

## Глава 1 - введение в технологию

1. Укажите отличительные особенности технологии БД выделяющие ее среди технологий вообще и информационных технологий в частности.

База данных – это технология регламентирующая процесс создания и эксплуатации информационных систем в которых для долговременного хранения информации используется база данных а управляет ею некоторая система управления базами данных (СУБД).

Информационную технологию баз данных отличает использование СУБД той или иной модели данных (МД) – концепции представления в ней объектов мира и их взаимосвязей. Развитые модели, обеспечивающие адекватное представление данных, относящихся к различным предметным областям, в конечном итоге определяют степень соответствия СУБД целям их создания.

2. Каковы предпосылки использования технологии БД?

- 1) Необходимость в долговременном хранении информации;
- 2) Необходимость обеспечения взаимодействия нескольких пользователей
- 3) Нужно хорошо структурированное и непротиворечивое хранилище
- 4) Необходимость вносить изменения
- 5) Необходимость в простом и удобном извлечении информации
- 6) Производная информация должна получаться из первичной без разработки спец. алгоритмов
- 7) Хранение информации в алфавитно-цифровой форме

3. В чем особенности термина модель в технологии БД?

Модель в технологии БД – теория моделирования а сама БД – результат моделирования. В других отраслях знаний моделью называют результат.

4. Каково основное назначение моделей данных?

Основное назначение МД – обеспечение процесса интерпретации (осмысления) данных и превращения их в информацию.

5. Перечислите составные части любой модели данных.

Для этого в каждой модели данных имеются следующие компоненты:

- правила порождения допустимых **структур** данных;
- правила порождения **ограничений целостности** данных;
- набор **операций** над данными.

6. Каково назначение структур данных ограничений целостности и операций над данными?

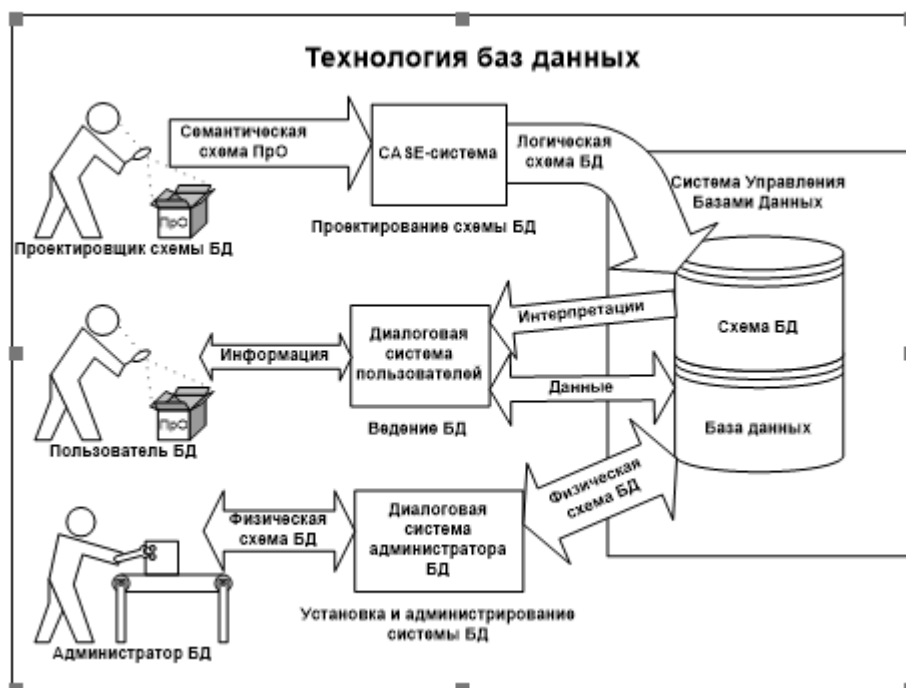
На основании правил определяется схема БД (основные понятия предметной области и их взаимосвязи) а операции необходимы чтобы первоначально наполнить схему непротиворечивыми данными и поддерживать их в дальнейшем в адекватном предметной информации.

Правила структуризации определяют БД как совокупность отношений (таблиц). Каждое отношение соответствует одному типу объектов (смысл их представляется именем отношения) и задает одинаковый для всех объектов этого типа набор атрибутов (столбцов). В схеме отношения для каждого атрибута определено его имя представленное в таблице в заголовке столбца. Таким образом структуру отношения может олицетворять поименованная шапка пустой таблицы.

Ограничения целостности связаны с поддержанием БД в целостном непротиворечивом состоянии максимально соответствующем состоянию моделируемой предметной области.

До тех пор пока не определена как минимум структура данных система не позволит вводить эти данные. Это происходит потому что при создании изменении и удалении данных обязательно должно быть указано элементом какой структуры они являются. Так при определении в системе нового человека нужно с помощью соответствующей операции создать строку именно в таблице ЧЕЛОВЕК причем значения в строке должны четко соответствовать шапке таблицы.

7. Укажите основные процессы в системах БД кто и какие задачи в них реализует.



К **необязательным** но часто важным процессам создания информационных систем (ИС) по этой технологии **относятся процессы разработки специализированного программного обеспечения (ПО) реализующего специфическую для предметной области (ПрО) бизнес-логику.**

В двух словах **этот процесс состоит** в выполнении следующих действий. **Проектировщик схемы БД изучает самостоятельно и с помощью экспертов процессы ПрО структуры и закономерности ее информационного описания и формализует полученные знания в виде семантической схемы ПрО.** Эта схема строится с использованием той или иной семантической модели данных концепция которой максимально близка человеческому восприятию мира.

