Fondamenti di Cybersecurity

Introduction to Cryptography:

- History of Cryptography, Basic Ciphers, Rotor machines, Enigma
- One-time pad, Stream Ciphers (classic and real world) and Pseudo Random Generators
- Secret key cryptographic systems, Public key cryptographic systems
- Basics of DES protocols, AES
- Electronic Signatures, Public-key Infrastructure, Certificates and Certificate Authorities
- Sharing of secrets; User authentication; Passwords

What is cryptography?

- **Cryptography**
  - The art and science of using mathematics to obscure the meaning of data by applying transformations to the data that are impractical or impossible to reverse without the knowledge of some key.
  - The term comes from the Greek for “hidden writing”
  - *Kryptós*: hidden
  - *Graphía: writing*

- **Cryptanalysis**
  - The art and science of breaking encryption/secret codes/secret messages (recovering plaintext from ciphertext when the key is unkown).

- **Cryptology: Cryptography + Cryptanalysis**

Cryptography is everywhere

Secure communication:
- web traffic: HTTPS
- wireless traffic: Wireless/cellular Networks

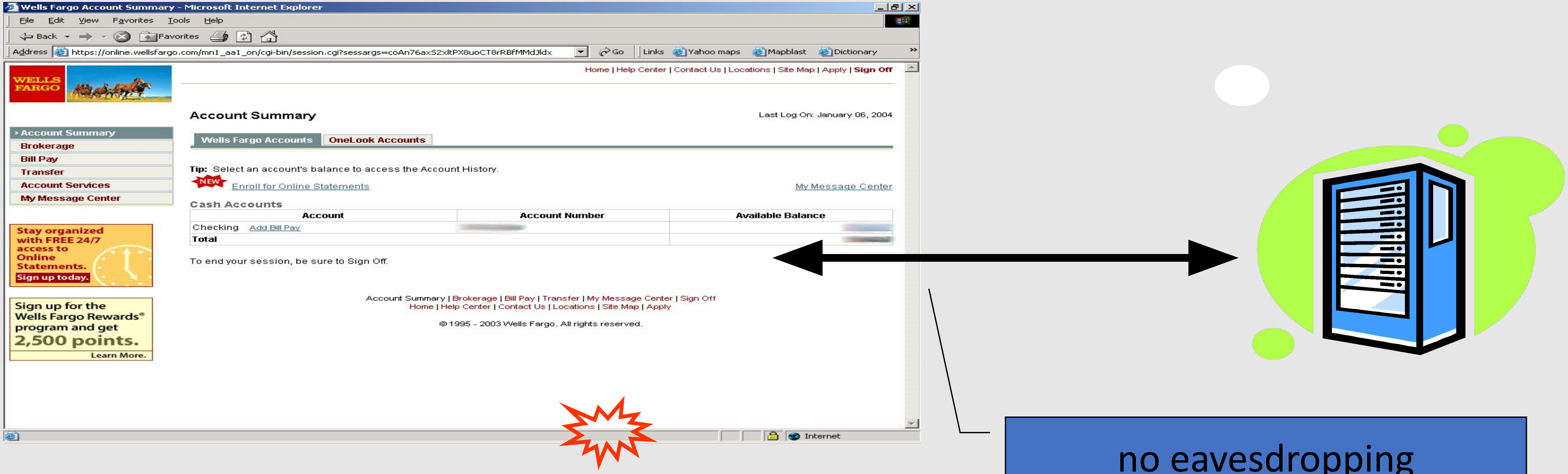
Encrypting files on disk

Content protection (e.g., DVD, Blu-ray)

User authentication

… and much much more (more “magical” applications later…)

Secure communication



no eavesdropping
no tampering

Approaches to secure communication

Steganography:

“covered writing”

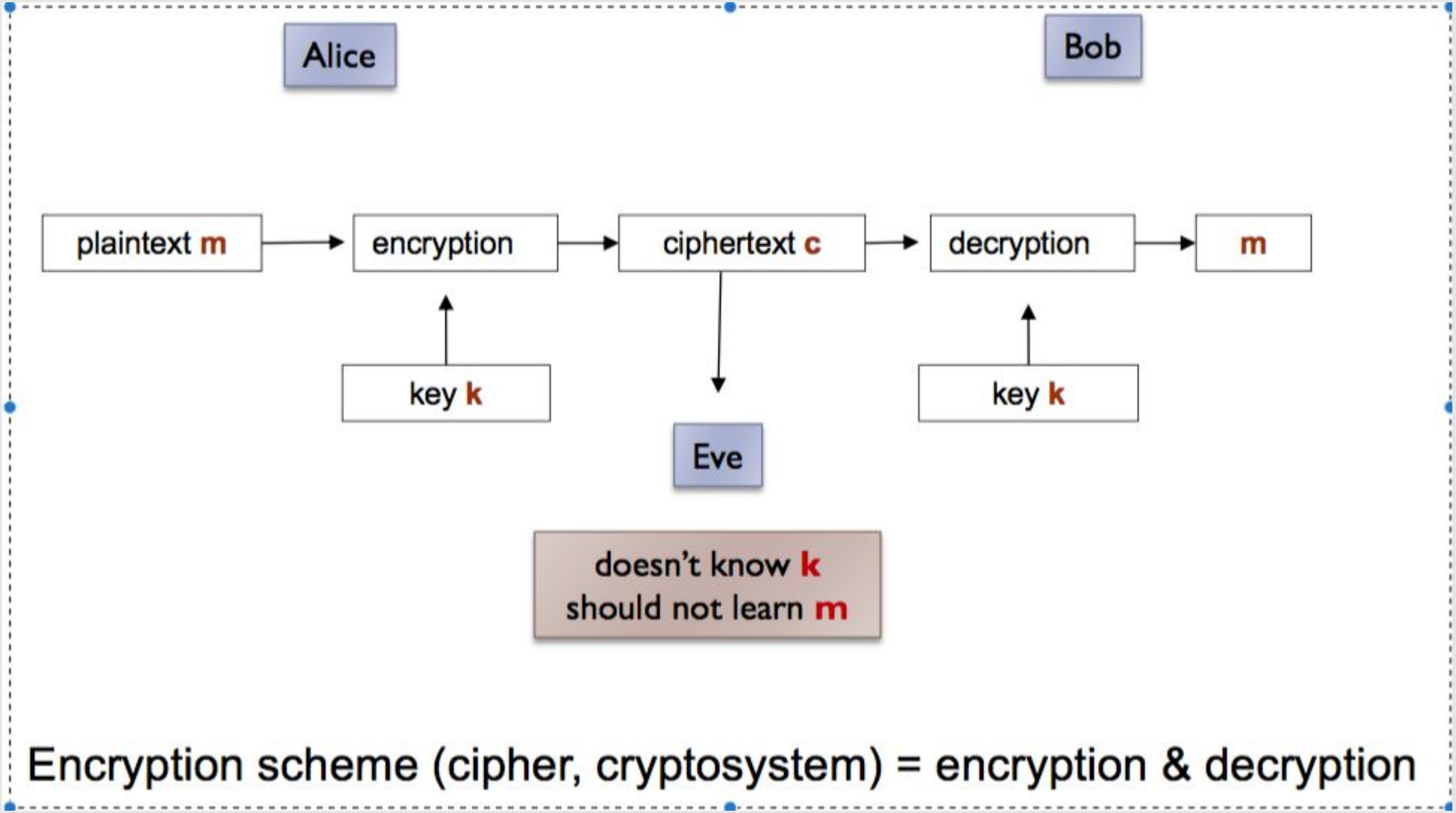
hides the existence of a message

Cryptography:

“hidden writing”

hide the meaning of a message

Encryption terminology



https://cnitarot.github.io/courses/fc_Fall_2022/index.html

Goals and objectives

Objectives: Ensure security of communication between parties over an insecure medium

Basic security goals:

- **privacy (secrecy, confidentiality)**: only the intended recipient can see the communication
- **authenticity**: the communication is generated by the alleged sender
- **Integrity**: no unauthorized modifications to messages
- **Non-repudiation**: no disclaiming of authorship

Cryptographic protocols

- Protocols that
  - Enable parties to … communicate securely
  - Achieve goals to … protect message confidentiality and integrity
  - In an environment where boundaries and interaction with it are well defined
  - Overcome adversaries

- Need to understand
  - Who are the parties and the context in which they act?
  - What are the security goals of the protocols?
  - What is the trusted computing base, i.e. what is trusted
  - What are the capabilities of the adversaries? Threat model

Kerckhoff‘s principle

The security of a protocol should rely only on the secrecy of the keys, while protocol designs should be made public (1883)

– <u>security by obscurity does not work</u>

(there are many examples, WEP, voting machines…)

Auguste Kerckhoffs (19 January 1835 – 9 August 1903) was a Dutch linguist and cryptographer who was professor of languages at the School of Higher Commercial Studies in Paris in the late 19th century.

