|  |
| --- |
|  |

|  |
| --- |
| Информационные технологии в обработке текстов |
| [Введите подзаголовок документа] |
|  |
|  |
|  |
| **Кадырова Залина** |
| **15.10.2018** |
|  |

*Оглавление:*

Представление текста………………………………………………………………..…..2-3

Текстовые файлы……………………………………………………………………3-4

Правила машинописного набора текста……………………………….4-5

Оформление текста…………………………………………………………….….6-8 Автоматизированная обработка текста……………………………………8 Поиск……………………………………………………………………………………….…9 Расшифровка или уточнение значений слова……………………..9-10 Системы автоматизированной доработки текста…………………..10 Возможности текстовых процессоров……………………………..…11-12 Текстовые редакторы………………………………………………….……...12-16

**Представление текста**

Представление информации в виде текста стало одним из первых доступных для обработки с помощью ЭВМ и до сих пор остается одним из наиболее универсальных. Энциклопедический словарь дает такое определение понятию “текст”: “**Текст** -- это*упорядоченный набор слов, предназначенный для того, чтобы выразить некий смысл. В лингвистике термин используется в широком значении, включая в себя и устную речь*”. Марчук Ю.Н. Основы компьютерной лингвистики. М., 2000.

Представление информации в виде текста при обработке с помощью вычислительной техники близко к этому определению. Под “текстовым” понимают такое представление информации, в котором она представлена в виде записи слов (логических элементов) некоторого языка и доступна для чтения человеком.

Язык для такого представления характеризуется некоторым алфавитом -- т.е. допустимым набором символов. Поскольку компьютер работает только с двоичным кодом, то для записи и обработки требуется взаимно-однозначно сопоставить символы и двоичные коды. *Правило сопоставления кодов и символов, входящих в алфавит, называется* *кодировкой.*

Первый широко распространенный стандарт кодирования -- таблица (т.е. прямое сопоставление кодов символам) кодировки ASСII (***A****merican****S****tandard****C****ode for****I****nformation****I****nterchange,* американский стандартный код для обмена информацией) -- был разработан в 1963 году. Стандарт предполагал использование не только в вычислительной технике, но и в телеграфии. В нем для кодирования каждого символа отводилось 7 бит. Восьмой бит использовался для служебных целей -- контроля четности при передаче.

Эта часть таблицы кодировки содержит символы латинского алфавита, цифры, некоторые знаки препинания и набор управляющих символов (возврат каретки, перевод строки, конец файла, сигнал и т.п.).

Позже восьмой бит стали использовать для представления символов национальных алфавитов: первая часть таблицы -- US-ASCII -- использовалась по-прежнему, а содержание второй менялось в зависимости от исходного естественного языка. Каждый вариант этой второй половины (расширенной таблицы) исходной таблицы получил название “кодовой страницы” языка (code page).

Для русского языка таких расширений несколько. Наиболее известны: CP866 (DOS), KOI-8R (UNIX), CP1251 (Windows) и MacCyr.

Применение такого способа кодирования сильно затрудняет передачу текстовых сообщений между разными странами, объединение в сообщении текста на нескольких языках, а в случае с русским языком -- и обмен файлами между разными ОС (для русского языка до сих пор активно применяется 4 разных кодовых таблицы). Для решения этих проблем в 1991 году некоммерческим объединением был предложен стандарт кодирования Юникод (Unicode).

Стандарт состоит из двух частей: универсального набора символов (Universal Character Set) и правил трансформации (Unicode Transformation Format). Универсальный набор символов предполагает описание всех возможных при записи текстов символов в виде общей таблицы кодов. Правила трансформации определяют способ записи этих кодов. Потапова Р.К. Новые информационные технологии и лингвистика. М., 2002.

Первая версия стандарта предполагала использование двух байтов для кодирования каждого символа. В дальнейшем это кодовое пространство было расширено.

Сейчас чаще всего применяется способ трансформации UTF-8, обеспечивающий совместимость с предыдущими реализациями и стандартами. В частности, коды менее 128 записываются одним байтом, что автоматически превращает их в коды ASCII.

