

Реляционная алгебра

Реляционная алгебра – это коллекция операций, которые принимают отношения (таблицы) в качестве операндов и возвращают в качестве результатов.

Операции реляционной алгебры относятся к следующим классам:

1. Обычные операции над множествами – *объединение, пересечение и разность*.
2. Операции удаления частей отношения (таблицы) – *выбора и проекции*.
3. Операции сочетания строк двух отношений – *декартово произведение и соединение*.
4. Операция *переименования* атрибутов (столбцов) или отношений (таблиц) целиком.

Реляционная замкнутость – результатом любой реляционной операции с отношениями (таблицами) является отношение (таблица).

Теоретико-множественные операции над таблицами

Если A и B представляют собой множества записей, пару отношений, то они должны удовлетворять следующим условиям:

- Обладать схемами с идентичным набором атрибутов (столбцов), типы которых обязаны попарно совпадать.
- Атрибуты должны следовать в одном и том же порядке.

Отношение A

Группа	Фамилия	Имя	Адрес	ДатРожд
КБ-301	Алексеев	Иван	Екатеринбург	27.11.1995
КН-301	Иванов	Сергей	Снежинск	23.06.1996
КН-302	Сергеев	Алексей	Первоуральск	29.10.1995

Отношение В

Группа	Фамилия	Имя	Адрес	ДатРожд
КБ-301	Петров	Иван	Екатеринбург	27.11.1995
КН-301	Иванов	Сергей	Снежинск	23.06.1996
КН-302	Сергеев	Макар	Челябинск	29.10.1995

Объединение отношений А и В

Группа	Фамилия	Имя	Адрес	ДатРожд
КБ-301	Алексеев	Иван	Екатеринбург	27.11.1995
КН-301	Иванов	Сергей	Снежинск	23.06.1996
КН-302	Сергеев	Алексей	Первоуральск	29.10.1995
КБ-301	Петров	Иван	Екатеринбург	27.11.1995
КН-302	Сергеев	Макар	Челябинск	29.10.1995

Пересечение отношений A и B

Группа	Фамилия	Имя	Адрес	ДатРожд
КН-301	Иванов	Сергей	Снежинск	23.06.1996

Разность отношений А и В (A\B)

Группа	Фамилия	Имя	Адрес	ДатРожд
КБ-301	Алексеев	Иван	Екатеринбург	27.11.1995
КН-302	Сергеев	Алексей	Первоуральск	29.10.1995

Проекция

Оператор **проекции** применяется к отношению A для получения нового отношения, содержащее только часть атрибутов отношения A . Значением выражения $\pi_{A_1, A_2, \dots, A_n}(A)$ является отношение содержащее атрибуты A_1, A_2, \dots, A_n отношения A .

Π Γруппа, Фамилия, Адрес (A)

Группа	Фамилия	Адрес
КБ-301	Алексеев	Екатеринбург
КН-301	Иванов	Снежинск
КН-302	Сергеев	Первоуральск

$\sigma_{\text{Группа}=\text{'КН-302'}}(A \cup B)$

Группа	Фамилия	Имя	Адрес	ДатРожд
КН-302	Сергеев	Алексей	Первоуральск	29.10.1995
КН-302	Сергеев	Макар	Челябинск	29.10.1995

Декартово произведение

Если в операции $R \times S$ участвуют отношения ее суть не меняется.

Поскольку членами отношений являются записи, то в результате операции $R \times S$ получатся более длинные записи с компонентами обеих исходных записей.

Если отношения обладают одноименными атрибутами, то требуется переименования по меньшей мере одного из них.

Отношение R

A	B
1	2
3	4

Отношение S

B	C	D
2	5	6
4	7	8
9	10	11

Отношение $R \times S$

A	R.B	S.B	C	D
1	2	2	5	6
1	2	4	7	8
1	2	9	10	11
3	4	2	5	6
3	4	4	7	8
3	4	9	10	11

Естественное соединение

В отличие от декартова произведения, гораздо чаще возникает потребность в соединении (join) двух отношений, то есть формировании пар записей, удовлетворяющих определенному критерию.

Естественное соединение $R \bowtie S$ включает те записи из R и S , которые совпадают в общих атрибутах.

Определение.

Пусть A_1, A_2, \dots, A_n общие атрибуты для схем отношений R и S . Записи $r \in R$ и $s \in S$ допускают соединение тогда и только тогда, когда совпадают в каждом из атрибутов A_1, A_2, \dots, A_n .

