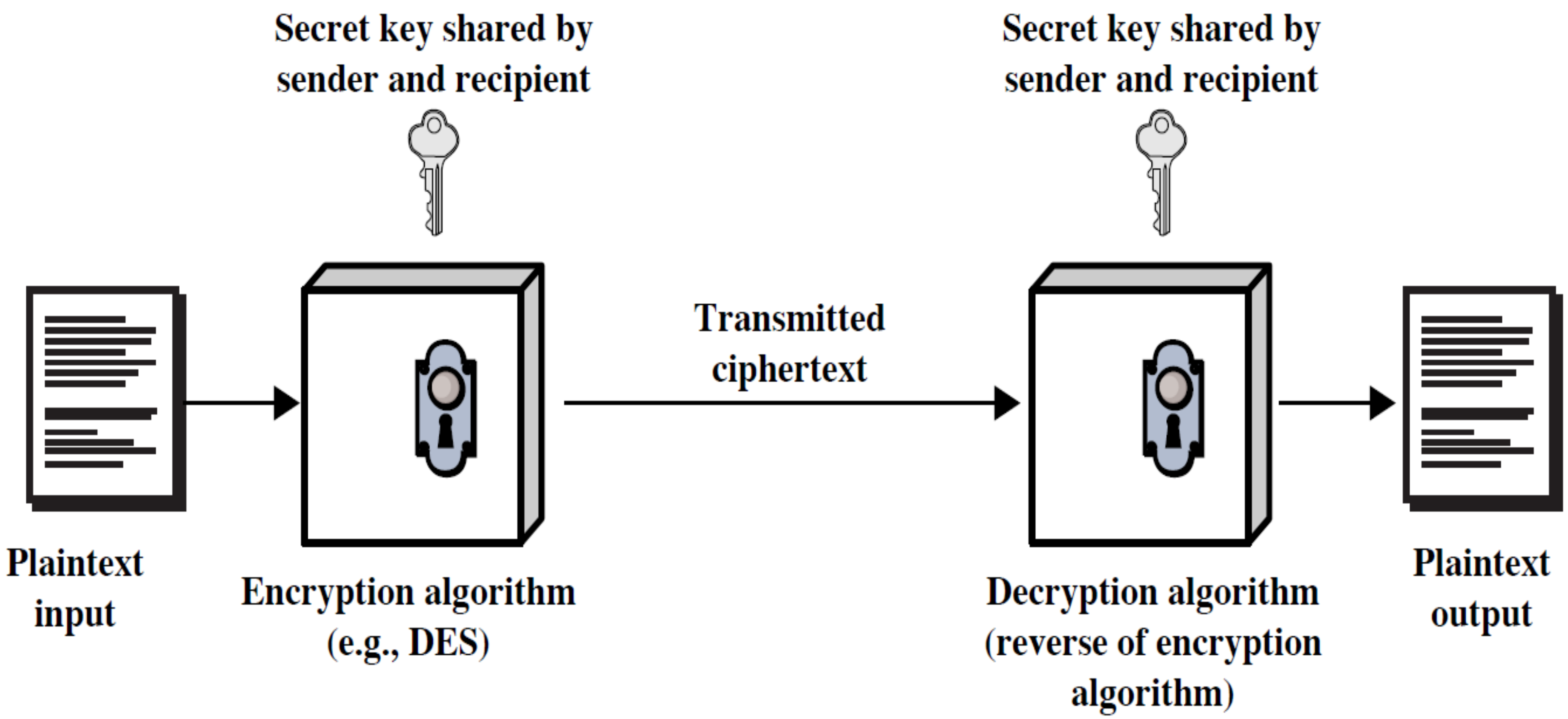
Classical Encryption Techniques

- As opposed to modern cryptography
- Symmetric-key

Symmetric Cipher Model



Brute-Force Attack

- Try every key to decipher the ciphertext.
- On average, need to try half of all possible keys
- Time needed proportional to size of key space

