

Основы реляционной алгебры

Лекция 7

Основная идея *реляционной алгебры* состоит в том, что **отношения** являются **множествами**, средства манипулирования отношениями могут базироваться на *традиционных теоретико-множественных операциях*, дополненных некоторыми специальными операциями, специфичными для реляционных баз данных.

Рассмотрим начальный вариант *алгебры*, который был предложен Коддом ("*алгебра Кодда* ").

В этом варианте набор основных алгебраических операций состоит из восьми операций, которые делятся на два класса —

- теоретико-множественные операции
- специальные реляционные операции.

В состав теоретико-множественных операций входят операции:

- *объединения отношений ;*
- *пересечения отношений ;*
- *взятия разности отношений ;*
- *взятия декартова произведения отношений.*

Специальные реляционные операции включают:

- *ограничение отношения ;*
- *проекцию отношения ;*
- *соединение отношений ;*
- *деление отношений.*

Кроме того, в состав *алгебры* включается:

- *операция присваивания*, позволяющая сохранить в базе данных результаты вычисления алгебраических выражений
- *операция переименования атрибутов*, дающая возможность корректно сформировать заголовки (схему) результирующего отношения.

Интерпретация реляционных операций

- При выполнении операции объединения (*UNION*) двух отношений с одинаковыми заголовками производится отношение, включающее все кортежи, которые входят хотя бы в одно из отношений-операндов.
- Операция пересечения (*INTERSECT*) двух отношений с одинаковыми заголовками производит отношение, включающее все кортежи, которые входят в оба отношения-операнда.
- Отношение, являющееся разностью (*MINUS*) двух отношений с одинаковыми заголовками, включает все кортежи, входящие в отношение-первый операнд, такие, что ни один из них не входит в отношение, которое является вторым операндом.

Интерпретация реляционных операций

- При выполнении декартова произведения (*TIMES*) двух отношений, пересечение заголовков которых пусто, производится отношение, кортежи которого производятся путем объединения кортежей первого и второго операндов.
- Результатом ограничения (*WHERE*) отношения по некоторому условию является отношение, включающее кортежи отношения-операнда, удовлетворяющие этому условию.
- При выполнении проекции (*PROJECT*) отношения на заданное подмножество множества его атрибутов производится отношение, кортежи которого являются соответствующими подмножествами кортежей отношения-операнда.
- При соединении (*JOIN*) двух отношений по некоторому условию образуется результирующее отношение, кортежи которого производятся путем объединения кортежей первого и второго отношений и удовлетворяют этому условию.

Интерпретация реляционных операций

- У операции реляционного деления (*DIVIDE BY*) два операнда – бинарное и унарное отношения. Результирующее отношение состоит из унарных кортежей, включающих значения первого атрибута кортежей первого операнда таких, что множество значений второго атрибута (при фиксированном значении первого атрибута) включает множество значений второго операнда.
- Операция переименования (*RENAME*) производит отношение, тело которого совпадает с телом операнда, но имена атрибутов изменены.
- Операция присваивания (*:=*) позволяет сохранить результат вычисления реляционного выражения в существующем отношении БД.

Приоритеты операций:

- RENAME >= WHERE = PROJECT >= TIMES = JOIN = INTERSECT = DIVIDE BY >= UNION = MINUS

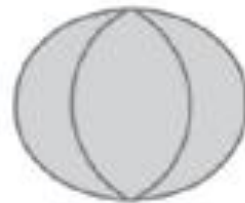
В другой форме
приоритеты операций

| Операция | Приоритет |
|-----------|-----------|
| RENAME | 4 |
| WHERE | 3 |
| PROJECT | 3 |
| TIMES | 2 |
| JOIN | 2 |
| INTERSECT | 2 |
| DIVIDE BY | 2 |
| UNION | 1 |
| MINUS | 1 |

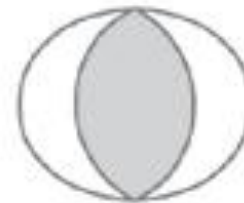
- Вычисление выражения производится слева направо с учетом приоритетов операций и скобок.

