
# More Ciphers

Based on slides from the book
Classical & Contemporary Cryptology
By Richard Spillman

Classical & Contempory Cryptology

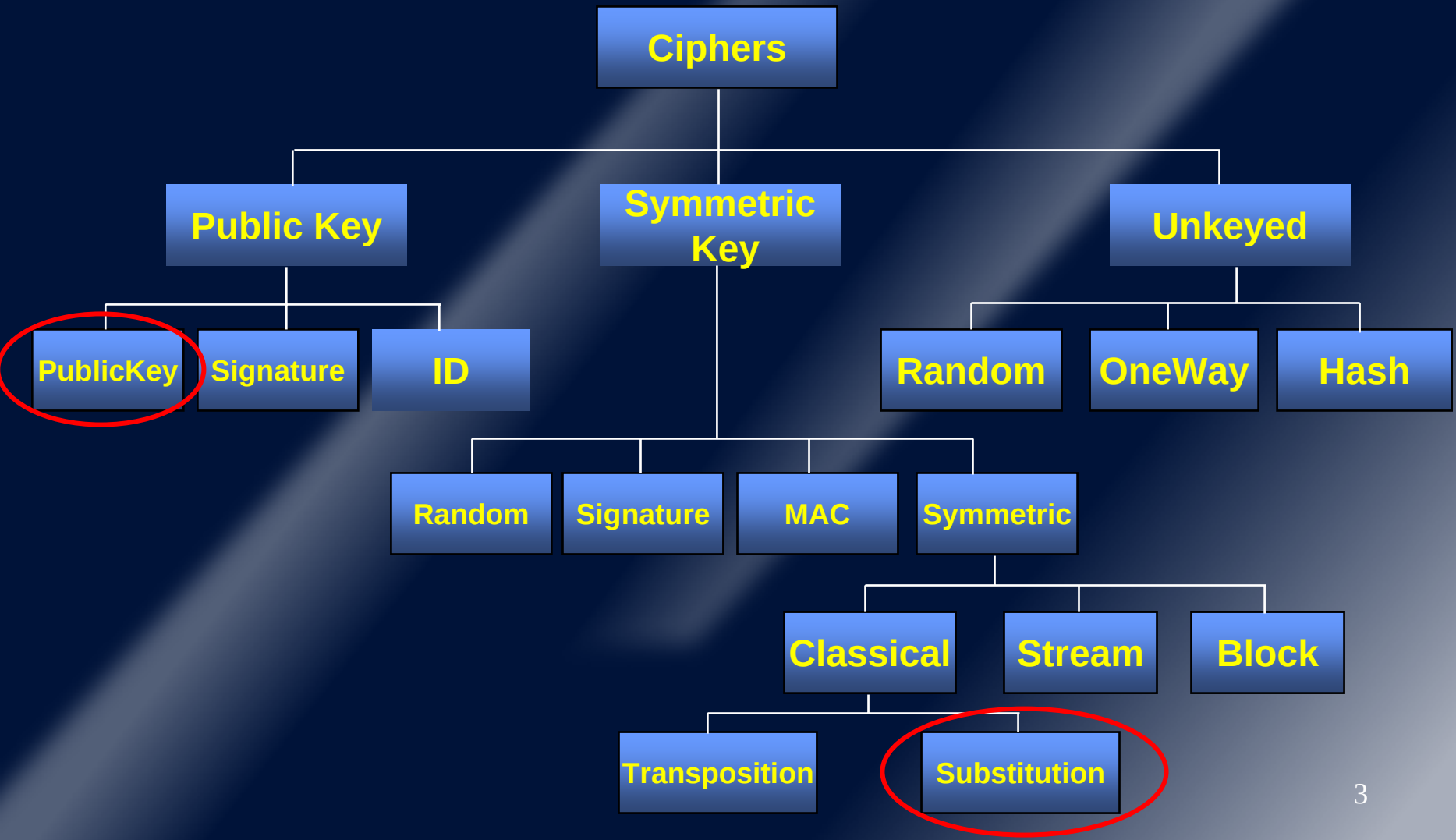
# A Good Cipher

- Enciphering and deciphering should be efficient for all keys - it should not take forever to get message.
- Easy to use. The problem with hard to use cryptosystems is that mistakes tend to be made
- The strength of the system should not lie in the secrecy of your algorithms. The strength of the system should only depend the secrecy of your key.

