Лекция 11

Нормальная форма Бойса-Кодда

Четвертая и пятая нормальные формы

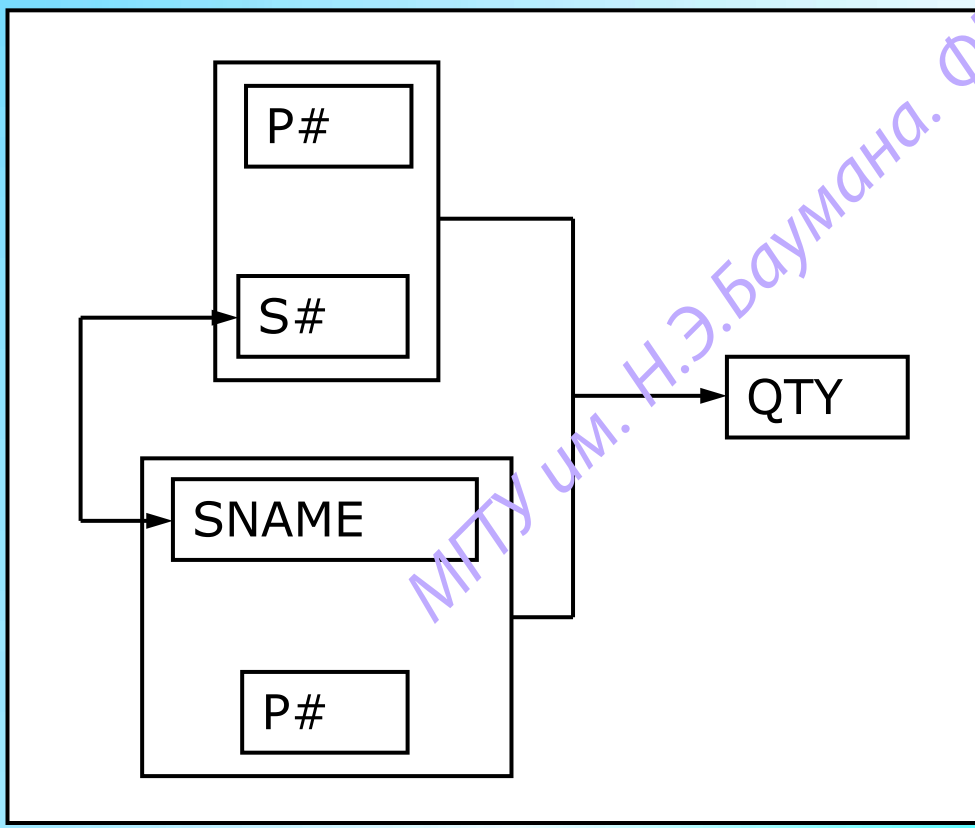
Перекрывающиеся возможные ключи и нормальная форма Бойса-Кодда

Пусть дано отношение R:

- удовлетворяет требованиям *3НФ,*
- имеет несколько потенциальных ключей;
- эти потенциальные ключи составные;
- некоторые из этих возможных ключей «перекрываются», т. е. содержат общие атрибуты.

Аномалии обновлений, связанные с наличием перекрывающихся возможных ключей

Пример 8.1 Переменная-отношение SSP{S#,SNAME,P#,QTY}, все поставщики имеют уникальные имена.
Потенциальные ключи {S#, P#} и {SNAME, P#}.



ФЗ:
S#→SNAME
SNAME→S#,

{S#,P#}→SNAME
{S#,P#}→QTY

{SNAME,P#}→S#,
{SNAME,P#}→QTY

Рис. 8.1. Диаграмма ФЗ переменной-отношения SSP

S#	SNAME	P#	QTY
S_1	Иванов	P_1	300
S_1	Иванов	P_2	100
S_1	Иванов	P_3	200
S_1	Иванов	P_4	350
S_1	Иванов	P_5	400
S_2	Петров	P_3	300
...	...	...	...

Таблица 8.1 Возможное тело значения переменной-отношения SSP

В случае изменения имени поставщика требуется обновить атрибут SNAME во всех кортежах переменной-отношения SSP, соответствующих данному служащему. Иначе ФЗ S#$\rightarrow$SNAME будет нарушена. БД окажется в несогласованном состоянии.

Нормальная форма Бойса-Кодда

Переменная отношения находится в ***нормальной форме Бойса-Кодда*** (**НФБК** или ***BCNF***) в том и только в том случае, когда любая выполняемая для этой переменной- отношения нетривиальная и неприводимая ФЗ имеет в качестве детерминанта некоторый потенциальный ключ данного отношения.

Переменная-отношение SSP{S#,SNAME,P#,QTY} должна быть разделена на две проекции.

Существуют два варианта декомпозиции.

1 вариант

SS{S#, SNAME }

SP{S#, P#, QTY}

2 вариант

SS{S#, SNAME}

SP{SNAME, P#, QTY}

Пример 8.2

Имеется переменная-отношение SNew{S#,SNAME,STATUS,CITY}.

