

Лекция №5

Отношение – представляет собой подмножество декартова произведения доменов.

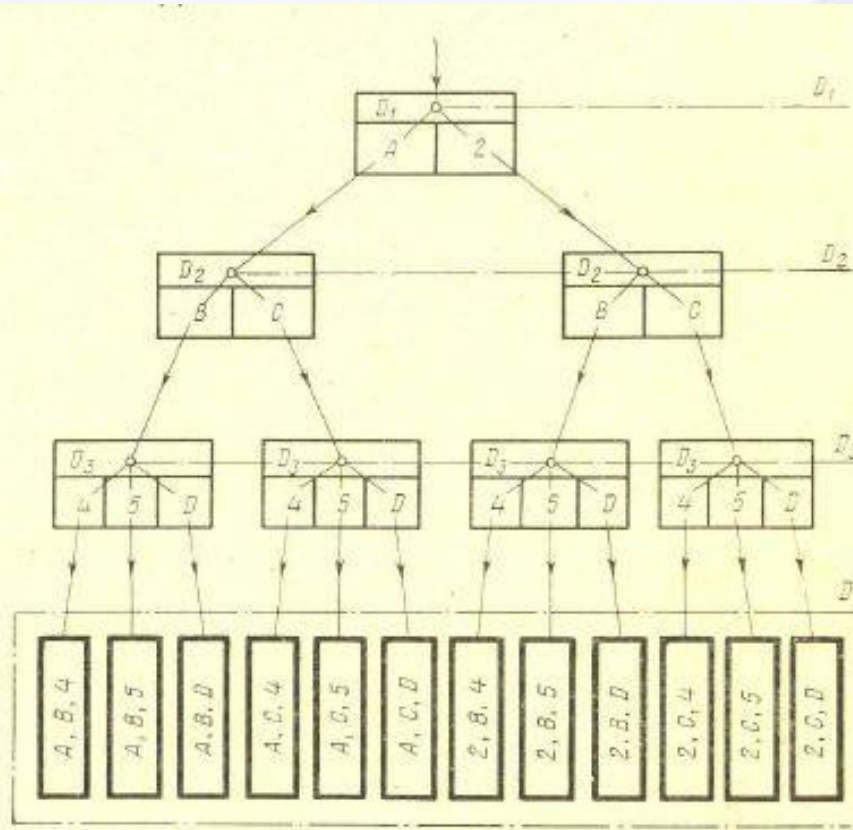
Домен – это некоторое множество элементов или допустимых значений, которое может принимать объект по некоторому свойству.

Декартовым произведением множеств D_1, D_2, \dots, D_n

где $D = D_1 + D_2 + \dots + D_n$ называется множество всех **кортежей** длины n , т.е. состоящих из n элементов – по одному из каждого множества D_i на i -м месте.

$$D = \left\{ \begin{array}{l} d_{1,1}, d_{1,2}, \dots, d_{1,k-1}, d_{1,k} \\ d_{2,1}, d_{2,2}, \dots, d_{2,k-1}, d_{2,k} \\ \vdots \\ d_{k,1}, d_{k,2}, \dots, d_{k,k-1}, d_{k,k} \end{array} \right\}$$

Декартово произведение позволяет получить все возможные комбинации элементов исходных множеств — элементов рассматриваемых



Отношением R – на множествах D_1, D_2, \dots, D_n
называется подмножество декартова
произведения $D_1 \times D_2 \times \dots \times D_n$

Отношение R , определенное на множествах
 D_1, D_2, \dots, D_n
есть (некоторое) множество кортежей n -ности
 $(a_{1,i_1}, a_{2,i_2}, \dots, a_{n,i_n})$, таких что $a_{1,i_1} \in D_1$
 $a_{2,i_2} \in D_2$, $a_{n,i_n} \in D_n$ принадлежит,
и т.д.:

$$R \subseteq D_1 \times D_2 \times \dots \times D_n$$

Элементами отношений являются кортежи.

Арность кортежа определяется **арностью отношения**. Отношение арностью 1 называют – унарными, арности 2 – бинарными, арности 3 – тернарными, арности n – n -арными.

Замечание: поскольку отношение – это множество, то:

1. не должны встречаться одинаковые кортежи;
2. порядок кортежей в отношении не существен.

Отношения удобно представлять через **таблицу**, где каждая строка это – **кортеж**, а каждый столбец соответствует одному и тому же компоненту декартова произведения, т.е. в нем могут появляться только элементы из соответствующего домена.

Таблица представляющая n -арное отношение, обладает следующими свойствами:

1. каждая строка представляет собой кортеж из n значений, принадлежащих n столбцам;
2. порядок столбцов фиксирован: $(1, 2, \dots, n)$;
3. порядок строк безразличен;
4. любые две строки различаются хотя бы одним элементом;
5. строки и столбцы могут обрабатываться в любой последовательности, определяемой применяемыми операциями обработки.

При определении теоретико-множественного отношения предполагает работу с **линейными списками** при обработке данных. Такая форма удобна для операций реляционной алгебры.

Однако она не всегда целесообразна из-за фиксированного порядка столбцов в отношении, т.к. в ряде приложений возникает необходимость перестановки столбцов в отношении в любом порядке.

Необходимости фиксированного порядка столбцов в отношении столбцы именуют что делает их порядок в отношении не существенным .

Столбцы отношения, с присвоенными им именами, называют **атрибутами**.

Список имен атрибутов отношения называют **схемой отношения**.

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

Существует аналогия между схемой отношения и форматом записи, между кортежем и записью, между отношением и файлом.

