

*27 января.
Классная работа.*

**Тексты в
компьютерной
памяти.**



Тексты в компьютерной памяти

1. Работа с текстами – одна из областей применения компьютера.
2. Преимущества компьютерного документа по сравнению с бумажным.
3. Представление текстов в памяти компьютера.
4. Гипертекст.



ПОДГОТОВКА ТЕКСТОВЫХ ДОКУМЕНТОВ



На любой каллиграфии есть
есть текстовый редактор.
Каллиграф - лучший
инструмент для подготовки
текстов.



ВВОД

На любой каллиграфии есть
есть ¹⁰ текстовый редактор.
Каллиграф - лучший
инструмент для подготовки
текстов.

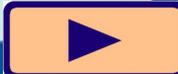
РЕДАКТИРОВАНИЕ

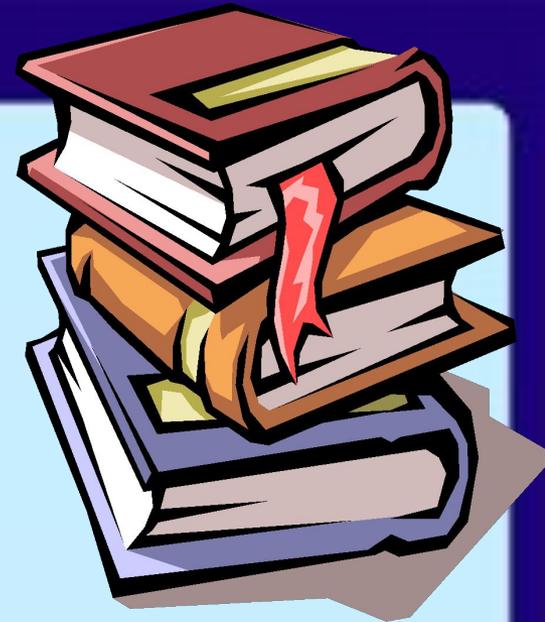
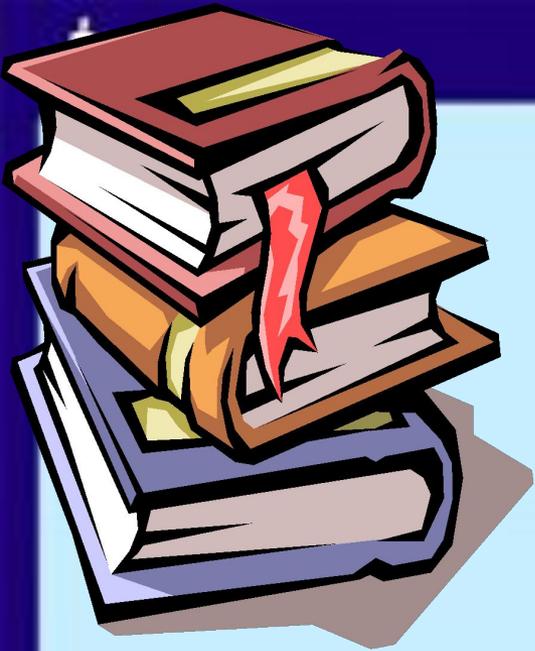
Каллиграф - лучший
инструмент для подготовки
текстов.
На любой каллиграфии
есть текстовый редактор.

ФОРМАТИРОВАНИЕ

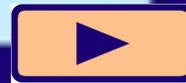
фрагмент

- СИМВОЛ
- СЛОВО
- СТРОКА
- ПРЕДЛОЖЕНИЕ
- АБЗАЦ
- ДОКУМЕНТ
- ПОСЛЕДОВАТЕЛЬНОСТЬ СИМВОЛОВ





Текст на внешних носителях
сохраняется в виде файла.

