



Базы данных

**Преподаватель
Токаревская Светлана Анатольевна**

Литература:

- Голицына О.Л., Максимов Н.В., Попов И.И. Базы данных. 2007.
- Ямова Т.Н. Практикум по базам данных. 2009.

План работы:

1. Модели данных.
2. Проектирование баз данных.
3. Разработка запросов к базе данных.
4. СУБД MS Access.
5. СУБД MS Visual Fox Pro.

Основные понятия

- **База данных (БД)** – совместно используемый набор логически связанных данных (и их описание), предназначенный для удовлетворения информационных потребностей организации.
- **Система управления базами данных (СУБД)** – программное обеспечение, с помощью которого пользователи могут определять, создавать и поддерживать базу данных, а также получать к ней контролируемый доступ.

История развития компьютеризации

- ▣ **Технический период** (≈ с 1946 по 1964 г.)
- ▣ **Программный период** (с 1954 по 1970 г.)
- ▣ **Информационный период** (с 1970 по наст. время)
- ▣ **Гуманитарный период** (с начала 90-х гг. прошлого века)

Модели данных

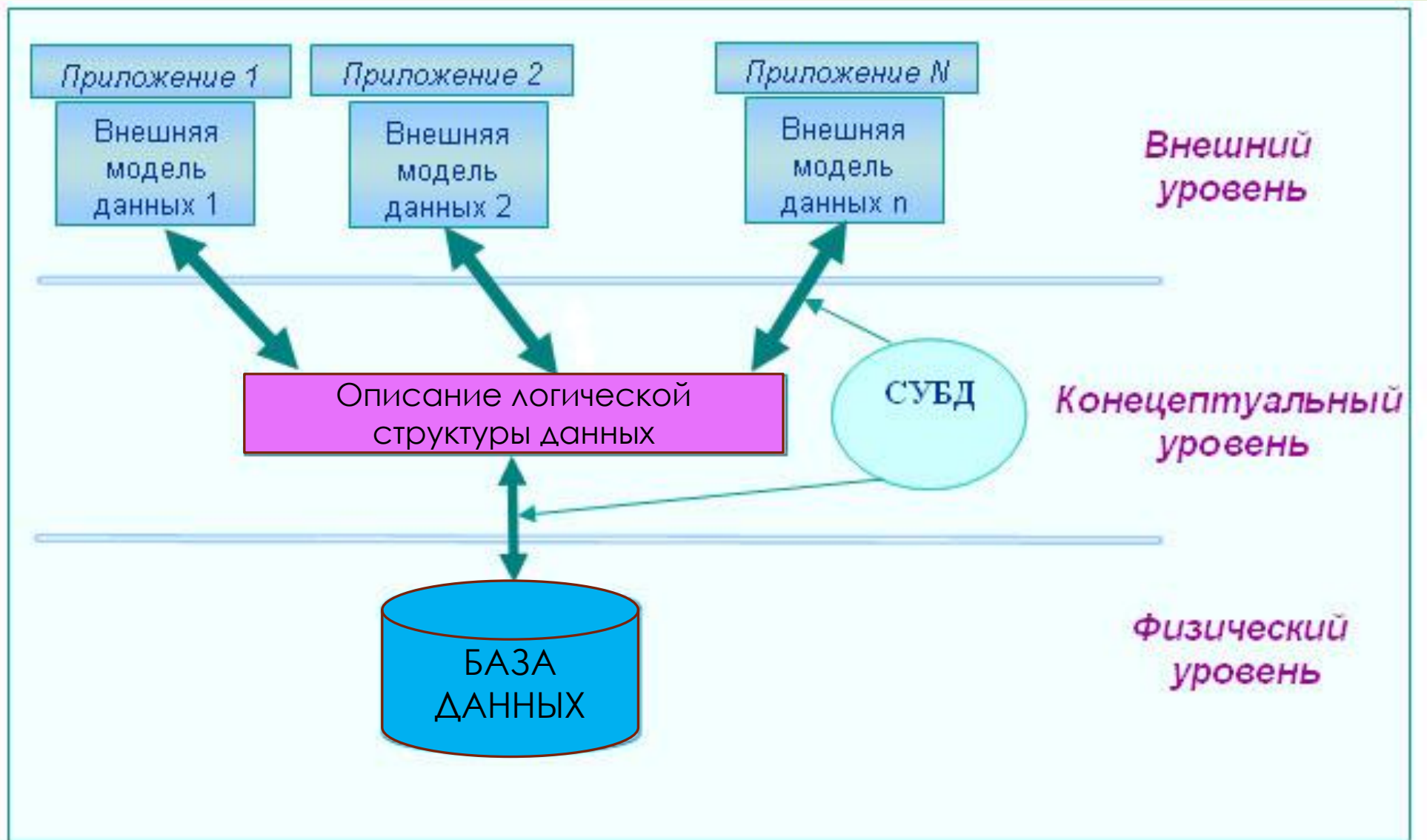
Уровни представления данных:

- Концептуальный;
- Внутренний (физический);
- Внешний.

6

Модель данных (МД) – это некоторая абстракция, которая позволяет пользователям баз данных(БД) и разработчикам трактовать конкретные данные как информацию, то есть как сведения, содержащие не только данные, но и описание связей между ними.

Модели данных (ANSI)

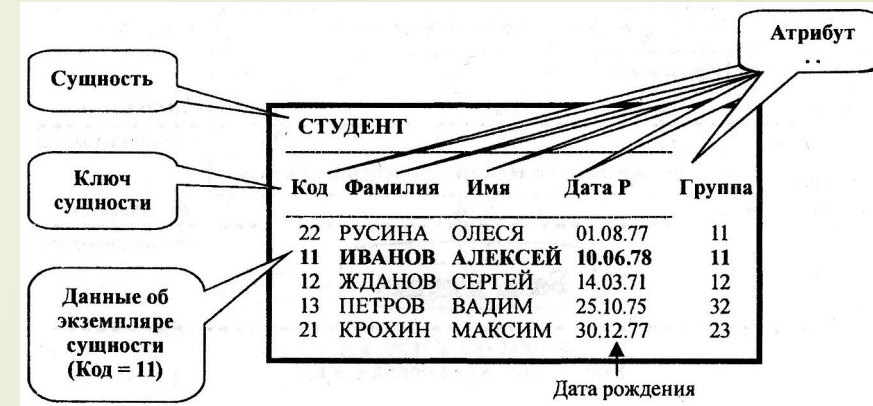


Моделирование:

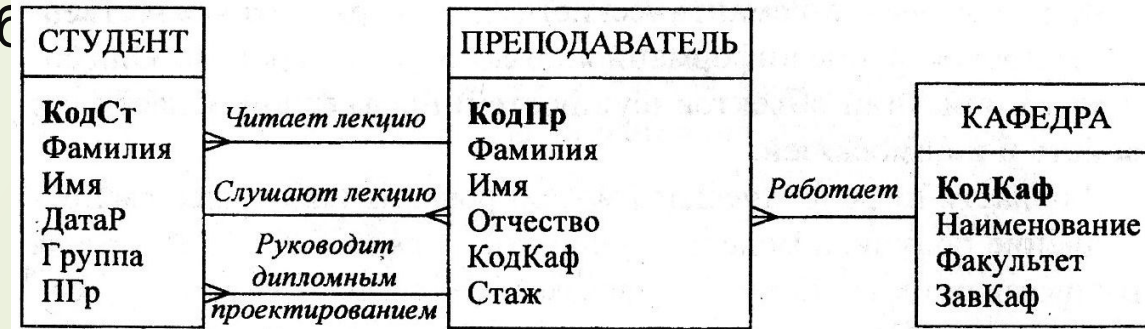
- Инфологическое (семантическое)

Модель «сущность-связь» (метод ER-диаграмм):

- выделяют объекты (сущности),
- формируют список основных свойств (атрибутов), описывающую каждую объект рассматриваемой сущности.



8



- Документальное

- дискрипторные модели;
- тезаурусные модели;
- форматные модели.

- Фактографическое

- (относится к концептуальному уровню, поддерживаются конкретными СУБД)
- иерархические;
- сетевые;
- реляционные.

Иерархическая модель данных

✓ Эффективна для работы с иерархически упорядоченной информацией.

