

Введение в базы данных

3 января 2022 г.

Содержание

1	Теория	2
1.1	Реляционная алгебра. Предназначение и свойства	3
1.2	Реляционная алгебра. Унарные и множественные операции	5
1.2.1	Проекция	5
1.2.2	Фильтрация	6
1.2.3	Переименование	8
1.2.4	Множественные операции	8
2	Практика	11

1 Теория

1.1 Реляционная алгебра. Предназначение и свойства

Базы данных нужно уметь не только проектировать, но и использовать. Существует несколько способов формулировать запросы. Первый из рассматриваемых – *реляционная алгебра*.

Мотивация Действительно, в базах данных можно не только хранить данные, но и делать выборки, изменять их каким-либо образом. Для этого вводится понятие запроса. При первом рассмотрении, запросы нужны как минимум для выполнения следующих действий:

- Выборка данных: получить данные из базы, чтобы тем или иным способом обрабатывать их уже извне.
- Область действия обновлений: запросы позволят указывать область действия тех или иных операций, что крайне полезно. Например, к таким операциям относятся операции удаления или изменения данных: хочется указывать, на какие именно записи эти операции подействуют.
- Ограничения целостности: до сих пор было только два вида ограничений (ключи и внешние ключи). Некоторые базы данных позволяют создавать произвольные ограничения целостности, заданные на поддерживаемом языке. В рамках этих ограничений очень удобно пользоваться запросами.
- Ограничения доступа.

Определение. *Реляционная алгебра* – алгебра над множеством всех отношений.

Далее будут определены некоторые из операций (которые по определению должны быть замкнуты над носителем), и ограничения, которые им соответствуют. В целом, реляционная алгебра – императивный язык для работы с отношениями, который позволяет в явном виде, по действиям, описать, каким именно образом должен быть получен результат.

Примеры Рассмотрим несколько простых примеров операций в рамках реляционной алгебры.

- Проекция отношения на множество атрибутов: $\pi_A(R)$;
- Естественное соединение $R_1 \bowtie R_2$.

Замечание. Как уже говорилось, все операции в рамках алгебры замкнуты по определению. Это означает, что их можно комбинировать произвольным образом (при сохранении условий на возможность исполнения операции). Например: $\pi_A(R_1 \bowtie \pi_B(R_2)) \bowtie R_3$.

Операции В текущем контексте полезно уточнить, что именно понимается под операцией над отношениями в рамках реляционной алгебры. А именно, для того, чтобы определить операцию, необходимо определить следующее:

- Правило построения заголовка по заданным отношениям;
- Правило построения тела по заданным отношениям;
- Условия, при которых операция выполнима, то есть ограничения на отношения, к которым она применяется.

1.2 Реляционная алгебра. Унарные и множественные операции

В этом разделе будут описаны унарные операции в рамках реляционной алгебры. В соответствии с определением, для определения каждой операции нужно указать способ построения заголовка, тела отношения, а также условия применимости, если такие есть.

1.2.1 Проекция

Определение. *Проекцией* отношения R на множество атрибутов $A = \{a_1, a_2, \dots, a_n\}$ называется отношение, полученное из исходного путем удаления атрибутов не из A . Обозначается $\pi_A(R)$.

Данная операция может быть полезна для следующего:

- Привести отношение к виду, в котором над ним можно будет осуществить другую операцию (например, объединение);
- Выбрать из отношения только нужные данные (для выборки).

На рисунке 1 приведена иллюстрация к определению $\pi_{A_2, A_4, A_5}(A)$.

