

621.382.2

К 172

47 797 у

**А.Ж. Каливанов**

**Основы проектирования реляционных баз данных**

*Руководство к практическим занятиям*



---

---

**ВОЕННО-КОСМИЧЕСКАЯ АКАДЕМИЯ  
имени А.Ф. МОЖАЙСКОГО**

Санкт-Петербург

УДК 681.3.06 (075.8)

Руководство к практическим занятиям содержит методические рекомендации по решению ряда основных задач проектирования реляционных баз данных, а также к оформлению результатов их решения. Данные рекомендации рассмотрены на примере одной предметной области, что позволяет получить представление о взаимосвязи и целенаправленности этапов создания информационной системы.

Руководство к практическим занятиям предназначено для курсантов, изучающих дисциплину “Системы управления базами данных”. Темы практических занятий даны в рамках учебного плана и подготовлены по материалам занятий, которые автором проводились в течение 2000÷2015 гг. Руководство может быть использовано профессорско-преподавательским и инженерно-техническим составом академии при подготовке и проведении курсового и дипломного проектирования, а также при разработке автоматизированных информационных систем.

© КАЛИВАНОВ А.Ж.

© ВКА имени А.Ф. Можайского

## Содержание

	<u>Стр.</u>
<b><u>Введение</u></b> .....	4
<b>Нормализация отношений</b>	
<u>Практическое занятие № 1</u> .....	8
<u>Практическое занятие № 2</u> .....	22
<b>Формирование канонической структуры отношений</b>	
<u>Практическое занятие № 3</u> .....	37
<u>Практическое занятие № 4</u> .....	43
<b>Формирование концептуальной схемы базы данных</b>	
<u>Практическое занятие № 5</u> .....	56
<b>Порядок расчета затрат на реализацию логической схемы базы данных</b>	
<u>Практическое занятие № 6</u> .....	66
<u>Практическое занятие № 7</u> .....	77

## **ВВЕДЕНИЕ**

База данных (БД), как основа современных информационных систем, позволяет накапливать, хранить, и реализовать запросы на корректирование и выдачу информации.

Логика построения современных информационных систем определяется переходом от индивидуального программирования на языке машин к созданию операционных систем и языков высокого уровня, а от них – к выделению процедур ведения массивов данных и созданию систем управления базами данных (СУБД).

Создание БД стало возможным благодаря тому, что была найдена возможность разделения физического и логического уровня представления данных.

Логический уровень – уровень представления данных пользователя в СУБД.

Физический уровень – уровень представления того, как эти данные хранятся на носителях информации.

Понятие структуры данных – основывается на взаимосвязи совокупности некоторых элементов данных (ЭД), конструкций, составленных из этих ЭД, и определении правил композиции этих конструкций. Примерами таких конструкций являются массивы, структуры, наборы данных и т.д.

Назначение структур данных состоит не в том, чтобы добиться однозначного представления смешанных данных, а в том, чтобы ими можно было управлять с помощью пакета стандартных программ, входящих в конкретную СУБД.

Таким образом, концепция создания современной информационной системы предполагает создание БД, подразумевая, что данные с некоторой семантической (смысловой) структурой укладываются в заданные структуры, заданного логического уровня. Кроме того, предполагается,

что для наиболее полного изучения и осмысления предметной области, имеется возможность обобщения семантических структур (представлений) всех пользователей данной БД на концептуальном уровне. Это позволяет рассматривать структуры данных независимо от приложений (программ).

Независимость от приложения необходима потому, что обычно в момент проектирования БД не известны все возможные способы использования данных. Поэтому необходимо, чтобы структура (или схема) БД оставалась стабильной (работоспособной) даже при возникновении у пользователей новых требований к данным. Таким образом, нужно создать концептуальную схему БД, т.е. абстрактную структуру логического уровня, не зависящую от аппаратного обеспечения, операционной системы, СУБД, языка программирования, пользователя и т.д.

Поэтому суть проектирования БД состоит в последовательном создании и анализе схем концептуального, логического и физического уровня.

Основная цель проектирования БД – это сокращение избыточности хранимых данных, а следовательно, экономия объема используемой памяти, уменьшение затрат на многократные операции обновления избыточных копий и устранение возможности возникновения противоречий целостности данных из-за хранения в разных местах сведений об одном и том же объекте.

Существует взаимобратная зависимость между затратами памяти и времени, которое расходуется на выполнение операций над данными в БД. Так, введением избыточности данных можно исключить ряд физических операций ввода-вывода, а отсутствие избыточности сохраняет память, но увеличивает временные затраты.

Нахождение разумного компромисса при выборе накладных расходов памяти и/или времени и является задачей проектирования БД.

Необходимо отметить, что проектирование БД скорее искусство, чем просто наука. Конечно, существуют научные принципы проектирования схем БД, и именно они рассматриваются в данном руководстве. Хотя, при построении БД возникает множество других задач, которые не всегда можно охватить этими принципами, тем не менее, методические рекомендации данного руководства даны на основе классических подходов к проектированию БД.

В руководстве рассматривается решение следующих задач:

- нормализация отношений;
- формирование канонической структуры отношений;
- формирование концептуальной схемы.

а также даны примеры описания результатов и формулировки выводов.

Отчет по практическому занятию рекомендуется оформлять в виде пояснительной записки. При оформлении отчета по практическому занятию следует выполнять общие требования ЕСКД, которые предъявляются к разработке текстовых документов. Примеры оформления основных пунктов пояснительных записок приведены в данном руководстве. Они обобщают перечень требований, который содержится в следующей литературе:

1. ГОСТ 2.105–95 ЕСКД. Общие требования к текстовым документам. –М: Госстандарт России. Изд. Официальное. 1995 г.
2. Информационный бюллетень № 96: Оформление текстовых конструкторских документов при курсовом и дипломном проектировании/ Сост. Н.Н. Лазуткова, Н.А.

Иванов. ВИКА им. А.Ф. Можайского. –Санкт-Петербург, 1998 г. –60с.

3. Информационный бюллетень № 104: Методические рекомендации по разработке и исполнению выпускной квалификационной работы/ Сост. А.Г. Ломако, С.В.Баушев, В.И.Якунин, И.А.Ткач. ВИКУ им.А.Ф.Можайского. –Санкт-Петербург, 2002 г. –88с.
4. Каливанов А.Ж. Введение в операционные системы, вычислительные сети и средства разработки документов. Учебное пособие, –СПб., 2001 г., ВИКУ, -109с.

## ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 1

**Тема занятия:** Способы нормализации отношений.

**Цель занятия:** Приобретение навыков приведения реляционных отношений к 1, 2 и 3-ей формам нормализации.

**Задание:** Нормализовать исходные отношения между атрибутами заданной предметной области.

### Методические указания по выполнению задания № 1

#### *Накануне занятия:*

- ознакомиться с темой, целью и заданием;
- изучить лекционный материал и рекомендуемую литературу по заданной теме;
- подготовить подсобный материал для составления письменного отчета по заданию;
- получить в библиотеке и иметь на занятии рекомендованную литературу.

#### *В ходе занятия:*

- принять участие в разборе примера выполнения задания;
- получить у преподавателя вариант задания;
- самостоятельно выполнить нормализацию реляционного отношения заданной предметной области;
- подготовить письменный отчет о выполнении задания.

### Порядок выполнения задания № 1

1. Изучение предметной области. Предметная область может быть задана текстовым описанием, овал-диаграммой или таблицей значений атрибутов. Пусть имеется текстовое описание некоторого фрагмента учебной базы данных. Из этого описания необходимо выделить предложения,



указывающие на взаимосвязь между элементами (атрибутами) предметной области. Например:

- a) Составлен список, в котором каждой фамилии и инициалам (ФИО) курсанта соответствует номер по порядку (№ п.п.).
- b) Каждый курсант (ФИО) имеет одну зачетную книжку (ЗачКнижка№).
- c) По зачетной книжке (ЗачКнижка№) курсанта определяется средний балл (СрБалл).
- d) Таким образом, каждый курсант (ФИО) учится со средним баллом (СрБалл).
- e) Каждый курсант (ФИО) изучает множество дисциплин (УчДисциплина) в течение множества семестров (Семестр№).
- f) В свою очередь, в течение каждого семестра (Семестр№) множество курсантов (ФИО) изучают множество дисциплин (УчДисциплина).
- g) А также, учебная дисциплина (УчДисциплина) может быть запланирована для изучения в течение нескольких семестров (Семестр№), и изучается множеством курсантов (ФИО).
- h) Каждая учебная дисциплина (УчДисциплина) содержит множество заданий (Задание№) и состоит из множества занятий (Занятие№).
- i) Множество заданий (Задание№) по каждой учебной дисциплине (УчДисциплина) выдано курсанту (№ п.п.).
- j) Курсант (ФИО) по каждой учебной дисциплине (УчДисциплина) выполняет множество заданий (Задание№).

## 2. Выявление функциональных зависимостей.

Подготовленное описание определяет следующие функциональные (ФЗ, или FD) и многозначные

зависимости (МЗ, или MVD) между атрибутами предметной области:

- a)  $\PhiЗ : \PhiИО \rightarrow \text{№ п.п.}$
- b)  $\PhiЗ : \PhiИО \rightarrow \text{ЗачКнижка№}$
- c)  $\PhiЗ : \text{ЗачКнижка№} \rightarrow \text{СрБалл}$
- d)  $\PhiЗ : \PhiИО \rightarrow \text{СрБалл}$
- e)  $\text{МЗ} : \PhiИО \rightarrow \rightarrow \text{УчДисциплина}$   
 $\PhiИО \rightarrow \rightarrow \text{Семестр№}$
- f)  $\text{МЗ} : \text{Семестр№} \rightarrow \rightarrow \PhiИО$   
 $\text{Семестр№} \rightarrow \rightarrow \text{УчДисциплина}$
- g)  $\text{МЗ} : \text{УчДисциплина} \rightarrow \rightarrow \text{Семестр№}$   
 $\text{УчДисциплина} \rightarrow \rightarrow \PhiИО$
- h)  $\text{МЗ} : \text{УчДисциплина} \rightarrow \rightarrow \text{Задание№}$   
 $\text{УчДисциплина} \rightarrow \rightarrow \text{Занятие№}$
- i)  $\PhiЗ : \text{Задание№}, \text{УчДисциплина} \rightarrow \text{№ п.п.}$
- j)  $\text{МЗ} : \PhiИО, \text{УчДисциплина} \rightarrow \rightarrow \text{Задание№}$

**Обратите внимание** на то, что фразы каждого из пунктов 2.а) – 2.ж) соответствуют предложениям из пунктов 1.а) – 1.ж), соответственно.

3. Построение овал-диаграммы. Овал-диаграмма формируется в соответствии с теми зависимостями, которые были выявлены в пункте 2. Таким образом, заданную предметную область можно представить в виде отношения R1 (см. Рисунок 1.1).
4. Нормализация отношений предметной области. Каждая реляционная таблица должна соответствовать, по крайней мере, **первой нормальной форме** (1НФ, или 1NF; см. Таблицу 1.1).

**Обратите внимание** на то, что в Таблице 1.1 нет, и не может быть пустых ячеек. Кроме того, в Таблице 1.1 нет, и не может быть повторяющихся строк.

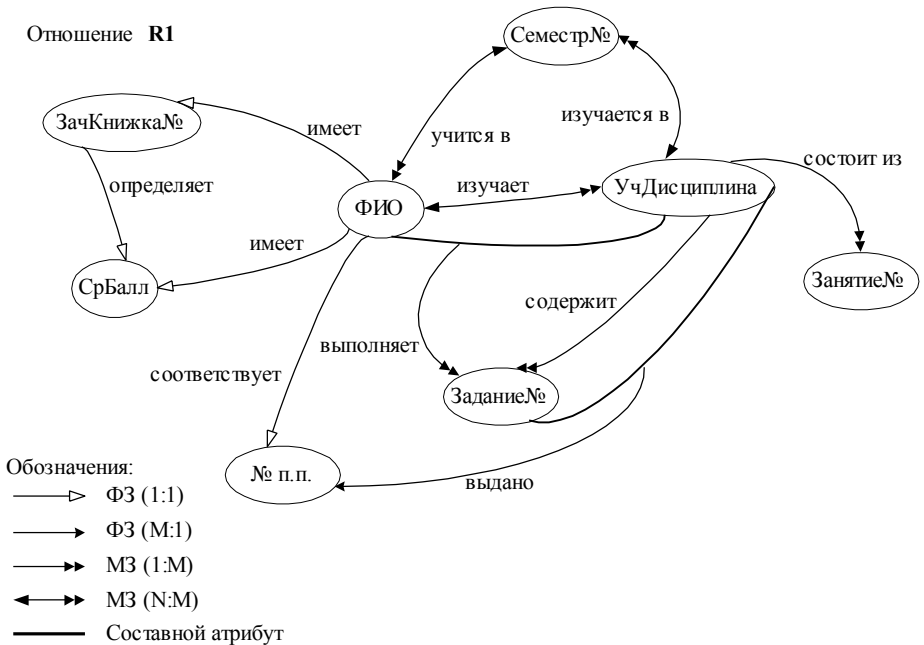


Рисунок 1.1. Овал-диаграмма отношения R1

5. Выявление возможных ключей. Для выявления возможных ключей в ИФФ необходимо выявить атрибуты (в том числе и составные, связанные ФЗ или МЗ), имеющие уникальные значения в данном отношении.

В Таблице 1.1 таких атрибутов два:

- а) ФИО, УчДисциплина, Занятие№ – 16 значений;
- б) УчДисциплина, Занятие№, Задание№ – 16 значений.

Кроме того, в качестве возможных ключей могут быть выбраны атрибуты, связанные многозначными зависимостями. В данном примере таким атрибутом является

- с) Семестр№.

Таблица 1.1. Отношение RI

Зач Книжка №	Ср Балл	ФИО	№ п.п	Семестр №	Уч Дисциплина	Занятие №	Задание №
ЗК_1	4,3	Ивлев Б.Б.	1	С1	МАнализ	М31	М1
ЗК_1	4,3	Ивлев Б.Б.	1	С1	МАнализ	М32	М1
ЗК_1	4,3	Ивлев Б.Б.	1	С2	МАнализ	М321	М3
ЗК_1	4,3	Ивлев Б.Б.	1	С2	МАнализ	М322	М3
ЗК_2	4,5	Пикалин В.В.	2	С1	МАнализ	М31	М2
ЗК_2	4,5	Пикалин В.В.	2	С1	АГеометрия	А31	А1
ЗК_2	4,5	Пикалин В.В.	2	С2	АГеометрия	А318	А3
ЗК_3	4,4	Сигалов А.А.	3	С3	МЛогика	Л31	Л1
ЗК_3	4,4	Сигалов А.А.	3	С3	СУБД	С31	Б1
ЗК_4	4,8	Суровов Е.Е.	4	С3	СУБД	С31	Б2
ЗК_4	4,8	Суровов Е.Е.	4	С3	МЛогика	Л31	Л2
ЗК_5	4,5	Сусанов Г.Г.	5	С1	АГеометрия	А31	А2
ЗК_5	4,5	Сусанов Г.Г.	5	С2	АГеометрия	А318	А4
ЗК_6	5,0	Ярцев Д.Д.	6	С3	МЛогика	Л31	Л3
ЗК_6	5,0	Ярцев Д.Д.	6	С4	МЛогика	Л316	Л4
ЗК_6	5,0	Ярцев Д.Д.	6	С4	СУБД	С317	Б3

Для наглядности овал-диаграммы атрибуты, входящие в состав возможных ключей необходимо подчеркнуть (см. Рисунок 1.2).

