Cryptography: Overview and Classical Ciphers

Alex Corner

Sheffield Hallam University

Cryptography

- Cryptography: From Greek word ‘kryptos’ meaning ‘hidden’, and ‘graph’ meaning ‘writing’.

Cryptography

- Cryptography: From Greek word ‘kryptos’ meaning ‘hidden’, and ‘graph’ meaning ‘writing’.
- Originally the study of techniques to disguise data.

Cryptography

- Cryptography: From Greek word ‘kryptos’ meaning ‘hidden’, and ‘graph’ meaning ‘writing’.
- Originally the study of techniques to disguise data.
- Now it is more widely used to describe the subject area associated with hiding information.

Cryptography

- Cryptography: From Greek word ‘kryptos’ meaning ‘hidden’, and ‘graph’ meaning ‘writing’.
- Originally the study of techniques to disguise data.
- Now it is more widely used to describe the subject area associated with hiding information.
- **Cryptanalysis**: Breaking the disguise.

Cryptography

- Cryptography: From Greek word ‘kryptos’ meaning ‘hidden’, and ‘graph’ meaning ‘writing’.
- Originally the study of techniques to disguise data.
- Now it is more widely used to describe the subject area associated with hiding information.
- **Cryptanalysis**: Breaking the disguise.
- **Cryptology**: Study of secret codes and ciphers.

Why do we need cryptography?

- Originally the preserve of the military and diplomatic services, now we all use it.

Why do we need cryptography?

- Originally the preserve of the military and diplomatic services, now we all use it.
- ATM cards, online banking, electronic commerce, computer passwords, email.

Why do we need cryptography?

- Originally the preserve of the military and diplomatic services, now we all use it.
- ATM cards, online banking, electronic commerce, computer passwords, email.
- Mobile phones, DVD players, pay-TV decoders, game consoles, car keys, burglar alarms.

Intended content

- Classical ciphers

Intended content

- Classical ciphers
- Symmetric (secret) keys

Intended content

- Classical ciphers
- Symmetric (secret) keys
- Public keys

Intended content

- Classical ciphers
- Symmetric (secret) keys
- Public keys
- Modern ciphers

Some terminology

- **Plaintext**: The original message to be sent.

Some terminology

- **Plaintext**: The original message to be sent.
- **Encryption**: The process of disguising the plaintext.

Some terminology

- **Plaintext**: The original message to be sent.
- **Encryption**: The process of disguising the plaintext.
- **Cipher**: The method of encryption

Some terminology

- **Plaintext**: The original message to be sent.
- **Encryption**: The process of disguising the plaintext.
- **Cipher**: The method of encryption
- **Encrypt** (or encipher): To disguise the text.

Some terminology

- **Plaintext**: The original message to be sent.
- **Encryption**: The process of disguising the plaintext.
- **Cipher**: The method of encryption
- **Encrypt** (or encipher): To disguise the text.
- **Ciphertext**: The disguised text.

Some terminology

- **Plaintext**: The original message to be sent.
- **Encryption**: The process of disguising the plaintext.
- **Cipher**: The method of encryption
- **Encrypt** (or encipher): To disguise the text.
- **Ciphertext**: The disguised text.
- **Decrypt** (or decipher): To remove the disguise and find the plaintext.

Some terminology

- **Plaintext**: The original message to be sent.
- **Encryption**: The process of disguising the plaintext.
- **Cipher**: The method of encryption
- **Encrypt** (or encipher): To disguise the text.
- **Ciphertext**: The disguised text.
- **Decrypt** (or decipher): To remove the disguise and find the plaintext.
- **Key**: This keeps the message secret. Used to encrypt or decrypt a message.

Some terminology

- **Plaintext**: The original message to be sent.
- **Encryption**: The process of disguising the plaintext.
- **Cipher**: The method of encryption
- **Encrypt** (or encipher): To disguise the text.
- **Ciphertext**: The disguised text.
- **Decrypt** (or decipher): To remove the disguise and find the plaintext.
- **Key**: This keeps the message secret. Used to encrypt or decrypt a message.
- **Brute force attack**: Try every possible key.

Kerchoff's Principle

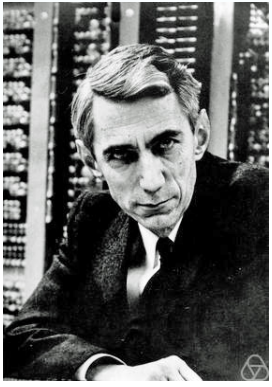
"It must not be required to be secret, and it must be able to fall into the hands of the enemy without inconvenience."

— **Auguste Kerckhoffs**, 1883

Shannon's Maxim

"The enemy knows the system."

— **Claude Shannon**, 1916-2001



Cryptography: an overview

Cryptography: an overview

Classical ciphers:

- Using methods of transposition and substitution.
- Substitution ciphers
- Caesar Cipher
- Vigenère cipher

Cryptography: an overview

Classical ciphers:

- Using methods of transposition and substitution.
- Substitution ciphers
- Caesar Cipher
- Vigenère cipher

Modern ciphers:

- Symmetric key and Public key
- Symmetric Key (Secret Key): DES and AES (Rijndael)
- Public Key: RSA named after its inventors Rivest, Shamir and Adleman.

