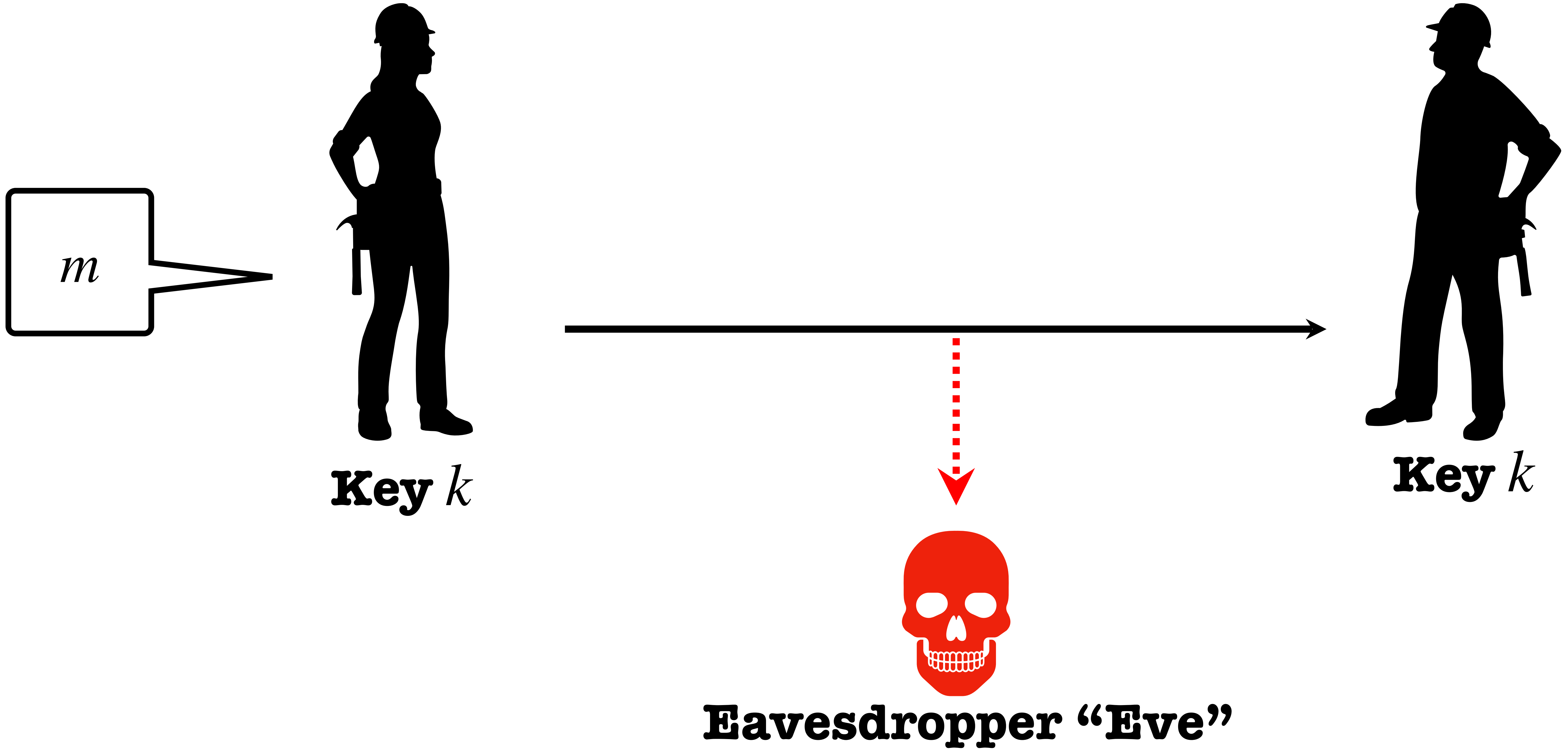
CIS 5560

Cryptography
Lecture 2

Announcements

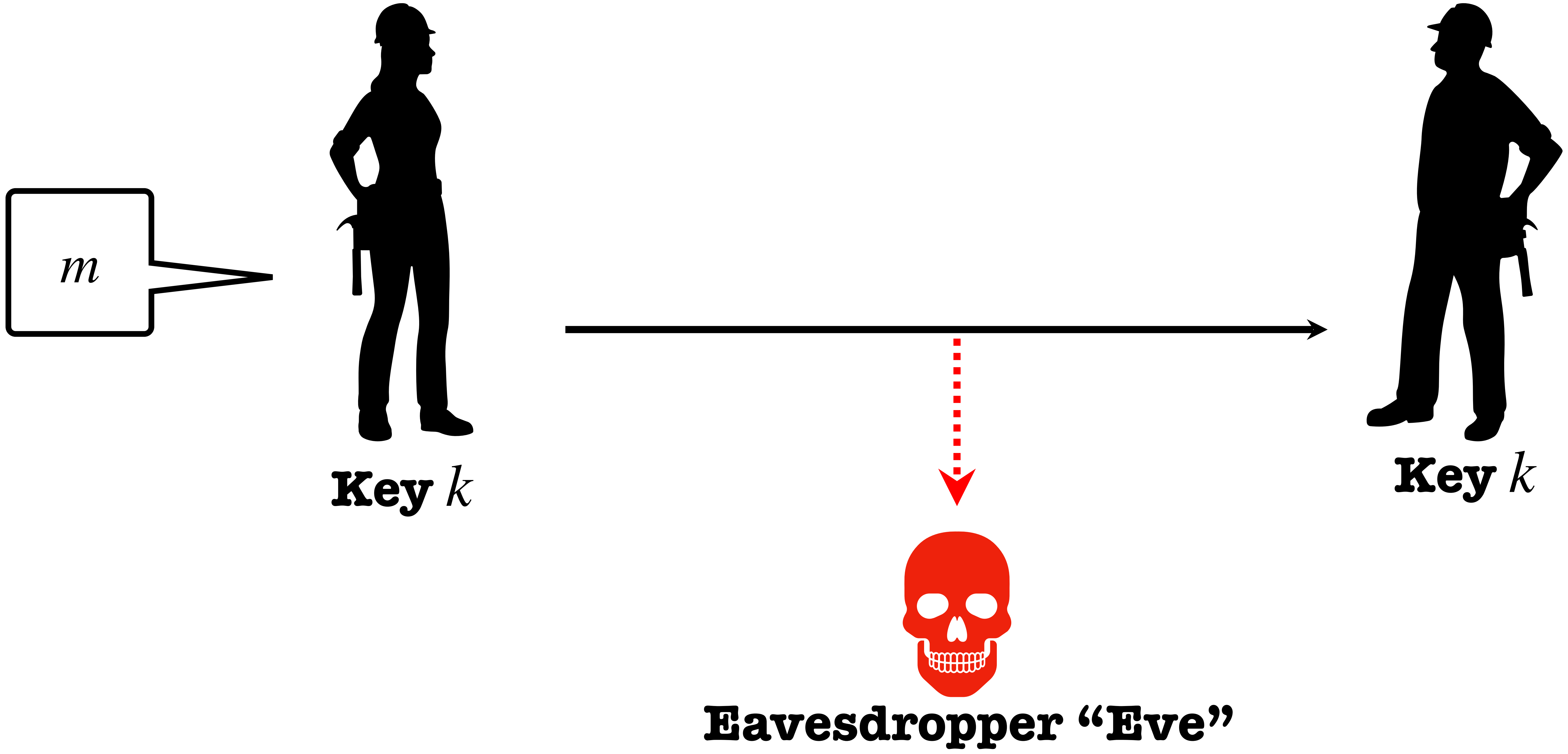
- **HW 1 will be released tomorrow Wed Jan 21**
  - **Due Friday Jan 30** at 5PM on Gradescope
  - Recap on probability and mathematical background
  - Get started ASAP and make use of office hours!
  - Will have homework “party” Wednesdays 4:30-6PM
- **For HW2** onwards, we will experiment with a new format for homework:
  - Instead of offline written submissions, in-person “homework-writing” sessions on Friday
- Course website is up: pratyushmishra.com/classes/cis-5560/s26!

