

Базы данных и СУБД

Михайлова Елена Георгиевна,
кафедра ИАС, доцент

Задачи курса

- Введение в теорию баз данных
- Практические навыки работы с СУБД на базе MS SQL Server
- Лекции + практика = зачет + экзамен

План курса

- Основные определения и понятия
- Реляционная модель
- Реляционная алгебра
- Индексы
- Процедуры/функции
- Триггеры
- Транзакции
- Оптимизация запросов
- Распределенные БД
- NoSQL базы данных

Рекомендуемая литература по курсу

- Гарсиа-Молина, Ульман, Уидом. Системы баз данных. Полный курс. "Вильямс", 2003.
- Кристофер Дейт — «Введение в системы баз данных»
- С. Кузнецов. Основы баз данных.
- Б.А. Новиков, Г.Р. Домбровская: Настройка приложений баз данных. БХВ: СПб. 2006.
- Джеффри Ульман, Дженнифер Уидом. Введение в системы баз данных, Лори, 2006 г.
- Joe Celko's. Thinking in sets. Auxiliary, Temporal, and Virtual Tables in SQL
- Joe Celko's. SQL for smarties: advanced sql programming
- msdn.microsoft.com

MS SQL Server-2012-14-16

- <https://dspark.spbu.ru/>
- Зная свои Log –Pass (st00XXXX)
- Устанавливаем с сайта СПбГУ Visual Studio
- Устанавливаем с сайта СПбГУ SQL Server

https://dspark.spbu.ru/



Санкт-Петербургский государственный
университет

Программа Microsoft DreamSpark

Краткая инструкция по работе с системой

Шаг 0. Убедитесь, что вы являетесь студентом или сотрудником СПбГУ и имеете на руках параметры своей единой учетной записи. Подробная информация о единой учетной записи находится на сайте [УСИТ СПбГУ](#). Никаких дополнительных логинов и паролей для доступа к DreamSpark не требуется, и запрашивать их не нужно!

Шаг 1. Выберите подразделение, в котором вы числитесь. Вы будете перенаправлены на соответствующий интернет-магазин.

Шаг 2. На сайте интернет-магазина в правом верхнем углу перейдите по ссылке "Вход", а потом "Единый вход". Остальные разделы нужно проигнорировать.

Шаг 3. Вам будет предложено представиться системе. В качестве имени пользователя нужно ввести единую учетную запись.

Шаг 4. В случае успешной авторизации вы будете перенаправлены на сайт интернет-магазина, где можно будет совершать покупки.

Если возникают вопросы по работе с системой, вы можете написать на адрес support@spbu.ru. Пожалуйста, используйте почтовые адреса СПбГУ, а не @gmail.com или @mail.ru.

Выберите подразделение

Академическая гимназия
Биологический факультет
Восточный факультет

DreamSpark: Microsoft software for learning, teaching and research

St. Petersburg State University - Direction of Mathematics and Mechanics Faculty of Applied Math - DreamSpark Premium

Поиск продукта



DreamSpark Premium

Категории

- новое (2)
- Популярное (19)
- Операционные системы (21)
- Инструменты разработчика (62)
- Серверы (69)
- Приложения (65)
- Другое (1)
- Все (200)

Microsoft Corporation



Microsoft Windows 10
Technical Preview



Microsoft Windows
8.1



Microsoft Windows 8



Microsoft Windows 7



Microsoft Visual
Studio 2013



Microsoft Visual
Studio 2012



Microsoft Project
2013



Microsoft Visio 2013



Microsoft Visual
Studio 2010



Microsoft Project
2010



Microsoft Visio 2010



Microsoft SQL Server
2012

Приступайте к работе
прямо сейчас

Есть учетная запись?
Нажмите кнопку "Войти" выше.

Нет учетной записи?
[Запросите учетную запись](#)



Web Site



Travel reservations



Databases



E-Commerce



Social Networking



Information System



Online banking



iTunes

Accounting Information Systems



Customer Relationship Management



Payroll



Вы пользуетесь информационной системой, когда:

- Используете ВКонтакте или Facebook
- Делаете звонок с мобильного телефона
- Совершаете online покупки
- Бронируете авиабилеты
- Резервируете номер в гостинице
- Расплачиваетесь кредитной картой
- Используете банкомат
- Ищете что-то в поисковой системе
- ...

Что такое база данных?

- Набор порций информации, существующий в течение длительного времени



Причины появления БД

- Сверхмощные вычислительные системы
- ПК высокой производительности
- Рост объемов носителей информации (до 60х гг – последовательные)
- Особенности устройств прямого доступа:
 - Выборочное обновление
 - Сложные структуры данных
- Мировая сеть Internet
- Разнообразиие языков программирования
- Методы анализа данных, основанные на достижениях теории искусственного интеллекта

Откуда берутся данные?