Потенциальные ключи —атрибуты S#, SNAME.

Атрибуты STATUS и CITY являются независимыми.

Множество ФЗ показано на рис. 8.2.

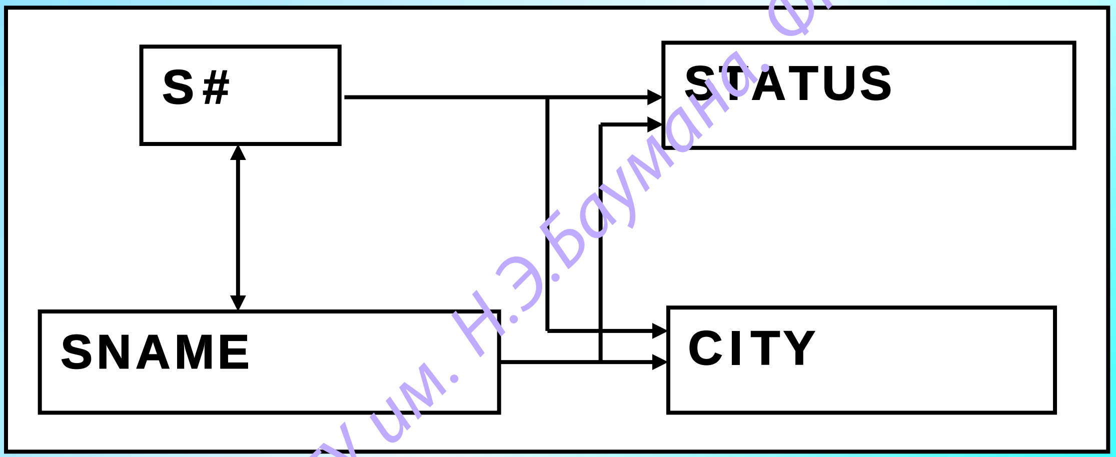


Рис. 8.2. Диаграмма ФЗ переменной-отношения SNew.

Переменная-отношение SNew находится в н. ф. Бойса-Кодда.

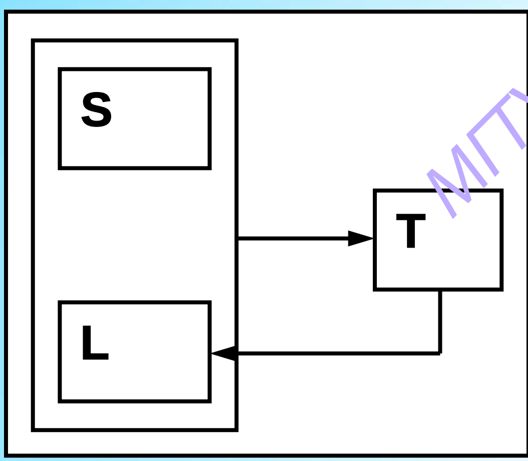
Всегда ли следует стремиться к НФБК?

В переменной-отношение SLT{S,L,T} хранится информация о студентах S, лекциях L и преподавателях T. Кортеж (s,l,t) означает, что студент *s* слушает лекции *l* у преподавателя *t*.

Ограничения:

1. Каждый студент изучает определенный предмет только у одного преподавателя; ФЗ {S,L}→{T};
2. Каждый преподаватель ведет только один предмет, каждый предмет может преподаваться несколькими преподавателями; ФЗ {T}→{L}.

Множество ФЗ переменной-отношения SLT {S,L}→{T} ,{T}→{L}.



Существуют два перекрывающихся потенциальных ключа: {S,L} и {S,T} (∃{T}→{L}⇒{S,T}→{S,L}⇒{S,T}→{L}).

Рис. 8.3. Множество ФЗ переменной-отношения.

SLT		
S	**L**	**T**
Алешин	Мат. анализ	Проф. Иванов
Борисов	Лин. алгебра	Проф. Петров
Васькин	Мат. анализ	Проф. Иванов
Гришин	Лин. алгебра	Проф. Сидоров

Таблица 8.2 Возможное тело значения переменной-отношения SLT.

Переменная отношение SLT удовлетворяет требованиям *3НФ*.

Отношение не находится в НФБК (наличие ФЗ $\{T\} \rightarrow \{L)\}$).

Атрибут T является детерминантом ФЗ $\{T\} \rightarrow \{L\}$, но не является ПК (T входит в состав потенциального ключа).

SLT свойственны *аномалии обновления*.

Например, удалив данные о студенте Гришине, утратим информацию о том, что пр. Сидоров преподает линейную алгебру(т.к. S является компонентом обоих возможных ключей).

