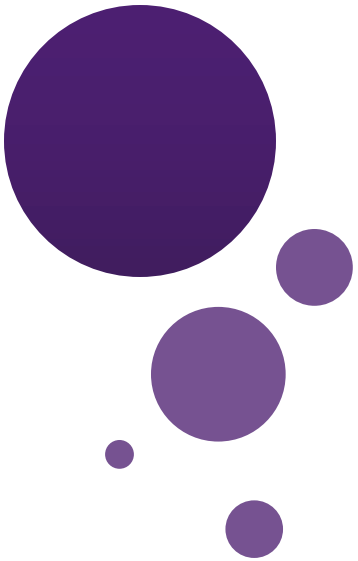
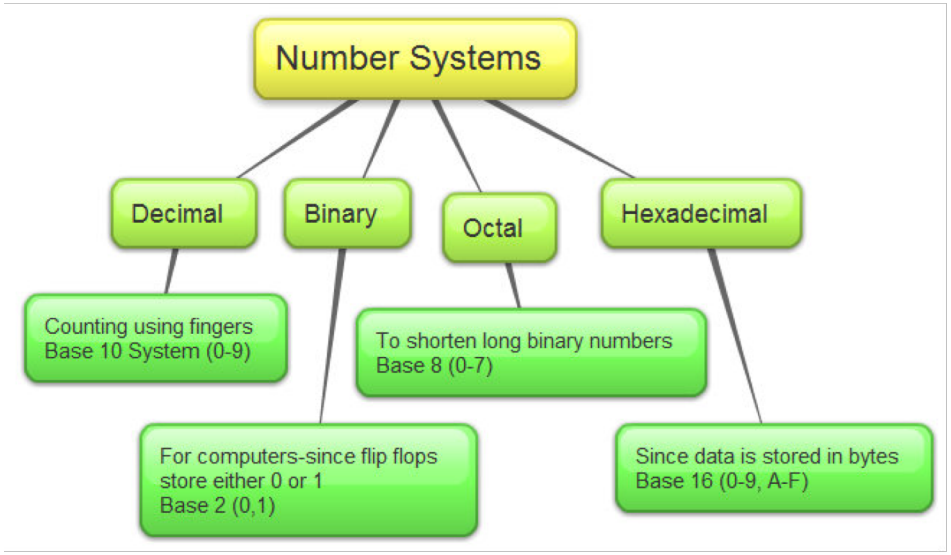


Lecture 6: Cryptograph and Basics of Counting

Dr. Chengjiang Long
Computer Vision Researcher at Kitware Inc.
Adjunct Professor at SUNY at Albany.
Email: **clong2@albany.edu**

Recap Previous Lecture

- Integer Representation, Binary Modular Exponentiation.
- Prime Factorization, Great Common Divisor, Least Common Multiple, Euclidean Algorithm



procedure $gcd(a, b$: positive integers)
$x := a$
$y := b$
while $y \neq 0$
    $r := x$ **mod** y
    $x := y$
    $y := r$
return $x\{\gcd(a, b)$ is $x\}$

Recap Previous Lecture

- Chinese Reminder Theory (one equation, two equations, three equations)

$$ax \equiv b \pmod{n}$$

Theorem. Let a, b, n be given integers. The above equation has a solution if and only if gcd(a,n) | b.

Furthermore, if the condition gcd(a,n) | b is satisfied,

then the solutions are of the form y mod (n/gcd(a,n)) for some integer y.

Theorem: If $n_1, n_2, \ldots, n_k$ are relatively prime and $a_1, a_2, \ldots, a_k$ are integers, then

$$x \equiv a_1 \pmod{n_1}$$
$$x \equiv a_2 \pmod{n_2}$$
$$\ldots$$
$$x \equiv a_k \pmod{n_k}$$

have a simultaneous solution x that is **unique** modulo n, where $n = n_1 n_2 \ldots n_k$.

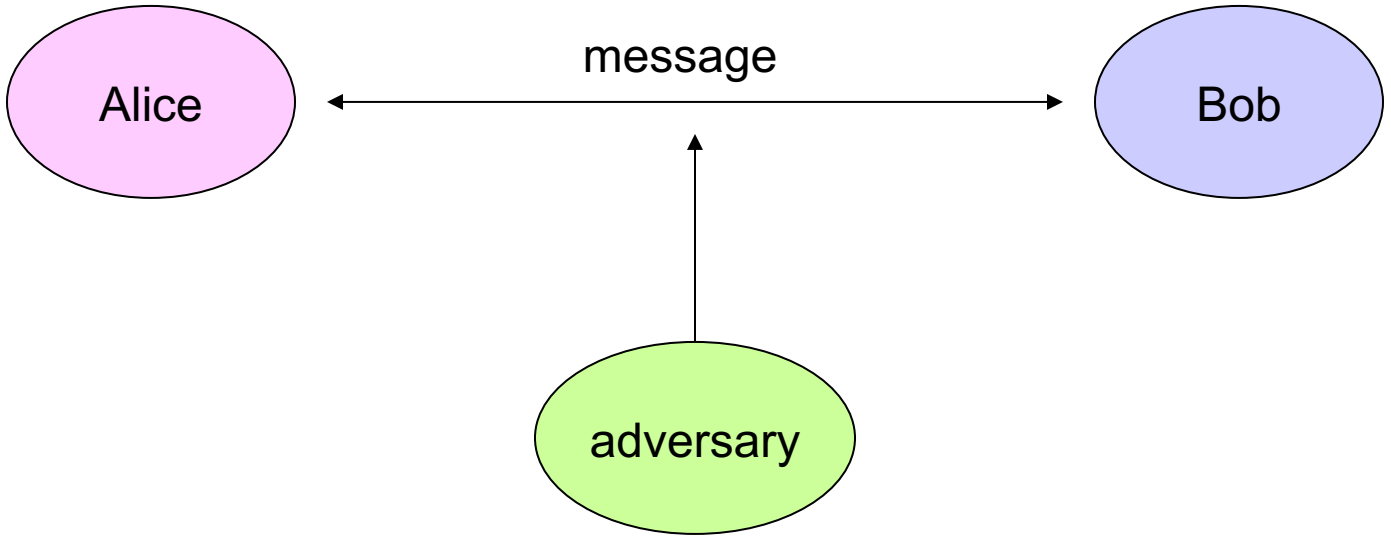
Outline

- Cryptograph
- Basic Counting Rules
- The Pigeonhole Principle
- Permutations and Combination

Cryptograph

Cryptograph

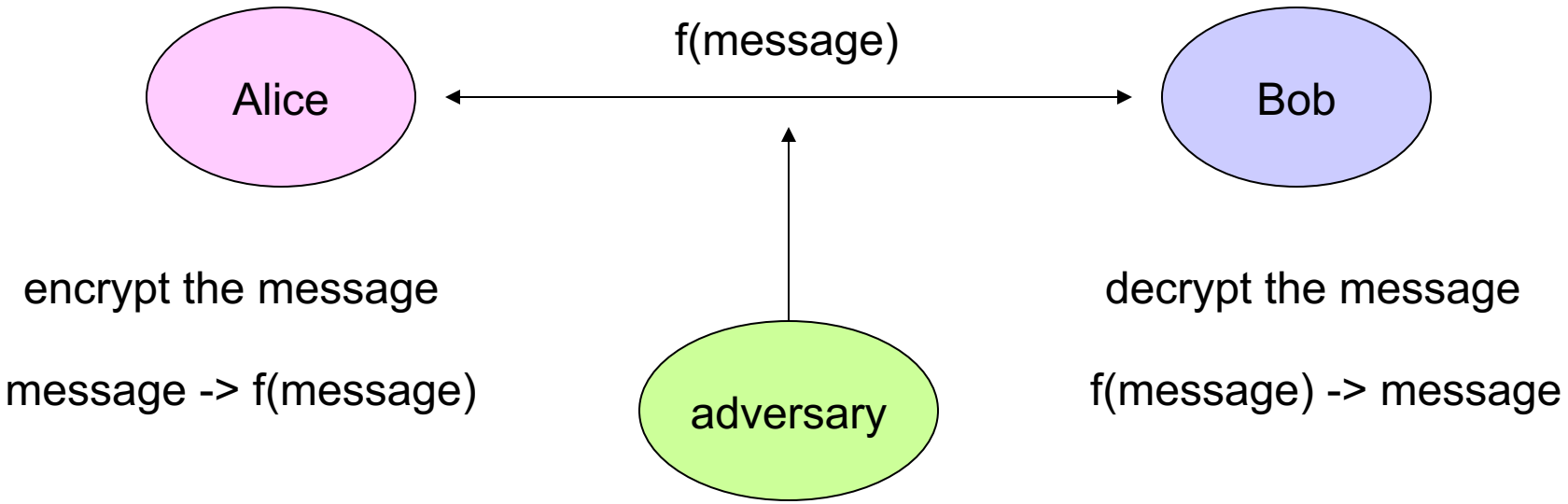
Cryptography is the study of methods for sending and receiving secret messages.



Goal: Even though an adversary can listen to your conversation, the adversary can not learn what the message was.

Cryptograph

Goal: Even though an adversary can listen to your conversation, the adversary can not learn what the message was.

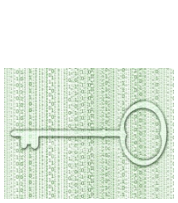


But the adversary has no clue how to obtain message from f(message)

A difficult goal!

Cryptograph: Key

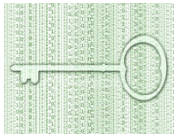
Goal: Even though an adversary can listen to your conversation, the adversary can not learn what the message was.



Alice

f(message, key)

Bob



encrypt the message using the key

message -> f(message,key)

adversary

decrypt the message using the key

f(message,key) -> message

But the adversary can not decrypt f(message,key) without the key

Use number theory!

Turing's Code (Version 1.0)

The first step is to translate a message into a number

"v i c t o r y"

-> 22 09 03 20 15 18 25

Beforehand The sender and receiver agree on a secret key, which is a large number k.