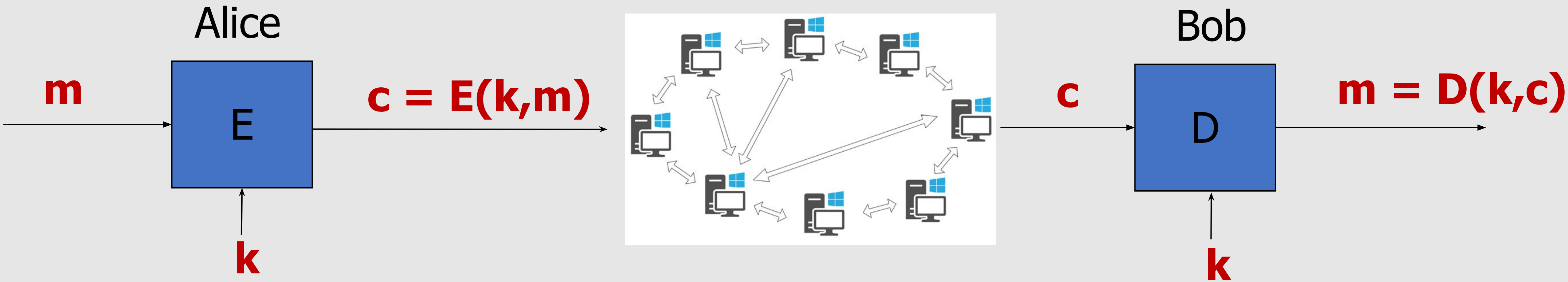
Attacker threat model (1/2)

- Knowledge about the cipher (cryptosystem)
  - **Kerchhoff's Principle**
    - A cryptosystem should be secure even if everything about the system, except the key, is public knowledge
  - Attacker is assumed to have full knowledge of the chosen cryptographic algorithm; ***No security through obscurity***
- Interaction with messages and the protocol
  - **Passive**: only observes and attempts to decrypt messages
    - Only threatens confidentiality
  - **Active**: observes, modifies, injects, or deletes messages
    - Threatens confidentiality, integrity, and authenticity

Attacker threat model (2/2)

- Interaction with the encryption algorithm
  - Ciphertext-only attack: attacker only sees encrypted messages
  - Chosen-plaintext attack (CPA): Attacker may choose a number of messages and obtain the ciphertexts for them
  - Chosen-ciphertext attack (CCA): Attacker may choose a number of ciphertexts and obtain the plaintexts
  - Both CPA and CCA attacks may be adaptive: Choices may change based on results of previous requests
- Resources available (storage and/or ***computation***)
  - Unlimited resources
  - Finite resources – Computational security
    - to calculate, typically polynomial running time
    - to store things

Symmetric Encryption (confidentiality)



- **k**: **secret key (A SHARED SECRET KEY)**
- **m**: **plaintext**
- **c**: **ciphertext**
- **E**: **Encryption algorithm**
- **D**: **Decryption algorithm**
- **E, D**: **Cipher**

- **Confidentiality scenario**
- Other scenarios are possible, with the secret key used differently…
  - e.g., **MACs** (for **integrity**)

Algorithms are publicly known, never use a proprietary cipher

Use Cases

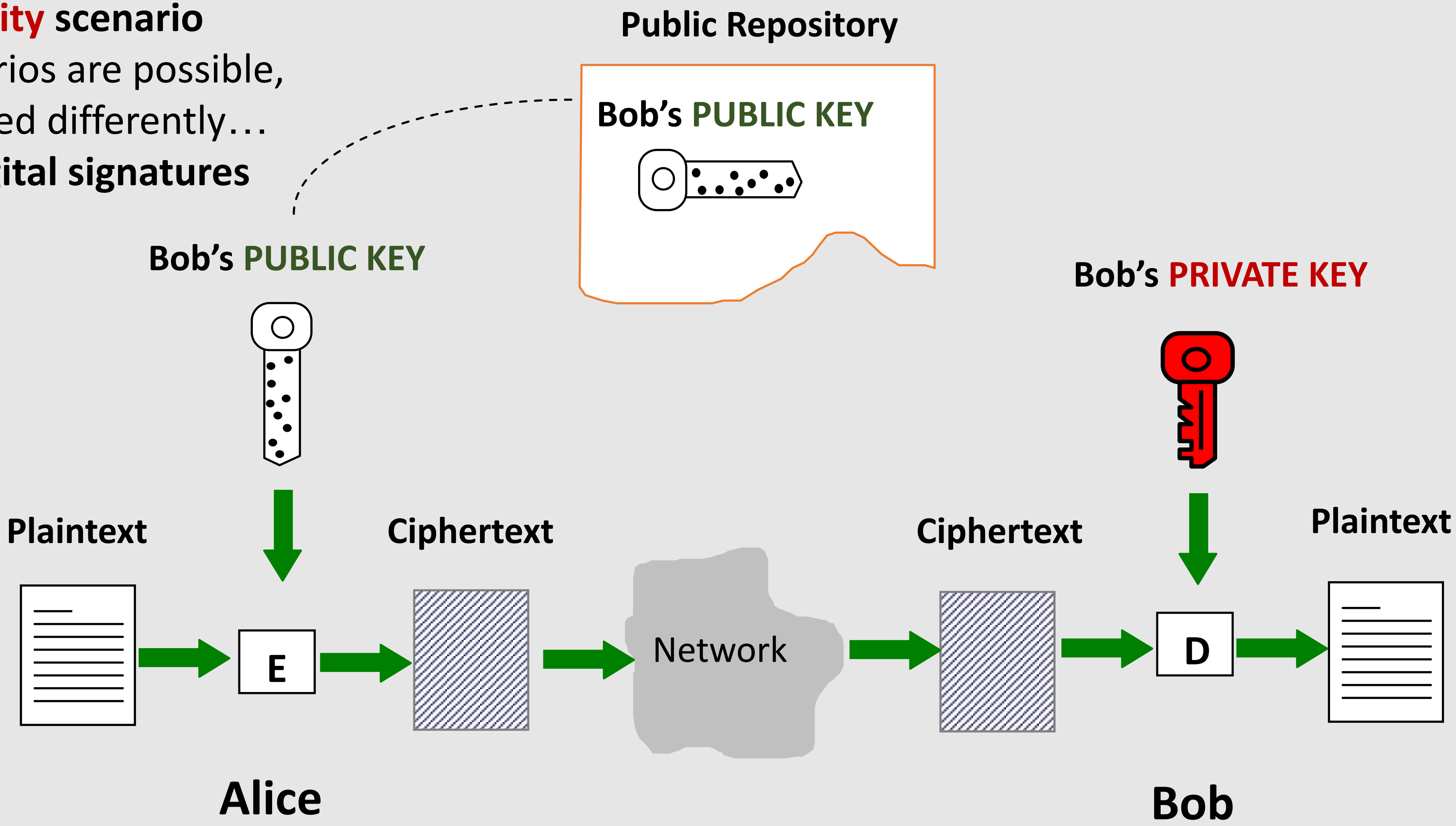
- **Single-use key**: (or **one-time key**):
  Key is only used to encrypt **one and only one message**
  - encrypted email: new key generated for every email

- **Multi-use key**: (or **many-time key**):
  Same key used to encrypt **multiple messages**
  - encrypted files: same key used to encrypt many files

  Need more machinery than for one-time key

Asymmetric Encryption

- **Confidentiality scenario**
- Other scenarios are possible, with keys used differently…
  - e.g., **Digital signatures**



Things to remember

Cryptography is:

- A tremendous tool
- The basis for many security mechanisms

Cryptography is **not**:

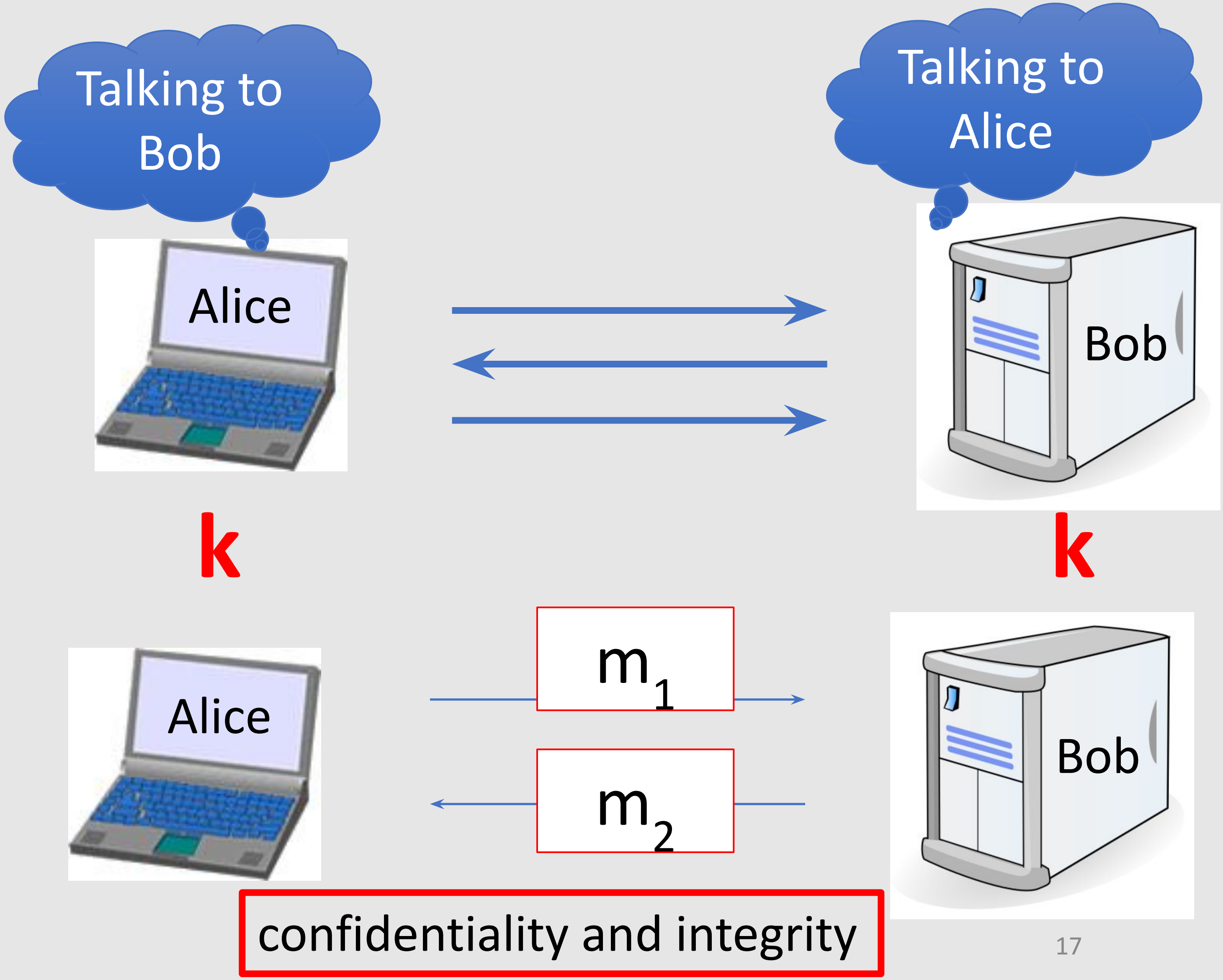
- The solution to all security problems
- Reliable unless implemented and used properly
- Something you should try to invent yourself
  - many many examples of broken ad-hoc designs