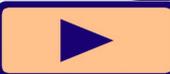
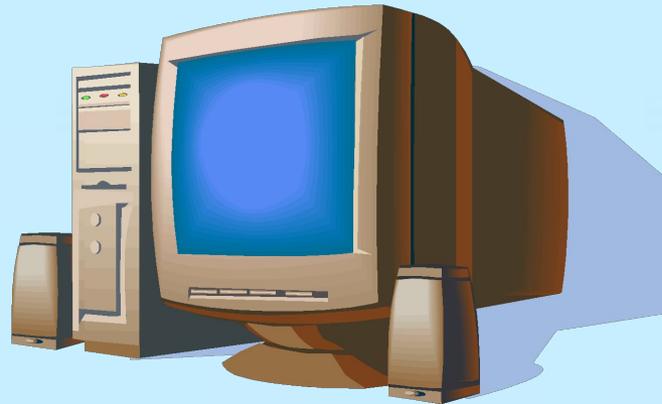


Преимущества компьютерного документа по сравнению с бумажным.

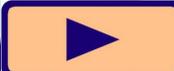
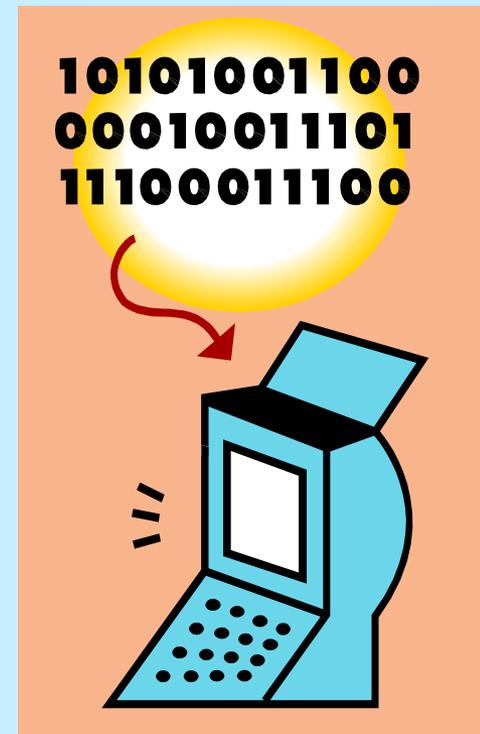
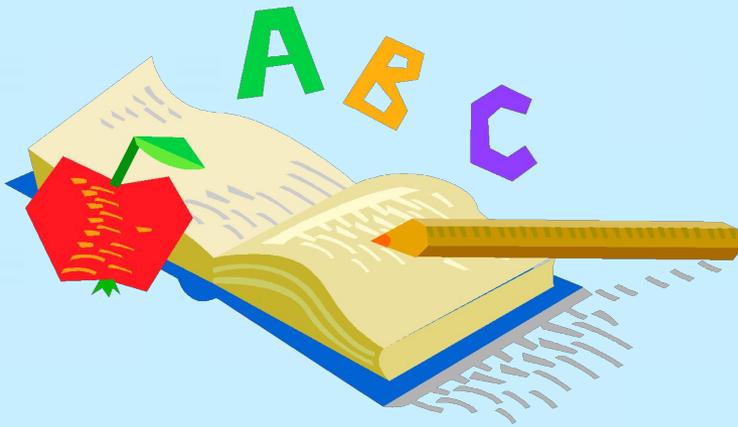
1. Экономия бумаги.
2. Компактное размещение текста и хранение на носителях внешней памяти в виде файлов.
3. Возможность быстрого исправления ошибок.
4. Возможность быстрой копии и печати.
5. Оперативная передача текста по линиям компьютерной связи.

Недостаток компьютерного документа:

- Документ можно прочитать, только с помощью компьютера.



Представление текста в памяти компьютера.





Алфавит – это множество всех символов , с помощью которых записывается текст.

Код - система условных знаков для представления информации.

Кодирование - это перевод информации в удобную для передачи, обработки или хранения форму с помощью некоторого кода.

Декодирование - это перевод закодированной информации в форму понятную человеку.

