

Проектирование реляционных баз данных

Цели и проблемы проектирования

Проектирование информационных систем, включающих базы данных,

осуществляется на физическом и логическом уровнях. Решение проблем проектирования на *физическом уровне* во многом зависит от используемой СУБД, зачастую автоматизировано и скрыто от пользователя.

Логическое проектирование

заключается в определении числа и структуры таблиц, формировании запросов к БД, определении типов отчетных документов, разработке алгоритмов обработки информации, создании форм для ввода и редактирования данных в базе и решении ряда других задач.

Классический подход при проектировании *структур данных*

Сбор информации об объектах решаемой задачи в рамках одной таблицы (одного отношения) и последующая декомпозиция ее на несколько взаимосвязанных таблиц на основе процедуры нормализации отношений.

Избыточное дублирование данных и аномалии

- Различают простое (неизбыточное) и избыточное дублирование данных. Наличие первого из них допускается в базах данных, а избыточное дублирование данных может приводить к проблемам при обработке данных.

Неизбыточное дублирование

С_Т

Сотрудник	телефон
Неделин	3721
Федосеев	4328
Старова	4328
Ермошина	4328

Избыточное дублирование

С_Т_К

сотрудник	телефон	№ кабинета
Неделин	3721	105
Федосеев	4328	111
Старова	4328	111
Ермошина	4328	111

Решение проблемы избыточности

T_K

Телефон	№_каб
3721	105
4328	111

C_K

Сотрудник	№_каб
Неделин	105
Федосеев	111
Старова	111
Ермошина	111

Вывод

- Процедура декомпозиции отношения S_{TK} на два отношения T_K и S_T является основной процедурой нормализации отношений.

Э. Кодд:

- Избыточное дублирование данных при обработке кортежей отношения создает проблемы, называемые «аномалиями обновления отношения». Эти проблемы возникают при попытке удаления, добавления или редактирования их кортежей.

Определение

- **Аномалиями** будем называть такую ситуацию в таблицах БД, которая приводит к противоречиям в БД, либо существенно усложняет обработку данных.

Три основные вида аномалий:

- аномалии модификации (или редактирования),
- 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.**

Формирование исходного отношения

- Проектирование БД начинается с определения всех объектов, сведения о которых будут включены в базу, и определения их атрибутов. Затем атрибуты сводятся в одну таблицу - исходное отношение.

Универсальное (исходное) отношение

- **Универсальным отношением** называется отношение, включающее все представляющие интерес атомарные атрибуты.

Пример

- Для учебной части факультета создается БД о преподавателях. определены содержащиеся в базе сведения о том, как она должна использоваться и какую информацию заказчик хочет получать в процессе ее эксплуатации. В результате устанавливаются атрибуты, которые должны содержаться в отношениях БД, и связи между ними.

•
Имена атрибутов и их краткие характеристики:

- **ФИО** - фамилия и инициалы преподавателя. Исключаем возможность совпадения фамилии и инициалов у преподавателей.
- **Должн** - должность, занимаемая преподавателем.
- **Оклад** - оклад преподавателя.
- **Стаж** - преподавательский стаж.
- **Д_Стаж** - надбавка за стаж.

Имена атрибутов и их краткие характеристики: (продолжение)

- **Каф** - номер кафедры, на которой числится преподаватель.
- **Предм** - название предмета (дисциплины), читаемого преподавателем.
- **Группа** - номер группы, в которой преподаватель проводит занятия.
- **ВидЗан** - вид занятий, проводимых преподавателем в учебной группе.