Fundamental Principle of Cryptology

Classical & Contempory Cryptology

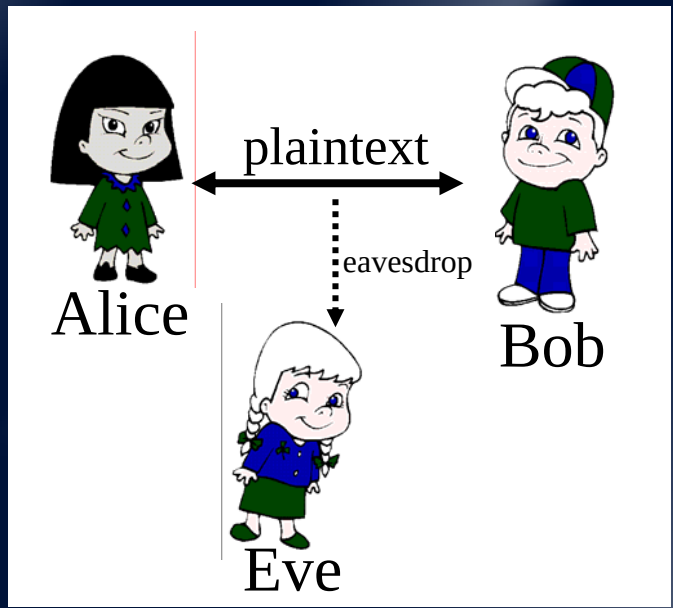
# Cipher Classification



Classical & Contempory Cryptology

# Cipher Environment

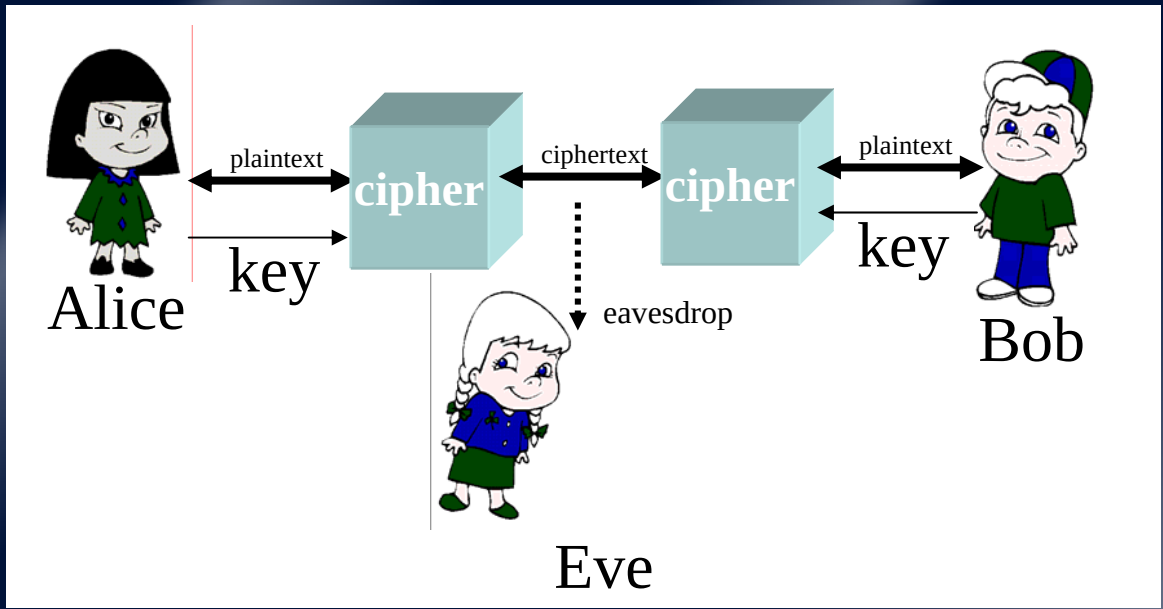
- The typical communication environment for discussing ciphers is



Classical & Contempory Cryptology

# Cipher System

- If Alice and Bob use a cipher system, this environment becomes:



Classical & Contemporary Cryptology

# Caesar Ciphers

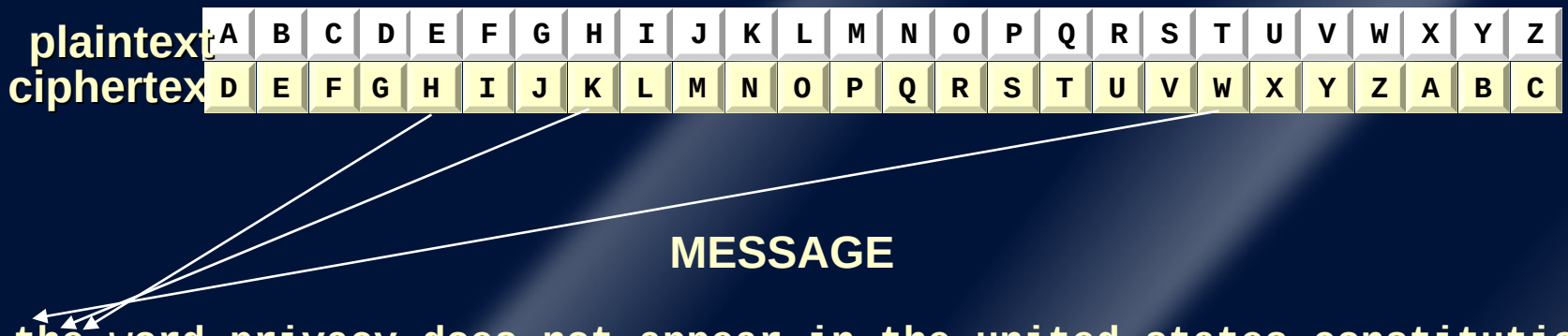
- A *substitution* cipher is one in which each character in the plaintext is substituted for another character in the ciphertext

- The Caesar Cipher replaces each plaintext character by the character k positions to the right. In this example, k=3.

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

Classical & Contemporary Cryptology

# Example Operation




**MESSAGE**

the word privacy does not appear in the united states constitution

wkh zrug sulydfb grhv qrw dsshdu lq wkh xqlwhg vwdwhv frqvwlwxwlrq

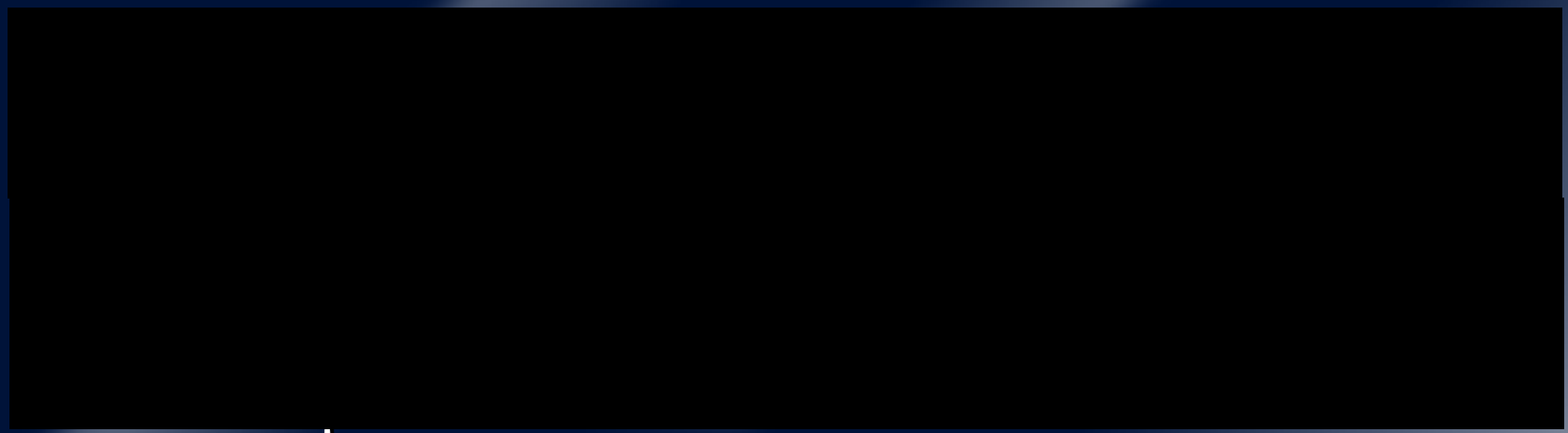
**NOTE: the shift could be any value from 1 to 25**

**NOTE: It helps to remove spaces and make blocks of letters- WHY??**

wkhzr ugsul ydfbg rhvqr wdssh dulqw khxql whgvw dwhvf rqvwl wxwlr q

# Cryptanalysis

- How would you break the Caesar cipher?

