

25.11.2016

Измерение информации



Усольцева Э.М-А.
преподаватель информатики
ГОУ НПО «Качканарское ПУ»

Алфавитный подход к измерению информации



Центральный процессор (ЦП) -

**управляет работой ПК и
преобразует информацию
(«МОЗГ» ПК)**





ЦП может представлять собой один кремниевый кристалл.

В таких случаях ЦП называют **микропроцессором (МП)**.

Компьютер работает от электрической сети в которой может быть реализована система, основанная на 2-х состояниях:

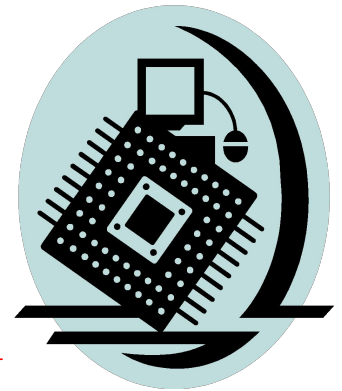
Есть ток – нет тока

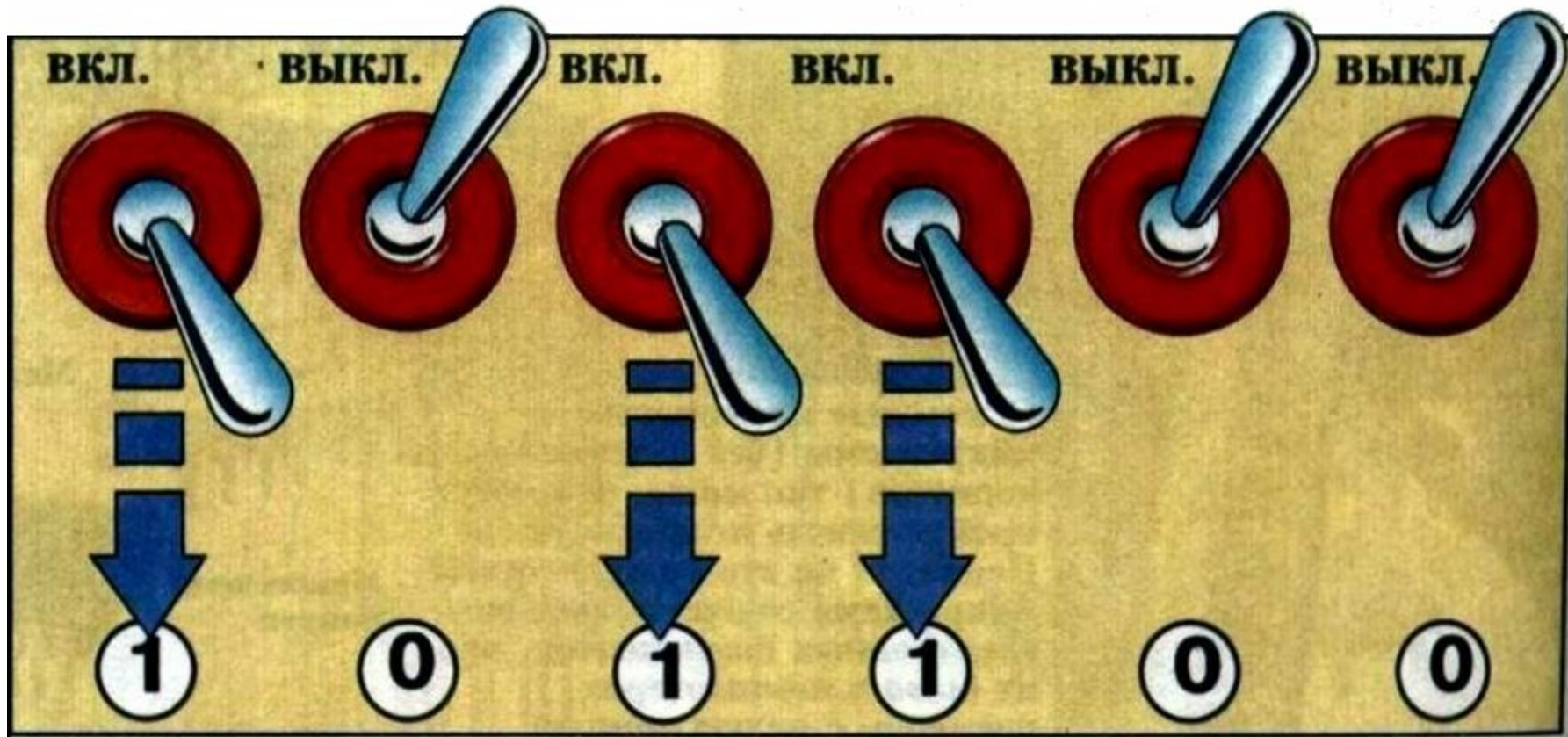
Есть напряжение – нет напряжения

На этом и базируется работа ЦП

0 – нет тока, «ложь»

1 – есть ток, «истина»





0 или 1 = 1 бит информации

Вопрос:

Хватит ли 0 и 1, чтобы закодировать все символы, которые мы вводим в компьютер с клавиатуры?



