

КОНСПЕКТ ЗАНЯТИЯ № 1 ТРЕТЬЕЙ НЕДЕЛИ КУРСА «БАЗЫ ДАННЫХ»

1. ОСНОВНЫЕ ОПЕРАЦИИ РЕЛЯЦИОННОЙ АЛГЕБРЫ

На предыдущем занятии упоминались три составляющих реляционной модели данных: *структурная* часть, *целостная* часть и *манипуляционная* часть. Две из них – структурную и целостную части – мы рассмотрели. В этой лекции пойдет разговор о **манипуляционной** части реляционной модели данных.

Как мы уже отмечали, Эдгар Кодд при разработке своей модели определил для манипуляционной составляющей два базовых механизма манипулирования реляционными данными: – основанная на теории множеств *реляционная алгебра* и базирующееся на математической логике (точнее, на исчислении предикатов) *реляционное исчисление*.

Эти два метода эквивалентны друг другу по выразительной силе, т.е. для любого допустимого *выражения* реляционной алгебры можно построить эквивалентную (т. е. производящую такой же результат) *формулу* реляционного исчисления и наоборот. Но различия между этими методами всё же есть: рассмотрим каждый механизм более конкретно.

Реляционная алгебра (РА) основана на теории множеств, с её помощью мы можем описать КАК получить желаемое отношение, перечислив в нужном порядке определённые операции. Например, если потребуется узнать названия отделов, в которых работают сотрудники по фамилии Васильчиков, перечислим следующие операции:

- 1) выборка из отношения с информацией о сотрудниках всех кортежей, в которых значение атрибута “Фамилия” - Васильчиков;
- 2) затем соединение полученного отношения и отношения со сведениями об отделах по совпадению номеров отделов;
- 3) затем усечение или проекция результирующего отношения по одному атрибуту “Название отдела”.

Другими словами, запрос, представленный на языке реляционной алгебры, может быть вычислен на основе выполнения элементарных алгебраических операций с учетом их приоритетности.

Реляционное исчисление (РИ) основано на теории исчисления предикатов, с его помощью можно указать, ЧТО мы хотим получить, описав условия, которым должно удовлетворять желаемое отношения:

например, получить значение атрибута “Номер отдела” таких отделов, для которых в отношении с информацией о сотрудниках существуют кортежи с таким же значением атрибута “Номер отдела”, что и у рассматриваемых кортежей из первого отношения и плюс со значением атрибута “Фамилия” - Васильчиков. Как именно получить результат в этом случае нас не интересует.

Другими словами, формула только ставит условия, которым должны удовлетворять кортежи результирующего отношения.

В рамках курса мы рассмотрим методы манипулирования данными с помощью реляционной алгебры.

Реляционная алгебра - это набор операций, обладающих свойством замкнутости.

Свойство замкнутости: на вход любой операции подаются отношения, на выходе получаем – отношение. Другими словами, результатом любой операции над отношением является отношение:

$$R = f(R_1, R_2, \dots, R_n) \rightarrow R = f(f_1(R_{11}, R_{12}, \dots), f_2(R_{21}, R_{22}, \dots))$$

Это означает, что выражения реляционной алгебры (и формулы реляционного исчисления) определяются над отношениями реляционных БД и результатом их «вычисления» также являются отношения. В результате любое выражение (или формула) может интерпретироваться как отношение, что позволяет использовать его в других выражениях (или формулах).

Второе свойство, которое должны поддерживать отношения (когда речь идет о реляционной алгебре) – **свойство совместимости**. Отношения называются **совместимыми по типу**, если они имеют идентичные заголовки, а именно:

- для каждого атрибута в одном отношении должен найтись одноимённый в другом, т.е. два отношения имеют *одно и то же множество имен атрибутов*;
- атрибуты с одинаковыми именами *определены на одних и тех же доменах*.

Основная идея реляционной алгебры состоит в том, что раз отношения являются множествами, то средства манипулирования отношениями могут базироваться на традиционных теоретико-

множественных операциях, дополненных некоторыми специальными операциями, специфичными для реляционных баз данных.

