

ТЕМА 7. НОРМАЛИЗАЦИЯ РЕЛЯЦИОННЫХ ОТНОШЕНИЙ. НОРМАЛЬНЫЕ ФОРМЫ

В качестве одной из методологий проектирования баз данных может быть рассмотрен подход, при котором весь процесс осуществляется в терминах реляционной модели данных методом последовательных приближений к удовлетворительному набору схем отношений. Исходной точкой является представление предметной области в виде одного или нескольких отношений, и на каждом шаге проектирования производится некоторый набор схем отношений, обладающих «улучшенными» свойствами. Процесс проектирования представляет собой процесс нормализации схем отношений, причем каждая следующая нормальная форма обладает свойствами, в некотором смысле, лучшими, чем предыдущая.

Каждой нормальной форме соответствует определенный набор ограничений, и отношение находится в некоторой нормальной форме, если удовлетворяет свойственному ей набору ограничений. Примером может служить ограничение первой нормальной формы – значения всех атрибутов отношения атомарны. Поскольку требование первой нормальной формы является базовым требованием классической реляционной модели данных, мы будем считать, что исходный набор отношений уже соответствует этому требованию.

В теории реляционных баз данных обычно выделяется следующая последовательность нормальных форм:

- первая нормальная форма (1NF);
- вторая нормальная форма (2NF);
- третья нормальная форма (3NF);
- нормальная форма Бойса-Кодда (BCNF);
- четвертая нормальная форма (4NF);
- пятая нормальная форма, или нормальная форма проекции-соединения (5NF или PJ/NF).

Основные свойства нормальных форм состоят в следующем:

- каждая следующая нормальная форма в некотором смысле лучше предыдущей нормальной формы;
- при переходе к следующей нормальной форме свойства предыдущих нормальных форм сохраняются.

В основе процесса проектирования лежит метод нормализации, т. е. декомпозиции отношения, находящегося в предыдущей нормальной форме, на два или более отношений, которые удовлетворяют требованиям следующей нормальной формы.

Минимальные функциональные зависимости и вторая нормальная форма

Пусть имеется переменная отношения `СЛУЖАЩИЕ_ПРОЕКТЫ_ЗАДАНИЯ` {`СЛУ_НОМ`, `СЛУ_УРОВ`, `СЛУ_ЗАРП`, `ПРО_НОМ`, `СЛУ_ЗАДАН`}. Новые атрибуты `СЛУ_УРОВ` и `СЛУ_ЗАДАН` содержат, соответственно, данные о разряде служащего и о задании, которое выполняет служащий в данном проекте. Будем считать, что разряд служащего определяет размер его заработной платы и что каждый служащий может участвовать в нескольких проектах, но в каждом проекте он выполняет только одно задание. Тогда очевидно, что единственным возможным ключом отношения `СЛУЖАЩИЕ_ПРОЕКТЫ_ЗАДАНИЯ` является составной атрибут {`СЛУ_НОМ`, `ПРО_НОМ`}. Диаграмма минимального множества FD показана на [рис. 1](#), а возможное тело значения отношения – на [рис. 2](#).

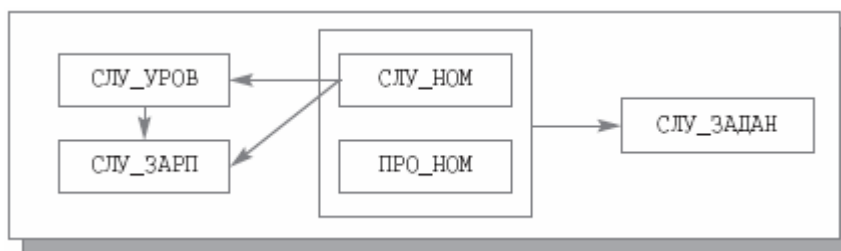


Рис. 1. Диаграмма множества FD отношения СЛУЖАЩИЕ_ПРОЕКТЫ_ЗАДАНИЯ

СЛУ_НОМ	СЛУ_УРОВ	СЛУ_ЗАРП	ПРО_НОМ	СЛУ_ЗАДАН
2934	2	22400.00	1	А
2935	3	29600.00	1	В
2936	1	20000.00	1	С
2937	1	20000.00	1	Д
2934	2	22400.00	2	Д
2935	3	29600.00	2	С
2936	1	20000.00	2	В
2937	1	20000.00	2	А