Исходное отношение

ПРЕПОДАВАТЕЛЬ

ПРЕПОДАВАТЕЛЬ

ФИО	Должн	Оклад	Стаж	Д_Стаж	Каф	Предм	Группа	ВидЗан
Ежова И.М.	преп	5000	5	100	25	БД	256	Практ
Ежова И.М.	преп	5000	5	100	25	БД	123	Лекция
Бобров М.И.	ст. преп	8000	7	120	25	БД	256	Лекция
Бобров М.И.	ст. преп	8000	7	120	25	Паскаль	256	Практ
Волков Н.Г.	преп	5000	10	150	25	БД	123	Практ
Волков Н.Г.	преп	5000	10	150	25	Паскаль	256	Лекция
Зайцев В. В.	преп	5000	5	100	24	ТСИ	244	Лекция


Этапы проектирования БД

- **Этап 1.** Обследование (анализ) предметной области.
- **Этап 2.** Выявление объектов, сведения о которых будут включаться в БД и определение перечня атрибутов. Формирование исходного (универсального) отношения.
- **Этап 3.** Построение инфологической модели, проектируемой БД на языке ER – диаграммы с учётом всех сущностей, атрибутов и связей.
- **Этап 4.** Формирование набора предварительных отношений

Этапы проектирования БД

(продолжение)

- **Этап 5.** Нормализация отношений.
- **Этап 6.** Внешнее кодирование . Оно заключается в замене длинных текстовых значений атрибутов короткими кодами.
- **Этап 7.** Пересмотр и редактирование ER – диаграммы.
- **Этап 8.** Построение схемы БД на языке «Таблицы – связи».
- **Этап 9.** Выбор СУБД для программой реализации.

- 
- Проектирование БД завершается проверкой корректности и полноты полученного проекта. Оно состоит в проверке возможности выполнения всех запросов пользователей к БД.

Зависимости между атрибутами

- Атрибут ***В функционально зависит*** от атрибута ***А***, если каждому значению ***А*** соответствует в точности одно значение ***В***.

$A \rightarrow B$

- Это означает, что во всех кортежах с одинаковым значением атрибута ***А*** атрибут ***В*** будет иметь также одно и то же значение.
- ***А*** и ***В*** могут быть составными - состоять из двух и более атрибутов.

ПРИМЕРЫ

ФИО-> Должн

Должн -> Оклад

Стаж-> Д_Стаж

- Наличие функциональной зависимости в отношении определяется природой вещей, информация о которых представлена кортежами отношения.

Функциональная взаимозависимость

- Если существует функциональная зависимость вида $A \rightarrow B$ и $B \rightarrow A$, то между A и B имеется взаимно однозначное соответствие, или функциональная взаимозависимость. Наличие функциональной взаимозависимости между атрибутами A и B обозначим как $A \leftrightarrow B$ или $B \leftrightarrow A$.

Частичная зависимость

- **Частичной зависимостью (частичной функциональной зависимостью)** называется зависимость неключевого атрибута от части составного ключа.

Частичные ФЗ

- ФИО-> Должн
- ФИО-> Оклад
- ФИО-> Стаж
- ФИО-> Д_Стаж
- ФИО-> Каф

Ключ – ФИО, Предм, Группа

ФИО – часть ключа

Полная зависимость

Полная функциональная зависимость – это зависимость неключевого атрибута от всего составного ключа.

В нашем примере:

Полная зависимость

ФИО, Предм, Группа -> ВидЗан

Транзитивные зависимости

Атрибут С зависит от атрибута А **транзитивно** (существует **транзитивная зависимость**), если для атрибутов А, В, С выполняются условия $A \rightarrow B$ и $B \rightarrow C$, но обратная зависимость отсутствует.