Реляционная алгебра Кодда представляет собой набор основных алгебраических операций (если точнее – восьми операций), которые делятся на два класса – *теоретико-множественные операции* и *специальные реляционные операции*.

В состав теоретико-множественных операций входят операции:

- ✓ декартово произведение множеств;
- ✓ объединение множеств;
- ✓ пересечения множеств;
- ✓ разность множеств.

Специальные реляционные операции включают:

- ✓ проекцию отношения;
- ✓ соединение отношений;
- ✓ деление отношений;
- ✓ выборка отношения.

Также отдельно выделяют две дополнительные операции: присваивания и переименования.

Приоритет операций РА показан в таблице 4.1. Для операций, имеющих одинаковый приоритет, вычисление производится слева направо. Порядок вычисления по умолчанию можно переопределить с помощью скобок.

Таблица 4.1. Приоритет операций реляционной алгебры

Операция	Приоритет
Переименование	1
Проекция	2
Выборка	
Декартово произведение	3
Пересечение	
Соединение	
Деление	
Объединение	4
Разность	
Присваивание	5

Рассмотрим для начала определения дополнительного набора операций: присваивания и переименования.

Присваивание - позволяет сохранить в базе данных результаты вычисления алгебраических выражений (т.е. позволяет задать имя отношению, полученному в результате выполнения некоторой операции).

$$T = F \langle \text{операция} \rangle S$$

Переименование - позволяет создать новое отношение, тело которого совпадает с телом исходного, но имена атрибутов изменены.

$$A(A_1, A_2) \text{ RENAME } B(A_1, B_1),$$

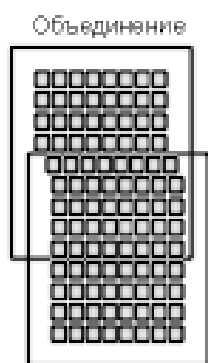
где A – исходное имя отношения, B – новое.

В данном примере отношение с именем A , состоящее из двух атрибутов: A_1 и A_2 , было переименовано в B . Одновременно имя второго атрибута этого отношения - A_2 - было изменено на B_1 , а имя первого оставлено прежним.

Теоретико-множественные операции реляционной алгебры

Начнем с операции *объединение* отношений.

При выполнении операции *объединения* (UNION) двух отношений с одинаковыми заголовками производится отношение, включающее все кортежи, которые входят хотя бы в одно из отношений-операндов.



Иначе говоря: *объединением* двух совместимых по типу отношений:

F , состоящего из атрибутов $\{A_1, A_2, \dots, A_n\}$ и

S , состоящего из атрибутов $\{A_1, A_2, \dots, A_n\}$

называется отношение R с той же схемой¹, что и у отношений F и S и телом¹, состоящим из кортежей¹, принадлежащих либо F , либо S , либо обоим сразу.

Варианты синтаксиса:

$$R = F \text{ UNION } S$$

$$R = F \cup S$$

Если два отношения совместимы по типу, то при выполнении над ними операции объединения результатом будет отношение с корректно определенным заголовком, совпадающим с заголовком каждого из

¹ Все понятия отмеченные сноской ранее были определены в конспекте второй недели.

отношений-операндов. В случае, если атрибуты отличны друг от друга, то перед выполнением объединения необходимо применить операцию переименования.

При проведении операций, требующих совместимости отношений по типу, степень получившегося отношения всегда равна степени исходных.

Если в исходных отношениях нет одинаковых кортежей кардинальность результата будет равна сумме кардинальностей объединяемых отношений (рис. 4.3).