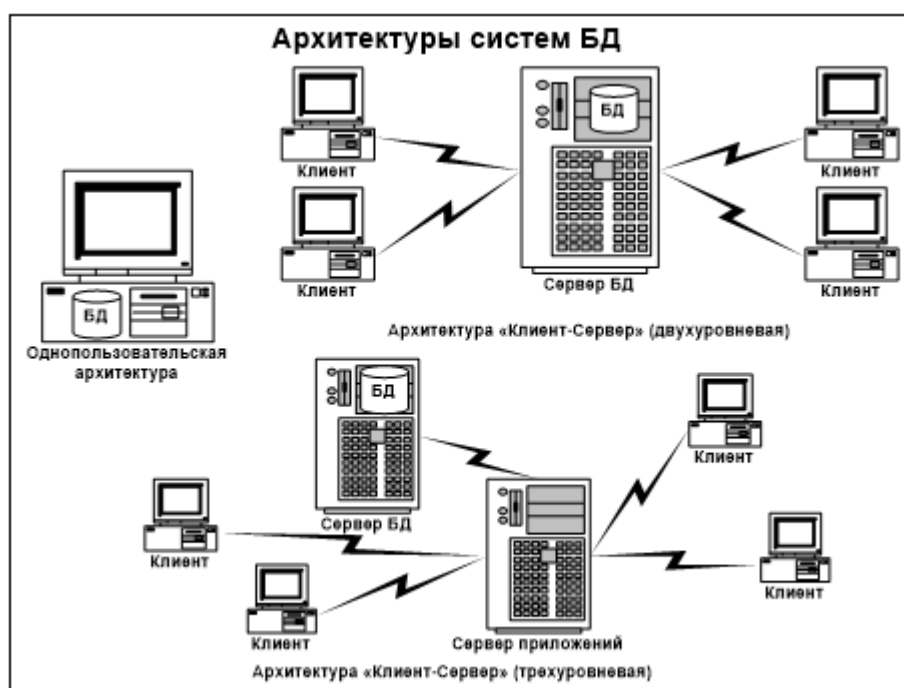
Проектирование схемы БД - Проектировщик БД - исследует ПрО и реализует ее семантическую схему

- Администрирование БД - Администратор БД - устанавливает БД на оборудование и администрирует ее. (физическая схема БД)  $\Rightarrow$  СУБД
- Использование БД - Пользователь БД - “пишет” или “читает” необходимую информацию. (это основной процесс ибо ради этого БД создается и поддерживается)

#### 8. Для чего предназначены CASE-системы?

Осуществить ввод семантической схемы а во-вторых **автоматически перевести ее на язык СУБД без потери информации.** Каждый элемент семантической схемы должен найти свое отражение в логической модели СУБД (в крайнем случае – в процедурной форме).

#### 9. Перечислите и кратко охарактеризуйте архитектуры систем БД.



В **однопользовательской архитектуре** вся система БД управляется одним компьютером на котором расположены и БД и СУБД и диалоговая система пользователя.

Клиент-сервер (двухуровневая) - обработка БД на сервере диалог с пользователем происходит через компьютер (клиент).

Наконец в настоящее время БД используются в приложениях для Интернета и интрасетей. Во многих приложениях БД применяющих интернет-технологии используется **трехуровневая архитектура** включающая сервер БД web-сервер и браузер. Между web-сервером и браузером передаются web-страницы клиентский код и данные между web-сервером и сервером БД - SQL-операторы и реляционные данные. Каждый из этих уровней может работать под управлением своей операционной системы на отдельном компьютере. Так же как и двухуровневую трехуровневую архитектуру программных продуктов технологии БД можно реализовать и на одном компьютере.

10. В чем основные различия OLTP -систем БД и хранилищ данных? Что такое OLAP-системы?

Первоначально СУБД проектировались для управления большим потоком транзакций каждая из которых сопровождалась внесением незначительных изменений в оперативные данные предприятия – т.е. в данные которые предприятие обрабатывало в процессе своей повседневной деятельности. Системы подобного типа называются системами **оперативной обработки транзакций** или OLTP-системами (англ. Online Transaction Processing – OLTP).

**!!! Отличие OLTP от хранилища данных:**

- 1) OLTP - позволяет редактировать\ удалять\ добавлять\ обновлять в онлайн
- 2) Хранилище просто хранит позволяет посмотреть инфу хоть где и хоть когда

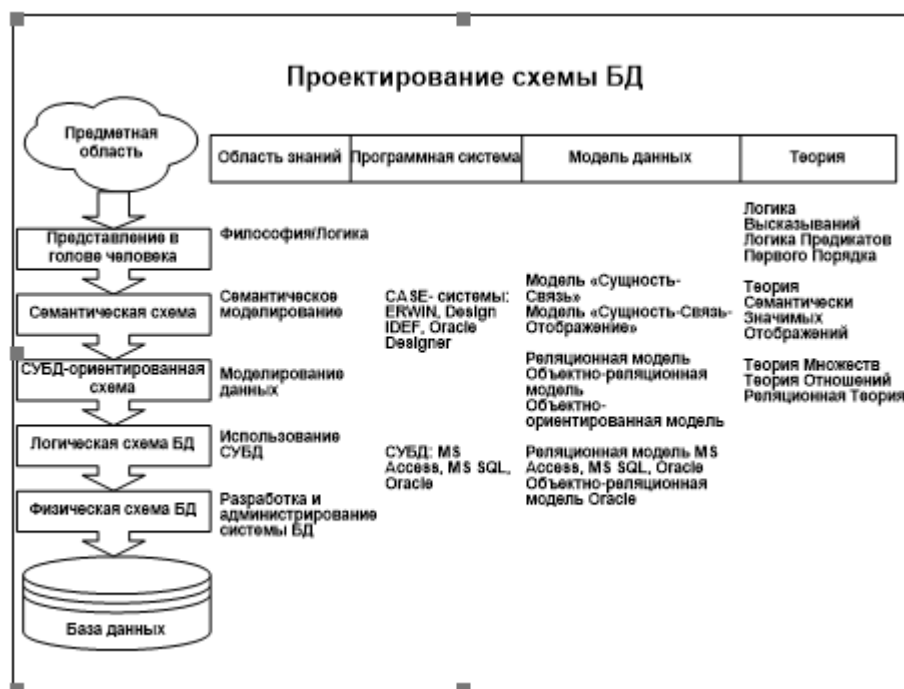
Основными средствами анализа в последние годы стали инструменты **оперативной аналитической обработки** (англ. Online Analytical Processing – OLAP). Термин OLAP определяет архитектуру которая поддерживает сложные аналитические приложения. Большинство OLAP-приложений создается на основе специализированных многомерных СУБД с ограниченным по структуре набором данных и настраиваемым пользовательским интерфейсом.

