Кодирование текстовой информации



АЗБУКА МОРЗЕ

A		И	P	Ш	
16 B		Й	C	Щ	
В		K	T -	Ъ	
Г		Л	У	ы	
Д		M	Φ	ь	
Z E	-3	H	X	Э	
Ж		0	Ц	Ю	
3		П	Ч	Я	1

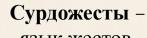
Информация, выраженная с помощью естественных и формальных языков в письменной форме, называется текстовой информацией.

Начиная с конца 60-х прошлого века, комг больше стали использоваться для обработки текстово

Историческая справка

Пайцике тсют т "камащамлтой чмароке" кайпонили, нмирепяшвейля ш Моллии ися цинсоракигелтой неменилти

Знаки кода	Бук	вы	Знаки кода	Циф-		Знаки препинания в				
Морзе	Pyc.	Анг.	Морзе	ры	Морзе	служебные сигналь				
	Α	Aa		1		(,) запятая				
	Б	Bb		2		(.) Точка				
	В	Ww		3		(:) точка с запятой				
	Г	Gg	****	4		Particular and Control of the Contro				
	Д	Dd	*****	5		(:) двоеточие				
-	E	Ee		6		(?) вопросит, знак				
****	ж	Vv		7						
	3	Zz		8		(№) номер				
	И	li		9		(") ковычки				
	K	Kk		0		(') апостроф				
	л	LI				() скобки				
	M	Mm				(!) восклицательный				
	н	Nn				знак				
	0	00				(-) тире				
	п	Pp			. —	Ждать				
	P	Rr				Понял				
	C	Ss								
_	T	Tt				(/) дробная черта				
	У	Uu				Знак раздела				
	Φ	Ff				Перебой (исправ-				
	×	Hh				ление ошибки)				
	ц	Cc				Сигнал о начале				
	ч					передачи (НП)				
	ш			1		Сигнал о готовности				
	щ	Qq		1		к приёму (ПО)				
	ы	Yy		1						
	Ю			1		Начало действия				
	Я			1		Знак окончания				
		Jj		1		передачи				
	ь,ъ	Xx		1						
** ***	Э	Ee	_		ı	ı				



язык жестов, используемый людьми с нарушениями слуха

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)



Кодирование текстовой информации - преобразование входной информации в форму, воспринимаемую компьютером, т.е. данные (двоичный код).

Декодирование текстовой информации -

преобразование данных из двоичного кода в форму, понятную человеку.

Для кодирования одного символа требуется один байт информации

1 символ — 1 байт (8 бит)

Учитывая, что каждый бит принимает значение 1 или 0, получаем, что с помощью 1 байта можно закодировать 256 различных символов.

N=2^I N - мощность алфавита

28=256 І – информационный вес

Кодирование и декодирование текстовой информации Кодирование заключается в том, что каждому

символу ставится в соответствие уникальный десятичный код от О до 255 или соответствующий ему двоичный код от 0000000 до 11111111. Таким образом, человек различает символы по их начертанию, а компьютер - по их коду.

Процессор О10000 Оперативная память

0100000





Таблица кодов таблица, в которой всем символам компьютерного алфавита поставлены в соответствие порядковые номера (коды)

Таблица кодировки <u>ASCII</u> является стандартной, и ее понимают абсолютно все программы, работающие с текстами.

