



Chapter 3

Classical Encryption Techniques

Definitions

Plaintext

- An original message

Ciphertext

- The coded message

Enciphering/encryption

- The process of converting from plaintext to ciphertext

Deciphering/decryption

- Restoring the plaintext from the ciphertext

Cryptography

- The area of study of the many schemes used for encryption

Cryptographic system/cipher

- A scheme

Cryptanalysis

- Techniques used for deciphering a message without any knowledge of the enciphering details

Cryptology

- The areas of cryptography and cryptanalysis

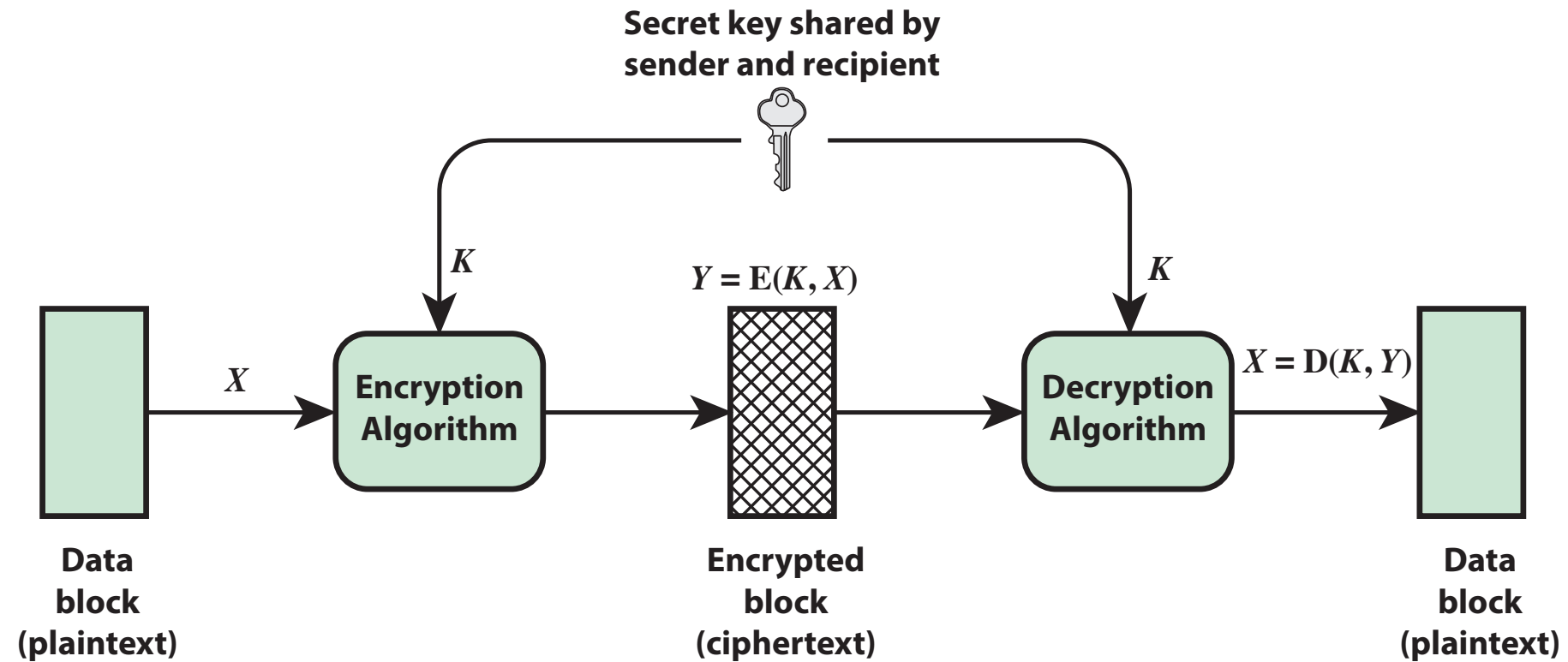
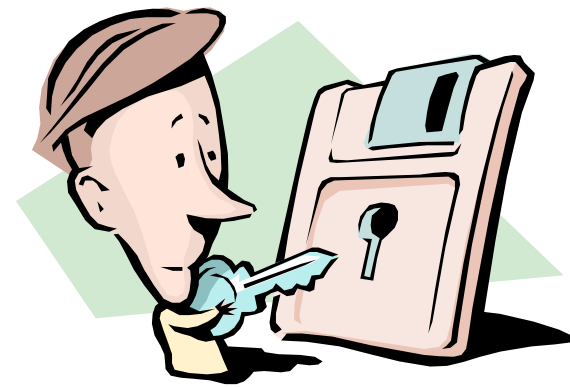


Figure 3.1 Simplified Model of Symmetric Encryption

Symmetric Cipher Model

- There are two requirements for secure use of conventional encryption:
 - A strong encryption algorithm
 - Sender and receiver must have obtained copies of the secret key in a secure fashion and must keep the key secure



