UNIVERSIDAD
NACIONAL
DE COLOMBIA

Introduction to Cryptography and Information Security

Symmetric Ciphers

Jeisson Andrés Vergara Vargas, M.Sc.
Departamento de Ingeniería de Sistemas e Industrial
http://javergarav.academy/
javergarav@unal.edu.co

2018-II

©

Notations

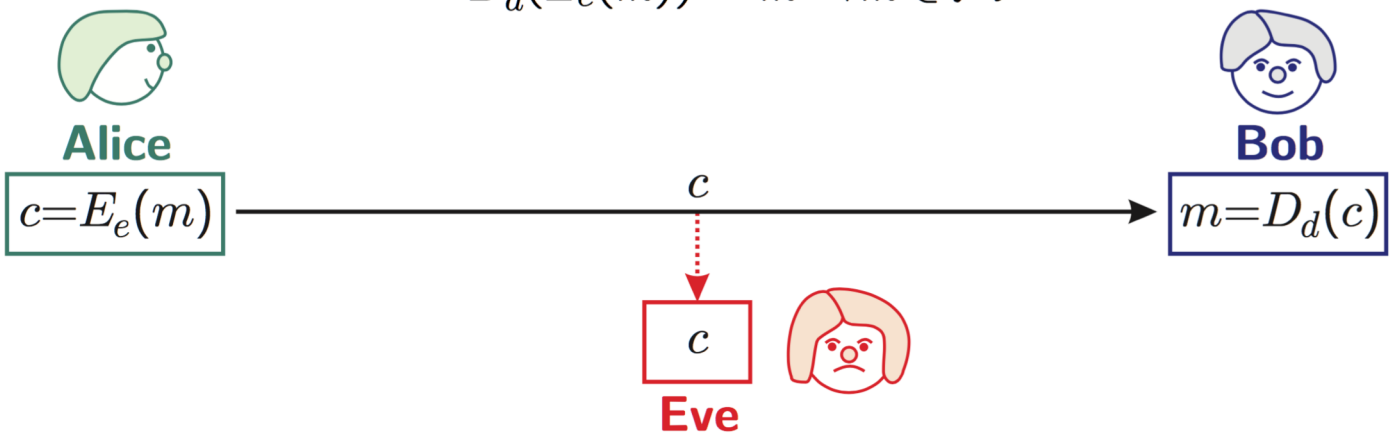
Notations

- A, the **alphabet of definition**, eg. $A = \{0,1\}$.
- M is the **message space**; set of strings over A, $M \subseteq A^*$.
- C is the **ciphertext space**; set of strings over A'.
- K denotes the **key space**; each $e \in K$ uniquely determines a bijection.
- m is a **plaintext**, $m \in M$.
- $E_e: M \to C$ is the **encryption function** (bijection).
- $\mathrm{D_d}: \mathrm{C} \to \mathrm{M}$ is the **decryption function** (bijection).
- Applying E_e is called **encryption**.
- Applying D_d is called **decryption**.

Notations

An **encryption scheme** (or **cipher**) consist of a set $\{E_e : e \in K\}$ and a corresponding set $\{D_d : d \in K\}$ with the property that $\forall e \in K \; \exists$ a unique $d \in K$ s.t. $D_d : E_e^{-1}$; i.e.,

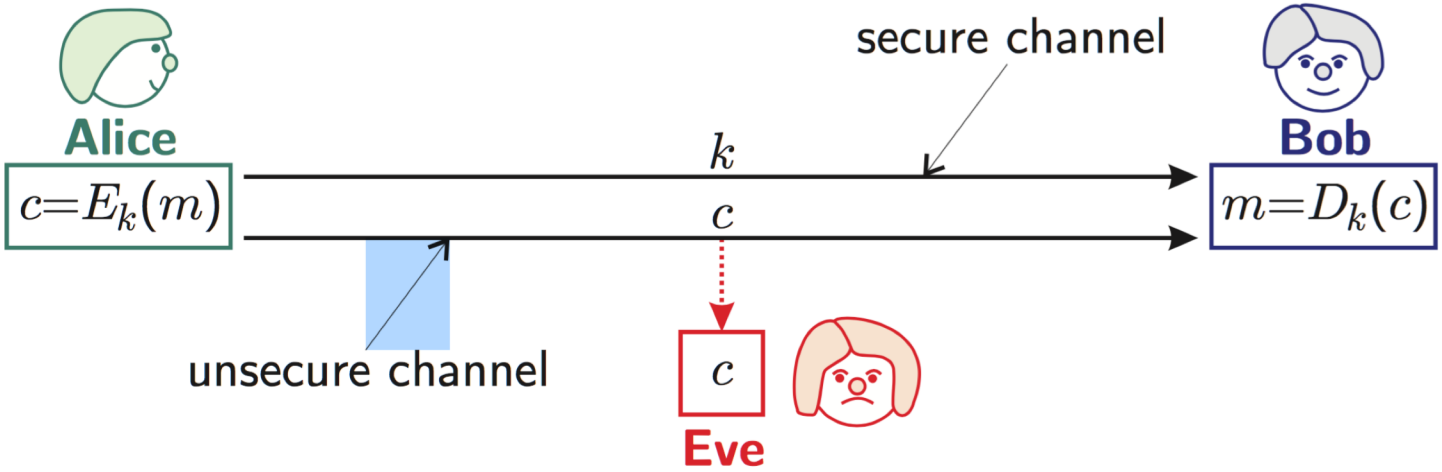
$$D_d(E_e(m)) = m \quad \forall m \in \mathcal{M}$$



To **construct** an encryption scheme requires fixing a message space M, a cipher space C, and a key space K, as well as encryption transformation $\{E_e : e \in K\}$ and corresponding $\{D_d : d \in K\}$.