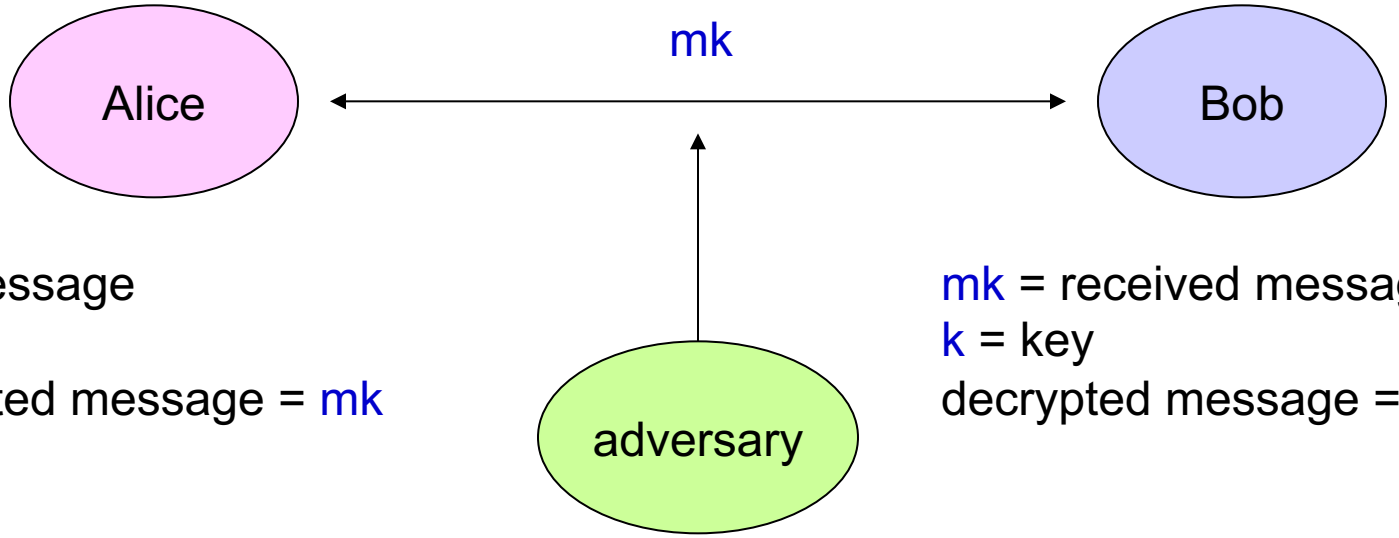
Encryption The sender encrypts the message m by computing:

$$m^* = m \cdot k$$

Decryption The receiver decrypts m by computing:

$$m^*/k = m \cdot k/k = m$$

Turing's Code (Version 1.0)



m = message
k = key
encrypted message = mk

mk = received message
k = key
decrypted message = mk/k=m

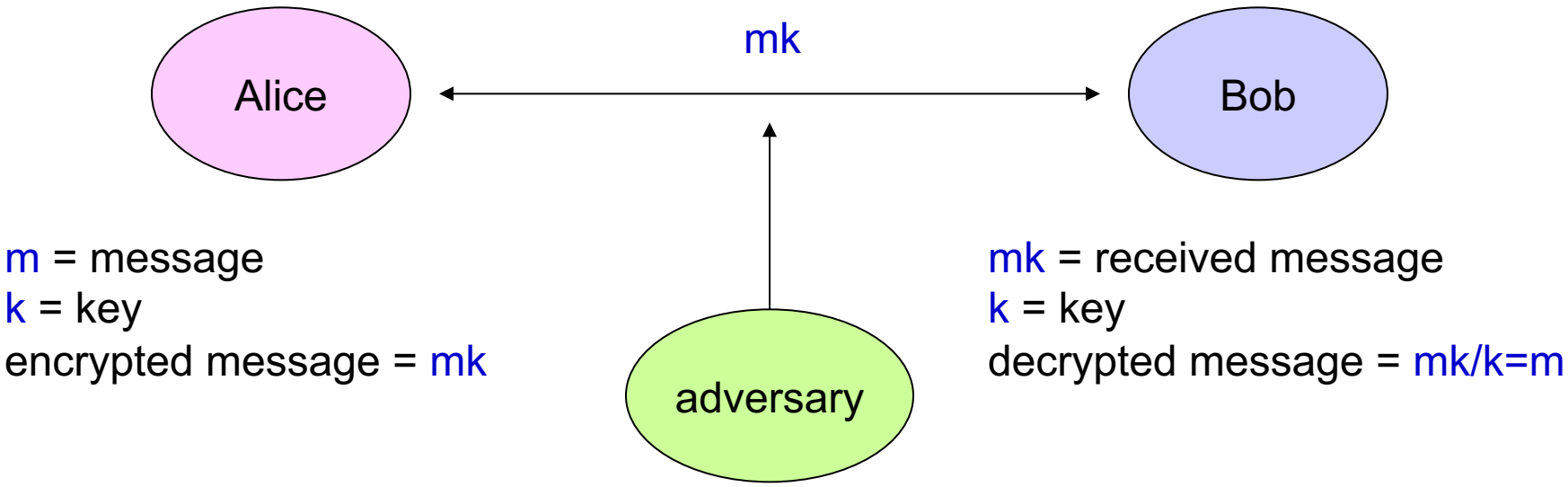
Why the adversary cannot figure out m?

The adversary doesn't have the key k,

and so can only factor mk to figure out m,

but factoring is a difficult task to do.

Turing's Code (Version 1.0)



So why don't we use this Turing's code today?

Major flaw: if you use the same key to send two messages m and m',

then from mk and m'k,

we can use gcd(mk,m'k) to figure out k,

and then decrypt every message.

Turing's Code (Version 2.0)

Beforehand The sender and receiver agree on a large prime p, which may be made public. (This will be the modulus for all our arithmetic.) They also agree on a secret key k in $\{1, 2, \ldots, p-1\}$.

Encryption The message m can be any integer in the set $\{0, 1, 2, \ldots, p-1\}$. The sender encrypts the message m to produce m^* by computing:

$$m^* = mk \bmod p$$

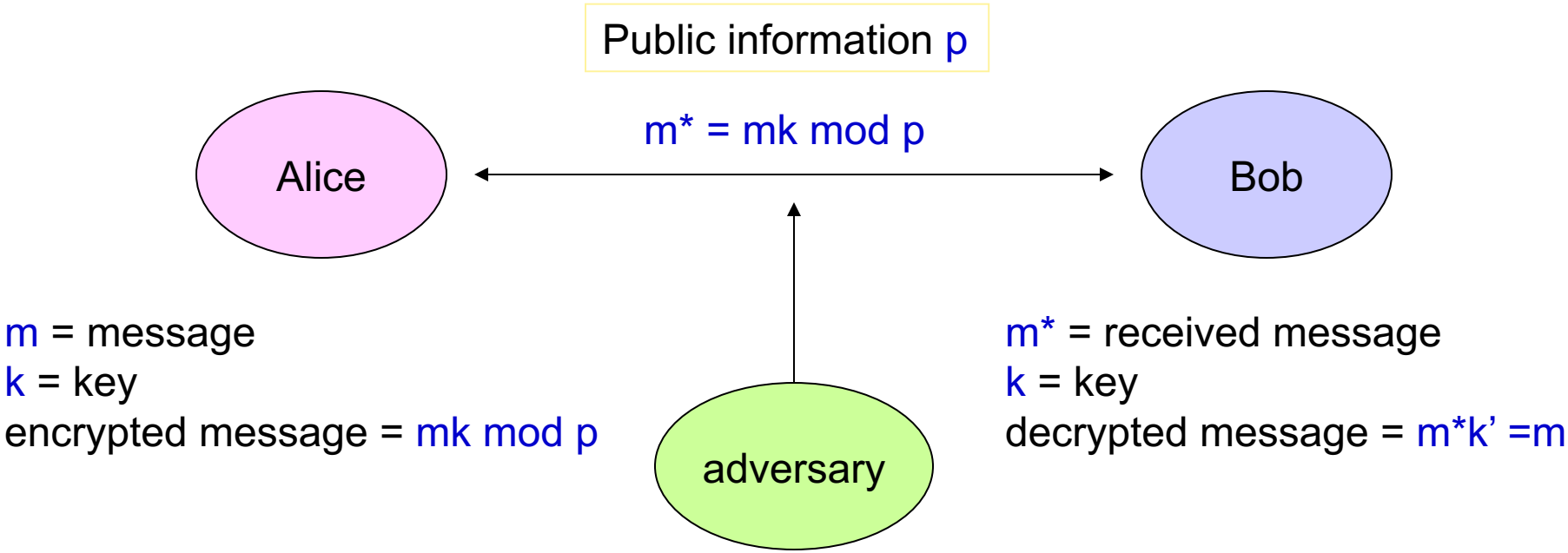
Decryption Let k' be the multiplicative inverse of k under modulo p.

$$m^* \equiv mk \pmod{p}$$

$$m^*k' \equiv m \pmod{p}$$

$$m^*k' \equiv m$$

Turing's Code (Version 2.0)

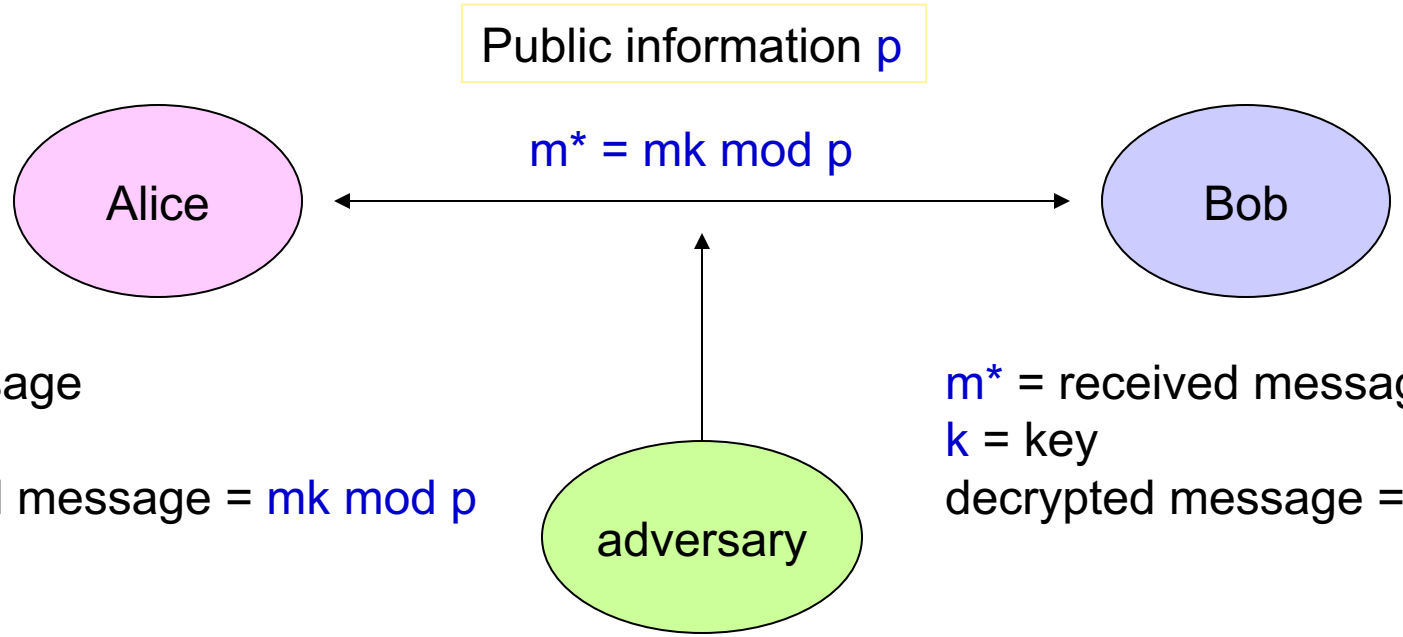


Why the adversary cannot figure out m?

