

Кодирование текстовой информации

А 128	Б 129	В 130	Г 131	Д 132	Е 133	Ж 134	З 135	И 136	Й 137	К 138	Л 139	М 140	Н 141	О 142	П 143
Р 144	С 145	Т 146	У 147	Ф 148	Х 149	Ц 150	Ч 151	Ш 152	Щ 153	Ъ 154	Ы 155	Ь 156	Э 157	Ю 158	Я 159
† 160	° 161	Ы 162	£ 163	§ 164	• 165	¶ 166	Ы 167	® 168	© 169	™ 170	Á 171	á 172	ê 173	à 174	è 175
Ё 176	± 177	≤ 178	≥ 179	э 180	μ 181	г 182	э 183	Ю 184	ю 185	Я 186	я 187	É 188	é 189	Й 190	ò 191
Э 192	Ю 193	¬ 194	√ 195	f 196	≈ 197	Δ 198	« 199	» 200	… 201	nbsp 202	Ó 203	ó 204	Й 205	Й 206	Я 207
- 208	— 209	“ 210	” 211	‘ 212	’ 213	÷ 214	” 215	ù 216	Ы 217	У 218	У 219	№ 220	Ё 221	ё 222	я 223
а 224	б 225	в 226	г 227	д 228	е 229	ж 230	з 231	и 232	й 233	к 234	л 235	м 236	н 237	о 238	п 239
р 240	с 241	т 242	у 243	ф 244	х 245	ц 246	ч 247	ш 248	щ 249	ъ 250	ы 251	ь 252	э 253	ю 254	я 255

[illegible]

Историческая справка

*Пайцике тсут т "камацамлтой чмароке" —
кайтонили, нмирепяшвейля ш Молли ися
цинсоракигелтой неменилти*

Знаки кода Морзе	Буквы		Знаки кода Морзе	Циф- ры	Знаки кода Морзе	Знаки препинания и служебные сигналы
	Рус.	Анг.				
· —	А	Aa	· — · —	1	· — · —	(,) запятая
· · —	Б	Bb	· — · — · —	2	· — · — · —	(.) Точка
· · · —	В	Vv	· — · — · — · —	3	· — · — · — · —	(;) точка с запятой
· · · · —	Г	Gg	· — · — · — · — · —	4	· — · — · — · — · —	(:) двоеточие
· · · · · —	Д	Dd	· — · — · — · — · — · —	5	· — · — · — · — · — · —	(?) вопросит. знак
· · · · · · —	Е	Ee	· — · — · — · — · — · — · —	6	· — · — · — · — · — · — · —	(№) номер
· · · · · · · —	Ж	Vv	· — · — · — · — · — · — · — · —	7	· — · — · — · — · — · — · — · —	(") кавычки
· · · · · · · · —	З	Zz	· — · — · — · — · — · — · — · — · —	8	· — · — · — · — · — · — · — · — · —	(') апостроф
· · · · · · · · · —	И	Ii	· — · — · — · — · — · — · — · — · — · —	9	· — · — · — · — · — · — · — · — · — · —	() скобки
· · · · · · · · · · —	К	Kk	· — · — · — · — · — · — · — · — · — · — · —	0	· — · — · — · — · — · — · — · — · — · — · —	(!) восклицательный знак
· · · · · · · · · · · —	Л	Ll	· — · — · — · — · — · — · — · — · — · — · — · —		· — · — · — · — · — · — · — · — · — · — · — · —	(-) тире
· · · · · · · · · · · · —	М	Mm	· — · — · — · — · — · — · — · — · — · — · — · — · —		· — · — · — · — · — · — · — · — · — · — · — · — · —	Ждать
· · · · · · · · · · · · · —	Н	Nn	· — · — · — · — · — · — · — · — · — · — · — · — · — · —		· — · — · — · — · — · — · — · — · — · — · — · — · — · —	Понял
· · · · · · · · · · · · · · —	О	Oo	· — · — · — · — · — · — · — · — · — · — · — · — · — · — · —		· — · — · — · — · — · — · — · — · — · — · — · — · — · — · —	(/) дробная черта
· · · · · · · · · · · · · · · —	П	Pp	· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —		· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —	Знак раздела
· · · · · · · · · · · · · · · · —	Р	Rr	· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —		· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —	Перебой (исправ- ление ошибки)
· · · · · · · · · · · · · · · · · —	С	Ss	· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —		· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —	Сигнал о начале передачи (НП)
· · · · · · · · · · · · · · · · · · —	Т	Tt	· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —		· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —	Сигнал о готовности к приему (ПО)
· · · · · · · · · · · · · · · · · · · —	У	Uu	· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —		· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —	Начало действия
· —	Ф	Ff	· — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · — · —		· — · —	Знак окончания передачи
· —	Х	Hh	· — · —		· — · —	
· —	Ц	Cc	· — · —		· — · —	
· —	Щ	Qq	· — · —		· — · —	
· —	Ы	Yy	· — · —		· — · —	
· —	Ю	Jj	· — · —		· — · —	
· —	Я	Xx	· — · —		· — · —	
· —	Ь	ЬЬ	· — · —		· — · —	
· —	Э	Ee	· — · —		· — · —	