Key Size (bits)	Number of Alternative Keys	Time required at 1 decryption/µs	Time required at 10^6 decryptions/µs
32	$2^{32} = 4.3 \times 10^9$	2^{31} µs = 35.8 minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	2^{55} µs = 1142 years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	2^{127} µs = 5.4×10^{24} years	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	2^{167} µs = 5.9×10^{36} years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	2×10^{26} µs = 6.4×10^{12} years	6.4×10^6 years

Classical Ciphers

- Plaintext is viewed as a sequence of elements (e.g., bits or characters)
- Substitution cipher: replacing each element of the plaintext with another element.
- Permutation cipher: rearranging the order of the elements of the plaintext.
- Product cipher: using multiple stages of substitutions and permutations

Caesar Cipher

- Earliest known substitution cipher
- Invented by Julius Caesar
- Each letter is replaced by the letter three positions further down the alphabet.
- 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: 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
- Example: ohio state → RKLR VWDWH

Caesar Cipher

- Mathematically, map letters to numbers:

```
a, b, c, ..., x,  y,  z
0, 1, 2, ..., 23, 24, 25
```

- Then the general Caesar cipher is:

$$c = E_K(p) = (p + k) \bmod 26$$

$$p = D_K(c) = (c - k) \bmod 26$$

- Can be generalized with any alphabet.

Cryptanalysis of Caesar Cipher

- Key space: {0, 1, ..., 25}
- Vulnerable to brute-force attacks.
- E.g., break ciphertext "UNOU YZGZK“

- Need to recognize it when have the plaintext
- What if the plaintext is written in Swahili?

Monoalphabetic Substitution Cipher

- Shuffle the letters and map each plaintext letter to a different random ciphertext letter:

Plain letters: abcdefghijklmnopqrstuvwxyz

Cipher letters: DKVQFIBJWPESCXHTMYAUOLRGZN

Plaintext: ifwewishtoreplaceletters

Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA

- What does a key look like?