В OLAP-технологии серверы БД для хранения данных и связей между ними используют многомерные структуры. Многомерные структуры лучше всего представлять как многомерные кубы данных. В диалоговых средствах OLAP многомерный куб визуально представляется в виде динамично изменяемых таблиц. Данные откладываются по двум осям которыми являются строки и столбцы. Иногда используется третья ось значениям которой соответствуют закладки. В таком случае на каждой закладке представлена своя таблица значений иногда называемая слоем.

Позволяют работать с многомерными хранилищами.

Диалоговые OLAP-инструменты позволяют динамично строить всевозможные слои и проекции куба с использованием техник drag and drop (перетаскивание элементов диалога с одного места экрана на другое) и drill-down and drill-up (перемещение по уровням иерархии членов измерения с агрегированного на более детальный и наоборот). Тем самым легко и быстро без помощи программистов пользователь может получать различные интересующие его взгляды на объект анализа.

11. Какие формы представлений о ПрО выделяют в технологии БД?  
Охарактеризуйте каждую из них.



- 1) Представление в голове человека формируется на основании личного знакомства с ПрО
- 2) Семантическая схема формальное представление
- 3) СУБД-ориентированная схема ПрО на диалекте СУБД-ориентированной МД
- 4) Логическая схема ПрО обращена к человеку и предоставляет языковые и диалоговые методы общения
- 5) Физическая обращена к средствам хранения данных в оперативной и внешней памяти

## Глава 2.1 - основные понятия

12. Объясните почему человека интересуют не данные а информация.

Сами данные ничего не говорят человеку а информация - это приращение знаний человека полученное на основе некоторых данных. Данные представлены в формализованном виде удобном для передачи или их обработки а информация представлена в доступной для усвоения человеком форме.

13. Что кроме данных необходимо для получения информации?

Кроме данных необходимы интерпретации (см кг) и знания человека. Интерпретация данных дает человеку представление о чем именно свидетельствуют эти данные. На основе знаний человека может быть получена дополнительная информация во время этапа прагматического анализа.

Информация – приращение знаний человека которое может быть получено на основе данных.

14. К какому классу информационных систем относятся системы БД с точки зрения принципов получения информации?

1. Интерпретациями владеет только человек.
2. Интерпретации указаны в программе
3. **Интерпретации хранятся на диске в виде специальным образом организованных данных. Интерпретации хранятся на диске в виде структурированных данных.**

**Заданные однократно интерпретации ассоциируются с данными при их вводе и в дальнейшем при всех манипуляциях с данными эта связь остается неразрывной обеспечивая таким образом получение информации**

15. Перечислите и охарактеризуйте три этапа процесса образования информации из данных.

Синтаксический анализ семантический анализ прагматический анализ

- 1) Синтаксический анализ – получение из данных информации путем анализа знаков с точки зрения синтаксиса;
- 2) Семантический анализ – получение из данных информации с помощью интерпретации данных;
- 3) Прагматический анализ - получение из данных дополнительной информации на основе тех знаний читателя которые у него были ранее.

**Синтактика – раздел семиотики изучающий внутренние свойства систем знаков безотносительно к интерпретации** (синтаксис – правила

построения знаков и знакосочетаний в рамках знаковой системы). Просто знаки с точки зрения синтаксиса без адресатов и интерпретаций.

**Семантика** – раздел семиотики рассматривающий отношение знаков к обозначаемому (содержание знаков) то есть соотношения между знаками и их интерпретациями независимо от того кто служит адресатом (интерпретатором).

Термин семантически значимый который мы в дальнейшем будем часто использовать предполагает значимость с точки зрения семантики изучающей семантические отношения которые образуются между объектами и знаками представляющими эти объекты в знаковой системе. Таким образом семантически значимыми могут быть как объекты (если они представлены некоторыми знаками в семиотической системе) так и знаки (если они определяют некоторые реальные объекты ПрО).

**Прагматика** – раздел семиотики изучающий восприятие выражений знаковой системы в соответствии с разрешающими способностями воспринимающего. Прагматика исследует связь знаков с адресатом т. е. проблемы интерпретации знаков теми кто их использует их полезность и ценность для интерпретатора.

16. Какие разделы семиотики изучают эти этапы?

Процесс получения информации в знаковых системах проходит в три этапа, каждому из которых соответствует свой раздел семиотики.

Синтактика, семантика и прагматика – разделы семиотики, изучающие этапы процесса образования информации из данных.

**Синтактика** – раздел семиотики, изучающий внутренние свойства систем знаков безотносительно к интерпретации (синтаксис – правила построения знаков и знакосочетаний в рамках знаковой системы).