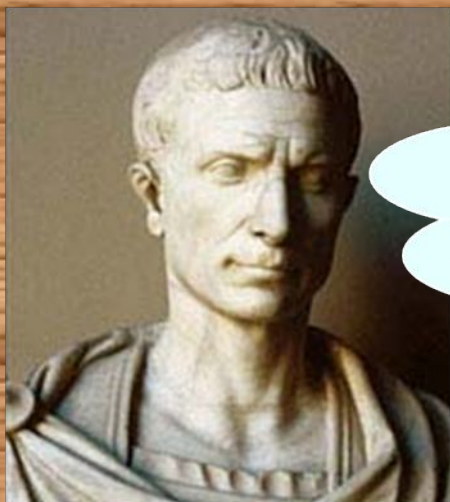
Криптография - это тайнопись, система изменения письма с целью сделать текст непонятным для непосвященных лиц

Азбука Морзе или неравномерный телеграфный код, в котором каждая буква или знак представлены своей комбинацией коротких элементарных посылок электрического тока (точек) и элементарных посылок утроенной продолжительности (тире)



Сурдожесты - язык жестов, используемый людьми с нарушениями слуха

Код Цезаря



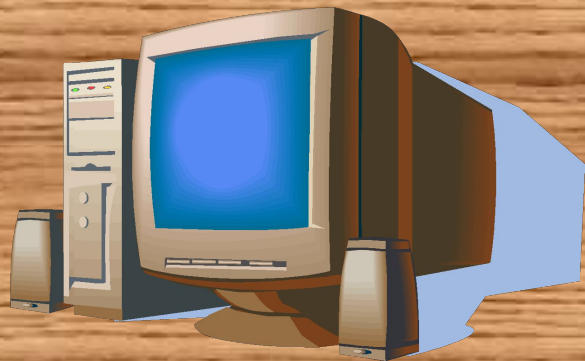
Юлий Цезарь
(I век до н.э.)

Замени каждую букву шифруемого текста на другую путем смещения в алфавите от исходной буквы на фиксированное количество символов!

Закодируем Б А Й Т –
сместим буквы
на 2 символа вправо

Получим: Г В Л Ф

А Б В Г Д Е Ё Ж З И Й К Л М Н О П Р С Т У
Ф Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я



ЗАДАНИЕ:

Расшифруйте слово

З Л Й М Щ Ы П В Н,

закодированное с помощью

шифра Цезаря.

**А Б В Г Д Е Ё Ж З И Й К Л М Н О П Р С Т У
Ф Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я**

ЗНАКОВЫЕ СИСТЕМЫ

```
graph TD; A[ЗНАКОВЫЕ СИСТЕМЫ] --> B[Естественные языки]; A --> C[Формальные языки]; A --> D[Генетический алфавит]; A --> E[Двоичная знаковая система]; B --- F{ }; C --- F; F --- G[Текстовая информация];
```

ЕСТЕСТВЕНН
ЫЕ
ЯЗЫКИ

ФОРМАЛЬНЫ
Е
ЯЗЫКИ

ГЕНЕТИЧЕСК
ИЙ АЛФАВИТ

ДВОИЧНАЯ
ЗНАКОВАЯ
СИСТЕМА

(В ПИСЬМЕННОЙ ФОРМЕ)

**ТЕКСТОВАЯ
ИНФОРМАЦИЯ**

Для представления текстовой информации в компьютере достаточно **256** различных знаков.

Воспользуемся формулой Хартли, чтобы вычислить количество информации для кодирования одного знака:

$$N=2^i \quad 256=2^8 \quad i=8 \text{ битов}$$

Значит, длина кода для одного символа составляет **8** двоичных знаков (**0** или **1**).

