

# Cryptography CS 555

## Lecture 8



### Semantic Security of Symmetric Ciphers

# What Properties Should A Block Cipher Has?

- Given a number of plaintext/ciphertext pairs,
  - should be difficult to recover the key
  - should be difficult to know how any other ciphertext block is decrypted
  - these should be true even if the pairs are obtained by choosing the plaintext blocks, or choosing ciphertext blocks
- Satisfied perfectly by random permutations
- Thus, a block cipher should be computationally indistinguishable from random permutations

# Symmetric Encryption Schemes

- A block cipher operates on one block
- An encryption scheme encrypts much longer messages
- Randomized vs. deterministic schemes
  - CBC is randomized
- Stateful vs. stateless schemes
  - CTR is stateful

# What Does Security Mean?

- What does insecurity mean?
  - from a few ciphertexts, can recover the encryption key
  - from a few ciphertexts, can recover the plaintext of some ciphertexts
  - from a few ciphertexts, can recover partial information of some ciphertexts

# What Does Security Mean?

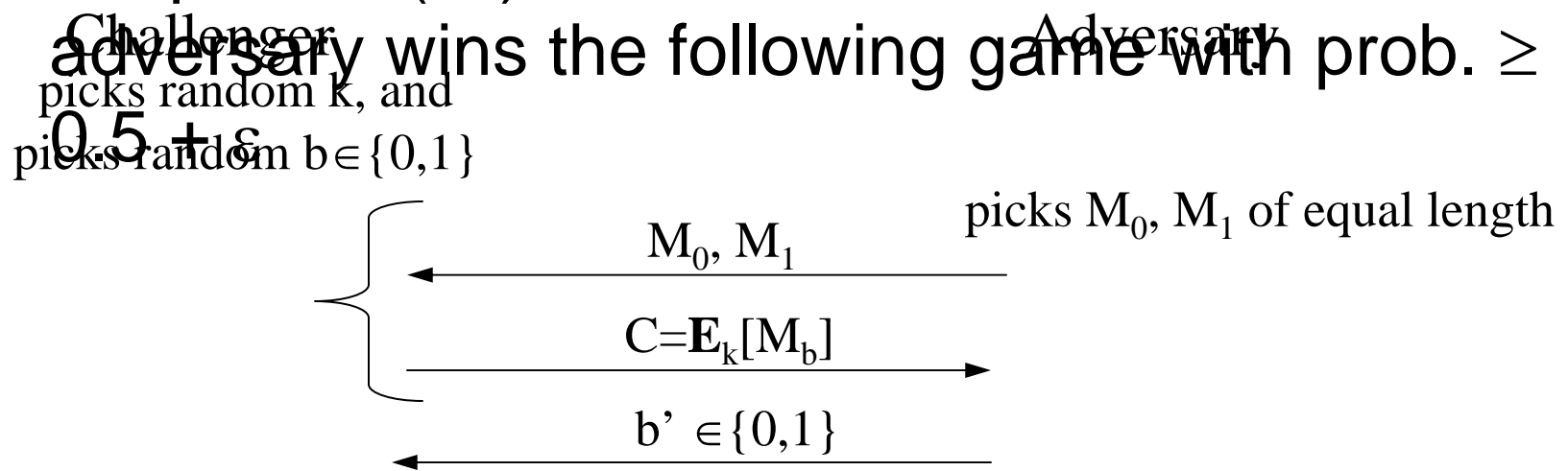
- Perfect secrecy
  - Given ciphertexts, cannot learn anything (other than the length of the message) about the plaintext
  - not very useful as requires long keys
- Approximate perfect secrecy?
  - with limited computing resources, it is extremely unlikely one can learn anything (other than the length) about the plaintexts from the ciphertexts
- How to formalize this?

