Introduction

Course Overview

Welcome

Course objectives:

- Learn how crypto primitives work

- Learn how to use them correctly and reason about security

My recommendations:

- Take notes

- Pause video frequently to think about the material

- Answer the in-video questions

Cryptography is everywhere

Secure communication:

- web traffic: HTTPS

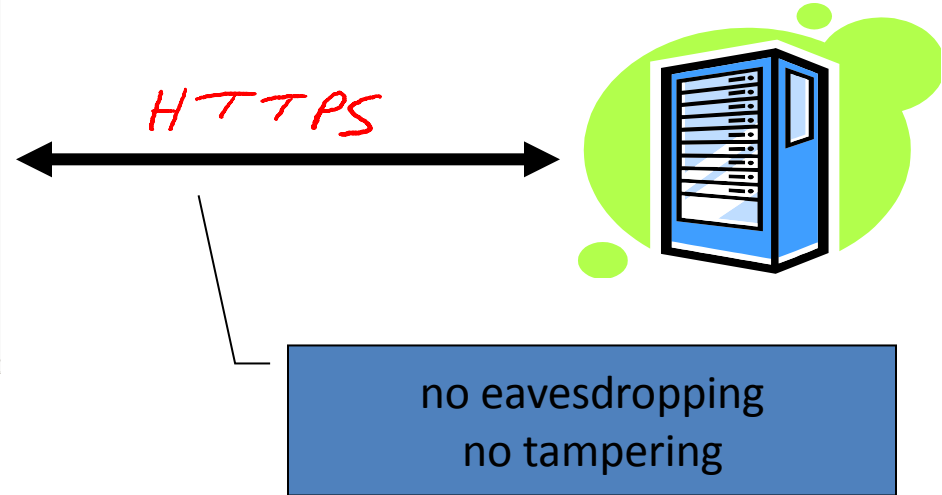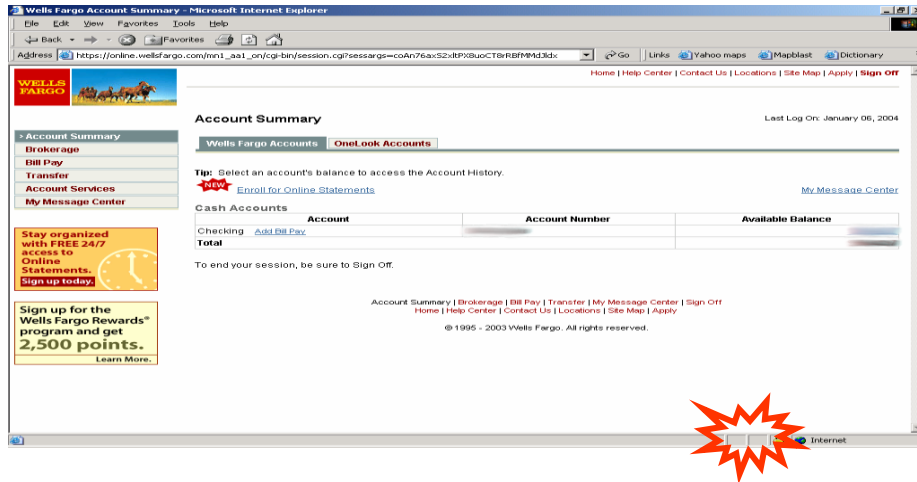- wireless traffic: 802.11i WPA2 (and WEP),   GSM,   Bluetooth

Encrypting files on disk: EFS, TrueCrypt

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

User authentication

…   and much much more

Secure communication


HTTPS

no eavesdropping
no tampering

Secure Sockets Layer / TLS

Two main parts

1. Handshake Protocol: **Establish shared secret key using public-key cryptography**   (2nd part of course)

2. Record Layer: **Transmit data using shared secret key**

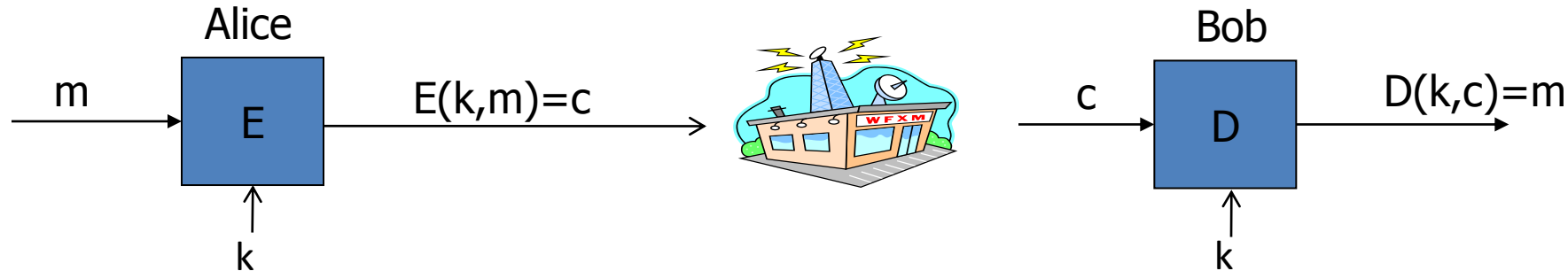   Ensure confidentiality and integrity   (1st part of course)

