

Файлы и файловая система



ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Что такое файл?



Все программы и данные хранятся в долговременной (внешней) памяти компьютера в виде файлов.

Файл — это определенное количество информации (программа или данные), имеющее имя и хранящееся в долговременной (внешней) памяти.

Имя файла



Имя файла состоит из двух частей, разделенных точкой: собственно *имя файла* и *расширение*, определяющее его тип (программа, данные и т. д.).

Собственно имя файлу дает пользователь, а тип файла обычно задается программой автоматически при его создании.

Пример:

proba.txt

Единицы измерения информации.doc

Соглашение 8.3



До появления операционной системы Windows 95 на большинстве компьютеров IBM PC работала операционная система MS-DOS, в которой действовали весьма строгие правила присвоения имен файлам. Эти правила называют *соглашением 8.3*

По соглашению 8.3 имя файла может состоять из двух частей, разделенных точкой. Первая часть может иметь длину до 8 символов, а вторая часть (после точки) — до 3 символов. Вторая часть, стоящая после точки, называется расширением имени.

При записи имени файла разрешается использовать только буквы английского алфавита и цифры. Начинаться имя должно с буквы. Пробелы и знаки препинания не допускаются, за исключением восклицательного знака (!), тильды (~) и символа подчеркивания (_).

Длинные имена файлов



После введения в действие операционной системы Windows 95 требования к именам файлов стали существенно мягче. Они действуют и во всех последующих версиях операционных систем Windows.

1. Разрешается использовать до 255 символов.
2. Разрешается использовать символы национальных алфавитов, в частности русского.
3. Разрешается использовать пробелы и другие ранее запрещенные символы, за исключением следующих девяти: / \ : * ? " < > |.
4. В имени файла можно использовать несколько точек. Расширением имени считаются все символы, стоящие за последней точкой.

Тип файла	Расширение
Исполняемые программы	exe, com
Текстовые файлы	txt, rtf, doc
Графические файлы	bmp, gif, jpg, png, pds
Web-страницы	htm, html
Звуковые файлы	wav, mp3, midi, kar, ogg
Видеофайлы	avi, mpeg
Код (текст) программы на языках программирования	bas, pas, cpp

Файловая система



На каждом носителе информации (гибком, жестком или лазерном диске) может храниться большое количество файлов. Порядок хранения файлов на диске определяется установленной файловой системой.

Файловая система - это система хранения файлов и организации каталогов.

Одноуровневая файловая система



Для дисков с небольшим количеством файлов (до нескольких десятков) удобно применять *одноуровневую файловую систему*, когда каталог (оглавление диска) представляет собой линейную последовательность имен файлов. Для отыскания файла на диске достаточно указать лишь имя файла.



Многоуровневая иерархическая файловая система



Если на диске хранятся сотни и тысячи файлов, то для удобства поиска файлы организуются в *многоуровневую иерархическую файловую систему*, которая имеет «древовидную» структуру (имеет вид перевернутого дерева).



Путь к файлу

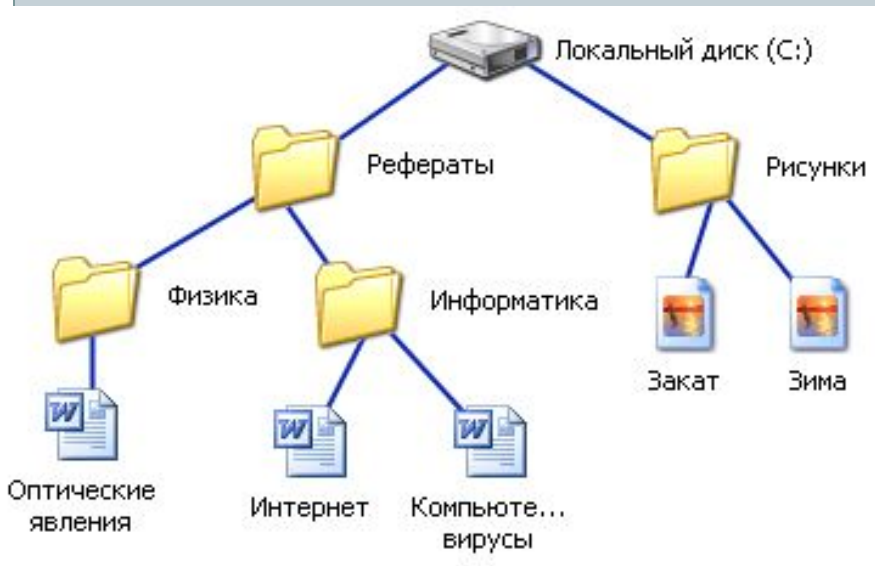


Для того чтобы найти файл в иерархической файловой структуре необходимо указать путь к файлу. В путь к файлу входят записываемые через разделитель "\" логическое имя диска и последовательность имен вложенных друг в друга каталогов, в последнем из которых находится данный нужный файл.



**C:\Рефераты\
C:\Рефераты\Физика\
C:\Рефераты\Информатика\
C:\Рисунки**

Полное имя файла



Путь к файлу вместе с именем файла называют полным именем файла.