Примеры:

ФИО \rightarrow Должн \rightarrow Оклад

ФИО \rightarrow Стаж \rightarrow Д_Стаж

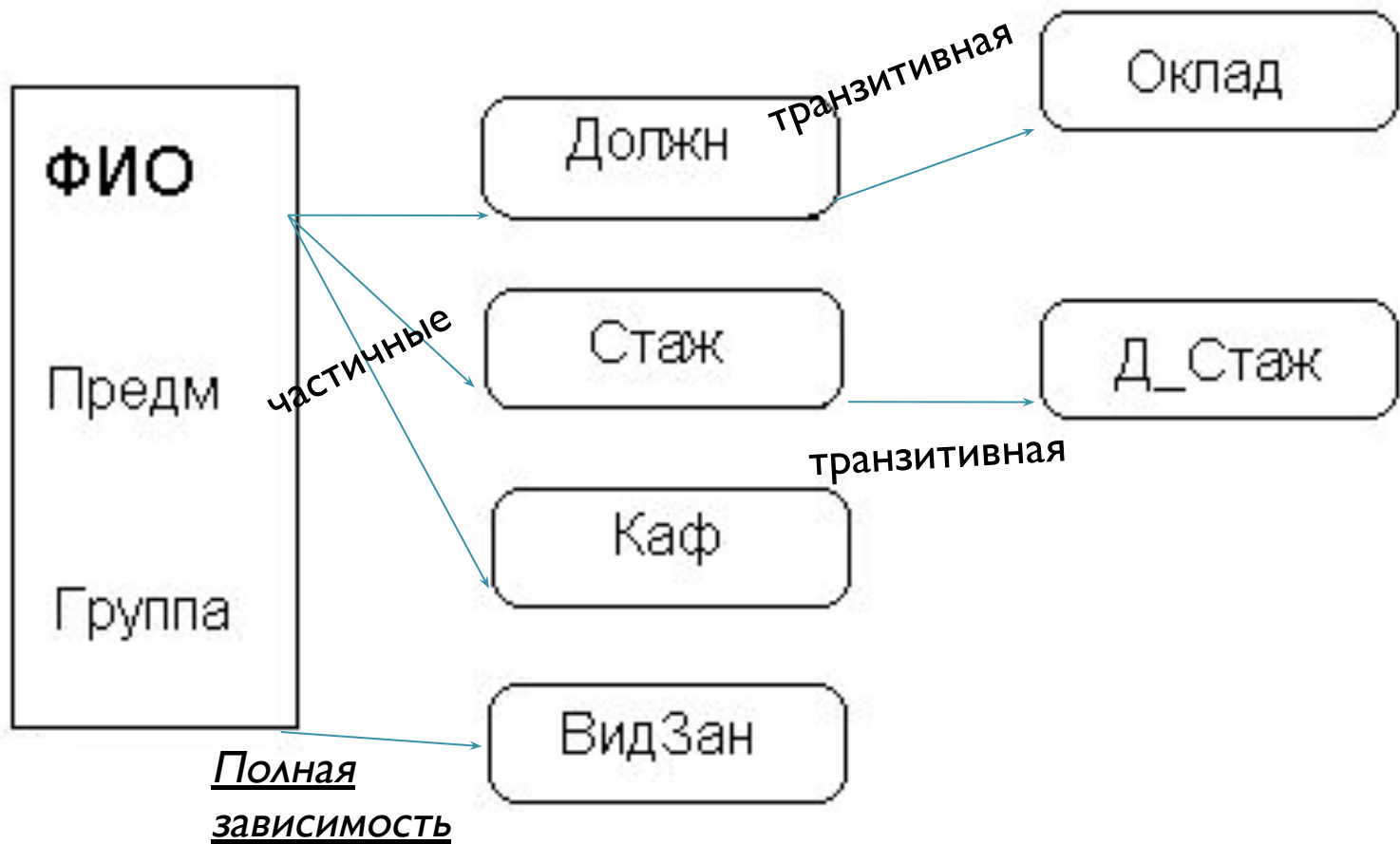
Многозначная зависимость

- В отношении R атрибут B **многозначно зависит** от атрибута A , если каждому значению A соответствует множество значений B , не связанных с другими атрибутами из R .
- Многозначные зависимости могут быть «один ко многим» ($1:M$), «многие к одному» ($M:1$) или «многие ко многим» ($M:M$), обозначаемые соответственно: $A \Rightarrow B$, $A \Leftarrow B$ и $A \Leftrightarrow B$.