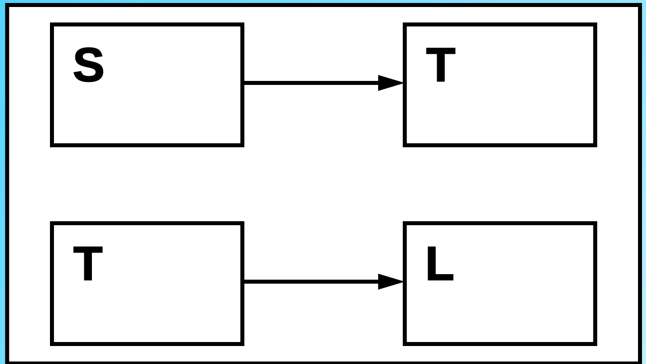


Рис. 8.4. Множество ФЗ переменных-отношений TS и TL.

ST	
S	**T**
Алешин	Проф. Иванов
Борисов	Проф. Петров
Васькин	Проф. Иванов
Гришин	Проф. Сидоров

TL	
T	**L**
Проф. Иванов	Мат. анализ
Проф. Петров	Лин. алгебра
Проф. Сидоров	Лин. алгебра

Таблица 8.3 Возможные тела значений переменных-отношений ST и TL.

Трудности: проекции TS и TL не являются **независимыми**.

ФЗ $\{S,L\} \rightarrow \{T\}$ не выводится из ФЗ $\{T\} \rightarrow \{L\}$

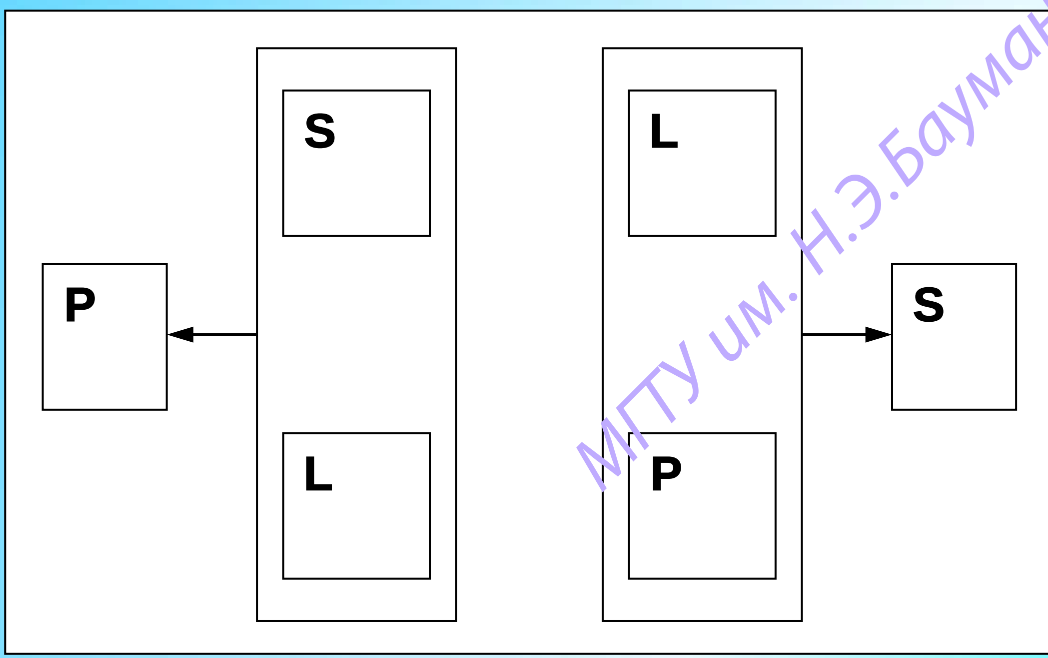
Отношение SLT — *атомарно*, хотя и не находится в НФБК.

Переменная-отношение называется **атомарной**, если она не может быть подвергнута декомпозиции с получением независимых проекций.

Пример.
Задана переменная-отношение EXAM{S,L,P}.
(S — студент, L — предмет,P — позиция).

Кортеж (***s***, ***l***, ***p***) означает, что студент ***s*** экзаменуется по предмете ***l*** и занимает позицию ***p*** в экзаменационной ведомости.

Ограничение: никакие два студента не могут занимать одну и ту же позицию в экзаменационной ведомости по одному и тому же предмету.



Перекрывающиеся ПК {S,L} и {L,P}.

Переменная-отношение находится в НФБК, ключи являются единственными детерминантами.

Рис. 8.5 Множество ФЗ переменной-отношения EXAM

Многозначные зависимости и четвертая нормальная форма

Пусть имеется отношение с повторяющимися наборами данных. Декомпозиция, основанная на ФЗ, не приводит к исключению такой избыточности.

В этом случае используют декомпозицию, основанную на *многозначных зависимостях (МЗЗ)*.

МЗЗ является обобщением ФЗ и рассматривает соответствия между множествами значений атрибутов.

Пример. Переменная-отношение ПРЕПОДАВАТЕЛЬ_ОБЩ (КУРСЫ, ИМЕНА, УЧЕБНОЕ_ПОСОБИЕ).