**Семантика** – раздел семиотики, рассматривающий отношение знаков к обозначаемому (содержание знаков) или, что то же, соотношения между знаками и их интерпретациями, независимо от того, кто служит «адресатом» (интерпретатором).

**Прагматика** – раздел семиотики, изучающий восприятие выражений знаковой системы в соответствии с разрешающими способностями воспринимающего. Прагматика исследует связь знаков с «адресатом», т. е. проблемы интерпретации знаков теми, кто их использует, их полезность и ценность для интерпретатора.

17. Опишите как протекает процесс передачи информации о ПрО с использованием систем БД.

I этап. **Проектировщик производит у себя в голове первичную формализацию** ситуации (представления) в частности **определяет знаки для идентификации объектов и связей.**

II этап. **Обращаемся к знаковой системе** (создание новых объектов типов и их связей).

III этап. **Читатель воссоздает представление** писателя о ПрО обращаясь к системе.

#### 18. В чем заключается основное назначение модели данных?

Модель данных (МД) обеспечивает процесс интерпретации данных и превращает их в информацию.

#### 19. Из каких компонентов состоит атомарная единица информации (АЕИ)?

Она определяется четверкой – <Идентификатор объекта Наименование признака Значение признака [Время]>; Каждая АЕИ задает истинность следующего факта: объект на который указывает идентификатор имеет определенное значение признака заданного именем в конкретное время .

#### 20. Объясните почему первые три компонента АЕИ являются обязательными.

Нельзя сформировать информацию без первых 3 компонентов.

Последний элемент четверки не напрасно заключен в квадратные скобки означающие что он может быть опущен. Действительно все остальные элементы являются обязательными. Если опустить хотя бы один из них информация не образуется. Что же касается времени то не существует моделей данных в которых механизмы работы с так называемыми темпоральными данными были бы доведены до идеала. Более того подавляющее большинство моделей и БД предполагают хранение информации об одном (текущем) состоянии ПрО. Такие БД еще называют оперативными подчеркивая тот факт что они меняются синхронно с изменением состояния ПрО. Предыдущие состояния данных в них не сохраняются. Поэтому в дальнейшем под АЕИ будем понимать тройку – <Идентификатор объекта Наименование признака Значение признака>. Примеры АЕИ приведены на слайде.

#### 21. Синтезируйте категориальную модель.

Если есть несколько таблиц описывающие объект одной категории то категориальная модель предполагает объединение этих таблиц в одну единую.

Понятие категории является основным структурным понятием одноименной модели – категориальной модели. Она предполагает разбиение всех объектов ПрО по категориям. Для каждой категории определяется набор признаков значения которых характеризуют объекты данной категории.

Совокупность именованных категорий и их признаков а также ограничений на допустимые данные называется схемой БД.

Совокупность данных структура и значения которых соответствуют конкретной схеме называется базой данных (БД).

{{ Сначала построим таблицы признаков (*Рост* и *Вес*) а затем соберем в одну таблицу значения всех признаков однотипных объектов. Полученная таблица описывает все объекты одной категории – *ЧЕЛОВЕК*. Понятие категории является основным структурным понятием одноименной модели – категориальной модели. Она предполагает разбиение всех объектов



ПрО по категориям. Для каждой категории определяется набор признаков значения которых характеризуют объекты данной категории. }}

## 22. Дайте структурное определение модели данных.

Модель данных определяется 2 множествами:

- 1) Множество правил порождения схем
  - a. Правила порождения структур данных
  - b. Правила порождения ограничения целостности
- 2) Множество операций над данными

## 23. Проиллюстрируйте компоненты модели данных на примере категориальной модели.

Множество порождений допустимых структуры данных:

- ☐ БД – это совокупность таблиц.
- ☐ Каждая таблица предназначена для хранения информации об объектах одной категории. Имя таблицы – это имя категории.
- ☐ Для каждой категории определяется набор признаков представляющих интерес для объектов этой категории. Имена признаков составляют шапку соответствующей таблицы.
- ☐ Каждый объект категории представляется в виде строки таблицы в столбце признака указывается его значение для данного объекта.

Множество правил порождения ограничений целостности:

- ☐ указанием их типа (символьные числовые даты и т.д.)
- ☐ перечислением этих значений
- ☐ сравнением значений с константой.

Множество операций над данными:

- ☐ операция INSERT для добавления новой строки в таблицу
- ☐ операция UPDATE для изменения значений одного или нескольких признаков в строке таблицы
- ☐ операция DELETE для удаления строки из таблицы
- ☐ операция SELECT для поиска строк таблицы удовлетворяющих определенному условию.

## 24. С чем ассоциируются понятия схема БД и база данных при табличном представлении данных?

Совокупность именованных категорий и их признаков а также ограничений на допустимые данные называется **схемой БД**.

Совокупность данных структура и значения которых соответствуют конкретной схеме называется **базой данных (БД)**.

Схема БД ассоциируется с совокупностью категорий и ограничениями на корректность данных.  
БД ассоциируется с совокупностью таблиц и внесенных в них данных

По поводу этих определений следует сделать несколько замечаний.

Во-первых как мы увидим в дальнейшем каждая модель данных предполагает свой набор понятий используя которые мы структурируем ПрО. Наши определения даны для категориальной модели поэтому в них используются понятия категория и признак. Тем не менее лучше использовать конкретные но конструктивные определения чем давать размытые определения типа база данных – это некоторый набор перманентных данных используемых прикладными системами какого-либо предприятия. Во-вторых в этих определениях указано основное содержание понятий схема и база данных. Иногда это содержание так или иначе расширяют. Так в схему БД могут быть включены программы обработки данных определения запросов диалоговых форм и т.д. БД тоже может представляться как все что хранится на диске помимо собственно СУБД. Мы и сами иногда будем поступать так. В частности на приведенной схеме БД вобрала в себя помимо собственно данных еще и их схему.

