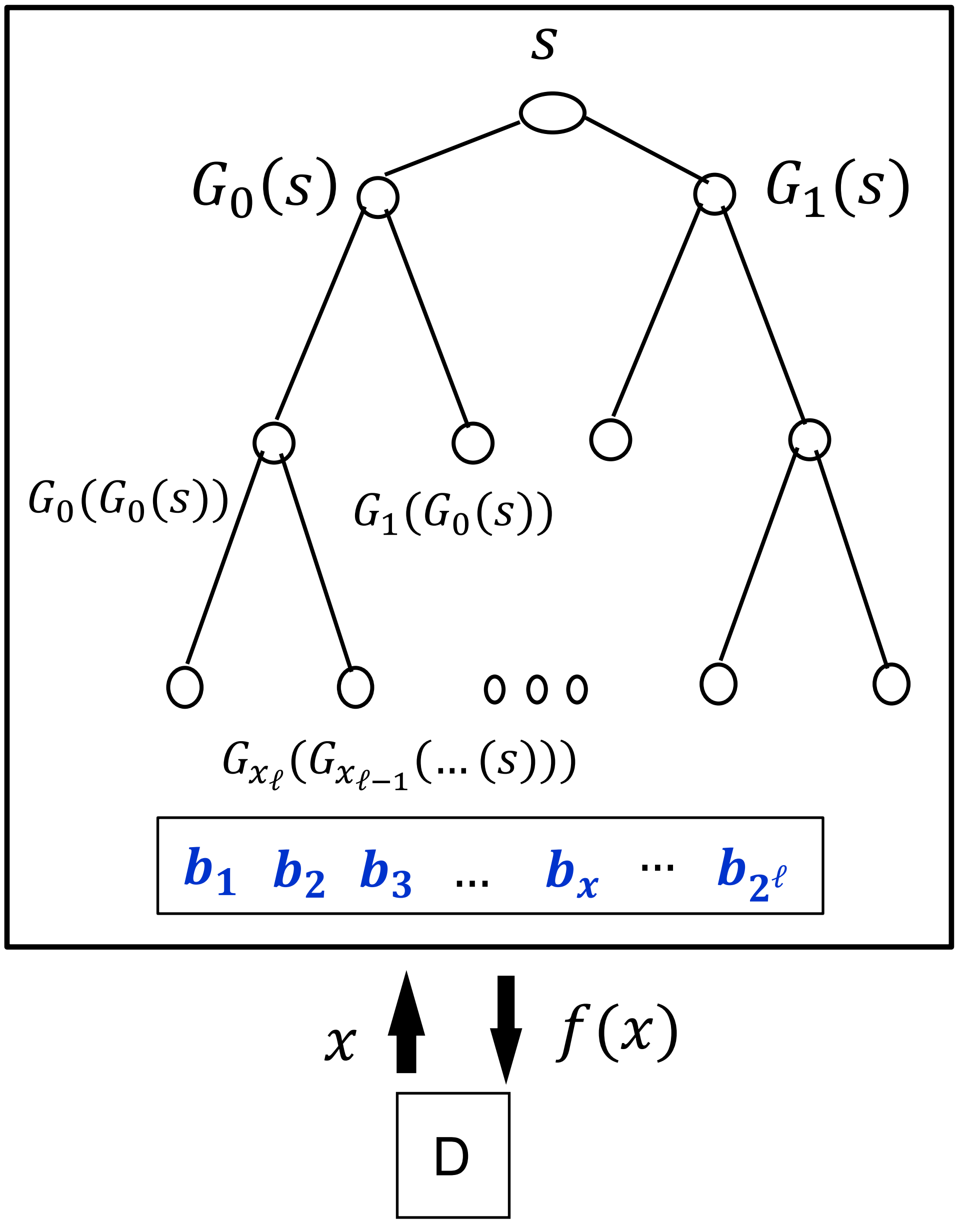


The pseudorandom world:
Hybrid 0

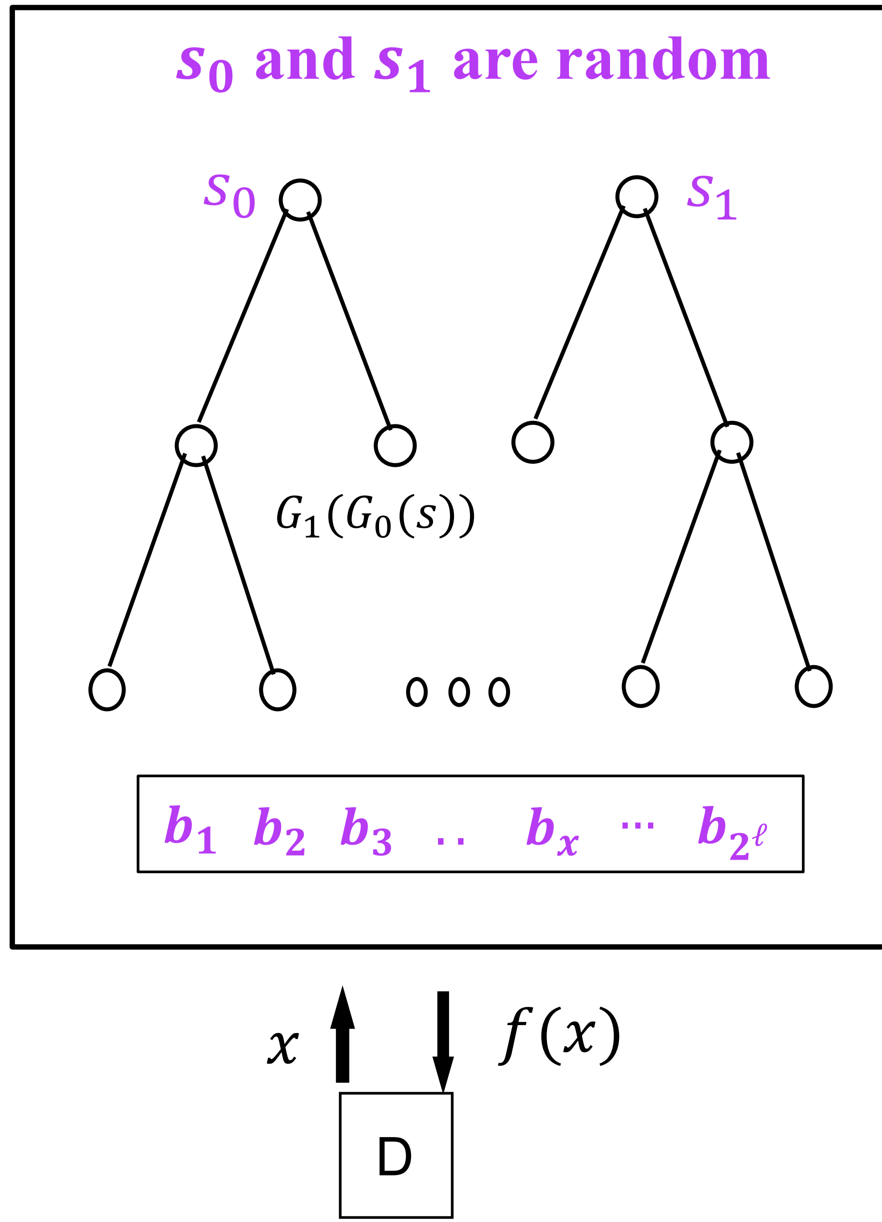


