Лекция 12. Нормальные формы высших порядков. Нормальная форма Бойса-Кодда, 4-я, 5-я нормальные формы.

12.1 Нормальная форма Бойса –Кодда........1
12.2 MV-зависимости и 4-я нормальная форма........2
12.3 Зависимости соединения и 5НФ........3

В лекции 10 мы познакомились с этапами нормализации схем реляционных отношений до 3-ей нормальной формы. Однако, между атрибутами схем могут существовать гораздо более неочевидные зависимости.
Рассмотрим наиболее известные из таких зависимостей, и соответствующие им нормальные формы.

12.1 Нормальная форма Бойса –Кодда

При определении 3НФ было сделано допущение, что ни один непервичный атрибут не является транзитивно зависимым от первичного ключа, причем предполагалось, что такой ключ – всего один.
Пусть отношение имеет несколько потенциальных ключей, что чаще всего и происходит на практике.
Тогда

> **Схема отношения *R* находится в нормальной форме Бойса –Кодда** (НФБК), тогда и только тогда когда каждая *неприводимая слева* F-зависимость имеет левую часть, равную некоторому потенциальному ключу.
> Т.е. все непервичные атрибуты полностью зависят от **всех** потенциальных ключей схемы *R*.

> **Неприводимая слева** F-зависимость – минимальна в том смысле, что убрав один из атрибутов в левой части , зависимость выполняться не будет.

НФБК – более сильная, по сравнению с 3НФ.
Однако, отношение, находящееся в НФБК, также может быть избыточным.

Пример1.
Рассмотрим отношение r1.

Каждый курс может прочесть любой преподаватель соответствующей группы, с использованием всех указанных для этого курса учебников.
Предполагается, что независимо от того, кто читает курс, используется один и тот же набор учебников. Допускается, что один и тот же преподаватель читает разные курсы, и один учебник может использоваться на разных курсах.
После приведения этого отношения в 1НФ, несмотря на то, что r1 формально будет в НФБК, очевидна избыточность, повторение данных.
Способ разрешения проблемы показан на отношениях r2 и r3. Отношения формально находятся в НФБК, отношение r3 «унаследовало» избыточность данных. Теперь при добавлении нового преподавателя на курс в отношение r3 нужно добавить несколько кортежей, по числу учебников для этого курса.

Использование только F-зависимостей и процедуры нормализации до 3НФ не дает решения проблемы, т.к. отношения, находящиеся в НФБК, имеют только «правильные» функциональные зависимости.

Пример 2.
Интуитивно понятно, что *r3* удобно разбить на два отношения, *r4* и *r5*, т.е. разделить все повторяющиеся группы.

Декомпозицию такого отношения можно осуществить на основе **многозначной зависимости** (Multi-Valued dependency, MV - зависимости)

Понятие MV – зависимости является обобщением понятия F-зависимости.

12.2 MV-зависимости и 4-я нормальная форма

Пусть *R* – схема отношения *r*, множества $A \subseteq R, B \subseteq R$ – подмножества атрибутов схемы, $A \cap B = \varnothing$, множество $C = R \setminus (A \cup B)$.

> Отношение *r(R)* удовлетворяет **многозначной зависимости** *A->>B* тогда и только тогда когда для любого значения атрибута *A* существует **множество** (т.е. несколько) значений атрибута *B*.

Другими словами,пусть в отношении *r* со схемой *R*(*ABC*) кортежи *t* имеют вид *<a,b,c>*.

> Тогда, отношение *r(R)* удовлетворяет MV-зависимости *A->>B* только в таком случае:
> если в отношении есть кортежи *<a1,b1,c1>* и *<a1,b2,c2>*, то обязательно должны быть и кортежи *<a1,b2,c1>* и *<a1,b1,c2>*.

MV-зависимости всегда образуют пары, т.к. MV-зависимость *A->>B* выполняется тогда и только тогда когда выполняется MV-зависимость *A->>C* (в силу симметрии при переобозначении), поэтому обычно MV-зависимости записываются так: *A->>B/C*.

Рассмотрим свойства MV-зависимостей .

Для любых подмножеств $A \subseteq R, B \subseteq R, A \cap B = \varnothing, C = R \setminus (A \cup B)$ схемы *R* верно:

1. Если отношение *r(R)* удовлетворяет F-зависимости *A->B*, то *r(R)* удовлетворяет MV-зависимости *A->>B*
2. Если отношение *r(R)* удовлетворяет нетривиальной (т.е. *B* не является подмножеством *A*) MV-зависимости *A->>B*, то r разлагается без потерь на отношения со схемами *R1=AB*, *R2=AC*.

Пусть *r(R)* – реляционное отношение, *X, Y, Z* – попарно непересекающиеся подмножества атрибутов из схемы *R*.