25. Что представляет собой СУБД? Какие классы операций необходимо реализовать в любой СУБД?

СУБД (система управления баз данных) - специализированные программные средства управления БД на ЭВМ каждая из которых предлагает свои языковые и диалоговые формы для множеств G и O: язык определения данных (ЯОД) и язык манипулирования данными (ЯМД). Иногда выделяют отдельный язык определения ограничений целостности (ЯООЦ) но чаще ограничения целостности задаются вместе со структурой в командах ЯОД.

## Глава 2.2 - Структуры

26. Что представляют собой элементарные единицы данных и интерпретаций. Для чего они используются в технологии БД?

Знак – это данное (строка символов число дата или их агрегат) определяющее конкретный объект связь объектов или значение характеристики объекта.

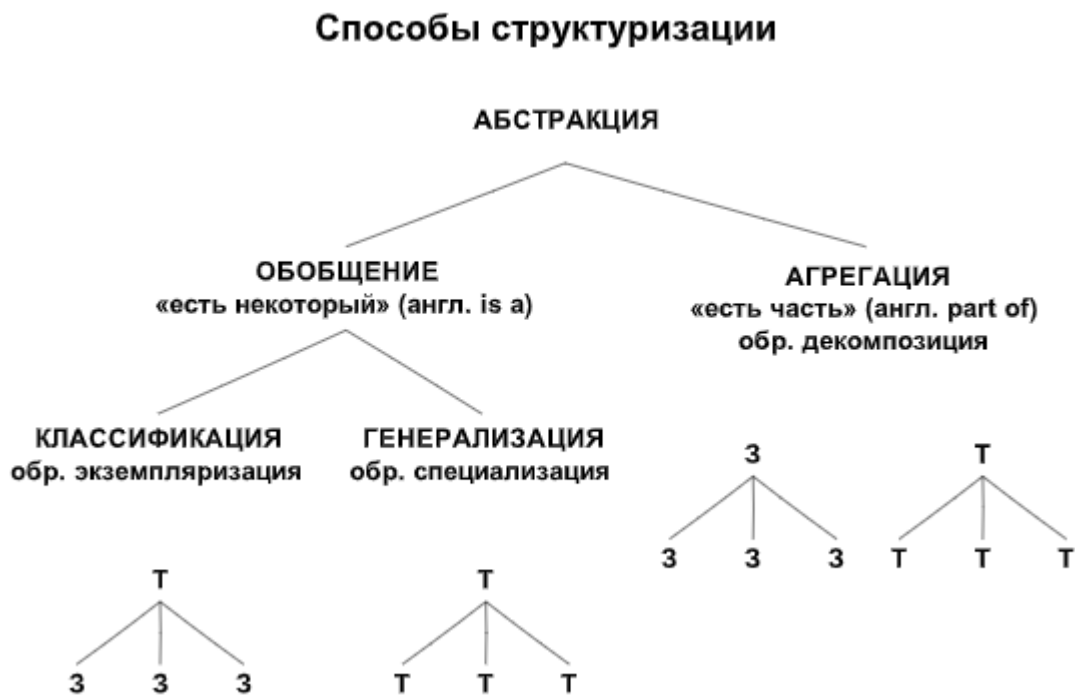
Тип – это именованный класс подобных знаков. Имя типа определяет интерпретацию принадлежащих типу знаков.

27. В каких частях системы БД представлены элементарные единицы данных и интерпретаций?

Структурный компонент схемы БД представляет собой определения типов (типов объектов ПрО типов их связей и их характеристик). Типы в схеме не изолированы друг от друга а образуют взаимосвязанную систему типов.

Задачей структуризации ПрО в конкретной модели данных и является построение такой системы типов с использованием понятий предлагаемых правилами структуризации данных этой модели.

28. Какие мыслительные процедуры используются при структуризации данных?  
Уточните когда применяется каждая из них.



Абстракция используется для образования категорий данных. Обобщение – абстракция из множеств типов и знаков (есть некоторый) множество знаков или типов соотносятся с одним типом. Агрегация – абстракция объект конструируется из базовых объектов (в отношении есть часть)

Классификация – это обобщение знаков до типа (отношение экземпляр-класс)  
а генерализация – это обобщение типов до типа (отношение подкласс-суперкласс).

29. Поясните как абстракция используется для интерпретации данных?

Абстракция используется для придания множествам объектов некой семантики. В дальнейшем если в ПрО обнаруживается объект подпадающий под некоторое понятие мы строим его абстрактное представление определяя значения фиксированных для понятия признаков.

30. Перечислите традиционно используемые в моделях данных формы данных.

Комплекс множество кортеж домен отношение атрибут

Комплекс – это собрание правильно идентифицированных объектов удовлетворяющих условию  $i$ -принадлежности. Отношение  $i$ -принадлежности между элементом  $x$  и комплексом  $Y$  наблюдается тогда и только тогда когда  $x$  находится в  $i$ -ой позиции комплекса  $Y$ .

Множество – это собрание правильно идентифицированных объектов удовлетворяющих условию принадлежности. (это комплекс все элементы которого находятся в одной позиции).

Кортеж (n-местный) – это комплекс который имеет по одному элементу в каждой позиции от 1 до n (строка таблицы).

Домен - множество элементы которого синтаксически однородны. # {1 2 3 4 5}

Атрибут - именованный домен представляющий семантически значимые объекты. Отношение - множество кортежей.

