ТРИ ЭТАПА РАБОТЫ ПОИСКОВОГО УКАЗАТЕЛЯ

- Сбор первичной базы данных
- Индексация базы данных
- Рафинирование результирующего списка

1

СБОР ПЕРВИЧНОЙ БАЗЫ ДАННЫХ

- Сканирование пространства *World Wide Web* осуществляется специальными агентскими программами – *червями* (не следует путать с сетевыми компьютерными вирусами)
- Их задача – автоматически разыскивать Web-ресурсы, следуя по гиперссылкам, и, убедившись, что этот ресурс системе ещё не известен, копировать его в свою базу данных

2

ИНДЕКСАЦИЯ БАЗЫ ДАННЫХ

- Чтобы не задерживать клиента при поиске собранные базы данных проходят предварительную обработку, называемую *индексацией*
- На этапе индексации создаются специализированные документы – *поисковые указатели* или *индексы*
- **Индексация** документов сайта (*Web-индексация*) – одна из функций комплекса программ, называемого *поисковой машиной*
- При построении поискового индекса используются достижения и методы лингвистики, когнитивной психологии, математики и информатики
- Основная задача – *индексация текстовых документов*, но видео, аудио и графика также могут быть упорядочены и проиндексированы для последующего быстрого поиска
- Метапоисковые машины используют индексы других машин не храня собственного индекса, в то время как собственно поисковые машины хранят индекс наряду с копиями страниц.
- В отличие от полнотекстовых индексов полного текста, частичные индексы ограничивают размер индекса
- Поиск может быть более полным с гарантированным временем ответа либо «мгновенным»

ИНДЕКСЫ

- **Индекс поисковой системы** – в Интернет - компонента поисковой системы; информационный массив, в котором хранятся специальным образом преобразованные текстовые составляющие всех посещенных и проиндексированных роботом веб-страниц и текстовых файлов

4

ОСНОВНЫЕ ЗАДАЧИ ПРИ КОНСТРУИРОВАНИИ ПОИСКОВЫХ ИНДЕКСОВ

- **Совместный доступ** - *синхронизация* индексации с поиском и организация *взаимодействия множественных* агентов индексации
- **Хранение** - выбор оптимальной структуры и *сжатие vs фильтрация*
- **Размер индекса** - использование и ограничения размеров разных видов *компьютерной памяти*
- **Скорость поиска** vs скорость изменения или удаления
- **Процедуры** долговременной корректной **поддержки**
- **Устойчивость к отказам оборудования**, определение процедур изоляции и восстановления повреждённых данных

5

РАФИНИРОВАНИЕ РЕЗУЛЬТИРУЮЩЕГО СПИСКА

- *Рафинирование* заключается в *фильтрации* и *ранжировании* результатов поиска
- Фильтрация – отсев ссылок, которые пользователю выдавать нецелесообразно
- Ранжирование – создание специального порядка представления результирующего списка, при котором наиболее «полезные» результаты попадают в начало списка

6

социология окружающей среды - Поиск в Google - Mozilla Firefox

Файл Правка Вид Журнал Закладки Инструменты Справка

http://www.google.ru/#hl=ru&source=hp&q=социология+окружающей+среды&btnG=Поиск+в+Google&lr=&aq=

Яндекс

Самые популярные Начальная страница Лента новостей Настройка ссылок

социология окружающей ... | Расписание занятий (обно... | социология окружающ_ | Scirus - for scientific informati... | Home - SciTopics

Веб Картинки Видео Карты Новости Переводчик Gmail ещё ▾ История веб-поиска | Настройки поиска | Войти

Google социология окружающей среды Поиск Расширенный поиск

Поиск в Интернете Поиск страниц на русском

Веб Показать настройки... Результаты **1** - **10** из примерно **215 000** для **социология окружающей среды**. (**0,27** секунд)

Социология окружающей среды (OCW Университета Аликанте)
Социология окружающей среды (OCW Университета Аликанте). Описание курса: Изучение связей между обществом и **окружающей средой**, со специальной ориентацией на **...**
ocw.universia.net/.../environmental-sociology-(ocw-university-of-alicante)/ - Сохранено в кэше