6. Приведение отношений ко второй нормальной форме.  
Чтобы привести отношение ко второй нормальной форме (2НФ, или 2NF) необходимо выполнить следующие действия:

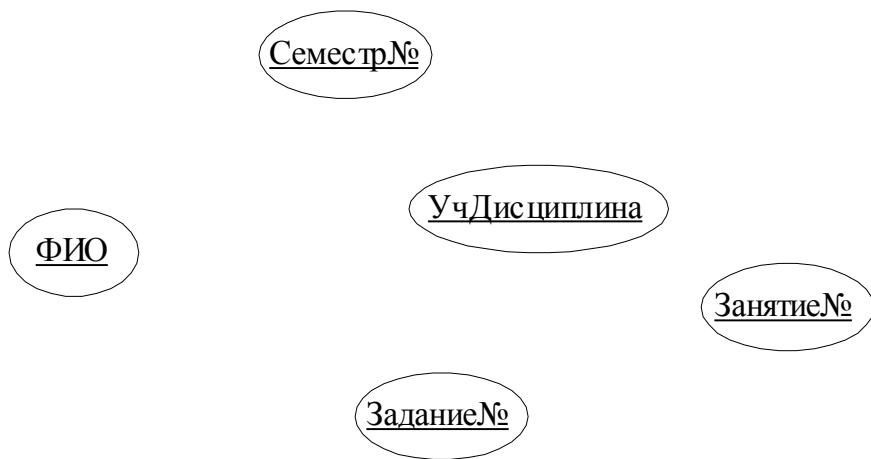


Рисунок 1.2. Диаграмма возможных ключей

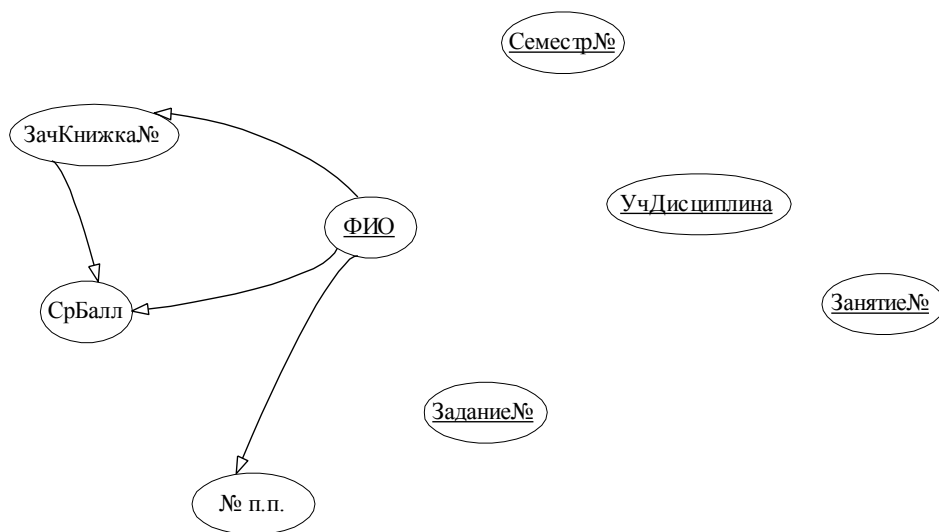


Рисунок 1.3. Диаграмма простых ФЗ

- а) разместить отдельно на диаграмме каждый из атрибутов, входящих в состав возможных ключей (см. Рисунок 1.2);
- б) добавить к каждому из атрибутов на Рисунке 1.2 только те атрибуты, которые *зависят* от данных ключевых по ФЗ, и восстановить между ними *все* связи (см. Рисунок 1.3);
- с) дополнить полученную овал-диаграмму теми отношениями, которые выпали в результате выполнения действий 6.а) – 6.б), но имели место на диаграмме 1НФ (см. Рисунки 1.1, 1.4).

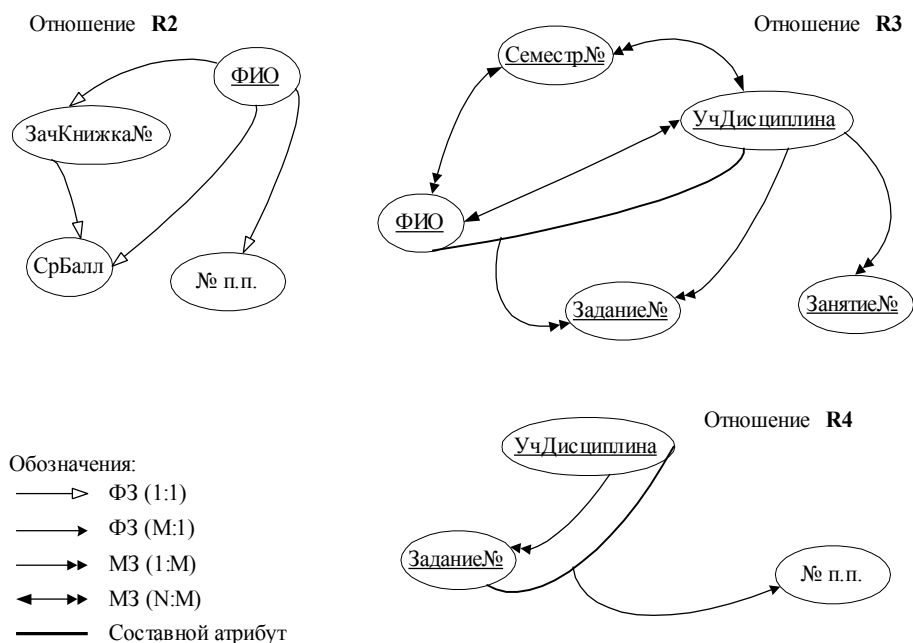


Рисунок 1.4. Овал-диаграмма совокупности отношений в 2НФ

В результате получим совокупность отношений в 2НФ (см. Рисунок 1.4).

**Обратите внимание** на следующие моменты:

- вторая нормальная форма строится относительно возможных ключей;
- совокупность отношений в 2НФ избыточна по числу атрибутов, но эта избыточность необходима в связи со следующим:
  - для восстановления связи без потерь между отношениями R2, R4 по атрибуту № п.п.;
  - для восстановления связи без потерь между отношениями R3, R4 по составному атрибуту <УчДисциплина, Задание№>. Кроме того, отношение R4, в отличие от отношения R1 (см. Рисунок 1.1), позволяет хранить ценные и полезные задания, которые, тем не менее, оказались не выданными по различным причинам ни одному из курсантов. Такое *состояние отношения R1* называется *аномалией добавления данных*;
- атрибут № п.п. не может быть добавлен к отношению R3, так как он не зависит ни от одного из ключевых атрибутов Занятие№ и Семестр№.

7. Приведение отношений к третьей нормальной форме.  
 Чтобы привести отношение к третьей нормальной форме (3НФ, или 3NF) необходимо в совокупности отношений 2НФ выявить и устранить транзитивные зависимости между атрибутами. Поскольку транзитивная зависимость

определена *только* для  $\Phi 3$ , то в данном примере (см. Рисунок 1.4) транзитивной является связь между атрибутами ФИО и СрБалл. Таким образом, отношение R2 может быть декомпозировано на два отношения R5 и R6 (см. Рисунок 1.5).

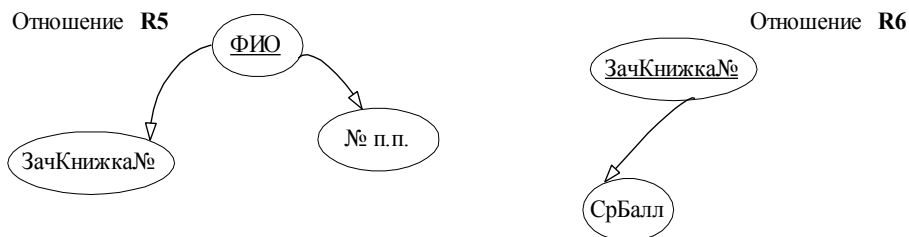


Рисунок 1.5. Овал-диаграмма декомпозирующих отношений в ЗНФ

В результате получим совокупность отношений в ЗНФ (см. Рисунок 1.5).

**Обратите внимание** на следующие моменты:

- поскольку в отношении между атрибутами ФИО и СрБалл связь была функциональной, то и декомпозиция отношения R2, и соединение отношений R5, R6 по атрибуту ЗачКнижка№ будет происходить без потерь;
- атрибут ЗачКнижка№ становится ключевым в отношении R6. Кроме того, отношение R6 позволяет устранить *аномалию добавления данных*, которая характерна для отношения R2. Она выражается в том, что если курсант не сдал еще ни одного экзамена или зачета с оценкой, то невозможно вычислить средний балл. И если в отношении R6, в данном случае, для



соответствующей зачетной книжки не требуется вносить никаких данных, то в отношении R2 потребуется что-то внести. А это нельзя делать по определению реляционного отношения;

- поскольку в отношениях R3, R4 (см. Рисунок 1.4) нет транзитивных связей (по определению), то эти отношения также находятся в 3НФ.

8. Результаты нормализации предметной области, изображенной на Рисунке 1.1 могут быть представлены в виде совокупности реляционных Таблиц 1.2–1.5.

***Обратите внимание*** на следующие моменты:

- в Таблицах 1.3, 1.4 и 1.6 произошло сокращение числа строк, однако, исходное отношение R1 будет восстанавливаться без потери данных (проверьте самостоятельно) в соответствии с правилами выполнения операций над реляционными отношениями;
- в заголовках Таблиц 1.2, 1.3, 1.4 и 1.6 ключевые атрибуты выделены шрифтом с подчеркиванием. Таблицы 1.2 и 1.3 содержат составные ключи;
- в рассматриваемых таблицах данные упорядочены по значениям ключевых атрибутов, при этом в Таблице 1.3 произошло переупорядочивание строк в соответствии с определением составного ключа: <УчДисциплина, Задание№>.

Таблица 1.2. Отношение R3

<u>ФИО</u>	<u>Семестр</u> <u>№</u>	<u>Уч</u> <u>Дисциплина</u>	<u>Занятие</u> <u>№</u>	<u>Задание№</u>
Ивлев Б.Б.	С1	МАнализ	М31	М1
Ивлев Б.Б.	С1	МАнализ	М32	М1
Ивлев Б.Б.	С2	МАнализ	М321	М3
Ивлев Б.Б.	С2	МАнализ	М322	М3
Пикалин В.В.	С1	МАнализ	М31	М2
Пикалин В.В.	С1	АГеометрия	А31	А1
Пикалин В.В.	С2	АГеометрия	А318	А3
Сигалов А.А.	С3	МЛогика	Л31	Л1
Сигалов А.А.	С3	СУБД	С31	Б1
Суровов Е.Е.	С3	СУБД	С31	Б2
Суровов Е.Е.	С3	МЛогика	Л31	Л2
Сусанов Г.Г.	С1	АГеометрия	А31	А2
Сусанов Г.Г.	С2	АГеометрия	А318	А4
Ярцев Д.Д.	С3	МЛогика	Л31	Л3
Ярцев Д.Д.	С4	МЛогика	Л316	Л4
Ярцев Д.Д.	С4	СУБД	С317	Б3

Таблица 1.3. Отношение R4

№ п.п.	Уч Дисциплина	Задание№
1	МАнализ	М1
1	МАнализ	М3
2	АГеометрия	А1
2	АГеометрия	А3
2	МАнализ	М2
3	МЛогика	Л1
3	СУБД	Б1
4	МЛогика	Л2
4	СУБД	Б2
5	АГеометрия	А2
5	АГеометрия	А4
6	МЛогика	Л3
6	МЛогика	Л4
6	СУБД	Б3

Таблица 1.4. Отношение R5

Зач Книжка№	ФИО	№ п.п.
ЗК_1	Ивлев Б.Б.	1
ЗК_2	Пикалин В.В.	2
ЗК_3	Сигалов А.А.	3
ЗК_4	Суровов Е.Е.	4
ЗК_5	Сусанов Г.Г.	5
ЗК_6	Ярцев Д.Д.	6

Таблица 1.5. Отношение R6

Зач Книжка№	Ср Балл
ЗК_1	4,3
ЗК_2	4,5
ЗК_3	4,4
ЗК_4	4,8
ЗК_5	4,5
ЗК_6	5,0

9. Выводы. Исходная предметная область (см. Рисунок 1.1) может быть представлена совокупностью отношений R3, R4, R5 и R6, находящихся в ЗНФ. Для каждого из них определены возможные (и/или первичные) ключи, соответственно: <ФИО, Семестр№, УчДисциплина, Занятие№, Задание№>, <УчДисциплина, Задание№>, <ФИО>, <ЗачКнижка№>.

Отношения в ЗНФ устраняют аномалии 1 и 2НФ (см. пп. 6, 7).

В данном примере для отношения R3 в качестве первичного ключа достаточно принять составной атрибут <ФИО, Занятие№>, (или же другой; *определите самостоятельно, сколько их может быть ?*).

Таким образом, Таблицы 1.2, 1.3, 1.4 и 1.5 представляют собой отношения, которые могут быть размещены на носителе информации в виде реляционной базы данных.

### **Оформление и содержание отчета**

Отчет по практическому занятию должен быть написан аккуратным почерком или распечатан на принтере. При подготовке материалов отчета необходимо выполнять общие требования ЕСКД, которые предъявляются к разработке текстовых документов.

В отчете должны быть указаны следующие данные и представлены следующие материалы:

- фамилия и номер учебной группы курсанта;
- тема, цель и номер задания;
- содержание задания: исходные данные и задача;
- результаты выполнения по каждому (с 1 по 8) пункту задания: ФЗ и МЗ, овал-диаграммы, таблицы;
- выводы, содержащие пояснения к дальнейшему использованию результатов.

### **Литература к практическому занятию № 1**

1. Дейт К. Введение в системы баз данных. : Пер. с англ. –6-е изд –К.: Наука, 1998 г. –784с.
2. Хомоненко А.Д., Цыганков В.М., Мальцев М.Г. Базы данных: Учебник для высших учебных заведений/ Под ред. проф. А.Д. Хомоненко. –СПб: КОРОНА принт, 2000 г. –416с.
3. Озкарахан Э. Машины баз данных и управление базами данных –М: Мир, 1989 г. –696 с.
4. Атре Ш. Структурный подход к организации баз данных. –М.: Финансы и статистика, 1983 г. –317с.

## ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 2

**Тема занятия:** Способы приведения отношений к наивысшим формам нормализации.

**Цель занятия:** Приобретение навыков приведения реляционных отношений к наивысшим формам нормализации.

**Задание:** Привести отношения, находящиеся в 3НФ отношений к наивысшим формам нормализации.

### Методические указания по выполнению задания № 2

#### ***Накануне занятия:***

- ознакомиться с темой, целью и заданием;
- изучить лекционный материал и рекомендуемую литературу по заданной теме;
- выполнить задание по практическому занятию № 1;
- подготовить подсобный материал для составления письменного отчета по заданию;
- получить в библиотеке и иметь на занятии рекомендованную литературу.

#### ***В ходе занятия:***

- принять участие в разборе примера выполнения задания;
- самостоятельно привести отношения, нормализованные при выполнении задания по практическому занятию № 1 к наивысшим формам нормализации;
- подготовить письменный отчет о выполнении задания.

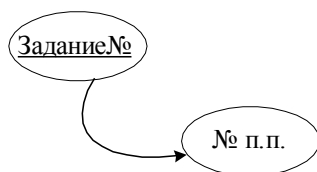
### Порядок выполнения задания № 2

1. Приведение отношений к нормальной форме Бойса-Кодда. Чтобы привести отношение к нормальной форме Бойса-Кодда (БКНФ, или BCNF) необходимо в каждом из отношений в 3НФ выявить детерминанты и атрибуты,

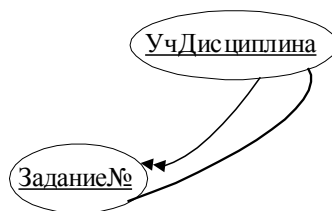
которые находятся в полной функциональной зависимости от детерминантов. Если в каком-либо отношении окажется несколько детерминантов и не все из них будут являться его возможными ключами, то это отношение потребуется декомпозировать так, чтобы в декомпозирующих отношениях не было атрибутов, не зависящих от каждого из детерминантов.