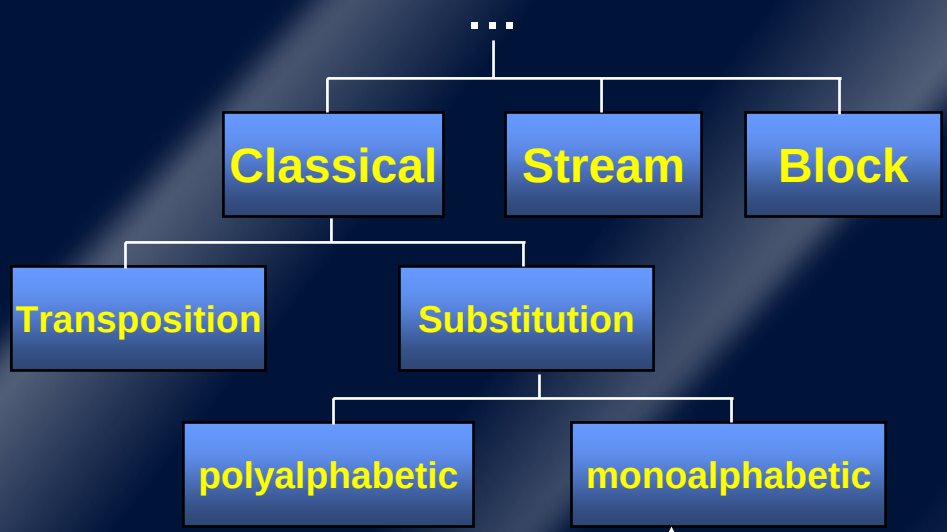


# *Monoalphabetic Ciphers*

Classical & Contemporary Cryptology

# New Cipher Types

- **Further subdivisions:**




Each plaintext character is mapped to several ciphertext characters

Each plaintext character is mapped to one and only one ciphertext character

Classical & Contemporary Cryptology

# Monoalphabetic Ciphers we've seen

- **Caesar (Additive) Cipher** (only 26 keys)

  $c = p + k \pmod{26}$

  p is plaintext, c is ciphertext, k is key

- **Multiplicative Ciphers** (only 12 keys)

  $c = p * k \pmod{26}$ (gcd(k,26) = 1)

- **Affine Ciphers** (only 26*12 = 312 keys)

  $c = a*p + b \pmod{26}$ (gcd(a,26) = 1)

Classical & Contempory Cryptology

# Breaking Ciphers

- The Caesar cipher is easy to break because there are only 26 possible keys, so we need a stronger cipher. The multiplicative cipher has even fewer keys.
- What about the affine cipher?
  It has 312 possible keys so it might seem a bit stronger.

  With modern computers, 312 is very small
  All keys can be checked.

  Even worse, with some simple frequency analysis, there are even easier ways to find the key

Classical & Contemporary Cryptology

# Affine Example

- **Select the key (a,b) = (7, 3)**
  - gcd(7,26) = 1
  - $7^{-1}$ mod 26 = 15 (since 7 x 15 = 105 and 105 mod 26 = 1)

The general encipher/decipher equations are:

$c = 7p + 3 \bmod 26$ $\quad$ $p = 15(c - 3) \bmod 26$

"hot" is 7 14 19

c(h) = 7*7 + 3 mod 26 = 52 mod 26 = 0 = a
c(o) = 7*14 + 3 mod 26 = 98 mod 26 = 23 = x
c(t) = 7*19 + 3 mod 26 = 133 mod 26 = 6 = g

Classical & Contemporary Cryptology

# Breaking an Affine Cipher

- How would you break the affine cipher?
  - Check all 312 (a,b) combinations
  - or, take advantage of the mathematical relationship

    $c = a*p + b \pmod{26}$
- Given this ciphertext from an affine cipher find the key and plaintext by using frequency analysis to guess two (a,b) pairs.