Кодирование одной и той же информации разными способами

КНИГА – русский язык

BOOK – английский язык

BUCH – немецкий язык



Кодирование одной и той же информации разными способами

КОМПЬЮТЕР

русский язык

COMPUTER

английский
язык

· - - - - - · - - - - - · - - - - - · - - - - - · - - - - - · - - - - - · - - - - - · - - - - -

код Морзе

Кодирование символов

1. Для представления текстовой информации в компьютере используется алфавит мощностью
2. Один символ алфавита несёт 8 битов информации **256**

$$256 = 2^8 \implies 8 \text{ бит на символ}$$

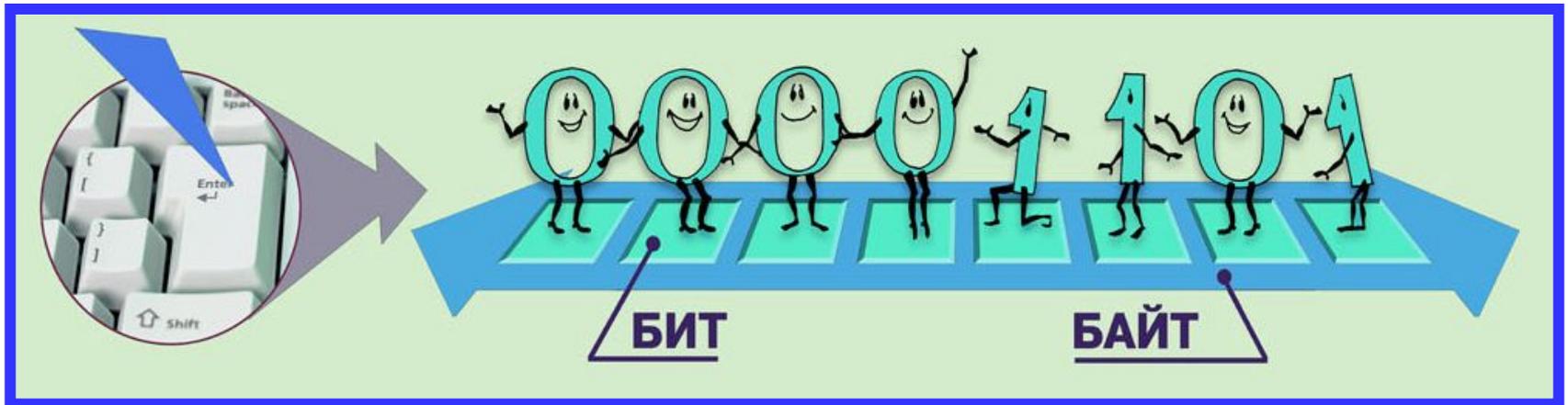
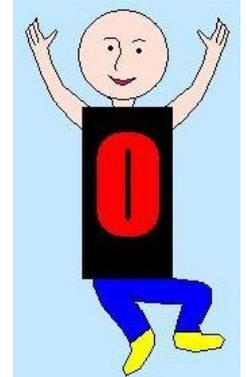
3. Каждому символу – **уникальный десятичный код 0..255**

Таблица символов:

	коды	65	66	67	68		
4. Коды – в	...	A	B	C	D	...	



1 байт – это 8 битов.



Двоичное кодирование

– это кодирование всех видов информации с помощью двух знаков (обычно 0 и 1).



A	1100 0000
Б	1100 0001
В	1100 0010

10101001100
00010011101
11100011100



Двоичное кодирование



- в такой форме можно закодировать **все виды** информации
- нужны только устройства с **двумя состояниями**
- практически **нет ошибок** при передаче
- **компьютеру легче** обрабатывать данные