Secure Communication



Alice wants to send a message m to Bob without revealing it to Eve.

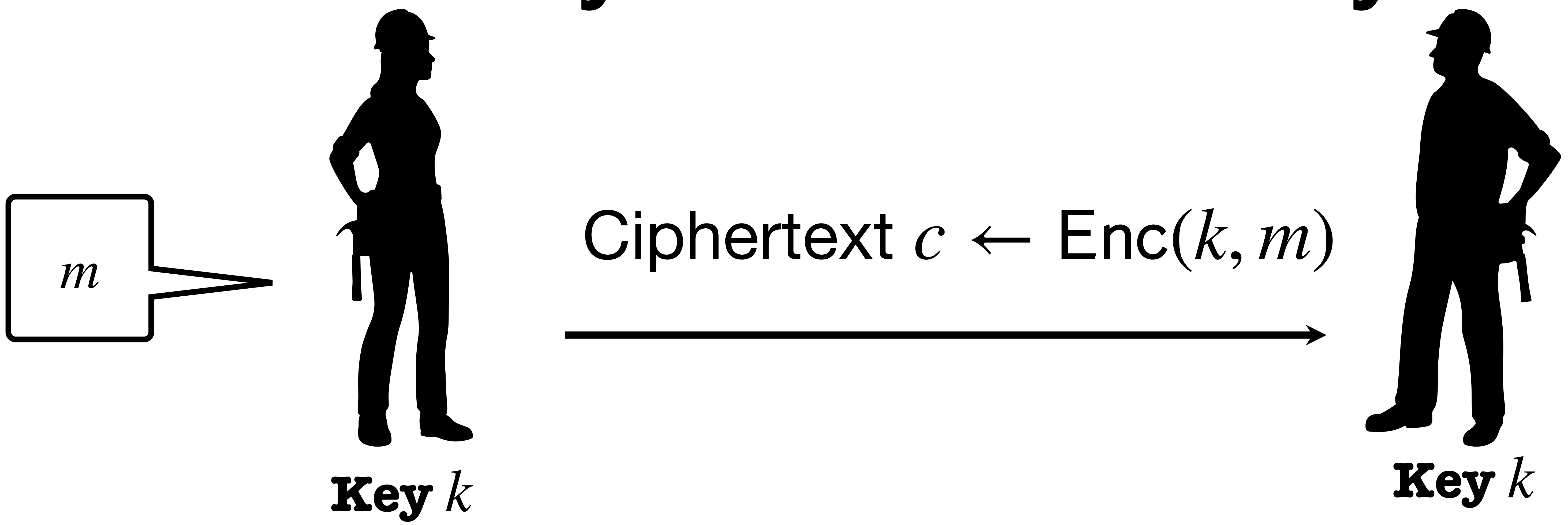
Secure Communication



Alice wants to send a message m to Bob without revealing it to Eve.

SETUP: Alice and Bob meet beforehand to agree on a secret key k.

Key notion: Symmetric-Key Encryption



<u>Three (possibly randomized) polynomial-time algorithms:</u>

Key Generation Algorithm: $\mathsf{Gen}(1^\lambda) \to k$

Has to be randomized (why?)

Encryption Algorithm: $\mathsf{Enc}(k, m) \to c$

Decryption Algorithm: $\mathsf{Dec}(k, c) \to m$

Property 1: Correctness

- $\forall k \in \mathscr{K}, \forall m \in \mathscr{M}$, $\mathsf{Dec}(k, \mathsf{Enc}(k, m)) = m$
- **Most basic property: if Bob gets incorrect answer, scheme is useless!**

Property 2: Security?

The Worst-case Adversary

An arbitrary computationally *unbounded* algorithm **EVE**.*

Knows Alice and Bob's algorithms Gen, Enc and Dec but does not know the key nor their internal randomness.