Пример многозначной зависимости

- Пусть преподаватель ведет несколько предметов, а каждый предмет может вестись несколькими преподавателями, тогда имеет место зависимость ФИО \leftrightarrow Предмет.
- В нашем примере:
- преподаватель Бобров М. И. ведет занятия по двум предметам, а дисциплина БД - читается тремя преподавателями: Ежовой И. М., Бобровым М.И. и Волковым Н.Г.

Схема зависимостей



Нормальные формы

- Процесс проектирования БД с использованием **метода нормальных форм** заключается в последовательном переводе отношений из первой нормальной формы в нормальные формы более высокого порядка по определенным правилам. Каждая следующая нормальная форма ***устраняет соответствующие аномалии*** при выполнении операций над отношениями БД и ***сохраняет свойства предшествующих нормальных форм.***

Последовательность нормальных форм:

- первая нормальная форма (1НФ);
- вторая нормальная форма (2НФ);
- третья нормальная форма (3НФ);
- усиленная третья нормальная форма, или нормальная форма Бойса-Кодда (НФБК).

Первая нормальная форма

- Отношение находится в 1НФ, если все его атрибуты являются простыми (имеют единственное значение).

Исходное отношение строится таким образом, чтобы оно было в 1НФ.

- Перевод отношения в следующую нормальную форму осуществляется методом «декомпозиции без потерь».
- Основной операцией метода является операция проекции.

Частичная зависимость от ключа приводит к следующему:

1. В отношении присутствует явное и неявное избыточное дублирование данных
2. Избыточное дублирование данных порождает проблемы их редактирования.

Часть избыточности устраняется при переводе отношения в 2НФ.

Вторая нормальная форма

- Отношение находится в 2НФ, если оно находится в 1НФ и каждый неключевой атрибут функционально полно зависит от первичного ключа (составного).

Правила перевода отношения в 2НФ

Для устранения частичной зависимости необходимо, используя операцию проекции, разложить его на несколько отношений следующим образом:

1. построить проекции на части составного первичного ключа и атрибуты, зависящие от этих частей;
2. построить проекцию без атрибутов, находящихся в частичной зависимости от первичного ключа.

Сведения

2НФ

ФИО *	Должн	Оклад	Стаж	Д_Стаж	Каф
Ежова И.М.	преп	5000	5	100	25
Бобров М.И.	ст. преп	8000	7	120	25
Волков Н.Г.	преп	5000	10	150	25
Зайцев В. В.	преп	5000	5	100	24

Нагрузка

ФИО *	Предм*	Группа*	ВидЗан
Ежова И.М.	БД	256	Практ
Ежова И.М.	БД	123	Лекция
Бобров М.И.	БД	256	Лекция
Бобров М.И.	Паскаль	256	Практ
Волков Н.Г.	БД	123	Практ
Волков Н.Г.	Паскаль	256	Лекция
Зайцев В. В.	ТСИ	244	Лекция

- Исследование отношений **Сведения** и **Нагрузка** показывает, что переход к 2НФ позволил исключить явную избыточность данных в таблице **Сведения** - повторение строк со сведениями о преподавателях.

Но в нем по-прежнему имеет место неявное дублирование данных.

- Для дальнейшего совершенствования отношения необходимо преобразовать его в 3НФ.

Третья нормальная форма

- **Определение 1.** Отношение находится в 3НФ, если оно находится в 2НФ и каждый неключевой атрибут нетранзитивно зависит от первичного ключа.
- **Определение 2.** Отношение находится в 3НФ в том и только в том случае, если все неключевые атрибуты отношения взаимно независимы и полностью зависят от первичного ключа.