Symmetric-Key Encryption

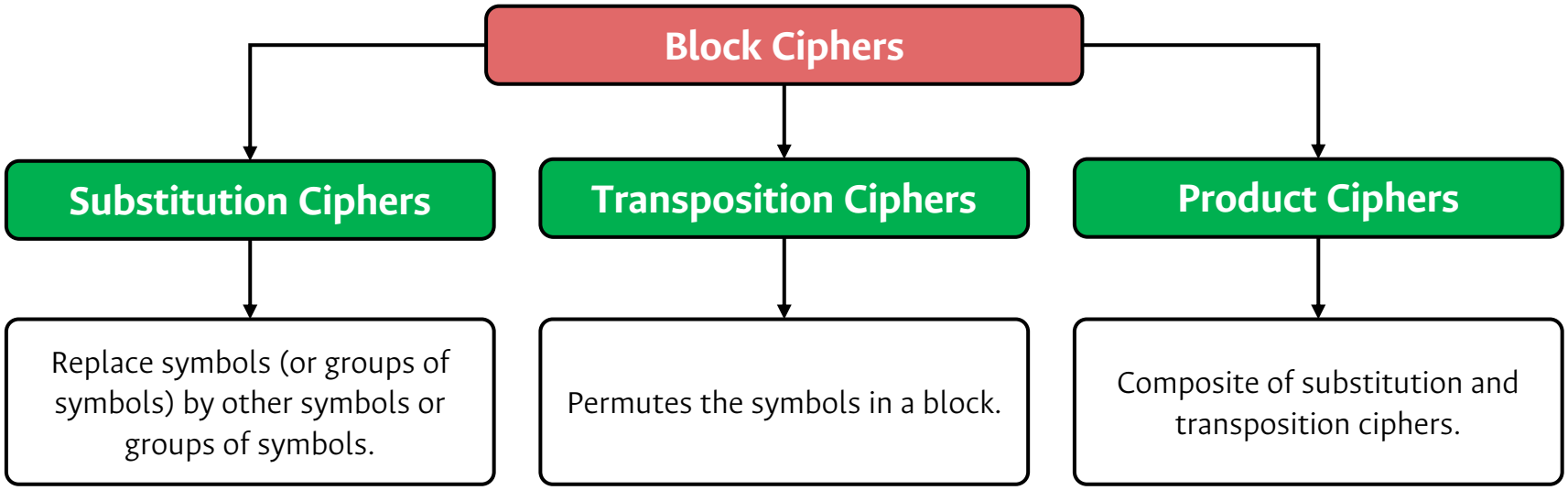
Symmetric-key encryption uses the same key (or are easily derived from each other) to encrypt and decrypt. $e = d = k$.



Block Ciphers

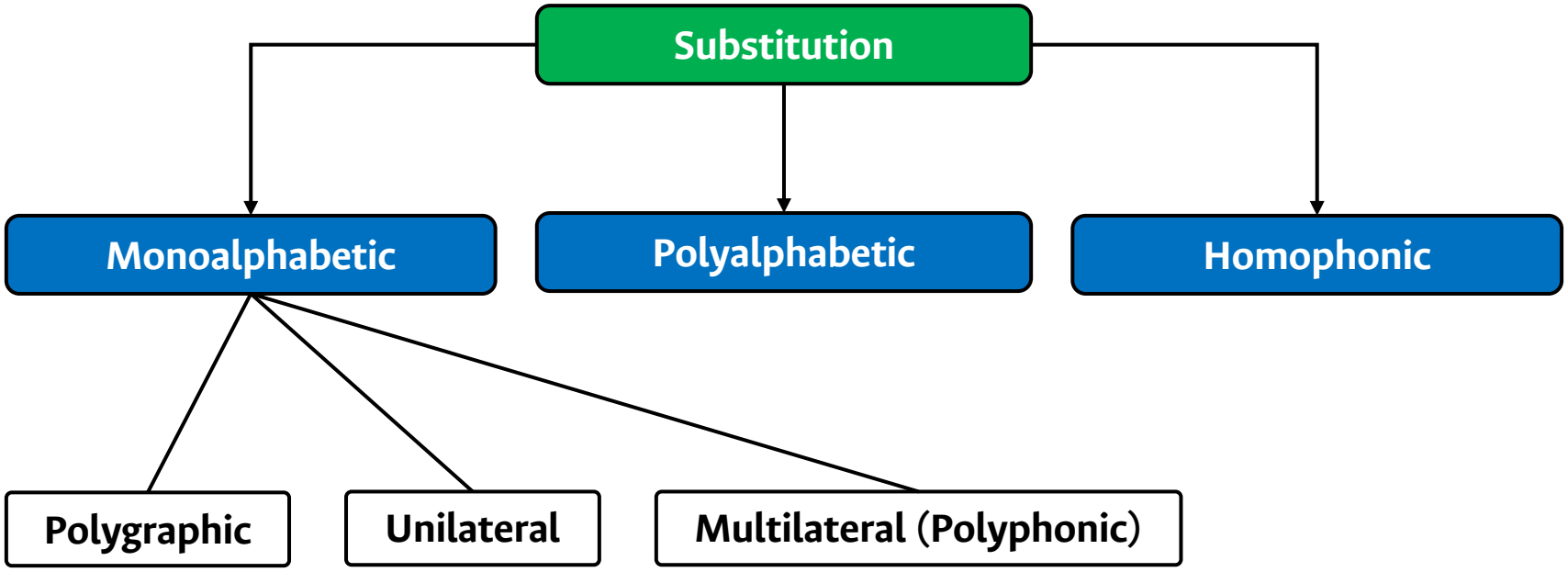
Block Ciphers

A **block cipher** is a symmetric-key encryption scheme that breaks up the plaintext message into strings (**blocks**) of a fixed length t over an alphabet A and encrypts one block at a time.



Substitution Ciphers

A **substitution cipher** is a **cipher** that replaces each plaintext symbol (or group of symbols) with another ciphertext symbol.



Monoalphabetic Ciphers

(Substitution Ciphers)

The main idea of these ciphers is an one by one substitution of a symbol from plaintext to corresponding symbol in ciphertext.

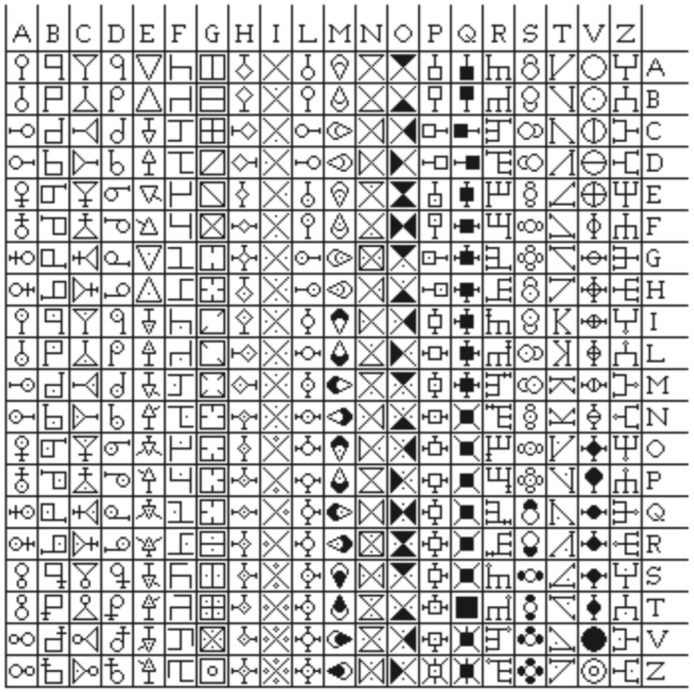
There are three types of monoalphabetic ciphers:

- **Polygraphic:** the ciphertext units are consistently more than one plaintext letter long.
- **Unilateral:** the ciphertext unit is always one character long.
- **Multilateral (Polyphonic):** the ciphertext unit is more than one character in length. The ciphertext characters may be letters, numbers, or special characters.

