COS433/Math 473: Cryptography

Mark Zhandry

Princeton University

Spring 2020

Announcements/Reminders

Mark's Office hours on 2/21: Moved to 11am

HW1 Due TODAY
HW2 Due Feb 27th

PR1 Due March 10th

Previously on COS 433…

Security Experiment/Game

(One-time setting)

b

Challenger

$m_0, m_1 \in M_\lambda$

c

$k \leftarrow K_\lambda$

$c \leftarrow Enc(k, m_b)$

b'

$\text{IND-Exp}_b($  $, \lambda)$

Security Definition

(One-time setting, asymptotic)

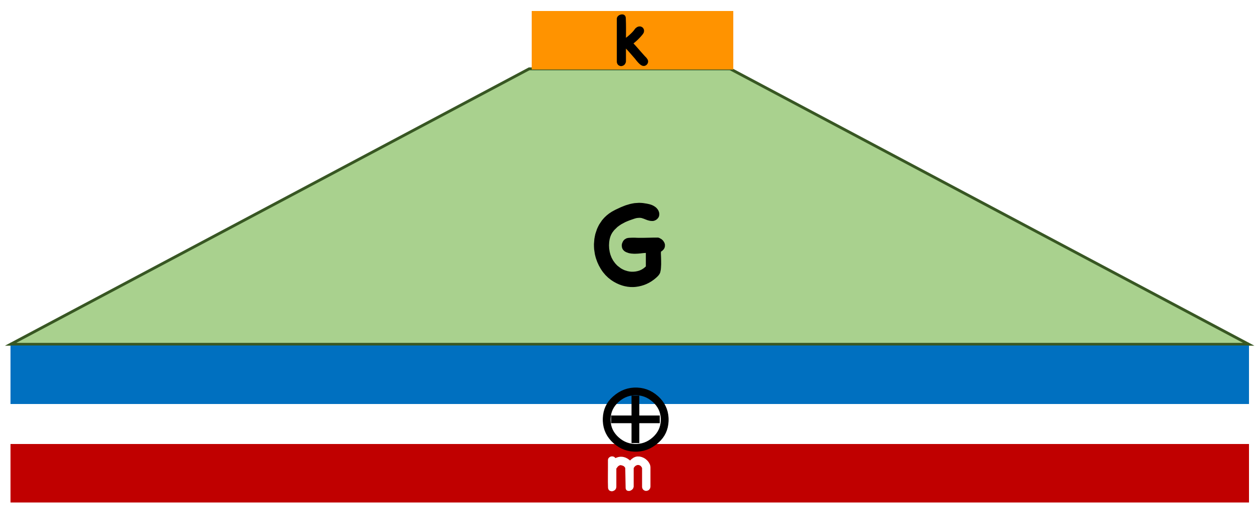
Definition: (Enc, Dec) has **ciphertext indistinguishability** if, for all  running in polynomial time, $\exists$ negligible ε s.t.

| Pr[1←IND-Exp$_0$(,λ)]

− Pr[1←IND-Exp$_1$(,λ)] | ≤ ε(λ)

Construction with |k| << |m|

Idea: use OTP, but have key generated by some expanding procedure **G**



What do we want out of **G**?

Defining Pseudorandom Generator (PRG)

Syntax:

- Seed space $\mathbf{S}_\lambda$
- Output space $\mathbf{X}_\lambda$
- $\mathbf{G}$: $\mathbf{S}_\lambda \rightarrow \mathbf{X}_\lambda$ (deterministic)

Correctness:

- $|s|=\log|S_\lambda|$, $|x|=\log|X_\lambda|$ polynomial in $\boldsymbol{\lambda}$,
- $|X_\lambda| > 2\times|S_\lambda|$
- Running time of $\mathbf{G}$ polynomial in $\boldsymbol{\lambda}$

Security of PRGs

Definition: $G:S_\lambda \rightarrow X_\lambda$ is a **secure pseudorandom generator** (PRG) if:

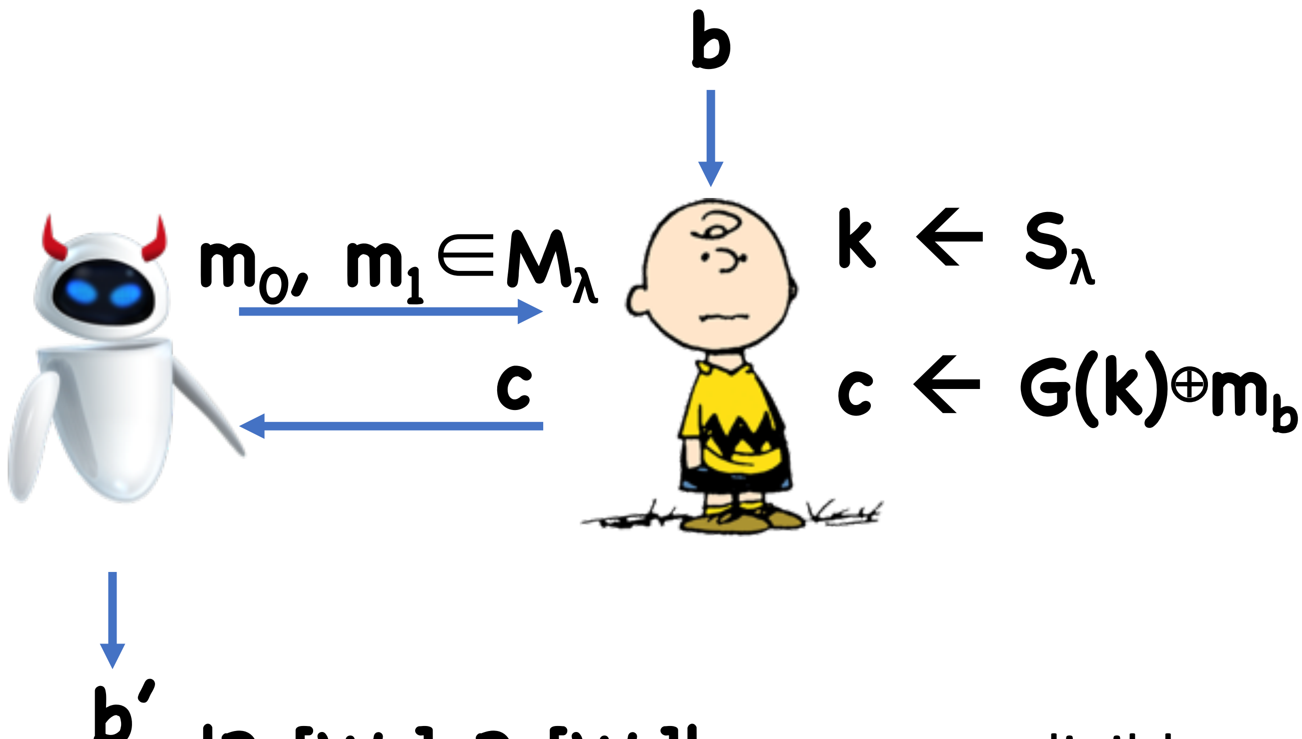
- For all  running in polynomial time, $\exists$ negl **ε,**