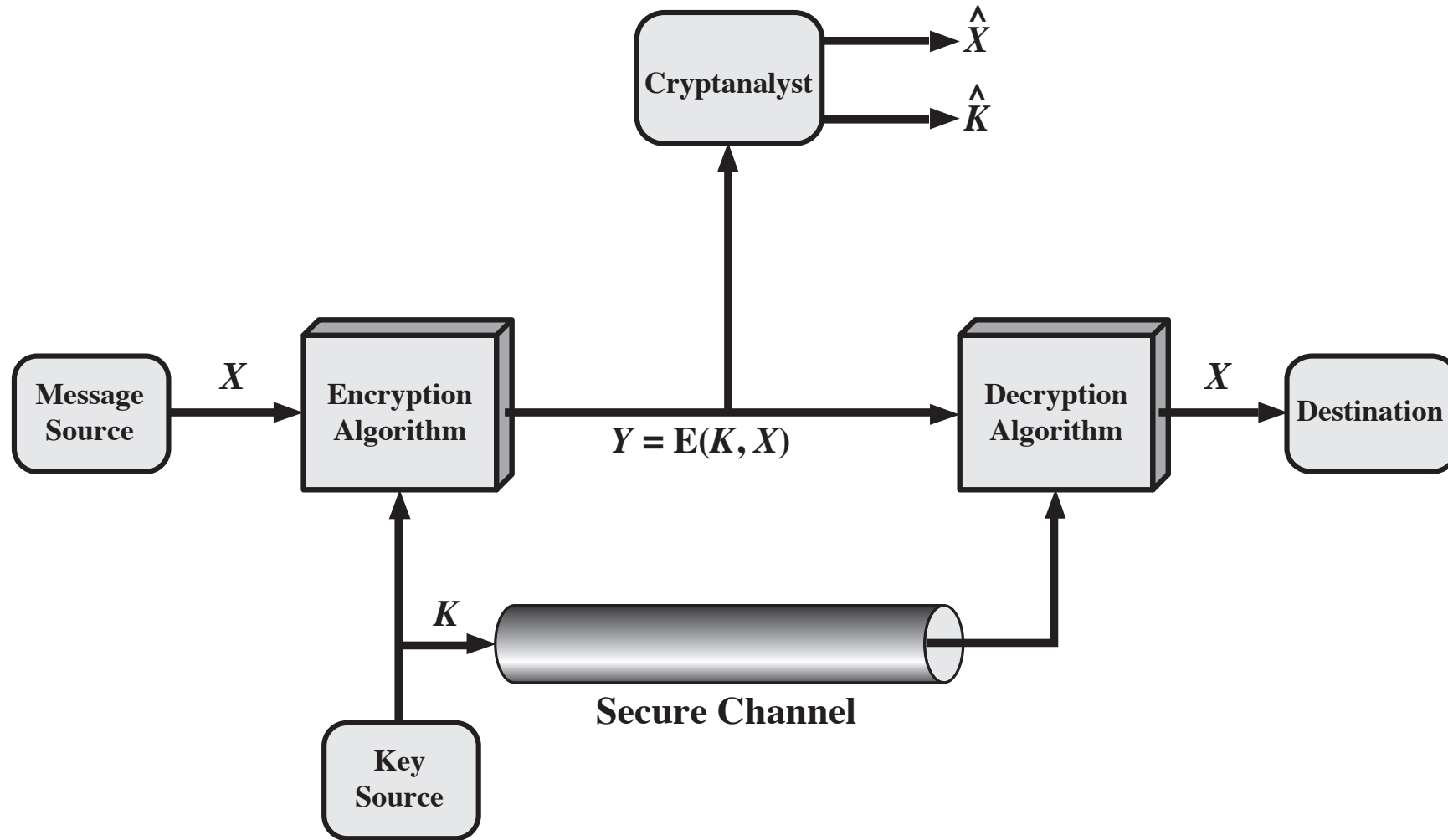
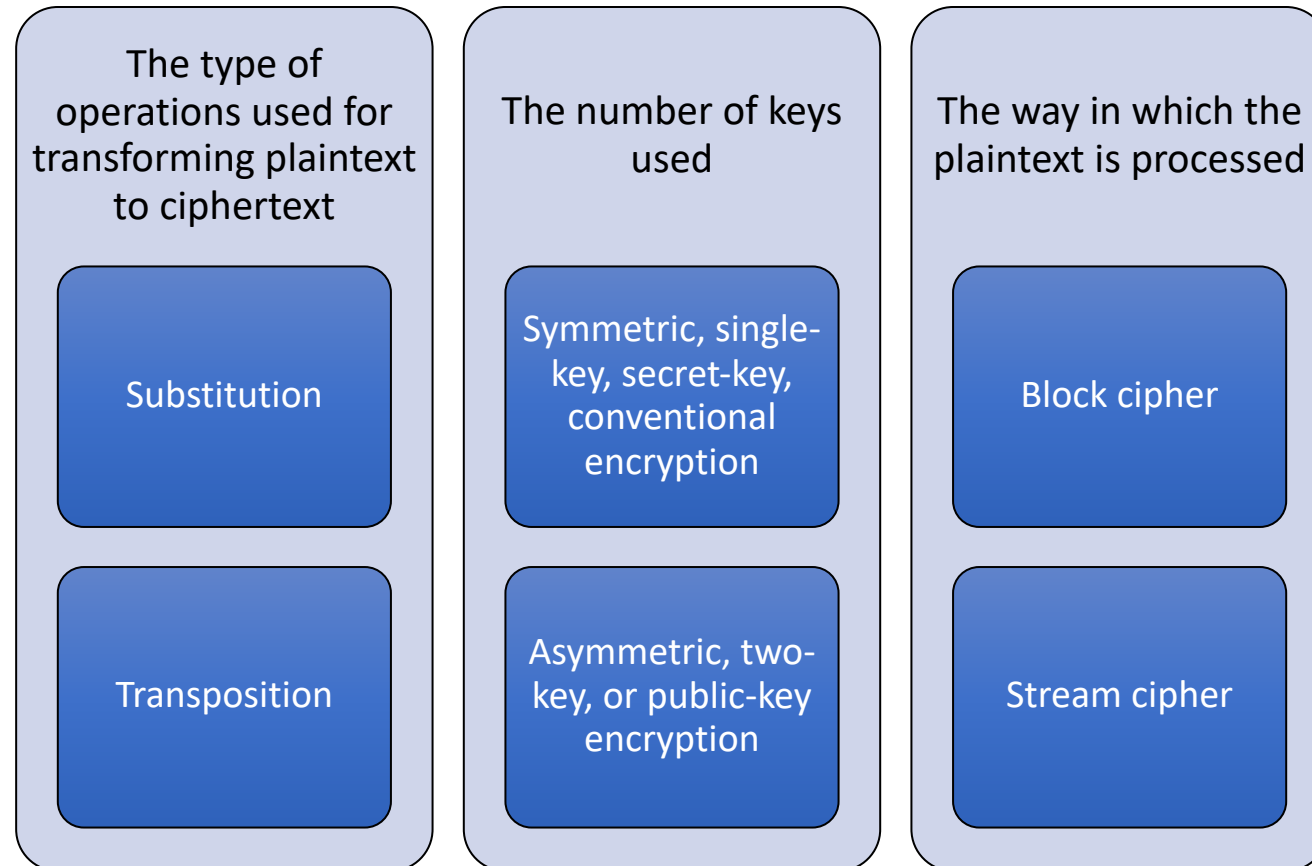