Porta Cipher

Substitution + Monoalphabetic + Polygraphic

Based on the following **20×20 tableau** filled with **400** unique glyphs:



Playfair Cipher

Substitution + Monoalphabetic + Polygraphic

To **encipher**, pick a keyword and write it into a **5×5 square**, omitting repeated letters and combining **I** and **J** in one cell.

We need to break the plaintext up into two-letter groups. If letters in a pair are the same, insert **X** between them. If there is only one letter in the last group, add **X** to it. To encrypt find each two-letter group in the square and if they are:

- **In the same column:** use the letter below it as the cipher text.
- **In the same row:** use the letter to the right as the cipher text.
- **Neither:** each letter is exchanged with the letter at the intersection of its own row and the other column.

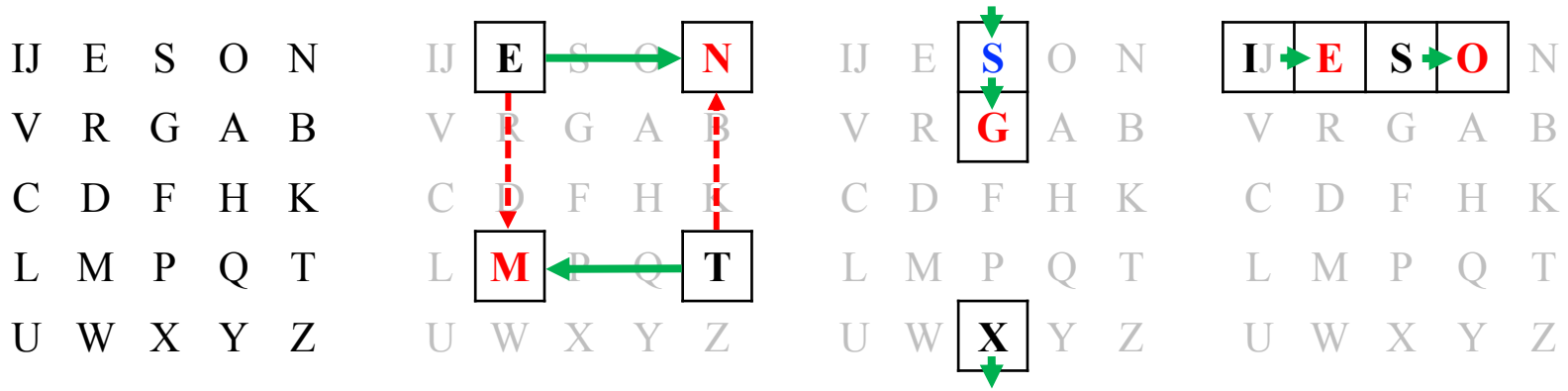
For **deciphering**, the rules are the exact opposite.

Playfair Cipher

Substitution + Monoalphabetic + Polygraphic

Example: encrypt the message "THIS SECRET MESSAGE IS ENCRYPTED".

TH IS SE CR ET ME SS AG EI SE NC RY PT ED



TH IS SE CR ET ME SX SA GE IS EN CR YP TE DX
QK EO OS DV NM WR GS OG RS EO SI DV XQ MN FW

Caesar's Cipher

Substitution + Monoalphabetic + Unilateral

- $A = \{A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z\}$
- $M = C =$ strings of length **5**
- $k =$ permutations of **3**

Example: cipher the message "RETURN TO ROME".

Plaintext	=	RETUR NTORO ME
Ciphertext	=	UHWXU QWRUR PH

Rot13 Cipher

Substitution + Monoalphabetic + Unilateral

Replaces each letter with the letter thirteen places down the alphabet. ***A*** becomes ***N***, ***B*** becomes ***O*** and so on.

Examples:

aha → nun	ant → nag	balk → onyx	bar → one
barf → ones	be → or	bin → ova	ebbs → roof
envy → rail	er → re	errs → reef	flap → sync
fur → she	gel → try	gnat → tang	irk → vex

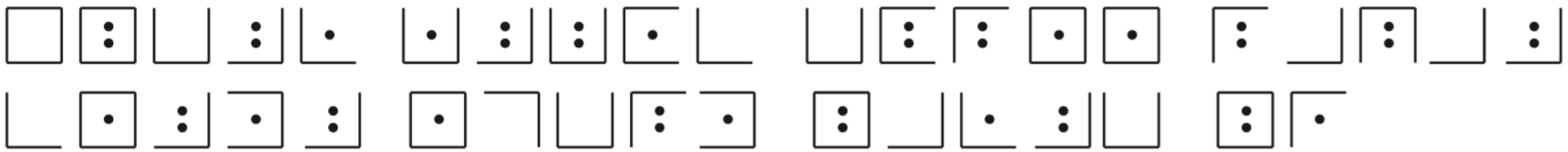
Dots Cipher

Substitution + Monoalphabetic + Unilateral

Symbols in the plaintext will be replaced by **squares** with or without points and with or without **surrounding** lines using the following rules:

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

Example: encrypt the message "WE TALK ABOUT FINNISH SAUNA MANY TIMES LATER".

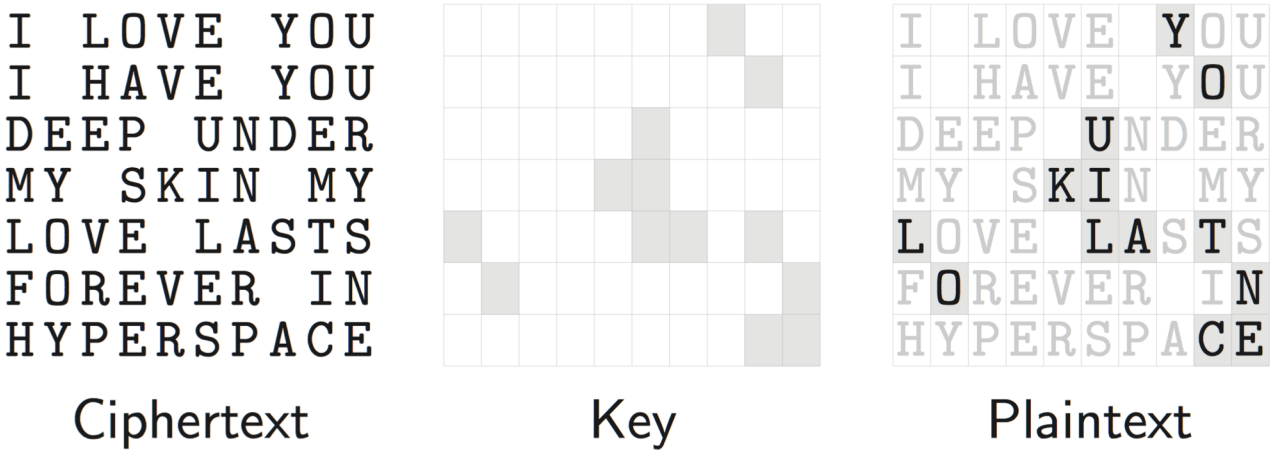


Garbage-in-Between Cipher

Substitution + Monoalphabetic + Multilateral

The message is written in positions determined by the key. After that, empty positions are filled by **arbitrary letters**, in a misleading way so that the whole message looks like a different story.

