

Модуль 2.

Основы реляционных БД

12.05.2016

Введение в базы данных. Мяснов А., Савин И., 2016 г.

Содержание модуля

- Введение в реляционную модель данных, преимущества
- Ахитектура ANSI/SPARC
- Запросы к SQL-базам данных: UNION, JOIN
- Триггеры и хранимые процедуры

Понятие модели данных

- Модель данных
 - Понятия (объекты)
 - Операторы
 - Описание структуры и доступа к данным
- Общая характеристика модели данных (Кристофер Дейт)
 - Структурная часть
 - Манипуляционная часть
 - Ограничения целостности
- Смешение и разделение понятий
 - Реляционная и SQL модели данных

Реляционная модель данных

- Сложно и просто
 - Математический аппарат
 - Простые формализмы
- Структурная часть
 - Набор понятий, родовая структура
- Манипуляционная часть
 - Операции на данными
- Ограничения целостности
 - Средства поддержания структуры и содержимого БД в соответствии модели данных

РМД. Структурная часть

- Тип данных
 - Все элементы типизированы
 - Аналогично понятию типа в ЯП:
 - Множество значений
 - Операции
 - Литералы
 - Целочисленные типы данных:
 - Данные с дробной частью:
 - Строковые:
 - Календарные:
 - Другие:
 - NULL

РМД. Структурная часть

- Домен
 - Базовый тип данных
 - Логическое выражение над элементом базового типа
- Подмножество значений типа данных
- Примеры:

РМД. Структурная часть

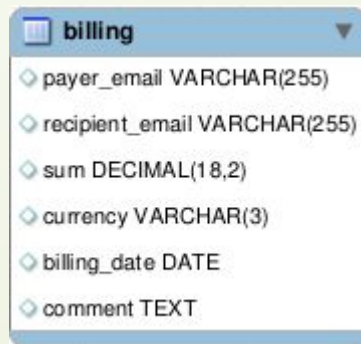
- Атрибут
 - Именованная характеристика – свойство сущности
 - Определяется на некотором типе данных или домене
- Примеры:

РМД. Структурная часть

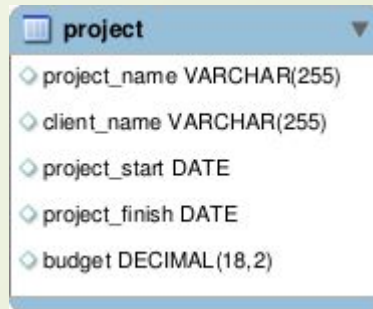
- Кортеж
 - Множество упорядоченных триплетов: имя атрибута, типа данных или домен, значение
 - Примеры:

РМД. Структурная часть

- Схема отношения, схема БД
 - Схема отношения – именованное множество упорядоченных пар: имя атрибута, тип данных или домен
 - Примеры:



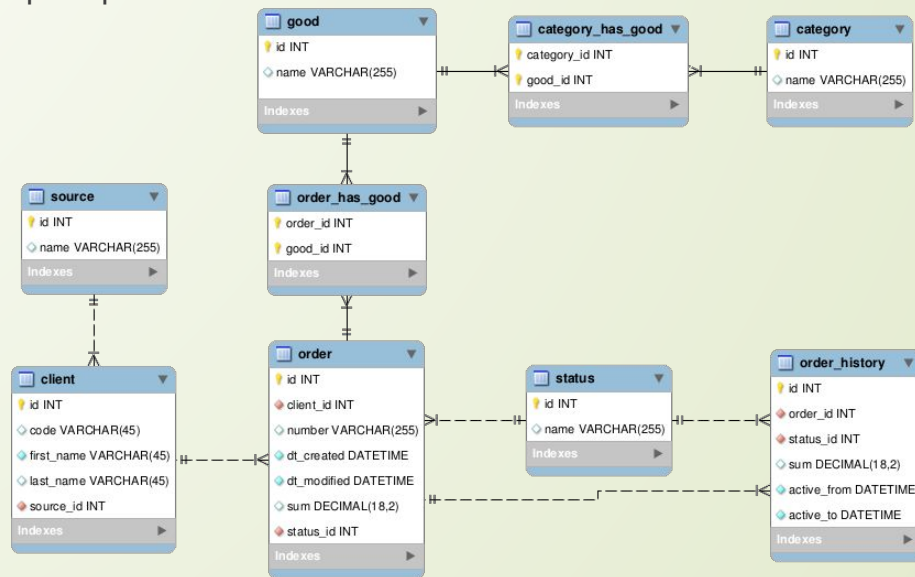
billing	
◇ payer_email	VARCHAR(255)
◇ recipient_email	VARCHAR(255)
◇ sum	DECIMAL(18,2)
◇ currency	VARCHAR(3)
◇ billing_date	DATE
◇ comment	TEXT