Десятичный код	Двоичный код
0	00000000
...	
...	
...	
255	11111111

ЧЕЛОВЕК РАЗЛИЧАЕТ ЗНАКИ ПО ИХ
НАЧЕРТАНИЮ, А КОМПЬЮТЕР – ПО ИХ
ДВОИЧНЫМ КОДАМ



Кодовая таблица – соответствие СИМВОЛОВ И ИХ ДВОИЧНЫХ КОДОВ.

Коды 0 - 32 соответствуют операциям (перевод строки, удаление символа слева от курсора, и т. п.)

Коды 33 – 127 – международный алфавит (латинские буквы, цифры, знаки препинания)

Коды 128 – 255 - национальный алфавит (для русских букв – 5 кодовых таблиц: **Windows, MS-DOS, КОИ-8, Mac, ISO**)

Таблицы кодировки русскоязычных символов

КОИ8-Р

—		Г	г	Л	л	Т	т	Т	т	■	■	■	■	■
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142
▤	▥	▦	Г	■	•	√	≈	≤	≥	nbsp	Ј	°	²	•
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158
=		ƒ	ё	п	Г	г	П	п	Е	Ц	Ц	Ј	Ј	Ј
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174
		ƒ	Ё			Т	П	Т	±	Ш	Ш	Т	Т	Т
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190
Ю	а	б	ц	д	е	ф	г	х	и	й	к	л	м	н
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206
П	я	р	с	т	у	ж	в	ь	ы	з	ш	э	щ	ч
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222
Ю	А	Б	Ц	Д	Е	Ф	Г	Х	И	Й	К	Л	М	Н
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238
П	Я	Р	С	Т	У	Ж	В	Ь	Ы	З	Ш	Э	Щ	Ч
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254

CP1251

Á	à	,	è	»	...	†	‡	€	%	É	<	й	Й	ó	ý
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
á	‘	’	“	”	•	—	—	€	™	é	>	ò	й	ó	ý
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
nbsp	û	Ы	Э	»	Ы	!	§	€	©	Ю	«	¬	shy	©	Я
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
°	±	Ы	Э	‘	µ	¶	•	ё	№	Ю	»	Э	Ю	Я	Я
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255

CP866

А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
▤	▥	▦		└	┌	┐	┌	┐		┐	┐	┐	┐	┐	┐
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
└	└	└	└	—	└	└	└	└	└	└	└	└	└	└	└
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
Ш	Т	П	Ц	Е	Р	П	└	└	└	Г	■	■	■	■	■
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
Ё	ё	Є	є	Ї	ї	Ў	ў	°	•	•	√	№	»	■	nbsp
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255

Mac

А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
†	°	Ы	£	§	•	¶	Ы	©	©	™	Á	á	è	à	è
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
Ё	±	≤	≥	з	μ	г	ó	ю	ю	я	я	é	é	й	ò
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
э	ю	¬	√	f	≈	Δ	«	»	...	nbsp	ó	ó	й	й	я
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
–	—	“	”	‘	’	÷	„	ý	Ы	ý	ý	№	Ё	ё	я
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255

ISO

І	І	І	І	І	І	І	І	І	І	І	І	І	І	І	І
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
І	І	І	І	І	І	І	І	І	І	І	І	І	І	І	І
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
nbsp	Ё	Ъ	Ѓ	Є	Ѕ	І	Ї	Ј	Љ	Њ	Ћ	Ќ	shy	Ў	Ў
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
№	ё	ђ	ѓ	є	ѕ	і	ї	ј	љ	њ	ћ	ќ	ѕ	ў	ұ
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255

В последние годы широкое распространение получил новый международный стандарт кодирования текстовых символов **Unicode**, который отводит на каждый символ не один, а два байта.

По формуле Хартли определим количество символов, которые можно закодировать по этому стандарту:

$$N=2^i \quad 2^{16}=65536$$

Такого количества символов оказалось достаточно, чтобы закодировать русский, латинский, греческий, арабский, иврит и другие алфавиты, цифры, знаки и математические символы.

Ответьте на вопросы:

1. Каков информационный объем текста, содержащего фразу

КОД СИМВОЛА, в кодировке Windows?

В кодировке Unicode?

2. Автоматическое устройство осуществило перекодировку информационного сообщения из кодировки **Windows** в кодировку **Unicode**. Во сколько раз увеличился информационный объём сообщения?

Решите задачу:

Информационное сообщение записано на компьютере на 4 строках по 40 символов в строке. Каков его информационный объём в кодировке Windows? В кодировке Unicode?

Домашнее задание:

По учебнику Угриновича § 3.1,
задания
для самостоятельного выполнения
№ 3.1, 3.2 (письменно).