ПРЕПОДАВАТЕЛЬ_ОБЩ		
КУРСЫ	**ИМЕНА**	**УЧЕБНОЕ_ПОСОБИЕ**
Линейная алгебра	**ИМЯ** Иванов Сидоров	**УЧЕБНИК** «Линейная алгебра» Крищенко А.П., Канатников А.Н. «Линейная алгебра» Воеводин В.В.
Математический анализ	**ИМЯ** Сидоров	**УЧЕБНИК** «Курс математического анализа» Кудрявцев Л.Д. «Курс дифференциального и интегрального исчисления» Фихтенгольц Г.М.

Таблица 8.4 Набор значений данных в переменной-отношения ПРЕПОДАВАТЕЛЬ_ОБЩ.

атрибут КУРС содержит название курса;

атрибут-отношение ИМЯ содержит имена преподавателей;

атрибут-отношение УЧЕБНОЕ_ПОСОБИЕ содержит название используемых учебников и фамилии авторов.

Ограничения:

Соответствующий курс может преподаваться любым из указанных преподавателей с использованием всех указанных учебников. Для заданного курса может быть определено произвольное количество соответствующих преподавателей и учебников.

Преподаватели и учебники независимы друг от друга (всегда исполь-зуется один и тот же набор учебников). Определенный преподаватель или определенный учебник может быть связан с любым количеством курсов.

Необходимо привести отношение в 1НФ, т.е. исключить атрибуты, принимающие в качестве значений отношения.

Заменим переменную-отношение ПРЕПОДАВАТЕЛЬ_ОБЩ (ИМЕНА, КУРСЫ, УЧЕБНОЕ_ПОСОБИЕ) переменной-отношением ПРЕПОДАВАТЕЛЬ (ИМЯ, КУРС, УЧЕБНИК) с тремя скалярными атрибутами.

Из каждого кортежа исходной переменной-отношения ПРЕПОДАВАТЕЛЬ_ОБЩ получаем $m*n$ кортежей в переменной-отношении ПРЕПОДАВАТЕЛЬ, где m и n количество строк в отношениях ИМЯ и УЧЕБНИК.

Все атрибуты результирующей переменной-отношения ПРЕПОДАВАТЕЛЬ входят в состав ключа. В переменной-отношении ПРЕПОДАВАТЕЛЬ_ОБЩ ключ состоял из одного атрибута КУРСЫ.

ПРЕПОДАВАТЕЛЬ		
КУРС	**ИМЯ**	**УЧЕБНИК**
Линейная алгебра	Иванов	«Линейная алгебра» Крищенко А.П., Канатников А.Н.
Линейная алгебра	Иванов	«Линейная алгебра» Воеводин В.В.
Линейная алгебра	Сидоров	«Линейная алгебра» Крищенко А.П., Канатников А.Н.
Линейная алгебра	Сидоров	«Линейная алгебра» Воеводин В.В.
Математический анализ	Сидоров	«Курс математического анализа» Кудрявцев Л.Д.
Математический анализ	Сидоров	«Курс дифференциального и интегрального исчисления» Фихтенгольц Г.М.

Таблица 8.5 Набор значений данных в переменной-отношения ПРЕПОДАВАТЕЛЬ, эквивалентный набору данных таблицы 8.4.

Данные, помещаемые в переменную-отношение ПРЕПОДАВАТЕЛЬ, имеют следующий смысл: кортеж (КУРС:к, ИМЯ:и, УЧЕБНИК:у) появляется в переменной-отношении ПРЕПОДАВАТЕЛЬ тогда и только тогда, когда курс ***к*** читается преподавателем ***и*** с использованием учебника ***у***. Для каждого курса указаны все возможные комбинации имени преподавателей и названий учебников.

Можно утверждать, что для переменной-отношения ПРЕПОДАВАТЕЛЬ верно следующее ограничение:

ЕСЛИ кортежи (***к***,***и***1,***у***1) и (***к***,***и***2,***у***2) присутствуют одновременно,

ТО кортежи (***к***,***и***1,***у***2) и (***к***,***и***2,***у***1) также присутствуют одновременно.

Аномалии обновлений при наличии многозначных зависимостей и возможная декомпозиция

Аномалии обновления — связаны с избыточностью.

Например, для добавления информации о том, что курс математического анализа может читаться новым преподавателем, необходимо создать два новых кортежа, по одному для каждого используемого учебника.

Проблемы возникают в результате того, что преподаватели и учебники не зависят друг от друга.

Декомпозиция переменной-отношения ПРЕПОДАВАТЕЛЬ на проекции К_Имя{КУРС,ИМЯ} и К_Учебник{КУРС,УЧЕБНИК}.

К_Имя

КУРС	**ИМЯ**
Линейная алгебра	Иванов
Линейная алгебра	Сидоров
Математический анализ	Сидоров

К_Учебник

КУРС	**УЧЕБНИК**
Линейная алгебра	«Линейная алгебра» Крищенко А.П., Канатников А.Н.
Линейная алгебра	«Линейная алгебра» Воеводин В.В.
Математический анализ	«Курс математического анализа» Кудрявцев Л.Д.
Математический анализ	«Курс дифференциального и интегрального исчисления» Фихтенгольц Г.М.