Many m and k can produce m* as output,

just impossible to determine m without k.

Turing's Code (Version 2.0)



m = message
k = key
encrypted message = mk mod p

m* = received message
k = key
decrypted message = m*k' =m

So why don't we use this Turing's code today?

If the adversary somehow knows m,
then first compute m' := multiplicative inverse of m

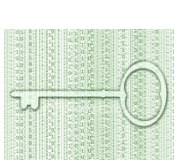
$m^* \equiv mk \pmod{p}$

$m^*m' \equiv k \pmod{p}$

So the adversary can figure out k.

plain-text attack

Private Key Cryptosystem



Alice

f(message, key)

Bob

encrypt the message using the key

message -> f(message,key)

adversary

decrypt the message using the key

f(message,key) -> message

But the adversary can not decrypt f(message,key) without the key

Two parties have to agree on a **secret key**, which may be difficult in practice.

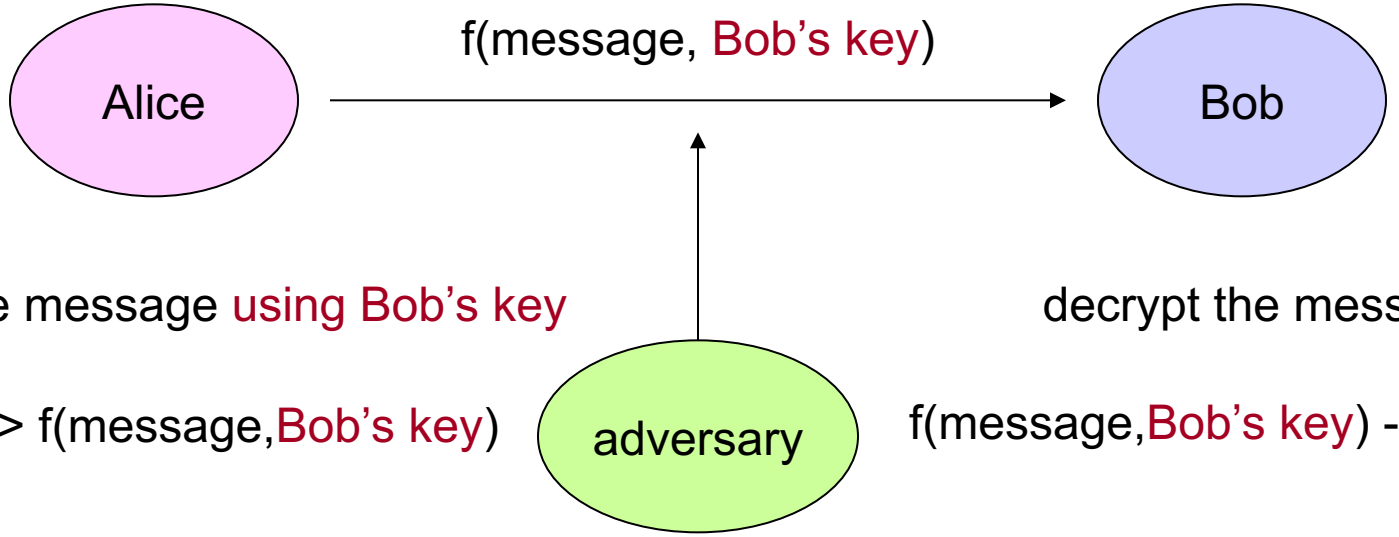
If we buy books from Amazon, we don't need to exchange a secret code.

Why is it secure?

Private Key Cryptosystem

Public information: Key for Alice

Public information: Key for Bob



encrypt the message using Bob's key

decrypt the message

message -> f(message,Bob's key)

f(message,Bob's key) -> message

But the adversary can not decrypt f(message, Bob's key)!

Only Bob can decrypt the message sent to him!

There is no need to have a secret key between Alice and Bob.

How is it possible???

RSA Cryptosystem

RSA are the initials of three Computer Scientists, Ron Rivest, Adi Shamir and Len Adleman, who discovered their algorithm when they were working together at MIT in 1977.

Generating Public Key

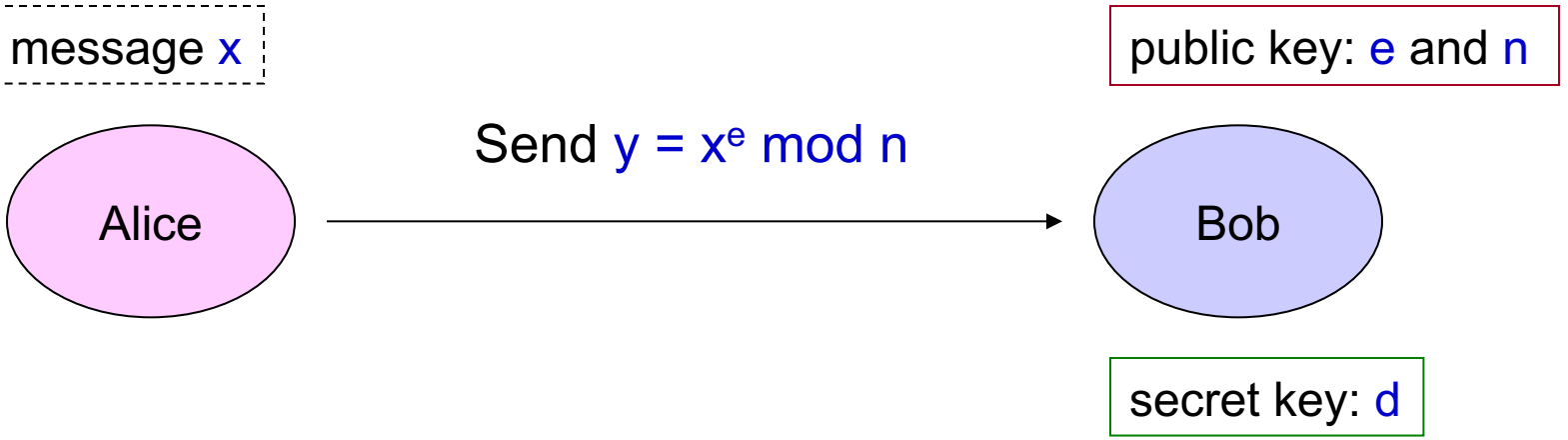
How Bob creates his public keys?

Secret key only known to Bob.

- Choose 2 large prime numbers p and q.
- Set n = pq and T = (p-1)(q-1)
- Choose e ≠1 so that gcd(e,T)=1
- Calculate d so that de ≡ 1 (mod T)
- Publish e and n as public keys
- Keep d as secret key

> 150 digits

Encrypting Message

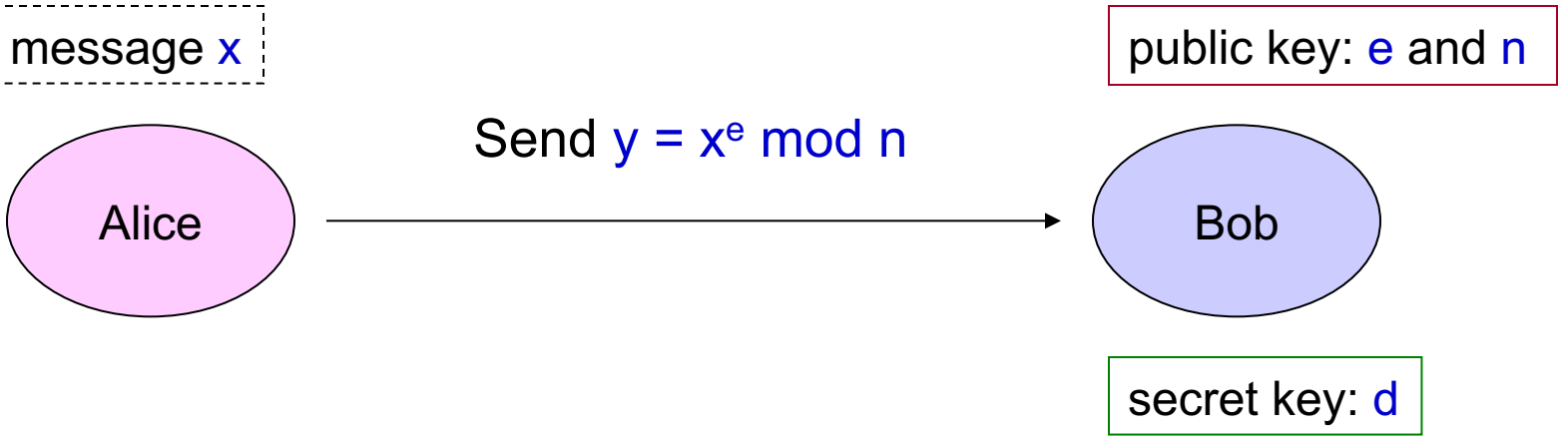


How Alice sends a message to Bob?

- Look at Bob's homepage for e and n.
- Send $y = x^e \bmod n$

Alice does not need to know Bob's secret key to send the message.