(*Kerckhoff's principle or Shannon's maxim*)

Can see the ciphertexts going through the channel
(but cannot modify them… we will come to that later)

Security Definition: What is she trying to learn?

What is a secure encryption scheme?

- Attacker's abilities: **CT only attack** (for now)
- Possible security requirements:
- attempt #1: **attacker cannot recover secret key**
  - $\text{Enc}(k, m) = m$ would be secure
- attempt #2: **attacker cannot recover all of plaintext**
  - $\text{Enc}(k, (m_1, m_2)) = \text{Enc}(k, m_1) \,||\, m_2$ would be secure
- Shannon's idea: **CT should reveal no "info" about PT**

Attempt 1: Caesar cipher

- Idea: shift each letter over by a specific amount N.
- Example: A → D, B → E, …, Z → C
  Encrypt "HELLO CLASS" → "KHOOR FODVV"

- Keyspace $\mathscr{K} = ?$
  - Answer: "shifts by $N \in \{0,\ldots,25\}$"
- Gen: Sample $k = N \leftarrow \{0,\ldots,25\}$
- $\text{Enc}(k, m)$: replace each character ch in m with $\text{ch} + N$
- $\text{Dec}(k, c)$: replace each character ch in c with $\text{ch} - N$

Attempt 1: Caesar cipher

- Question: Is this secure? Can adversary recover message?
- Answer: Yes!
  - Just iterate over 26 possible keys, and see which one decrypts!
- Example: “KHOOR FODVV”
  - Try with shift 1 → “LIPPS GPEWW”
  - Try with shift 2 → “IFMMP DMBTT”
  - Try with shift 3 → “HELLO CLASS”

