Unit 6

Cryptography

Outline of Unit 6

- 6.1 Chinese Remainder Theorem
- 6.2 Symmetric Key Cryptography
- 6.3 Public Key Cryptography

Class Activity

- Pick a natural number smaller than 100.
- Divide it by 3 and tell me the remainder.
- Divide it by 5 and tell me the remainder.
- Divide it by 7 and tell me the remainder.
- Then I can tell you what the number is.

A Chinese Poem (just for fun)

3 三人同行七十稀， 70

5 五樹梅花廿一枝， 21

7 七子團圓正半月， 15

除百零五便得知。 105

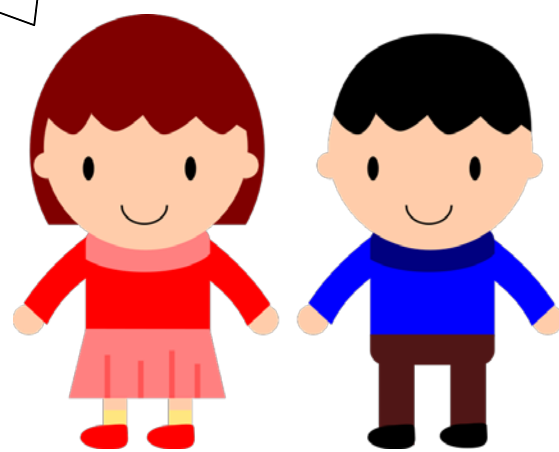
Unit 6.1

Chinese Remainder Theorem

Problem about Last Digit

When x is divided by 2,
the remainder is 1.
When x is divided by 5,
the remainder is 3.
What is the last digit of x?

That's simple…



Modulo mn

- In the previous problem,
$$x \equiv 1 \pmod{2},$$
$$x \equiv 3 \pmod{5},$$
$$x \equiv ? \pmod{10}.$$

- It is easy to find that
$$x \equiv 3 \pmod{10}.$$

- In general,
$$x \equiv a \pmod{m},$$
$$x \equiv b \pmod{n},$$
$$x \equiv c \pmod{mn}.$$

1. Given c, can we always determine a and b?
2. Given a and b, can we always determine c?

Modulo mn

- Consider
$$x \equiv 1 \pmod{2},$$
$$x \equiv 3 \pmod{4},$$
$$x \equiv ? \pmod{8}.$$

- The solution is *not* unique:
  - e.g. x can be 3 or 8.

- Assume m and n are co-prime.
$$x \equiv a \pmod{m},$$
$$x \equiv b \pmod{n},$$
$$x \equiv c \pmod{mn}.$$

- Given a and b, can we uniquely determine c?

Definition of the Function f

- Define a function $f: R_{mn} \longrightarrow R_m \times R_n$ as follows:

$$f(c) = (a, b),$$

where

$$c \equiv a \pmod{m}, \qquad c \equiv b \pmod{n}.$$

and

$$R_{mn} \triangleq \{0, 1, 2, \ldots, mn - 1\},$$
$$R_m \triangleq \{0, 1, 2, \ldots, m - 1\},$$
$$R_n \triangleq \{0, 1, 2, \ldots, n - 1\}.$$

Example

- Consider $m = 3, n = 5$.
- $f(c) = (c \bmod 3, c \bmod 5)$, where $0 \le c \le 14$.

$x \equiv a \pmod{3}$,
$x \equiv b \pmod{5}$,
$x \equiv c \pmod{15}$.

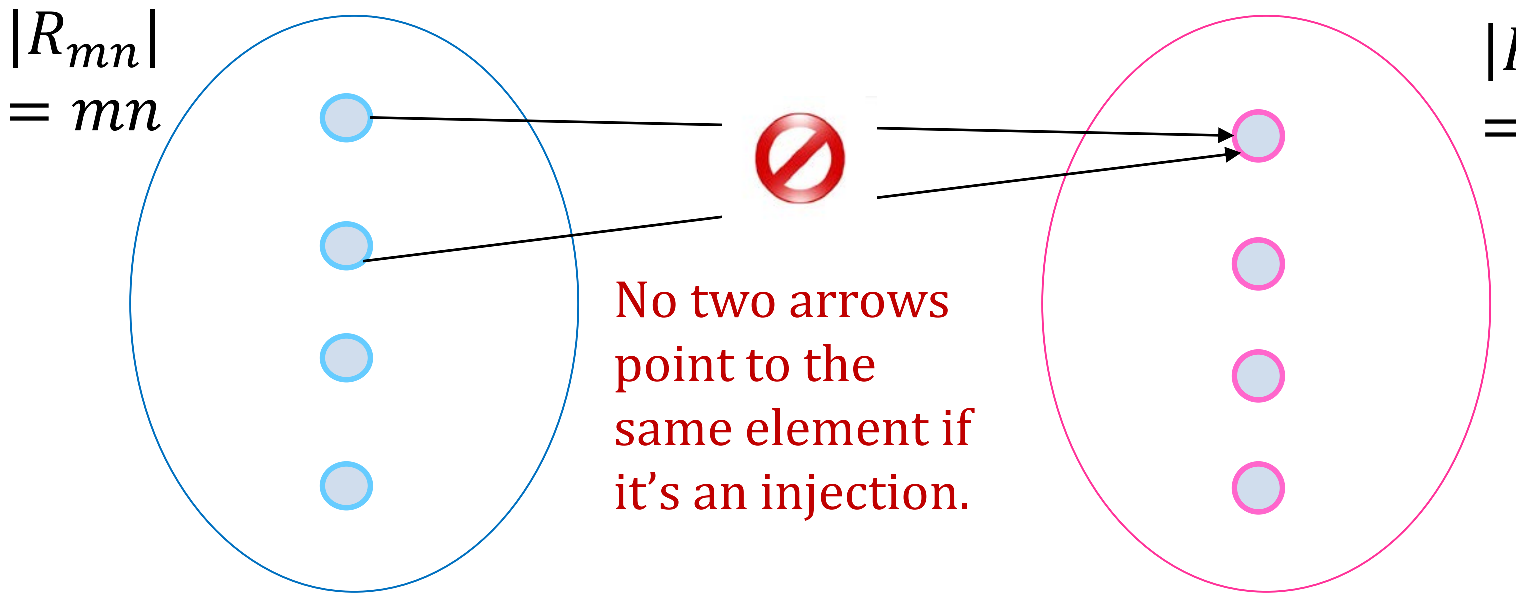
	$b = 0$	1	2	3	4
$a = 0$	0	6	12	3	9
$a = 1$	10	1	7	13	4
$a = 2$	5	11	2	8	14