Protected files on disk

Disk

Alice ➡ File 1 ➡ Alice

File 2

No eavesdropping
No tampering
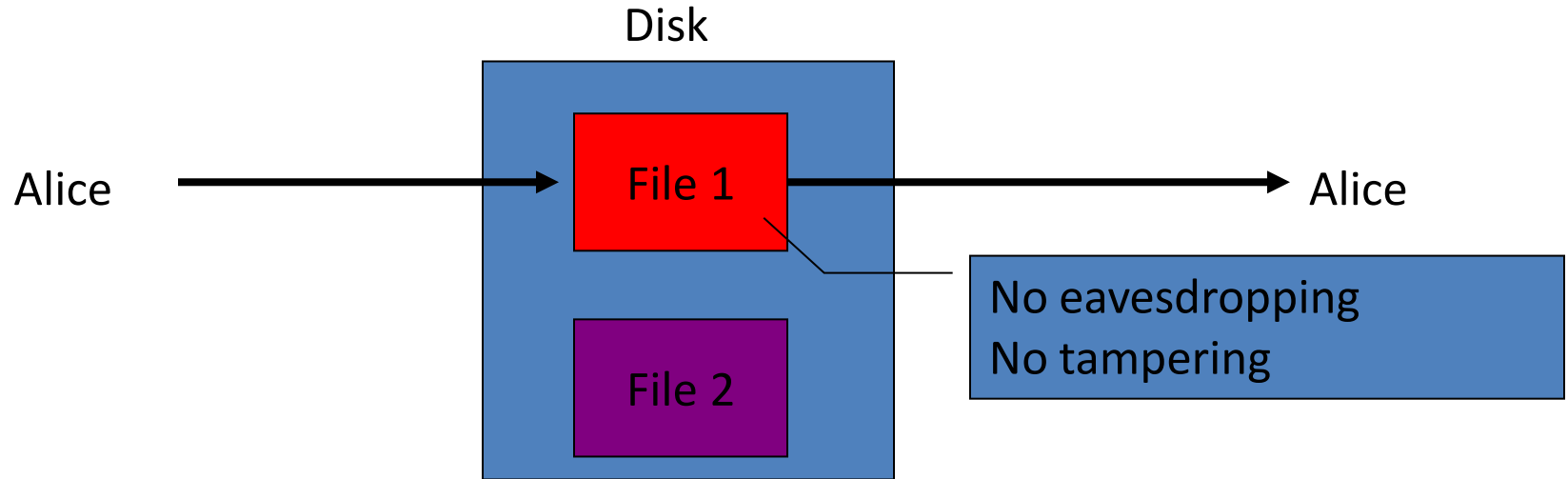
Analogous to secure communication:

Alice today sends a message to Alice tomorrow

Building block: sym. encryption

E, D: cipher   k: secret key (e.g. 128 bits)

m, c: plaintext, ciphertext

Encryption algorithm is publicly known
- Never use a proprietary cipher

Dan Boneh

Use Cases

Single use key: (one time key)

- Key is only used to encrypt one message

  - encrypted email: new key generated for every email

Multi use key: (many time key)

- Key used to encrypt multiple messages

  - encrypted files: same key used to encrypt many files

- Need more machinery than for one-time key

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
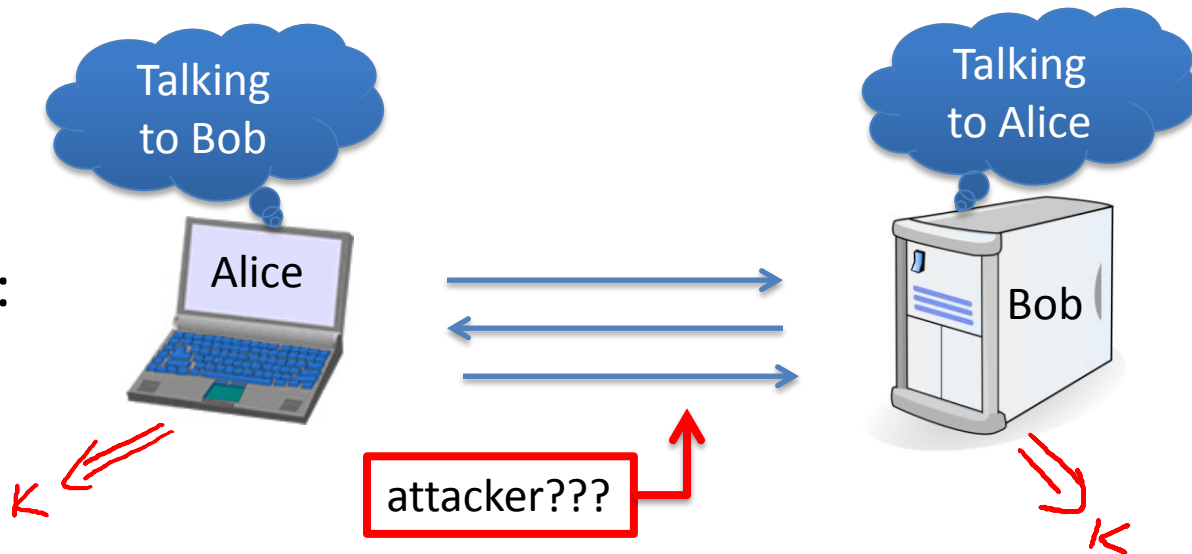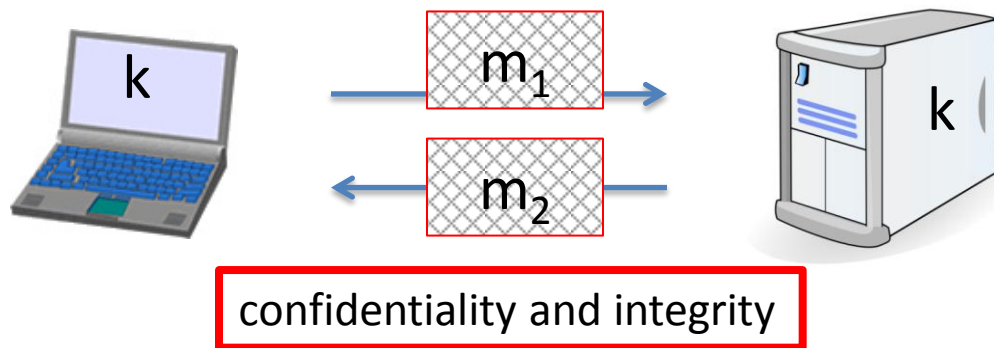
End of Segment