Some Applications

Secure communication

1. <u>Secret key</u> establishment:

2. Secure communication:



But crypto can do much more

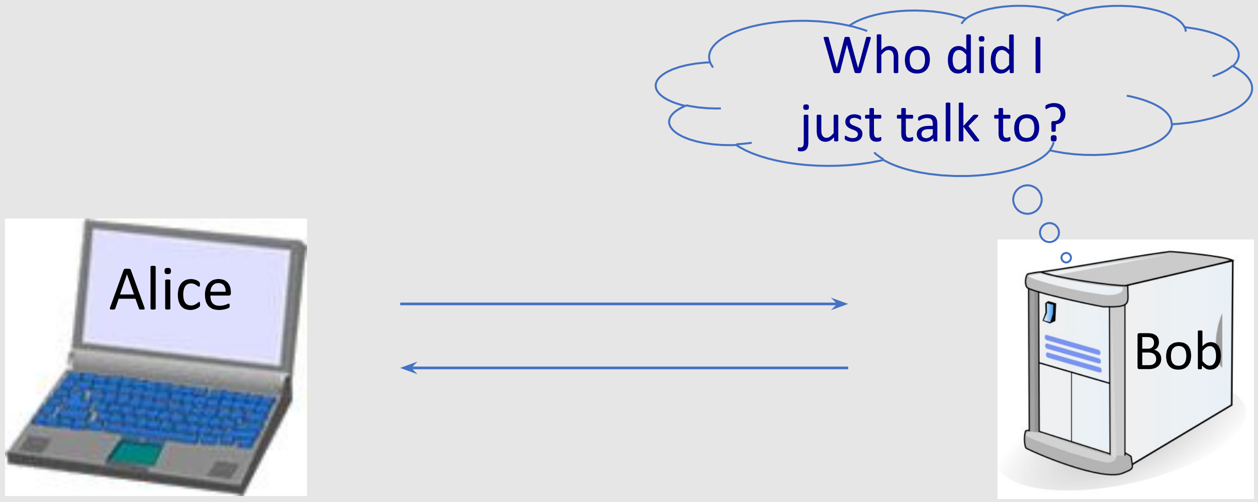
- Digital signatures

- Signatures of the same person change over different documents

- Asymmetric Cryptography is used

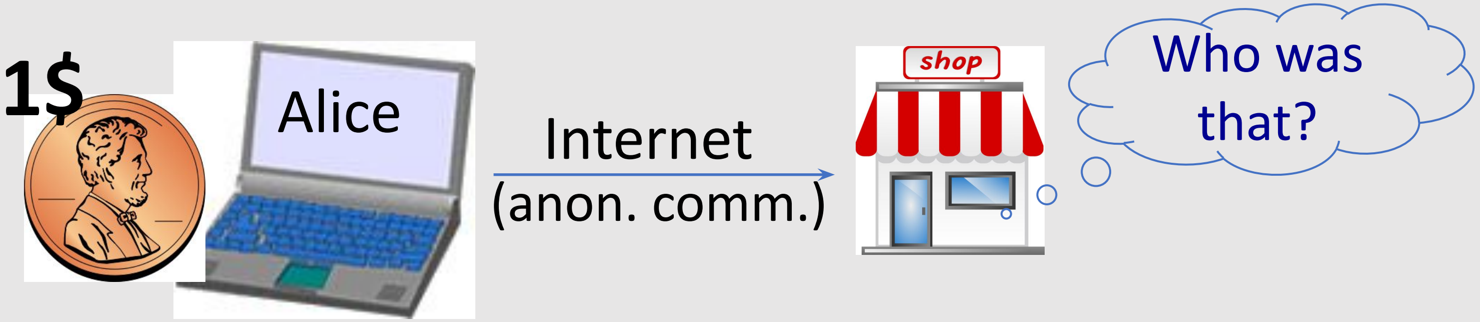
But crypto can do much more

- Anonymous communication
  (e.g., mix networks)



But crypto can do much more

- Anonymous **digital** cash
  - Can I spend a “digital coin” without anyone knowing who I am?
  - How to prevent double spending?

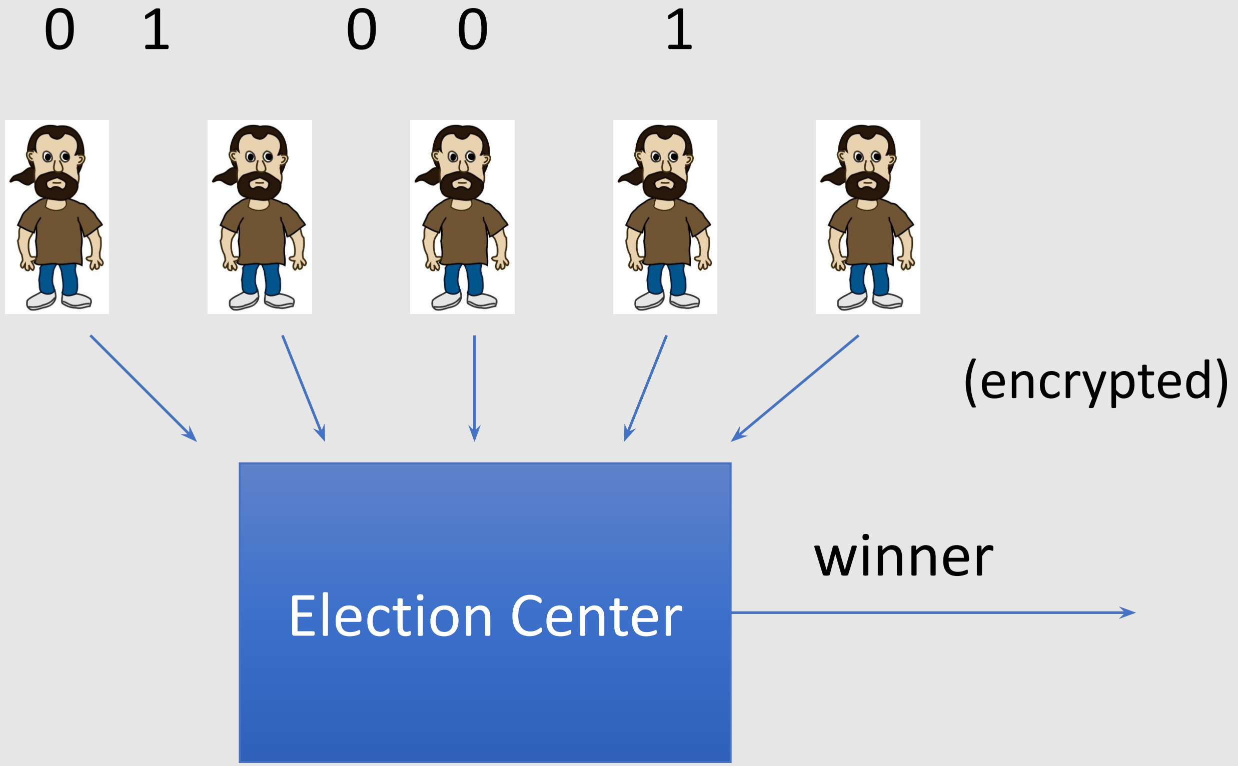


Protocols

- Elections
- Private auctions

winner= majority [votes]

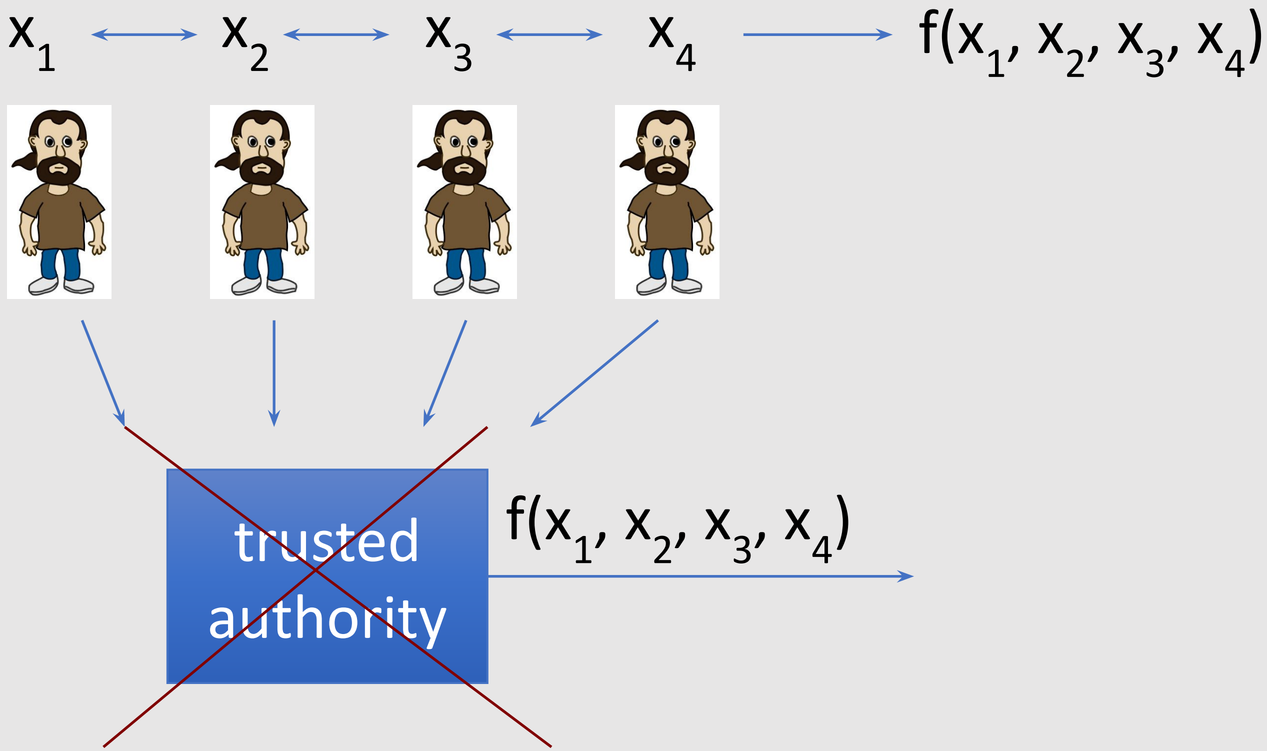
(Vickrey Auction)
Auction winner = highest bidder
pays 2^{nd} highest bid



Election Center must determine the winner
without knowing the individual votes!