$f(14)$
$= (14 \bmod 3,\ 14 \bmod 5)$
$= (2,\ 4)$

Bijection

Lemma: f is a bijection if m, n are co-prime.

This result immediately implies CRT (to be discussed next).



proof idea

Proof

- Since the domain and co-domain have the same size, if f is an injection, it must be a surjection.
- It is sufficient to prove that f is an injection.
  - Suppose $f(c_1) = f(c_2)$.
  - Then $c_1 \equiv c_2 \pmod{m} \Longrightarrow m \mid (c_1 - c_2)$.
  - Similarly, $c_1 \equiv c_2 \pmod{n} \Longrightarrow n \mid (c_1 - c_2)$.
  - By Unique Factorization Theorem, both m and n can be uniquely factorized into prime factors.
  - Since m, n are co-prime, they have no common prime factors.
  - Therefore, all their prime factors are contained in $(c_1 - c_2)$, so $mn \mid (c_1 - c_2)$.
  - $c_1 \equiv c_2 \pmod{mn}$ *Q.E.D.*

Chinese Remainder Theorem (CRT)

Theorem: Let m and n be co-primes. Consider the system of two linear congruences:

$$x \equiv a \pmod{m},$$
$$x \equiv b \pmod{n},$$

where $0 \leq a < m,$ and $0 \leq b < n$.

There exists a unique solution $0 \leq c < mn$ such that

$$x \equiv c \pmod{mn}.$$

The result can be generalized to more than two congruences.

Problem Statement

- We use another notation, which can be easily generalized to arbitrary number of equations:

$$x \equiv a_1 \pmod{m_1},$$
$$x \equiv a_2 \pmod{m_2},$$
$$x \equiv c \pmod{m_1 m_2}.$$

- Given m_1, m_2 that are co-primes, we want to find the value of c.

Solution

- Since m_1, m_2 are co-primes, $\gcd(m_1, m_2) = 1$.
- $m_1\alpha_2 + m_2\alpha_1 = 1$ for some integers α_1, α_2.
  - α_1, α_2 can be found by extended Euclidean algorithm.
- Note that
  - $m_1\alpha_2 \equiv 1 \pmod{m_2},\ m_2\alpha_1 \equiv 1 \pmod{m_1}$

$$\boxed{x \equiv c \equiv a_1 m_2 \alpha_1 + a_2 m_1 \alpha_2 \pmod{m_1 m_2}}$$

- It is easy to verify that
  - $x \equiv a_1 \pmod{m_1}$ and $x \equiv a_2 \pmod{m_2}$

Corollary

Consider the special case where $a_1 = a_2 = a$.

$$x \equiv a \pmod{m_1},$$
$$x \equiv a \pmod{m_2}.$$

If m_1, m_2 are co-primes, then

$$x \equiv a \pmod{m_1 m_2}.$$

Proof: In the previous slide,

$$c = a_1 m_2 \alpha_2 + a_2 m_1 \alpha_1 = a(m_2 \alpha_2 + m_1 \alpha_1) = a$$

Q.E.D.

Useful for proving the correctness of RSA.

Example

$$x \equiv 3 \pmod{19},$$
$$x \equiv 8 \pmod{11}$$

We use extended Euclidean algorithm to find

$$19\alpha_2 + 11\alpha_1 = 1,$$

where $\alpha_2 = -4$ and $\alpha_1 = 7$.

