Introduction

Random Numbers

Symmetric Encryption

Hashes

Asymmetric Encryption

Certificates

Signatures

SSL/TLS

SSH

VPN

Email

Disk Encryption

Attacks

Introduction

Applied Cryptography

Fundamental Concepts

WHAT IF I TOLD YOU
Your mother was a hamster and your father smelt of elderberries

What is Cryptography?

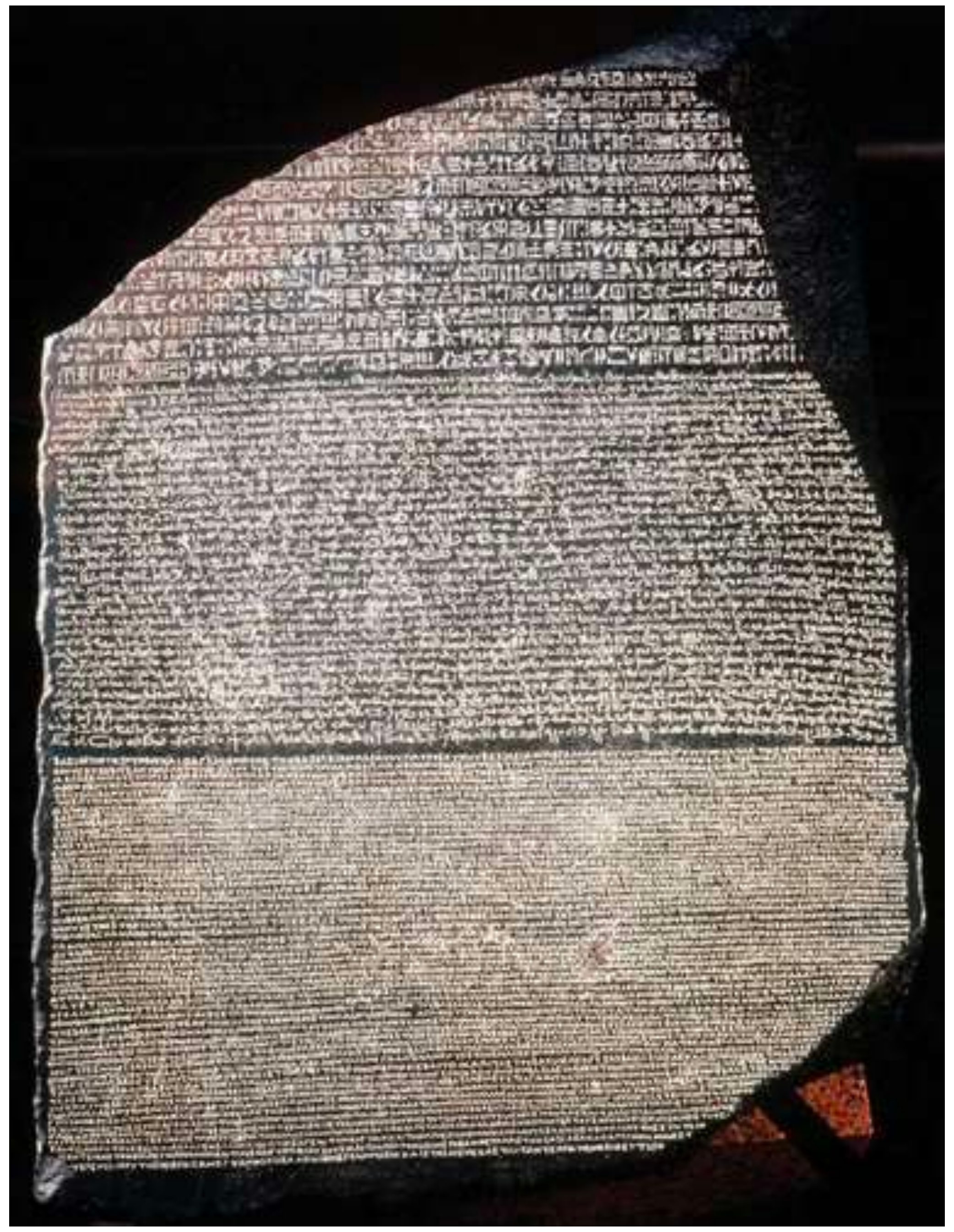
- Cryptography is the theory and practice of communicating securely between co-operating parties
- Cryptography is often thought of as another term for encryption, but it is much more
- Cryptography's goal is to add the following properties, known as the CIA properties, to information:
  - Confidentiality
  - Integrity
  - Authentication

- Many schemes have been devised to communicate securely, adding one or more of these properties with varying degrees of success

Applied Cryptography

- Putting cryptographic theory into practice requires designing and implementing processes and tools that embody those theories
- In this course, we are examining applications of cryptography to protect digital data at rest and in flight
- Modern digital cryptography is built on a set of mathematical theories which are beyond the scope of this course - we are engaging in using cryptographic tools and techniques, not designing them
- Four cryptographic primitives are used to create all other cryptographic capabilities and we will examine these and some tools for working with them
- We will also examine related technologies such as creating authentication codes