Применение этого стандарта кодирования позволяет объединять в одном тексте слова на различных языках (без ограничений на их количество), использовать устаревшие языки, дополнительные символы.

Наиболее переносимым и легко используемым с технической точки зрения способом хранения и передачи текста являются текстовые файлы. По сути, эти файлы представляют собой последовательности символов, разбитых на абзацы

**Текстовые файлы**

Понятие “текстового файла” не предусматривает строго заданного формата или расширения. Тем не менее, помимо характерной для той или иной ОС таблицы кодировки, в текстовых файлах могут применяться три основных способа деления текста на строки (абзацы):

Текстовые файлы применяются для самых различных целей и часто оказываются формой хранения данных, описанных более сложными формальными языками. Эти файлы часто используются для записи конфигурации ПО, документирования, переноса данных, описания HTML- или XML-кода. текст обработка шрифт автоматизированный.

**Правила машинописного набора текста**

Для облегчения анализа и последующего преобразования текста при его наборе в самых различных случаях рекомендуется соблюдать общие правила машинописного набора:

**1**. Все слова разделяются пробелом, и только одним пробелом.

**2.** Знаки препинания примыкают к предыдущему слову.

**3.** Скобки и кавычки всех видов примыкают к первому и последнему слову заключенного в них текста.

**4.** Текст разрывается только в конце абзаца.

**5.** Большие форматированные пробелы делаются вставкой символа табуляции, а не несколькими пробелами подряд.

Соблюдение этих правил позволяет легко использовать текст при подготовке более сложных документов, в которые он входит как важнейший элемент, или при организации автоматической обработки.

Текст может появиться из самых разных источников. Чаще всего текстовую информацию вводят с помощью клавиатуры. Стандартная клавиатура и программа, принимающая от нее информацию о нажатых клавишах, позволяют вводить текст, указывать место ввода в уже введенном тексте (перемещая маркер места ввода клавишами перемещения курсора либо с помощью мыши) и удалять неверно введенные символы слева или справа от курсора (с помощью клавиш и ).

Возможность исправлять ошибки и набирать текст постепенно стала одной из существенных причин, по которым подготовка текстовой информации практически повсеместно была переведена с бумажной на компьютерную основу.

Текстовые редакторы с развитыми возможностями предоставляют пользователям возможность протоколировать и сохранять наборы действий -- создавать макрокоманды, или **макросы**. Использование макросов позволяет ускорить выполнение частых простых задач обработки.

Специализированные программы, основной задачей которых является обеспечение набора текста, разделяют на **текстовые редакторы**, т.е. *программы, которые помогают именно подготовить тот или иной специфический текст, но не оформить его для печати*, и **текстовые процессоры** -- более сложные *программные комплексы, позволяющие выполнить оформление текста, точно задать его расположение, сопроводить его графическими материалами* и т.д. Степанов М.Ф. Машинный перевод и общение на естественном языке.

**Оформление текста**

**Шрифты**

Чаще всего текстовая информация используется при подготовке различных печатных материалов. Конечная цель подготовки такого материала -- его печать или точное изображение печатной страницы на экране. В отличие от простой подготовки текстового файла, при подготовке печатного материала важно, как отображается текст. Практически все основные элементы и приемы оформления текстовых материалов заимствованы у давно существующих технологий -- печатных, оттуда же пришла и основная часть терминологии.

Основным и наиболее важным средством определения внешнего вида текста является **шрифт**(*schreiben*, от нем. -- “писать”). Шрифт -- это графический рисунок букв, цифр и символов, обладающий общими для всех символов стилистическими особенностями изображения.

Отдельный символ контурного шрифта с обозначенными элементами рисунка.

