# Part I: Crypto
# Chapter 2

# Chapter 2: Crypto Basics

MXDXBVTZWVMXNSPBQXLIMSCCSGXSCJXBOVQXCJZMOJZCVC
TVWJCZAAXZBCSSCJXBQCJZCOJZCNSPOXBXSBTVWJC
JZDXGXXMOZQMSCSCJXBOVQXCJZMOJZCNSPJZHGXXMOSPLH
JZDXZAAXZBXHCSCJXTCSGXSCJXBOVQX

□ plaintext from Lewis Carroll, *Alice in Wonderland*

The solution is by no means so difficult as you might
be led to imagine from the first hasty inspection of the characters.
These characters, as any one might readily guess,
form a cipher □ that is to say, they convey a meaning…

□ Edgar Allan Poe, *The Gold Bug*

# Crypto

- **Cryptology** □ The art and science of making and breaking "secret codes"
- **Cryptography** □ making "secret codes"
- **Cryptanalysis** □ breaking "secret codes"
- **Crypto** □ all of the above (and more)

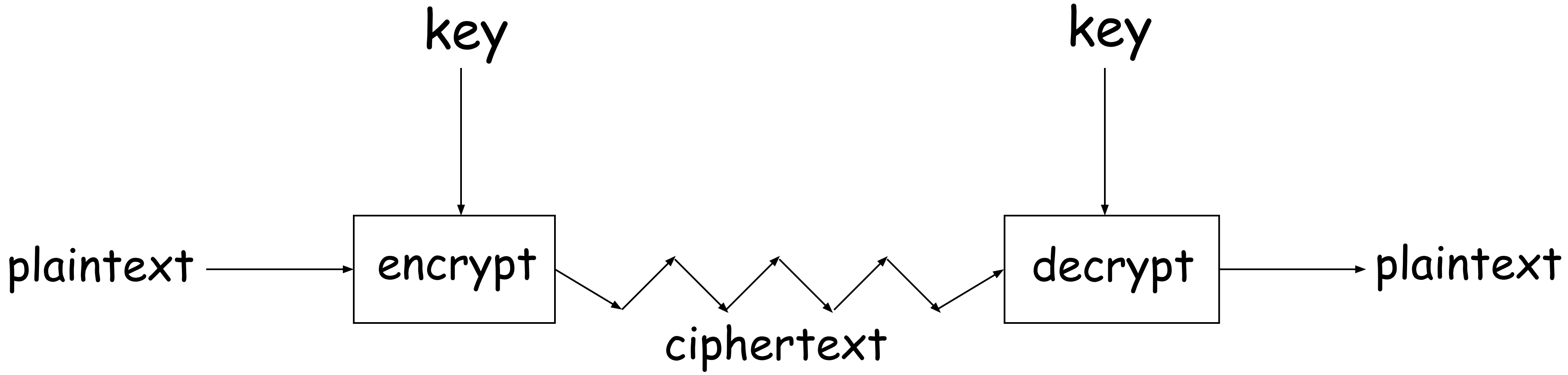
# How to Speak Crypto

- A *cipher* or *cryptosystem* is used to *encrypt* the *plaintext*
- The result of encryption is *ciphertext*
- We *decrypt* ciphertext to recover plaintext
- A *key* is used to configure a cryptosystem
- A *symmetric key* cryptosystem uses the same key to encrypt as to decrypt
- A *public key* cryptosystem uses a *public key* to encrypt and a *private key* to decrypt

# Crypto

- Basic assumptions
  - The system is completely known to the attacker
  - Only the key is secret
  - That is, crypto algorithms are not secret
- This is known as **Kerckhoffs' Principle**
- Why do we make such an assumption?
  - Experience has shown that secret algorithms tend to be weak when exposed
  - Secret algorithms never remain secret
  - Better to find weaknesses beforehand

# Crypto as Black Box



A generic view of symmetric key crypto

# Simple Substitution

- Plaintext: fourscoreandsevenyearsago
- Key:

