

Управление данными

Лекция 7

Проектирование базы данных

Свечников С.В.

Содержание занятия

01. Моделирование данных
02. Разработка ER-диаграммы
03. Нормализация данных

01. Моделирование данных

Моделирование ситуации

Проектирование БД представляет собой итерационный процесс.

Перед началом проектирования БД организации необходимо:

- изучить ее предметную область
- изучить процессы, проходящие в ней
- изучить документацию организации
- опросить сотрудников организации
- изучить существующие формы, накладные и пр.

Необходимо разобраться в структуре данных и тех правилах, которые определяют их взаимосвязь.

Для проектирования модели данных необходимо выявить сущности и связи между ними.

ER-моделирование

ER-диаграмма (диаграмма сущность-связь) - модель данных, позволяющая описывать концептуальные схемы предметной области.

Первая ER-модель была разработана Питером Ченом в 1976 г.

ER-модель оперирует следующими понятиями:

- сущность
- связь
- атрибут сущности

Сущности

Сущность – класс однотипных объектов, информация о которых должна быть учтена в модели. Это реально существующие элементы окружающего мира.

Примеры

Студенты (Person)
Группы (Gruppa)
Курсы (Course)

Сущность представляет собой логическую категорию, а не конкретный предмет.

Экземпляр сущности – конкретный представитель сущности.

Примеры

Иванов
ИТС-1-10
Управление данными

Атрибуты

Атрибут сущности – это именованная характеристика, являющаяся некоторым свойством сущности.

Атрибуты делятся на:

- однозначные – имеющие одно значение для каждого экземпляра сущности
- многозначные – имеющие несколько значений

Примеры

Возраст

Пол

Телефон

Работа

Ключ сущности – это избыточный набор атрибутов, значения которых в совокупности являются уникальными для каждого экземпляра сущности.

СВЯЗИ

Между сущностями существуют связи. **Связь** – это некоторая ассоциация между двумя сущностями.

Связи обладают следующими свойствами:

- тип связи
- модальность связи

Типы связей

Существует 3 типа связей

1:1

Связь **один-к-одному**: когда один экземпляр сущности А связан с одним экземпляром сущности В.

1:∞

Связь **один-ко-многим**: когда один экземпляр сущности А может быть связан с несколькими экземплярами сущности В.

∞:∞

Связь **многие-ко-многим**: когда каждый экземпляр сущности А может быть связан с несколькими экземплярами сущности В и каждый экземпляр сущности В может быть связан с несколькими экземплярами сущности А.

Модальности связей

Существует две модальности связей.

Модальность «**может**» - означает, что экземпляр сущности А может быть связан с одним или несколькими экземплярами сущности В, а может быть и не связан ни с одним экземпляром.

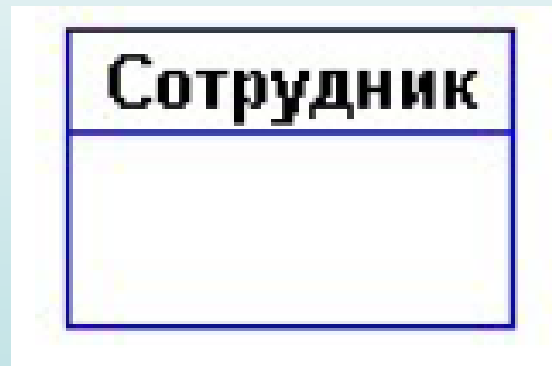
Модальность «**должен**» - означает, что экземпляр сущности А обязан быть связан не менее чем с одним экземпляром сущности В.

02. Разработка ER-диаграммы

Наглядность диаграммы

Диаграммы сущность-связь основываются на одной идее – рисунок всегда нагляднее текстового описания. Все ER-диаграммы используют графическое изображение сущностей предметной области, их свойств – атрибутов и взаимосвязей между сущностями.

Каждая сущность в модели изображается в виде прямоугольника с наименованием.



Изображаем атрибуты

Атрибуты сущности изображаются в пределах прямоугольника, определяющего его сущность.



Ключевые атрибуты изображаются на диаграмме подчеркиванием.