Classical ciphers: Transposition

- Letters in a message are rearranged like an anagram.

Classical ciphers: Transposition

- Letters in a message are rearranged like an anagram.
- First used 2400 years ago in Sparta: scytale ciphers.

Classical ciphers: Transposition

- Letters in a message are rearranged like an anagram.
- First used 2400 years ago in Sparta: scytale ciphers.
- 'Scramble the letters of this message'

Classical ciphers: Transposition

- Letters in a message are rearranged like an anagram.
- First used 2400 years ago in Sparta: scytale ciphers.
- 'Scramble the letters of this message'
- There are $123,921,541,253,404,749,254,400,000$ ways!

Classical ciphers: Transposition

- Letters in a message are rearranged like an anagram.
- First used 2400 years ago in Sparta: scytale ciphers.
- 'Scramble the letters of this message'
- There are $123,921,541,253,404,749,254,400,000$ ways!
- 'aabceeeeeefghhillmmorrssssstttt'

Classical ciphers: Transposition

- Letters in a message are rearranged like an anagram.
- First used 2400 years ago in Sparta: scytale ciphers.
- 'Scramble the letters of this message'
- There are $123,921,541,253,404,749,254,400,000$ ways!
- 'aabceeeeeefghhillmmorrssssstttt'
- Disadvantage: how do you unscramble them?

Classical ciphers: Transposition

- Letters in a message are rearranged like an anagram.
- First used 2400 years ago in Sparta: scytale ciphers.
- 'Scramble the letters of this message'
- There are $123,921,541,253,404,749,254,400,000$ ways!
- 'aabceeeeeefghhillmmorrssssstttt'
- Disadvantage: how do you unscramble them?
- 'Fiberglass thermostats melt cheese'

Classical ciphers: Rail Fence

m		t		e		a		i		a	
	a		h		m		t		c		l

Classical ciphers: Rail Fence

m		t		e		a		i		a	
	a		h		m		t		c		l

- We generate the ciphertext by reading the top rail first, followed by the second rail.

Classical ciphers: Rail Fence

m		t		e		a		i		a	
	a		h		m		t		c		l

- We generate the ciphertext by reading the top rail first, followed by the second rail.
- For the plaintext '`mathematical`', this would generate the ciphertext '`MTEAIAAHMTCL`'.

Classical ciphers: Rail Fence

- We can also increase the number of rails in the cipher. For example, with three rails, the word 'mathematical' becomes 'MEIAHMTCLTAA'.

m				e				i			
	a		h		m		t		c		l
		t				a				a	

Classical ciphers: Rail Fence

- We can also increase the number of rails in the cipher. For example, with three rails, the word '`mathematical`' becomes '`MEIAHMTCLTAA`'.

m				e				i			
	a		h		m		t		c		l
		t				a				a	

- Rail ciphers are quite similar to another type of transposition cipher, called a Scytale cipher.

Classical ciphers: Substitution

- Each letter is disguised as another.

Classical ciphers: Substitution

- Each letter is disguised as another.
- The first documented use of a substitution cipher is from the 1st century BC.

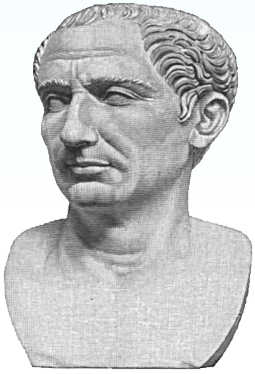
Classical ciphers: Substitution

- Each letter is disguised as another.
- The first documented use of a substitution cipher is from the 1st century BC.
- It's called Caesar's Cipher, attributed to Julius Caesar (100BC – 44 BC).

Classical ciphers: Substitution

- Each letter is disguised as another.
- The first documented use of a substitution cipher is from the 1st century BC.
- It's called Caesar's Cipher, attributed to Julius Caesar (100BC – 44 BC).
- During the Gallic Wars secrecy was vital, so Caesar wanted to disguise written messages and devised a Shift Cipher.

Classical ciphers: Caesar's Cipher



- The plain alphabet is written above the cipher alphabet but shifted by a number of places.
- The original Caesar cipher was shifted by 3 places, as shown below.

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext:

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `L`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQ`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQY`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYD`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDV`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVL`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLR`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse:

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `s`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `se`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `sec`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `secu`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `secur`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `securi`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `securit`

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `security`.

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `security`.
- This is a **monoalphabetic** cipher, since it uses *one* alphabet.

Classical ciphers: Caesar's Cipher

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Using a shift of 3, encrypt the plaintext `invasion`.
- We read off each of the letters in turn from the table to create the ciphertext: `LQYDVLRQ`.
- If we receive the ciphertext `VHFXULWB`, then we decrypt by using the table in reverse: `security`.
- This is a **monoalphabetic** cipher, since it uses *one* alphabet.
- There are 25 possible keys for this cipher (different shift values).