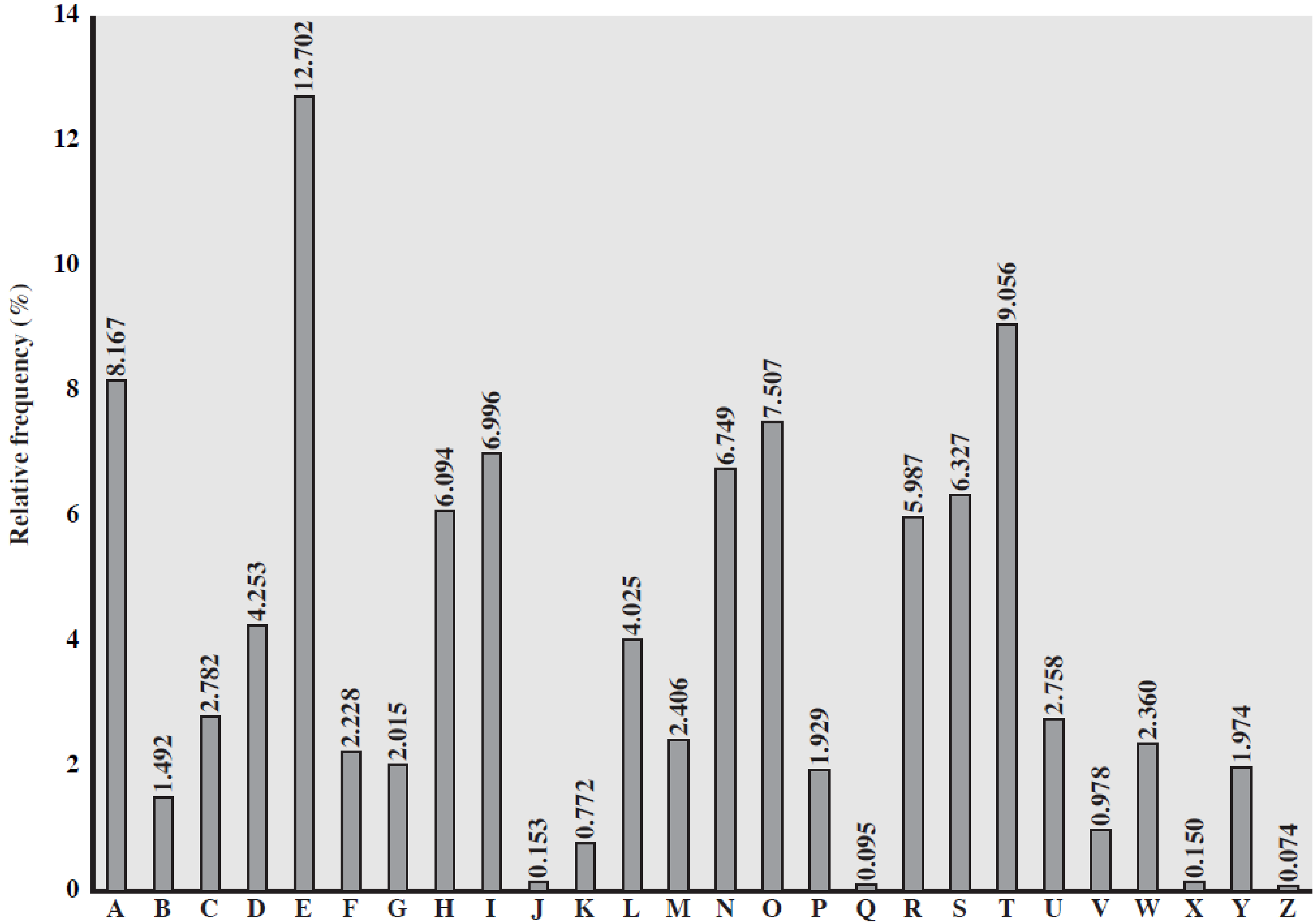
Monoalphabetic Cipher Security

- Now we have a total of 26! = 4 x 10^{26} keys.
- With so many keys, it is secure against brute-force attacks.
- But not secure against some cryptanalytic attacks.
- Problem is language characteristics.

Language Statistics and Cryptanalysis

- Human languages are not random.
- Letters are not equally frequently used.
- In English, E is by far the most common letter, followed by T, R, N, I, O, A, S.
- Other letters like Z, J, K, Q, X are fairly rare.
- There are tables of single, double & triple letter frequencies for various languages

English Letter Frequencies



Statistics for double & triple letters

- In decreasing order of frequency
- Double letters:

  th he an in er re es on, …
- Triple letters:

  the and ent ion tio for nde, …

Use in Cryptanalysis

- Key concept: monoalphabetic substitution does not change relative letter frequencies

- To attack, we
  - calculate letter frequencies for ciphertext
  - compare this distribution against the known one

Example Cryptanalysis

- Given ciphertext:

  `UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ`
  `VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX`
  `EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ`

- Count relative letter frequencies (see next page)
- Guess {P, Z} = {e, t}
- Of double letters, ZW has highest frequency, so guess ZW = th and hence ZWP = the
- Proceeding with trial and error finally get:

  `it was disclosed yesterday that several informal but`
  `direct contacts have been made with political`
  `representatives of the viet cong in moscow`

Letter frequencies in ciphertext

P 13.33	H 5.83	F 3.33	B 1.67	C 0.00
Z 11.67	D 5.00	W 3.33	G 1.67	K 0.00
S 8.33	E 5.00	Q 2.50	Y 1.67	L 0.00
U 8.33	V 4.17	T 2.50	I 0.83	N 0.00
O 7.50	X 4.17	A 1.67	J 0.83	R 0.00
M 6.67				

What kind of attack?

- Ciphertext-only attack
- Known-plaintext attack
- Chosen-plaintext attack
- Chosen-ciphertext attack

Playfair Cipher

- Not even the large number of keys in a monoalphabetic cipher provides security.
- One approach to improving security is to encrypt multiple letters at a time.
- The **Playfair Cipher** is the best known such cipher.
- Invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair.