Отношение F: степень = 4, кардинальность = 3				Отношение R = F ∪ S: степень = 4, кардинальность = 3+2 = 5 (в исходных отношениях нет одинаковых кортежей)			
A ₁	A ₂	A ₃	A ₄				
f ₁₁	f ₂₁	f ₃₁	f ₄₁	f ₁₁	f ₂₁	f ₃₁	f ₄₁
f ₁₂	f ₂₂	f ₃₂	f ₄₂	f ₁₂	f ₂₂	f ₃₂	f ₄₂
f ₁₃	f ₂₃	f ₃₃	f ₄₃	f ₁₃	f ₂₃	f ₃₃	f ₄₃
Отношение S: степень = 4, кардинальность = 2				s ₁₁	s ₂₁	s ₃₁	s ₄₁
A ₁	A ₂	A ₃	A ₄	s ₁₂	s ₂₂	s ₃₂	s ₄₂
s ₁₁	s ₂₁	s ₃₁	s ₄₁				
s ₁₂	s ₂₂	s ₃₂	s ₄₂				

Рис. 4.3. Объединение отношений, не имеющих общих кортежей

В ином случае - кардинальность результата будет меньше из-за необходимости удалить возникшие дубликаты (рис. 4.4)

Отношение F: степень = 4, кардинальность = 3				Отношение R = F ∪ S: степень = 4, кардинальность = 3+2-1 = 4 (один из кортежей присутствовал в обоих исходных отношениях)			
A ₁	A ₂	A ₃	A ₄				
a ₁₁	a ₂₁	a ₃₁	a ₄₁	a ₁₁	a ₂₁	a ₃₁	a ₄₁
f ₁₂	f ₂₂	f ₃₂	f ₄₂	f ₁₂	f ₂₂	f ₃₂	f ₄₂
f ₁₃	f ₂₃	f ₃₃	f ₄₃	f ₁₃	f ₂₃	f ₃₃	f ₄₃
Отношение S: степень = 4, кардинальность = 2				a₁₁	a₂₁	a₃₁	a₄₁
A ₁	A ₂	A ₃	A ₄	s ₁₂	s ₂₂	s ₃₂	s ₄₂
a ₁₁	a ₂₁	a ₃₁	a ₄₁				
s ₁₂	s ₂₂	s ₃₂	s ₄₂				

Рис. 4.4. Объединение отношений, имеющих общие кортежи

При объединении двух отношений, содержащих данные о студентах из двух групп, обучающихся по разным направлениям, мы получим общий список студентов. При этом студент Савицкий, входящий в состав обеих групп, в результат попадёт только один раз. Фамилия Иванова

встречается в результате дважды, т.к. принадлежит разным студенткам - с разными номерами зачётных книжек (рис. 4.5).

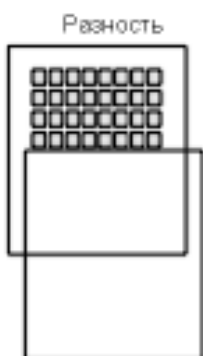
Отношение Студент 1	
Номер зачётной книжки	Фамилия
123456	Савицкий
123457	Иванова
123458	Паулайнен

Отношение Студент 2	
Номер зачётной книжки	Фамилия
123456	Савицкий
123463	Иванова
123462	Кленов

Отношение R = Студент1 ∪ Студент2	
Номер зачётной книжки	Фамилия
123456	Савицкий
123457	Иванова
123458	Паулайнен
123463	Иванова
123462	Кленов

Рис. 4.5. Пример объединения отношений «Студент1» и «Студент2»

Следующая рассматриваемая операция – *разность*. Отношение, являющееся *разностью* (MINUS) двух отношений с одинаковыми заголовками, включает все кортежи, входящие в отношение первого операнда, такие, что ни один из них не входит в отношение, которое является вторым операндом.



Иначе говоря, *разностью* двух совместимых по типу отношений

F, состоящего из атрибутов $\{A_1, A_2, \dots, A_n\}$ и

S, состоящего из атрибутов $\{A_1, A_2, \dots, A_n\}$

называется отношение R с той же схемой, что и у отношений F и S и телом, состоящим из кортежей, принадлежащих отношению-уменьшаемому F и не принадлежащих отношению-вычитаемому S.

