Lab work no. 1

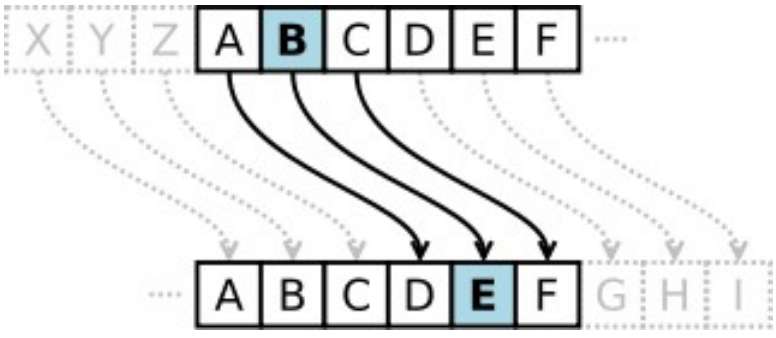
1.1 Caesar's cipher

Caesar's cipher: In this cipher, each letter of the plaintext is replaced by a new letter obtained by an alphabetic shift. The secret key k, which is the same for encryption as for decryption, consists of the number indicating the alphabetic shift, i.e. $k \in \{1, 2, 3,..., n-1\}$, where n is the length of the alphabet. The encryption and decryption of the Caesar cipher message can be defined by the formulas

$$c = e_k(x) = x + k \pmod{n},$$
$$m = d_k(y) = y - k \pmod{n},$$

where x and y are the numeric representation of the respective character of the plaintext. The function called Modulo (a mod b) returns the remainder of the integer a divided by the integer b. This encryption method is named after Julius Caesar, who used it to communicate with his generals using the key k = 3 (Table 5.1).

For example, for k = 3 we have



$$e_k(\text{S}) = 18 + 3 \pmod{26} = 21 = \text{V}$$
$$d_k(\text{V}) = 21 - 3 \pmod{26} = 18 = \text{S}$$

In this case for *m* =„*cifrul cezar*”, we get *c* =„*fliuxo fhcdu*”.

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

Table 1. Caesar cipher with key k=3

The Caesar cipher is very easy to crack, so it is a very weak cipher. So a cryptanalyst can get the plaintext by trying all 25 keys. It is not known how useful the Caesar cipher was at the time when it was used by its namesake, but it is likely that it was quite secure, as long as only a few of Caesar's enemies were able to write and read, let alone know cryptanalytic concepts.

Task 1.1. ***Implement the Caesar algorithm for the English alphabet in one of the programming languages. Use only the letter encodings as shown in Table 1 (encodings specified in the programming language, e.g. ASCII or Unicode, are not allowed to be used). Key values shall be between 1 and 25 inclusive and no other values are allowed. Text character values shall be between 'A' and 'Z', 'a' and 'z' and no other values are prefixed. If the user enters other values - the user will be prompted for the correct slide. Before encryption the text will be converted to upper case and spaces will be removed. The user will be able to choose the operation - encryption or decryption, enter the key, message or cryptogram and get the cryptogram or decrypted message respectively.***

1.1. Caesar's number + permutation

Given the low encryption strength of the Caesar cipher, primarily due to the keyspace, which consists of only 25 different keys for the Latin alphabet, it can be cracked by trying all the keys consecutively. If the message has been encrypted with the Caesar cipher, then one of the keys will give readable text in the language in which the message was written.

For example, if

$\boldsymbol{m}$ = BRUTE FORCE ATTACK

is a message written in English and has been encrypted with the key

$$\boldsymbol{k} = \mathbf{17},$$

get the cryptogram

$$\boldsymbol{c} = \text{SILKVWFITVRKKRTB}$$

If the cryptanalyst intercepts the encrypted message and traverses all keys 1, 2, ..., 25 - it will get the following:

1	RHKJUVEHSUQJJQSA
2	QGJITUDGRTPIIPRZ
3	PFIHSTCFQSOHHOQY
4	OEHGRSBEPRNGGNPX
5	NDGFQRADOQMFFMOW
6	MCFEPQZCNPLEELNV
7	LBEDOPYBMOKDDKMU
8	KADCNOXALNJCCJLT
9	JZCBMNWZKMIBBIKS
10	IYBALMVYJLHAAHJR
11	HXAZKLUXIKGZZGIQ
12	GWZYJKTWHJFYYFHP
13	FVYXIJSVGIEXXEGO
14	EUXWHIRUFHDWWDFN
15	DTWVGHQTEGCVVCEM
16	CSVUFGPSDFBUUBDL
17	BRUTEFORCEATTACK
18	AQTSDENQBDZSSZBJ
19	ZPSRCDMPACYRRYAI
20	YORQBCLOZBXQQXZH
21	XNQPABKNYAWPPWYG
22	WMPOZAJMXZVOOVXF
23	VLONYZILWYUNNUWE
24	UKNMXYHKVXTMMTVD
25	TJMLWXGJUWSLLSUC

As you can see - only the text obtained by using the key k=17 is meaningful in English, so the corresponding cryptogram message is

m = BRUTEFORCEATTACK.

To enhance the crypto-resistance of the Caesar cipher, a permutation of the alphabet can be applied by applying a keyword (not to be confused with the cipher's base key). This key can be any letter sequence of the alphabet - either a vocabulary word or a nonsense word.

Let the second key be **k2=cryptography**. We apply this key to the alphabet

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

and we get:

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

We obtained this new order by placing the letters of k2 first, then the other letters of the alphabet follow in their natural order. We will keep in mind that the letters will not repeat, i.e. if the letter occurs a few times, it is placed only once.

The Caesar cipher is then applied, taking into account the new alphabetical order:

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
C	**R**	**Y**	**P**	**T**	**O**	**G**	**A**	**H**	B	D	E	F	I	J	K	M	L	N	Q	S	U	V	W	X	Z
3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	0	1	2
P	**T**	**O**	**G**	**A**	**H**	B	D	E	F	I	J	K	M	L	N	Q	S	U	V	W	X	Z	**C**	**R**	**Y**

Table 2. Caesar cipher with key k1 =3 and k2 =cryptography

Since there are 26!= 403291461126605635584000000, the number of keys for this version of the algorithm will be

$$26! \cdot 25 = 10082286528165140889600000000,$$

which complicates exhaustive cracking, but does not save us from frequency analysis attacks.

Task 1.2. Implement the Caesar algorithm with 2 keys, preserving the conditions expressed in Task 1.1. In addition, key 2 must contain only letters of the Latin alphabet, and have a length of not less than 7.