Так, в данном примере (см. Рисунок 1.4) отношения R5, R6 находятся в БКНФ (по определению; *проверьте !*), а наибольший интерес представляют отношения R4 и R3. В отношении R4, судя по данным из Таблицы 1.3, детерминантом является атрибут Задание№. Вследствие этого наблюдается ФЗ атрибута №п.п. от Задание№. В свою очередь атрибут Задание№ многозначно зависит от УчДисциплины, но атрибут УчДисциплина не является возможным

Отношение R7



Отношение R8



Обозначения:

————→ ФЗ (М:1)

————→ МЗ (1:М)

———— Составной атрибут

Рисунок 2.1. Овал-диаграмма отношений, декомпозирующих R4 в БКНФ

ключом в R4. Тогда отношение R4 необходимо декомпозировать следующим образом (см. Рисунок 2.1).

В отношении R3, судя по данным из Таблицы 1.2, детерминантами являются составные атрибуты <ФИО, Занятие№>, <Задание№, Занятие№>. Однако, в каждом из них нет “прямой” связи между указанными атрибутами. Напротив, они связаны через атрибут УчДисциплина. Кроме того, атрибут Семестр№ не находится в полной функциональной зависимости ни от одного из данных детерминантов. Следовательно, необходимо произвести декомпозицию отношения R3.

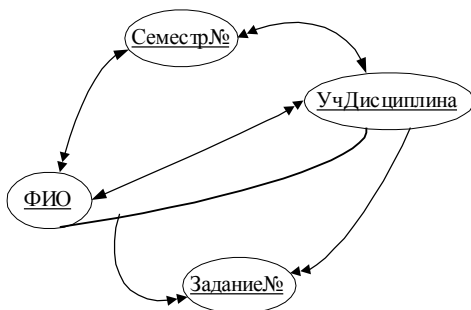
Рассматривая детерминант <ФИО, Занятие№>, необходимо отметить, что в полной функциональной зависимости от него находится только один атрибут: УчДисциплина. А в полной функциональной зависимости от детерминанта <Задание№, Занятие№> находятся атрибуты УчДисциплина и ФИО (см. Рисунок 1.4). Поэтому декомпозируемое отношение R3 может быть представлено следующими двумя вариантами декомпозиции (см. Рисунки 2.2, 2.3).

**Обратите внимание** на следующие моменты:

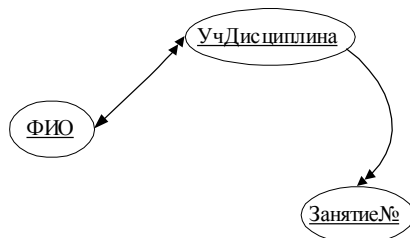
- в каждом из отношений R5, R6 (см. Таблицы 1.4, 1.5), R7, R8, R9, R10, R11, R12 присутствуют только детерминанты и зависимые от них атрибуты (на основе данных из Таблиц 1.2, 1.3 составьте таблицы отношений R7, R8, R9, R10, R11, R12 и *проверьте !*), поэтому они находятся в БКНФ. При этом отношение R9 представляет собой *журнал учета выполнения заданий*;



Отношение R9



Отношение R10



Обозначения:

ФЗ (M:1)

МЗ (1:M)

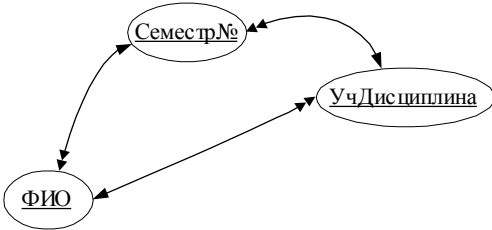
МЗ (N:M)

Составной атрибут

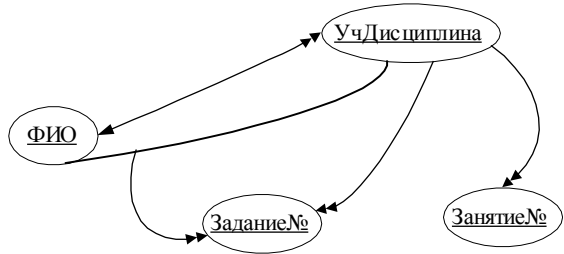
Рисунок 2.2. Первый вариант отношений, декомпозирующих R3 в БКНФ

- в отношениях R10, R11 имеется по одному детерминанту в составе следующих атрибутов: <ФИО, Занятие№> и <ФИО, Семестр№, УчДисциплина>, соответственно. При этом, в таблице значений отношения R11 произойдет сокращение числа строк (проверьте !);
- в отношении R9 имеется два детерминанта в составе следующих атрибутов: <ФИО, УчДисциплина> и <Задание№>. При этом, в таблице значений отношения R9 произойдет сокращение числа строк (проверьте !);

Отношение R11



Отношение R12



Обозначения:

→ ФЗ (M:1)

→ MЗ (1:M)

↔ MЗ (N:M)

— Составной атрибут

Рисунок 2.3. Второй вариант отношений, декомпозирующих R3 в БКНФ

- в отношении R12 также два детерминанта:  $\langle \text{ФИО}, \text{Занятие№} \rangle$  и  $\langle \text{Занятие№}, \text{Задание№} \rangle$ ;
- каждый из детерминантов может быть назначен первичным ключом рассмотренных отношений;
- поскольку в предметной области (см. Рисунок 1.1) не было зафиксировано отношение между атрибутами Семестр№ и Занятие№, Семестр№ и Задание№, то не следует ожидать от декомпозирующих отношений R9, R10 или R11, R12 отсутствия потери данных при восстановлении кортежей значений  $\langle \text{Семестр№}, \text{Занятие№} \rangle$ ,  $\langle \text{Семестр№}, \text{Задание№} \rangle$ ;
- сравнивая отношения R8, R9 и R8, R12 необходимо отметить избыточность атрибутов,

входящих в отношение R8. Ее можно устранить двумя способами:

- d) вычеркнуть отношение R8 из дальнейшего рассмотрения или
- e) удалить атрибут Задание№ из отношений R9 и R12. При этом отношение R9 преобразуется в R11, а отношение R12 – в R10. Тогда, окончательная совокупность отношений в БКНФ будет представлена отношениями R5, R6, R7, R8, R10, R11.

2. Приведение отношений к четвертой нормальной форме.  
Чтобы привести отношение к четвертой нормальной форме (4НФ, или 4NF) необходимо в каждом из отношений в БКНФ выявить и устранить не тривиальные МЗ. С учетом результатов пункта 1.b) данного занятия не тривиальная МЗ присутствует лишь в отношении R10 между атрибутом УчДисциплина и ФИО | Занятие№. Тогда декомпозиция отношения R10 может быть выполнена следующим образом (см. Рисунок 2.4).

**Обратите внимание** на следующие моменты:

- первичными ключами отношений R13 и R14 являются их детерминанты: <ФИО, УчДисциплина> в R13 и <Занятие№> в R14;
- отталкиваясь от результатов пункта а) данного занятия необходимо отметить, что в отношениях R9 и R12 содержатся тривиальные МЗ (<ФИО, УчДисциплина> →→ Семестр№ – Занятие№) (УчДисциплина →→ Занятие№ – <ФИО, Задание№>). После декомпозиции (приведения к 4НФ), с учетом того, что декомпозирующее отношение <ФИО, УчДисциплина, Задание№> для устранения избыточности и без потери смысла может быть замещено отношением R4 (а

далее R7 и R8; см. Рисунки 1.4, 2.1), а также после устранения одинаковых (по составу атрибутов и связей между ними) отношений получится та же совокупность декомпозирующих отношений R13, R14 и R11, находящихся в 4НФ (см. Рисунки 2.3, 2.4);



Обозначения:

→→ M3 (1:M)

↔→ M3 (N:M)

— Составной атрибут

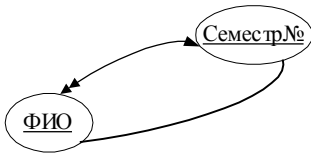
Рисунок 2.4. Овал-диаграмма декомпозирующих отношений в 4НФ

- декомпозиция отношений R10, R12 и выделение из них отношения R14 устраняет аномалии (определите в чем они выражаются ?) добавления, изменения и удаления данных из R10, R12. При этом отношение R14 представляет собой порядок изучения (структуру) учебной дисциплины.

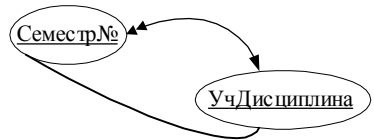
3. Приведение отношений к пятой нормальной форме. Чтобы привести отношение к пятой нормальной форме (5НФ, или 5NF, PJ/NF) необходимо в каждом из отношений в 4НФ выявить зависимость по соединению

без потерь его не пустых подмножеств. Рассматривая отношения R11, R13, R14, находящиеся в 4НФ необходимо отметить, что указанная зависимость подмножеств их атрибутов присутствует только в отношении R11. Эти подмножества определяют совокупность декомпозирующих отношений. Они представлены на Рисунке 2.5.

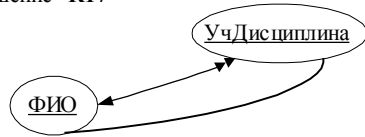
Отношение R15



Отношение R16



Отношение R17



Обозначения:

↔ M3 (N:M)

— Составной атрибут

Рисунок 2.5. Овал-диаграмма декомпозирующих отношений в 5НФ

**Обратите внимание** на следующие моменты:

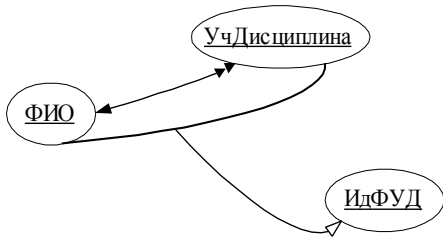
- отношения R13 (см. Рисунок 2.4) и R17 идентичны, поэтому одно из них необходимо исключить из окончательной совокупности НФ по заданной предметной области;
- в каждом из отношений R15, R16, R17 наблюдается сокращение строк по сравнению с исходным отношением R11 (см. Рисунок 2.3). Однако, при попытке восстановить отношение R11 путем соединения R15, R16, R17 в

результатирующем отношении появятся две новые (по сравнению с отношением R11) строки (*определите какие !*);

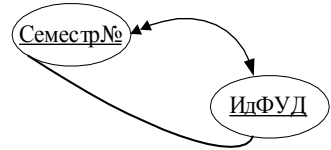
- первичными ключами отношений R15, R16 и R17 являются их детерминанты: <ФИО, Семестр№>, <Семестр№, УчДисциплина> и <ФИО, УчДисциплина>, соответственно;
- полученная совокупность отношений R15, R16 и R17 могла быть получена и ранее: при формировании 2НФ. Для этого необходимо было воспользоваться правилом формирования составных атрибутов, при котором каждая МЗ типа (М:N) преобразуется в составной атрибут. Тогда из отношения R3 сразу получится совокупность R12, R15 и R16, что значительно сократит число этапов нормализации;
- декомпозиция отношения R11 устраняет аномалии (*определите в чем они выражаются ?*) добавления, изменения и удаления данных из таблицы, соответствующей этому отношению;
- на Рисунке 2.5 представлен классический способ приведения отношений к 5НФ. Здесь предлагается другой способ (см. Рисунок 2.6), который декомпозирует R11 на два отношения в БКНФ и 5НФ. Восстановление отношения R11 путем соединения этих двух отношений будет производиться без потерь и лишних строк (*проверьте !*). Кроме того, он лучше предыдущего способа, поскольку, несмотря на то, что введен новый атрибут ИдФУД (счетчик кортежей значений <ФИО, УчДисциплина>), этот способ характеризуется меньшим числом атрибутов и их значений. Еще одним преимуществом данного способа является то,

что он пригоден для декомпозиции отношений типа R11, атрибуты которых связаны МЗ не только как (М:N), но и (1:М) (*проверьте !*).

Отношение R18



Отношение R19



Обозначения:

—> ФЗ (1:1)

↔ МЗ (N:M)

— Составной атрибут

Рисунок 2.6. Декомпозиция R11 на отношения в БКНФ и 5НФ

Суть такой декомпозиции состоит в том, что в отношении R11 выделяется любая пара атрибутов. В данном примере это <ФИО, УчДисциплина>, которые формируют новое отношение R18. Затем в R18 с помощью атрибута ИдФУД фиксируются все присутствующие на данный момент кортежи значений <ФИО, УчДисциплина>. Между составным атрибутом <ФИО, УчДисциплина> и ИдФУД устанавливается ФЗ типа (1:1). Далее формируется отношение R19: <Семестр№, ИдФУД>. Между атрибутами Семестр№ и ИдФУД устанавливается МЗ типа (М:N).

Первичными ключами отношений R18 и R19 являются их детерминанты: <ФИО, УчДисциплина> и <Семестр№, ИдФУД>, соответственно. При этом отношение R13 также является избыточным и должно быть исключено из окончательной совокупности отношений.

4. Результаты нормализации предметной области, изображенной на Рисунке 1.1, окончательно, могут быть представлены в виде следующих реляционных таблиц: Таблицы 1.4, 1.5 (см. пп. 8 практического занятия № 1), а также Таблицы 2.1 – 2.6.

Таблица 2.1.Отношение R7

<u>Задание№</u>	№ п.п.
A1	2
A2	5
A3	2
A4	5
B1	3
B2	4
B3	6
Л1	3
Л2	4
Л3	6
Л4	6
M1	1
M2	2
M3	1

Таблица 2.2.Отношение R8

<u>Уч Дисциплина</u>	<u>Задание №</u>
АГеометрия	A1
АГеометрия	A2
АГеометрия	A3
АГеометрия	A4
МАнализ	M1
МАнализ	M2
МАнализ	M3
МЛогика	Л1
МЛогика	Л2
МЛогика	Л3
МЛогика	Л4
СУБД	Б1
СУБД	Б2
СУБД	Б3



Таблица 2.3.Отношение R14

<u>Уч</u> <u>Дисциплина</u>	<u>Занятие</u> <u>№</u>
АГеометрия	А31
АГеометрия	А318
МАнализ	М31
МАнализ	М32
МАнализ	М321
МАнализ	М322
МЛогика	Л31
МЛогика	Л316
СУБД	С31
СУБД	С317

Таблица 2.4.Отношение R15

<u>ФИО</u>	<u>Семестр</u> <u>№</u>
Ивлев Б.Б.	С1
Ивлев Б.Б.	С2
Пикалин В.В.	С1
Пикалин В.В.	С2
Сигалов А.А.	С3
Суровов Е.Е.	С3
Сусанов Г.Г.	С1
Сусанов Г.Г.	С2
Ярцев Д.Д.	С3
Ярцев Д.Д.	С4

Таблица 2.5.Отношение R16

<u>Семестр</u> <u>№</u>	<u>Уч</u> <u>Дисциплина</u>
С1	АГеометрия
С1	МАнализ
С2	АГеометрия
С2	МАнализ
С3	МЛогика
С3	СУБД
С4	МЛогика
С4	СУБД

Таблица 2.6.Отношение R17

<u>ФИО</u>	<u>Уч</u> <u>Дисциплина</u>
Ивлев Б.Б.	МАнализ
Пикалин В.В.	АГеометрия
Пикалин В.В.	МАнализ
Сигалов А.А.	МЛогика
Сигалов А.А.	СУБД
Суровов Е.Е.	МЛогика
Суровов Е.Е.	СУБД
Сусанов Г.Г.	АГеометрия
Ярцев Д.Д.	МЛогика
Ярцев Д.Д.	СУБД

Таблица 2.7. Отношение R18

<b><u>ФИО</u></b>	<b><u>Уч Дисциплина</u></b>	<b><u>ИдФУД</u></b>
Ивлев Б.Б.	МАнализ	1002
Пикалин В.В.	АГеометрия	2001
Пикалин В.В.	МАнализ	2002
Сигалов А.А.	МЛогика	3003
Сигалов А.А.	СУБД	3004
Суровов Е.Е.	МЛогика	4003
Суровов Е.Е.	СУБД	4004
Сусанов Г.Г.	АГеометрия	5001
Ярцев Д.Д.	МЛогика	6003
Ярцев Д.Д.	СУБД	6004

Таблица 2.8. Отношение R19

<b><u>Семестр №</u></b>	<b><u>ИдФУД</u></b>
С1	1002
С1	2001
С1	2002
С1	5001
С2	1002
С2	2001
С2	5001
С3	3003
С3	3004
С3	4003
С3	4004
С3	6003
С4	6003
С4	6004

В данном примере для отношения R11 может быть предложено два варианта декомпозиции (см. Рисунки 2.5, 2.6). Наилучшим из них по наиболее точному восстановлению предметной области (см. Рисунок 1.1) является декомпозиция на отношения R18, R19.