Protocols

- Elections
- Private auctions



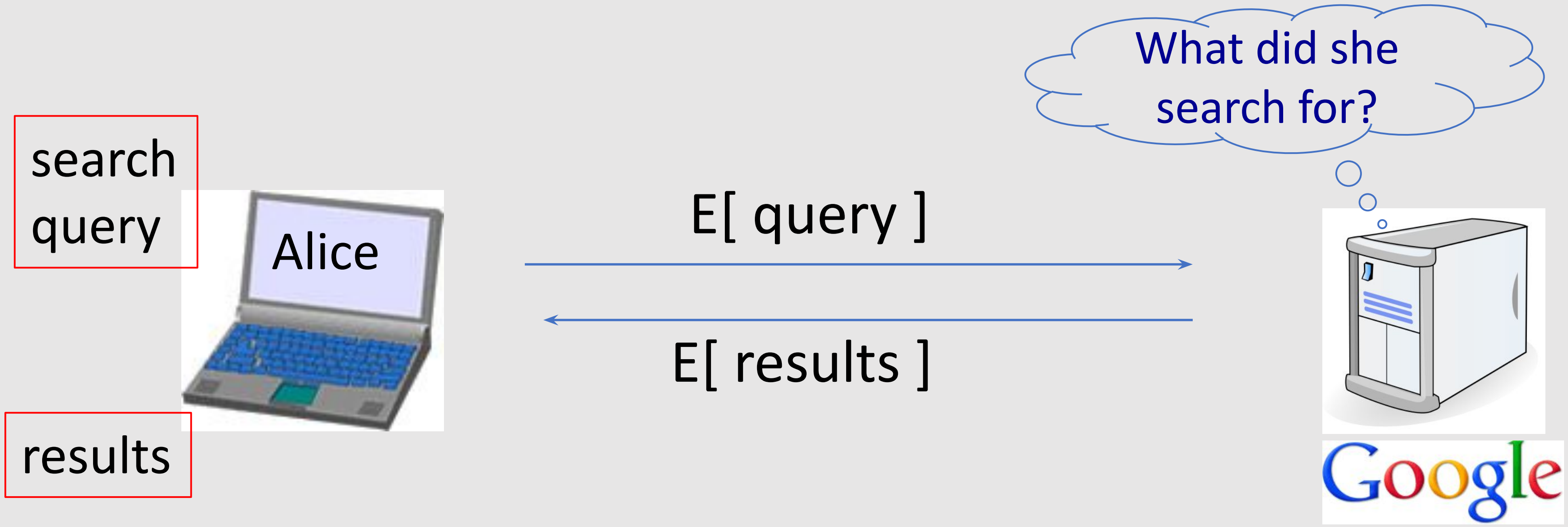
Secure multi-party computation

Goal: compute $f(x_1, x_2, x_3, x_4)$

"Thm:" anything that can be done with trusted auth. can also be done without

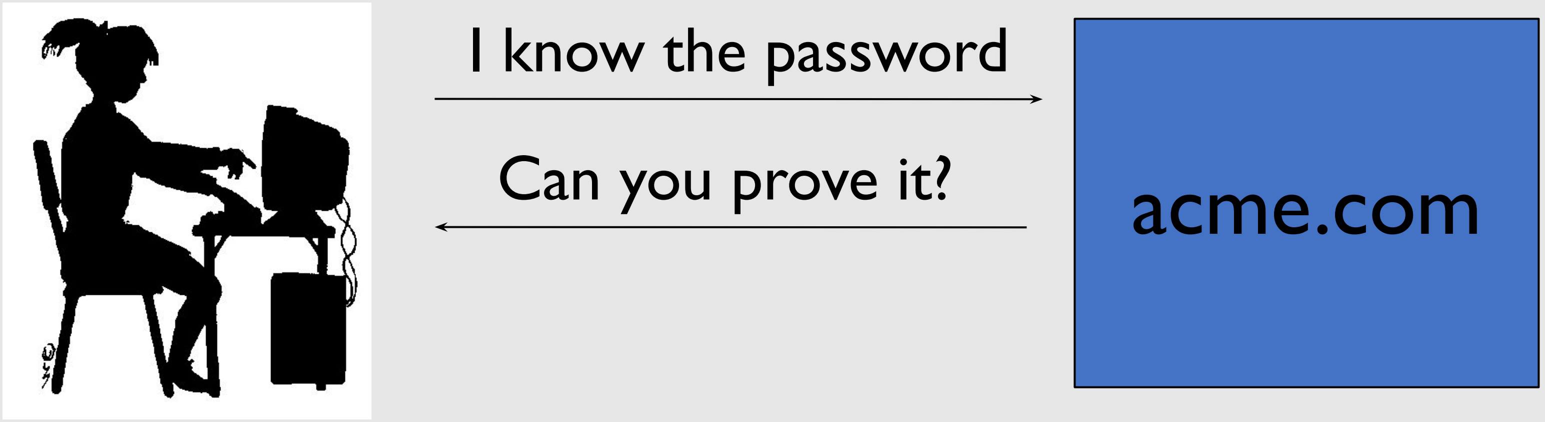
Crypto magic

- Privately outsourcing computation



Crypto magic

- Zero knowledge (proof of knowledge)



A rigorous science

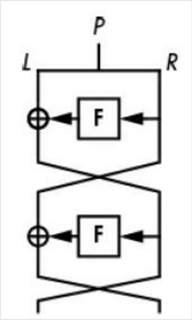
The three steps in cryptography:

- Precisely specify the threat model
- Propose a construction
- Prove that breaking construction under the threat model will solve an underlying hard problem

Brief History of Crypto

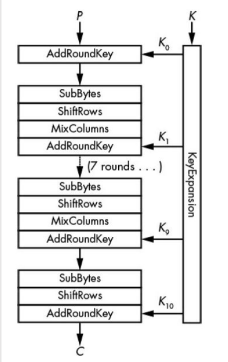
Che cos’è la Crittografia?

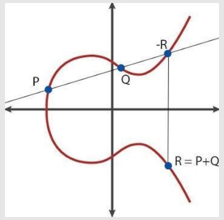
- Metodi per **memorizzare**, **elaborare** e **trasmettere** informazioni in maniera **sicura** in presenza di agenti ostili
- **Crittografia:** *Kryptós*: nascosto + *Graphía:* scrittura



$n = p \times q$

$p, q?$

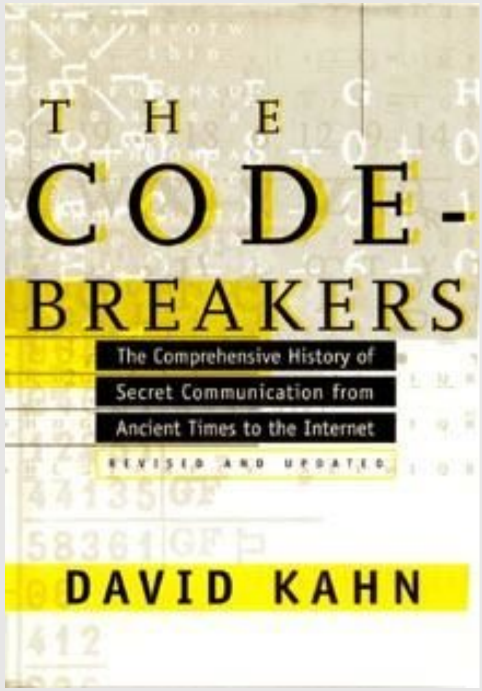




Scitala	Cifrario di Cesare	Enigma	DES	RSA	AES	Crittografia ellittica
400 aC	50 aC	1918	1975	1977	2001	2005

History

David Kahn, "The code breakers" (1996)