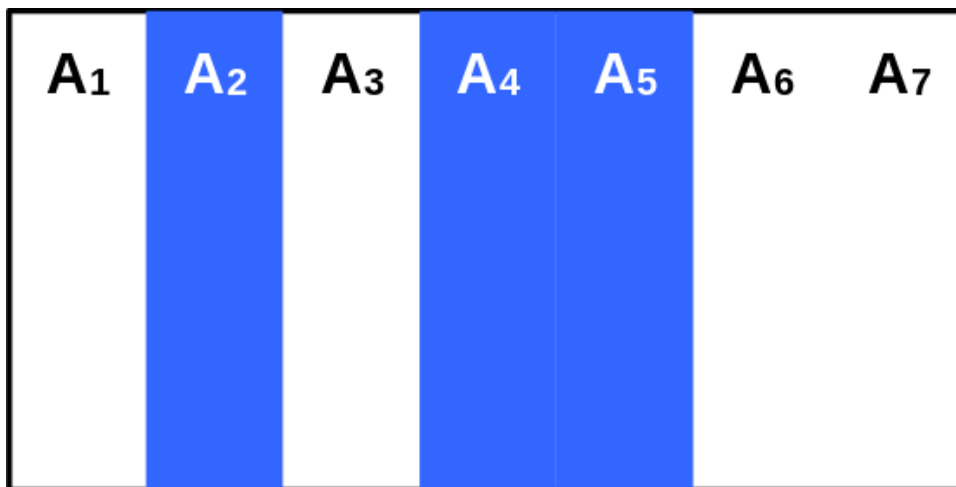


Рис. 1: Иллюстрация к определению проекции

Синим здесь обозначены столбцы, которые есть в результате операции. Остальные столбцы не используются, и результат никак не зависит от их содержимого.

Примеры Приведем несколько тривиальных примеров применения проекции.

- $\pi_{FirstName, LastName}$

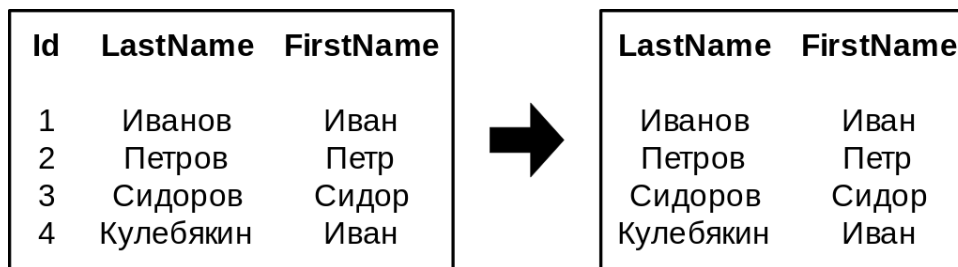


Рис. 2: Проекция. Пример 1

- $\pi_{FirstName}$

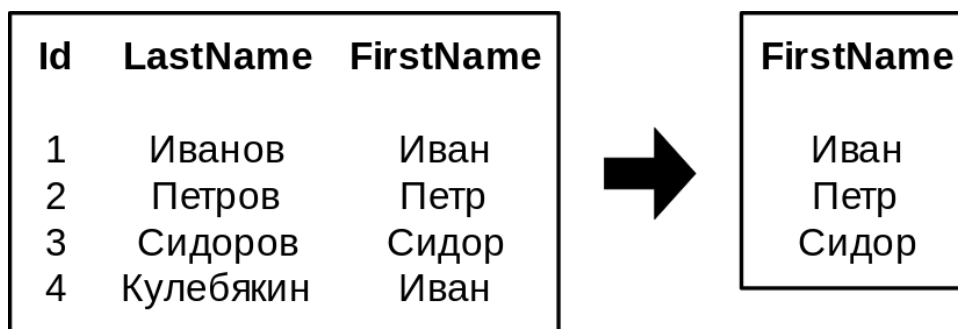


Рис. 3: Проекция. Пример 2

1.2.2 Фильтрация

Определение. *Фильтрацией (селекцией, выборкой из)* отношения R называется отношение, чей заголовок полностью совпадает с заголовком R , но тело содержит только кортежи, удовлетворяющее условию s . Обозначение: $\sigma_c(R)$.

Операция часто используется для

- Ограничения области действия изменяющих запросов;
- Получения выборки данных, соответствующих определенному условию.

На рисунке 4 приведена иллюстрация к определению $\sigma_c(R)$.



Рис. 4: Иллюстрация к определению фильтрации

Синим здесь обозначены столбцы, которые есть в результате операции. Остальные столбцы не используются, и результат никак не зависит от их содержимого.

Примеры Приведем несколько тривиальных примеров применения фильтрации.