> Пусть *r1(XY)=*$\pi_{XY}(r)$, *r2(XZ)=*$\pi_{XZ}(r)$– проекции отношения r на схемы *XY* и *XZ*.
> Если отношение r можно получить естественным соединением этих двух проекций, то говорят, что ***r* разлагается без потерь** на отношения со схемами *XY* и *XZ*.

Для нас важно именно второе свойство, т.к. оно показывает возможность декомпозиции отношения на основании многозначных зависимостей.

> **Отношение *r(R)* находится в четвертой нормальной форме** (4НФ), если все MV-зависимости, которым оно удовлетворяет, являются F-зависимостями от потенциальных ключей (F-зависимостями с левой частью в виде потенциального ключа).

Пример 3.
Отношение r3 из примера 1 не находится в 4НФ, т.к. содержит MV-зависимости {код_курса}->>{преподаватель} и {код_курса}->>{учебники}, не являющиеся F-зависимостями(т.е. не выполняется условие «каждому курсу соответствует единственный учебник»).

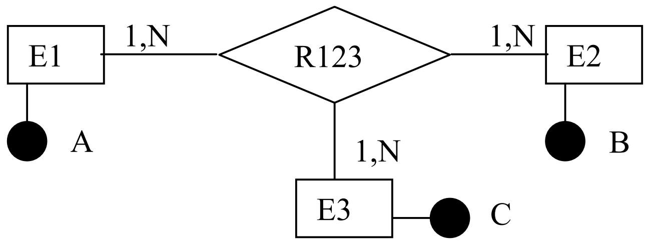
Замечание: 4НФ всегда можно достигнуть.

12.3 Зависимости соединения и 5НФ

До сих пор предполагалось, что при декомпозиции без потерь вплоть до 4НФ нужно разбивать исходное отношение на две его проекции. Однако существуют отношения, для которых ***нельзя*** выполнить декомпозицию без потерь на ***две*** проекции, но ***можно*** нормализовать без потерь на ***три или более*** проекций. Такое отношение называют **n-декомпозируемым** (n>2).
Выясним, каким должно быть отношение, чтобы оно было 3-декомпозируемым.

Пример 4. Ситуация с 3-декомпозируемостью часто возникает при рассмотрении тернарных связей.

r	*A*	*B*	*C*
	A1	*B1*	*C1*
	A1	*B1*	*C2*
	A1	*B2*	*C1*
	A2	*B1*	*C1*



Отношение r декомпозируется на три таких проекции:

r1	*A*	*B*
	a1	*b1*
	a1	*b2*
	a2	*b1*

r2	*B*	*C*
	b1	*c2*
	b2	*c1*
	b1	*c1*

r3	*C*	*A*
	c2	*a1*
	c1	*a1*
	c1	*a2*

Следовательно, отношение r является 3-декомпозируемым без потерь на r1, r2, r3, что эквивалентно:

$$r = r1 \bowtie r2 \bowtie r3 .$$

Для каждого кортежа <a1, b1> из r1, <b1,c2> из r2, <c2,a1> из r3 верно, что <a1,b1,c2>находится в r.

Данное в примере 4 отношение удовлетворяет некоторому ограничению на значения, такому, что это отношение является 3-декомпозируемым в любой момент времени (!). Это ограничение можно записать в виде:

Если *<a1,b1,c2>*, *<a1,b2,c1>*, *<a2,b1,c1>* находятся в отношении *r*
То *<a1,b1,c2>* также находится в отношении *r*

Ограничение это имеет **циклическую структуру,** т.е. если *a1* связано с *b1, b1* связано с *c2, c2* связано с *a1*, то *a1,b1,c1* должны находиться в одном кортеже.

> Отношение будет **3-декомпозируемым**, если оно удовлетворяет этому циклическому ограничению (3д-ограничению)

3д-ограничение выполняется тогда и только тогда, когда отношение равносильно соединению 3-х проекций.
В общем виде, такое ограничение называется **зависимостью соединения** (Join-dependency, J-зависимость)

Пусть *r(R)* – реляционное отношение.
Пусть {*R_1, ..., R_n* } – набор его проекций, $R = R_1 \cup ... \cup R_n$

> Отношение *r(R)* удовлетворяет **зависимости соединения** *[*R_1, ..., R_n*] тогда и только тогда, когда соединение всех его указанных проекций дает отношение *r*
> $r = \pi_{R_1} \bowtie \pi_{R_2} \bowtie ... \bowtie \pi_{Rn}$

Отношения, в которых есть J-зависимости, имеют много аномалий.