| Pr[ $(G(s))=1 : s \leftarrow S_\lambda]$

$-$ Pr[ $(x)=1 : x \leftarrow X_\lambda]$ | $\leq \varepsilon(\lambda)$

Security

Assume towards contradiction that there is a  and non-negligible **ε** such that

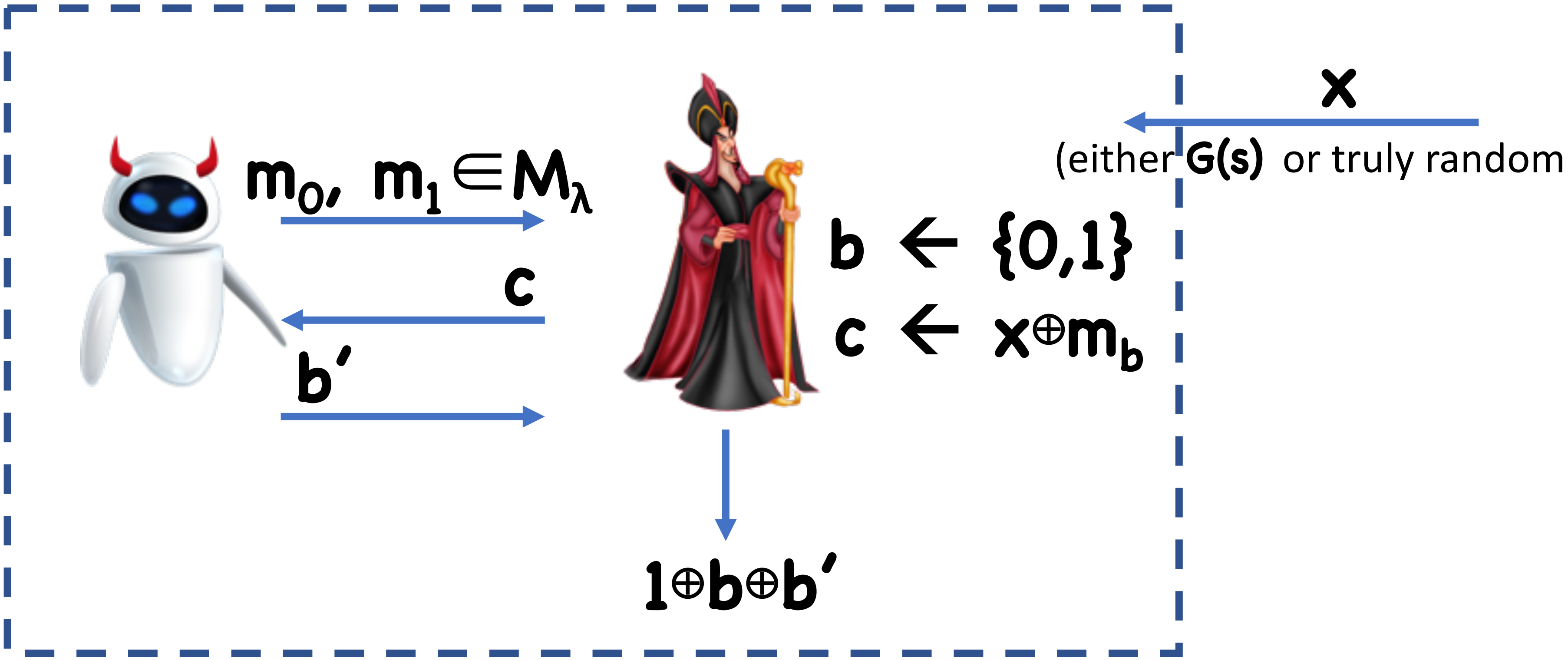


$|\Pr[W_0]-\Pr[W_1]| \geq \varepsilon$, non-negligible

W_b: $b' = 1$ in IND-Exp_b

Security

Use  to build .  will run  as a subroutine, and pretend to be 



Security

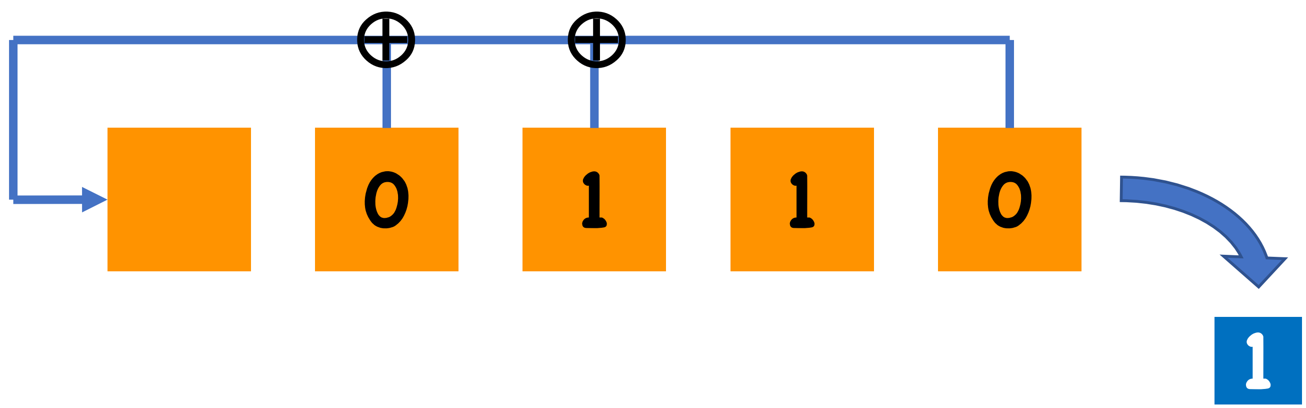
Thm: If **G** is a secure PRG, then **(Enc,Dec)** is has ciphertext indistinguishability

Today: Constructing PRGs

Linear Feedback Shift Registers

In each step,

- last bit of state is removed and outputted
- Rest of bits are shifted right
- First bit is XOR of subset of remaining bits



Linear Feedback Shift Registers

Are LFSR's secure PRGs?

No!

First $\mathbf{n}$ bits of output = initial state

$\mathbf{x}$

Write $\mathbf{x} = \mathbf{x_1,\dots,x_n, x'}$

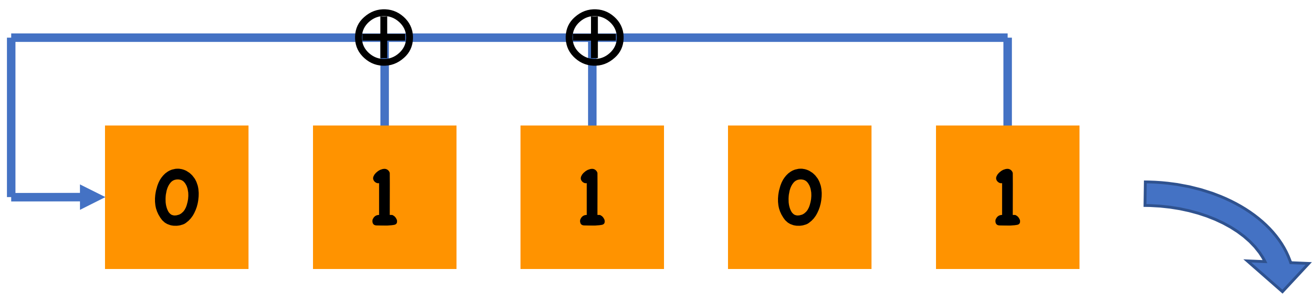
Initialize LFSB to have state $\mathbf{x_1,\dots,x_n}$

Run LFSB for $\mathbf{|x|}$ steps, obtaining $\mathbf{y}$

Check if $\mathbf{y} = \mathbf{x}$

Linearity

LFSR's are linear:



$$\text{state}' = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \bullet \text{state}$$

$$\text{output} = (0\ 0\ 0\ 0\ 1) \bullet \text{state}$$

Linearity

LFSR's are linear:

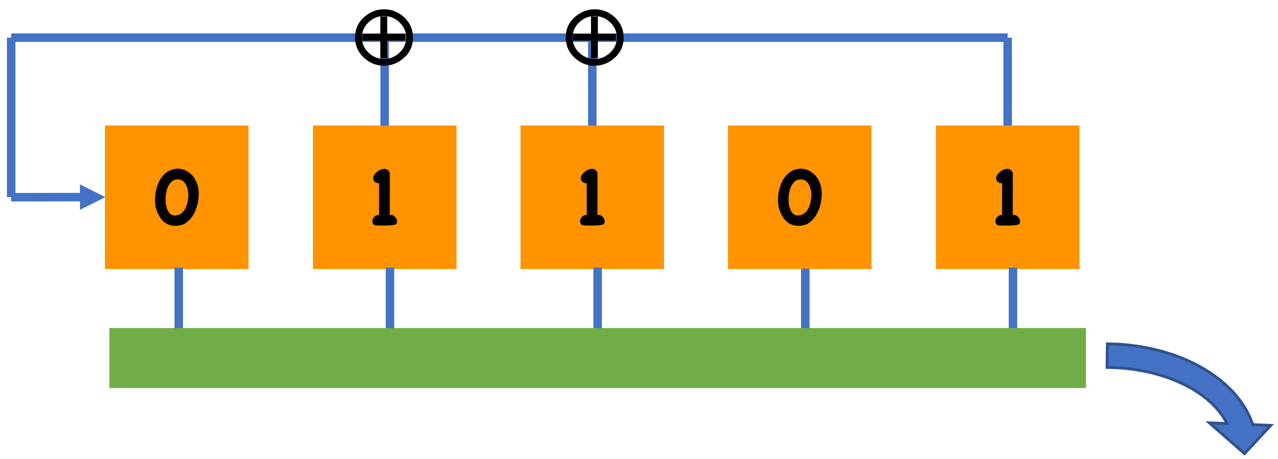
- Each output bit is a linear function of the initial state (that is, $\mathbf{G(s) = A \bullet s \pmod 2}$)

Any linear $\mathbf{G}$ <u>cannot</u> be a PRG

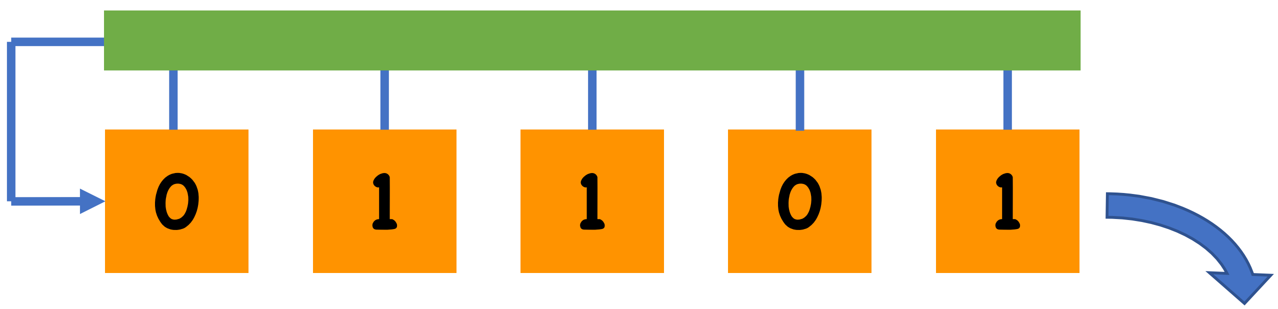
- Can check if $\mathbf{x}$ is in column-span of $\mathbf{A}$ using linear algebra