Варианты синтаксиса:

$R = F \text{ MINUS } S$

$R = F \setminus S$

Если в исходных отношениях нет одинаковых кортежей, то вычитать нечего и кардинальность результата будет равна кардинальности отношения-уменьшаемого (рис. 4.6).

Отношение F:
 степень = 4, кардинальность = 3

A ₁	A ₂	A ₃	A ₄
f ₁₁	f ₂₁	f ₃₁	f ₄₁
f ₁₂	f ₂₂	f ₃₂	f ₄₂
f ₁₃	f ₂₃	f ₃₃	f ₄₃

Отношение S:
 степень = 4, кардинальность = 2

A ₁	A ₂	A ₃	A ₄
s ₁₁	s ₂₁	s ₃₁	s ₄₁
s ₁₂	s ₂₂	s ₃₂	s ₄₂

Отношение R = F \ S:
 степень = 4, кардинальность = 3-0 = 3
 (в исходных отношениях нет одинаковых кортежей)

A ₁	A ₂	A ₃	A ₄
f ₁₁	f ₂₁	f ₃₁	f ₄₁
f ₁₂	f ₂₂	f ₃₂	f ₄₂
f ₁₃	f ₂₃	f ₃₃	f ₄₃

Рис. 4.6. Разность отношений, не имеющих общих кортежей

В ином случае (когда есть что вычитать) - будет равна разности между кардинальностью отношения-уменьшаемого и количеством общих для исходных отношений кортежей (рис. 4.7).

Отношение F:
 степень = 4, кардинальность = 3

A ₁	A ₂	A ₃	A ₄
a ₁₁	a ₂₁	a ₃₁	a ₄₁
f ₁₂	f ₂₂	f ₃₂	f ₄₂
f ₁₃	f ₂₃	f ₃₃	f ₄₃

Отношение S:
 степень = 4, кардинальность = 2

A ₁	A ₂	A ₃	A ₄
a ₁₁	a ₂₁	a ₃₁	a ₄₁
s ₁₂	s ₂₂	s ₃₂	s ₄₂

Отношение R = F \ S:
 степень = 4, кардинальность = 3-1 = 2
 (один из кортежей присутствовал в обоих исходных отношениях)

A ₁	A ₂	A ₃	A ₄
a₁₁	a₂₁	a₃₁	a₄₁
f ₁₂	f ₂₂	f ₃₂	f ₄₂
f ₁₃	f ₂₃	f ₃₃	f ₄₃

Рис. 4.7. Разность отношений, имеющих общие кортежи

Если мы из отношения, содержащего информацию о всех студентах группы, вычтем отношение, содержащее информацию о студентах, получивших хотя бы одну тройку на сессии, то в результате получим отношение со сведениями о студентах, которым будет начислена стипендия, т.к. по результатам сессии они не получили ни одной тройки (рис. 4.8).

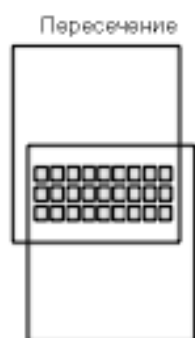
Отношение Студент	
Номер зачётной книжки	Фамилия
123456	Савицкий
123457	Иванова
123458	Паулайнен

Отношение Студент-троечник	
Номер зачётной книжки	Фамилия
123456	Савицкий
123457	Иванова

Отношение R = Студент \ Студент-троечник	
Номер зачётной книжки	Фамилия
123458	Паулайнен

Рис. 4.8. Пример разности отношений «Студент» и «Студент-троечник»

Операция *пересечения* (INTERSECT) двух отношений с одинаковыми заголовками производит отношение, включающее все кортежи, которые входят в оба отношения-операнда.



Иными словами, *пересечением* двух совместимых по типу отношений

F , состоящего из атрибутов $\{A_1, A_2, \dots, A_n\}$ и

S , состоящего из атрибутов $\{A_1, A_2, \dots, A_n\}$

называется отношение R с той же схемой, что и у отношений F и S и телом, состоящим из кортежей, принадлежащих обоим отношениям одновременно.