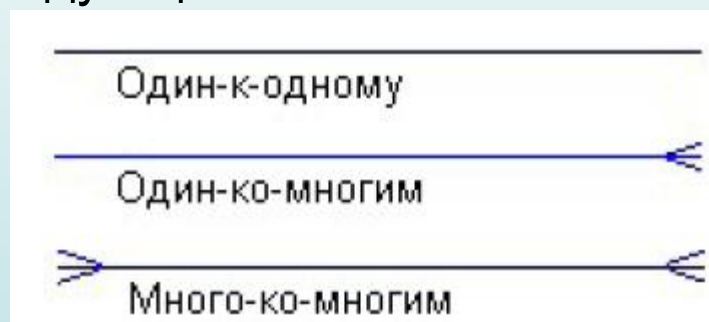


Изображаем связь

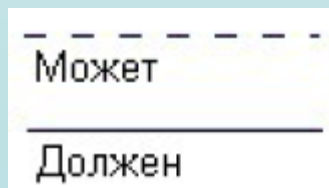
Графически связь изображается линией, соединяющей две сущности.



Каждая связь может иметь один из следующих типов



Каждая связь может иметь одну из модальностей, причем может иметь разную модальность с разных концов.



Разработка ER-модели

При разработке ER-модели необходимо получить следующую информацию о предметной области:

1. Список сущностей предметной области
2. Список атрибутов сущностей
3. Описание взаимосвязей между сущностями

Сперва определяем действия, которые должна выполнять проектируемая система.

Пример

1. Хранить информацию о покупателях
2. Печатать накладные на отпущенные товары
3. Следить за наличием товаров на складе
4. ...

Определение сущностей и связей

После того, как были определены действия, определяем возможные сущности.

Пример

Покупатель – явный кандидат на сущность

Товар – явный кандидат на сущность

Склад – возможный кандидат на сущность

Определяем связи между сущностями.