Introduction

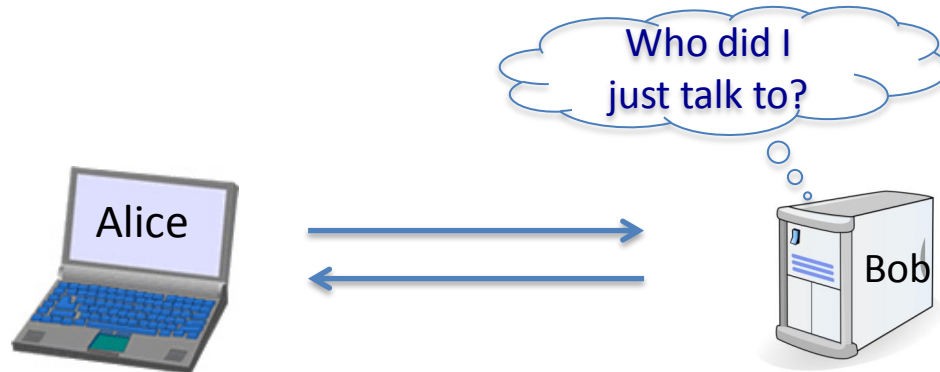## What is cryptography?

Crypto core

Secret key establishment:

Secure communication:

confidentiality and integrity

Dan Boneh

But crypto can do much more

- Digital signatures

- Anonymous communication

Who did I just talk to?
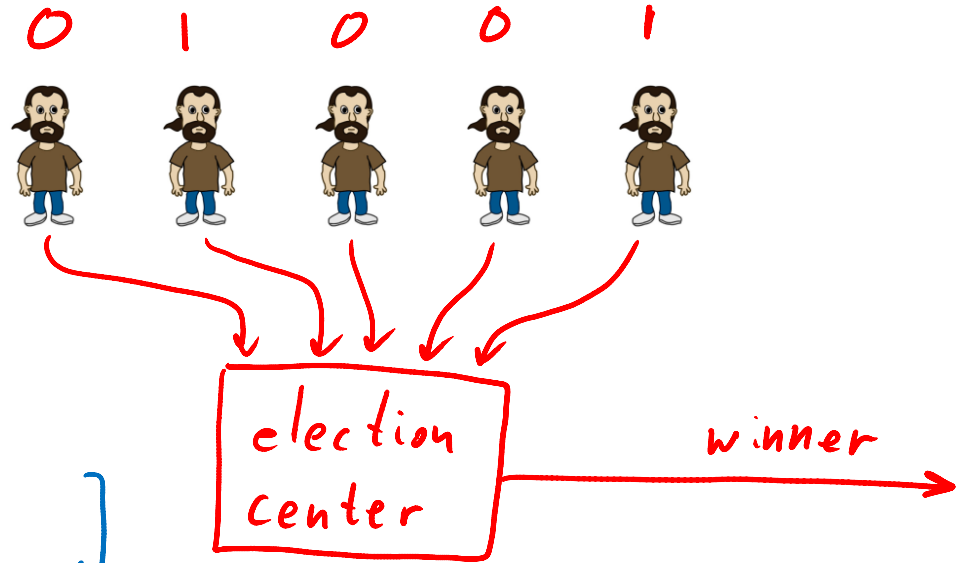
Alice

Bob

Alice signature

But crypto can do much more

- Digital signatures

- Anonymous communication

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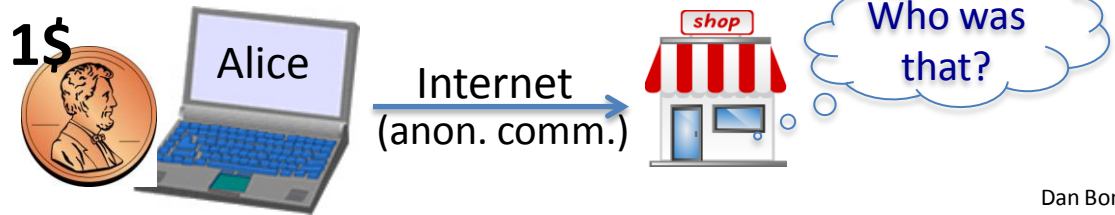
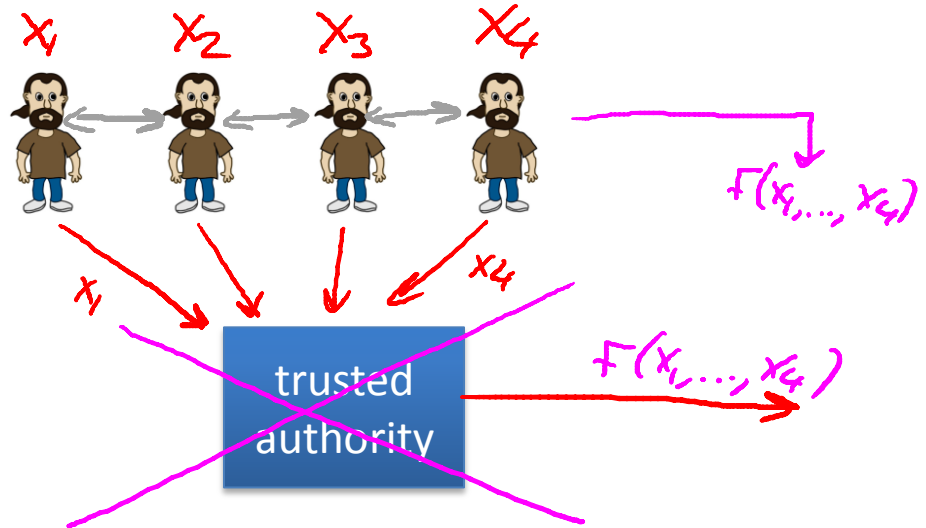# Protocols

- Elections
- Private auctions

winner = MAJ [votes]

auction winner = [highest bidder, pays 2nd highest bid]

$0 \quad 1 \quad 0 \quad 0 \quad 1$

election center

winner

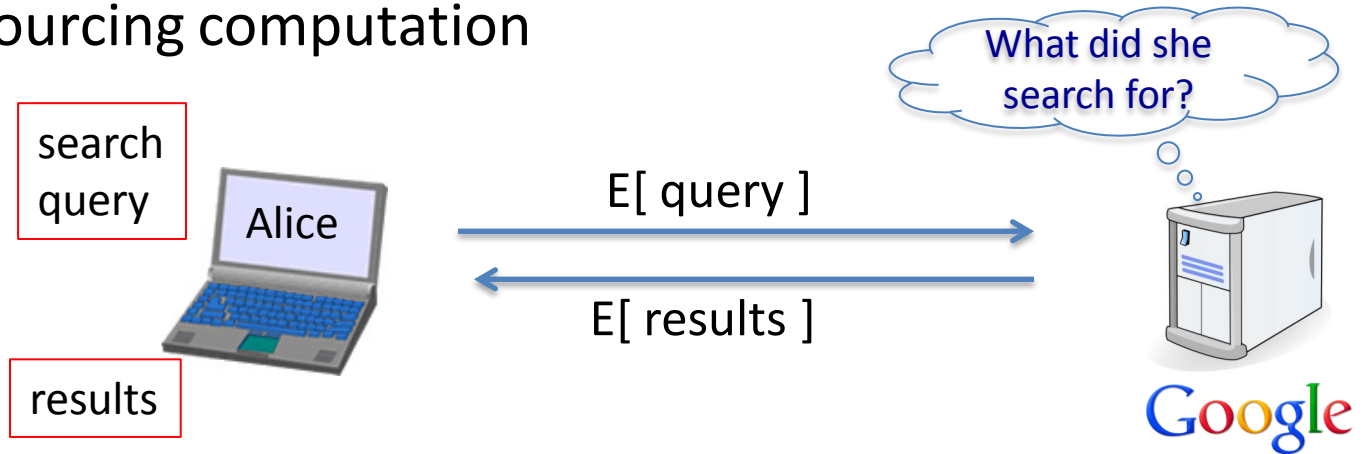# Protocols

- Elections

- Private auctions

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

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

- Secure multi-party computation

x_1 x_2 x_3 x_4

x_1 x_4

trusted authority

$f(x_1, ..., x_4)$

$f(x_1, ..., x_4)$

Crypto magic

- Privately outsourcing computation

search query

Alice

E[query]

E[results]

results

What did she search for?

Google

- Zero knowledge (proof of knowledge)

N=p·q

Alice

I know the factors of N !!

proof π

???