- $\sigma_{Id>2}$

Id	LastName	FirstName
1	Иванов	Иван
2	Петров	Петр
3	Сидоров	Сидор
4	Кулебякин	Иван

➔

Id	LastName	FirstName
3	Сидоров	Сидор
4	Кулебякин	Иван

Рис. 5: Фильтрация. Пример 1

- Можно писать более сложные условия. $\sigma_{Id>2 \wedge FirstName=Иван}$

Id	LastName	FirstName
1	Иванов	Иван
2	Петров	Петр
3	Сидоров	Сидор
4	Кулебякин	Иван

➔


Id	LastName	FirstName
4	Кулебякин	Иван

Рис. 6: Фильтрация. Пример 2

- Можно использовать функции, доступные в используемой БД.

$\sigma_{\text{length}(FirstName)+2 \geq \text{length}(LastName)}$

Id	LastName	FirstName
1	Иванов	Иван
2	Петров	Петр
3	Сидоров	Сидор
4	Кулебякин	Иван



Id	LastName	FirstName
1	Иванов	Иван
2	Петров	Петр
3	Сидоров	Сидор

Рис. 7: Фильтрация. Пример 3

1.2.3 Переименование


Определение. *Переименованием* называется операция, при которой меняются названия атрибутов отношения. Тело при этом остается неизменным.

Операция часто применяется для того, чтобы отношение можно было использовать в рамках другой операции (например, при объединении с другим отношением).

Примеры Ниже приведен тривиальный пример-пояснение для операции переименования.

- $\rho_{Name=FirstName, Surname=LastName}$

Id	LastName	FirstName
1	Иванов	Иван
2	Петров	Петр
3	Сидоров	Сидор
4	Кулебякин	Иван



Id	Surname	Name
1	Иванов	Иван
2	Петров	Петр
3	Сидоров	Сидор
4	Кулебякин	Иван

Рис. 8: Переименование. Пример

1.2.4 Множественные операции

Из теории множеств в реляционную алгебру естественным образом переходят операции:

- $R_1 \cup R_2$ – объединение.
- $R_1 \cap R_2$ – пересечение.
- $R_1 \setminus R_2$ – разность.

Эти операции по определению применимы только к отношениям с одинаковыми заголовками. В результате получается отношение с таким же заголовком и телом, полученным в соответствии с множественной операцией. Иначе говоря, заголовок остается тем же, а над телами отношений производится соответствующая множественная операция (объединение, пересечение, вычитание и прочие).

