



ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ

ОСНОВЫ ПРОЕКТИРОВАНИЯ БАЗ ДАННЫХ

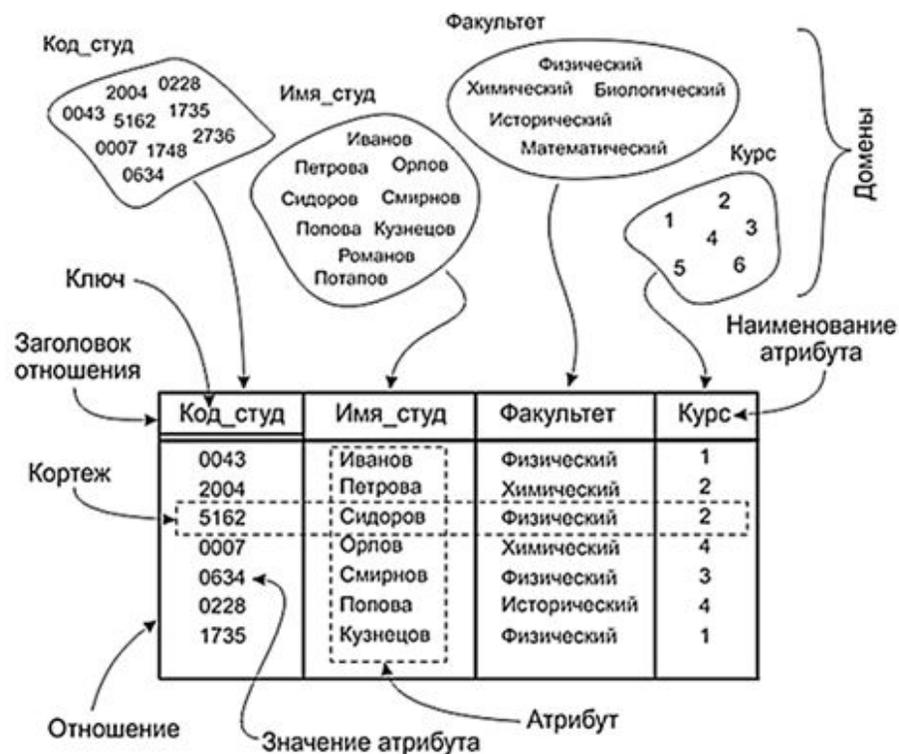


ОСНОВНЫЕ ПОНЯТИЯ РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ

- В основе реляционных систем лежит *реляционная модель данных*.
- Принципы реляционной модели были заложены в 1969–1970 гг. американским ученым *Е. Ф. Коддом* (E. F. Codd), в то время работавшим в корпорации IBM. Будучи математиком по образованию, он привнес в область управления базами данных строгие математические принципы и точность, которых не хватало ранним системам. Хотя реляционный подход утвердился не сразу, можно отметить, что почти все созданные с конца 70-х гг. продукты баз данных основаны именно на реляционном подходе.
- Подавляющее большинство научных исследований в области баз данных в течение последних 35 лет также проводилось именно в этом направлении.
- *Рассматривая и постепенно уточняя основные понятия реляционной модели, будем иметь в виду три компоненты модели данных:*
 - структуры данных,
 - операции, которые можно выполнять над данными, и
 - ограничения, связанные с обеспечением целостности данных.

ОСНОВНЫЕ ПОНЯТИЯ РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ

- Основной структурой данных в реляционной модели являются **таблицы**, называемые в реляционной теории отношениями. Собственно от термина **отношение** (по-английски relation) и произошло само название модели – **реляционная**. На рисунке приведен пример такой таблицы-отношения и пояснение основных терминов реляционной модели – **кортеж**, **кардинальное число**, **атрибут**, **степень**, **домен**, **первичный ключ**.