Introducing Non-linearity

Non-linearity in the output:



Non-linear feedback:



LFSR period

Period = number of bits before state repeats

After one period, output sequence repeats

Therefore, should have extremely long period

- Ideally almost 2^λ
- Possible to design LFSR's with period $2^\lambda-1$

Hardware vs Software

PRGs based on LFSR's are very fast in hardware

Unfortunately, not easily amenable to software

RC4

Fast software based PRG

Resisted attack for several years

No longer considered secure, but still widely used

RC4

State = permutation on **[256]** plus two integers
- Permutation stored as **256**-byte array **S**

Init(16-byte **k)**:
- For **i=0,…,255**
  S[i] = i
- **j = 0**
- **For i=0,…,255**
  j = j + S[i] + k[i mod 16] (mod 256)
  Swap **S[i]** and **S[j]**
- Output **(S,0,0)**

RC4

GetBits(S,i,j):

- **i++ (mod 256)**
- **j+= S[i] (mod 256)**
- Swap **S[i]** and **S[j]**
- **t = S[i] + S[j] (mod 256)**
- Output **(S,i,j), S[t]**

New state   Next output byte

Insecurity of RC4

Second byte of output is slightly biased towards 0

- **Pr[**second byte **= 0^8] ≈ 2/256**
- Should be **1/256**

Means RC4 is not secure according to our definition

-  outputs **1** iff second byte is equal to **0^8**
- Advantage: **≈ 1/256**

Not a serious attack in practice, but demonstrates some structural weakness

Insecurity of RC4

Possible to extend attack to actually recover the input **k** in some use cases

- The seed is set to **(IV, k)** for some initial value **IV**
- Encrypt messages as $\mathbf{RC4(IV,k)} \oplus \mathbf{m}$
- Also give **IV** to attacker
- Cannot show security assuming RC4 is a PRG

Can be used to completely break WEP encryption standard

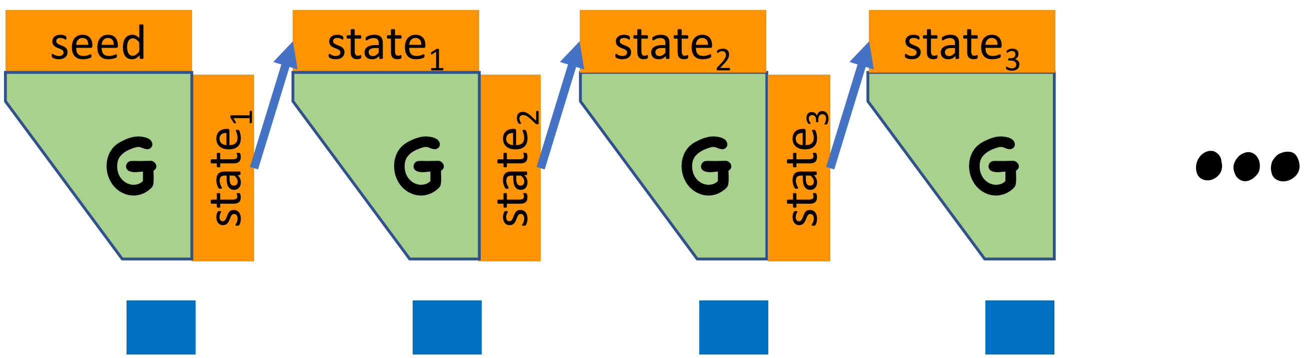
Length Extension for PRGs

Suppose I give you a PRG $\mathbf{G:\{0,1\}^{\lambda} \rightarrow \{0,1\}^{\lambda+1}}$

On it's own, not very useful: can only compress keys by 1 bit

But, we can use it to build PRGs with *arbitrarily-long* outputs!