Таким образом, окончательно, для размещения на носителе информации в виде реляционной базы данных предлагается выбрать Таблицы 1.4, 1.5, 2.1–2.3, а также Таблицы 2.7, 2.8.

### **Оформление и содержание отчета**

Отчет по практическому занятию должен быть написан аккуратным почерком или распечатан на принтере. При подготовке материалов отчета необходимо выполнять общие требования ЕСКД, которые предъявляются к разработке текстовых документов.

В отчете должны быть указаны следующие данные и представлены следующие материалы:

- фамилия и номер учебной группы курсанта;
- тема, цель и номер задания;
- содержание задания: исходные данные и задача;
- результаты выполнения по каждому (с 1 по 4) пункту задания: ФЗ и МЗ, овал-диаграммы, таблицы;
- выводы, содержащие пояснения к дальнейшему использованию результатов.

### **Литература к практическому занятию № 2**

1. Дейт К. Введение в системы баз данных.: Пер. с англ. –6-е изд –К.: Наука, 1998 г. –784с.
2. Хомоненко А.Д., Цыганков В.М., Мальцев М.Г. Базы данных: Учебник для высших учебных заведений/ Под

ред. проф. А.Д. Хомоненко. –СПб: КОРОНА принт, 2000 г. –416с.

1. Озкарахан Э. Машины баз данных и управление базами данных –М: Мир, 1989 г. –696 с.
2. Атре Ш. Структурный подход к организации баз данных. –М.: Финансы и статистика, 1983 г. –317с.

### ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 3

**Тема занятия:** Порядок формирования матрицы функциональных и многозначных отношений.

**Цель занятия:** Приобретение навыков матричного представления отношений в предметной области.

**Задание:** Сформировать матрицы функциональных и многозначных отношений.

#### Методические указания по выполнению задания № 3

##### *Накануне занятия:*

- ознакомиться с темой, целью и заданием;
- изучить лекционный материал и рекомендуемую литературу по заданной теме;
- выполнить задание по практическому занятию № 2;
- подготовить подсобный материал для составления письменного отчета по заданию;
- получить в библиотеке и иметь на занятии рекомендованную литературу.

##### *В ходе занятия:*

- принять участие в разборе примера выполнения задания;
- получить у преподавателя вариант задания;
- самостоятельно выполнить задание;
- подготовить письменный отчет о выполнении задания.

#### Порядок выполнения задания № 3

1. Анализ исходных данных. Исходные данные для выполнения задания № 3, как правило, задаются в виде овал-диаграммы предметной области. Отношения предметной области должны быть нормализованы и представлены, по крайней мере, в ЗНФ (см. практическое занятие № 1). Далее необходимо перенумеровать все

атрибуты. Например, для атрибутов предметной области, нормализованной на практическом занятии № 2, можно задать следующие номера:

- |               |                 |             |
|---------------|-----------------|-------------|
| 1. ЗачКнижка№ | 4. СрБалл       | 7. Занятие№ |
| 2. ФИО        | 5. Задание№     | 8. ИдФУД    |
| 3. № п.п.     | 6. УчДисциплина | 9. Семестр№ |

2. Формирование матрицы функциональных отношений.

Рассматриваемая матрица является квадратной, размерность которой соответствует числу атрибутов предметной области. В каждую ячейку на главной диагонали, а также в каждую ячейку с индексом  $\langle \text{№строки}A1, \text{№столбца}A2 \rangle$ , соответствующую ФЗ  $A1 \rightarrow A2$  или  $M3 \ A2 \rightarrow A3$ , заносится 1. Остальные ячейки могут оставаться пустыми, или заполняться нулями.

Для совокупности отношений R5, R6, R7, R8, R14, R18, R19 (см. практическое занятие № 2) матрица функциональных отношений (**MF**) примет вид Таблицы 3.1.

**Обратите внимание** на то, что, если для каждого **зависимого** атрибута в МЗ в Таблице 3.1 вместо единицы записать М, то матрица **MF** совместно с рассчитанным числом единиц в каждой строке может быть использована для выявления ключевых атрибутов (см. пп.5 практического занятия № 1).

3. Перенумерация атрибутов является первым шагом процедуры преобразования матрицы **MF** к треугольному виду. Чтобы перенумеровать атрибуты, необходимо подсчитать число единиц в каждой строке. Затем, присвоить первый номер атрибуту с наибольшим числом

единиц. Далее, номера присваиваются атрибутам в порядке убывания числа единиц (см. Таблицу 3.1).

Таблица 3.1. Исходная  $MF$

		1	2	3	4	5	6	7	8	9	Число ед.	Новый номер
ЗачКнижка№	1	1			1						2	3
ФИО	2	1	1	1			1		1		5	1
№ п.п.	3			1							1	7
СрБалл	4				1						1	8
Задание№	5			1		1					2	4
УчДисциплина	6		1			1	1	1	1		5	2
Занятие№	7							1			1	9
ИдФУД	8								1	1	2	5
Семестр№	9								1	1	2	6

4. Преобразование матрицы  $MF$  к треугольному виду выполняется следующим образом. Формируется новая квадратная матрица  $\Delta MF$  с пустыми ячейками. Главная диагональ заполняется единицами. Строки и столбцы в  $\Delta MF$  соответствуют атрибутам с новыми номерами (см. предыдущий пункт). Используя рассматриваемые отношения необходимо заполнить матрицу  $\Delta MF$  единицами так же, как было указано в пп.2 данного практического занятия. Эти же манипуляции можно выполнить, перемещая каждую единицу из ячейки  $\langle \text{№строки}A1, \text{№столбца}A2 \rangle$  матрицы  $MF$  в ячейку с индексом  $\langle \text{новый№атрибута}A1, \text{новый№атрибута}A2 \rangle$  матрицы  $\Delta MF$ . Так, например, единица из ячейки  $\langle 1,4 \rangle$  переместится в ячейку  $\langle 3,8 \rangle$ , и т.д. (см. Таблицу 3.2).





Таблица 3.3. Матрица  $\Delta MV$ 

		1	2	3	4	5	6	7	8	9
ФИО	1	1	М	1		1		1		
УчДисциплина	2	М	1		М	1				М
ЗачКнижка№	3			1					1	
Задание№	4				1			1		
ИдФУД	5					1	М			
Семестр№	6					М	1			
№ п.п.	7							1		
СрБалл	8								1	
Занятие№	9									1

7. Выводы. Полученная матрица  $\Delta MF$  по внешнему виду близка к треугольной. Единицы в ячейках  $\langle 2,1 \rangle$  и  $\langle 6,5 \rangle$  обусловлены наличием МЗ типа М:N между соответствующими атрибутами, что не позволяет получить классическую треугольную матрицу. Матрица  $\Delta MV$  учитывает наличие многозначных зависимостей между атрибутами предметной области, и может быть использована для дальнейшего анализа отношений заданной предметной области.

### Оформление и содержание отчета

Отчет по практическому занятию должен быть написан аккуратным почерком или распечатан на принтере. При подготовке материалов отчета необходимо выполнять общие требования ЕСКД, которые предъявляются к разработке текстовых документов.

В отчете должны быть указаны следующие данные и представлены следующие материалы:

- фамилия и номер учебной группы курсанта;

- тема, цель и номер задания;
- содержание задания: исходные данные и задача;
- результаты выполнения по каждому (с 1 по 5) пункту задания: таблицы, матрицы;
- выводы, содержащие пояснения к дальнейшему использованию результатов.

### **Литература к практическому занятию № 3**

1. Мартин Дж. Организация баз данных в вычислительных системах –М.: Мир, 1980 г. –662 с.
2. Озкарахан Э. Машины баз данных и управление базами данных –М.: Мир, 1989 г. –696 с.

## ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 4

**Тема занятия:** Порядок синтеза канонической структуры отношений.

**Цель занятия:** Приобретение навыков минимизации числа элементов и отношений в схеме базы данных.

**Задание:** Построить каноническую структуру отношений предметной области.

### Методические указания по выполнению задания № 4

#### *Накануне занятия:*

- ознакомиться с темой, целью и заданием;
- изучить лекционный материал и рекомендуемую литературу по заданной теме;
- выполнить задание по практическому занятию № 3;
- подготовить подсобный материал для составления письменного отчета по заданию;
- получить в библиотеке и иметь на занятии рекомендованную литературу.

#### *В ходе занятия:*

- принять участие в разборе примера выполнения задания;
- получить у преподавателя вариант задания;
- самостоятельно построить каноническую структуру отношений для заданной предметной области;
- подготовить письменный отчет о выполнении задания.

### Порядок выполнения задания № 4

1. Анализ исходных данных и инициализация. Исходными данными для синтеза канонической структуры отношений предметной области являются результаты выполнения заданий №№ 2 и 3.

$N_0$  – исходное число атрибутов предметной области;

$MF_{[N_0]}$ ,  $\Delta MF_{[N_0]}$ ,  $\Delta MV_{[N_0]}$  – матрицы отношений;

Для выполнения задания № 4 необходимо ввести следующие переменные:

$i$  – номер яруса (уровня структуры);

$N_i$  – общее число атрибутов на  $i$ -ом ярусе;

$K_i$  – число ключей на  $i$ -ом ярусе;

$P_l$  – отношение (сегмент структуры  $S$  предметной области);

$M_l$  – множество покрытых атрибутов в отношении  $P_l$ .

Далее необходимо задать начальные значения следующим переменным:

$i = 0$ ,  $K_i = 0$ ,  $\{P_l\} = \{\emptyset\}$ ,  $l = \overline{1, N_0}$ ;

$N_i = N_0 = 9$  (в соответствии с результатами практического занятия № 3);

2. Формирование матрицы функциональных отношений текущего яруса. Для яруса  $i = 0$  матрица  $MF_{[N_0]}$  была построена на практическом занятии № 3, поэтому можно перейти на шаг 3.

Для ярусов  $i > 0$  матрица  $MF_{[N_i - K_i]}$  формируется путем вычеркивания строк и столбцов соответствующих поглощаемому элементу, т.е. вычеркиваются все элементы с номером  $l \geq K_i$ .

3. Построение треугольной матрицы отношений. Если для  $i$ -го яруса матрица  $MF_{[N_i]}$  не является единичной ( $MF_{[N_i]} \neq E_{[N_i]}$ ), то производится построение матриц  $\Delta MF_{[N_i]}$ ,  $\Delta MV_{[N_i]}$  (см. практическое занятие № 3). Затем производится инициализация переменной  $j = 0$ .

Если  $MF_{[N_i]} = E_{[N_i]}$ , то переходим на шаг 4.

4. Построение очередного яруса структуры. Если  $i \neq 0$ , то из множества атрибутов имеющих одну единицу в строке

матрицы  $\Delta MF_{[Ni]}$  формируется  $i$ -ый ярус канонической структуры  $S$  отношений предметной области. При этом каждое отношение из списка  $\{P_l\}$  связывается с сегментами предыдущих ярусов. Затем, список  $\{P_l\}$  обнуляется.

Поскольку на первом проходе этого шага  $i = 0$ , то переходим на шаг 5.

5. Окончание алгоритма. Если матрица  $\Delta MF_{[Ni]}$  является единичной ( $\Delta MF_{[Ni]} = E_{[Ni]}$ ), то формируем окончательный вид канонической структуры  $S$ .

Матрица  $\Delta MF_{[N0]}$  (см. Таблицу 3.2) не является единичной, поэтому – переходим на шаг 6.

6. Определение ближайших связанных отношений. Для атрибута с номером  $N_i - j$  ищем множество  $M_l$  ближайших по номеру атрибутов, которые являются независимыми (как следствие они имеют не нулевые отметки в правых треугольниках матриц  $\Delta MF_{[N0]}$  и  $\Delta MV_{[N0]}$ ) и покрываются одним и тем же атрибутом.

**Обратите внимание** на следующие моменты:

- на первом проходе этого шага  $N_0 - j = 9$ ;
- атрибут 9 (Занятие№) покрывается атрибутом УчДисциплина;
- атрибут УчДисциплина покрывает множество  $M_9 = \{\text{Задание№, ИдФУД, Занятие№}\}$  независимых атрибутов (см. Таблицы 3.2, 3.3).

7. Формирование списков связей. Независимо от значения мощности множества  $|M_l|$  покрытых атрибутов в отношении  $P_l$ , для элемента с наибольшим номером  $l$  из множества  $M_l$  формируется отношение  $P_l := \{P_l \cup N_i - j\}$ .

Если  $|M_l| \leq 1$ , то выполняем переход на шаг 8.

Если же  $|M_l| > 1$ , то производим следующие действия:

- к) Увеличиваем число ключей на  $i$ -ом ярусе:  $K_i := K_i + 1$ .
- л) Меняем местами строку  $N_i - j$  и столбец  $K_i$  таким образом, что значения ячеек в строке  $N_i - j$ , находящиеся справа от столбца  $K_i$ , переносятся в ячейки столбца  $K_i$ , которые находятся выше строки  $N_i - j$ . Аналогично, старые значения столбца  $K_i$ , находящиеся выше строки  $N_i - j$  переносятся в ячейки строки  $N_i - j$ , находящиеся справа от столбца  $K_i$ .

**Обратите внимание** на следующие моменты:

- поскольку для  $l = N_0 - j = 9$  отношение  $P_9$  еще пусто, то  $P_9 := \{P_9 \cup 9\} = \{9\}$ ;
- поскольку  $|M_9| = 3$ , то  $K_0 := K_0 + 1 = 1$ , а в матрице  $\Delta MF_{[N_0]}$  ячейка  $\langle 2, 1 \rangle$  обращается в нуль.

8. Переход к обработке следующего атрибута этого же яруса. Сравниваем значения  $N_i - j$  и  $K_i + 1$ .

Если  $N_i - j > K_i + 1$ , то увеличиваем значение  $j$  на единицу, и возвращаемся на шаг 6.

Если  $N_i - j \leq K_i + 1$ , то увеличиваем значение  $i$  на единицу, и возвращаемся на шаг 2.

**Обратите внимание** на следующий момент:

- поскольку  $N_0 - j = 9$ , а  $K_0 + 1 = 2$ , следовательно  $j := j + 1 = 1$ , и необходимо вернуться на шаг 6.

9. Далее, в соответствии с данным алгоритмом, трижды повторяем шаги 6, 7, 8 при одних и тех же условиях.