ОСНОВНЫЕ ПОНЯТИЯ РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ

- *Отношение это таблица, подобная приведенной на рисунке, состоящая из строк и столбцов и имеющая вверху строку, называемую **заголовком отношения**.*
- Строки таблицы-отношения называются *кортежами* (tuple), а столбцы *атрибутами* (attribute).
- *Количество кортежей в отношении называется кардинальным числом отношения, а количество атрибутов называется степенью отношения.*
- Каждый атрибут в отношении имеет *наименование*, которое указывается в заголовочной части отношения.
- *Ключ отношения* – это атрибут или набор атрибутов отношения такие, что в любой момент времени в отношении не существует строк, для которых значение или комбинация значений ключевых атрибутов являются одинаковыми. *Ключ*, таким образом, является *уникальным идентификатором кортежей* отношения (на рисунке ключевой атрибут выделен жирным шрифтом).
- *Домен отношения* – это множество значений, из которого могут браться значения конкретного атрибута. То есть конкретный набор значений атрибута в любой момент времени должен быть подмножеством множества значений домена, на котором определен этот атрибут. Значения атрибута, которые отсутствуют в множестве, задаваемом доменом, являются недопустимыми.

ОСНОВНЫЕ ПОНЯТИЯ РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ

- Понятие домена является важным для реляционной модели. Домен фактически задает ограничения, которым должны удовлетворять значения соответствующего атрибута.
- Как уже отмечалось, приведенные выше определения не являются строгими. Такие термины как таблица, строка, столбец, строго говоря, не являются полностью эквивалентными используемым в реляционной модели математическим понятиям отношение, кортеж, атрибут соответственно. Однако на практике их часто используют именно как синонимы, что, в общем, допустимо, если при этом понимать, какой действительный смысл вкладывается в эти термины.

ЭТАПЫ И ОСНОВНЫЕ ПРИНЦИПЫ ПРОЕКТИРОВАНИЯ БАЗЫ ДАННЫХ

■ **Основные задачи проектирования баз данных:**

- Обеспечение хранения в БД всей необходимой информации.
- Обеспечение возможности получения данных по всем необходимым запросам.
- Сокращение избыточности и дублирования данных.
- Обеспечение целостности данных (правильности их содержания): исключение противоречий в содержании данных, исключение их потери и т.д.

■ **Основные этапы проектирования баз данных:**

- **1) Концептуальное (инфологическое) проектирование** – построение формализованной модели предметной области. Такая модель строится с использованием стандартных языковых средств, обычно графических, например **ER-диаграмм** (диаграмм «Сущность-связь»). Такая модель строится без ориентации на какую-либо конкретную СУБД.

- Основные элементы данной модели:
- Описание объектов предметной области и связей между ними.
- Описание информационных потребностей пользователей (описание основных запросов к БД).
- Описание алгоритмических зависимостей между данными.
- Описание ограничений целостности, т.е. требований к допустимым значениям данных и к связям между ними.

ЭТАПЫ И ОСНОВНЫЕ ПРИНЦИПЫ ПРОЕКТИРОВАНИЯ БАЗЫ ДАННЫХ

- **2) Логическое (даталогическое) проектирование** – отображение инфологической модели на модель данных, используемую в конкретной СУБД, например на реляционную модель данных. Для реляционных СУБД **даталогическая модель** – набор таблиц, обычно с указанием ключевых полей, связей между таблицами. Если инфологическая модель построена в виде **ER-диаграмм** (или других формализованных средств), то даталогическое проектирование представляет собой построение таблиц по определённым формализованным правилам, а также нормализацию этих таблиц. Этот этап может быть в значительной степени автоматизирован.
- **3) Физическое проектирование** – реализация даталогической модели средствами конкретной СУБД, а также выбор решений, связанных с физической средой хранения данных: выбор методов управления дисковой памятью, методов доступа к данным, методов сжатия данных и т.д. – эти задачи решаются в основном средствами СУБД и скрыты от разработчика БД.
- **На этапе инфологического проектирования в ходе сбора информации о предметной области требуется выяснить:**
 - основные объекты предметной области (объекты, о которых должна храниться информация в БД);
 - атрибуты объектов;
 - связи между объектами;
 - основные запросы к БД.