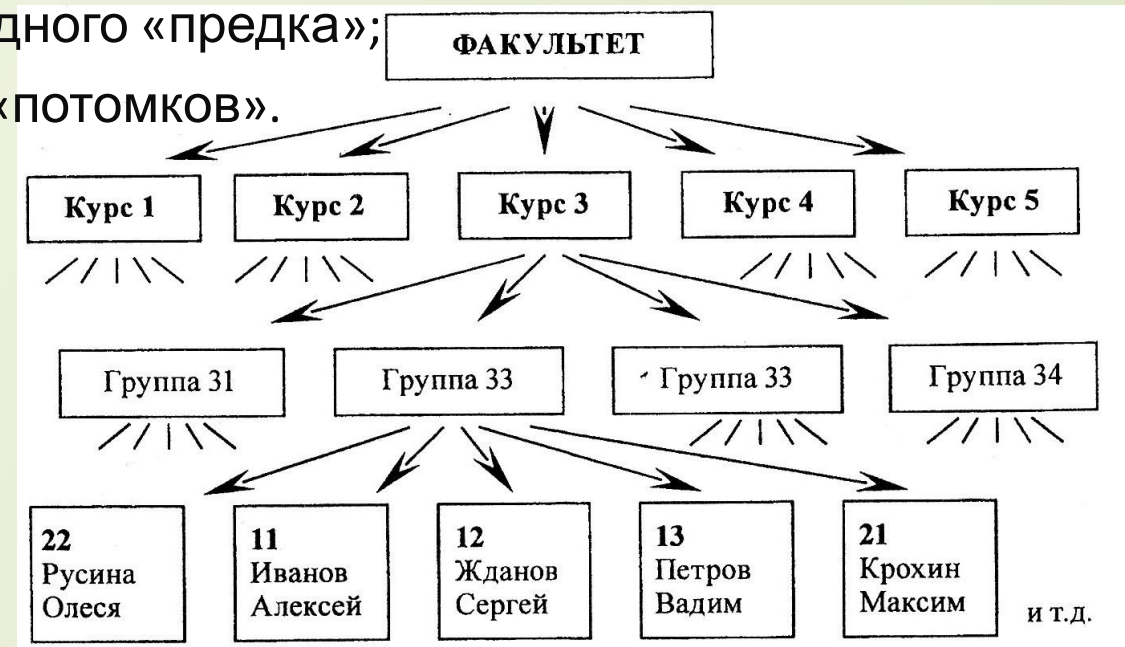
✓ Имеет структуру «дерево», может быть «лесом».

Структура записи для каждого уровня одинакова, на разных уровнях могут быть различные структуры. Обход сверху вниз, слева направо.

✓ Основные принципы проектирования:

1. Каждый «потомок» имеет только одного «предка»;
2. «Предок» может иметь несколько «потомков».

✓ Основными СУБД являются
IMS, PC/Focus, Team-Up, Data Edge,
Ока, ИНЭС, МИРИС.



Сетевая модель данных

- ✓ Является обобщенным вариантом иерархической модели.
- ✓ Структура - произвольный граф.

Два вида описателей: описание записи (группы записей) и описание связей двух типов (для записей «предка» и «потомка»).

- ✓ Основные принципы проектирования:

1. Каждый «потомок» может иметь несколько «предков»;
2. «Предок» может иметь несколько «потомков».

- ✓ Основными СУБД являются IDMS, db_Vistal, СЕТЬ, СЕТОР, КОМПАС.



Реляционная модель данных

✓ 1970-1971 гг. Эдгар Кодд ввел понятие реляции (relation - отношение) и реляционные языки обработки данных.

✓ Структура – двумерные таблицы.

Кортеж – Строка (сведения об одном экземпляре объекта).

Атрибут – Столбец (представляет свойство объекта).

Атрибут, идентифицирующий кортеж, называется *ключом*.

Домен – набор атрибутов одного столбца.

Степень реляции – число атрибутов в реляции.

✓ Основными СУБД являются:

MS Access, MS FoxPro, PARADOX, MySQL и др.

Среди более мощных (типа «клиент-сервер») широкое распространение получили Oracle, MS SQL Server, Informix.