Key Idea:
Hybrid argument by levels of the tree

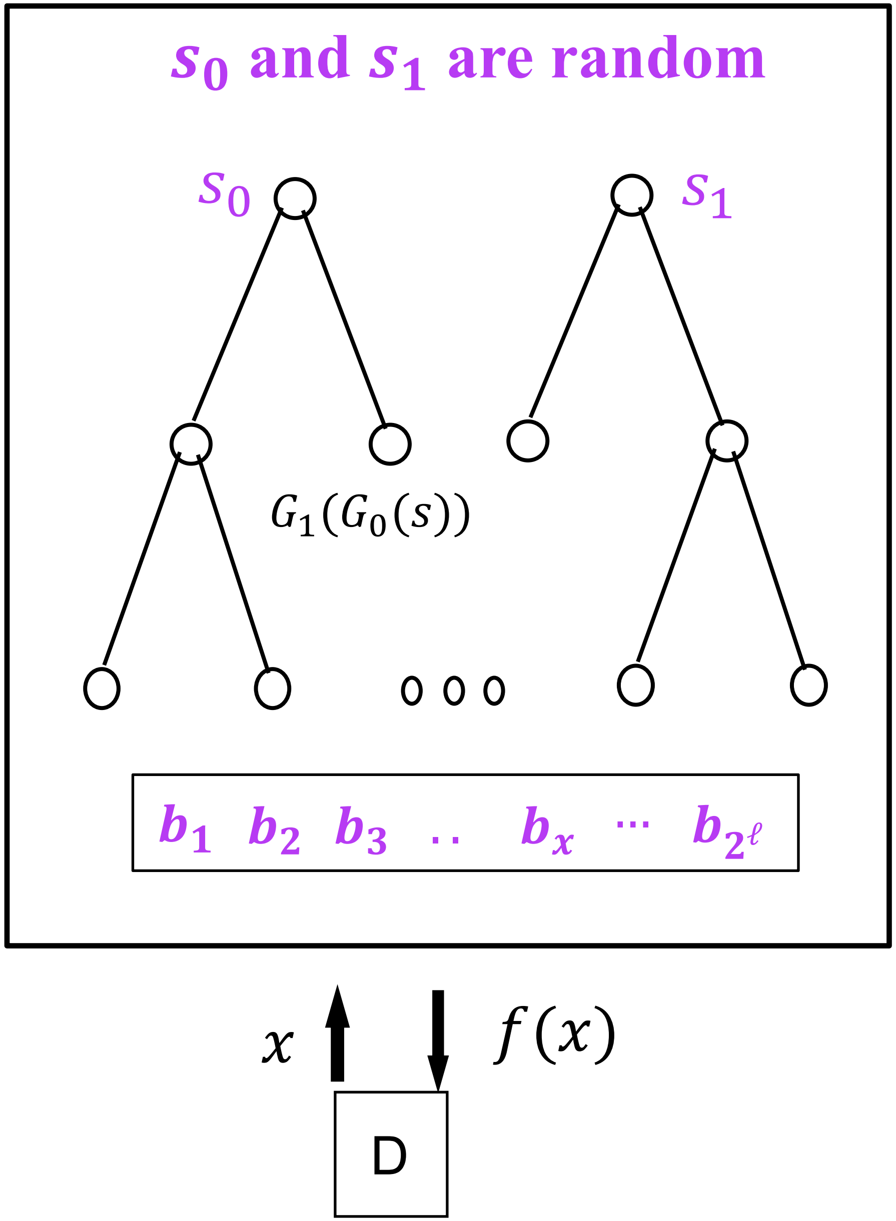