- Google увеличивает объем на 20 PB в день (2008)
- Facebook - 2.5 PB данных + 15 TB/в день (2009)
- eBay -s 6.5 PB данных + 50 TB/в день (2009)



640К должно хватить
каждому



1 NEW DEFINITION IS ADDED ON **urban**

1,600+ READS ON **Scribd**

13,000+ HOURS **MUSIC** STREAMING ON **PANDORA**

12,000+ NEW ADS POSTED ON **craigslist**

370,000+ MINUTES VOICE CALLS ON **skype**

98,000+ **TWEETS**



320+ NEW **twitter** ACCOUNTS

100+ NEW **LinkedIn** ACCOUNTS

1 associatedcontent **NEW** ARTICLE IS PUBLISHED

THE WORLD'S LARGEST COMMUNITY CREATED CONTENT!!

20,000+ NEW POSTS ON **tumblr**

13,000+ **iPhone** APPLICATIONS DOWNLOADED



QUESTIONS ASKED ON THE INTERNET...

100+ Answers.com

40+ YAHOO! ANSWERS



600+ NEW VIDEOS

25+ HOURS TOTAL DURATION

70+ DOMAINS REGISTERED

60+ NEW BLOGS

1,500+ BLOG POSTS

168 MILLION EMAILS ARE SENT

694,445 SEARCH QUERIES

1,700+ **Firefox** DOWNLOADS

695,000+ **facebook** STATUS UPDATES

50+ **WORDPRESS** DOWNLOADS



125+ **PLUGIN** DOWNLOADS

79,364 **WALL** POSTS

510,040 **COMMENTS**



Google Search



СУБД

система управления базой данных

- Высокая стоимость данных в важнейших областях применения
- Дублирование средств управления данными в каждом приложении
- Высокая сложность создания приложений
- Чрезмерная сложность организации доступа к общим (разделяемым между приложениями) данным

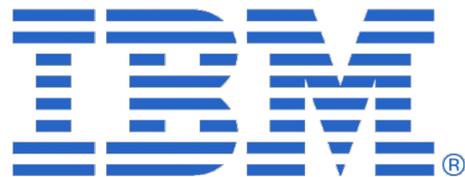
Лидеры на рынке СУБД



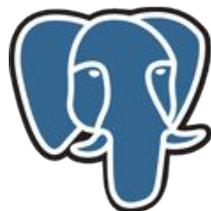
ORACLE



Microsoft®
SQL Server®



IBM®



PostgreSQL



MySQL®

Функции СУБД

- Средства постоянного хранения данных
- Безопасность данных и защита от несанкционированного доступа
- Защита целостности
- Согласованность данных
- Независимость данных и программ

Возможности СУБД

- Создание БД и описание схемы
- Сложные структуры данных
- Высокоуровневые языки запросов
- Масштабирование
- Параллельный/конкурентный доступ к общим данным