Социология городской **среды**
Социология городской среды. № 1 (2006) Количество просмотров: 2093 **...** содержание строительно-коммунальных отношений **окружающей среды** накладываются все те **...**
www.rostovstroy.ru/archive/.../251.html - Сохранено в кэше

Среда и общество - Журнал "**Социологические** исследования"
«Среда и общество» или «Инвайронментальная **социология**»? **...** Публикации с ключевыми словами: **социология окружающей среды** - инвайронментальная **социология ...**
www.ecsocman.edu.ru/socis/.../242382.html - Сохранено в кэше

ИССЛЕДОВАТЕЛЬСКИЕ ПРОЕКТЫ -- ЭКОЛОГИЧЕСКАЯ **СОЦИОЛОГИЯ**
Публикация специального выпуска журнала "**Социология** и социальная антропология" на **.....** Информирование принимающих решения в области **окружающей среды** о **...**
www.cisr.ru/projects_ecology.html - Сохранено в кэше

eLIBRARY.RU - Уайленис М. **Социология**, современность и глобальные **...**
публикации, УАЙЛЕНИС М. **СОЦИОЛОГИЯ**, СОВРЕМЕННОСТЬ И ГЛОБАЛЬНЫЕ ИЗМЕНЕНИЯ **ОКРУЖАЮЩЕЙ СРЕДЫ** (РЕФЕРАТ). Sociology, modernity and the globalization of **...**

Готово

Буфер обмена (16 из 24)
Объект добавлен в буфер.

Пуск [s... Co... Mi... П... с... Д... Te... S... Ko... nero SEARCH RU 12:11

АВТОМАТИЧЕСКАЯ КАТАЛОГИЗАЦИЯ

- Перспективные направления развития основаны на внедрении так называемых *SMART*-технологий автоматической каталогизации
- Например - присвоение весовых коэффициентов терминам относительно рубрик
- Тогда, получив новую Web-страницу, поисковая система сможет построить математический вектор, основанный на формальном анализе содержания, и сравнить его с векторами для различных поисковых категорий

8

ПОИСКОВЫЕ СИСТЕМЫ РЕАЛЬНОГО ВРЕМЕНИ

- Клиентская программа предаёт серверу копии всех Web-страниц, которые посещает пользователь
- Отличия от автоматических *червей*
  - Сервер получает копии только тех ресурсов, которые заинтересовали кого-либо из клиентов
  - При наличии многих клиентов индексация Web-пространства происходит целенаправленнее и быстрее
- На какой бы странице ни находился пользователь, система всегда готова предложить ему список других страниц, имеющих близкое **(с её точки зрения)** по тематике содержание

