CIS 5371, Fall 2024

Introduction

Viet Tung Hoang

Agenda

1. Crypto Usage & Goals

2. Classical Crypto

3. One-time Pad & Perfect Secrecy

4.Modern Crypto

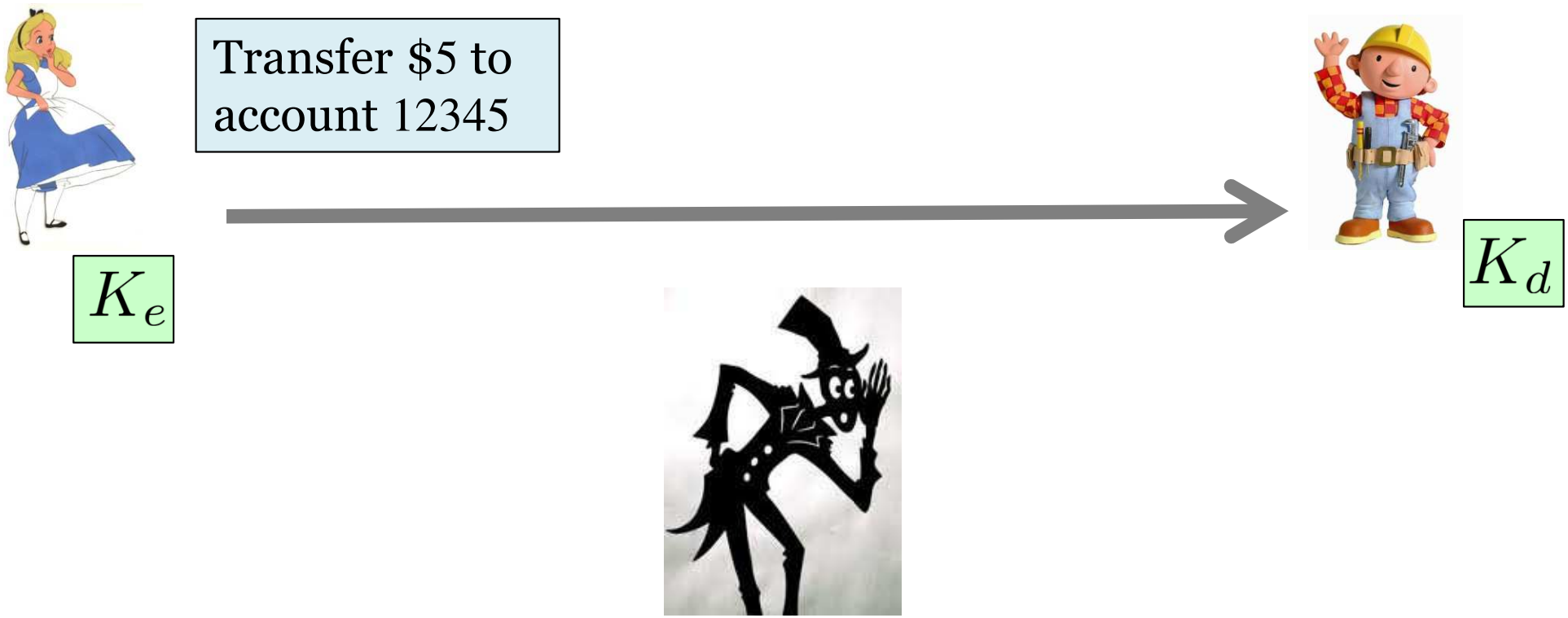
Crypto Use Is Ubiquitous

HTTPS

Secure messaging

Bitcoin

A Classical Crypto Goal: Privacy



Privacy: Adversary can't learn anything from the content that it eavesdrops.

A Classical Crypto Goal: Privacy

K_e

1100000100101
0101011110111

K_d

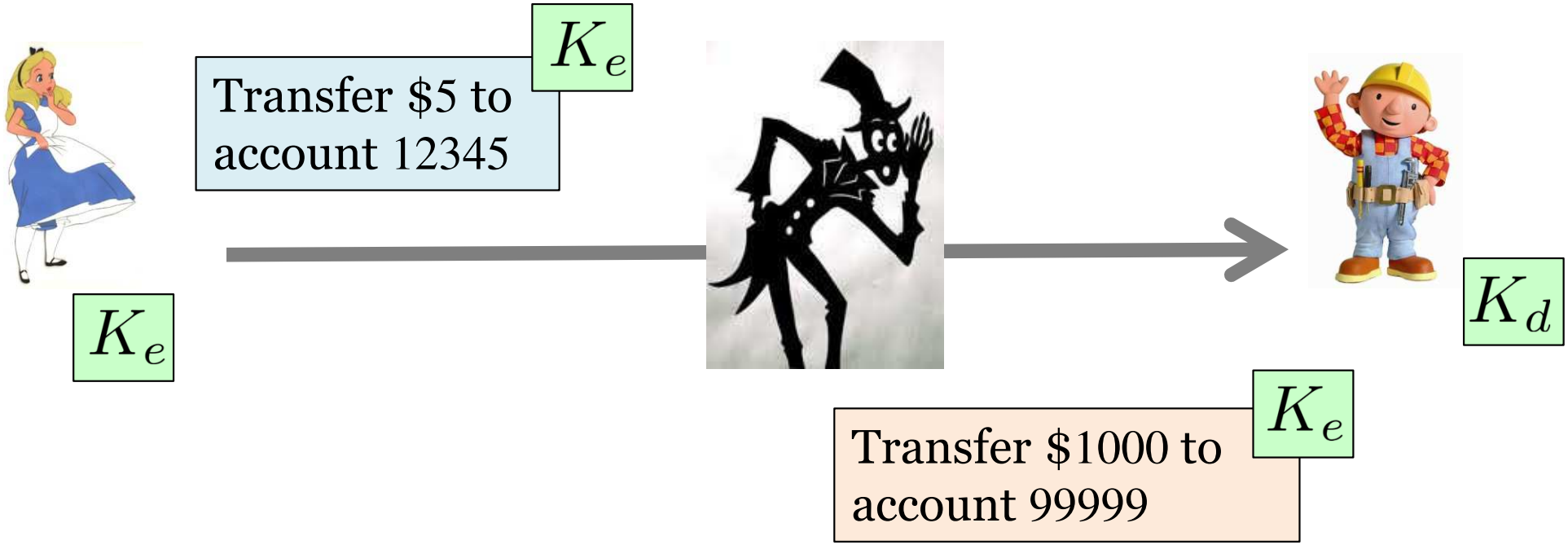
Private-key setting: $K_e = K_d$

secret

Public-key setting: $K_e \neq K_d$

public secret

But Privacy Alone Is <u>Not</u> Enough



Authenticity: Adversary can't forge valid ciphertexts

Four Fundamental Cryptographic Problems

	Privacy	Authenticity
Private key	**Private-key Enc**	**MAC**
Public key	**Public-key Enc**	**Digital Signature**

Agenda

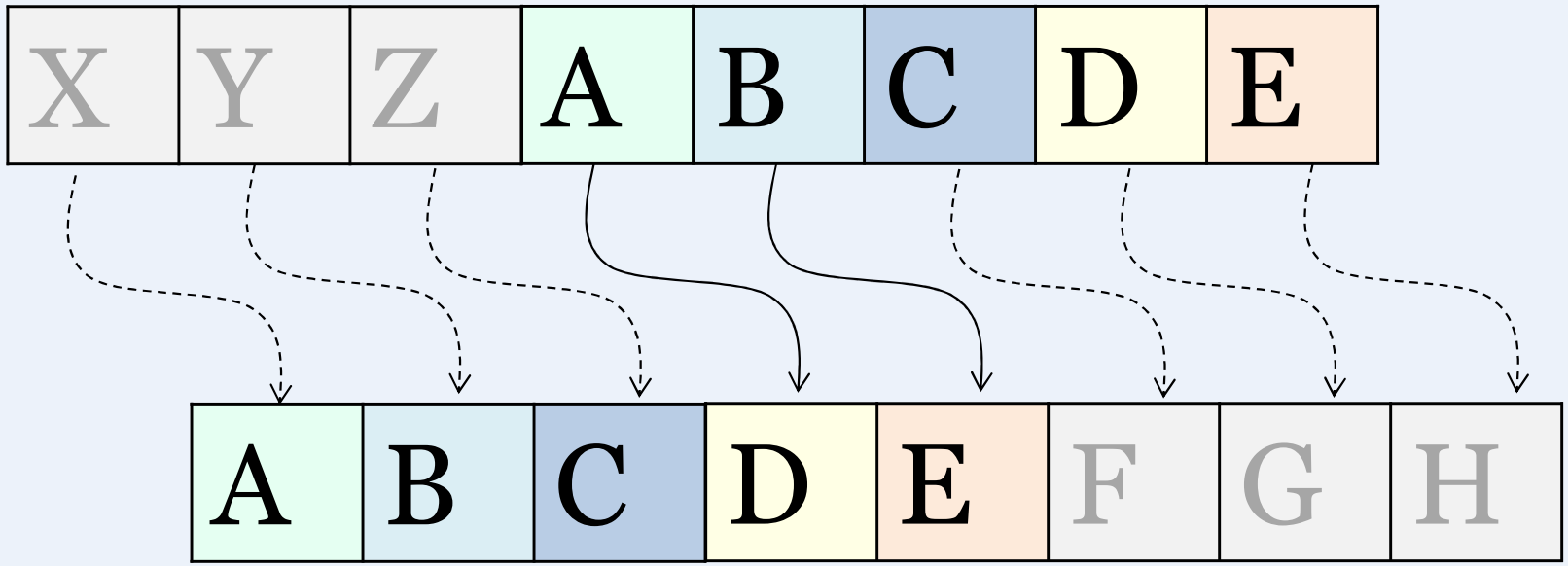
1. Crypto Usage & Goals

2. **Classical Crypto**

3. One-time Pad & Perfect Secrecy

4. Modern Crypto

Caesar Cipher



No key

Broken once scheme is known

Caesar Cipher In The Wild

The Register®

This article is more than **1 year old**

Mafia boss undone by clumsy crypto

Little Caesar

John Leyden

Wed 19 Apr 2006 // 14:14 UTC

Clues left in the clumsily encrypted notes of a Mafia don have helped Italian investigators to track his associates and ultimately contributed to his capture after years on the run.

The Register®

This article is more than **1 year old**

BA jihadist relied on Jesus-era encryption

30 years for airline bomb plot

Team Register

Tue 22 Mar 2011 // 11:52 UTC

An IT worker from British Airways jailed for 30 years for terrorism offences used encryption techniques that pre-date the birth of Jesus.

Shift Cipher

Use a secret key $K \in \{0, \ldots, 25\}$

Same as Caesar cipher, but shift K positions, instead of 3.