Таблица 8.6. Значения данных в проекциях К_Имя и К_Учебник, соответствующие содержанию табл. 8.5.

В переменной-отношения ПРЕПОДАВАТЕЛЬ существуют **многозначные зависимости** (**МЗЗ** *multi-valued dependency – MVD*).

КУРС$\rightarrow\rightarrow$ИМЯ КУРС$\rightarrow\rightarrow$УЧЕБНИК

A→→B B многозначно зависит от A   или
A многозначно определяет B

Зависимость КУРС→→ИМЯ, означает, что, хотя для каждого курса не существует одного соответствующего только ему преподавателя, т.е. не выполняется ФЗ КУРС→ИМЯ, каждый курс имеет вполне определенное множество соответствующих преподавателей.

Для данного курса ***к*** и данного учебника ***у*** множество преподавателей ***и***, соответствующее паре (***к***,***у***) переменной-отношения ПРЕПОДАВАТЕЛЬ, зависит от значения ***к*** и не зависит от значения ***у***.

Зависимость КУРС→→УЧЕБНИК можно интерпретировать аналогично.

Из теории отношений известно понятие соответствие из множества А в множество В, когда одному элементу множества А может сопоставляться несколько элементов множества В. Понятие соответствие обобщает понятие ФЗ между множествами А и В.

В реляционной теории возникает ситуация, когда устанавливается <u>соответствие</u> из множества атрибутов А в множество атрибутов В и из множества атрибутов А в множество атрибутов С. При этом множества В и С между собой независимы.

Сечением отношения R с атрибутами A,B,C по значению $\boldsymbol{a}$ атрибуту A называется отношение $S(R,\boldsymbol{a})$, заголовок которого совпадает с заголовком отношения R и телом состоящим из кортежей (a,x,y), таких что $x \in B$, $y \in C$ и $(a,x,y) \in R$.

Определение

Пусть A, *B* и *C* являются произвольными подмножествами множества атрибутов переменной-отношения *R*.

Подмножество В ***многозначно зависит*** от подмножества А (А→→В), в том и только в том случае, когда множество значений атрибута В, соответствующее заданной паре значений атрибутов *А* и *С* переменной-отношения *R*, зависит от *значения А*, но не зависит от *значения С*.

Для $\forall a \in A$

S(R,a) PROJECT[B,C]= S(R,a) PROJECT[B] TIMES S(R,a) PROJECT[C]

Многозначные зависимости обладают свойством "двойственности".

Лемма Фейджина

Для переменной-отношения R{A, B, C} многозначная зависимость $A \to\to B$ выполняется тогда и только тогда, когда выполняется многозначная зависимость $A \to\to C$.

Доказательство

- RV — некоторое удовлетворяющее МЗЗ значение R,
- ***a*** — значение атрибута A в некотором кортеже тела RV,
- {b} и {c } — множества значений атрибута B и атрибута C, взятых из всех кортежей тела RV, в которых значением атрибута A является ***a***;
- ***C*** — множество допустимых значений атрибута C;
- ***B*** — множество допустимых значений атрибута B;

1. Пусть выполняется МЗЗ $A \to\to B$.

Пусть для этого значения ***a*** МЗЗ $A \to\to C$ не выполняется $\Rightarrow \exists$ ***c***$\in$***C*** & ***b***$\in${b}, что (a, b, c)$\notin$RV. Противоречит МЗЗ $A \to\to B$ (по определению МЗЗ
$\forall$ a$\in$A S(R,a) PROJECT[B,C] = S(R,a) PROJECT[B] TIMES S(R,a) PROJECT[C]).

Следовательно, если выполняется МЗЗ $A \to\to B$, то выполняется МЗЗ $A \to\to C$.

2. Пусть выполняется МЗЗ $A \to\to C$,
Пусть для этого значения ***a*** МЗЗ $A \to\to B$ не выполняется $\Rightarrow \exists$ ***c***$\in${c} & ***b***$\in$***B***, что (a, b, c)$\notin$RV. Противоречит МЗЗ $A \to\to C$ (по определению МЗЗ). **#**

Таким образом, многозначные зависимости всегда образуют связанные пары, поэтому обычно их представляют вместе в виде $A \rightarrow\rightarrow B|C$.

Для примера такая запись будет иметь следующий вид

КУРС→→ИМЯ | УЧЕБНИК

МЗЗ являются обобщениями ФЗ.

ФЗ— это МЗЗ, в которой множество зависимых значений, соответствующее заданному значению детерминанта, всегда является одноэлементным множеством.

Если $A \rightarrow B$, то $A \rightarrow\rightarrow B$.

Обратное утверждение неверно.

Теорема Фейджина.