### 31. Укажите свойства множеств

- множества необязательно должны иметь фиксированное количество элементов они могут быть и бесконечными;
- на элементах множества отсутствует какой-либо порядок;
- элементы-дубликаты во множестве не имеют смысла.

### 32. Дайте определения интенционала и экстенционала множества. Почему в моделировании данных уместно говорить о нескольких реализациях множества?

Интенционал множества определяет свойства общие для всех элементов всех реализаций множества (в математике его называют условием принадлежности).

Экстенционал множества определяет актуальную реализацию множества путем явного указания его элементов. Обратите внимание в условии интенционала фигурируют сами объекты ПрО в экстенционале же представлены знаки идентифицирующие эти объекты.

При моделировании в БД динамичных ПрО одному и тому же множеству в различные моменты времени могут соответствовать различные реализации.

В качестве примера можно рассмотреть множество студентов сидящих в конкретной аудитории. Его экстенционал будет меняться чуть ли не на каждой паре занятий при неизменном интенционале.

### 33. Чем отличается комплекс от множества?

В отличие от множества элементы комплекса распределены по позициям определяемым целыми положительными числами. Причем в каждой позиции может быть любое (в т.ч. и бесконечное) число различных элементов.

В пределах одной позиции порядок элементов не определен.

В разных позициях могут быть совпадающие элементы. Позиция элемента в комплексе записывается с помощью верхнего индекса.

Множество – это собрание правильно идентифицированных объектов удовлетворяющих условию принадлежности. Данное определение отражает

специфику моделирования данных в частности в нем фигурируют правильно идентифицированные объекты .

Комплекс – это собрание правильно идентифицированных объектов удовлетворяющих условию  $i$ -принадлежности. Отношение  $i$ -принадлежности между элементом  $x$  и комплексом  $Y$  (записывается  $x \in Y$ ) наблюдается тогда и только тогда когда  $x$  находится в  $i$ -ой позиции комплекса  $Y$ .

34. Покажите что понятия множество и кортеж являются специализациями понятия комплекс .

Множество – это комплекс все элементы которого находятся в одной позиции.  $n$ -местный кортеж или просто

кортеж – это комплекс который имеет по одному элементу в каждой позиции от 1 до  $n$ .

35. В чем принципиальное отличие элементов принадлежащих домену и атрибуту?

Существуют некоторые множества элементы которых более или менее синтаксически однородны например множество целых чисел от 10 до 20

строки букв длиной до 20 символов и т.д. Подобные однородные множества в моделировании данных носят название доменов. Домены могут быть однородными как чисто синтаксически (все их элементы удовлетворяют одному и тому же синтаксису) так и семантически (в их условии принадлежности фигурируют особенности конкретной ПРО). Примерами доменов первого вида являются домен целых положительных чисел (ЦПЧ) домен строк символов компьютерного алфавита домен дат и т.п.

Именованные домены представляющие семантически значимые объекты называются атрибутами. Атрибуты являются интерпретацией объектов реального мира и их характеристик. Вводя атрибуты мы даем интерпретацию абстрактным понятиям таким как числа и строки.

Домен можно рассматривать как обобщение атрибутов (смотрите иерархию обобщения в левой нижней части последнего слайда). Атрибуты определенные на общем домене наследуют его свойства. И наоборот домен обладает всеми общими свойствами определенных на нем атрибутов. Обобщение в применении к атрибутам позволяет соотнести структуру данных с доменами и установить общность различных атрибутов.

Атрибут - это именованный домен представляющий семантически значимые объекты.

Домен – обобщение атрибутов (атрибут без имени). Атрибуты не существуют сами по себе они являются компонентами других объектов БД и посредством агрегации они ассоциируются с другими атрибутами.

36. Дайте определения и уточните различия определений понятия отношение в математике логике и моделировании данных.

Математическое отношение – это множество выражающее соответствие между двумя или более множествами. Соответствие между двумя множествами  $S_1$  и  $S_2$  называется бинарным отношением если оно является подмножеством прямого (Декартова) произведения  $S_1$  и  $S_2$  т.е. подмножеством

$S_1 \times S_2$ . Бинарное отношение задается множеством упорядоченных пар которые удовлетворяют некоторому критерию.

**Логическое отношение** – отображение произведения  $n$  на множество  $\{0, 1\}$  (ложь правда).

**Определение 2.2.14.** Пусть  $E^n = E \times \dots \times E$  есть произведение  $n$  множеств  $E$ , т.е. множество всех кортежей  $\langle x_1, \dots, x_n \rangle$ ,  $x_i \in E (i=1, \dots, n)$ . Отображение  $P : E^n \rightarrow \{0, 1\}$  называется  **$n$ -местным отношением (предикатом, логической функцией) над  $E$** . Множество  $A \subseteq E^n$  всех кортежей, для которых  $P(x_1, \dots, x_n) = 1$ , определяет «свойство» кортежей:  $x_1, \dots, x_n$  состоят в отношении  $P$  тогда и только тогда, когда  $\langle x_1, \dots, x_n \rangle \in A$ .

Отношение в моделировании – множество из заголовка и тела. Заголовок — это множество отображений названия атрибута в имя типа а тип — это множество множеств отображений названия атрибута в значение атрибута

Отличия: логическое отношение хранит истинность или ложность факта а в моделировании данных принято хранить только истинные факты.

От математического отношения отношение в моделировании данных отличается отсутствием порядка в элементах кортежа.

В математике отношение - это просто множество без семантической интерпретации и неизменным экстенсионалом в отличие от отношения в смысле моделирования данных.

37. В каком виде задаются в БД интерпретации данных?