The pseudorandom world: Hybrid 0



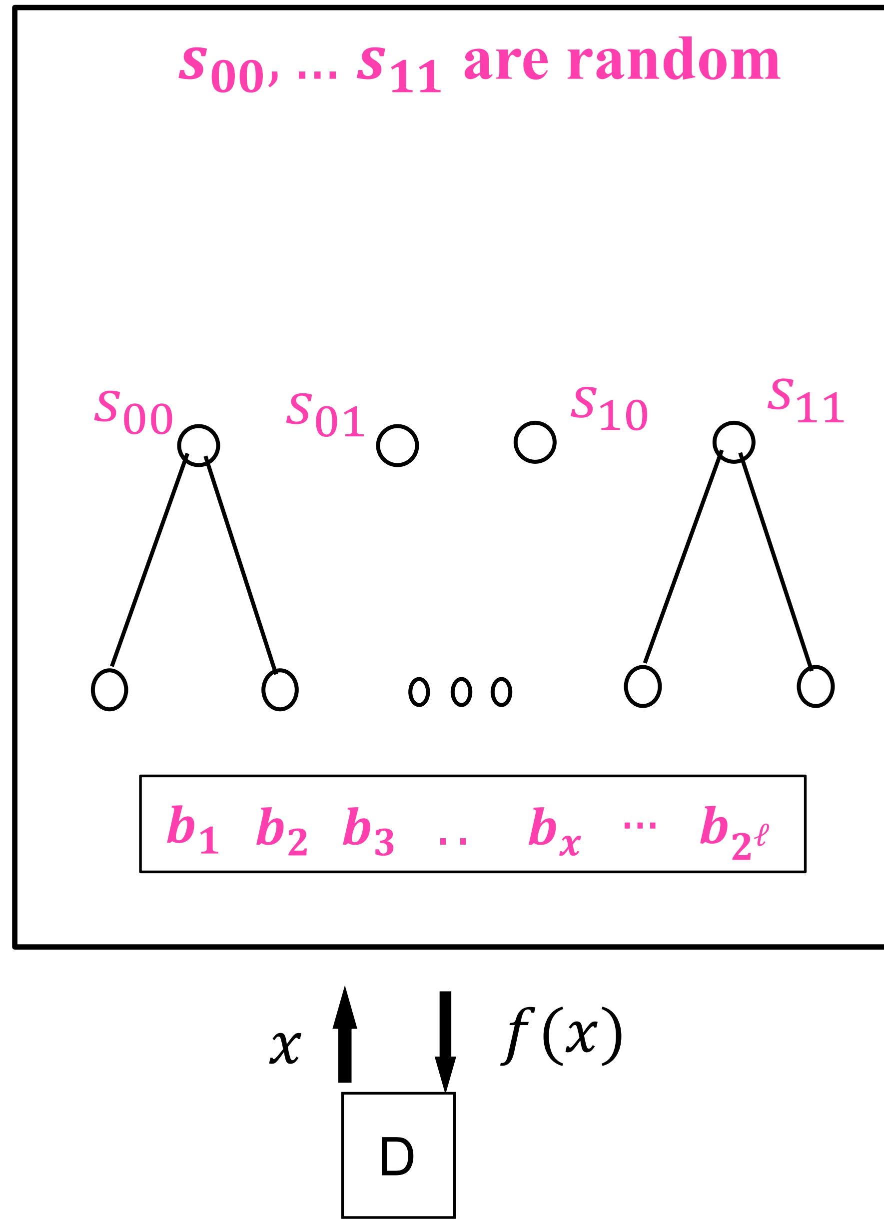
Hybrid 1



Hybrid 1

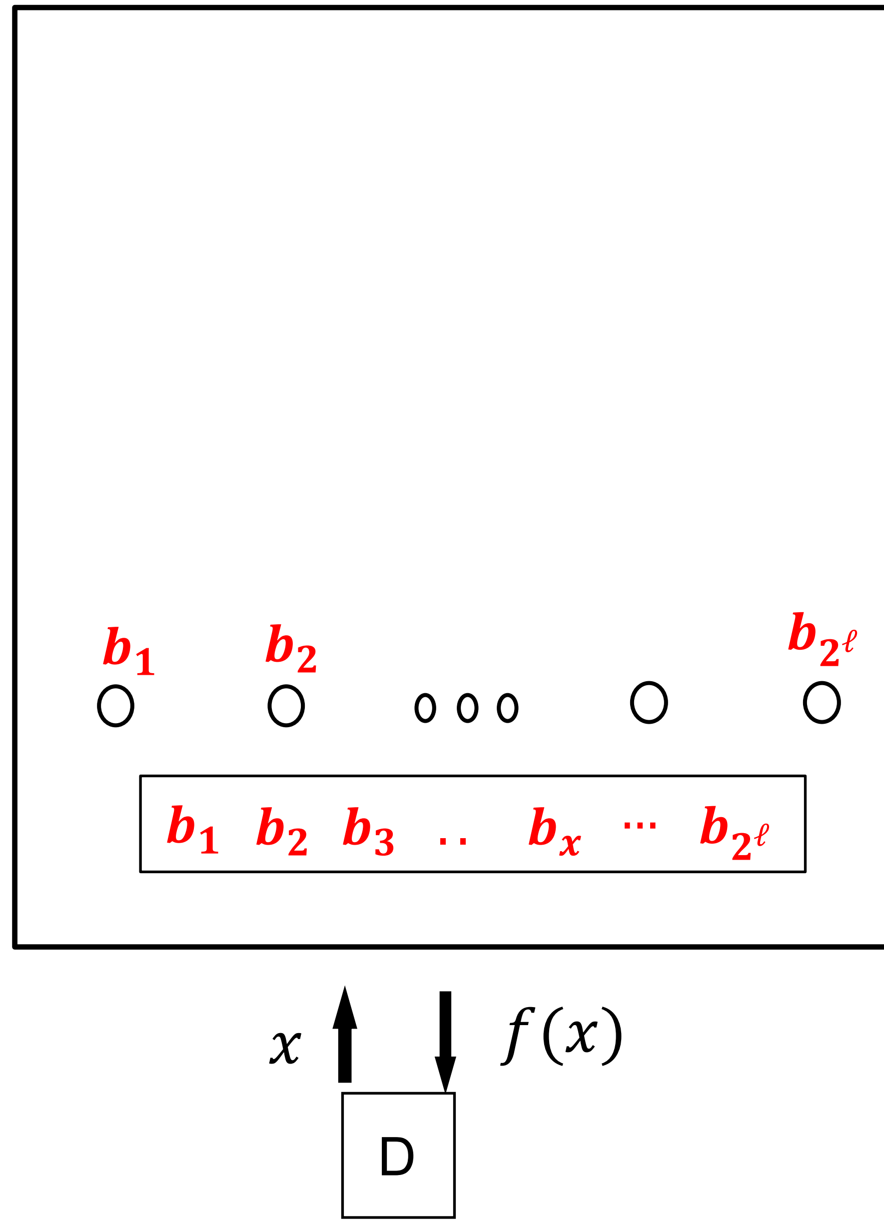