19	11		
1	0	19	(a)
0	1	11	(b)
1	–1	8	(c) = (a) – 1(b)
–1	2	3	(d) = (b) – 1(c)
3	–5	2	(e) = (c) – 2(d)
–4	7	1	(f) = (d) – 1(e)

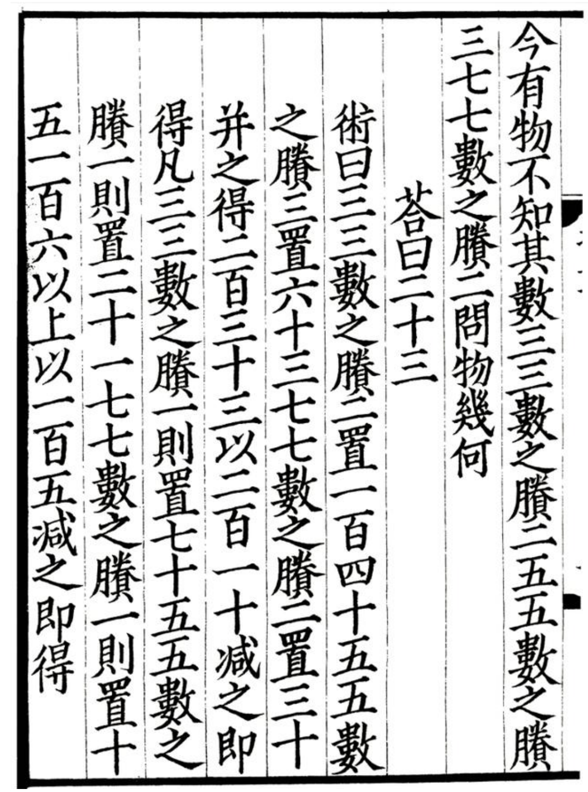
Hence, $c = a_1 m_2 \alpha_1 + a_2 m_1 \alpha_2 \pmod{m_1 m_2}$

$$= 3(11)(7) + 8(19)(-4) \pmod{209}$$
$$= 41$$

A Problem from Sunzi Suanjing (孫子算經)

- There are certain things whose number is unknown.
- Repeatedly divided by 3, the remainder is 2; by 5 the remainder is 3; and by 7 the remainder is 2.
- What will be the number?

$$x \equiv 2 \pmod{3}$$
$$x \equiv 3 \pmod{5}$$
$$x \equiv 2 \pmod{7}$$



今有物不知其數三三數之賸二五五數之賸
三七七數之賸二問物幾何
荅曰二十三
術曰三三數之賸二置一百四十五五數
之賸三置六十三七七數之賸二置三十
并之得二百三十三以二百一十減之即
得凡三三數之賸一則置七十五五數之
賸一則置二十一七七數之賸一則置十
五一百六以上以一百五減之即得

Sunzi Suanjing (孫子算經), an ancient Chinese math book (circa 300 A.D.)

CRT (General Case)

Theorem: Let m_i be pairwise co-primes. Consider the system of linear congruences:

$$x \equiv a_i \pmod{m_i},$$

where $0 \leq a_i < m_i$.

There exists a unique solution $0 \leq c < M$, such that

$$x \equiv c \pmod{M},$$

where $M = \prod_i m_i$

the product of all m_i's

Solution

- Define $M_i = \frac{M}{m_i}$.
  - i.e. the product of all moduli excluding m_i.
- Define $\alpha_i \equiv M_i^{-1} \pmod{m_i}$.
- The solution is given by

M_i^{-1} exists because M_i and m_i are co-prime.

$$\boxed{c = \sum_i a_i M_i \alpha_i \pmod{M}.}$$

- It can be verified that for each congruence j,

$$\sum_i a_i M_i \alpha_i \pmod{m_j} = a_j M_j \alpha_j \pmod{m_j} = a_j.$$

Solution to the Standard Problem

$$x \equiv a_1 (\text{mod } 3)$$
$$x \equiv a_2 (\text{mod } 5)$$
$$x \equiv a_3 (\text{mod } 7)$$