Figure 3.2 Model of Symmetric Cryptosystem

Cryptographic Systems

- Characterized along three independent dimensions:



Cryptanalysis and Brute-Force Attack

Cryptanalysis

- Attack relies on the nature of the algorithm plus some knowledge of the general characteristics of the plaintext
- Attack exploits the characteristics of the algorithm to attempt to deduce a specific plaintext or to deduce the key being used

Brute-force attack

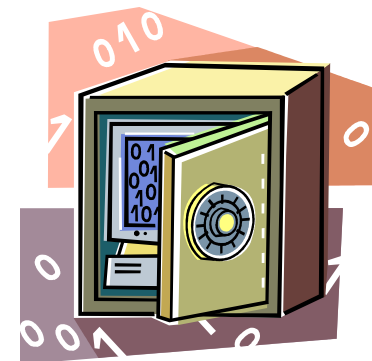
- Attacker tries every possible key on a piece of ciphertext until an intelligible translation into plaintext is obtained
- On average, half of all possible keys must be tried to achieve success

Types of Attacks on Encrypted Messages

Type of Attack	Known to Cryptanalyst
Ciphertext Only	<ul style="list-style-type: none"> • Encryption algorithm • Ciphertext
Known Plaintext	<ul style="list-style-type: none"> • Encryption algorithm • Ciphertext • One or more plaintext-ciphertext pairs formed with the secret key
Chosen Plaintext	<ul style="list-style-type: none"> • Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key
Chosen Ciphertext	<ul style="list-style-type: none"> • Encryption algorithm • Ciphertext • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
Chosen Text	<ul style="list-style-type: none"> • Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key

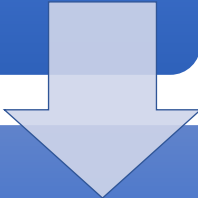
Encryption Scheme Security

- Unconditionally secure
 - No matter how much time an opponent has, it is impossible for him or her to decrypt the ciphertext simply because the required information is not there
- Computationally secure
 - The cost of breaking the cipher exceeds the value of the encrypted information
 - The time required to break the cipher exceeds the useful lifetime of the information

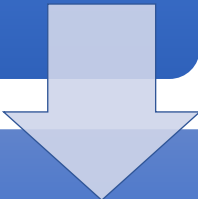


Brute-Force Attack

Involves trying every possible key until an intelligible translation of the ciphertext into plaintext is obtained



On average, half of all possible keys must be tried to achieve success



To supplement the brute-force approach, some degree of knowledge about the expected plaintext is needed, and some means of automatically distinguishing plaintext from garble is also needed

Strong Encryption

- The term *strong encryption* refers to encryption schemes that make it impractically difficult for unauthorized persons or systems to gain access to plaintext that has been encrypted
- Properties that make an encryption algorithm strong are:
 - Appropriate choice of cryptographic algorithm
 - Use of sufficiently long key lengths
 - Appropriate choice of protocols
 - A well-engineered implementation
 - Absence of deliberately introduced hidden flaws

Substitution Technique

- Is one in which the letters of plaintext are replaced by other letters or by numbers or symbols
- If the plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns



Caesar Cipher



- Simplest and earliest known use of a substitution cipher
- Used by Julius Caesar
- Involves replacing each letter of the alphabet with the letter standing three places further down the alphabet
- Alphabet is wrapped around so that the letter following Z is A

plain: meet me after the toga party

cipher: PHHW PH DIWHU WKH WRJD SDUWB

Caesar Cipher Algorithm



- Can define transformation as:

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
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

- Mathematically give each letter a number

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- Algorithm can be expressed as:

$$c = E(3, p) = (p + 3) \bmod (26)$$

- A shift may be of any amount, so that the general Caesar algorithm is:

$$C = E(k, p) = (p + k) \bmod 26$$

- Where k takes on a value in the range 1 to 25; the decryption algorithm is simply:

$$p = D(k, C) = (C - k) \bmod 26$$

Figure 3.3

Brute-Force Cryptanalysis of Caesar Cipher

(This chart can be found on page 71 in the textbook)

KEY	PHHW	PH	DIWHU	WKH	WRJD	SDUWB
1	oggv	og	chvgt	vjg	vqic	rctva
2	nffu	nf	bgufs	uif	uphb	qbsuz
3	meet	me	after	the	toga	party
4	ldds	ld	zesdq	sgd	snfz	ozqsx
5	kccr	kc	ydrpc	rfe	rmey	nyprw
6	jbbq	jb	xcqbo	qeb	qldx	mxoqv
7	iaap	ia	wbpan	pda	pkcw	lwnpu
8	hzzo	hz	vaozm	ocz	objv	kvmot
9	gyyn	gy	uznyl	nby	niau	julns
10	fxxm	fx	tymxk	max	mhzt	itkmr
11	ewwl	ew	sxlwj	lzw	lgys	hsjlg
12	dvvk	dv	rwkvi	kyv	kfxr	grikp
13	cuuj	cu	qvjuh	jxu	jewq	fghjo
14	btti	bt	puitg	iwt	idvp	epgin
15	assh	as	othsf	hvs	hcuo	dofhm
16	zrrg	zr	nsgr	gur	gbtn	cnegl
17	yqqf	yq	mrfqd	ftq	fasm	bmdfk
18	xppe	xp	lqepc	esp	ezrl	alcej
19	wood	wo	kpdob	dro	dyqk	zkbdi
20	vnnc	vn	jocna	cqn	cxpj	yjach
21	ummb	um	inbmz	bpm	bwoi	xizbg
22	tlla	tl	hmaly	aol	avnh	whyaf
23	skkz	sk	glzcx	znk	zumg	vgxze
24	rjjy	rj	fkyjw	ymj	ytlf	ufwyd
25	qiix	qi	ejxiv	xli	xske	tevx