Extending the Stretch of a PRG

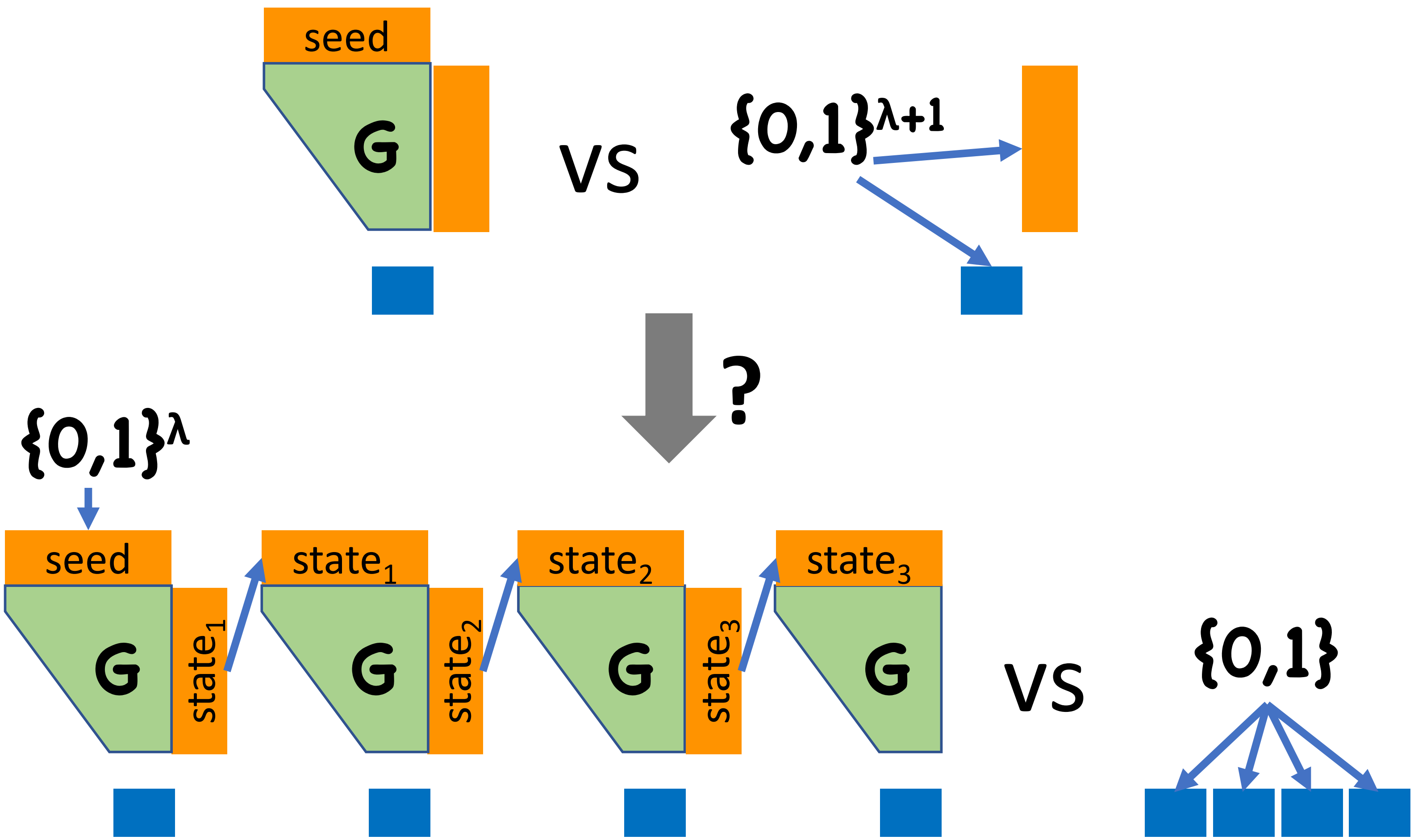


Security Proof

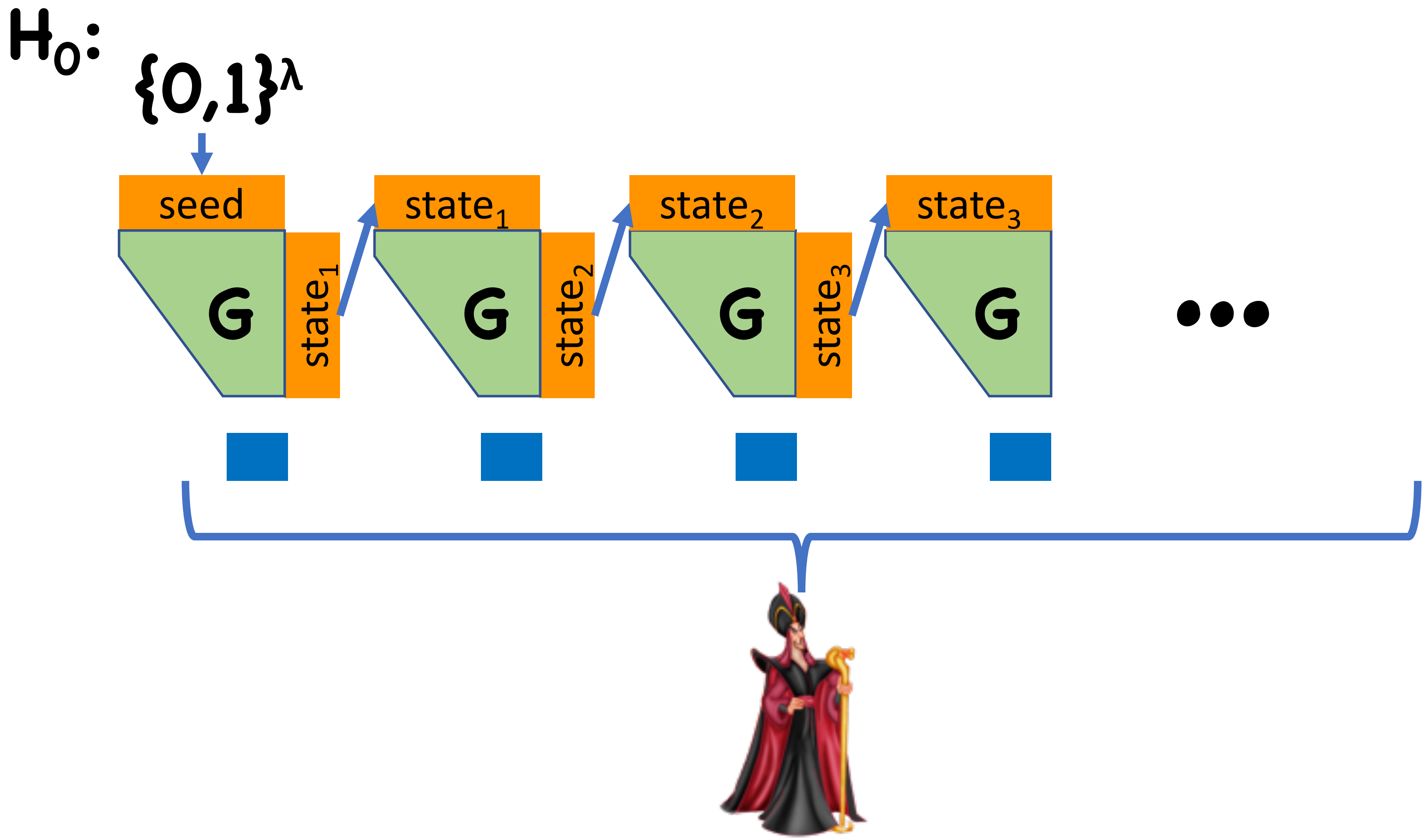
Assume towards contradiction  that breaks big PRG

Goal: build adversary  that breaks **G**

Problem?

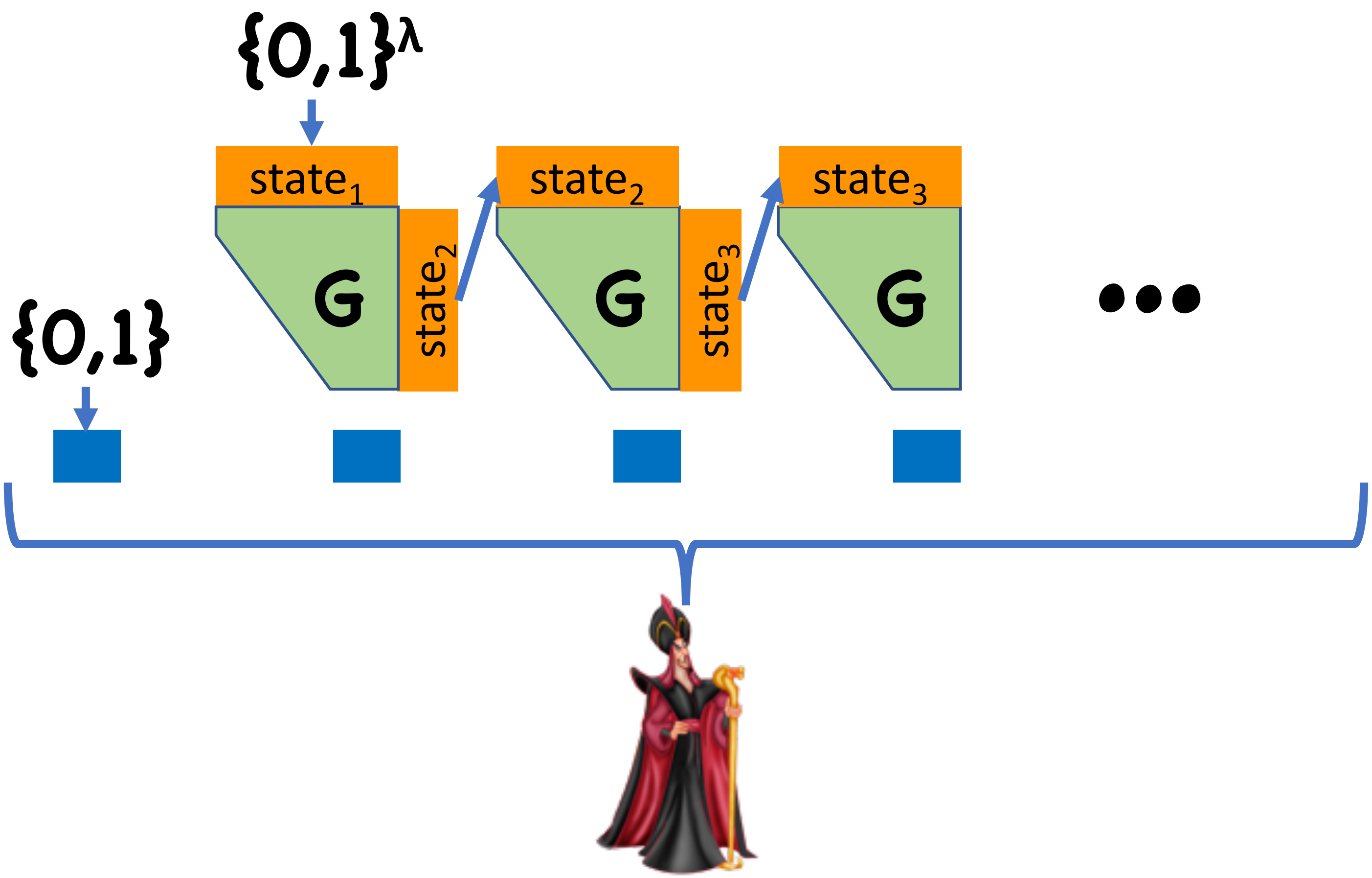

seed
G
vs
$\{0,1\}^{\lambda+1}$
?
$\{0,1\}^{\lambda}$
seed
$state_1$
$state_1$
$state_2$
$state_2$
$state_3$
$state_3$
G
G
G
G
vs
$\{0,1\}$

Proof by Hybrids

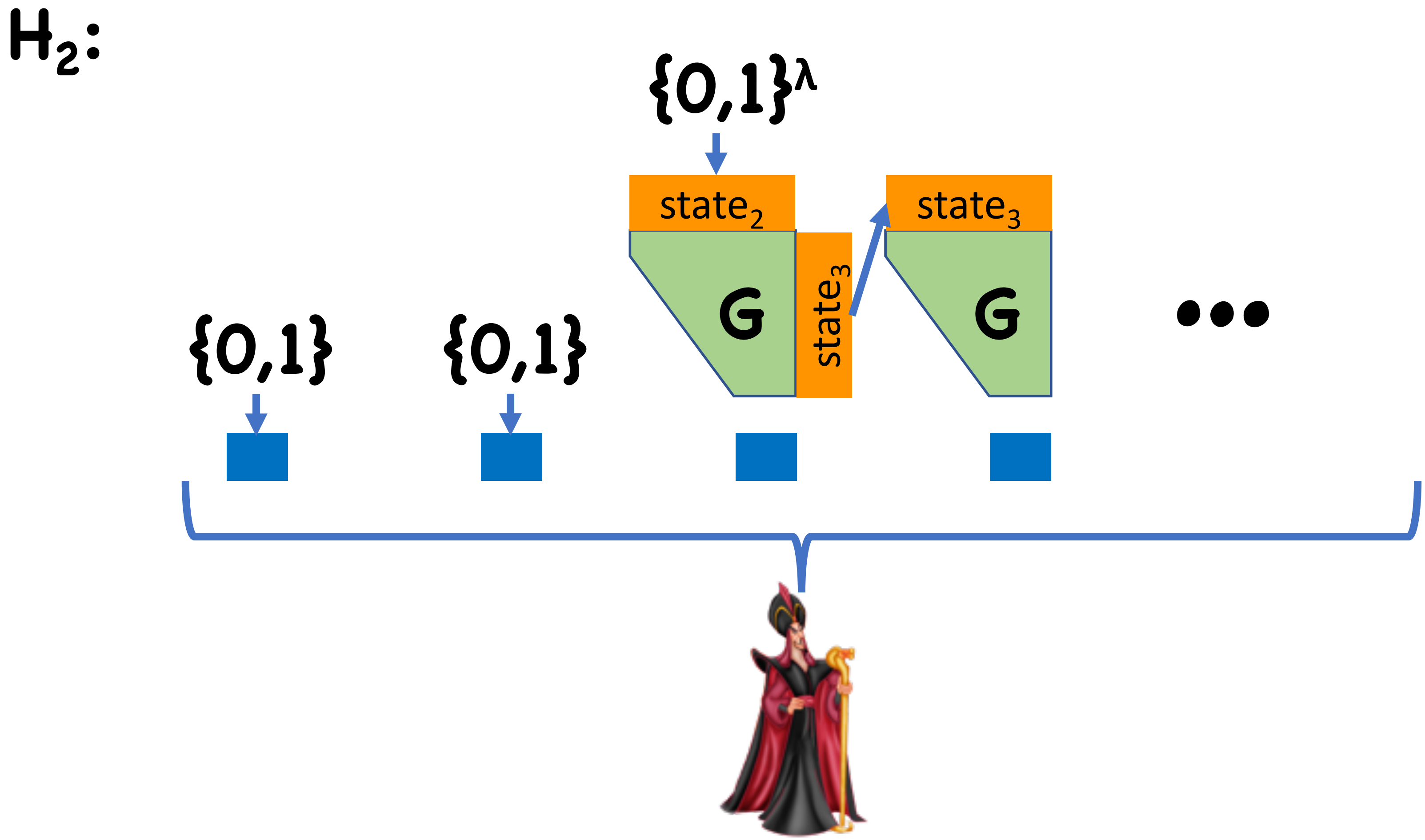


Security Proof

H_1:

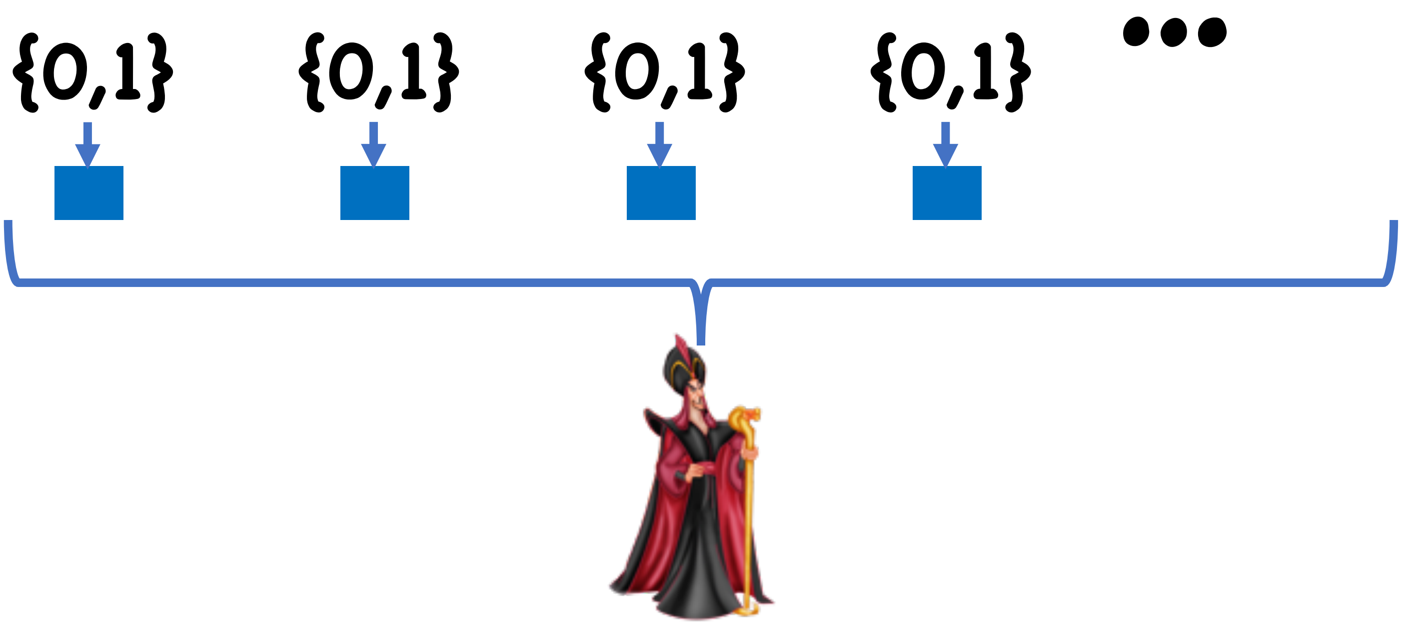


Security Proof



Security Proof

H_t:



Security Proof

H_0 corresponds to pseudorandom x

H_t corresponds to truly random x

Let $q_i = \Pr[$  $(x)=1 : x \leftarrow H_i]$

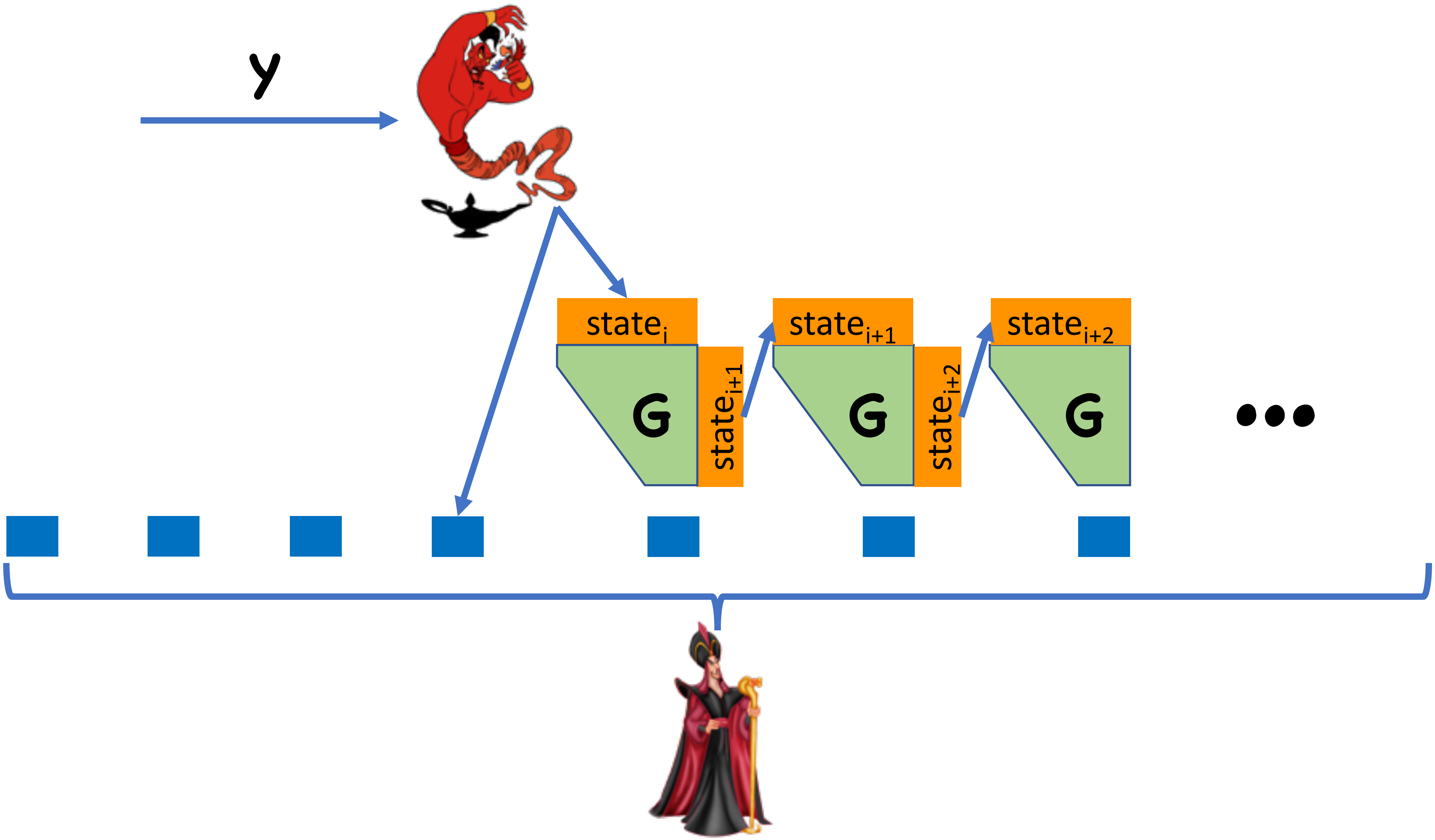
By assumption, $|q_t - q_0| > \varepsilon$

