

10 класс

Кодирование и обработка информации

Учитель информатики
МОУ СОШ №8 с.Ульяновка

Голованева Елена Владимировна

Сегодня на уроке

- Проверим наши знания
- Изучим новую тему
- Потренируемся, закрепим новые знания
- Разберём домашнее задание
- Подведём итоги

Кодирование и обработка текстовой информации

Уже с 60-х годов прошлого столетия, компьютеры всё больше стали использовать для обработки текстовой информации. Для кодирования текстовой информации в компьютере применяется двоичное кодирование, т.е. представление текста в виде последовательности 0 и 1 (Эти два символа называются двоичными цифрами, по-английски – binary digit или сокращённо **bit**).

Почему же двоичное кодирование ?

Такой способ легко реализовать технически:

1 – есть сигнал,

0 – нет сигнала.

Каждому символу алфавита сопоставили определённое количество и последовательность нулей и единиц.

Сколько же бит необходимо для кодирования символов?

Посчитаем примерное достаточное количество символов для алфавита компьютера и по формуле вычислим необходимое количество бит.

33 русских прописных буквы +
33 русских строчных букв +
26 английских строчных букв +
26 прописных английских букв +
10 цифр + знаки препинания +
скобки и знаки математических операций +
специальные символы (@, #, \$, %, &, *) +
знаки псевдографики

≈ 256 символов.

$$N = 2^l$$
$$256 = 2^8$$

В чём суть кодирования?

Кодирование заключается в том, что каждому символу ставится в соответствие уникальный десятичный код от 0 до 255 или соответствующий ему двоичный код от 00000000 до 11111111.

ASCII (American Standard Code for Information Interchange).

- Первые 33 кода (с 0 по 33) этой таблице соответствует не символам, а операциям (ввод пробела, перевод строки и т.д.)
- Коды с 33 по 127 являются интернациональными и соответствуют символам латинского алфавита, цифрам, знакам арифметических операций и знакам препинания.

ASCII

символ	10- Б код	2-Б код	символ	10- Б код	2-Б код	символ	10-Б код	2-Б код	символ	10-Б код	2-Б код
	32	00100000	8	56	00111000	P	80	01010000	h	104	01101000
!	33	00100001	9	57	00111001	Q	81	01010001	i	105	01101001
"	34	00100010	:	58	00111010	R	82	01010010	j	106	01101010
#	35	00100011	;	59	00111011	S	83	01010011	k	107	01101011
\$	36	00100100	<	60	00111100	T	84	01010100	l	108	01101100
%	37	00100101	=	61	00111101	U	85	01010101	m	109	01101101
&	38	00100110	>	62	00111110	V	86	01010110	n	110	01101110
'	39	00100111	?	63	00111111	W	87	01010111	o	111	01101111
(40	00101000	@	64	01000000	X	88	01011000	p	112	01110000
)	41	00101001	A	65	01000001	Y	89	01011001	q	113	01110001
*	42	00101010	B	66	01000010	Z	90	01011010	r	114	01110010
+	43	00101011	C	67	01000011	[91	01011011	s	115	01110011
,	44	00101100	D	68	01000100	\	92	01011100	t	116	01110100
-	45	00101101	E	69	01000101]	93	01011101	u	117	01110101
.	46	00101110	F	70	01000110	^	94	01011110	v	118	01110110
/	47	00101111	G	71	01000111	_	95	01011111	w	119	01110111
0	48	00110000	H	72	01001000	`	96	01100000	x	120	01111000
1	49	00110001	I	73	01001001	a	97	01100001	y	121	01111001
2	50	00110010	J	74	01001010	b	98	01100010	z	122	01111010
3	51	00110011	K	75	01001011	c	99	01100011	{	123	01111011
4	52	00110100	L	76	01001100	d	100	01100100		124	01111100
5	53	00110101	M	77	01001101	e	101	01100101	}	125	01111101
6	54	00110110	N	78	01001110	f	102	01100110	~	126	01111110
7	55	00110111	O	79	01001111	g	103	01100111	□	127	01111111

ASCII

- Коды с 128 по 255 являются национальными.