Playfair Key Matrix

- Use a 5 x 5 matrix.
- Fill in letters of the key (w/o duplicates).
- Fill the rest of matrix with other letters.
- E.g., key = MONARCHY.

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

Encrypting and Decrypting

Plaintext is encrypted two letters at a time.

1. If a pair is a repeated letter, insert filler like 'X’.
2. If both letters fall in the same row, replace each with the letter to its right (circularly).
3. If both letters fall in the same column, replace each with the the letter below it (circularly).
4. Otherwise, each letter is replaced by the letter in the same row but in the column of the other letter of the pair.

Security of Playfair Cipher

- Equivalent to a monoalphabetic cipher with an alphabet of 26 x 26 = 676 characters.
- Security is much improved over the simple monoalphabetic cipher.
- Was widely used for many decades
  - eg. by US & British military in WW1 and early WW2
- Once thought to be unbreakable.
- Actually, it **can** be broken, because it still leaves some structure of plaintext intact.

Polyalphabetic Substitution Ciphers

- A sequence of monoalphabetic ciphers (M_1, M_2, M_3, ..., M_k) is used in turn to encrypt letters.
- A key determines which sequence of ciphers to use.
- Each plaintext letter has multiple corresponding ciphertext letters.
- This makes cryptanalysis harder since the letter frequency distribution will be flatter.

Vigenère Cipher

- Simplest polyalphabetic substitution cipher
- Consider the set of all Caesar ciphers:

$$\{ C_a, C_b, C_c, \ldots, C_z \}$$

- Key: e.g. security
- Encrypt each letter using C_s, C_e, C_c, C_u, C_r, C_i, C_t, C_y in turn.
- Repeat from start after C_y.
- Decryption simply works in reverse.

Security of Vigenère Ciphers

- There are multiple (how many?) ciphertext letters corresponding to each plaintext letter.
- So, letter frequencies are obscured but not totally lost.
- To break Vigenere cipher:

1. Try to guess the key length. How?
2. If key length is N, the cipher consists of N Caesar ciphers. Plaintext letters at positions k, N+k, 2N+k, 3N+k, etc., are encoded by the same cipher.
3. Attack each individual cipher as before.

Example of Vigenère Cipher

- Keyword: *deceptive*

key:

`deceptivedeceptivedeceptive`

plaintext:

`wearediscoveredsaveyourself`

ciphertext: `ZICVTWQNGRZGVTWAVZHCQYGLMGJ`

Guessing the Key Length

- Main idea: Plaintext words separated by multiples of the key length are encoded in the same way.
- In our example, if plaintext = "…thexxxxxxthe…" then "the" will be encrypted to the same ciphertext words.
- So look at the ciphertext for repeated patterns.
- E.g. repeated "VTW" in the previous example suggests a key length of 3 or 9:

  `ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ`

- Of course, the repetition could be a random fluke.

Rotor Cipher Machines

- Before modern ciphers, rotor machines were most common complex ciphers in use.
- Widely used in WW2.
- Used a series of rotating cylinders.
- Implemented a polyalphabetic substitution cipher of period K.
- With 3 cylinders, $K = 26^3 = 17{,}576$.
- With 5 cylinders, $K = 26^5 = 12 \times 10^6$.