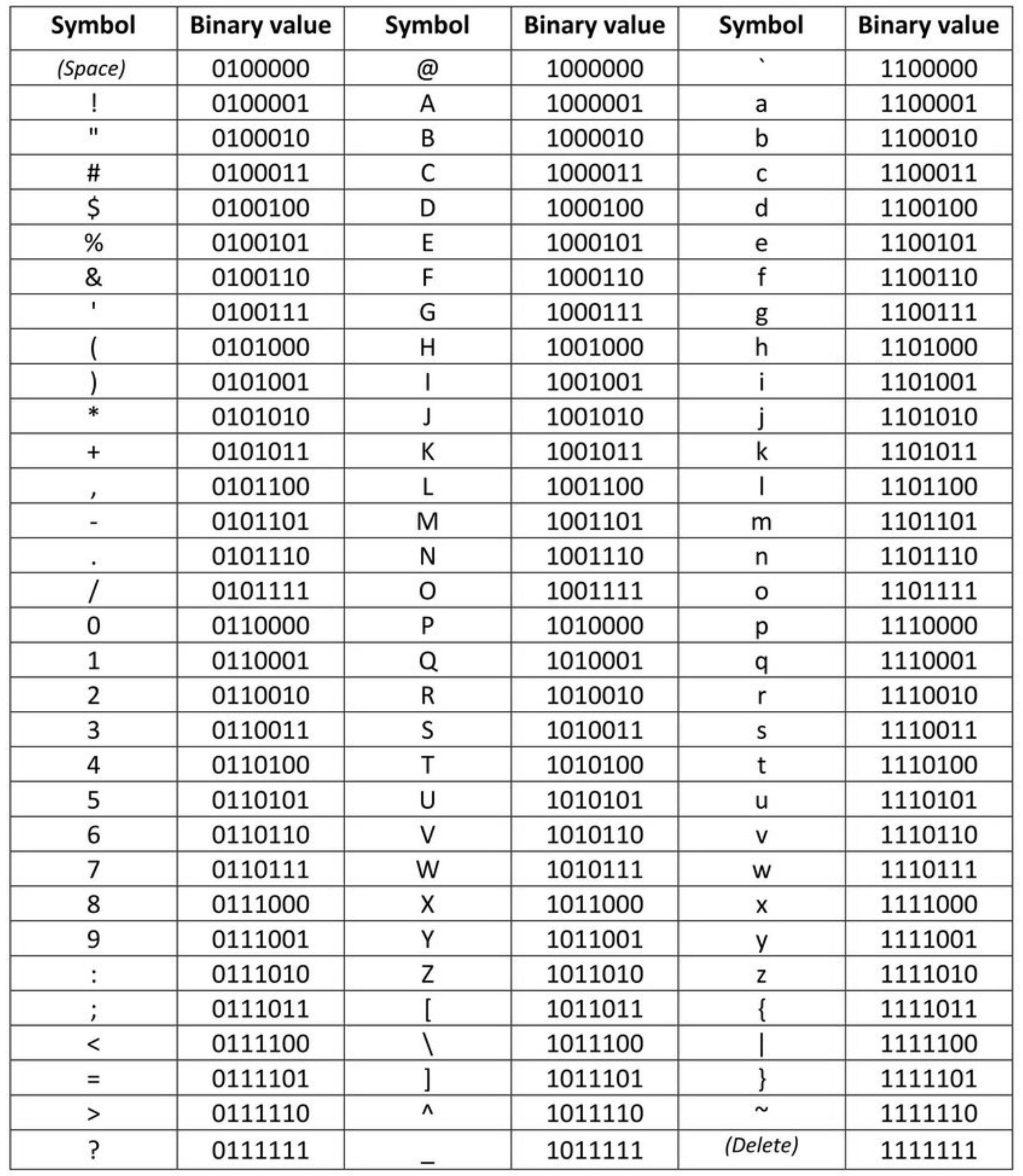
Encoding, Hiding, Compressing



https://cdn.britannica.com/37/116537-050-69BA573B/Rosetta-Stone-basalt-slab-Fort-Saint-Julien-Egypt-196-bce.jpg

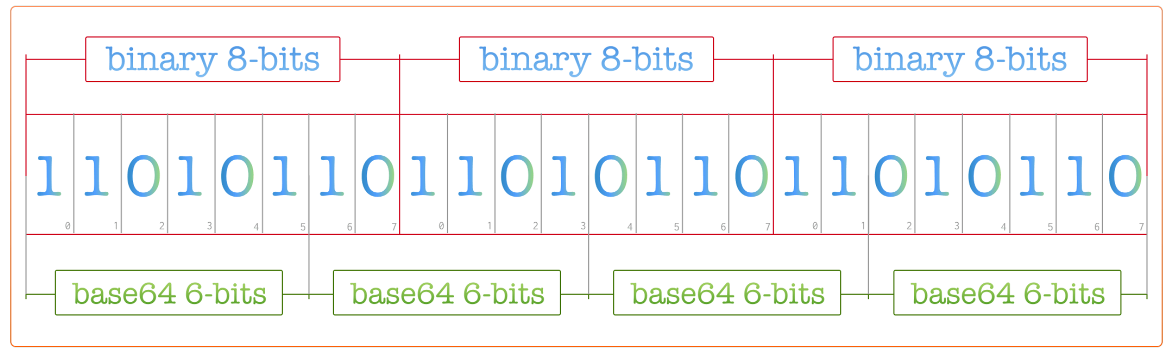
Encoding vs. Encrypting

- Data can be encoded (converted from one form to another) to enable storage or transmission of that data (e.g. Morse Code, ASCII, UTF-16, Base64, compression, etc.)
- Encoding uses publicly available algorithms or tables and does not use any privileged or secret information in the encoding/decoding process
- Encryption uses publicly available algorithms to convert between unencrypted data and encrypted data, but also requires one or more pieces of secret information for the data converter to complete the conversion