Рис. 2. Возможное значение переменной отношения СЛУЖАЩИЕ_ПРОЕКТЫ_ЗАДАНИЯ

Аномалии обновления, возникающие из-за наличия неминимальных функциональных зависимостей

Во множество FD отношения СЛУЖАЩИЕ_ПРОЕКТЫ_ЗАДАНИЯ входит много FD, в которых детерминантом является не возможный ключ отношения (соответствующие стрелки в диаграмме начинаются не с {СЛУ_НОМ, ПРО_НОМ}, т. е. некоторые функциональные зависимости атрибутов от возможного ключа не являются минимальными). Это приводит к так называемым аномалиям обновления. Под аномалиями обновления понимаются трудности, с которыми приходится сталкиваться при выполнении операций добавления кортежей в отношение (INSERT), удаления кортежей (DELETE) и модификации кортежей (UPDATE). Обсудим сначала аномалии обновления, вызываемые наличием FD СЛУ_НОМ → СЛУ_УРОВ (эти аномалии связаны с избыточностью хранения значений атрибутов СЛУ_УРОВ и СЛУ_ЗАРП в каждом кортеже, описывающем задание служащего в некотором проекте).

- **Добавление кортежей.** Мы не можем дополнить отношение СЛУЖАЩИЕ_ПРОЕКТЫ_ЗАДАНИЯ данными о служащем, который в данное время еще не участвует ни в одном проекте (ПРО_НОМ является частью первичного ключа и не может содержать неопределенных значений). Между тем часто бывает, что сначала служащего принимают на работу, устанавливают его разряд и размер зарплаты, а лишь потом назначают для него проект.

- **Удаление кортежей.** Мы не можем сохранить в отношении СЛУЖАЩИЕ_ПРОЕКТЫ_ЗАДАНИЯ данные о служащем, завершившем участие в своем последнем проекте (по той причине, что значение атрибута ПРО_НОМ для этого служащего становится неопределенным). Между тем характерна ситуация, когда между проектами возникают перерывы, не приводящие к увольнению служащих.

- **Модификация кортежей.** Чтобы изменить разряд служащего, мы будем вынуждены модифицировать все кортежи с соответствующим значением атрибута СЛУ_НОМ. В противном случае будет нарушена естественная FD СЛУ_НОМ → СЛУ_УРОВ (у одного служащего имеется только один разряд).

Возможная декомпозиция

Для преодоления этих трудностей можно произвести декомпозицию переменной отношения СЛУЖАЩИЕ_ПРОЕКТЫ_ЗАДАНИЯ на две переменных отношений – СЛУЖ {СЛУ_НОМ, СЛУ_УРОВ, СЛУ_ЗАРП} и СЛУЖ_ПРО_ЗАДАН {СЛУ_НОМ, ПРО_НОМ, СЛУ_ЗАДАН}. На основании теоремы Хита эта декомпозиция является декомпозицией без потерь, поскольку в исходном отношении имела FD {СЛУ_НОМ, ПРО_НОМ} → СЛУ_ЗАДАН. На [рис. 3](#) показаны диаграммы множеств FD этих отношений, а на [рис. 4](#) – их значения.



Рис. 3. Диаграммы FD в переменных отношениях СЛУЖ и СЛУЖ_ПРО_ЗАДАН

Теперь мы можем легко справиться с операциями обновления.