Bob

N

Dan Boneh
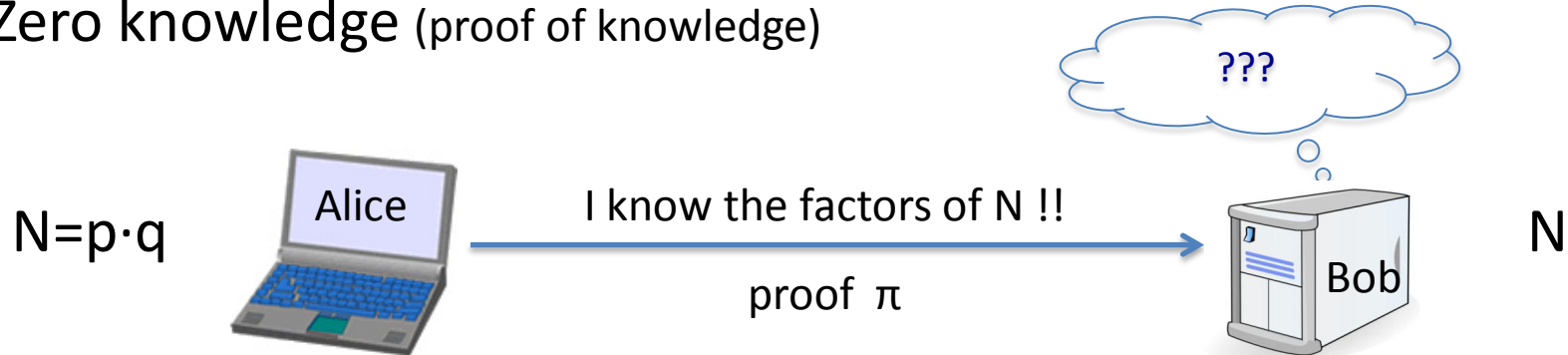
A rigorous science

The three steps in cryptography:

- Precisely specify threat model

- Propose a construction

- Prove that breaking construction under
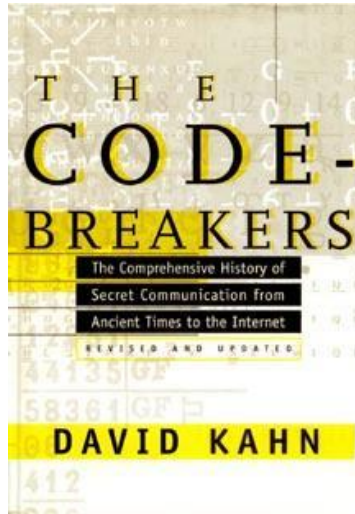 threat mode will solve an underlying hard problem

End of Segment

Introduction

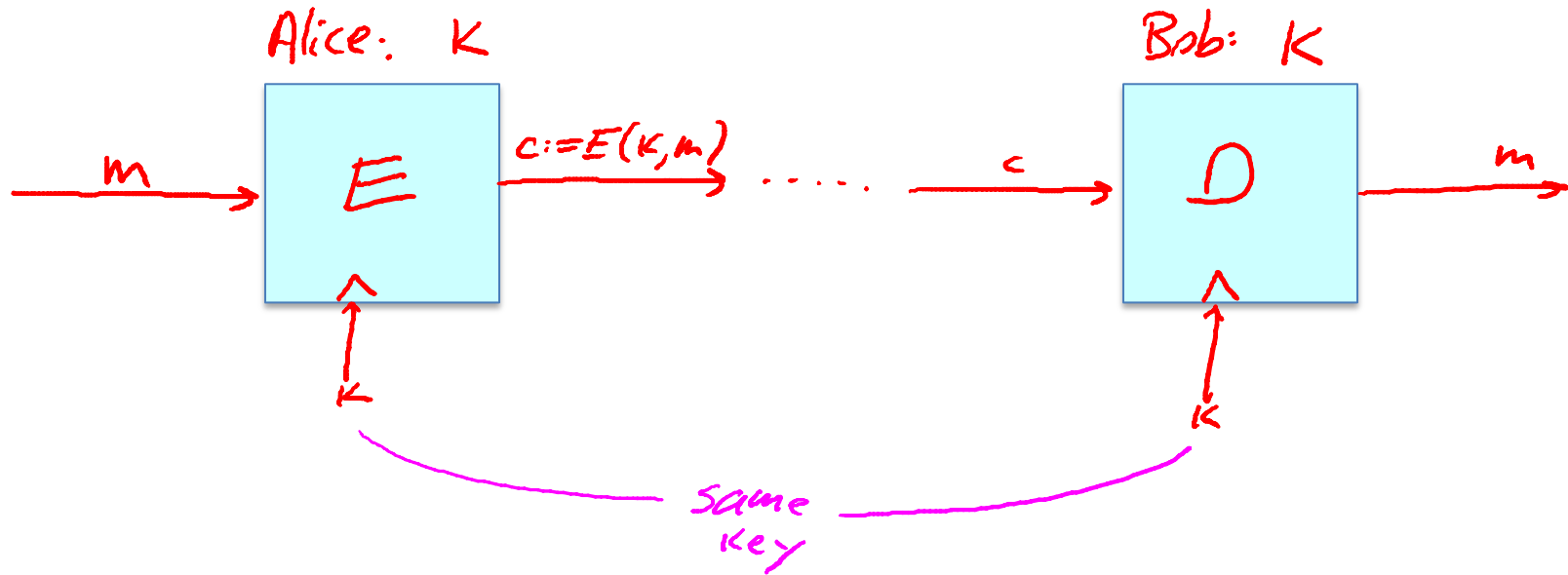## History

History

David Kahn, "The code breakers" (1996)

Symmetric Ciphers

Few Historic Examples   (all badly broken)

1. Substitution cipher

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

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

k :=

$a \longrightarrow c$
$b \longrightarrow w$
$c \longrightarrow n$
$\vdots$
$z \longrightarrow a$

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?