ЭТАПЫ И ОСНОВНЫЕ ПРИНЦИПЫ ПРОЕКТИРОВАНИЯ БАЗЫ ДАННЫХ

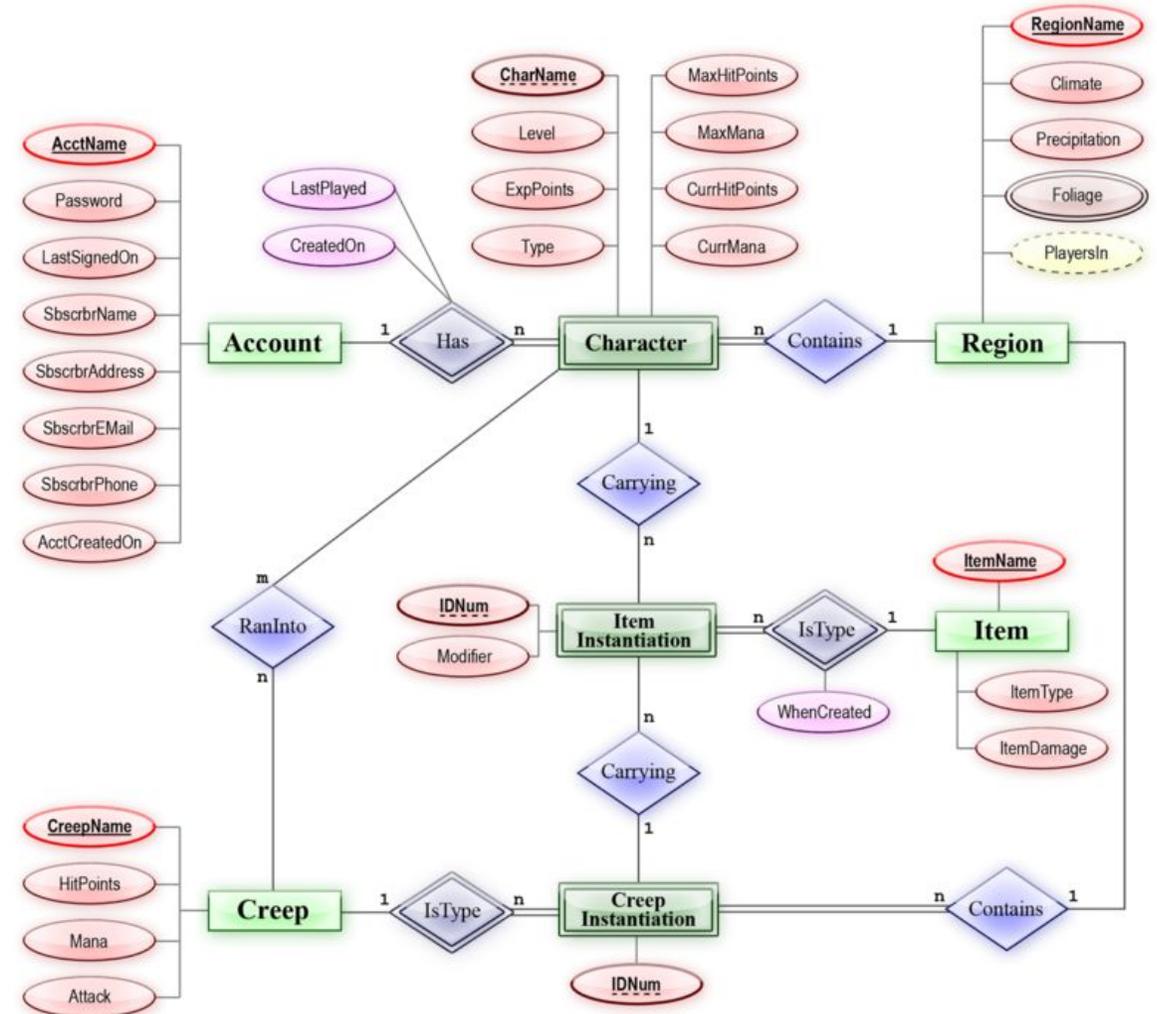
- **Принципы разработки многопользовательских баз данных** должны сводиться к соблюдению двух обязательных условий: *системного подхода* и *стандартизации*.
- **Системный подход.** Системный подход к разработке информационной системы означает, что такая система рассматривается как большая система, состоящая из некоторого множества взаимосвязанных и взаимодействующих между собой элементов. При проектировании информационных систем необходимо соблюдать следующие принципы:
 - учет интересов всех потенциальных пользователей систем;
 - модульный принцип разработки и внедрения.
- **Стандартизация.** Стандартизация разработки информационных систем, учитывая их многопользовательский характер, имеет *следующие аспекты*:
 - информационный;
 - программный;
 - аппаратный.
- Стандартизация *информационного* обеспечения обусловлена принципами компьютерной обработки символьной информации, так как объекты баз данных должны однозначно распознаваться компьютером.