ASCII

символ	10-В код	2-В код	символ	10-В код	2-В код	символ	10-В код	2-В код	символ	10-В код	2-В код
Т	128	10000000		160	10100000	А	192	11000000	а	224	11100000
Г	129	10000001	Ў	161	10100001	Б	193	11000001	б	225	11100001
,	130	10000010	ў	162	10100010	В	194	11000010	в	226	11100010
г	131	10000011	Ј	163	10100011	Г	195	11000011	г	227	11100011
„	132	10000100	о	164	10100100	Д	196	11000100	д	228	11100100
...	133	10000101	Г	165	10100101	Е	197	11000101	е	229	11100101
†	134	10000110		166	10100110	Ж	198	11000110	ж	230	11100110
‡	135	10000111	§	167	10100111	З	199	11000111	з	231	11100111
€	136	10001000	€	168	10101000	И	200	11001000	и	232	11101000
%.	137	10001001	©	169	10101001	Й	201	11001001	й	233	11101001
Љ	138	10001010	€	170	10101010	К	202	11001010	к	234	11101010
<	139	10001011	«	171	10101011	Л	203	11001011	л	235	11101011
Њ	140	10001100	¬	172	10101100	М	204	11001100	м	236	11101100
К	141	10001101	-	173	10101101	Н	205	11001101	н	237	11101101
Т	142	10001110	@	174	10101110	О	206	11001110	о	238	11101110
Ц	143	10001111	Ї	175	10101111	П	207	11001111	п	239	11101111
ђ	144	10010000	°	176	10110000	Р	208	11010000	р	240	11110000
‘	145	10010001	±	177	10110001	С	209	11010001	с	241	11110001
’	146	10010010	І	178	10110010	Т	210	11010010	т	242	11110010
“	147	10010011	і	179	10110011	У	211	11010011	у	243	11110011
”	148	10010100	г	180	10110100	Ф	212	11010100	ф	244	11110100
•	149	10010101	μ	181	10110101	Х	213	11010101	х	245	11110101
—	150	10010110	¶	182	10110110	Ц	214	11010110	ц	246	11110110
—	151	10010111	·	183	10110111	Ч	215	11010111	ч	247	11110111
□	152	10011000	ë	184	10111000	Ш	216	11011000	ш	248	11111000
™	153	10011001	№	185	10111001	Щ	217	11011001	щ	249	11111001
Љ	154	10011010	€	186	10111010	Ъ	218	11011010	ъ	250	11111010
>	155	10011011	»	187	10111011	Ы	219	11011011	ы	251	11111011
Њ	156	10011100	j	188	10111100	Ь	220	11011100	ь	252	11111100
К	157	10011101	š	189	10111101	Э	221	11011101	э	253	11111101
ћ	158	10011110	s	190	10111110	Ю	222	11011110	ю	254	11111110
ц	159	10011111	ï	191	10111111	Я	223	11011111	я	255	11111111

