

Вводная лекция

Курс «Базы данных»

Вадим Цесько

Санкт-Петербургский государственный политехнический университет

9 февраля 2012 г.

Содержание

- 1 Цели и задачи курса
- 2 Настоящий курс
- 3 Информация vs данные
- 4 Информационные системы
- 5 Базы данных
- 6 Модели данных

Цель курса

Получение навыков проектирования и разработки информационных систем (ИС)

- Освоить основные понятия и определения

Задачи курса

- Освоить основные понятия и определения
- Познакомиться с базовым математическим аппаратом

- Освоить основные понятия и определения
- Познакомиться с базовым математическим аппаратом
- Изучить современные технологии и средства разработки

- Освоить основные понятия и определения
- Познакомиться с базовым математическим аппаратом
- Изучить современные технологии и средства разработки
- Получить практические навыки проектирования и разработки ИС

- Научиться ориентироваться в предметной области

- Научиться ориентироваться в предметной области
- Научиться самостоятельно изучать новые технологии и понимать их место

- Научиться ориентироваться в предметной области
- Научиться самостоятельно изучать новые технологии и понимать их место
- Ознакомиться с основными проблемами и их решениями

- Научиться ориентироваться в предметной области
- Научиться самостоятельно изучать новые технологии и понимать их место
- Ознакомиться с основными проблемами и их решениями
- Почувствовать тенденции в предметной области

- Цесько Вадим Александрович
- tsesko@kspt.ftk.spbstu.ru
- <http://kspt.ftk.spbstu.ru/people/tsesko/>

Заимствования:

- Курс «Базы данных» Моисеева М. Ю.
- Стэнфордский курс «Introduction to Databases»:
<http://www.db-class.org>
- Курс «Базы данных» Ильи Тетерина:
<http://www.lektorium.tv/course/?id=22828>

Рекомендуемые материалы по курсу:

- <http://www.db-class.org/course/resources/index>

Требования к слушателям

Технологии:

- OS Linux: <http://www.mandriva.com/en/>
- RDBMS Firebird: <http://www.firebirdsql.org/>
- Java: <http://www.oracle.com/technetwork/java/index.html>
- IDE NetBeans: <http://netbeans.org/>
- Trac: <http://trac.edgewall.org/>
- VCS Subversion: <http://subversion.tigris.org/>

Требования к слушателям

Технологии:

- OS Linux: <http://www.mandriva.com/en/>
- RDBMS Firebird: <http://www.firebirdsql.org/>
- Java: <http://www.oracle.com/technetwork/java/index.html>
- IDE NetBeans: <http://netbeans.org/>
- Trac: <http://trac.edgewall.org/>
- VCS Subversion: <http://subversion.tigris.org/>

Лабораторные работы:

- <http://tiger.ftk.spbstu.ru/trac/edu-db-2012/>
- <http://tiger.ftk.spbstu.ru/svn/edu-db-2012/>

- Задаёте вопросы по ходу лекции
- Отвечаете на вопросы в зал
- Планируются письменные коллоквиумы и домашние задания
- Презентации выкладывать **не планируется**

Информация vs данные

- **Информация** — любые сведения о каком-либо событии
- **Данные** — информация, фиксированная в определенной форме, удобной для передачи, хранения и обработки

Инфологический аспект:

- Какие объекты реального мира принимаются в рассмотрение
- Какие характеристики и взаимосвязи этих объектов учитываются

Инфологический аспект:

- Какие объекты реального мира принимаются в рассмотрение
- Какие характеристики и взаимосвязи этих объектов учитываются

Даталогический аспект:

- Разрабатываются формы представления информации
- Приводятся модели и методы преобразования данных

Классификация ИС (как пример)

Масштаб: индивидуальные, корпоративные, глобальные, ...

Классификация ИС (как пример)

Масштаб: индивидуальные, корпоративные, глобальные, ...

Область применения: информационно-справочные (поисковые),
управление контентом, ГИС, ...

Классификация ИС (как пример)

Масштаб: индивидуальные, корпоративные, глобальные, ...

Область применения: информационно-справочные (поисковые),
управление контентом, ГИС, ...

Способ построения:

- Файловые
- Клиент-серверные
- Трёх- и более уровневые
- ...

Классификация ИС (как пример)

Масштаб: индивидуальные, корпоративные, глобальные, ...

Область применения: информационно-справочные (поисковые), управление контентом, ГИС, ...

Способ построения:

- Файловые
- Клиент-серверные
- Трёх- и более уровневые
- ...

Модель данных:

- Реляционная
- Объектно-ориентированная (объектно-реляционная)
- Сетевая
- Иерархическая
- ...