Пусть A, B и C — множества атрибутов переменной-отношения R. Переменная-отношение R будет равна соединению ее проекций [A,B] и [A,C] тогда и только тогда, когда для переменной-отношения R выполняется МЗЗ $A \to\to B|C$.

$$R=(R\ PROJECT[A,B])\ JOIN\ (R\ PROJECT[A,C]) \Leftrightarrow A \to\to B|C$$

Доказательство

1. <u>Достаточность</u>.

$A \to\to B|C \Rightarrow R=R_1$, где $R_1=(R\ PROJECT[A,B])\ JOIN\ (R\ PROJECT[A,C])$

a значение атрибута A в некоторых кортежах R.

{b}, {c} — множества значений атрибута B и атрибута C, соответственно, взятых из всех кортежей R, где значением атрибута A является ***a***.

$\mathbf{R \subseteq R_1}$ $\quad \forall b_i \in \{b\}\ \&\ \forall c_i \in \{c\} \Rightarrow (a,b_i,c_i) \in R$ (по определению МЗЗ $\forall\, a \in A$ $\boldsymbol{S(R,a)\ PROJECT[B,C] = S(R,a)\ PROJECT[B]\ TIMES\ S(R,a)\ PROJECT[C]}$) $\Rightarrow$

$$(a,b_i) \in R\ PROJECT[A,B]\ \&\ (a,c_i) \in R\ PROJECT[A,C] \Rightarrow (a,b_i,c_i) \in R_1$$

$\mathbf{R_1 \subseteq R}$ $\quad (a,b_i,c_j) \in R_1 \Rightarrow (a,b_i) \in R\ PROJECT[A,B]\ \&\ (a,c_j) \in R\ PROJECT[A,C]$

$$b_i \in \{b\}\ \&\ c_j \in \{c\} \Rightarrow (\text{по определению МЗЗ}) \Rightarrow (a,b_i,c_j) \in R$$

2. Необходимость.

Пусть R=(R PROJECT[A,B]) JOIN (R PROJECT[A,C]), $\Rightarrow$ A$\rightarrow\rightarrow$B|C.

A$\rightarrow\rightarrow$B|C$\Rightarrow$ IF **(**$((a,b_1,c_1)\in R)$ AND $((a,b_2,c_2)\in R)$**)**

THEN **(**$((a,b_1,c_2)\in R)$ AND $((a,b_2,c_1)\in R)$**)**

Докажем **от противного**.

Пусть $((a,b_1,c_1)\in R \ \&\ (a,b_2,c_2)\in R) \ \&\ ((a,b_1,c_2)\notin R \vee (a,b_2,c_1)\notin R)$.

По условию $(a,b_1)\in$(R PROJECT[A,B]) & $(a,b_2)\in$(R PROJECT[A,B])

$(a,c_1)\in$(R PROJECT[A,C]) & $(a,c_2)\in$(R PROJECT[A,C])

$\Rightarrow$ $(a,b_1,c_2)\in$(R PROJECT[A,B]) JOIN (R PROJECT[A,C])

$(a,b_2,c_1)\in$(R PROJECT[A,B]) JOIN (R PROJECT[A,C])

предположение об отсутствии по крайней мере одного из этих кортежей в R противоречит R=(R PROJECT[A,B]) JOIN (R PROJECT[A,C])

T*еорема* Фейджина доказана. **#.**

Четвертая нормальная форма.

Переменная отношения R находится в **четвертой нормальной форме** (4НФ) в том и только в том случае, когда она находится в НФБК, и все МЗЗ R являются ФЗ с детерминантами – потенциальными ключами отношения R.

Иначе говоря, в переменной-отношении R могут находиться только нетривиальные зависимости (функциональные или многозначные) вида $K \rightarrow X$, т.е. некоторый атрибут X *функционально* зависит от суперключа K.

МЗЗ $A \rightarrow \rightarrow B$ называется **тривиальной,** если либо A является супермножеством B, либо объединение A и B образует весь заголовок отношения.

Эквивалентное определение 4НФ:

Отношение находится в 4НФ, если оно находится в НФБК и в нем отсутствуют МЗЗ, не являющиеся ФЗ.

Замечание. Теорема Риссанена (см. лекцию 10) сформулирована с использованием ФЗ..

Переменную-отношение R{A,B,C}, удовлетворяющую ФЗ $A \rightarrow B$ и $B \rightarrow C$, необходимо разбивать на проекции [A,B] и [B,C], а не на проекции [A,B] и [A,C].

Теорема Риссанена справедлива в отношении МЗЗ.

Переменную-отношение R{A,B,C}, удовлетворяющую МЗЗ $A \rightarrow\rightarrow B$ и $B \rightarrow\rightarrow C$, необходимо разбивать на проекции [A,B] и [B,C], а не на проекции [A,B] и [A,C].

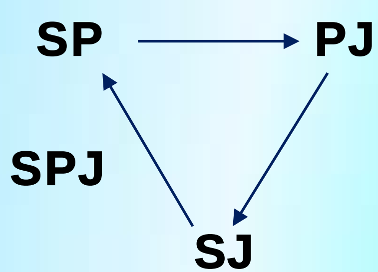
Зависимости соединения и пятая нормальная форма

n*-декомпозируемым** отношение — это отношение, которое может быть декомпозировано без потерь на ***n проекций.