project	
◇ project_name	VARCHAR(255)
◇ client_name	VARCHAR(255)
◇ project_start	DATE
◇ project_finish	DATE
◇ budget	DECIMAL(18,2)

РМД. Структурная часть

- Схема отношения, схема БД
 - Схема БД – множество именованных схем отношений
- Примеры:



РМД. Структурная часть

- Отношение
 - Тело отношения – множество кортежей, соответствующих схеме отношения
 - Значение отношения – пара множеств: схема отношения, тело отношения

РМД. Структурная часть

- Интуитивная интерпретация
 - Таблицы, столбцы, строки, поля

Атрибуты

Типы данных

Домены?

Схема отношения

Кортежи

Тело отношения

#	Field	Type	Null	Key	Default	Extra
1	project_name	varchar(255)	YES			
2	client_name	varchar(255)	YES			
3	project_start	date	YES			
4	project_finish	date	YES			
5	budget	decimal(8,2)	YES			

#	project_name	client_name	project_start	project_finish	budget
1	Accidentally Railroad	Bradly Noonan	2027-04-20	2021-10-20	64741.00
2	Oyster Stormy	Lina Southey	2027-02-20	2029-12-20	297752.00
3	Subtle Poseidon	Liliana Payton	2005-01-20	2002-08-20	831756.00
4	Accidentally Skilled Pottery	Lina Southey	2031-05-20	2013-10-20	157639.00
5	Liquid Panther	Poppy Whitney	2018-02-20	2021-08-20	359697.00
6	Rapid Monkey	Tavia Chatham	2006-02-20	2022-12-20	969210.00
7	Insane Bulldozer	Casey Copeland	2018-01-20	2027-09-20	122786.00
8	Rapid Limousine	Lina Southey	2018-06-20	2010-07-20	498218.00
9	Confidential Metaphor	Copper Smith	2005-05-20	2002-07-20	254164.00
10	Hollow Eternal Nitrogen	Liliana Payton	2007-04-20	2018-10-20	716153.00
11	Essential Peasant	Catalina Barlow	2012-05-20	2025-12-20	795797.00
12	Restless Monkey	Tavia Chatham	2021-02-20	2006-11-20	561243.00
13	Aberrant Insane Eagle	Xavier Abram	2002-06-20	2028-08-20	839598.00
14	Hot Pure Mountain	Fearn Hayden	2009-02-20	2019-09-20	919429.00
15	Sleepy Uranium	Pell Dudley	2011-05-20	2004-12-20	991093.00
16	Dreaded Steady Weather	Ballard Ridley	2003-02-20	2001-12-20	670338.00
17	Helpless Shower	Blade Padley	2023-05-20	2024-10-20	78214.00
18	Stormy Flannel	Ballard Ridley	2017-02-20	2026-07-20	842305.00
19	Dead Sapphire	Copper Smith	2019-06-20	2003-11-20	455783.00
20	Appropriate Swallow	Bradly Noonan	2010-06-20	2023-12-20	702889.00
21	Omega Accidentally	Krista Stamper	2008-03-20	2019-07-20	493309.00
22	Green Storm	Elmore Shelly	2008-01-20	2013-10-20	589506.00
23	Rare Emerald	Merry Ramsay	2002-03-20	2026-12-20	630439.00
24	Maroon Snake	Fearn Hayden	2023-05-20	2005-07-20	656082.00
25	Grim Sledgehammer	Liliana Payton	2009-06-20	2003-10-20	287998.00
26	Heavy Cobra	Xavier Abram	2002-06-20	2005-07-20	246467.00
27	Silver Rhinestone	Blade Padley	2011-02-20	2003-08-20	499913.00