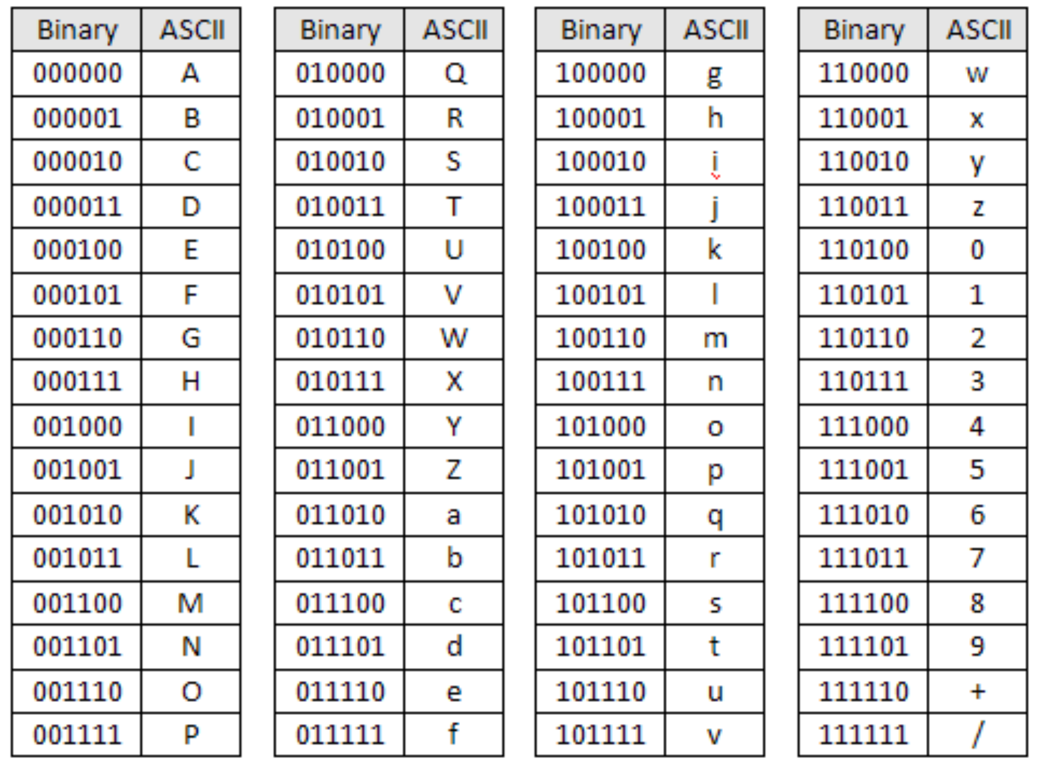


Symbol	Binary value	Symbol	Binary value	Symbol	Binary value
(Space)	0100000	@	1000000	`	1100000
!	0100001	A	1000001	a	1100001
"	0100010	B	1000010	b	1100010
#	0100011	C	1000011	c	1100011
$	0100100	D	1000100	d	1100100
%	0100101	E	1000101	e	1100101
&	0100110	F	1000110	f	1100110
'	0100111	G	1000111	g	1100111
(	0101000	H	1001000	h	1101000
)	0101001	I	1001001	i	1101001
*	0101010	J	1001010	j	1101010
+	0101011	K	1001011	k	1101011
,	0101100	L	1001100	l	1101100
-	0101101	M	1001101	m	1101101
.	0101110	N	1001110	n	1101110
/	0101111	O	1001111	o	1101111
0	0110000	P	1010000	p	1110000
1	0110001	Q	1010001	q	1110001
2	0110010	R	1010010	r	1110010
3	0110011	S	1010011	s	1110011
4	0110100	T	1010100	t	1110100
5	0110101	U	1010101	u	1110101
6	0110110	V	1010110	v	1110110
7	0110111	W	1010111	w	1110111
8	0111000	X	1011000	x	1111000
9	0111001	Y	1011001	y	1111001
:	0111010	Z	1011010	z	1111010
;	0111011	[	1011011	{	1111011
<	0111100	\	1011100	|	1111100
=	0111101	]	1011101	}	1111101
>	0111110	^	1011110	~	1111110
?	0111111	_	1011111	*(Delete)*	1111111

https://www.researchgate.net/figure/Printable-characters-in-ASCII-set-and-their-binary-values_fig2_317351196

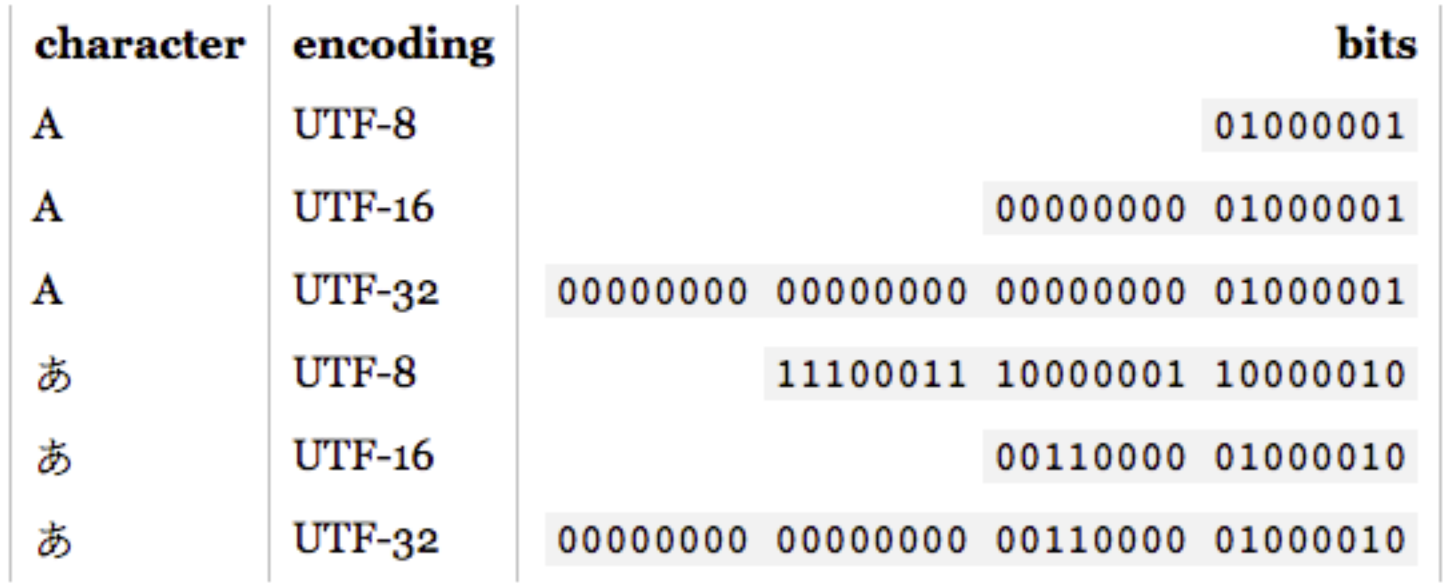


https://time2hack.com/convert-encode-decode-text-files-to-base64-in-browser/



Binary	ASCII	Binary	ASCII	Binary	ASCII	Binary	ASCII
000000	A	010000	Q	100000	g	110000	w
000001	B	010001	R	100001	h	110001	x
000010	C	010010	S	100010	i	110010	y
000011	D	010011	T	100011	j	110011	z
000100	E	010100	U	100100	k	110100	0
000101	F	010101	V	100101	l	110101	1
000110	G	010110	W	100110	m	110110	2
000111	H	010111	X	100111	n	110111	3
001000	I	011000	Y	101000	o	111000	4
001001	J	011001	Z	101001	p	111001	5
001010	K	011010	a	101010	q	111010	6
001011	L	011011	b	101011	r	111011	7
001100	M	011100	c	101100	s	111100	8
001101	N	011101	d	101101	t	111101	9
001110	O	011110	e	101110	u	111110	+
001111	P	011111	f	101111	v	111111	/