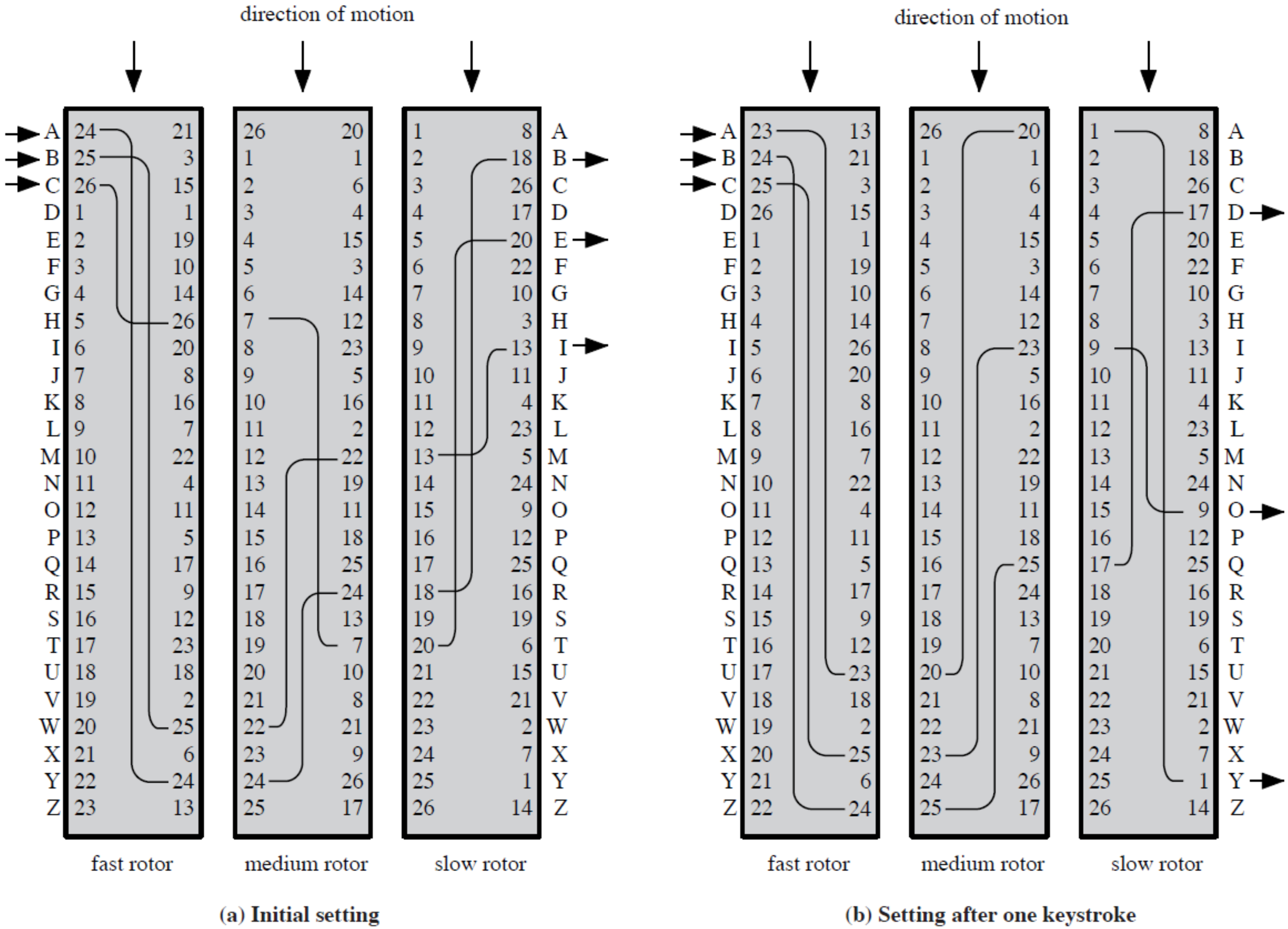


Figure 2.7 Three-Rotor Machine With Wiring Represented by Numbered Contacts

German secret setting sheets

Geheim! Secret indeed! This is an example of the setting shee

Geheim!

Nicht im Flugzeug mitnehmen!

Sonder-Maschinenschlüssel BGT

Datum	Walzenlage			Ringstellung			Steckerverbindungen									
31.	I	V	III	06	20	24	UA	PF	RQ	SO	NI	EY	BG	HL	TX	ZJ
30.	V	II	III	01	07	12	GF	KV	JM	[illegible]B	UW	LX	TD	QS	NA	ZH
29.	IV	I	V	11	17	26	CI	OK	PV	ZL	HX	NB	AW	DJ	FE	ST