СТУДЕНТ				
Код	Фамилия	Имя	Дата Р	Группа
22	РУСИНА	ОЛЕСЯ	01.08.77	11
11	ИВАНОВ	АЛЕКСЕЙ	10.06.78	11
12	ЖДАНОВ	СЕРГЕЙ	14.03.71	12
13	ПЕТРОВ	ВАДИМ	25.10.75	32
21	КРОХИН	МАКСИМ	30.12.77	23

Структуры и виды связей реляционной модели

данных

- ✓ Основные информационные *объекты* – двумерные таблицы (отношения или реляции).
- ✓ Между ними устанавливаются *связи* (бинарные ассоциации).
- ✓ Процесс трансформации данных в реляционную форму – *нормализация* – удаление избыточных данных из каждой таблицы.
- ✓ *Главная цель нормализации* – получение такого проекта базы данных, в котором каждый факт появляется лишь в одном месте.

12



Пять форм (стадий) нормализации

1. Первая нормальная форма (1НФ) –
 - каждое поле таблицы должно быть неделимым и
 - не должно содержать повторяющихся групп.
2. Вторая нормальная форма (2НФ) –
 - Требования 1НФ,
 - Все не ключевые поля полностью зависят от первичного ключа.
3. Третья нормальная форма (3НФ) –
 - Требования 2НФ,
 - Все не ключевые поля не зависят друг от друга.

Пять форм (стадий) нормализации

Усиленная третья форма (НФБК) –

- Требования 3НФ,
- Отсутствуют зависимости ключей от не ключевых атрибутов.

4. Четвёртая нормальная форма (4НФ) –

- Запрещает хранить независимые элементы в одной и той же таблице, когда между этими элементами существует связь типа «многие-ко-многим».

5. Пятая нормальная форма (5НФ) –

- Имеется возможность перестраивать данные в нормализованных таблицах, в которые они были переведены (сохранение всех элементов в базе в процессе нормализации).

Пример нормализации (до 3-ей формы)

Отношения:

- НомЗачКн
- ФИО
- Группа
- Специальность
- Квалификация

Пример нормализации (до 3-ей формы): 1НФ

16

1. ФИО →
 - **Фамилия**
 - **Имя**
 - **Отчество**

Отношения:

- **НомЗачКн**
- **ФИО**
- **Группа**
- **Специальность**
- **Квалификация**

1Н

Отношения:

- **НомЗачКн**
- **Фамилия**
- **Имя**
- **Отчество**
- **Группа**
- **Специальность**
- **Квалификаци**

Пример нормализации (до 3-ей формы):

2НФ

2. Первичный ключ :

- **НомЗачКн**

17

Отношения:

- НомЗачКн
- Фамилия
- Имя
- Отчество
- Группа
- Специальность
- Квалификаци

2Н

Отношения:

- **НомЗачКн**
- Фамилия
- Имя
- Отчество
- Группа
- Специальность
- Квалификаци

Пример нормализации (до 3-ей формы):

3НФ

3. Атрибуты **Специальность** и **Квалификация** полностью зависят от атрибута **Группа**, который не является ключевым. Разделим данное отношение на два:

- **СТУДЕНТЫ** - **НомЗачКн** (ключ), **Группа**, **Фамилия**, **Имя**, **Отчество**
- **ГРУППЫ** - **Группа** (ключ), **Специальность**, **Квалификация**

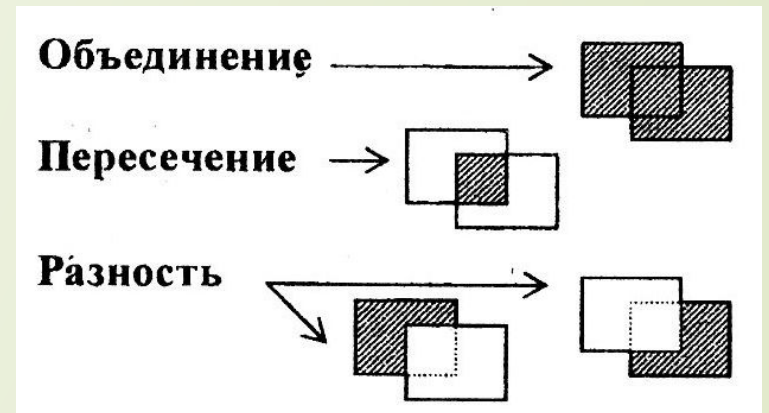
18



Операции реляционной алгебры

Над эквивалентными отношениями!