Triangle ineq:

$$|q_t - q_0| \le |q_1-q_0| + |q_2-q_1| + \ldots + |q_t-q_{t-1}|$$

$$\Rightarrow \exists i \text{ s.t. } |q_i - q_{i-1}| > \varepsilon/t$$

Security Proof



Security Proof

Analysis

- If $\mathbf{y} = \mathbf{G(s)}$, then  sees $\mathbf{H_{i-1}}$

  $\Rightarrow$ **Pr[**  **outputs 1] =** $\mathbf{q_{i-1}}$

  $\Rightarrow$ **Pr[**  **outputs 1] =** $\mathbf{q_{i-1}}$

- If $\mathbf{y}$ is random, then  sees $\mathbf{H_i}$

  $\Rightarrow$ **Pr[**  **outputs 1] =** $\mathbf{q_i}$

  $\Rightarrow$ **Pr[**  **outputs 1] =** $\mathbf{q_i}$

Summary So Far

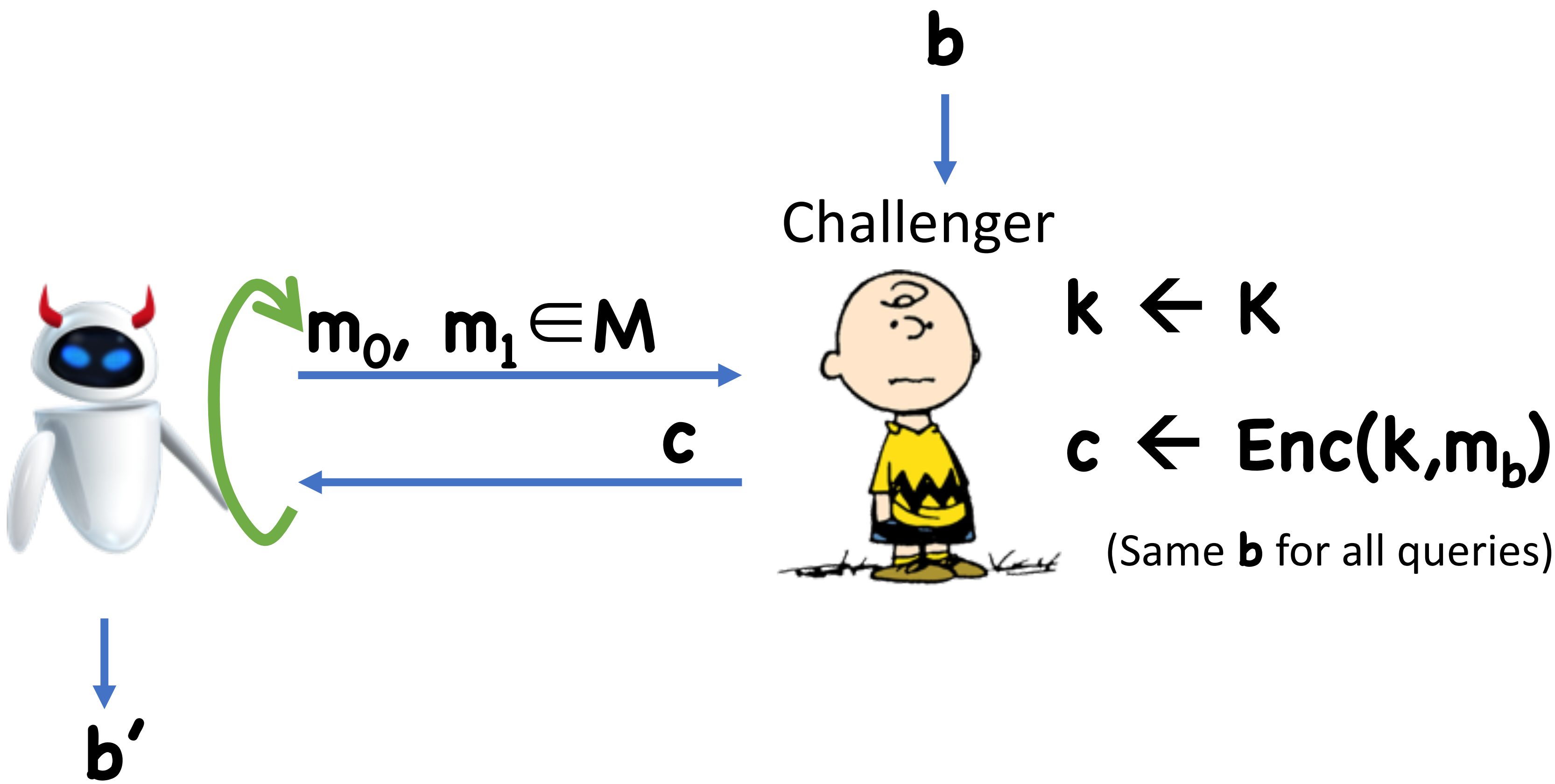
Stream ciphers = Encrytpion with PRG

- Secure encryption for arbitrary length, number of messages (though we did not completely prove it)

However, implementation difficulties due to having to maintaining state

Multiple Message Security

Left-or-Right Experiment



$LoR\text{-}Exp_b($  $, \lambda)$

LoR Security Definition

Definition: (Enc, Dec) has **Left-or-Right indistinguishability** if, for all  running in polynomial time, $\exists$ negligible **ε** such that:

| Pr[1←LoR-Exp$_0$(**, λ)]**

− Pr[1←LoR-Exp$_1$(**, λ)] | ≤ ε(λ)**

Alternate Notion: CPA Security

What if adversary can additionally learn encryptions of messages of her choice?

Examples:

- Midway Island, WWII:
  - US cryptographers discover Japan is planning attack on a location referred to as “AF”
  - Guess that “AF” meant Midway Island
  - To confirm suspicion, sent message in clear that Midway Island was low on supplies
  - Japan intercepted, and sent message referencing “AF”

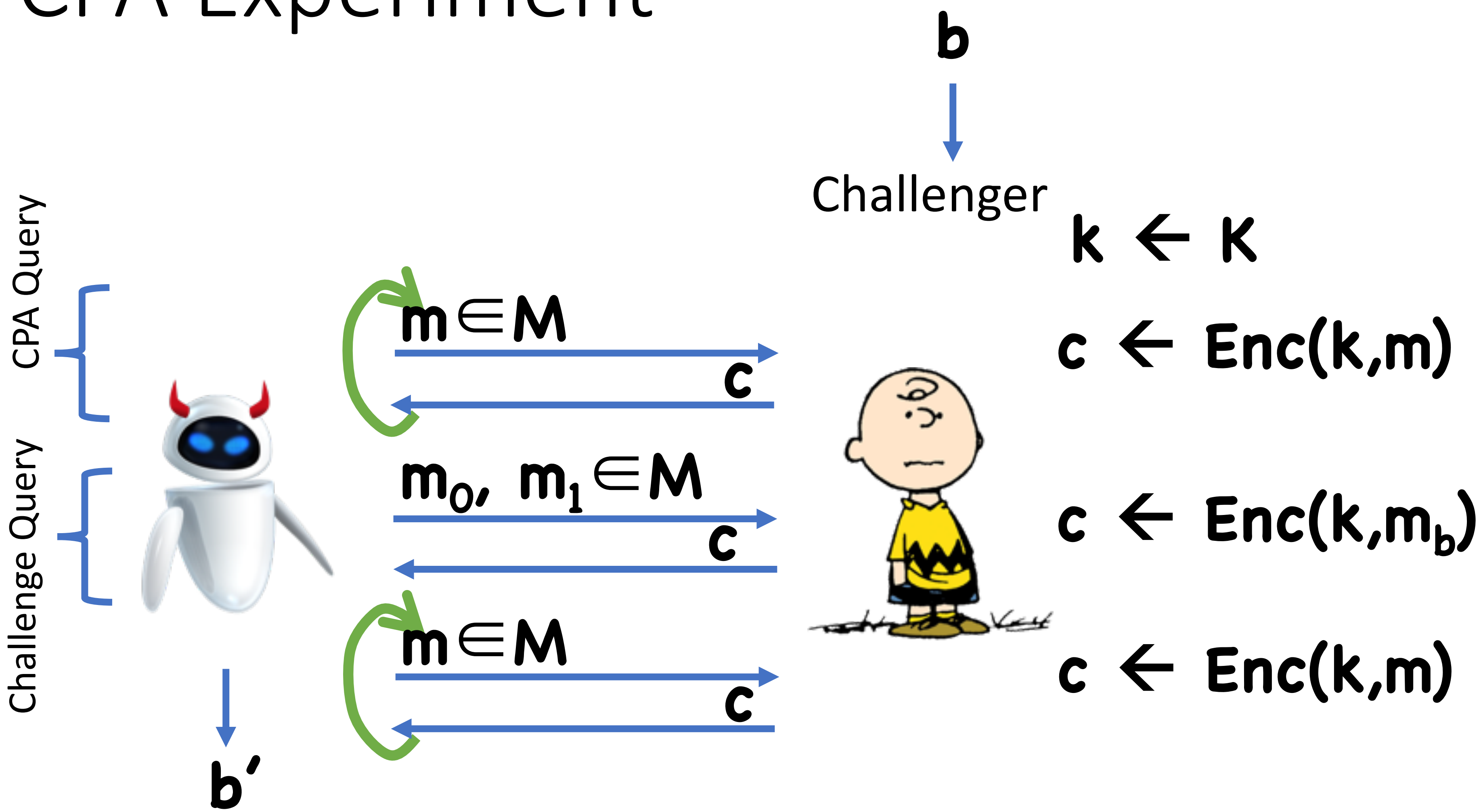
Alternate Notion: CPA Security

What if adversary can additionally learn encryptions of messages of her choice?