- Файловые

История развития ИС

- Файловые
- Появление первой базы данных (БД) — 1960-е гг.

- Файловые
- Появление первой базы данных (БД) — 1960-е гг.
- Первое поколение БД — иерархические и сетевые

- Файловые
- Появление первой базы данных (БД) — 1960-е гг.
- Первое поколение БД — иерархические и сетевые
- Второе поколение БД — реляционные

- Файловые
- Появление первой базы данных (БД) — 1960-е гг.
- Первое поколение БД — иерархические и сетевые
- Второе поколение БД — реляционные
- Третье поколение БД — ООБД и ОРБД

Файловая ИС

Набор прикладных программ, хранящих данные в файлах в некотором формате

Недостатки:

- Разделение и изоляция данных

Файловая ИС

Набор прикладных программ, хранящих данные в файлах в некотором формате

Недостатки:

- Разделение и изоляция данных
- Дублирование данных

Файловая ИС

Набор прикладных программ, хранящих данные в файлах в некотором формате

Недостатки:

- Разделение и изоляция данных
- Дублирование данных
- Зависимость данных и программ

Файловая ИС

Набор прикладных программ, хранящих данные в файлах в некотором формате

Недостатки:

- Разделение и изоляция данных
- Дублирование данных
- Зависимость данных и программ
- Несовместимость форматов

Файловая ИС

Набор прикладных программ, хранящих данные в файлах в некотором формате

Недостатки:

- Разделение и изоляция данных
- Дублирование данных
- Зависимость данных и программ
- Несовместимость форматов
- Дублирование кода

База данных

Разделяемый набор логически связанных данных вместе с описанием данных

База данных

Разделяемый набор логически связанных данных вместе с описанием данных

Система управления БД (СУБД) обеспечивает

- эффективный

доступ и хранение

База данных

Разделяемый набор логически связанных данных вместе с описанием данных

Система управления БД (СУБД) обеспечивает

- эффективный
- надёжный

доступ и хранение

База данных

Разделяемый набор логически связанных данных вместе с описанием данных

Система управления БД (СУБД) обеспечивает

- эффективный
- надёжный
- удобный

доступ и хранение

База данных

Разделяемый набор логически связанных данных вместе с описанием данных

Система управления БД (СУБД) обеспечивает

- эффективный
- надёжный
- удобный
- безопасный

доступ и хранение

База данных

Разделяемый набор логически связанных данных вместе с описанием данных

Система управления БД (СУБД) обеспечивает

- эффективный
- надёжный
- удобный
- безопасный
- многопользовательский

доступ и хранение

База данных

Разделяемый набор логически связанных данных вместе с описанием данных

Система управления БД (СУБД) обеспечивает

- эффективный
- надёжный
- удобный
- безопасный
- многопользовательский

доступ и хранение

- больших объёмов

База данных

Разделяемый набор логически связанных данных вместе с описанием данных

Система управления БД (СУБД) обеспечивает

- эффективный
- надёжный
- удобный
- безопасный
- многопользовательский

доступ и хранение

- больших объёмов
- персистентных данных

- Стандартизированный формат хранения данных

Преимущества использования СУБД

- Стандартизированный формат хранения данных
- Доступ к данным через единый интерфейс

Преимущества использования СУБД

- Стандартизованный формат хранения данных
- Доступ к данным через единый интерфейс
- Высокоуровневый доступ к данным

Преимущества использования СУБД

- Стандартизированный формат хранения данных
- Доступ к данным через единый интерфейс
- Высокоуровневый доступ к данным
- Независимость данных и программ

Преимущества использования СУБД

- Стандартизированный формат хранения данных
- Доступ к данным через единый интерфейс
- Высокоуровневый доступ к данным
- Независимость данных и программ
- Эффективные алгоритмы работы с данными

Преимущества использования СУБД

- Стандартизированный формат хранения данных
- Доступ к данным через единый интерфейс
- Высокоуровневый доступ к данным
- Независимость данных и программ
- Эффективные алгоритмы работы с данными
- Автоматическая поддержка и контроль целостности данных

Преимущества использования СУБД

- Стандартизированный формат хранения данных
- Доступ к данным через единый интерфейс
- Высокоуровневый доступ к данным
- Независимость данных и программ
- Эффективные алгоритмы работы с данными
- Автоматическая поддержка и контроль целостности данных
- Ускоряется и упрощается разработка конечных ИС

- Стандартизированный формат хранения данных
- Доступ к данным через единый интерфейс
- Высокоуровневый доступ к данным
- Независимость данных и программ
- Эффективные алгоритмы работы с данными
- Автоматическая поддержка и контроль целостности данных
- Ускоряется и упрощается разработка конечных ИС
- Гибкость разработанных решений (возможность перехода на другую платформу или другую СУБД)