C:\Рефераты\Физика\Оптические явления.doc

C:\Рефераты\Информатика\Интернет.doc

C:\Рефераты\Информатика\Компьютерные вирусы.doc

C:\Рисунки\Закат.jpg

C:\Рисунки\Зима.jpg

Папка

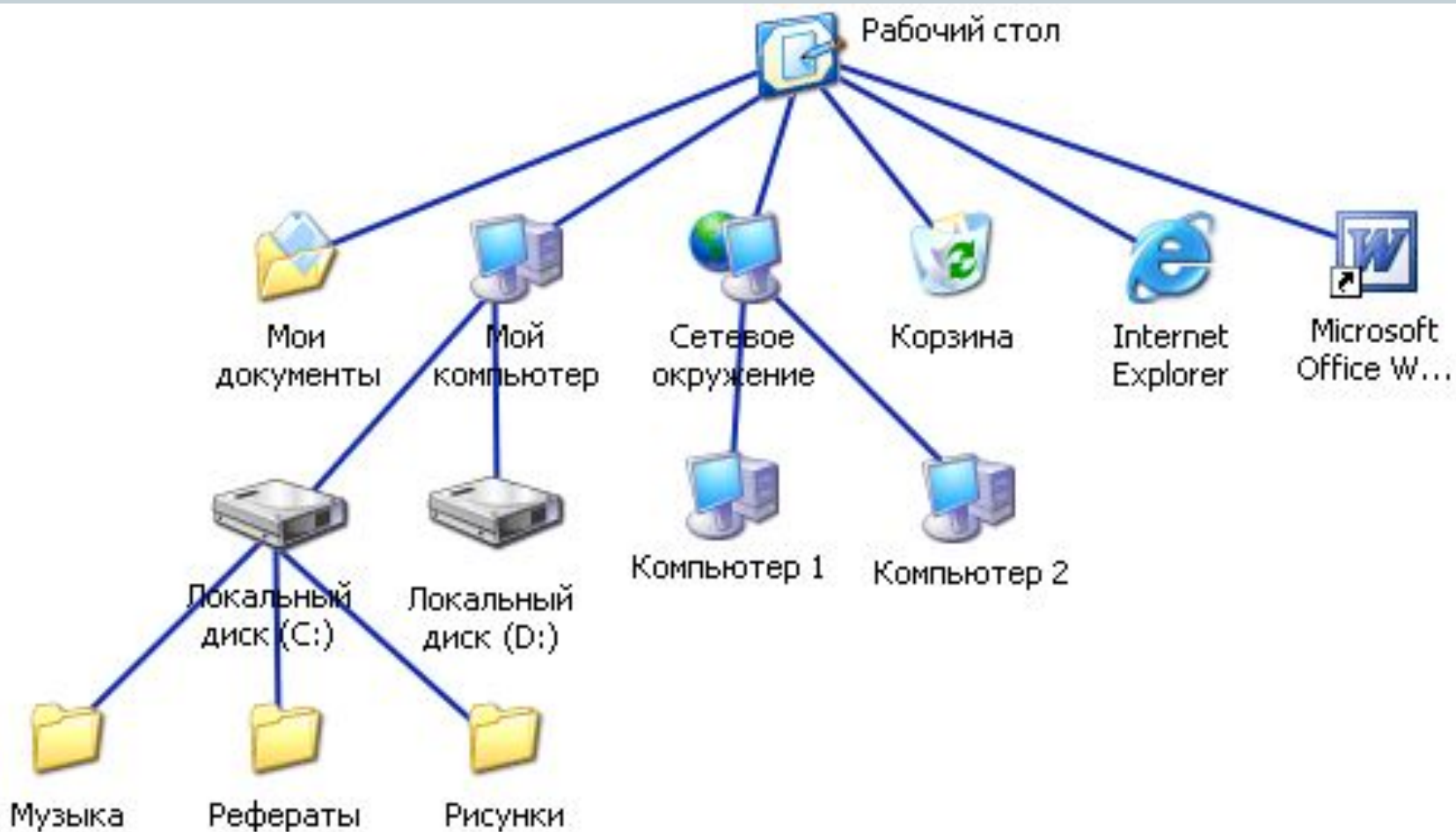


В операционной системе Windows вместо каталогов используется понятие «папка».

Папка – это объект Windows, предназначенное для объединения файлов и других папок в группы.

Понятие папки шире, чем понятие «каталог». В Windows на вершине иерархии папок находится папка Рабочий стол. (Следующий уровень представлен папками Мой компьютер, Корзина и Сетевое окружение (если компьютер подключен к локальной сети)).

Иерархии папок Windows



Стандартные действия с файлами



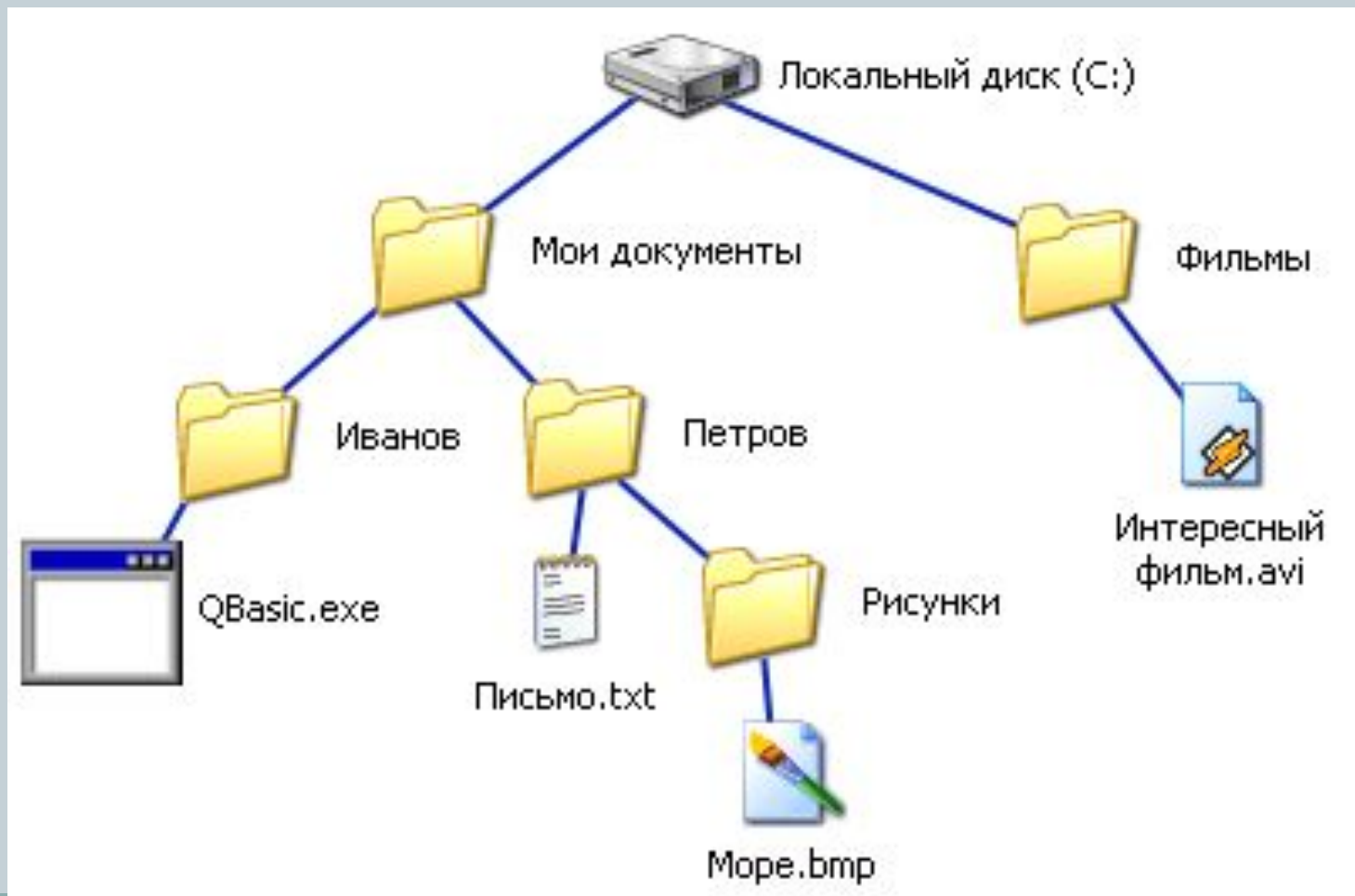
- Такие действия с файлами, как «создать», «сохранить», «заккрыть» можно выполнить только в прикладных программах («Блокнот», «Paint», ...).
- Действия «открыть», «переименовать», «переместить», «копировать», «удалить» можно выполнить в системной среде.