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 | Y | Z | A | B | C |

- Ciphertext:

  IRXUVFRUHDQGVHYHQBHDUVDJR
- Shift by 3 is "Caesar's cipher"

# Ceasar's Cipher Decryption

- Suppose we know a Caesar's cipher is being used:

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 | Y | Z | A | B | C |

- Given ciphertext:
  VSRQJHEREVTXDUHSDQWV
- Plaintext: spongebobsquarepants

# Not-so-Simple Substitution

- Shift by $n$ for some $n \in \{0,1,2,\ldots,25\}$
- Then key is $n$
- Example: key $n = 7$

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 | 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 |

# Cryptanalysis I: Try Them All

- A simple substitution (shift by $n$) is used
  - But the key is unknown
- Given ciphertext: CSYEVIXIVQMREXIH
- How to find the key?
- Only 26 possible keys  try them all!
- **Exhaustive key search**
- Solution: key is $n = 4$

# Simple Substitution: General Case

- In general, simple substitution key can be any **permutation** of letters
  - Not necessarily a shift of the alphabet
- For example

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

- Then $26! > 2^{88}$ possible keys

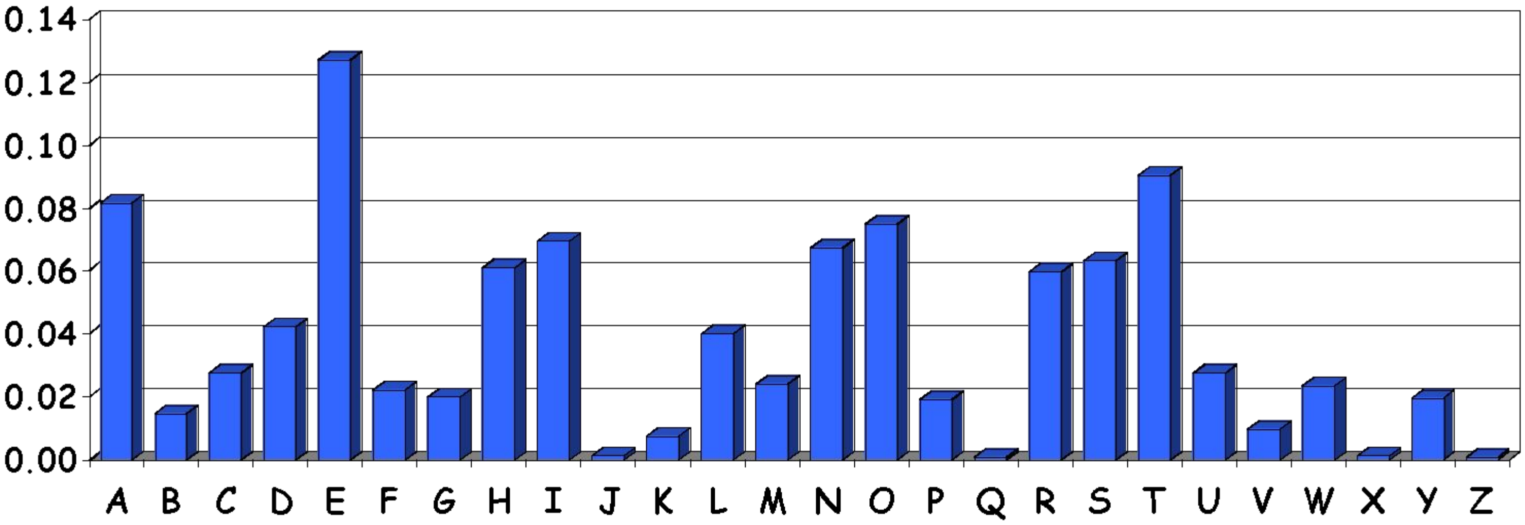
# Cryptanalysis II: Be Clever

- We know that a simple substitution is used
- But not necessarily a shift by n
- Find the key given the ciphertext:

```
 PBFPVYFBQXZTYFPBFEQJHDXXQVAPTPQJKTOYQWIPBVWLXTOX
BTFXQWAXBVCXQWAXFQJVWLEQNTOZQGGQLFXQWAKVWLXQ