1. Объединение
2. Пересечение
3. Разность



ПРИМЕР

R1 Знатоки английского языка

Код	Группа	Фамилия	Имя
1	14	Иванова	Анна
2	11	Петров	Иван
4	12	Андреев	Иван

R2 Знатоки немецкого языка

Код	Группа	Фамилия	Имя
2	11	Петров	Иван
3	12	Сидоров	Степан
4	12	Андреев	Иван
5	14	Павлова	Нина

R3 = R1 ∪ R2 Объединение

Код	Группа	Фамилия	Имя
1	14	Иванова	Анна
2	11	Петров	Иван
3	12	Сидоров	Степан
4	12	Андреев	Иван
5	14	Павлова	Нина

R4 = R1 ∩ R2 Пересечение

Код	Группа	Фамилия	Имя
2	11	Петров	Иван
4	12	Андреев	Иван

R5 = R1 / R2 Разность

Код	Группа	Фамилия	Имя
1	14	Иванова	Анна

R6 = R2 / R1 Разность

Код	Группа	Фамилия	Имя
3	12	Сидоров	Степан
5	14	Павлова	Нина

Операции реляционной алгебры

4. **Расширенное декартово произведение** отношения R1 степени N (3) и отношения R2 степени M (2) называется отношением R3 степени N+M (5), которое содержит кортежи, полученные сцеплением кортежей R1 и R2.

21



Операции реляционной алгебры

5. Горизонтальный выбор (операция фильтрации)

R1

Код	Группа	Фамилия	Имя
1	14	Иванова	Анна
2	11	Петров	Иван
3	12	Сидоров	Степан
4	14	Павлова	Нина

R2 = R1 [Группа = 14]

Код	Группа	Фамилия
1	14	Иванова
5	14	Павлова

6. Вертикальный выбор (операция проектирования)

R1

Код	Группа	Фамилия	Имя
1	14	Иванова	Анна
2	11	Петров	Иван
3	12	Сидоров	Степан
5	14	Павлова	Нина

R3 = R1 [Группа]

Группа
14
11
12

22

7. Операция условного соединения

R1

Кинотеатр	Фильм
Русь	Матрица
Модерн	Звонок
Эдисон	Матрица
Мир	Бум

Кинотеатры

R2

Фильм	Жанр	Цена
Матрица	Боевик	120
Бум	Мелодрама	100
Звонок	Триллер	110

Фильмы

R3

Кинотеатр
Русь
Модерн
Эдисон

8. Операция деления

R1

Кино-театр	Дата	Фильм	Жанр
Русь	20.01	Матрица	Боевик
Модерн	21.01	Матрица	Боевик
Эдисон	22.01	Матрица	Боевик
Русь	21.01	Звонок	Триллер
Модерн	22.01	Бум	Разное
Эдисон	23.01	Звонок	Триллер

Сеансы

R2

Фильм	Жанр
Матрица	Боевик
Звонок	Триллер

Фильмы

R3

Кино-театр
Русь
Эдисон

Модели данных физического уровня

Предназначены для управления файловыми и бесфайловыми структурами данным на уровне физических устройств.

1. Файловая структура

- ✓ **Файл** - это линейная последовательность записей, расположенных на внешнем носителе.

Файлы последовательного доступа имеют переменную длину записи.

Файлы прямого доступа имеют переменную длину записи.

Файлы индексной структуры – дополнение к файлам с собственно данными и содержат списки номеров записей таблицы в порядке сортировки по заданному полю (по ключу).

- Плотный индекс
- Неплотный индекс
- В-деревья

Модели данных физического уровня

2. Бесфайловая структура

- ✓ Защита данных от злоумышленников и конкурентов.
- ✓ Данные в виде *страниц данных*:
 - заголовок страницы;
 - данные;
 - слоты с информацией об адресе, длине страницы и т.д.
- ✓ Данные объединены в экстенды (от 4-х страниц, 2 Гбайта).