Caesar's Cipher: Modular Arithmetic

- We can see the Caesar cipher as a simple application of modular arithmetic.

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

Caesar's Cipher: Modular Arithmetic

- We can see the Caesar cipher as a simple application of modular arithmetic.

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- The cipher relies on a shift number k, which acts as the key.

Caesar's Cipher: Modular Arithmetic

- We can see the Caesar cipher as a simple application of modular arithmetic.

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- The cipher relies on a shift number k, which acts as the key.
- This shift k is an element of the set $\{1, 2, 3, 4, \ldots, 25\}$.

Caesar's Cipher: Modular Arithmetic

- We can see the Caesar cipher as a simple application of modular arithmetic.

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- The cipher relies on a shift number k, which acts as the key.
- This shift k is an element of the set $\{1, 2, 3, 4, \ldots, 25\}$.
- If α represents the corresponding number from the table above of a single letter in our message, then the encryption function is given by

$$E(\alpha, k) = \alpha + k \mod 26.$$

Caesar's Cipher: Modular Arithmetic

- We can see the Caesar cipher as a simple application of modular arithmetic.

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- The cipher relies on a shift number k, which acts as the key.
- This shift k is an element of the set $\{1, 2, 3, 4, \ldots, 25\}$.
- If α represents the corresponding number from the table above of a single letter in our message, then the encryption function is given by

$$E(\alpha, k) = \alpha + k \mod 26.$$

- The use of mod 26 is what allows the shift to 'wrap around' when it gets to z.

Caesar's Cipher: Modular Arithmetic

- We can see the Caesar cipher as a simple application of modular arithmetic.

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- The cipher relies on a shift number k, which acts as the key.
- This shift k is an element of the set $\{1, 2, 3, 4, \ldots, 25\}$.
- If α represents the corresponding number from the table above of a single letter in our message, then the encryption function is given by

$$E(\alpha, k) = \alpha + k \mod 26.$$

- The use of mod 26 is what allows the shift to 'wrap around' when it gets to Z.
- Julius Caesar's version of the cipher had a shift value of $k = 3$. Obviously, a trivial shift value of $k = 0$ won't be very effective!

Caesar's Cipher: Modular Arithmetic

- We can see the Caesar cipher as a simple application of modular arithmetic.

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- The cipher relies on a shift number k, which acts as the key.
- This shift k is an element of the set $\{1, 2, 3, 4, \ldots, 25\}$.
- If α represents the corresponding number from the table above of a single letter in our message, then the encryption function is given by

$$E(\alpha, k) = \alpha + k \mod 26.$$

- The use of mod 26 is what allows the shift to 'wrap around' when it gets to Z.
- Julius Caesar's version of the cipher had a shift value of $k = 3$. Obviously, a trivial shift value of $k = 0$ won't be very effective!
- The decryption key works in much the same way but using subtraction.

Classical ciphers: Substitution Ciphers

- The security of a Caesar cipher is quite limited - there are only 25 keys, so a brute force attack would quickly reveal the message.

Classical ciphers: Substitution Ciphers

- The security of a Caesar cipher is quite limited - there are only 25 keys, so a brute force attack would quickly reveal the message.
- We could use **general substitution**: a random unused letter is chosen for each letter of the plaintext.

Classical ciphers: Substitution Ciphers

- The security of a Caesar cipher is quite limited - there are only 25 keys, so a brute force attack would quickly reveal the message.
- We could use **general substitution**: a random unused letter is chosen for each letter of the plaintext.
- This method produces $26! \approx 4.03 \times 10^{26}$ keys.

Classical ciphers: Substitution Ciphers

- The security of a Caesar cipher is quite limited - there are only 25 keys, so a brute force attack would quickly reveal the message.
- We could use **general substitution**: a random unused letter is chosen for each letter of the plaintext.
- This method produces $26! \approx 4.03 \times 10^{26}$ keys. But a random selection would be far too difficult to remember in practice, so we can use a **key phrase**.

Classical ciphers: Substitution Ciphers

- The security of a Caesar cipher is quite limited - there are only 25 keys, so a brute force attack would quickly reveal the message.
- We could use **general substitution**: a random unused letter is chosen for each letter of the plaintext.
- This method produces $26! \approx 4.03 \times 10^{26}$ keys. But a random selection would be far too difficult to remember in practice, so we can use a **key phrase**.
- E.g., using `ALEXANDER` as the key we first remove any repeated letters: `ALEXNDR`.

Classical ciphers: Substitution Ciphers

- The security of a Caesar cipher is quite limited - there are only 25 keys, so a brute force attack would quickly reveal the message.
- We could use **general substitution**: a random unused letter is chosen for each letter of the plaintext.
- This method produces $26! \approx 4.03 \times 10^{26}$ keys. But a random selection would be far too difficult to remember in practice, so we can use a **key phrase**.
- E.g., using `ALEXANDER` as the key we first remove any repeated letters: `ALEXNDR`. And then fill in the rest of the alphabet in order following `R`.