WAEBIPBFXFQVXGTVJVWLBTPQWAEBFPBFHCVLXBQUFEVWLXGD
PEQVPQGVPPBFTIXPFHXZHVFAGFOTHFEFBQUFTDHZBQPOTHXTY
FTODXQHFTDPTOGHFQPBQWAQJJTODXQHFOQPWTBDHHIXQV
APBFZQHCFWPFHPBFIPBQWKFABVYYDZBOTHPBQPQJTQOTOGHF
QAPBFEQJHDXXQVAVXEBQPEFZBVFOJIWFFACFCCFHQWAUVWF
LQHGFXVAFXQHFUFHILTTAVWAFFAWTEVOITDHFHFQAITIXPFH
XAFQHEFZQWGFLVWPTOFFA
```

# Cryptanalysis II

- Cannot try all $2^{88}$ simple substitution keys
- Can we be more clever?
- English letter frequency counts…



# Cryptanalysis II

- Ciphertext:

```
 PBFPVYFBQXZTYFPBFEQJHDXXQVAPTPQJKTOYQWIPBVWLXTOXBTFXQ
WAXBVCXQWAXFQJVWLEQNTOZQGGQLFXQWAKVWLXQWAEBIPBFXFQ
VXGTVJVWLBTPQWAEBFPBFHCVLXBQUFEVWLXGDPEQVPQGVPPBFTIXPFH
XZHVFAGFOTHFEFBQUFTDHZBQPOTHXTYFTODXQHFTDPTOGHFQPBQW
AQJJTODXQHFOQPWTBDHHIXQVAPBFZQHCFWPFHPBFIPBQWKFABVYY
DZBOTHPBQPQJTQOTOGHFQAPBFEQJHDXXQVAVXEBQPEFZBVFOJIWFF
ACFCCFHQWAUVWFLQHGFXVAFXQHFUFHILTTAVWAFFAWTEVOITDHFH
FQAITIXPFHXAFQHEFZQWGFLVWPTOFFA
```

- Analyze this message using statistics below

Ciphertext frequency counts:

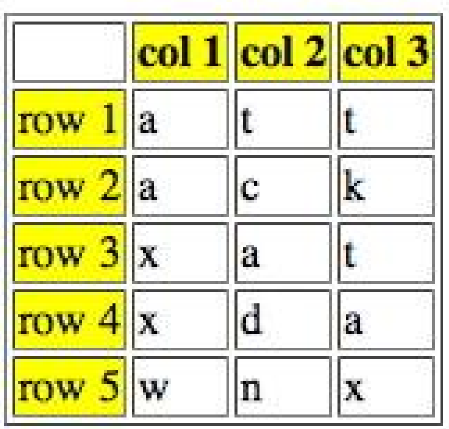
| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 21 | 26 | 6 | 10 | 12 | 51 | 10 | 25 | 10 | 9 | 3 | 10 | 0 | 1 | 15 | 28 | 42 | 0 | 0 | 27 | 4 | 24 | 22 | 28 | 6 | 8 |

# Cryptanalysis: Terminology

- Cryptosystem is **secure** if best know attack is to try all keys
  - Exhaustive key search, that is
- Cryptosystem is **insecure** if ***any*** shortcut attack is known
- But then insecure cipher might be harder to break than a secure cipher!
  - What the … ?