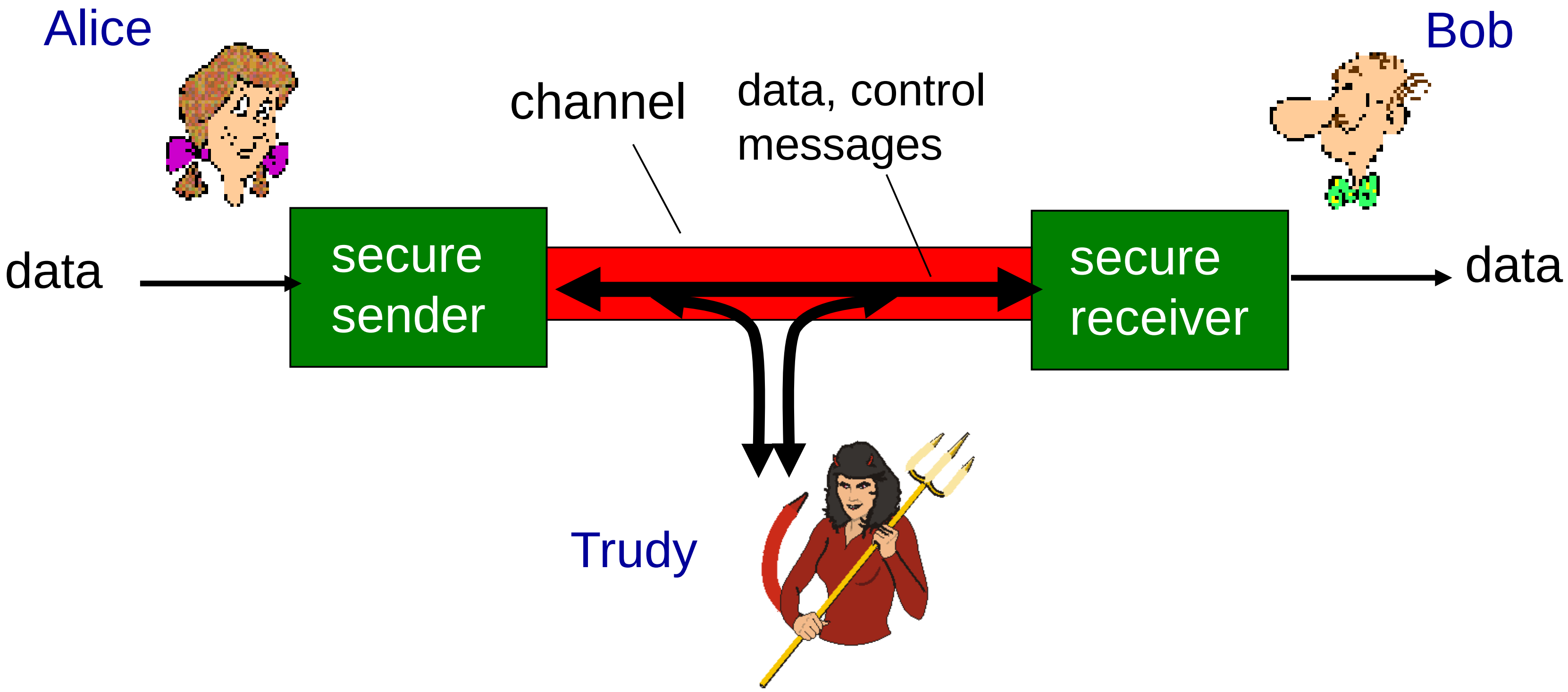
- $M_1 = 5 \times 7 = 35,\ \alpha_1 \equiv 35^{-1} \equiv 2^{-1} \equiv 2\ (\text{mod } 3)$
- $M_2 = 3 \times 7 = 21,\ \alpha_2 \equiv 21^{-1} \equiv 1^{-1} \equiv 1\ (\text{mod } 5)$
- $M_3 = 3 \times 5 = 15,\ \alpha_3 \equiv 15^{-1} \equiv 1^{-1} \equiv\ 1\ (\text{mod } 7)$
- $c = a_1 M_1 \alpha_1 + a_2 M_2 \alpha_2 + a_3 M_3 \alpha_3\ (\text{mod } M)$
  $= 70a_1 + 21a_2 + 15a_3 (\text{mod } M)$

Unit 6.2

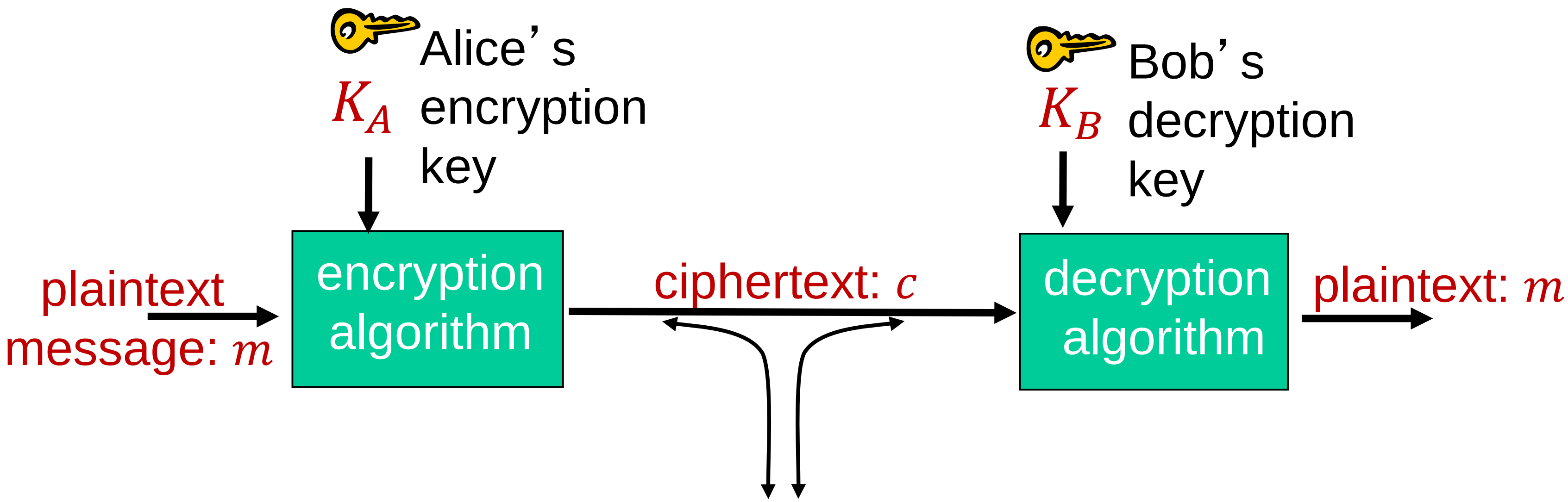
Symmetric Key Cryptography

Secure Communications

- Bob & Alice want to communicate "securely"
- Trudy (intruder) may intercept, delete, add messages