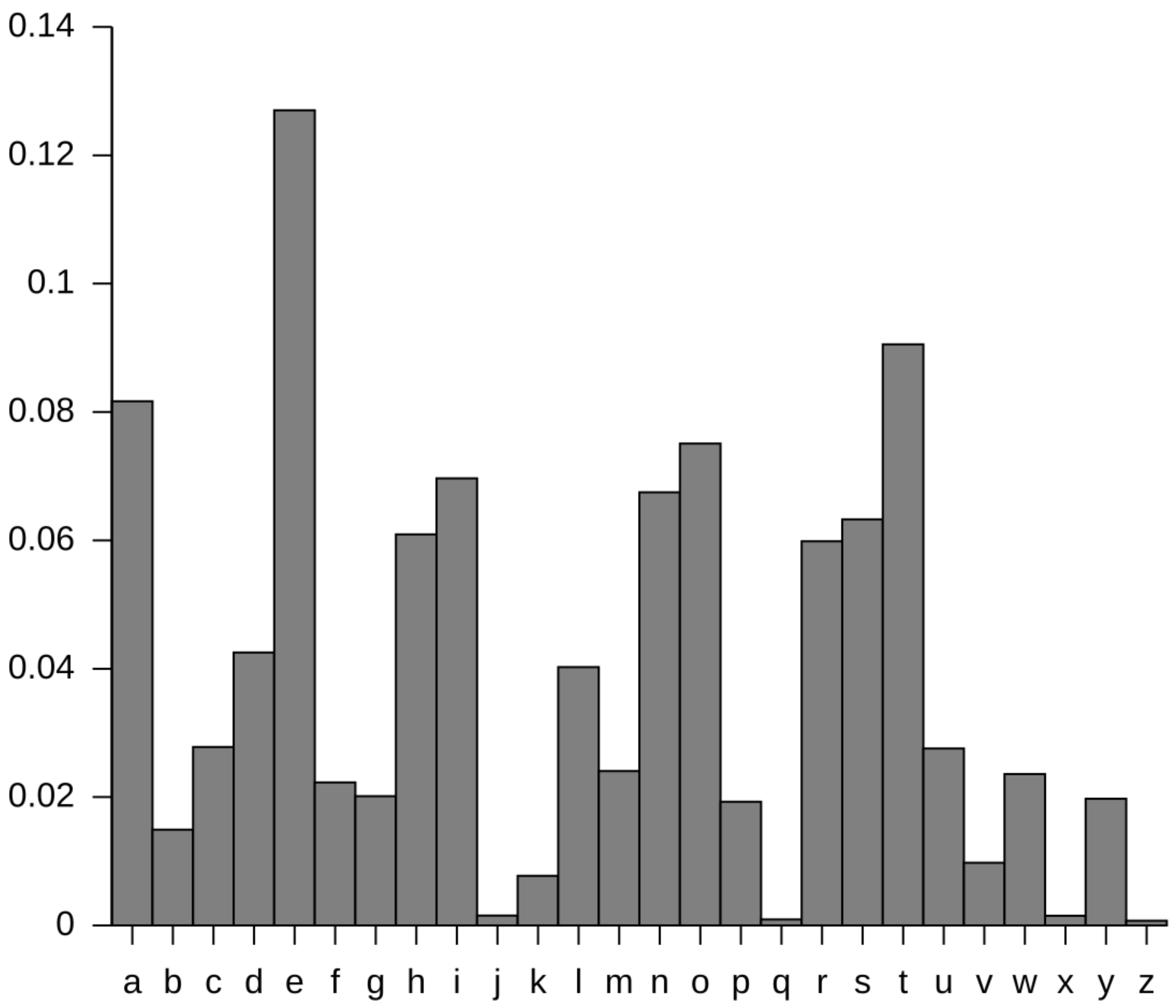
Attempt 2: Substitution cipher

- Idea: Caesar cipher maps letters to other letters in a simple way (shifts)
- Can we use an arbitrary mapping?
- Example: A → E, B → C, …, Z → D

- Keyspace $\mathcal{K} = ?$
  - Answer: “all permutations over $\{0,\ldots,25\}$”
- Gen: Sample a random permutation $k = \pi$
- Enc(k, m) : replace each character ch in m with π(ch)
- Dec(k, c) : replace each character ch in c with π^{-1}(ch)

Attempt 2: Substitution cipher

- Question: Does the old attack work?
- Answer: No!
  - Number of permutations = $26! \approx 2^{88}$, can't try each one!
- Question: Is this secure?
- Answer: Also no!
  - Idea: how many times does "X" show up in a message?
  - How many times does "E" show up in a message?
  - E is much more common!



Can count number of times letters shows up in ciphertext, match with frequency table

Perfect Secrecy [Shannon]

What Eve knows *after* looking at c

=

What Eve knew *before* looking at c

- **Probability distribution** P over a finite set S is a function $P : S \to [0,1]$ such that $\sum_{x \in S} P(x) = 1$
- **An event** is a set $A \subseteq S$; $\Pr[A] = \sum_{x \in A} P(x) \in [0,1]$
- **Union bound:** For events A_1 and A_2, $\Pr[A_1 \cup A_2] \leq \Pr[A_1] + \Pr[A_2]$
- A **random variable** X is a fn $X : S \to V$ that induces a dist. on V
- Events A and B are **independent** if $\Pr[A \text{ and } B] = \Pr[A] \cdot \Pr[B]$
- RVs X and Y are **ind.** if $\Pr[X = a \text{ and } Y = b] = \Pr[X = a] \cdot \Pr[Y = b]$

- $S = \{0,1\}^2$
- **Example distribution:** Uniform: for all $x \in S, P(x) = 1/|S|$
- **Example event:** $A = \{x \in S \mid \text{lsb}(x) = 1\}$. $\Pr[A] = 1/2$
- **Example RV:** $X = \text{lsb}$. Here $V = \{0,1\}$, and induced distribution is $\Pr[X = 0] = 1/2 \ ; \ \Pr[X = 1] = 1/2$
- **Example independent RVs:** $X = \text{lsb}$ and $Y = \text{msb}$

$$\Pr[X(x) = 0 \text{ and } Y(x) = 0] = \Pr[x = 00] = \frac{1}{4} = Pr[X(x) = 0] \Pr[Y(x) = 0]$$