Example:



Polyalphabetic Ciphers
(Substitution Ciphers)

The main idea behind the **polyalphabetic** cipher is that a single symbol can be encrypted to several different symbols instead of just one.

The reason behind using polyalphabetic ciphers is to flatten frequency distributions.

Vigènere Cipher

Substitution + Polyalphabetic

Based on the **tableau** shown below and the use of a **keyword**.

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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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

Vigènere Cipher

Substitution + Polyalphabetic

Example: encrypt the message "TO BE OR NOT TO BE THAT IS THE QUESTION", using the keyword "RELATIONS" and $t = 5$.

Keyword	=	RELAT	IONSR	ELATI	ONSRE	LATIO	NSREL
Plaintext	=	TOBEO	RNOTT	OBETH	ATIST	HEQUE	STION
Ciphertext	=	KSMEH	ZBBLK	SMEMP	OGAJX	SEJCS	FLZSY

Hill Cipher

Substitution + Polyalphabetic

Each letter is treated as a digit in base 26: $A = 0$, $B = 1$, and so on.

Let $M = C = (\mathbb{Z}^{26})^t$

The idea is to take t linear combinations of the t alphabetic characters in one plaintext element, thus producing the m alphabetic characters in one cipher element.

For example, if $t = 2$, we could write the plaintext element as $m = (m_1, m_2)$ and a ciphertext element as $c = (c_1, c_2)$. Here, c_1 would be a linear combination of m_1 and m_2, as would c_2. We might take:

$$c_1 = Wm_1 + Xm_2$$
$$c_2 = Ym_1 + Zm_2$$

Hill Cipher

Substitution + Polyalphabetic

Wich can be written more succinctly in matrix notation as follows:

$$(c_1, c_2) = (m_1, m_2)\begin{pmatrix} W & Y \\ X & Z \end{pmatrix}$$

We use the inverse matrix K^{-1} to decrypt. The ciphertext is decrypted using the formula $m = cK^{-1}$.

Hill Cipher

Substitution + Polyalphabetic

Example: encrypt the message "JULY" using $K = \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix}$

$(9 \quad 20) \rightarrow$ "JU" and $(11 \quad 24) \rightarrow$ "LY", then

$$(9 \quad 20)\begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} = (99 + 60 \quad 72 + 140) = (159 \quad 212) = (3 \quad 4)$$

and

$$(11 \quad 24)\begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} = (121 + 72 \quad 88 + 168) = (193 \quad 256) = (11 \quad 22)$$

Hence the encrypted message of "JULY" is "DELW".

Class Activity

Instructions

Decrypt the message previously encrypted with the **Hill Cipher**.

Time: 20 minutes.

Homophonic Ciphers

(Substitution Ciphers)

Homophonic substitution ciphers are ciphers in which several different ciphertext letters can be used to stand for some plaintext letters.

Typically in a homophonic substitution cipher, the cipher alphabet (the symbols which can appear in a ciphertext) is larger than the plaintext alphabet.

Coding symbols are assigned to each plain letter based on their frequency.

Homophonic Cipher

Substitution + Homophonic

Example: encrypt the message "CRYPTO IS FUN".

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
9,12,33,47,53,67,78,92	48,81	13,41,62	1,3,45,79	14,16,24,44,46,55,57,64,74,82,87,98	10,31	6,25	23,39,50,56,65,68	32,70,73,83,88,93	15	4	26,37,51,84	22,27	18,58,59,66,71,91	0,5,7,54,72,90,99	38,95	94	29,35,40,42,77,80	11,19,36,76,86,96	17,20,30,43,49,69,75,85,97	8,61,63	34	60,89	28	21,52	2

Ciphertext : 62 40 21 95 69 90 32 19 31 61 91

Transposition Ciphers

A **transposition ciphers** are where the letters are jumbled up together. Instead of replacing characters with other characters, this cipher just changes the order of the characters.

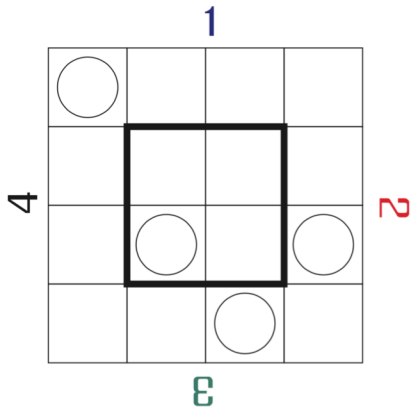
Mathematically a **bijective function** *is used on the characters' positions to encrypt and an inverse function to decrypt.*

Turning Grille Cipher

Transposition

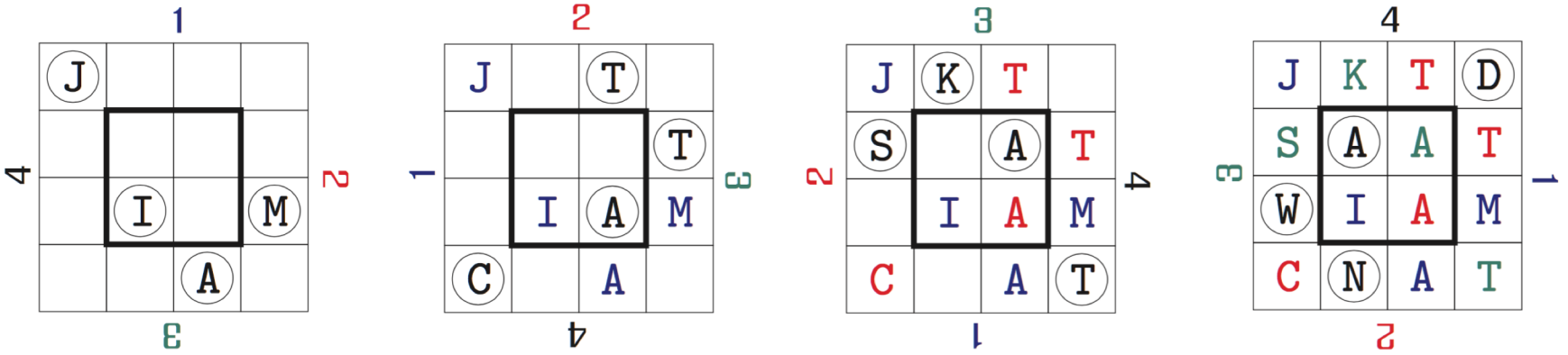
This is a square piece of **cardboard** with holes in it such that each cell in the square appears in **no more than one position** when the grille is rotated to each of its four positions.