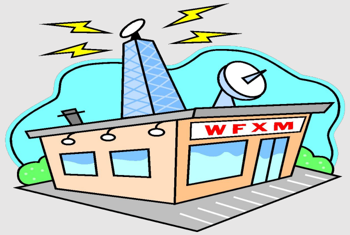


Symmetric Ciphers

Alice

Bob

m → E → $E(k,m) = \mathbf{c}$



c → D → $D(k,c) = \mathbf{m}$

k

k

Same key

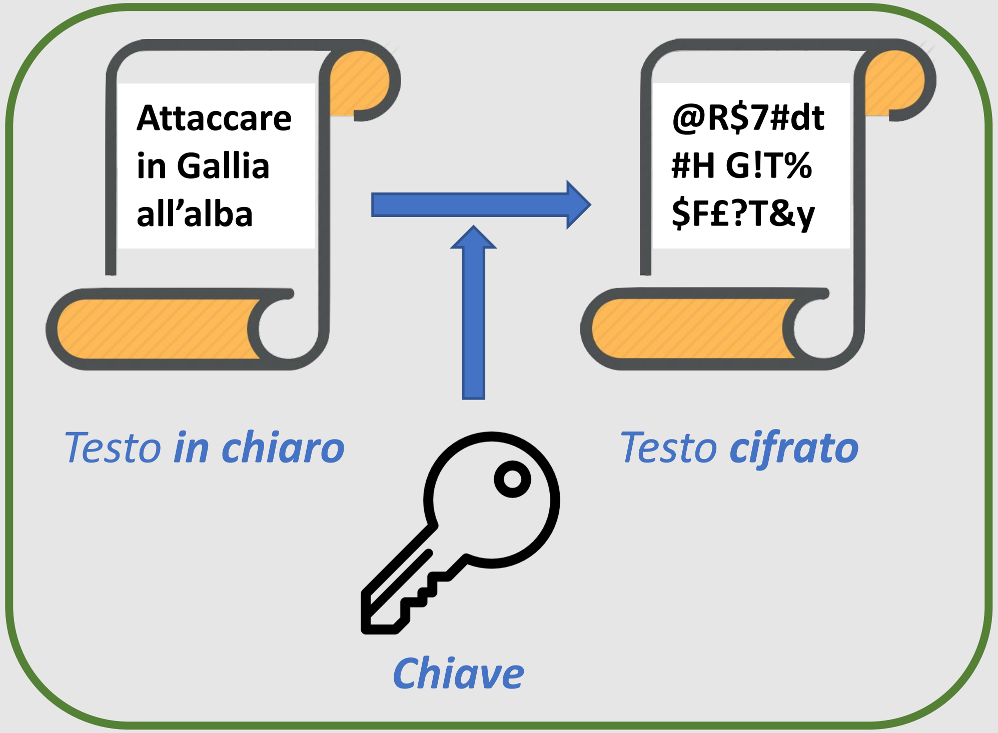
Cypher: (E, D)

Un classico scenario

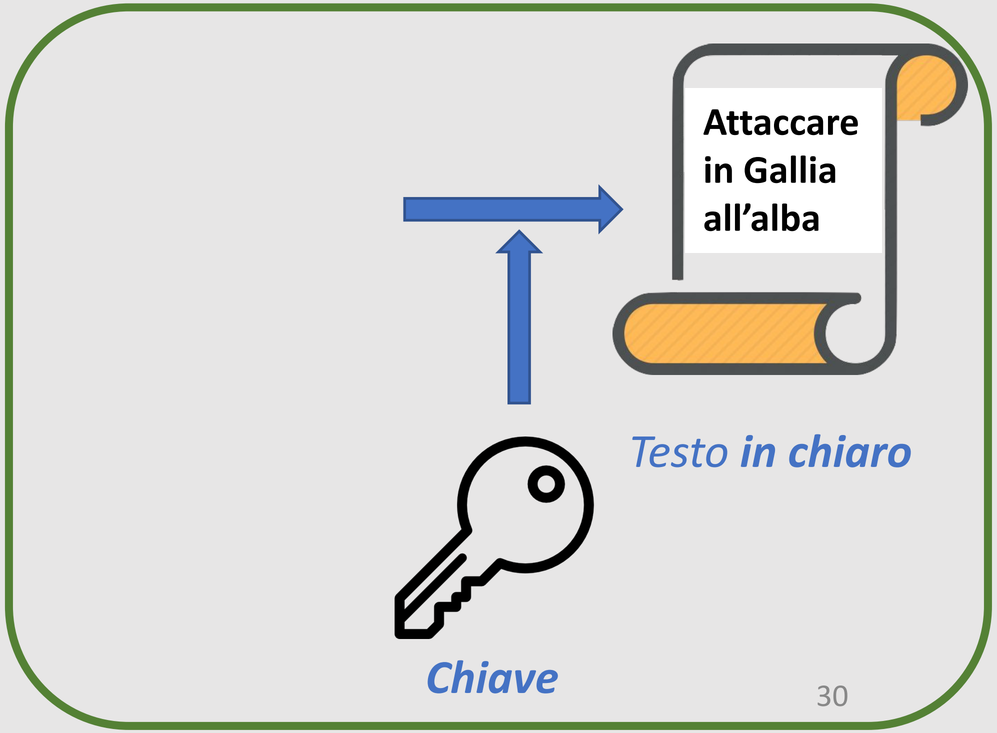
*Algoritmi di cifratura e decifratura: **pubblici***

Crittografia **simmetrica** e **asimmetrica**

Cifratura



Decifratura



Cifrario di Cesare

Chiave	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
	D	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**	**A**	**B**	**C**

Attaccare in Gallia all’alba

*Testo **in chiaro***

Dwwdffduh lq Jdoold doo’doed

*Testo **cifrato***

(Cifrario a sostituzione)

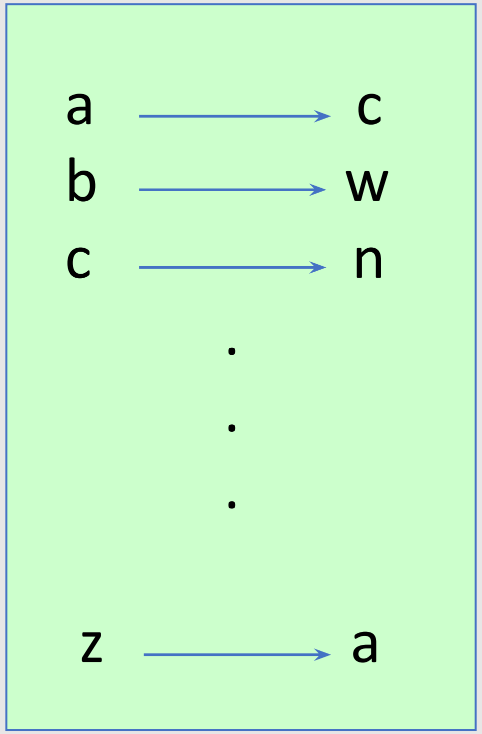
Symmetric substitution cipher

- Key is a **number *k***; (for Caesar shift k=3)
- To encrypt, “shift” each letter by k positions
- To decrypt, “shift” each letter back by k positions

c:= E(k, “bcza”) = “wnac”

D(k,c) = “bcza”

k :=



Caesar Cipher

Shift by 3

$a \rightarrow d$
$b \rightarrow e$
$c \rightarrow f$
.
.
.
$y \rightarrow b$
$z \rightarrow c$

Shift cipher: Mathematical View

- The plaintext P (the messages): words over alphabet {A,...,Z} ≈ {0,...,25}
- The Key Space K: {0,...,25}
- **Encryption** given a key k: each letter in the plaintext P is replaced with the k'th letter following the corresponding number (**shift right**)
- **Decryption** given K: **shift left**

Formally:

Let $P=C=K=Z_{26}$ For $0 \leq k \leq 25$

$$e_k(m) = m+k \mod 26 \text{ and}$$

$$d_k(c) = c-k \mod 26$$

$(m, c \in Z_{26})$

Shift Cipher: an example

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- P = CRYPTOGRAPHYISFUN
- K = 11
- C = NCJAVZRCLASJTDQFY
- C $\longrightarrow$ 2; 2+11 mod 26 = 13 $\longrightarrow$ N
- R $\longrightarrow$ 17; 17+11 mod 26 = 2 $\longrightarrow$ C
- …
- N $\longrightarrow$ 13; 13+11 mod 26 = 24 $\longrightarrow$ Y

Note that punctuation is often eliminated

Security of the Shift Cipher

- Can an attacker find the key k?
  - **YES**: exhaustive search (a **brute force attack**); the key space is small (<= 26 possible keys).
  - The decryption of the ciphertext “makes sense”
  - Once k is found, very easy to decrypt

*Mono*alphabetic Substitution Cipher

- The key space: all possible permutations of $\Sigma = \{A, B, C, \ldots, Z\}$
- Encryption, given a key (permutation) π:
  - each letter X in the plaintext P is replaced with $\pi(X)$
- Decryption, given a key π:
  - each letter Y in the ciphertext C is replaced with $\pi^{-1}(Y)$
- Example

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
π	B	A	D	C	Z	H	W	Y	G	O	Q	X	S	V	T	R	N	M	S	K	J	I	P	E	F	U

BECAUSE → **AZDBJSZ**

What is the size of key space in the monoalphabetic substitution cipher assuming 26 letters?

$$|\mathcal{K}| = 26$$

$$|\mathcal{K}| = 26!$$

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

$$|\mathcal{K}| = 2^{26}$$

$$|\mathcal{K}| = 26^2$$

How to break a substitution cipher?