Операции с файлами и папками

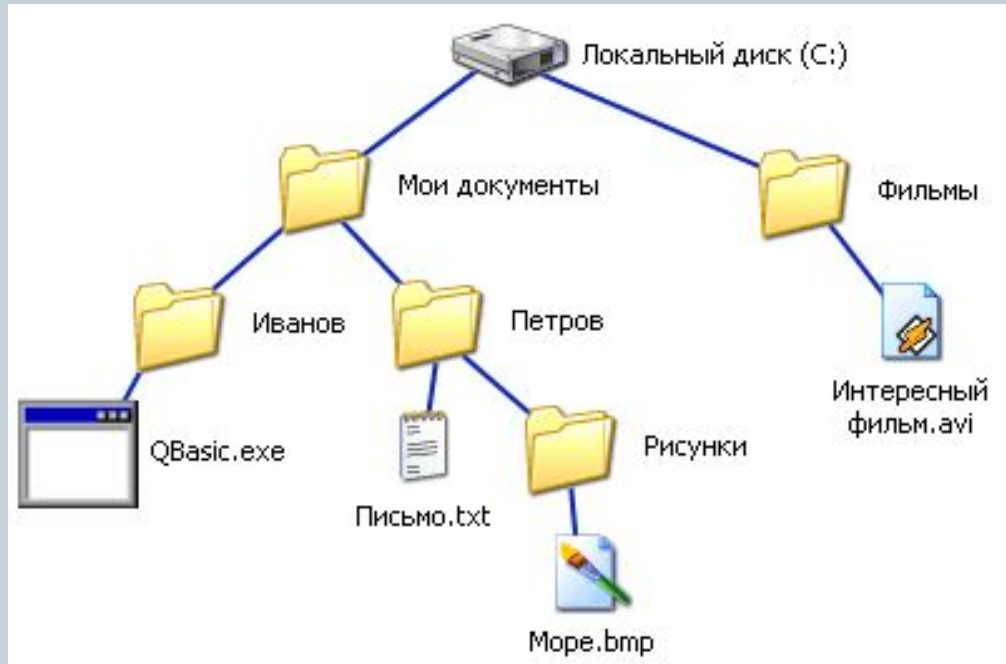


- **Копирование** (копия файла помещается в другой каталог);
- **Перемещение** (сам файл перемещается в другой каталог);
- **Удаление** (запись о файле удаляется из каталога);
- **Переименование** (изменяется имя файла).

Запишите полные имена всех файлов ?



Запишите полные имена всех файлов



C:\Мои документы\Иванов\QBasic.doc

C:\Мои документы\Петров\Письмо.txt

C:\Мои документы\Петров\Рисунки\Море.bmp

C:\Фильмы\Интересный фильм.avi

Постройте дерево каталогов



C:\Рисунки\Природа\Небо.bmp

C:\Рисунки\Природа\Снег.bmp

C:\Рисунки\Компьютер\Монитор.bmp

C:\Мои документы\Доклад.doc

Структура данных на магнитных дисках



ФАЙЛЫ И ФАЙЛОВАЯ СИСТЕМА

Какие единицы измерения информации вы знаете?



Бит – наименьшая единица измерения информации и обозначается двоичным числом.

1 байт = 8 битов.

- 1 Кбайт = 2^{10} байт = 1024 байт;
- 1 Мбайт = 2^{20} байт = 1024 Кбайт;
- 1 Гбайт = 2^{30} байт = 1024 Мбайт.
- Дискета – 1.44 Мбайт
- CD-ROM – 700 Мбайт
- DVD-ROM – 4.7 Гбайт
- Винчестер – 40 ~ 200 Гбайт



Полное имя файла



Путь к файлу вместе с именем файла называют полным именем файла.

C:\Рефераты\Физика\Оптические явления.doc

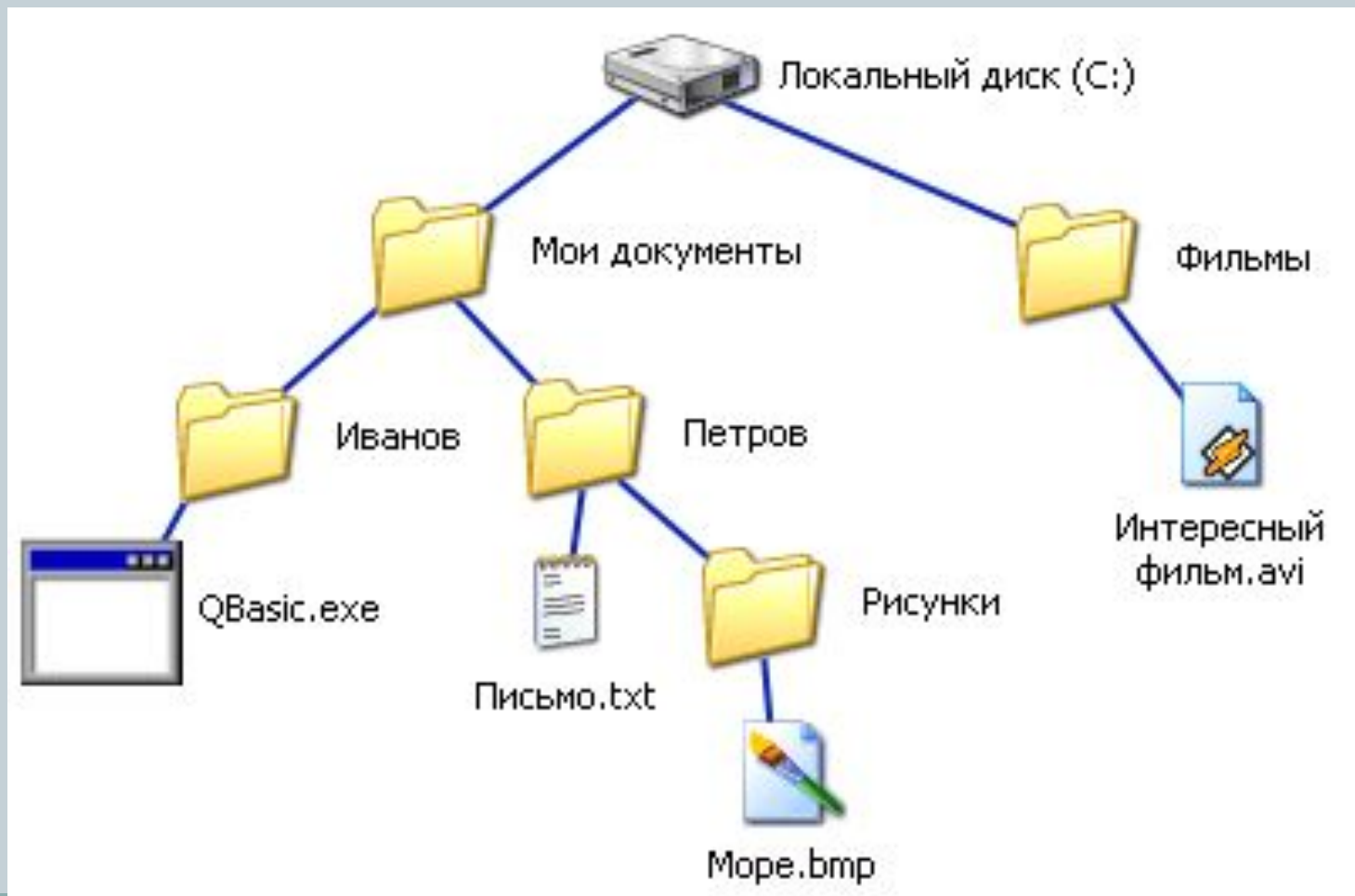
C:\Рефераты\Информатика\Интернет.doc

C:\Рефераты\Информатика\Компьютерные вирусы.doc

C:\Рисунки\Закат.jpg

C:\Рисунки\Зима.jpg

Запишите полные имена всех файлов ?