- **Добавление кортежей.** Чтобы сохранить данные о принятом на работу служащем, который еще не участвует ни в каком проекте, достаточно добавить соответствующий кортеж в отношение СЛУЖ.
- **Удаление кортежей.** Если кто-то из служащих прекращает работу над проектом, достаточно удалить соответствующий кортеж из отношения СЛУЖ_ПРО_ЗАДАН. При увольнении служащего нужно удалить кортежи с соответствующим значением атрибута СЛУ_НОМ из отношений СЛУЖ и СЛУЖ_ПРО_ЗАДАН.
- **Модификация кортежей.** Если у служащего меняется разряд (и, следовательно, размер зарплаты), достаточно модифицировать один кортеж в отношении СЛУЖ.

Значение переменной отношения СЛУЖ		
СЛУ_НОМ	СЛУ_УРОВ	СЛУ_ЗАРП
2934	2	22400.00
2935	3	29600.00
2936	1	20000.00
2937	1	20000.00

Значение переменной отношения СЛУЖ_ПРО_ЗАДАН		
СЛУ_НОМ	ПРО_НОМ	СЛУ_ЗАДАН
2934	1	A
2935	1	B
2936	1	C
2937	1	D
2934	2	D
2935	2	C
2936	2	B
937	2	A

Рис. 4. Значения переменных отношений

Вторая нормальная форма

Как видно, на [рис. 3](#) отсутствуют FD, не являющиеся **минимальными**. Наличие таких FD на [рис. 1](#) вызывало аномалии обновления. Проблема заключалась в том, что атрибут СЛУЖ_УРОВ относился к сущности служащий, в то время как первичный ключ идентифицировал сущность задание_служащего_в_проекте.

Переменная отношения находится во второй нормальной форме (2NF) тогда и только тогда, когда она находится в первой нормальной форме, и каждый неключевой атрибут минимально функционально зависит от первичного ключа.

Нетранзитивные функциональные зависимости и третья нормальная форма

В произведенной декомпозиции переменной отношения СЛУЖАЩИЕ_ПРОЕКТЫ_ЗАДАНИЯ множество FD переменной отношения СЛУЖ_ПРО_ЗАДАН предельно просто – в единственной нетривиальной функциональной зависимости детерминантом является возможный ключ. При использовании этой переменной отношения какие-либо аномалии обновления не возникают. Однако переменная отношения СЛУЖ не является такой же совершенной.

Аномалии обновлений, возникающие из-за наличия транзитивных функциональных зависимостей

Функциональные зависимости переменной отношения СЛУЖ по-прежнему порождают некоторые аномалии обновления. Они вызываются наличием транзитивной FD $СЛУ_НОМ \rightarrow СЛУ_ЗАРП$ (через $СЛУ_НОМ \rightarrow СЛУ_УРОВ$ и $СЛУ_УРОВ \rightarrow СЛУ_ЗАРП$). Эти аномалии связаны с избыточностью хранения значения атрибута СЛУ_ЗАРП в каждом кортеже, характеризующем служащих с одним и тем же разрядом.

- **Добавление кортежей.** Невозможно сохранить данные о новом разряде (и соответствующем ему размере зарплаты), пока не появится служащий с новым разрядом. (Первичный ключ не может содержать неопределенные значения.)

- **Удаление кортежей.** При увольнении последнего служащего с данным разрядом мы утратим информацию о наличии такого разряда и соответствующем размере зарплаты.

- **Модификация кортежей.** При изменении размера зарплаты, соответствующей некоторому разряду, мы будем вынуждены изменить значение атрибута СЛУ_ЗАРП в кортежах всех служащих, которым назначен этот разряд (иначе не будет выполняться FD $СЛУ_УРОВ \rightarrow СЛУ_ЗАРП$).

Возможная декомпозиция

Для преодоления этих трудностей произведем декомпозицию переменной отношения СЛУЖ на две переменных отношений – СЛУЖ1 {СЛУ_НОМ, СЛУ_УРОВ} и УРОВ {СЛУ_УРОВ, СЛУ_ЗАРП}. По теореме Хита, это снова декомпозиция без потерь по причине наличия, например, FD $СЛУ_НОМ \rightarrow СЛУ_УРОВ$. На [рис. 7.5](#) показаны диаграммы FD этих переменных отношений, а на [рис. 7.6](#) – их возможные значения.