Example: encrypt "JIM ATTACKS AT DAWN" using the 4×4 grille.



Turning Grille Cipher

Transposition

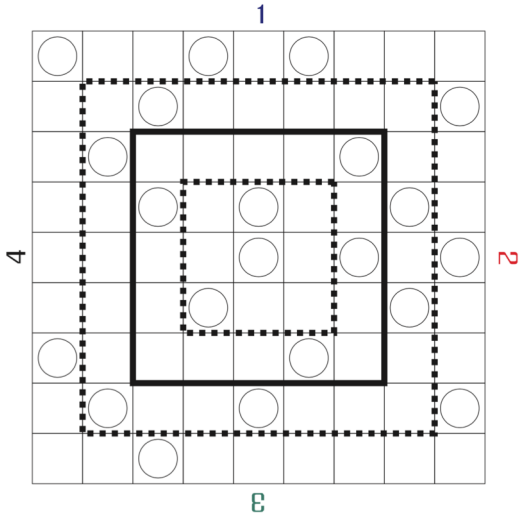


JIMA TTAC KSAT DAWN
JKTD SAAT WIAM CNAT

Class Activity

Instructions

Decrypt the ciphertext "**TESHN INCIG LSRGY LRIUS PITSA TLILM REENS ATTOG SIAWG IPVER TOTEH HVAEA XITDT UAIME RANPM TLHIE I**", using the following grille:



Time: 30 minutes.

Product Ciphers

A **product cipher** is a composite of substitution (confusion) and transposition (diffusion) ciphers.

- **Diffusion:** refers to the dissipation of the statistical properties of the plaintext.
- **Confusion:** makes relationship between ciphertext and key as complex as possible.

- 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**.

Lucifer Cipher

Product

This system uses **permutations** (**transpositions**) on large blocks for the mixing transformation, and substitution on small blocks for confusion. Note that there is no key, thus the system is insecure.

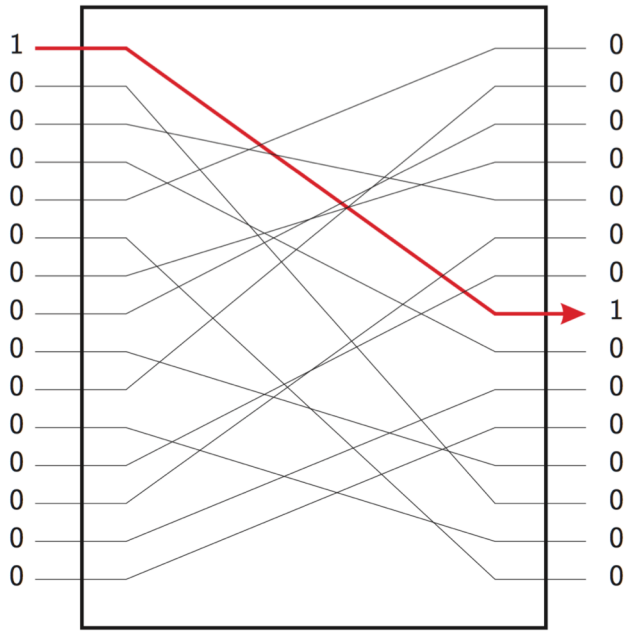
Since this system was set up in hardware, they called the chips which did the permutations **P-Boxes**, and those that did the substitution **S-Boxes**.

Lucifer Cipher

Product

P-Box Example

A **P-Box** (**Permutation-Box**) is a box which permutes the bits of its input.