Структура дисков



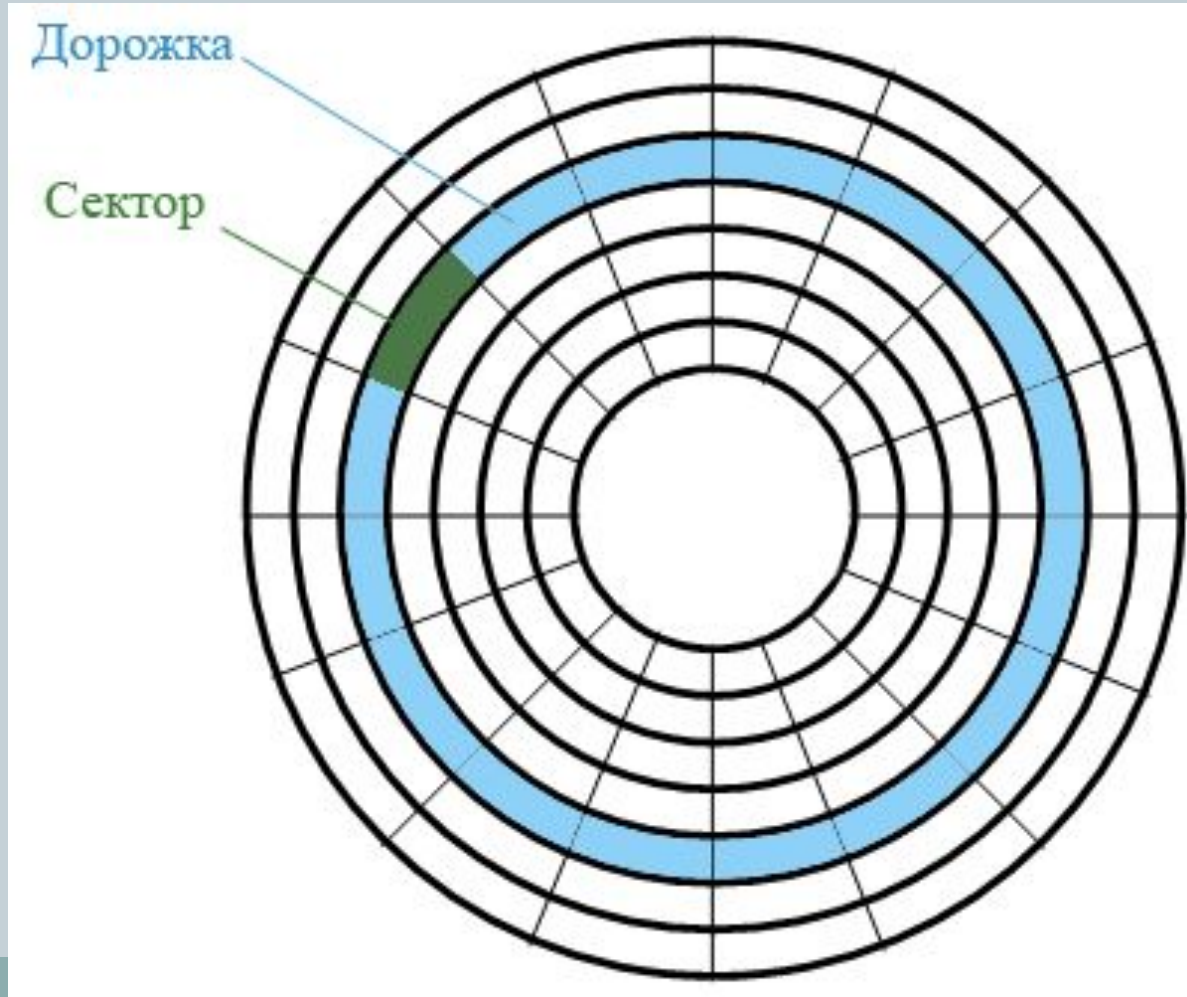
- Для того чтобы на диске можно было хранить информацию, диск должен быть **отформатирован**, то есть должна быть создана **физическая и логическая структура диска**.
- Формирование физической структуры диска состоит в создании на диске **концентрических дорожек**, которые, в свою очередь, делятся на **секторы**. Для этого в процессе форматирования магнитная головка дисководов расставляет в определенных местах диска метки дорожек и секторов.

Форматирование дисков



- **Форматирование диска** - процесс разметки диска на сектора и дорожки для записи данных. Форматирование создает структуру диска, обеспечивающую запись/чтение файлов и программ операционной системой.
- Форматирование выполняется служебными программами. Форматирование диска чем-то похоже на разлиновывание тетради.