**Обратите внимание** на следующие моменты:

- формируются отношения  $P_8 := \{8\}$ ,  $P_7 := \{7\}$ ,  $P_6 := \{6\}$ ;
- поскольку для атрибутов 8 и 6  $|M_l| = 1$ , то замены строк и столбцов не происходит. По

этой же причине число ключей  $K_i$  не увеличивается;

- для атрибута 7  $|M_l| = 3$ , поэтому число ключей увеличивается:  $K_i = 2$ ;
- по той же причине производится замена строк и столбцов. При этом ячейка  $\langle 1,2 \rangle$  обращается в нуль;
- поскольку всегда на этом шаге  $N_i - j > K_i + 1$ , то значение переменной  $j$  последовательно увеличивается до 4 и вновь возвращаемся на шаг 6.

10. На шаге 6 для атрибута  $N_0 - j = 5$  (ИдФУД) покрывающими являются атрибуты  $\langle \text{Семестр№} \rangle$  и  $\langle \text{ФИО, УчДисциплина} \rangle$  (см. Рисунок 2.6 и Таблицу 3.3).

**Обратите внимание** на следующие моменты:

- наличие МЗ ИдФУД  $\leftarrow \rightarrow \rightarrow \text{Семестр№}$  свидетельствует о том, что независимых покрытых атрибутов в ней нет. Поэтому  $M_5 = \{\text{ЗачКнижка№}, \text{Задание№}, \text{ИдФУД}, \text{№ п.п.}, \text{Занятие№}\}$ ;
- наибольшим номером элемента в множестве  $M_5$  является  $l=9$ , соответствующий атрибуту  $\langle \text{Занятие№} \rangle$ , поэтому продолжается формирование отношения  $P_9$   $P_9 = \{9, 5\}$ ;
- так как  $|M_5| = 5$ , то  $K_0 := K_0 + 1 = 3$ , а в матрице  $\Delta MF_{[N_0]}$  происходит замена строк и столбцов. При этом ячейка  $\langle 1,3 \rangle$  обращается в нуль, а ячейка  $\langle 2,3 \rangle$  обращается в единицу;
- поскольку  $(N_i - j = 5) > (K_i + 1 = 4)$ , то значение переменной  $j$  увеличивается до 5 и производится переход на шаг 6.

11. Далее повторяем шаги 6, 7, 8.





Таблица 4.2. Матрица  $\Delta MV$  после преобразований на 11 шаге

		1	2	3	4	5	6	7	8	9
ФИО	1	1				1		1		
УчДисциплина	2		1		М	1				М
ЗачКнижка№	3			1						
Задание№	4				1					
ИдФУД	5					1	М			
Семестр№	6					М	1			
№ п.п.	7							1		
СрБалл	8								1	
Занятие№	9								М	1

12. После шага 11,  $K_0 = 4$ , а  $i = 1$ . Поскольку значение  $i$  больше нуля, то, необходимо сформировать новую матрицу  $MF_{[Ni]}$  (см. Таблицу 4.3). При этом  $N_i \approx K_0 = 3$ . Для этого необходимо вычеркнуть строки и столбцы соответствующие поглощаемому элементу, т.е. вычеркиваются все элементы с номером  $l \geq K_0$ , при этом  $l = N_{i-1} - j$ .

Таблица 4.3. Матрица  $MF_{[Ni]}$ 

		1	2	3
ФИО	1	1		
УчДисциплина	2		1	1
ЗачКнижка№	3			1

13. Выполняется шаг 3, на котором формируются матрицы  $\Delta MF_{[Ni]}$ ,  $\Delta MV_{[Ni]}$  (см. Таблицы 4.4, 4.5).

Таблица 4.4. Матрица  $\Delta MF_{[Ni]}$ 

		1	2	3
УчДисциплина	1	1		1
ФИО	2		1	
ЗачКнижка№	3			1

Таблица 4.5. Матрица  $\Delta MV_{[Ni]}$ 

		1	2	3
УчДисциплина	1	1		М
ФИО	2		1	
ЗачКнижка№	3			1

Затем производится инициализация переменных:  $j = 0$ ,  $K_i = 0$ .

14. Выполняется шаг 4. Поскольку  $i = 1$ , то из множества атрибутов имеющих одну единицу в строке матрицы  $\Delta MF_{[Ni]}$  формируем первый ярус канонической структуры  $S$  отношений предметной области и связываем его с сегментами из списка  $\{ P_l \}$  нулевого яруса (см. Рисунок 4.1).

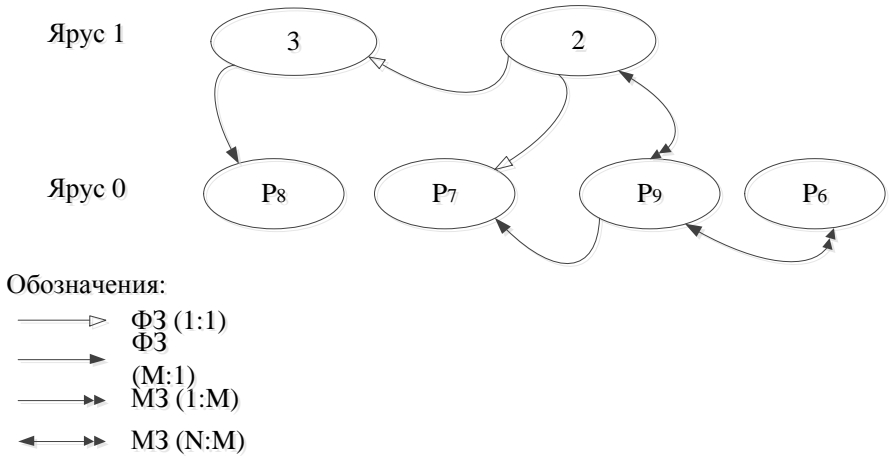


Рисунок 4.1. Первый и нулевой ярусы структуры  $S$

15. Выполняется шаг 5. Матрица  $\Delta MF_{[N_i]}$  (см. Таблицу 4.4) не является единичной, поэтому – переходим на шаг 6.
16. Трижды выполняются шаги 6, 7, 8 при одних и тех же условиях.

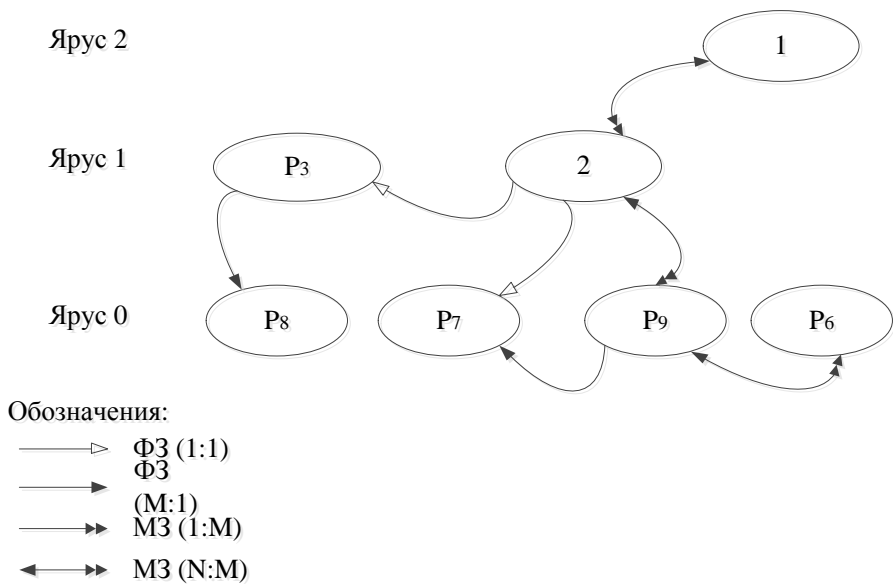
**Обратите внимание** на следующие моменты:

- формируется отношение  $P_3 := \{3\} = < \text{ЗачКнижка№} >$ ;
- поскольку  $|M_i| \leq 1$ , то замены строк и столбцов не происходит. По этой же причине число ключей  $K_i$  не увеличивается;
- поскольку для элементов 3 и 2 на этом шаге  $N_i - j > K_i + 1$ , то значение переменной  $j$  последовательно увеличивается до 2;
- для элемента 1  $(N_i - j = 1) = (K_i + 1 = 1)$ , то  $i$  увеличивается до 2 и вновь возвращаемся на шаг 3.

17. После шага 16 установились следующие значения переменных:  $K_i = 0$ , а  $i = 2$ . Поскольку значение  $i$  больше нуля, все элементы предметной области проанализированы и, кроме элементов, имеющих в  $\Delta MF_{[Ni]}$  одну единицу в строке остался только элемент 1 (УчДисциплина), то формируем окончательный вид канонической структуры  $S$  (см. Рисунок 4.2).

**Обратите внимание** на следующие моменты:

- между отношением  $P_3$  (ЗачКнижка№) и атрибутом ФИО (№ 2 в Таблице 4.5) имеется горизонтальная (на одном и том же ярусе структуры  $S$ ) функциональная связь поэтому отношение  $P_3$  и атрибутом ФИО целесообразно объединить в один сегмент (см. Рисунок 4.3);
- отношения  $P_6$ ,  $P_7$  и  $P_9$  находятся на одном и том же ярусе структуры  $S$ , имеют горизонтальные связи по атрибуту ИдФУД (и атрибуты, которые он представляет; см. Рисунок 2.6) и связаны с атрибутом ФИО (№ 2 на Рисунке 4.2). Поэтому отношения  $P_6$ ,  $P_7$  и  $P_9$  целесообразно объединить в один сегмент (см. Рисунок 4.3).



*Рисунок 4.2. Каноническая структура отношений предметной области*



*Рисунок 4.3. Состав сегментов канонической структуры*

18. Результат синтеза канонической структуры отношений предметной области представлен на Рисунке 4.3.
19. Выводы. Построенная каноническая структура характеризуется минимальным (по сравнению с совокупностью R5, R6, R7, R8, R14, R18, R19; см. практическое занятие № 2) набором отношений (сегментов), атрибутов в отношениях в целом, а также минимальным (на том же множестве отношений) числом связей между ними. При этом структура  $S$  (см. Рисунок 4.2) обладает той же представительностью заданной предметной области, что и результат практического занятия № 2.

Каноническая структура входит в состав исходных данных для оптимизации логической структуры базы данных.

### **Оформление и содержание отчета**

Отчет по практическому занятию должен быть написан аккуратным почерком или распечатан на принтере. При подготовке материалов отчета необходимо выполнять общие требования ЕСКД, которые предъявляются к разработке текстовых документов.

В отчете должны быть указаны следующие данные и представлены следующие материалы:

- фамилия и номер учебной группы курсанта;
- тема, цель и номер задания;
- содержание задания: исходные данные и задача;
- результаты выполнения по каждому (с 1 по 18) пункту задания: таблицы, матрицы, овал-диаграммы;
- выводы, содержащие пояснения к дальнейшему использованию результатов.

**Литература к практическому занятию № 4**

1. Мартин Дж. Организация баз данных в вычислительных системах: Пер. с англ. Издание второе дополненное. –М.: Мир, 1980 г. –662 с.
2. Дейт К. Введение в системы баз данных.–М.: Наука, 1980 г.
3. Материалы диссертации ктн Семенкова В.В. –Л: ВИКИ им. А.Ф.Можайского, 1982 г.

## ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 5

**Тема занятия:** Порядок формирования концептуальной схемы БД.

**Цель занятия:** Приобретение навыков описания информационных требований к БД.

**Задание:** Дать описание концептуальной схемы БД.

### Методические указания по выполнению задания № 5

#### *Накануне занятия:*

- ознакомиться с темой, целью и заданием;
- изучить лекционный материал и рекомендуемую литературу по заданной теме;
- выполнить задание по практическим занятиям №№ 1,2,4;
- подготовить подсобный материал для составления письменного отчета по заданию;
- получить в библиотеке и иметь на занятии рекомендованную литературу.

#### *В ходе занятия:*

- принять участие в разборе примера выполнения задания;
- получить у преподавателя вариант задания;
- самостоятельно сформировать описание концептуальной схемы БД в виде информационных требований к БД;
- подготовить письменный отчет о выполнении задания.

### Порядок выполнения задания № 5

1. Анализ и обобщение представлений пользователей о предметной области. Ранее, на практическом занятии №1 было отмечено, что предметная область может быть задана текстовым описанием, овал-диаграммой или таблицей значений атрибутов. Различные описания одной



и той же заданной предметной области были получены в ходе практических занятий №№ 1, 2, 4. Поскольку, в ходе построения указанных схем не рассматривались требования какого-либо конкретного пользователя или групп пользователей, то все эти схемы являются абстрактными, т.е. концептуальными. Единственной оговоркой данного заявления служит ориентация табличного представления (см. результаты практических занятий №№ 1, 2) предметной области на реляционную модель БД.

2. Описание элементов в информационных требованиях к БД должно содержать имя каждого элемента данных в схеме БД и его числовые характеристики: размер отводимого поля и мощность множества значений. Числовые характеристики элемента заключаются в скобки.

Составим описание элементов данных заданной предметной области:

- a) ФИО (12 – размер отводимого поля,  
6 – мощность множества значений);

далее аналогично:

- b) УчДисциплина (10, 4);
- c) ЗачКнижка№ (4, 6);
- d) Задание№ (2, 14);
- e) ИдФУД (4, 10);
- f) Семестр№ (2, 4);
- g) № п.п. (2, 6);
- h) СрБалл (3, 5);
- i) Занятие№ (4, 10).

**Обратите внимание** на то, что значения числовых характеристик (особенно мощность множества значений) элементов данных представляют собой срез данных в БД на конкретный момент времени (см. Таблицы 1.4, 1.5, 2.1 – 2.3, 2.7, 2.8).

3. Формирование отношений на основе канонической структуры БД. Каноническая структура заданной предметной области (см. Рисунок 4.2) является независимой от программного обеспечения и аппаратных средств. Поэтому для того, чтобы построить на ее основе реляционную структуру, необходимо преобразовать эту каноническую структуру в соответствующую ей логическую схему БД.

Для отображения канонической структуры в логическую необходимо в качестве основного использовать правило, обратное правилу формирования полного сцепленного ключа (ПСК) заданного сегмента. Хотя, правило формирования ПСК применяется для иерархических структур, тем не менее, оно не противоречит и принципам реляционного подхода. Для формирования ПСК необходимо дополнить **исходные** сегменты (сегменты верхнего уровня; см. Рисунок 4.3) ключевыми атрибутами (см. Рисунки 1.5, 2.1, 2.4, 2.6) сегментов нижнего уровня. В результате получим множество отношений  $RCS_1 - RCS_5$ :

- a)  $RCS_1 : <УчДисциплина, ФИО, ИдФУД>$ ;
- b)  $RCS_2 : <ЗачКнижка№, ФИО, №п.п.>$ ;
- c)  $RCS_3 : <СрБалл, ЗачКнижка№>$ ;
- d)  $RCS_4 : <Задание№, ИдФУД, Занятие№, Семестр№>$ .

**Обратите внимание** на следующие моменты:

- отношение  $RCS_1$  дополнено ключевым атрибутом ФИО. Формирование отношения  $RCS_1$  будет продолжено далее;

- по правилу формирования ПСК, ключем отношения  $P_8$  является ПСК отношения  $RCS_2$ . Поэтому, как исключение,  $RCS_3$  сформировано путем дополнения  $P_8$  атрибутом ЗачКнижка№. Обойтись без отношения  $RCS_3$  и внести атрибут СрБалл в отношение  $RCS_2$  нельзя, т.к. в этом случае восстановится транзитивная связь между атрибутами ФИО и СрБалл (см. Рисунок 1.3);
- в отношении  $RCS_4$  имеется составной ключ  $\langle \text{Задание№}, \text{ИдФУД}, \text{Занятие№} \rangle$ . Но переносить его на верхний уровень в отношение  $RCS_2$  нецелесообразно, так как однозначная связь между этими отношениями может быть установлена по функциональной зависимости (1:1) между атрибутами ФИО и №п.п. (№ 2 и  $P_7$  на Рисунке 4.2);
- вследствие последнего, отношение  $RCS_1$  дополнено ключевым атрибутом ИдФУД, минуя отношение  $RCS_2$ . Это обусловлено предложенной оптимизацией декомпозирующих отношений, приведенной на Рисунке 2.6, и не противоречит правилу формирования ПСК;
- а также, вследствие функциональной зависимости (1:1) между атрибутами ФИО и №п.п., и совокупности множественных зависимостей между остальными атрибутами в сегменте  $\langle P_9, P_6 \rangle$  атрибут №п.п. может быть перенесен наверх в  $RCS_2$ . при этом связь между  $RCS_2$  и  $RCS_4$  восстанавливается через  $RCS_1$ .