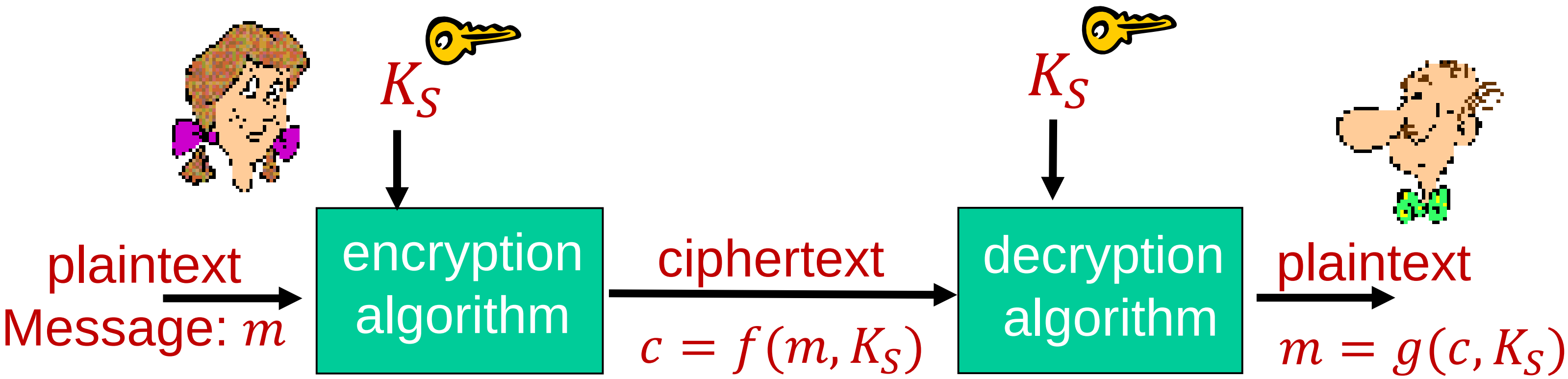


The Language of Cryptography



- Encryption: $c = f(m, K_A)$
- Decryption: $m = g(c, K_B)$:

Symmetric Key Cryptography



- The encryption and decryption algorithms (i.e. the functions f and g) are assumed known to the public.
  - For many applications, it is difficult to keep the algorithms as a secret.
- Alice and Bob share the same (symmetric) key: K_S
  - The key is private (known only by Alice and Bob).

Caesar Cipher (~58 BC)

- Named after Julius Caesar, who used it in his private correspondence.
- Each letter in the plaintext is (cyclically) shifted by a fixed number of positions down the alphabet.
  - e.g. if the shift is 3, a→d, b→e, …, z→c
- The symmetric key K_S is the number of shift positions.
- Decrypt it without knowing K_S!

```
k ecog k ucy k eqpswgtgf
```

Julius Caesar, arguably the greatest of the dictators of Rome, ruling from 49 BC to 44 BC.

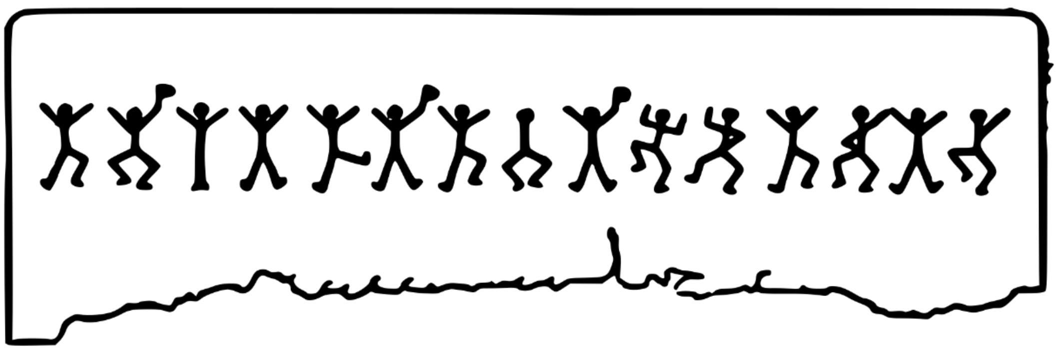
Substitution Cipher

- Substitution cipher: replace one thing by another.
  - Caesar cipher is a special case of substitution cipher.
- Example: (replace each letter by another)
  - plaintext: `abcdefghijklmnopqrstuvwxyz`
  - ciphertext: `mnbvcxzasdfghjklpoiuytrewq`
- The symmetric key K_S is the mapping.
- Decrypt the ciphertext below using the above key:

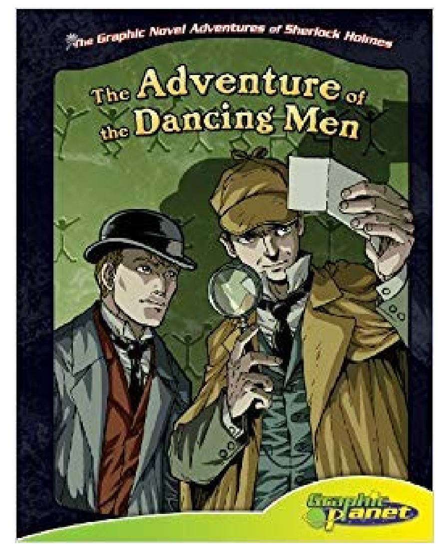
`nkn s gktc wky mgsbc`

Dancing Men Cipher (1892)

Holmes held up the paper so that the sunlight shone full upon it. It was a page torn from a note-book. The markings were done in pencil, and ran in this way:—



An extract from *The Adventure of the Dancing Men*



How to crack the code?
https://www.youtube.com/watch?v=oKMAiIL_1V8&t=240s (first 3.5 min)

Hill Cipher (1929)

- A polygraphic substitution cipher, invented by Lester S. Hill.
- Each letter is represented by a number modulo 26 (i.e., A = 0, B = 1, …, Z = 25).
- A vector of n letters is encrypted by multiplication with an $n \times n$ invertible matrix (mod 26), which is the secret key.
- Decryption is done by multiplication with the inverse of the encryption matrix.

Example: Encryption Matrix

The plaintext is "ACT".

$$\begin{bmatrix} 0 \\ 2 \\ 19 \end{bmatrix}$$

The encryption matrix.

$$\begin{bmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{bmatrix}$$

It must be invertible.

```
R = IntegerModRing(26)
E = Matrix(R,[[6,24,1],[13,16,10],[20,17,15]])
E.is_invertible()
```

You can use SageMath to verify that the matrix is invertible.
https://sagecell.sagemath.org/

Example: Encryption

The ciphertext is obtained by

$$\begin{bmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 19 \end{bmatrix} = \begin{bmatrix} 67 \\ 222 \\ 319 \end{bmatrix} \equiv \begin{bmatrix} 15 \\ 14 \\ 7 \end{bmatrix} \pmod{26}$$

```
R = IntegerModRing(26)
E = Matrix(R,[[6,24,1],[13,16,10],[20,17,15]])
m = vector(R,[0,2,19])
E * m
```

Example: Decryption Matrix

The decryption matrix.

$$\begin{bmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{bmatrix}^{-1} \equiv \begin{bmatrix} 8 & 5 & 10 \\ 21 & 8 & 21 \\ 21 & 12 & 8 \end{bmatrix} \pmod{26}$$

```
R = IntegerModRing(26)
E = Matrix(R,[[6,24,1],[13,16,10],[20,17,15]])
E.inverse()
```

Example: Decryption

The plaintext is obtained by

$$\begin{bmatrix} 8 & 5 & 10 \\ 21 & 8 & 21 \\ 21 & 12 & 8 \end{bmatrix} \begin{bmatrix} 15 \\ 14 \\ 7 \end{bmatrix} \equiv \begin{bmatrix} 260 \\ 574 \\ 539 \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 2 \\ 19 \end{bmatrix} \pmod{26}$$

```
R = IntegerModRing(26)
E = Matrix(R,[[6,24,1],[13,16,10],[20,17,15]])
m = vector(R,[0,2,19])
c = E * m
D = E.inverse()
D * c
```

Remarks on Hill Cipher

- Encryption matrix must be invertible.
- According to Cramer's rule, matrix inverse is computed by "division by determinant".
  - Determinant is not equal to zero.
  - Determinant is co-prime with the modular base (which can be guaranteed if the base is chosen as a prime number).
- Matrix multiplication alone is not secure.
  - Vulnerable to known-plaintext attack because a system of linear equations is easy to solve.
  - It is still useful when combining with other non-linear operations.

The One-Time Pad (OTP) (1882)

- In the substitution cipher, every occurrence of an object is replaced by the same object.
  - Easy to break (via statistical analysis) for long messages.
- Let *n* be the length of the plaintext.
- The symmetric key K_S is a list of *n* random shifts.
- Example:
  - `number theory is the queen of mathematics`
  - `354123 ...... .. ... ..... .. ...........`
  - `qzqcgu ...... .. ... ..... .. ...........`

  Note: only the first word is shown.
- The scheme is perfectly secure, but the size of the key is as large as the message.