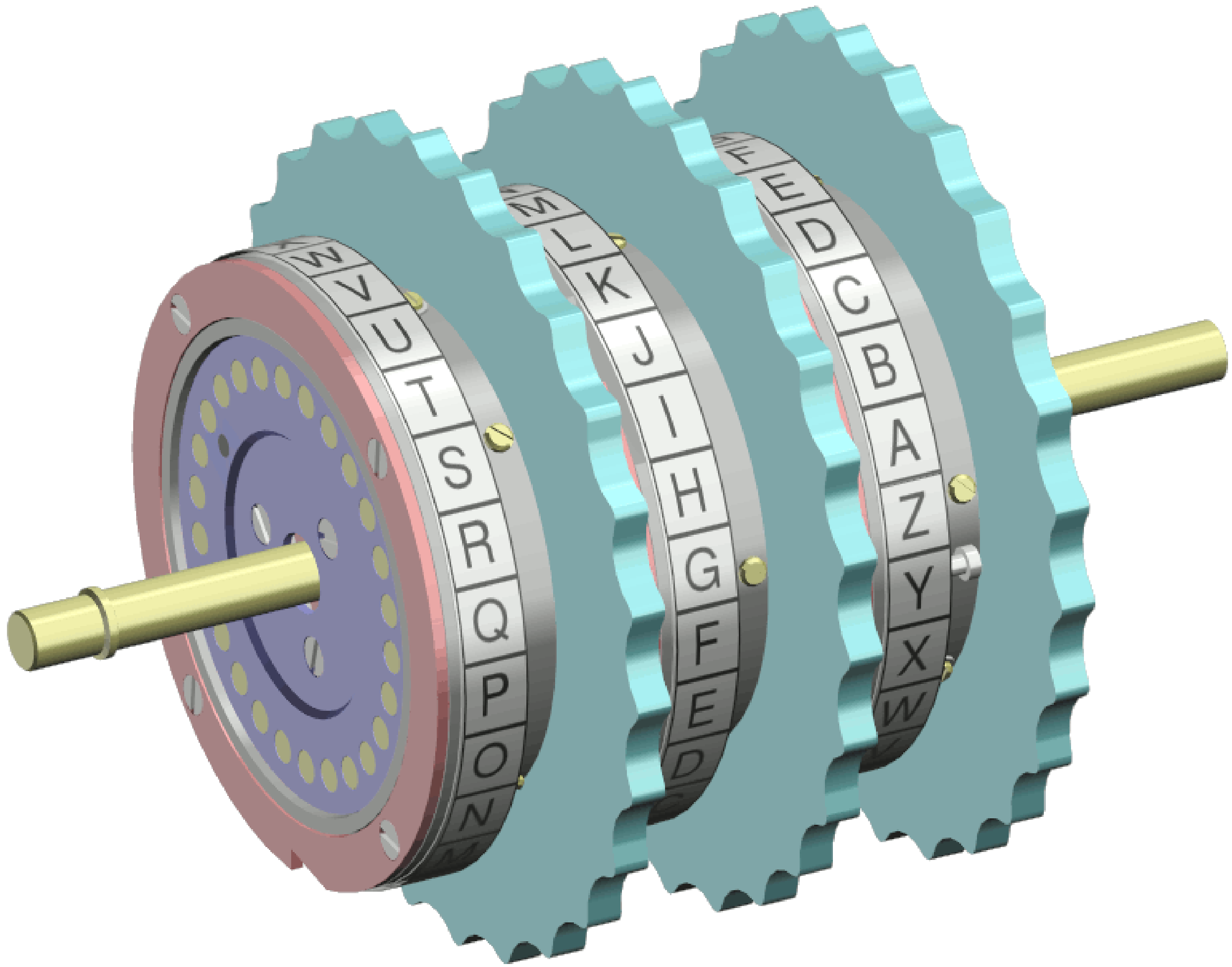
Date

Which rotors to use (there were 10 rotors)