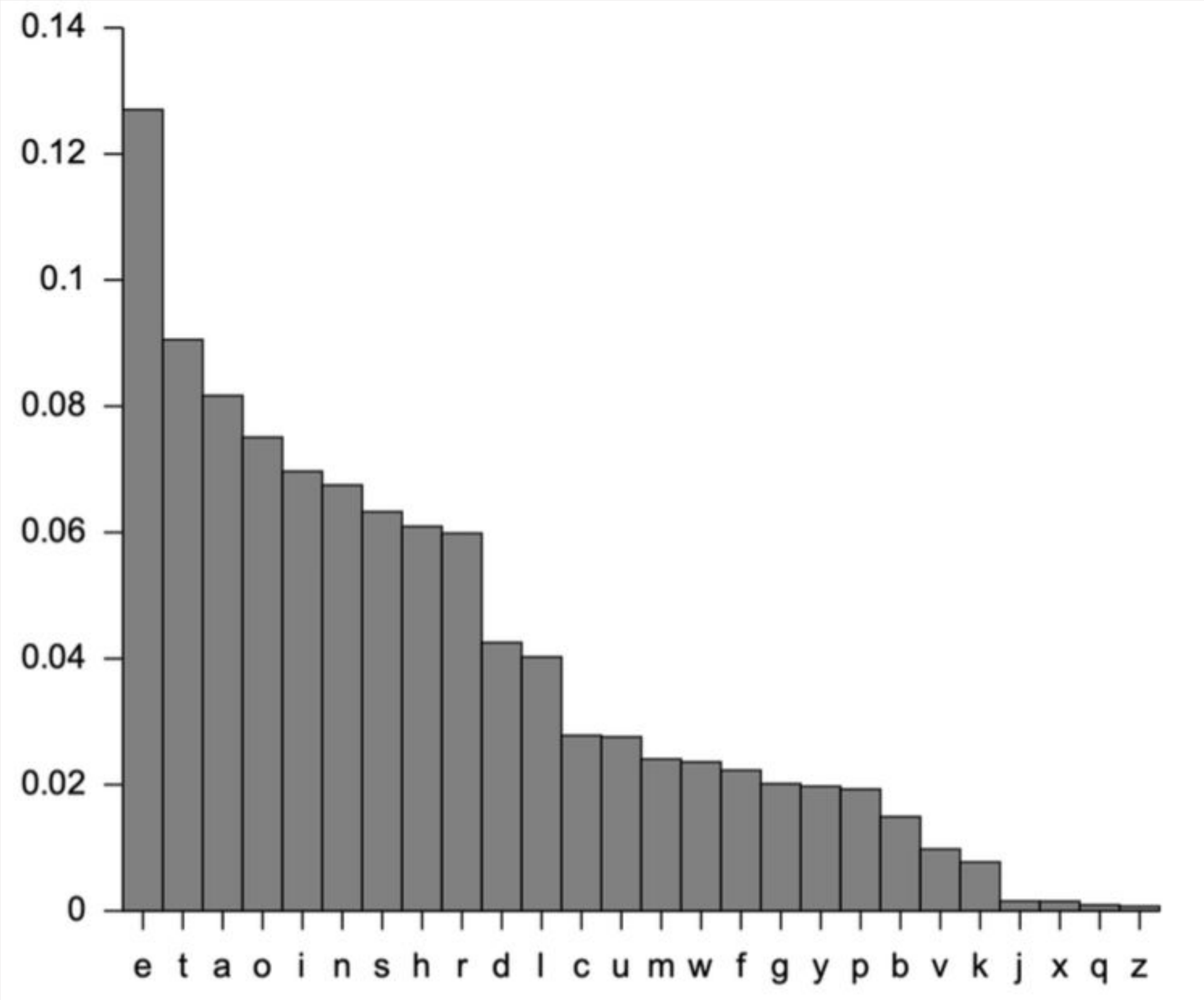
What is the most common letter in English text?

"X"

"L

"E"

"H"



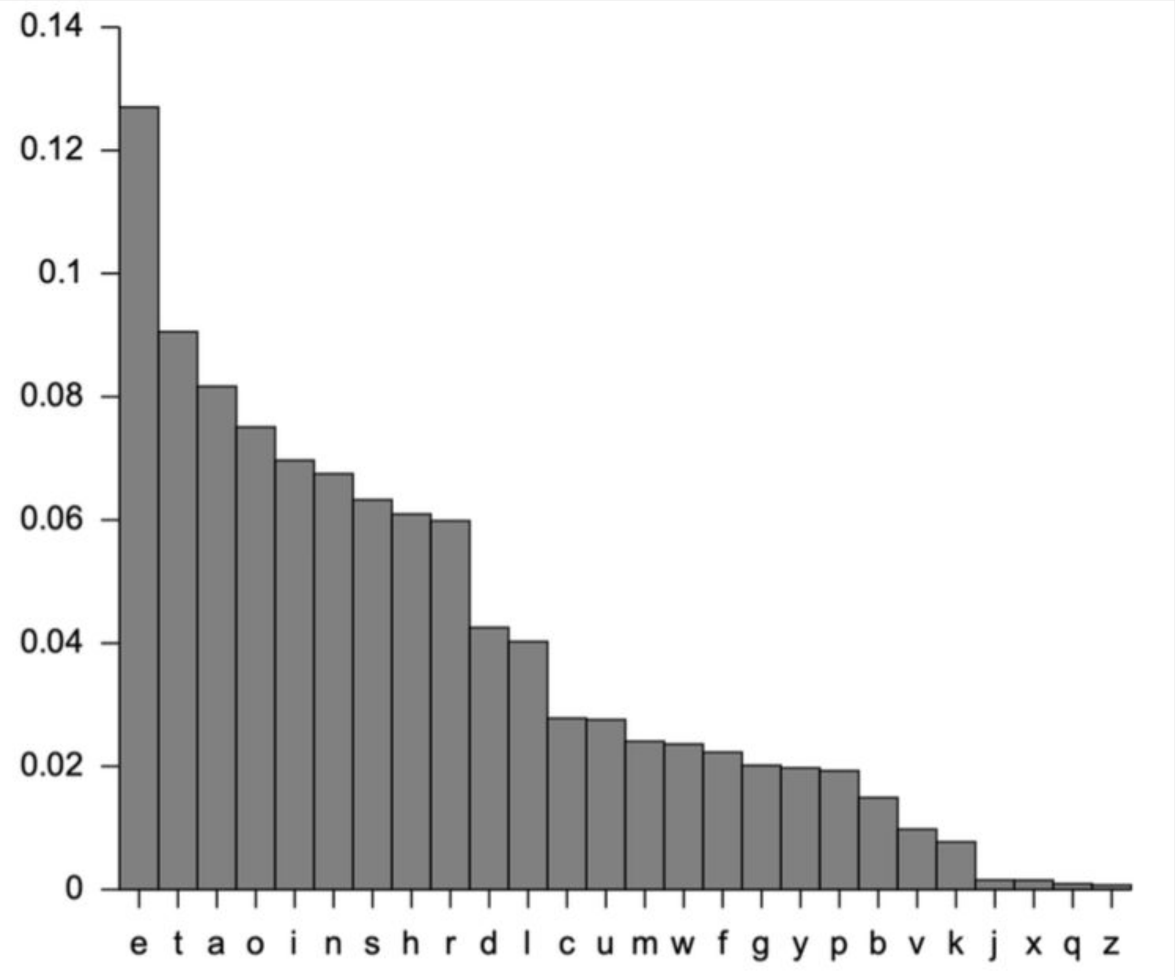
How to break a substitution cipher?

(1) Use frequency of English letters

e: 12,7%

t: 9,1%

a: 8,1%



(2) Use frequency of pairs of letters

(digrams): **he**, **an**, **in**, **th**

How to break a substitution cipher?

Basic ideas:

- Each language has certain features: frequency of letters, or of groups of two or more letters.
- Substitution ciphers preserve the language features.
- *Substitution ciphers are vulnerable to frequency analysis attacks.*
    - Al-Kindi 800 AD

How to break a substitution cipher?

- The number of different ciphertext characters or combinations (digrams, trigrams) are counted to determine the frequency of usage.
- The ciphertext is examined for patterns, repeated series, and common combinations.
- Replace ciphertext characters with possible plaintext equivalents using known language characteristics

An Example

UKBYBIPOUZBCUFEEBORUKBYBHOBBRFESPVKBWFOFERVNBCVBZPRUBOFERVNBCVBPCYYFV
UFOFEIKNWFRFIKJNUPWRFIPOUNVNIPUBRNCUKBEFWWFDNCHXCYBOHOPYXPUBNCUBOYNR
VNIWNCPOJIOFHOPZRVFZIXUBORJRUBZRBCHNCBBONCHRJZSFWNVRJRUBZRPCYZPUKBZPUN
VPWPCYVFZIXUPUNFCPWRVNBCVBRPYYNUNFCPWWJUKBYBIPOUZBCUIPOUNVNIPUBRNCHO
PYXPUBNCUBOYNRVNIWNCPOJIOFHOPZRNCRVNBCUNENVVFZIXUNCHPCYVFZIXUPUNFCPWZ
PUKBZPUNVR

B	36	
N	34	
U	33	=> T
P	32	=> A
C	26	

(B 36 => E)

NC	11	
PU	10	=> AT
UB	10	
UN	9	

(NC 11 => IN)

digrams

UKB	6	
RVN	6	
FZI	4	

(UKB 6 => THE)

trigrams

Vigenère cipher (16’th century, Rome)

- Main weakness of monoalphabetic substitution ciphers:
  - Each letter in the ciphertext corresponds to <u>only one letter</u> in the plaintext

- *Poly*alphabetic substitution cipher
  - Given a key $k = (k_1, k_2, ..., k_l)$,
  - Shift each letter p in the plaintext by k_i, where i is modulo m
  - Somewhat resistant to frequency analysis

Vigenère cipher (16'th century, Rome)

k = $\boxed{\text{C R Y P T O}}$ C R Y P T O C R Y P T

(+ mod 26)

m = W H A T A N I C E D A Y T O D A Y

c = Y Y Y I T B K T C S T M V F B P R

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Vigenère cipher (16'th century, Rome)

k = C R Y P T O C R Y P T O C R Y P T

m = W H A T A N I C E D A Y T O D A Y

(+ mod 26)

c = Y Y Y I T B K T C S T M V F B P R

Polyalphabetic cypher

plaintext m

key k

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Vigenère cipher (16'th century, Rome)

k = $\boxed{\text{C R Y P T O}}$ C R Y P T O C R Y P T

(+ mod 26)

m = W H A T A N I C E D A Y T O D A Y

c = Y Y Y I T B | K T C S T M | V F B P R

Suppose the most common letter is "G" ⟶ It is likely that "G" corresponds to "E"

⟶ **First letter of key = "G" – "E" = "C"** (c[i] = m[i] + k[i] ⇒ k[i] = c[i] – m[i])

Cryptanalysis of Vigenère cipher

- A collection of Shift ciphers
  - as many as the length of the key (the number of characters/letters in the key)
- One letter in the ciphertext corresponds to multiple letters in the plaintext: in the previous example: Y Y Y ⇔ W H A
- this makes the use of frequency analysis more difficult