9

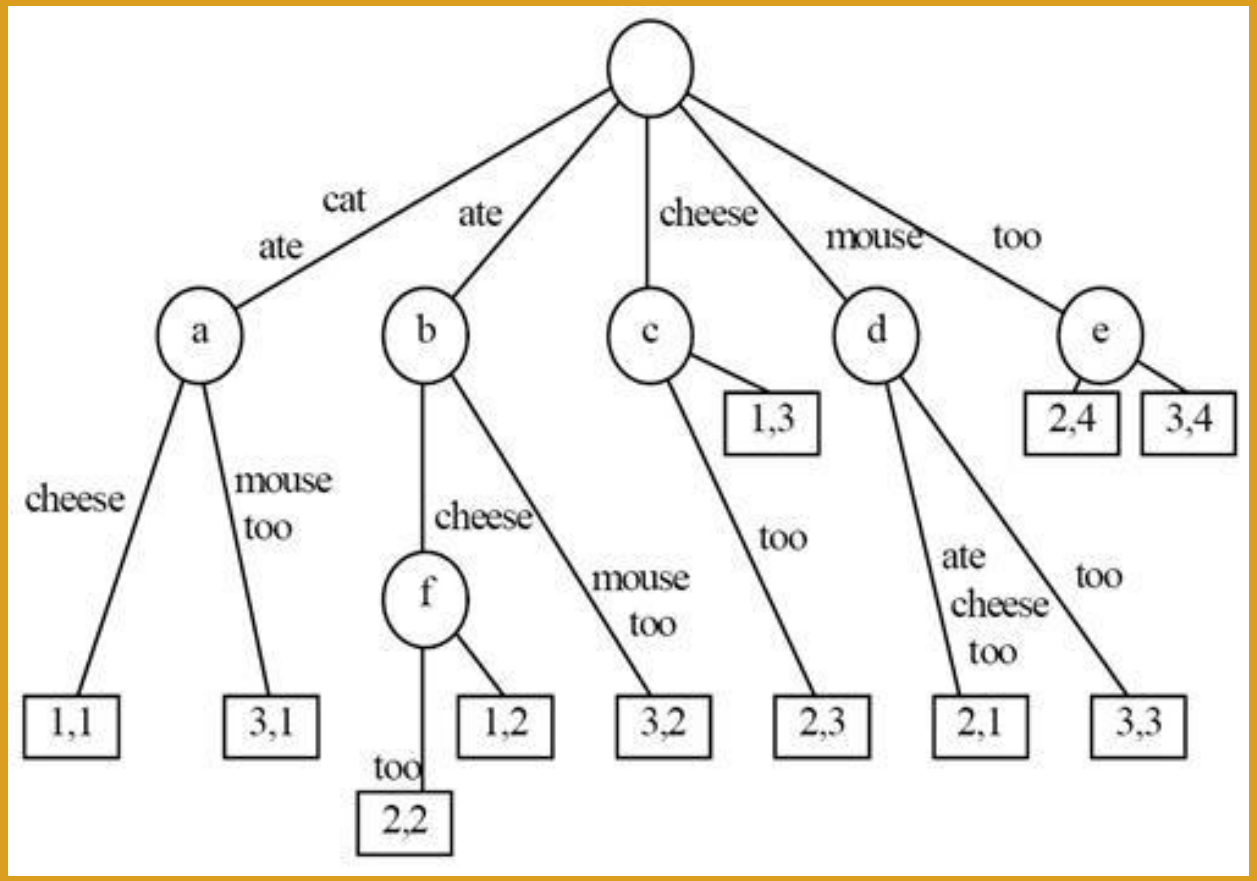
ТИПЫ ПОИСКА

- *Простой поиск* – по одному или нескольким ключевым словам
- *Расширенный поиск* – ключевые слова связывают между собой операторами логических отношений
- *Контекстный поиск* – поиск по точной фразе
- *Специальный поиск* применяют при розыске Web-страниц, содержащих ссылки на заданные адреса URL, а также содержащих заданные слова в специальных полях, например, поле заголовка

10

СУФФИКСНОЕ ДЕРЕВО ДЛЯ ТРЁХ ПРЕДЛОЖЕНИЙ

- Cat ate cheese. Mouse ate cheese too. Cat ate mouse too.



ДРУГИЕ ВСПОМОГАТЕЛЬНЫЕ ТЕХНИКИ

- Инвертированные индексы
- Индексы цитируемости
- Индексы n-грамм
- Матрицы *термин-документ*

12

ЛЕКЦИЯ №4

Тема №3. Microsoft Word и культура создания текстовых документов с его помощью

ОПРЕДЕЛЕНИЕ

- Microsoft Word - текстовый процессорMicrosoft Word - текстовый процессор, предназначенный для создания, просмотра и редактирования текстовых документов.

НАСТРОЙКА ПАРАМЕТРОВ СТРАНИЦЫ

- Работа с программой начинается с настройки печатной страницы. Команда Файл->Параметры страницы

1. На вкладке Размер бумаги выбрать размер бумаги. Наиболее ходовой – А4
2. На вкладке Поля задать размеры полей. Зависят от необходимости брошюровки, полей для ручной правки, планируемого наличия колонтитулов. Номер страницы (**колонцифра**) также является элементом колонтитула
3. На вкладке Поля выбирают также режим ориентации печатной страницы: Книжная либо Альбомная и – через выпадающий список Несколько страниц – режим представления страниц на листе бумаги: обычный, зеркальные поля, 2 страницы на листе, брошюра
4. На вкладке Источник бумаги можно также установить флажки Различать колонтитулы первой страницы и Различать колонтитулы чётных и нечётных страниц, назначить нумерацию строк и вид границы

