

Резюме предыдущей лекции

Необходимо сохранять следующие данные о студентах

Сном	— номер студбилета студента;
Сфам	— фамилия, имя, отчество студента;
Кном	— номер комнаты в общежитии, в которой проживает студент;
Тел	— номер телефона в комнате;
Спорт	— вид спорта, которым занимается студент;
Сем	— семестр, в котором студент сдал экзамен по дисциплине;
Дис	— дисциплина, сданная студентом;
Преп	— фамилия преподавателя, принявшего экзамен;
Оц	— оценка, полученная студентом на экзамене по дисциплине.

Выявлены правила предметной области (бизнес-правила)

1. Номера студбилетов **уникальны**.
3. Студент проживает **в одной** комнате.
5. Номера комнат **уникальны**.
6. В каждой комнате установлен **точно один** телефон.
7. Номера телефонов **уникальны**.
8. Студент может увлекаться **несколькими** видами спорта.
10. Студент может получить **только одну** оценку по дисциплине в семестре.
11. Каждый преподаватель принимает экзамен **только по одной** дисциплине.

Из правил 10, 11 следует

12. Преподаватель может поставить студенту **не более одной** оценки в семестре.

Для хранения данных создана таблица УНИВЕРСУМ

Сном	Сфам	Кном	Тел	Спорт	Сем	Дис	Преп	Оц
------	------	------	-----	-------	-----	-----	------	----

Из правила 8 следует

В столбце Спорт таблицы УНИВЕРСУМ могут встречаться значения-списки.
УНИВЕРСУМ не является отношением в смысле РМД.

Таблица УНИВЕРСУМ представлена в виде двух отношений

УВЛЕЧЕНИЕ

Сном	Спорт
------	-------

УНИВЕРСУМ_1

Сном	Сем	Дис	Сфам	Кном	Тел	Преп	Оц
------	-----	-----	------	------	-----	------	----

Отношение УВЛЕЧЕНИЕ полноключевое.

```
CREATE BASE RELATION УВЛЕЧЕНИЕ
{Сном  DOMAIN S_numbers NOT NULL,
 Спорт DOMAIN Sport      NOT NULL,
  PRIMARY KEY (Сном, Спорт) ,
};
```

РСУБД может поддерживать уникальность кортежей этого отношения.

Каждый кортеж отношения УВЛЕЧЕНИЕ ссылается на студента Сном,
но определить эту ссылку в нашей модели НЕВОЗМОЖНО, так как
не определено отношение с первичным ключом Сном.

Из правила 10 следует

Группа {Сном, Сем, Дис} есть **первичный ключ** отношения УНИВЕРСУМ_1.