Пример

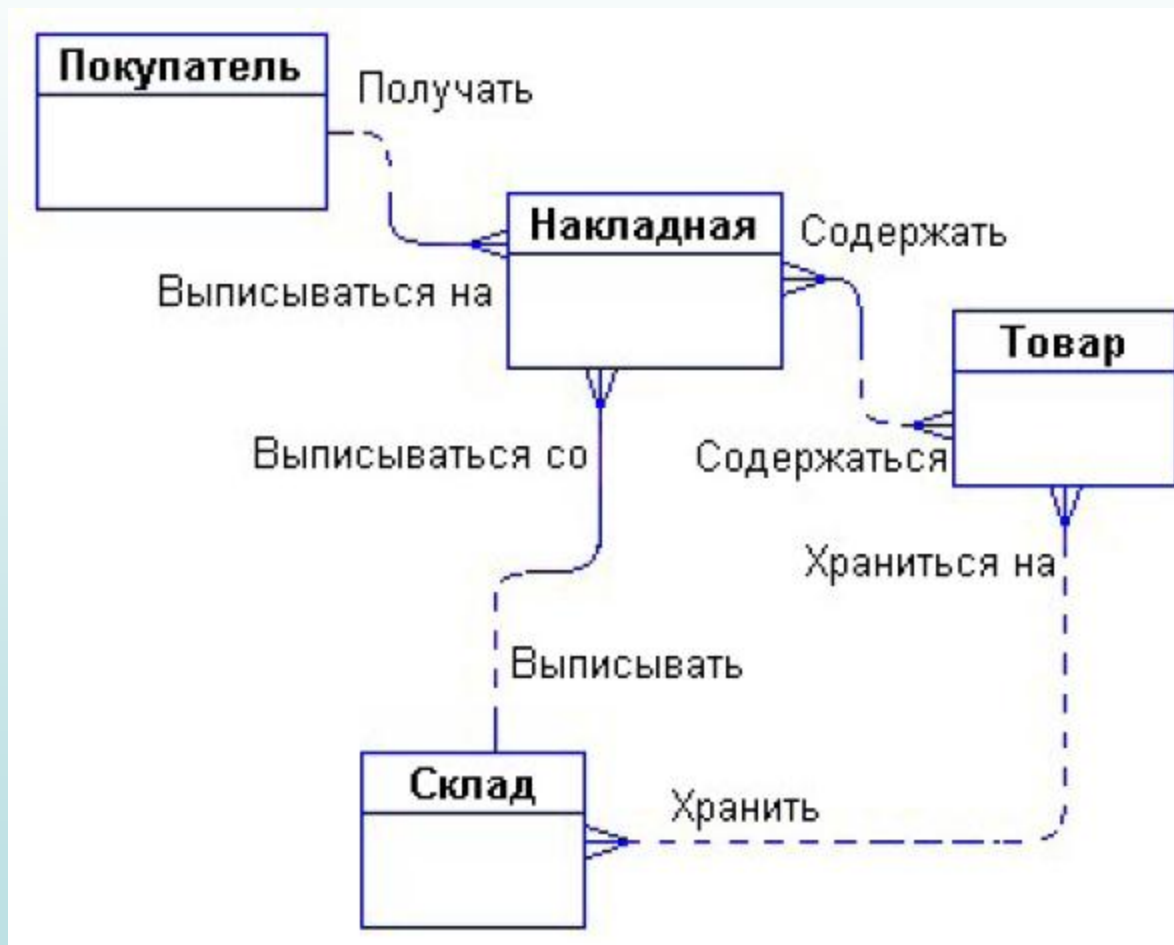
Покупатели могут покупать много товаров

Товары могут продаваться многим покупателям