В виде таблиц графов или их симбиоза.

Таблица: Объекты одного типа представляются отдельной таблицей олицетворяющей отношение. Таким же образом могут быть представлены типы связей. Заголовок – общее отношение и характер объектов таблицы строка – объект столбец – атрибут.

Граф: граф схемы представляется графом типов граф базы данных - графом знаков. Между вершинами и ребрами подразумевается отношение экземплярзации.

38. Каким формам данных приписываются интерпретации?

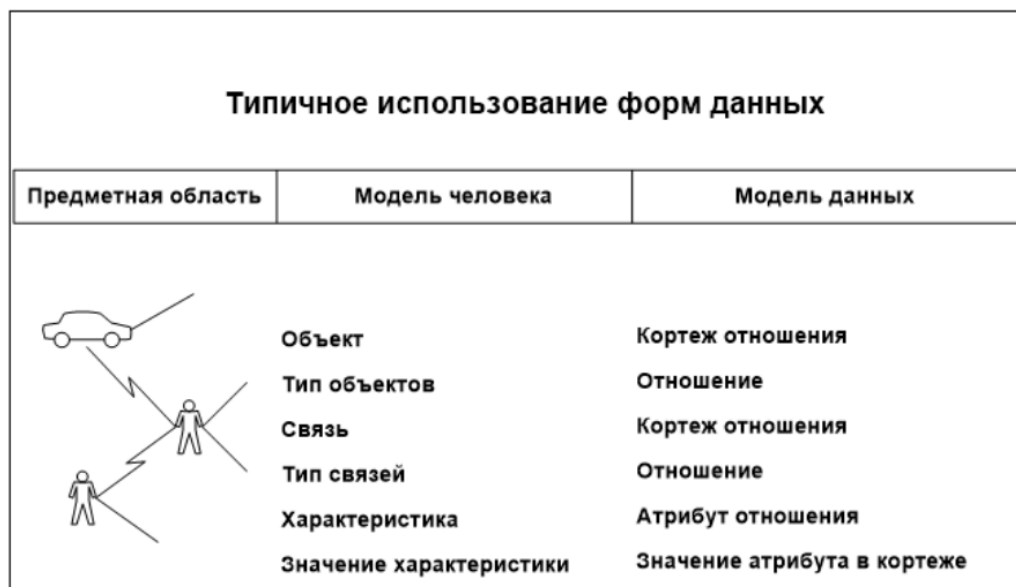
Атрибутам и отношениям.

Памятуя об основной функции любой модели данных – обеспечении интерпретации данных, переходим к рассмотрению традиционных и наиболее естественных способов интерпретации данных, представленных в приведенных ранее формах. Основной принцип очень прост – интерпретации наследуются в соответствии с иерархией обобщения. Возможно, простейшим и широко распространенным способом представления данных и их интерпретаций являются таблицы. С ними люди привыкли работать повседневно. Табличные формы представления характерны и для большинства моделей данных.

39. Как они используются для интерпретации знаков (данных)?

Кортеж – 180 см (данные + интерпретация)

Объект обладающий совокупностью индивидуальных значений характеристик представляется в виде кортежа этих значений. Для всех объектов одного типа выделяется одинаковый набор характеристик который составляет множество атрибутов отношения представляющего этот тип объектов. Каждая связь конкретного типа задается в БД в виде кортежа элементами которого являются знаки соответствующих объектов. Используя атрибуты происходит интерпретация абстрактных понятий таким как числа и строки.



Именованные вершины графа типов представляют атрибуты (*№ Служащего* *Фамилия* *Адрес* *Пол* *Название* *Город*) и типы объектов (*СЛУЖАЩИЙ* *КОМПАНИЯ*). Совокупность не имеющих имен ребер графа типов соединяющих атрибуты с их типом объектов соответствует отношению агрегации типов. Именованные ребра этого графа олицетворяют типы связей (*СЛУЖБА*) – агрегаты типов объектов.

40. Установите соответствие между формами данных и их представлениями в виде таблиц и графов.

### Таблицы

Соответствие между формами данных и таблицами: СМ КАРТИНКУ ВЫШЕ

- 1) Объекты одного типа представляются таблицей и показывают соответствующее отношение
- 2) Заголовок таблицы - общее имя и характеризует соответствующее понятие о предметах
- 3) Каждая строка таблицы - это кортеж соответствующий конкретному объекту и задающий его значения однозначных характеристик общих для всех объектов одного типа (атрибутов)
- 4) Каждый столбец - один атрибут имя которого помещается в шапку таблицы

5) Типы связи также могут быть представлены таблицами олицетворяющими соответствующие отношения. В табличной форме связи задаются дублированием значений атрибутов соответствующих объектов.

## ГРАФ

Соответствие между формами данных и графами:

Граф схемы - граф типов граф БД - граф знаков

- 1) Именованные вершины графа типов - атрибуты и типы объектов
- 2) Совокупность не имеющих имен ребер графа типов соединяющих атрибуты с их типом объектов - соответствуют отношению агрегации типов
- 3) Именованные ребра этого графа - типы связей
- 4) Вершины графа знаков - объекты и значения
- 5) Совокупность ребер знаков соединяющих значения с их объектом соответствует отношению агрегации знаков. Остальные ребра графа знаков на концах которых расположены вершины-объекты олицетворяют связи между этими объектами.

Агрегаты типов объектов – типы связей – также могут быть представлены таблицами олицетворяющими соответствующие отношения. В табличной форме связи задаются дублированием значений атрибутов соответствующих объектов.

### 41. Что ассоциируется с понятиями интенционал БД и экстенционал БД ?

Интенционал БД – (структура) схема БД включающая определения структур данных и ограничения целостности.