Classical ciphers: Substitution Ciphers

- The security of a Caesar cipher is quite limited - there are only 25 keys, so a brute force attack would quickly reveal the message.
- We could use **general substitution**: a random unused letter is chosen for each letter of the plaintext.
- This method produces $26! \approx 4.03 \times 10^{26}$ keys. But a random selection would be far too difficult to remember in practice, so we can use a **key phrase**.
- E.g., using `ALEXANDER` as the key we first remove any repeated letters: `ALEXNDR`. And then fill in the rest of the alphabet in order following `R`.
- This reduces the number of keys, but makes the cipher more practical to implement. It can also lead to letters being insecurely encrypted, e.g., 'a' as 'A'.

plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m
ciphertext	A	L	E	X	N	D	R	S	T	U	V	W	Y
plaintext	n	o	p	q	r	s	t	u	v	w	x	y	z
ciphertext	Z	B	C	F	G	H	I	J	K	M	O	P	Q

The Dancing Men

- A clever technique to further disguise a substitution cipher is to not substitute plaintext letters for ciphertext letters, but to use symbols instead.

The Dancing Men

- A clever technique to further disguise a substitution cipher is to not substitute plaintext letters for ciphertext letters, but to use symbols instead.
- A Sherlock Holmes story features this prominently as a plot point.

The Dancing Men

- A clever technique to further disguise a substitution cipher is to not substitute plaintext letters for ciphertext letters, but to use symbols instead.
- A Sherlock Holmes story features this prominently as a plot point.

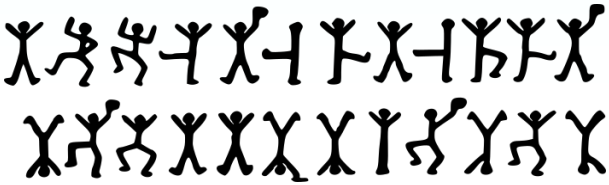
The Dancing Men

- A clever technique to further disguise a substitution cipher is to not substitute plaintext letters for ciphertext letters, but to use symbols instead.
- A Sherlock Holmes story features this prominently as a plot point.

- The men with the flags represent the ends of words and this deciphers to: `Elsie prepare to meet thy god`.

The Dancing Men

- A clever technique to further disguise a substitution cipher is to not substitute plaintext letters for ciphertext letters, but to use symbols instead.
- A Sherlock Holmes story features this prominently as a plot point.



- The men with the flags represent the ends of words and this deciphers to: `Elsie prepare to meet thy god`.
- Even using symbols doesn't make this method secure, as Mary Queen of Scots found out as part of the Babington Plot. Similarly, Anne Lister's diaries were eventually decoded after being discovered by a distant relative.

Frequency Analysis

- For hundreds of years the general substitution cipher was thought to be secure because of the huge number of possible keys, so alternative ciphers were not developed.

Frequency Analysis

- For hundreds of years the general substitution cipher was thought to be secure because of the huge number of possible keys, so alternative ciphers were not developed.
- However, a shortcut to cracking the cipher is to exploit the variation in the frequency of letters.

Frequency Analysis

- For hundreds of years the general substitution cipher was thought to be secure because of the huge number of possible keys, so alternative ciphers were not developed.
- However, a shortcut to cracking the cipher is to exploit the variation in the frequency of letters.
- The earliest record of using this technique to break a cipher was in the 9th century.

Frequency Analysis

- For hundreds of years the general substitution cipher was thought to be secure because of the huge number of possible keys, so alternative ciphers were not developed.
- However, a shortcut to cracking the cipher is to exploit the variation in the frequency of letters.
- The earliest record of using this technique to break a cipher was in the 9th century.
- It is also much easier if the spaces are kept between words.

Frequency Analysis: Common Letters

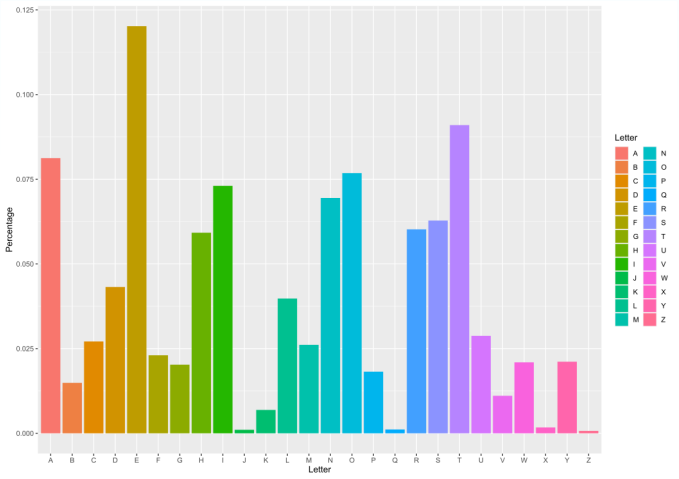
In the English language, the most commonly occurring letter is e.