Decrypting Message

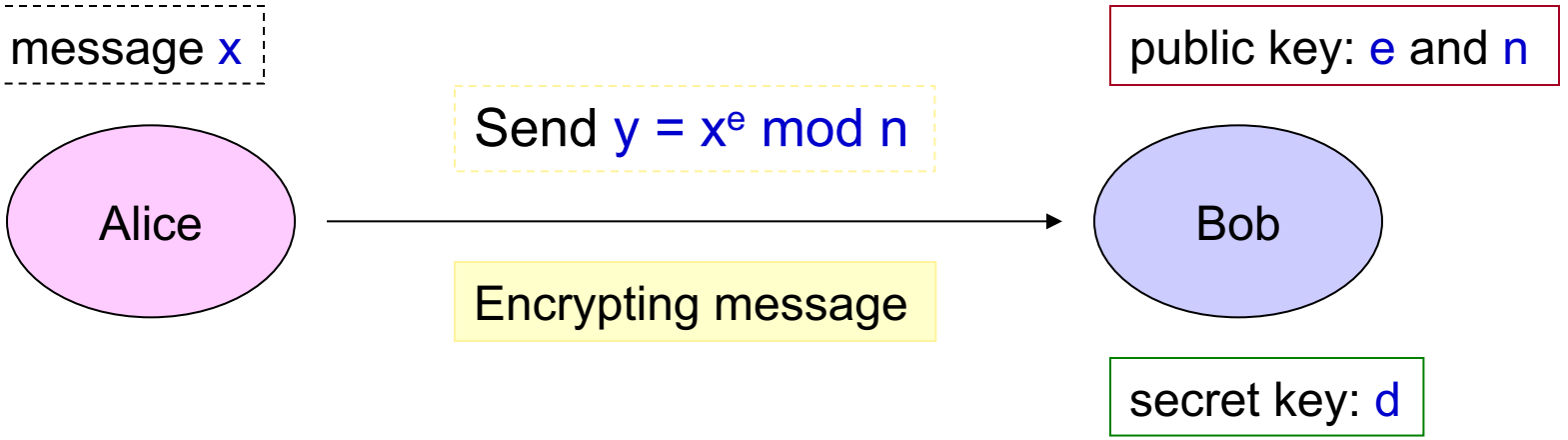


How Bob recover Alice's message?

- Receive $y = x^e \bmod n$
- Compute $z = y^d \bmod n$

Bob uses z is the original message that Alice sent.

RSA Cryptosystem



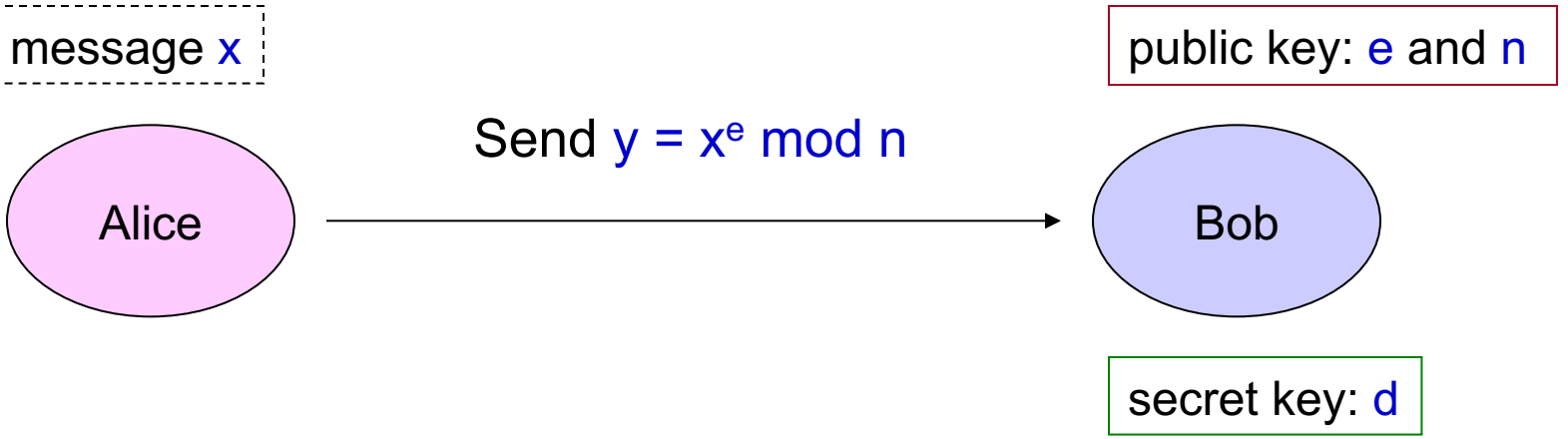
Key generation

- Choose 2 large prime numbers p and q.
- Set n = pq and T = (p-1)(q-1)
- Choose e ≠1 so that gcd(e,T)=1
- Calculate d so that $de \equiv 1 \pmod{T}$
- Publish e and n as public keys
- Keep d as secret key

Decrypting message

Compute $z = y^d \bmod n$

RSA Cryptosystem



For the RSA cryptosytem to work,
we need to show:

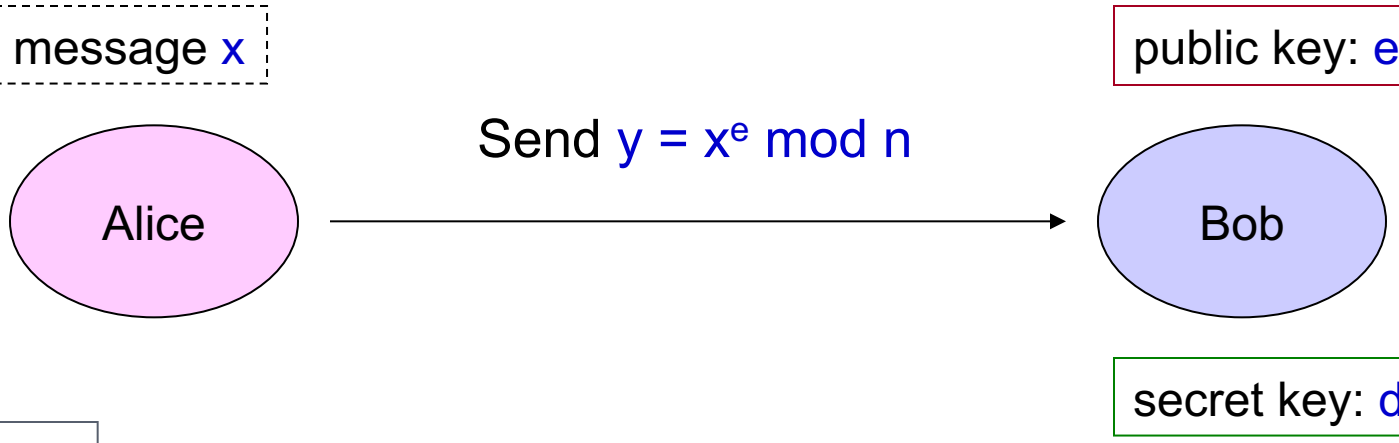
1) $z = x$

2) Without the secret key d,
we can not compute the original message
before the sun burns out.

Compute $z = y^d \bmod n$

with additional assumptions…

Correctness



1) z = x

Compute $z = y^d \bmod n$

Note that $z = y^d \bmod n = x^{ed} \bmod n$.

Therefore we need to prove $x = x^{ed} \bmod n$.

(a) $x \bmod p = x^{ed} \bmod p$

(b) $x \bmod q = x^{ed} \bmod q$

(c) $x \bmod n = x^{ed} \bmod n$

p, q prime

n = pq

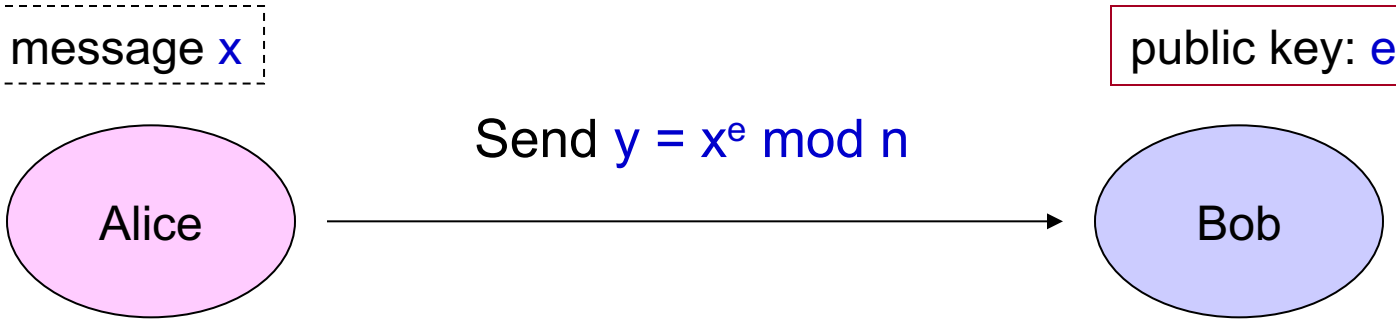
T = (p-1)(q-1)

e s.t. gcd(e,T)=1

$de \equiv 1 \pmod T$

Therefore, if Alice sends x < n, then Bob can recover correctly.