Пример 5.
Рассмотрим отношение r0(ABC).

r	*A*	*B*	*C*
	a1	*b1*	*c2*
	a1	*b2*	*c1*

*Известно, что в отношении r0 выполняется J-зависимость *[AB,BC, AC]. Тогда рассмотрим его проекции.*

r1	*A*	*B*
	a1	*b1*
	a1	*b2*

r2	*B*	*C*
	b1	*c2*
	b2	*c1*

r3	*C*	*A*
	c2	*a1*
	c1	*a1*

При добавлении кортежа <a1,b1,c1> в отношение r0 кортеж <b1,c1> будет добавлен в r2. Соединение 3-х полученных проекций даст в точности r0.

r1	*A*	*B*
	a1	*b1*
	a1	*b2*

r2	*B*	*C*
	b1	*c2*
	b2	*c1*
	b1	*c1*

r3	*C*	*A*
	c2	*a1*
	c1	*a1*

r=r1 JOIN r2 JOIN r3=r0	*A*	*B*	*C*
	a1	*b1*	*c2*
+	***a1***	***b1***	***c1***
	a1	*b2*	*c1*

При добавлении кортежа <a2,b1,c1> в отношение r0, получим такое отношение r:

r	*A*	*B*	*C*
	a1	*b1*	*c2*
	a1	*b2*	*c1*
+	***a2***	***b1***	***c1***

Его проекциями будут

r1	*A*	*B*
	a1	*b1*
	a1	*b2*
	a2	*b1*

r2	*B*	*C*
	b1	*c2*
	b2	*c1*
	b1	*c1*

r3	*C*	*A*
	c2	*a1*
	c1	*a1*
	c1	*a2*

r=r1 JOIN r2 JOIN r3 ***не равно*** *r0*	*A*	*B*	*C*
	a1	*b1*	*c2*
+	***a1***	***b1***	***C1***
	a1	*b2*	*c1*
+	*A2*	*B1*	*C1*

Как видно, естественное соединение трех проекций не равно исходному отношению r, т.к. появился кортеж <a1,b1,c1>, который должен быть добавлен одновременно с <a2,b1,c1> для выполнения J-зависимости. Следовательно, в исходном отношении есть аномалии обновления.

Из определения зависимости соединения ясно, что отношение с J-зависимостью можно декомпозировать на соответствующие проекции.
Оказывается, что J-зависимость – это обобщение MV-зависимости (т.е. любая многозначная зависимость является зависимостью соединения, и обратное утверждение неверно).

Сформулируем **теорему Фейгина** (Fagin)

> Отношение *r(ABC)* удовлетворяет зависимости соединения *[AB, AC] тогда и только тогда, когда *r* удовлетворяет многозначной зависимости *A->>B / C*

J-зависимость – наиболее общая форма зависимости по отношению к операциям проекции и соединения.

> **Отношение *r(R)* находится в 5-й нормальной форме** (5НФ, нормальная форма «проекция-соединение») тогда и только тогда, когда каждая зависимость соединения в отношении *r* определяется (только) потенциальными ключами отношения

Или

> **Отношение *r(R)* находится в 5-й нормальной форме** тогда и только тогда, когда все зависимости соединения, которым оно удовлетворяет, является F-зависимостями с левыми частями в виде потенциальных ключей.

Для того, чтобы утверждать, что *r(R)* находится в 5НФ, нужно знать **все** потенциальные ключи и **все** J-зависимости отношения *r.*

Пример 6.
Отношение r0 из примера 5 не находится в 5НФ, т.к. зависимость соединения не определяется потенциальным ключом.

5НФ является высшей нормальной формой, которая может быть получена по отношению к операциям проекции и соединения.
Из всего вышесказанного следует, что
1. 5НФ всегда достижима.
2. Отношение в 5НФ не содержит аномалий, которые могут быть исключены разбиением на проекции.
3. Если отношение *r(R)* находится в 3НФ и все потенциальные ключи в *r* – одноатрибутные (т.е. состоят из одного атрибута), то *r(R)* автоматически находится в 5НФ.
4. Если отношение *r(R)* находится в НФБК и хотя бы один потенциальный ключ в *r* – одноатрибутный, то *r(R)* находится в 4НФ, но не обязательно в 5НФ.

Таким образом, получаем **полный процесс декомпозиции**:
1. Получить 1НФ.
2. Разбить отношения, находящиеся в1НФ на проекции, чтобы исключить частичные зависимости. Получена 2НФ.
3. Разбить отношения, находящиеся в 2НФ на проекции, чтобы исключить транзитивные зависимости. Получена 3НФ.
4. Исключить из отношений, находящихся в 3НФ, те F-зависимости, левые части которых не являются потенциальными ключами. Получена НФБК.

5. Отношения в НФБК разбить на проекции, лишенные MV-зависимостей, которые не являются F-зависимостями. Получена 4НФ.
6. Отношения в 4НФ разбить на проекции, лишенные J-зависимостей, которые не обусловлены потенциальными ключами(если J-зависимости вообще можно выявить). Получена 5НФ.