Frequency Analysis: Common Letters

In the English language, the most commonly occurring letter is `e`. It is then followed by `t` and `a` as the second and third most common.

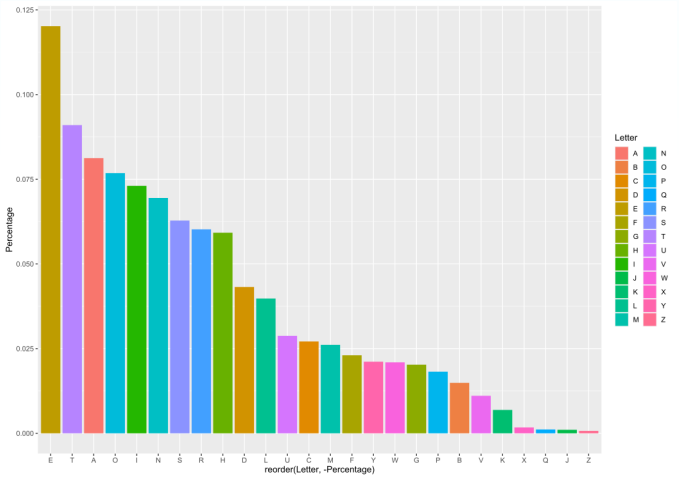
Frequency Analysis: Common Letters

In the English language, the most commonly occurring letter is `e`. It is then followed by `t` and `a` as the second and third most common.



Frequency Analysis: Common Letters

In the English language, the most commonly occurring letter is `e`. It is then followed by `t` and `a` as the second and third most common.



Frequency Analysis: Language Patterns

- Other patterns in language can often be detected.

Frequency Analysis: Language Patterns

- Other patterns in language can often be detected.
- Common pairs of letters are: `ss`, `ee`, `tt`, `ff`.

Frequency Analysis: Language Patterns

- Other patterns in language can often be detected.
- Common pairs of letters are: `ss`, `ee`, `tt`, `ff`.
- Common two-letter combinations (**digraphs**) are: `th`, `er`, `on`, `an`, `re`, `he`, `in`.

Frequency Analysis: Language Patterns

- Other patterns in language can often be detected.
- Common pairs of letters are: `ss`, `ee`, `tt`, `ff`.
- Common two-letter combinations (**digraphs**) are: `th`, `er`, `on`, `an`, `re`, `he`, `in`.
- Common three-letter combinations (**trigraphs**) are: `the`, `and`, `tha`, `ent`, `ion`.

Frequency Analysis: Language Patterns

- Other patterns in language can often be detected.
- Common pairs of letters are: `ss`, `ee`, `tt`, `ff`.
- Common two-letter combinations (**digraphs**) are: `th`, `er`, `on`, `an`, `re`, `he`, `in`.
- Common three-letter combinations (**trigraphs**) are: `the`, `and`, `tha`, `ent`, `ion`.
- If spaces are left in it becomes easier to identify the small words.

Frequency Analysis: Language Patterns

- Identify common small words first.

Frequency Analysis: Language Patterns

- Identify common small words first.
- One-letter words: `I`, `a`.

Frequency Analysis: Language Patterns

- Identify common small words first.
- One-letter words: `I`, `a`.
- Two-letter words: `of`, `to`, `in`, `it`, `is`, `be`, `as`, `at`, `so`.

Frequency Analysis: Language Patterns

- Identify common small words first.
- One-letter words: `I`, `a`.
- Two-letter words: `of`, `to`, `in`, `it`, `is`, `be`, `as`, `at`, `so`.
- Three-letter words: `the`, `and`, `for`, `are`, `but`, `not`, `you`.

Frequency Analysis: Language Patterns

- Identify common small words first.
- One-letter words: `I`, `a`.
- Two-letter words: `of`, `to`, `in`, `it`, `is`, `be`, `as`, `at`, `so`.
- Three-letter words: `the`, `and`, `for`, `are`, `but`, `not`, `you`.
- Four-letter words: `that`, `with`, `have`, `this`, `will`, `your`.

Frequency Analysis: Language Patterns

- Identify common small words first.
- One-letter words: `I`, `a`.
- Two-letter words: `of`, `to`, `in`, `it`, `is`, `be`, `as`, `at`, `so`.
- Three-letter words: `the`, `and`, `for`, `are`, `but`, `not`, `you`.
- Four-letter words: `that`, `with`, `have`, `this`, `will`, `your`.
- Initial letters: `t`, `o`, `a`, `w`, `b`, `c`, `d`.