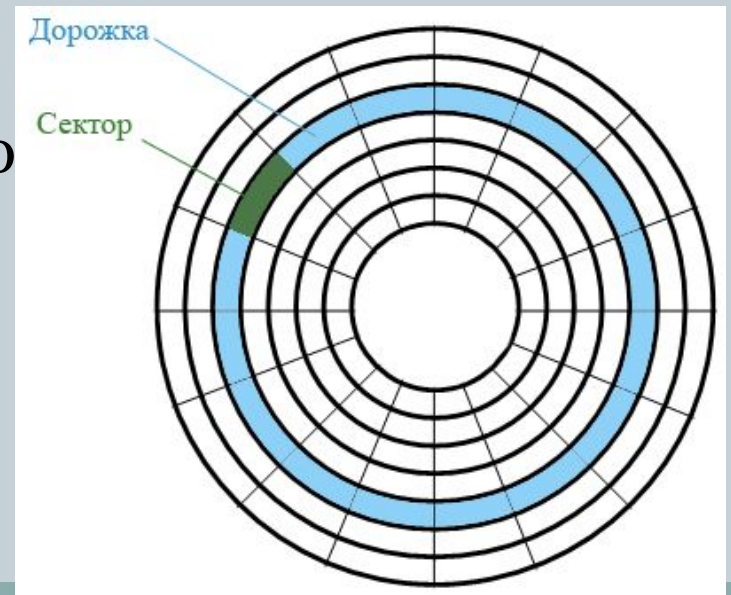
Логическая структура гибких дисков



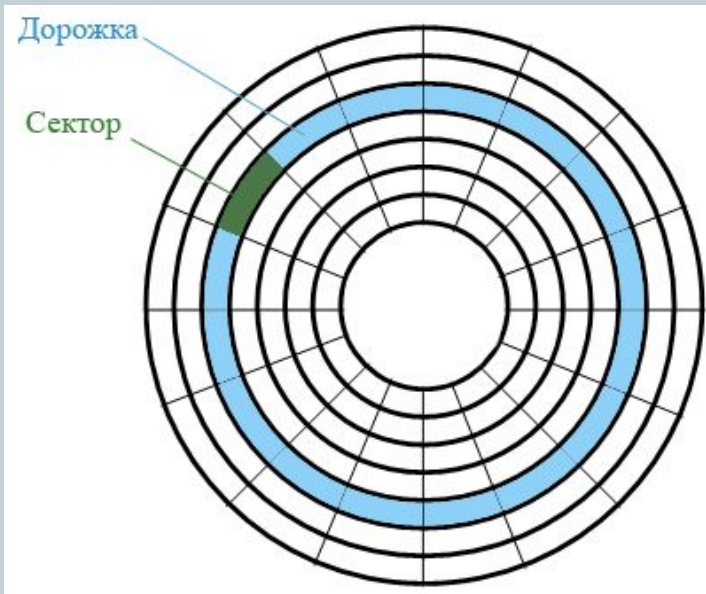
Логическая структура гибких дисков



- Логическая структура магнитного диска представляет собой совокупность секторов (емкостью 512 байтов), каждый из которых имеет свой порядковый номер (например, 100). Сектора нумеруются в линейной последовательности от первого сектора нулевой дорожки до последнего сектора последней дорожки.



Логическая структура гибких дисков



У гибкого диска две стороны, на которых создается по 80 дорожек. На каждой дорожке по 18 секторов. Общая емкость гибкого диска составляет $2 * 80 * 18 * 512 = 1474560$ байт \approx 1.44 Мбайт.

Таблица размещения файлов



- Для того чтобы можно было найти файл по его имени, на диске имеется каталог, представляющий собой базу данных. Запись о файле содержит имя файла, адрес первого сектора, с которого начинается файл, объем файла, а также дату и время его создания.
- Полная информация о секторах, которые занимают файлы, содержится в таблице размещения файлов (**FAT** — File Allocation Table).

Для размещения каталога — базы данных и таблицы FAT на гибком диске отводятся секторы со 2 по 33. Первый сектор отводится для размещения загрузочной записи операционной системы. Сами файлы могут быть записаны, начиная с 34 сектора.

Виды форматирования



- **полное форматирование** включает в себя как физическое форматирование (проверку качества магнитного покрытия диска и ее разметку на дорожки и секторы), так и логическое форматирование (создание каталога и таблицы размещения файлов). После полного форматирования вся хранившаяся на диске информация будет уничтожена.
- **быстрое форматирование** производит лишь очистку корневого каталога и таблицы размещения файлов. Информация, то есть сами файлы, сохраняется и в принципе возможно восстановление файловой системы.

Логическая структура жестких дисков



- Минимальным адресуемым элементом жесткого диска является кластер, который может включать в себя несколько секторов. Размер кластера зависит от типа используемой таблицы FAT и от емкости жесткого диска.
- *На жестком диске минимальным адресуемым элементом является кластер, который содержит несколько секторов.*

Дефрагментация дисков




- Замедление скорости обмена данными может происходить в результате **фрагментации файлов**. Фрагментация файлов (фрагменты файлов хранятся в различных, удаленных друг от друга кластерах) возрастает с течением времени, в процессе удаления одних файлов и записи других.
- Так как на диске могут храниться сотни и тысячи файлов в сотнях тысяч кластеров, то фрагментированность файлов будет существенно замедлять доступ к ним (магнитным головкам придется постоянно перемещаться с дорожки на дорожку) и в конечном итоге приводить к преждевременному износу жесткого диска. Рекомендуется периодически проводить **дефрагментацию** диска, в процессе которой файлы записываются в кластеры, последовательно идущие друг за другом.

Дефрагментация - процесс перезаписи частей файла в соседние сектора на жестком диске для ускорения доступа и загрузки.

Физические и логические диски

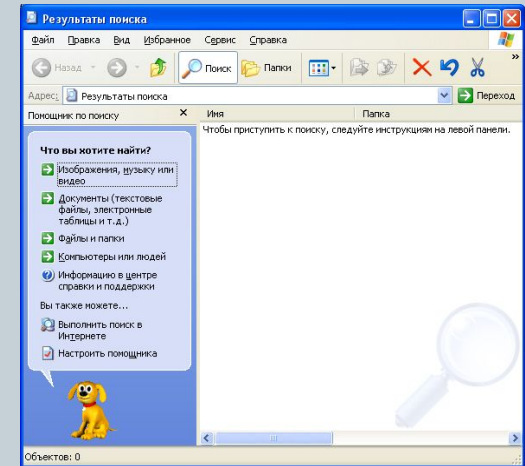
Для борьбы с нерациональными потерями или, просто, для удобства, часто жесткий диск разбивают на несколько разделов. Каждый такой раздел можно рассматривать как отдельный логический жесткий диск.

 Диск 0 Основной 111,79 ГБ Подключен	System (C:) 14,65 ГБ NTFS Исправен (Система)	Temp (D:) 4,88 ГБ NTFS Исправен (Файл подка'	Home (E:) 92,26 ГБ NTFS Исправен
--	---	---	---

Поиск файлов и папок



- Для отображения на экране окна поиска нажмите Пуск→Поиск. Окно поиска можно также активировать, нажав кнопку Поиск на панели инструментов в окне Мой компьютер или проводника.
- Для облегчения поиска пользователю предлагается нажатием кнопки выбрать, что он хочет найти: изображение, музыку или видео, файл или папку, компьютеры или людей и т.д.
- Чтобы найти файл или папку, в окне Результаты поиска на панели помощника по поиску щелкните ссылку Файлы и папки. Вы можете найти файл по его имени или части имени или некоторым другим критериям.



Шаблоны поиска



При поиске файла по имени можно использовать шаблон: звездочка «*» и знак вопроса «?».

Символ шаблона звездочка «*» заменяет любое количество символов, знак вопроса «?» - один символ.

Например, после ввода для поиска «Мыш?.doc» будут найдены файлы: «Мышь.doc» и «Мыши.doc».

После ввода «*.jpg» будут найдены все рисунки в формате jpg.

Архивация файлов



ФАЙЛЫ И ФАЙЛОВАЯ СИСТЕМА

Избыточность



- Редакторы, работающие с текстовой, графической, звуковой и другой информацией, **кодируют** ее наиболее **естественным**, но **не самым экономичным способом**.
- Действительно, если внимательно посмотреть любой текст, то можно заметить, что такие буквы «а» и «о», встречаются в нем гораздо чаще чем «ю» и «у». То же самое можно отнести и к сочетаниям букв.
- На рисунках цвета соседних точек в большинстве случаев близки по оттенку. Подобно этому в любой последовательности информации некоторые сочетания встречаются намного чаще других.

Избыточность



- Все это приводит к тому, что в файлах, хранящих эту информацию, некоторые **комбинации** из 0 и 1 **встречаются гораздо чаще**, чем другие. В таких случаях говорят, что информация обладает **избыточностью**, и есть возможность перекодировать содержание файла, уменьшив его размер.
- Для **сжатия** достаточно придерживаться правила: чем чаще встречается комбинация, тем более коротким сочетанием из 0 и 1 ее можно **перекодировать**. Разумеется, делать это должна программа.