ВВОД И РЕДАКТИРОВАНИЕ ТЕКСТА В ДОКУМЕНТЕ WORD

- Основной набор – с клавиатуры
- Варианты выравнивания абзаца: слева, справа, посередине, по ширине.
- Отступы от края страницы
- Межстрочные расстояния
- Курсив, полужирный, подчёркнутый

ВСТАВКА СПЕЦИАЛЬНЫХ СИМВОЛОВ

- Посредством меню
- Alt-33 !
- Alt-21 §
- Alt-64 @
- 03BE, Alt-X → ξ

ЗАКЛАДКИ

- Чтобы в следующий раз не искать данный текст во всем документе, можно перейти к нему с помощью диалогового окна **Закладка**.
- Выделите элемент, которому следует назначить закладку, или щелкните в документе место для ее вставки.
- В меню **Вставка** выберите команду **Закладка**.
- В поле **Имя закладки** введите или выберите нужное имя. Имя закладки должно начинаться с буквы; в нем могут использоваться цифры, но не должно быть пробелов. Если требуется разделить слова в имени закладки, пользуйтесь знаком подчеркивания, например «Первый_заголовок».
- Нажмите кнопку **Добавить**.

СНОСКИ, ГРАНИЦЫ, ЗАЛИВКА

- Обычные и концевые сноски – возможность преобразования документа
- Границы и заливка устанавливаются командой **Формат – Границы и заливка**
  В Word 2007-2010: ***Главная - Абзац - Границы и заливка - Границы***

ПОЛЕЗНЫЕ СОВЕТЫ

- Список соответствия команд Word 2003 и Word 2007: http://www.intuit.ru/department/office/msword2007b/13/
- Пользуйтесь возможностями справки (клавиша F1)

ФОРМАТИРОВАНИЕ ТЕКСТА. АТРИБУТЫ ФОРМАТА СИМВОЛОВ И АБЗАЦЕВ

- Абзац – основная единица оформления текста документа. Спецсимвол «Конец абзаца»

АВТОМАТИЧЕСКОЕ ФОРМАТИРОВАНИЕ ДОКУМЕНТОВ

- Сервис->Параметры автозамены
- С помощью автоформата можно быстро отформатировать заголовки, **маркированные** и нумерованные списки, границы, числа, знаки и дроби. Дополнительные возможности Microsoft Word включают:
- автоматическое форматирование адресов Интернета, сетей и электронной почты в виде **гиперссылок**;
- применение полужирного или курсивного шрифта к фрагментам текста, выделенным звездочками (*) или знаками подчеркивания (_);
- замена двух дефисов (--) длинным тире (—).
- Можно автоматически форматировать документ либо при вводе, либо после его создания. В обоих случаях можно отслеживать изменения, внесенные Microsoft Word. При необходимости возможность автоматического форматирования можно отключить.
- При этом проводится анализ использования каждого абзаца в документе (например, является ли абзац заголовком или элементом нумерованного списка), а затем к каждому элементу применяется соответствующий **стиль**.

ФОРМАТНАЯ КИСТЬ

- Кнопка “Копировать формат” в панели Основные

СТИЛИ АБЗАЦЕВ

- Совокупность настроек оформления абзаца, имеющая уникальное имя, называется **стилем абзаца**
- Различают следующие категории абзацев:
- *Абзацы основного текста*
- *Заголовки*
- *Маркированные списки*
- *Нумерованные списки*
- *Табличные стили*

БИБЛИОТЕКИ СТИЛЕЙ

- В меню **Формат** выберите команду **Тема.**
- Нажмите кнопку **Библиотека стилей.**
- В списке **Шаблон** выберите **шаблон**, содержащий нужные стили. Если нажать кнопку **OK**, то все выделенные стили из шаблона копируются в документ.
- Совет

Чтобы получить представление о том, как будет выглядеть текущий документ при использовании различных стилей, установите переключатель **Образец** в положение **Документ**. Чтобы просмотреть образец документа со стилями из выбранного шаблона, выберите вариант **Пример**. Для ознакомления со списком стилей, используемых в выбранном шаблоне, выберите вариант **Образцы стилей.**