Frequency Analysis: Language Patterns

- Identify common small words first.
- One-letter words: `I`, `a`.
- Two-letter words: `of`, `to`, `in`, `it`, `is`, `be`, `as`, `at`, `so`.
- Three-letter words: `the`, `and`, `for`, `are`, `but`, `not`, `you`.
- Four-letter words: `that`, `with`, `have`, `this`, `will`, `your`.
- Initial letters: `t`, `o`, `a`, `w`, `b`, `c`, `d`.
- Final letters: `e`, `s`, `t`, `d`, `d`, `n`, `r`, `y`.

Frequency Analysis: An Example

- Below are 655 characters of ciphertext encrypted using a substitution cipher. Spaces are left in to make it slightly easier to decrypt.

```
TKNKQBX GCUHT CJ QBH ICTPQ KY QBH ANHDQ FDJPDP LNDCNCHP VCQB SJRGH BHJNX VBK
VDP D YDNIHN DJT DSJQ HI VBK VDP QBH YDNIHNP VCYH QBHCN BKSPH VDP PIDGG YKN
QBH GSIOHN QK OSCGT CQ BDT QK OH RDNNCHT OX VDAKJ IDJX ICGHP QBHNH VHNH YKSN
VDGGP D YGKKN DJT D NKKY VBCRB IDTH KJH NKKI DJT QBCP NKKI RKJQDCJHT D NSPQX
GKKFCJA RKKFPQKUH D RSLOKDNT YKN QBH TCPBHP D QDOGH QBNHH KN YKSN RBDCNP DJT
QBH OHTP SJRGH BHJNX DJT DSJQ HI BDT D OCA OHT CJ KJH RKNJHN DJT TKNKQBX D
GCQQGH OHT CJ DJKQBHN RKNJHN QBHNH VDP JK ADNNHQ DQ DGG DJT JK RHGGDN HWRHLQ
D PIDGG BKGH TSA CJ QBH ANKSJT RDGGHT D RXRGKJH RHGGDN VBHNH QBH YDICGX
RKSGT AK CJ RDPH KJH KY QBKPH ANHDQ VBCNGVCJTP DNKPH ICABQX HJKSAB QK RNSPB
DJX OSCGTCJA CJ CQP LDQB CQ VDP NHDRBHT OX D QNDL TKKN CJ QBH ICTTGH KY QBH
YGKKN YNKI VBCRB D GDTTHN GHT TKVJ CJQK QBH PIDGG TDNF BKGH
```

Frequency Analysis: An Example

TKNKQBX GCUHT CJ QBH ICTPQ KY QBH ANHDQ FDJPDP LNDCNCHP VCQB SJRGH BHJNX VBK
VDP D YDNIHN DJT DSJQ HI VBK VDP QBH YDNIHNP VCYH QBHCN BKSPH VDP PIDGG YKN
QBH GSIOHN QK OSCGT CQ BDT QK OH RDNNCHT OX VDAKJ IDJX ICGHP QBHNH VHNH YKSN
VDGGP D YGKKN DJT D NKKY VBCRB IDTH KJH NKKI DJT QBCP NKKI RKJQDCJHT D NSPQX
GKKFCJA RKKFPQKUH D RSLOKDNT YKN QBH TCPBHP D QDOGH QBNHH KN YKSN RBDCNP DJT
QBH OHTP SJRGH BHJNX DJT DSJQ HI BDT D OCA OHT CJ KJH RKNJHN DJT TKNKQBX D
GCQQGH OHT CJ DJKQBHN RKNJHN QBHNH VDP JK ADNNHQ DQ DGG DJT JK RHGGDN HWRHLQ
D PIDGG BKGH TSA CJ QBH ANKSJT RDGGHT D RXRGKJH RHGGDN VBHNH QBH YDICGX
RKSGT AK CJ RDPH KJH KY QBKPH ANHDQ VBCNGVCJTP DNKPH ICABQX HJKSAB QK RNSPB
DJX OSCGTCJA CJ CQP LDQB CQ VDP NHDRBHT OX D QNDL TKKN CJ QBH ICTTGH KY QBH
YGKKN YNKI VBCRB D GDTTHN GHT TKVJ CJQK QBH PIDGG TDNF BKGH

Frequency Analysis: An Example

TKNKQBX GCUeT CJ QBe ICTPQ KY QBe ANeDQ FDJPDP LNDCNCeP VCQB SJRGe BeJNX VBK
VDP D YDNIeN DJT DSJQ eI VBK VDP QBe YDNIeNP VCYe QBeCN BKSPe VDP PIDGG YKN
QBe GSIOeN QK OSCGT CQ BDT QK Oe RDNNCeT OX VDAKJ IDJX ICGeP QBeNe VeNe YKSN
VDGGP D YGKKN DJT D NKKY VBCRB IDTe KJe NKKI DJT QBCP NKKI RKJQDCJeT D NSPQX
GKKFCJA RKKFPQKUe D RSLOKDNT YKN QBe TCPBeP D QDOGe QBNee KN YKSN RBDCNP DJT
QBe OeTP SJRGe BeJNX DJT DSJQ eI BDT D OCA OeT CJ KJe RKNJeN DJT TKNKQBX D
GCQQGe OeT CJ DJKQBeN RKNJeN QBeNe VDP JK ADNNeQ DQ DGG DJT JK ReGGDN eWReLQ
D PIDGG BKGe TSA CJ QBe ANKSJT RDGGeT D RXRGKJe ReGGDN VBeNe QBe YDICGX
RKSGT AK CJ RDPe KJe KY QBKPe ANeDQ VBCNGVCJTP DNKPe ICABQX eJKSAB QK RNSPB
DJX OSCGTCJA CJ CQP LDQB CQ VDP NeDRBeT OX D QNDL TKKN CJ QBe ICTTGe KY QBe
YGKKN YNKI VBCRB D GDTTeN GeT TKVJ CJQK QBe PIDGG TDNF BKGe