Отношение R

A	B
1	2
3	4

Отношение S

B	C	D
2	5	6
4	7	8
9	10	11

Отношение $R \bowtie S$

A	B	C	D
1	2	5	6
3	4	7	8

Отношение U

A	B	C
1	2	3
6	7	8
9	7	8

Отношение V

B	C	D
2	3	4
2	3	5
7	8	10

Отношение $U \bowtie V$

A	B	C	D
1	2	3	4
1	2	3	5
6	7	8	10
9	7	8	10

Тета-соединение

Естественное соединение основано на условии равенства содержимого в общих атрибутах. Часто возникает потребность соединения отношений на основании иных критериев.

Такая операция называется тета-соединением (theta-join).

Операция тета-соединения отношений R и S в соответствии с условием C обозначим так $R \theta(C) S$.

Выполняется операция следующим образом:

- 1) вычисляется декартово произведение R и S ;
- 2) из результирующего набора выбираются записи, удовлетворяющие условию C .

Рассмотрим операцию $U \Theta(A < D) V$.

Отношение U

A	B	C
1	2	3
6	7	8
9	7	8

Отношение V

B	C	D
2	3	4
2	3	5
7	8	10

Отношение $U \theta(A < D) V$

A	U.B	U.C	V.B	V.C	D
1	2	3	2	3	4
1	2	3	2	3	5
1	2	3	7	8	10
6	7	8	7	8	10
9	7	8	7	8	10

Рассмотрим операцию $U \Theta (A < D \& U.B \neq V.B) V$.

Отношение $W=U \theta(A < D \& U.B \neq V.B) V$

A	U.B	U.C	V.B	V.C	D
1	2	3	7	8	10

Для переименования отношения R
используется оператор

$$\rho_S(A_1, A_2, \dots, A_n)(R).$$

Для переименования отношения R
используется оператор

$$\rho_{S(A_1, A_2, \dots, A_n)}(R).$$

Переименуем отношение W из последнего примера

$$\rho_{R(A,B,C,E,F,D)}(W).$$

Отношение R

A	B	C	E	F	D
1	2	3	7	8	10

Отношение R

A	B
1	2
3	4

Отношение S

B	C	D
2	5	6
4	7	8
9	10	11

Отношение $R \times \rho_{S(x,c,D)}(S)$.

A	B	X	C	D
1	2	2	5	6
1	2	4	7	8
1	2	9	10	11
3	4	2	5	6
3	4	4	7	8
3	4	9	10	11

Зависимые и независимые операции

Шесть операций реляционной алгебры – *объединения, разности, выбора, проекции, декартова произведения и переименования* – являются независимыми.

Зависимые и независимые операции

Операции пересечения и соединения могут быть выражены через другие операции.

$$R \cap S = R \setminus (R \setminus S)$$

Операция тета-соединения

$$R \theta(C) S = \sigma_C(R \times S)$$

Реляционные операции над мультимножествами

Множество, допускающее наличие дубликатов элементов, принято называть мультимножеством (bag).

Реляционные операции над мультимножествами

Множество, допускающее наличие дубликатов элементов, принято называть мультимножеством (bag).

Результат - мультимножество

A	B
1	2
1	2
1	2
3	4
3	4

Итоговое мультимножество, несмотря на большой размер может быть получено значительно быстрее множества.

Если проекция выполняется с целью последующего вычисления, например, среднего арифметического атрибута A , то множество просто даст неверный результат.

Объединение, пересечение и разность мультимножеств

При объединении двух отношений-мультимножеств количество вхождений каждой строки складываются.

Пусть в R строка t содержится n раз, а в S — m раз, тогда в $R \cup S$ — ?

В $R \cap S$ строка t будет встречаться $\min(n, m)$,
а в $R \setminus S$ — $\max(0, n - m)$ раз.

Отношение R

A	B
1	2
1	2
1	2
3	4
3	4

Отношение S

A	B
1	2
5	6
3	4
3	4

Отношение $R \cup S$

A	B
1	2
1	2
1	2
1	2
3	4
3	4
3	4
3	4
5	6

Проекция мультимножеств

При проекции мультимножеств каждая строка обрабатывается независимо от других.

