Understanding the definition:

Can we learn the least significant bit of a message given only the ciphertext (assuming a semantically-secure cipher)

No! Suppose we could. Then, adversary can choose two messages m_0, m_1 that differ in their least significant bit and distinguish with probability 1.

This generalizes to <u>any</u> efficiently-computable property of the two messages.

How does semantic security relate to perfect secrecy?

<u>Theorem</u>. If a cipher satisfies perfect secrecy, then it is semantically secure.

<u>Proof</u>. Perfect secrecy means that $\forall\, m_0, m_1 \in \mathcal{M}$, $c \in \mathcal{C}$:

$$\Pr[k \xleftarrow{R} \mathcal{K} : \text{Encrypt}(k, m_0) = c] = \Pr[k \xleftarrow{R} \mathcal{K} : \text{Encrypt}(k, m_1) = c]$$

Equivalently, the distributions

$$\underbrace{\{k \xleftarrow{R} \mathcal{K} : \text{Encrypt}(k, m_0)\}}_{D_0} \quad \text{and} \quad \underbrace{\{k \xleftarrow{R} \mathcal{K} : \text{Encrypt}(k, m_1)\}}_{D_1}$$

are <u>identical</u> ($D_0 \equiv D_1$). This means that the adversary's output b' is identically distributed in the two experiments, and so

$$\text{SSAdv}[\mathcal{A}, \Pi_{SE}] = |W_0 - W_1| = 0.$$

<u>Corollary</u>. The one-time pad is semantically secure.

encryption key (PRG seed)

$$c \leftarrow G(s) \oplus m$$
$$m \leftarrow G(s) \oplus c$$

seems straightforward, but takes some care to prove

<u>Theorem</u>. Let G be a secure PRG. Then, the resulting stream cipher constructed from G is semantically secure.

<u>Proof</u>. Consider the semantic security experiments:

Experiment 0: Adversary chooses m_0, m_1 and receives $c_0 = G(s) \oplus m_0$

Experiment 1: Adversary chooses m_0, m_1 and receives $c_1 = G(s) \oplus m_1$

} Want to show that adversary's output in these two experiments are indistinguishable

Let $W_0 = \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Experiment } 0]$

$W_1 = \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Experiment } 1]$

<u>Goal</u>: Show that if G is a secure PRG, then for all efficient adversaries $\mathcal{A}$, $|W_0 - W_1| = \text{negl}(\lambda)$.

<u>Idea</u>: If $G(s)$ is uniform random string (i.e., one-time pad), then $W_0 = W_1$. But $G(s)$ is like a one-time pad!

Define Experiment 0': Adversary chooses m_0, m_1 and receives $c_0 = t \oplus m_0$ where $t \xleftarrow{R} \{0,1\}^n$

Experiment 1': Adversary chooses m_0, m_1 and receives $c_1 = t \oplus m_1$ where $t \xleftarrow{R} \{0,1\}^n$

} called "hybrid experiments"

Define W_0', W_1' accordingly.

Now we can write

$$|W_0 - W_1| = |W_0 - W_0' + W_0' - W_1' + W_1' - W_1|$$
$$\leq |W_0 - W_0'| + \underbrace{|W_0' - W_1'|}_{W_0' = W_1' \text{ (for all adversaries } \mathcal{A})} + |W_1' - W_1| \quad \text{by triangle inequality}$$

since OTP satisfies perfect secrecy

Suffices to show that for all efficient adversaries, $|W_0 - W_0'| = \text{negl}(\lambda)$ and $|W_1 - W_1'| = \text{negl}(\lambda)$.

Show. If G is a secure PRG, then for all efficient A, $|W_0 - W_0'| = $ negl.

Common proof technique: prove the contrapositive.

Contrapositive: If A can distinguish Experiments 0 and 0', then G is not a secure PRG.

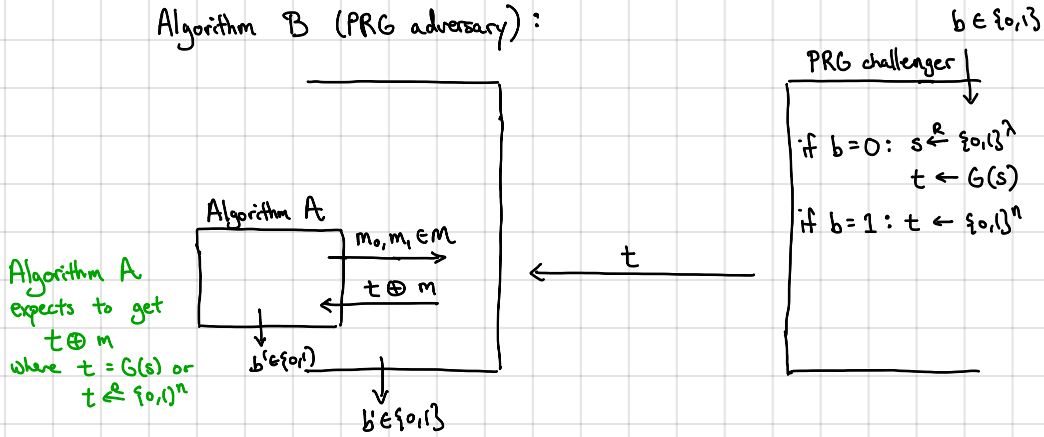
Suppose there exists efficient A that distinguishes Experiment 0 from 0'

$\Rightarrow$ We use A to construct efficient adversary B that breaks security of G.

↳ this step is a reduction

[we show how adversary (i.e., algorithm) for distinguishing Exp. 0 and 0' $\Rightarrow$ adversary for PRG]

Algorithm B (PRG adversary):



Running time of B = running time of A = efficient

Compute PRGAdv[B, G].

$\Pr[B \text{ outputs } 1 \text{ if } b=0] = W_0$ ← if $b=0$, then A gets $G(s) \oplus m$ which is precisely the behavior in Exp. 0

$\Pr[B \text{ outputs } 1 \text{ if } b=1] = W_0'$ ← if $b=1$, then A gets $t \oplus m$ which is precisely the behavior in Exp. 0'

$\Rightarrow$ PRGAdv[B, G] $= |W_0 - W_0'|$, which is non-negligible by assumption. This proves the contrapositive.

Important note: Security of above schemes shown assuming message space is $\{0,1\}^n$ (i.e., all messages are n-bits long)

In practice: We have variable-length messages. In this case, security guarantees indistinguishability from other messages of the same length, but length itself is leaked [inevitable if we want short ciphertexts]

↳ can be problematic — see traffic analysis attacks!

So far, we have shown that if we have a PRG, then we can encrypt messages efficiently (stream cipher).