# Towards Semantic Security

- Suppose that we give an adversary a ciphertext and tell the adversary that this ciphertext is from one of two possible plaintexts, the adversary should not be able to tell that one plaintext is more likely than the other.
- To give the adversary more control, we let the adversary choose the two possible plaintexts.
  - maybe the adversary can more easily distinguish some pairs of plaintexts

# IND-CPA

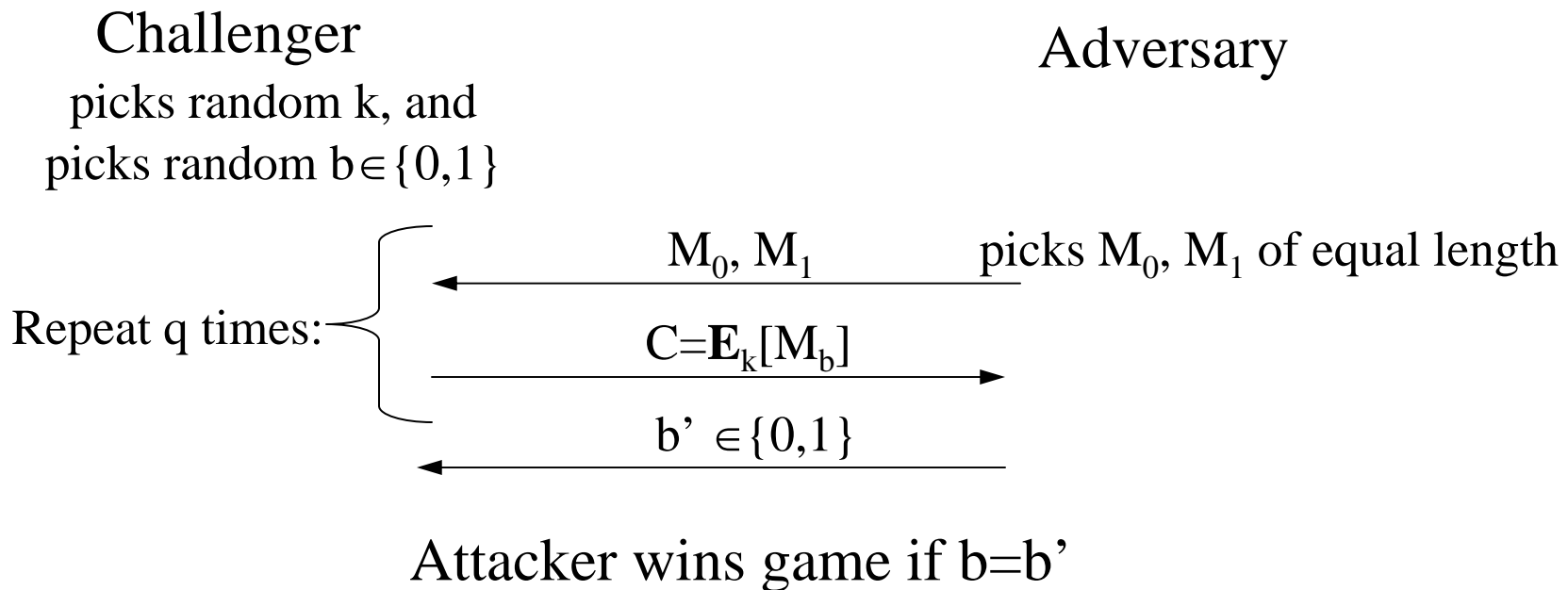
- a.k.a Semantic Security, or left-or-right indistinguishability under a chosen-plaintext attack
- A cipher is  $(t, \epsilon)$  IND-CPA secure if no  $t$ -time adversary wins the following game with prob.  $\geq 0.5 + \epsilon$