Варианты синтаксиса:

$R = F \text{ INTERSECT } S$

$R = F \cap S$

Если кортежи одного из исходных отношений полностью входят в состав второго, кардинальность результата будет равна кардинальности меньшего отношения (рис. 4.9).

Отношение F:			
степень = 4, кардинальность = 3			
A ₁	A ₂	A ₃	A ₄
a ₁₁	a ₂₁	a ₃₁	a ₄₁
a ₁₂	a ₂₂	a ₃₂	a ₄₂
f ₁₃	f ₂₃	f ₃₃	f ₄₃

Отношение S:			
степень = 4, кардинальность = 2			
A ₁	A ₂	A ₃	A ₄
a ₁₁	a ₂₁	a ₃₁	a ₄₁
a ₁₂	a ₂₂	a ₃₂	a ₄₂

Отношение R = F ∩ S:			
степень = 4, кардинальность = 2			
(все кортежи второго отношения входят в состав первого)			
A ₁	A ₂	A ₃	A ₄
a ₁₁	a ₂₁	a ₃₁	a ₄₁
a ₁₂	a ₂₂	a ₃₂	a ₄₂

Рис. 4.9. Пересечение отношений,

Выполнив пересечение отношений с информацией о студентах, которым нужно общежитие и студентах-контрактниках, мы получим отношение, содержащее информацию о студентах, обладающих обоими признаками сразу, т.е. контрактниках, нуждающихся в общежитии (рис. 4.10).

Отношение Нуждающийся в общежитии

Номер зачётной книжки	Фамилия
123456	Савицкий
123457	Иванова

Отношение Контрактник

Номер зачётной книжки	Фамилия
123457	Иванова
123458	Паулайнен

Отношение $R = \text{Нуждающийся в общежитии} \cap \text{Контрактник}$

Номер зачётной книжки	Фамилия
123457	Иванова

Рис. 4.10. Пример пересечения отношений
«Нуждающийся в общежитии» и «Контрактник»

Декартовым произведением этих отношений называется отношение R :

- схема которого является результатом сцепления схем исходных отношений: $\{F_1, F_2, \dots, F_n, S_1, S_2, \dots, S_m\}$,
- а тело состоит из множества кортежей $\{f_1, f_2, \dots, f_n, s_1, s_2, \dots, s_m\}$, таких, что $(f_1, f_2, \dots, f_n) \in F, (s_1, s_2, \dots, s_m) \in S$.

Варианты синтаксиса:

$R = F \times S$

$R = F \text{ TIMES } S$

Кардинальность, т.е. количество кортежей получившегося отношения, будет равна произведению кардинальностей исходных, а степень, т.е. количество атрибутов - всегда равна сумме степеней исходных (рис. 4.1).