Архиваторы



- Программы, осуществляющие сжатие (упаковку файлов), называют архиваторами.

При сжатии можно уменьшить размер файла в несколько раз, что дает заметную экономию памяти.

Например: WinRAR и WinZip



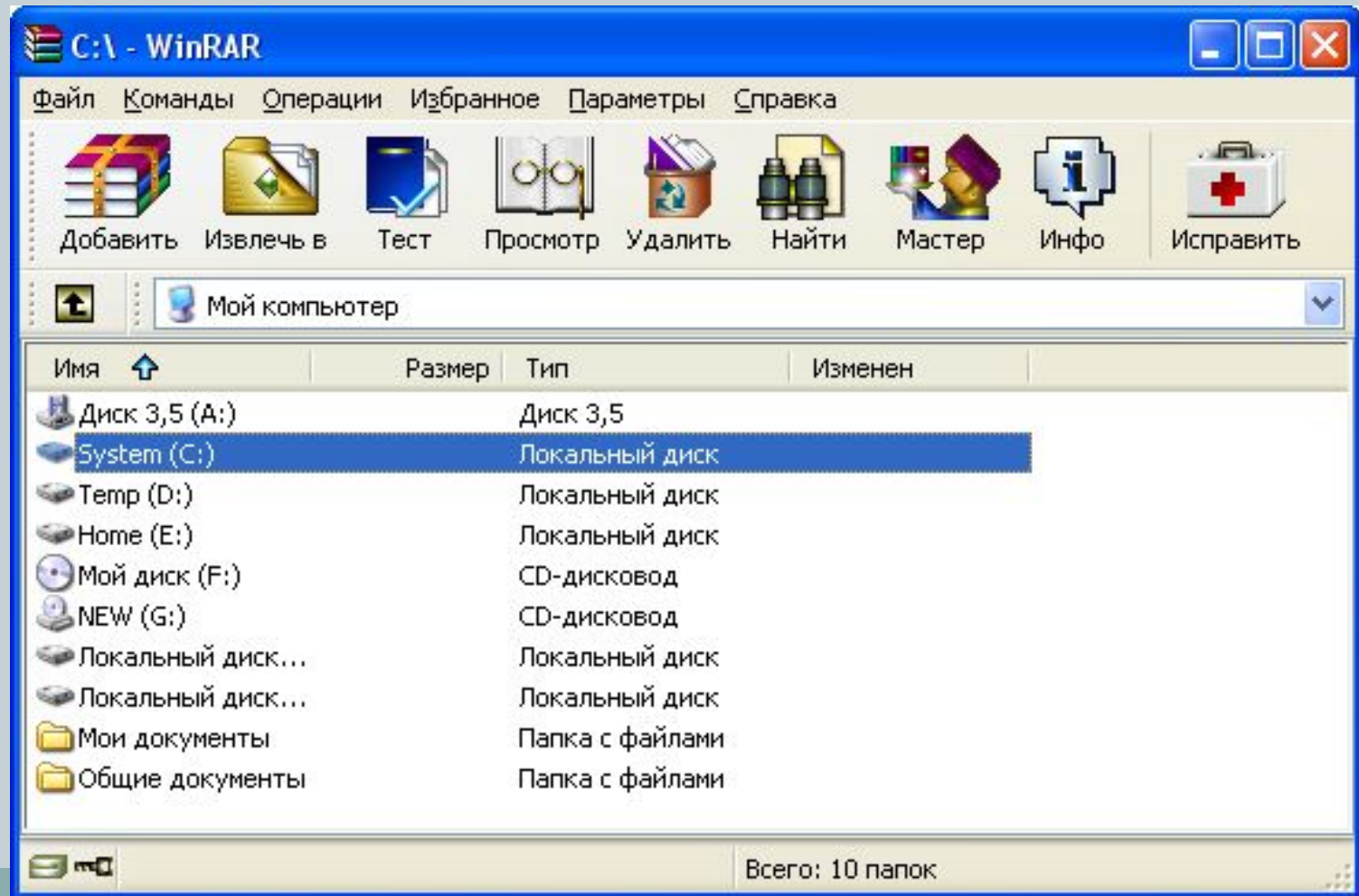
Пуск → Все программы → Архиваторы → WinRAR → WinRAR

Основные действия при работе с архивами:

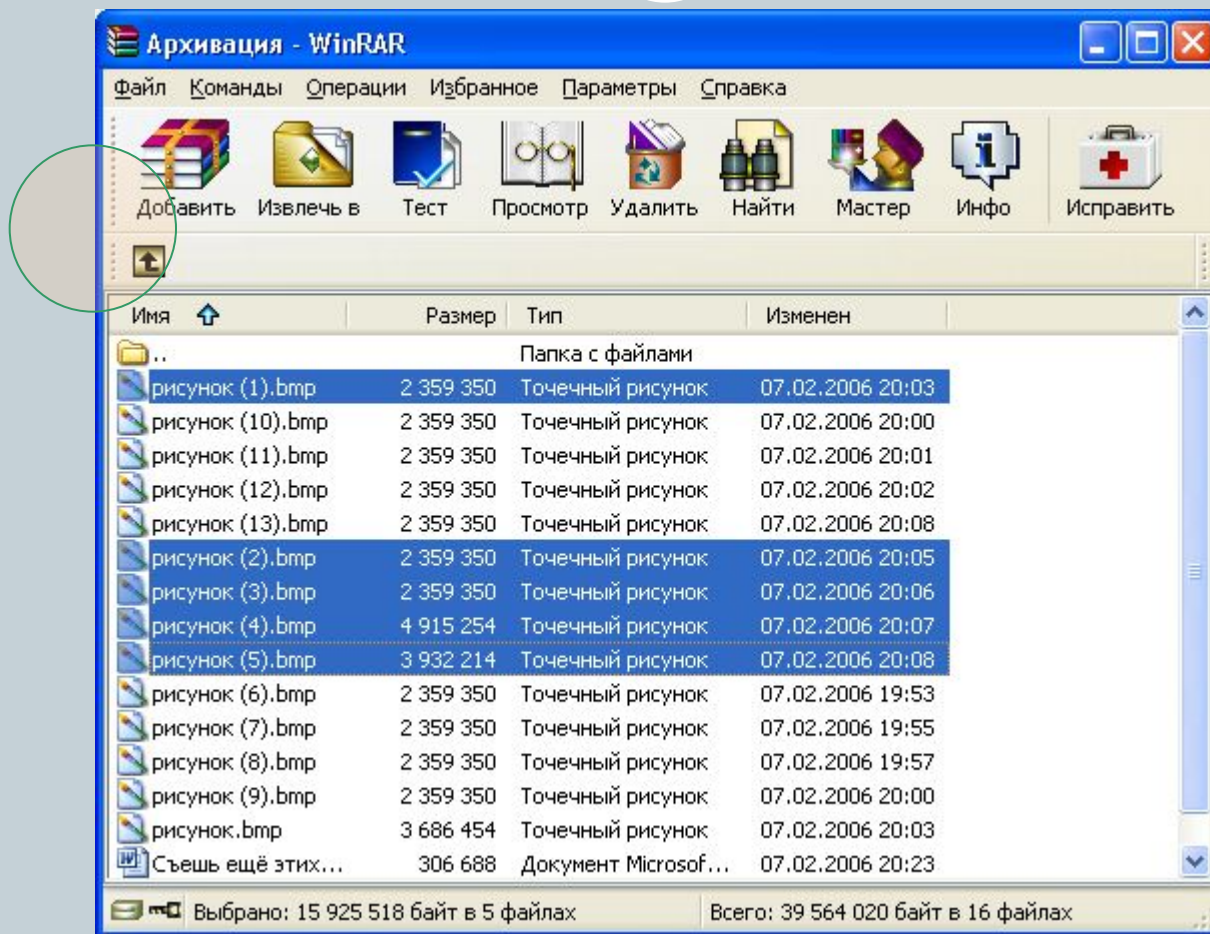


- создание нового архива;
- добавление файлов в архив;
- просмотр содержимого архива;
- извлечение файлов из архива;
- просмотр файла в архиве;
- удаление файлов из архива.

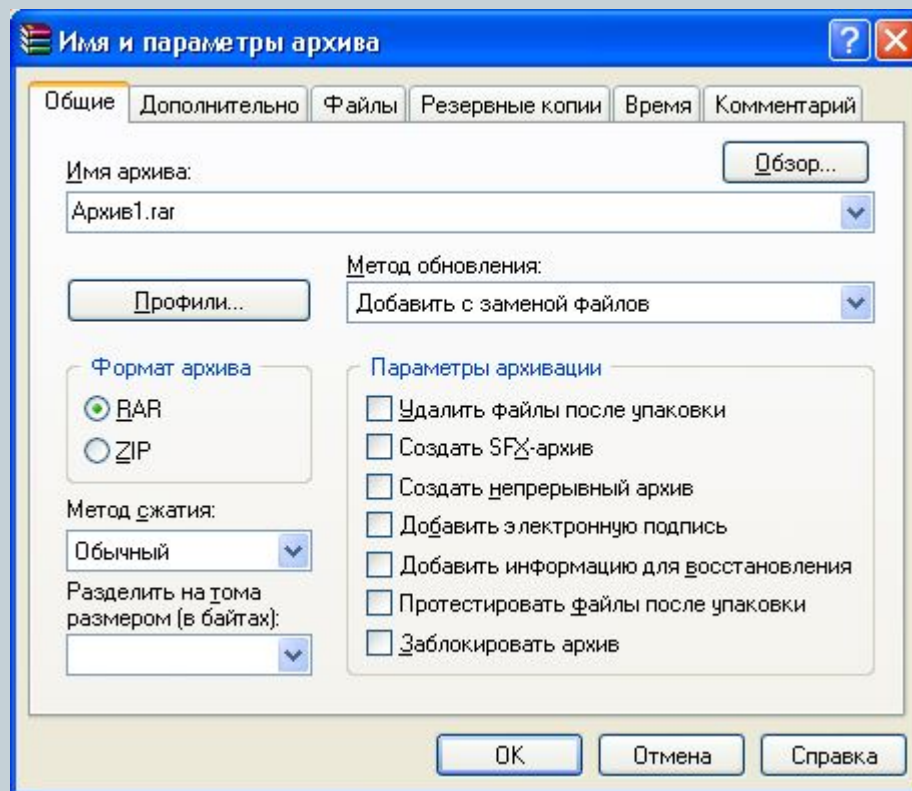
Оболочка WinRAR



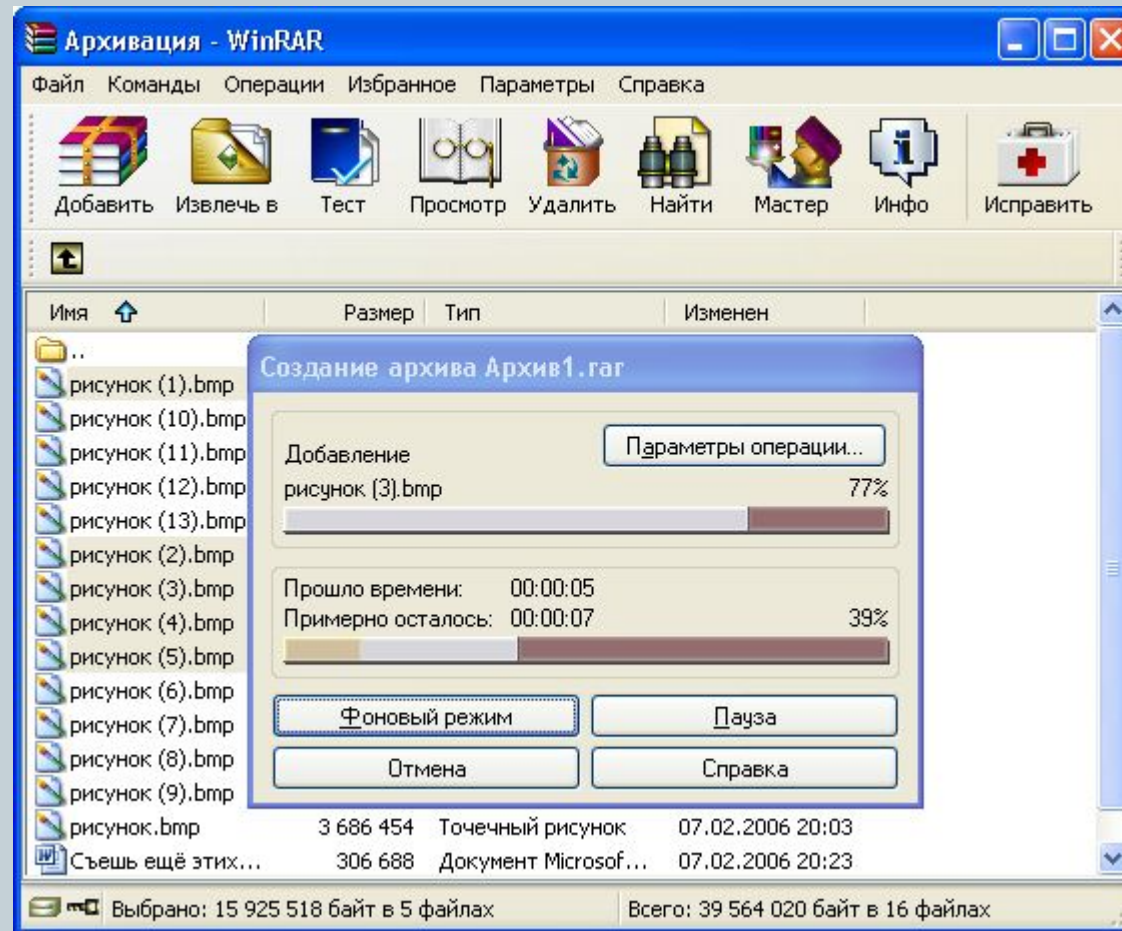
Архивация с помощью оболочки WinRAR



Архивация с помощью оболочки WinRAR



Архивация с помощью оболочки WinRAR



Информация об архиве



Архив Архив1.rar

Информация | Параметры | Комментарий | SFX

RAR архив

Версия для извлечения:	2.9
Базовая ОС:	Windows

Всего файлов:	5
Общий размер:	15 925 518
Размер в архиве:	5 715 310
Степень сжатия:	35%

Размер SFX-модуля:	0 байт
Главный комментарий:	Нет
Пароли:	Нет

Размер словаря:	4096 Кб
Информация для восстановления:	Нет
Блокировка архива от изменений:	Нет

Электронная подпись:	Нет
----------------------	-----

35%

OK Отмена Справка

Архив Съешь ещё этих мягких французск...