ПРИНЦИП НАСЛЕДОВАНИЯ СТИЛЕЙ

- **Для изготовления нового стиля** – нажать кнопку Создать стиль в области задач Стили и форматирование
- Надо иметь в виду, что новый стиль создаётся на основе уже существующего (список Основан на стиле).
- Если поменять что-то в родительском стиле, то соответственно поменяются все производные. Но только тогда, когда соответствующий параметр оформления не был задан в новом стиле

ШАБЛОНЫ ДОКУМЕНТОВ

- Параметры настройки печатной страницы можно сохранить вместе со всеми стилями абзацев, созданными и настроенными для данного документа
- Сохранение происходит в виде файла, который называется **шаблоном документа**
- Для разных типов документов, имеющих разное форматирование, целесообразно заранее заготовить базовые шаблоны
- Стандартная папка шаблонов - *drive*:\Documents and Settings*User*\Application Data\Microsoft\Шаблоны, набор шаблонов – личный для каждого пользователя

ПРИНЦИП НАСЛЕДОВАНИЯ ШАБЛОНОВ

- Шаблоны, как и стили, обладают наследованием свойств
- При загрузке какого-либо шаблона вместе с ним загружаются и все родительские вплоть до «корневого» Normal.dot. Цепочка загруженных шаблонов называется **активными шаблонами**

ПРИНЦИПЫ КОРРЕКТНОЙ РАБОТЫ

Корректное создание документа подразумевает следующий порядок действий:

1. Выбор либо создание шаблона, на базе которого готовится документ
2. Ввод текста в шаблон
3. Оформление документа с помощью функциональных стилей, заготовленных в шаблоне

СТРУКТУРА ДОКУМЕНТА

- Режим просмотра текста Структура документа (Ctrl-Alt-O) очень удобен для работы с большим текстом – дипломным проектом, художественным романом либо научной монографией
- Даёт возможность Выбора уровня показа, повышения и понижения уровня заголовка, перемещения вверх и вниз по тексту

СХЕМА ДОКУМЕНТА

- Чтобы отобразить **схему документа**, необходимо все заголовки документа отформатировать с помощью одного из встроенных **стилей заголовков**
- Чтобы отобразить все заголовки не ниже определенного уровня, щелкните правой кнопкой мыши заголовок на схеме документа, а затем выберите соответствующий уровень заголовков в **контекстном меню**. Например, чтобы отобразить заголовки с 1 по 3 уровни, выберите команду **Показать заголовки 1-3**

СОЗДАНИЕ ОГЛАВЛЕНИЯ

РАЗМЕЩЕНИЕ ТЕКСТА В КОЛОНКАХ

- Команда Формат->Колонки
- Можно создать кнопку Колонки на панели Форматирование
- Принудительный разрыв колонки

СОЗДАНИЕ ТАБЛИЦ

СПИСКИ

МАРКИРОВАННЫЕ И НУМЕРОВАННЫЕ СПИСКИ

МНОГОУРОВНЕВЫЕ СПИСКИ

ВСТАВКА НОМЕРОВ СТРАНИЦ И КОЛОНТИТУЛОВ

- Вставка – Номера страниц (*Вставка - Колонтитулы - Номер страницы* в Word 2007-2010)
- Вид – Колонтитулы (*Вставка - Колонтитулы* в Word 2007-2010)

ПРЕДВАРИТЕЛЬНЫЙ ПРОСМОТР И ПЕЧАТЬ ДОКУМЕНТОВ

- Файл – Предварительный просмотр
  Файл – Печать
- *Кнопка Office-Печать-Печать*
  Кнопка Office-Печать-Предварительный просмотр
  (Word 2007)

ИСПОЛЬЗОВАНИЕ ПОЛЕЙ

- Отображение сведений о документе, таких как имя автора, размер файла или число страниц. Для этого используются поля AUTHOR, FILESIZE, NUMPAGES или DOCPROPERTY.
- Выполнение вычислений. Для этого используется поле «= (Формула)».
- Работа с документами слияния. Например, вставьте поля ASK и FILLIN, чтобы выводить запрос, который заполняется при слиянии каждой записи данных с основным документом.