***Шрифт характеризуется рядом параметров:***

**1.** Рисунок шрифта -- графические особенности, определяющие общность шрифта и его отличие от всех других.

**2.** Кегль -- размер шрифта -- предельная высота большой буквы и окружающих ее пробелов (термин введен для описания высоты площадки литеры при наборе с помощью типографской кассы). Чаще всего задается в типографских пунктах (1 пункт = 1/72 дюйма = 0,375 мм). По историческим причинам некоторые размеры имеют собственные названия: 8 пт -- “петит”, 9 пт -- “боргес”, 10 пт -- “корпус”, 12 пт -- “цицеро”.

**3.** Начертание -- шрифт с общим рисунком, но какими-либо отличительными признаками: более жирный, наклонный, разреженный. Иногда параметр плотности шрифта (светлый, полужирный, жирный) отделяют от начертания.

**4.** Часто как параметр задается подчеркивание или зачеркивание шрифта, или его написание как индекса -- с уменьшением размера и подъемом/спуском относительно текущей строки.

Совокупность всех возможных размеров и вариантов написания шрифта называется **гарнитурой**. Гарнитуры имеют имена, по которым часто называют и конкретный шрифт.

По общим чертам рисунка различают три основных вида шрифтов:

Шрифт задается для набранного текста и не изменяет самих символов -- он только определяет написание каждого символа, исходя из эталонного изображения. Библиотека таких изображений называется просто “шрифтом”.

Существует несколько основных способов описания шрифтов (точнее -- гарнитуры шрифта):

При таком способе каждая буква описывается отдельно, как некоторая матрица точек. Способ позволяет максимально ускорить обработку, но сильно затрудняет изменение размеров или начертаний. Для достижения качества каждый символ такой гарнитуры должен быть отредактирован вручную и должен храниться отдельно.

*Векторные шрифты.* При таком способе описания шрифт задается с помощью некоторых математических кривых, совокупность которых и составляет рисунок символов. Такой шрифт может изменять размеры без потери качества, но с помощью примитивов трудно добиться прорисовывания заполняемых элементов.

*Контурные шрифты*. Аналогично векторным, описываются с помощью некоторых математических кривых, но они определяют не символ, а его контур, который заполняется по определенным правилам. Именно этот тип шрифтов и является наиболее популярным.

Для использования векторных и контурных шрифтов необходимо выполнение операции, “создающей” шрифт (заданного рисунка, размера и начертания), годного для отображения. Такая операция называется “растеризацией”. В состав графических оболочек современных операционных систем входят программы -- растеризаторы шрифтов определенного формата.

Наиболее популярные форматы шрифтов -- это TrueTypeFonts (TTF, поддерживается ОС Windows и MacOS) и PostScript (разработан фирмой Adobe, для использования необходима программа Adobe TypeManager). Сейчас на смену этим форматам приходит совместно разработанный этими компаниями формат OpenType.

Растеризация шрифта -- достаточно ресурсоемкая операция, поэтому контурные шрифты получили распространение только с началом массового применения достаточно мощных компьютеров.

Структурирование теста

Помимо внешнего вида букв, важное значение имеет пространственное расположение текста. Единицей пространственного размещения служит **абзац**. Как и в литературе, в компьютерном тексте абзацем называется выделенный по смыслу участок.

Для оформления абзаца используют несколько параметров:

***1. Выравнивание*** (выключка) -- *правило расположения букв в строке абзаца*. Видов выравнивания четыре: по левому краю, центральное, по правому краю и по ширине полосы набора.

***2. Отступы*** от краев полосы набора.

***3. Абзацный отступ*** (красная строка) -- *положение первой строки абзаца*.

***4. Интервалы.***Различают межстрочное расстояние-- задается множителем размера шрифта (одинарный, полуторный, двойной интервал) -- и промежутки до и после абзаца.

***5. Буквица*** -- *крупная выступающая первая буква абзаца*. Часто задается не просто более крупным размером буквы, но и буквой другого рисунка.

Абзацы размещаются в рамках **полосы** -- *выделенного участка страницы, как правило, прямоугольной формы, в котором размещаются текст и иллюстрации*. На листе может быть либо одно такое место (одна колонка), либо несколько -- тогда говорят о многоколоночном тексте.

Как правило, текстовые процессоры не дают появляться *висячим строкам*--*отдельным строкам абзацев в начале или конце полосы*.