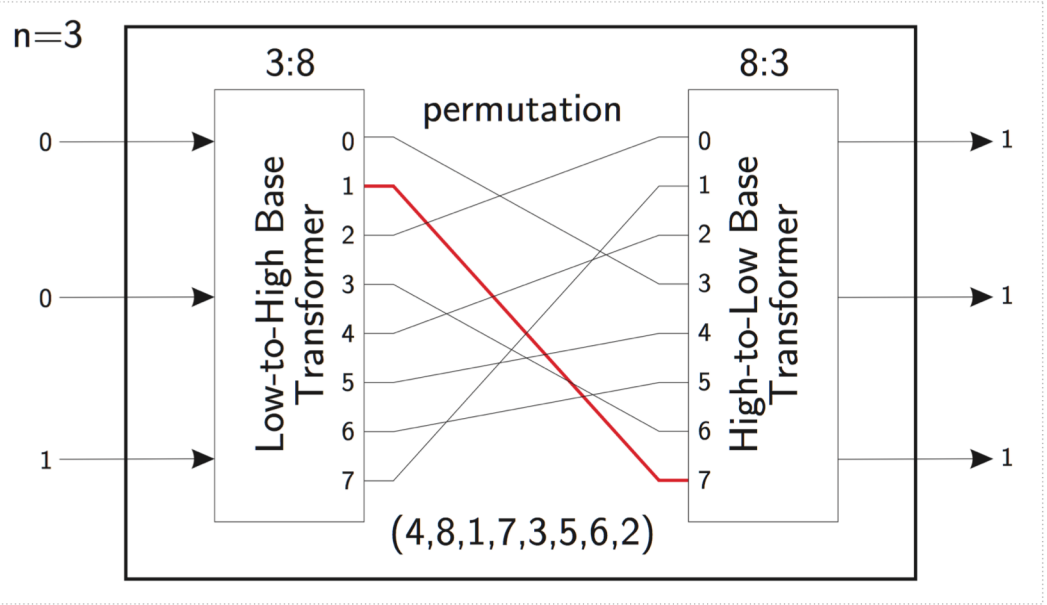


Lucifer Cipher

Product

S-Box Example

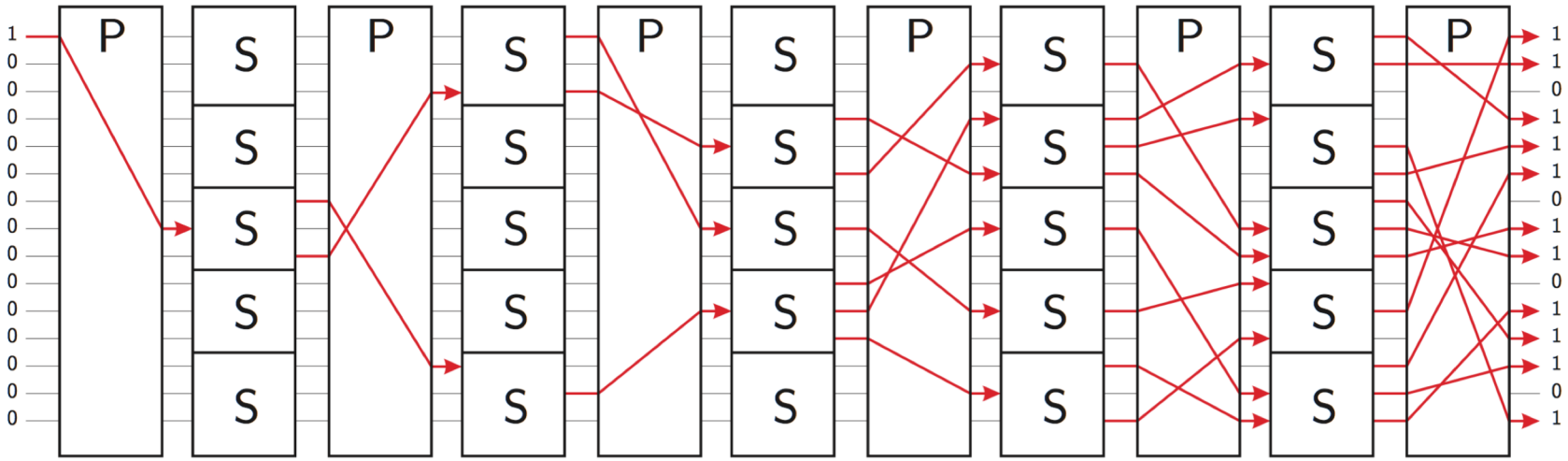
The **S-Box** (**Substitution-Box**) is a hardware device which encodes n bit numbers to other n bit numbers.



Lucifer Cipher

Product

Example



Hayhanen Cipher

Product

Example: encrypt "**Feb/4/have contacted fred. Will consult before investing in coverbusiness.nat**" using the key '**oriental**'.

Feb/4/havecontactedfred.Willconsultbeforeinvestingincoverbusiness.nat

First you set up a keysquare, some letters code to 1 digit, some to 2. Note that the numbered rows use the numbers that are not used in the first row. We substitute all the letters according to this keysquare:

	3	9	6	4	0	8	5	1	7	2
	o	r	i	e	n	t	a	l		
7	b	c	d	f	g	h	j	k	m	p
2	q	s	u	v	w	x	y	z	/	.

Hayhanen Cipher

Product

	3 9 6 4 0 8 5 1 7 2
	o r i e n t a l
7	b c d f g h j k m p
2	q s u v w x y z / .

F eb / 4 / h av ec ontac ted f red . W illc ons u ltb ef oreinv es
744732744427785244793085798476749476222061179302926187347439460244 29

ting inc ov erb u s ines s . nat
860706079324497326296042929220 58

Hayhanen Cipher

Product

Next, we write these out in a block for transposition as follows:

4	7	5	8	2	9	6	1	3
7	4	4	7	3	2	7	4	4
4	2	7	7	8	5	2	4	4
7	9	3	0	8	5	7	9	8
4	7	6	7	4	9	4	7	6
2	2	2	0	6	1	1	7	9
3	0	2	9	2	6	1	8	7
3	4	7	4	3	9	4	6	0
2	4	4	2	9	8	6	0	7
0	6	0	7	9	3	2	4	4
9	7	3	2	6	2	9	6	0
4	2	9	2	9	2	2	0	5
8								

Next, we read down each column in turn, starting with column '**1**'. We copy the numbers into another table, using a route cipher: We fill in the area to the left of the line (in **red**) first, then the areas to the right.

Hayhanen Cipher

Product

Note how the line is formed: we start to the left of column **1**, and **zig-zag** down to the right. When we reach the edge, we jump to the left of column **2** and **zig-zag** again, and so on.

4	7	5	8	2	9	6	1	3
7	4	4	7	3	2	7	4	4
4	2	7	7	8	5	2	4	4
7	9	3	0	8	5	7	9	8
4	7	6	7	4	9	4	7	6
2	2	2	0	6	1	1	7	9
3	0	2	9	2	6	1	8	7
3	4	7	4	3	9	4	6	0
2	4	4	2	9	8	6	0	7
0	6	0	7	9	3	2	4	4
9	7	3	2	6	2	9	6	0
4	2	9	2	9	2	2	0	5
8								

4	6	2	1	5	3
4	4	9	6	2	9
7	7	8	6	2	4
0	4	6	0	3	2
8	8	4	6	2	3
9	9	9	7	2	0
6	9	4	4	4	6
4	8	6	9	7	2
7	0	7	4	0	7
5	7	4	7	4	2
3	3	2	0	9	7
4	8	4	7	3	6
0	7	0	9	4	2
2	7	2	2	2	5
2	7	5	9	1	6
4	0	3	9	8	3
9	7	2	7	2	2
4	1	1	4		

Hayhanen Cipher

Product

Once we have completed this step, we transcribe the numbers by columns in groups of 5:

4	6	2	1	5	3
4	4	9	6	2	9
7	7	8	6	2	4
0	4	6	0	3	2
8	8	4	6	2	3
9	9	9	7	2	0
6	9	4	4	4	6
4	8	6	9	7	2
7	0	7	4	0	7
5	7	4	7	4	2
3	3	2	0	9	7
4	8	4	7	3	6
0	7	0	9	4	2
2	7	2	2	2	5
2	7	5	9	1	6
4	0	3	9	8	3
9	7	2	7	2	2
4	1	1	4		

```
66067 49470 79299 74986 49467 42402 53219 42306 27276 25632
47089 64753 40224 94223 22470 49342 18247 48998 07387 77071
```

Modes of Operation

Modes of Operation

A **block cipher** (such as DES) encrypts plaintext in fixed-size n-bit blocks (often $n = 64$). For messages exceeding n bits we should use the following standard methods (**modes of operation**):

1. **ECB**: Electronic Code Block
2. **CBC**: Cipher-Block Chaining
3. **CFB**: Cipher FeedBack
4. **OFB**: Output FeedBack
5. **CTR**: CounTeR