# Double Transposition

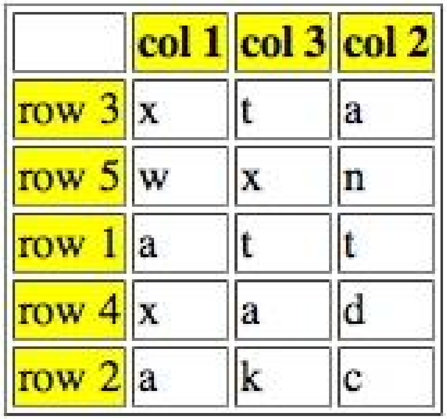
- Plaintext: attackxatxdawn



| | col 1 | col 2 | col 3 |
|---|---|---|---|
| row 1 | a | t | t |
| row 2 | a | c | k |
| row 3 | x | a | t |
| row 4 | x | d | a |
| row 5 | w | n | x |

Permute rows and columns

⇒



| | col 1 | col 3 | col 2 |
|---|---|---|---|
| row 3 | x | t | a |
| row 5 | w | x | n |
| row 1 | a | t | t |
| row 4 | x | a | d |
| row 2 | a | k | c |

- Ciphertext: xtawxnattxadakc
- Key is matrix size and permutations: (3,5,1,4,2) and (1,3,2)

# One-Time Pad: Encryption

**e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111**

Encryption: Plaintext ⊕ Key = Ciphertext

| | h | e | i | l | h | i | t | l | e | r |
|---|---|---|---|---|---|---|---|---|---|---|
| Plaintext: | 001 | 000 | 010 | 100 | 001 | 010 | 111 | 100 | 000 | 101 |
| Key: | 111 | 101 | 110 | 101 | 111 | 100 | 000 | 101 | 110 | 000 |
| Ciphertext: | 110 | 101 | 100 | 001 | 110 | 110 | 111 | 001 | 110 | 101 |
| | s | r | l | h | s | s | t | h | s | r |

# One-Time Pad: Decryption

**e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111**

Decryption: Ciphertext ⊕ Key = Plaintext

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | s | r | l | h | s | s | t | h | s | r |
| Ciphertext: | 110 | 101 | 100 | 001 | 110 | 110 | 111 | 001 | 110 | 101 |
| Key: | 111 | 101 | 110 | 101 | 111 | 100 | 000 | 101 | 110 | 000 |
| Plaintext: | 001 | 000 | 010 | 100 | 001 | 010 | 111 | 100 | 000 | 101 |
| | h | e | i | l | h | i | t | l | e | r |

# One-Time Pad

Double agent claims following “**key**” was used:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **s** | **r** | **l** | **h** | **s** | **s** | **t** | **h** | **s** | **r** |
| Ciphertext: | 110 | 101 | 100 | 001 | 110 | 110 | 111 | 001 | 110 | 101 |
| “**key**”: | 101 | 111 | 000 | 101 | 111 | 100 | 000 | 101 | 110 | 000 |
| “Plaintext”: | 011 | 010 | 100 | 100 | 001 | 010 | 111 | 100 | 000 | 101 |
| | **k** | **i** | **l** | **l** | **h** | **i** | **t** | **l** | **e** | **r** |

**e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111**

# One-Time Pad

Or claims the key is...

|  | s | r | l | h | s | s | t | h | s | r |
|---|---|---|---|---|---|---|---|---|---|---|
| Ciphertext: | 110 | 101 | 100 | 001 | 110 | 110 | 111 | 001 | 110 | 101 |
| "key": | 111 | 101 | 100 | 011 | 101 | 110 | 001 | 011 | 101 | 101 |
| "Plaintext": | 001 | 000 | 100 | 010 | 011 | 000 | 110 | 010 | 011 | 000 |
|  | h | e | l | i | k | e | s | i | k | e |

**e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111**

# One-Time Pad Summary

- Provably secure
  - Ciphertext gives **no** useful info about plaintext
  - All plaintexts are ***equally likely***
- BUT, only when be used correctly
  - Pad must be random, used only once
  - Pad is known only to sender and receiver
- Note: pad (key) is same size as message
- So, why not distribute msg instead of pad?