https://dev.to/neumaneuma/decoding-the-confusing-world-of-encodings-part-2-4lo



character	encoding	bits
A	UTF-8	01000001
A	UTF-16	00000000 01000001
A	UTF-32	00000000 00000000 00000000 01000001
あ	UTF-8	11100011 10000001 10000010
あ	UTF-16	00110000 01000010
あ	UTF-32	00000000 00000000 00110000 01000010

https://stackoverflow.com/questions/2241348/what-is-unicode-utf-8-utf-16

Data Hiding

- Sieves and other Steganographic techniques to hide data are not cryptography
- Data hiding primarily aims to avoid suspicion that there is any secret data at all
- Steganography does not alter the plaintext, it simply covers it using other reasonable-looking plaintext so that it is not possible to know the hidden data is there or extract it without knowledge of its presence and the technique used to hide it
- Steganography does not authenticate who created the stegofile, or verify the integrity of the extracted data
- Cryptographic capabilities have been added to some stego tools, but they do their work separately from the data hiding/recovery

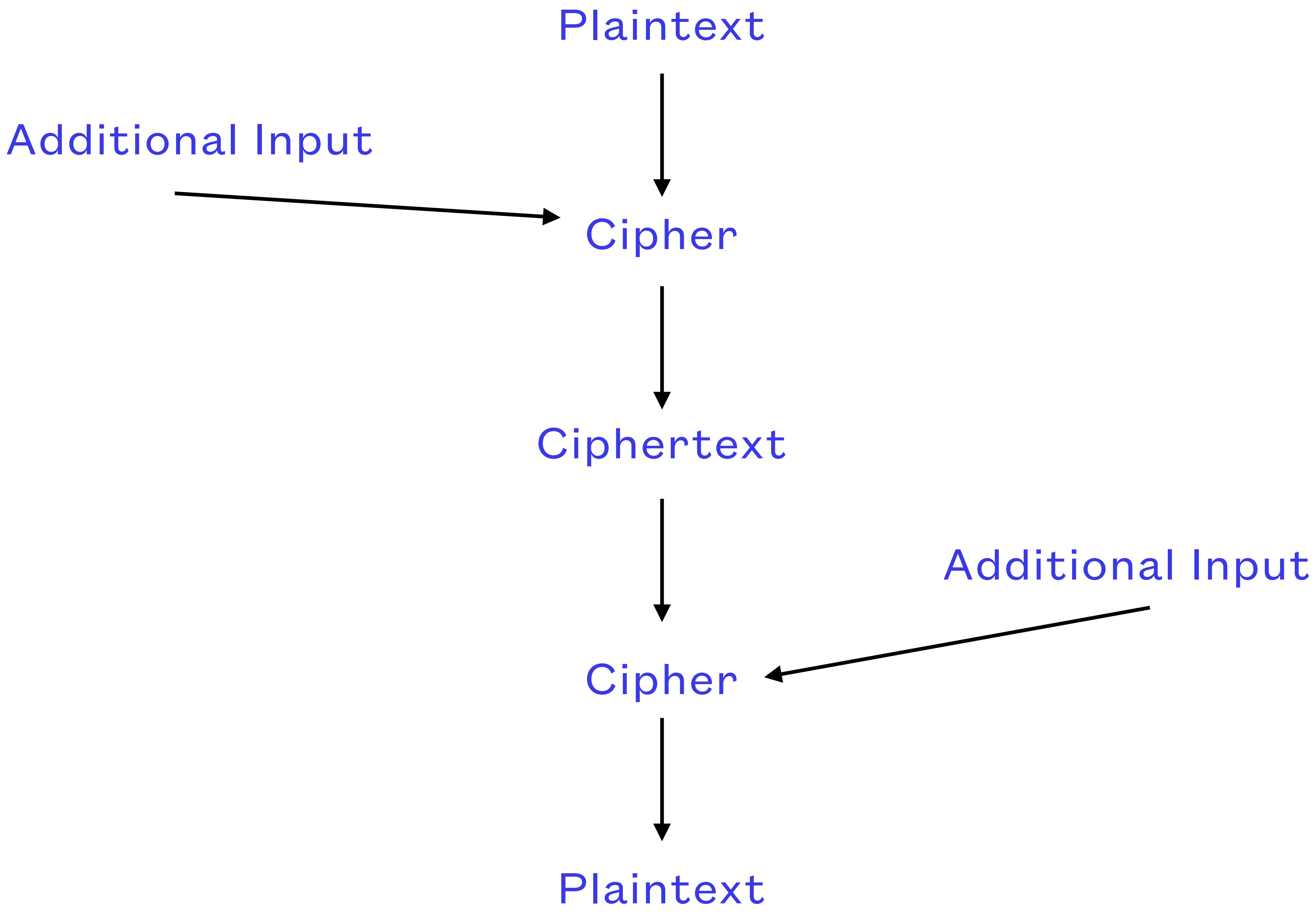
Compression

- Compression is not cryptography, but many crypto tools include compression capabilities
- Compression replaces data by encoding it using an encoding method designed to reduce the length of the data
- Compression does nothing to prevent anyone from decompressing the compressed data
- Compression does not authenticate the author of the compressed data, or verify the integrity of the extracted data
- It is possible with some compression algorithms to select algorithm parameters such that you cannot uncompress the data without knowing those parameters, but this is normally done for efficiency not security
- The number of parameter choices is too limited to provide security and would fall trivially to simple trial and error
- Cryptographic capabilities (encryption/hashes) have been added to some compression tools, but they do their work separately from the data compression/decompression

Crypto Basics

https://arstechnica.com/information-technology/2019/10/alexa-and-google-home-abused-to-eavesdrop-and-phish-passwords/

Terminology



Plaintext

- The original source data
- Does not imply that the source data is composed of text (e.g. letters, numbers, symbols)

