

Представление текстовой информации в ПК.

Начиная с шестидесятых годов компьютеры все больше стали использоваться для обработки текстовой информации, а в настоящее время основная доля персональных компьютеров занята обработкой именно текстовой информации.

Для кодирования одного символа используется количество информации, равное одному байту, т.е. $I = 1 \text{ байт} = 8 \text{ бит}$.

Если рассматривать символы как возможное событие, то можно вычислить, какое количество различных символов можно закодировать:

$$N = 2^I = 2^8 = 256.$$

Такое количество символов достаточно для представления текстовой информации, включая прописные и заглавные буквы русского и латинского алфавита, цифры, знаки, графические символы и т.д.

Кодирование заключается в том что каждому символу ставится в соответствие уникальный десятичный код от 0 до 255 или соответствующий ему двоичный код 00000000 до 11111111. Таким образом человек различает символы по их начертанию, а компьютер по их коду.

При вводе в компьютер текстовой информации изображение символа преобразуется в его двоичный код. Пользователь нажимает на клавиатуре клавишу с символом – и в компьютер поступает определенная последовательность из восьми электрических импульсов (двоичный код символа). Код символа хранится в оперативной памяти компьютера, где занимает одну ячейку.

В процессе вывода символа на экран происходит обратный процесс – преобразование кода символа в его изображение.

Существует соглашение которое фиксируется в кодовой таблице (ASCII). Первые 33 кода (0..32) обозначают операции перевод строки, ввод пробела и т.д.

Коды 33..127 – интернациональные и соответствуют символам латинского алфавита, цифрам, знакам арифметических операций и знакам препинания.

Коды 128..255 являются национальными, предназначены для кодировки национальных алфавитов, символов псевдографики и т.д.