Figure 3.3 Brute-Force Cryptanalysis of Caesar Cipher

Sample of Compressed Text

~+Wμ"- Ω-0)≤4{∞‡, ë~Ω%ràu·-í ◇-z-
Ú≠2Ò#Åæð æ«q7,Ωn·®3NÔÚ ℄z'Y-f∞Í[±Ũ_ èΩ,<NO¬±«˘xă Åä£èü3Å
x}ö§k°Â
_yÍ ^ΔÉ] ,κ J/°iTê&1 'c<uΩ-
ÄD(G WÄC~y_iöÄW PÔ1«ÎÜ†ç],κ;˘Î^üÑπ˘≈˘L˘9OgflO˘&℄≤ ¬≤ ØÔ§":
˘℄!SGqèvo^ ú\,S>h<-*6ø‡%x'˘|fiÓ#≈˘my%˘≥ñP<,fi Áj ÅÔ¿˘Zù-
Ω˘Ö˘6℄ÿ{% „ΩÊó ,i π÷Áî°ú02çSÿ'O-
2Äflßi /@^"ΠK°ªP℄π,úé^'3Σ˘ö˘ÔZÎ˘Y¬ÿΩæY> Ω+eô/ ·<K£¿*÷˘"≤û˘
B ZøK~Qßÿüf, !òflîzssS/]>ÈQ ü

Figure 3.4 Sample of Compressed Text

Monoalphabetic Cipher

- Permutation
 - Of a finite set of elements S is an ordered sequence of all the elements of S , with each element appearing exactly once
- If the “cipher” line can be any permutation of the 26 alphabetic characters, then there are $26!$ or greater than 4×10^{26} possible keys
 - This is 10 orders of magnitude greater than the key space for DES
 - Approach is referred to as a *monoalphabetic substitution cipher* because a single cipher alphabet is used per message

Monoalphabetic Cipher Security

- now have a total of $26! = 4 \times 10^{26}$ keys
- with so many keys, might think is secure
- but would be **!!!WRONG!!!**
- problem is language characteristics

Language Redundancy and Cryptanalysis

- human languages are **redundant**
- eg "th lrd s m shphrd shll nt wnt"
- letters are not equally commonly 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
- have tables of single, double & triple letter frequencies for various languages

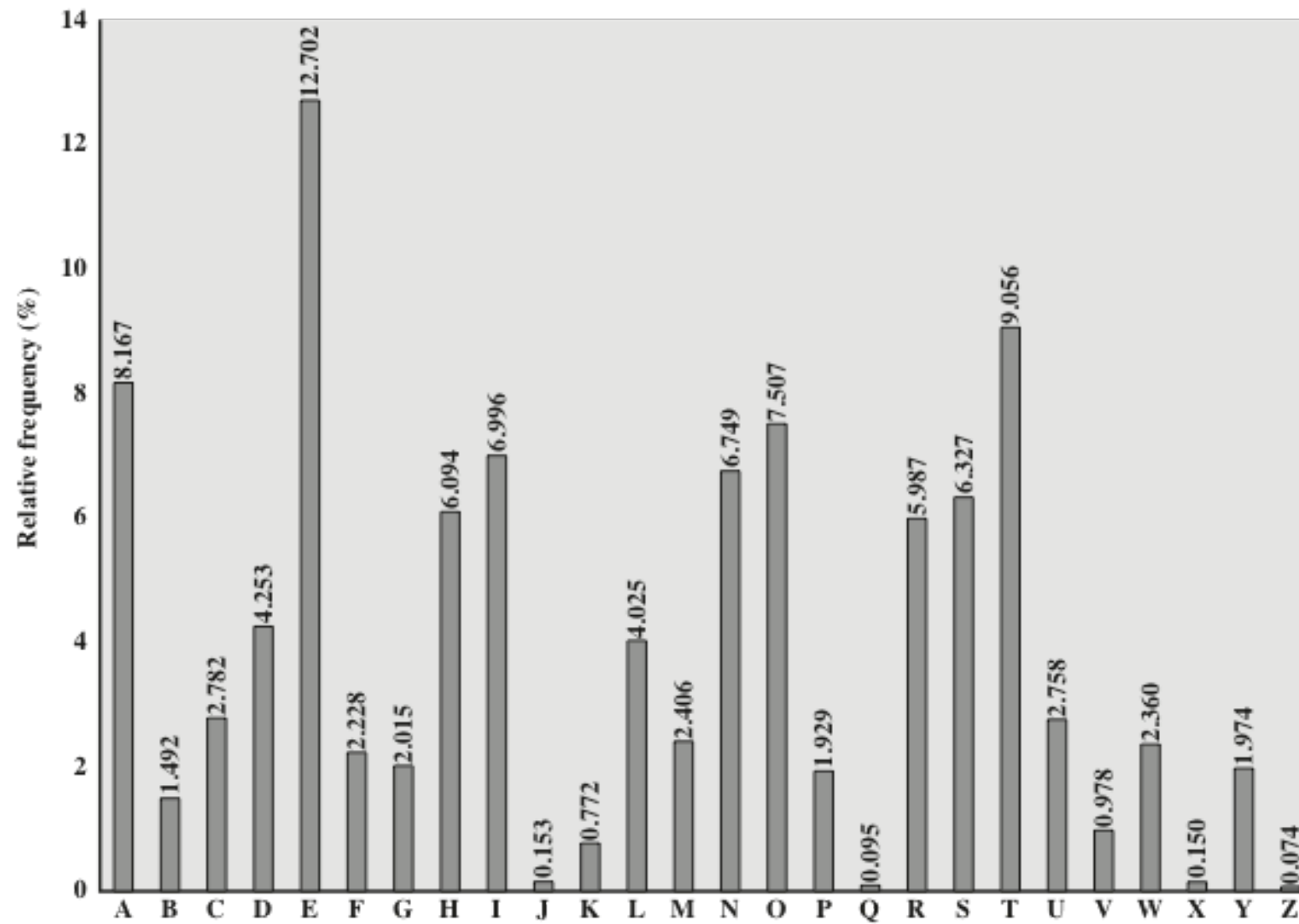
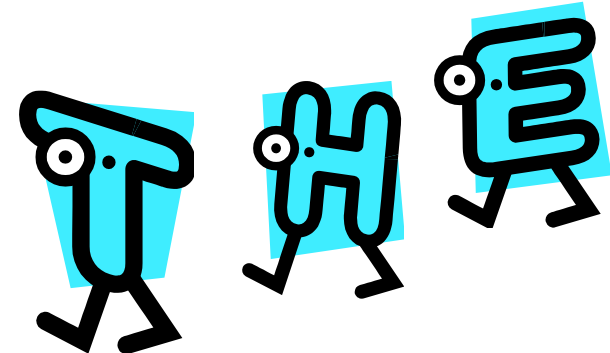
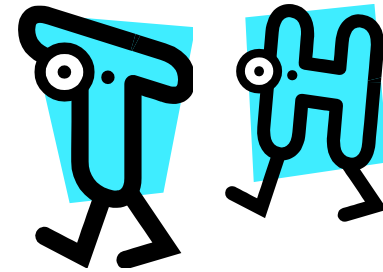