Кодовая таблица АЅСП

American Standard Code for Information Interchange

1	«	#	\$	%	&		()	*	+	,	-		1
33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
1	2	3	4	5	6	7	8	9	:		\	Ш	^	?
49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
Α	В	С	D	Е	F	G	Н	- 1	J	K	L	M	N	0
35	66	67	68	69	70	71	72	73	74	75	76	77	78	79
Q	R	S	T	U	٧	W	X	Υ	Z]	1]	٨	
31	82	83	84	85	86	87	88	89	90	91	92	93	94	95
a	b	С	d	е	f	g	h	i	i	k	1	m	n	0
97	98	99	100	101	102	103	104	105	106	107	108	109	110	111
q	r	S	t	u	٧	W	X	у	Z	{	n o	}	٧	ڤ
C 10 10 10 10	114	115	116	117	118	119	120	121	122	123	124	125	126	127
4 / 6	1 19 A 65 Q 31 a 97	33 34 1 2 19 50 A B 55 66 Q R 31 82 a b 97 98	33 34 35 1 2 3 19 50 51 A B C 55 66 67 Q R S 31 82 83 a b c 97 98 99 q r s	33 34 35 36 1 2 3 4 19 50 51 52 A B C D 15 66 67 68 10 R S T 11 82 83 84 12 B C D 13 84 S T 14 82 83 84 15 C d 17 98 99 100 17 98 99 100	33 34 35 36 37 1 2 3 4 5 19 50 51 52 53 A B C D E 15 66 67 68 69 Q R S T U 16 82 83 84 85 17 98 99 100 101 Q r s t u	33 34 35 36 37 38 1 2 3 4 5 6 19 50 51 52 53 54 A B C D E F 65 66 67 68 69 70 Q R S T U V 31 82 83 84 85 86 a b c d e f 97 98 99 100 101 102 q r s t u v	33 34 35 36 37 38 39 1 2 3 4 5 6 7 19 50 51 52 53 54 55 A B C D E F G 65 66 67 68 69 70 71 Q R S T U V W 31 82 83 84 85 86 87 a b c d e f g 97 98 99 100 101 102 103 q r s t u v w	33 34 35 36 37 38 39 40 1 2 3 4 5 6 7 8 19 50 51 52 53 54 55 56 A B C D E F G H 65 66 67 68 69 70 71 72 Q R S T U V W X 31 82 83 84 85 86 87 88 a b c d e f g h 97 98 99 100 101 102 103 104 q r s t u v w x	33 34 35 36 37 38 39 40 41 1 2 3 4 5 6 7 8 9 19 50 51 52 53 54 55 56 57 A B C D E F G H I 65 66 67 68 69 70 71 72 73 Q R S T U V W X Y 31 82 83 84 85 86 87 88 89 a b c d e f g h i 97 98 99 100 101 102 103 104 105 q r s t u v w x y	33 34 35 36 37 38 39 40 41 42 1 2 3 4 5 6 7 8 9 : 19 50 51 52 53 54 55 56 57 58 A B C D E F G H I J 35 66 67 68 69 70 71 72 73 74 Q R S T U V W X Y Z 31 82 83 84 85 86 87 88 89 90 a b c d e f g h i j 97 98 99 100 101 102 103 104 105 106 q r s t u v w x y z	33 34 35 36 37 38 39 40 41 42 43 11 2 3 4 5 6 7 8 9 : ; 19 50 51 52 53 54 55 56 57 58 59 A B C D E F G H I J K 65 66 67 68 69 70 71 72 73 74 75 Q R S T U V W X Y Z [31 82 83 84 85 86 87 88 89 90 91 a b c d e f g h i j k 97 98 99 100 101 102 103 104 105 106 107 q r s t u v w x y z {	33 34 35 36 37 38 39 40 41 42 43 44 1 2 3 4 5 6 7 8 9 : ; <	33 34 35 36 37 38 39 40 41 42 43 44 45 1 2 3 4 5 6 7 8 9 : ; <	33 34 35 36 37 38 39 40 41 42 43 44 45 46 1 2 3 4 5 6 7 8 9 : ; <

коды от 0 до 31

коды от 32 до 127

коды от 128 до 255



функциональные клавиши



буквы английского алфавита, знаки математических операций и т.д.



национальный алфавит

Кодовая таблица в системе Windows

Символ	Десятичный код	Двоичный код	Символ	Десятичный код	Двоичный код
Пробел	32	00100000	0	48	00110000
1	33	00100001	0 1 2 3 4 5 6 7	49	00110001
*	42	00101010	2	50	00110010
+	43	00101011	3	51	00110011
•	44	00101100	4	52	00110100
100	45	00101101	5	53	00110101
- 4	46	00101110	6	54	00110110
/ = ?	47	001011110	7	55	00110111
=	61	00111101	8 9	56	00111000
3	63	00111111	9	57	00111001
7/20		00000000000	8820	- Charles	7000000000
A	192	11000000	P	208	11010000
А Б В Г	193	11000001	P C T	209	11010001
В	194	11000010	1	210	11010010
	195	11000011	y	211	11010011
Д E Ж	196	11000100	Φ	212	11010100
E	197	11000101	A	213	11010101
ж	198	11000110	у Ф Х Ц	214	11010110
3 N	199	11000111	ш	215	11010111
N N	200	11001000		216	11011000
	201	11001001	Щ	217	11011001
K	202	11001010	Ъ	218	11011010
Л	203	11001011	ы	219	11011011
M H	204 205	11001100 11001101	ь Э	220 221	11011100 11011101
	206	11001110	10		11011110
о П	207	1100 1111	Ю Я	222 223	11011111
19037	201	ALVO ALLE	SM	-223	

Кодировки русского

В настоящее врега про ределиоте нение получил Новый международный стандарт Unicode, который отводит на каждый символ не один байт, а два, и потому с его помощью можно закодировать не 256 символов, а N = 2^{16} = 65 536 различных символов. Такого количества символов достаточно, чтобы закодировать не только русский и латинский алфавиты, цифры, знаки и математические символы, но и греческий, арабский, иврит и другие алфавиты.