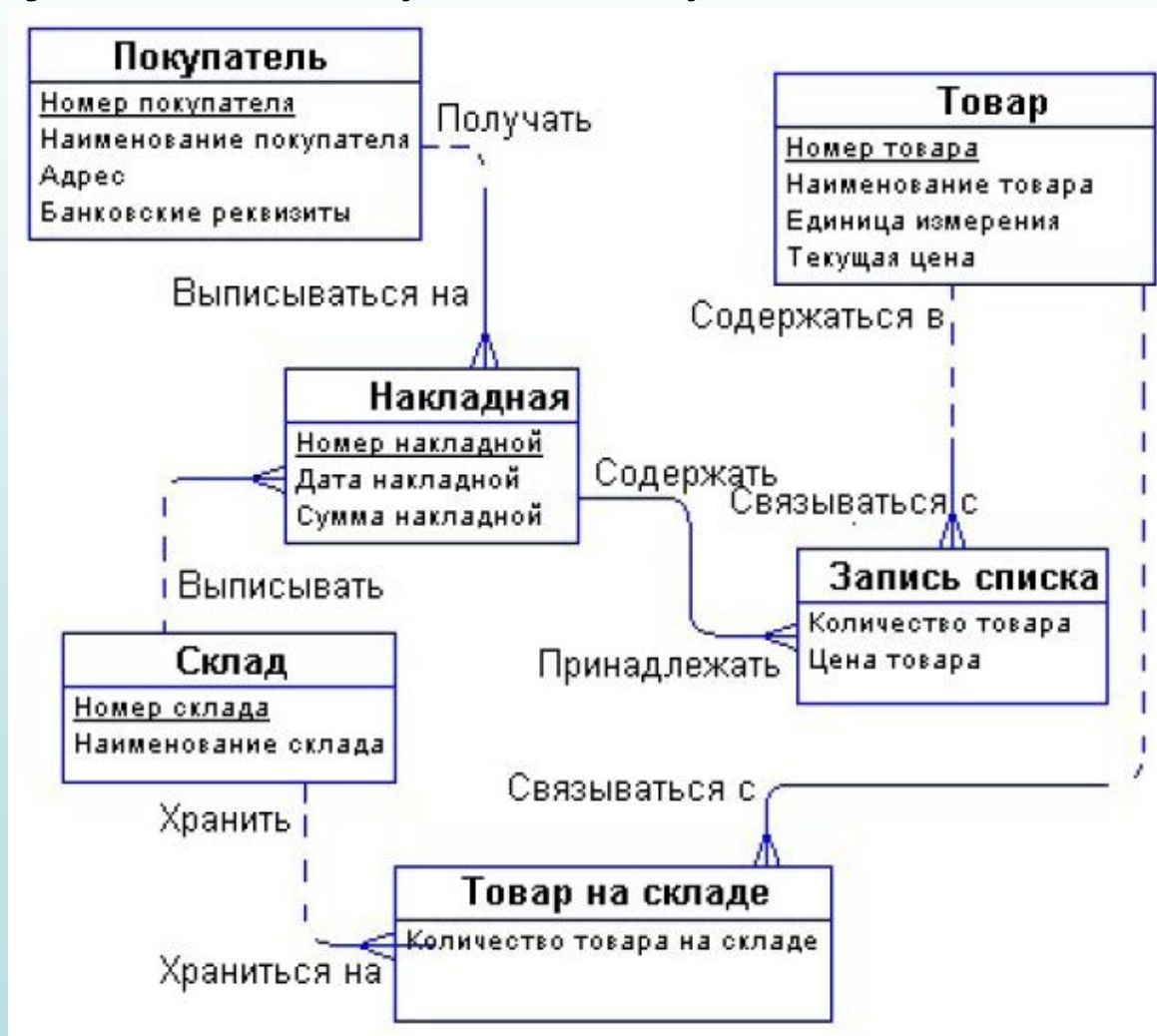
ER-модель без атрибутов

После того, как были определены сущности и связи получается следующая структура.