Uniform RV

- A **Uniform RV** is $R : S \to S$ that induces a uniform dist on S.
- That is, for all $x \in S$, $\Pr[R = x] = 1/|S|$

Randomized algorithms

- Deterministic algorithm: $y \leftarrow A(m)$
- Randomized algorithm: $y \leftarrow A(m; R)$ where $R \xleftarrow{\$} \{0,1\}^n$
  - Output is a random variable $y \xleftarrow{\$} A(m)$

An important property of XOR

<u>Thm</u>: Y is an RV over $\{0,1\}^n$, X is a uniform ind. RV over $\{0,1\}^n$

Then $Z := Y \oplus X$ is uniform var. on $\{0,1\}^n$

Perfect Secrecy

$\forall m \in \mathscr{M}, \forall c \in \mathscr{C}, M$ is adversary's guess

$$\Pr[M = m \,|\, \mathsf{Enc}(\mathscr{K}, m) = c] = \Pr[M = m]$$

after

before

Probability that c encrypts the particular message m

Shannon's Perfect Secrecy Definition

$\forall m \in \mathscr{M}, \forall c \in \mathscr{C}, M$ is adversary's guess

$$\Pr[M = m \,|\, \text{Enc}(\mathscr{K}, m) = c] = \Pr[M = m]$$

after before

✓ **CT reveals no info about PT**

But this def is difficult to work with:

How to prove that ciphertext reveals no info?

Alternate Def: Perfect Indistinguishability

For every m, m'

Probability that c encrypts m (with random key k)

=

Probability that c encrypts m' (with diff. key k')

Hence every ciphertext is equally likely to decrypt to a given message

$$\forall m, m' \in \mathcal{M}, c \in \mathcal{C}$$
$$\Pr_{k \leftarrow \mathcal{K}}[\text{Enc}(k, m) = c] = \Pr_{k' \leftarrow \mathcal{K}}[\text{Enc}(k', m') = c]$$

The Two Definitions are Equivalent

THEOREM: An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ satisfies perfect secrecy IFF it satisfies perfect indistinguishability.

Intuition:

SEC → IND: If a ciphertext reveals no information about plaintext, it can equally likely be an encryption for m or m'

IND → SEC: If for any m, m', ciphertext is equally likely to decrypt to either m or m', then it reveals no "distinguishing" information about m or m'. Since this works for any m, m', ciphertext reveals no information about *any* message.

Perfect Secrecy is Achievable

The One-time Pad Construction:

Gen: Choose an n-bit string k at random, i.e. $k \leftarrow \{0,1\}^n$

Enc(k, m) with $\mathscr{M} = \{0,1\}^n$: Output $c = m \oplus k$

Dec(k, c): Output $m = c \oplus k$

Perfect Secrecy is Achievable

The One-time Pad Construction:

Gen: Choose an n-bit string k at random, i.e. $k \leftarrow \{0,1\}^n$

Enc(k, m) with $\mathscr{M} = \{0,1\}^n$: Output $c = m \oplus k$

Dec(k, c): Output $m = c \oplus k$

Correctness: $c \oplus k = m \oplus k \oplus k = m$

Perfect Secrecy is Achievable

The One-time Pad Construction:

Gen: Choose an n-bit string k at random, i.e. $k \leftarrow \{0,1\}^n$

Enc(k, m) with $\mathscr{M} = \{0,1\}^n$: Output $c = m \oplus k$

Dec(k, c): Output $m = c \oplus k$

Claim: One-time Pad achieves Perfect Indistinguishability (and therefore perfect secrecy).

Proof: For any $m, c \in \{0,1\}^n$,

$$\Pr_{k \leftarrow \mathscr{K}} [\text{Enc}(k, m) = c] = \Pr[k \oplus m = c] = \Pr[k = c \oplus m] = 1/2^n$$

Perfect Secrecy is Achievable

The One-time Pad Construction:

Gen: Choose an n-bit string k at random, i.e. $k \leftarrow \{0,1\}^n$

Enc(k, m) with $\mathscr{M} = \{0,1\}^n$: Output $c = m \oplus k$

Dec(k, c): Output $m = c \oplus k$

Claim: One-time Pad achieves Perfect Indistinguishability (and therefore perfect secrecy).