Рис. 5. Диаграммы FD в отношениях СЛУЖ1 и УРОВ

Как видно из [рис. 6](#), это преобразование обратимо, т. е. любое допустимое значение исходной переменной отношения СЛУЖ является естественным соединением значений отношений СЛУЖ1 и УРОВ. Также можно заметить, что мы избавились от трудностей при выполнении операций обновления.

- **Добавление кортежей.** Чтобы сохранить данные о новом разряде, достаточно добавить соответствующий кортеж к отношению УРОВ.
- **Удаление кортежей.** При увольнении последнего служащего, обладающего данным разрядом, удаляется соответствующий кортеж из отношения СЛУЖ1, и данные о разряде сохраняются в отношении УРОВ.
- **Модификация кортежей.** При изменении размера зарплаты, соответствующей некоторому разряду, изменяется значение атрибута СЛУ_ЗАРП ровно в одном кортеже отношения УРОВ.

Третья нормальная форма

Значение переменной отношения СЛУЖ1	
СЛУ_НОМ	СЛУ_УРОВ
2934	2
2935	3
2936	1
2937	1

Значение переменной отношения УРОВ	
СЛУ_УРОВ	СЛУ_ЗАРП
2	22400.00
3	29600.00
1	20000.00

Рис. 6. Тела отношений СЛУЖ1 и УРОВ

Трудности, которые мы испытывали, были связаны с наличием транзитивной FD $СЛУ_НОМ \rightarrow СЛУ_ЗАРП$. Наличие этой FD на самом деле означало, что атрибут СЛУ_ЗАРП характеризовал не сущность служащий, а сущность разряд.

Переменная отношения находится в третьей нормальной форме (3NF) в том и только в том случае, когда она находится во второй нормальной форме, и каждый неключевой атрибут нетранзитивно⁴¹ функционально зависит от первичного ключа⁴².

Перекрывающиеся возможные ключи и нормальная форма Бойса-Кодда

До сих пор в определениях нормальных форм мы предполагали, что у декомпозируемого отношения имеется только один возможный ключ. На практике чаще всего бывает именно так. Но имеется один частный случай, который (почти) удовлетворяет требованиям 2NF и 3NF, но, тем не менее, порождает аномалии обновления. Это тот случай, когда у отношения имеется несколько возможных ключей, и некоторые из этих возможных ключей «перекрываются», т. е. содержат общие атрибуты.

Аномалии обновлений, связанные с наличием перекрывающихся возможных ключей

Например, пусть имеется переменная отношения СЛУЖ_ПРО_ЗАДАН1 {СЛУ_НОМ, СЛУ_ИМЯ, ПРО_НОМ, СЛУ_ЗАДАН} с множеством FD, показанным на [рис. 7](#).