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

$$|\mathcal{K}| = 26! \qquad \text{(26 factorial)}$$

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

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

$26! \approx 2^{88}$

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"

⟹ CT only attack!! 💬

An Example

UKBYBIPOUZBCUFEEBORUKBYBHOBBRFESPVKBWFOFERVNBCVBZPRUBOFERVNBCVBPCYYFVUFO
FEIKNWFRFIKJNUPWRFIPOUNVNIPUBRNCUKBEFWWFDNCHXCYBOHOPYXPUBNCUBOYNRVNIWN
CPOJIOFHOPZRVFZIXUBORJRUBZRBCHNCBBONCHRJZSFWNVRJRUBZRPCYZPUKBZPUNVPWPCYVF
ZIXUPUNFCPWRVNBCVBRPYYNUNFCPWWJUKBYBIPOUZBCUIPOUNVNIPUBRNCHOPYXPUBNCUB
OYNRVNIWNCPOJIOFHOPZRNCRVNBCUNENVVFZIXUNCHPCYVFZIXUPUNFCPWZPUKBZPUNVR

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. Vigener 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 = **Z Z Z J U C L U D T U N W G C Q S**

suppose most common = "H" ⟹ first letter of key = "H" − "E" = "C"
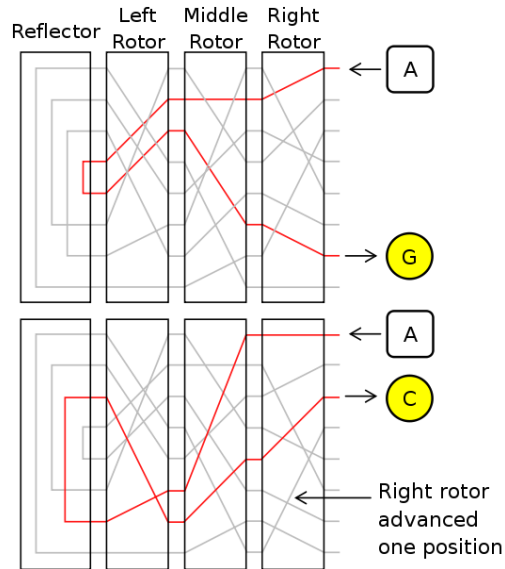
3. Rotor Machines   (1870-1943)

Early example: the Hebern machine   (single rotor)
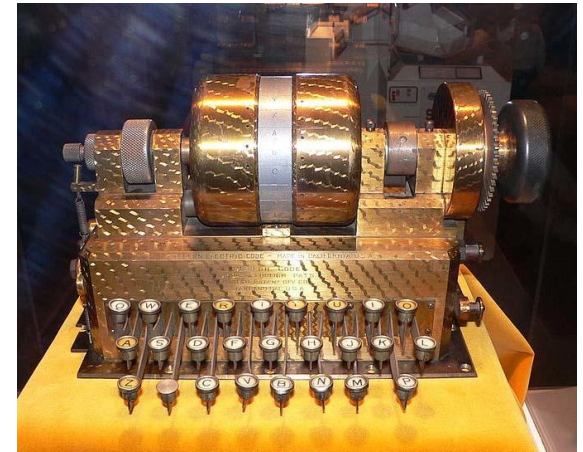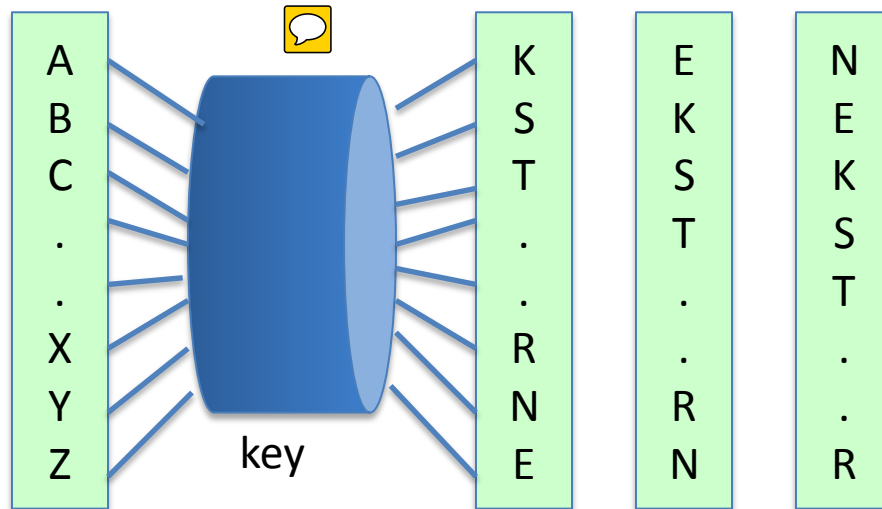
Rotor Machines (cont.)

Most famous: the Enigma (3-5 rotors)

keys = $26^4 = 2^{18}$ (actually 2^{36} due to plugboard)

4. Data Encryption Standard (1974)

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

Today: AES (2001),   Salsa20 (2008)   (and many others)

End of Segment

See also: http://en.wikibooks.org/High_School_Mathematics_Extensions/Discrete_Probability

Introduction

Discrete Probability
(crash course, cont.)