ER-МОДЕЛЬ ДАННЫХ

- **Модель сущность-связь** (ERM) — модель данных, позволяющая описывать концептуальные схемы предметной области.
- ER-модель используется при высокоуровневом (концептуальном) проектировании баз данных. С её помощью можно выделить **ключевые сущности** и **обозначить связи**, которые могут устанавливаться между этими сущностями.
- Во время проектирования баз данных происходит преобразование ER-модели в конкретную схему базы данных на основе выбранной модели данных (реляционной, объектной, сетевой или др.).
- ER-модель представляет собой **формальную конструкцию**, которая сама по себе не предписывает никаких графических средств её визуализации.
- **Модель «сущность-связь»** была предложена в 1976 году Питером Пин-Шен Ченом (англ. Peter Pin-Shen Chen), американским профессором компьютерных наук в университете штата Луизиана.

ER-МОДЕЛЬ ДАННЫХ

- **Нотация Питера Чена**
- Множества **сущностей** изображаются в виде прямоугольников, **множества отношений** изображаются в виде ромбов.
- Если сущность участвует в **отношении**, они связаны линией. Если отношение не является обязательным, то линия пунктирная.
- **Атрибуты** изображаются в виде овалов и связываются линией с одним отношением или с одной сущностью



ПРЕОБРАЗОВАНИЕ КОНЦЕПТУАЛЬНОЙ МОДЕЛИ В РЕЛЯЦИОННУЮ

- Преобразование концептуальной модели в реляционную состоит в следующем:
 - Построить набор предварительных таблиц и указать первичные ключи.
 - Провести процесс нормализации.
- Первый пункт мы рассматривали в третьем уроке, со вторым мы пока не знакомы, но ознакомимся на практике. Итак, нам надо построить набор таблиц.
- Сделать это несложно, т.к. таблицы - это наши объекты, а поля таблиц - атрибуты объектов. Набор предварительных таблиц, исходя из нашей концептуальной модели, выглядит так:

Покупатель	Поставщик	Покупка	Поставка
Id покупателя (PK)	Id поставщика (PK)	Id покупки (PK)	Id поставки (PK)
ФИО	Наименование	Id покупателя (FK)	Id поставщика (FK)
E-mail	Адрес	Дата	Дата

Товар	Журнал покупок	Журнал поставок
Id товара (PK)	Id покупки (FK)	Id поставки (FK)
Наименование	Id товара (FK)	Id товара (FK)
Цена	Количество	Количество

Рис. 1. Набор предварительных таблиц, исходя из нашей концептуальной модели.

ПРЕОБРАЗОВАНИЕ КОНЦЕПТУАЛЬНОЙ МОДЕЛИ В РЕЛЯЦИОННУЮ

- Таким образом, у нас определены *таблицы, поля, первичные ключи* (PK) и *связи* (FK).
- В таблицах Журнал поставок и Журнал покупок первичные ключи - *составные*, т.е. состоят из *двух полей*.
- Теоретически бывают таблицы, в которых все поля являются одним *составным ключом*.
- **Нормализация** - это пошаговый, обратимый процесс замены исходной схемы другой схемой, в которой таблицы имеют более простую и логичную структуру. Это нужно для устранения избыточности данных.

Id сообщения	Тема	Текст	Автор
1	О рыбалке	Думаю надо сделать так...	Кирилл
2	О рыбалке	Согласен	Вася
3	О рыбалке	А еще можно сделать так...	Семен
4	Велосипеды	Согласен	Вася