Fondamenti di Cybersecurity – Modulo I

- 20h circa
- Docente: **Riccardo Treglia**
  - Email: **riccardo.treglia@unibo.it**

Piattaforma didattica

- Virtuale

  e verrà costantemente aggiornato con:
  - Informazioni
  - **Materiale didattico (slides)**
  - **Annunci**

Materiale didattico

- **Slide** caricate su Virtuale del corso
- Testi consigliati:
  - Jean-Philippe Aumasson,
    Serious Cryptography: A Practical Introduction to Modern Encryption.
  - Bruce Schneier,
    Applied Cryptography: Protocols, Algorithms, and Source Code in C.
  - Mark Stamp,
    Information Security: Principles and Practice.
  - William Stallings
    Crittografia
  - Dan Boneh, Victor Shoup,
    A Graduate Course in Applied Cryptography. (approccio matematico)

Esame

- Prova scritta
  - Voto finale = Scritto + Successo laboratori

Scritto: 24/25 pt
Laboratori: max 8 pt
NO orali

- Date esami: consultare il sito del Dipartimento
  Due appelli a **Giugno**, uno a **Luglio** e uno a **Settembre**

Roadmap

0. What is Cryptography - History of Cryptography

1. Introduction Mathematics: Modular Arithmetic - Discrete Probability

2. One-time pad, Stream Ciphers and Pseudo Random Generators

3. Attacks on Stream Ciphers and The One-Time Pad

4. Real-World Stream Ciphers (weak(RC4), eStream,nonce, Salsa20)

5. Secret key cryptographic systems;

6. Public key cryptographic systems

7. DES protocols (just as an introduction), AES

8. Electronic Signatures, Public-key Infrastructure, Certificates and Certificate Authorities

9. Sharing of secrets; User authentication; Passwords

10. Tutor Training

Bonus. Legislation, Ethics and Management

Introduction

Welcome

Course **objectives**:

- Learn how crypto primitives work
- Learn how to use them correctly and reason about security

Che cos’è la Crittografia?

- **Crittografia**
  - *Kryptós*: nascosto
  - *Graphía:* scrittura
  - Metodi che consentano di **memorizzare**, **elaborare** e **trasmettere** informazioni in presenza di agenti ostili

- **Crittoanalisi**
  - Analisi di un testo cifrato nel tentativo di decifrarlo senza possedere la chiave

- **Crittologia:** Crittografia + Crittoanalisi

Cryptography is everywhere

Secure communication:

- web traffic: HTTPS
- wireless traffic: Wireless Network, GSM, Bluetooth

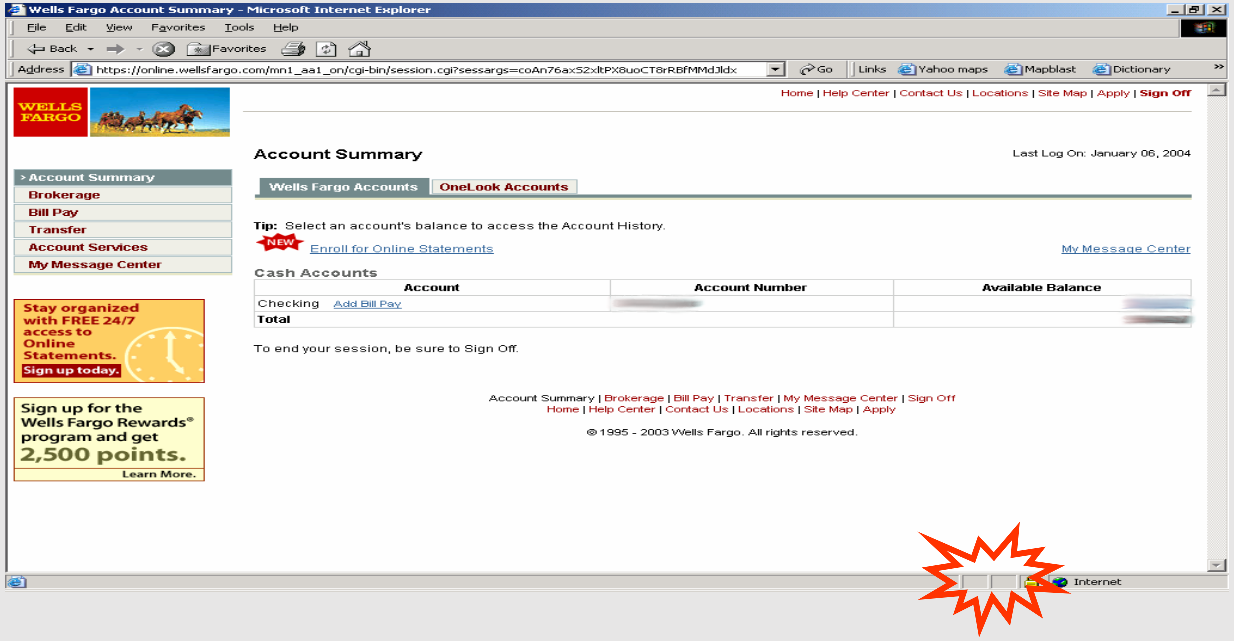
Encrypting files on disk

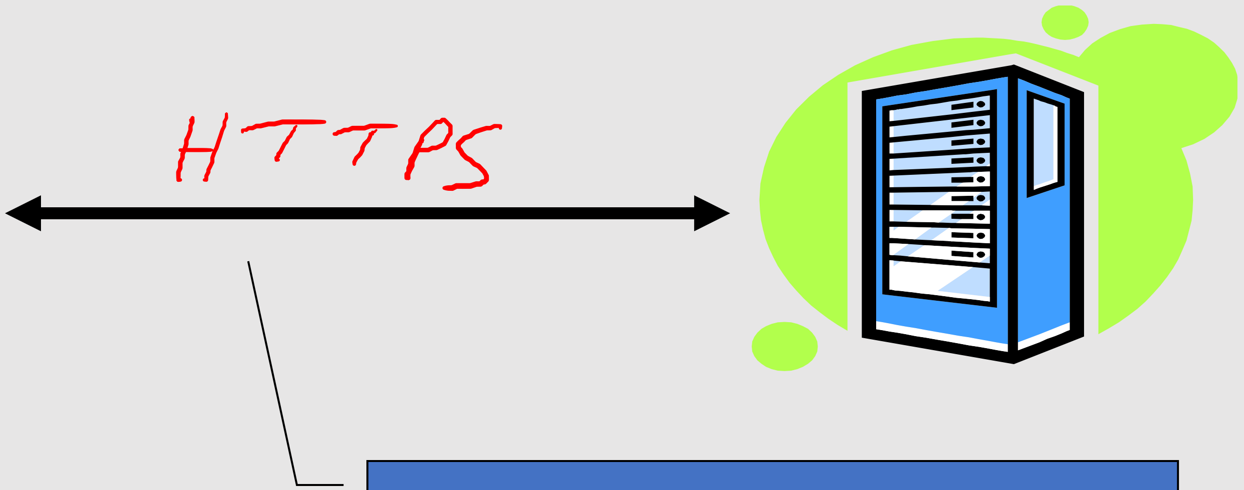
Content protection (e.g., DVD, Blu-ray)