В теории множеств:

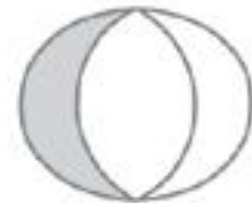
- результатом **объединения** двух множеств $A \{a\}$ и $B \{b\}$ является такое множество $C \{c\}$, что для каждого c либо существует такой элемент a , принадлежащий множеству A , что $c=a$, либо существует такой элемент b , принадлежащий множеству B , что $c=b$;
- **пересечением** множеств A и B является такое множество $C \{c\}$, что для любого c существуют такие элементы a , принадлежащий множеству A , и b , принадлежащий множеству B , что $c=a=b$;
- **разностью** множеств A и B является такое множество $C \{c\}$, что для любого c существует такой элемент a , принадлежащий множеству A , что $c=a$, и не существует такой элемент b , принадлежащий B , что $c=b$.



Объединение



Пересечение



Взятие разности

Два отношения *совместимы по объединению* в том и только в том случае, когда обладают одинаковыми заголовками.

В развернутой форме это означает, что в заголовках обоих отношений содержится один и тот же набор имен атрибутов, и одноименные атрибуты определены на одном и том же домене.

Предположим, что в базе данных имеются два отношения

- СЛУЖАЩИЕ_В_ПРОЕКТЕ_1
- СЛУЖАЩИЕ_В_ПРОЕКТЕ_2

с одинаковыми схемами {СЛУ_НОМЕР, СЛУ_ИМЯ, СЛУ_ЗАРП, СЛУ_ОТД_НОМЕР} (имена доменов опущены по причине очевидности).

Каждое из отношений содержит данные о служащих, участвующих в соответствующем проекте.

СЛУЖАЩИЕ В ПРОЕКТЕ_1

| СЛУ_НОМЕР | СЛУ_ИМЯ | СЛУ_ЗАРП | СЛУ_ОТД_НОМЕР |
|-----------|---------|----------|---------------|
| 2934 | Иванов | 22000.00 | 310 |
| 2935 | Петров | 30000.00 | 310 |
| 2936 | Сидоров | 18000.00 | 313 |
| 2937 | Федоров | 20000.00 | 310 |
| 2938 | Иванова | 22000.00 | 315 |

СЛУЖАЩИЕ В ПРОЕКТЕ_2

| СЛУ_НОМЕР | СЛУ_ИМЯ | СЛУ_ЗАРП | СЛУ_ОТД_НОМЕР |
|-----------|-----------|----------|---------------|
| 2934 | Иванов | 22000.00 | 310 |
| 2935 | Петров | 30000.00 | 310 |
| 2939 | Сидоренко | 18000.00 | 313 |
| 2940 | Федоренко | 20000.00 | 310 |
| 2941 | Иваненко | 22000.00 | 315 |

1. Тогда выполнение операции

СЛУЖАЩИЕ_В_ПРОЕКТЕ_1 UNION СЛУЖАЩИЕ_В_ПРОЕКТЕ_2

позволит получить информацию обо всех служащих, участвующих в обоих проектах.

2. Выполнение операции

СЛУЖАЩИЕ_В_ПРОЕКТЕ_1 INTERSECT СЛУЖАЩИЕ_В_ПРОЕКТЕ_2

позволит получить данные о служащих, которые одновременно участвуют в двух проектах.

3. Наконец, операция

СЛУЖАЩИЕ_В_ПРОЕКТЕ_1 MINUS СЛУЖАЩИЕ_В_ПРОЕКТЕ_2

выработает отношение, содержащее кортежи служащих, которые участвуют только в первом проекте.

СЛУЖАЩИЕ_В_ПРОЕКТЕ_1 UNION СЛУЖАЩИЕ_В_ПРОЕКТЕ_2

| СЛУ_НОМЕР | СЛУ_ИМЯ | СЛУ_ЗАРП | СЛУ_ОТД_НОМЕР |
|-----------|-----------|----------|---------------|
| 2934 | Иванов | 22000.00 | 310 |
| 2935 | Петров | 30000.00 | 310 |
| 2939 | Сидоренко | 18000.00 | 313 |
| 2940 | Федоренко | 20000.00 | 310 |
| 2941 | Иваненко | 22000.00 | 315 |
| 2936 | Сидоров | 18000.00 | 313 |
| 2937 | Федоров | 20000.00 | 310 |
| 2938 | Иванова | 22000.00 | 315 |