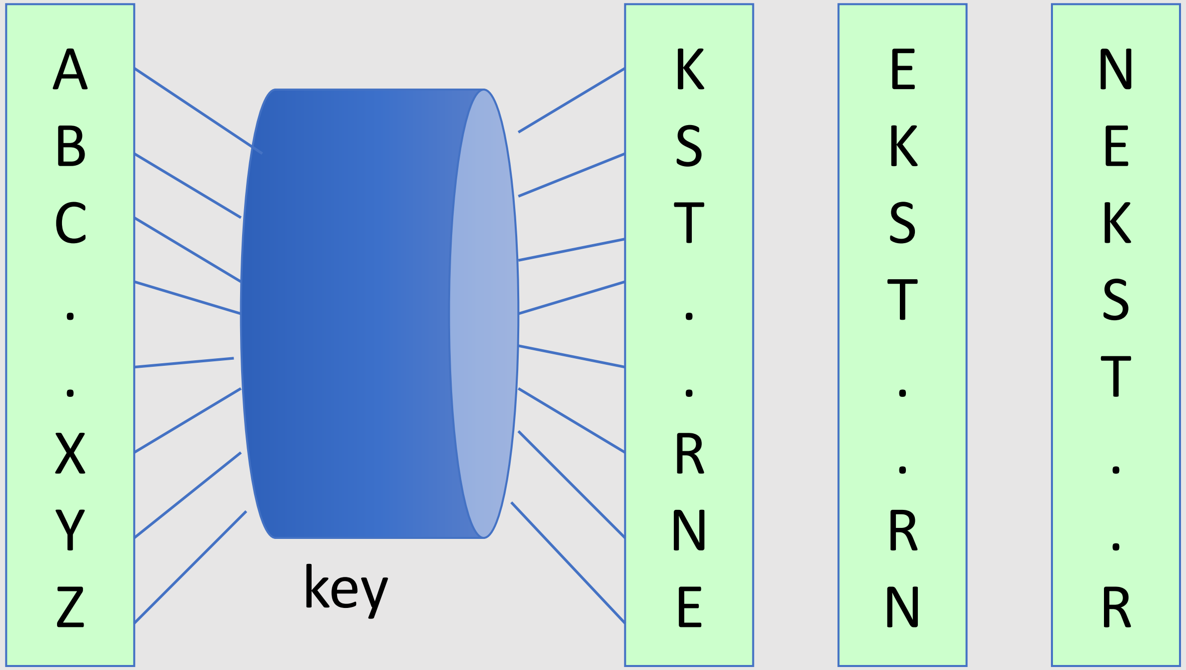
- **How to break the Vigenère cipher?**
  - Guess the length of the key l using some methods
  - Divide the ciphertext into l shift cipher encryptions
  - Use frequency analysis on each shift cipher

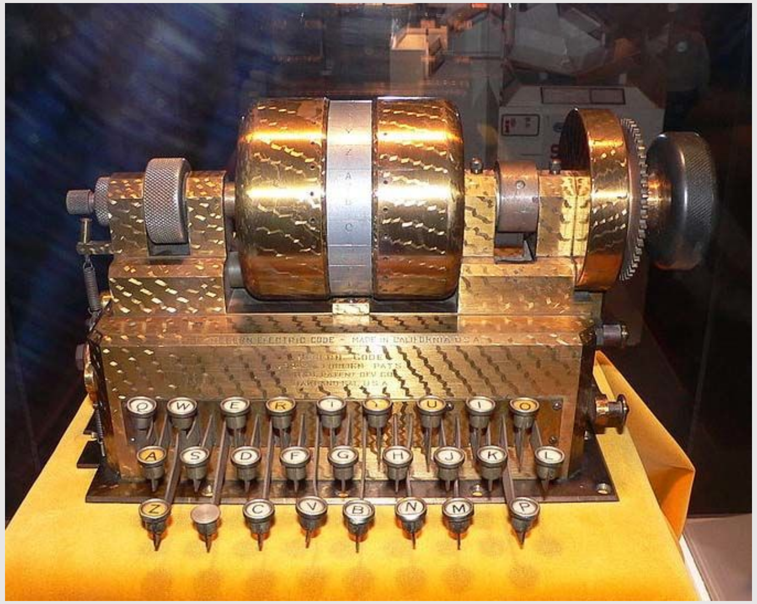
Rotor Machines (1870-1943?)

- Vigenere can be broken once somebody finds the key length
- How to have a longer key?
- Idea:
  - Multiple rounds of substitution, encryption consists of mapping a letter many times
  - Mechanical/electrical wiring to automate the encryption/decryption process
- A machine consists of multiple cylinders (rotors) that map letters several times

Rotor Machines (1870-1943)

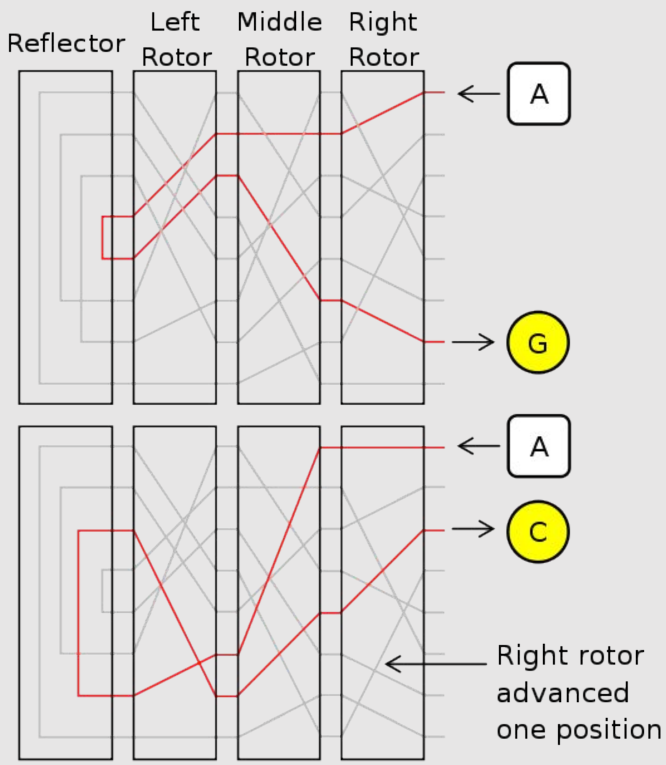
Early example: the Hebern machine (single rotor)





Rotor Machines (cont.)

Most famous: the Enigma (3-5 rotors)



Data Encryption Standard (1974)

~~DES: # keys = 2^{56} , block size = 64 bits~~

<u>Today</u>: AES (2001), Salsa20 (2008) (and many others)

Discrete Probability (crash course)

Probability distribution

- **U**: **finite set**, called **Universe** or **Sample space**

  Examples:
  - Coin flip: **U = { heads, tail }** or **U = { 0, 1 }**
  - Rolling a dice: **U = { 1, 2, 3, 4, 5, 6 }**

- A **Probability distribution** P over U is a function $\mathrm{P} : \mathrm{U} \longrightarrow [0,1]$

  such that $\sum_{\mathrm{x} \in \mathrm{U}} \mathrm{P}(\mathrm{x}) = 1$

  Examples:
  - Coin flip: **P(heads) = P(tail) = 1/2**
  - Rolling a dice: **P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6**

Probability distribution

- **U**: **finite set**, called **Universe** or **Sample space**
- A **Probability distribution** P over U is a function **P : U ⟶ [0,1]**

  such that $\sum_{x \in U} P(x) = 1$

- Notation: U = $\{0,1\}^n$
- **Example:**

  Universe **U** = $\{0,1\}^2$ = {00, 01, 10, 11}

  Probability distribution **P** defined as follows:

  P(00)= 1/2 P(01)= 1/8 P(10)= 1/4 P(11)= 1/8

Probability distributions

Examples:

1. Uniform distribution:for all $x \in U$: $P(x) = 1/|U|$
2. Point distribution at x_0: $P(x_0) = 1$, $\forall x \neq x_0$: $P(x) = 0$

… and many others

Events

Let us consider a universe **U** and a probability distribution **P** over U.

- An **event is a subset A of U,** that is, A ⊆ U
- The **probability of A** is $\mathbf{Pr}[\mathbf{A}] = \sum_{\mathbf{x} \in \mathbf{A}} \mathbf{P}(\mathbf{x})$

Note: Pr[U] = 1

Example

- Universe U = { 1, 2, 3, 4, 5, 6 }
- Probability distribution P s.t. P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6
- **A = {1, 3, 5}**
- **P[A] = 1/6 + 1/6 + 1/6 = 1/2**

Events

Let us consider a universe **U** and a probability distribution **P** over U.

- An **event is a subset A of U,** that is, $A \subseteq U$
- The **probability of A** is $\mathbf{Pr}[\mathbf{A}] = \sum_{\mathbf{x} \in \mathbf{A}} \mathbf{P}(\mathbf{x})$

Example

- Universe $U = \{0,1\}^8$
- Uniform distribution P over U, that is, $P(x) = 1/2^8$ for every $x \in U$
- **A =** $\left\{ \textbf{all x in U such that } \mathbf{lsb_2(x)=11} \right\} \subseteq \mathbf{U}$
- **Pr[A] = ¼**

Hints: $Pr[A] = 1/2^8 \times |A|$
each element in A is of the form _ _ _ _ _ _ 1 1

Union of Events

Given events $\mathbf{A_1}$ and $\mathbf{A_2}$,
$\mathbf{A_1} \cup \mathbf{A_2}$ is an event.