User authentication

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

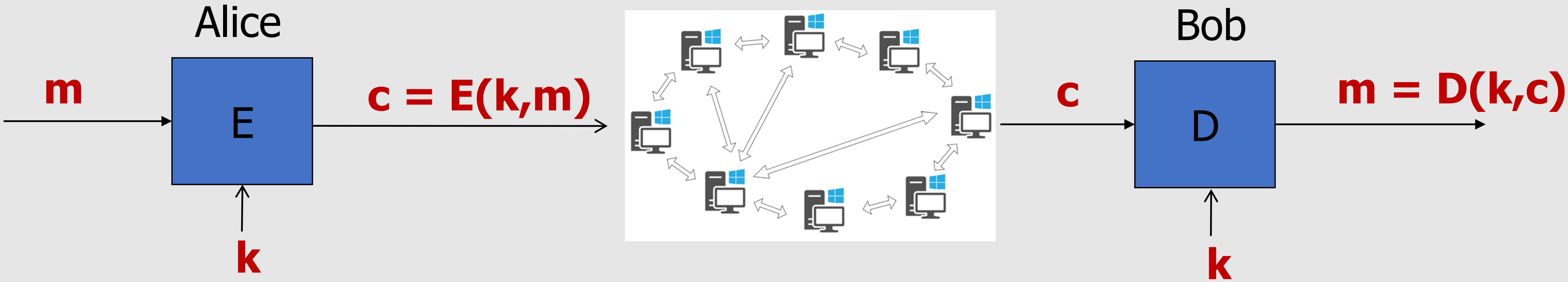
Secure communication





no eavesdropping
no tampering

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

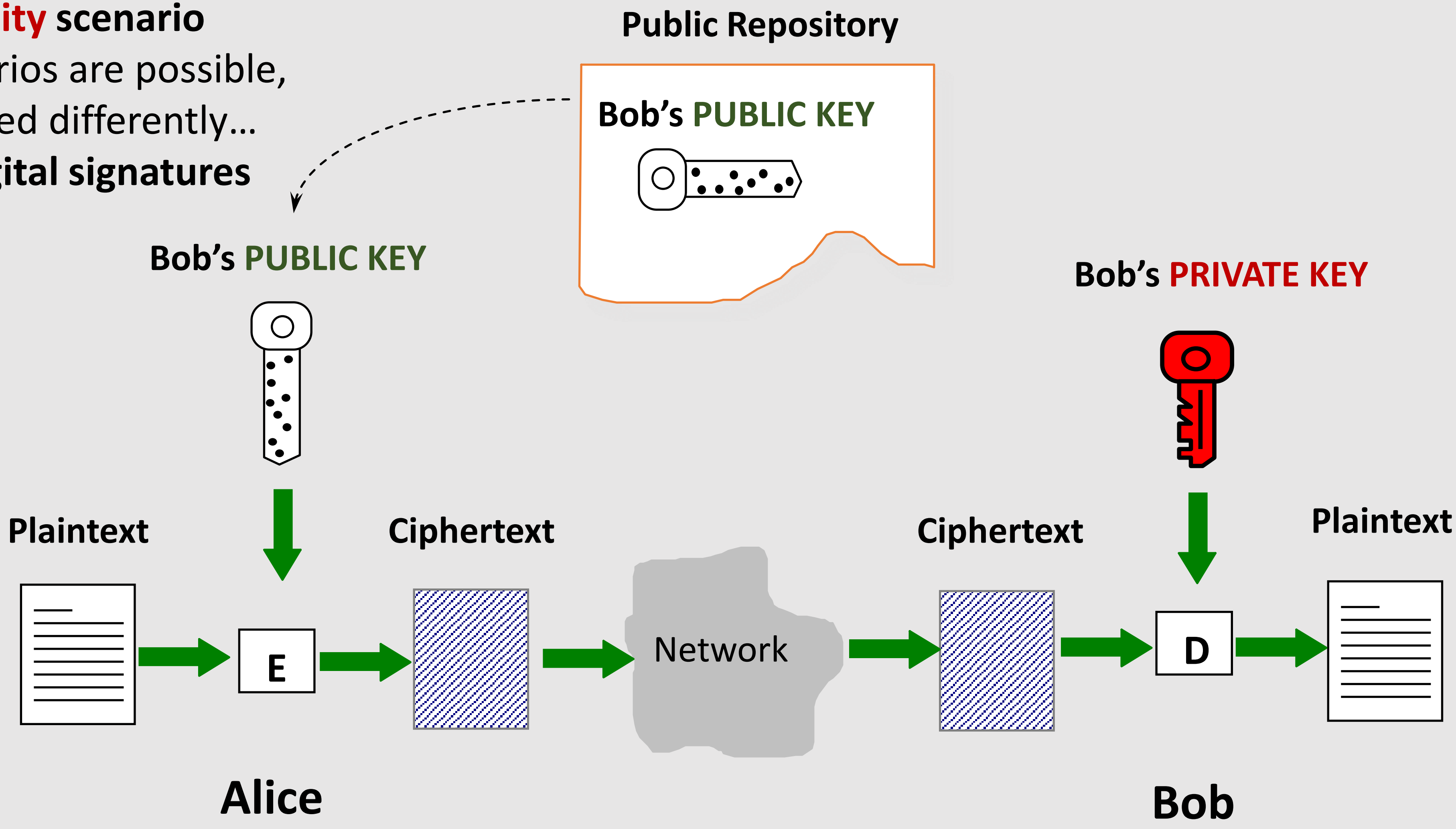
- **Single-use key**: (or **one-time key**):
  Key is only used to encrypt **one message**
  - encrypted email: new key generated for every email

- **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**:

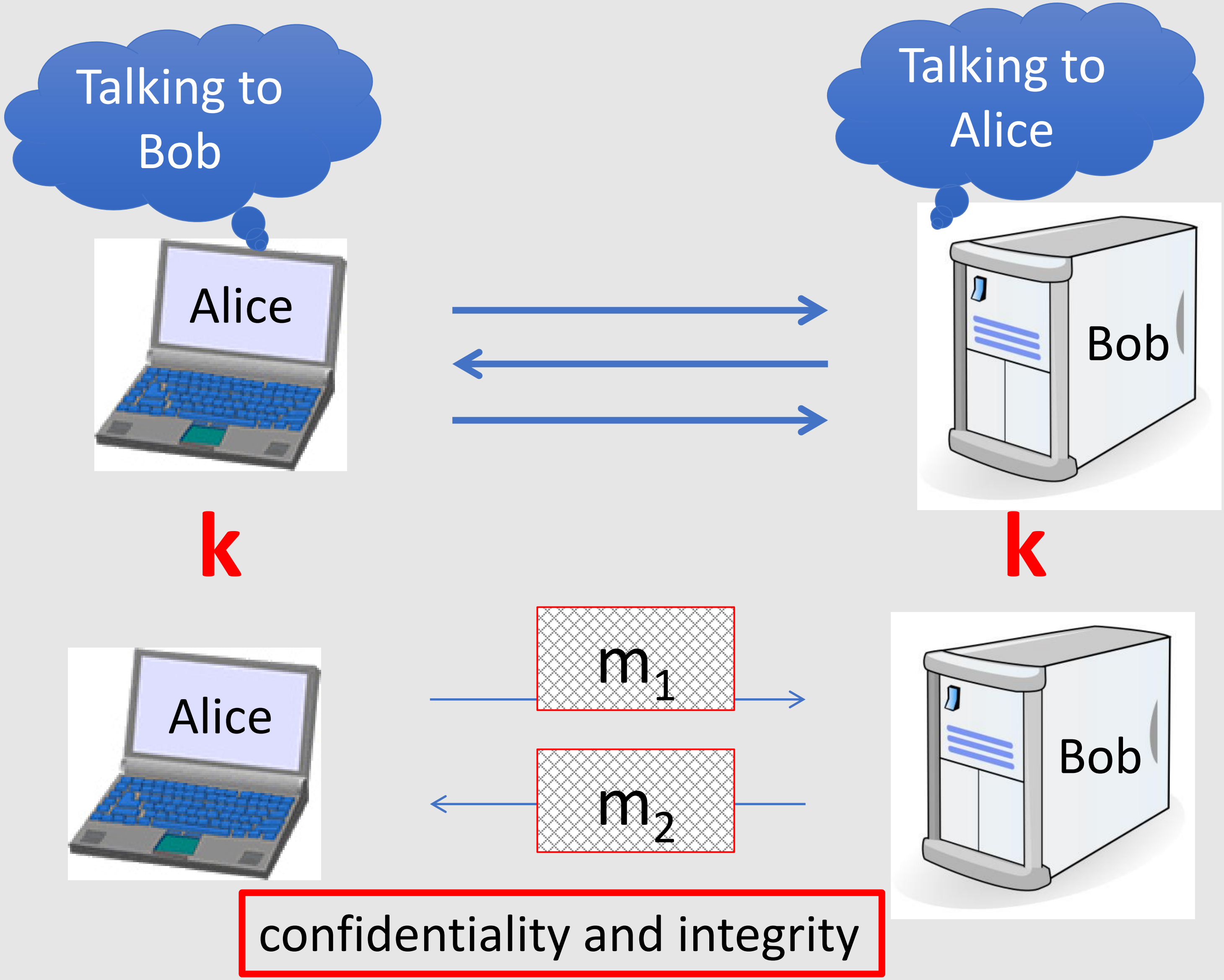
- The solution to all security problems
- Reliable unless implemented and used properly
- Something you should try to invent yourself
    - many many examples of broken ad-hoc designs

Some Applications

Secure communication

1. Secret key establishment:

2. Secure communication:



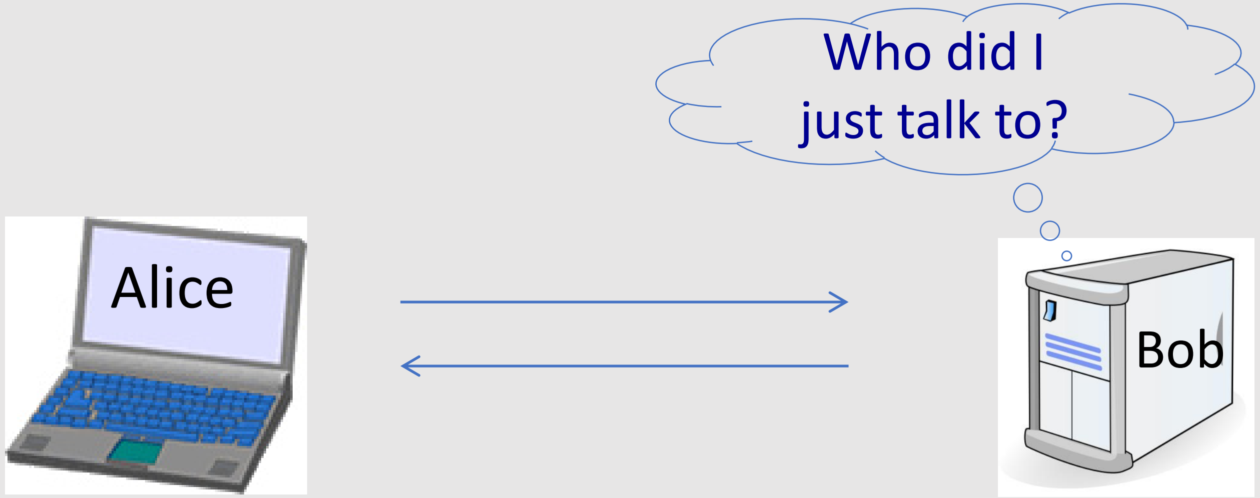
But crypto can do much more

- Digital signatures

- Signatures of the same person change over different documents
- Asymmetric Cryptography is used