Frequency Analysis: An Example

TKNKQBX GCUeT CJ QBe ICTPQ KY QBe ANeaQ FaJPaP LNaCNCeP VCQB SJRGe BeJNX VBK
VaP a YaNIeN aJT aSJQ eI VBK VaP QBe YaNIeNP VCYe QBeCN BKSPe VaP PIaGG YKN
QBe GSIOeN QK OSCGT CQ BaT QK Oe RaNNCeT OX VaAKJ IaJX ICGeP QBeNe VeNe YKSN
VaGGP a YGKKN aJT a NKKY VBCRB IaTe KJe NKKI aJT QBCP NKKI RKJQaCJeT a NSPQX
GKKFCJA RKKFPQKUe a RSLOKaNT YKN QBe TCPBeP a QaOGe QBNee KN YKSN RBaCNP aJT
QBe OeTP SJRGe BeJNX aJT aSJQ eI BaT a OCA OeT CJ KJe RKNJeN aJT TKNKQBX a
GCQQGe OeT CJ aJKQBeN RKNJeN QBeNe VaP JK AaNNeQ aQ aGG aJT JK ReGGaN eWReLQ
a PIaGG BKGe TSA CJ QBe ANKSJT RaGGeT a RXRGKJe ReGGaN VBeNe QBe YaICGX
RKSGT AK CJ RaPe KJe KY QBKPe ANeaQ VBCNGVCJTP aNKPe ICABQX eJKSAB QK RNSPB
aJX OSCGTCJA CJ CQP LaQB CQ VaP NeaRBeT OX a QNaL TKKN CJ QBe ICTTGe KY QBe
YGKKN YNKI VBCRB a GaTTeN GeT TKVJ CJQK QBe PIaGG TaNF BKGe

Frequency Analysis: An Example

TKNKthX GCUeT CJ the ICTPt KY the ANeat FaJPaP LNaCNCeP VCth SJRGe heJNX VhK
VaP a YaNIeN aJT aSJt eI VhK VaP the YaNIeNP VCYe theCN hKSPe VaP PIaGG YKN
the GSIOeN tK OSCGT Ct haT tK Oe RaNNCeT OX VaAKJ IaJX ICGeP theNe VeNe YKSN
VaGGP a YGKKN aJT a NKKY VhCRh IaTe KJe NKKI aJT thCP NKKI RKJtaCJeT a NSPtX
GKKFCJA RKKFPtKUe a RSLOKaNT YKN the TCPheP a taOGe thNee KN YKSN RhaCNP aJT
the OeTP SJRGe heJNX aJT aSJt eI haT a OCA OeT CJ KJe RKNJeN aJT TKNKthX a
GCttGe OeT CJ aJKtheN RKNJeN theNe VaP JK AaNNet at aGG aJT JK ReGGaN eWReLt
a PIaGG hKGe TSA CJ the ANKSJT RaGGeT a RXRGKJe ReGGaN VheNe the YaICGX
RKSGT AK CJ RaPe KJe KY thKPe ANeat VhCNGVCJTP aNKPe ICAhtX eJKSAh tK RNSPh
aJX OSCGTCJA CJ CtP Lath Ct VaP NeaRheT OX a tNaL TKKN CJ the ICTTGe KY the
YGKKN YNKI VhCRh a GaTTeN GeT TKVJ CJtK the PIaGG TaNF hKGe

Frequency Analysis: An Example

dKNKthX GCUed Cn the ICdPt KY the ANeat FanPaP LNaCNCeP VCth SnRGe henNX VhK
VaP a YaNIeN and aSnt eI VhK VaP the YaNIeNP VCYe theCN hKSPe VaP PIaGG YKN
the GSIOeN tK OSCGd Ct had tK Oe RaNNCed OX VaAKn IanX ICGeP theNe VeNe YKSN
VaGGP a YGKKN and a NKKY VhCRh Iade Kne NKKI and thCP NKKI RKntaCned a NSPtX
GKKFCnA RKKFPtKUe a RSLOKaNd YKN the dCPheP a taOGe thNee KN YKSN RhaCNP and
the OedP SnRGe henNX and aSnt eI had a OCA Oed Cn Kne RKNneN and dKNKthX a
GCttGe Oed Cn anKtheN RKNneN theNe VaP nK AaNNet at aGG and nK ReGGaN eWReLt
a PIaGG hKGe dSA Cn the ANKSnd RaGGed a RXRGKne ReGGaN VheNe the YaICGX
RKSGd AK Cn RaPe Kne KY thKPe ANeat VhCNGVCndP aNKPe ICAhtX enKSAh tK RNSPh
anX OSCGdCnA Cn CtP Lath Ct VaP NeaRhed OX a tNaL dKKN Cn the ICddGe KY the
YGKKN YNKI VhCRh a GaddeN Ged dKVn CntK the PIaGG daNF hKGe