# Real-World One-Time Pad

- Project VENONA
  - Soviet spies encrypted messages from U.S. to Moscow in 30's, 40's, and 50's
  - Nuclear espionage, etc.
  - Thousands of messages
- Spy carried one-time pad into U.S.
- Spy used pad to encrypt secret messages
- Repeats within the "one-time" pads made cryptanalysis possible

# VENONA Decrypt (1944)

[C% Ruth] learned that her husband [v] was called up by the army but he was not sent to the front. He is a mechanical engineer and is now working at the ENORMOUS [ENORMOZ] [vi] plant in SANTA FE, New Mexico. [45 groups unrecoverable]

detain VOLOK [vii] who is working in a plant on ENORMOUS. He is a FELLOWCOUNTRYMAN [ZEMLYaK] [viii]. Yesterday he learned that they had dismissed him from his work. His active work in progressive organizations in the past was cause of his dismissal. In the FELLOWCOUNTRYMAN line LIBERAL is in touch with CHESTER [ix]. They meet once a month for the payment of dues. CHESTER is interested in whether we are satisfied with the collaboration and whether there are not any misunderstandings. He does not inquire about specific items of work [KONKRETNAYa RABOTA]. In as much as CHESTER knows about the role of LIBERAL's group we beg consent to ask C. through LIBERAL about leads from among people who are working on ENOURMOUS and in other technical fields.

- "Ruth" == Ruth Greenglass
- "Liberal" == Julius Rosenberg
- "Enormous" == the atomic bomb

# Codebook Cipher

- Literally, a book filled with "codewords"
- Zimmerman Telegram encrypted via codebook

| | |
|---|---|
| Februar | 13605 |
| fest | 13732 |
| finanzielle | 13850 |
| folgender | 13918 |
| Frieden | 17142 |
| Friedenschluss | 17149 |
| : | : |

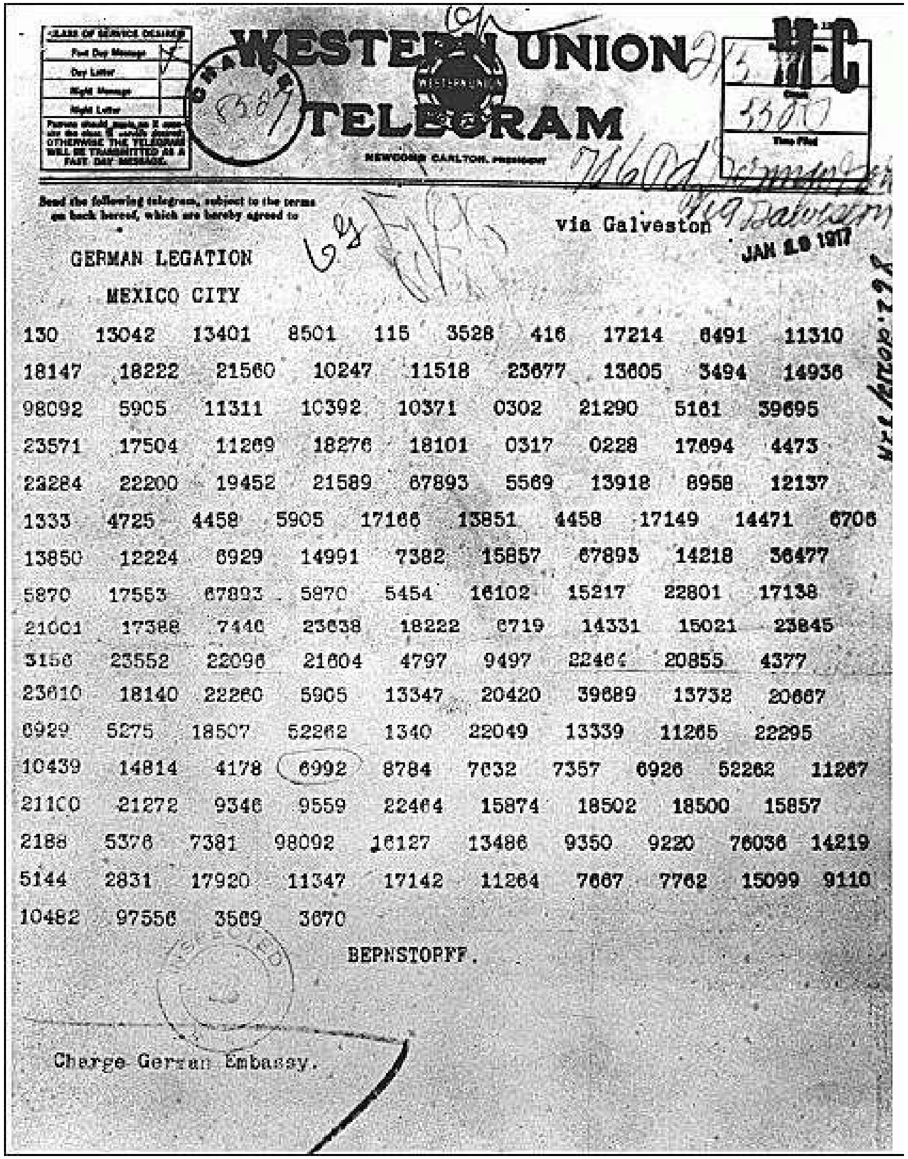
- Modern block ciphers are codebooks!
- More about this later…