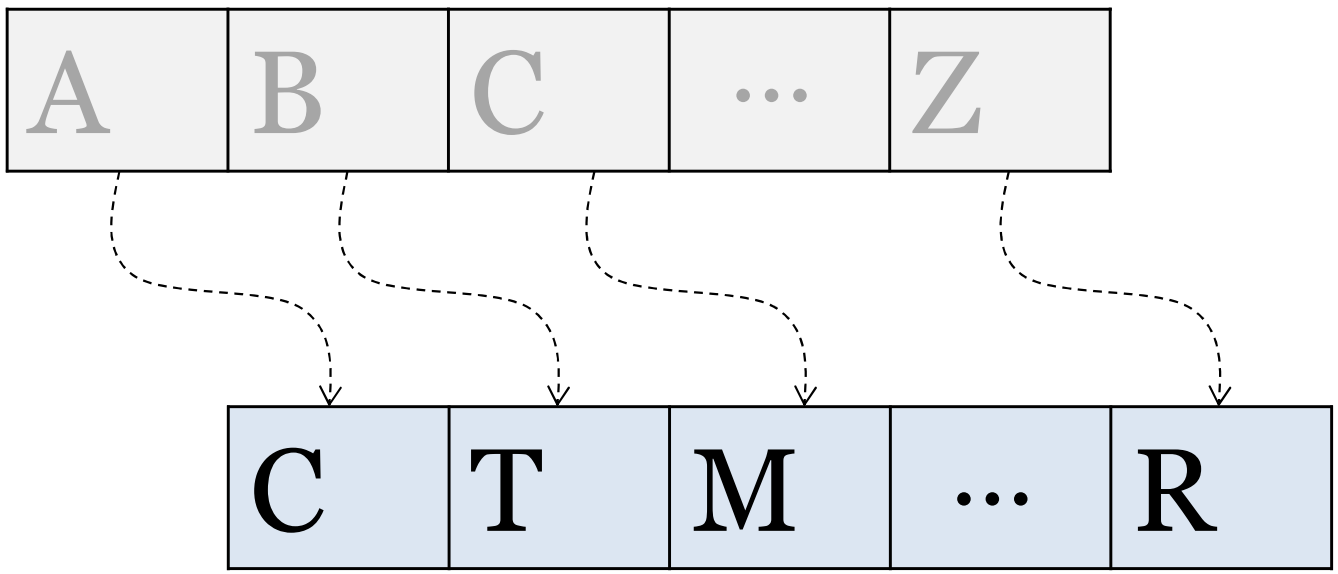
Small Keyspace

Broken by brute force

Substitution Cipher

Key: a permutation $\pi : \Sigma \to \Sigma$

Example: $\Sigma = \{\mathrm{A}, \mathrm{B}, \mathrm{C}, \ldots, \mathrm{Z}\}$

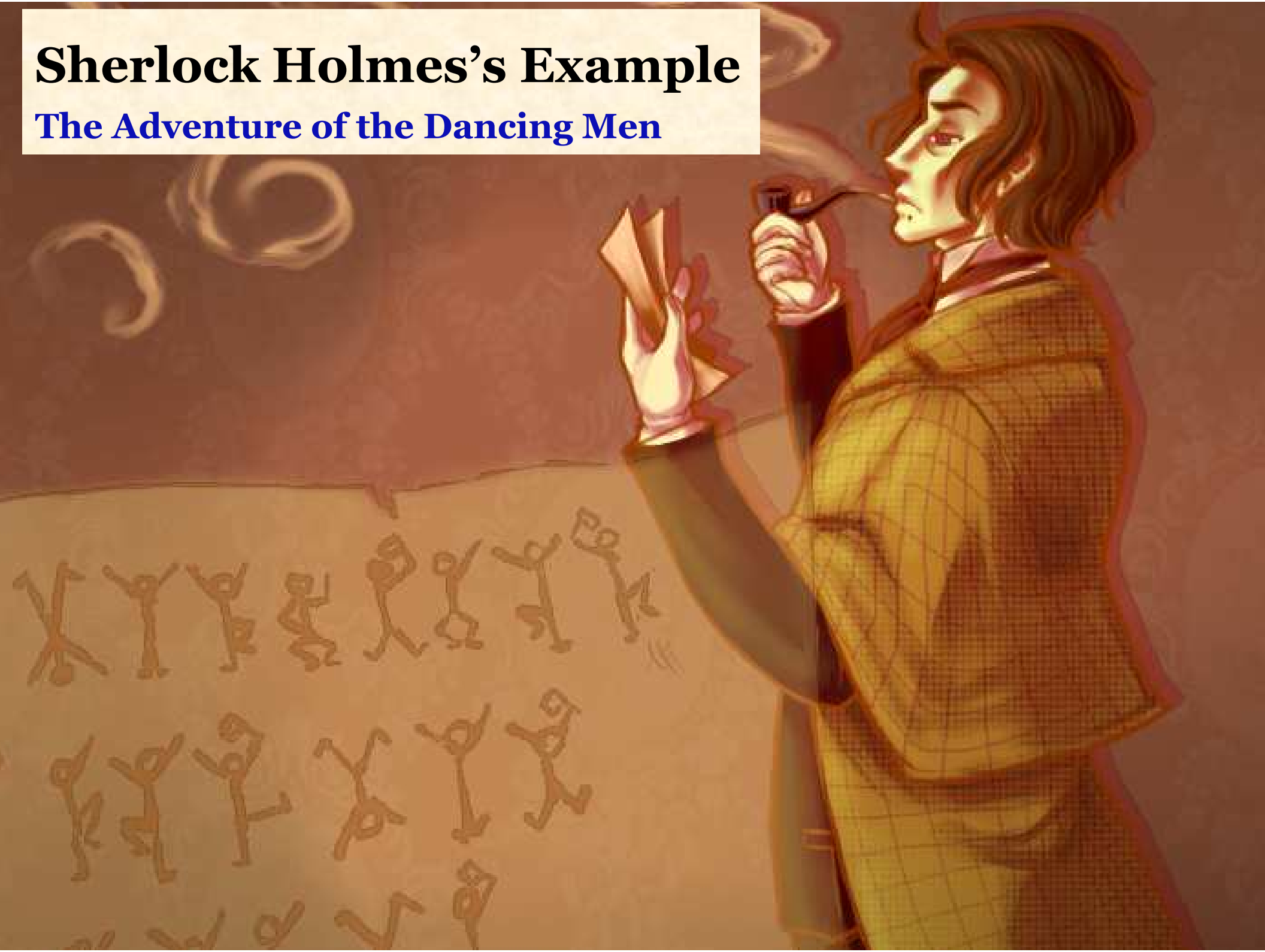


$$26! \approx 2^{88}$$

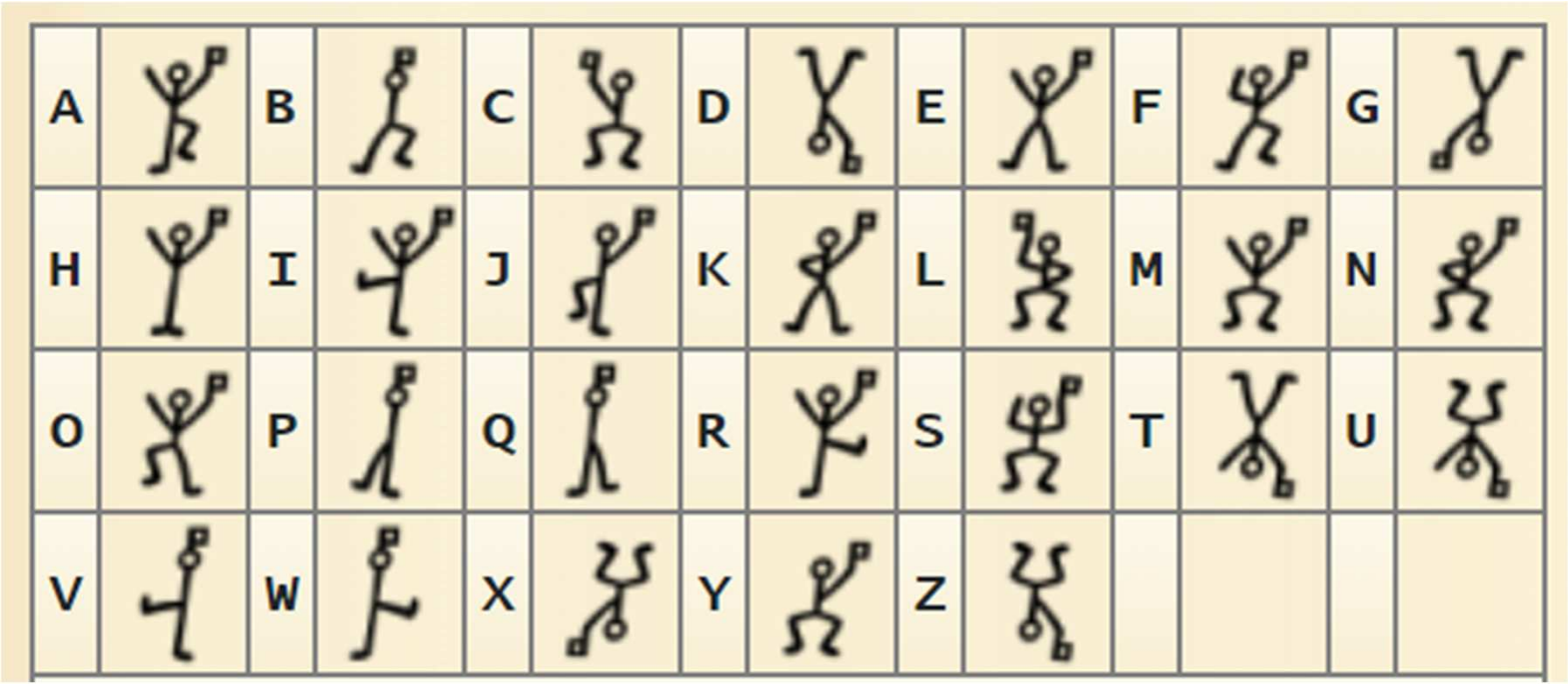
keys for the English alphabet

Sherlock Holmes's Example

The Adventure of the Dancing Men

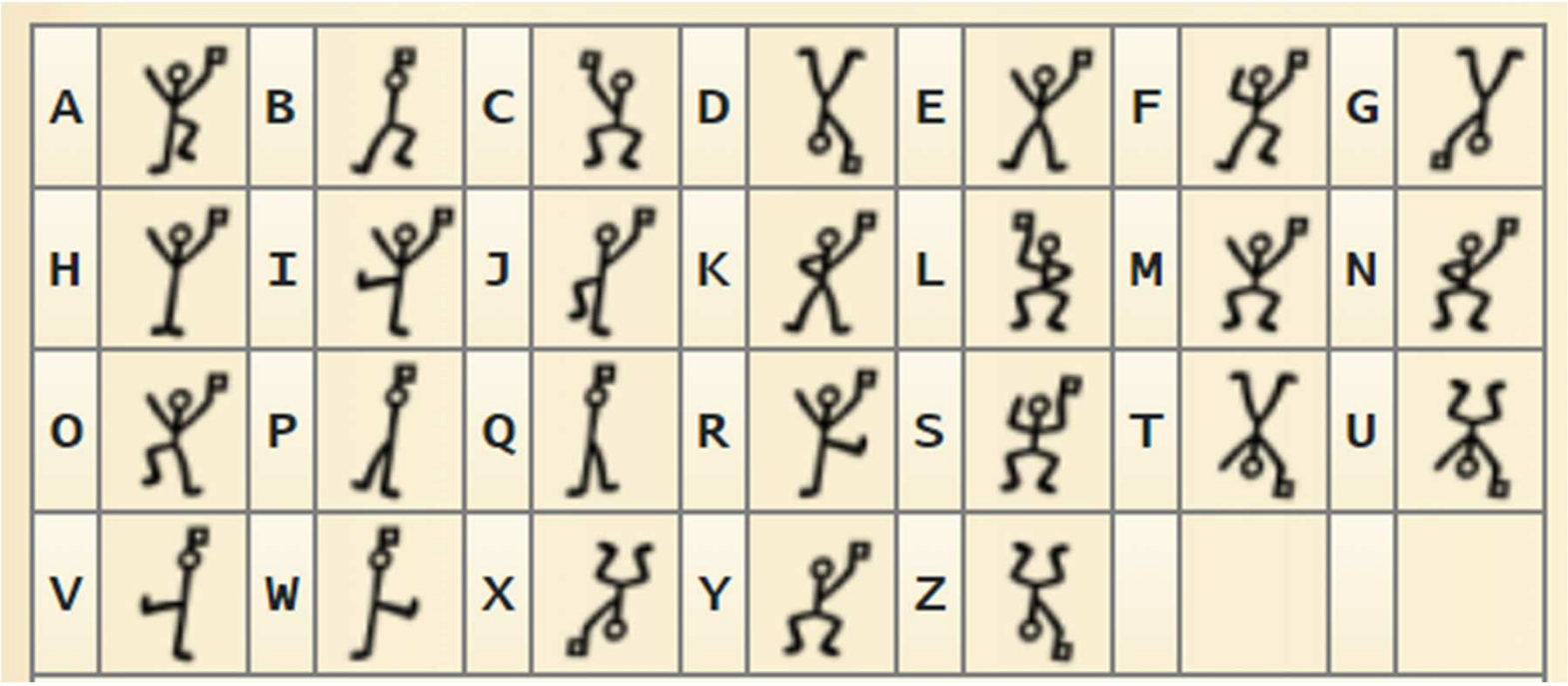


Breaking The Dancing-Men Code



Problem: DM code is **not** a substitution cipher, as its output is not in English alphabet.