U: finite set   (e.g. $U = \{0,1\}^n$)

Def: **Probability distribution** P over U is a function $P: U \longrightarrow [0,1]$

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

Examples:

1. Uniform distribution: for all $x \in U$: $P(x) = 1/|U|$

2. Point distribution at x_0: $P(x_0) = 1$, $\forall x \neq x_0$: $P(x) = 0$

Distribution vector: $\big($ $P(000), P(001), P(010), \ldots , P(111)$ $\big)$

Events

- For a set $A \subseteq U$: $\quad \Pr[A] = \displaystyle\sum_{x \in A} P(x) \quad \in [0,1]$

note: $\Pr[U] = 1$

- The set A is called an **event**

Example: $\quad U = \{0,1\}^8$

- $A = \{$ all x in U such that $\text{lsb}_2(x) = 11 \} \subseteq U$
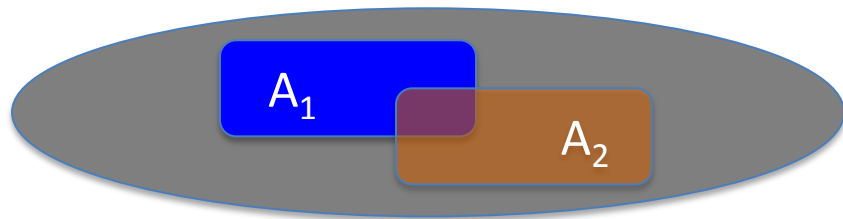
for the uniform distribution on $\{0,1\}^8$: $\quad \Pr[A] = 1/4$

The union bound

- For events A_1 and A_2

$$\Pr[\ A_1 \cup A_2\ ] \leq \Pr[A_1] + \Pr[A_2]$$

$A_1 \cap A_2 = \phi \Rightarrow \Pr[A_1 \cup A_2] = \Pr[A_1] + \Pr[A_2]$

Example:

$A_1 = \{\ $ all x in $\{0,1\}^n$ s.t $lsb_2(x)=11\ \}$; $A_2 = \{\ $ all x in $\{0,1\}^n$ s.t. $msb_2(x)=11\ \}$

$$\Pr[\ lsb_2(x)=11 \text{ or } msb_2(x)=11\ ] = \Pr[A_1 \cup A_2] \leq \tfrac{1}{4}+\tfrac{1}{4} = \tfrac{1}{2}$$

Random Variables

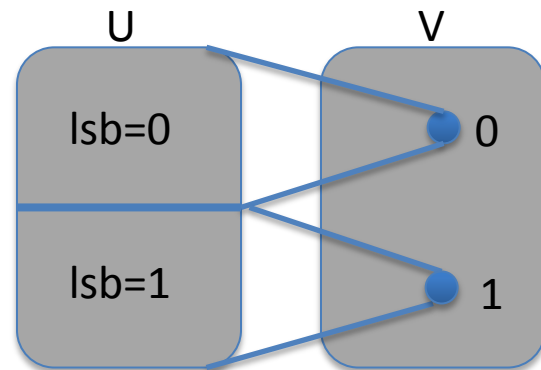Def: a random variable X is a function $X: U \longrightarrow V$

Example: $X: \{0,1\}^n \longrightarrow \{0,1\}$; $X(y) = lsb(y)$ $\in \{0,1\}$

For the uniform distribution on U:

$$Pr[\ X=0\ ] = 1/2 \quad , \quad Pr[\ X=1\ ] = 1/2$$

More generally:

rand. var. X induces a distribution on V: $Pr[\ X=v\ ] := Pr[\ X^{-1}(v)\ ]$

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 $x \in U$)

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

inputs   outputs
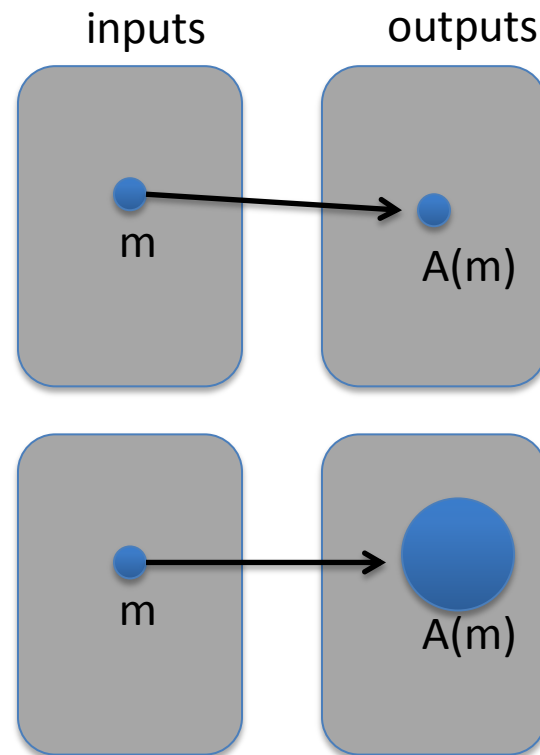
- Deterministic algorithm: $y \longleftarrow A(m)$

- Randomized algorithm

    $y \longleftarrow A(m ; r)$ where $r \xleftarrow{R} \{0,1\}^n$

  output is a random variable

    $y \xleftarrow{R} A(m)$

Example: $A(m ; k) = E(k, m)$, $y \xleftarrow{R} A(m)$

End of Segment

See also: http://en.wikibooks.org/High_School_Mathematics_Extensions/Discrete_Probability

Introduction

Discrete Probability
(crash course, cont.)

Recap

U: finite set   (e.g. $U = \{0,1\}^n$)

Prob. distr. P over U is a function $P: U \longrightarrow [0,1]$ s.t. $\sum_{x \in U} P(x) = 1$

$A \subseteq U$ is called an **event**   and $Pr[A] = \sum_{x \in A} P(x) \in [0,1]$

A **random variable** is a function $X: U \longrightarrow V$.