# Codebook Cipher: Additive

- Codebooks also (usually) use **additive**
- Additive □ book of “random” numbers
  - Encrypt message with codebook
  - Then choose position in additive book
  - Add in additives to get ciphertext
  - Send ciphertext and additive position (MI)
  - Recipient subtracts additives before decrypting
- Why use an additive sequence?

# Zimmerman Telegram

- Perhaps most famous codebook ciphertext ever
- A major factor in U.S. entry into World War I



CLASS OF SERVICE DESIRED

Fast Day Message

Day Letter

Night Message

Night Letter

Patrons should mark an X opposite the class of service desired; OTHERWISE THE TELEGRAM WILL BE TRANSMITTED AS A FAST DAY MESSAGE.

WESTERN UNION TELEGRAM

NEWCOMB CARLTON, PRESIDENT

Check

Time Filed

Send the following telegram, subject to the terms on back hereof, which are hereby agreed to

via Galveston

JAN 19 1917

GERMAN LEGATION

MEXICO CITY

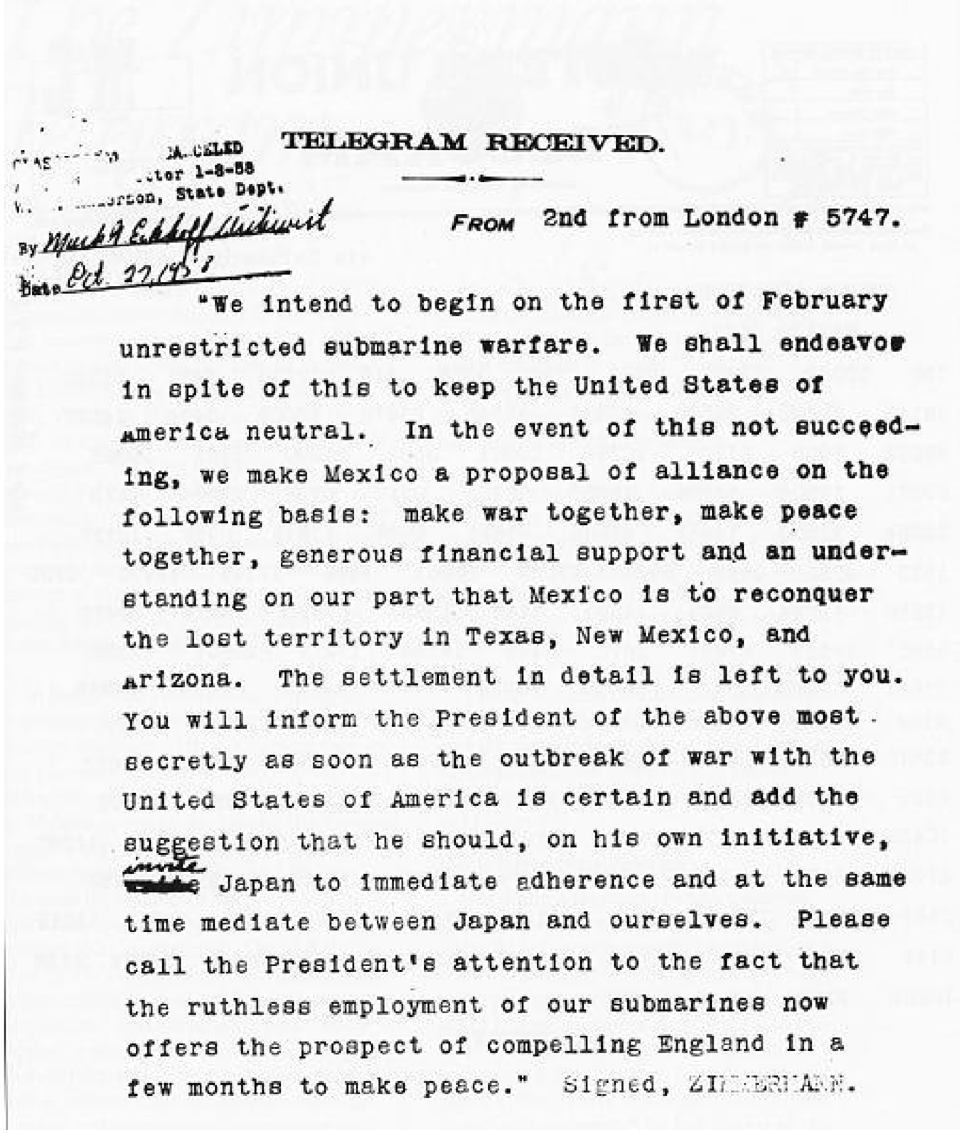
```
130    13042  13401  8501   115    3528   416    17214  6491   11310
18147  18222  21560  10247  11518  23677  13605  3494   14936
98092  5905   11311  10392  10371  0302   21290  5161   39695