4. Описание отношений элементов данных в информационных требованиях к БД должно содержать следующие характеристики: имя и тип каждого отношения в схеме БД, имя исходного атрибута (как

правило, верхнего уровня) в отношении и его числовые характеристики, имена зависимых атрибутов и их числовые характеристики; среднее количество значений исходного атрибута, отображающихся в одно значение зависимого атрибута; числовые характеристики отношения: кардинальное число, определяющее тип отношения, размер отводимого поля и мощность множества значений. Исходные и зависимые атрибуты, а также числовые характеристики отношения заключаются в скобки. Кроме того, рекомендуется подчеркнуть исходные атрибуты.

Составим описание отношений, определенных на основе канонической структуры БД:

- а) **RCS<sub>I</sub>** (1:1 – кардинальное число, определяющее тип отношения,

26 – размер отводимого поля,

10 – мощность множества значений):

(УчДисциплина (10 – размер поля,  
отводимого под  
зависимый атрибут,

4 – мощность множества  
значений)

ФИО (12 – размер поля, отводимого под  
зависимый атрибут,

6 – мощность множества  
значений)

(ИдФУД\_(4 – размер поля, отводимого под исходный  
атрибут,

10 – мощность множества значений,

1 – среднее количество значений исходного  
атрибута, отображающихся в одно значение зависимого  
атрибута)))));

далее аналогично:

- b) **RCS<sub>2</sub>** (1:1, 18, 6):  
(ФИО (12, 6) (ЗачКнижка№ (4, 6, 1), №  
п.п. (2, 6, 1)));
- c) **RCS<sub>3</sub>** (M:1, 7, 6):  
(ЗачКнижка№ (4, 6) (СрБалл (3, 5, 6/5)));  
;
- d) **RCS<sub>4</sub>** (M:N, 12, 14):  
(Задание№ (2, 14) ИдФУД (4, 10)  
Занятие№ (4, 10)  
(Семестр№ (2, 4, 7/2))).

***Обратите внимание*** на следующие моменты:

- размер поля, отводимого под отношение, определяется суммой длин полей, отводимых под входящие в него атрибуты;
- в отношении ***RCS<sub>4</sub>***, которое определяются кардинальным числом M:N, возможным ключом является вся совокупность атрибутов, за исключением атрибута Семестр№. Поэтому мощность множества значений данного отношения определяется числом, которое не менее наибольшей мощности одного из атрибутов, входящих в возможный ключ отношения ***RCS<sub>4</sub>***;
- последнее справедливо и для отношений, которые определяются кардинальным числом 1:M.

5. В результате отображения канонической структуры в логическую схему реляционной БД определены информационные требования, и сформирована совокупность отношений ***RCS<sub>1</sub>*** – ***RCS<sub>4</sub>*** (см. Таблицы 5.1–5.4). Отношение ***RCS<sub>1</sub>*** находится в 4НФ, ***RCS<sub>4</sub>*** – в БКНФ, а ***RCS<sub>2</sub>*** и ***RCS<sub>3</sub>*** – в 3НФ (*проверьте !*).

6. Выводы. Описание информационных требований к составу БД, представленное в пп. а) – д) пункта 4 служит дальнейшему осмыслению предметной области. Оно является формализованным представлением концептуальной схемы, которое позволяет применять автоматизированные средства для проектирования логической структуры БД.

Информационные требования входят в состав ограничений для оптимизации логической структуры базы данных.

*Таблица 5.6. Отношение  $RCS_I$*

<u>Уч</u> <u>Дисциплина</u>	<u>ФИО</u>	ИдФУД
МАнализ	Ивлев Б.Б.	1002
МАнализ	Пикалин В.В.	2002
АГеометрия	Пикалин В.В.	2001
МЛогика	Сигалов А.А.	3003
СУБД	Сигалов А.А.	3004
СУБД	Суровов Е.Е.	4004
МЛогика	Суровов Е.Е.	4003
АГеометрия	Сусанов Г.Г.	5001
МЛогика	Ярцев Д.Д.	6003
СУБД	Ярцев Д.Д.	6004

Таблица 5.2. Отношение  $RCS_2$ 

Зач Книжка№	<u>ФИО</u>	№ п.п.
ЗК_1	Ивлев Б.Б.	1
ЗК_2	Пикалин В.В.	2
ЗК_3	Сигалов А.А.	3
ЗК_4	Суровов Е.Е.	4
ЗК_5	Сусанов Г.Г.	5
ЗК_6	Ярцев Д.Д.	6

Таблица 5.3. Отношение  $RCS_3$ 

Ср Балл	<u>Зач</u> <u>Книжка№</u>
4,3	ЗК_1
4,5	ЗК_2
4,4	ЗК_3
4,8	ЗК_4
4,5	ЗК_5
5,0	ЗК_6

Таблица 5.47. Отношение  $RCS_4$ 

<u>Задание №</u>	<u>ИдФУД</u>	<u>Занятие №</u>	<u>Семестр №</u>
M1	1002	M31	C1
M1	1002	M32	C1
M3	1002	M321	C2
M3	1002	M322	C2
M2	2002	M31	C1
A1	2001	A31	C1
A3	2001	A318	C2
Л1	3003	Л31	C3
Б1	3004	С31	C3
Б2	4004	С31	C3
Л2	4003	Л31	C3
A2	5001	A31	C1
A4	5001	A318	C2
Л3	6003	Л31	C3
Л4	6003	Л316	C4
Б3	6004	С317	C4

### Оформление и содержание отчета

Отчет по практическому занятию должен быть написан аккуратным почерком или распечатан на принтере. При подготовке материалов отчета необходимо выполнять общие требования ЕСКД, которые предъявляются к разработке текстовых документов.



В отчете должны быть указаны следующие данные и представлены следующие материалы:

- фамилия и номер учебной группы курсанта;
- тема, цель и номер задания;
- содержание задания: исходные данные и задача;
- результаты выполнения по каждому (с 1 по 5) пункту задания: описания и таблицы;
- выводы, содержащие пояснения к дальнейшему использованию результатов.

### **Литература к практическому занятию № 5**

1. Хаббард Дж. Автоматизированное проектирование баз данных: Пер. с англ. –М.: Мир, 1984 г. –296с.
2. Тиори Т., Фрай Дж. Проектирование структур баз данных: В 2-х кн. Кн. 1. Пер. с англ. –М.: Мир, 1985 г. – 287 с.
3. Материалы диссертации ктн Семенкова В.В. –Л: ВИКИ им. А.Ф.Можайского, 1982 г.

## ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 6

**Тема занятия:** Порядок расчета затрат внешней памяти на реализацию логической схемы БД.

**Цель занятия:** Приобретение навыков оптимизации по затратам внешней памяти логической схемы БД.

**Задание:** Оценить варианты схем БД по затратам внешней памяти.

### Методические указания по выполнению задания № 6

#### *Накануне занятия:*

- ознакомиться с темой, целью и заданием;
- изучить лекционный материал и рекомендуемую литературу по заданной теме;
- выполнить задание по практическому занятию № 5;
- подготовить подсобный материал для составления письменного отчета по заданию;
- получить в библиотеке и иметь на занятии рекомендованную литературу.

#### *В ходе занятия:*

- принять участие в разборе примера выполнения задания;
- получить у преподавателя вариант задания;
- самостоятельно оценить варианты схем заданной БД и определить наилучшую из них;
- подготовить письменный отчет о выполнении задания.

### Порядок выполнения задания № 6

1. Описание ресурсов. Обычно рассматривают две основные группы ресурсов, которые будут израсходованы при

эксплуатации БД, созданной по тому или иному варианту схемы для заданной предметной области. В первую группу входят ресурсы внешней памяти на носителе информации, а во вторую – затраты времени на выполнение информационных запросов

$TR_i [i = \overline{1, Q}]$  и затраты на путь селекции данных при их выполнении.

В общем случае для выбора оптимальной схемы необходимо решить следующую систему неравенств:

$$\begin{cases} \min_{j \in \{S_1, S_2, \dots, S_m\}} \{V^{(j)}\} > 0 \\ \min_{j \in \{S_1, S_2, \dots, S_m\}} \{T^{(j)}\} > 0 \end{cases} \quad (6.1)$$

где  $V^{(j)}$  – ресурсы из первой группы;

$T^{(j)}$  – ресурсы из второй группы.

В результате решения системы (6.1) стремятся найти (если удастся) аргумент  $j \in \{S_1, S_2, \dots, S_m\}$  соответствующий оптимальной по затратам ресурсов схеме БД. Таким образом, в общем виде, требуется решить следующую задачу:

$$S^{(opt)} = \arg_{\{S_1, S_2, \dots, S_m\}} \left\{ \begin{array}{l} \min_{j \in \{S_1, S_2, \dots, S_m\}} \{V^{(j)}\} > 0 \\ \min_{j \in \{S_1, S_2, \dots, S_m\}} \{T^{(j)}\} > 0 \end{array} \right\} \quad (6.2)$$

Затраты ресурсов из первой группы  $V_\Sigma$  зависят от длин полей  $v_l [l = \overline{1, G}]$  отведенных для хранения отношений, а также от состава дополнительных индексов  $I_{\{l\}}$  и другой служебной информации, характеризующей схему. Поэтому

для расчета затрат внешней памяти необходимо использовать информационные требования к БД, сформированные на практическом занятии №5.

Затраты ресурсов из второй группы зависят от точки входа  $P_i$  (отдельного атрибута, или отношения в целом) в схему БД. При этом также необходимо использовать информационные требования к БД.

2. Описание информационных запросов. Поскольку БД, как информационная система, предназначена для реализации информационных запросов пользователей, то сами запросы должны содержаться (явно или косвенно) в текстовом описании данной предметной области (см. практическое занятие №1), или же прилагаться к нему. Пусть множество  $TR_{\{i\}}$  представлено следующими запросами (в скобках указана частота их использования в единицу времени):

- a)  $TR_{\{1\}}$  (4): Определить номер зачетной книжки (ЗачКнижка№) и средний балл (СрБалл) заданного курсанта (ФИО);
- b)  $TR_{\{2\}}$  (4): Какие задания (Задание№) были выданы заданному курсанту (ФИО) по данной дисциплине (УчДисциплина).
- c)  $TR_{\{3\}}$  (2): По заданной дисциплине (УчДисциплина) выдать список занятий (Занятие№) и семестров (Семестр№), в течение которых она изучается.

**Обратите внимание** на следующие моменты:

- подчеркнутые линией атрибуты и заданные значения характеризуют запрашиваемую информацию. Они формируют условия выборки данных, и определяют состав дополнительных индексов;

- все последующие расчеты выполняются для зафиксированных информационных запросов  $TR_{\{i\}}$ . При дополнении множества запросов, или изменении их сути, или же их частоты, выбранный вариант схемы БД может оказаться не самым лучшим. Тем не менее, приведенные ниже расчеты полезно проводить, так как в противном случае имеется риск всегда использовать не оптимальную схему БД, или же отдельного запроса.

3. Описание вариантов схем БД. В процессе отображения канонической структуры БД в логическую может быть получено множество вариантов схем. На практическом занятии №5 на основании схемы, изображенной на Рисунке 4.2, была получена схема  $S_5$ , представленная отношениями  $RCS_1 - RCS_4$ . На основании схемы, изображенной на Рисунке 4.1, может быть получена другая схема БД (*постройте самостоятельно !*) для той же предметной области.

Кроме того, можно учесть и то, что для реляционных БД в качестве вариантов могут рассматриваться схемы, полученные в процессе нормализации отношений заданной предметной области. Так на практическом занятии №2 была получена схема  $S_2$ , представленная отношениями  $R_5, R_6, R_7, R_8, R_{14}, R_{18}, R_{19}$  (см. Таблицы 1.4, 1.5, 2.1–2.3, 2.7, 2.8), а также схема с потерями информации  $R_5, R_6, R_7, R_8, R_{14}, R_{15}, R_{16}, R_{17}$ .

Для вычисления и сравнения затрат ресурсов будем рассматривать схемы  $S_5$  и  $S_2$ . Информационные требования (см. практическое занятие №5) для схемы  $S_2$  и другие варианты схем *оцените самостоятельно !*

4. Расчеты затрат внешней памяти для схемы S<sub>5</sub>. Затраты внешней памяти вычисляются по следующей формуле:

$$V_{\Sigma} = \sum_{I=1}^G v_I \cdot |RCS_I|, \quad (6.3)$$

где  $|RCS_I|$  – мощность множества значений кортежей отношения.

Используя формулу (6.1), оценим схемы S<sub>5</sub> и S<sub>2</sub>.

$$\begin{aligned} V_{S5} &= 26 \cdot |RCS_1| + 18 \cdot |RCS_2| + 7 \cdot |RCS_3| + \\ &+ 12 \cdot |RCS_4| = \\ &= 26 \cdot 10 + 18 \cdot 6 + 7 \cdot 6 + 12 \cdot 14 = 578; \end{aligned} \quad (6.4)$$

$$\begin{aligned} V_{S2} &= 18 \cdot |R5| + 7 \cdot |R6| + 4 \cdot |R7| + 12 \cdot |R8| + \\ &+ 14 \cdot |R14| + 26 \cdot |R18| + 6 \cdot |R19| = \\ &= 18 \cdot 6 + 7 \cdot 6 + 4 \cdot 14 + 12 \cdot 14 + 14 \cdot 10 + 26 \cdot 10 + 6 \cdot 14 = \\ &= 858 \end{aligned}$$

(см. информационные требования к  $RCS_1 - RCS_5$  и Таблицы 1.4, 1.5, 2.1–2.3, 2.7, 2.8);

Рассчитаем затраты памяти для хранения дополнительных индексов  $I_{\{I\}}$ . Для этого необходимо проанализировать информационные запросы  $TR_{\{1\}}$ ,  $TR_{\{2\}}$  и  $TR_{\{3\}}$ .

В схеме S<sub>5</sub> информация, необходимая для выполнения  $TR_{\{1\}}$ , находится в отношениях  $RCS_2$  и  $RCS_3$ . Соединение отношений  $RCS_2$  и  $RCS_3$  необходимо выполнить по атрибуту ЗачКнижка№. Атрибуты ФИО и ЗачКнижка№, определяющие условие поиска и связь между  $RCS_2$  и  $RCS_3$  являются первичными ключами указанных отношений. Следовательно, для выполнения  $TR_{\{1\}}$  на  $RCS_2$  и  $RCS_3$  дополнительной информации не требуется. Однако, для обеспечения целостности представления

данных в  $RCS_2$  необходимо сформировать внешний ключ на основе атрибута ЗачКнижка№. Это влечет за собой создание индекса в  $RCS_2$  по значениям атрибута ЗачКнижка№ и дополнительные затраты памяти. А также, так как между составным атрибутом ФИО и №п.п. установлена ФЗ типа (1:1), как показано на Рисунке 1.1, то на основе атрибута №п.п. требуется построить дополнительный уникальный индекс в  $RCS_2$ .