Correctness



1) z = x

(a) x mod p = x^{ed} mod p

Compute z = y^d mod n

Note that de = 1 + kT = 1 + k(p-1)(q-1)

Hence, x^{ed} mod p = $x^{1+k(p-1)(q-1)}$ mod p

= $x \cdot x^{k(p-1)(q-1)}$ mod p

= $x \cdot (x^{k(q-1)})^{(p-1)}$ mod p

p, q prime

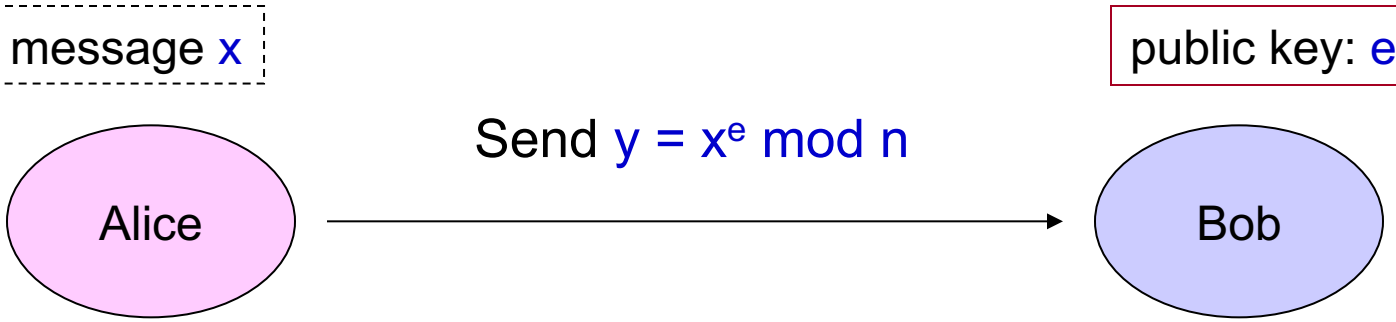
n = pq

T = (p-1)(q-1)

e s.t. gcd(e,T)=1

de ≡ 1 (mod T)

Correctness



1) $z = x$

(a) $x \bmod p = x^{ed} \bmod p$

Compute $z = y^d \bmod n$

Fermat's little theorem: If $p \nmid a$, then $a^{p-1} \equiv 1 \bmod p$

Hence, $x^{ed} \bmod p = x^{1+k(p-1)(q-1)} \bmod p$

$= x \cdot x^{k(p-1)(q-1)} \bmod p$

$= x \cdot (x^{k(q-1)})^{(p-1)} \bmod p$

a

$= x \bmod p$

p, q prime

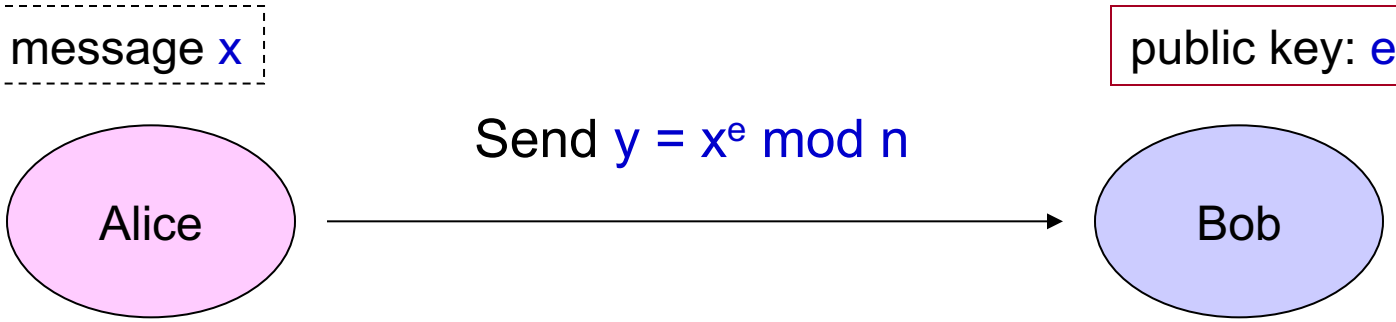
$n = pq$

$T = (p-1)(q-1)$

e s.t. $\gcd(e,T)=1$

$de \equiv 1 \pmod T$

Correctness



1) z = x

(a) x mod p = x^{ed} mod p

Compute z = y^d mod n

This means p | $x^{k(q-1)}$, implying p | x, since p is prime

Hence, x^{ed} mod p = $x^{1+k(p-1)(q-1)}$ mod p

= $x \cdot x^{k(p-1)(q-1)}$ mod p

= $x \cdot (x^{k(q-1)})^{(p-1)}$ mod p

a

What if p | a?

Since p | x, we have x^{ed} mod p = x mod p = 0.

p, q prime

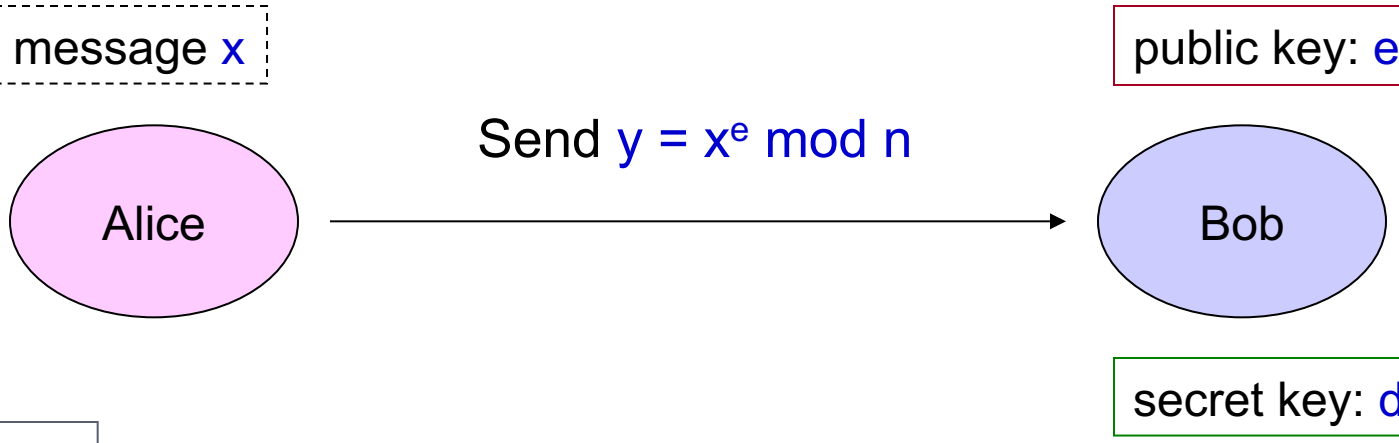
n = pq

T = (p-1)(q-1)

e s.t. gcd(e,T)=1

de ≡ 1 (mod T)

Correctness



1) z = x

Compute $z = y^d \bmod n$

Note that $z = y^d \bmod n = x^{ed} \bmod n$.

Therefore we need to prove $x = x^{ed} \bmod n$.

(a) $x \bmod p = x^{ed} \bmod p$

(b) $x \bmod q = x^{ed} \bmod q$ ← The same proof.

(c) $x \bmod n = x^{ed} \bmod n$

p, q prime

n = pq

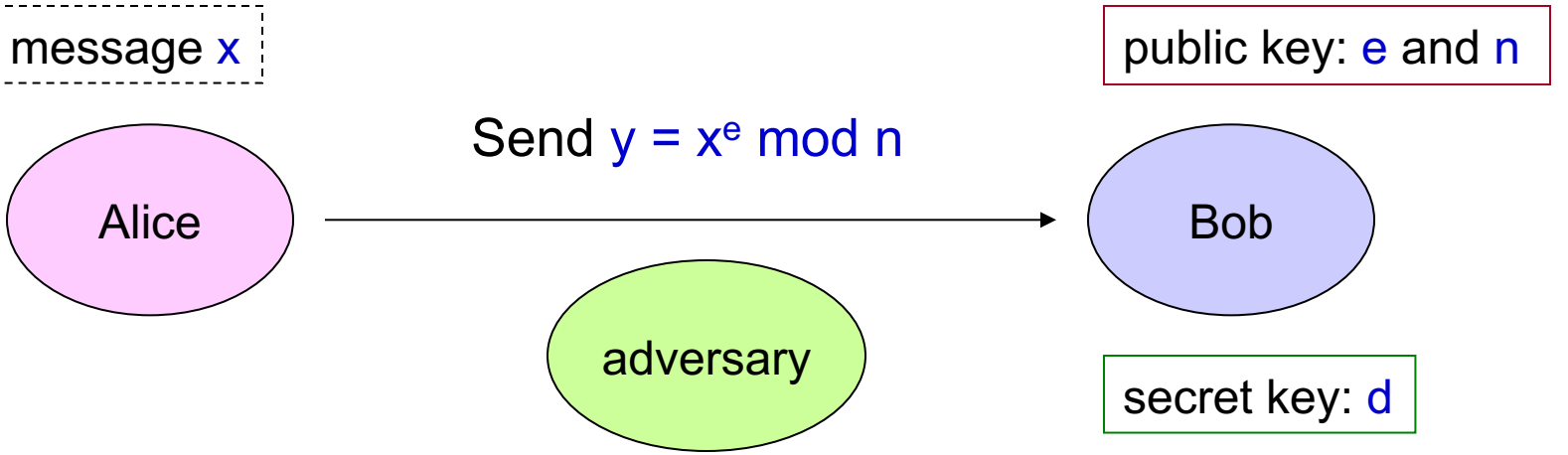
T = (p-1)(q-1)

e s.t. gcd(e,T)=1

$de \equiv 1 \pmod T$

(c) can be proved directly, also follows from Chinese Remainder theorem.

Why is This Secure?



2) Without the secret key d,
we can not compute the original message
before the sun burns out.

Compute $z = y^d \bmod n$

Method 1:

From $y=x^e \bmod n$, don't know how to compute x.

Thus not possible to work backward.

It is an example of an "one-way" function.

p, q prime

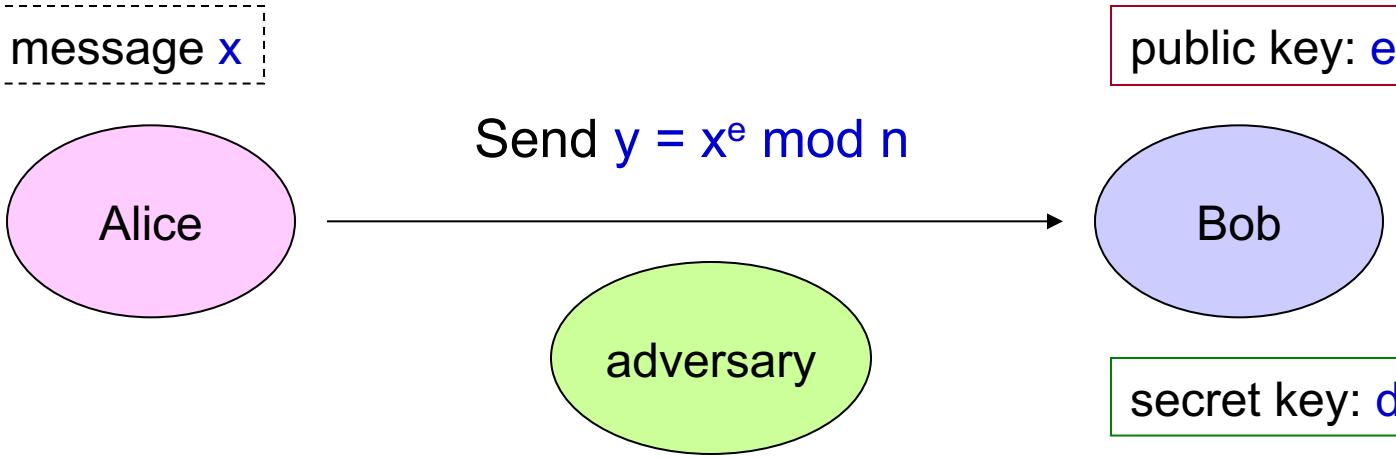
n = pq

T = (p-1)(q-1)

e s.t. gcd(e,T)=1

$de \equiv 1 \pmod T$

Why is This Secure?