23571  17504  11269  18276  18101  0317   0228   17694  4473
23284  22200  19452  21589  67893  5569   13918  8958   12137
1333   4725   4458   5905   17166  13851  4458   17149  14471  6706
13850  12224  6929   14991  7382   15857  67893  14218  36477
5870   17553  67893  5870   5454   16102  15217  22801  17138
21001  17388  7440   23638  18222  6719   14331  15021  23845
3156   23552  22096  21604  4797   9497   22464  20855  4377
23610  18140  22260  5905   13347  20420  39689  13732  20667
6929   5275   18507  52262  1340   22049  13339  11265  22295
10439  14814  4178   6992   8784   7632   7357   6926   52262  11267
21100  21272  9346   9559   22464  15874  18502  18500  15857
2188   5376   7381   98092  16127  13486  9350   9220   76036  14219
5144   2831   17920  11347  17142  11264  7667   7762   15099  9110
10482  97556  3569   3670
```

BEPNSTOPFF.

Charge German Embassy.

# Zimmerman Telegram Decrypted

- British had recovered partial codebook
- Then able to fill in missing parts



A..CELED
..tor 1-8-58
...orson, State Dept.
By
Date Oct. 22/1957

**TELEGRAM RECEIVED.**

FROM 2nd from London # 5747.