Proof: For any m, m'

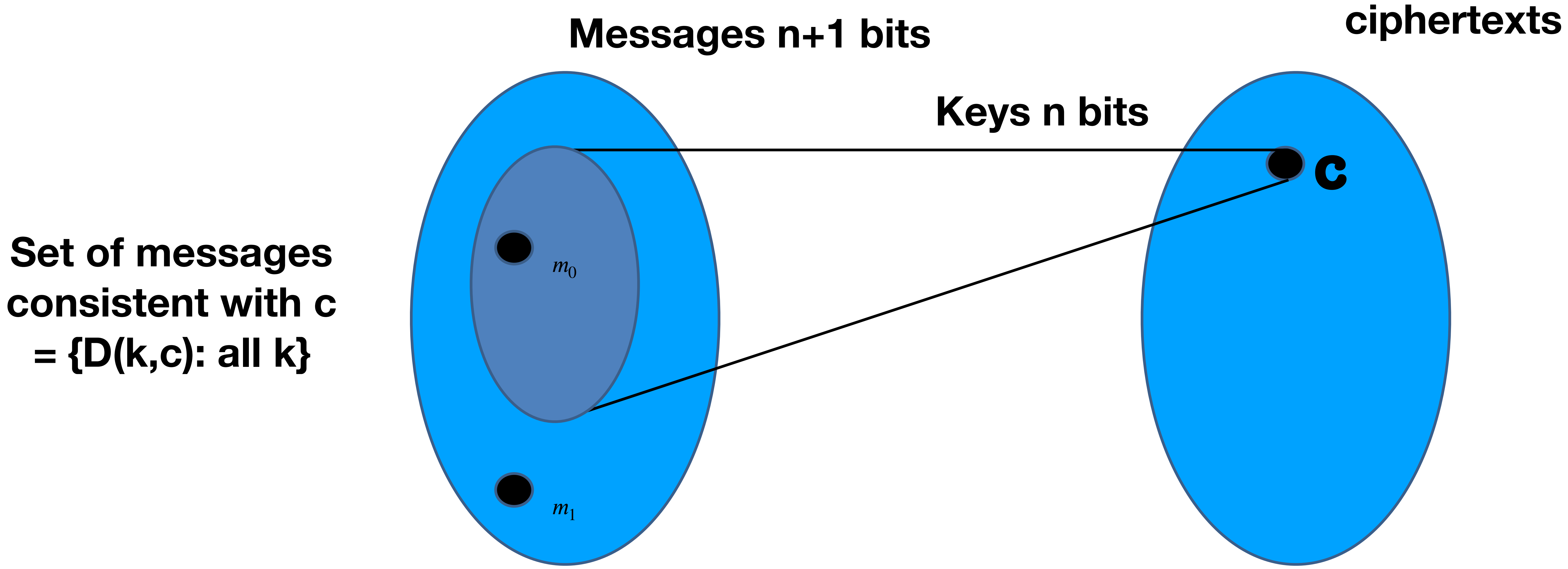
$$\Pr[\text{Enc}(k, m) = c] = 1/2^n = \Pr[\text{Enc}(k, m') = c]$$

Perfect Secrecy has its Price

THEOREM: For any perfectly secure encryption scheme,

$$|\mathcal{K}| \geq |\mathcal{M}|$$

Shannon's impossibility!



Each cipher text can correspond to at most 2^n messages, but message space contains 2^{n+1} possible messages!

So it is possible (and likely!) that a given cipher text can *never* decrypt to m_1!

$$\Pr[\mathsf{Enc}(\mathscr{K}, m_1) = c] = 0$$

Why is this bad?

- Exchanging large keys is difficult
- Need to keep large keys secure for a long time
- Generating truly random bits is kinda expensive!

So what can we do?