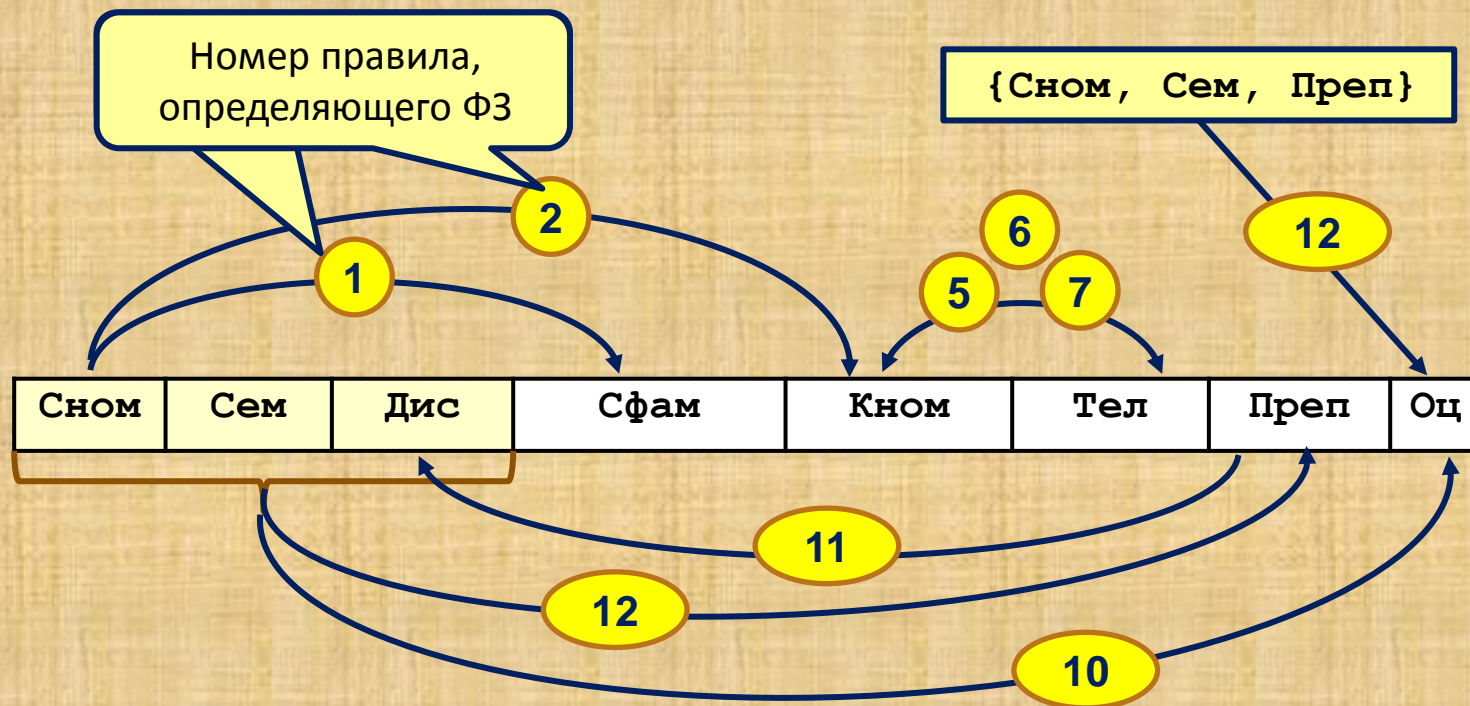
Отношение УНИВЕРСУМ_1

```
CREATE BASE RELATION УНИВЕРСУМ_1
{  Сном DOMAIN S_numbers NOT NULL,
   Сфам DOMAIN Surnames   NOT NULL,
   Кном DOMAIN R_numbers  NOT NULL,
   Тном DOMAIN T_numbers  NOT NULL,
   Сем  DOMAIN Semesters  NOT NULL,
   Дис  DOMAIN Subjects   NOT NULL,
   Преп DOMAIN Teachers   NOT NULL,
   Оц   DOMAIN E_marks    NOT NULL,
      PRIMARY KEY (Сном, Сем, Дис)
};
```

РСУБД гарантирует поддержку ФЗ с детерминантом {Сном, Сем, Дис}
И ТОЛЬКО ЭТИХ ФЗ.

Этим обусловлены проблемы обновления УНИВЕРСУМ_1.

Набор неприводимых ФЗ на отношении УНИВЕРСУМ_1



НОРМАЛЬНЫЕ ФОРМЫ ОТНОШЕНИЙ

Первая нормальная форма (1НФ)

1НФ

Отношение находится в 1НФ, если и только если каждый его атрибут определён на домене простого типа данных.

Любое отношение РМД находится в 1НФ ПО ОПРЕДЕЛЕНИЮ.

ФЗ, возможные в 1НФ

R(A, B, C, D, E, F)

A, B, C, ...
подмножества атрибутов

Первичный ключ

РСУБД может поддерживать ТОЛЬКО эти ФЗ!

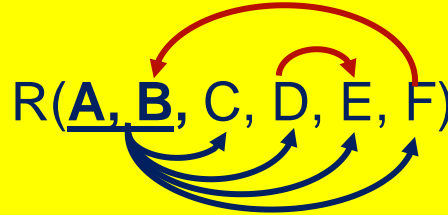
Отношение, находящееся в 1НФ, НЕЖЕЛАТЕЛЬНО в структуре РБД.

Вторая нормальная форма (2НФ)

2НФ

Отношение находится в 2НФ, если и только если оно находится в 1НФ и каждый его неключевой атрибут зависит от полного первичного ключа.

ФЗ, возможные в 2НФ



Любое отношение, находящееся в 1НФ, можно представить без потерь информации в виде двух проекций, находящихся в 2НФ.

Теорема Хеза. Пусть $R(A, B, C)$ отношение, A, B, C — его атрибуты.
Если R удовлетворяет ФЗ $A \rightarrow B$, то $R = R[A, B] \text{ JOIN } R[A, C]$.

ПОЯСНЕНИЕ. Эта декомпозиция порождает ограничение ссылочной целостности.

В проекции $R[A, B]$ детерминант ФЗ является **первичным ключом**,

В проекции $R[A, C]$ — **внешним ключом**, ссылающимся на первичный ключ $R[A, B]$.

Исходное отношение может быть восстановлено путём естественного соединения проекций.

Соединение будет содержать **все** кортежи исходного отношения R и

не будет содержать кортежей, которых **не было** в отношении R .