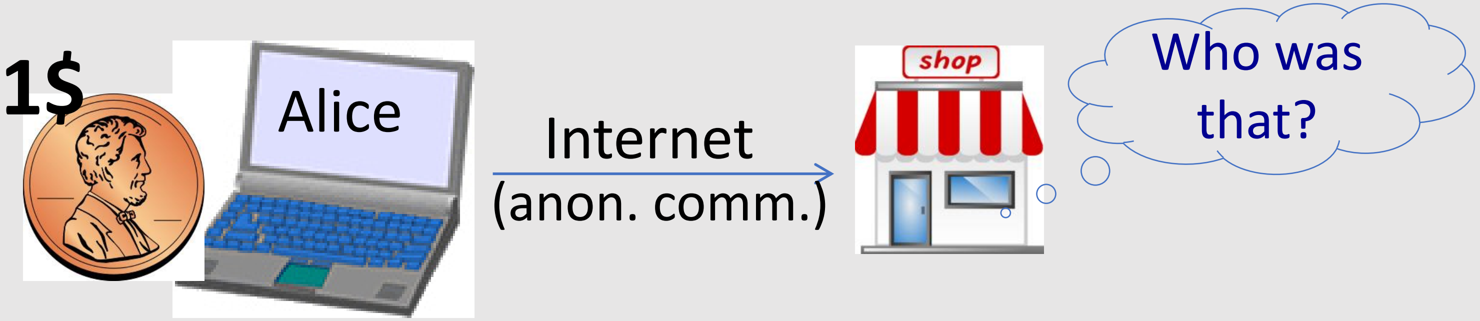
But crypto can do much more

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



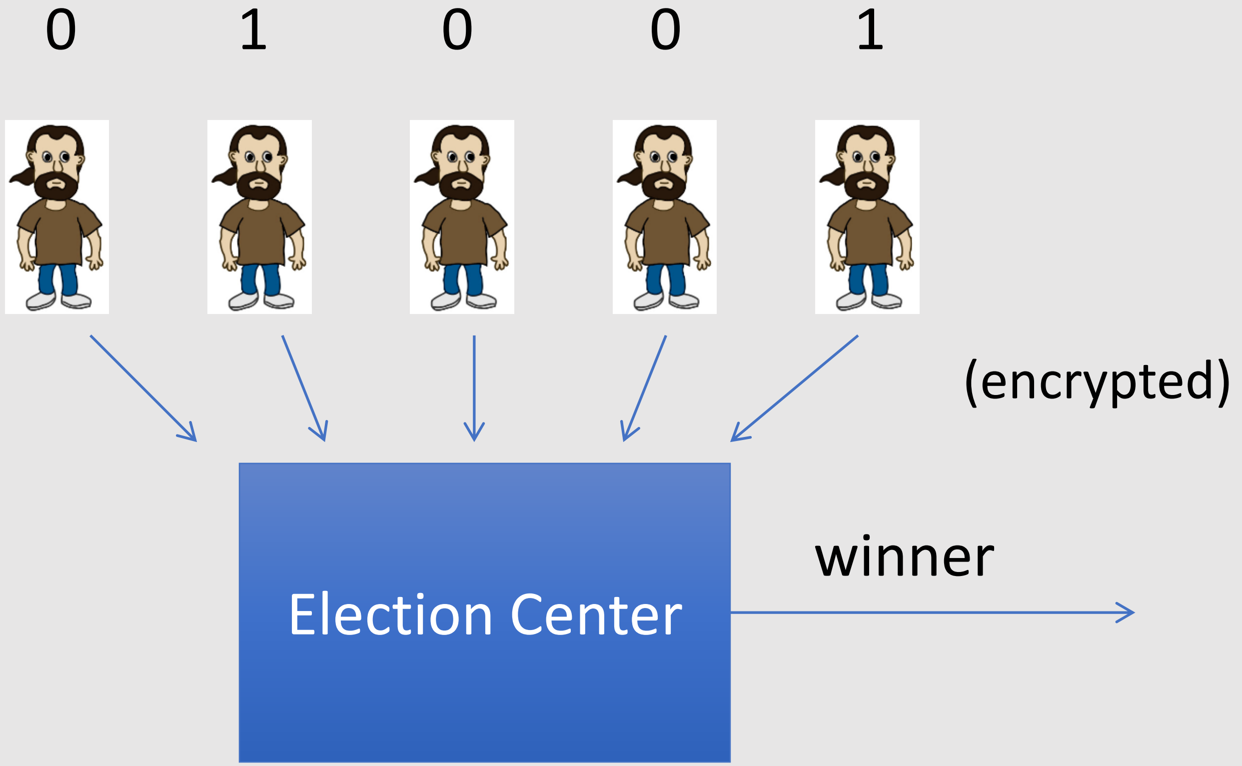
But crypto can do much more

- 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 2nd highest bid

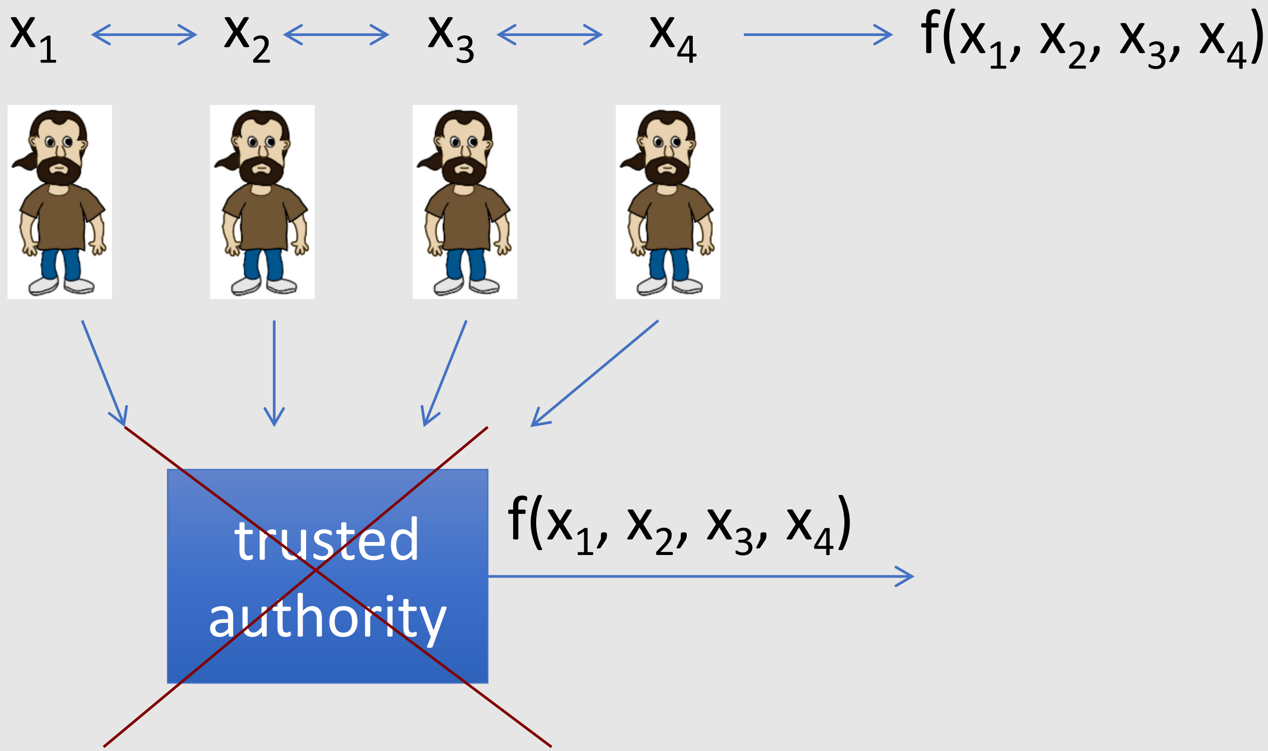
Election Center must determine the winner
without knowing the individual votes!