Hybrid 2

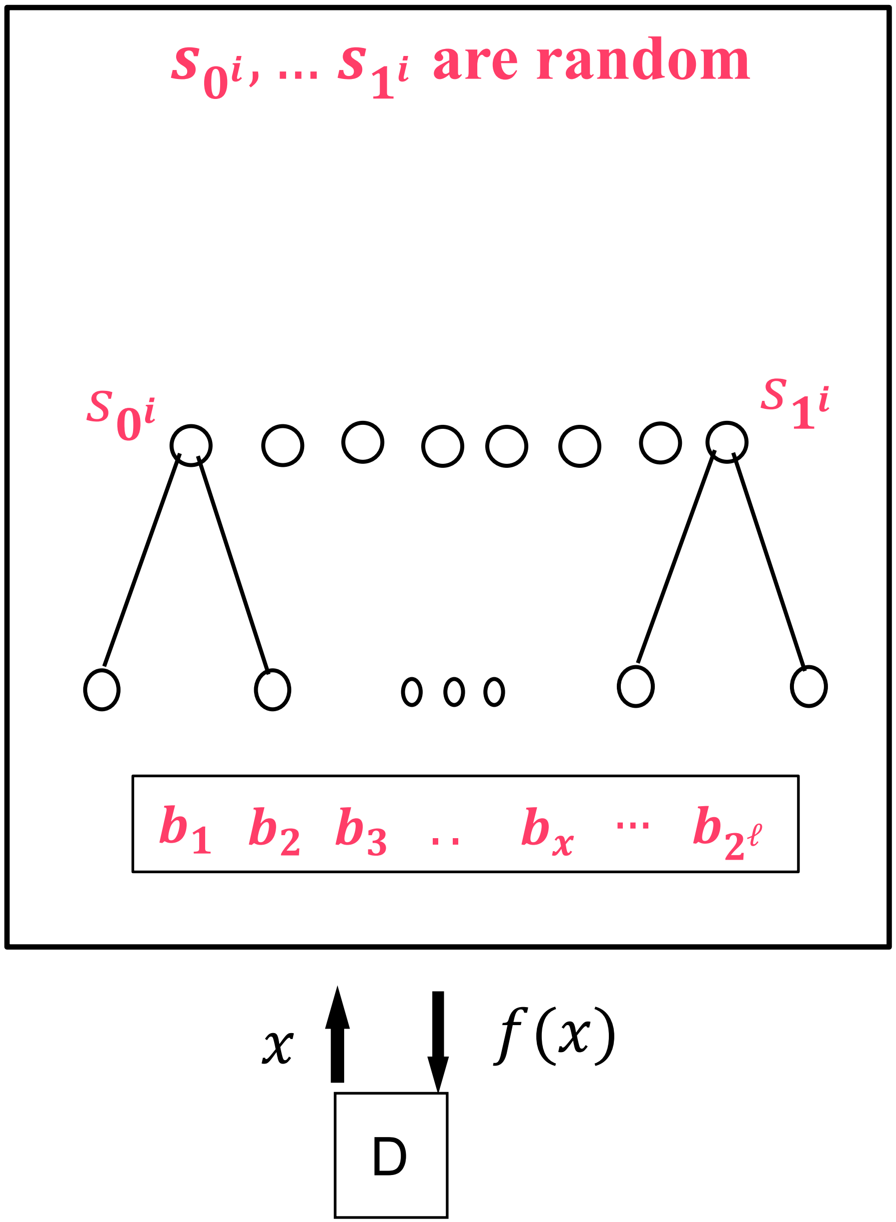


The random world:
Hybrid ℓ

...