Важным элементом оформления текста на странице являются **поля** -- пробелы вдоль края страницы и интервалы между колонками. Для удобного чтения, в силу особенностей восприятия, такие пробелы должны быть обязательно.

Как правило, в достаточно большом (больше нескольких страниц) тексте выделяется несколько смысловых блоков (разделов) и видов содержательного текста -- обычный текст, примечания, ссылки и т.п.

Для оформления таких элементов создаются *стили* -- *определенные наборы параметров оформления шрифта и абзацев*. Применение стилей позволяет ускорить набор, автоматизировать оформление (например, автоматически создавать оглавления) и изменять внешний вид различных элементов, не разыскивая их по всему тексту. Практически все современные текстовые процессоры опираются на стили, даже если пользователь не использует их. Единство оформления -- одно из условий удобочитаемости и красоты печатного издания.

Для решения некоторых типовых задач оформления текстов существующие текстовые процессоры предусматривают два мощных средства автоматизации.

**1.** **Списки**. При оформлении текста это *набор визуально выделенных элементов перечисления*. Элементы выделяют с помощью символа-маркера (**маркированные списки**) либо номером -- в упорядоченных списках. При оформлении списка чаще всего также предусматривают форматирование абзацев -- так, чтобы они не выступали за маркер. Автоматизация оформления позволяет автоматически маркировать и выделять новые элементы списков.

**2.** **Таблицы**. Современные текстовые процессоры предусматривают средства для создания двухмерной структуры размещения информации. Применение таких средств позволяет редактировать структуру и содержание таблицы, добавлять строки и столбцы, изменять их линейные размеры, выделять их с помощью сетки или фона. Фактически каждая ячейка таблицы становится листом в миниатюре.

Стоит отметить, что файлы текстового процессора содержат массу дополнительных (по отношению к тексту) данных об оформлении и текстовыми очень часто не являются.

Как и текстовые редакторы, текстовые процессоры обладают средствами создания макрокоманд. Современные процессоры реализуют их на развитом языке сценариев, позволяющем решать довольно сложные задачи преобразования и оформления публикаций.

**Автоматизированная обработка текста**

Текстовое представление информации -- одно из наиболее удобных для организации автоматической обработки. Связано это с тем, что в этой форме информация представляется в виде близком к исходному языку, что позволяет выполнять преобразования, связанные со смыслом текста.

Существует несколько наиболее распространенных автоматизированных операций, связанных с текстовым представлением.

**Поиск**

Задача поиска необходимой информации чаще всего формулируется как поиск фрагментов, содержащих некоторые понятия, в достаточно большом массиве. Большое значение этот вид автоматической обработки получил с ростом популярности межсетевой среды Интернет. Существует несколько подходов к организации такого поиска.

Первый подход опирается на поиск фрагмента текста, соответствующего некоторому образцу. Наиболее популярная форма задания этого образца -- так называемые **регулярные выражения**. По сути, это *описание фрагмента текста, удовлетворяющего некоторым условиям*, по тем частям, которые в нем содержатся, и их порядку. Таким способом в большом текстовом массиве можно находить упоминания тех или иных слов, адреса, номера телефонов и т.п. шаблонные элементы.

Достоинство этого подхода -- возможность применять его к массиву текста без предварительной обработки. Например, сразу при посимвольном получении текста.

Второй подход предусматривает предварительное создание *специального вида базы для ускорения поиска* -- *индекса*. Такой способ применяется для ускорения поиска, если некоторые типовые поисковые запросы повторяются часто и нет возможности формировать/хранить весь массив текста. Например, при организации поисковой машины в среде Интернет. Романенко В. Н., Никитина Г. В. Сетевой информационный поиск: Информация в Интернете; Поисковые машины; Электронные каталоги библиотек; Как формулировать запросы: Практическое пособие.

**Расшифровка или уточнение значений слова**

Для решения такой задачи в самых разных видах применяют *словари*-- *базы информационных фрагментов, связанных с некоторыми ключевыми словами или словосочетаниями*. Примером таких баз могут быть словари различных языков: англо-русский, русско-английский, толковый и другие виды словарей. Одно из самых распространенных применений словарей -- проверка правописания слов при наборе.