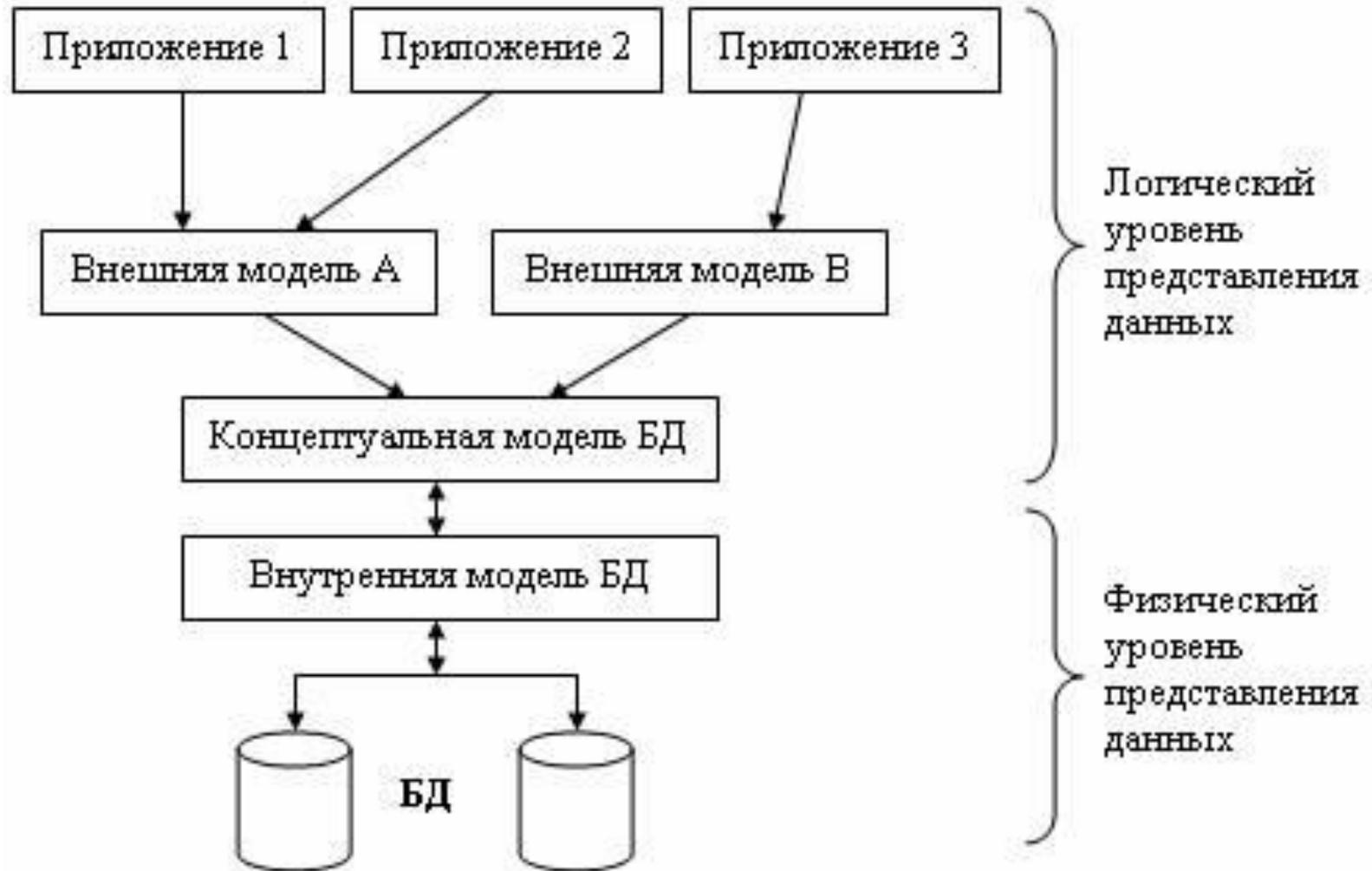
Проектирование БД

- Любая база данных является составной частью некой информационной системы
- Необходимо не только хранение данных, но и их обработка.
- Предметная область - часть реального мира, подлежащая изучению с целью автоматизации
 - *объекты,*
 - *процессы,*
 - *пользователи.*

Уровни абстракции

- Внешняя схема (описывает структуры данных, предоставляемые приложению, опр. совокупностью требований к данным со стороны некоторой конкретной функции)
- Концептуальный уровень (описывает полную логическую структуру хранимых данных)
- Схема хранения (описывает представление концептуальной схемы на носителях данных)

Уровни абстракции



Схема

- Перечень всех объектов называется схемой.
- Конкретное содержание – экземпляр.

Физическая и логическая независимость данных

- Если изменяется физическое место или структура хранения данных, это не отражается на пользователе и его программах
- Если изменяется подсхема для одного пользователя, то это не должно отразиться на других
- Если меняется все концептуальная схема, это не должно отразиться на пользователях, чьи данные изменения не касаются

Диаграммы «сущность-связь» (Entity-relationship):

База данных может быть представлена как совокупность объектов и связей между ними

- Сущность (прямоугольники)
- Атрибуты (овалы)
- Связи (ромбы)

Entity Relationship Diagram (ERD)

- Модель ER позволяет эскизам базе
- Графический инструмент для моделирования данных
- Широко используется в проектировании баз данных
- Графическое представление логической структуры базы данных
- Модель, которая идентифицирует понятия или объекты, которые существуют в системе, и связи между этими сущностями

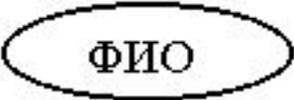
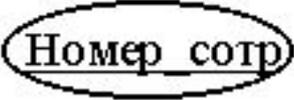
Назначение ERD

- Проектировщик/аналитик базы данных получает лучшее понимание информации, содержащейся в базе данных через процесс построения диаграмм.
- ERD служат инструментом документации.
- ERD используется для передачи логической структуры базы данных для пользователей.

Компоненты ERD

1. Entity Сущность
2. Relationship Связь
3. Cardinality Кардинальность
4. Attribute Атрибут

Базовые понятия (нотация Чена)

Сущность (объект)	
Атрибут сущности (свойство, характеризующее объект)	
Ключевой атрибут (атрибут, входящий в первичный ключ)	
СВЯЗЬ	

Модель «сущность-связь»

сущность

- Сущность, это единичное, обладающее самостоятельностью, в отличие от его состояний и отношений, которые являются изменчивыми и зависят от времени, места, от связей с другими сущностями

Аристотель

Модель «сущность-связь»

сущность

- Сущность – это абстрактный объект определенного вида. Любой предмет или понятие, информацию о которых мы будем хранить.
- Набор однородных сущностей образует множество сущностей.
- ER-модель имеет дело с со структурами данных, но НЕ С ОПЕРАЦИЯМИ над данными.