ЧТО ДОЛЖНЫ ЗНАТЬ СТУДЕНТЫ ПО MS WORD

- Установка параметров страниц
- Основные параметры настройки (форматирования) абзаца
- Способы вставки спецсимволов
- Использование закладок
- Оформление границ и заливки

ЧТО ДОЛЖНЫ ЗНАТЬ СТУДЕНТЫ ПО MS WORD (СТИЛИ)

- Использование стилей
  - Способы форматирования по образцу
  - Автоформатирование
- Принципы наследования стилей и шаблонов
- Использование шаблонов!!!
- Принципы корректной подготовки документов!!!

ЧТО ДОЛЖНЫ ЗНАТЬ СТУДЕНТЫ ПО MS WORD

- Режимы просмотра документов
- Создание оглавлений
- Разбиение на колонки
- Работа с таблицами
- Использование списков
- Номера страниц и колонтитулы

А ЕЩЁ WORD УМЕЕТ

- Создавать указатели
- Готовить письма для массовых рассылок
- Работать с графикой
- ...

ЛЕКЦИЯ №4

Тема №4. Представление и обработка графической и звуковой информации в компьютере

ОСНОВЫ КОДИРОВАНИЯ АНАЛОГОВОЙ ИНФОРМАЦИИ

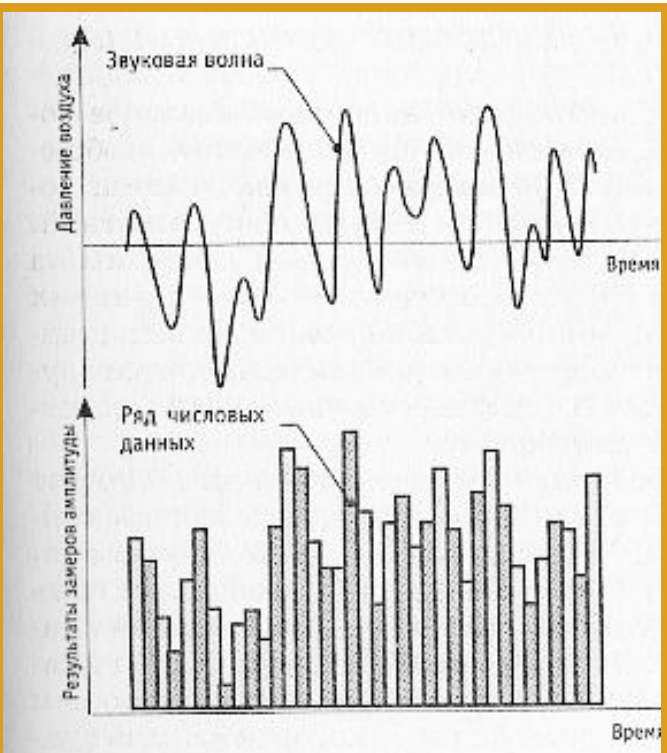


Рисунок 18. Числовые значения ряда данных пропорциональны мгновенным значениям амплитуды записываемого сигнала

- В основе кодирования звуковой и визуальной информации лежит **аналого-цифровое преобразование**: согласно этой схеме кодирования элементы информации записываются числами, значения которых – с точностью до округления – пропорциональны величине зарегистрированного сигнала
- Через краткие интервалы времени (с определённой частотой, называемой **частотой дискретизации**) производят замер амплитуды регистрируемого сигнала и запись её в числовом виде
- Применяется **правило Котельникова-Найквиста** не менее, чем двукратного превышения этой частоты над частотой (верхней гармоники) регистрируемого сигнала

ФОРМАТЫ ЗАПИСЕЙ

- Последовательность чисел, полученная в результате многократных замеров амплитуд сигнала, называется **выборкой данных**
- Запись выборки данных называют **записью в формате *RAW*** (это не сокращение, а английское слово *raw* – сырой)
- Рабочие **форматы записи** используют дополнительные соглашения о способах записи и обязательно конкретизируют такие данные как
- Частота дискретизации
- Количество каналов записи и взаимное расположение точек записи
- Глубина кодирования (в двоичных разрядах)<-> размер шага дискретизации