Особым видом словарей являются *тезаурусы* -- *словари, в которых слова связываются на основе каких-либо лексических отношений*. Например: слова являются синонимами (смысловыми аналогами), антонимами (противоположны по смыслу) и т.п.

Этот вид словарей важен не только потому, что может помочь при подготовке текстов, но и потому, что это отразит смысл слов -- для систем, моделирующих отдельные аспекты мышления человека.

**Системы автоматизированной доработки текста**

Используя закономерности естественного языка и описанные выше средства выполнения некоторых операций и выявления зависимостей, с помощью ЭВМ автоматизируют некоторые операции по смысловому преобразованию текста. Современные системы обработки позволяют создавать краткие обзоры текстов (рефераты) или готовить перевод с одного естественного языка на другой. Приходится отметить, что точного решения эти задачи не имеют, поскольку зачастую трудно подобрать адекватное слово или выражение, учитывая не только формальный перевод, но и грамматические особенности, и культурные. Тем не менее с применением специализированных по областям знания словарей современные системы автоматизированного перевода создают подстрочник, который может дать представление о смысле текста и в дальнейшем помочь переводчику в переводе документа.

**Возможности текстовых процессоров**

Редакторы текстов программ рассчитаны на редактирование программ на том или ином языке программирования. Редакторы текста, и рассчитаны на тексты программ, и выполняют следующие функции:

- диалоговый просмотр текста;

- редактирование строк программы;

- копирование и перенос блоков текста из одного места в другое;

- копирование одной программы или её части в указанное место другой программы;

- контекстный поиск и замену подстрок текста;

- автоматический поиск строки, содержащей ошибку;

- распечатку программы или её необходимой её части;

**Редакторы документов** - программы для обработки документов, ориентированные на работу с текстами, имеющие структуру документа, т. е. состоящими из разделов, страниц, абзацев, предложений, слов и т.д. Следовательно, редакторы для обработки документов обеспечивают такие функции, как:

- возможность использования различных шрифтов символов;

- задание произвольных межстрочных промежутков;

- автоматический перенос слов на следующую строку;

- автоматическую нумерацию страниц;

- обработку и нумерацию строк;

- печать верхних и нижних заголовков страниц (колонтитулов);

- выравнивание краев абзаца;

- набор текста в несколько столбцов;

- создание таблиц и построение диаграмм;

- проверку правописания и подбор символов;

Редакторы текстов программ, можно использовать для создания и корректировки небольших документов. Однако при необходимости серьезной работы с документами лучше использовать редакторы, ориентированные на работу с документами.

Современные текстовые процессоры предоставляют пользователю широкие возможности по подготовке документов. Это и функции редактирования, допускающие возможность любого изменения, вставки, замены, копирования и перемещения фрагментов в рамках одного документа и между различными документами, контекстного поиска, функции форматирования символов, абзацев, страниц, разделов документа, верстки, проверки грамматики и орфографии, использования наряду с простыми текстовыми элементами списков, таблиц, рисунков, графиков и диаграмм.

Значительное сокращение времени подготовки документов обеспечивают такие средства автоматизации набора текста, как автотекст и автозамена, использование форм, шаблонов и мастеров типовых документов.

Наличие внешней памяти компьютера обеспечивает удобное длительное хранение подготовленных ранее документов, быстрый доступ к ним в любое время.

Существенно упрощают процедуру ввода данных сканеры и голосовые устройства. Существующие системы распознавания текстов, принимаемых со сканера, включают функцию экспорта документа в текстовые редакторы.

Широкий спектр печатающих устройств в сочетании с функциями подготовки документа к печати, предварительного просмотра, обеспечивает получение высококачественных черно-белых и цветных копий на бумаге и прозрачной пленке.

Таким образом, современные программы предусматривают множество функций, позволяющих готовить текстовую часть документа на типографическом уровне. Кроме того, современные программы позволяют включать в текст графические объекты: рисунки, диаграммы, фотографии.