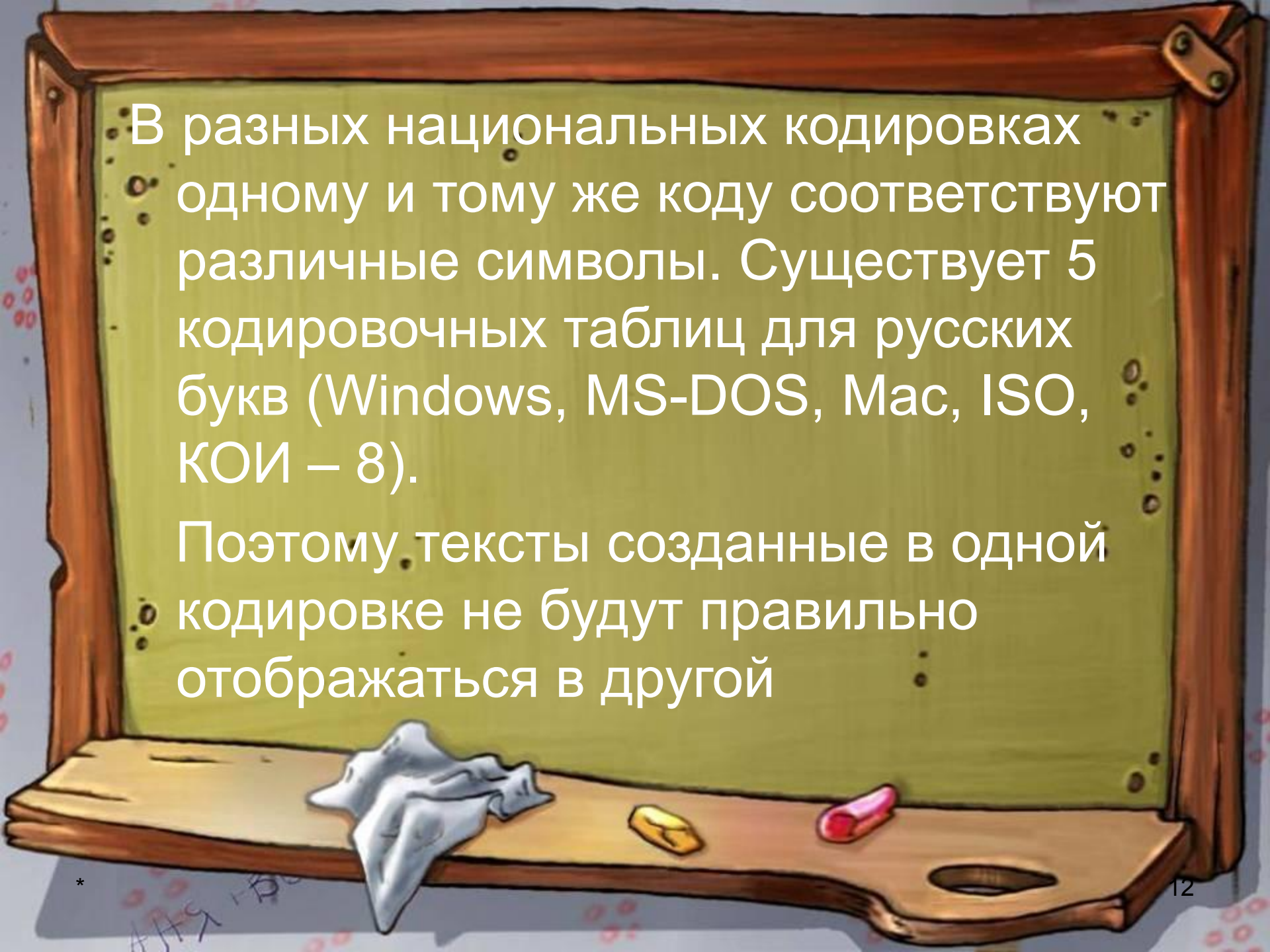
ASCII

- Какое слово закодировано?
(С помощью кодовой таблицы ASCII декодируйте слово)

- 11001010 11101110 11100100

- 202 238 228

- **К** **О** **Д**



В разных национальных кодировках одному и тому же коду соответствуют различные символы. Существует 5 кодировочных таблиц для русских букв (Windows, MS-DOS, Mac, ISO, KOI – 8).

Поэтому тексты созданные в одной кодировке не будут правильно отображаться в другой

Двоичный код	Десятичный код	КОИ8	CP1251	CP866	Mac	ISO
0000 0000	0					
.....						
0000 1000	8	Удаление последнего символа (клавиша Backspace)				
.....						
0000 1101	13	Перевод строки (клавиша Enter)				
.....						
0010 0000	32	Пробел				
0010 0001	33	!				
.....						
0101 1010	90	Z				
.....						
0111 1111	127	□				
.....	128	-	Ъ	А	А	К
.....						
1100 0010	194	Б	В	-	-	Т
.....						
1100 1100	204	Л	М	:	:	Ь
.....						
1101 1101	221	Щ	Э	-	Ё	Н
.....						

*

Windows 1251

Двоичное кодирование текстовой информации: кодовые таблицы символов

В качестве основного средства для обмена информацией с другими людьми человек использует естественные языки. Таковыми являются, например, русский, английский, японский и др. Они характеризуются тем, что носят национальный характер. Естественные языки существуют в устной (фонетика) и письменной форме (грамматика) (сейчас практически невозможно встретить язык, который не имеет письменной формы). В основе языка лежит алфавит, т.е. набор символов, которые используются для построения более крупных конструкций языка. В разных языках алфавит составляет от нескольких десятков до нескольких десятков тысяч символов. Вообще естественные языки являются предметом изучения филологии. В информатике гораздо большее внимание уделяется формальным языкам.

Основой *формальных языков* также служит алфавит. Множество всех символов, с помощью которых записывается текст, называется **алфавитом**, а число символов в алфавите — его **мощностью**. Но, в отличие от естественных, в формальных языках он довольно жестко фиксирован. Кроме того, правила грамматики и синтаксиса здесь более строгие, формализованные, фиксированные, существует ряд ограничений. В искусственных знаковых системах отсутствует многозначность. Каждая лексическая единица — слово — имеет ровно один смысл, и наоборот.