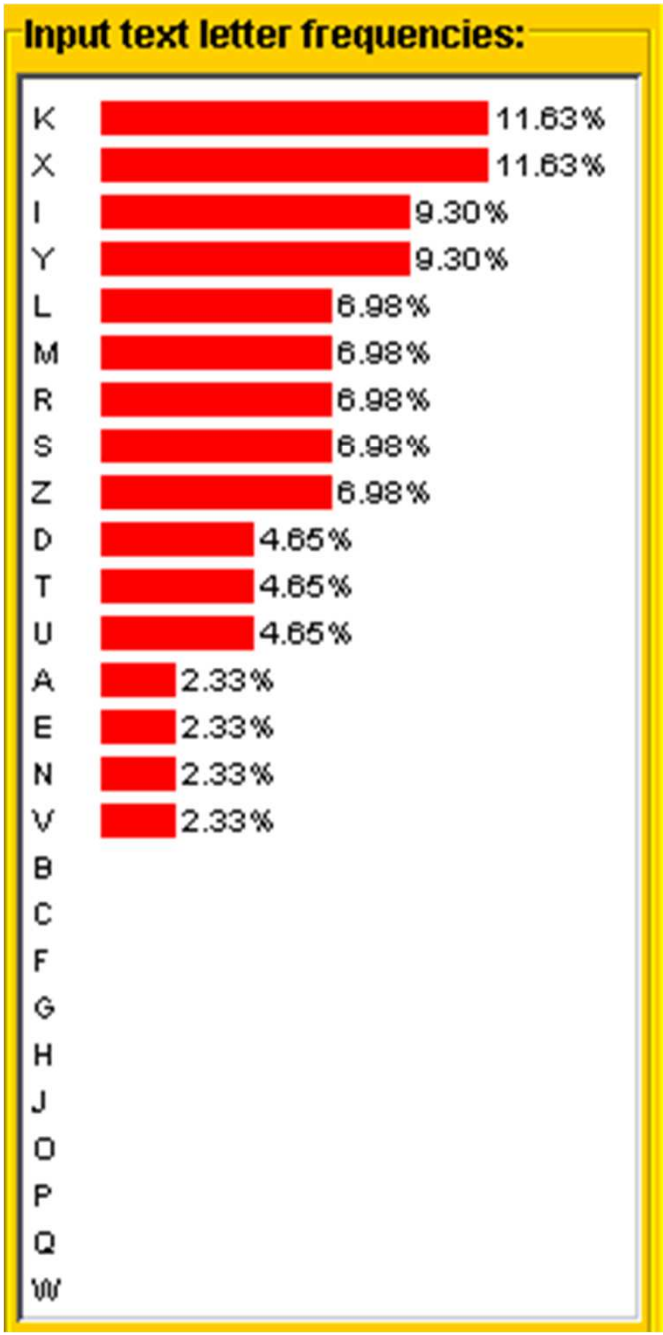
Exercise

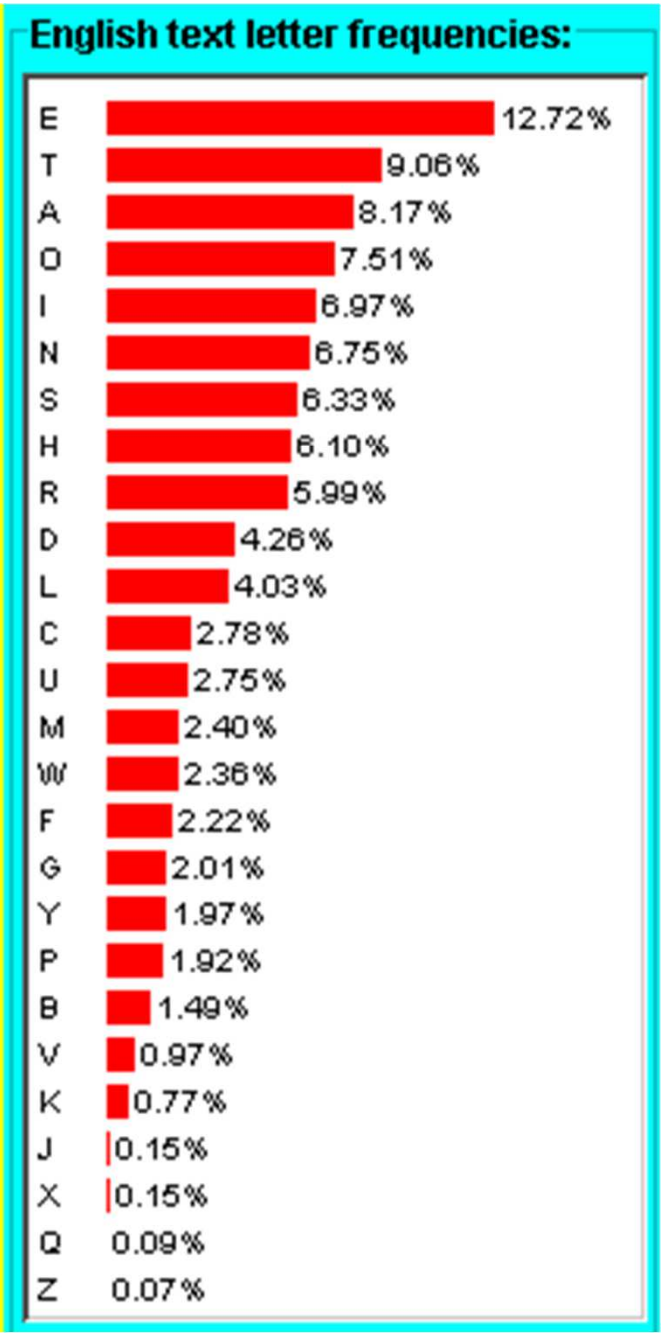


Given a machine Π to break substitution cipher,

find a way to use Π to break DM code

Break Substitution Cipher: Frequency Analysis

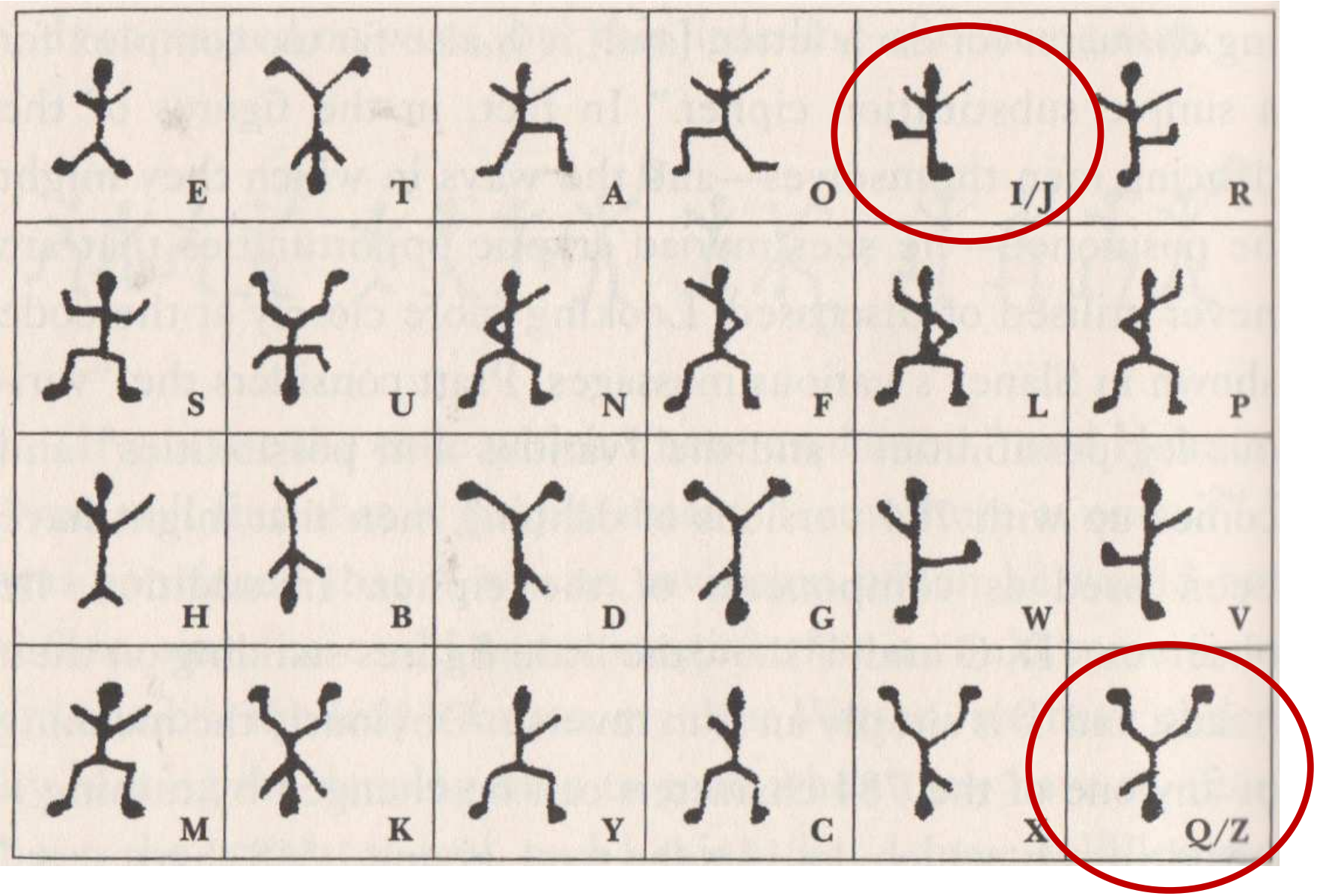




Only work for

large ciphertexts

Real-world Issues: Many-to-one Mapping



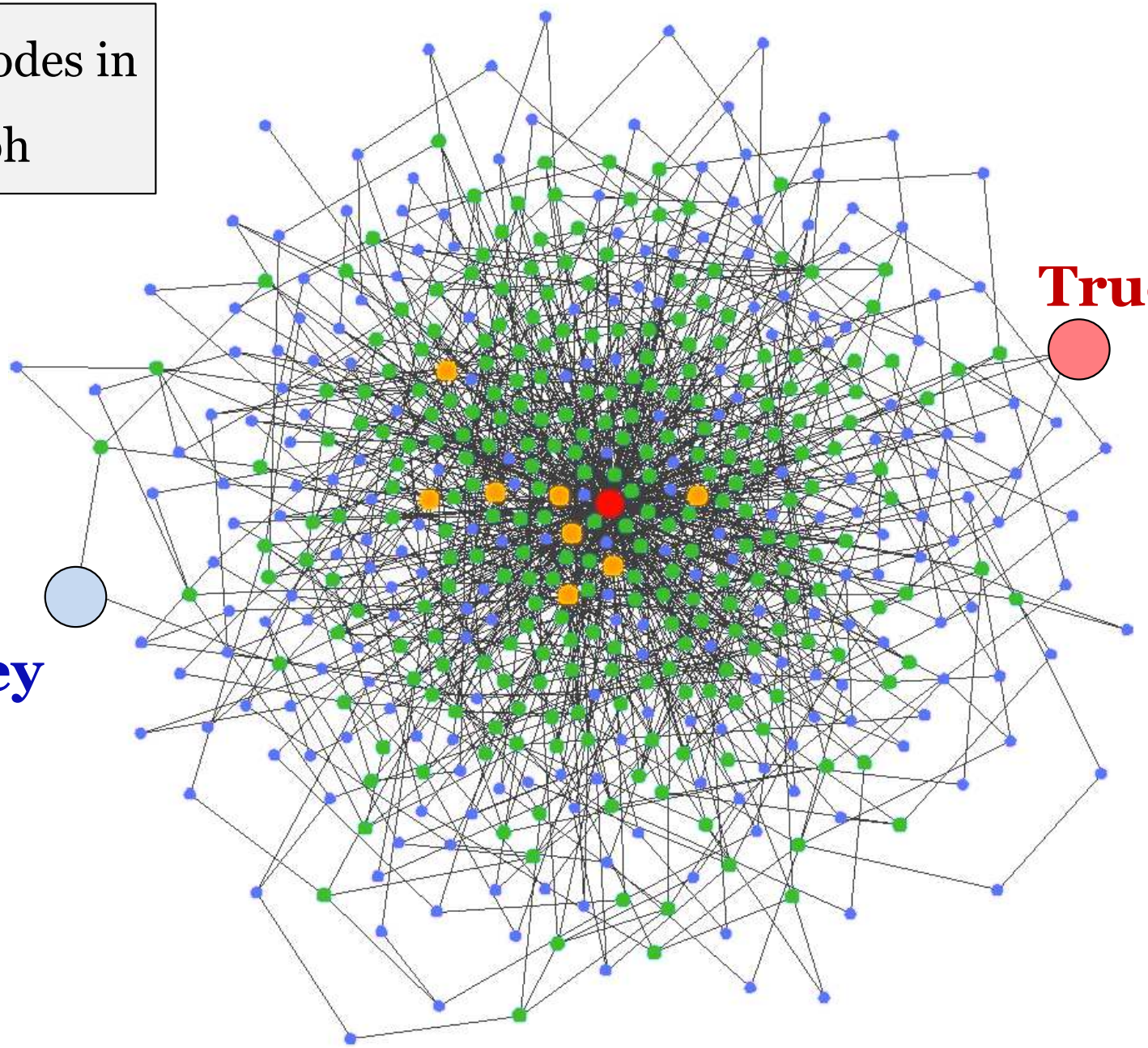
Real-world Issues: Imperfect English

to bat-rb. con todo mi respeto. I was sitting down playing chess with danny de emf and boxer de el centro was sitting next to us. Boxer was making loud and loud voices so I tell him por favor can you kick back homie

A decrypted jail letter

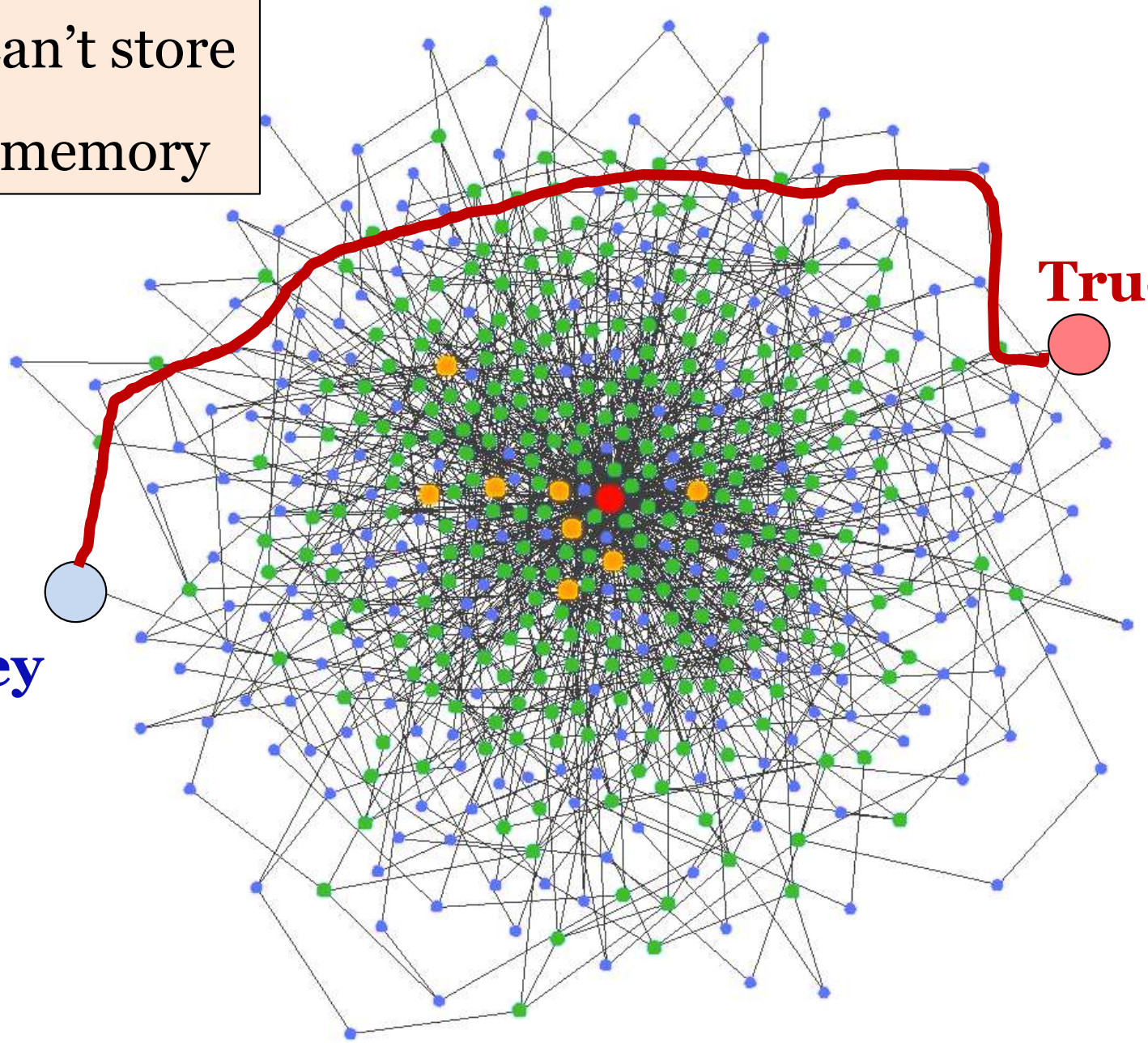
A Better Attack via Monte Carlo Markov Chain

View keys as nodes in a complex graph



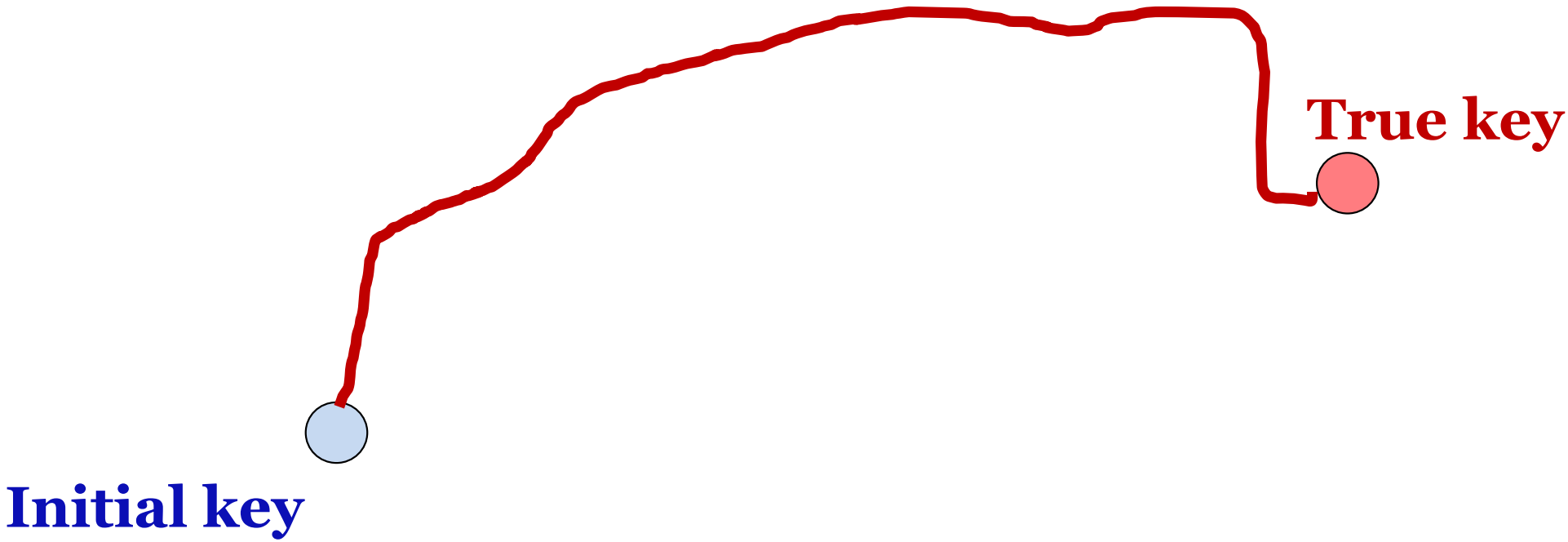
Our Goal Is To Find A Path To The Destination