Таким образом, затраты памяти на все индексы, построенные для заданных отношений, можно вычислить по следующей формуле:

$$I_l = |RCS_l| \cdot (4 + \sum_{p=1}^{K_p} a_p) + \sum_{e=1}^{K_e} (a_e + 4) \cdot |a_e|, \quad (6.5)$$

где  $a_p$  – длина поля, отводимая, под ключевой атрибут;

$K_p$  – число атрибутов, входящих в состав первичного ключа;

4 – длина прямого адреса значения кортежа соответствующего отношения;

$a_e$  – длина поля, отводимая, под атрибут, на основе которого построен дополнительный индекс;

$|a_e|$  – мощность множества значений атрибута  $a_e$ ;

$K_e$  – число дополнительных индексов данного отношения.

Тогда

$$\begin{aligned} I_2 &= 6 \cdot (4+12) + (4+4) \cdot 6 + (4+2) \cdot 6 = 180, \\ I_3 &= 6 \cdot (4+4) = 48. \end{aligned} \quad (6.6)$$

При выполнении  $TR_{[2]}$  информацию в схеме  $S_5$  необходимо извлекать из отношений  $RCS_1$  и  $RCS_4$ . Атрибуты ФИО и УчДисциплина, определяющие условие

поиска данных, входят в состав первичного ключа отношения  $RCS_I$ . Однако, при поиске информации в  $RCS_I$  и  $RCS_4$  индексы, построенные по значениям составных первичных ключей не будут использоваться СУБД. Первичный ключ отношения предназначен, прежде всего, для поддержания целостности данных. А, для ускорения поиска на основе атрибута ФИО требуется построить дополнительный индекс в  $RCS_I$ . Кроме того, так как между составным атрибутом <ФИО, УчДисциплина> и ИдФУД устанавливается ФЗ типа (1:1), как показано на Рисунке 2.6, то на основе атрибута ИдФУД требуется построить дополнительный уникальный индекс в  $RCS_I$ .

Соединение отношений  $RCS_I$  и  $RCS_4$  необходимо выполнить по атрибуту ИдФУД. Этот же атрибут должен использоваться для объявления внешнего ключа в  $RCS_4$ .

Тогда, при построении дополнительных индексов для отношений  $RCS_I$  и  $RCS_4$  будут следующими:

$$I_1 = 6 \cdot (4+12) + 4 \cdot (4+10) + (4+12) \cdot 6 + (4+4) \cdot 10 = 328;$$

$$I_4 = 14 \cdot (4+2) + 10 \cdot (4+4) + 10 \cdot (4+4) + (4+4) \cdot 10 = 324. \quad (6.7)$$

При выполнении  $TR_{\{3\}}$  информацию в схеме  $S_5$  также необходимо извлекать из отношений  $RCS_I$  и  $RCS_4$ . Соединение этих отношений по атрибуту УчДисциплина ранее уже было рассмотрено, поэтому, дополнительных затрат памяти для его формирования не потребуется.

5. Расчеты затрат внешней памяти для схемы  $S_2$ . Информация, необходимая для выполнения  $TR_{\{1\}}$ , находится в отношениях R5 и R6. Атрибуты ФИО и ЗачКнижка№, определяющие условие поиска и связь между отношениями R5 и R6 являются первичными ключами указанных отношений. Так же, как и в схеме



$S_5$ , для выполнения  $TR_{[1]}$  на R5 и R6, дополнительной информации не требуется. Однако, для обеспечения целостности представления данных в R5 (для обеспечения ФЗ 1:1 между атрибутами ФИО и ЗачКнижка№) необходимо сформировать дополнительный уникальный индекс, который будет участвовать в соединении R5 и R6, на основе атрибута ЗачКнижка№.

Тогда, по формуле (6.5):

$$\begin{aligned} I_{R5} &= 6 \cdot (4+12) + (4+4) \cdot 6 = 144, \\ I_{R6} &= 6 \cdot (4+4) = 48. \end{aligned} \quad (6.8)$$

При выполнении  $TR_{[2]}$  необходимую информацию в схеме  $S_2$  необходимо извлекать из отношений R5, R7 и R8. Отношение R5 уже, рассматривалось при выполнении  $TR_{[1]}$ , однако, в данном случае для обеспечения целостности представления данных в R5 (для обеспечения ФЗ 1:1 между атрибутами ФИО и № п.п.) и формирования связи между отношениями R5 и R7 потребуется создать дополнительный уникальный индекс, который будет участвовать в их соединении, на основе атрибута № п.п. Для формирования связи между отношениями R7 и R8 дополнительных затрат не требуется, так как она формируется на основании атрибута Задание№, который является первичным ключем в каждом из указанных отношений.

Тогда, по формуле (6.5) с учетом (6.8):

$$\begin{aligned} I'_{R5} &= 144 + (2+4) \cdot 6 = 180, \\ I_{R7} &= 14 \cdot (4+2) = 84, \\ I_{R8} &= 14 \cdot (4+2) = 84. \end{aligned} \quad (6.9)$$

Выполняя  $TR_{[3]}$ , информацию в схеме  $S_2$  необходимо извлекать из отношений R14, R18 и R19. При

этом отношения R14 и R18 требуется соединить по атрибуту УчДисциплина и объединить (операция UNION, т.к. в предметной области – отношение R1, Рисунок 1.1, не задана связь между атрибутами СеместрN и ЗанятиеN) полученный результат с выборкой данных из соединения отношений R18 и R19 по атрибуту ИдФУД. Атрибут УчДисциплина входит в состав составных первичных ключей R14 и R18, а атрибут ИдФУД – в составной первичный ключ отношения R19, поэтому при поиске информации эти индексы (на основе составных первичных ключей) не будут использоваться СУБД. Для ускорения поиска на основе атрибута УчДисциплина требуется построить дополнительный индекс в R14. Кроме того, так как в R18 между составным атрибутом <ФИО, УчДисциплина> и ИдФУД устанавливается ФЗ типа (1:1), как показано на Рисунке 2.6, то на основе атрибута ИдФУД требуется построить дополнительный уникальный индекс в R18. Далее, отношения R18 и R19 требуется соединить по атрибуту ИдФУД, который входит в состав первичного ключа отношения R19. Поэтому дополнительных затрат памяти в R18 и R19 не потребуется.

Тогда, по формуле (6.5):

$$\begin{aligned}
 I_{R14} &= 10 \cdot (4+10+4) + (10+4) \cdot 4 = 236, \\
 I_{R18} &= 10 \cdot (4+12+10) + (4+4) \cdot 10 = 340, \\
 I_{R19} &= 14 \cdot (4+2+4) = 140.
 \end{aligned}
 \tag{6.10}$$

6. Результаты расчета затрат внешней памяти для рассмотренных вариантов схем БД состоят в следующем:
  - для схемы S5 с учетом (6.4, 6.6, 6.7):

$$V^{(S5)} = V_{S5} + I_1 + I_2 + I_3 + I_4 + I_5 = 1458 \text{ (байт)};$$

- для схемы  $S_2$  с учетом (6.4, 6.8, 6.9, 6.10):

$$V^{(S2)} = V_{S2} + I'_{R5} + I_{R6} + I_{R7} + I_{R8} + I_{R14} + I_{R18} + I_{R19} = 1970 \text{ (байт)}.$$

## 7. Выводы.

- по затратам внешней памяти оптимальной является схема  $S_5$ ;
- оценку затрат времени на выполнение информационных запросов необходимо провести в соответствии с практическим занятием № 7.

## **Оформление и содержание отчета**

Отчет по практическому занятию должен быть написан аккуратным почерком или распечатан на принтере. При подготовке материалов отчета необходимо выполнять общие требования ЕСКД, которые предъявляются к разработке текстовых документов.

В отчете должны быть указаны следующие данные и представлены следующие материалы:

- фамилия и номер учебной группы курсанта;
- тема, цель и номер задания;
- содержание задания: исходные данные и задача;
- результаты выполнения по каждому (с 1 по 6) пункту задания: расчеты, итоговые значения;
- выводы, содержащие пояснения к дальнейшему использованию результатов.

**Литература к практическому занятию № 6**

1. Хаббард Дж. Автоматизированное проектирование баз данных: Пер. с англ. –М.: Мир, 1984 г. –296с.
2. Тиори Т., Фрай Дж. Проектирование структур баз данных: В 2-х кн. Кн. 1. Пер. с англ. –М.: Мир, 1985 г. – 287 с.

## ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 7

**Тема занятия:** Порядок расчета затрат времени на выполнение информационных запросов к БД.

**Цель занятия:** Приобретение навыков оптимизации по затратам времени на выполнение информационных запросов к БД.

**Задание:** Оценить варианты схем БД по затратам времени на выполнение информационных запросов с использованием среды MS SQL Server Management Studio.

### Методические указания по выполнению задания № 7

#### *Накануне занятия:*

- ознакомиться с темой, целью и заданием;
- изучить лекционный материал и рекомендуемую литературу по заданной теме;
- выполнить задание по практическому занятию № 6;
- подготовить подсобный материал для составления письменного отчета по заданию;
- получить в библиотеке и иметь на занятии рекомендованную литературу.

#### *В ходе занятия:*

- принять участие в разборе примера выполнения задания;
- получить у преподавателя вариант задания;
- самостоятельно оценить варианты схем заданной БД и определить наилучшую из них;
- подготовить письменный отчет о выполнении задания.

### Порядок выполнения задания № 7

1. Описание задачи. В условиях задачи 6.1 – 6.3 (см. практическое занятие №5) необходимо с помощью среды MS SQL Server Management Studio оценить затраты времени на выполнение информационных запросов к БД, реализованной по схемам **S<sub>5</sub>** и **S<sub>2</sub>**.
2. Тексты информационных запросов для схемы **S<sub>5</sub>**. Пусть множество **TR<sub>{i}</sub>** (см. практическое занятие №6) представлено на языке SQL двумя способами («прямое соединение» и «с вложенным запросом»):

d) **TR<sub>{1}</sub>** :

– прямое соединение (ПС)

```
SELECT RCS2.ЗачКнижкаN, RCS3.СрБалл
FROM RCS3, RCS2
WHERE RCS3.ЗачКнижкаN = RCS2.ЗачКнижкаN
and RCS2.ФИО='Сусанов Г.Г.';
```

– с вложенным запросом (ВЗ)

```
SELECT RCS3.ЗачКнижкаN, RCS3.СрБалл
FROM RCS3
WHERE RCS3.ЗачКнижкаN in
( SELECT RCS2.ЗачКнижкаN
FROM RCS2
WHERE RCS2.ФИО='Сусанов Г.Г.');
```

e) **TR<sub>{2}</sub>** :

– прямое соединение

```
SELECT RCS4.ЗаданиеN
FROM RCS1, RCS4
WHERE RCS1.ИдФУД = RCS4.ИдФУД
and RCS1.ФИО='Суровов Е.Е.'
and RCS1.УчДисциплина='МЛогика';
```

– с вложенным запросом

```
SELECT RCS4.ЗаданиеN
FROM RCS4
WHERE RCS4.ИдФУД in
( SELECT RCS1.ИдФУД
  FROM RCS1
  WHERE RCS1.ФИО='Суровов Е.Е.'
    and RCS1.УчДисциплина='МЛогика');
```

f) **TR<sub>{3}</sub>** :

– прямое соединение

```
SELECT RCS4.ЗанятиеN, RCS4.СеместрN
FROM RCS1, RCS4
WHERE RCS1.ИдФУД = RCS4.ИдФУД
and RCS1.УчДисциплина='СУБД';
```

– с вложенным запросом

```
SELECT RCS4.ЗанятиеN, RCS4.СеместрN
FROM RCS4
WHERE RCS4.ИдФУД in
( SELECT RCS1.ИдФУД
  FROM RCS1
```

WHERE RCS1.УчДисциплина='СУБД');

**Обратите внимание** на следующие моменты:

- значения атрибутов, заданные константами ('Сусанов Г.Г.', 'МЛогика' и др.) указывают на точку входа (отношение, атрибут) в схему БД. Они формируют последовательность обработки таблиц (отношений) в БД;
- все последующие расчеты выполняются для зафиксированных информационных запросов  $TR_{\{i\}}$ . При дополнении множества запросов, или изменении их сути, или же их частоты (см. практическое занятие №6), выбранный вариант схемы БД может оказаться не самым лучшим. Тем не менее, приведенные ниже расчеты полезно проводить, так как в противном случае имеется риск всегда использовать не оптимальную схему БД, или же отдельного запроса.

### 3. Тексты информационных запросов для схемы S<sub>2</sub> :

#### а) $TR_{\{I\}}$ :

– прямое соединение

```
SELECT R5.ЗачКнижкаN, R6.СрБалл
FROM R5, R6
WHERE R5.ЗачКнижкаN = R6.ЗачКнижкаN
and R5.ФИО='Сусанов Г.Г.';
```

– с вложенным запросом

```
SELECT R6.ЗачКнижкаN, R6.СрБалл
FROM R6
WHERE R6.ЗачКнижкаN in
```



```
( SELECT R5.ЗачКнижкаN
  FROM R5
 WHERE R5.ФИО='Сусанов Г.Г.');
```

b) **TR<sub>{2}</sub>** :

– прямое соединение

```
SELECT R8.ЗаданиеN
FROM R5, R7, R8
WHERE R5.ФИО='Суровов Е.Е.'
and R5.Нпп = R7.Нпп
and R7.ЗаданиеN = R8.ЗаданиеN
and R8.УчДисциплина='МЛогика';
```

– с вложенным запросом

```
SELECT R8.ЗаданиеN
FROM R8
WHERE R8.УчДисциплина='МЛогика'
and R8.ЗаданиеN in
  (SELECT R7.ЗаданиеN
   FROM R7
   WHERE R7.Нпп in
     (SELECT R5.Нпп
      FROM R5
      WHERE R5.ФИО='Суровов Е.Е.'));
```

c) **TR<sub>{3}</sub>** :

– прямое соединение

```

SELECT R19.СеместрN AS Перечень, 'Семестр' AS
Расшифровка
FROM R18, R19
WHERE R18.УчДисциплина='СУБД'
AND R18.ИдФУД=R19.ИдФУД

```

union

```

SELECT R14.ЗанятиеN , 'Занятие'
FROM R18, R14
WHERE R18.УчДисциплина='СУБД'
AND R18.УчДисциплина=R14.УчДисциплина;

```

– с вложенным запросом

```

SELECT R19.СеместрN AS Перечень, 'Семестр' AS
Расшифровка
FROM R19
WHERE R19.ИдФУД in
( SELECT R18.ИдФУД
FROM R18
WHERE R18.УчДисциплина='СУБД')

```

union

```

SELECT R14.ЗанятиеN , 'Занятие'
FROM R14
WHERE R14.УчДисциплина in
( SELECT R18.УчДисциплина
FROM R18
WHERE R18.УчДисциплина='СУБД');

```

**Обратите внимание** на следующие моменты:

- в запросе  $TR_{fz}$  список атрибутов «SELECT» (выше) для схемы  $S_2$  и, соответственно, результирующая Таблица 7.3 дополнена фиктивным столбцом «Расшифровка» и значениями «Семестр, Занятие». При этом значения, полученные из БД сведены в столбец «Перечень». Это выполнено лишь для форматирования, для удобства наблюдения результата;
- при выполнении одного и того же запроса  $TR_{fz}$  на любой схеме:  $S_5$  или  $S_2$ , в любом виде: ПС или ВЗ, результат выполнения должен быть одним и тем же с точностью до множества значений по  $TR_{fz}$  (см. Таблицы 7.1 – 7.4).

4. Подготовка среды MS SQL Server Management Studio для определения затрат времени на выполнение информационных запросов к БД. Для этого необходимо установить и запустить на компьютере MS SQL Server Management Studio. После запуска и установления соединения с БД, созданной и заполненной данными в соответствии со схемой  $S_5$ , а в дальнейшем и со схемой  $S_2$ , необходимо выполнить команды из меню «Query» (запрос) в следующей последовательности:

- «New Query» (открыть вкладку «SQL» для написания, редактирования и отладки информационного запроса, см. Рисунок 7.1);
- Include Actual Execution Plan (при выполнении запроса использовать его фактический план).