Код	Сим-вол	Код	Сим-вол	Код	Сим-вол	Код	Сим-вол
00100000	пробел	00110000	0	01000000	@	01010000	P
00100001	!	00110001	1	01000001	A	01010001	Q
00100010	"	00110010	2	01000010	B	01010010	R
00100011	#	00110011	3	01000011	C	01010011	S
00100100	\$	00110100	4	01000100	D	01010100	T
00100101	%	00110101	5	01000101	E	01010101	U
00100110	&	00110110	6	01000110	F	01010110	V
00100111	'	00110111	7	01000111	G	01010111	W
00101000	(00111000	8	01001000	H	01011000	X
00101001)	00111001	9	01001001	I	01011001	Y
00101010	*	00111010	:	01001010	J	01011010	Z
00101011	+	00111011	;	01001011	K	01011011	[
00101100	,	00111100	<	01001100	L	01011100	\
00101101	—	00111101	=	01001101	M	01011101]
00101110	.	00111110	>	01001110	N	01011110	^
00101111	/	00111111	?	01001111	O	...	

Коду 00100000 в этой таблице соответствует *пробел* — пустой промежуток величиной в один символ, который используется для отделения одного слова от другого.

Коды русских букв отличаются от кодов латинских. Например, большая русская буква "М" имеет код 11101101, буква "И" — код 11101001, буква "Р" — код 11110010, буква "У" — код 11110101. Таким образом, слово "МИР" кодируется последовательностью из 24 бит

111011011110100111110010,

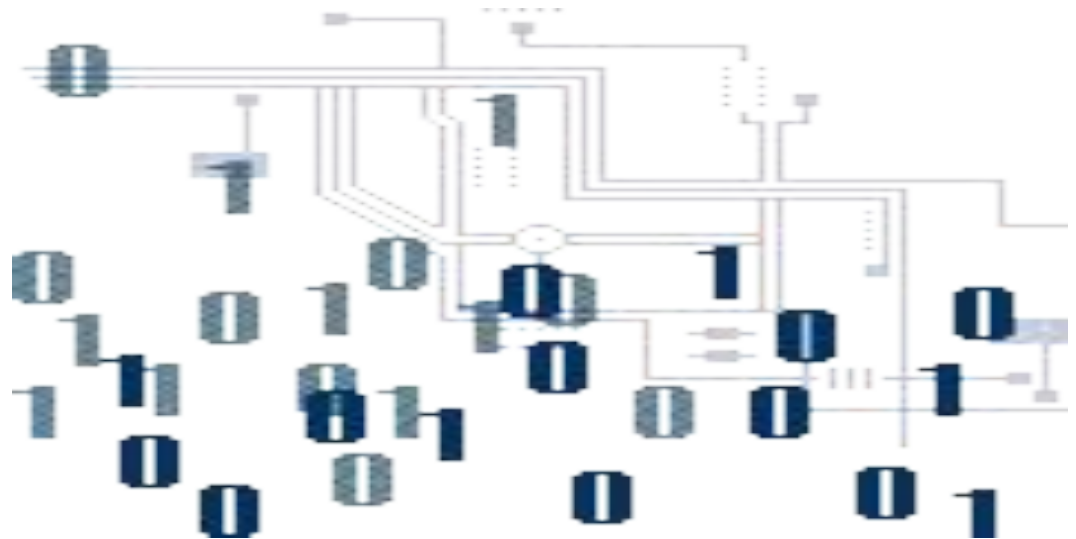


Для кодировки текстовой информации одним из первых стандартов был код **КОИ-8** (код обмена информацией 8-битный)

Эта кодировка применяется в современных компьютерах с операционной системой **UNIX**


Одному символу присваивается код из 8 двоичных разрядов

- М русская большая – 11101101
- М латинская большая – 01001101



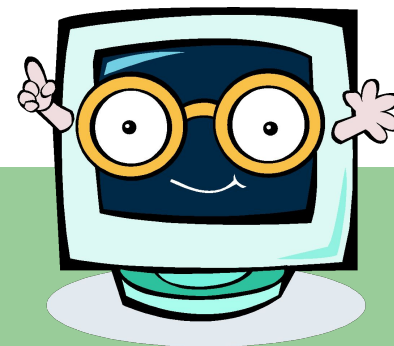
Для 8 бит существует 256
комбинаций (2^8)

от 0000 0000 до 1111 1111



Для каждой цифры, буквы,
символа, а так же пробела
существует своя
комбинация из 8 бит

Единицы измерения количества информации



8 бит = 1 байт


1 Кбайт = 2^{10} байт = 1024 байт

1 Мбайт = 2^{10} Кбайт = 1024^2 байт = 1 048 576 байт

1 Гбайт = 2^{10} Мбайт = 1024^3 байт \approx 1 млрд. байт

Физический объем информации -

это число символов в сообщении,
содержащем информацию,
умноженное на информационную
емкость одного символа



Физический объем информации определяет объем памяти компьютера или дискового пространства, необходимый для ее хранения

Задание:

Посчитайте количество бит и байт в следующих выражениях:

Мир

3 байта = 24 бит

Миру мир!

9 байт = 72 бит

Vile, vide, vice

16 байт = 128 бит



С помощью стандартной программы операционной среды Windows КАЛЬКУЛЯТОР вычислите:

- Сколько бит в 1 Кбайт
- Сколько байт в 1 Гбайт

Ответ:

1 Кбайт=1 024 байт=8 192 бит

1 Гбайт= 1 024³ байт= 1 073 741 824 байт



Содержательный подход к измерению информации



Неопределенность знаний о некотором событии – это количество возможных результатов события.

Сообщение, уменьшающее неопределенность в 2 раза, несет 1 бит информации.

Количество информации i ,
содержащейся в сообщении о том, что
произошло одно из равновероятных
событий, определяется из решения
показательного уравнения:

Формула Р.
Хартли

$$2^i = N \text{ или } I = \log_2 N$$