Problem: Can't store the graph in memory



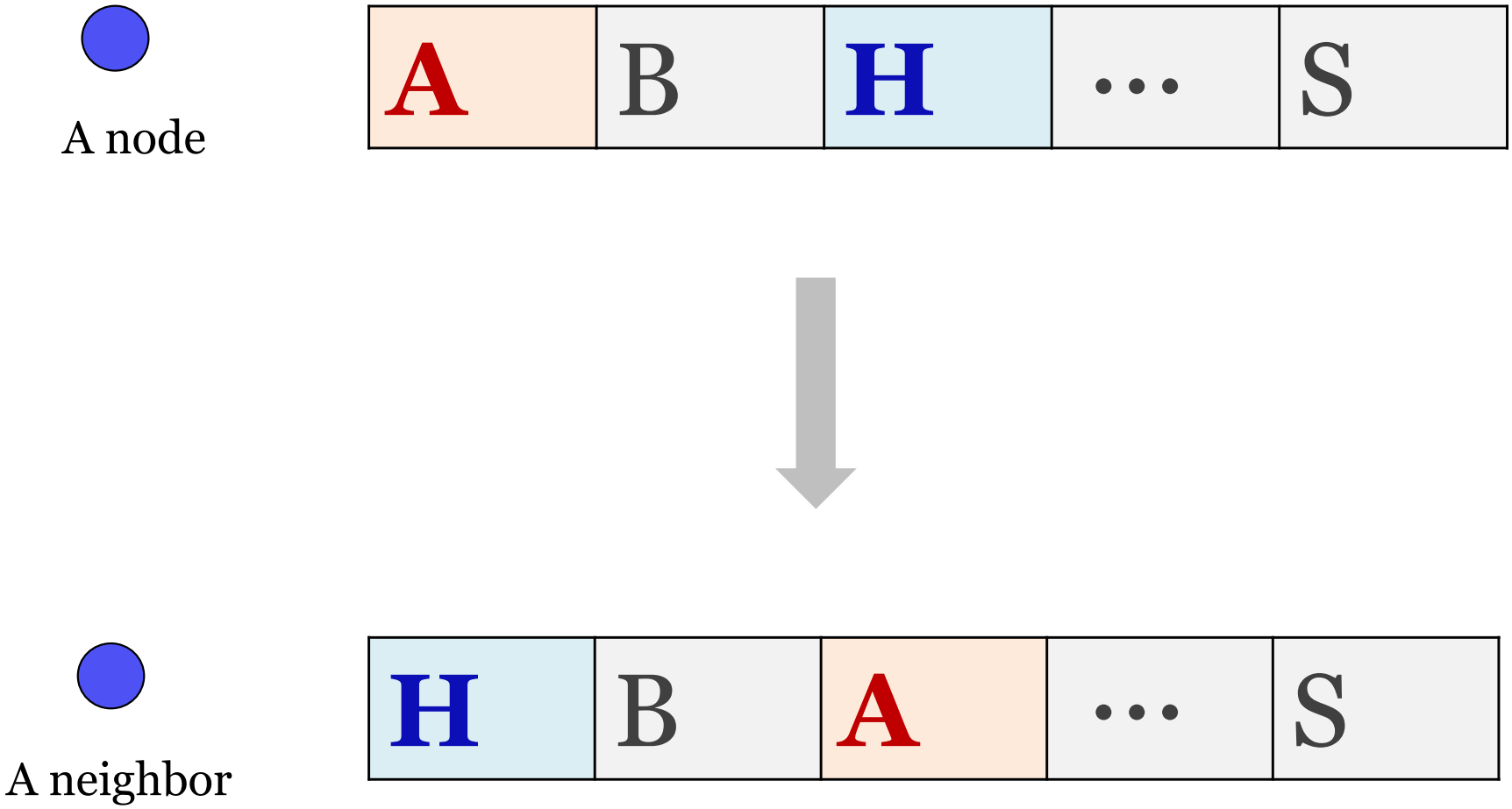
How Far Is The Destination?

Is It Short Enough To Travel In Reasonable Time?



Claim: By properly defining how nodes are connected, the distance will be **short**.

Define Connectivity: Swap Two Entries

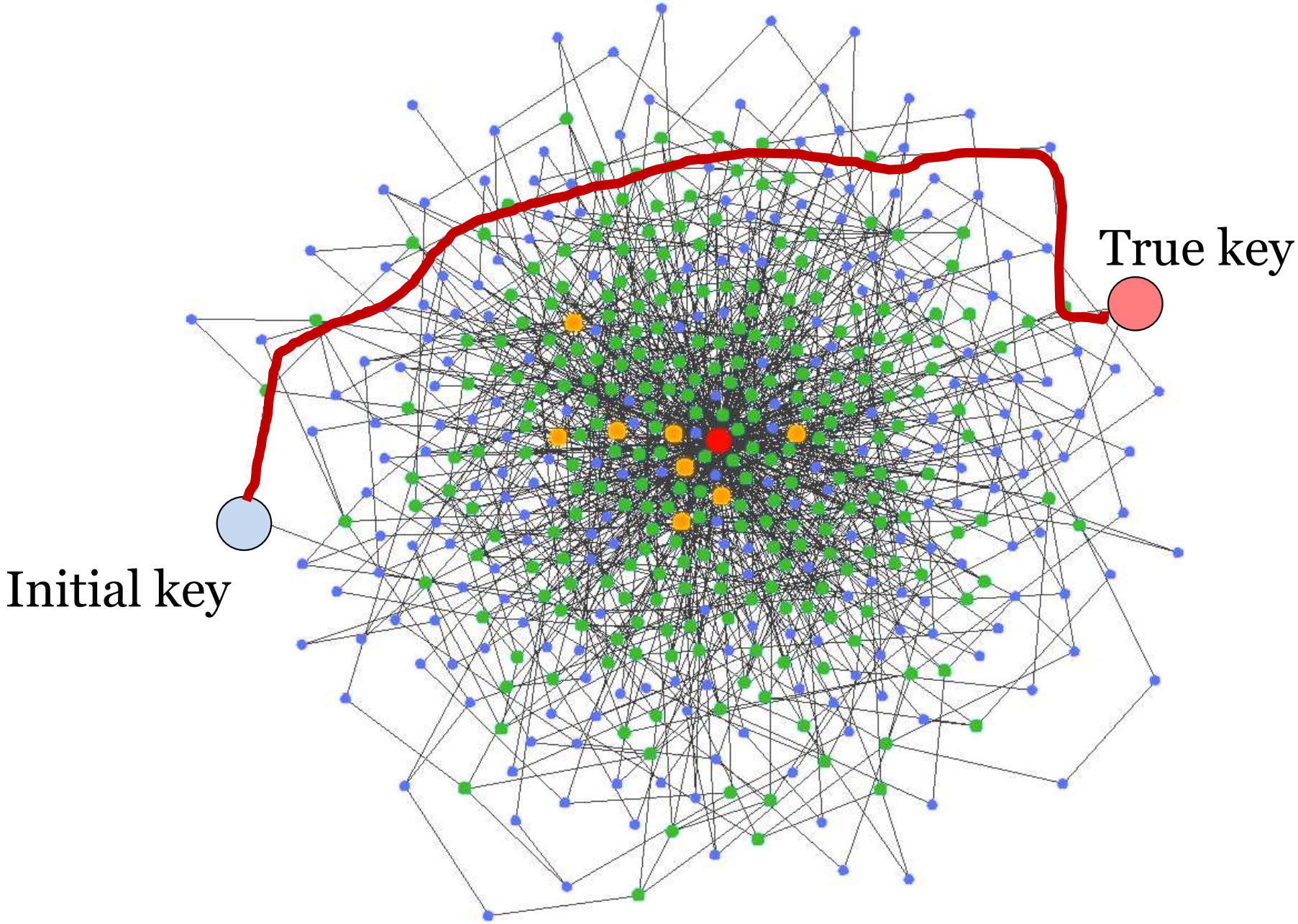


Quiz: Find The Distance Between These Nodes

A	B	C	D	E

D	B	A	E	C

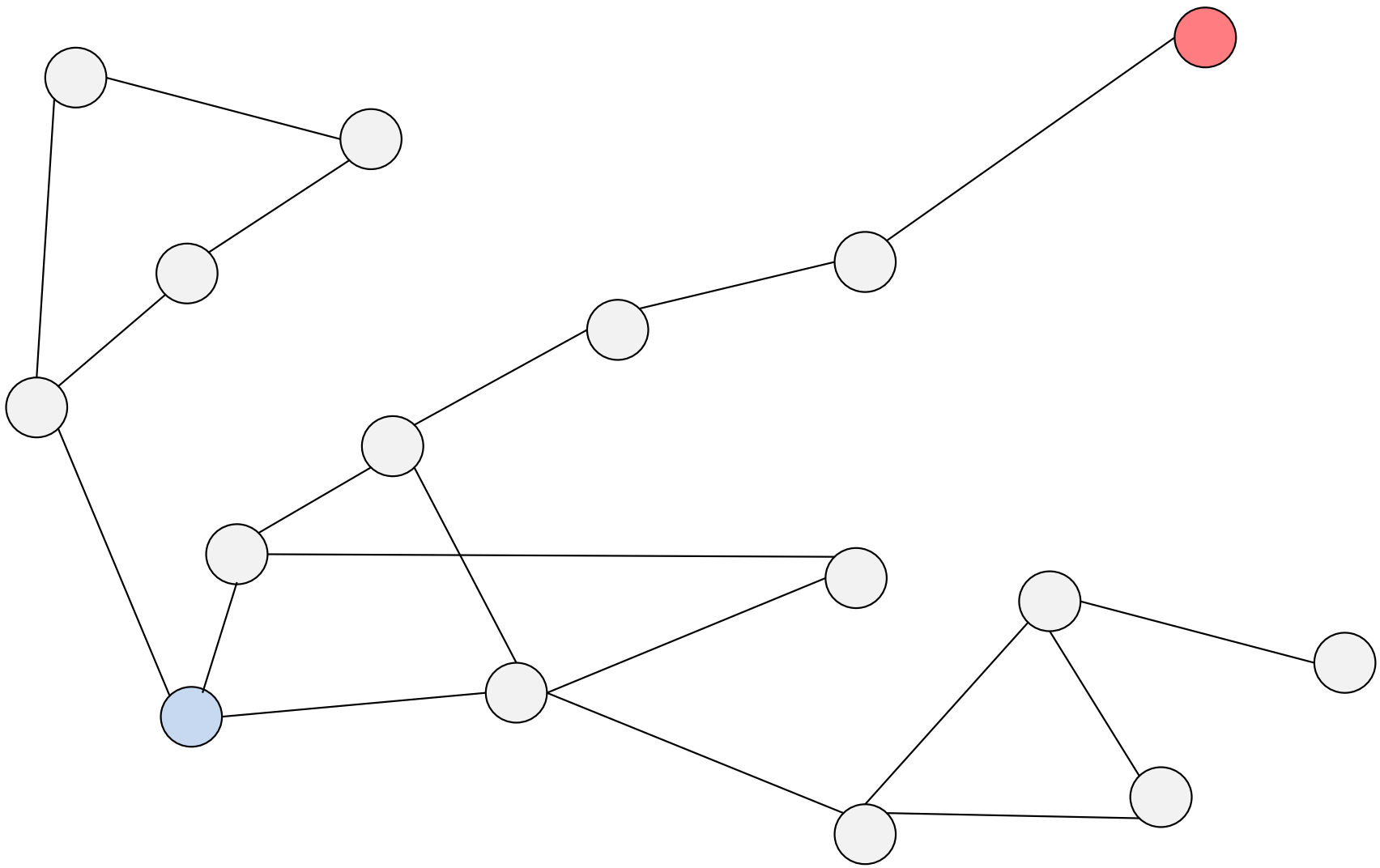
How Far Is The Distance For This Huge Graph?



25

steps to go from source to destination

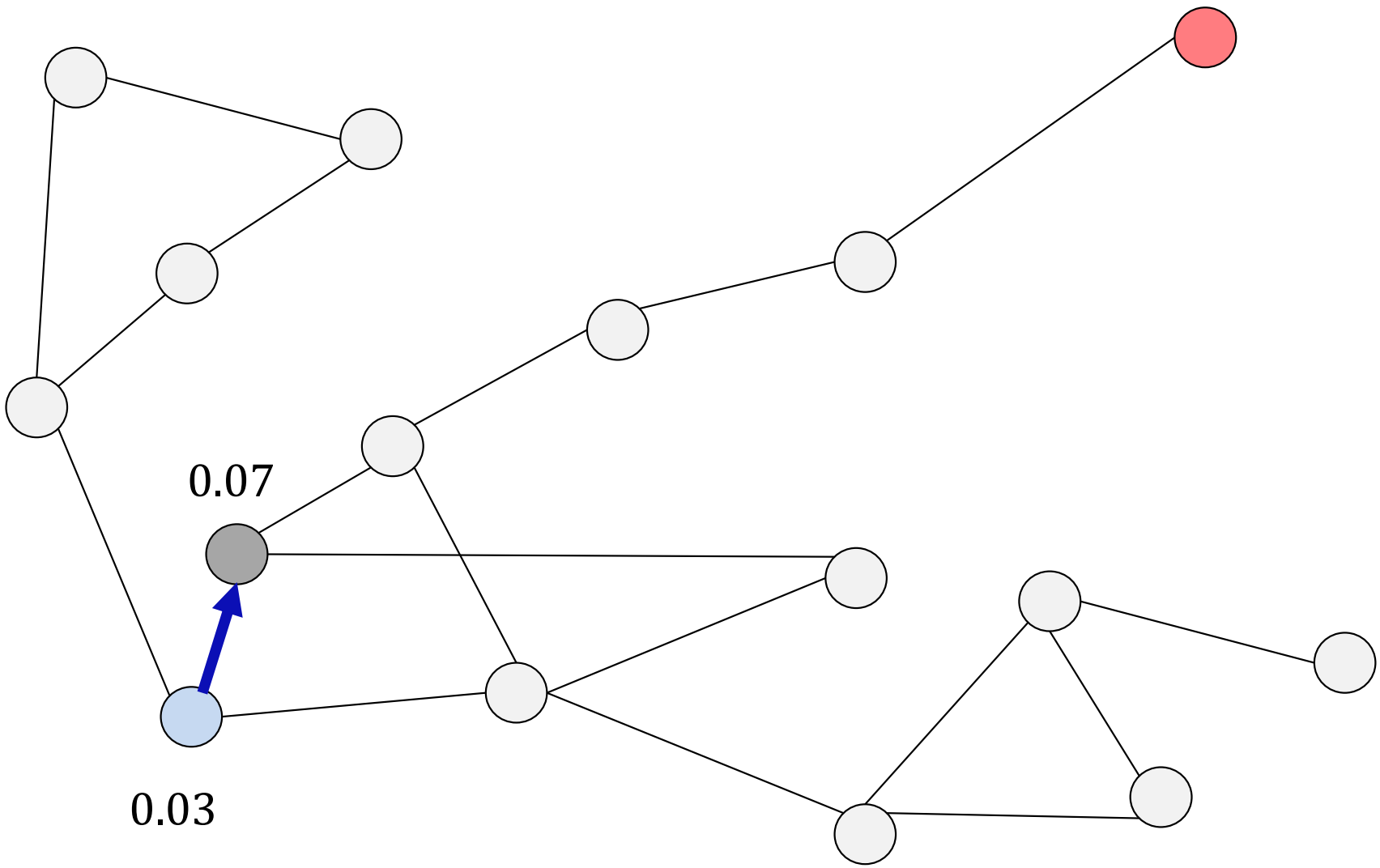
How To Find A Way When Can’t Store Graph?