- Дополнительные затраты на преобразование данных

Недостатки использования СУБД

- Дополнительные затраты на преобразование данных
- Производительность при обработке больших объёмов слабоструктурированных данных

Недостатки использования СУБД

- Дополнительные затраты на преобразование данных
- Производительность при обработке больших объёмов слабоструктурированных данных
- Дополнительная стоимость

- Внешние (с точки зрения пользователя)

Уровни абстракции представления данных

- Внешние (с точки зрения пользователя)
- Концептуальный (семантическая информация, язык модели данных)

Уровни абстракции представления данных

- Внешние (с точки зрения пользователя)
- Концептуальный (семантическая информация, язык модели данных)
- Внутренний (формат хранения данных в БД)

Уровни абстракции представления данных

- Внешние (с точки зрения пользователя)
- Концептуальный (семантическая информация, язык модели данных)
- Внутренний (формат хранения данных в БД)
- Физический

- Модель данных
- Схема vs данные
- Data Definition Language
- Data Manipulation Language

Модель данных

Набор понятий, описывающих данные и способы обработки данных, связи между ними и ограничения, накладываемые на данные

Модель данных

Набор понятий, описывающих данные и способы обработки данных, связи между ними и ограничения, накладываемые на данные

Модель данных определяет:

- Возможные структуры
- Ограничения
- Операции

- Модель Entity-relationship (ER) предложена Питером Ченом в 1976 г.

- Модель Entity-relationship (ER) предложена Питером Ченом в 1976 г.
- Концептуальный уровень (семантическая информация о предметной области)

- Модель Entity-relationship (ER) предложена Питером Ченом в 1976 г.
- Концептуальный уровень (семантическая информация о предметной области)
- Определяет значения данных в контексте их взаимосвязи с другими данными

- Модель Entity-relationship (ER) предложена Питером Ченом в 1976 г.
- Концептуальный уровень (семантическая информация о предметной области)
- Определяет значения данных в контексте их взаимосвязи с другими данными
- Является наиболее общей — из нее могут быть порождены все существующие модели данных (иерархическая, сетевая, реляционная, объектная)

- Модель Entity-relationship (ER) предложена Питером Ченом в 1976 г.
- Концептуальный уровень (семантическая информация о предметной области)
- Определяет значения данных в контексте их взаимосвязи с другими данными
- Является наиболее общей — из нее могут быть порождены все существующие модели данных (иерархическая, сетевая, реляционная, объектная)
- Не определяет операций над данными и ограничивается лишь описанием их логической структуры

Модель сущность-связь (2)

- **Модель** — множество **сущностей**, между которыми определено множество **связей**

Модель сущность-связь (2)

- **Модель** — множество **сущностей**, между которыми определено множество **связей**
- **Сущность** — объект, который может быть идентифицирован неким способом, отличающим его от других объектов

Модель сущность-связь (2)

- **Модель** — множество **сущностей**, между которыми определено множество **связей**
- **Сущность** — объект, который может быть идентифицирован неким способом, отличающим его от других объектов
- **Набор сущностей** — множество сущностей одного типа

Модель сущность-связь (2)

- **Модель** — множество **сущностей**, между которыми определено множество **связей**
- **Сущность** — объект, который может быть идентифицирован неким способом, отличающим его от других объектов
- **Набор сущностей** — множество сущностей одного типа
- Сущность определяется значениями **атрибутов**, которые описывают свойства всех членов данного набора сущностей

Модель сущность-связь (2)

- **Модель** — множество **сущностей**, между которыми определено множество **связей**
- **Сущность** — объект, который может быть идентифицирован неким способом, отличающим его от других объектов
- **Набор сущностей** — множество сущностей одного типа
- Сущность определяется значениями **атрибутов**, которые описывают свойства всех членов данного набора сущностей
- Множество значений (область определения) атрибута называется **доменом**

Модель сущность-связь (3)

- **Связь** — ассоциация, установленная между двумя наборами сущностей

Модель сущность-связь (3)

- **Связь** — ассоциация, установленная между двумя наборами сущностей
- **Степень связи** — число сущностей, которое ассоциировано с помощью данной связи: $1 : 1$, $1 : N$, $N : N$ (обратите внимание на реализацию)

Модель сущность-связь (3)

- **Связь** — ассоциация, установленная между двумя наборами сущностей
- **Степень связи** — число сущностей, которое ассоциировано с помощью данной связи: $1 : 1$, $1 : N$, $N : N$ (обратите внимание на реализацию)
- **Обязательные и необязательные связи**

Пример

На доске.