Информация | Параметры | Комментарий | SFX

RAR архив

Версия для извлечения:	2.9
Базовая ОС:	Windows

Всего файлов:	1
Общий размер:	306 688
Размер в архиве:	4 276
Степень сжатия:	1%

Размер SFX-модуля:	0 байт
Главный комментарий:	Нет
Пароли:	Нет

Размер словаря:	512 Кб
Информация для восстановления:	Нет
Блокировка архива от изменений:	Нет

Электронная подпись:	Нет
----------------------	-----

1%

OK Отмена Справка

Распаковка файлов

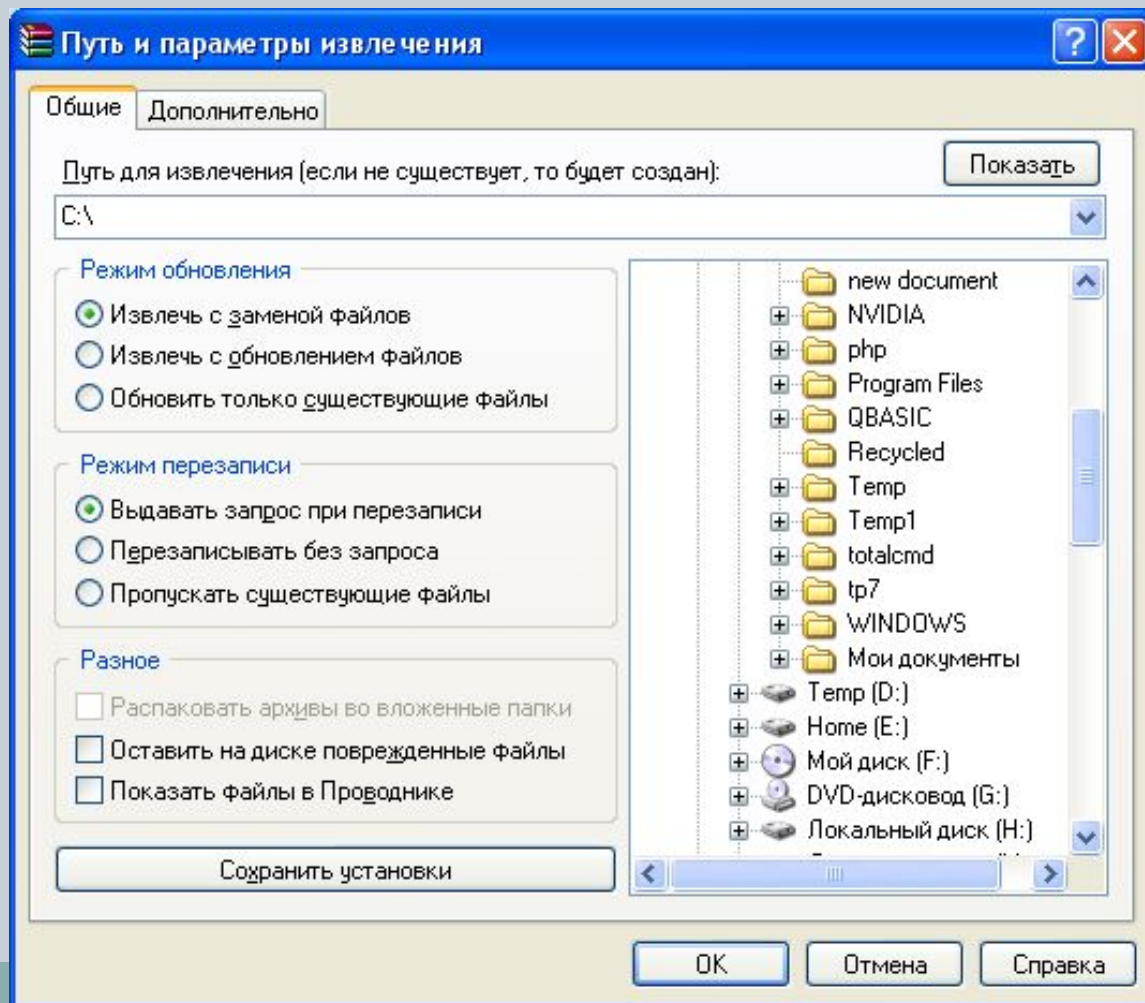


The screenshot shows the WinRAR application window titled "архив1.rar - WinRAR". The menu bar includes "Файл", "Команды", "Операции", "Избранное", "Параметры", and "Справка". The toolbar contains icons for "Добавить", "Извлечь в" (highlighted with a green circle), "Тест", "Просмотр", "Удалить", "Найти", "Мастер", "Инфо", and "Вирусы". The address bar shows the current file: "архив1.rar - RAR архив, размер исходных файлов 16 232 206 байт".

Имя	Размер	Сжат	Тип	Изменен
..			Папка	
рисунок (1).bmp	2 359 350	778 106	Точечный рисунок	07.02.20
рисунок (2).bmp	2 359 350	729 971	Точечный рисунок	07.02.20
рисунок (3).bmp	2 359 350	825 756	Точечный рисунок	07.02.20
рисунок (4).bmp	4 915 254	2 342 958	Точечный рисунок	07.02.20
рисунок (5).bmp	3 932 214	1 038 519	Точечный рисунок	07.02.20
Съешь ещё этих мягких французски...	306 688	4 276	Документ Microsof...	07.02.20

At the bottom, the status bar indicates: "Выбрано: 7 581 292 байт в 3 файлах" and "Всего: 16 232 206 байт в 6 файлах".

Распаковка файлов



Вопросы:



- Что такое файл?
- Имя файла: состав?
- Типы файла?
- Что такое файловая система?
- Как записать полное имя файла?
- Правила записи имен файлов?
- Что такое папка?
- Операции с файлами и папками?

Вопросы:



- Что такое форматирование диска?
- Почему отличаются реальный информационный объем файла и объем, который он занимает на диске?
- Чем отличается полное и быстрое форматирование?
- Для чего необходимо проводить дефрагментацию диска?
- Что такое логический диск?

Вопросы:



- Почему есть возможность уменьшать размер файлов?
- Что такое архивация?
- Какие файлы не имеет смысла архивировать?
- Почему перед пересылкой текстового файла по электронной почте имеет смысл предварительно его упаковать в архив?