     X takes values in V and defines a distribution on V

Independence

<u>Def</u>:   events A and B are **independent** if   Pr[A and B] = Pr[A] · Pr[B]

random variables X,Y taking values in V are **independent** if

∀a,b∈V: Pr[X=a and Y=b] = Pr[X=a] · Pr[Y=b]

<u>Example</u>:   U = {0,1}^{2} = {00, 01, 10, 11}   and   r $\xleftarrow{R}$ U

Define r.v. X and Y as: X = lsb(r)   ,   Y = msb(r)

Pr[X=0   and Y=0] = Pr[r=00] = ¼ = Pr[X=0] · Pr[Y=0]

Review: XOR

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

An important property of XOR

Thm: Y a rand. var. over $\{0,1\}^n$,   X an indep. uniform var. on $\{0,1\}^n$

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

Proof: (for n=1)

$\Pr[\ Z=0\ ] = \Pr\left[(x,y)=(0,0) \text{ or } (x,y)=(1,1)\right] =$

$= \Pr\left[(x,y)=(0,0)\right] + \Pr\left[(x,y)=(1,1)\right] =$

$= \dfrac{P_0}{2} + \dfrac{P_1}{2} = \dfrac{1}{2}$

Y	Pr
0	P_0
1	P_1

X	Pr
0	½
1	½

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 indep. identically distributed random vars.

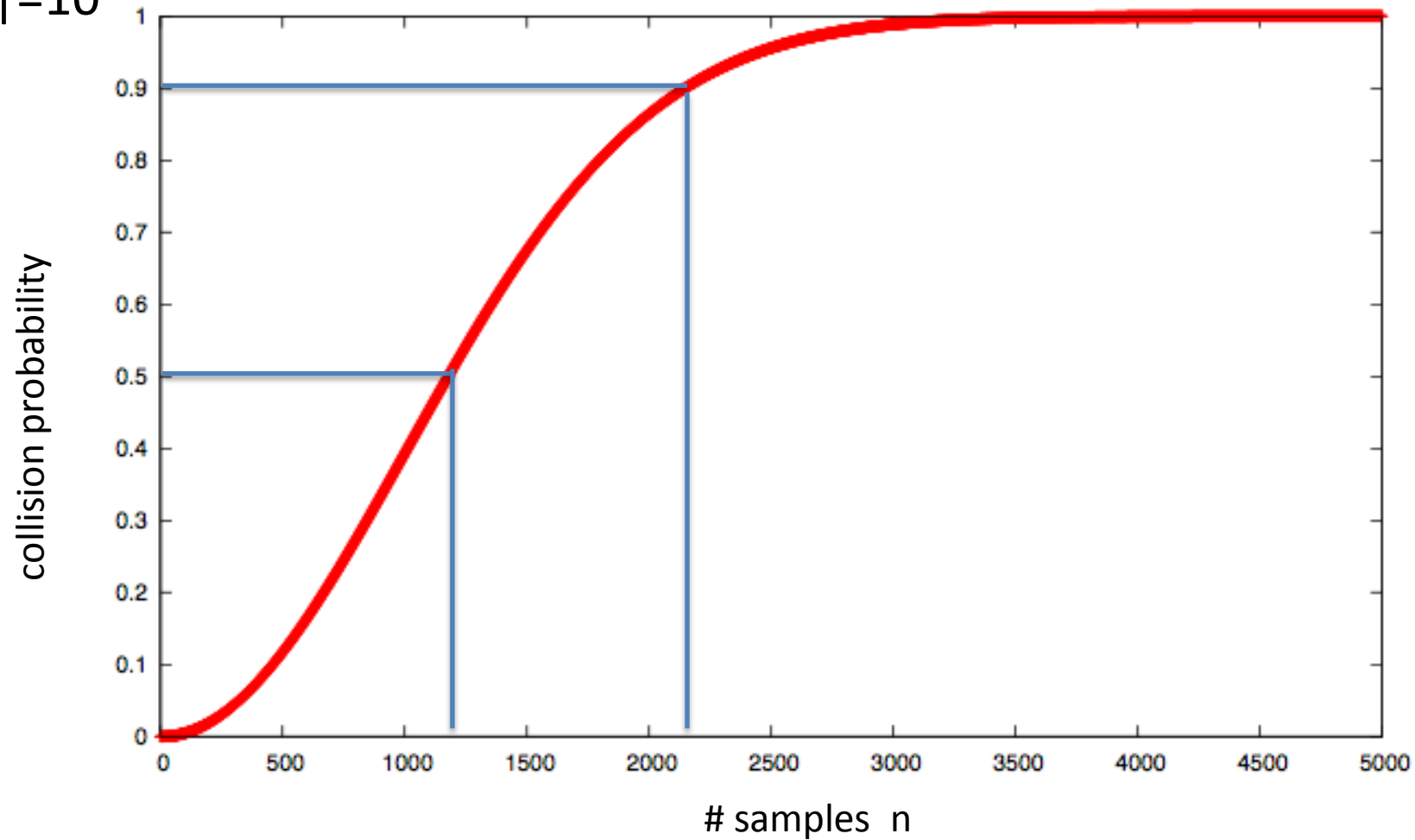**Thm**: when $n = 1.2 \times |U|^{1/2}$ then $\Pr\left[\, \exists i \neq j: \ r_i = r_j \,\right] \geq \tfrac{1}{2}$

notation: $|U|$ is the size of U

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

samples n

Dan Boneh

End of Segment