Благодаря этим возможностям файл, представляющий собой текстовый документ, может содержать, помимо алфавитно-цифровых символов, обширную двоичную информацию о форматировании текста, а также графические объекты.

**Текстовые редакторы**

**Текстовый редактор Лексикон.**

Текстовый редактор Лексикон предназначен для обработки несложных документов с текстом на русском и английском языках. У ЛЕКСИКОНА можно выделить следующие функции редактирования текстов:

- просмотр и корректировку текста;

- автоматическое форматирование текста;

- автоматическое разбиение текста на страницы;

- перенос фрагментов текста из одного места в другое;

- создание оглавления разделов документа;

- одновременное редактирование нескольких документов;

Особенностями Лексикона, отличающим его от других текстовых редакторов, являются:

- многооконность (ТП Лексикон имеет 10 окон, в каждом из которых может находиться независимый текст, что позволяет выполнять комплексную обработку сложной системы документации или совокупности программных модулей).

- макропоследовательности (макроклавиши) - программирование макрокоманд позволяет определить часто используемые текстовые конструкции или последовательности команд путем одного нажатия клавиши.

- калькулятор (встроенный экранный калькулятор позволяет производить вычисления с числами по разным основаниям, вычислять арифметические выражения, использовать элементарные функции).

**Текстовый редактор Word Pad.**

Редактор Word Pad - текстовый редактор, позволяющий создавать и редактировать достаточно длинные тексты.

Редактор Word Pad, входящий в состав Windows, не является достаточно мощным, но с большинством повседневных задач - с написанием письма, студенческого реферата, изготовлением поздравительной открытки - Word Pad справляется вполне успешно.

С помощью Word Pad можно:

- работать со шрифтами, использовать разнообразное начертание и цвет шрифтов;

- сохранять тексты как в своем собственном формате, так и в других популярных форматах (в том числе в формате Microsoft Word);

- вставлять в текст картинки разнообразных форматов.

**Текстовый редактор Word Perfect.**

Текстовый редактор Word Perfect - один из самых популярных текстовых редакторов в компьютерном мире. Word Perfect обладает довольно необычным набором свойств. Используя одни и те же положенные в его основу универсальные принципы, можно с одной стороны довольно просто обрабатывать несложные документы, а с другой - готовить тексты со сложной структурой, иллюстрациями, таблицами и т.п.

Перечислим некоторые основные характерные черты редактора Word Perfect:

- содержит все традиционные действия по редактированию текста

- постоянно поддерживает формат текста

- обладает широким набором встроенных символов

- имеется возможность автоматически строить содержание, списки рисунков и литературы, различные указатели, проставлять колонтитулы и т.д.

- стоит математические формулы, вставляет в текст рисунки, графики, таблицы и т.д.

**Текстовый редактор Microsoft Word 6.0.**

Microsoft Word 6.0 - мощный текстовой процессор, предназначенный для выполнения всех процессов обработки текста: от набора и верстки, до проверки орфографии, вставки в текст графики в стандарте \*.pcx или \*.bmp, распечатки текста. Он работает со многими шрифтами, как с русским, так и с любым из двадцати одного языка мира. В одно из многих полезных свойств Word входит автоматическая коррекция текста по границам, автоматический перенос слов, а также проверка правописания слов, сохранение текста в определенный устанавливаемый промежуток времени, наличие мастеров текстов и шаблонов, позволяющих в считанные минуты создать деловое письмо, факс, автобиографию, расписание, календарь и многое другое. Word обеспечивает поиск заданного слова или фрагмента текста, замену его на указанный фрагмент, удаление, копирование во внутренний буфер или замену по шрифту, гарнитуре или размеру шрифта, а так же по надстрочным или по подстрочным символам. Наличие закладки в тексте позволяет быстро перейти к заложенному месту в тексте. Можно так же автоматически включать в текст дату, время создания, обратный адрес и имя написавшего текст. При помощи макрокоманд Word позволяет включать в текст базы данных или объекты графики, музыкальные модули в формате \*.wav. Для ограничения доступа к документу можно установить пароль на текст, который Word будет спрашивать при загрузке текста для выполнения с ним каких-либо действий. Word позволяет открывать много окон для одновременной работы с несколькими текстами, а так же разбить одно активное окно по горизонтали на два и выровнять их.