How To Find A Way When Can't Store Graph?

Estimate scores for current node and a **random** neighbor

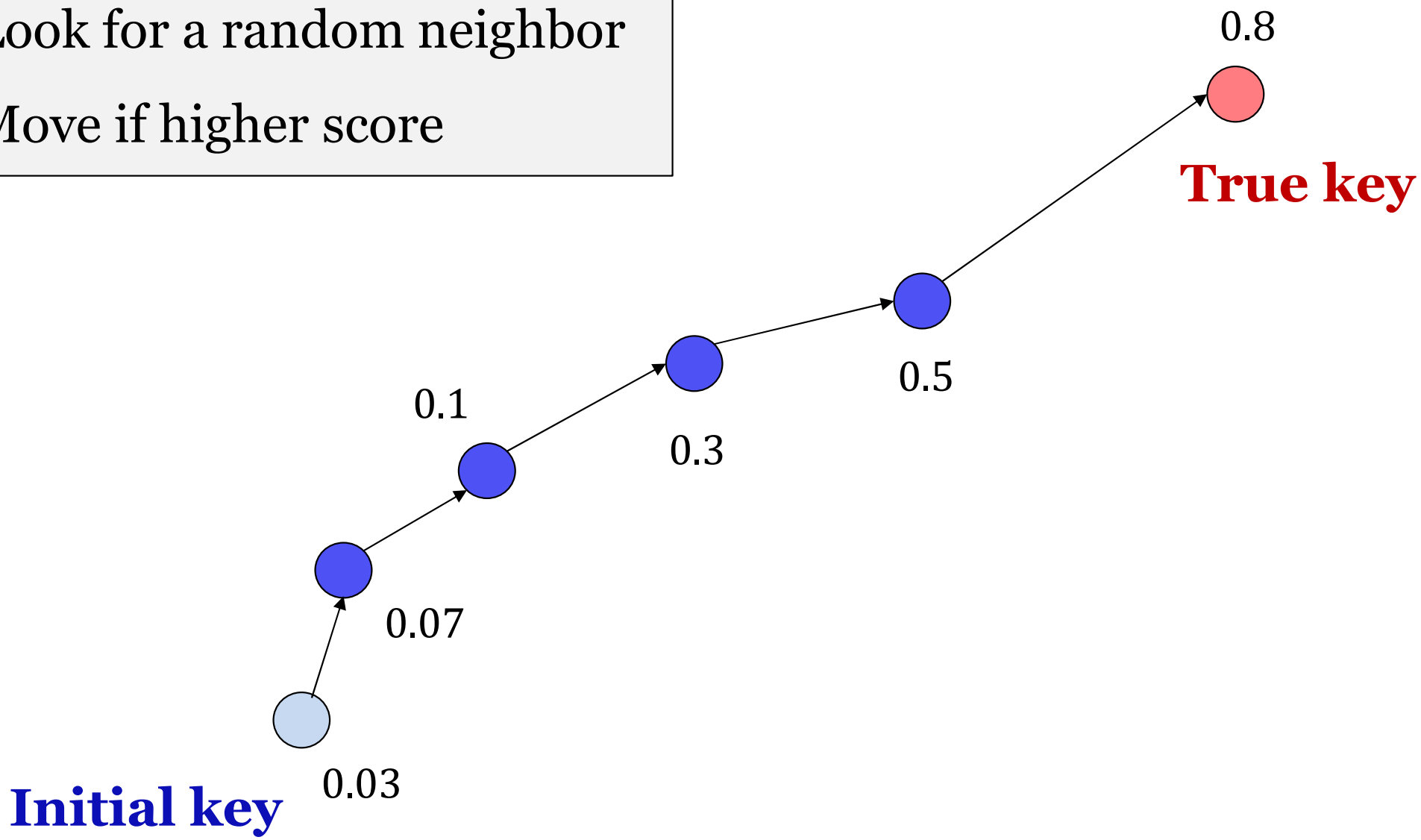
Higher score → Closer to destination → Move



Graph Traversing: Hill Climbing

-Look for a random neighbor

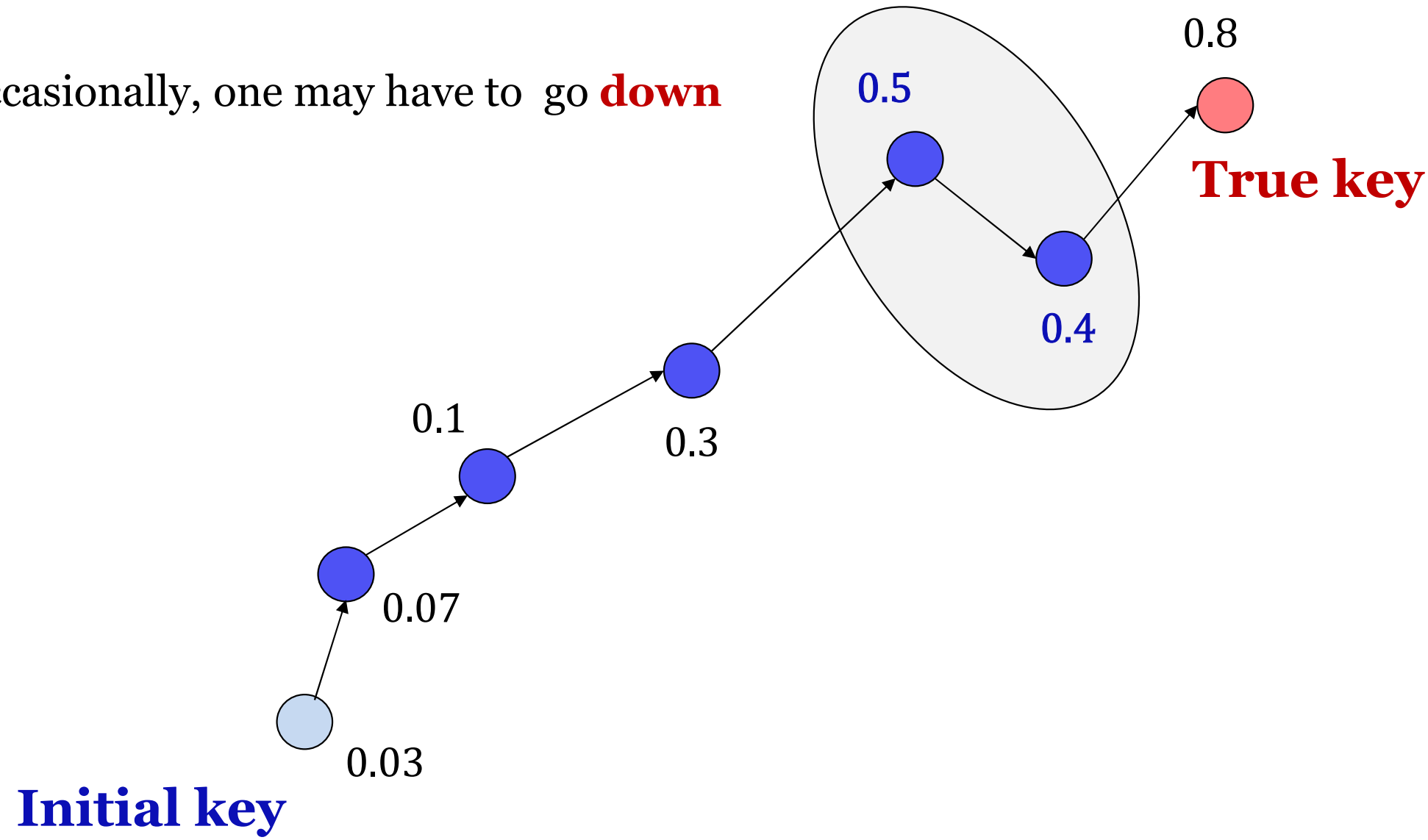
-Move if higher score



Implementing Hill Climbing

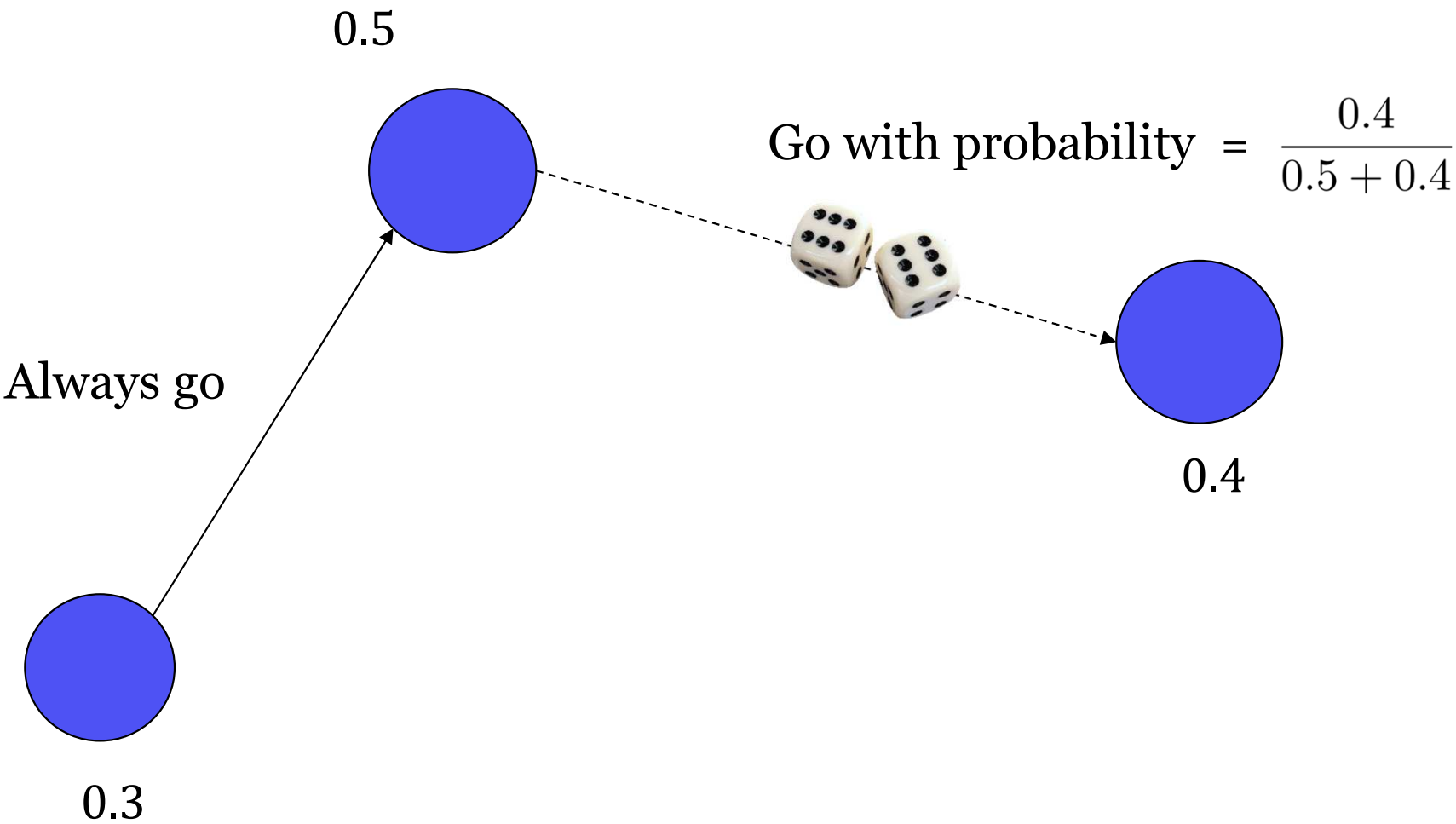
Get Around Local Maximum in Hill Climbing

Occasionally, one may have to go **down**



Implementing Hill Climbing

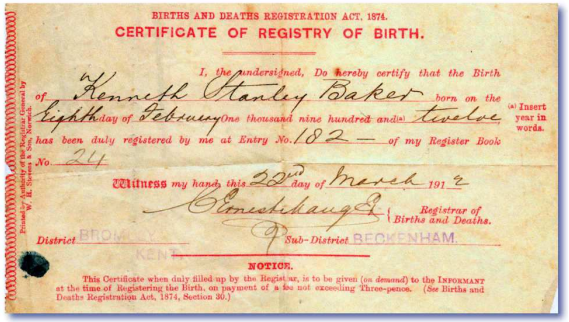
To Go or Not To Go



Estimating Score

Ciphertext

Candidate key K



BIRTHS AND DEATHS REGISTRATION ACT, 1874.

CERTIFICATE OF REGISTRY OF BIRTH.

I, the undersigned, Do hereby certify that the Birth of Kenneth Stanley Baker born on the eighth day of February one thousand nine hundred and (a) twelve has been duly registered by me at Entry No. 182 of my Register Book No. 244

(a) Insert year in words.

Witness my hand, this 22nd day of March 1912

Registrar of Births and Deaths.

District BROMLEY KENT Sub-District BECKENHAM

NOTICE.

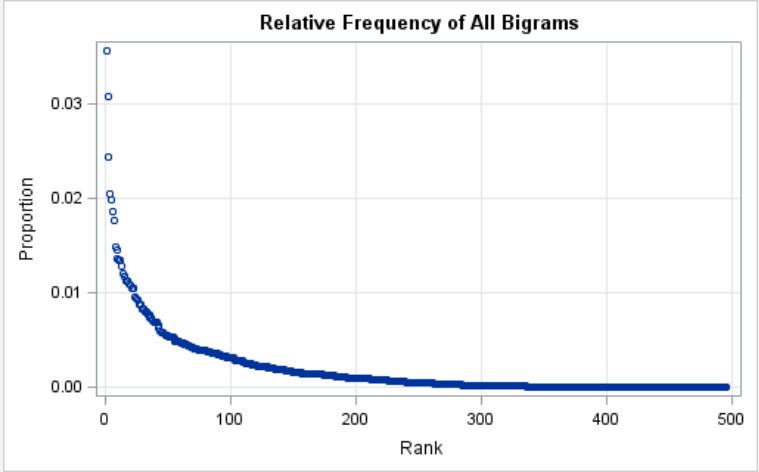
This Certificate when duly filled up by the Registrar, is to be given (on demand) to the INFORMANT at the time of Registering the Birth, on payment of a fee not exceeding Three-pence. (See Births and Deaths Registration Act, 1874, Section 30.)