Examples:

- Mines, WWII:
  - Allies would lay mines at specific locations
  - Wait for Germans to discover mine
  - Germans would broadcast warning message about the mines, encrypted with Enigma
  - Would also send an “all clear” message once cleared

CPA Experiment



$\mathbf{CPA\text{-}Exp_b(}$  $\mathbf{)}$

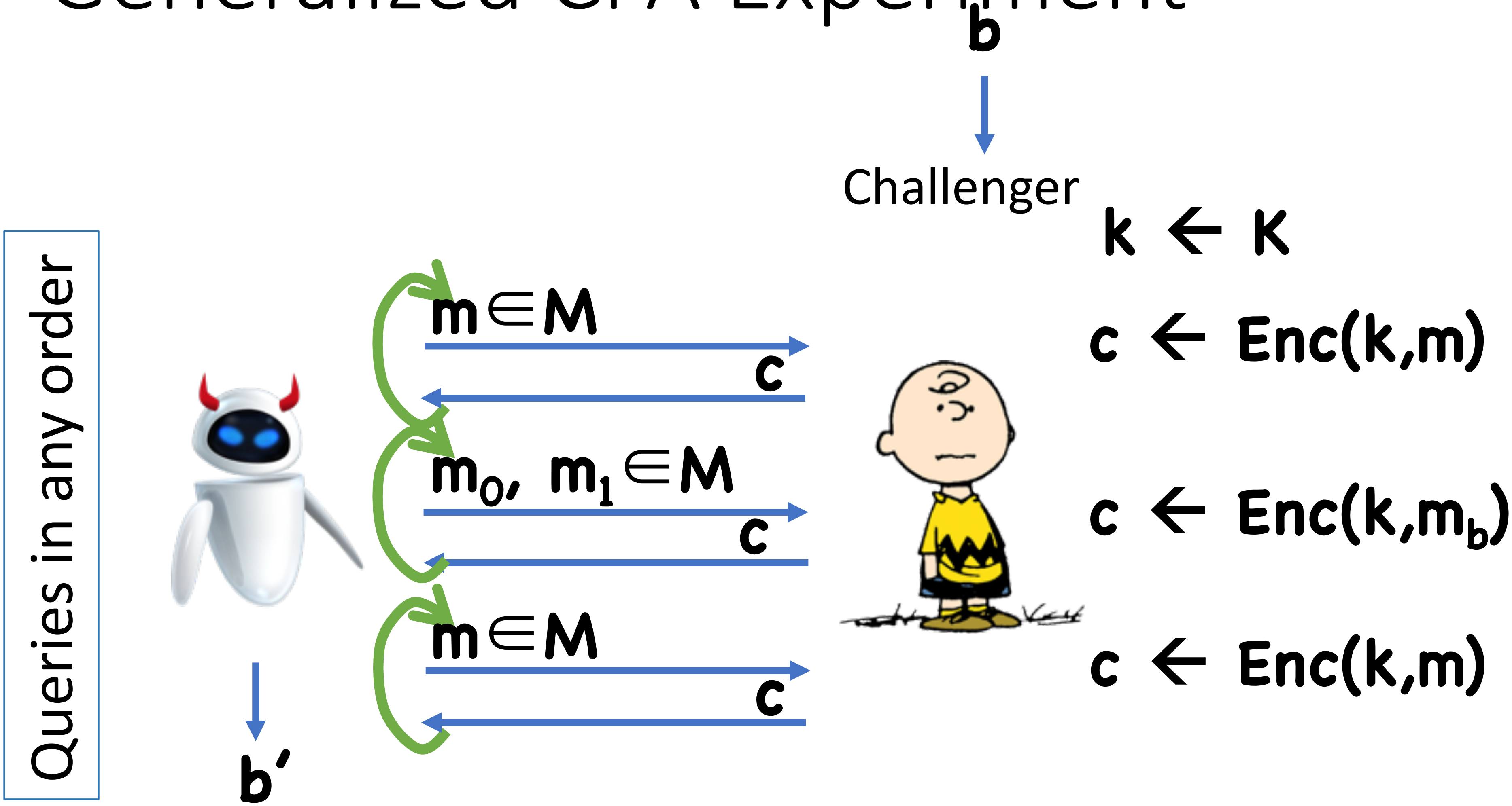
CPA Security Definition

Definition: (Enc, Dec) is **CPA Secure** if, for all  running in polynomial time, $\exists$ negligible **ε** such that:

| Pr[1←CPA-Exp$_0$( **, λ)]**

– Pr[1←CPA-Exp$_1$( **, λ)] | ≤ ε(λ)**

Generalized CPA Experiment



GCPA-Exp$_b$(, λ)

GCPA Security Definition

Definition: (Enc, Dec) is **Generalized CPA Secure** if, for all running in polynomial time, $\exists$ negligible $\boldsymbol{\varepsilon}$ such that:

$$\Big|\ \mathbf{Pr[1 \leftarrow GCPA\text{-}Exp_0(}$$  $$\mathbf{,\ \lambda)\ ]}$$

$$\mathbf{-\ Pr[1 \leftarrow GCPA\text{-}Exp_1(}$$  $$\mathbf{,\ \lambda)\ ]}\ \Big|\ \leq\ \boldsymbol{\varepsilon(\lambda)}$$

Equivalences

Theorem:

Left-or-Right indistinguishability

$\Updownarrow$

CPA-security

$\Updownarrow$

Generalized CPA-security

Proof

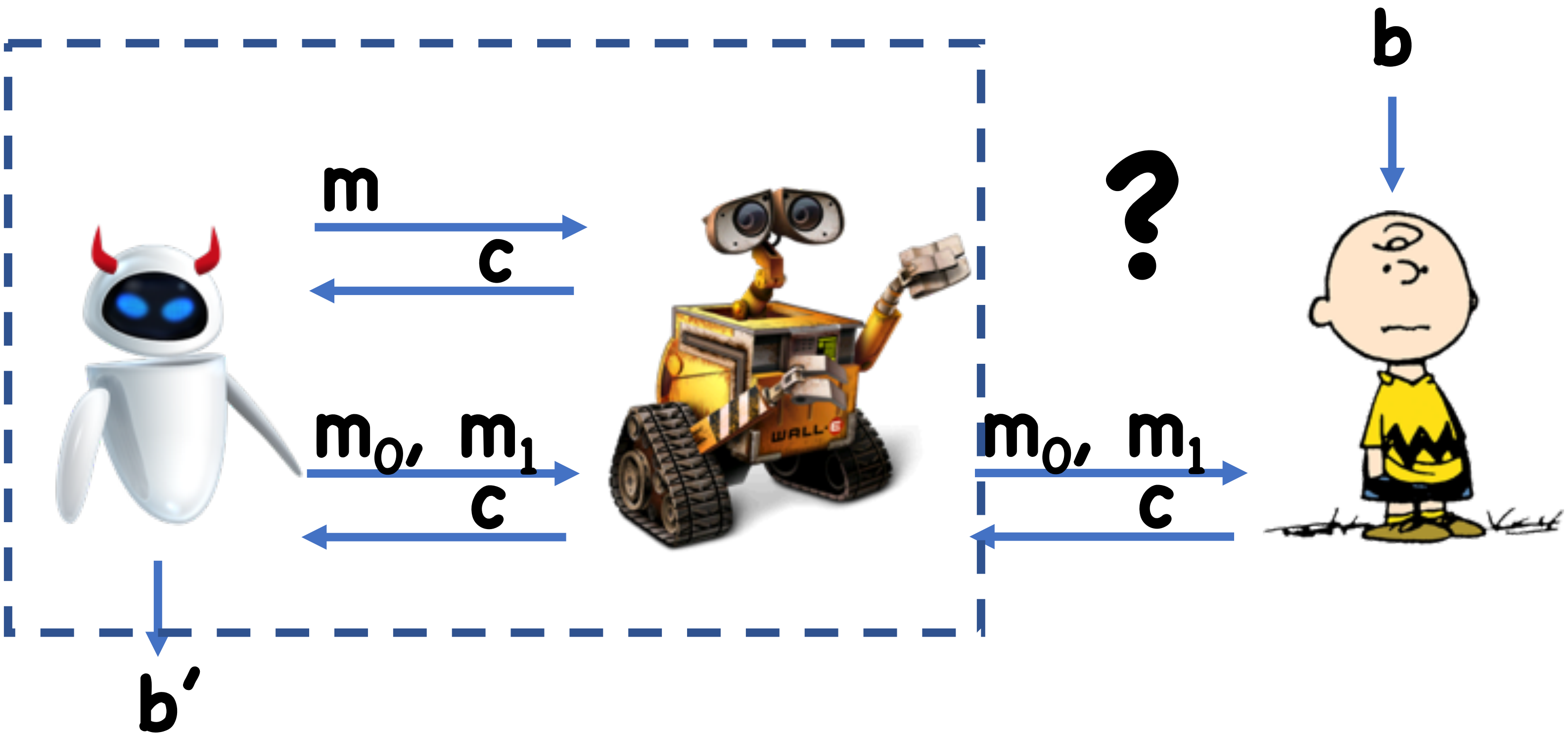
Generalized CPA-security → CPA-security

- Trivial: any adversary in the CPA experiment is also an adversary for the generalized CPA experiment that just doesn't take advantage of the ability to make multiple challenge/LoR queries