Figure 3.5 Relative Frequency of Letters in English Text

Monoalphabetic Ciphers

- Easy to break because they reflect the frequency data of the original alphabet
- Countermeasure is to provide multiple substitutes (homophones) for a single letter
- Digram
 - Two-letter combination
 - Most common is *th*
- Trigram
 - Three-letter combination
 - Most frequent is *the*



Example Cryptanalysis

- given ciphertext:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAI Z
VUEPHZHMDZSHZOWSFPAPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFPUPZHMDJUDTMOHMQ

- count relative letter frequencies (see text)
- guess P & Z are e and t
- guess ZW is th and hence ZWP is 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

Playfair Cipher



- Best-known multiple-letter encryption cipher
- Treats digrams in the plaintext as single units and translates these units into ciphertext digrams
- Based on the use of a 5 x 5 matrix of letters constructed using a keyword
- Invented by British scientist Sir Charles Wheatstone in 1854
- Used as the standard field system by the British Army in World War I and the U.S. Army and other Allied forces during World War II

Playfair Key Matrix

- Fill in letters of keyword (minus duplicates) from left to right and from top to bottom, then fill in the remainder of the matrix with the remaining letters in alphabetic order
- Using the keyword 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'
 - eg. "balloon" encrypts as "ba lx lo on"
 2. if both letters fall in the same row, replace each with letter to right (wrapping back to start from end)
 - eg. "ar" encrypts as "RM"
 3. if both letters fall in the same column, replace each with the letter below it (again wrapping to top from bottom)
 - eg. "mu" encrypts to "CM"
 4. otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair
 - eg. "hs" encrypts to "BP", and "ea" to "IM" or "JM" (as desired)

Security of Playfair Cipher

- security much improved over monoalphabetic
- since have $26 \times 26 = 676$ digrams
- would need a 676 entry frequency table to analyse (verses 26 for a monoalphabetic)
- and correspondingly more ciphertext
- was widely used for many years
 - eg. by US & British military in WW1
- it **can** be broken, given a few hundred letters
- since still has much of plaintext structure

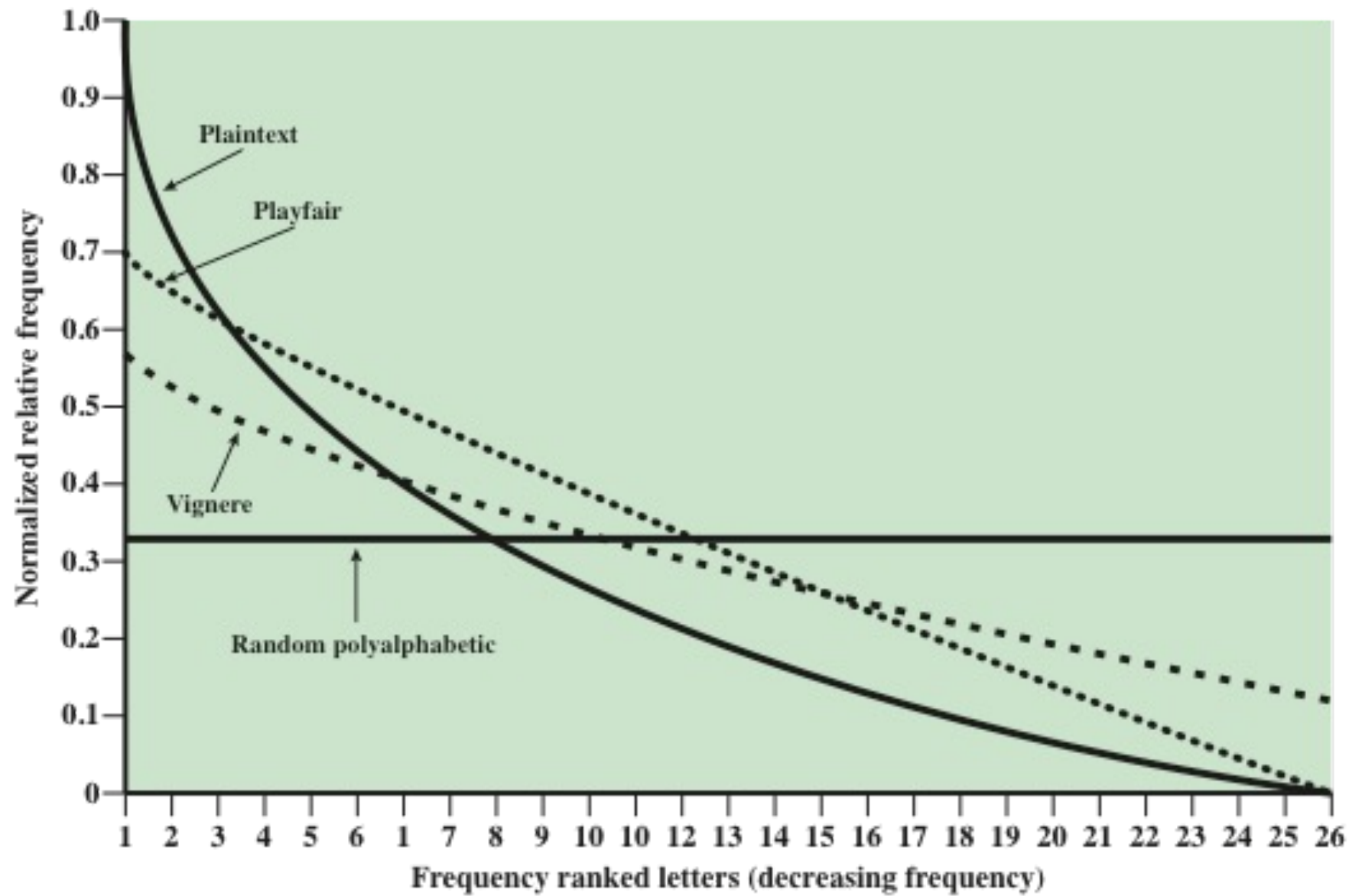
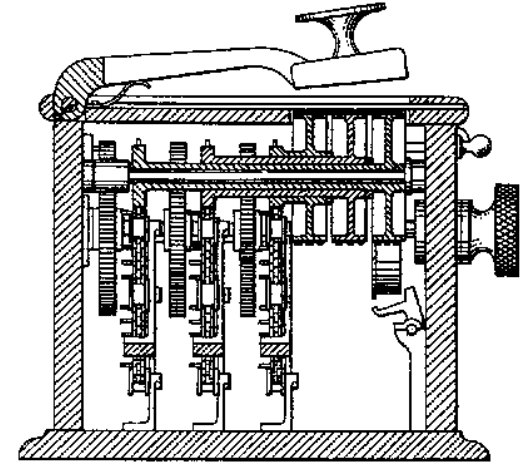