Алгебраические законы для мультимножеств

Закон коммутативности для объединения
множеств

$$R \cup S = S \cup R$$

Закон дистрибутивности для объединения
множеств

$$(R \cup S) \setminus T = (R \setminus T) \cup (S \setminus T)$$

Справедливы ли они для мультимножеств?

Выбор из мультимножеств

При выполнении операции выбора строк из отношения – мультимножества, условие выбора применяется к каждой строке независимо от других.

Отношение R

A	B	C
1	2	5
1	2	7
1	2	7
3	4	6

Результат операции выбора
 $\sigma_{C \geq 6}(R)$

A	B	C
1	2	7
1	2	7
3	4	6

Декартово произведение мультимножеств

При выполнении операции декартового произведения каждая строки одного отношения образует пару с каждой строкой другого отношения - независимо от того, повторяются строки или нет.

Операции соединения мультимножеств

Каждая строки одного отношения сопоставляется с каждой строкой другого отношения, и если строки способны образовывать пару, то такая пара составляется и включается в итоговое отношение.

Повторяющиеся строки из результата не удаляются.

Отношение R

A	B
1	2
1	2

Отношение S

B	C
2	3
4	5
4	5

Отношение $R \times S$

A	R.B	S.B	C
1	2	2	3
1	2	2	3
1	2	4	5
1	2	4	5
1	2	4	5
1	2	4	5

Отношение $R \bowtie S$

A	B	C
1	2	3
1	2	3

Рассмотрим операцию $R \ominus (R.B < S.B) S$.

A	R.B	S.B	C
1	2	4	5
1	2	4	5
1	2	4	5
1	2	4	5

Дополнительные операторы реляционной алгебры

1. Оператор удаления строк-дубликатов δ .
2. Операторы агрегирования.
3. Оператор группировки γ .
4. Оператор сортировки τ .
5. Оператор проекции π .
6. Оператор внешнего соединения.

Операторы агрегирования

1. SUM
2. AVG
3. MIN и MAX
4. COUNT

Вычисляют соответствующие выражения в определенном столбце.

Оператор группировки $\gamma_L(R)$

Отношение, возвращаемое оператором $\gamma_L(R)$, строится следующим образом:

1. Множество строк отношения R разбивается на группы. Каждая группа включает все строки, содержащие определенные значения группирующих атрибутов из L . Если группирующие атрибуты не заданы, то все отношение рассматривается как единая группа.

2. Для каждой группы создается по одной строке, состоящей из:
 - a. значений группирующих атрибутов, соответствующих группе;
 - b. результатов агрегирования, вычисленных по агрегируемым атрибутам всех строк группы, заданным в списке L.

Оператор расширенной проекции

В операторе расширенной проекции $\pi_L(R)$ список L может содержать элементы следующих типов:

- 1) отдельный атрибут отношения R ;
- 2) выражения вида $x \rightarrow y$, где x и y – наименование атрибутов; элемент $x \rightarrow y$ списка L предусматривает *переименование* атрибута x отношения R в y ;

3) выражения вида $E \rightarrow z$, где E - выражение, включающее наименование атрибутов отношения R , константы, арифметические и строковые операторы, а z - новое имя атрибута, получаемого в результате вычисления выражения E .

Отношение R

A	B	C
0	1	2
0	1	2
3	4	5

Результат операции $\pi_{A, B+C \rightarrow X}(R)$

A	X
0	3
0	3
3	9

Результат операции $\pi_{B-A \rightarrow X, C-B \rightarrow Y}(R)$

X	Y
1	1
1	1
1	1

Оператор сортировки

Часто встречаются ситуации, когда строки отношения необходимо сортировать по значениям одного или нескольких атрибутов.

Оператор сортировки $\tau_L(R)$.

R – некоторое отношение, а $L = A_1, A_2, \dots, A_n$ – список выбранных атрибутов R .

Внешние соединения

Оператор соединения может оставлять некоторые строки отношения висящими (dangling), если они не совпадают в атрибутах, общих для двух отношений, ни с одной строкой другого отношения.

Информация, содержащаяся в висячих строках никак не отражается в результирующем отношении. Чтобы учесть эти строки применяется специальная разновидность соединения – внешнее соединение (outer join).

Немного истории

В 1974 – 1975 годах на первом этапе выполнения проекта System/R был создан минимальный прототип реляционной системы управления базами данных. В это же время вышла первая статья с описанием языка запросов к базе данных, названного **SEQUEL – Structured English Query Language** (структурированный английский язык запросов).