- $\Pr[A_1 \cup A_2] = \Pr[A_1] + \Pr[A_2] - \Pr[A_1 \cap A_2]$
- $\Pr[A_1 \cup A_2] \leq \Pr[A_1] + \Pr[A_2]$ ("Union bound")
- $A_1 \cap A_2 = \varnothing \Rightarrow \Pr[A_1 \cup A_2] = \Pr[A_1] + \Pr[A_2]$

Random Variables

Def: a **random variable** X is a function **X : U ⟶ V**

Example (Rolling a dice):
U = { 1, 2, 3, 4, 5, 6 }
Uniform distribution P over U: P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6

Random variable **X : U ⟶ { "even", "odd" }**
X(2) = X(4) = X(6) = "even"
X(1) = X(3) = X(5) = "odd"

Pr[X="even"] = 1/2 , Pr[X="odd"] = 1/2

More generally: **X *induces* a distribution on V**

The **uniform** random variable

Let S be some set, e.g. $S = \{0,1\}^n$

We write $r \longleftarrow S$ to denote a **uniform random variable** over S

for all $a \in S$: $\Pr[\, r=a \,] = 1/|S|$

The **uniform** random variable

Let U be some set, e.g. $U = \{0,1\}^n$

We write $r \xleftarrow{R} U$ to denote a **uniform random variable** over U

for all $a \in U$: $\Pr[\, r=a \,] = 1/|U|$

(formally, r is the identity function: $r(x)=x$ for all $x \in U$)

Defining a random variable in terms of another

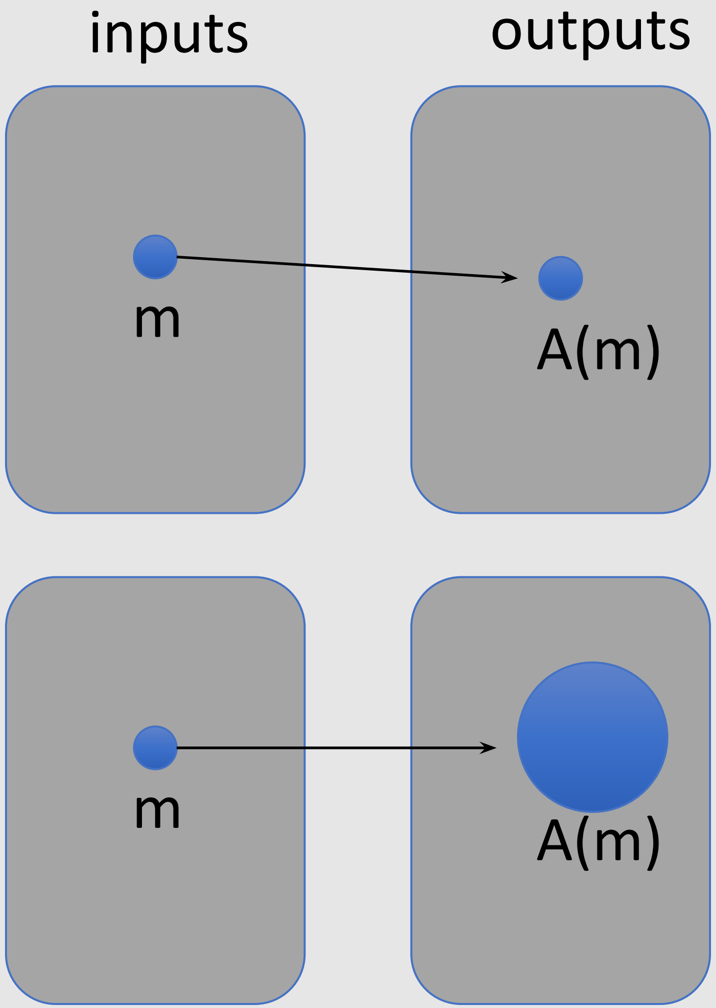
- Let r be a uniform random variable on $\{0,1\}^2$
- Define the random variable $X = r_1 + r_2$

- Then $\Pr[X=2] = ¼$

- Hint: $\Pr[X=2] = \Pr[\, r=11 \,]$

Randomized algorithms

- **Deterministic** algorithm: $y \longleftarrow A(m)$
- **Randomized** algorithm
  output is a random variable $y \longleftarrow A(m)$



Recap

- U: Universe or Sample space (e.g., $U = \{0,1\}^n$)
- A Probability distribution P over U is a function $P : U \longrightarrow [0,1]$ such that $\sum_{x \in U} P(x) = 1$
- An event is a subset A of U, that is, $A \subseteq U$
- The probability of event A is $\Pr[A] = \sum_{x \in A} P(x)$
- A random variable is a function $X : U \longrightarrow V$
  X takes values in V and defines a distribution on V

Independence

Definition. **Independent events**

Events A and B are **independent** if

$\Pr[\ A \cap B\ ] = \Pr[A] \cdot \Pr[B]$

Definition. **Independent random variables**

Random variables X and Y taking values in V are **independent** if

$\forall a,b \in V: \quad \Pr[\ X=a \text{ and } Y=b] = \Pr[X=a] \cdot \Pr[Y=b]$

XOR

XOR of two strings in $\{0,1\}^n$ is their bit-wise addition mod 2

X	Y	X $\oplus$ Y
0	0	0
0	1	1
1	0	1
1	1	0

$$
\begin{array}{r}
0\;1\;1\;0\;1\;1\;1 \\
1\;0\;1\;1\;0\;1\;0 \quad \oplus \\
\hline
1\;1\;0\;1\;1\;0\;1
\end{array}
$$

An important property of XOR

Theorem:

1. **X:** a random variable over $\{0,1\}^n$ with a **uniform distribution**
2. **Y:** a random variable over $\{0,1\}^n$ with an **arbitrary distribution**
3. **X** and **Y** are **independent**

- Then **Z := Y⊕X** is a **UNIFORM** random variable over $\{0,1\}^n$

Proof: (for n=1)

Pr[Z=0] =

Pr[(X,Y)=(0,0) or (X,Y)=(1,1)] =

Pr[(X,Y)=(0,0)] + Pr[(X,Y)=(1,1)] =

$p_0/2 + p_1/2 = ½$

Therefore Pr[Z=1] = ½

Y	Pr
0	p_0
1	p_1

X	Pr
0	1/2
1	1/2

X	Y	Pr
0	0	$p_0/2$
0	1	$p_1/2$
1	0	$p_0/2$
1	1	$p_1/2$

The birthday paradox

Let $r_1, \ldots, r_n \in U$ be **independent identically distributed** random variables

Theorem: when $\mathbf{n} = 1.2 \times \mathbf{|U|^{1/2}}$ then $\Pr[\,\exists i \neq j: \ r_i = r_j\,] \geq ½$

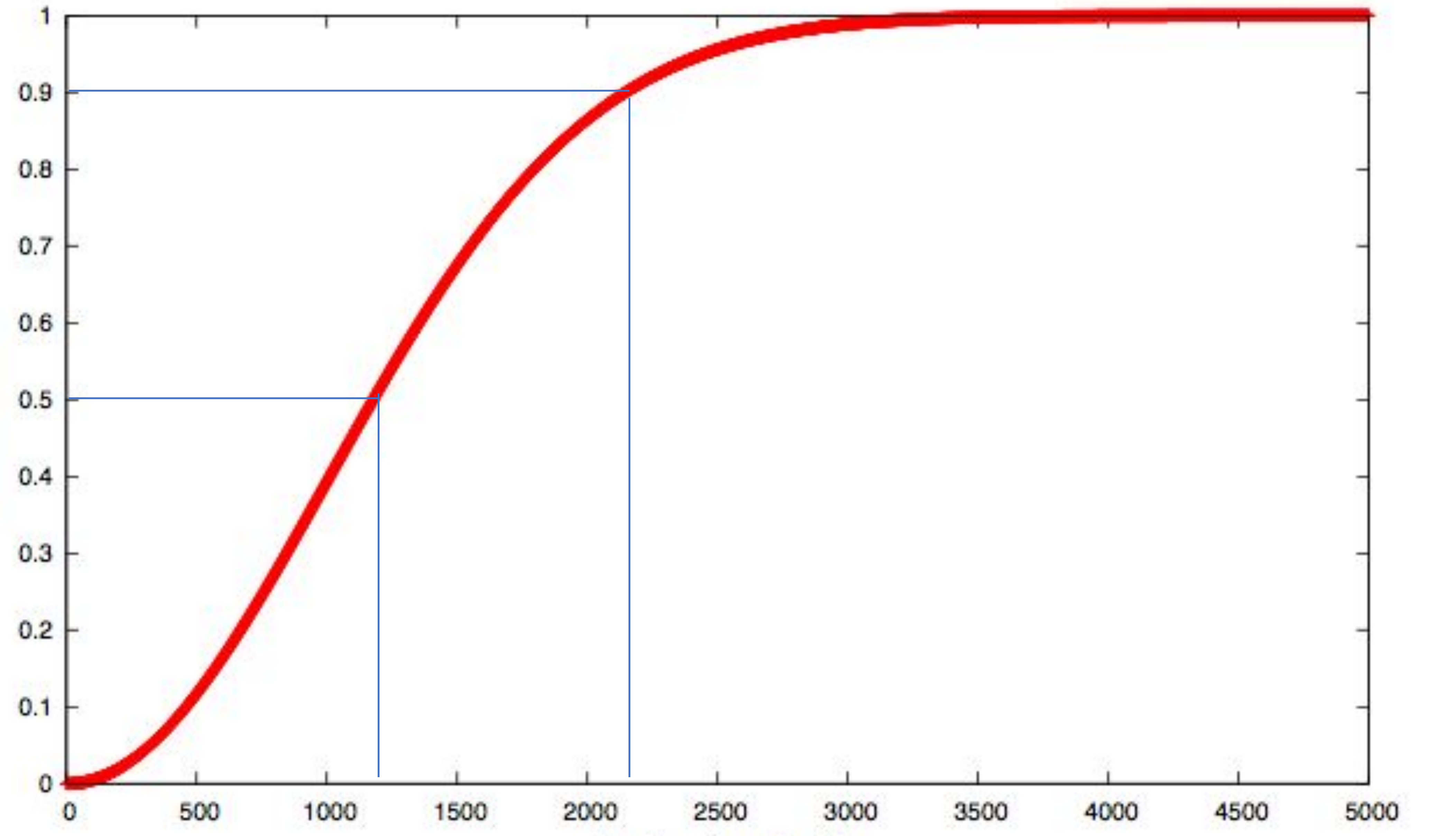
Example:

- $U = \{1, 2, 3, \ldots, 366\}$
- When $\boldsymbol{n} = 1.2 \times \sqrt{366} \approx \mathbf{23}$, two people have the same birthday with probability $\geq ½$

Example:

- Let $U = \{0,1\}^{128}$
- After sampling about 2^{64} random messages from U, some two sampled

$|U|=10^6$



collision probability

samples n