В отличие от естественных, искусственные языки ориентированы в основном на письменное представление. Примерами таких языков могут служить язык математики (математическая символика), физики, химии, музыки (ноты) и т.д. Языки программирования (и другие средства записи алгоритмов) тоже относятся к формальным. При автоматизированной обработке информации нет возможности использования естественных языков.

Кодирование информации подразумевает преобразование знаков одной знаковой системы в знаки или группы знаков другой знаковой системы. Обратное преобразование называют **декодированием**.

КОИ -8

ДБНХВМНЕ ЙНДХПНБЮМХЕ РЕЙАРНБНИ ХМТНПЛЮЖХХ: ЙНДНБШЕ РЮАКХЖШ ЯХЛБНКНБ

Б ЙЮВЕАРБЕ НАМНБМНЦН ЯПЕДЯРБЮ ДКЪ НАЛЕМЮ ХМТНПЛЮЖХЕИ Я ДПСЦХЛХ КЧДЭЛХ ВЕКНБЕЙ ХЯОНКЭГСЕР ЕЯРЕАРБЕММШЕ ЁГШЙХ. РЮЙНБШЛХ ЁБКЪЧРЯЪ, МЮОПХЛЕП, ПСЯЯЙХИ, ЮМЦКХИЯЙХИ, ЁОНМЯЙХИ Х ДП. НМХ УЮПЮЙРЕПХГСЧРЯЪ РЕЛ, ВРН МНЯЪР МЮЖХНМЮКЭМШИ УЮПЮЙРЕП. ЕЯРЕАРБЕММШЕ ЁГШЙХ ЯСЫЕАРБСЧР Б СЯРМНИ (ТНМЕРХИЮ) Х ОХЯЭЛЕММНИ ТНПЛЕ (ЦПЮЛЛЮРХИЮ) (ЯЕВИЮЯ ОПЮЙРХВЕЯЙХ МЕБНГЛНФМН БЯРПЕРХРЭ ЁГШЙ, ЙНРНПШИ МЕ ХЛЕЕР ОХЯЭЛЕММНИ ТНПЛШ). Б НАМНБЕ ЁГШИЮ КЕФХР ЮКТЮБХР, Р.Е. МЮАНП ЯХЛБНКНБ, ЙНРНПШЕ ХЯОНКЭГСЧРЯЪ ДКЪ ОНЯРПНЕМХЪ АНКЕЕ ЙПСОМШУ ЙНМАРПСИЖХИ ЁГШИЮ. Б ПЮГМШУ ЁГШИЮЮ ЮКТЮБХР ЯНЯРЮБКЪЕР НР МЕЯЙНКЭЙХУ ДЕЯЪРИНБ ДН МЕЯЙНКЭЙХУ ДЕЯЪРИНБ РШАЪВ ЯХЛБНКНБ. БННАЫЕ ЕЯРЕАРБЕММШЕ ЁГШЙХ ЁБКЪЧРЯЪ ОПЕДЛЕРНЛ ХГСВЕМХЪ ТХКНКЦХХ. Б ХМТНПЛЮРХИЕ ЦНПЮГДН АНКЭБЕЕ БМХЛЮМХЕ СДЕКЪЕРЯЪ ТНПЛЮКЭМШЛ ЁГШИЮЛ.

НАМНБНИ ТНПЛЮКЭМШУ ЁГШИНБ РЮЙФЕ ЯКСФХР ЮКТЮБХР. ЛМНФЕАРБН БАЕУ ЯХЛБНКНБ, Я ОНЛНЫЭЧ ЙНРНПШУ ПЮОХЯШБЮЕРЯЪ РЕЙАР, МЮГШБЮЕРЯЪ ЮКТЮБХРНЛ, Ю ВХЯКН ЯХЛБНКНБ Б ЮКТЮБХРЕ ≈ ЕЦН **ЛНЫМНЯРЭЧ**. МН, Б НРКХВХЕ НР ЕЯРЕАРБЕММШУ, Б ТНПЛЮКЭМШУ ЁГШИЮЮ НМ ДНБНКЭМН ФЕЯРИН ТХИЯХПНБЮМ. ЙПНЛЕ РНЦН, ОПЮБХКЮ ЦПЮЛЛЮРХИХ Х ЯХМРЮЙЯХЮ ГДЕЯЭ АНКЕЕ ЯРПНЦХЕ, ТНПЛЮКХГНБЮММШЕ, ТХИЯХПНБЮММШЕ, ЯСЫЕАРБСЕР ПЪД НЦПЮМХВЕМХИ. Б ХЯЙСЯАРБЕММШУ ГМЮЙНБШУ ЯХЯРЕЛЮУ НРЯСРЯРБСЕР ЛМНЦНГМЮВМНЯРЭ. ЙЮФДЮЪ КЕЙЯХВЕЯЙЮЪ ЕДХМХЖЮ ≈ ЯКНБН ≈ ХЛЕЕР ПНБМН НДХМ ЯЛШАК, Х МЮНАНПНР.