Ранее мы имели дело с 2-декомпозируемыми отношениями.

Пример. Переменная-отношение SPJ из БД поставщиков, деталей и проектов (рис. 8.6).

Переменная-отношение состоит только из ключевых атрибутов, не содержит нетривиальных ФЗ и МЗЗ, находится в 4НФ.



SPJ		
S#	**P#**	**J#**
S_1	P_1	J_2
S_1	P_2	J_1
S_2	P_1	J_1
S_1	P_1	J_1

рис. 8.6

SP	
S#	**P#**
S_1	P_1
S_1	P_2
S_2	P_1

PJ	
P#	**J#**
P_1	J_2
P_2	J_1
P_1	J_1

JS	
J#	**S#**
J_2	S_1
J_1	S_1
J_1	S_2

Соединяем по атрибуту P#

SPJ_1		
S#	**P#**	**J#**
S_1	P_1	J_2
S_1	P_1	J_1
S_1	P_2	J_1
S_2	*P_1*	*J_2*
S_2	P_1	J_1

Соединяем по Комбинации атрибутов J# и S#

SPJ		
S#	**P#**	**J#**
S_1	P_1	J_2
S_1	P_2	J_1
S_2	P_1	J_1
S_1	P_1	J_1

SPJ		
S#	**P#**	**J#**
S_1	P_1	J_2
S_1	P_2	J_1
S_2	P_1	J_1
S_1	P_1	J_1

Рис. 8.7.

Утверждение 1

«Переменная-отношение SPJ равна соединению трех своих проекций SP,JS и PJ» эквивалентно следующему утверждению

ЕСЛИ пара (s_1,p_1) присутствует в SP
И пара (p_1,j_1) присутствует в PJ,
И пара (j_1,s_1) присутствует в JS,
ТО тройка (s_1,p_1,j_1) присутствует в SPJ.

$$((s_1,p_1)\in SP \ \&\ (p_1,j_1)\in PJ \ \&\ (j_1,s_1)\in JS) \Rightarrow (s_1,p_1,j_1)\in SPJ$$

Верно и обратное $(s_1,p_1,j_1)\in SPJ \Rightarrow ((s_1,p_1)\in SP \ \&\ (p_1,j_1)\in PJ \ \&\ (j_1,s_1)\in JS)$

Можно записать

$$((s_1,p_1)\in SP \ \&\ (p_1,j_1)\in PJ \ \&\ (j_1,s_1)\in JS) \Leftrightarrow (s_1,p_1,j_1)\in SPJ$$

Так как пара (s_1,p_1) присутствует в отношении SP тогда и только тогда, когда тройка (s_1,p_1,j_2) присутствует в отношении SPJ для некоторого значения j_2 (аналогично для (p_1,j_1) и (j_1,s_1)), утверждение 1 можно переписать в виде ограничения, накладываемого на отношение SPJ.

ЕСЛИ кортежи (s_1,p_1,j_2), (s_2,p_1,j_1), (s_1,p_2,j_1) присутствуют в SPJ

ТО кортеж (s_1,p_1,j_1) также присутствует в SPJ

$$(((s_1,p_1,j_2), (s_2,p_1,j_1), (s_1,p_2,j_1)) \in SPJ) \Rightarrow (s_1,p_1,j_1) \in SPJ$$

Это ограничение имеет **циклическую структуру** и называется 3-декомпозируемым ограничение (*3Д-ограничением*)

«Если значение s_1 связано с p_1 и p_1 связано с j_1, а j_1 связано опять с s_1, то s_1, p_1 и j_1 должны находиться в одном кортеже»

Определение. Переменная-отношение будет n-декомпозируемой для n>2 в том и только в том случае, когда она удовлетворяет некоторому циклическому ограничению.

Пример.

Пусть переменная-отношение SPJ действительно удовлетворяет этому не зависящему от времени ограничению.

(пример, представленный на рис. 8.7 соответствует такой гипотезе)

С практической точки зрения 3Д-ограничение означает :

если в реальном мире для переменной-отношения SPJ верны утверждения

а) Смит поставляет гаечные ключи,

б) Гаечные ключи используются в Манхэттенском проекте,

в) Смит является поставщиком для Манхэттенского проекта,

то

г) Смит поставляет гаечные ключи для Манхэттенского проекта.

ЗД-ограничение называется **зависимостью соединения** (ЗС), поскольку ЗД-ограничение удовлетворяется тогда и только тогда, когда переменная-отношение равносильна соединению некоторых ее проекций и является таким же ограничением для данной переменной-отношения, как МЗЗ и ФЗ.

Определение

Дано:

R — переменная отношения;

A, B, ..., Z — произвольные подмножества заголовка R (составные, перекрывающиеся атрибуты).