Protocols

- Elections
- Private auctions

Secure multi-party computation

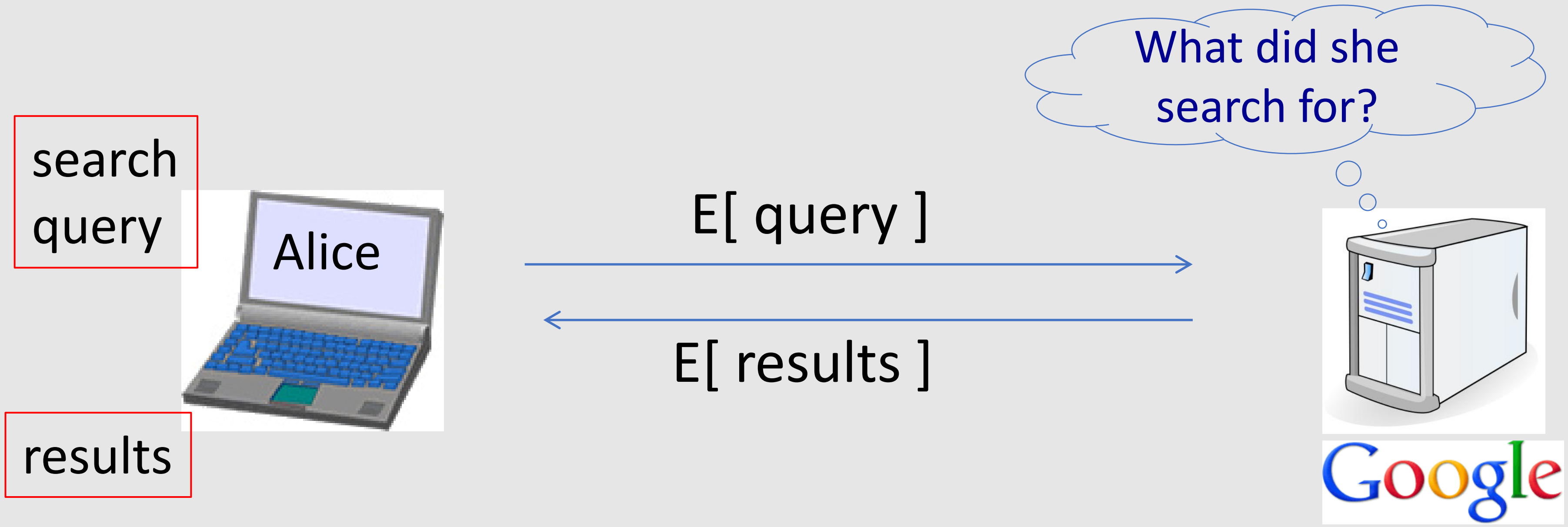
Goal: compute $f(x_1, x_2, x_3, x_4)$



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

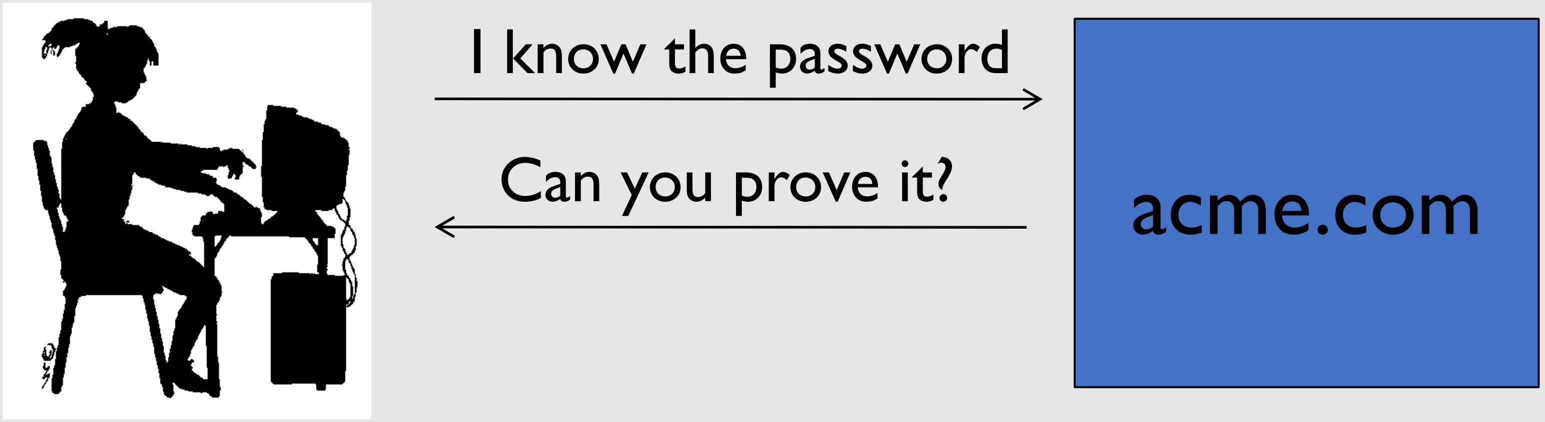
Crypto magic

- Privately outsourcing computation



Crypto magic

- Zero knowledge (proof of knowledge)



A rigorous science

The three steps in cryptography:

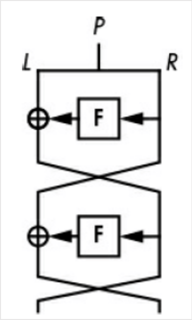
- Precisely specify threat model
- Propose a construction
- Prove that breaking construction under
  threat model will solve an underlying hard problem

Brief History of Crypto

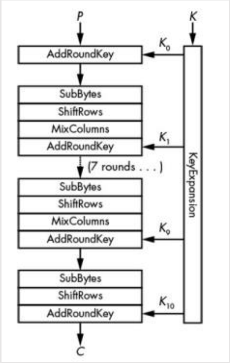
Che cos'è la Crittografia?

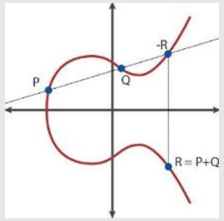
- Metodi per **memorizzare**, **elaborare** e **trasmettere** informazioni in maniera **sicura** in presenza di agenti ostili
- **Crittografia:** *Kryptós*: nascosto + *Graphía:* scrittura



$n = p \times q$
$p, q?$

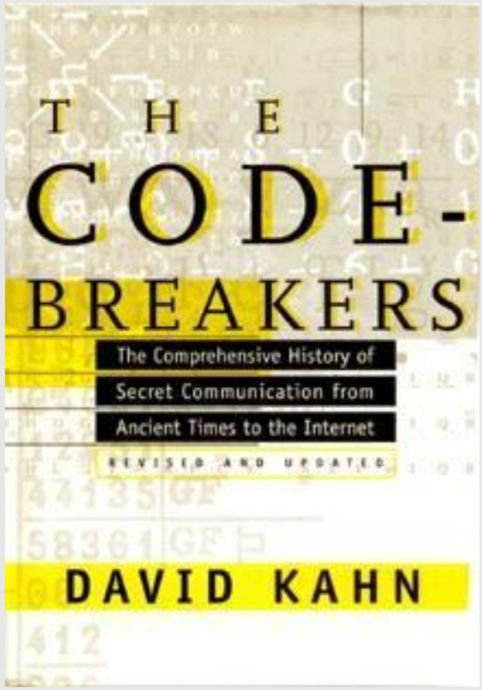




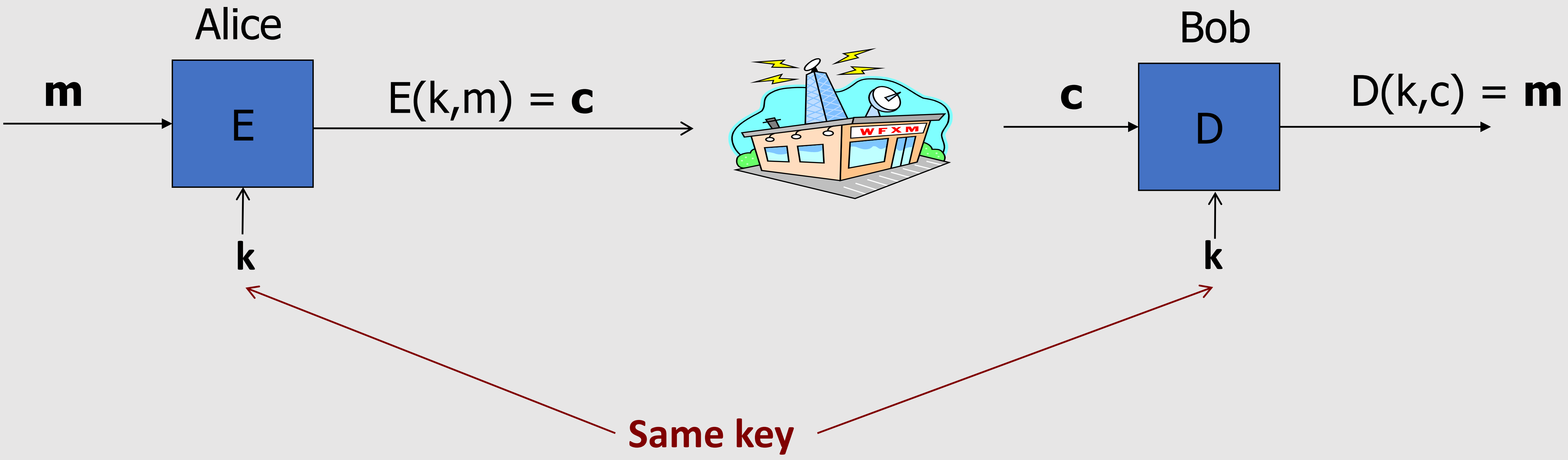
Scitala	Cifrario di Cesare	Enigma	DES	RSA	AES	Crittografia ellittica
400 aC	50 aC	1918	1975	1977	2001	2005