Attacker wins game if  $b=b'$

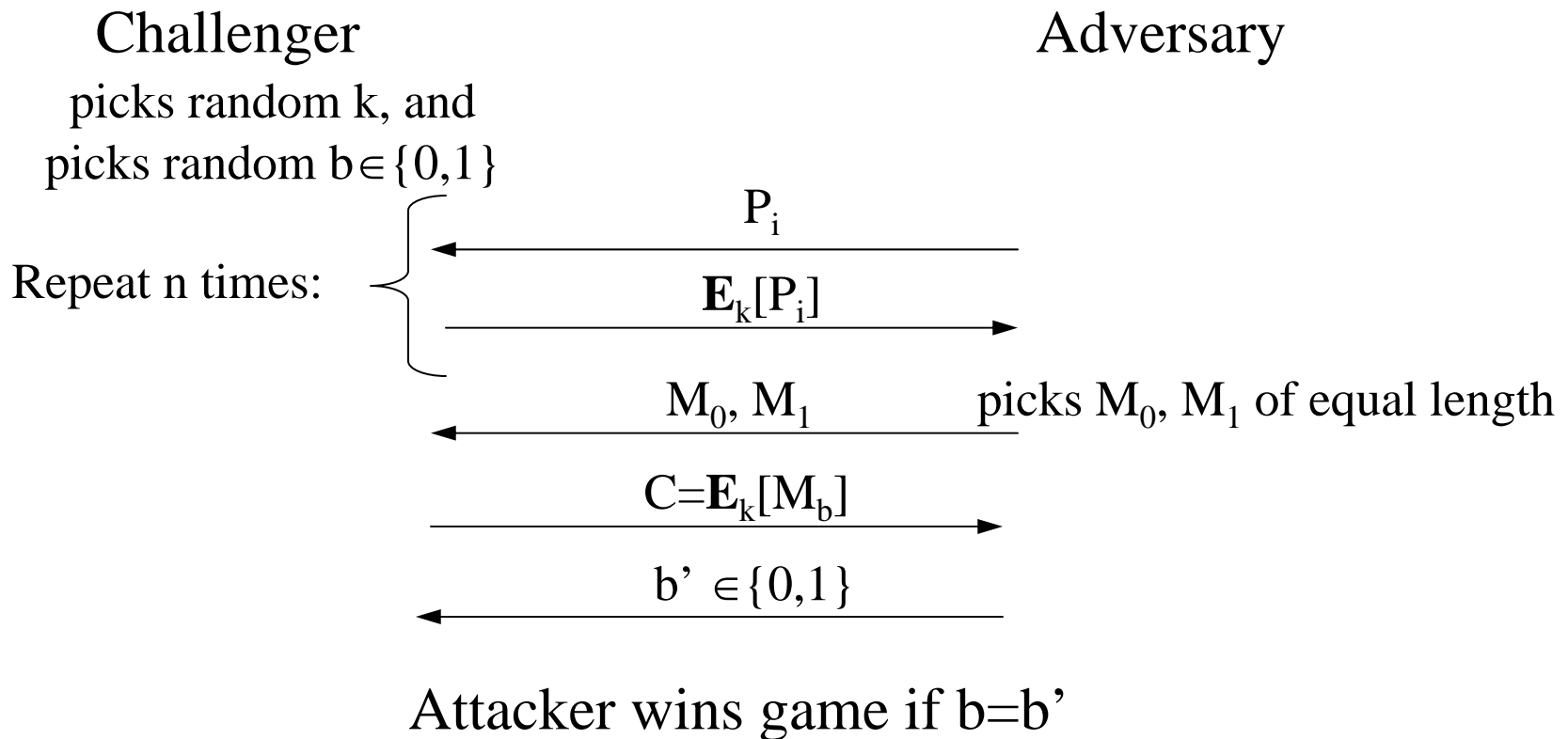
# A Variant of IND-CPA Security

- Allows the adversary to choose a sequence of messages to be encrypted.
  - relevant when the encryption scheme is stateful



# Another Variant of IND-CPA

- Add a training phase



# Why IND-CPA (Semantic) Security?

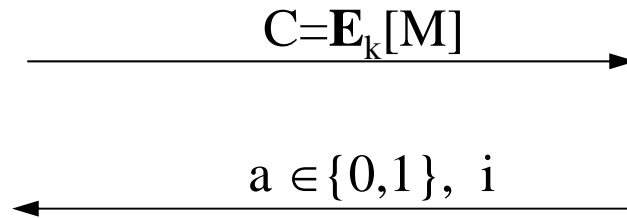
- Introduce another notion of security
- A cipher is  $(t, \epsilon)$  bit secure if no  $t$ -time adversary wins the following game with prob.  $\geq 0.5 + \epsilon$