2) Without the secret key d,
we can not compute the original message
before the sun burns out.

Compute $z = y^d \bmod n$

p, q prime
n = pq
T = (p-1)(q-1)
e s.t. gcd(e,T)=1
$de \equiv 1 \pmod{T}$

Method 2:

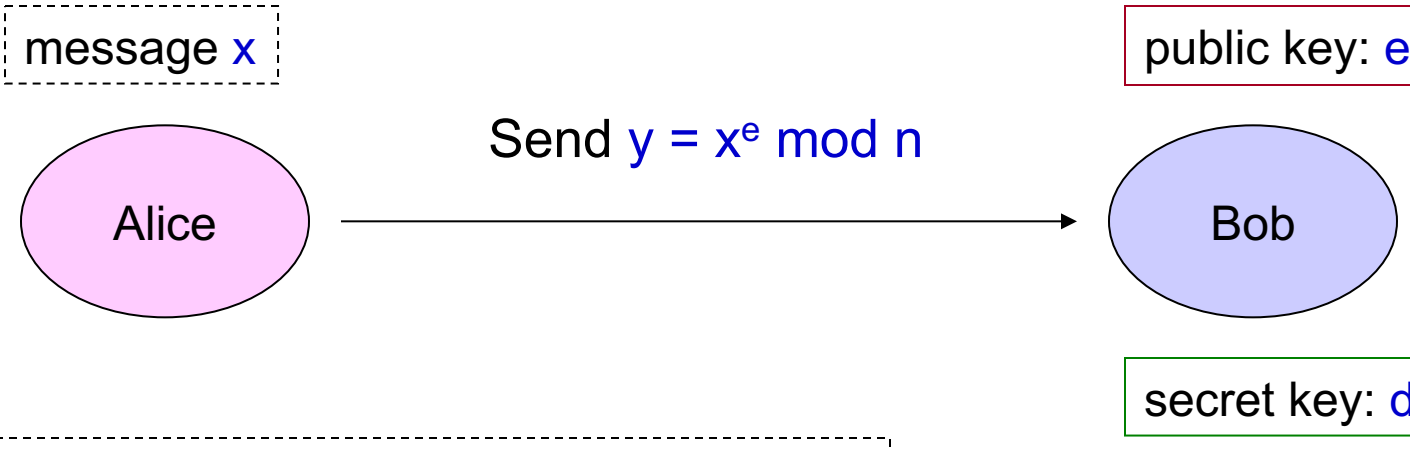
Factor n = pq. Compute secrete key d.

Then decrypt everything!

No one knows an efficient way to do factoring.

The security is based on assumptions that some computational problems are hard.

RSA Example



Then Alice sends the encrypted message.

Compute $z = y^d \bmod n$

$x=33$ $y = 33^{23} \bmod 55$

$y = 842986495178819225397387346633991 37 \bmod 55$

How to compute it efficiently?

p=5 q=11
n = 55
T = 40
e = 7
d = 23

p, q prime
$n = pq$
$T = (p-1)(q-1)$
e s.t. gcd(e,T)=1
$de \equiv 1 \pmod T$

First Bob generated his keys.

Exponentiation

To compute exponentiation mod n

$144^4 \bmod 713$

$= 144 * 144 * 144 * 144 \bmod 713$

$= 20736 * 144 * 144 \bmod 713$

$= 59 * 144 * 144 \bmod 713$

$= 8496 * 144 \bmod 713$

$= 653 * 144 \bmod 713$

$= 94032 \bmod 713$

$= 629 \bmod 713$

This still takes too long when the exponent is large.

$20736 * 20736 \bmod 713$

$= 59 * 59 \bmod 713$

$= 3481 \bmod 713$

$= 629 \bmod 713$

This is much more efficient.

Repeated Squaring

Note that $50 = 32 + 16 + 2$

$144^{50} \bmod 713$

$= 144^{32}\ 144^{16}\ 144^{2} \bmod 713$

$= 648 \cdot 485 \cdot 59 \bmod 713$

$= 242$

$144^{2} \bmod 713 = 59$

$144^{4} \bmod 713$
$= 144^{2} \cdot 144^{2} \bmod 713$
$= 59 \cdot 59 \bmod 713$
$= 629$

$144^{8} \bmod 713$
$= 144^{4} \cdot 144^{4} \bmod 713$
$= 629 \cdot 629 \bmod 713$
$= 639$

$144^{16} \bmod 713$
$= 144^{8} \cdot 144^{8} \bmod 713$
$= 639 \cdot 639 \bmod 713$
$= 485$

$144^{32} \bmod 713$
$= 144^{16} \cdot 144^{16} \bmod 713$
$= 485 \cdot 485 \bmod 713$
$= 648$

Remarks

- We have derived everything from basic principle.

- RSA cryptosystem is one of the most important achievements in compute science. (The researchers won the Turing award for their contribution.)

- Number theory is also very useful in coding theory (e.g. compression).

- Mathematics is very important in computer science.

Remarks

Theorem: if n is composite, for more than half of $a < n$, the strong primality test will say n is composite!

The proof uses Chinese Remainder theorem and some elementary number theory. (Introduction to Algorithms, MIT press)

Conjecture: It is enough to try a to up to roughly log(n).

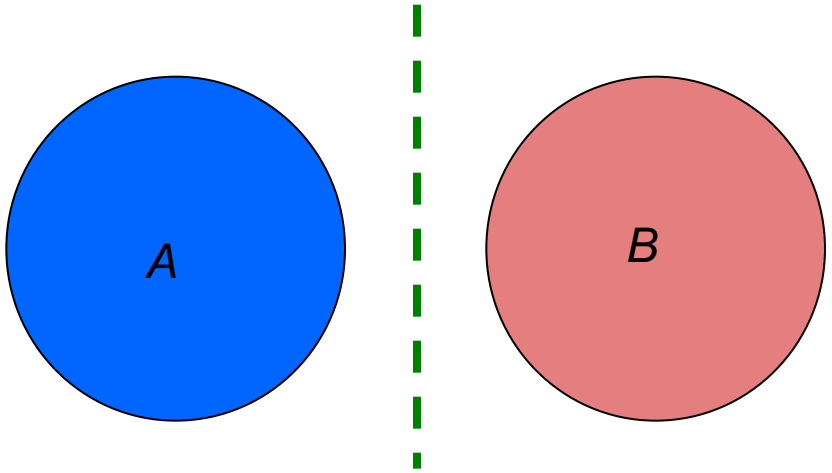
Theroem (Primes is in P, 2004)
There is an efficient and deterministic primality test.

Major Open Problem:
Is there an efficient algorithm to compute the prime factorization?

Basic Counting Rules

Sum Rule

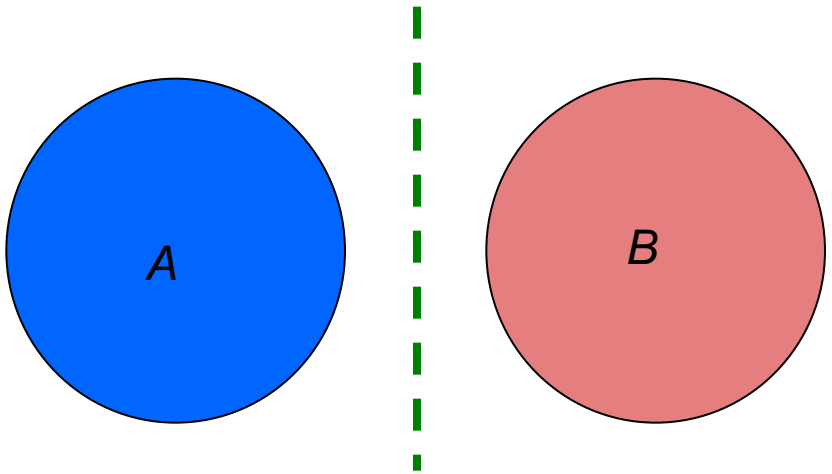
|S|: the number of elements in a set S.



If sets A and B are disjoint, then

$$|A \cup B| = |A| + |B|$$

Sum Rule



If sets A and B are disjoint, then
$|A \cup B| = |A| + |B|$

- Class has 43 women, 54 men, so total enrollment = 43 + 54 = 97
- 26 lower case letters, 26 upper case letters, and 10 digits,
  so total characters = 26+26+10 = 62

Product Rule

Recall that, given two sets A and B, the Cartisean product

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}$$

Fact: If |A| = n and |B| = m, then |AxB| = mn.

A = {a, b, c, d}, B = {1, 2, 3}

$A \times B$ = {(a,1),(a,2),(a,3),
(b,1),(b,2),(b,3),
(c,1),(c,2),(c,3),
(d,1),(d,2),(d,3) }

Example: If there are 4 men and 3 women, there are

$4 \times 3 = 12$ possible married couples.

Product Rule

Fact: If |A| = n and |B| = m, then |AxB| = mn.

In general let $A = \{a_1, a_2, a_3, \ldots, a_m\}$ and $B = \{b_1, b_2, \ldots, b_n\}$.

We can arrange the elements into a table as follows.

$A \times B = \{(a_1,b_1), (a_1,b_2),\ldots, (a_1,b_n),$

$(a_2,b_1), (a_2,b_2),\ldots, (a_2,b_n),$

$(a_3,b_1), (a_3,b_2),\ldots, (a_3,b_n),$

$\ldots$

$(a_m,b_1), (a_m,b_2),\ldots, (a_m,b_n), \}$

There are m rows, and each row has n elements,

and so there are a total of mn elements.