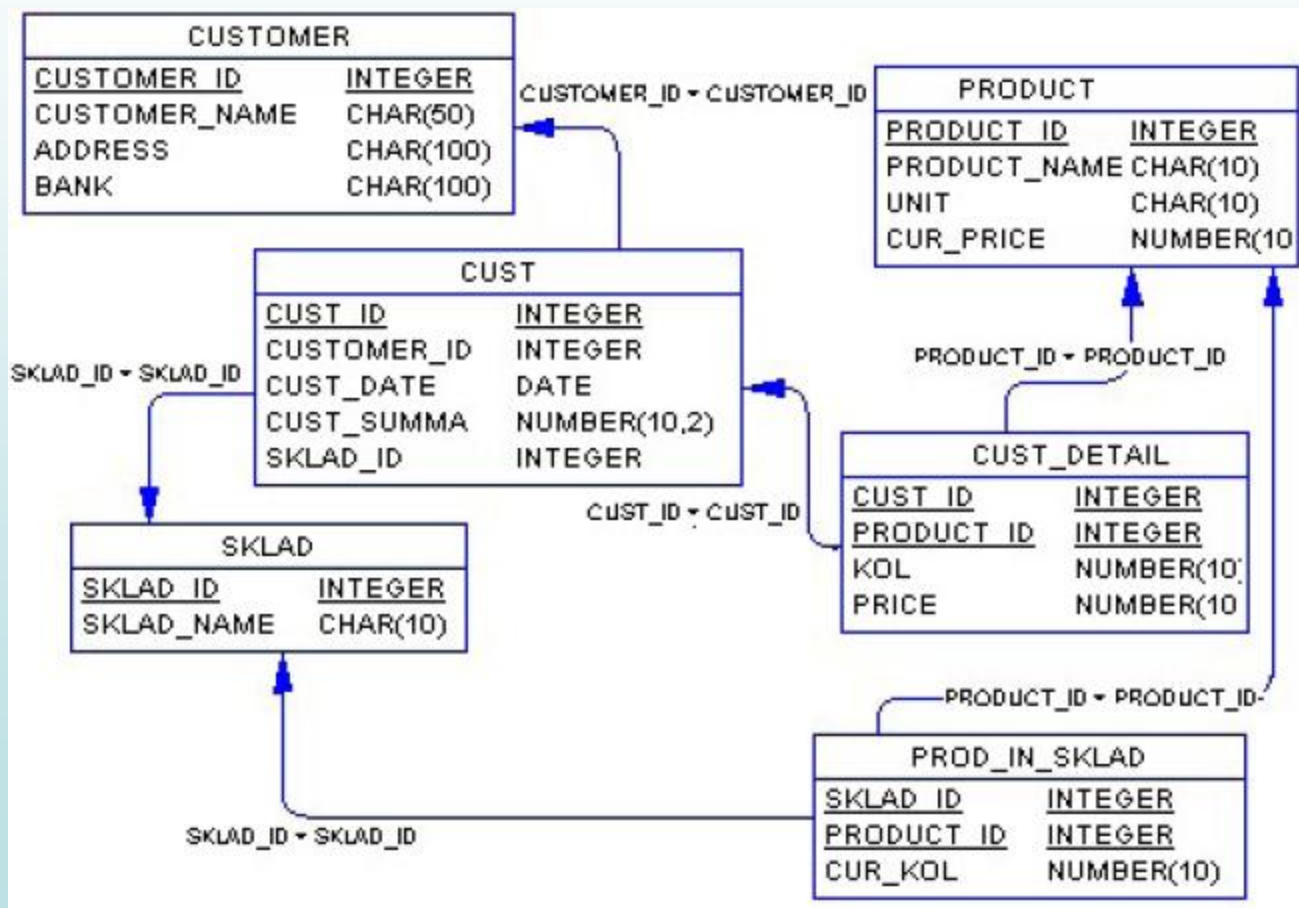
Концептуальная модель

Анализируя предметную область определяем атрибуты сущностей. Вносим их в модель и получаем **концептуальную диаграмму** – модель, не учитывающую особенности конкретной СУБД.



Физическая модель

После того, как реализована концептуальная модель, разрабатываем **физическую диаграмму** – учитывающую особенности конкретной СУБД.



03. Нормализация данных

Нормализация

Нормализация – процесс позволяющий распределить атрибуты по таблицам, исключая их избыточность.

Подход с нормальными формами был разработан Эдгаром Коддом в 1970-х гг.

Существуют следующие стандартные нормальные формы:

- 1-я нормальная форма
- 2-я нормальная форма
- 3-я нормальная форма
- нормальная форма Бойса-Кодда
- 4-я нормальная форма
- 5-я нормальная форма

Нормализация желательна, однако нужно понимать, чем выше степень нормализации, тем больше получается таблиц, что усложняет схему и снижает производительность БД.

1НФ

Первая нормальная форма

Таблица находится в 1НФ, если каждый ее атрибут атомарен и все строки различны.

Таблица не отвечает требованиям 1НФ

Сотрудник	Номер телефона
Иванов И. И.	283-56-82 390-57-34
Петров П. П.	708-62-34

Приведение к 1НФ

Сотрудник	Номер телефона
Иванов И. И.	283-56-82
Иванов И. И.	390-57-34
Петров П. П.	708-62-34

2НФ

Вторая нормальная форма

Таблица находится в 2НФ, если она находится в 1НФ и все ее атрибуты, помимо входящих в первичный ключ, зависят только от этого ключа.

Таблица не отвечает требованиям 2НФ

<u>Сотрудник</u>	<u>Должность</u>	Зарплата	Компьютер
Гришин	Кладовщик	20000	Нет
Васильев	Программист	40000	Есть
Иванов	Кладовщик	25000	Нет

Приведение к 2НФ

<u>Сотрудник</u>	<u>Должность</u>	Зарплата
Гришин	Кладовщик	20000
Васильев	Программист	40000
Иванов	Кладовщик	25000

<u>Должность</u>	Компьютер
Кладовщик	Нет
Программист	Есть

3НФ

Третья нормальная форма

Таблица находится в 3НФ, если она находится в 2НФ и при этом любой ее неключевой атрибут функционально зависит только от первичного ключа.

Таблица не отвечает требованиям 3НФ

<u>Сотрудник</u>	Отдел	Телефон
Гришин	Бухгалтерия	11-22-33
Васильев	Бухгалтерия	11-22-33
Петров	Снабжение	44-55-66

Приведение к 3НФ

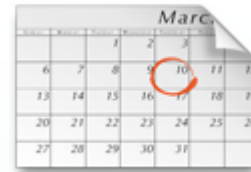
<u>Отдел</u>	Телефон
Бухгалтерия	11-22-33
Снабжение	44-55-66

<u>Сотрудник</u>	Отдел
Гришин	Бухгалтерия
Васильев	Бухгалтерия
Петров	Снабжение

Вопросы



Домашнее задание



- Изучить главу 19 книги М.Грабера Введение в SQL
- Изучить лекцию «Элементы модели Сущность-связь»
- Подготовиться к контрольной работе по лекции
- Выполнить лабораторную работу №2 и №3