Challenger

1. picks random  $k$
2. picks random  $M$

Adversary

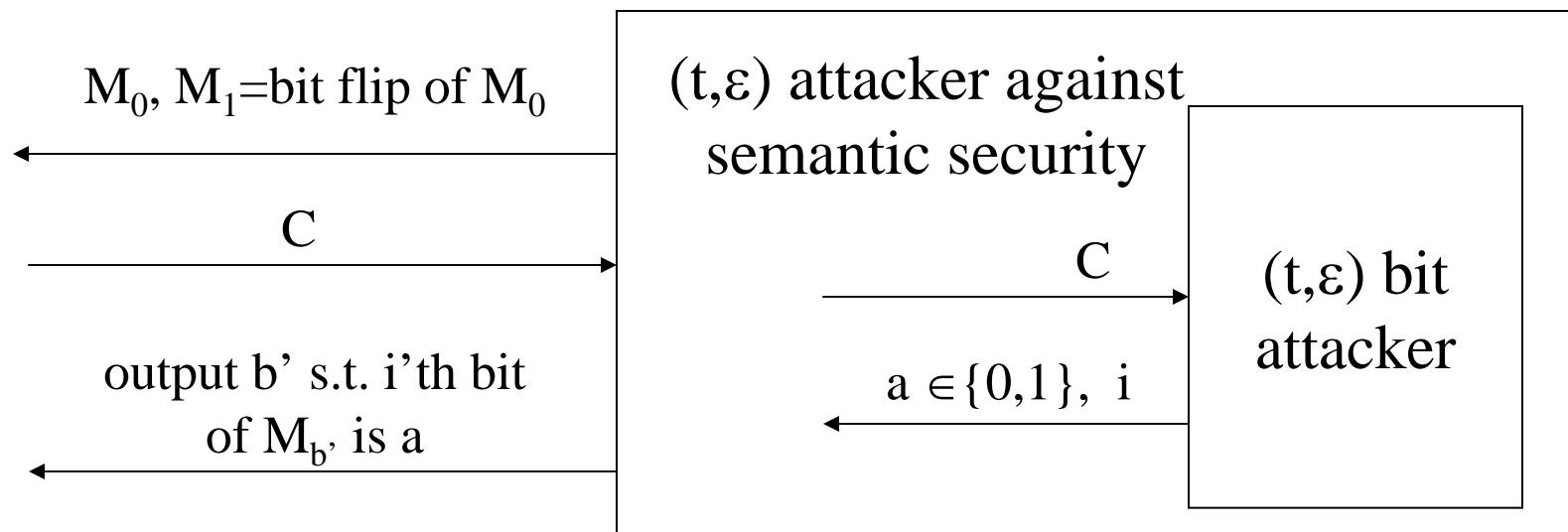
3. Pick  $i$



Attacker wins game if  $a=i$ 'th bit of  $M$

# Justification for semantic security

- Any cipher that is  $(t, \epsilon)$  semantically secure against eavesdroppers is also  $(t, \epsilon)$  bit secure
- Proof. Given a  $(t, \epsilon)$  attacker against bit security, build a  $(t, \epsilon)$  attacker against semantic security.



# More Justification for Semantic Security

- Semantic security implies that plaintext recovery attack is difficult
  - how to show it?

# ECB is not semantically secure

- **Claim:** There exists fast attacker that wins the semantic security game with prob. close to 1 (advantage close to 0.5)
- **Proof:** the attacker sends  $M_0$ ="hello hello " and  $M_1$ ="hello world ", then checks whether the two blocks in the ciphertext are the same or not.
- **Fact:** Any deterministic, stateless scheme is insecure

# CBC With Predictable IV is Insecure

- Recall that in CBC,  $C_0 = \mathbf{IV}$ ,  $C_1 = \mathbf{E}_k (M_1 \oplus C_0)$ ,
- If two messages are encrypted using CBC with their IV's related, and first block of the ciphertexts are the same, then one knows that the two messages are related in a particular way

# Security of Block Cipher Encryption Modes

- CBC (with random IV's), OFB (with random IV's), and CTR can be shown to be semantically secure, assuming that the underlying block cipher is a pseudo-random permutation.