Модель «сущность-связь»

сущность

- Сущности должны быть отличимы от других сущностей – нужны идентификаторы.
- Кроме идентификаторов есть другие свойства - атрибуты.
- Назовем сущности с одинаковым набором атрибутов множеством сущностей (аналог класса).
- Множества не содержат дубликатов

Модель «сущность-связь» атрибуты

- Отдельные характеристики объекта называются атрибутами
- Каждый атрибут имеет уникальное имя.
- Каждый атрибута имеет свой тип данных.
- Объект может обладать любым количеством атрибутов.
- Значение атрибута атомарно.

Представление сущностей



Атрибуты

- Атрибут - характеристика объекта. Не может быть множеством. Каждый атрибут имеет определенный тип данных.
- Объект и его атрибуты соединяются ненаправленными дугами.

Идентификация сущностей

- Отличимость сущностей означает возможность идентификации
- Сущности идентифицируются выделенными атрибутами – ключами
- Искусственные (суррогатные) или естественные ключи?

3 вида идентификации

- Естественные ключи
- «По положению» (географическое, по порядку, во времени)
- Суррогатные

Суррогатные

- Необходимы системе
- Не связаны с реальным миром
- Например,
 - номер паспорт (для УФМС – суррогатный, для нас – естественный)
 - номер ИНН

Естественные

- «А ростом он мал, грудь широкая, одна рука короче другой, глаза голубые, волоса рыжие, на щеке бородавка, на лбу другая»
(Пушкин, Борис Годунов)
- Ни один естественный идентификатор не может быть абсолютно надежен

По положению

- Два тюбика зубной пасты – отличить невозможно
- При допущении дубликатов их невозможно пересчитать
- Если добавить количество, то это будет другой набор сущностей

Ключи

- Атрибуты бывают полезные для поиска и бесполезные
- Наиболее важные образуют ключ
- Ключ – это один ли несколько атрибутов объекта, по которому объект можно однозначно определить
- Если нет естественного ключа, придумывают искусственный – «суррогатный»

Модель «сущность-связь» СВЯЗИ

- Связи – это соединение между двумя или большим числом множества сущностей. Это отношения между объектами.
- Связь – упорядоченный набор сущностей
- У связей тоже есть идентификатор – совокупность идентификаторов сущностей, в них участвующих

Связи: количество «наборов» объектов

- бинарные
- тернарные
- n-арные
- рекурсивные

Как увидеть связь

- Если хочется типом какого-нибудь свойства сделать сущность или список сущностей
- Если хочется записать в одну сущность идентификатор другой
 - => Значит вам хочется сделать связь

Модель «сущность-связь»

СВЯЗИ

- Связи могут иметь собственные атрибуты
- Подобные связи объединяются в множества
- Отличие связей от сущностей: связи не могут существовать без связываемых сущностей
- Идентификация связей: ключ связи включает ключи связываемых сущностей и, возможно, выделенные атрибуты связи

Характеристики связи

Осознав факт наличия связи, полезно выяснить некоторые характеристики

- Название связи. Иногда удастся подобрать короткий и ясный термин, но часто название составляют из названий сущностей-участников
- Роли участников: “метки”, означающие смысл участия каждой сущности. Это часто бывают глаголы

Бинарные связи – кардинальность (степень связи)

Делятся на три вида в зависимости от количества участвующих в них объектов.

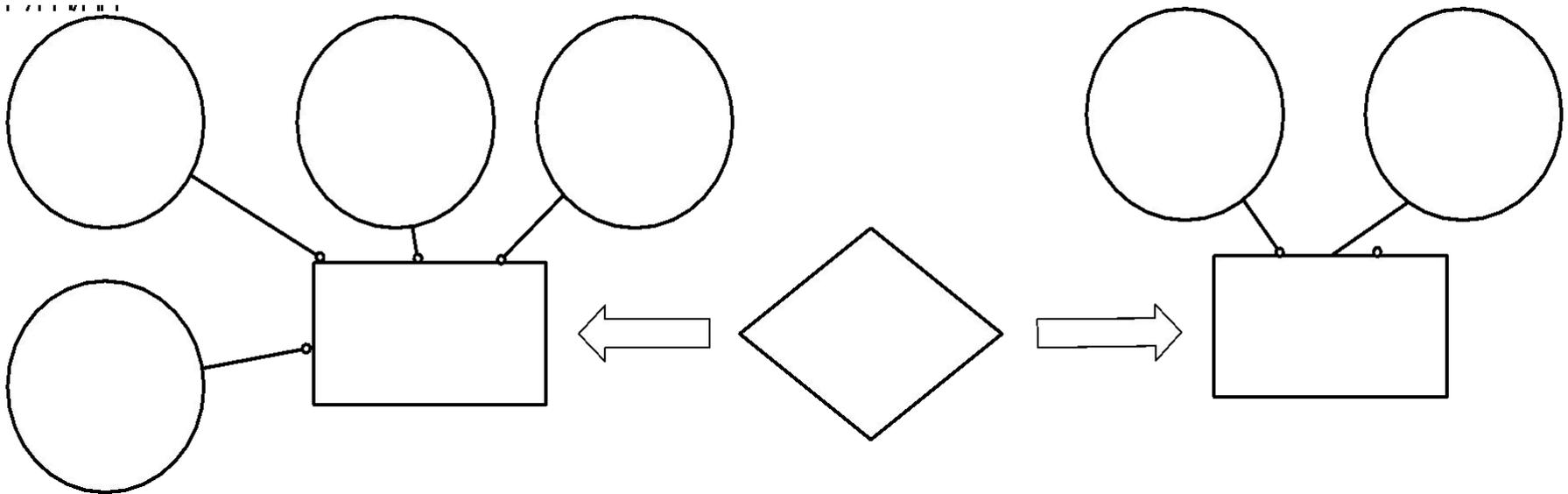
- Один к одному 1:1
- Один ко многим 1:N
- Многие ко многим M:N