**FMXVE DKAPH FERBN DKRXR SREFM ORUDS DKDVS**
**HVUFE DKAPR KDLYE VLRHH RH**

Classical & Contemporary Cryptology

# Analysis 1

- Start with a frequency analysis of the ciphertext
  - the most frequent letters in order are:
    R D E H K F S V
  - Assuming that R is “e”and D is “t” implies:

c(4) = 17
c(19) = 3

4a + b = 17
19a + b = 3

⟶ a = 6, b = 19

WRONG, since gcd(6,26) = 2 so try another combination

Classical & Contemporary Cryptology

# Analysis 2

- Try other possible combinations:

R D E H K F S V

e t

| R = e, E = t | R = e, H = t | R = e, K = t |
|---|---|---|
| a = 13 | a = 8 | a = 3 |
| gcd(13,26) = 13 | gcd(8,26) = 2 | gcd(3,26) = 1<br>b = 5 |

Try on the ciphertext:

Algorithmsarequitegeneraldefinitions
ofarithmeticprocesses

Classical & Contempory Cryptology

# Keyword Cipher

- Caesar, multiplicative and affine ciphers can be easily broken by just checking all possible keys. We now introduce a monoalphabetic substitution cipher that can not be broken this way
- There will be many keys but still easy to remember
- Keyword cipher:

1. select a keyword – if any letters are repeated, drop the second and all other occurrences from the keyword
2. write the keyword below the alphabet, fill in the rest of the space with the remaining letters in the alphabet in their standard order

Classical & Contemporary Cryptology

# Example

- The keyword is COUNT

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

So a goes to c, b goes to o, . . .

Classical & Contemporary Cryptology

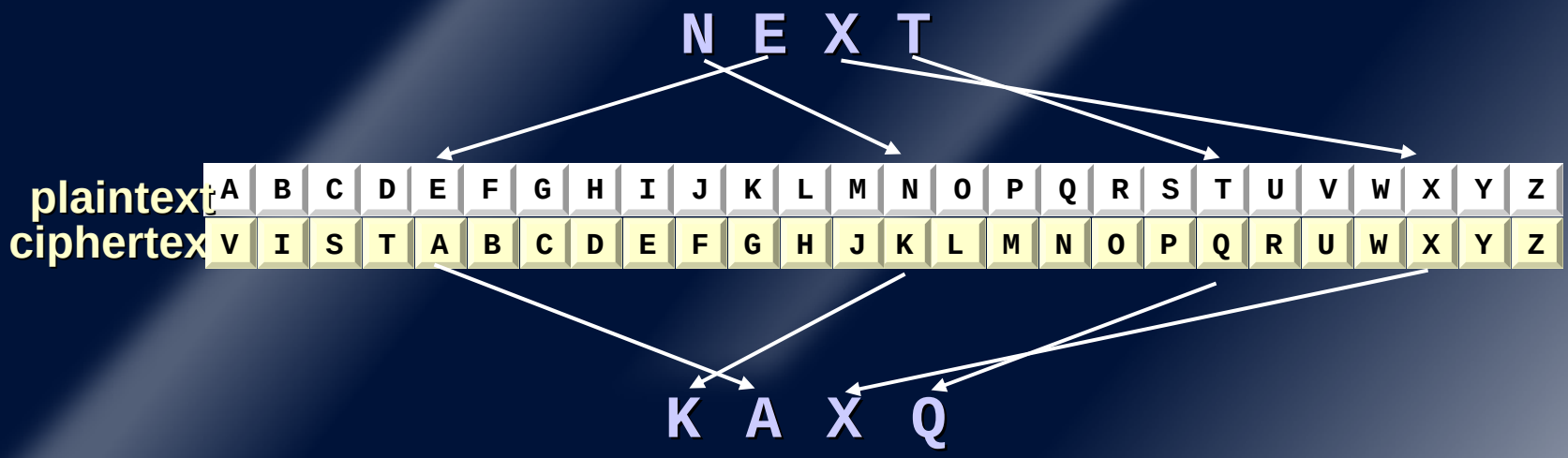
# Starting Position

- The keyword does not have to start at the beginning of the plaintext alphabet
  - it could start at any letter
  - for example, “count” could start at “k”

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

# Example

- If the keyword is “visit” (note, the second “i” in visit is dropped below) starting at “a” and the plaintext is “next”, the application is:



# Breaking a Keyword

- Surprisingly, the keyword cipher is *not* secure; in fact it is easy to break
- One reason why it is useful to study such a cipher is that in order to break this cipher you must use some of the most fundamental tools of cryptanalysis

Classical & Contemporary Cryptology

# Challenge

- **Consider the following ciphertext outputted by a simple monoalphabetic keyword substitution cipher:**

```
GJXXN GGOTZ NUCOT WMOHY JTKTA MTXOB YNFGO GINUG JFNZV QHYNG
NEAJF HYOTW GOTHY NAFZN FTUIN ZANFG NLNFU TXNXU FNEJC INHYA
ZGAEU TUCQG OGOTH JOHOA TCJXK HYNUV OCOHQ UHCNU GHHAF NUZHY
NCUTW JUWNA EHYNA FOWOT UCHNP HOGLN FQZNG OFUVC NZJHT AHNGG
NTHOU CGJXY OGHTN ABNTO TWGNT HNTXN AEBUF KNFYO HHGIU TJUCE
AFHYN GACJH OATAE IOCOH UFOXO BYNFG
```

**How would you go about breaking it?**

**We know that the plaintext is standard English and that each character in the ciphertext stands in for another character**

**So, what do we know about English that can help us?**

# Basic Cryptanalysis

- The most basic observation of cryptanalysis is that every letter of a language has its own personality.
  - if every plaintext t is changed to a ciphertext m, then in the ciphertext, m assumes the personality of t
  - to the trained observer, the personality of a letter gives away its identity
- Some of these personality characteristics are:
  - frequency of occurrence
  - contact with other letters (digrams, trigrams)
  - position within words

Classical & Contemporary Cryptology

# Letter Frequency

- What is the most frequent letter in English?
- Actually the frequency depends on the type of text. A widely used frequency table of 400 letters of standard English:

| Letter: | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Count: | 32 | 6 | 12 | 16 | 42 | 8 | 6 | 24 | 26 | 2 | 2 | 14 | 12 | 28 | 32 | 8 | 1 |

| Letter: | R | S | T | U | V | W | X | Y | Z |
|---|---|---|---|---|---|---|---|---|---|
| Count: | 26 | 24 | 36 | 12 | 4 | 6 | 2 | 8 | 1 |

In Order: ETAONIRSHDLUCMPFYWGBVJKQXZ

Classical & Contempory Cryptology

# Frequency Analysis

- **The frequency count for the challenge text is:**

| 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 17 | 4 | 13 | 0 | 7 | 17 | 23 | 26 | 5 | 12 | 3 | 2 | 2 | 36 | 25 | 1 | 5 | 0 | 0 | 23 | 20 | 3 | 6 | 9 | 13 | 8 |

**We could compare this with the expected frequency:**

```
Standard: ETAONIRSHDLUCMPFYWGBVJKQXZ
  Cipher: NHOGTUAFCYJXZEWIQBKVLMPDRS
```

**Result:** OLUUE OOANC EIHAN PJATD . . .

**This is not surprising since the two text items are based on different words**

**However, while relative frequencies may shift slightly, (i may be more frequent than a), they do not stray far from their area in the frequency table**

Classical & Contempory Cryptology

# Frequency Groups

- High Frequency Group
  - E T A O N I R S H
- Medium Frequency Group
  - D L U C M
- Low Frequency Group
  - P F Y W G B V
- Rare Group
  - J K Q X Z

There is usually a sharp break between the high and medium groups. That is, H is usually 6% and D is usually 4%

SO: look for the break between the high and medium group

Classical & Contemporary Cryptology

# Single Frequency Reasoning

- Things to look for in a frequency report
  - If there are hills and valleys similar to standard English then the cipher is most likely a substitution, so:
    - Find the break between high frequency and medium frequency (look for a 2% drop between two letters)
    - The most frequent letter is probably “e” or at least “t” or “a”

WARNING: this is only useful if you have enough text to maintain the “average” picture of frequency distributions

Classical & Contemporary Cryptology

# Challenge Frequency Report

- Again consider the frequency count for the challenge:

| N | H | O | G | T | U | A | F | C | Y | J | X | Z | E | W | I | Q | B | K | V | L | M | P | D | R | S |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 36 | 26 | 25 | 23 | 23 | 20 | 17 | 17 | 13 | 13 | 12 | 9 | 8 | 7 | 6 | 5 | 5 | 4 | 3 | 3 | 2 | 2 | 1 | 0 | 0 | 0 |