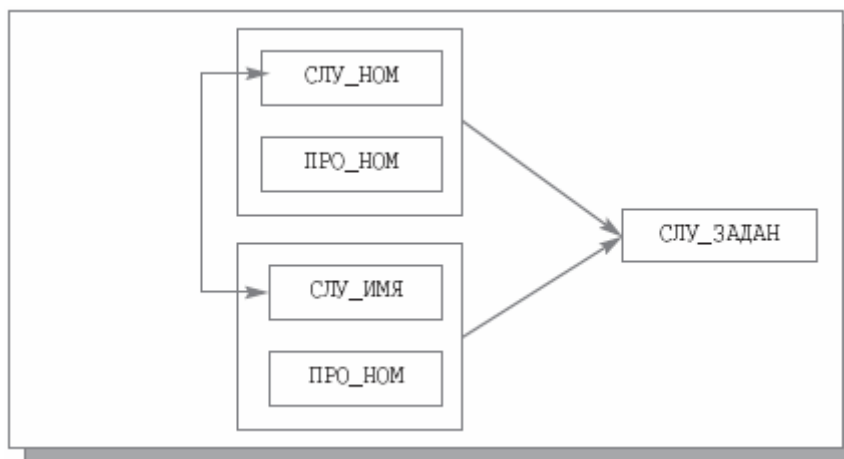


Рис. 7. Диаграмма FD отношения СЛУЖ_ПРО_ЗАДАН1

В отношении СЛУЖ_ПРО_ЗАДАН1 служащие уникально идентифицируются как по номерам удостоверений, так и по именам. Следовательно, существуют FD $\text{СЛУ_НОМ} \rightarrow \text{СЛУ_ИМЯ}$ и $\text{СЛУ_ИМЯ} \rightarrow \text{СЛУ_НОМ}$. Но один служащий может участвовать в нескольких проектах, поэтому возможными ключами являются $\{\text{СЛУ_НОМ}, \text{ПРО_НОМ}\}$ и $\{\text{СЛУ_ИМЯ}, \text{ПРО_НОМ}\}$. На [рис. 8](#) показано возможное значение переменной отношения СЛУЖ_ПРО_ЗАДАН1.

СЛУ_НОМ	СЛУ_ИМЯ	ПРО_НОМ	СЛУ_ЗАДАН
2934	Иванов	1	А
2941	Иваненко	2	В
2934	Иванов	2	В
2941	Иваненко	1	А

Рис. 8. Возможное значение переменной отношения СЛУЖ_ПРО_ЗАДАН1

Очевидно, что, хотя в отношении СЛУЖ_ПРО_ЗАДАН1 все FD неключевых атрибутов от возможных ключей являются минимальными и транзитивные FD отсутствуют, этому отношению свойственны аномалии обновления. Например, в случае изменения имени служащего требуется обновить атрибут СЛУ_ИМЯ во всех кортежах отношения СЛУЖ_ПРО_ЗАДАН1, соответствующих данному служащему. Иначе будет нарушена FD $\text{СЛУ_НОМ} \rightarrow \text{СЛУ_ИМЯ}$, и база данных окажется в несогласованном состоянии.

Нормальная форма Бойса-Кодда

Причиной отмеченных аномалий является то, что в требованиях 2NF и 3NF не требовалась минимальная функциональная зависимость от первичного ключа атрибутов, являющихся компонентами других возможных ключей. Проблему решает нормальная форма, которую исторически принято называть нормальной формой Бойса-Кодда и которая является уточнением 3NF в случае наличия нескольких перекрывающихся возможных ключей.

Переменная отношения находится в нормальной форме Бойса-Кодда (BCNF) в том и только в том случае, когда любая выполняемая для этой переменной отношения нетривиальная и минимальная FD имеет в качестве детерминанта некоторый возможный ключ данного отношения.

Переменная отношения СЛУЖ_ПРО_ЗАДАН1 может быть приведена к BCNF путем одной из двух декомпозиций: СЛУЖ_НОМ_ИМЯ {СЛУ_НОМ, СЛУ_ИМЯ} и СЛУЖ_НОМ_ПРО_ЗАДАН {СЛУ_НОМ, ПРО_НОМ, СЛУ_ЗАДАН} с множеством FD и значениями, показанными на [рис. 9](#), и СЛУЖ_НОМ_ИМЯ {СЛУ_НОМ, СЛУ_ИМЯ} и СЛУЖ_ИМЯ_ПРО_ЗАДАН {СЛУ_ИМЯ, ПРО_НОМ, СЛУ_ЗАДАН} (FD и значения результирующих переменных отношений выглядят аналогично).

Очевидно, что каждая из декомпозиций устраняет трудности, связанные с обновлением отношения СЛУЖ_ПРО_ЗАДАН1.



Рис. 9. Диаграммы FD и значения переменных отношений СЛУЖ_НОМ_ИМЯ и СЛУЖ_НОМ_ПРО_ЗАДАН