Decrypted message

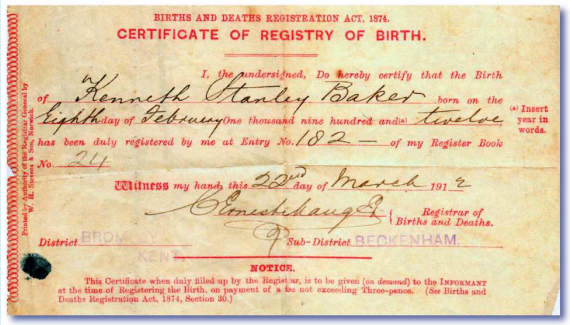
Need to tell how decrypted message looks like an English text

Estimating Score: Overview

Training phase



Testing phase


BIRTHS AND DEATHS REGISTRATION ACT, 1874.

CERTIFICATE OF REGISTRY OF BIRTH.

I, the undersigned, Do hereby certify that the Birth of Kenneth Stanley Baker born on the Eighth day of February One thousand nine hundred and twelve has been duly registered by me at Entry No. 182 — of my Register Book No. 24

Witness my hand this 22nd day of March 1912

Registrar of Births and Deaths.

District BROMLEY Sub-District BECKENHAM

NOTICE.

This Certificate when duly filled up by the Registrar, is to be given (on demand) to the INFORMANT at the time of Registering the Birth, on payment of a fee not exceeding Three-pence. (See Births and Deaths Registration Act, 1874, Section 30.)

Tentative decrypted message

Frequency histogram

Score

Training Phase: Bigram Analysis

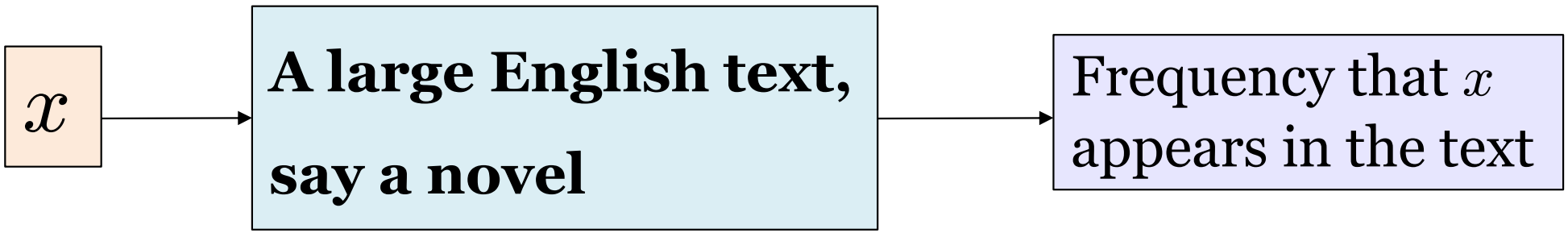
"*the cat is chasing the mouse*"

Bigrams: th | he | e□ | □c | ca | at | t□ | …

must consider white space

Frequency of Bigrams in English

$$\overbrace{\text{For each bigram } x}^{27 \times 27 \text{ iterations}} \text{ do:}$$



Store the frequencies in an array F

An Example

repetition included

"ca" appears 1235 times out of tottal 437907 bigrams

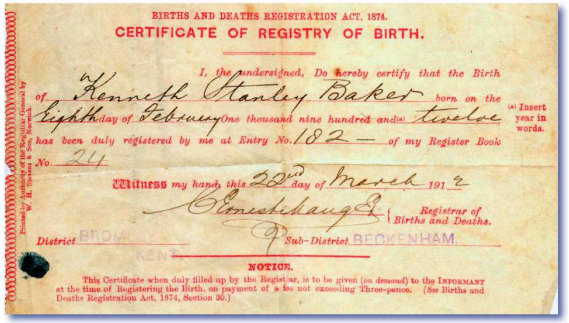
$F[\text{“ca”}] = \frac{1235}{437907}$: frequency of bigram "ca" in English

Some bigrams will have frequency 0, as they don't appear in English

Testing Phase: Assess Englishness

Ciphertext

Candidate key K



BIRTHS AND DEATHS REGISTRATION ACT, 1874.

CERTIFICATE OF REGISTRY OF BIRTH.

I, the undersigned, Do hereby certify that the Birth of Kenneth Stanley Baker born on the eighth day of February one thousand nine hundred and (a) twelve has been duly registered by me at Entry No. 182 — of my Register Book No. 24

(a) Insert year in words.

Witness my hand, this 28th day of March 1912

Registrar of Births and Deaths.

District BROMLEY KENT Sub-District BECKENHAM.

NOTICE.

This Certificate when duly filled up by the Registrar, is to be given (on demand) to the INFORMANT at the time of Registering the Birth, on payment of a fee not exceeding Three-pence. (See Births and Deaths Registration Act, 1874, Section 30.)

Decrypted message M

$$\text{Score}[K] = \prod F[x]$$

bigram x in M (repetition included)

An Example

Decrypted message = "*ale lens*"

Bigrams: al | le | e□ | □l | le | en | ns

$$\textbf{Score} = F[\text{al}] \cdot F[\text{le}] \cdot F[\text{e}\square] \cdot F[\square\text{l}] \cdot F[\text{le}] \cdot F[\text{en}] \cdot F[\text{ns}]$$

Programming Issues: When To Stop?

Loop forever, but keep track of the best node so far

User decide when to terminate, typically after 5000 steps

Programming Issues: Numerical Errors

$$\text{Score}[K] = \prod_{\text{bigram } x \text{ in } M} F[x]$$

Example: 1000 bigrams, and each of frequency ~ 0.01

$$\text{Score} \approx 2^{-6643}$$

Impossible to represent score via floating point

Maneuvering Scores: Keep Track of Score-Log

$$\log(\text{Score}[K]) = \sum_{\text{bigram } x \text{ in } M} \log(F[x])$$

Example: 1000 bigrams, and each of frequency ~ 0.01

Score-log ≈ -6643

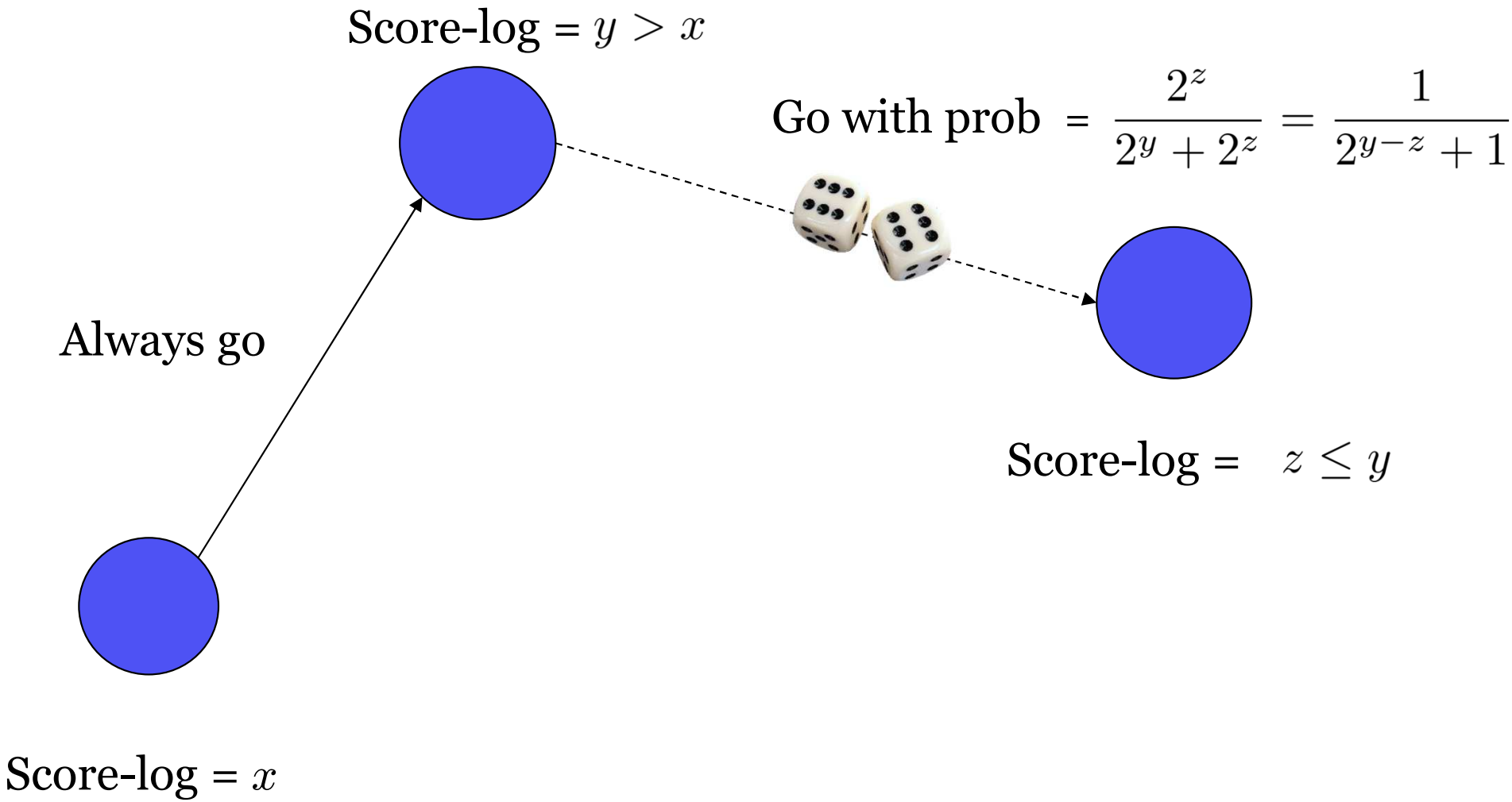
Represent score-log with reasonable accuracy

Maneuvering Scores: Dealing with Zero

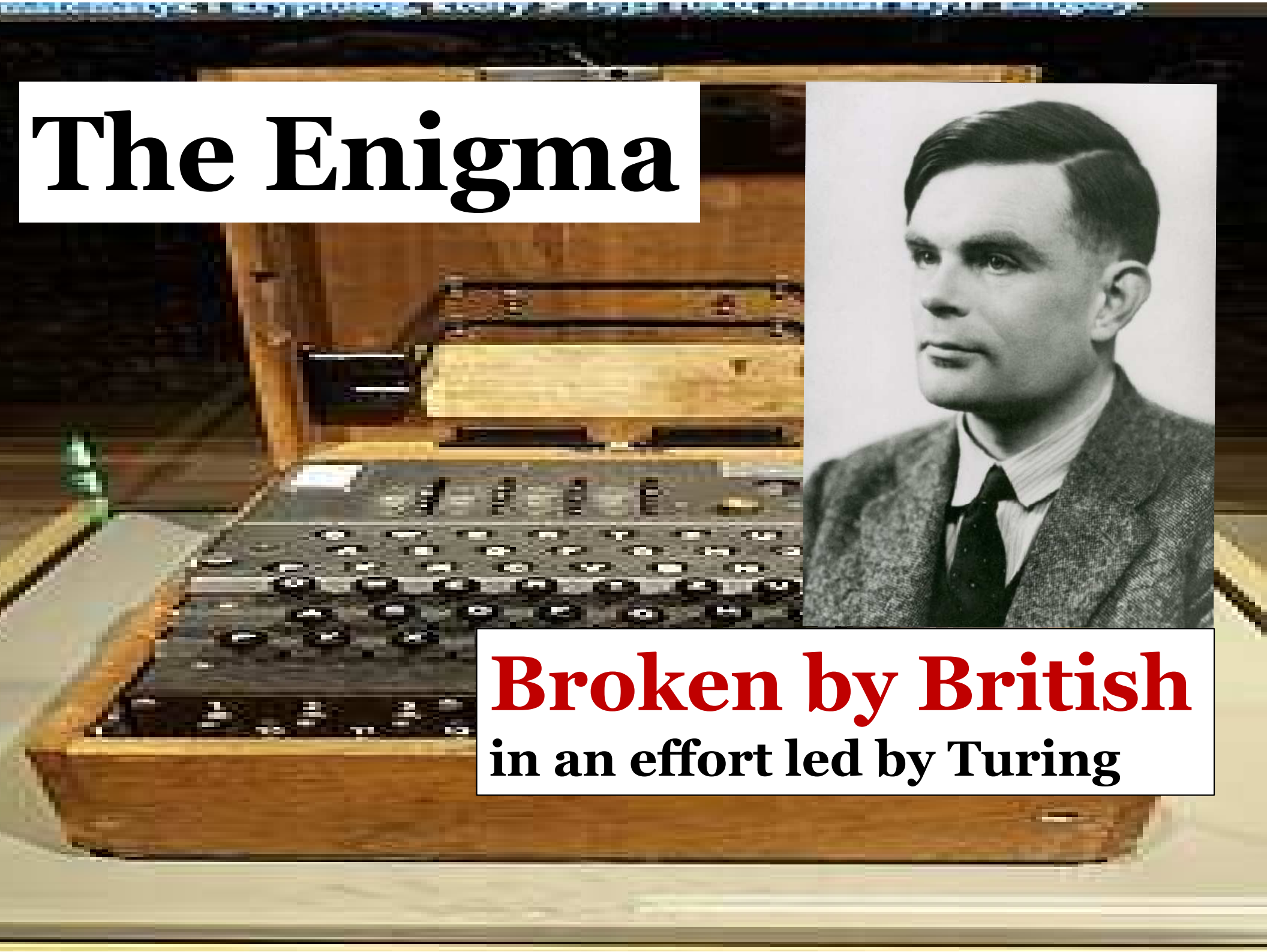
$$F[r] = 0 \quad \Rightarrow \quad \log(F[r]) = ?$$