Следствием этого является возможная реализация отношения в виде файла записей, формат которых соответствует схеме отношения.

Реляционная база данных – это набор экземпляров конечных отношений.

Схема реляционной БД представляется в виде совокупности схем отношений

$$\left\{ \begin{array}{l} R_1(A_{1.1}, A_{1.2}, \dots, A_{1.k_1}), \\ R_2(A_{2.1}, A_{2.2}, \dots, A_{2.k_2}), \\ \vdots \\ R_m(A_{m.1}, A_{m.2}, \dots, A_{m.k_m}). \end{array} \right.$$

Нормализация

Нормализация – процесс реорганизации данных путем ликвидации повторяющихся групп и иных противоречий в хранении данных с целью приведения таблиц к виду позволяющему осуществлять непротиворечивое и корректное редактирование данных.

1ая нормальная форма

Для того чтобы таблица соответствовала 1НФ все значения ее полей должны быть **атомарными** а записи **уникальными**.

Значения доменов рассматриваются как не делимые, а не как множества или кортежи из более элементарных доменов.

Покупатель	Покупка	Дата покупки	Сумма	ТелПокупателя
ООО "Пупкин"	Иванов "Микрософт Офис" - 2 экз	05.01.2010	400	111-11-11
"Пупкин" ООО	Иванов "Микрософт Офис" - 1 экз Дерк "Справочник по PHP" - 1 экз Дерк "Справочник по JScript" - 1 экз	06.01.2010	800	111-11-11
Петрова	Донцова "Сборник уличных романов" - 1 экз	30.12.2009	50	222-22-22

Покупатель	Покупка	Автор	Колво	Цена	Дата покупки	Сумма	ТелПокупателя
ООО "Пупкин"	"Микрософт Офис"	Иванов	2	200	05.01.2010	400	111-11-11
"Пупкин" ООО	"Микрософт Офис"	Иванов	1	200	06.01.2010	200	111-11-11
"Пупкин" ООО	"Справочник по PHP"	Дерк	1	300	06.01.2010	300	111-11-11
"Пупкин" ООО	"Справочник по JScript"	Дерк	1	300	06.01.2010	300	111-11-11
Петрова	"Сборник уличных романов"	Донцова	1	50	30.12.2009	50	222-22-22

2ая нормальная форма

Таблица находится во 2НФ, если она находится в 1НФ и ее не ключевые поля полностью зависят от всего первичного ключа, не являясь его частью.

Покупатель	ТелПокупателя
ООО "Пупкин"	111-11-11
"Пупкин" ООО	111-11-11
Петрова	222-22-22

Автор
Иванов
Дерк
Донцова

Книга	Цена	ISBN
Микрософт Офис	200	1-111-11111-1
Справочник по PHP	300	2-222-22222-2
Справочник по JScript	300	3-333-33333-3
Сборник уличных романов	50	4-444-44444-4

таб. Покупатели

ID_Customer	Покупатель	ТелПокупателя
1	ООО "Пупкин"	111-11-11
2	Петрова	222-22-22

таб. Авторы

ID_Autor	Автор
1	Иванов
2	Дерк
3	Донцова

таб. Книги

ID_Book	AutorID	Книга	Цена	ISBN
1	1	Микрософт Офис	200	1-111-11111-1
2	2	Справочник по PHP	300	2-222-22222-2
3	2	Справочник по JScript	300	3-333-33333-3
4	3	Убить лысого	50	4-444-44444-4

таб. Orders

ID_Order	CustomerID	Дата покупки	Итого
1	1	05.01.2010	400
2	1	06.01.2010	800
3	2	30.12.2009	50

таб. OrdersDetail

OrderID	BookID	Колво	Сумма
1	1	2	400
2	1	1	200
2	2	1	300
2	3	1	300
3	4	1	50

Зья нормальная форма

Таблица находится в 3НФ, если она находится во 2 НФ и ни одно из неключевых полей не определяет другое не ключевое поле.

Отношение имеет только один ключ и другие зависимости, в том числе многозначные

Мы видим, что в таблице OrdersDetail поле Сумма зависит от поля Колво. Аналогичное расчетное поле - Итого таблицы Orders. Убираем их. Наша таблица продаж теперь выглядит следующим образом:

таб. Orders

ID_Order	CustomerID	Дата покупки
1	1	05.01.2010
2	1	06.01.2010
3	2	30.12.2009

таб. OrdersDetail

OrderID	BookID	Колво
1	1	2
2	1	1
2	2	1
2	3	1
3	4	1

Нормальная форма Бойса-Кода

Детерминант a – это атрибут (или комбинация атрибутов) от которого другой атрибут зависит функционально (полно).

В этом случае отношение R находится в БКНФ, если каждая детерминанта является ВОЗМОЖНЫМ КЛЮЧОМ.

Существует отношение:

(Номер зачетной книжки, идентификатор студента, дисциплина, дата, оценка)

БКНФ:

(Идентификатор студента, дисциплина, дата, оценка)

(Номер зачетной книжки, идентификатор студента)

Или

(Номер зачетной книжки, дисциплина, дата, оценка)

(Номер зачетной книжки, идентификатор студента)

4ая нормальная форма

4НФ запрещает хранить независимые компоненты в одной таблице, когда между этими компонентами существуют отношение многие-ко-многим.

**Существует отношение:
(Номер зачетной книжки, группа, дисциплина)**

**4НФ:
(группа - дисциплина)
(группа – номер зачетной книжки)**

Спасибо за внимание!!!