(Декомпозиция без потерь информации)

УНИВЕРСУМ_1 (1НФ)

Сном	Сфам	Кном	Тел	Сем	Дис	Преп	Оц
0221	Иванков П.С.	923	41-12-13	1	РЯ	Ломов	5
0223	Латушов К.Ж.	923	41-12-13	1	Физ.	Пикин	3
0112	Дубов Вася	123	11-21-32	1	Физ.	Зубов	3
0221	Иванков П.С.	923	41-12-13	2	РЯ	Ломов	4
0223	Латушов К.Ж.	923	41-12-13	1	ОТУ	Пакин	4

2НФ

СТУДЕНТ_КОМНАТА

Сном	Сфам	Кном	Тел
0221	Иванков П.С.	923	41-12-13
0223	Латушов К.Ж.	923	41-12-13
0112	Дубов Вася	123	11-21-32
5678	Колов Петя	123	55-32-78

ОЦЕНКА_1

Сном	Сем	Дис	Преп	Оц
0221	1	РЯ	Ломов	5
0223	1	Физ.	Пикин	3
0112	1	Физ.	Зубов	3
0221	2	РЯ	Ломов	4
0223	1	ОТУ	Пакин	4

СТУДЕНТ_КОМНАТА JOIN ОЦЕНКА_1 эквивалентно УНИВЕРСУМ_1

Не поддерживается ФЗ Кном ↔ Тел

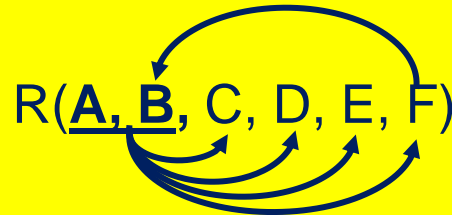
Отношение, находящееся в 2НФ, **нежелательно** в структуре РБД.

Третья нормальная форма (3НФ)

3НФ

Отношение находится в 3НФ, если и только если оно находится в 2НФ и нет ФЗ между неключевыми атрибутами.

ФЗ, возможные в 3НФ



Любое отношение 2НФ, не находящееся в 3НФ, можно представить без потерь информации в виде двух проекций, находящихся в 3НФ.

СТУДЕНТ_КОМНАТА (2НФ)

Сном	Сфам	Кном	Тел
0221	Иванков П.С.	923	41-12-13
0223	Латушов К.Ж.	923	41-12-13
0112	Дубов Вася	123	11-21-32

=

СТУДЕНТ (3НФ)

Сном	Сфам	Кном
0221	Иванков П.С.	923
0223	Латушов К.Ж.	923
0112	Дубов Вася	123
5678	Колов Петя	123

JOIN

КОМНАТА (3НФ)

Кном	Тел
923	41-12-13
123	11-21-32

СТУДЕНТ JOIN КОМНАТА эквивалентно СТУДЕНТ_КОМНАТА

Нормальная форма Бойса-Кодда (НФБК)

Проблемы обновления отношения, находящегося в 3НФ

ОЦЕНКА_1

Сном	Сем	Дис	Преп	Оц
0221	1	РЯ	Ломов	5
0223	1	Физ.	Пикин	3
0112	1	Физ.	Зубов	3
0221	2	РЯ	Ломов	4
0223	1	ОТУ	Пакин	4
0223	1	РЯ	Пакин	5

Невозможно
объявить

Недопустимо.
Нарушена ФЗ
Преп → Дис

Нормальная форма
Бойса-Кодда (НФБК)

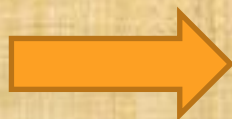
Отношение находится в НФБК, если и только если
каждый его детерминант является потенциальным ключом.

Отношение, находящееся в НФБК, не имеет аномалий обновления,
обусловленных необъявленными ФЗ.

Любое отношение ЗНФ, не находящееся в НФБК, можно представить без потерь информации в виде двух проекций, находящихся в НФБК.

ОЦЕНКА_1

Сном	Сем	Дис	Преп	Оц
0221	1	РЯ	Ломов	5
0223	1	Физ.	Пикин	3
0112	1	Физ.	Зубов	3
0221	2	РЯ	Ломов	4
0223	1	ОТУ	Пакин	4



ОЦЕНКА

Сном	Сем	Преп	Оц
0221	1	Ломов	5
0223	1	Пуров	3
0112	1	Зубов	3
0221	2	Ломов	4
0223	1	Пакин	4
0223	1	Зубов	4

ПРЕП_ДИС

Преп	Дис
Ломов	РЯ
Пуров	Физ.
Зубов	Физ.
Кукин	Физ.
Пакин	ОТУ

Потеряны ФЗ с детерминантом {Сном, Дис, Сем}.