Use a negative constant of large magnitude

Maneuvering Scores: Graph Traversing



The Enigma



Broken by British

in an effort led by Turing

Agenda

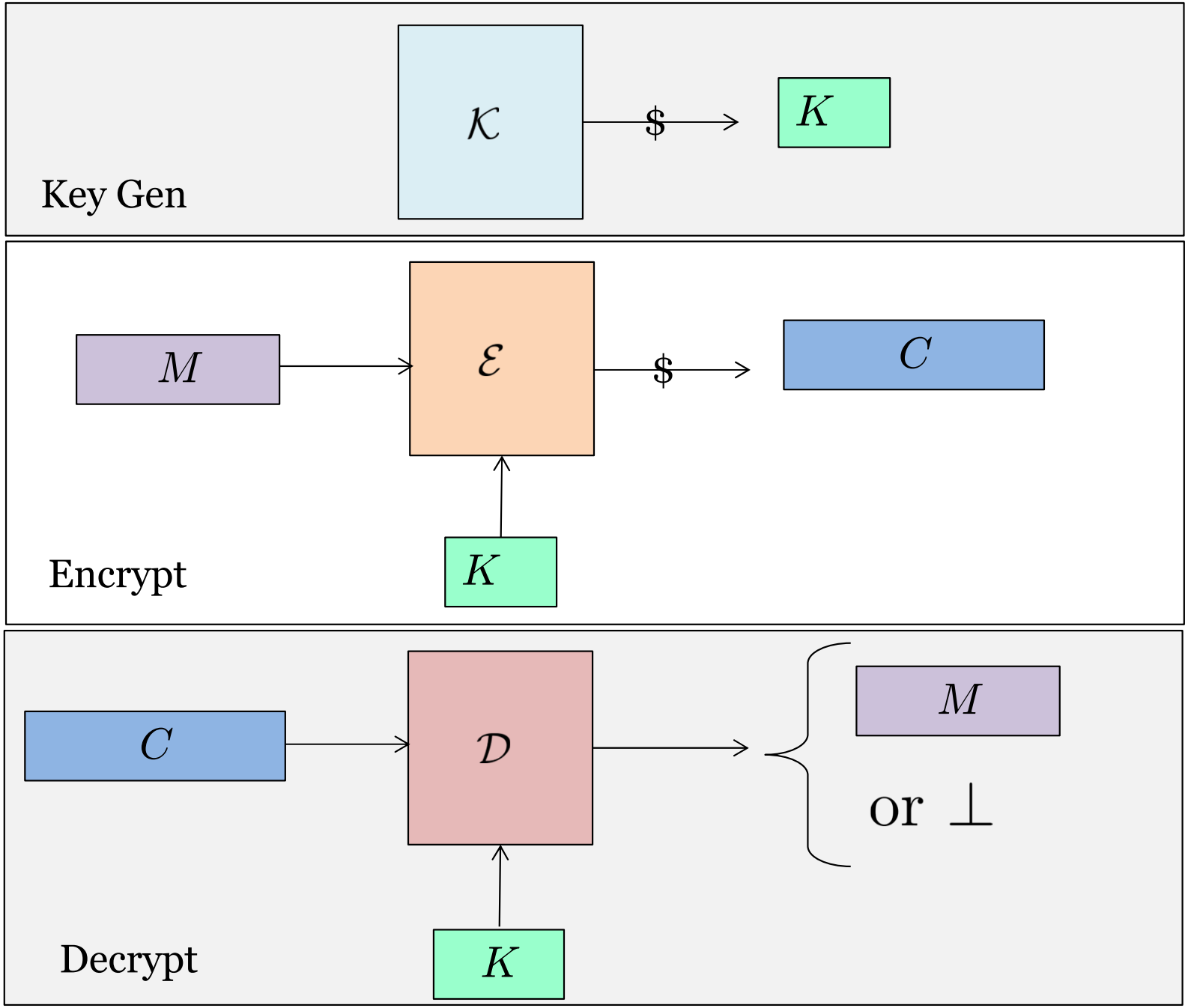
1. Crypto Usage & Goals

2. Classical Crypto

3. One-time Pad & Perfect Secrecy

4. Modern Crypto

Encryption Syntax



Define security?

Perfect Secrecy

Intuition: Ciphertext should reveal **no additional info** about plaintext

For every m and c:

$$\Pr_{K \xleftarrow{\$} \mathcal{K}}[\mathrm{Msg} = m \mid \mathcal{E}_K(\mathrm{Msg}) = c] = \Pr[\mathrm{Msg} = m]$$

Common case: Ciphertext is uniformly random, independent of msg

An Example

$

Msg

{cat, dog, hot}

Ctx

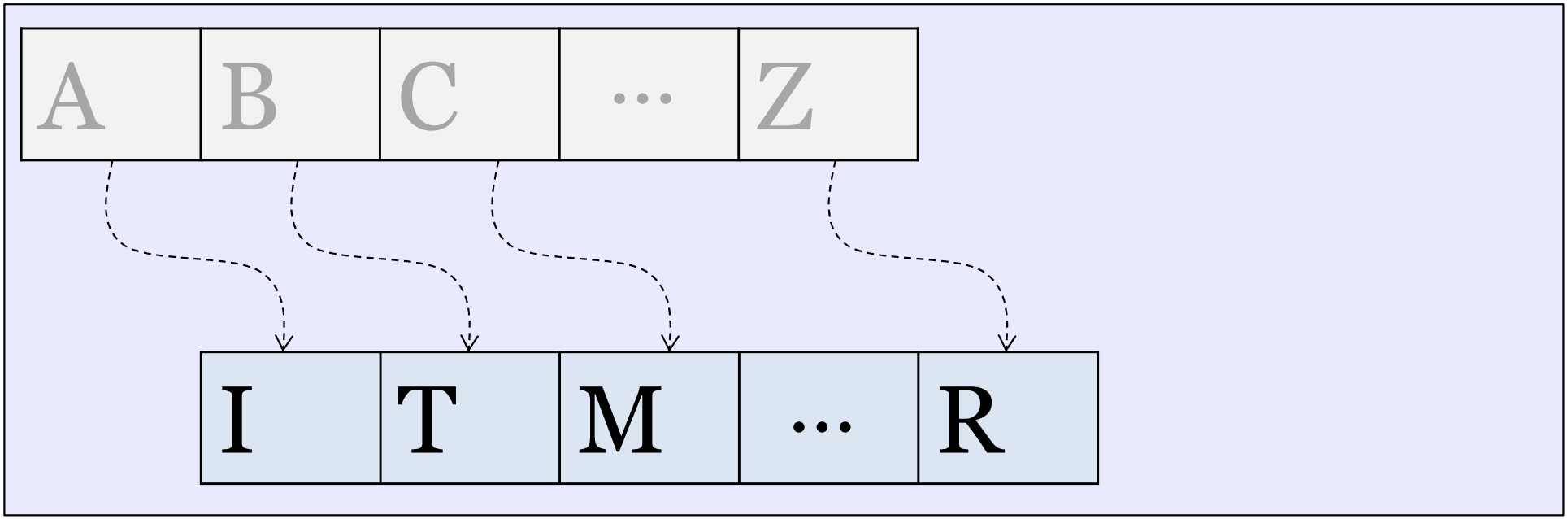
cracked

Msg = *o*

Not perfectly secure

$$\Pr[\text{Msg} = \text{dog} \mid \text{Ctx}] = \frac{1}{2} \neq \Pr[\text{Msg} = \text{dog}] = \frac{1}{3}$$

Substitution Cipher Is Not Perfectly Secret



Msg $ **{hot day, bad boy}**

Ctx = “tic tok”

$$\Pr[\text{Msg} = \text{bad boy} \mid \text{Ctx}] = 1 \neq \Pr[\text{Msg} = \text{bad boy}] = 1/2$$

An Application of Perfect Secrecy

Bob's Privacy in the 5-card Trick for The Dating Problem

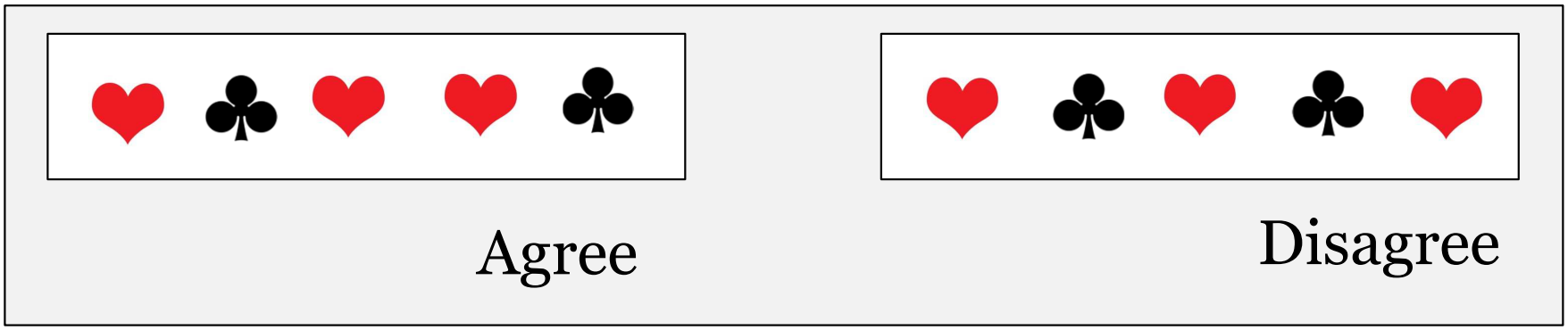
Disagree

Otherwise, Bob's privacy is moot

Assume that Alice disagrees

Want: prove 5-card trick reveals nothing about Bob's input

5-card Trick As Encryption

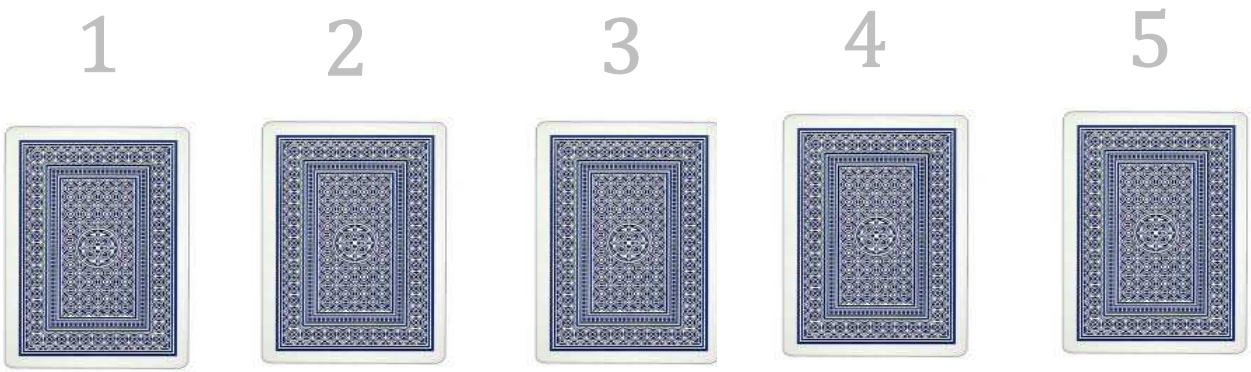


Initial configuration = Bob's message

Assume nothing about message distribution

Cutting as Cyclic Shift

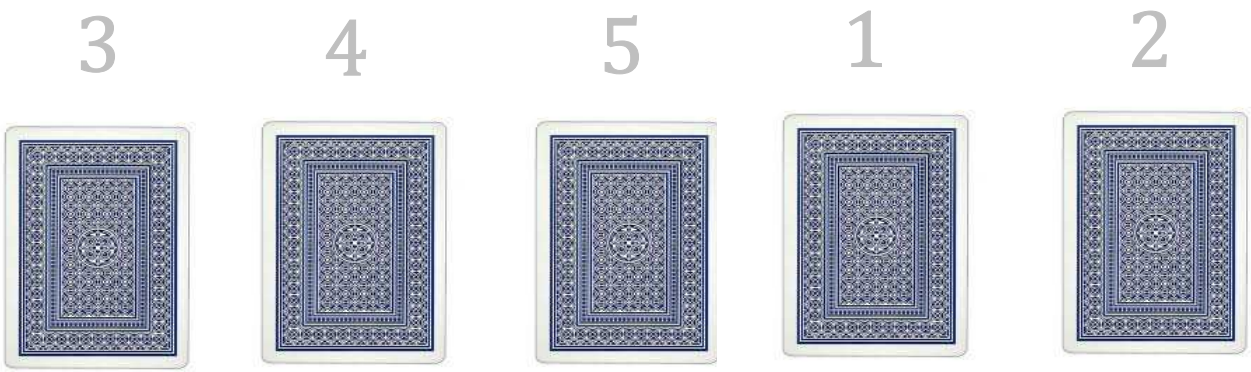
Alice's cut



Alice's cyclic shift by $\underbrace{a \in \{0, 1, 2, 3, 4\}}_{\textbf{known by Alice}}$ positions