РМД. Структурная часть

- Свойства отношений
 - Отсутствие кортежей дубликатов – вытекает из определения тела отношения как множества кортежей
 - Отсутствие упорядоченности кортежей
 - Формально - из множественной природы отношений
 - На практике - эффективно
 - Отсутствие упорядоченности атрибутов
 - Аналогично упорядоченности кортежей
 - Атомарность значений атрибутов
 - Условность
 - Схема отношения принадлежит первой нормальной форме

РМД. Манипуляционная часть

- Механизмы
 - Реляционная алгебра
 - Реляционное исчисление
 - DML SQL
 - Примеры:

РМД. Ограничения целостности

- Целостность сущности
- Целостность ссылок

РМД. Целостность сущности

- Кортежи должны быть отличимы. Почему?
 - Понятие потенциального ключа:
 - Минимальное подмножество множества атрибутов отношения такое, что для любого кортежа отношения набор значений отличается от другого
- Свойства потенциального ключа:
- Уникальность
 - Минимальность
- Виды ключей по количеству атрибутов
 - Простые
 - Составные

РМД. Целостность сущности

- Практическое значение
 - В таблице не должно быть повторяющихся записей
 - SQL?
 - Ключ идентифицирует кортеж/строку таблицы
 - SQL?
 - Атрибуты, входящие в состав ключа не могут содержать неопределенных значений
 - SQL?

РМД. Внешний ключ

- Специальный атрибут (набор атрибутов) отношения такой, что значения данного атрибута (набора) в некотором отношении будут соответствовать значениям атрибутов первичного ключа некоторого отношения
- Можно сказать, что множество значений внешнего ключа некоторого отношения R2 является нестрогим подмножеством множества значений потенциального ключа некоторого отношения R1
- Терминология: в приведенном выше примере отношение R1 называют главным/целевым/родительским (master), R2 – подчиненным (slave)
 - Примеры:

РМД. Внешний ключ

- Примеры:

РМД. Ссылочная целостность

- Для каждого указанного значения внешнего ключа должен найтись кортеж в отношении, на которое ссылается набор атрибутов внешнего ключа
- SQL: Или значения внешнего ключа должны быть неопределенными (NULL)
- Способы поддержания ссылочной целостности:
 - Блокировка операции;
 - Указание NULL;
 - Каскадное удаление.

РМД. Преимущества

- Контроль целостности данных;
- Гибкость: выборка, изменение схемы;
- Физическая и логическая независимость;
- Разработка прикладных программ;
- Основа – небольшое число интуитивно понятных абстракций;
- Теоретическим базисом реляционного подхода к организации баз данных служит простой и мощный математический аппарат теории множеств и математической логики;
- Декларативные языки определения и манипулирования данными;
- Средства разграничения доступа.

РМД. Недостатки

- Производительность;
- Большое количество таблиц в реальных БД;
- Сложность поддержки;
- Некоторые предметные области плохо представляются в форме отношений;
- Разработка прикладных программ
- Репликация.

Архитектура ANSI SPARC

- Определяет принцип организации СУБД (не обязательно РМД)
- Цель: введение уровней абстракции для пользователей и архитекторов БД
 - Разные группы пользователей ИС должны иметь разные представления данных
 - Пользователи должны быть абстрагированы от физической и логической организации данных
 - Администраторы БД должны иметь возможность вносить изменения в схему данных без учета особенностей физического хранения данных
 - Внутренняя структура БД не должна зависеть от физических аспектов хранения информации

Архитектура ANSI SPARC

- Уровни и функции
 - 1. Внешний уровень – пользовательские представления
 - 2. Концептуальный уровень – логическая структура БД
 - 3. Внутренний уровень – аспекты физической реализации
- Логическая независимость – защищенность внешних схем от изменений, вносимых в концептуальную схему. Сохранение представлений внешнего уровня при изменении схемы данных.
- Физическая независимость – защищенность концептуальной схемы от изменений, вносимых во внутреннюю схему. Файловые системы и системные функции, организация хранения данных, настройки и внутренние структуры данных.