**Where is the break - that is, which set of characters are in the high frequency group?**

**Out of the possible high frequency group which is E?**

Classical & Contemporary Cryptology

# Next Step

- **Contact information will help**
  - **every letter has a cluster of preferred associations as part of its personality**
  - **these are called** digrams
  - **What are some of the most frequent digrams?**

**There are a number of characteristics of letter contacts**

**R forms digrams with more different letters more often than any other letter**

**The 3 vowels A, I, O avoid each other except for IO**

**EA is the most frequent digram involving vowels**

**80% of the letters which precede N are vowels**

**H frequently appears before E and almost never after it**

Classical & Contempory Cryptology

# Challenge Digrams

- This chart lists the digrams formed by the most frequent letters in the ciphertext:

**First task - identify (or confirm) E**

N is a good possibility by frequency counts

N also forms digrams with more characters than any other (17 - look at the full digram table)

| | N | H | O | G | T | U | A | F |
|---|---|---|---|---|---|---|---|---|
| N | 0 | 3 | 0 | 4 | 1 | 0 | 1 | 3 |
| H | 1 | 2 | 4 | 2 | 4 | 1 | 1 | 2 |
| O | 0 | 4 | 0 | 6 | 1 | 0 | 0 | 1 |
| G | 5 | 1 | 4 | 2 | 0 | 2 | 0 | 3 |
| T | 4 | 2 | 7 | 0 | 0 | 4 | 2 | 1 |
| U | 5 | 1 | 1 | 0 | 3 | 0 | 0 | 2 |
| A | 5 | 1 | 2 | 2 | 3 | 0 | 0 | 0 |
| F | 7 | 0 | 1 | 0 | 0 | 3 | 4 | 0 |

Classical & Contemporary Cryptology

# Consonants

- The easiest to spot is N because 80% of the letters that precede N are vowels
  - look for a high frequency letter which most often follows a vowel
  - for the challenge text, T follows one of the vowels (N, O, U, A) 17 out of 23 times
- H frequently appears before E and almost never after it
  - in the challenge, the pair YN occurs frequently but NY never occurs
- TH is common
  - if Y is really H, then H must be T because HY is common

# Current Status

- Using our best guess, the key looks like

```
plain:  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
cipher: U       N     Y O         T A         H
```

Evidence:

Frequency count suggests N is E

Contact data suggests that O is I

Contact data also suggests that A is O

So the remaining vowel suggests that U is A

Contact data suggests that Y is H

The common TH pair suggests that H is T

Contact with vowels suggests that T is N

Remember: this is only a best guess based on our observation some may be correct and some may be wrong

Classical & Contemporary Cryptology

# Challenge Text

- The challenge text looks like:

```
G J X X N G G O T Z N U C O T W M O H Y J T K T A M T X O B Y N F G O G I N U G
      E    I N  E A  I N   I T H  N  N O  N  I  H E    I    E A

J F N Z V Q H Y N G N E A J F H Y O T W G O T H Y N A F Z N F T U I N Z A N F G
   E      T H E  E  O   T H I N   I N T H E O    E  N A  E  O E

N L N F U T X N X U F N E J C I N H Y A Z G A E U T U C Q G O G O T H J O H O A
E  E  A   E  A  E      E T H O   O  A N A     I  I N T  I T I O

T C J X K H Y N U V O C O H Q U H C N U G H H A F N U Z H Y
N       T H E A  I  I T  A T  E A  T T O  E A  T H
```

**Wheel of Fortune Time - are there any words?**

Classical & Contemporary Cryptology

# Update

- **Work with both the text and the key:**

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

order — Keyword?

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

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

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

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

**We also know that the cipher key has some letters in order . . .**

# Summary

- Introduction to Ciphers
- Breaking Caesar, Multiplicative and Affine Ciphers
- Keyword Ciphers
- Breaking KeyWord Ciphers
  This shouldn't be done by hand.
  There are lots of good computer tools available, e.g.,
  http://www.cs.plu.edu/pub/faculty/spillman/CAP/index.htm **(associated with these slides)**
  http://www.cryptool.org/ **(freeware)**