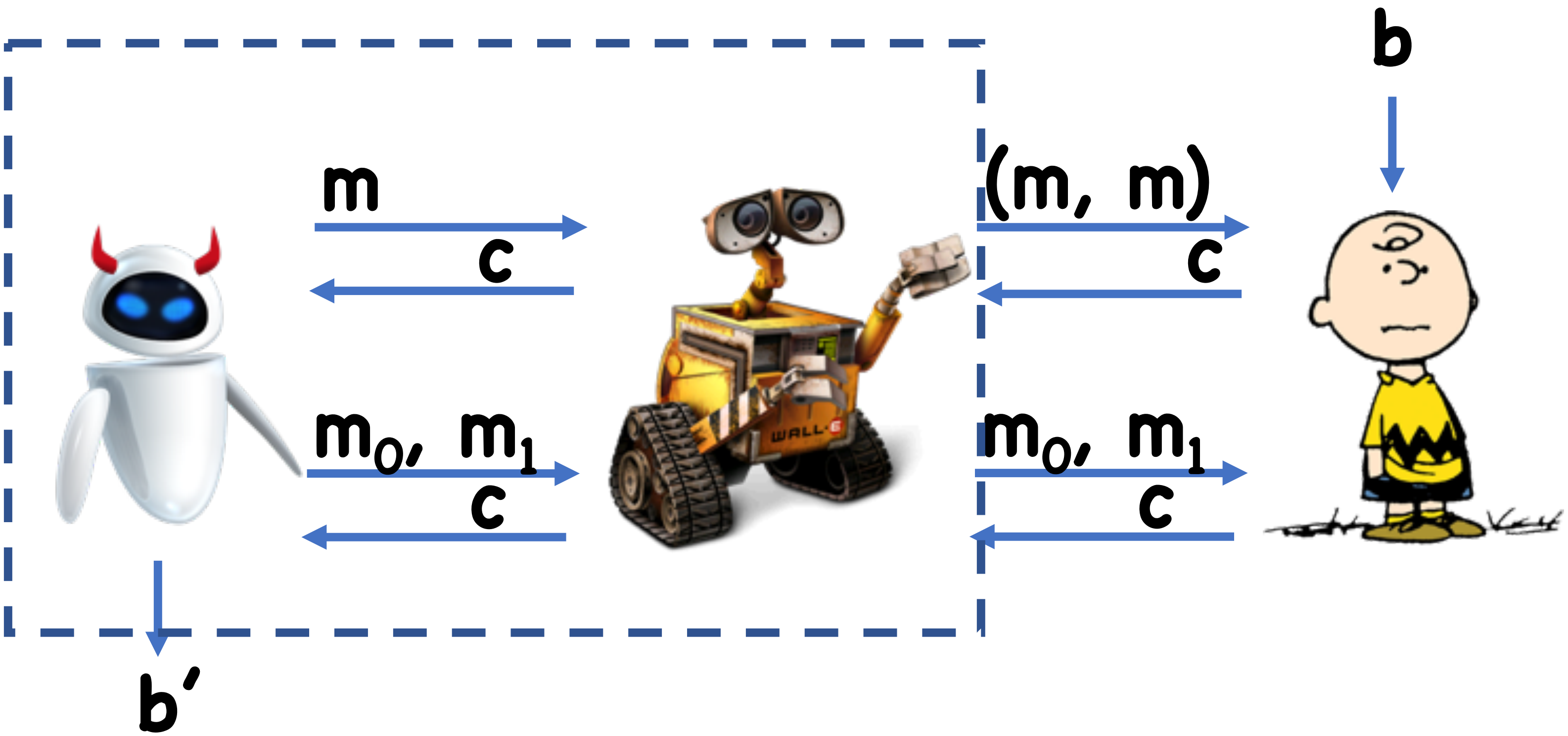
Proof

Left-or-Right → Generalized CPA

- Assume towards contradiction that we have an adversary  for the generalized CPA experiment
- Construct an adversary  that runs  as a subroutine, and breaks the Left-or-Right indistinguishability



$\Pr[1 \leftarrow \text{LoR-Exp}_b($  $, \lambda)] = \Pr[1 \leftarrow \text{GCPA-Exp}_b($  $, \lambda)\ ]$



Pr[1←LoR-Exp$_b$(, λ)] = Pr[1←GCPA-Exp$_b$(, λ)]

Proof

Left-or-Right → Generalized CPA

$\Big|$ Pr[1←LoR-Exp$_0$(, λ)]

− Pr[1←LoR-Exp$_1$(, λ)] $\Big|$

= $\Big|$ Pr[1←GCPA-Exp$_0$(, λ)]

− Pr[1←GCPA-Exp$_1$(, λ)] $\Big|$ = ε

Proof

(regular) CPA → Left-or-Right

- Assume towards contradiction that we have an adversary  for the **LoR Indistinguishability**

- Hybrids!

Hybrid **i**:

$m_0, m_1 \rightarrow$

$\leftarrow c$

$\downarrow$

b'

$k \leftarrow K$

If at most **i** queries so far,

$c \leftarrow Enc(k,m_0)$

If more than **i** queries so far,

$c \leftarrow Enc(k,m_1)$

Proof

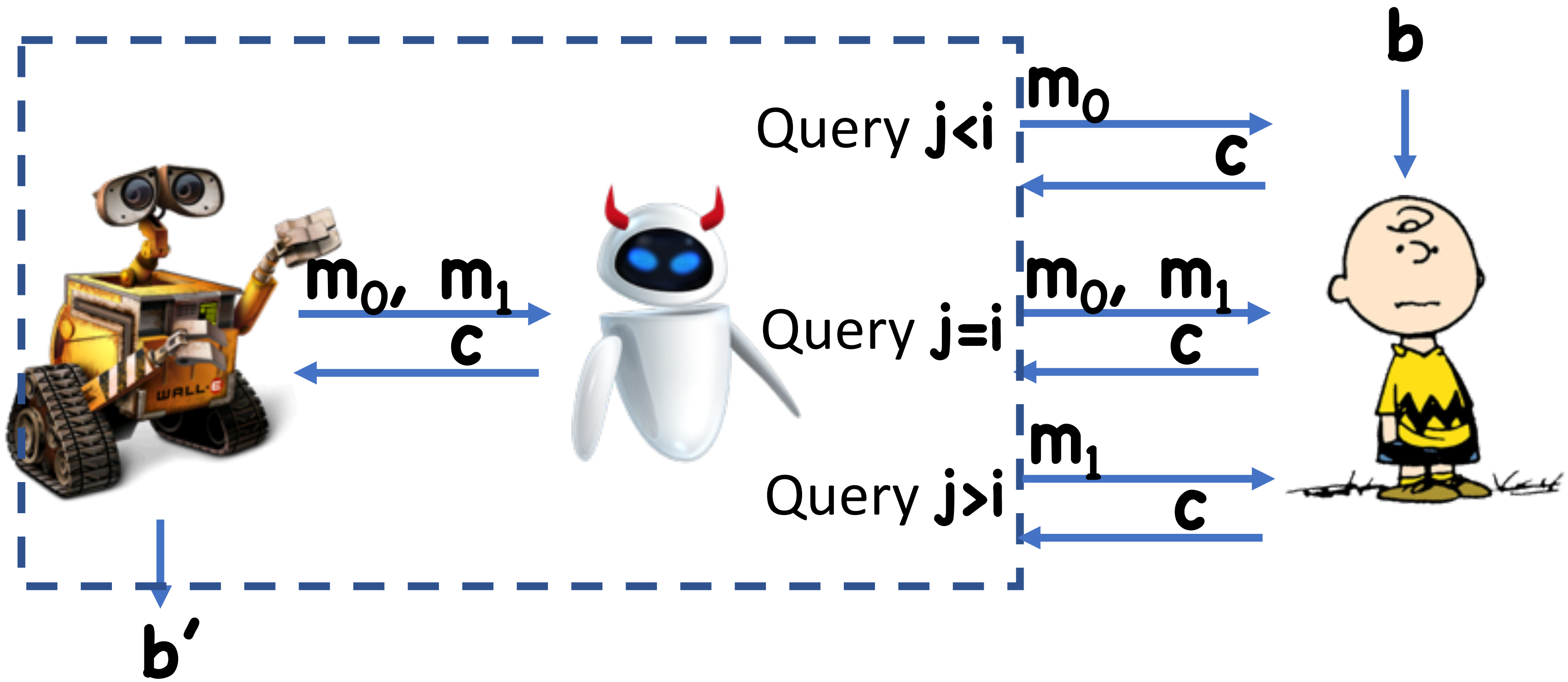
(regular) CPA → Left-or-Right

- Hybrid **0** is identical to **LoR-Exp**$_1$(, λ)

- Hybrid **q** is identical to **LoR-Exp**$_0$(, λ)

- We know that  distinguishes Hybrid **q** and Hybrid **0** with advantage ε

  $\Rightarrow$ $\exists$ **i** s.t.  distinguishes Hybrid **i** and Hybrid **i–1** with advantage ε/q



Pr[1←CPA-Exp$_b$(**, λ)] = Pr[1←**  in Hybrid **i-b]**

Proof

(regular) CPA → Left-or-Right

$\Big|$ $\Pr[1 \leftarrow \text{CPA-Exp}_0($  $, \lambda)\ ]$

$- \Pr[1 \leftarrow \text{CPA-Exp}_1($  $, \lambda)\ ]$ $\Big|$

$= \Big|$ $\Pr[1 \leftarrow$  in Hybrid $i]$

$- \Pr[1 \leftarrow$  in Hybrid $i-1]$ $\Big|$ $\geq \varepsilon / q$

Equivalences

Theorem:

Left-or-Right indistinguishability

$\Updownarrow$

CPA-security

$\Updownarrow$

Generalized CPA-security

Therefore, you can use whichever notion you like best

Next time: how to construct

Reminders

Mark's Office hours on 2/21: Moved to 11am

HW1 Due TODAY

HW2 Due Feb 27th

PR1 Due March 10th