Product Rule

Fact: $|A_1 x A_2 x \ldots x A_k| = |A_1| x |A_2| x \ldots x |A_k|$.

The formal proof uses mathematical induction.

But the proof idea is not difficult.

We think of $A_1 x A_2 x \ldots x A_k$ as $(\ldots((A_1 x A_2) x A_3) \ldots x A_k)$.

That is, we first construct $A_1 x A_2$, and it is a set of size $|A_1| x |A_2|$.

Then, we construct $(A_1 x A_2) x A_3$, the product of $A_1 x A_2$ and A_3,

and it is a set of size $(|A_1| x |A_2|) x |A_3|$ by the product rule on two sets.

Repeating the argument we can see that $|A_1 x A_2 x \ldots x A_k| = |A_1| x |A_2| x \ldots x |A_k|$.

Example: Counting Strings

What is the number of 10-bit strings?

Let B={0,1}.

The set of 2-bit strings is just BxB.

The set of 10-bit strings is just BxBxBxBxBxBxBxBxBxB, denoted by B^{10}.

By the product rule, |BxB| = |B|x|B| = 2x2 = 4, and

$|B^{10}| = |B|x|B|x|B|x|B|x|B|x|B|x|B|x|B|x|B|x|B| = |B|^{10} = 2^{10} = 1024$.

Example: IP Addresses

What is the number of IP addresses?

An IP address is of the form 192.168.0.123.

There are four numbers, each is between 0 and 255.

Let $B=\{0,1,\ldots,255\}$.

Then the set of IP addresses is just B^4.

By the product rule, $|B^4| = |B|^4 = 256^4 = 4294967296$.

Example: Product Rule

In general we have:

The number of length-n *strings* from an *alphabet* of size m is m^n.

That is, $|B^n| = |B|^n$.

e.g. the number of length-n binary strings is 2^n

the number of length-n strings formed by capital letters is 26^n

Defective Dollars

A dollar is defective if some digit appears more than once in the 6-digit serial number.

How common are nondefective dollars?

Example: Counting Passwords

How many passwords satisfy the following requirements?

- between 6 & 8 characters long
- starts with a letter
- case sensitive
- other characters: digits or letters

First we define the set of letters and the set of digits.

$L = \{a,b,\ldots,z,A,B,\ldots,Z\}$

$D = \{0,1,\ldots,9\}$

Example: Counting Passwords

L ::= {a,b,…,z,A,B,…,Z}

D ::= {0,1,…,9}

We first count the number of passwords with a specific length.

Let P_n be the set of passwords with length n.

$$P_6 = L \times (L \cup D) \times (L \cup D) \times (L \cup D) \times (L \cup D) \times (L \cup D)$$

$$= L \times (L \cup D)^5$$

$$P_n ::= \text{length } n \text{ passwords}$$

$$= L \times (L \cup D)^{n-1}$$

Example: Counting Passwords

$$\left|L \times \left(L \cup D\right)^{n-1}\right| = |L| \cdot |L \cup D|^{n-1} \qquad \text{by product rule}$$

$$= |L| \cdot \left(|L| + |D|\right)^{n-1} \qquad \text{by sum rule}$$

$$= 52 \cdot 62^{n-1}$$

The set of Passwords:

$$P = P_6 \cup P_7 \cup P_8$$

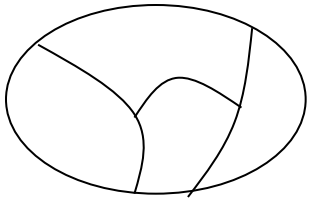
$$|P| = |P_6| + |P_7| + |P_8|$$

$$= 52 \cdot 62^5 + 52 \cdot 62^6 + 52 \cdot 62^7$$

$$= 186125210680448$$

$$\approx 19 \cdot 10^{14}$$

counting by partitioning



This is a common technique.

Divide the set into disjoint subsets.

Count each subset and add the answers.

At Least One Seven

How many # 4-digit numbers with at least one 7?

Method 1: count by *1st occurrence* of 7:

7xxx + o7xx + oo7x + ooo7

where x represents any digit from 1 to 10,

while o represent any digit from 1 to 10 except 7.

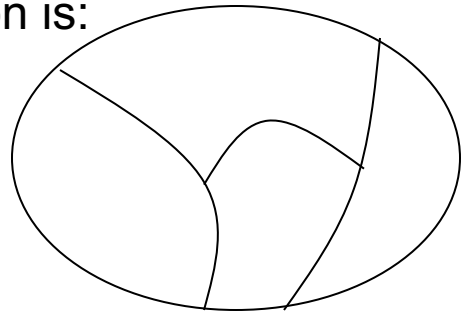
Clearly, each number containing at least one 7 is in one of the above sets, and these sets are disjoint.

Therefore, the answer to the question is:

$10^3 + 9 \cdot 10^2 + 9^2 \cdot 10 + 9^3 = 3439$

The set of 4-digit numbers with 7 in the first digit.

The set of 4-digit numbers with 7 in the second digit, but the first digit is not 7, and so on.



(counting by partitioning)

At Least One Seven

How many # 4-digit numbers with at least one 7?

Method 2:

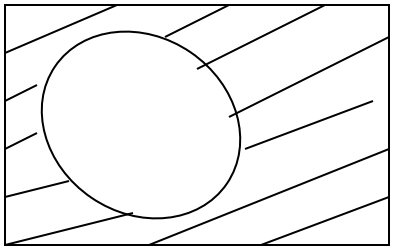
|4-digit numbers with at least one 7|=

|4-digit numbers| − |those with no 7s|

$= 10^4 - 9^4$

= 3439

(counting the complement)



Counting the complement is a useful technique.

Defective Dollars

How common are nondefective dollars?

10 possible choices for the first digit,

9 possible choices for the second digit, and so on…

So, there are 10x9x8x7x6x5

= 151200 serial number with all its digit different

There are totally 10^6 = 1000000 serial numbers.

So, only about 15% of dollars are nondefective.

Generalized Product Rule

Q a set of length-*k* sequences. If there are:

n_1 possible 1st elements in sequences,

n_2 possible 2nd elements for each first entry,

n_3 possible 3rd elements for each 1st & 2nd,

…

then, $|Q| = n_1 \cdot n_2 \cdot n_3 \cdot \ldots \cdot n_k$

The Pigeonhole Principle

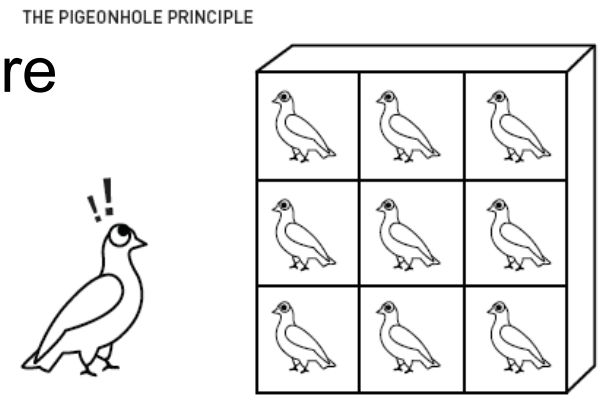
Pigeonhole Principle

Motivation:
The mapping of n objects to m buckets
E.g. Hashing.

The principle is used for proofs of certain complexity derivation.

Pigeonhole Principle

Pigeonhole Principle: If n pigeonholes are occupied by n + 1 or more pigeons, then at least one pigeonhole is occupied by more than one pigeon.



Remark: The principle is obvious. No simpler fact or rule to support or prove it.

Generalized Pigeonhole Principle: If n pigeonholes are occupied by kn + 1 pigeons, then at least one pigeonhole is occupied by k + 1 or more pigeons.

Example 1: Birthmonth

- In a group of 13 people, we have 2 or more who are born in the same month.

# pigeons	# holes	At least # born in the same month
13	12	2 or more
20	12	2 or more
121	12	11 or more
65	12	6 or more
111	12	10 or more
≥kn+1	n	k+1 or more

Example 2: Handshaking

Given a group of n people (n>1), each shakes hands with some (a nonzero number of) people in the group. We can find at least two who shake hands with the same number of people.

Proof:
Number of pigeons (number of people): n
Number of pigeonholes (range of number of handshakes): n-1

Example 3: Cast in theater

A theater performs 7 plays in one season. There are 15 women. Then some play has at least 3 women in its cast.

Number of pigeons: 15
Number of pigeonholes: 7
k*n+1=2*7+1
3 or more pigeons in the same pigeonhole

Example 4: Pairwise difference

Given 8 different natural numbers, none greater than 14. Show that at least three pairs of them have the same difference.

Try a set: 1, 2, 3, 7, 9, 11, 12, 14
Difference of 12 and 14 = 2.
Same for 9 and 11; 7 and 9; 1 and 3.
In this set, there are four pairs that all have the same difference.

Example 4: Pairwise difference

Given 8 different natural numbers, none greater than 14. Show that at least three pairs of them have the same difference.

Try a set: 1, 2, 3, 7, 9, 11, 12, 14
Difference of 12 and 14 = 2.
Same for 9 and 11; 7 and 9; 1 and 3.
In this set, there are four pairs that all have the same difference.

Proof:
pigeons (different pairs: C(8,2) = 8*7/2): 28
pigeonholes (14-1): 13
Since 28≥k*n+1=2*13+1, we have
3 or more pigeons in the same pigeonhole.

Remark

- Pigeonhole principle has applications to assignment and counting.
- The usage of the principle relies on the identification of the pigeons and the pigeonholes.

Permutations and Combinations

Permutations

Definition: A **permutation** of a set S is a sequence that contains every element of S exactly once.

For example, here are all six permutations of the set {a, b, c}:

(a, b, c) (a, c, b) (b, a, c)

(b, c, a) (c, a, b) (c, b, a)

Ordering is important here.

How many permutations of an n-element set are there?

You can think of a permutation as a ranking of the elements.

So the above question is asking how many rankings of an n-element set.

Permutations

How many permutations of an n-element set are there?

- There are n choices for the first element.
- For each of these, there are n − 1 remaining choices for the second element.
- For every combination of the first two elements, there are n − 2 ways to choose the third element, and so forth.
- Thus, there are a total of

$$n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1 = n!$$

permutations of an n-element set.

← This is called n **factorial**.