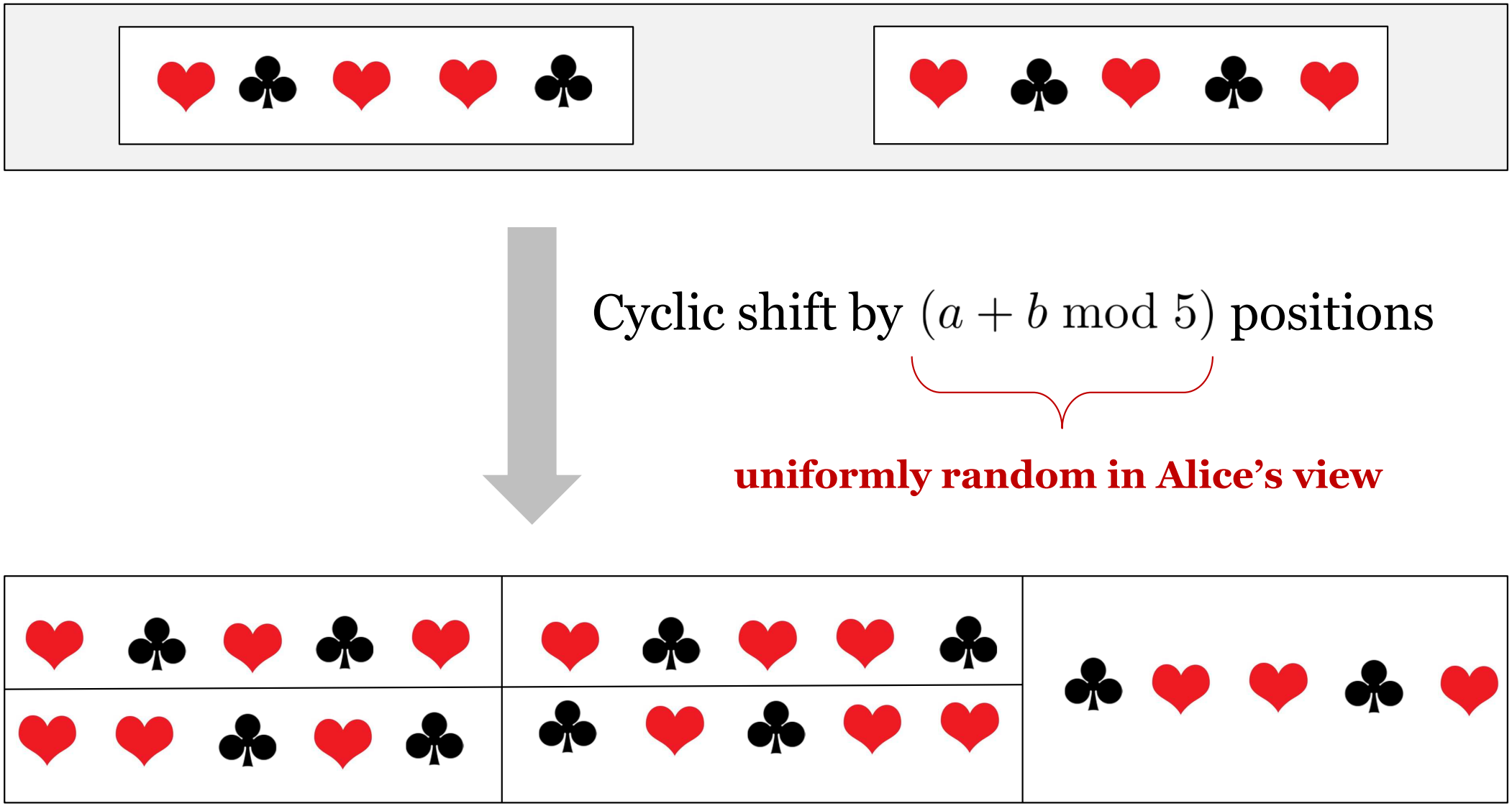
Cutting as Cyclic Shift

Bob's cut



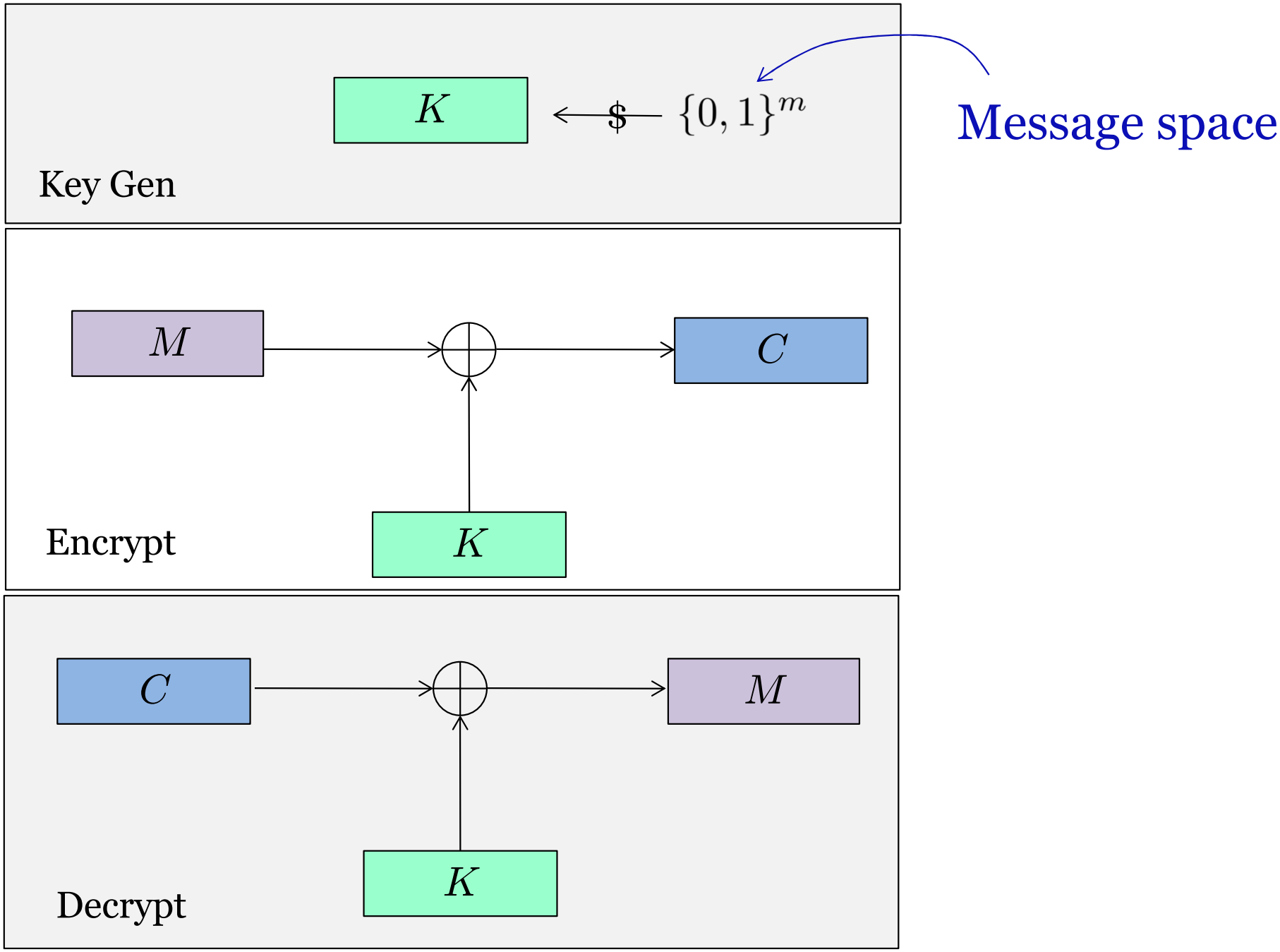
Bob's cyclic shift by $\underbrace{b \leftarrow\$ \{0, 1, 2, 3, 4\}}_{\textbf{uniformly random in Alice's view}}$ positions

Composition of Cuts



Ctx is uniformly random, independent of msg → **Perfect secrecy**

Achieving Perfect Secrecy: One-time Pad



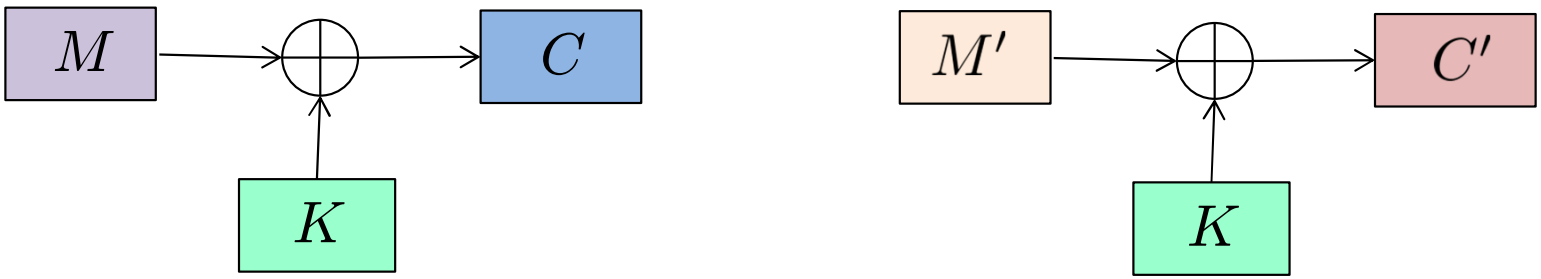
Behind Every Notion, There Is An Assumption

For every m and c:

$$\Pr_{K \xleftarrow{\$} \mathcal{K}}[\mathrm{Msg} = m \mid \mathcal{E}_K(\mathrm{Msg}) = c] = \Pr[\mathrm{Msg} = m]$$

It's **assumed** that you pick a fresh key for each encryption

Reusing One-time Pad Breaks Security



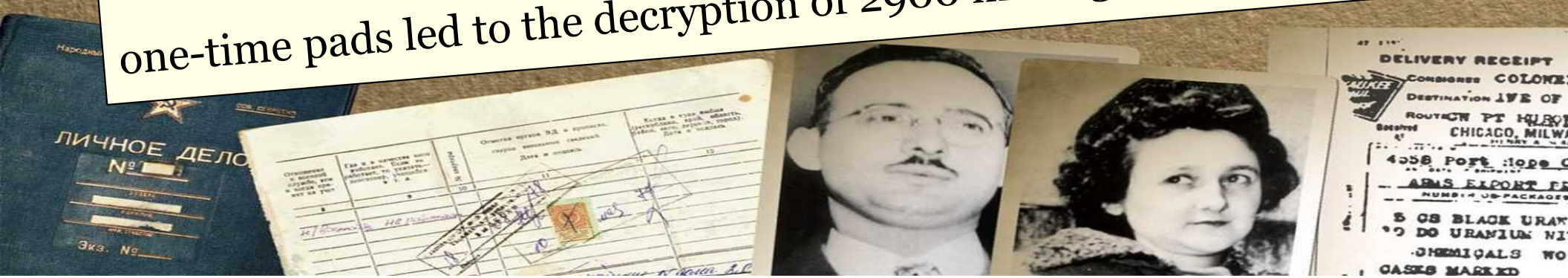
One can obtain $M \oplus M'$ via $C \oplus C'$

Can recover both M and M' if the messages are English texts and long enough

THE VENONA SECRETS

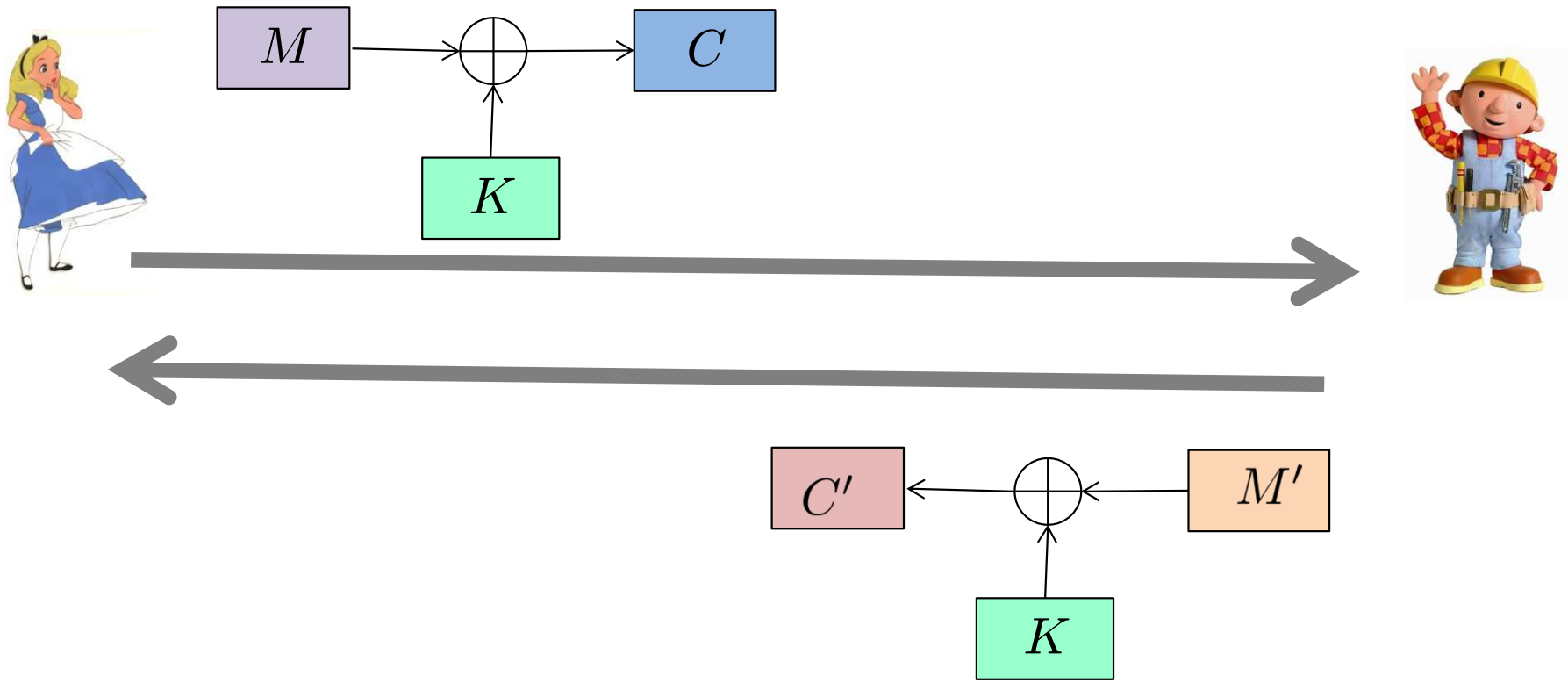
The Definitive Exposé of Soviet Espionage in America

Bad Usage of One-time Pad: USSR's reusing of one-time pads led to the decryption of 2900 messages.



Bad Usage of One-time Pads:

PPTP protocol in Windows NT



Nov 26, 2019

FORTINET PRODUCTS USED HARDCODED ENCRYPTION KEY

By Dennis Fisher

Fortinet's blunder led them to reuse a one-time pad several times

Limitation of Perfect Secrecy

If $|\mathcal{M}| > |\mathcal{K}|$ then **no** scheme is perfectly secret

Impractical

Agenda

1. Crypto Usage & Goals

2. Classical Crypto

3. One-time Pad & Perfect Secrecy

4. Modern Crypto

Lego Approach

Computational Science

Modern Crypto

Provable Security

Modern Crypto: A Lego Approach

Primitives: AES SHA-2 Factoring ...

Transformers

Applications: Encryption MAC Digital Signature

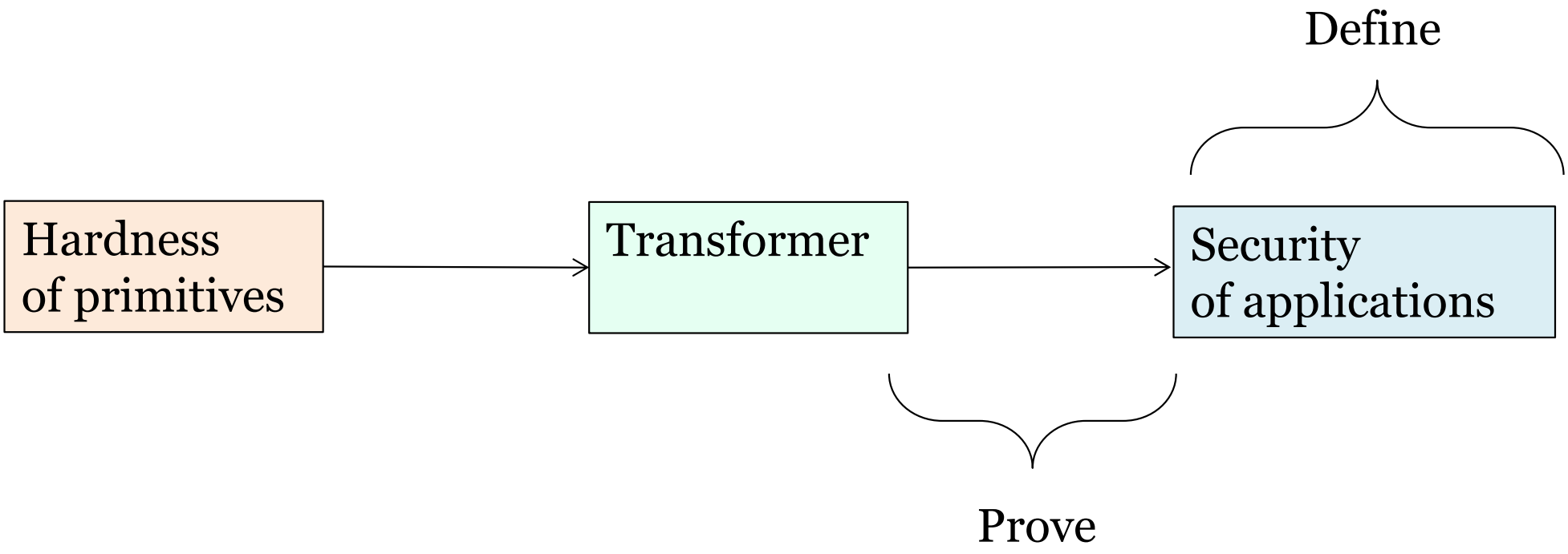
Modern Crypto: A Computational Science

- Assume **computational** hardness of **a few** primitives

AES | SHA-2 | Factoring | ...

- Confidence by cryptanalysis

Modern Crypto: Provable Security

-**Define** security notions for applications

-**Prove** the transformer meets the notions