

3.20. Дискретные модели данных в компьютере. Представление текста, графики и звука

В этом параграфе обсудим способы компьютерного кодирования текстовой, графической и звуковой информации. С текстовой и графической информацией конструкторы ЭВМ «научили» работать машины, начиная с третьего поколения (1970-е годы). А работу со звуком «освоили» лишь машины четвертого поколения, современные персональные компьютеры. С этого момента началось распространение технологии мультимедиа.

Что принципиально нового появлялось в устройстве компьютеров с освоением ими новых видов информации? Главным образом, это периферийные устройства для ввода и вывода текстов, графики, видео, звука. Процессор же и оперативная память по своим функциям изменились мало. Существенно возросло их быстродействие, объём памяти. Но, как это было на первых поколениях ЭВМ, так и осталось на современных ПК - основным навыком процессора в обработке данных является умение выполнять вычисления с двоичными числами. Обработка текста, графики и звука представляет собой тоже обработку числовых данных. Если сказать ещё точнее - то это *обработка целых чисел*. По этой причине компьютерные технологии называют *цифровыми технологиями*.



О том, как тексты, графика и звук сводятся к целым числам, будет рассказано дальше. Предварительно отметим, что здесь мы снова встретимся с *главной формулой информатики*:

$$2^i=N$$

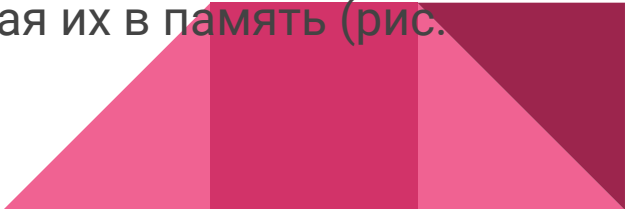
Смысл входящих в неё величин следующий: i - разрядность ячейки памяти (в битах), N - количество различных целых положительных чисел, которые можно записать в эту ячейку.



Текстовая информация

Принципиально важно, что текстовая информация уже дискретна - состоит из отдельных знаков. Поэтому возникает лишь технический вопрос, как разместить её в памяти компьютера.

Напомним о байтовом принципе организации памяти компьютера, обсуждавшемся в базовом курсе информатики. Вернёмся к рис. 4.7. Каждая клеточка на нём обозначает бит памяти. Восемь подряд расположенных битов образуют байт памяти. Байты пронумерованы. Порядковый номер байта определяет его адрес в памяти компьютера. Именно по адресам процессор обращается к данным, читая или записывая их в память (рис. 4.12).





байт № 1

байт № 2

Рис. 4.12. Байтовая организация памяти

Модель представления текста в памяти весьма проста. За каждой буквой алфавита, цифрой, знаком препинания и иным общепринятым при записи текста символом закрепляется определённый двоичный код, длина которого фиксирована. В популярных системах кодировки (ASCII, KOI8 и др.) каждый символ заменяется на 8-разрядное целое положительное двоичное число; оно хранится в одном байте памяти. Это число является порядковым номером символа в кодовой таблице. Согласно главной формуле информатики, определяем, что размер алфавита, который можно закодировать, равен: $2^8 = 256$. Этого количества вполне достаточно для размещения двух алфавитов естественных языков (английского и русского) и всех необходимых дополнительных символов.