Ring setting

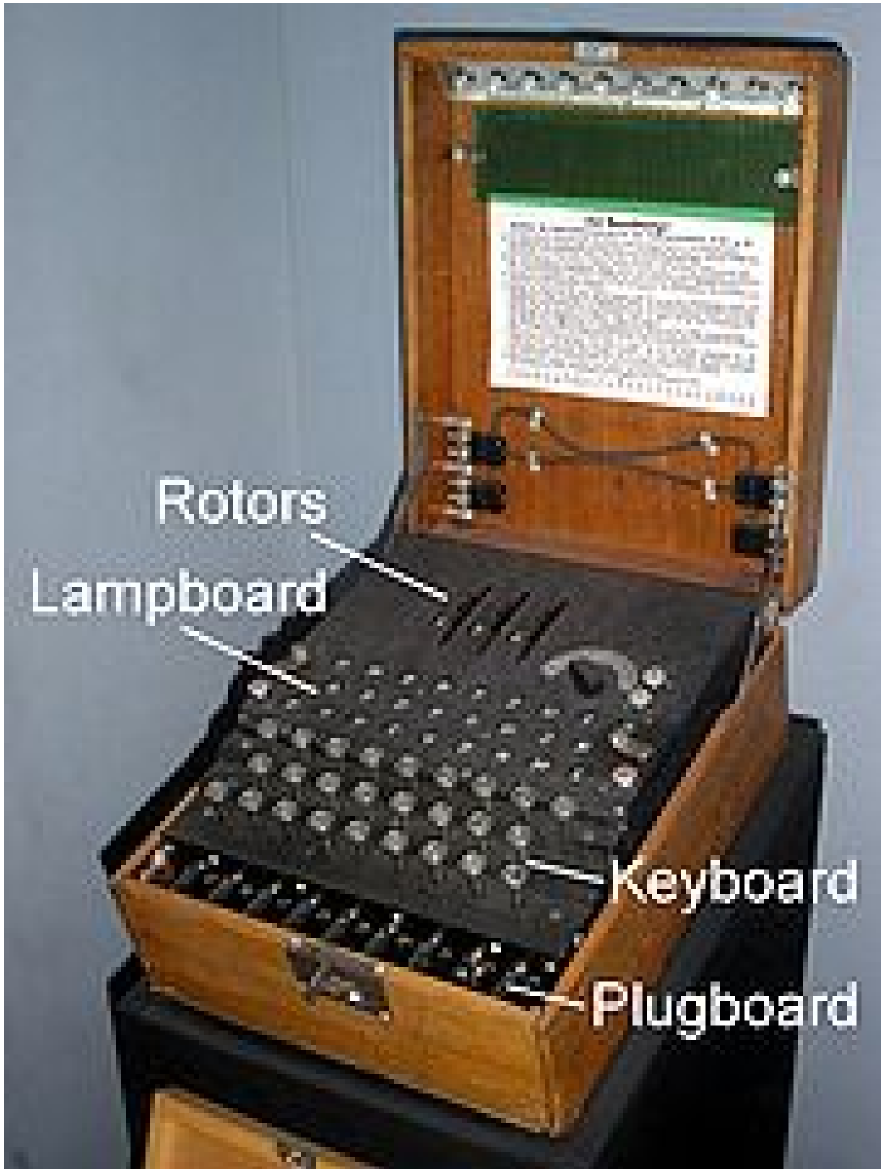
Plugboard setting

The Rotors



Enigma Rotor Machine

Enigma Rotor Machine



Transposition Ciphers

- Also called **permutation** ciphers.
- Shuffle the plaintext, without altering the actual letters used.
- Example: Row Transposition Ciphers

Row Transposition Ciphers

- Plaintext is written row by row in a rectangle.
- Ciphertext: write out the columns in an order specified by a key.

Key: 3 4 2 1 5 6 7

Plaintext:

a	t	t	a	c	k	p
o	s	t	p	o	n	e
d	u	n	t	i	l	t
w	o	a	m	x	y	z

Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ

Product Ciphers

- Uses a sequence of substitutions and transpositions
  - Harder to break than just substitutions or transpositions
- This is a bridge from classical to modern ciphers.