При вставке строки в отношение **ОЦЕНКА** требуется проверка.

Вставить 0223 | 1 | Зубов | 4

1) (ПРЕП_ДИС WHERE Дис = "Физ.") [Преп] (Пуров, Кукин, Зубов)

2) Если в **ОЦЕНКА** есть строка вида

| 0223 | 1 | кто-либо из (Пуров, Кукин, Зубов) |

то запретить вставку.

Независимость проекций

Проекции отношения называются *независимыми*, если при их обновлении достаточно поддерживать только целостность ссылок.

Теорема Риссанена.

Пусть R — отношение, а $R1, R2$ - его проекции.

Проекции независимы, если и только если каждая $\Phi3$, содержащаяся в R , является следствием определений потенциальных ключей проекций.

Любое нормализованное отношение можно привести к совокупности независимо обновляемых проекций, находящихся не ниже, чем в **3НФ**.

Если отношение находится в 3НФ и содержит детерминант, функционально определяющий часть первичного ключа,

то его декомпозиция по Хезу до **НФБК** приведёт к потере $\Phi3$.

Проекции не будут независимыми.

Альтернативная декомпозиция.

Сном	Сем	Дис	Преп	Оц
0221	1	РЯ	Ломов	5
0223	1	Физ.	Пуров	3
0112	1	Физ.	Зубов	3
0221	2	РЯ	Ломов	4
0223	1	ОТУ	Пакин	4

Сном	Сем	Дис	Оц
0221	1	РЯ	5
0223	1	Физ.	3
0112	1	Физ.	3
0221	2	РЯ	4
0223	1	ОТУ	4

JOIN

Дис	Преп
РЯ	Ломов
Физ.	Пуров
Физ.	Зубов
ОТУ	Пакин



Сном	Сем	Дис	Преп	Оц
0221	1	РЯ	Ломов	5
0223	1	Физ.	Пуров	3
0223	1	Физ.	Зубов	3
0112	1	Физ.	Зубов	3
0112	1	Физ.	Пуров	3
0221	2	РЯ	Ломов	4
0223	1	ОТУ	Пакин	4

**Потеря информации.
Кому сдали Физ. студенты 0223 и 0112**

Четвёртая нормальная форма (4НФ)

СТУД-ДИС-СЕКЦ

Сном	Дис	Секц
0112	БД	Бокс
0112	БД	Теннис
0112	Физ	Бокс
0112	Физ	Теннис
0113	Хим	Лыжи
0113	Физ	Лыжи
0113	БД	Лыжи

Студ, Дис и Секц взаимно независимы.
Фиксированному значению **Сном** соответствует *множество* значений **Дис** и *множество* значений **Секц**.

**Многозначная
зависимость**

Пусть **A**, **B** и **C** — произвольные подмножества атрибутов отношения **R**.
Говорят, что $A \twoheadrightarrow B$, если и только если для любой реализации **R**
множество значений **B**, соответствующее заданной паре значений **A** и **C**,
зависит только от значения **A** и не зависит от значения **C**.

Многозначная зависимость $A \twoheadrightarrow B$ имеет место, если и только если $A \twoheadrightarrow C$.
Обозначение: $A \twoheadrightarrow B \mid C$

Теорема Фейджина.

Пусть $R(A, B, C)$ отношение, **A**, **B**, **C** — подмножества его атрибутов,
 $R[A, B]$ и $R[A, C]$ — его проекции.

Если **R** удовлетворяет многозначной зависимости $A \twoheadrightarrow B \mid C$, то
 $R = R[A, B] \text{ JOIN } R[A, C]$.

СТУД-ДИС-СЕКЦ

Сном	Дис	Секц
0112	БД	Бокс
0112	Физ	Теннис
0112	БД	Теннис
0112	Физ	Бокс
0113	Хим	Лыжи
0113	Физ	Лыжи
0113	БД	Лыжи



СТУД-ДИС

Сном	Дис
0112	БД
0112	Физ
0113	Хим
0113	Физ
0113	БД

СТУД-СЕКЦ

Сном	Секц
0112	Бокс
0112	Теннис
0113	Лыжи

**СТУД-ДИС JOIN СТУД-СЕКЦ
ЭКВИВАЛЕНТНО
СТУД-ДИС-СЕКЦ**

4НФ

Отношение находится в **4НФ**, если и только если оно находится в **НФБК** и не содержит многозначных зависимостей.

**ОБЯЗАТЕЛЬНО ЧИТАТЬ
УЧ. ПОСОБИЕ
ПРОЕКТИРОВАНИЕ БАЗ ДАННЫХ,
РАЗДЕЛ 4, СТР. 80 - 119**