Переменная-отношение R удовлетворяет **зависимости соединения**

***{A, B, ... , Z}**

в том и только в том случае, когда любое допустимое значение переменной-отношения R эквивалентно соединению ее проекций на атрибуты A, B, ..., Z.

Зависимость соединения в англоязычной литературе называется зависимостью проекции/соединения Project-Join Dependency – PJD.

Наличие ЗС в переменной-отношения R характеризуется аномалиями обновления, устранить которые можно лишь с помощью 3-декомпозиции. Переменная-отношение SPJ удовлетворяет ЗС *{SP,PJ,JS} и может быть 3-декомпозируемой.

Примеры аномалий обновления в переменной-отношении S

SPJ		
S#	**P#**	**J#**
S_1	P_1	J_2
S_1	P_2	J_1

- Если вставляется кортеж (S_2,P_1,J_1), то также должен быть вставлен кортеж (S_1,P_1,J_1)

$\exists(S_1,P_1) \ \& \ \exists(P_1,J_1) \ \& \ \exists(S_1,J_1)$

$\Rightarrow \exists(S_1,P_1,J_1)$

- Обратное утверждение не является истинным

Проекции SPJ		
SP	**PJ**	**JS**
S_1P_1	P_1J_2	J_2S_1
S_1P_2	P_2J_1	J_1S_1
S_2P_1	P_1J_1	J_1S_2
S_1P_1	P_1J_1	J_1S_1

SPJ		
S#	**P#**	**J#**
S_1	P_1	J_2
S_1	P_2	J_1
S_2	P_1	J_1
S_1	P_1	J_1

- Кортеж (S_2,P_1,J_1) может быть удален без побочных эффектов
- Если удаляется кортеж (S_1,P_1,J_1), то также должен быть удален еще один кортеж.

Рис. 8.8. Примеры аномалий обновления в переменной-отношении SPJ

Теорема Фейджина : переменная-отношение R{A,B,C} может быть декомпозирована без потерь на проекции с атрибутами {A,B} и {A,C} тогда и только тогда, когда для переменной-отношения R выполняются МЗЗ $A \to\to B$ и $A \to\to C$.

Теперь теорему Фейджина можно сформулировать иначе.

Переменная-отношение R{A, B, C} удовлетворяет зависимости соединения *{AB, AC} тогда и только тогда, когда она удовлетворяет многозначной зависимости $A \to\to B|C$.

$$A \to\to B|C \Leftrightarrow *\{AB, AC\}$$

Зависимость соединения является обобщением понятия многозначной зависимости.

Пятая нормальная форма.

Переменная-отношение R находится в **пятой нормальной форме** (5НФ), или **проекционно-соединительной нормальной форме** (ПСНФ), в том и только в том случае, когда каждая нетривиальная зависимость соединения в переменной-отношении R определяется ее потенциальными ключами.

Зависимость соединения *{A, B, … , Z} в переменной отношения R называется **тривиальной**, если хотя бы один из составных атрибутов A, B, …, Z совпадает с заголовком R.

Заданная зависимость соединения *{A, B, ..., Z} определяется потенциальными ключами тогда и только тогда, когда каждое подмножество атрибутов A, B, ... , Z фактически является суперключом для данной переменной-отношения

Каждая проекция такого отношения содержит не менее одного ПК и не менее одного неключевого атрибута

Любая переменная -отношение может быть подвергнута декомпозиции без потерь на эквивалентный набор переменных-отношений в 5НФ, т.е. 5НФ всегда достижима.

Рассмотрим переменную-отношение S поставщиков с потенциальными ключами S# и SNAME. Такая переменная-отношение удовлетворяет нескольким зависимостям соединения, в частности следующей зависимости.

*{{S#, SNAME, STATUS}, {S#, CITY}}

Переменная-отношение S равносильна соединению ее проекций с атрибутами {S#, SNAME, STATUS} и {S #, CITY}. Поэтому она может быть подвергнута декомпозиции без потерь на указанные проекции.

Ее *не следует,* а лишь *можно* подвергать подобной декомпозиции. Существование данной ЗС предполагается на основании того факта, что атрибут {S#} является ПК данной переменной-отношения.

Переменная-отношение S удовлетворяет еще одной ЗС.

*{ {S#, SNAME}, {S#, STATUS}, {SNAME, CITY} }

Эта зависимость следует из того, что *оба* атрибута, { S#} и {SNAME}, являются потенциальными ключами.

Как следует из примера, заданная зависимость соединения *{A, B, ..., Z} определается потенциальными ключами *тогда и только тогда, когда каждое подмножество атрибутов A, B, ... , Z фактически является суперключом для данной переменной-отношения.*

Например, переменная-отношение поставщиков S находится в 5НФ. Эта переменная-отношение *может* быть подвергнута дальнейшей декомпозиции без потерь, причем в нескольких вариантах, но каждая проекция в любом из этих вариантов по-прежнему будет содержать один из исходных потенциальных ключей. Следовательно, подобная декомпозиция не даст никаких дополнительных преимуществ.