Текстовый процессор можно управлять с помощью команд основного меню, команд из контекстного меню и панели инструментов.

Каждый пункт основного меню содержит свое подменю, для выполнения различных действий в текстовом редакторе.

**Текстовый редактор Notepad++.**

Notepad++ это редактор текстовых файлов с поддержкой синтаксиса большого количества языков программирования. Программа располагает широким набором опций и отличается минимальным потреблением ресурсов процессора.

Среди продвинутых опций Notepad++ — опция подсветки текста и возможность сворачивания блоков, согласно синтаксису языка программирования. Пользователь может самостоятельно определить синтаксис языка программирования. Есть возможность настроить режим подсветки. Доступно выделение цветом директив и операторов языка программирования.

Notepad++ обеспечивает возможность одновременного просмотра и редактирования нескольких документов. Также Вы можете просматривать и редактировать в двух окнах отображения один и тот же документ в разных местах. Изменение документа в одном окне просмотра будет автоматически перемещено во второе окно просмотра (т.е. вы редактируете один документ, который имеет клона во втором окне просмотра).

Другие возможности Notepad++:

· Авто-завершение набираемого слова.

· Возможность создания собственного списка API функций (или скачать его со страницы загрузки).

· Поддержка регулярных выражений Поиска/Замены.

· Полная поддержка перетягивания фрагментов текста.

· Динамическое изменение окон просмотра.

· Автоматическое определение состояния файла (уведомление об изменении или удалении файла другой программой — с возможностью перезагрузить файл или удалить его из программы).

· Увеличение и уменьшение (масштабирование).

· Поддержка большого количества языков.

· Листинг программ можно печатать в цвете — так, как вы видите на экране в окне редактирования.

· Заметки о документе.

· Выделение скобок при редактировании текста.

· Запись макроса и его выполнение.

**Текстовый редактор Microsoft Word 2003.**

Текстовый редактор Microsoft Word на сегодняшний день является одной из лучших профессиональных программ для обработки текста. Он также представляет собой удобную издательскую систему для верстки и макетирования книг и периодических изданий.

**Microsoft Office** - пакет, предназначенный для выполнения различных операций с документами. В отличие от других аналогичных по функциям программ приложения, входящие в Microsoft Office, интегрированы в единую систему, что обеспечивает их эффективную работу с документами, включающими разные по типу элементы (например, документ Word может содержать таблицу Excel и часть базы данных Access).

Современные текстовые редакторы могут выполнять следующие функции:

- запоминать и вставлять символы или фрагменты текста;

- использовать большое количество различных шрифтов

- автоматически переносит по слогам слова, не умещающиеся в строке;

- проверять и исправлять орфографические и синтаксические ошибки;

- заменять повторяющие слова синонимами;

- вставлять иллюстрации, включая фотографии;

- автоматически заменять символ, слово или фрагмент во всем тексте;

- вставлять таблицы и диаграммы;

- автоматически нумеровать страницы;

- автоматически создавать сноски, алфавитный указатель и оглавление;

- форматировать текст в виде газетных колонок;

- устанавливать размер печатной страницы, полей и отступов и т.д.

В различных редакторах предусмотрены и другие возможности, в частности и встроенный язык программирования, позволяющий автоматизировать многие операции. Любой текстовый редактор позволяет сохранить информацию в виде документы и распечатать ее на бумаге, однако Word умеет гораздо больше.

**Word** - многофункциональная программа для редактирования текстов и графики, создания веб-страниц и обработки документов, в которые, кроме того, легко вставить доже аудио- и видеофрагменты. С помощью Word можно создавать практически любые документы, публиковать их в электронном виде или выводить на печать в виде бумажных копий.