



# РЕЛЯЦИОННАЯ АЛГЕБРА И НОРМАЛИЗАЦИЯ БАЗ ДАнных

ОСНОВЫ ПРОЕКТИРОВАНИЯ БАЗ ДАННЫХ

# АНОМАЛИИ И ИЗБЫТОЧНОЕ ДУБЛИРОВАНИЕ ДАННЫХ

- Наличие первого из них допускается в БД, а избыточное дублирование данных может приводить к проблемам при их об обработке.
- Приведём пример обоих вариантов **дублирования**.
- Пример не избыточного дублирования данных представляет отношение СГ (студент-группа) с атрибутами Студент и Группа:
- Для студентов, обучающихся в одной группе, номер группы **совпадает**. Для каждого студента номер группы - **уникальный атрибут**. Поэтому ни один из номеров групп не является избыточным.

Студент	Группа
Иванов	307
Петров	307
Сидоров	400

# АНОМАЛИИ И ИЗБЫТОЧНОЕ ДУБЛИРОВАНИЕ ДАННЫХ

- Пример **избыточного дублирования** (избыточности) представляет отношение СГК (студент-группа-куратор), которые, в отличии от отношения СГ, дополнено атрибутом Куратор.
- Вполне естественно, что студенты одной группы находятся под руководством одного и того же куратора. **В рассматриваемом отношении имеется избыточное дублирование данных:**

Студент	Группа	Куратор
Иванов	307	Епишина
Петров	307	Епишина
Сидоров	400	Прошкин

# АНОМАЛИИ И ИЗБЫТОЧНОЕ ДУБЛИРОВАНИЕ ДАННЫХ

- Возможный способ данных из этой ситуации в создании двух отношений СГ и ГК, полученных путём декомпозиции исходного отношения СГК:

Студент	Группа
Иванов	307
Петров	307
Сидоров	400

Группа	Куратор
307	Епишина
400	Прошкин

# ИСКЛЮЧЕНИЕ ИЗБЫТОЧНОСТИ ДУБЛИРОВАНИЯ

- Процедура декомпозиции отношения СГК на 2 отношения СГ и ГК является основной **процедурой нормализации отношений**.
- Избыточное дублирование создает проблемы при обработке кортежей отношениями, названной Эдгаром Коддом «Аномалии обновления отношений».
- **Аномалиями** будем называть такую ситуацию в таблица БД, которая приводит к противоречиям в БД, либо существенно усложняет обработку данных.
- **Выделяют 3 основных вида аномалий:**
  - Аномалии модификаций (редактирование)
  - Аномалии удаления
  - Аномалии добавления
- **Аномалии модификации** проявляются в том, что изменение значения одного данного может повлечь за собой просмотр всей таблицы и соответствующее изменение некоторых других записей таблицы.
- **Аномалии удаления** состоят в том, что при удалении какого-либо данного из таблицы может пропасть и другая информация, которая не связана напрямую с удаляемым данным.
- **Аномалии добавления** возникают в случаях, когда информацию в таблицу нельзя поместить до тех пор, пока она неполная, либо вставка новой записи требует дополнительного просмотра таблицы.

# НОРМАЛИЗАЦИЯ БАЗЫ ДАННЫХ

- **Первая нормальная форма (1НФ)**

- Отношение находится в 1НФ, если все его атрибуты являются простыми, все используемые домены должны содержать только скалярные значения.
- Не должно быть повторений строк в таблице.
- Например, есть таблица «Автомобили»:

Фирма	Модели
BMW	M5, X5M, M1
Nissan	GT-R

# НОРМАЛИЗАЦИЯ БАЗЫ ДАННЫХ

- Нарушение нормализации 1 НФ происходит в моделях BMW, т.к. в одной ячейке содержится список из 3 элементов: M5, X5M, M1, т.е. он не является атомарным.
- Преобразуем таблицу к 1НФ:

Фирма	Модели
BMW	M5
BMW	X5M
BMW	M1
Nissan	GT-R

# НОРМАЛИЗАЦИЯ БАЗЫ ДАННЫХ

- **Вторая нормальная форма (2НФ)**

- Отношение находится во 2НФ, если оно находится в 1НФ и каждый не ключевой атрибут неприводимо зависит от Первичного Ключа(ПК).

- **Неприводимость** означает, что в составе потенциального ключа отсутствует меньшее подмножество атрибутов, от которого можно также вывести данную функциональную зависимость.

- Например, дана таблица:

<u>Модель</u>	<u>Фирма</u>	Цена	Скидка
M5	BMW	5500000	5%
X5M	BMW	6000000	5%
M1	BMW	2500000	5%
GT-R	Nissan	5000000	10%

# НОРМАЛИЗАЦИЯ БАЗЫ ДАННЫХ

- Таблица находится в первой нормальной форме, но не во второй.
- Цена машины зависит от модели и фирмы.
- Скидка зависит от фирмы, то есть зависимость от первичного ключа **неполная**.  
Исправляется это путем декомпозиции на два отношения, в которых не ключ

<u>Модель</u>	<u>Фирма</u>	Цена
M5	BMW	5500000
X5M	BMW	6000000
M1	BMW	2500000
GT-R	Nissan	5000000

<u>Фирма</u>	Скидка
BMW	5%
Nissan	10%

# НОРМАЛИЗАЦИЯ БАЗЫ ДАННЫХ

- **Третья нормальная форма (3НФ)**

- Отношение находится в 3НФ, когда находится во 2НФ и каждый не ключевой атрибут нетранзитивно зависит от первичного ключа.
- Проще говоря, второе правило требует выносить все не ключевые поля, содержимое которых может относиться к нескольким записям таблицы в отдельные таблицы.
- Рассмотрим таблицу.