Frequency Analysis: An Example

dorothX GiUed in the IidPt oY the Areat FanPaP LrairieP Vith SncGe henrX Vho VaP a YarIer and aSnt eI Vho VaP the YarIerP ViYe their hoSPe VaP PIaGG Yor the GSIOer to OSiGd it had to Oe carried OX VaAon IanX IiGeP there Vere YoSr VaGGP a YGoor and a rooY Vhich Iade one rooI and thiP rooI contained a rSPtX GooFinA cooFPtoUe a cSLOoard Yor the diPheP a taOGe three or YoSr chairP and the OedP SncGe henrX and aSnt eI had a OiA Oed in one corner and dorothX a GittGe Oed in another corner there VaP no Aarret at aGG and no ceGGar eWceLt a PIaGG hoGe dSA in the AroSnd caGGed a cXcGone ceGGar Vhere the YaIiGX coSGd Ao in caPe one oY thoPe Areat VhirGVindP aroPe IiAhtX enoSAh to crSPh anX OSiGdinA in itP Lath it VaP reached OX a traL door in the IiddGe oY the YGoor YroI Vhich a Gadder Ged doVn into the PIaGG darF hoGe

Frequency Analysis: An Example

dorothy lived in the midst of the great kansas prairies with uncle henry who was a farmer and aunt em who was the farmers wife their house was small for the lumber to build it had to be carried by wagon many miles there were four walls a floor and a roof which made one room and this room contained a rusty looking cookstove a cupboard for the dishes a table three or four chairs and the beds uncle henry and aunt em had a big bed in one corner and dorothy a little bed in another corner there was no garret at all and no cellar except a small hole dug in the ground called a cyclone cellar where the family could go in case one of those great whirlwinds arose mighty enough to crush any building in its path it was reached by a trap door in the middle of the floor from which a ladder led down into the small dark hole

Weaknesses of Substitution

- Even with 400,000,000,000,000,000,000 or more keys, the general substitution cipher is still easily broken.

Weaknesses of Substitution

- Even with 400,000,000,000,000,000,000 or more keys, the general substitution cipher is still easily broken.
- This came as a big shock as it was thought to be safe for centuries.

Weaknesses of Substitution

- Even with 400,000,000,000,000,000,000 or more keys, the general substitution cipher is still easily broken.
- This came as a big shock as it was thought to be safe for centuries.
- The problem is when a letter is replaced by another it remains the same throughout. E.g., if a = J, then it stays as a J.

Weaknesses of Substitution

- Even with 400,000,000,000,000,000,000 or more keys, the general substitution cipher is still easily broken.
- This came as a big shock as it was thought to be safe for centuries.
- The problem is when a letter is replaced by another it remains the same throughout. E.g., if a = J, then it stays as a J.
- If we could use more than one alphabet, then a would not always end up as J.

Is there anything odd here...?

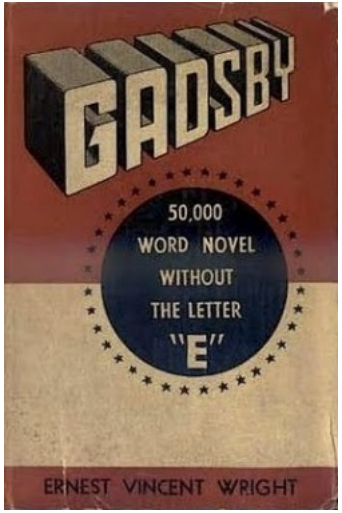
If Youth, throughout all history, had had a champion to stand up for it; to show a doubting world that a child can think; and, possibly, do it practically; you wouldn't constantly run across folks today who claim that ''a child don't know anything.'' A child's brain starts functioning at birth; and has, amongst its many infant convolutions, thousands of dormant atoms, into which God has put a mystic possibility for noticing an adult's act, and figuring out its purport.

Is there anything odd here...?

- Written in 1939, *Gadsby* is a novel written by Ernest Vincent Wright. It has 260 pages, over 50,000 words, but without using a single

Is there anything odd here...?

- Written in 1939, *Gadsby* is a novel written by Ernest Vincent Wright. It has 260 pages, over 50,000 words, but without using a single 'e'.



Tutorials

In the tutorial this week we will:

- Create a spreadsheet to perform encryption using a Caesar cipher.
- Create and use a spreadsheet to perform frequency analysis attacks on general substitution ciphers.