Таблица 7.1.  
Результат выполнения  
 $TR_{\{1\}}$  :

<i>Зач КнижкаN</i>	<i>СрБалл</i>
ЗК_5	4,5

Таблица 7.2.  
Результат выполнения  
 $TR_{\{2\}}$  :

<i>ЗаданиеN</i>
Л2

Таблица 7.3.  
Результат выполнения  
 $TR_{\{3\}}$  на схеме  $S_2$ :

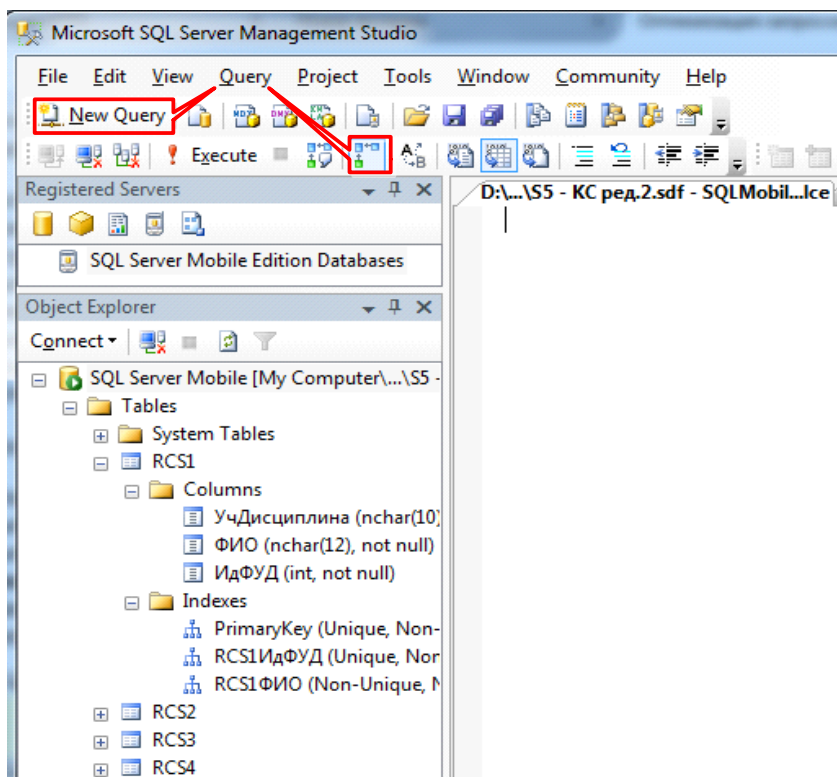
<i>Перечень</i>	<i>Расшифровка</i>
С3	Семестр
С4	Семестр
С31	Занятие
С317	Занятие

Таблица 7.4.  
Результат выполнения  
 $TR_{\{3\}}$  на схеме  $S_5$ :

<i>ЗанятиеN</i>	<i>СеместрN</i>
С31	С3
С31	С3
С317	С4

**Обратите внимание** на то, что

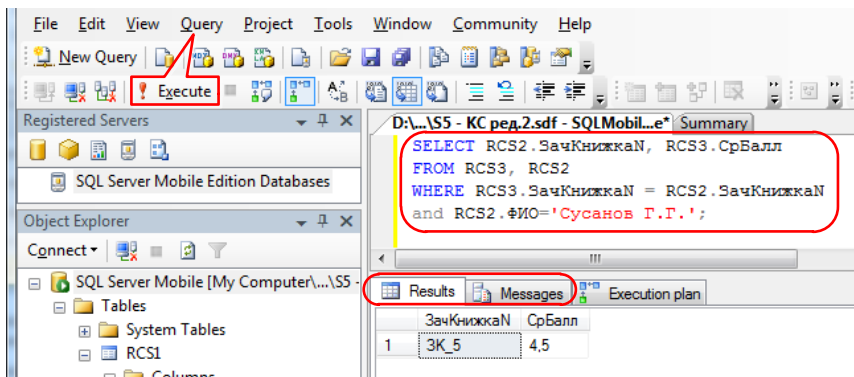
- курсор в это время должен находиться на вкладке «SQL» (см. Рисунок 7.1).



*Рисунок 7.1. Вид рабочего окна MS SQL Server Management Studio*

5. Для определения затрат времени на выполнение информационных запросов к БД необходимо последовательно выполнить следующие действия:

- а) на вкладке «SQL» написать требуемый информационный запрос (перечень запросов приведен в пункте 2, выше);
- б) выполнить команду «Execute» (см. Рисунок 7.2);
- с) убедиться в успешности выполнения информационного запроса:
  - на вкладке «Results» должен быть отображен результат выборки (результатирующая таблица, в заголовке которой отображен список атрибутов из строки «SELECT»),
  - а на вкладке «Messages» не должно быть никаких сообщений кроме записей «row(s) affected».



*Рисунок 7.2. Пояснения к порядку выполнения информационного запроса*

- д) перейти на вкладку «Execution Plan» и правой кнопкой «мышки» нажать элемент «SELECT»\_(см. Рисунок 7.3);
- е) в появившемся окне выполнить команду «Save Execution Plan As...» (сохранить план запроса в файле). В результате выполнения команды план запроса будет сохранен в формате xml;

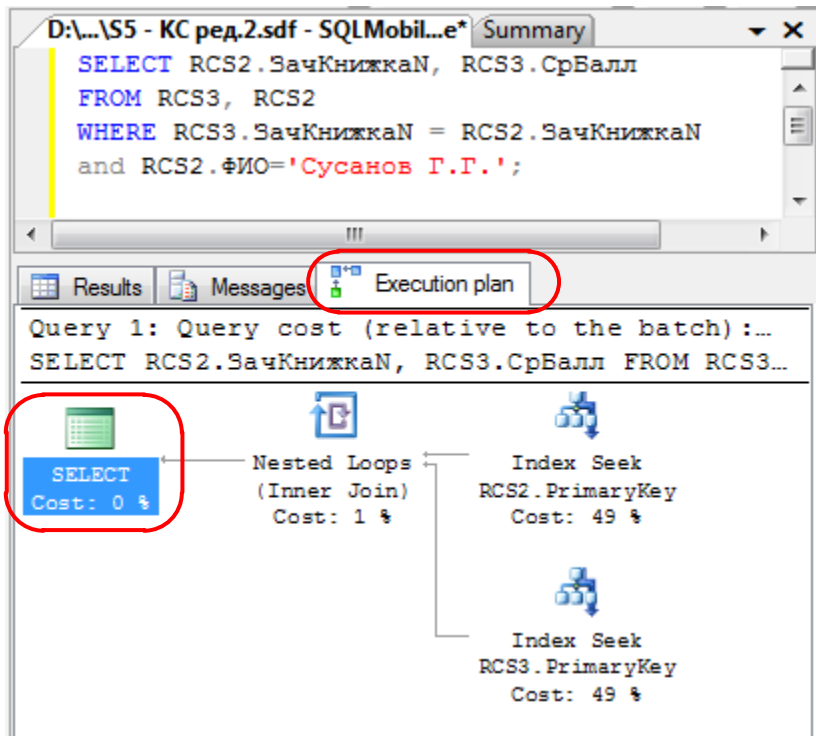


Рисунок 7.3. Вид плана выполнения информационного запроса

- f) для структурированного отображения xml-файла с планом запроса рекомендуется открыть его в интернет-браузере (см. Рисунок 7.4). Оценка искомых затрат временных ресурсов на выполнение информационного запроса к БД представлена значением параметра `EstimatedTotalSubtreeCost` (`StatementSubTreeCost`);

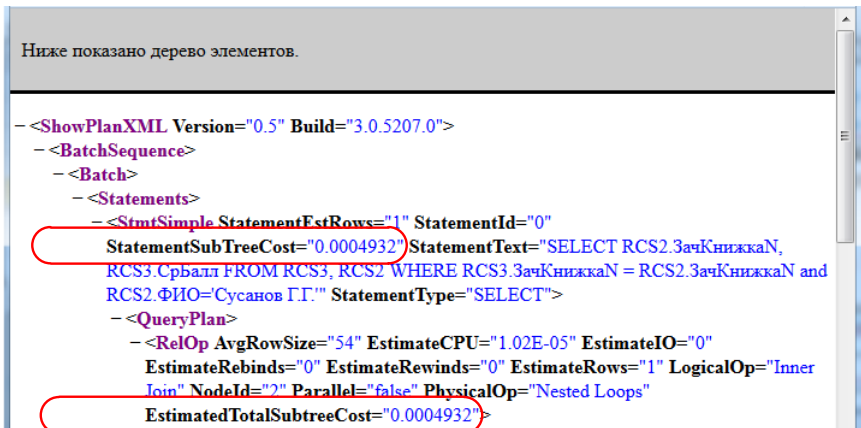


Рисунок 7.4. Вид xml-файла с планом запроса в интернет-браузере

- g) выполнить пункты 5а – 5f для каждого запроса, приведенного в пункте 2 (выше). Значения оценок затрат временных ресурсов записать в таблицу 7.1.
6. Выполнить пункты 4 и 5 для схемы **S<sub>2</sub>** по перечню запросов, приведенных в пункте 3 (выше). В результате получим таблицы 7.1 и 7.2.



Таблица 7.5. Оценки затрат временных ресурсов на реализацию схемы  $S_5$

Запрос	Способ выполнения запроса	EstimateCPU	EstimateIO	EstimatedTotalSubtreeCost
<b><math>TR_{\{1\}}</math></b>	ПС	1,02E-05	0,0004830	0,00049320
	non key ПС	0,0010404	0,0010260	0,00206640
	ВЗ	1,02E-05	0,0044080	0,00441820
	non key ВЗ	0,0005508	0,0058970	0,00644780
<b><math>TR_{\{2\}}</math></b>	ПС	0,0002101	0,0009720	0,00118212
	non key ПС	0,0030804	0,0011100	0,00419040
	ВЗ	6,12E-06	0,0052190	0,00522512
	non key ВЗ	0,0012852	0,0078250	0,00911020
<b><math>TR_{\{3\}}</math></b>	ПС	6,12E-06	0,0004830	0,00048912
	non key ПС	0,0020604	0,0011100	0,00317040
	ВЗ	6,12E-06	0,0044080	0,00441412
	non key ВЗ	0,0012852	0,0067050	0,00799020

Примечания: ПС – «прямое соединение»;

ВЗ – «с вложенным запросом»;

non key – запрос, выполненный на той же совокупности отношений, но реализованной без первичных ключей, индексов и без внешних ключей;

EstimateCPU – затраты процессорного времени;

EstimateIO – затраты на выполнение операций ввода/вывода;

EstimatedTotalSubtreeCost – общие затраты временных ресурсов на выполнение информационного запроса к БД.

Таблица 7.6. Оценки затрат временных ресурсов на реализацию схемы  $S_2$

Запрос	Способ выполнения запроса	EstimateCPU	EstimateIO	EstimatedTotalSubtreeCost
<b><u>TR<sub>{1}</sub></u></b>	ПС	1,02E-05	0,0004830	0,00049320
	non key ПС	0,0010404	0,0010260	0,00206640
	B3	1,02E-05	0,0044080	0,00441820
	non key B3	0,0005508	0,0058970	0,00644780
<b><u>TR<sub>{2}</sub></u></b>	ПС	0,0000619	0,0025332	0,00259513
	non key ПС	0,0023460	0,0031650	0,00551100
	B3	1,02E-04	0,0095900	0,00969204
	non key B3	0,0015504	0,0112494	0,01279980
<b><u>TR<sub>{3}</sub></u></b>	ПС	1,42E-03	0,0023850	0,00921860
	non key ПС	1,42E-03	0,0023850	0,01065320
	B3	1,42E-03	0,0023850	0,01447500
	non key B3	1,42E-03	0,0023850	0,01917280

Примечания те же, что и к таблице 7.1.

**Обратите внимание** на следующие моменты:

- параметры EstimateCPU, EstimateIO, EstimatedTotalSubtreeCost и их значения находятся в xml-файле плана выполнения запроса (см. Рисунок 7.4);
- в общем случае (за исключением **TR<sub>{3}</sub>** для схемы  $S_2$ ) можно принять, что 
$$\text{EstimateIO} = \text{EstimatedTotalSubtreeCost} - \text{EstimateCPU} .$$
- использование индексов (первичных ключей, дополнительных и внешних ключей) снижает

затраты временных ресурсов на выполнение информационных запросов к БД (см. значения EstimatedTotalSubtreeCost в таблицах 7.1 и 7.2). При этом процессор в общем случае быстрее обрабатывает вложенные запросы (см. значения EstimateCPU для ВЗ  $\underline{TR}_{\{2\}}$  в таблицах 7.1 и 7.2.), а подсистема ввода/вывода – медленнее;

- в целом на данной совокупности транзакций  $\underline{TR}_{\{1\}}$  –  $\underline{TR}_{\{3\}}$  при выполнении запросов в виде ПС (см. графу «EstimatedTotalSubtreeCost» таблиц 7.1 и 7.2) на схемах  $S_5$ ,  $S_2$  затраты временных ресурсов меньше.

7. Расчет общих затрат времени на выполнение информационных запросов к БД. С учетом частоты использования запросов в единицу времени (см. п. 2 практического занятия № 6) общие затраты времени на выполнение информационных запросов на заданной схеме БД рассчитываются по формуле:

$$T_{\Sigma} = \sum_{l=1}^G \psi_l \cdot |TR_{\{l\}}|, \quad (7.1)$$

где  $\psi_l$  – частота использования  $l$ -го запроса в единицу времени;

$|TR_{\{l\}}|$  – значение параметра EstimatedTotalSubtreeCost плана выполнения запроса.

8. Результаты расчета, по формуле (7.1) с учетом п. 2 практического занятия № 6 для запросов в виде ПС:

– для схемы  $S_5$ :

$$T_{\Sigma} \text{ }_{S_5} = 4 \cdot 0,00049320 + 4 \cdot 0,00118212 + 2 \cdot 0,00048912 = 0,00767952 \quad (7.2)$$

– для схемы  $S_2$ :

$$T_{\Sigma \text{ } s_2} = 4 \cdot 0,00049320 + 4 \cdot 0,00259513 + 2 \cdot 0,00911860 = 0,03059052 \text{ ,} \quad (7.3)$$

## 9. Выводы.

- по затратам времени на выполнение информационных запросов оптимальной является схема  $S_5$  ;
- в целом, с учетом результатов практического занятия № 6, удалось найти единственное решение задачи (6.2):  $S^{(opt)} = S_5$  .

## **Оформление и содержание отчета**

Отчет по практическому занятию должен быть написан аккуратным почерком или распечатан на принтере. При подготовке материалов отчета необходимо выполнять общие требования ЕСКД, которые предъявляются к разработке текстовых документов.

В отчете должны быть указаны следующие данные и представлены следующие материалы:

- фамилия и номер учебной группы курсанта;
- тема, цель и номер задания;
- содержание задания: исходные данные и задача;
- результаты выполнения по каждому (с 1 по 8) пункту задания: расчеты, итоговые значения;
- выводы, содержащие пояснения к дальнейшему использованию результатов.

## **Литература к практическому занятию № 7**

1. Дейт К.Дж. SQL и реляционная теория. Как грамотно писать код на SQL. [Е/Е]: «Символ-Плюс», 2014.
2. Бондарь А. Microsoft SQL Server 2012. [Е/Е]: БХВ-Петербург, 2013.

3. Грубер М. Понимание SQL: Пер. с англ. Лебедева В.Н., под ред. Булычева В.Н. –М., 1993 г. –420 с.
4. Михеев Р. [Microsoft SQL Server 2005 для администраторов](#) [Е/Е]: БХВ-Петербург, 2007 г. –544 с.
5. [SQL Server 2000](#): [Е/Е]: www.intuit.ru, 2003 г.