

*Проектирование
структуры РБД.
Нормализация*

Лекция 8

Например, мы построили отношение

Поставка(Товар, Поставщик, Адрес, Цена, Склад, Объем).

Предположим, что изменился адрес одного из поставщиков.

необходимо просмотреть все кортежи отношения и для соответствующего поставщика изменить его адрес.

- на выполнение этой операции потребуется много времени.
- возможно, что в каком-то кортеже адрес не будет изменен, тогда база данных будет содержать противоречивую информацию.

если какой-то поставщик временно прекратил поставки товаров и база данных не содержит кортежей с именем этого поставщика, то его адрес теряется.

Такая база данных содержит *избыточную информацию* (*дублирование*)

Решением этих проблем было бы разделение отношения **Поставка** на несколько других отношений:

- **Склад**(Номер, *Объем*),
- **Поставщик**(Название, *Адрес*),
- **Товар**(Наименование, Поставщик, *Цена*, *Склад*).

Действия по разделению одного отношения на несколько отношений называются **декомпозицией**.

Функциональные зависимости

Данные, которые будут храниться в отношении, взаимосвязаны между собой.

Эта взаимосвязь определяется множеством функциональных зависимостей между атрибутами отношения.

Зависимости эти вытекают из ограничений предметной области.

Например, в отношении **Поставки** существуют следующие ограничения:

- каждый поставщик имеет только один адрес,
- каждый поставщик поставяет товар по определенной цене,
- товары, поставленные разными поставщиками, могут быть распределены по разным складам, но товар одного наименования, поставяемый одним поставщиком, должен храниться только на одном складе,
- каждый склад имеет свой объем.

Эти ограничения являются зависимостями, которые можно сформулировать следующим образом:

- адрес функционально зависит от поставщика,
- цена функционально зависит от товара и поставщика,
- номер склада функционально зависит от товара и поставщика,
- объем функционально зависит от номера склада.

Функциональная зависимость имеет место, когда значения кортежа на одном множестве атрибутов однозначно определяют значения кортежа на другом множестве атрибутов (или на одном атрибуте).

Функциональную зависимость будем обозначать следующим образом:

- $\text{Поставщик} \rightarrow \text{Адрес},$
- $\{\text{Товар}, \text{Поставщик}\} \rightarrow \text{Цена},$
- $\{\text{Товар}, \text{Поставщик}\} \rightarrow \text{Склад},$
- $\text{Склад} \rightarrow \text{Объем}.$

А читаются они так:

- Поставщик определяет Адрес

Избыточное дублирование данных и аномалии

Следует различать

- *простое (неизбыточное)*
- *избыточное дублирование* данных.

Наличие первого из них *допускается* в базах данных, а избыточное дублирование данных может приводить к проблемам при обработке данных.

Пример *неизбыточного дублирования* данных представляет отношение С_Т с атрибутами Сотрудник и Телефон.

С_Т	
Сотрудник	Телефон
Иванов	3721
Петров	4328
Сидоров	4328
Егоров	4328

Пример *избыточного дублирования* (избыточности) представляет отношение С_Т_Н, которое, в отличие от отношения С_Т, дополнено атрибутом Н_комн (номер комнаты сотрудника).

С_Т_Н		
Сотрудник	Телефон	Н_комн
Иванов	3721	109
Петров	4328	111
Сидоров	4328	111
Егоров	4328	111

пример *неудачного отношения* С_Т_Н, в котором вместо телефонов Сидорова и Егорова поставлены прочерки (неопределенные значения).

С_Т_Н		
Сотрудник	Телефон	Н_комн
Иванов	3721	109
Петров	4328	111
Сидоров	----	111
Егоров	----	111

способ выхода из данной ситуации создать два отношения С_Н и Н_Т, полученные путем декомпозиции исходного отношения С_Т_Н.

С_Н

Сотрудник	Н_комн
Иванов	109
Петров	111
Сидоров	111
Егоров	111

Т_Н

Телефон	Н_комн
3721	109
4328	111

Процедура декомпозиции отношения С_Т_Н на два отношения С_Н и Н_Т является основной процедурой **нормализации отношений**.

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

Выделяют три основные вида аномалий:

- аномалии модификации (или редактирования),
- аномалии удаления
- аномалии добавления.

Аномалии модификации (или редактирования)

проявляются в том, что изменение значения одного данного может повлечь за собой просмотр всей таблицы и соответствующее изменение некоторых других записей таблицы.

Например, изменение номера телефона в комнате 111, что представляет собой один единственный факт, потребует просмотра всей таблицы С_Т_Н и изменения поля Н_комн согласно текущему содержимому таблицы в записях, относящихся к Петрову, Сидорову и Егорову

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.

В таблице С_Т_Н удаление записи о сотруднике Иванове (например, по причине увольнения или ухода на заслуженный отдых) приводит к исчезновению информации о номере телефона, установленного в 109-й комнате.

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.

Примером может служить операция добавления нового сотрудника все в ту же таблицу С_Т_Н. Очевидно, будет противоестественным хранение сведений в этой таблице только о комнате и номере телефона в ней, пока никто из сотрудников не помещен в нее. Более того, если в таблице С_Т_Н поле Служащий является ключевым, то хранение в ней неполных записей с отсутствующей фамилией служащего просто недопустимо из-за неопределенности значения ключевого поля.

Домашнее задание:

- Решить примеры реляционной алгебры