Stirling's formula (optional):

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

Example

Suppose each digit is an element in {1,2,3,4,5,6,7,8,9}.
How many 9-digit numbers are there where each nonzero digit appears once?

Each such number corresponds to a permutation of 123456789,

and each permutation corresponds to such a number.

So the numbers of such numbers is equal to

the number of permutations of {1,2,3,4,5,6,7,8,9}.

Hence there are exactly 9! such numbers.

Alternatively, one can use the generalized product rule directly to obtain the same result.

Combinations

How many subsets of size k of an n-element set?

Consider the set {1,2,3,4,5} where n=5.

If k=2, then there are 10 possible subsets of size 2,

i.e. {1,2}, {1,3}, {1,4}, {1,5}, {2,3}, {2,4}, {2,5}, {3,4}, {3,5}, {4,5}.

If k=3, then there are also 10 possible subsets of size 3,

i.e. {1,2,3}, {1,2,4}, {1,2,5}, {1,3,4}, {1,3,5}

{1,4,5}, {2,3,4}, {2,3,5}, {2,4,5}, {3,4,5}

Ordering is NOT important here.

Combinations

How many subsets of size k of an n-element set?

- There are n choices for the first element.
- For each of these, there are n − 1 remaining choices for the second element.
- There are n – k + 1 remaining choices for the last element.
- Thus, there are a total of
  $n \cdot (n - 1) \cdot (n - 2) \cdot \cdots (n - k + 1)$ to choose k elements.

So far we counted the number of ways to choose k elements, when the ordering is important.

e.g. {1,2,3}, {1,3,2}, {2,1,3}, {2,3,1}, {3,1,2}, {3,2,1} will be counted as 6 different ways.

Combinations

How many subsets of size k of an n-element set?

We form the subsets by picking one element at a time.

- There are n choices for the first element.
- For each of these, there are n − 1 remaining choices for the second element.
- There are n – k + 1 remaining choices for the last element.
- Thus, there are a total of
  $n \cdot (n - 1) \cdot (n - 2) \cdot \cdots (n - k + 1)$ to choose k elements.

So far we counted the number of ways to choose k elements, when the ordering is important.

Combinations

How many subsets of size k of an n-element set?

- Thus, there are a total of

  $n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot (n-k+1)$ ways to choose k elements,

  when the ordering is important.

How many different ordering of k elements are (over)-counted?

e.g. If we are forming subsets of size 3, then

(1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1)

are counted as 6 different ways if the ordering is important.

In general, each subset of size k has k! different orderings,
and so each subset is counted k! times in the above way of choosing k elements.

Combinations

How many subsets of size k of an n-element set?

- Thus, there are a total of

  $n \cdot (n - 1) \cdot (n - 2) \cdot \cdots \cdot (n - k + 1)$ ways to choose k elements,

  when the ordering is important.

- Each subset is counted, but is counted k! times, because

  each subset contributes k! different orderings to the above.

- So, when the ordering is not important, the answer is:

This is the shorthand for "n choose k"

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} = \frac{n!}{(n-k)!k!}$$

Example: Team Formation

There are m boys and n girls.

How many ways are there to form a team with 3 boys and 3 girls?

There are $\binom{m}{3}$ choices of 3 boys and $\binom{n}{3}$ choices for 3 girls.

So by the product rule there are $\binom{m}{3}\binom{n}{3}$ choices of such a team.

If m < 3 or n < 3, then the answer should be zero.

Don't worry. We don't like to trick you this way.

Example: Bit Strings with k Zeros

How many n-bit sequences contain k zeros and (n − k) ones?

We can think of this problem as choosing k positions (out of the n possible positions) and set them to zeroes and set the remaining positions to ones.

So the above question is asking the number of possible positions of the k zeros, and the answer is:

$$\binom{n}{k}$$

Example: Unbalanced Bit Strings

We say a bit string is unbalanced if there are more ones than zeroes or more zeros than ones.

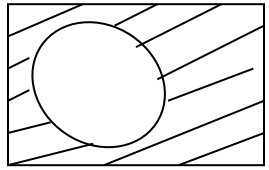
How many n-bit strings are unbalanced?

If n is odd, then every n-bit string is unbalanced, and the answer is 2^n.

If n is even, then the number of balanced strings is

$\binom{n}{n/2}$ by choosing n/2 positions to zeroes.

So the number of unbalanced n-bit strings is equal to the number of all n-bit strings minus the number of balanced strings, and so the answer is $2^n - \binom{n}{n/2}$ (counting the complement)



Poker Hands

There are 52 cards in a deck.

Each card has a suit and a value.

4 suits (♠ ♥ ♦ ♣)

13 values (2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A)

Five-Card Draw is a card game in which each player is initially dealt a hand, a subset of 5 cards.

How many different hands?

$$\binom{52}{5} = 2598960$$

Example 1: Four of a Kind

A Four-of-a-Kind is a set of four cards with the same value.

$$\{\ 8\spadesuit,\ 8\diamondsuit,\ Q\heartsuit,\ 8\heartsuit,\ 8\clubsuit\ \}$$
$$\{\ A\clubsuit,\ 2\clubsuit,\ 2\heartsuit,\ 2\diamondsuit,\ 2\spadesuit\ \}$$

How many different hands contain a Four-of-a-Kind?

One way to do this is to first map the problem into a problem of counting sequences.

Example 1: Four of a Kind

A hand with a Four-of-a-Kind is completely described by a sequence specifying:

1. The value of the four cards.
2. The value of the extra card.
3. The suit of the extra card.

$$(8, Q, \heartsuit) \leftrightarrow \{\ 8\spadesuit,\ 8\diamondsuit,\ 8\heartsuit,\ 8\clubsuit,\ Q\heartsuit\ \}$$
$$(2, A, \clubsuit) \leftrightarrow \{\ 2\clubsuit,\ 2\heartsuit,\ 2\diamondsuit,\ 2\spadesuit,\ A\clubsuit\ \}$$

There are 13 choices for (1), 12 choices for (2), and 4 choices for (3).
By generalized product rule, there are 13x12x4 = 624 hands.

Only 1 hand in about 4165 has a Four-of-a-Kind!

Example 2: Full House

A Full House is a hand with three cards of one value and two cards of another value.

$$\{\ 2\spadesuit,\ 2\clubsuit,\ 2\diamondsuit,\ J\clubsuit,\ J\diamondsuit\ \}$$
$$\{\ 5\diamondsuit,\ 5\clubsuit,\ 5\heartsuit,\ 7\heartsuit,\ 7\clubsuit\ \}$$

How many different hands contain a Full House?

Example 2: Full House

There is a bijection between Full Houses and sequences specifying:

1. The value of the triple, which can be chosen in 13 ways.
2. The suits of the triple, which can be selected in (4 3) ways.
3. The value of the pair, which can be chosen in 12 ways.
4. The suits of the pair, which can be selected in (4 2) ways.

$$(2, \{\spadesuit, \clubsuit, \diamondsuit\}, J, \{\clubsuit, \diamondsuit\}) \leftrightarrow \{\ 2\spadesuit,\ 2\clubsuit,\ 2\diamondsuit,\ J\clubsuit,\ J\diamondsuit\ \}$$
$$(5, \{\diamondsuit, \clubsuit, \heartsuit\}, 7, \{\heartsuit, \clubsuit\}) \leftrightarrow \{\ 5\diamondsuit,\ 5\clubsuit,\ 5\heartsuit,\ 7\heartsuit,\ 7\clubsuit\ \}$$

By generalized product rule, there are

$$13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2} = 3744$$

Only 1 hand in about 634 has a Full House!

Example 3: Two Pairs

How many hands have Two Pairs; that is,
two cards of one value, two cards of another value,
and one card of a third value?

$$\{\ 3\diamondsuit,\ 3\spadesuit,\ Q\diamondsuit,\ Q\heartsuit,\ A\clubsuit\ \}$$
$$\{\ 9\heartsuit,\ 9\diamondsuit,\ 5\heartsuit,\ 5\clubsuit,\ K\spadesuit\ \}$$

Example 3: Two Pairs

1. The value of the first pair, which can be chosen in 13 ways.
2. The suits of the first pair, which can be selected (4 2) ways.
3. The value of the second pair, which can be chosen in 12 ways.
4. The suits of the second pair, which can be selected in (4 2) ways
5. The value of the extra card, which can be chosen in 11 ways.
6. The suit of the extra card, which can be selected in 4 ways.

Number of Two pairs = $13 \cdot \binom{4}{2} \cdot 12 \cdot \binom{4}{2} \cdot 11 \cdot 4$

Double Count!

$(3, \{\diamondsuit, \spadesuit\}, Q, \{\diamondsuit, \heartsuit\}, A, \clubsuit)$ ↘

$(Q, \{\diamondsuit, \heartsuit\}, 3, \{\diamondsuit, \spadesuit\}, A, \clubsuit)$ ↗

$\{ \ 3\diamondsuit, \ 3\spadesuit, \ Q\diamondsuit, \ Q\heartsuit, \ A\clubsuit \ \}$

So the answer is $\frac{1}{2} \cdot 13 \cdot \binom{4}{2} \cdot 12 \cdot \binom{4}{2} \cdot 11 \cdot 4 = 123552$

Example 4: Every Suit

How many hands contain at least one card from every suit?

$$\{\ 7\diamondsuit,\ K\clubsuit,\ 3\diamondsuit,\ A\heartsuit,\ 2\spadesuit\ \}$$

1. The value of each suit, which can be selected in 13x13x13x13 ways.
2. The suit of the extra card, which can be selected in 4 ways.
3. The value of the extra card, which can be selected in 12 ways.

$$(7, K, A, 2, \diamondsuit, 3) \leftrightarrow \{\ 7\diamondsuit,\ K\clubsuit,\ A\heartsuit,\ 2\spadesuit,\ 3\diamondsuit\ \}$$

$$\begin{matrix}(7, K, A, 2, \diamondsuit, 3) & \searrow & \\ & & \{\ 7\diamondsuit,\ K\clubsuit,\ A\heartsuit,\ 2\spadesuit,\ 3\diamondsuit\ \} \\ (3, K, A, 2, \diamondsuit, 7) & \nearrow & \end{matrix}$$

Double count!

So the answer is 13^4x4x12/2 = 685464

Next class

- Topic: Probability and Applications
- Pre-class reading: Chap 6