В 1978 – 1979 годах System/R была установлена на компьютерах нескольких заказчиков фирмы IBM для опытной эксплуатации. Эта эксплуатация принесла первый опыт работы с реляционной СУБД и с языком базы данных, который был переименован в **SQL – Structured Query Language** (структурированный язык запросов). Несмотря на изменение названия, произношение «сиквел» сохранилось и по сегодняшний день.

В 1982 году началась работа над официальным стандартом SQL. В 1986 году стандарт был утвержден Американским институтом национальных стандартов (American National Standards Institute – ANSI), а в 1987 году – как стандарт Международной организации по стандартам (International Standards Organization – ISO).

Этот стандарт, пересмотренный в 1989 году, обычно называют стандартом «SQL-89» или «SQL1».

Многие из членов комитетов ANSI и ISO представляли различные фирмы – разработчик СУБД, в каждой из которых был реализован собственный диалект SQL.

Чтобы исправить эту ситуацию, комитет ANSI создал проект более жесткого стандарта «SQL2» и «SQL3».

SQL

Все существующие коммерческие продукты поддерживают, несмотря на стандарт SQL2, собственный диалект языка SQL. Однако ядро SQL стандартизировано достаточно жестко, и поставщики СУБД пытаются привести свои продукты в соответствие со стандартом.

Итак, *SQL* – структурированный язык команд для взаимодействия с сервером реляционных баз данных из любого инструментального пакета или приложения. Администратор базы данных использует язык запросов для формирования и обслуживания баз данных, а пользователь – для доступа к данным.

Возможности SQL

- Подходит разным пользователям, включая пользователей с самым малым опытом программирования или вообще без него
- Является непроцедурным языком
- Сокращает время, необходимое для создания и сопровождения систем

Основные команды SQL

Команды	Описание
SELECT	Извлекает данные из таблицы.
INSERT, UPDATE, DELETE	Команды языка манипулирования данными (DML). Предназначены для добавления, обновления и удаления строк из таблиц базы данных.
CREATE, ALTER, DROP	Команды языка определения данных (DDL). Используются для создания, изменения и удаления структур данных.
COMMIT, ROLLBACK, SAVEPOINT	Управляют изменениями, которые производятся помощью команд DML.

Правила ввода команд SQL

- Команда может занимать одну или несколько строк
- Для удобства чтения и редактирования можно использовать табуляцию и отступы
- Сокращения и перенос слов запрещены

Правила ввода команд SQL

- Ключевые слова и команды *рекомендуется* вводить в верхнем регистре, остальные слова – в нижнем
- Символы верхнего и нижнего регистров не различаются, если нет специальных указаний

Тип системы базы данных

- **OLTP** – Online Transaction Processing
(интерактивная обработка транзакций)
- **OLAP** - Online Analytical Processing
(интерактивная аналитическая обработка)

Проектирование и создание базы данных

System Databases

- master
- model
- msdb
- tempdb

master

База данных **master** лежит в основе SQL Server и содержит следующую основную информацию:

- все имена входа или роли, к которым принадлежат идентификаторы пользователя

- все системные параметры настройки (например, информация о сортировке данных, реализация обеспечения безопасности, язык по умолчанию)
- имена и информация о базах данных в сервере
- местоположение баз данных
- способ инициализации SQL Server

- специальные системные таблицы, содержащие в частности, следующее
 1. способ использования кэша
 2. доступные наборы символов
 3. список доступных языков

4. сообщения о системных ошибках и предупреждения
5. специальные объекты SQL Server, называемые *асSEMBЛИРОВАННЫМИ* программами

model – на ее основе строится пользовательская база данных

msdb – содержится необходимая информация для выполнения заданий под управлением SQL Server Agent (Служба Windows в SQL Server)

tempdb – временная база данных время жизни которой, продолжительность сеанса SQL Server

Создание базы данных

- CREATE DATABASE <database_name>
- [on
- ((NAME = logical_name,)
- FILENAME = physical_file_name
- [, FILESIZE = size]
- [, MAXSIZE = maxsize]
- [, FILEGROWTH = growth_increment]]]
- [LOG ON
- ((NAME = logical_name,)
- FILENAME = physical_file_name
- [, FILESIZE = size]
- [, MAXSIZE = maxsize]
- [, FILEGROWTH = growth_increment]]]
- [COLLATE collation_name]