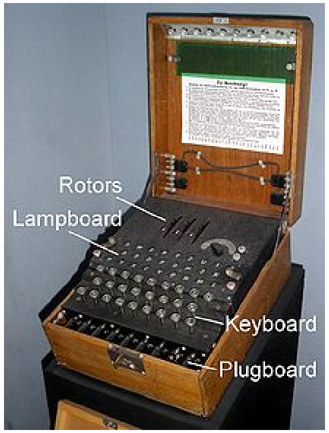
More on OTP

- Watch the following 3-min video:
  - https://www.youtube.com/watch?v=FlIG3TvQCBQ
- "One-time" because the key *cannot* be reused.
  - If reused, the scheme becomes insecure.
- If the plaintext m is encoded as a bit sequence of length n, then the symmetric key K_S is a random binary sequence of length n.
- Ciphertext c is obtained by bitwise-XOR between m and K_S, i.e.,

$$c = m \oplus K_S$$

The Enigma Machine (1918)

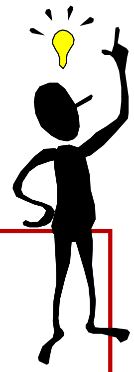
- Invented by Arthur Scherbius in 1918, right at the end of World War I.
- Early models were used commercially from the early 1920s.
- Adopted by Nazi Germany before and during World War II.
- Alan Turing, a mathematician, cracked the code, which shortened the war by more than two years.
- https://www.youtube.com/watch?v=G2_Q9FoD-oQ (12 min)



Unit 6.3

Public Key Cryptography

Symmetric Key vs Public Key



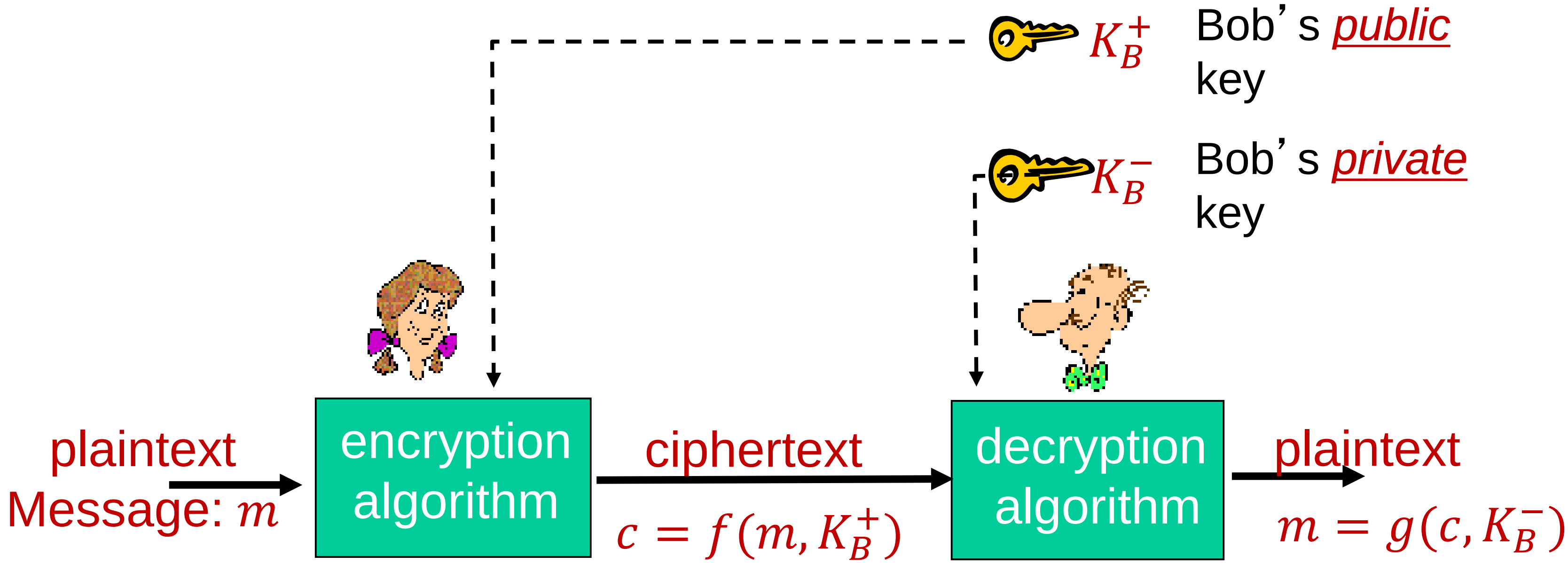
Symmetric key crypto

- requires sender and receiver know a shared secret key
- Q: how to agree on the key in the first place (particularly if never "met")?

Public key crypto

- radically different approach [Diffie-Hellman76, RSA78]
- sender, receiver do *not* share secret key
- *public* encryption key known to *all*
- *private* decryption key known only to receiver

Public Key Cryptography



Additional requirement:

- Given the public key K_B^+, it should be (almost) impossible to compute the private key K_B^-.

Practical Use

- The following two-step approach is commonly used (e.g., in HTTPS):

1) Use public key to privately share a session key (i.e. a symmetric key)
   - Public key crypto has a lot of overhead.
   - This step is done only at the beginning of a communication session.
2) Use symmetric key to encrypt data
   - Symmetric key crypto is quicker and uses less resource.

RSA Cryptosystem

- By Rivest, Shamir, Adleman of MIT in 1977.
- Best known and widely used public-key scheme.
- Use large integers (e.g. 1024 bits)
- Security due to the difficulty of factoring large numbers.

Is factorization difficult?

RSA Challenge

- Can you factorize the following number (which has 617 digits, or 2048 bits)?