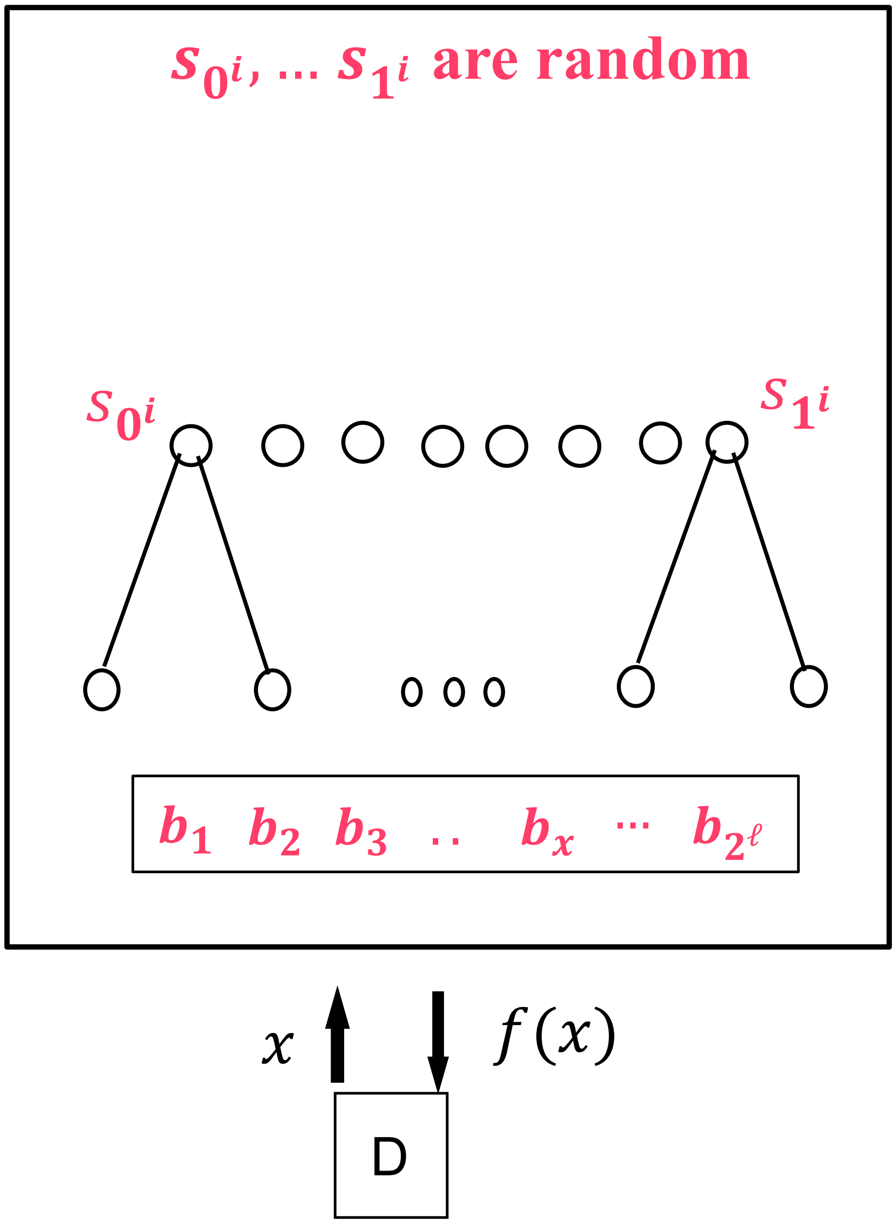
Hybrid i



Q: **Are the hybrids efficiently computable?**

A: **Yes! Lazy Evaluation.**

Hybrid i



Let $p_i = \Pr[f \leftarrow H_i : D^f(1^n) = 1]$

We know: $p_0 - p_\ell \geq \varepsilon$

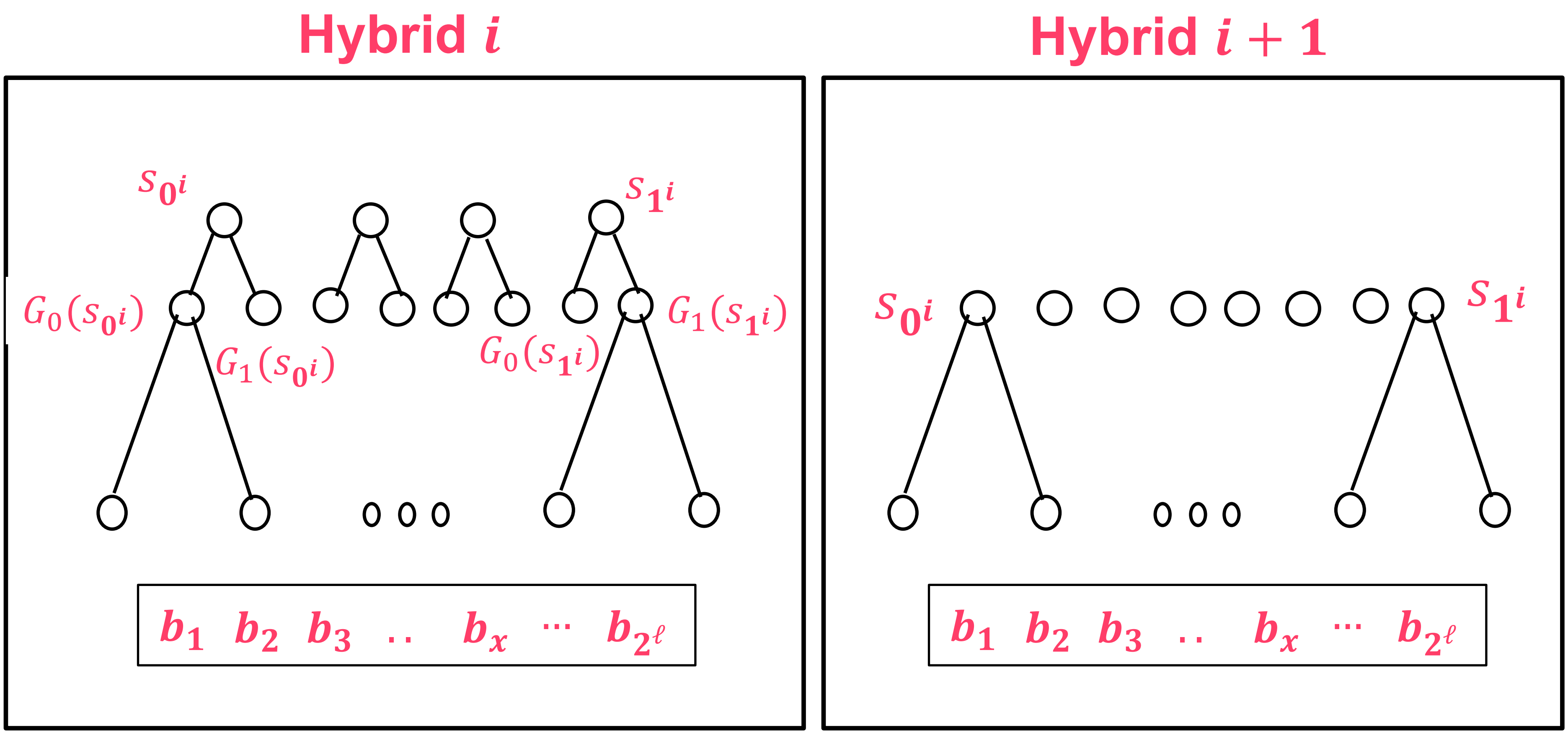
By a hybrid argument:

For some i: $p_i - p_{i+1} \geq \varepsilon/\ell$

(use the PRG repetition lemma)

***A distinguisher with advantage ε/ℓ between the hybrids implies a distinguisher with advantage $\geq \varepsilon/q\ell$* for the PRG.**

(where q is the number of queries that D makes)



GGM PRF

Theorem: Let G be a PRG. Then, for every polynomials ℓ, m, there exists a PRF family $\mathcal{F}_\ell = \{f_s\colon \{0,1\}^\ell \to \{0,1\}^m\}_{s\in\{0,1\}^n}$.

Some nits:

- *Expensive*: ℓ invocations of a PRG.
- *Sequential*: bit-by-bit, ℓ sequential invocations of a PRG.