Figure 3.6 Relative Frequency of Occurrence of Letters

Hill Cipher

- Developed by the mathematician Lester Hill in 1929
- Strength is that it completely hides single-letter frequencies
 - The use of a larger matrix hides more frequency information
 - A 3 x 3 Hill cipher hides not only single-letter but also two-letter frequency information
- Strong against a ciphertext-only attack but easily broken with a known plaintext attack



Polyalphabetic Ciphers

- Polyalphabetic substitution cipher
 - Improves on the simple monoalphabetic technique by using different monoalphabetic substitutions as one proceeds through the plaintext message

All these techniques have the following features in common:

- A set of related monoalphabetic substitution rules is used
- A key determines which particular rule is chosen for a given transformation

Vigenère Cipher

- Best known and one of the simplest polyalphabetic substitution ciphers
- In this scheme the set of related monoalphabetic substitution rules consists of the 26 Caesar ciphers with shifts of 0 through 25
- Each cipher is denoted by a key letter which is the ciphertext letter that substitutes for the plaintext letter a



Example of Vigenère Cipher

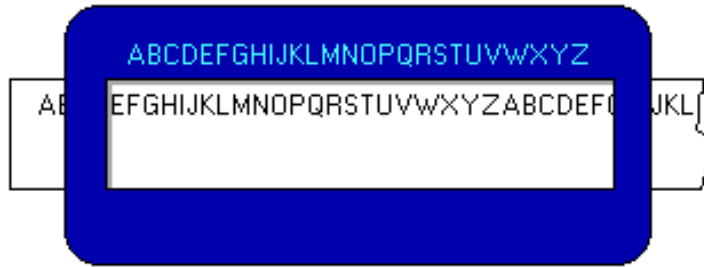
- To encrypt a message, a key is needed that is as long as the message
- Usually, the key is a repeating keyword
- For example, if the keyword is *deceptive*, the message “we are discovered save yourself” is encrypted as:

key: deceptivedeceptivedeceptive

plaintext: wearediscoveredsaveyourself

ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ

Aids



- simple aids can assist with en/decryption
- a **Saint-Cyr Slide** is a simple manual aid
 - a slide with repeated alphabet
 - line up plaintext 'A' with key letter, eg 'C'
 - then read off any mapping for key letter
- can bend round into a **cipher disk**
- or expand into a **Vigenère Tableau**

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

Security of Vigenère Ciphers

- have multiple ciphertext letters for each plaintext letter
- hence letter frequencies are obscured
- but not totally lost
- start with letter frequencies
 - see if look monoalphabetic or not
- if not, then need to determine number of alphabets, since then can attach each

Kasiski Method

- method developed by Babbage / Kasiski
- repetitions in ciphertext give clues to period
- so find same plaintext an exact period apart
- which results in the same ciphertext
- of course, could also be random fluke
- eg repeated “VTW” in previous example
- suggests size of 3 or 9
- then attack each monoalphabetic cipher individually using same techniques as before

Vigenère Autokey System

- A keyword is concatenated with the plaintext itself to provide a running key
- Example:
key: deceptivewearediscoveredsav
plaintext: wearediscoveredsaveyourself
ciphertext: ZICVTWQNGKZEIIGASXSTSLVWLA
- Even this scheme is vulnerable to cryptanalysis
 - Because the key and the plaintext share the same frequency distribution of letters, a statistical technique can be applied

Vernam Cipher

- ultimate defense is to use a key as long as the plaintext
- with no statistical relationship to it
- invented by AT&T engineer Gilbert Vernam in 1918
- originally proposed using a very long but eventually repeating key

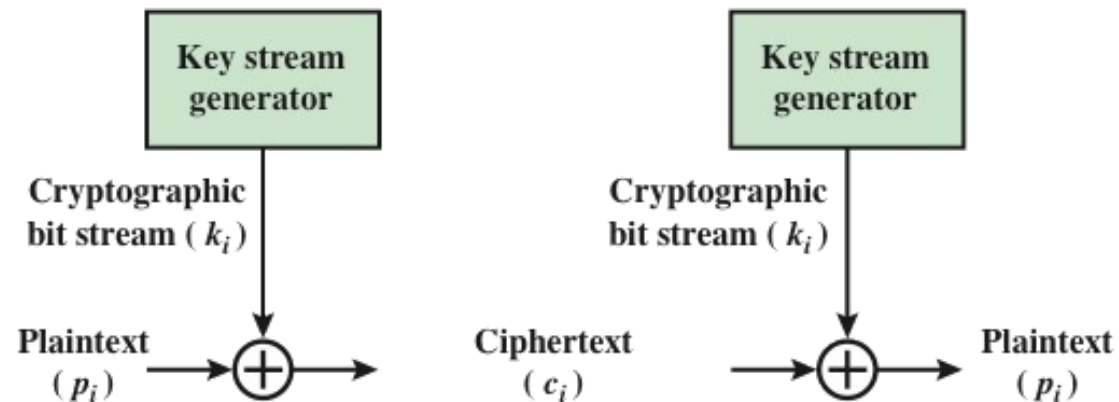


Figure 3.7 Vernam Cipher

One-Time Pad



- Improvement to Vernam cipher proposed by an Army Signal Corp officer, Joseph Mauborgne
- Use a random key that is as long as the message so that the key need not be repeated
- Key is used to encrypt and decrypt a single message and then is discarded
- Each new message requires a new key of the same length as the new message
- Scheme is unbreakable
 - Produces random output that bears no statistical relationship to the plaintext
 - Because the ciphertext contains no information whatsoever about the plaintext, there is simply no way to break the code