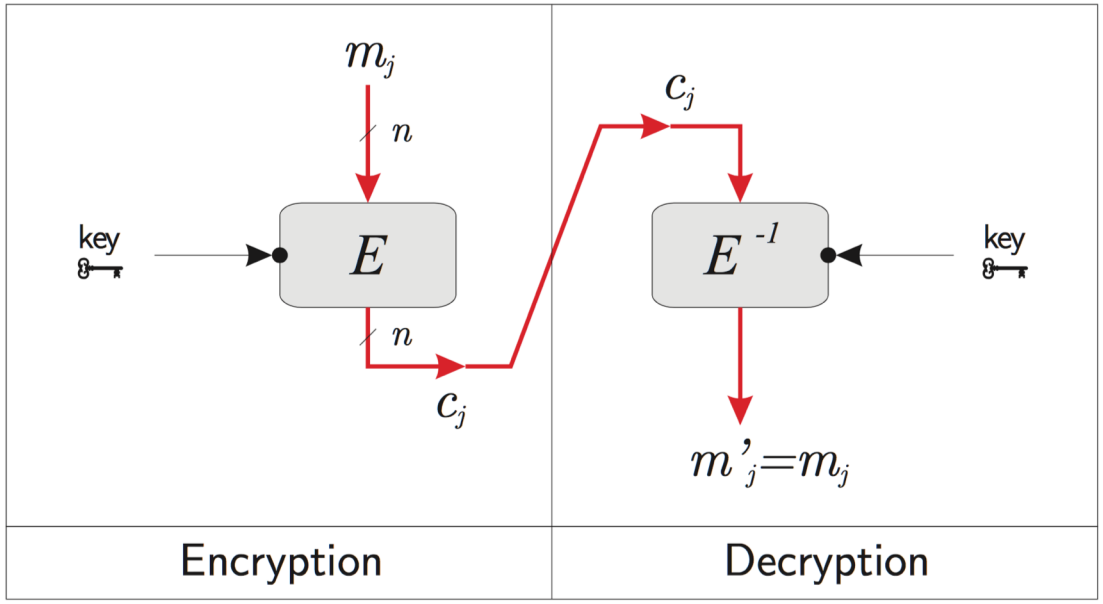
Modes of Operation

Notation

- $\boldsymbol{E_k}$: Encryption function E using key k
- $\boldsymbol{E_k^{-1}}$: Decryption function E^{-1} using key k
- $\boldsymbol{m}$: Plain message $m = m_1, m_2, \ldots, m_t$
- $\boldsymbol{k}$: The key
- $\boldsymbol{IV}$: Initialization vector
- $>>$: Shift to the right
- $(\rightsquigarrow)$: Encryption
- $(\leftsquigarrow)$: Decryption