ПРЕДСТАВЛЕНИЕ ГРАФИЧЕСКИХ ДАННЫХ

- Область информатики, изучающая методы и средства обработки изображений с помощью программно-аппаратных вычислительных комплексов называется *компьютерной графикой*
- В зависимости от способа формирования изображений компьютерную графику принято подразделять на *растровую, векторную и фрактальную*
- Отдельным направлением является *трёхмерная (3D) графика,* изучающая приёмы и методы построения объёмных моделей объектов в виртуальном пространстве
  - Трёхмерность 3D подразумевает *прежде всего* поддержание сведений о наборах цветных элементов («точек»), размещённых в трёхмерном геометрическом пространстве *и затем уже (и пока что)* отображение этого набора на двумерном экране согласно строгим законам геометрических проекций

ЦИФРОВОЕ КОДИРОВАНИЕ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ

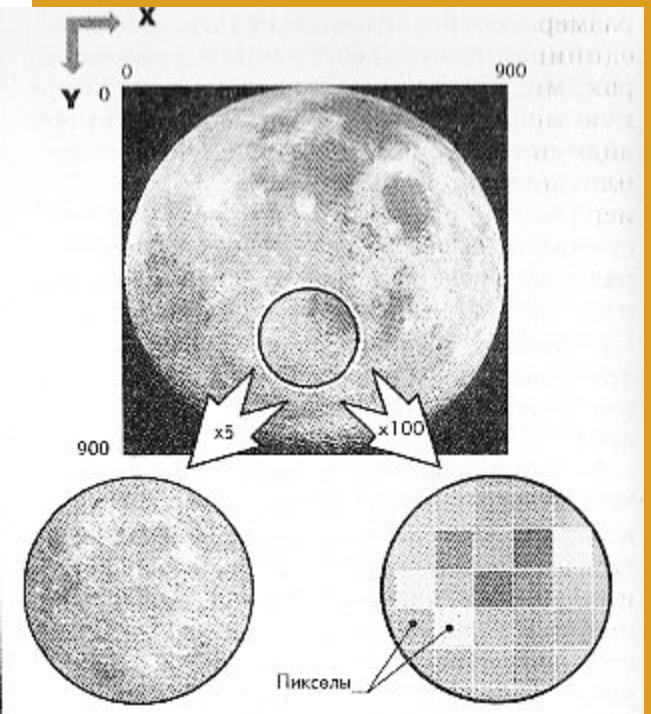


Рисунок 22. Математическая модель изображения

- Каждая точка двумерного изображения имеет некоторое положение, определяемое двумя координатами *X и Y,* и называется **пикселем** (сокращение от *picture cell* (*элемент изображения*))
- Параметр плотности точек на единицу длины называется **разрешающей способностью изображения** или просто **разрешением изображения**
- При большом увеличении изображения возникает дефект, называемый **эффектом пикселизации** **(см. рисунок)**
- Наиболее часто применяемой единицей разрешающей способности является количество точек на дюйм или *dpi (dots per inch)*

ПСИХОФИЗИОЛОГИЧЕСКИЕ МОДЕЛИ ЦВЕТОВОГО ПРОСТРАНСТВА

- *Цвет* чрезвычайно важен как средство усиления зрительного впечатления и повышения информационной насыщенности изображения
- В психофизиологических моделях визуальной информации каждая «точка» имеет размер и цвет
- Для излучающих объектов характерно *аддитивное цветовоспроизведение* (цветовые излучения суммируются), для отражающих - *субтрактивное*

МОДЕЛЬ RGB

- В основе цветовой модели *RGB* лежит положение о том, что любой произвольный цвет можно получить, смешивая три **основные цветовые составляющие**: красную *(Red, R)*, зелёную *(Green, G)* и синюю *(Blue, B)*
- Поскольку цвет получается сложением трёх цветов, эту модель называют **аддитивной** (*addition, англ. – сложение*). Белый цвет соответствует максимальным значениям составляющих: (255, 255, 255)
- В целом считается, что при использовании для представления одной точки 3 байтов, эта модель соответствует усреднённым психофизиологическим характеристикам человеческого зрения