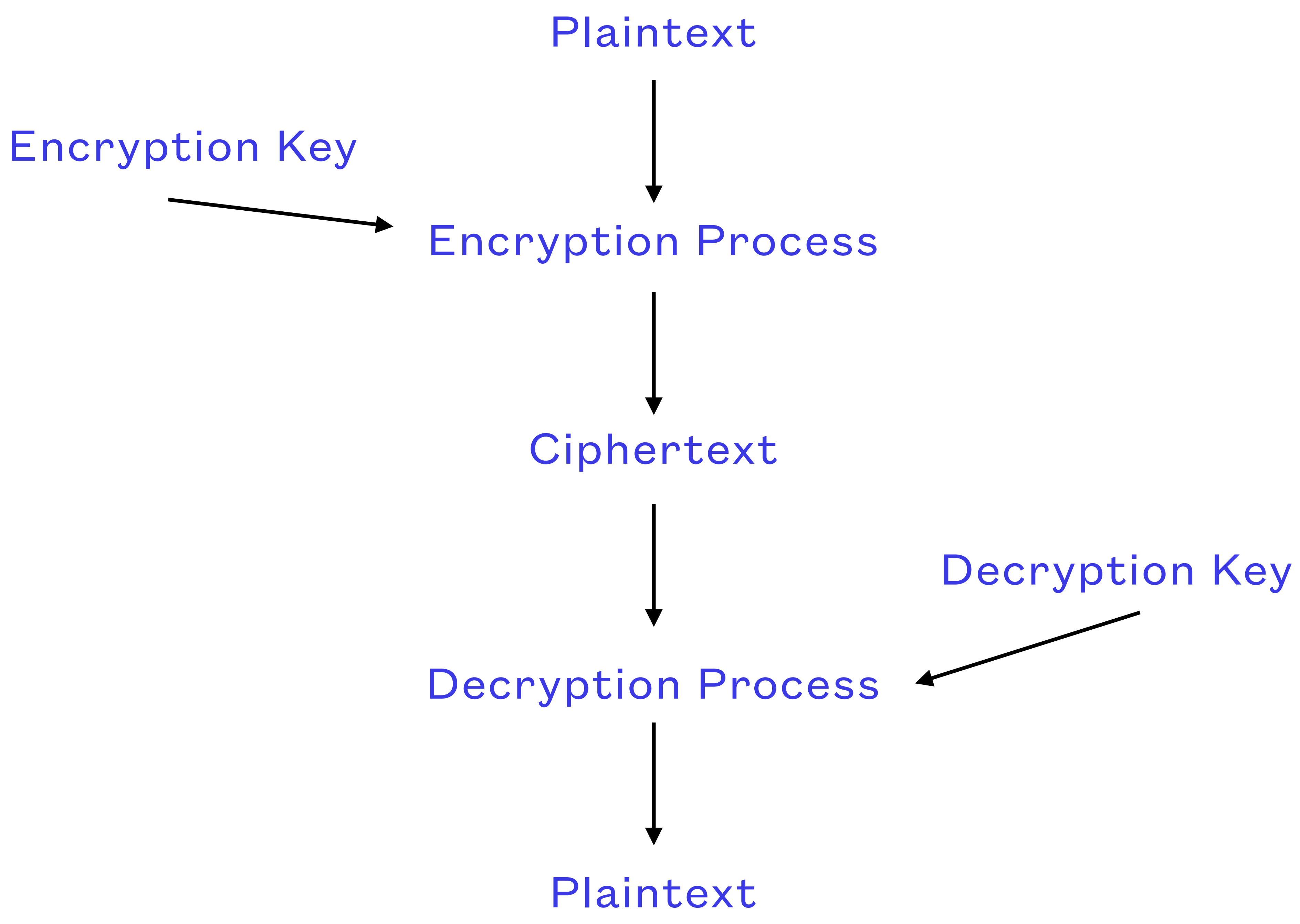
Ciphertext

- Data in an encrypted and unreadable form
- Requires secret knowledge to convert back into plaintext

Cipher

- The algorithm or process by which plaintext can be encrypted and the resulting ciphertext decrypted
- Requires additional inputs (secret knowledge), otherwise the data is being encoded/decoded, not encrypted/decrypted

Terminology



Encryption

- The act of converting plaintext into ciphertext (i.e. converted to a form which is unreadable by a third party)

Decryption

- The act of converting ciphertext into plaintext

Keys

- Data required by the encryption and/or decryption processes to successfully function