СЛУЖАЩИЕ_В_ПРОЕКТЕ_1 INTERSECT СЛУЖАЩИЕ_В_ПРОЕКТЕ_2

| СЛУ_НОМЕР | СЛУ_ИМЯ | СЛУ_ЗАРП | СЛУ_ОТД_НОМЕР |
|-----------|---------|----------|---------------|
| 2934 | Иванов | 22000.00 | 310 |
| 2935 | Петров | 30000.00 | 310 |

СЛУЖАЩИЕ_В_ПРОЕКТЕ_1 MINUS СЛУЖАЩИЕ_В_ПРОЕКТЕ_2

| СЛУ_НОМЕР | СЛУ_ИМЯ | СЛУ_ЗАРП | СЛУ_ОТД_НОМЕР |
|-----------|---------|----------|---------------|
| 2936 | Сидоров | 18000.00 | 313 |
| 2937 | Федоров | 20000.00 | 310 |
| 2938 | Иванова | 22000.00 | 315 |

1. Операция ограничения WHERE

требует наличия двух операндов:

- ограничиваемого отношения
- простого условия ограничения.

Простое условие ограничения может иметь вид:

- **a comp-op b**, где a и b – имена атрибутов ограничиваемого отношения; атрибуты a и b должны быть определены на одном и том же домене, для значений базового типа данных которого поддерживается операция сравнения `comp_op`, или на базовых типах данных, над значениями которых можно выполнять эту операцию сравнения;
- **a comp-op const**, где a – имя атрибута ограничиваемого отношения, а const – литерально заданная константа; атрибут a должен быть определен на домене или базовом типе, для значений которого поддерживается операция сравнения `comp_op`.

В результате выполнения *операции ограничения* производится отношение, заголовок которого совпадает с заголовком отношения-операнда, а в тело входят те кортежи отношения-операнда, для которых значением условия ограничения является true.

Результат выполнения операции

```
СЛУЖАЩИЕ_В_ПРОЕКТЕ_1 WHERE (СЛУ_ЗАРП > 20000.00  
AND (СЛУ_ОТД_НОМ = 310 OR СЛУ_ОТД_НОМ = 315))
```

| СЛУ_НОМЕР | СЛУ_ИМЯ | СЛУ_ЗАРП | СЛУ_ОТД_НОМЕР |
|-----------|---------|----------|---------------|
| 2934 | Иванов | 22000.00 | 310 |
| 2935 | Петров | 30000.00 | 310 |
| 2938 | Иванова | 22000.00 | 315 |

На интуитивном уровне *операцию ограничения* лучше всего представлять как взятие некоторой "горизонтальной" вырезки из отношения-операнда (выборки некоторых строк из таблицы).

2. Операция взятия проекции

требует наличия двух операндов

- проецируемого отношения A
- подмножества множества имен атрибутов, входящих в заголовок отношения A .

Результатом *проекции отношения A* на множество атрибутов $\{a_1, a_2, \dots, a_n\}$

PROJECT $A \{a_1, a_2, \dots, a_n\}$

является отношение с заголовком, определяемым множеством атрибутов $\{a_1, a_2, \dots, a_n\}$, и с телом, состоящим из кортежей вида $\langle a_1:v_1, a_2:v_2, \dots, a_n:v_n \rangle$ таких, что в отношении A имеется кортеж, атрибут a_1 которого имеет значение v_1 , атрибут a_2 имеет значение v_2 , ..., атрибут a_n имеет значение v_n .

Тем самым, при выполнении *операции проекции* выделяется "вертикальная" вырезка отношения-операнда с естественным уничтожением потенциально возникающих кортежей-дубликатов.

Результат операции

ПРОЕКТ СЛУЖАЩИЕ_В_ПРОЕКТЕ_1 {СЛУ_ОТД_НОМ}

(в каких отделах работают служащие, данные о которых содержатся в отношении СЛУЖАЩИЕ_В_ПРОЕКТЕ_1?)

| СЛУ_ОТД_НОМЕР |
|---------------|
| 310 |
| 313 |
| 315 |