

ТЕКСТОВАЯ ИНФОРМАЦИЯ В ПАМЯТИ КОМПЬЮТЕРА

Источник информатика и ИКТ: учебник для 7 класса
И. Г. Семакин

Оглавление

Преимущества компьютерного текста	3
Представление текста в памяти компьютера	4
Что такое гипертекст	7
Обработка текста с помощью компьютера.....	8
Коротко о главном	9

Тексты в компьютерной памяти

Преимущества компьютерного текста

При ручной записи часто неприятную проблему составляет необходимость исправлять ошибки или вносить какие-то изменения в текст. При этом приходится зачеркивать, стирать, заклеивать, что портит текст. Необходимость переписывать текст ведет к потере времени и лишнему расходу бумаги.

Имея компьютер, можно создавать тексты, не тратя на это лишнее время и бумагу. Носителем текста становится память компьютера. Конечно, для длительного его сохранения это должна быть внешняя память. Тексты на внешних носителях сохраняются в файлах.

Есть еще ряд преимуществ сохранения текстов в файлах на компьютерных носителях по сравнению с бумагой. Во-первых, это компактное размещение. Например, на компакт-диске (700 Мб) можно разместить тексты более сотни книг объемом в 500 страниц каждая. А если использовать специальные методы сжатия, то это количество можно увеличить в несколько раз.

Во-вторых, если данный текст становится ненужным, то с помощью компьютера его легко удалить с носителя, поместив на это место другой файл.

В-третьих, с помощью компьютера легко скопировать файлы в любом количестве на другие носители.

В-четвертых, файл с текстом можно быстро переслать другому человеку по электронной почте. Для этого ваш компьютер и компьютер адресата должны иметь связь через компьютерную сеть.

Главное неудобство хранения текстов в файлах состоит в том, что прочитать их можно только с помощью компьютера. Человек может просмотреть текст на экране монитора или напечатать, используя принтер.

Уже сейчас существуют издания, которые не печатаются на бумаге, а хранятся и распространяются в форме файлов. С распространением компьютеров число таких безбумажных изданий с каждым годом увеличивается. Представьте себе, что вся ваша личная библиотека разместится в коробке с дисками. Причем по объему информации она будет не меньше, чем сотни книг, собранных родителями. А экономя бумагу, мы сохраняем леса на нашей планете.

Представление текста в памяти компьютера

Разберемся как же представлена текстовая информация в памяти компьютера/файлах.

Текстовая информация состоит из символов: букв, цифр, знаков препинания, скобок и др. Мы уже говорили, что множество всех символов, с помощью которых записывается текст, называется алфавитом, а число символов в алфавите – его мощностью.

Широко распространенным способом представления текстовой информации в компьютере является использование алфавита мощностью 256 символов. Один символ такого алфавита несет 8 битов информации: $2^8 = 256$. 8 бит = 1 байт, следовательно, двоичный код каждого символа занимает 1 байт (8 бит) памяти компьютера.

Какой именно восьмиразрядный двоичный код поставить в соответствие тому или иному символу.

Все символы компьютерного алфавита пронумерованы от 0 до 255. Каждому номеру соответствует восьмиразрядный двоичный код от 00000000 до 11111111. Этот код – порядковый номер символа в двоичной системе счисления.

Таблица, в которой всем символам компьютерного алфавита поставлены в соответствие порядковые номера, называется **таблицей кодировки**.

На ЭВМ (электронно-вычислительная машина) первых поколений для разных типов машин использовались различные таблицы кодировки. С распространением персональных компьютеров типа IBM PC международным стандартом стала **таблица кодировки** под названием **ASCII** (American Standard Code for Information Interchange – американский стандартный код для обмена информацией). Точнее говоря, стандартной в этой таблице является только первая половина, т. е. символы с номерами от нуля (двоичный код 00000000) до 127 (01111111). Сюда входят буквы латинского алфавита, цифры, знаки препинания, скобки и некоторые другие символы. Остальные 128 кодов, от 10000000 до 11111111, составляют так называемую *кодovou страницу*. Например, кодовая страница номер 1251 (Windows-1251) содержит русский алфавит и используется в операционной системе Windows и её приложениях.

В первой таблице приведена стандартная часть кода ASCII (коды от 0 до 31 имеют особое назначение, не отражаются какими-либо знаками и в данную таблицу не включены). Здесь приведены десятичные номера символов, символы.

Обратите внимание на то, что в этой таблице латинские буквы (прописные и строчные) располагаются в алфавитном порядке. Расположение цифр также упорядочено по возрастанию значений. Это правило соблюдается и в других таблицах кодировки и называется принципом последовательного кодирования алфавитов. Благодаря этому понятие «алфавитный порядок» сохраняется и в машинном представлении символической информации. Для русского алфавита принцип последовательного кодирования соблюдается не всегда.