Б НРКХВХЕ НР ЕЯРЕАРБЕММШУ, ХЯЙСЯАРБЕММШЕ ЁГШИХ НПХЕМРХПНБЮМШ Б НАМНБМНЛ МЮ ОХЯЭЛЕММНЕ ОПЕДЯРЮБКЕМХЕ. ОПХЛЕПЮЛХ РЮЙХУ ЁГШИНБ ЛНЦСР ЯКСФХРЭ ЁГШИ ЛЮРЕЛЮРХИХ (ЛЮРЕЛЮРХВЕЯЙЮЪ ЯХЛБНКХИЮ), ТХГХИХ, УХЛХХ, ЛСГШИХ (МНРШ) Х Р.Д. ЁГШИХ ОПНЦПЮЛЛХПНБЮМХЪ (Х ДПСЦХЕ ЯПЕДЯРБЮ ПЮОХЯХ ЮКЦНПХРЛНБ) РНФЕ НРМНЯЪРЯЪ Й ТНПЛЮКЭМШЛ. ОПХ ЮБРНЛЮРХГХПНБЮММНИ НАПЮАНРИЕ ХМТНПЛЮЖХХ МЕР БНГЛНФМНЯРХ ХЯОНКЭГНБЮМХЪ ЕЯРЕАРБЕММШУ ЁГШИНБ.

ЙНДХПНБЮМХЕ ХМТНПЛЮЖХХ ОНДПЮГСЛЕБЮЕР ОПЕНАПЮГНБЮМХЕ ГМЮЙНБ НДМНИ ГМЮЙНБНИ ЯХЯРЕЛШ Б ГМЮИХ ХКХ ЦПСООШ ГМЮЙНБ ДПСЦНИ ГМЮЙНБНИ ЯХЯРЕЛШ. НАПЮРМНЕ ОПЕНАПЮГНБЮМХЕ МЮГШБЮЧР **ДЕЙНДХПНБЮМХЕЛ**.

ISO

Фтиошіэюх ъюфш№ютрэшх ѱхъёђютющ шэею№ьришш: ъюфютђх ѱрсышіћ ёшьтюыют

Т ѱрихёђтх юёэютэюю ё№ѳѳёђтр фыц юсьхэр шэею№ьришшц ё ф№ѓушьш ыўфкъш їхютхъ шёяюыќчѓћ хёђхёђтхээћх пчћш. врьютћш цтыцўђёц, зря№шьх№, №ѓёёшьщ, рзуыщёшьщ, цязёёшьщ ш ф№. Юэш ср№рђђх№шчѓўђёц ѱхъ, їюю зюёцђ эришюзрыќэћц ср№рђђх№. Хёђхёђтхээћх пчћш ёѓльхёђтѓўђ т řёђэющ (еюзхђшър) ш яшёќхээющ ею№ьх (у№рђђрђшър) (ёхцїрё я№рђђшїхёъш зхтючыюцэю тёђ№хђшђќ пчћ, ъюђю№ђщ зх шыхђ яшёќхээющ ею№ьћ). Т юёэютх пчћр ыхцшђ рьертшђ, ѱ.х. эрсю№ ёшьтюыют, ъюђю№ђш шёяюыќчѓўђёц фыц яюёђ№юхэшц суюхх ъ№ѓязћс ъюзёђ№ѓышщ пчћр. Т №рчэћс пчћрс зьертшђ ёюёђртыцхђ юђ зхёюыќъшс фхёцђют фю зхёюыќъшс фхёцђют ѱђёцї ёшьтюыют. Тююслъ хёђхёђтхээћх пчћш цтыцўђёц я№ѳѳёђтхё шчѓїхэшц зшыюююшш. Т шэею№ьрђшъх юю№рчфю суюќјхх тэшърэшх řфхыцхђёц ею№ьрыќэћ пчћрь.