History

David Kahn, "The code breakers" (1996)



Symmetric Ciphers



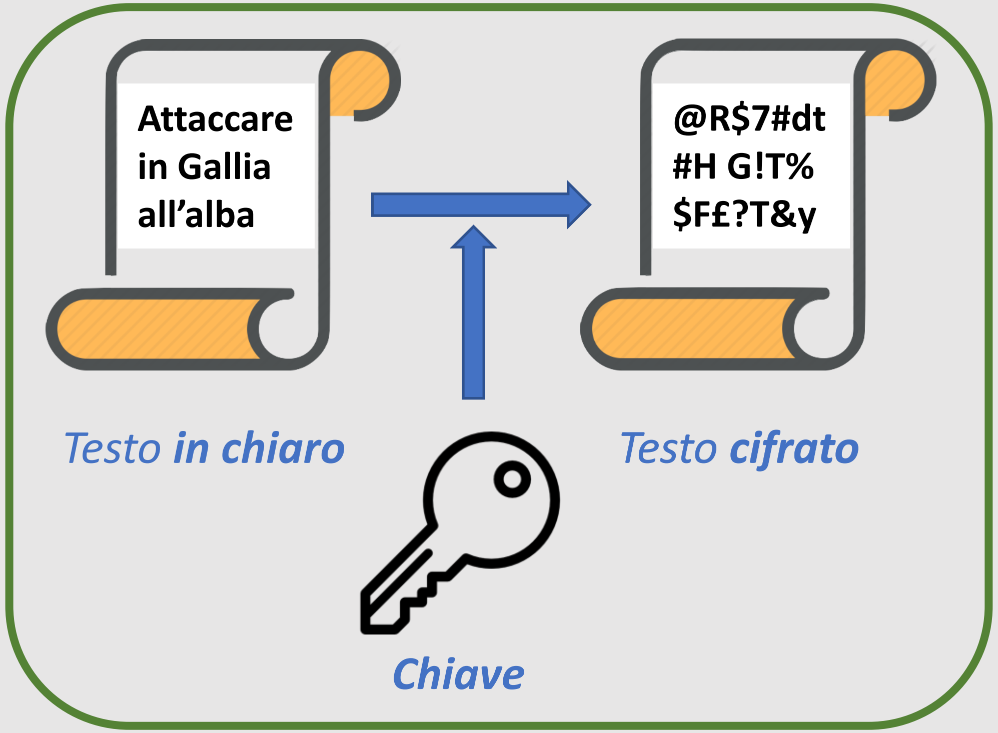
Cypher: (E, D)

Un classico scenario

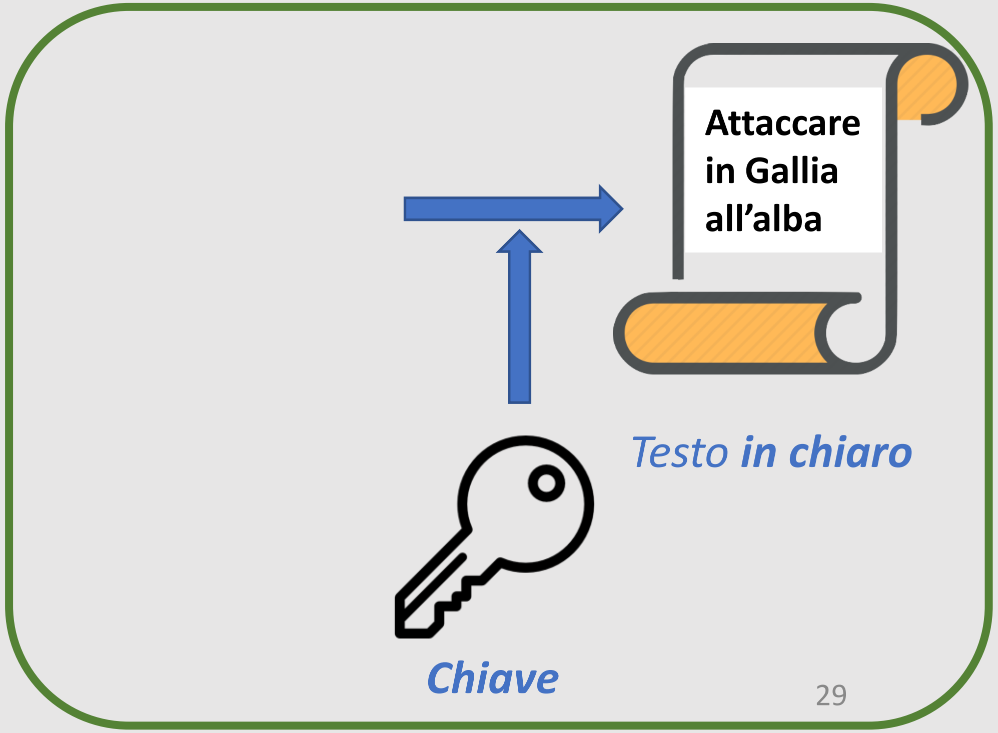
*Algoritmi di cifratura e decifratura: **pubblici***

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

Cifratura



Decifratura



Cifrario di Cesare

	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
Chiave	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)

Few Historic Examples (all badly broken)

1. Substitution cipher

$$c := E(k, \text{"bcza"}) = \text{"wnac"}$$

$$D(k, c) = \text{"bcza"}$$

k :=

a → c
b → w
c → n
.
.
.
z → a

Caesar Cipher (no key)

Shift by 3

a → d
b → e
c → f
.
.
.
y → b
z → c

What is the size of key space in the 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**

An Example

UKBYBIPOUZBCUFEEBORUKBYBHOBBRFESPVKBWFOFERVNBCVBZPRUBOFERVNBCVBPCYYFVU
FOFEIKNWFRFIKJNUPWRFIPOUNVNIPUBRNCUKBEFWWFDNCHXCYBOHOPYXPUBNCUBOYNRV
NIWNCPOJIOFHOPZRVFZIXUBORJRUBZRBCHNCBBONCHRJZSFWNVRJRUBZRPCYZPUKBZPUNV
PWPCYVFZIXUPUNFCPWRVNBCVBRPYYNUNFCPWWJUKBYBIPOUZBCUIPOUNVNIPUBRNCHOP
YXPUBNCUBOYNRVNIWNCPOJIOFHOPZRNCRVNBCUNENVVFZIXUNCHPCYVFZIXUPUNFCPWZP
UKBZPUNVR