Связь *один-к-одному*

- *Один - к одному*. Этот тип связи означает, что каждому объекту первого вида соответствует не более одного объекта второго вида, и наоборот.
- Например: сотрудник может руководить только одним отделом, и у каждого отдела есть только один руководитель.

Связь *один-к-одному*

Ромб связи и прямоугольник объекта соединяются ненаправленными дугами в сторону "ко многим" и направленными в сторону "к одному".



СВЯЗИ *ОДИН-КО-МНОГИМ*

- *Один - ко многим.* Этот тип связи означает, что каждому объекту первого вида может соответствовать более одного объекта второго вида, но каждому объекту второго вида соответствует не более одного объекта первого вида.
- Например: в каждом отделе может быть множество сотрудников, но каждый сотрудник работает только в одном отделе.

СВЯЗИ *ОДИН-КО-МНОГИМ*

Ромб связи и прямоугольник объекта соединяются ненаправленными дугами в сторону "ко многим" и направленными в сторону "к одному".



Для реализации связей вида «один ко многим» можно использовать отдельную таблицу, но чаще добавляют столбец с ключом «одного» в таблицу «КО МНОГИМ»

СВЯЗИ *МНОГИЕ-КО-МНОГИМ*

- *Многие - ко многим*. Этот тип связи означает, что каждому объекту первого вида может соответствовать более одного объекта второго вида, и наоборот.
- Например: каждый счет может включать множество товаров, и каждый товар может входить в разные счета.

СВЯЗИ *МНОГИЕ-КО-МНОГИМ*

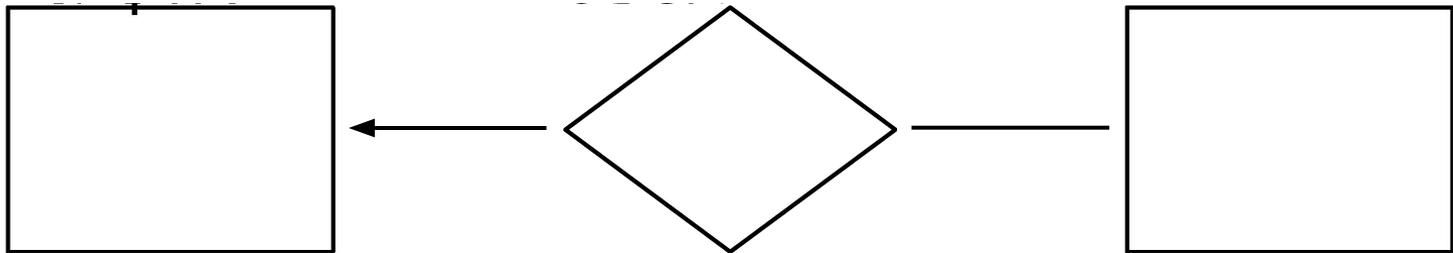
Ромб связи и прямоугольник объекта соединяются ненаправленными дугами в сторону "ко многим" и направленными в сторону "к одному".



Для реализации связей вида «многие ко многим» нужно использовать отдельную таблицу

Модальность

- Между двумя сущностями может быть определено несколько наборов связей:



Классификация связей

- **Необязательные (условные)**
 - Сотрудник может работать в отделе
 - Человек может иметь автомобиль
- **Обязательные**
 - Каждый курс лекций должен иметь преподавателя
 - У каждой кафедры должен быть заведующий

Бинарные связи – модальность (обязательность связи)

Обязаны ли экземпляры сущности участвовать в связи?