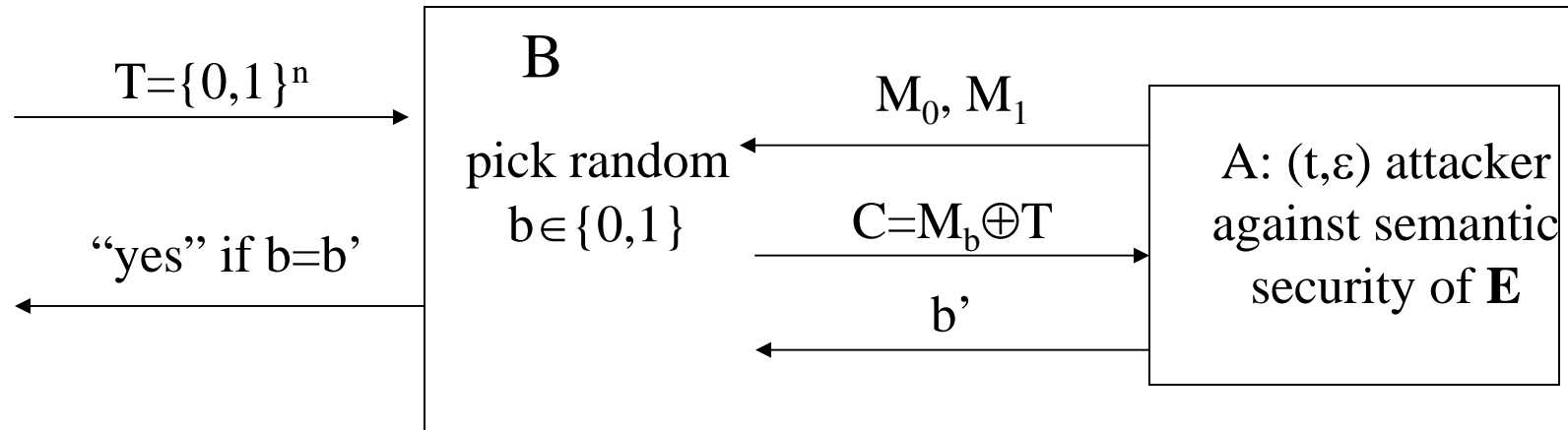
# PRNG

- Definition: a deterministic function  $G: \{0,1\}^s \rightarrow \{0,1\}^n$  ( $n \gg s$ ) is a  $(t, \epsilon)$ -PRNG if
  - there is an “efficient” algorithm to compute  $G$
  - $\forall$   $t$ -time algorithm  $A$ , we have
  - $|\Pr[A(G(S))=\text{“yes”}] - \Pr[A(R)=\text{“yes”}]| \leq \epsilon$ ,  
where  $S \in \{0,1\}^s$  is a random seed  
and  $R \in \{0,1\}^n$  is a length- $n$  random string
- E.g., RC4 with 128-bit key (seed) and  $2^{20}$  bytes of output is believed to be a  $(t, \epsilon)$ -PRNG for  $t=2^{80}$ ,  $\epsilon=1/2^{40}$

# A proof of semantic security

- Theorem: Suppose  $G: \{0,1\}^s \rightarrow \{0,1\}^n$  is a  $(t,\varepsilon)$  PRNG, then  $\mathbf{E}_k[M]=M\oplus G(k)$  is  $(t,\varepsilon)$  semantically secure.
- Proof: Contra-positive.
  - Suppose  $A$   $(t,\varepsilon)$ -breaks the semantic security of  $\mathbf{E}_k$ , build  $B$  that  $(t,\varepsilon)$ -breaks the PRNG security

# A proof of semantic security



- Claim: when  $T = G(S)$ , then  $\Pr[b = b'] > 0.5 + \epsilon$ , when  $T$  is random,  $\Pr[b = b'] = 1/2$ .
- Thus,  $|\Pr[A(G(S)) = \text{"yes"}] - \Pr[A(R) = \text{"yes"}]| > \epsilon$ .

# Next Lecture...

- AES & other block ciphers
- Recommended readings:
  - Stinson 3.6