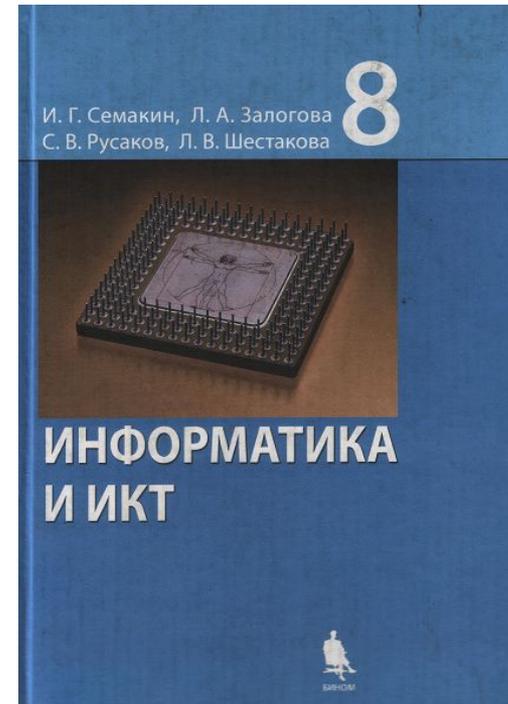


- **человеку сложно** воспринимать двоичные коды

**Двоичный код каждого символа
занимает 1 байт памяти компьютера.**

КОМПЬЮТЕР

9 символов - 9 байтов



**Символы всей книги
Примерно 400 Кбайт**

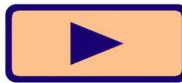


Таблица кодировки

- это таблица , в которой устанавливается соответствие между символами и их порядковыми номерами в компьютерном алфавите.
- это внутреннее представление символов в таблице.

Кодировка 1 байт на символ

0	1		127	128		254	255
		таблица ASCII (международная)			кодировка страница		

ASCII = *American Standard Code for Information Interchange*
(Американский стандартный код для информационного обмена)

0-31 управляющие символы:

7 – звонок, 10 – новая строка, 13 – возврат каретки

32 пробел

знаки препинания: . , : ; ! ?

специальные знаки: + - * / () { } []

48-57 цифры 0..9

65-90 заглавные латинские буквы **A-Z**

97-122 строчные латинские буквы **a-z**

Кодовая страница (расширенная таблица ASCII)

для русского языка:

CP-1251 для системы *Windows* (Интернет)

Стандартная часть кода ASCII

32		00100000	56	8	00111000	80	P	01010000	104	h	01101000
33	!	00100001	57	9	00111001	81	Q	01010001	105	i	01101001
34	"	00100010	58	:	00111010	82	R	01010010	106	j	01101010
35	#	00100011	59	;	00111011	83	S	01010011	107	k	01101011
36	\$	00100100	60	<	00111100	84	T	01010100	108	l	01101100
37	%	00100101	61	=	00111101	85	U	01010101	109	m	01101101
38	&	00100110	62	>	00111110	86	V	01010110	110	n	01101110
39	'	00100111	63	?	00111111	87	W	01010111	111	o	01101111
40	(00101000	64	@	01000000	88	X	01011000	112	p	01110000
41)	00101001	65	A	01000001	89	Y	01011001	113	q	01110001
42	*	00101010	66	B	01000010	90	Z	01011010	114	r	01110010
43	+	00101011	67	C	01000011	91	[01011011	115	s	01110011
44	,	00101100	68	D	01000100	92	\	01011100	116	t	01110100
45	-	00101101	69	E	01000101	93]	01011101	117	u	01110101
46	.	00101110	70	F	01000110	94	^	01011110	118	v	01110110
47	/	00101111	71	G	01000111	95	_	01011111	119	w	01110111
48	0	00110000	72	H	01001000	96	`	01100000	120	x	01111000
49	1	00110001	73	I	01001001	97	a	01100001	121	y	01111001
50	2	00110010	74	J	01001010	98	b	01100010	122	z	01111010
51	3	00110011	75	K	01001011	99	c	01100011	123	{	01111011
52	4	00110100	76	L	01001100	100	d	01100100	124		01111100
53	5	00110101	77	M	01001101	101	e	01100101	125	}	01111101
54	6	00110110	78	N	01001110	102	f	01100110	126	~	01111110
55	7	00110111	79	O	01001111	103	g	01100111	127	□	01111111