2519590847565789349402718324004839857142928212620
4032027777137836043662020707595556264018525880784
4069182906412495150821892985591491761845028084891
2007284499268739280728777673597141834727026189637
5014971824691165077613379859095700097330459748808
4284017974291006424586918171951187461215151726546
3228221686998754918242243363725908514186546204357
6798423387184774447920739934236584823824281198163
8150106748104516603773060562016196762561338441436
0383390441495263443219011465754445417842402092461
6515723350778707749817125772467962926386356373289
9121548314381678998850404453640235273819513786365
64391212010397122822120720357

RSA: Getting Ready

- Message: just a bit pattern
  - bit pattern can be uniquely represented by an integer
  - thus, encrypting a message is equivalent to encrypting a number.
- Example:
  - $M = 10010001$. This message is uniquely represented by the decimal number 145.
  - To encrypt M, we encrypt the corresponding number, which gives a new number (the ciphertext).

Key Generation

- Bob generates two large distinct random primes, p and q.
- Compute $N = pq$ and $\phi(N) = (p-1)(q-1)$.
- Choose at random e (with $1 < e < \phi(N)$) which is co-prime with $\phi(N)$.
- Solve the following equation to find d:

$$ed \equiv 1 \pmod{\phi(N)}$$

  - The inverse of e exists, since e and $\phi(n)$ are coprime.
- Bob publishes his public key K_B^+: (N, e).
- He keeps secret his private key K_B^-: (N, d).

Encryption and Decryption

1. To encrypt message M, Alice uses Bob's public key K_B^+: (N, e) to compute
$$C = M^e \pmod N$$
   - Note that M must be smaller than N (break down into blocks if necessary).
2. After receiving the ciphertext, Bob uses his private key K_B^-: (N, d) to compute
$$M' = C^d \pmod N$$

magic happens!	$M' = M$

RSA Toy Example

- Bob chooses $p = 5, q = 7$.
- Then $N = 35, \phi(n) = 24$.
- Suppose $e = 5$ is chosen (so $e, \phi(n)$ are co-prime)
- Compute $d = 29$ (by xgcd so that $ed \equiv 1 \pmod{\phi(n)}$)
- Encrypt 8-bit message

encrypt: $\underbrace{\text{bit pattern}}_{00001100}$ $\quad \underbrace{M}_{12}$ $\quad \underbrace{M^e}_{248832}$ $\quad \underbrace{C = M^e \bmod N}_{17}$

decrypt: $\underbrace{C}_{17}$ $\quad \underbrace{C^d}_{481968572106750915091411825223071697}$ $\quad \underbrace{M' = C^d \bmod N}_{12}$

Fast exponentiation should be used instead!

Why does RSA work?

- We want to show that

$$M' \equiv C^d \equiv M^{ed} \equiv M \pmod{N},$$

  or

$$M^{ed} \equiv M \pmod{pq}.$$

- By the corollary of CRT, it suffices to prove that

$$M^{ed} \equiv M \pmod{p},$$
$$M^{ed} \equiv M \pmod{q},$$

  for all M.

Case 1: $M \equiv 0 \pmod p$

- It implies $M^{ed} \equiv 0 \pmod p$.
- Hence, $M^{ed} \equiv M \pmod p$.

Case 2: $M \not\equiv 0 \pmod p \implies p \nmid M$

- Since $ed \equiv 1 \pmod{\phi(N)}$, we can write

$$ed = 1 + k\phi(N) \text{ for some integer } k.$$

- $M^{ed} \equiv M^{1+k\phi(n)} \equiv M(M^{p-1})^{k(q-1)} \pmod p$.
- Since $p \nmid M$, by Fermat's Little Theorem,

$$M^{p-1} \equiv 1 \pmod p.$$

- Hence, $M^{ed} \equiv M \pmod p$.

Similarly, we can show that $M^{ed} \equiv M \pmod q$.

Q.E.D.

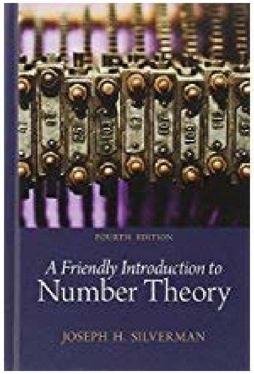
Why is RSA secure?

- Given the public key (N, e), how hard is it to determine the private key (N, d)?
- One needs to solve the following formula:

$$ed \equiv 1 \pmod{\phi(N)}$$

- But $\phi(N)$ is not known, since p, q are not known.
- If N is a large number, it is very hard to factorize it into pq.
- It is also very hard to find $\phi(N)$ directly.
  - Otherwise, N can be factorized easily, since p and q can be obtained easily from $\phi(N)$ and N.
    - $N = pq$
    - $\phi(N) = (p-1)(q-1) = pq - p - q + 1$

Recommended Reading



- Chapter 11, J. H. Silverman, *A Friendly Introduction to Number Theory*, 4th ed., Pearson, 2013.

- Section 8.2, J. Kurose and K. Ross, *Computer Networking: a top-down approach*, 6th ed., Prentice Hall, 2010.

Supplementary Materials on Enigma (optional)

- Flaw in the Enigma (11 min):
  - https://www.youtube.com/watch?v=V4V2bpZlqx8

- The Imitation Game
  - A movie based on the biography of Alan Turing
  - Trailer (2.5 min): https://www.youtube.com/watch?v=aG4-C4bGAw4