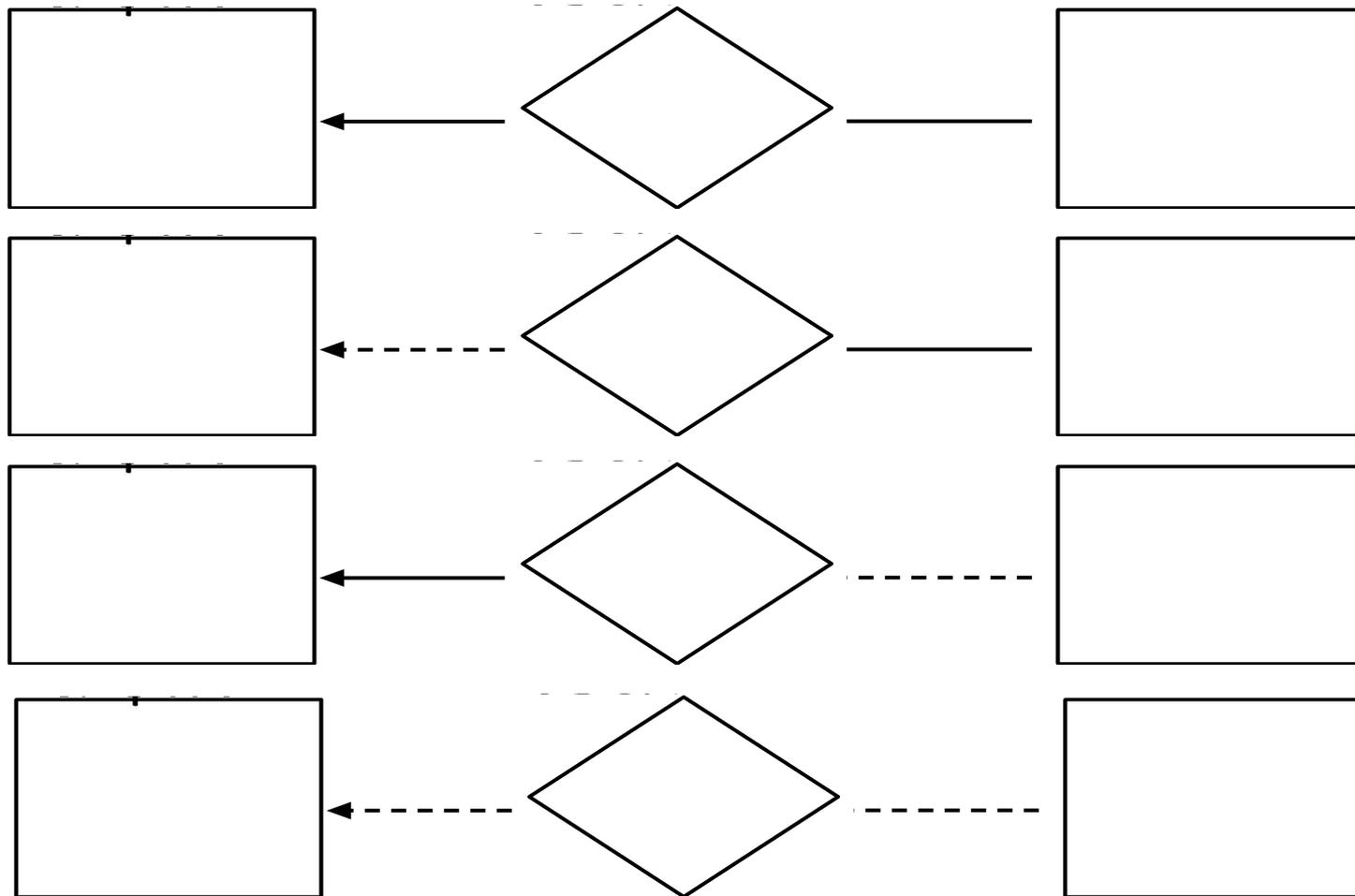
Не модальный тип -> не обязаны

Модальный тип -> обязаны

1:N

- (0,1) – (0,N)
- (0,1) – (1,N)
- (1,1) – (0,N)
- (1,1) – (1,N)