Запишем, например, внутреннее представление слова «file». В памяти компьютера оно займет 4 байта со следующим содержанием:

01100110 01101001 01101100 01100101.

А теперь попробуйте решить обратную задачу. Какое слово записано следующим двоичным кодом:

01100100 01101001 01110011 01101011?

Во второй таблице приведена кодовая страница Windows-1251. Видно, что в ней для букв русского алфавита соблюдается принцип последовательного кодирования.

Помимо восьмиразрядной кодировки символов все большее распространение получает шестнадцатиразрядная – двухбайтовая кодировка. Международный стандарт такой кодировки носит название UNICODE.

Таблица 1.1. Базовая таблица кодировки ASCII

32 пробел	48 0	64 @	80 P	96 `	112 p
33 !	49 1	65 A	81 Q	97 a	113 q
34 "	50 2	66 B	82 R	98 b	114 r
35 #	51 3	67 C	83 S	99 c	115 s
36 \$	52 4	68 D	84 T	100 d	116 t
37 %	53 5	69 E	85 U	101 e	117 u
38 &	54 6	70 F	86 V	102 f	118 v
39 ' .	55 7	71 G	87 W	103 g	119 w
40 (56 8	72 H	88 X	104 h	120 x
41)	57 9	73 I	89 Y	105 i	121 y
42 , * .	58 :	74 J	90 Z	106 j	122 z
43 +	59 ;	75 K	91 [107 k	123 {
44 ,	60 <	76 L	92 \	108 l	124
45 -	61 =	77 M	93]	109 m	125 }
46 .	62 >	78 N	94 ^	110 n	126 ~
47 /	63 ?	79 O	95 _	111 o	127

Тексты вводятся в память компьютера с помощью клавиатуры. На клавишах написаны привычные нам буквы, цифры, знаки препинания и другие символы. В оперативную память они попадают в форме двоичного кода.

Из памяти компьютера текст может быть выведен на экран или на печать в символьной форме. Но для длительного хранения его следует записать на внешний носитель в виде файла.

Что такое гипертекст

Наиболее существенное отличие компьютерного текста от бумажного вы почувствуете, если встретитесь с текстом, информация в котором организована по принципу гипертекста.

Гипертекст – это текст, организованный так, что его можно просматривать в последовательности смысловых связей между его отдельными фрагментами. Такие связи называются **гиперсвязями** (*гиперссылками*).

Чаще всего по принципу гипертекста организованы компьютерные справочники, энциклопедии, учебники. Такую «книгу» можно читать не только в обычном порядке, «листая страницы» на экране, но и перемещаясь по смысловым связям в произвольном порядке. Например, при изучении на уроке

физики темы «Второй закон Ньютона» с помощью компьютерного учебника ученик прочитал определение закона «Сила равна произведению массы на ускорение». Ему захотелось вспомнить определение массы. Указав в тексте на слово «масса» (связанные понятия обычно выделяются цветом или подчеркиванием, а указывать на них удобно с помощью мыши), он быстро перейдет к разделу учебника, где рассказывается о массе тел. Прочитав определение «Масса – мера инертности тела», ученик может пожелать уточнить, что такое инертность. По гиперссылке он быстро выйдет на нужный раздел.

После такой экскурсии вглубь материала ученик может вернуться в исходную точку, щелкнув мышью на кнопке «Назад», так как система запоминает весь маршрут продвижения по гиперссылкам.

Обработка текста с помощью компьютера

Для работы с текстовыми документами существуют прикладные программы, которые называются текстовыми редакторами.

Текстовый редактор (ТР) — это прикладная программа, позволяющая создавать текстовые документы, редактировать их, просматривать содержимое документа на экране, распечатывать документ.

По отношению к текстовым редакторам с широкими возможностями форматирования текста, включения графики, проверки правописания часто применяется название **текстовый процессор**.

Существует множество текстовых редакторов – от простейших учебных до мощных издательских систем, с помощью которых делают книги, газеты, журналы. Примеры: текстовые редакторы Microsoft Word (ОС Windows) и OpenOffice.org Writer (ОС Linux).

Коротко о главном

С помощью компьютера можно создавать текстовые документы и хранить их на носителях внешней памяти в виде файлов.

Преимущества файлового хранения текстов: возможность редактирования, быстрого копирования на другие носители, передачи текста по линиям компьютерной связи.

Для кодирования текстов используется 8-разрядный или 16-разрядный двоичный код. При 8-разрядном кодировании используемый алфавит содержит 256 символов.

В таблице кодировки каждому символу алфавита поставлен в соответствие порядковый номер и восьмиразрядный двоичный код.