- *Loss in security reduction*: break PRF with advantage $\varepsilon \Longrightarrow$ break PRG with advantage $\varepsilon/q\ell$, where q is an arbitrary polynomial = #queries of the PRF distinguisher.

  Tighter reduction? Avoid the loss?

TODAY

0. Finish up secret-key encryption.

1. **Theorem**: If there are PRGs, then there are PRFs.

   The Goldreich-Goldwasser-Micali (GGM) construction.

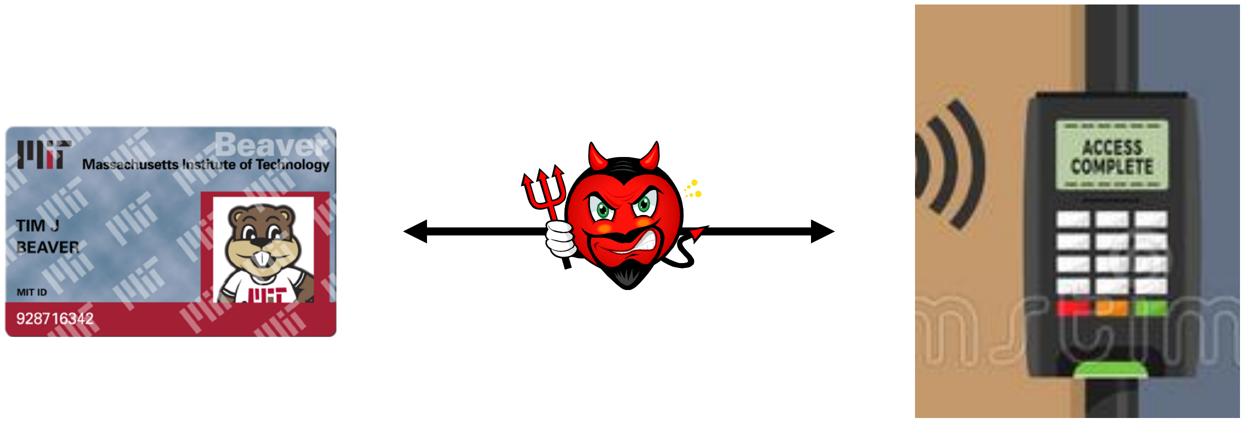
2. **More Applications of PRFs**:

   a. Identification Protocols

   b. Authentication

   c. Applications to Learning Theory

Friend-or-Foe Identification



- **Adversary**: person-in-the-middle.
- Can listen to / modify the communications. Wants to impersonate Tim.

A Simple Lemma about Unpredictability

Let $f_s\colon \{0,1\}^\ell \to \{0,1\}^m$ be a pseudorandom function.

- Consider an adversary who requests and obtains $f_s(x_1), \dots, f_s(x_q)$ for a polynomial $q = q(n)$.

