Лекция 7

Введение в реляционную алгебру

Общее определение реляционной алгебры: универсальной алгеброй называется пара или кортеж <m, Ω>, где М – некоторое не пустое множество элементов, а Ω – некоторое непустое множество операций, определённых на м.

Множество М называется носителем алгебры.

Реляционная алгебра – RA (T, O) – множество T реляционных таблиц, в совокупности множества О алгебраических операций над ними.

Реляционная алгебра является замкнутой. (Все операции над этим объектом приводят к объекту того же типа). Реляционная алгебра является языком последовательного использования отношений, т.е. циклы в ней не предполагаются.

Операции реляционной алгебры. Предложил Кодд

8 операций первоначально, потом добавлены ещё.

Четыре типа классификаций:

1) Теоритико-множественные: Union, intersection, difference, Cartesian product

Характерные только для бд: selection, join, projection и division

2) Одноместные и двуместные

Одноместные: selection и projection

Двуместные: все остальные

3) С параметрами: join и projection

Без параметров: все остальные

4) Общие алгебраические и процедурно-доопределяемые

Общие алгебраические: Union, intersection, difference, Cartesian product и projection

Процедурно-доопределяемые: selection, join и division

Унарные операции

Выборка (selection) – операция выборки применяется к одному отношению R, и определяет результирующее отношение, которое содержит только те кортежи, которые удовлетворяют заданному условию (предикату).

R’= σ предиката (R)

Более сложные предикаты могут быть созданы с помощью логических операций.

Если в предикат входит только операция сравнения, то операция выборки называется ограничение.

Проекция (projection) – применяется к одному отношению R и определяет новые отношения, содержащие вертикальные подмножества отношения R, создаваемые посредством извлечения значения указанных атрибутов и исключения из результата строк дубликата.

R’=П (а…аn) (R)

Union (объединение) – объединение двух отношений R и S определяет новое отношение, которое включает все кортежи, содержащиеся только в R, все кортежи, содержащиеся только в S, а также одновременно в R и S, причём все дубликаты исключается, а отношение R и S должны быть совместимы по объединению.

Совместимость по объединению – объединение возможно только если схемы двух отношений совпадают, т.е. состоят из одинаково количества атрибутов, причём каждая пара соответствующих атрибутов имеют одинаковы домен. Не указывается, что атрибуты должны иметь одинаковые имена.

В некоторых случаях для получения совместимости по объединению можно использовать операцию проекции.

R’=R S

R=>I

S=> J

К(R) <= I+J

Разность (difference) двух отношений R и S состоит из кортежей, которые присутствуют в R, но отсутствуют в S, при этом R и S должны быть совместимы.

R’=R-S

Операция пересечения (intersection) определяет отношения, которые содержат кортежи, присутствующие одновременно в R и S.

R’= R S

Декартово произведение (Cartesian product) – операция определяет новое отношение, которое является результатом конкатенации (сцепления) каждого кортежа из R с каждым кортежем из S.

R’=R\*S

Количество атрибутов складывается, количество кортежей перемножается.

Исходные отношения может содержать атрибуты с одинаковыми именами. В этом случае к именам атрибутов в виде префиксов добавляются имена исходных отношений.

Декомпозиция сложных операций

Для упрощения операций РА можно разбить, т.е. поделить на некоторые промежуточные. Используются присваивание и переименование.

Присваивание - действие аналогично в языках программирования

Переименование – задаётся произвольное имя любому атрибуту нового отношения.

Операции соединения

Как правило, пользователя интересуют только некоторая часть из кортежа декартового произведения. В этом случае используется операция соединения.

С точки зрения эффективности реализации в СУБД операция соединения является одной из самых сложных и вызывает наибольшие проблемы с производительностью.

Классификация различных соединений:

1) Соединение = декартово произведение всех имеющихся в R и S кортежей+ выборка некоторых кортежей по какому-то условию. И в зависимости от условий выделяются:

i) Theta join – тета соединение, в том числе соединение по эквивалентности (внутреннее соединение) и естественное соединение (natural join)

ii) Outter join - левое внешнее, правое внешнее и полное внешнее

iii) Semi-join (полусоединение) – тета полусоединение, полусоединение по эквивалентности и естественное полусоединение

Theta join – операция определяет отношение, которое содержит кортежи из декартово произведения R на S, удовлетворяющему предикату F

F=Rai ᶿ Sbj

ᶿ => (>, <, ≤, ≥, ∞)

Соединение по эквивалентности – если предикат F содержит только операции сравнения по равенству, то оно так называется

Естественное соединение – ещё одно соединение по эквивалентности двух отношений R и S, выполненное по всем общим атрибутам, из результатов которого исключаются по одному экземпляру каждого общего атрибута.

Степенью естественного соединения называется сумма степеней операндов отношений R и S за вычетом количества общих атрибутов.

Внешнее соединение

При соединении двух отношений часто возникает такая ситуация, что для кортежа одного отношения не находится соответствующий кортеж в другом отношении.

Но иногда может потребоваться, чтобы строка из одного отношения была представлена в результатах соединения, даже если в другом отношении нет совпадающего отношения. Достигается с помощью внешнего соединения.

Для обозначения отсутствующих отношений во втором отношении используется значения null.

Левое внешнее соединение – это такое соединение, при котором в результирующее отношение включаются также кортежи отношения R, не имеющие совпадающих отношений в общих столбцах отношения S. Правое – в обратную сторону.

Полное - внешнее соединение, в результирующее отношение которого помещаются все кортежи обоих отношений, при этом для обозначения несовпадающих используется null.

Полусоединение – операция полусоединения определяет отношения, содержащие те кортежи отношения R, которые входят в соединение отношений R и S.

Эта заготовка для операции соединения, особенно важная в распределённых системах.

Обобщение = содержательная классификация различных видов соединения.

ᶿ-соединение – соединение по условию

Условие = предикат ᶿ

R> <FS =σ F (R\*S)

1.1=1, где ᶿ - только «R.ключ=S.ключ»

1.2=1.1, где убраны повторы

2.1=1.1+добавлены кортежи, “где R.ключ=null.”

2.2= “Null=S.ключ»

2.3=2.1+2.2

3 выделяются те кортежи отношения R, которые входят в данный тип соединения R и S.

3 существует для пунктов 1, 1.1 и 1.2