

Тексты в памяти компьютера



6 класс

Двоичное кодирование
текстовой информации


Как компьютер распознает текстовые символы?

T → 210 → 11010010

- При двоичном кодировании текстовой информации каждому символу в соответствие ставится **числовой код**.

Как компьютер распознает текстовые символы?

T → 210 → 11010010

- При двоичном кодировании тестовой информации каждому символу в соответствие ставится 

Соответствие символов и кодов

- Соответствие символов и кодов задается с помощью специальных кодовых таблиц.


Символ	Десятичный код	Двоичный код
А	192	11000000
Б	193	11000001
В	194	11000010
...
М	204	11001100

Сколько всего нужно символов?

- **В текстах мы используем:**
- прописные и строчные русские буквы Аа Бб Вв ...
- прописные и строчные латинские буквы Аа Вв Сс ...
- знаки препинания ! , ?
- цифры 0 1 2 3 ...
- знаки арифметических операций + - × ...
- другие символы ([\ ...

Достаточно 256 различных символов.



Символы и байты

- Чаще всего для кодирования используется цепочка из 8 нулей и единиц, называемая 

Фрагмент кодовой таблицы СР 1251

Символ	Десятичный код	Двоичный код
Пробел	32	00100000
!	33	00100001
1	49	00110001
2	50	00110010
А	192	11000000
Б	193	11000001
В	194	11000010
К	202	11001010
Л	203	11001011

Соответствие символов и кодов

- Соответствие символов и кодов задается с помощью специальных 
- Достаточно  различных символов.

Как закодировать слово?

Л	У	Н	А
203	211	205	192
11001011	11010011	11001101	11000000

11001011110100111100110111000000

Как узнать, какое слово закодировано
двоичной последовательностью?

11010000110011101101000111000000

Как узнать, какое слово закодировано
двоичной последовательностью?

?	?	?	?
11010000	11001110	11010001	11000000

11010000.11001110.11010001.11000000

Как узнать, какое слово закодировано
двоичной последовательностью?

?	?	?	?
208	206	209	192
11010000	11001110	11010001	11000000

11010000.11001110.11010001.11000000

Как узнать, какое слово закодировано
двоичной последовательностью?

Р	О	С	А
208	206	209	192
11010000	11001110	11010001	11000000

11010000.11001110.11010001.11000000

Кодовые таблицы Windows и KOI8-R

Символ	Десятичный код	Двоичный код
А	192	11000000
Б	193	11000001
В	194	11000010
...
М	204	11001100

Символ	Десятичный код	Двоичный код
А	225	11000000
Б	226	11000001
В	247	11000010
...
М	237	11101101

Арабская (Windows)

Балтийская (ISO)

Балтийская (Windows)

Центральноевропейская (DOS)

Центральноевропейская (ISO)

Центральноевропейская (Windows)

Китайская упрощенная (GB2312)

Китайская упрощенная (HZ)

Китайская традиционная (Big5)

Кириллица (DOS)

Кириллица (ISO)

Кириллица (KOI8-R)





Кириллица (KOI8-U)

Греческая (ISO)

Греческая (Windows)

Иврит (DOS)

Самое главное

- При двоичном кодировании тестовой информации каждому символу в соответствие ставится 
- Чаще всего для кодирования используется цепочка из 8 нулей и единиц, называемая 
- Существует  таких цепочек.
- Соответствие символов и кодов задается с помощью специальной кодовой таблицы


Домашнее задание № 6

- § 1.3, стр. 21-23
- вопр. 5 с. 28 письм.