B	**36**	➔ E
N	**34**	
U	**33**	➔ T
P	**32**	➔ A
C	**26**	

NC	**11**	➔ IN
PU	**10**	➔ AT
UB	**10**	
UN	**9**	

digrams

UKB	**6**	➔ THE
RVN	**6**	
FZI	**4**	

trigrams

2. Vigenère cipher (16'th century, Rome)

k = 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

2. 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

	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

2. Vigenère cipher (16'th century, Rome)

k = 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

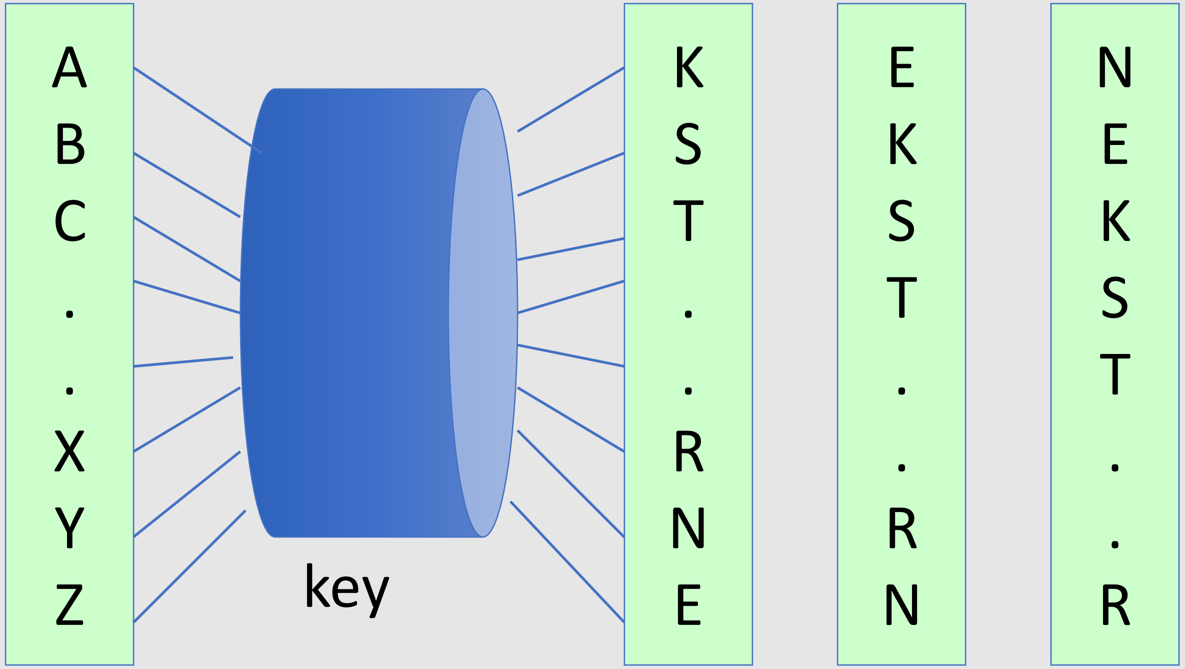
c = Y Y Y I T B | K T C S T M | V F B P R

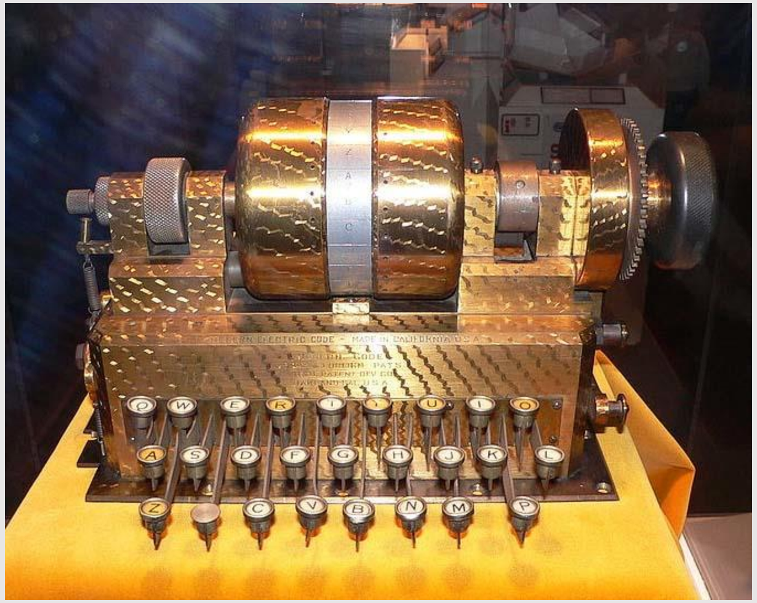
Suppose the most common letter is "G" → It is likely that "G" corresponds to "E"

→ **First letter of key = "G" – "E" = "C"** $(c[i] = m[i] + k[i] \Rightarrow k[i] = c[i] - m[i])$

3. Rotor Machines (1870-1943)

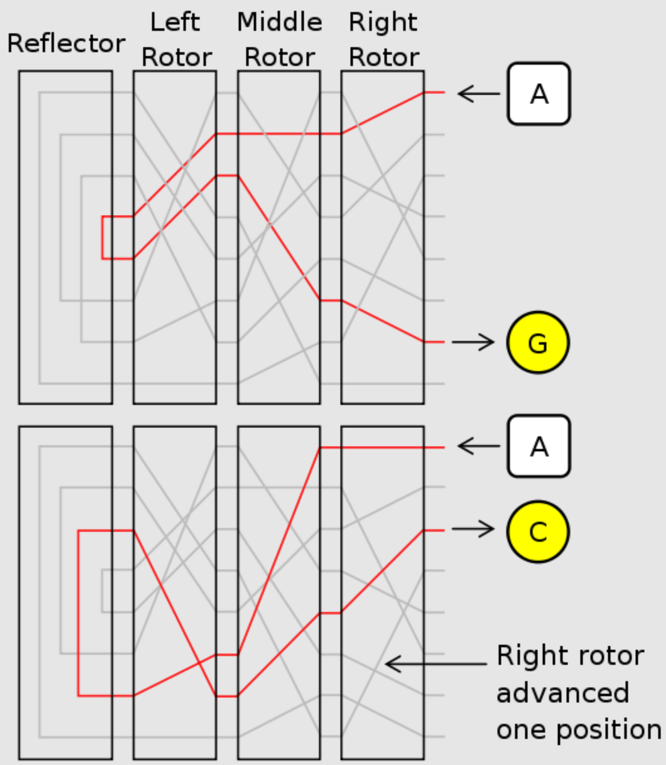
Early example: the Hebern machine (single rotor)





Rotor Machines (cont.)