- Can she predict $f_s(x^*)$ for some x^* of her choosing where $x^* \notin \{x_1,\dots, x_q\}$? How well can she do it?

Lemma: If she succeeds with probability $\frac{1}{2^m} + 1/\mathrm{poly}(n)$, then she broke PRF security. This is negligible in n if m is large enough, i.e. $\omega(\log n)$.

A Simple Lemma about Unpredictability

Let $f_s: \{0,1\}^\ell \to \{0,1\}^m$ be a pseudorandom function.

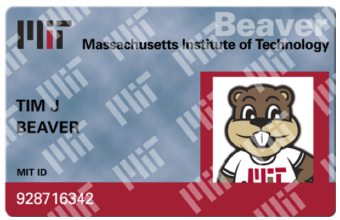
- Consider an adversary who requests and obtains $f_s(x_1), \dots, f_s(x_q)$ for a polynomial $q = q(n)$.

- Can she predict $f_s(x^*)$ for some x^* of her choosing where $x^* \notin \{x_1, \dots, x_q\}$? How well can she do it?

- Unpredictability $\equiv$ Indistinguishability *for bits* (lecture 3)
- Indistinguishability $\Longrightarrow$ Unpredictability (*but not vice versa*).

Challenge-Response Protocol

Random r



$(ID, f_s(r))$

PRF Key s

(ID number ID, PRF Key s)

"Proof": Adversary collects $(r_i, f_s(r_i))$ for poly many r_i (potentially of her choosing). She eventually has to produce $f_s(r^*)$ for a fresh random r^* when she is trying to impersonate.

This is hard as long as the input and output lengths of the PRF are long enough, i.e. $\omega(\log n)$.

TODAY

1. **Theorem**: If there are PRGs, then there are PRFs.

   The Goldreich-Goldwasser-Micali (GGM) construction.

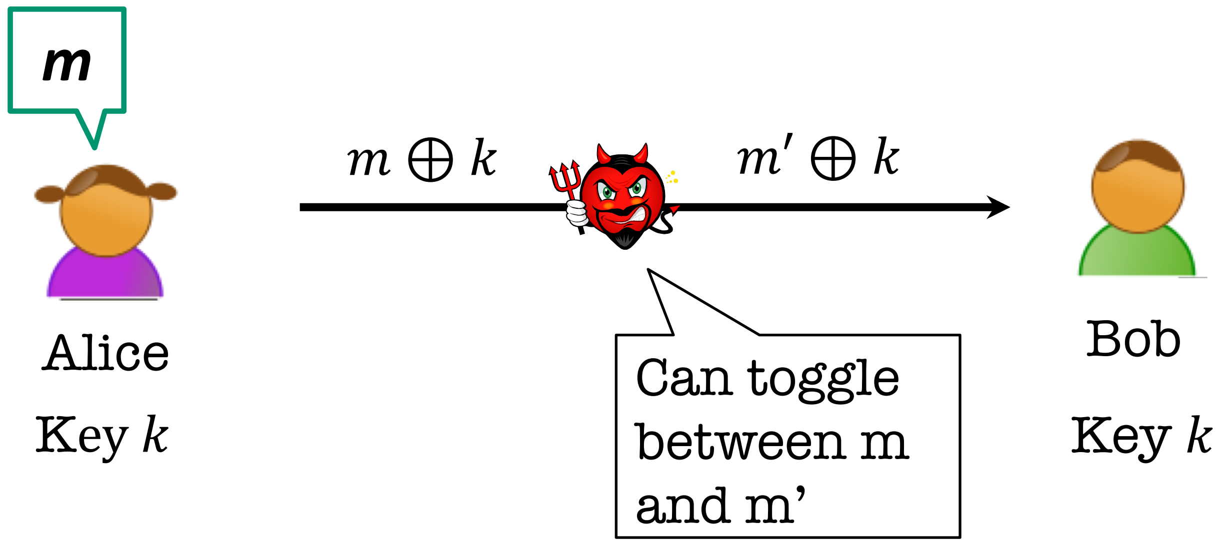
2. **More Applications of PRFs**:

   a. Identification Protocols

   b. Authentication

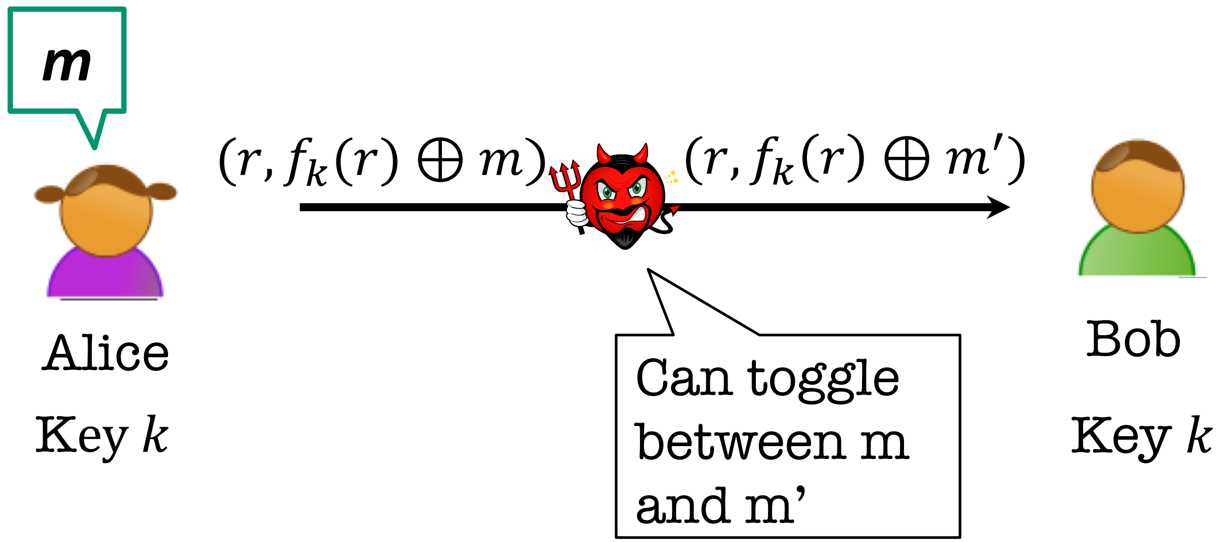
   c. Applications to Learning Theory

Secure Communication



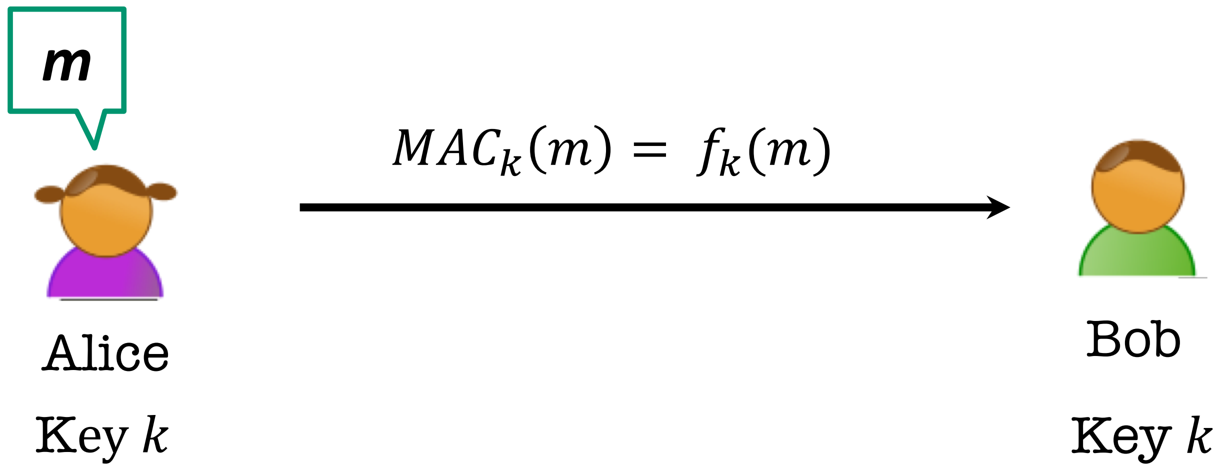
One-time pad (and encryption schemes in general) are *malleable*.

Secure Communication



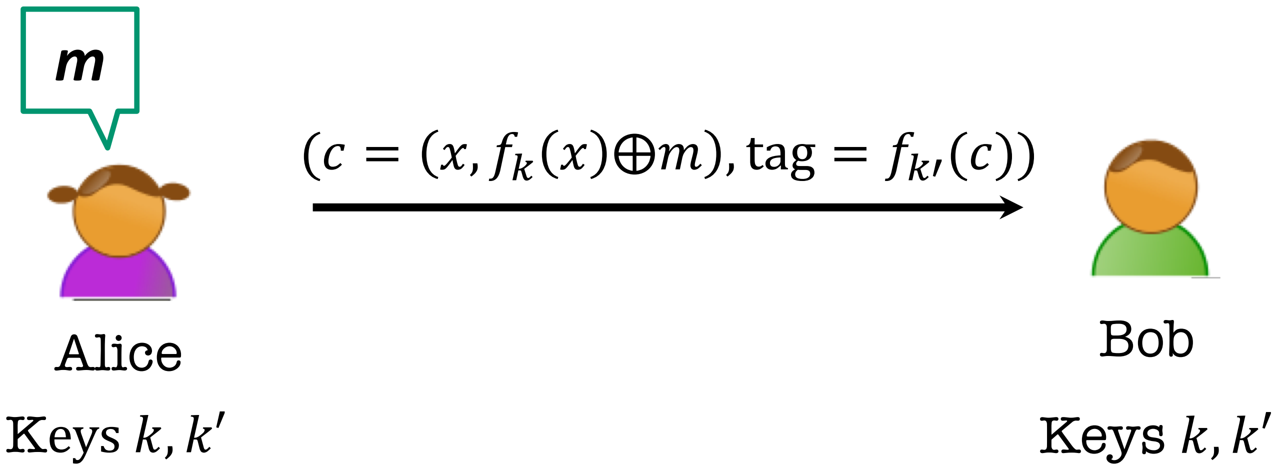
One-time pad (and encryption schemes in general) are ***malleable***.

Message Authentication Codes



MACs give us integrity, but not privacy.

Message Authentication Codes



MACs give us integrity, but not privacy.

Solution: Encrypt, then MAC (more in pset 3)

TODAY

1. **Theorem**: If there are PRGs, then there are PRFs.

   The Goldreich-Goldwasser-Micali (GGM) construction.

2. **More Applications of PRFs**:

   a. Identification Protocols

   b. Authentication

   c. Applications to Learning Theory

Negative Results in Learning Theory

Theorem [Kearns and Valiant 1994]:
Assuming PRFs exist, there are hypothesis classes that cannot be learned by polynomial-time algorithms.