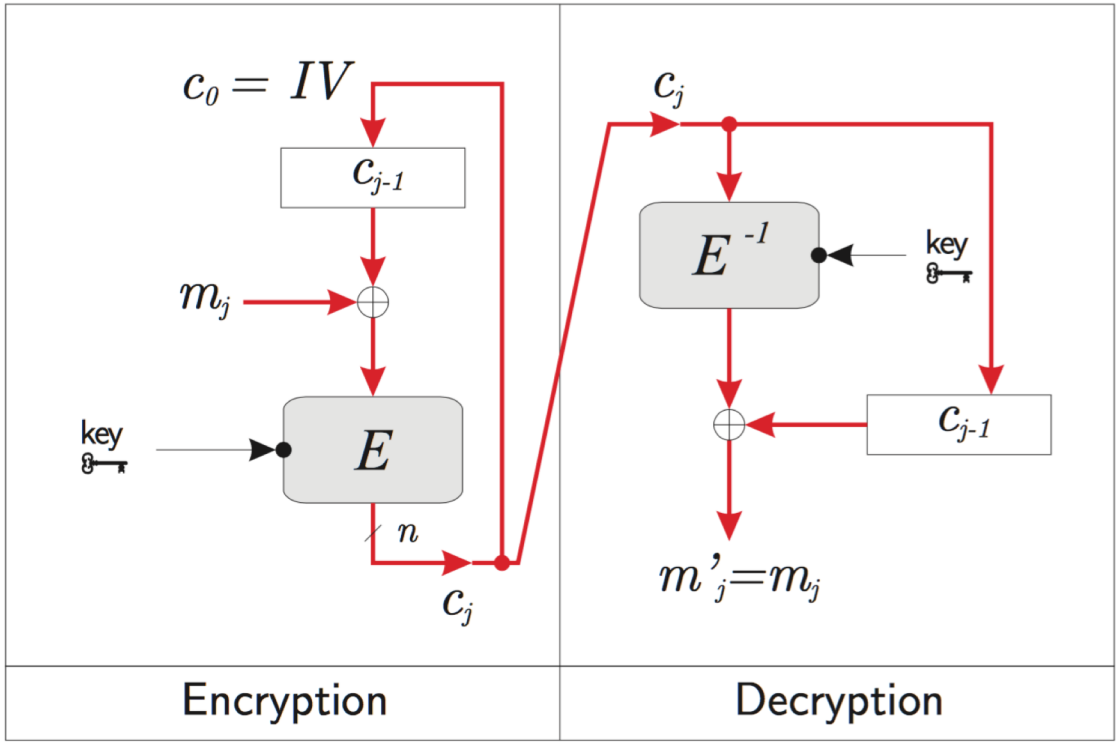
ECB: Electronic Code Block

Mode of Operation



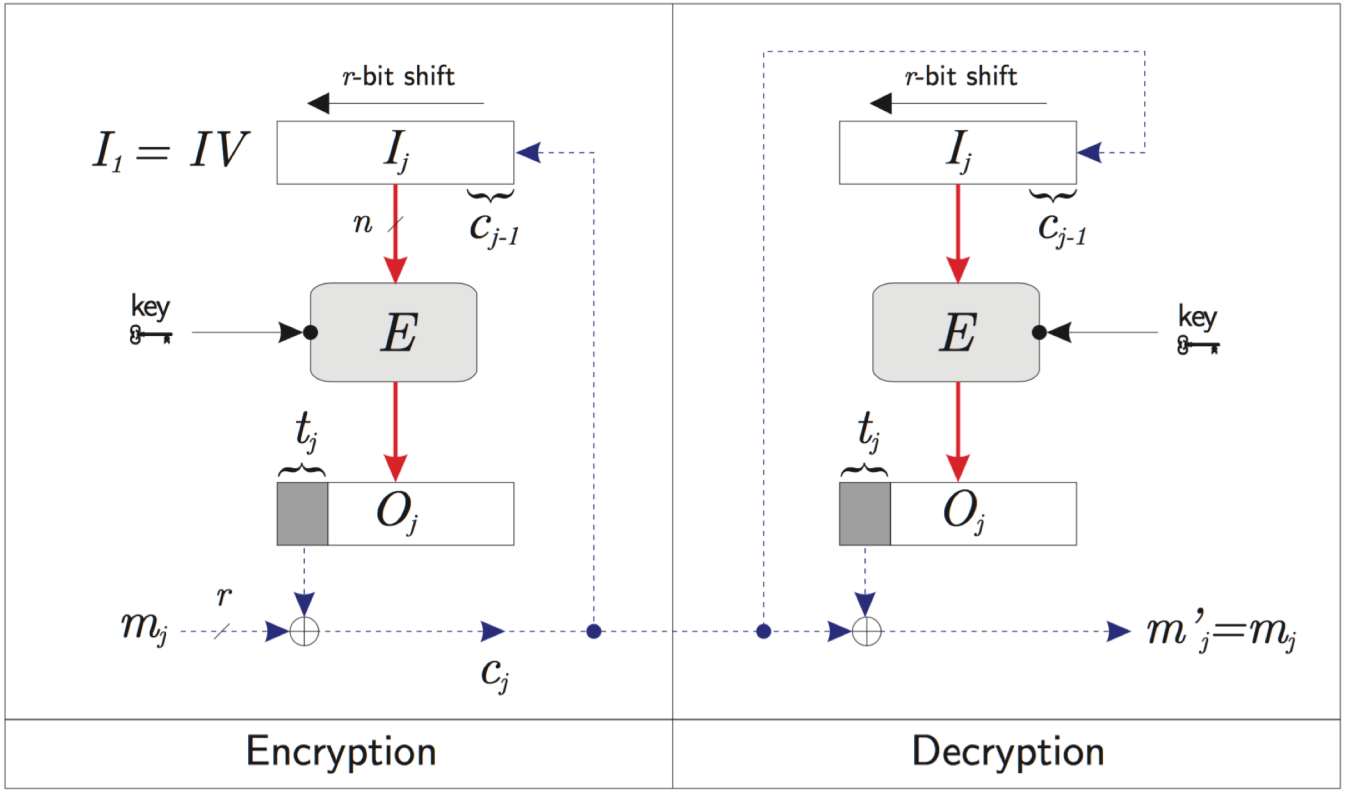
CBC: Cipher-Block Chaining

Mode of Operation



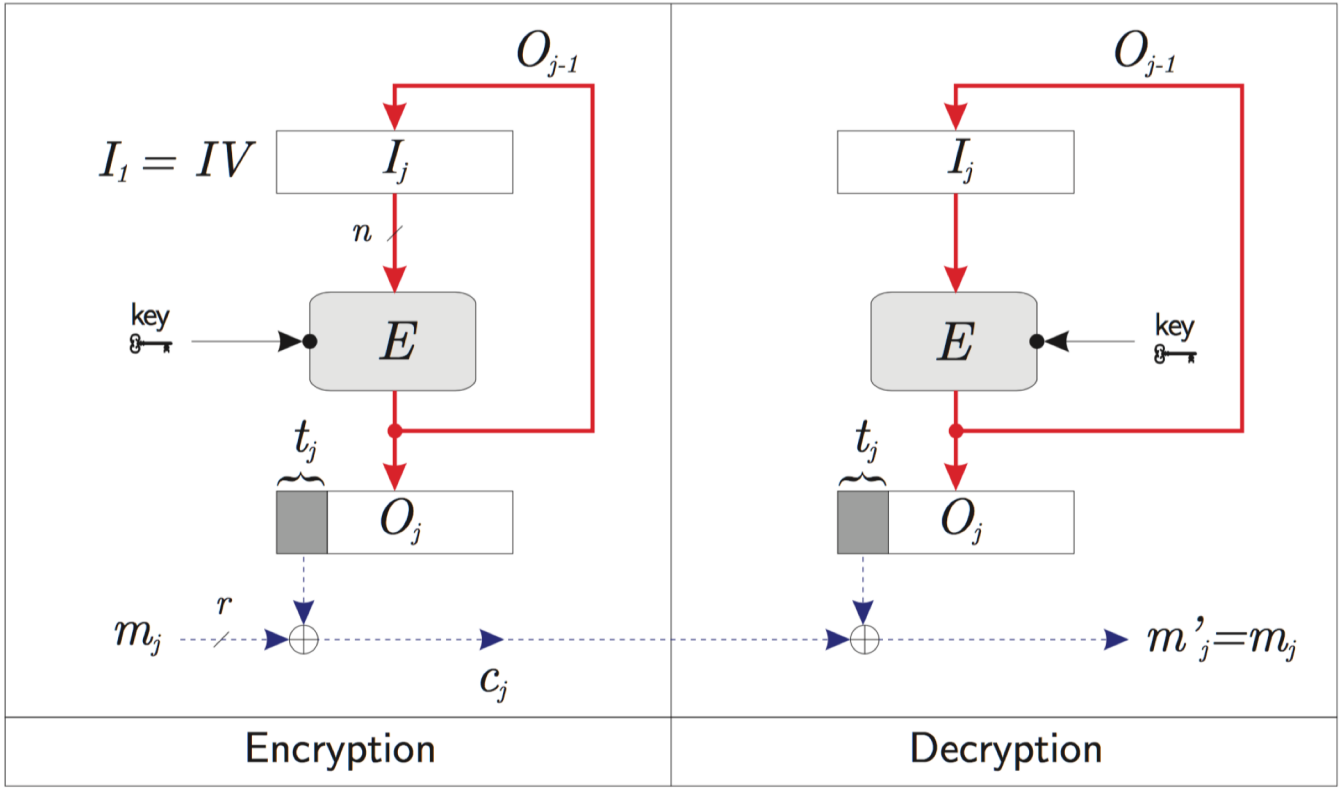
CFB: Cipher Feedback

Mode of Operation



OFB: Output Feedback

Mode of Operation



CTR: CounTeR

Mode of Operation

$\underline{\texttt{INPUT}}$: $m,\ k,\ IV = nonce \quad \{m = m_1, m_2, \ldots, m_t,\ |m_i| = n,\ |IV| = n\}$

1 $(\curvearrowright)$

$$\left.\begin{aligned} I_1 &\leftarrow IV \\ O_j &\leftarrow E_k(I_j) \\ c_j &\leftarrow m_j \oplus O_j \\ I_j &\leftarrow I_j + 1 \end{aligned}\right\} \quad \forall j \in \{1 \ldots u\}$$

2 $(\curvearrowleft)$

$$\left.\begin{aligned} I_1 &\leftarrow IV \\ O_j &\leftarrow E_k(I_j) \\ m_j &\leftarrow c_j \oplus O_j \\ I_j &\leftarrow I_j + 1 \end{aligned}\right\} \quad \forall j \in \{1 \ldots u\}$$

Bibliography

- [**PINZÓN**] Y. Pinzón, Cryptography. 2014.
- [**DELFS**] H. Delfs and H. Knebl, Introduction to Cryptography, 3rd ed. 2015.