Таблица кодировки *Unicode*

Стандарт кодирования Unicode отводит на каждый символ 2 байта, что позволяет закодировать многие алфавиты в одной таблице.

 $N=2^{I}=2^{16}=65536$

В настоящее время существует 5 кодовых таблиц для русских букв (Windows, MS-DOS, КОИ-8, Mac, ISO), поэтому тексты, созданные в одной кодировке, не будут правильно отображаться в другой.

Таблицы кодировки русскоязычных символов

КОИ8-Р

	-	129	٦ 130	7	L 132	133	134	1 135	T 136	137	+	139	140	161	142	143																						1		
	144	145	146	۲ 147	148	143	1	≈ 151	52	<u>}</u>	nbsp 154	J 155	156	2 157	158	÷ 159]	ר	1) <u> </u>)																
	=	161	F 962	ë 163	IT 964	F 965	7	TI *67	77	E 163	I	L	172	113	Á	à	,	è	,,		ŧ	İ	€	%		(ì	í	1 (5 5	7									
	- t76	177	178	Ë	110	1	T 182	TT 183	7F 184	185	186	<u>JL</u> 187	± 188	189	á	129	130	131	132	133	134	135	136 È	TM	é)	ò	í	1 6	ó ý	7							4		
	100	a 193	6	Ц 195	Д 136	e 197	ф 198	Г 133	X 200	И 201	303 N	K 203	л 204	M 205	nbsp	¥ 9	Ы	9	148 H	bi bi	150	§	152 Ë	(D)	ΙĆ	· «	-	sh	y 6	9 9	i				P	3		56		
	П 208	Я 200	P 210	C	T 212	y 213	Ж 214	B 215	b 216	bl 217	3 218	ш 210	3 220	Щ 221	950	161 ±	62 Ы	163 3	154	165 µ	166 ¶	167	168 ë	163 N	Α	Б	В	Г	Д	Е	ж	3	И	й	K	Л	М	Н	0	П
	1O	A 225	Б 226	Ц 227	Д	E 229	Ф 230	Г 231	X 232	И 233	Й 234	K 235	Л 206	M 237	176 A	5	В	173 F	Д	181 E	ж	3	184	Ü	128 P	129 C	130 T	131 y	132 Ф	133 X	134	135	136	137	138 Ъ	139 Ы	140 Ь	3	142 IO	143 Я
	П 240	Я 241	P 242	C 243	T 244	y 245	Ж 246	B 247	Ь 248	Ы 243	3	Ш 251	3	Щ 253	P	C	154 T	195 y	Ф	197 X	198	199	200	Ц	144 a	145 б	В	147 Γ	148 Д	149 e	150 Ж	151	152 M	153 Й	154 K	155 Л	156 M	н	158 O	153 П
															208 a	6	210 B	L 541	212	213 C	214	3	216 M	ü	160	161	162	163	164	165	166	167 TI	168	169	170	171	긔	Ш	174 -	175
				7	/	Γ_									P 240	C 241	226 T 242	227 y 243	ф 244	229 X 245	230 Ц 246	231 4 247	ш	ш	176 L 192	177 	178 T	179	180	181	182	183	184 L	185 [F	186 <u>JL</u>	187 TF	188 L r	=	190 JL Tr	191
				1	V	Lc	1(247	1	246	240	644	247	240	0.71	644		Ш	193	194	195 LL	196 L	197 F	198 IT 214	199	200 +	7	202 F					207
12	8	129		131	132 1	33 1	34 1	35 1	36 13	1 K	8 139	9 140	0 141	142	_										208 p	209 C	210 T	211 y	212 cp	213 X	ц	215 4	216 Ш	217 Щ	218	213 bl	220 b	э	Ю 222	я
		C			Ф		Ц '	4 I	ШП		э Ы	ΙЬ	Э	Ю	Я									-	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239

Символ	Windows	MS-DOS	КОИ-8	Mac	ISO	Unicode
A	192	128	225	128	176	1040
В	194	130	247	130	178	1042
M	204	140	237	140	188	1052
Э	221	157	252	157	205	1069
Я	255	239	241	223	239	1103

Декодировать текст с помощью кодовой таблицы **ASCII**:

99 111 109 112 117 116 101 114

computer

Работа в текстовом редакторе

Запустите текстрвый редактор

MS Word. Удерживая клавишу «ALT»,

наберите коды на дополнительной

цифровой клавиатуре:

152 170 174 171 160



Домашнее задание

п. 3.1.

Произвести кодирование стихотворения из 4-х строк (до 100 символов)