Модальность

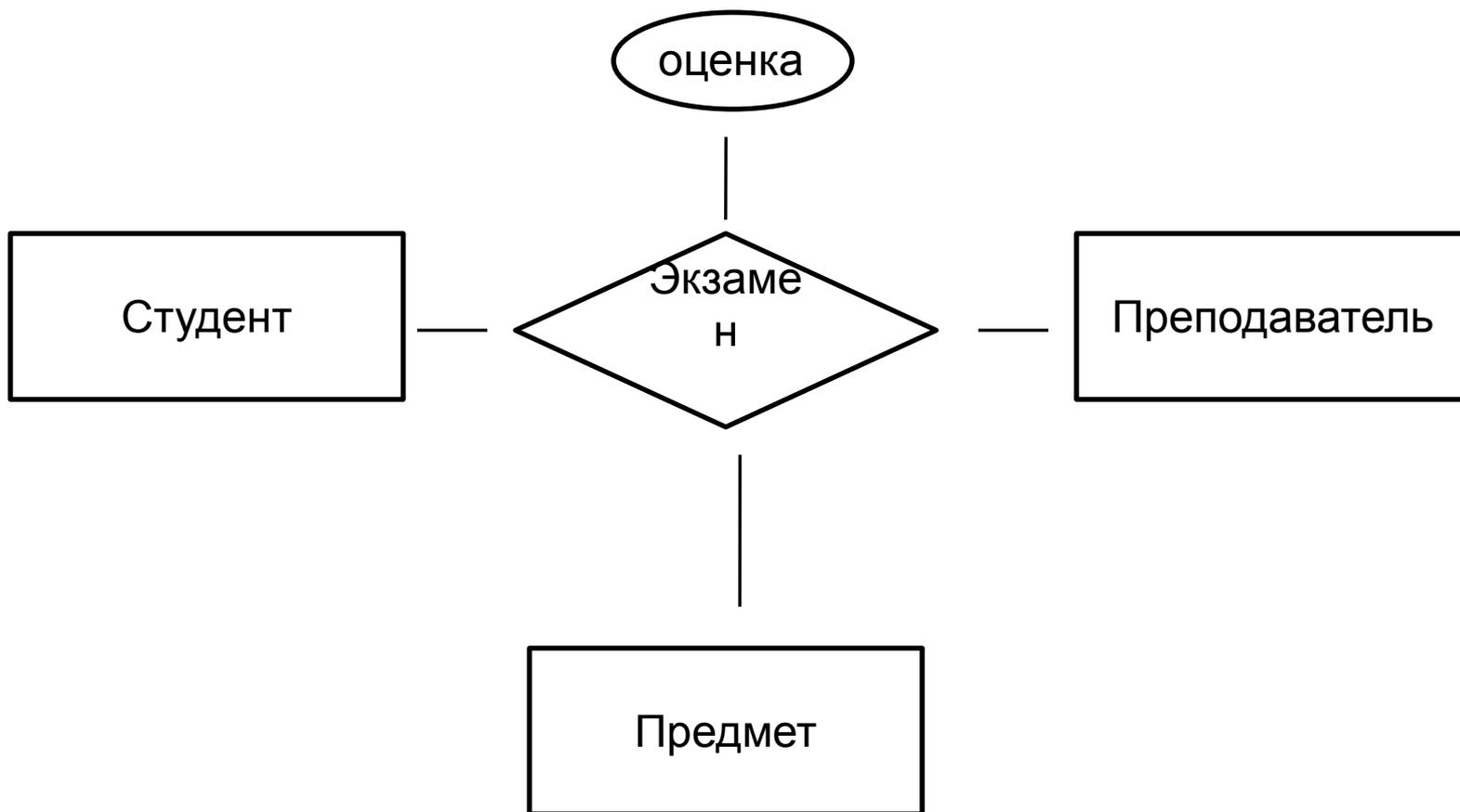


Примеры бинарных связей

- Между двумя сущностями может быть определено несколько наборов связей:



Многосторонние связи



Многосторонние связи

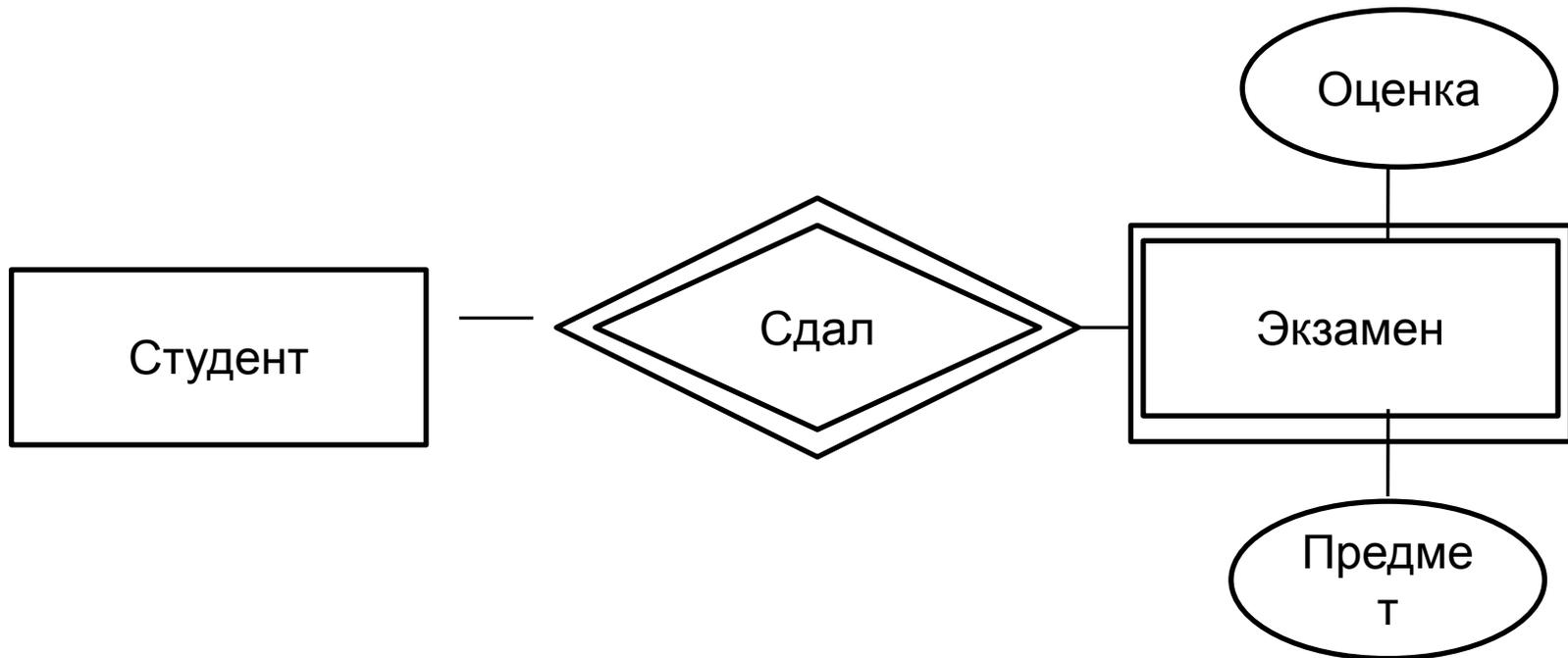
- Грузы
- Грузоотправитель
- Грузополучатель
- Перевозчик
- Владелец

Рекурсивная связь



Слабые сущности

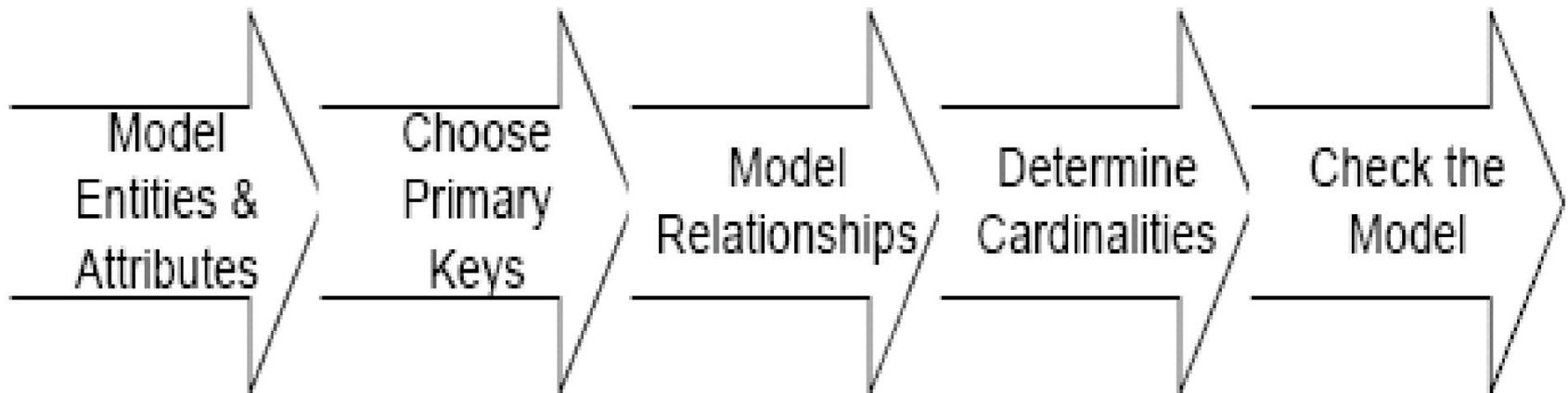
- Не могут существовать в БД без связи с другими (сильными) сущностями
- Не нужен суррогатный ключ



Шаги при создании ERD

- Определить сущности
- Определить атрибуты сущностей
- Определить первичный ключ
- Определить отношения между сущностями
- Определить кардинальность
- Нарисовать ERD
- Проверить ERD

Steps in building an ERD



Проектирование в модели сущность-связь

- Выбор между атрибутами с слабыми сущностями (адрес, номер телефона, ...)
- Выбор между сущностью и связью
- Бинарные или более сложные связи
- Ограничения целостности

Реляционная модель

- Похожа на таблицы
- Столбцы – атрибуты
- Строки – данные
- Шапка таблицы – имена атрибутов

Пример