- Если в отношении **Нагрузка** транзитивные зависимости отсутствуют, то в отношении **Сведения** они есть:
- ФИО □ Должн □ Оклад
- Ф И О □ Стажн □ Д_Стаж
- Транзитивные зависимости также порождают избыточное дублирование информации в отношении. Устраним их.
- Для этого используя операцию проекции на атрибуты, являющиеся причиной транзитивных зависимостей, преобразуем отношение **Сведения**, получив при этом отношения **Сведения I**, **Оклады** и **Стаж**, каждое из которых находится в ЗНФ

Сведения I

ФИО *	Должн	Стаж	Каф
Ежова И.М.	преп	5	25
Бобров М.И.	ст. преп	7	25
Волков Н.Г.	преп	10	25
Зайцев В. В.	преп	5	24

Оклады

Должн*	Оклад
преп	5000
ст. преп	8000

Стаж

Стаж*	Д_Стаж
5	100
7	120
10	150

База данных «Преподаватель»:

- Нагрузка
 - Сведения I
 - Оклады
 - Стаж
-
- Все отношения находятся в 3НФ.
Нормализация осуществлена.

Нормальная форма Бойса-Кодда

- Теоретики реляционных систем Кодд и Бойс обосновали и предложили более строгое определение для 3НФ, которое учитывает, что в таблице может быть несколько *ВОЗМОЖНЫХ* ключей.
- **Таблица находится в *нормальной форме Бойса-Кодда (НФБК)*, если и только если любая функциональная зависимость между его полями сводится к полной функциональной зависимости от *ВОЗМОЖНОГО* ключа.**

Торговля

Наименование магазина	ФИО директора	Адрес	телефон	Наименование товара	Кол-во	Цена (руб)
Осень	Котов	Гагарина,3	335463	Шоколад	30	60
Ясень	Яшкин	П.Зори,42	360215	Фанта	125	36
Ясень	Яшкин	П.Зори,42	360215	Шоколад	50	72
Клен	Нивин	Ленина,22		Набор мебели	5	72000
Клен	Нивин	Ленина,22		Меб. Гарнитур	4	56000
Бук	Гарин	Мира,28	244617	Электрочайник	50	1100
Бук	Гарин	Мира,28	244617	Холодильник	3	9000
Жасмин	Жуков	Марата,33	582214	Меб. Гарнитур	2	65000
Жасмин	Жуков	Марата,33	582214	Холодильник	5	9500

Первичный ключ:

Наим_магазина,Наим_товара

ЧФЗ:

- **Наим_магазина**-> ФИО_директора,
Адрес,Телефон

ПФЗ:

- **Наим_магазина,Наим_товара**-> Кол-во,
Цена

Магазины

ЗНФ

Наименование магазина *	ФИО директора	Адрес	телефон
Осень	Котов	Гагарина,3	335463
Ясень	Яшкин	П.Зори,42	360215
Клен	Нивин	Ленина,22	
Бук	Гарин	Мира,28	244617
Жасмин	Жуков	Марата,33	582214

Продажи

ЗН

Ф

Наименование магазина *	Наименование товара *	Кол-во	Цена (руб)
Осень	Шоколад	30	60
Ясень	Фанта	125	36
Ясень	Шоколад	50	72
Клен	Набор мебели	5	72000
Клен	Меб. Гарнитур	4	56000
Бук	Электрочайник	50	1100
Бук	Холодильник	3	9000
Жасмин	Меб. Гарнитур	2	65000
Жасмин	Холодильник	5	9500

База данных «Торговля»:

- Магазины 3НФ
- Продажи 3НФ

- Нормализация выполнена.