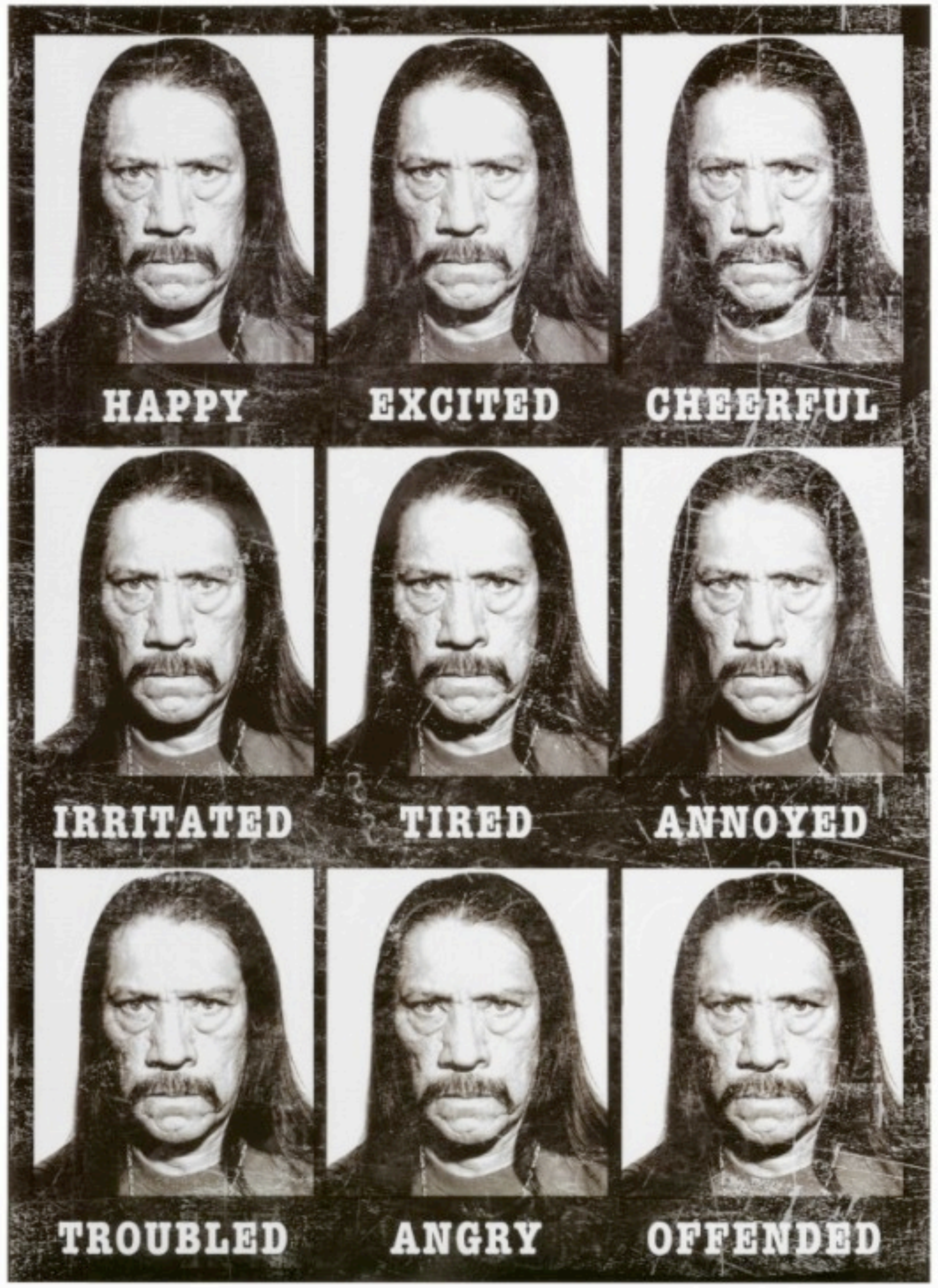
Symmetric Encryption

- The encryption key and decryption key are identical or trivially derived from each other

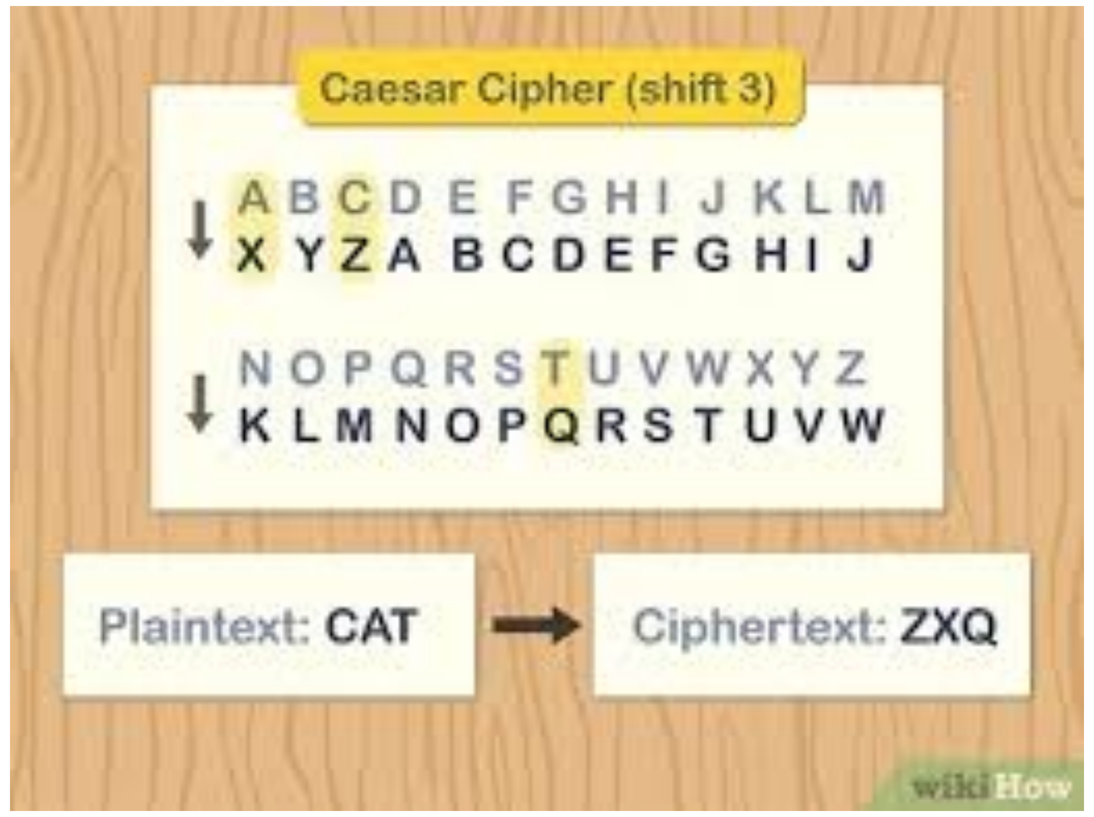
Asymmetric Encryption

- The encryption key and decryption key are neither identical or trivially derived from each other
- The encryption key and decryption key are mathematically matched to each other and only to each other