Most famous: the Enigma (3-5 rotors)



4. Data Encryption Standard (1974)

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

Today: 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 $\mathbf{Pr[A]} = \sum_{\mathbf{x \in A}} \mathbf{P(x)}$

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 $\mathbf{Pr[A]} = \sum_{\mathbf{x \in A}} \mathbf{P(x)}$

Example

- Universe $U = \{0,1\}^8$
- Uniform distribution P over U, that is, $P(x) = 1/2^8$ for every $x \in U$
- **A =** $\{$ **all x in U such that** $\mathbf{lsb_2(x)=11}$ $\}$ $\subseteq$ **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 $\mathbf{A_2}$,

$\mathbf{A_1 \cup A_2}$ 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")
- $A_1 \cap A_2 = \emptyset \Rightarrow \Pr[\, A_1 \cup A_2 \,] \;=\; \Pr[A_1] + \Pr[A_2]$

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

Defining a random variable in terms of another

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

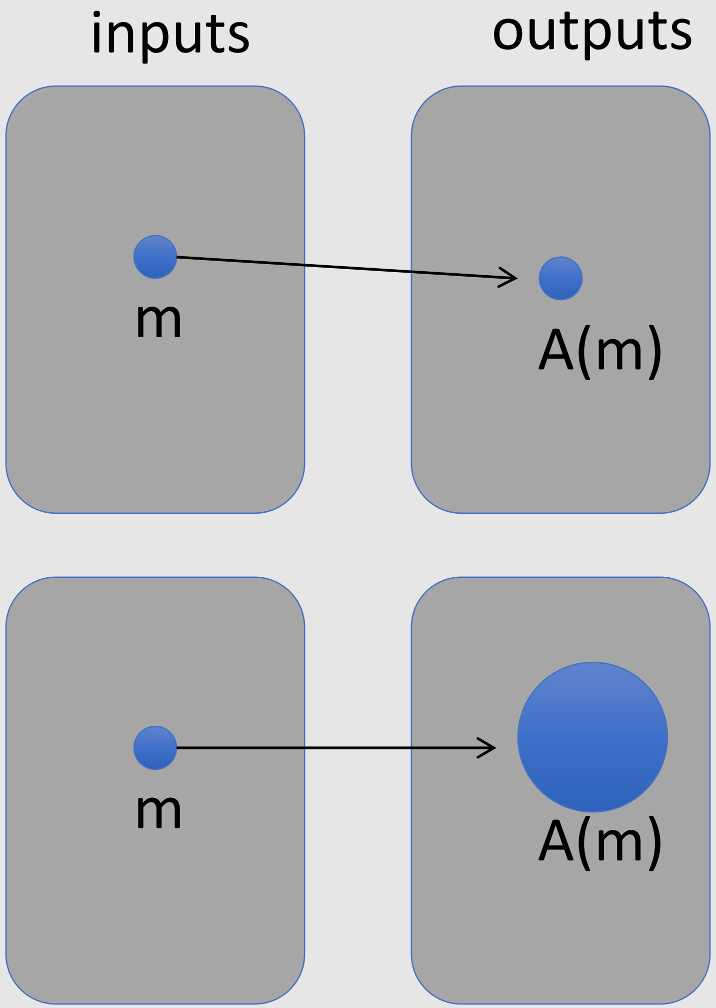
- Then $\Pr[X=2] = ¼$

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

Randomized algorithms

- **Deterministic** algorithm: y ⟵ A(m)

- **Randomized** algorithm
  output is a random variable y ⟵ 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 ⊕ Y
0	0	0
0	1	1
1	0	1
1	1	0

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

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 **$Z := Y \oplus X$** 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 **n** = 1.2 × **|U|**$^{1/2}$ then $\Pr[\, \exists i \neq j: \; r_i = r_j \,] \geq ½$

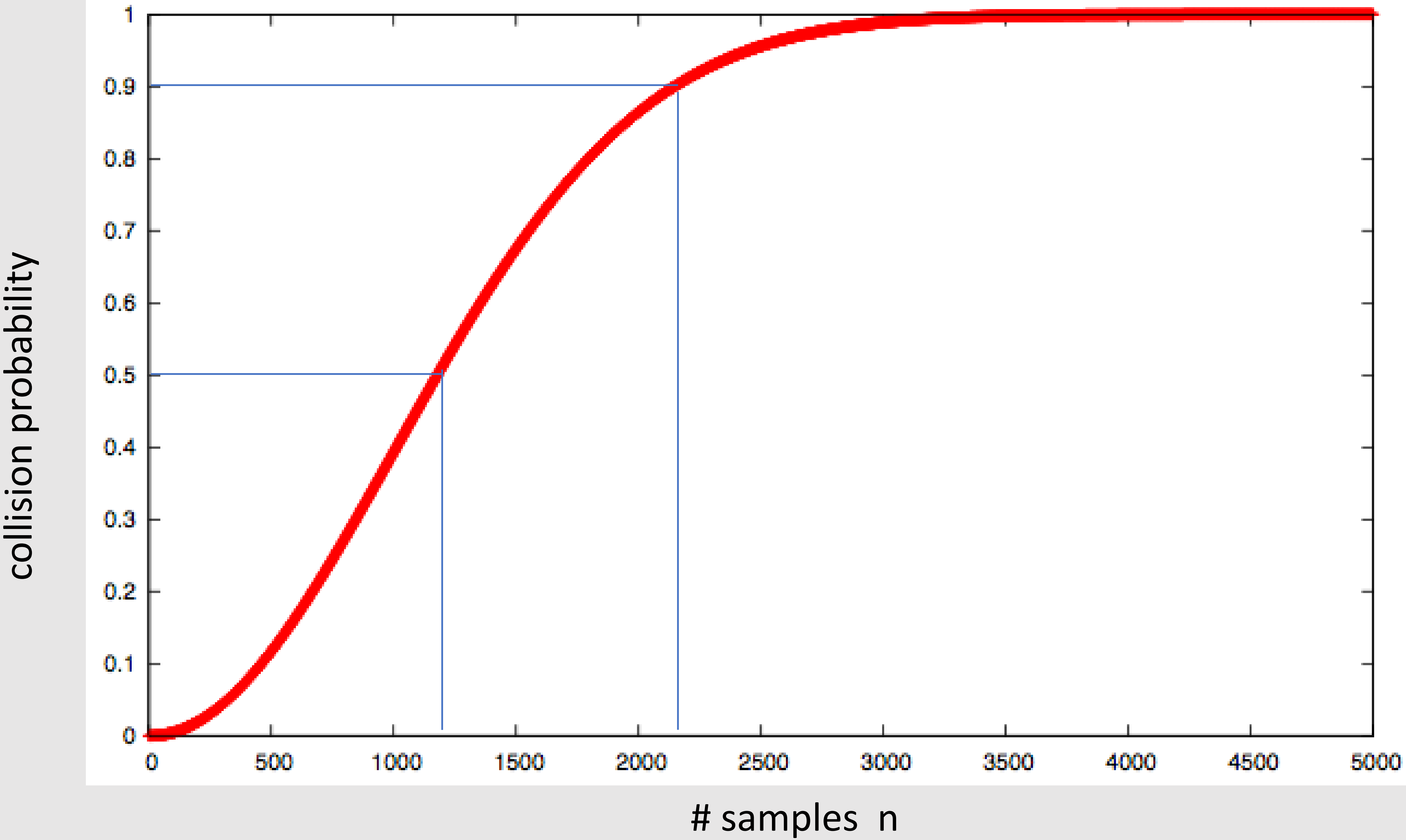
Example:

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

Example:

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

$|U|=10^6$



collision probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
samples n