Difficulties

- The one-time pad offers complete security but, in practice, has two fundamental difficulties:
 - There is the practical problem of making large quantities of random keys
 - Any heavily used system might require millions of random characters on a regular basis
 - Mammoth key distribution problem
 - For every message to be sent, a key of equal length is needed by both sender and receiver
- Because of these difficulties, the one-time pad is of limited utility
 - Useful primarily for low-bandwidth channels requiring very high security
- The one-time pad is the only cryptosystem that exhibits *perfect secrecy* (see Appendix F)

Transposition Ciphers

- now consider classical **transposition** or **permutation** ciphers
- these hide the message by rearranging the letter order
- without altering the actual letters used
- can recognise these since have the same frequency distribution as the original text

Rail Fence Cipher

- Simplest transposition cipher
- Plaintext is written down as a sequence of diagonals and then read off as a sequence of rows
- To encipher the message “meet me after the toga party” with a rail fence of depth 2, we would write:

m e m a t r h t g p r y

e t e f e t e o a a t

Encrypted message is:

MEMATRHTGPRYETEFETEOAAT



Row Transposition Cipher

- Is a more complex transposition
- Write the message in a rectangle, row by row, and read the message off, column by column, but permute the order of the columns
 - The order of the columns then becomes the key to the algorithm

Key: 4 3 1 2 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: T T N A A P T M T S U O A O D W C O I X K N L Y P E T Z

Product Ciphers

- ciphers using substitutions or transpositions are not secure because of language characteristics
- hence consider using several ciphers in succession to make harder, but:
 - two substitutions make a more complex substitution
 - two transpositions make more complex transposition
 - but a substitution followed by a transposition makes a new much harder cipher
- this is bridge from classical to modern ciphers

Rotor Machines

- before modern ciphers, rotor machines were most common complex ciphers in use
- widely used in WW2
 - German Enigma, Allied Hagelin, Japanese Purple
- implemented a very complex, varying substitution cipher
- used a series of cylinders, each giving one substitution, which rotated and changed after each letter was encrypted
- with 3 cylinders have $26^3=17576$ alphabets

Hagelin Rotor Machine



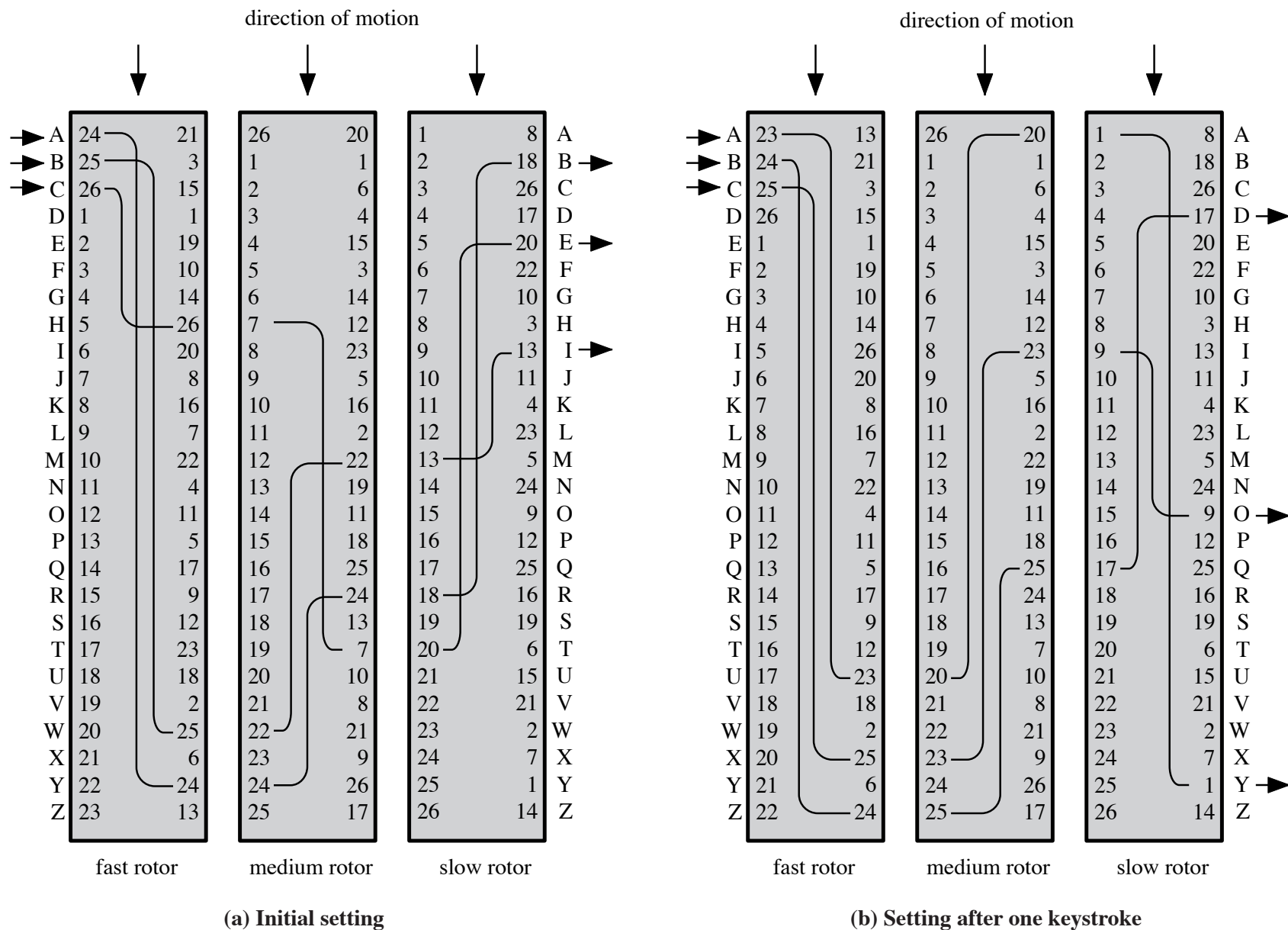


Figure 3.8 Three-Rotor Machine With Wiring Represented by Numbered Contacts

Steganography

- an alternative to encryption
- hides existence of message
 - using only a subset of letters/words in a longer message marked in some way
 - using invisible ink
 - hiding in LSB in graphic image or sound file
- has drawbacks
 - high overhead to hide relatively few info bits