Юёэютющ ею№ьрыќэћс пчћют ѱрцх ёыѓцшђ рьертшђ. Ьзюцхёђтю тёхс ёшьтюыют, ё яюыюльќў ъюђю№ђс чряшёђтрхђёц ѱхъёђ, эрчђтрхђёц *рьертшђюь*, з їшёюю ёшьтюыют т рьертшђх ¹⁰⁰ хую *ъюльзюёђќў*. Эю, т юђышїшх юђ хёђхёђтхээћс, т ею№ьрыќэћс пчћрс юз фютюыќэю цхёђюю ешьёш№ютрэ. Ь№юьх јююю, я№ртыр у№рђђрђшърш ш ёшэђрђёшёр чфхёќ суюхх ёђ№ююшх, ею№ьрышчютрээћх, ешьёш№ютрэзёћх, ёѓльхёђтѓћ №цф юу№рэшїхэшщ. Т шёѓёёђтхээћс чэрьютћс ёшёђхърс юђёђђёђтѓћ ѱзююючэриэюёђќ. Ьрцфрц ыхъёшїхёърц хфшэшїр ¹⁰⁰ ёыютю ¹⁰⁰ шыхђ №ютэю юфшэ ёьћёы, ш эрсюю№юђ.

Т юђышїшх юђ хёђхёђтхээћс, шёѓёёђтхээћх пчћш ю№шхзђш№ютрэћ т юёэютэю эр яшёќхээюх я№ѳѳёђртыхэшх. Я№шьх№реш ѱршс пчћют ъюуѓђ ёыѓцшђќ пчћ ѱђђхърђшъш (ѱђђхърђшїхёърц ёшьтюышър), ешчшш, сшьшш, ыѓчћш (зюђђ) ш ѱ.ф. пчћш я№ююу№рђшш№ютрэшц (ш ф№ѓушх ё№ѳѳёђтр чряшёш рыую№шђют) ѱюцх юђзюёцђёц ъ ею№ьрыќэћ. Я№ш рђђюрђшшчш№ютрэзющ юс№рсођђх шэею№ьришш зхђ тючыюцэюёђш шёяюыќчютрэшц хёђхёђтхээћс пчћют.

Ъюфш№ютрэшх шэею№ьришш яюф№рчѓхтрхђ я№юс№рчютрэшх чэрьют юфэющ чэрьютющ ёшёђхћ т чэрш шыш у№ѓязђћ чэрьют ф№ѓующ чэрьютющ ёшёђхћ. Юс№рђђэюх я№юс№рчютрэшх эрчђтрђђ *фхъюфш№ютрэшхъ*.

Закрепление изученного.

Практическая работа

2. Задание. А1

Считая, что каждый символ кодируется одним байтом, определите, чему равен информационный объём следующего высказывания Жан-Жака Руссо:

Тысячи путей ведут к заблуждению, к истине – только один.

- 1) 92 бита 2) 220 бит 3) 456 бит 4) 512 бит

Ответ : 3) 456 бит.

Закрепление изученного.

Практическая работа

- А2. Два текста содержат одинаковое количество символов. Первый текст составлен в алфавите мощностью 16 символов. Второй текст в алфавите мощностью 256 символов. Во сколько раз количество информации во втором тексте больше, чем в первом?
- 1) 12 2) 2 3) 24 4) 4

• Ответ: 2) 2.


Обобщение

- 1. Какой принцип кодирования текстовой информации используется в компьютере?
- 2. Как называется международная таблица кодировки символов?
- 3. Перечислите названия таблиц кодировок для русскоязычных символов.
- 4. В какой системе счисления представлены коды в перечисленных вами таблицах кодировок?

Домашнее задание

- Учебник Н.Д.Угринович 10 класс. § 1.1.1, ответить на вопрос.
- Выполнить задание из Демонстрационного варианта контрольных измерительных материалов единого государственного экзамена 2011 года по информатике и ИКТ

A2. Автоматическое устройство осуществило перекодировку информационного сообщения на русском языке длиной в 20 символов, первоначально записанного в 16-битном коде Unicode, в 8-битную кодировку КОИ-8. При этом информационное сообщение уменьшилось на
1) 320 бит 2) 20 бит 3) 160 байт 4) 20 байт



Спасибо за внимание!
Успехов в изучении
информатики!