КИНОПРОКАТ

Фильм	Исполнитель гл. роли	Жанр	Страна создания	Год выпуска	Кино театр	Адрес	Время сеанса
Форсаж 6	Уолкер	боевик	США	2013	Нева	Невский	19-00
Титаник	Ди Каприо	трагедия	США	1997	Галерея	Лиговский	18-00
Виноваты звезды	Вудли	мелодрама	США	2014	Москва	Хасанская	20-00
Илья Муромец и Соловей-разбойник	Соловьев	мультфильм	Россия	2007	Аврора	Садовая	10-00
Замуж на два дня	Крюгер	комедия	Франция	2012	Нева	Невский	16-00
Легенда № 17	Козловский	драма	Россия	2012	Аврора	Садовая	17-00
Титаник	Ди Каприо	трагедия	США	1997	ПИК	Сенная	15-00
Старые клячи	Гурченко	комедия	Россия	1999	Галерея	Лиговский	16-00
Форсаж 6	Уолкер	боевик	США	2013	Аврора	Садовая	20-00



Первичный ключ

Фильм, Кинотеатр

Зависимости:

● Частичные

Фильм -> исполнитель гл.роли, жанр,
страна, год выпуска

Кинотеатр-> адрес

Полная

Фильм, Кинотеатр-> время сеанса

Фильмы

ЗНФ

Фильм *	Исполнитель гл. роли	Жанр	Страна создания	Год выпуска
Форсаж 6	Уолкер	боевик	США	2013
Титаник	Ди Каприо	трагедия	США	1997
Виноваты звезды	Вудли	мелодрама	США	2014
Илья Муромец и Соловей-разбойник	Соловьев	мультфильм	Россия	2007
Замуж на два дня	Крюгер	комедия	Франция	2012
Легенда № 17	Козловский	драма	Россия	2012
Старые клячи	Гурченко	комедия	Россия	1999

Кинотеатр ЗНФ

Кино Театр *	Адрес
Нева	Невский
Галерея	Лиговский
Москва	Хасанская
Аврора	Садовая
ПИК	Сенная

Сеансы

ЗНФ

Фильм *	Кино * театр	Время сеанса
Форсаж 6	Нева	19-00
Титаник	Галерея	18-00
Виноваты звезды	Москва	20-00
Илья Муромец и Соловей-разбойник	Аврора	10-00
Замуж на два дня	Нева	16-00
Легенда № 17	Аврора	17-00
Титаник	ПИК	15-00
Старые клячи	Галерея	16-00
Форсаж 6	Аврора	20-00

БД «Кинопрокат»:

- Фильмы
- Кинотеатры
- Сеансы

Справочники

Жанры_с

Код_ж	Жанр
1	Боевик
2	Трагедия
3	мелодрама
4	Мультфильм
5	Комедия
6	Драма

Страны_с

Код_С	Страна
1	США
2	Россия
3	Франция

Фильмы _с

Код_ф	Фильм *
1	Форсаж 6
2	Титаник
3	Виноваты звезды
4	Илья Муромец и Соловей-разбойник
5	Замуж на два дня
6	Легенда № 17
7	Старые клячи

Кинотеатр_с

Код_К*	Кинотеатр
1	Нева
2	Галерея
3	Москва
4	Аврора
5	ПИК

Кинотеатр I

Код_К	Адрес
1	Невский
2	Лиговский
3	Хасанская
4	Садовая
5	Сенная

ФИЛЬМ_ I

Код_ф	Исполнитель гл. роли	Код_Ж	Код_С	Год выпуска
1	Уолкер	1	1	2013
2	Ди Каприо	2	1	1997
3	Вудли	3	1	2014
4	Соловьев	4	2	2007
5	Крюгер	5	3	2012
6	Козловский	6	2	2012
7	Гурченко	5	2	1999

Сеансы I

Код_ф	Код_К	Время сеанса
1	1	19-00
2	2	18-00
3	3	20-00
4	4	10-00
5	1	16-00
6	4	17-00
2	5	15-00
7	2	16-00
1	4	20-00

Окончательный вариант БД «Кинопрокат»

- Сеансы I,
- Фильм I,
- Кинотеатр I,
- Жанр_с,
- Страна_с
- Кинотеатр_с
- Фильм_с

Схема «таблицы-связи»