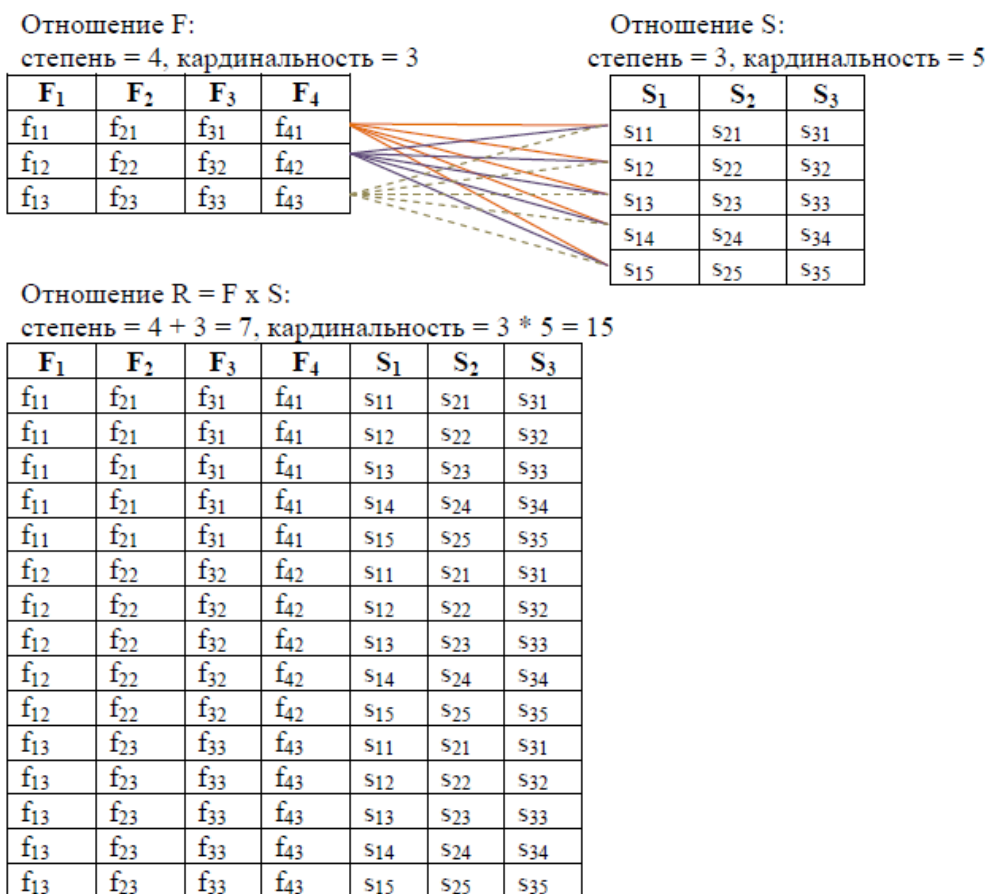


Рис. 4.1. Декартово произведение отношений в общем виде

Если в исходных отношениях встретились атрибуты с одинаковыми именами, в новом отношении возникнет проблема дублирования названий атрибутов, для решения которой предлагается два метода на выбор:

1. использовать префиксы в виде имён исходных отношений (см. пример на рис. 4.2);
2. произвести переименование атрибутов в одном из отношений.

С целью упрощения записи выражений в дальнейшем договоримся вместо переименования атрибутов использовать их полные имена в формате <имя отношения>.<имя атрибута>.

Следует заметить, что операция декартова произведения не является слишком осмысленной на практике. Во-первых, мощность тела ее результата очень велика, а во-вторых, результат операции не более информативен, чем взятые в совокупности операнды. Однако на основе нее определяется действительно полезная операция – *соединения*.

Однако порой можно задать ситуацию, когда декартово произведение может быть полезно. Например, если в одном отношении содержится информация о студентах группы, а в другом - об учебниках, которые понадобятся группе в этом семестре, то в результате декартова

произведения этих отношений мы получим список учебников для каждого студента из группы (рис. 4.2). Обратите внимание, на префикс в виде имени исходного отношения для различения атрибута «Фамилия» у студента и автора учебника.

Отношение Студент		Отношение Учебник	
Номер зачётной книжки	Фамилия	Фамилия	Название
123456	Савицкий	К.Дж. Дейт	Введение в системы баз данных
123457	Комарова	Н. Вирт	Алгоритмы и структуры данных
123458	Паулайнен		

Отношение Ведомость выдачи учебников = Студент x Учебник

Номер зачётной книжки	Студент.Фамилия	Учебник.Фамилия	Название
123456	Савицкий	К.Дж. Дейт	Введение в системы баз данных
123456	Савицкий	Н. Вирт	Алгоритмы и структуры данных
123457	Комарова	К.Дж. Дейт	Введение в системы баз данных
123457	Комарова	Н. Вирт	Алгоритмы и структуры данных
123458	Паулайнен	К.Дж. Дейт	Введение в системы баз данных
123458	Паулайнен	Н. Вирт	Алгоритмы и структуры данных

Рис. 4.2. Пример декартова произведения отношений «Студент» и «Учебник»