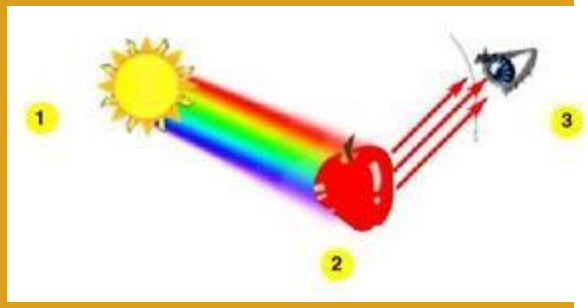
ЦВЕТОВАЯ МОДЕЛЬ CMYK

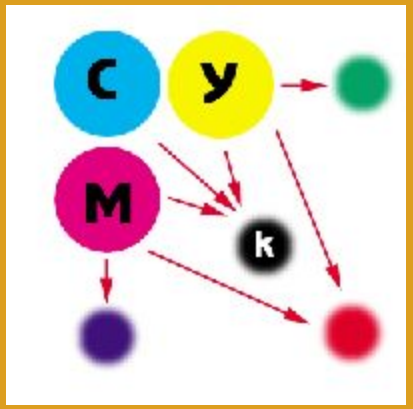
- Каждому из трёх основных цветов соответствует свой **дополнительный цвет:** голубой, пурпурный и жёлтый

  КРАСНЫЙ = БЕЛЫЙ – ГОЛУБОЙ
- При печати для воспроизведения цвета на странице вычитаются цветовые составляющие, «захваченные» тем или иным красителем:

  КРАСНЫЙ = ПУРПУРНЫЙ + ЖЁЛТЫЙ,

  а **голубой не используется**
- В основе модели *CMYK* лежат 4 составляющие: голубая (*Cyan, C*), пурпурная (*Magenta, M*), жёлтая (*Yellow, Y*) и чёрная (*BlacK, K*)
- Четвёртая, чёрная, составляющая добавлена для компенсации погрешностей, связанных с используемыми материалами. Максимальные значения составляющих (255, 255, 255, 255) соответствуют идеальному чёрному





МОДЕЛЬ ЦВЕТОВОГО КОЛЕСА

Модель цветового колеса

РАСТРОВАЯ, ВЕКТОРНАЯ И ФРАКТАЛЬНАЯ ГРАФИКА

- **Растровая компьютерная графика** оперирует попиксельными представлениями
- Растровую графику применяют при разработке электронных (мультимедийных) и полиграфических изданий. Иллюстрации, выполненные средствами растровой графики, редко создают вручную с помощью компьютерных программ. Чаще для этой цели используют сканированные иллюстрации, подготовленные художником на бумаге, или фотографии. В последнее время для ввода растровых изображений в компьютер нашли широкое применение цифровые фото- и видеокамеры. Соответственно, большинство графических растровых редакторов, предназначенных для работы с растровыми иллюстрациями, ориентированы не столько на создание изображений, сколько на их обработку. В Интернете пока применяются только растровые иллюстрации.

РАСТРОВАЯ, ВЕКТОРНАЯ И ФРАКТАЛЬНАЯ ГРАФИКА

- **Векторная компьютерная графика** работает на основании задания параметров фиксированного количества элементов (геометрических линий и фигур). Набором таких элементов компонуется целостная картинка
- Программные средства для работы с векторной графикой, в отличие от программ, работающих с растровой графикой, наоборот, предназначены, в первую очередь, для создания иллюстраций и в меньшей степени для их обработки. Такие средства широко используют в рекламных агентствах, дизайнерских бюро, редакциях и издательствах. Оформительские работы, основанные на применении шрифтов и простейших геометрических элементов, решаются средствами векторной графики намного проще. Существуют примеры высокохудожественных произведений, созданных средствами векторной графики, но они скорее исключение, чем правило, поскольку художественная подготовка иллюстраций средствами векторной графики чрезвычайно сложна

РАСТРОВАЯ, ВЕКТОРНАЯ И ФРАКТАЛЬНАЯ ГРАФИКА

- Программные средства для работы с ***фрактальной графикой*** предназначены для автоматической генерации изображений путем математических расчетов. *Создание фрактальной художественной композиции состоит не в рисовании или оформлении, а в программировании.* Фрактальную графику редко применяют для создания печатных или электронных документов, но ее часто используют в развлекательных программах