Запросы из нескольких таблиц. Виды соединений

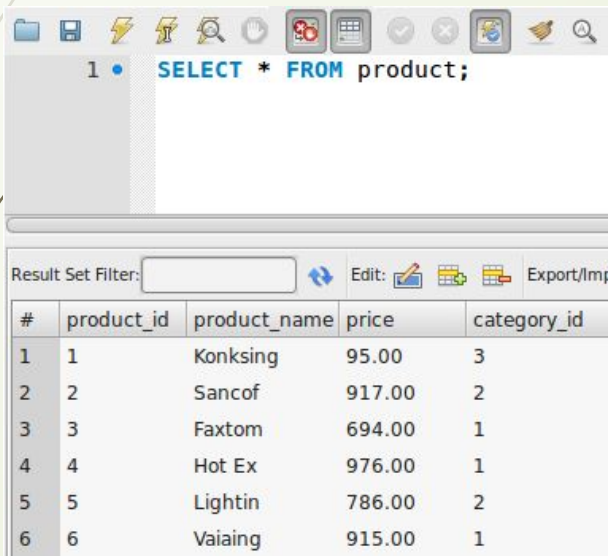
- Оператор SELECT возвращает результат соединения, если обращается к нескольким источникам данных
 - Таблицы
 - Представления
 - Результаты оператора SELECT – подзапрос
- Результат соединения нестрогое подмножество декартова произведения результатов выборок из нескольких источников

Запросы из нескольких таблиц. Виды соединений

- Декартово произведение: CROSS JOIN
- Внутреннее соединение: INNER JOIN
- Левое и правое внешнее соединение: LEFT OUTER JOIN, RIGHT OUTER JOIN
- Полное внешнее соединение: FULL OUTER JOIN
- Естественное соединение: NATURAL JOIN

Запросы из нескольких таблиц. Виды соединений

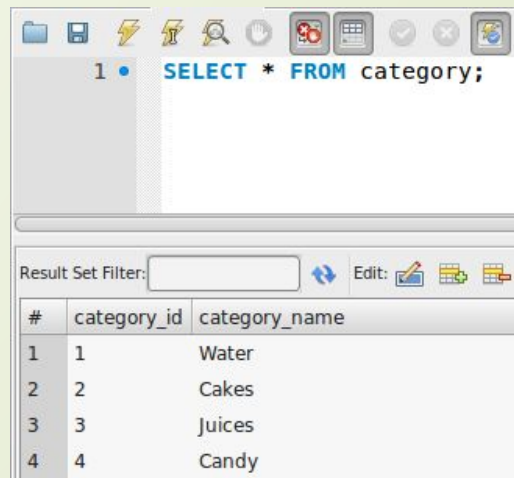
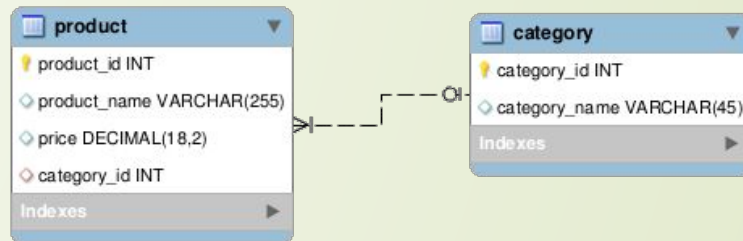
- Базовый пример



1 • **SELECT** * **FROM** product;

Result Set Filter: Edit: Export/Imp

#	product_id	product_name	price	category_id
1	1	Konksing	95.00	3
2	2	Sancof	917.00	2
3	3	Faxtom	694.00	1
4	4	Hot Ex	976.00	1
5	5	Lightin	786.00	2
6	6	Vaiaing	915.00	1



1 • **SELECT** * **FROM** category;

Result Set Filter: Edit:

#	category_id	category_name
1	1	Water
2	2	Cakes
3	3	Juices
4	4	Candy

Запросы из нескольких таблиц. Декартово произведение

- Количество строк в соединении равно произведению количества строк в соединяемых результатах
- Декартово произведение получается в результате любого запроса, содержащего более одного источника и не имеющего ограничивающих условий
 - Пример
- Зачем?