Примеры

- Объединение отношений: $R_1 \cup R_2$

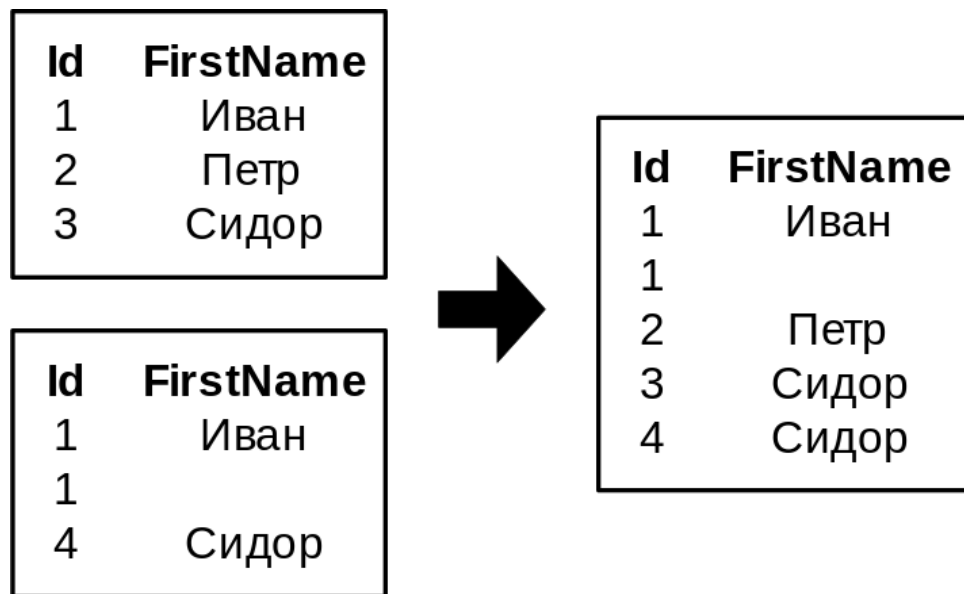


Рис. 9: Объединение отношений

- Пересечение отношений: $R_1 \cap R_2$

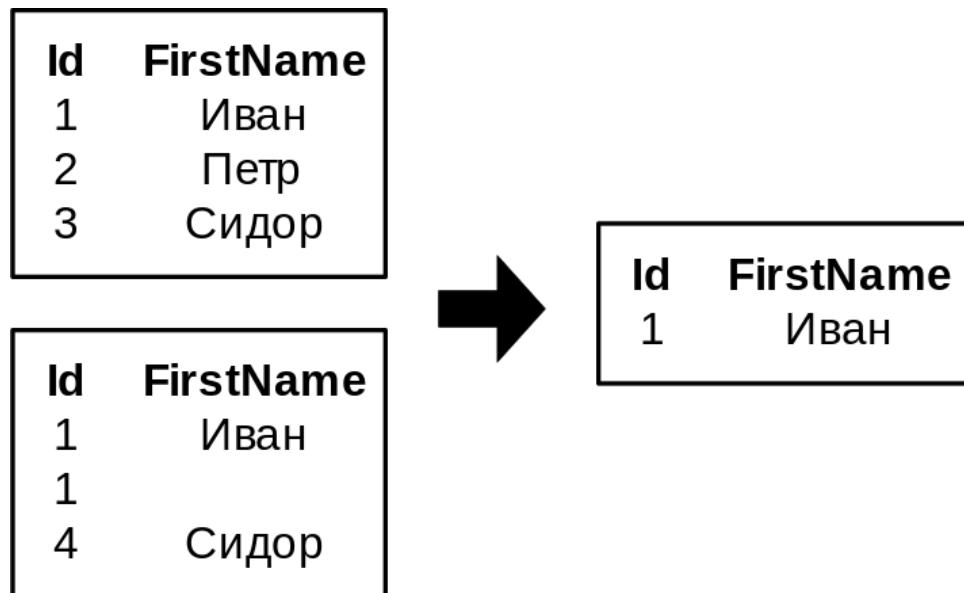


Рис. 10: Пересечение отношений

- Разность отношений: $R_1 \setminus R_2$

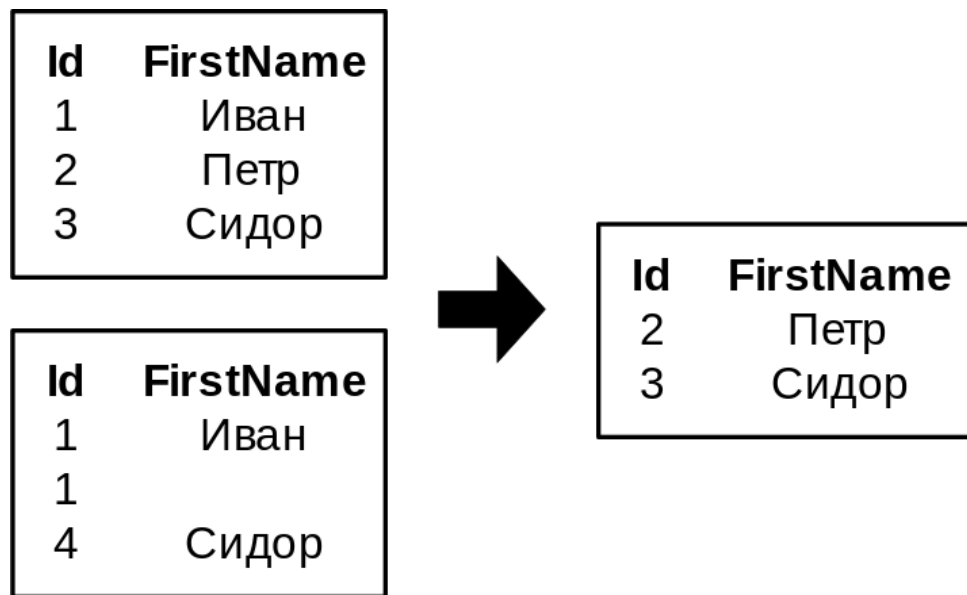


Рис. 11: Разность отношений

Стоит отметить, что для объединения отношений с различающимися именами атрибутов, но при равном их количестве, можно воспользоваться переименованием для того, чтобы привести заголовки к одному виду.

2 Практика