<u>Модель</u>	Магазин	Телефон
BMW	Риал-авто	87-33-98
Audi	Риал-авто	87-33-98
Nissan	Некст-Авто	94-54-12

# НОРМАЛИЗАЦИЯ БАЗЫ ДАННЫХ

- Таблица находится во 2НФ, но не в 3НФ.
- В отношении атрибут «Модель» является первичным ключом. Личных телефонов у автомобилей нет, и телефон зависит исключительно от магазина.
- Таким образом, в отношении существуют следующие функциональные зависимости: Модель → Магазин, Магазин → Телефон, Модель → Телефон.
- Зависимость Модель → Телефон является транзитивной, следовательно, отношение не находится в 3НФ.
- **В результате разделения исходного отношения получаются два отношения, находящиеся в 3НФ:**
  - Риал-авто 87-33-98
  - Риал-авто 87-33-98
  - Некст-Авто 94-54-12

<u>Модель</u>	Магазин
BMW	Риал-авто
Audi	Риал-авто
Nissan	Некст-Авто

# ВИДЫ СВЯЗЕЙ МЕЖДУ ТАБЛИЦАМИ

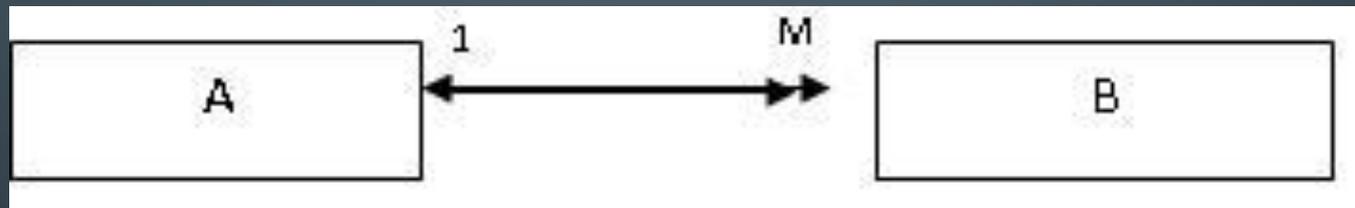
- Связь позволяет моделировать отношения между объектами предметной области. Наименование связи должно быть уникально во всей модели.
- **Существует 4 типа связей:**
- 1. **«Один-к-одному»** - любому экземпляру сущности А соответствует только один экземпляр сущности В, и наоборот.



- У любого конкретного ученика может быть только одна характеристика, и эта характеристика относится к единственному ученику.

# ВИДЫ СВЯЗЕЙ МЕЖДУ ТАБЛИЦАМИ

- 2. **«Один-ко-многим»** - любому экземпляру сущности А соответствует 0, 1 или несколько экземпляров сущности В, но любому экземпляру сущности В соответствует только один экземпляр сущности А.



- Ученику ставят много оценок; поставленная оценка принадлежит только одному ученику.

## ВИДЫ СВЯЗЕЙ МЕЖДУ ТАБЛИЦАМИ

- 3. **«Многие-к-одному»** - любому экземпляру сущности А соответствует только один экземпляр сущности В, но любому экземпляру сущности В соответствует 0, 1 или несколько экземпляров сущности А.



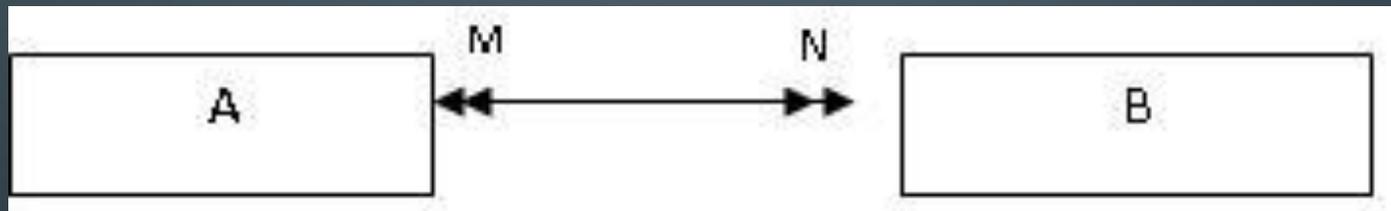
- Преподаватель работает только в одном кабинете, однако рабочий кабинет может быть закреплен за несколькими преподавателями.

# ВИДЫ СВЯЗЕЙ МЕЖДУ ТАБЛИЦАМИ

- Какая же разница между связями **«один-ко-многим»** и **«многие-к-одному»**?
- Такая же, как между фразами «портфель ученика» и «ученик портфеля».
- То есть важно, кто во взаимоотношении двух объектов главный - ученик или портфель.
- Суть отношений двух объектов отражается в **имени связи**.

## ВИДЫ СВЯЗЕЙ МЕЖДУ ТАБЛИЦАМИ

- 4. **«Многие-ко-многим»** - любому экземпляру сущности А соответствует 0, 1 или несколько экземпляров сущности В, и любому экземпляру сущности В соответствует 0, 1 или несколько экземпляров сущности А.



- Ученик Иванов учится у нескольких преподавателей. И каждый преподаватель работает со многими учениками.