Запросы из нескольких таблиц. Тета-соединение

- Результат тета-соединения содержит кортежи из декартова произведения, удовлетворяющие некоторому условию
- В языке SQL операция тета-соединения – INNER JOIN (внутреннее соединение), NATURAL JOIN
- Частный случай – соединение по эквивалентности – эквисоединение
 - Примеры

Запросы из нескольких таблиц. Левое, правое внешние соединения

- Результат левого внешнего соединения содержит кортежи из внутреннего соединения источников и не вошедшие во внутреннее соединение кортежи левого источника
- Атрибуты в кортежах, которые не имеют совпадений по общим столбцам заполняются неопределенными значениями
- Правые внешние соединения...
 - Примеры

Запросы из нескольких таблиц. Полное внешнее соединение

- Результат полного внешнего соединения содержит кортежи:
 - из внутреннего соединения источников,
 - не вошедшие во внутреннее соединение кортежи левого источника,
 - не вошедшие во внутреннее соединение кортежи правого источника.
- Атрибуты в кортежах, которые не имеют совпадений по общим столбцам заполняются неопределенными значениями
 - Примеры

Запросы из нескольких таблиц. Объединение

- Результат объединения кортежи из каждого источника
- Если добавлено ключевое слово ALL, то не выполняется проверка на уникальность кортежей в результирующем объединении
- Выборки из источников должны совпадать по атрибутам
 - Примеры

Процедурные расширения SQL

- SQL/PSM
 - Операторы SQL/DDDL, SQL/DML
 - Определения и вызовы функций и процедур
 - Управляющие конструкции, состояния
- Вычислительно полный язык

Процедурные расширения SQL

- Использование:
 - Хранимые процедуры
 - Триггеры
- Вариативность
 - Индивидуальность языка каждой СУБД
 - Типизация, назначение прав
 - Организация выполнения

Хранимые процедуры

- Принципы работы
 - Исполняются сервером БД
 - Могут осуществлять изменение данных и структуры БД
 - Могут возвращать работы данных аналогично оператору SELECT
 - Однократно компилируются
 - Могут кэшировать планы исполнения, результаты
- Терминология
 - Пользовательские функции
 - “Нативные” хранимые процедуры
 - Процедуры, реализуемые на языках общего назначения

Хранимые процедуры

- Преимущества:
 - Производительность
 - Организация поддержки целостности
 - Скрытие структуры данных
 - Модульность
 - Безопасность
 - Использование общего кода при работе с БД нескольких приложений

Хранимые процедуры

- Недостатки:
 - Разделение бизнес-логики
 - Меньшая языковая выразительность процедурных расширений SQL
 - Непереносимость
 - Погружение в особенности СУБД
 - Отладка
 - Версионирование

Триггеры

- Обработка событий
- Момент вызова: до и после операций с предметной таблицей
- Иницилирующие операции:
 - Вставка
 - Изменение
 - Удаление

Триггеры

- Особенности работы триггеров:
 - Могут быть инициированы изменением схемы или данных
 - Могут выполняться для каждой строки или для всего блока
 - Могут содержать дополнительные условия исполнения
 - Возможности использования старых и новых значений
 - Вызов процедур, инициирование триггеров
 - Транзакционность и сохранение состояния
 - Последовательность исполнения

Триггеры

- Варианты использования:
 - Проверка данных
 - Поддержка согласованности
 - Журналирование и аудит
 - Запуск процедур не связанных с обработкой данных
- Недостатки
 - Непрозрачность
 - Сложность отладки
 - Производительность
 - Глобальность