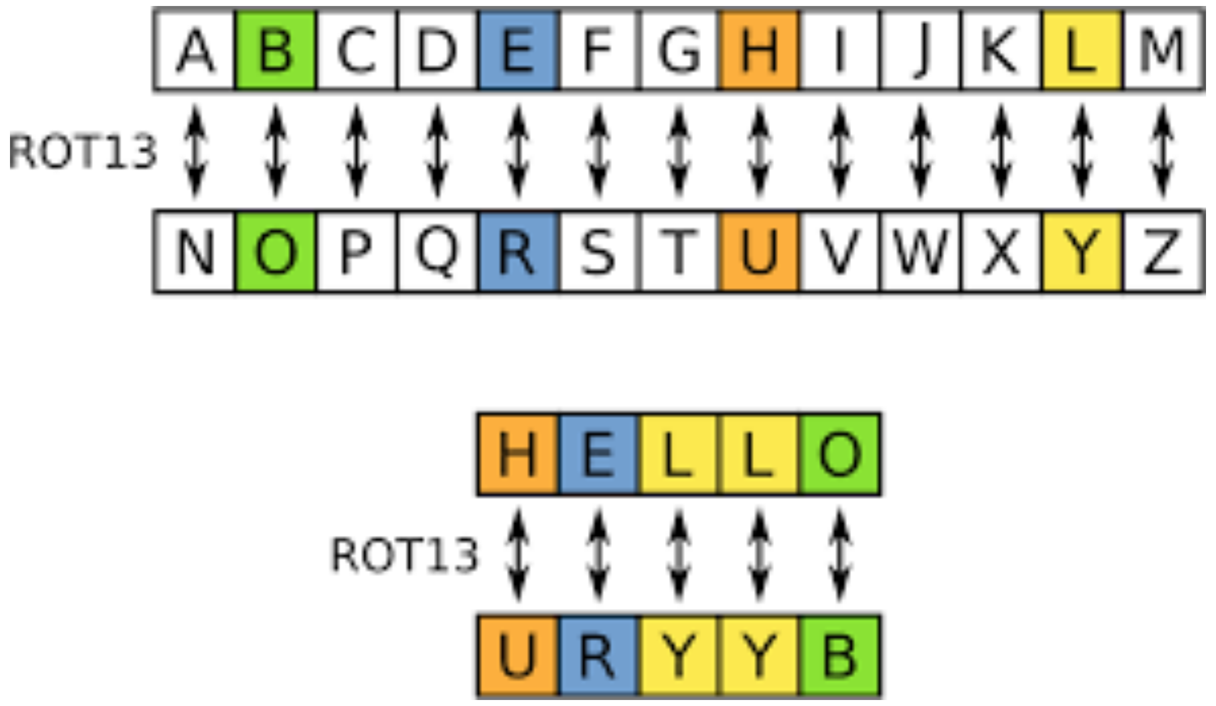
Substitution



- When we encrypt plaintext, we are substituting something we cannot read for something we can
- When we decrypt ciphertext, we are undoing the substitution which gives us back the original readable and usable content
- The rules for the substitution we call the cipher, and we use one or more secret keys as part of the rules, so that just knowing which cipher was used doesn't allow unauthorized actors to decrypt our data


Caesar Cipher (shift 3)
A B C D E F G H I J K L M
X Y Z A B C D E F G H I J
N O P Q R S T U V W X Y Z
K L M N O P Q R S T U V W
Plaintext: CAT
Ciphertext: ZXQ
wikiHow

https://www.wikihow.com/Create-Substitution-Ciphers



https://en.wikipedia.org/wiki/Substitution_cipher

Transposition Ciphers

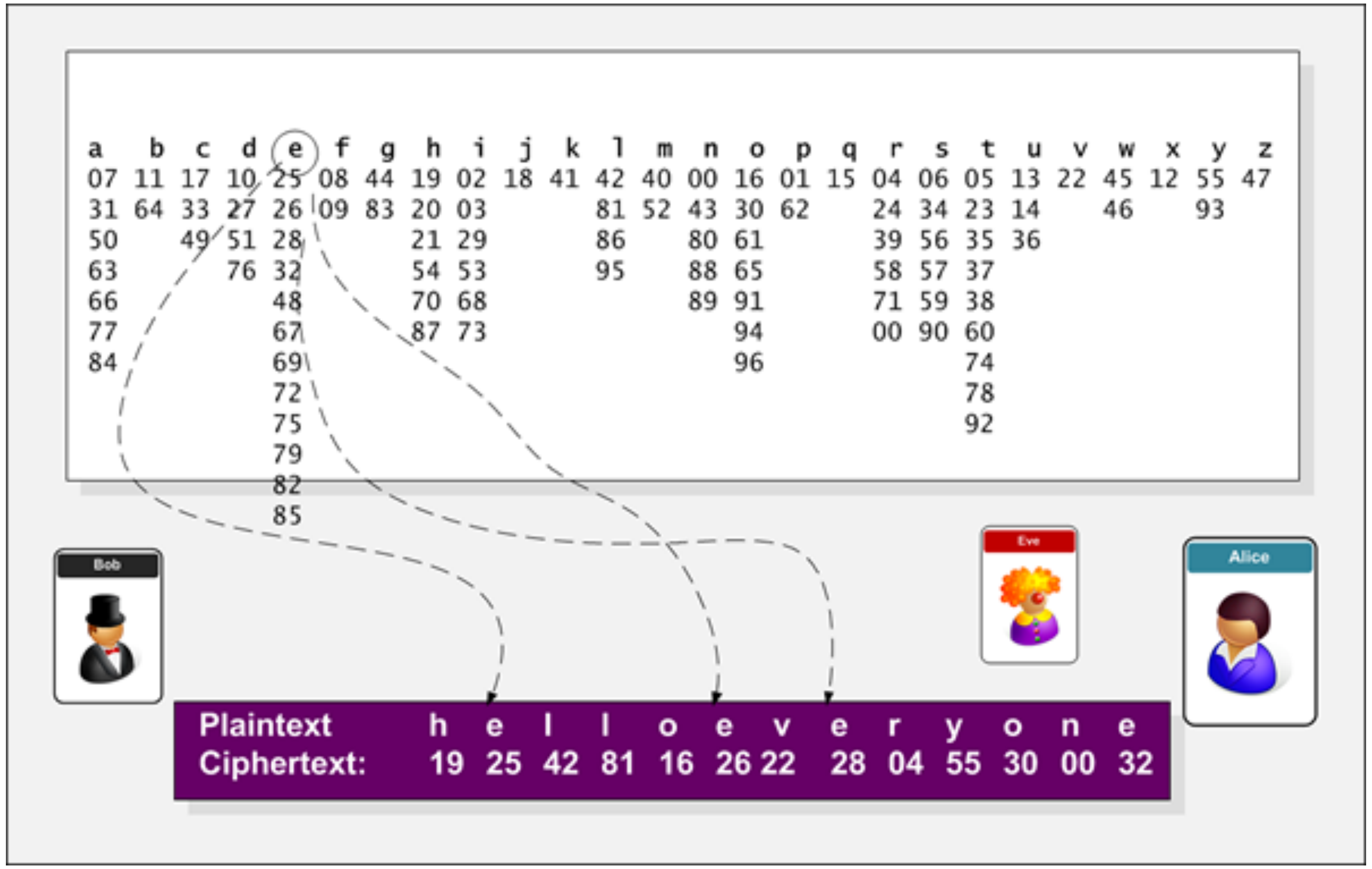
- Transposition Ciphers are among the simplest of ciphers and are a specific type of substitution
- A transposition cipher is one where we keep an ordered list of the possible characters in the plaintext, and generate the ciphertext by substituting using an offset in the list from the plaintext character
- Caesar cipher is a very common transposition cipher, with offset 3 and the offset does not change during the encrypting/decrypting process
- It is said Julius Caesar used this cipher
- The key in this kind of cipher is the value of the offset
- Only 25 possible keys means you can solve them with trial and error even on paper