"We intend to begin on the first of February unrestricted submarine warfare. We shall endeavor in spite of this to keep the United States of America neutral. In the event of this not succeeding, we make Mexico a proposal of alliance on the following basis: make war together, make peace together, generous financial support and an understanding on our part that Mexico is to reconquer the lost territory in Texas, New Mexico, and Arizona. The settlement in detail is left to you. You will inform the President of the above most secretly as soon as the outbreak of war with the United States of America is certain and add the suggestion that he should, on his own initiative, invite Japan to immediate adherence and at the same time mediate between Japan and ourselves. Please call the President's attention to the fact that the ruthless employment of our submarines now offers the prospect of compelling England in a few months to make peace." Signed, ZIMMERMANN.

# Random Historical Items

- Crypto timeline
- Spartan Scytale  transposition cipher
- Caesar's cipher
- Poe's short story: ***The Gold Bug***
- Election of 1876

# Election of 1876

- "Rutherfraud" Hayes vs "Swindling" Tilden
  - Popular vote was virtual tie
- Electoral college delegations for 4 states (including Florida) in dispute
- Commission gave all 4 states to Hayes
  - Voted on straight party lines
- Tilden accused Hayes of bribery
  - Was it true?

# Election of 1876

- Encrypted messages by Tilden supporters later emerged
- Cipher: Partial codebook, plus transposition
- Codebook substitution for important words

| **ciphertext** | **plaintext** |
|---|---|
| Copenhagen | Greenbacks |
| Greece | Hayes |
| Rochester | votes |
| Russia | Tilden |
| **Warsaw** | **telegram** |
| : | : |

# Election of 1876

- Apply codebook to original message
- Pad message to multiple of 5 words (total length, 10,15,20,25 or 30 words)
- For each length, a fixed permutation applied to resulting message
- Permutations found by comparing several messages of same length
- Note that the **same key** is applied to all messages of a given length

# Election of 1876

- Ciphertext: **Warsaw they read all unchanged last are idiots can't situation**
- Codebook: Warsaw == telegram
- Transposition: 9,3,6,1,10,5,2,7,4,8
- Plaintext: **Can't read last telegram. Situation unchanged. They are all idiots.**
- A weak cipher made worse by reuse of key
- Lesson? Don't overuse keys!

# Early 20th Century

- WWI □ Zimmerman Telegram
- "Gentlemen do not read each other's mail"
  - Henry L. Stimson, Secretary of State, 1929
- WWII □ golden age of cryptanalysis
  - Midway/Coral Sea
  - Japanese **Purple** (codename **MAGIC**)
  - German **Enigma** (codename **ULTRA**)

# Post-WWII History

- Claude Shannon  father of the science of information theory
- Computer revolution  lots of data to protect
- Data Encryption Standard (DES), 70's
- Public Key cryptography, 70's
- CRYPTO conferences, 80's
- Advanced Encryption Standard (AES), 90's
- The crypto genie is out of the bottle…

# Claude Shannon

- The founder of Information Theory
- 1949 paper: *Comm. Thy. of Secrecy Systems*
- Fundamental concepts
  - **Confusion** □ obscure relationship between plaintext and ciphertext
  - **Diffusion** □ spread plaintext statistics through the ciphertext
- Proved one-time pad is secure
- One-time pad is confusion-only, while double transposition is diffusion-only

# Taxonomy of Cryptography

- **Symmetric Key**
  - Same key for encryption and decryption
  - Modern types: Stream ciphers, Block ciphers
- **Public Key** (or “asymmetric” crypto)
  - Two keys, one for encryption (public), and one for decryption (private)
  - And digital signatures  nothing comparable in symmetric key crypto
- **Hash algorithms**
  - Can be viewed as “one way” crypto

# Taxonomy of Cryptanalysis

- From perspective of info available to Trudy…
  - Ciphertext only  Trudy's worst case scenario
  - Known plaintext
  - Chosen plaintext
    - "Lunchtime attack"
    - Some protocols will encrypt chosen data
  - Adaptively chosen plaintext
  - Related key
  - Forward search (public key crypto)
  - And others…