Steganography

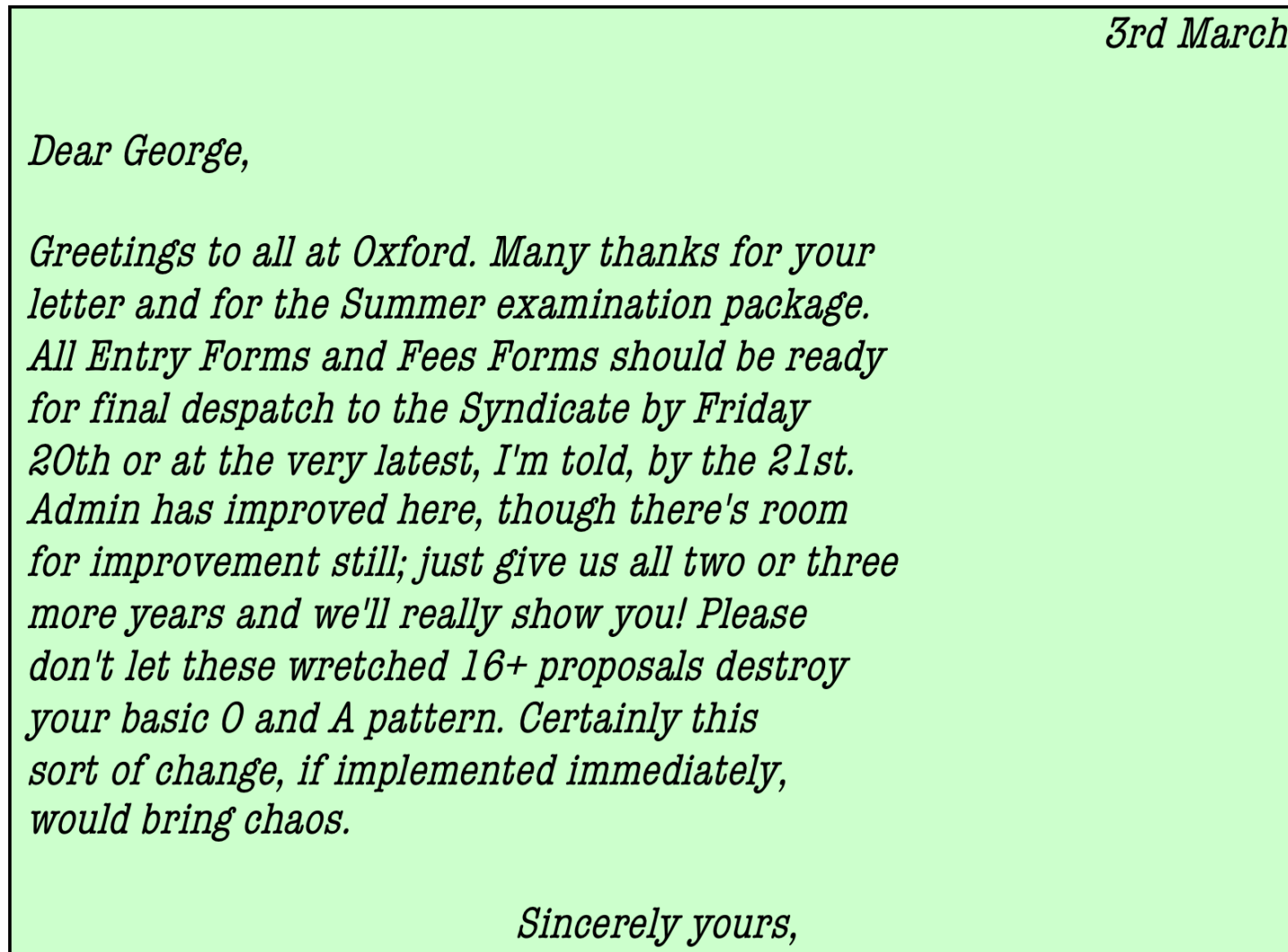


Figure 2.9 A Puzzle for Inspector Morse
(from *The Silent World of Nicholas Quinn*, by Colin Dexter)

Other Steganography Techniques




- **Character marking**
 - Selected letters of printed or typewritten text are over-written in pencil
 - The marks are ordinarily not visible unless the paper is held at an angle to bright light
- **Invisible ink**
 - A number of substances can be used for writing but leave no visible trace until heat or some chemical is applied to the paper
- **Pin punctures**
 - Small pin punctures on selected letters are ordinarily not visible unless the paper is held up in front of a light
- **Typewriter correction ribbon**
 - Used between lines typed with a black ribbon, the results of typing with the correction tape are visible only under a strong light

Steganography vs. Encryption

- **Steganography has a number of drawbacks when compared to encryption**
 - It requires a lot of overhead to hide a relatively few bits of information
 - Once the system is discovered, it becomes virtually worthless

- **The advantage of steganography**
 - It can be employed by parties who have something to lose should the fact of their secret communication (not necessarily the content) be discovered
 - Encryption flags traffic as important or secret or may identify the sender or receiver as someone with something to hide

Summary

- Present an overview of the main concepts of symmetric cryptography
 - Explain the difference between cryptanalysis and brute-force attack
 - Understand the operation of a monoalphabetic substitution cipher
- 
- Understand the operation of a polyalphabetic cipher
 - Present an overview of the Hill cipher