Кодовая страница CP1251

128	Ъ	10000000	160	"	10100000	192	А	11000000	224	а	11100000
129	Ҁ	10000001	161	Ҁ	10100001	193	Б	11000001	225	б	11100001
130	҂	10000010	162	҂	10100010	194	В	11000010	226	в	11100010
131	҄	10000011	163	҄	10100011	195	Г	11000011	227	г	11100011
132	҆	10000100	164	҆	10100100	196	Д	11000100	228	д	11100100
133	҈	10000101	165	҈	10100101	197	Е	11000101	229	е	11100101
134	Ҋ	10000110	166	Ҋ	10100110	198	Ж	11000110	230	ж	11100110
135	Ҍ	10000111	167	Ҍ	10100111	199	З	11000111	231	з	11100111
136	Ҏ	10001000	168	Е	10101000	200	И	11001000	232	и	11101000
137	Ґ	10001001	169	Ө	10101001	201	Й	11001001	233	й	11101001
138	Ќ	10001010	170	Є	10101010	202	К	11001010	234	к	11101010
139	Ў	10001011	171	•	10101011	203	Л	11001011	235	л	11101011
140	Ѓ	10001100	172	ґ	10101100	204	М	11001100	236	м	11101100
141	Ѕ	10001101	173	•	10101101	205	Н	11001101	237	н	11101101
142	Ї	10001110	174	Ө	10101110	206	О	11001110	238	о	11101110
143	Ў	10001111	175	Ї	10101111	207	П	11001111	239	п	11101111
144	Ҁ	10010000	176	•	10110000	208	Р	11010000	240	р	11110000
145	҂	10010001	177	±	10110001	209	С	11010001	241	с	11110001
146	҄	10010010	178	і	10110010	210	Т	11010010	242	т	11110010
147	҆	10010011	179	і	10110011	211	У	11010011	243	у	11110011
148	҈	10010100	180	Ҁ	10110100	212	Ф	11010100	244	ф	11110100
149	Ҋ	10010101	181	҂	10110101	213	Х	11010101	245	х	11110101
150	Ҍ	10010110	182	҄	10110110	214	Ц	11010110	246	ц	11110110
151	Ҏ	10010111	183	•	10110111	215	Ч	11010111	247	ч	11110111
152	Ґ	10011000	184	є	10111000	216	Ш	11011000	248	ш	11111000
153	Ғ	10011001	185	№	10111001	217	Щ	11011001	249	щ	11111001
154	҄	10011010	186	Є	10111010	218	Ъ	11011010	250	ъ	11111010
155	҆	10011011	187	•	10111011	219	Ы	11011011	251	ы	11111011
156	҈	10011100	188	і	10111100	220	Ь	11011100	252	ь	11111100
157	Ҋ	10011101	189	Ѕ	10111101	221	Э	11011101	253	э	11111101
158	Ҍ	10011110	190	є	10111110	222	Ю	11011110	254	ю	11111110
159	Ҏ	10011111	191	і	10111111	223	Я	11011111	255	я	11111111

Кодирование символов

Текстовый файл

- на экране (символы)
- в памяти – двоичные коды



1000001 ₂	1000010 ₂	1000011 ₂	1000100 ₂
65	66	67	68



В файле хранятся не изображения символов, а их числовые коды в двоичной системе!

Какое слово записано двоичным кодом?

01100100

01101001

01110011

01101011

d

i

s

k



Кодировка UNICODE

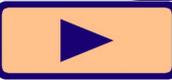
- *Windows, MS Office, ...*
- **16 бит на символ**
- **65536** или **2^{16}** СИМВОЛОВ в одной таблице



можно одновременно использовать символы разных языков (Интернет)



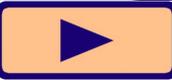
размер файла увеличивается в **2 раза**



Гипертекст

- это текст, организованный так, что его можно просматривать в последовательности смысловых связей между его отдельными фрагментами.

Такие связи называются ГИПЕРСВЯЗЯМИ
(ГИПЕРССЫЛКАМИ).





Тексты в компьютерной памяти

1. Работа с текстами – одна из областей применения компьютера.
2. Преимущества компьютерного документа по сравнению с бумажным.
3. Представление текстов в памяти компьютера.
4. Гипертекст.