rot13/caesar

CLI tool to do caesar ciphers

```
caesar <<<"Hello World"
caesar 13 <<<"Hello World"
rot13 <<<"Hello World"
caesar 13 <<<"Hello World" | rot13
caesar 13 <<<"Hello World" | caesar 3 | caesar 6 | caesar 4
```

- Does 13 character rotation when run as rot13
- Does 3 character rotation when run as caesar, but you can specify the rotation to use if you run as caesar
- Was a common tool to use in the early days of Usenet to conceal info from the uninformed
- Only handles english letters
- Preserves case
- Preserves non-alphabetic characters
- Man page has a table of character frequencies for english literature



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
07	11	17	10	25	08	44	19	02	18	41	42	40	00	16	01	15	04	06	05	13	22	45	12	55	47
31	64	33	27	26	09	83	20	03			81	52	43	30	62		24	34	23	14		46		93	
50		49	51	28			21	29			86		80	61			39	56	35	36					
63			76	32			54	53			95		88	65			58	57	37						
66				48			70	68					89	91			71	59	38						
77				67			87	73						94			00	90	60						
84				69										96					74						
				72															78						
				75															92						
				79																					
				82																					
				85																					

Plaintext	h	e	l	l	o	e	v	e	r	y	o	n	e
Ciphertext:	19	25	42	81	16	26	22	28	04	55	30	00	32

https://asecuritysite.com/challenges/ho

Polyalphabetic Substitution Ciphers

- Substitution ciphers can be created with more sophisticated substitutions than just transposition with the goal being to make cryptanalysis and brute forcing harder
- A polyalphabetic substitution cipher is one where we keep an ordered list of the possible characters in the plaintext, and generate the ciphertext by using multiple values from a lookup table for those characters
- The table content is the key for this kind of cipher
- Tables like this but significantly expanded are called S-boxes and form one component of advanced encryption algorithms

Vigenère Ciphers

Plaintext: vignere is not secure

Key: playdirty

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

Offsets For Key: 15 11 0 24 3 8 17 19 24

plaintext	v	i	g	n	e	r	e		i	s		n	o	t		s	e	c	u	r	e
key	p	l	a	y	d	i	r		t	y		p	l	a		y	d	i	r	t	y
offset	15	11	0	24	3	8	17		19	24		15	11	0		24	3	8	17	19	24
algorithm	v >> 15	i >> 11	g >> 0	n >> 24	e >> 3	r >> 8	e >> 17		i >> 19	s >> 24		n >> 15	o >> 11	t >> 0		s >> 24	e >> 3	c >> 8	u >> 17	r >> 19	e >> 24
ciphertext	k	t	g	l	h	z	v		b	q		c	z	t		q	h	k	l	k	c

Ciphertext: ktglhzv bq czt qhklkc

- A Vigenère cipher is one where we start with a plaintext message and a key and is a form of polyalphabetic cipher
- Each character in the key is assigned a numerical value which is used as the transposition offset for that plaintext character
- The key repeats as many times as needed to provide an offset value for each character of the plaintext
- The key to decrypt is the same as the one to encrypt - it is a symmetric cipher
- The key is needed to decrypt the ciphertext
- For a discussion with examples of substitution ciphers and why they aren't secure, see the dcode.fr website link on the course website

Breaking Simple Ciphers

- Simple ciphers are not used for anything but education because they are trivial to break
- The ciphertext is the same length as the plaintext, and the plaintext always produces the same ciphertext
- The search space for keys is small and easily subjected to brute forcing attacks, simply trying keys until you have found the right one
- Techniques to make it harder include longer keys, having multiple ciphertext outputs for the same plaintext, using multiple keys, re-encrypting multiple times in various ways, adding extra data to the plaintext (padding) prior to encryption, etc.
- Cracking ciphertext is always much easier if you have information about the characteristics of the plaintext (e.g. length, english language text, image data, microsoft documents, protocol encapsulations, etc.)

One-time Pad

- The methods for breaking simple ciphers involve taking advantage of keys that repeat and the small character sets being used along with knowledge about the statistically probable plaintext character distributions (information outside of the encryption algorithm that exposes secrets, sometimes referred to as a side channel - think of a stethoscope on the door of a safe)

- Making this task impossible requires that every message must have a unique key and the key must be random and as long as the plaintext - this is known as a one-time pad because in practice, the keys were recorded using pads of paper sort of like in the movie The Numbers Station

- Even though it cannot be decrypted without the key, all you have really done is move the problem of keeping the plaintexts secret to one of keeping the keys secret and have added the problem of generating keys

- Because of the logistical issues around generating and distributing one-time pad keys, existing keys get reused and then the uniqueness of the keys is lost

- Using machines to generate the one-time keys in such a way that a human cannot predict them but duplicate machines can produce the same sequences of keys was the basis for the german Enigma cipher and the japanese Purple cipher, both used and defeated in WW2 by cryptanalysis that took side channels into account

XOR Ciphers

- Many ciphers for secure communications use boolean exclusive-or (XOR) substitutions as part of their encryption/decryption algorithm
- Instead of using transposition or table-based substitutions, the key's binary representation is exclusive-or'ed with the plaintext's binary representation to produce the ciphertext's binary representation
- This is a symmetric key substitution, re-applying the same key the same way to the ciphertext will produce the plaintext
- Because it is a binary-based substitution it is suitable for literally any kind of digital data, not just human-readable messages, and any key content, not just human-readable characters
- It is very fast to do on a CPU, and is part of more or less all secure digital ciphers