Экстенционал БД – текущая реализация БД из реализации определенных форм данных - множеств и отношений в схеме БД.

## Глава 2.3 - ограничение целостности

### 42. Что такое ограничение целостности (ОЦ) и для чего они предназначены?

Ограничение целостности (ОЦ) можно представлять себе как логическое условие которое для реализации данного множества атрибута отношения или нескольких отношений либо истинно либо ложно.

ОЦ не всегда задаются в виде явного логического выражения синтаксически они могут представлять собой просто декларацию определенных фактов. Однако механизм проверки ОЦ



СУБД всегда использует их именно как логические условия поскольку его задача – обеспечивать чтобы для любого состояния БД все ОЦ декларированные в ее схеме были истинными.

ОЦ вводятся в модели данных в целях повышения ее семантичности и расширения возможностей поддержания целостности данных. Первый аспект связан с адекватностью отражения реального мира в схеме а второй – с возможностями СУБД обеспечивать соответствие порождаемых состояний БД требованиям выражаемым ОЦ.

43. Дайте определения понятий связанных с процессом верификации ОЦ.

Рассмотрим явное ограничение  $C_i$  указанное в схеме  $S$  и состояние БД. Это ограничение будет:

- 1) правильно построено если оно соответствует синтаксическим правилам задания ОЦ;
- 2) удовлетворено состоянием БД если оно истинно для него;
- 3) удовлетворяемо если существует некоторое состояние БД удовлетворяющее ограничению;
- 4) недостоверно если никакое состояние БД не удовлетворяет ограничению;
- 5) логическим следствием ограничений  $C_1 \dots C_n$  (т.е. избыточным ограничением) если удовлетворяют все состояния БД которые удовлетворяют  $C_1 \dots C_n$ ;
- 6) эквивалентно другому ограничению если они – логические следствия друг друга.

Состояние БД удовлетворяет схеме если это состояние удовлетворяет всем ее ограничениям. Схема удовлетворяема если существует некоторое состояние БД удовлетворяющее ей. Схема противоречива если никакое состояние БД ей не удовлетворяет. Работоспособная схема должна быть удовлетворяема.

44. Охарактеризуйте способы проверки ОЦ с точки зрения их предпочтительности.

ОЦ не всегда задаются в виде явного логического выражения синтаксически они могут представлять собой просто декларацию определенных фактов.

Внутренние: (узкий диапазон возможностей): связаны с правилами структуризации они определяются самими формами данных и не требуют деклараций.

Явные: Реализуются на языке описания ограничений. Алгоритмы проверяют их непротиворечивость и обеспечивают создание алгоритмов контроля.

45. Какие типы ОЦ выделяются в моделировании данных? Каковы их области действия?

Большинство ОЦ можно отнести к одному из следующих типов:

1. Ограничения на значения атрибутов.

Область действия: Так областью действия ОЦ первого типа являются отдельные атрибуты. Редко но встречаются случаи когда они действуют на несколько атрибутов одного отношения.СУБД. Ее задачей является обеспечение того чтобы для каждого кортежа отношения после подстановки значений указанных в таких ОЦ атрибутов эти ограничения выполнялись.

На русском: их задача – максимально сузить область допустимых значений атрибута и таким образом исключить возможность появления в БД недостоверных значений.

2. Ограничения на отображения:

а) между атрибутами одного отношения: Область действия является одно отношение как агрегат атрибутов.

б) между отношениями: Область действия: подтипа могут охватывать несколько отношений (чаще всего – два) поскольку они отражают закономерности взаимосвязей между кортежами этих отношений

46. Перечислите виды ОЦ на значения атрибутов.

### **1) Ограничения целостности на значения атрибутов**

**Множество допустимых значений атрибутов можно задавать:**

- 1) принадлежностью к определенному типу или домену  
Рост INTEGER
- 2) сравнением с константой или значением атрибута того же отношения (возможны более сложные выражения)  
Рост > 50 или Рост > Вес + 110
- 3) диапазоном  
Рост BETWEEN 50 AND 300
- 4) перечислением значений  
Пол IN {'м', 'ж'}
- 5) более сложным логическим выражением, включающим в виде атомов конструкции 2 – 4  
(Пол = 'ж' AND Рост > Вес + 110) OR  
(Пол = 'м' AND Рост > 180)

47. В чем особенность традиционного определения понятия отображение в математике?

### Отображение в математике

**Отображение** (функция, оператор) есть закон соответствия, сопоставляющий каждому элементу множества  $A$  некоторый (единственный) элемент множества  $B$ .  $\square: A \rightarrow B$  означает, что задано отображение  $A$  в  $B$ , называемое  $\square$ .

#### Классификация Кофмана

**Соответствие**  $\Gamma$  между множествами  $E_1$  и  $E_2$  определено, если задан обычный граф  $G \subseteq E_1 \times E_2$ . Тогда говорят, что  $G$  – граф соответствия  $\Gamma$ ,  $E_1$  – область определения, а  $E_2$  – область значений  $\Gamma$ . Соответствие, обратное  $\Gamma$ , обозначается  $\Gamma^{-1}$ , и  $E_2$  – область определения, а  $E_1$  – область значений  $\Gamma^{-1}$ .

**Отображением** множества  $E_1$  во множество  $E_2$  называется такое соответствие, которое любому  $x \in E_1$  сопоставляет, по крайней мере, один  $y \in E_2$ . Тогда говорят, что элемент  $y$  – образ элемента  $x$ , а  $x$  – переменная или аргумент.

**Функцией**  $E_1$  в  $E_2$  называется такое отображение, которое каждому  $x \in E_1$  сопоставляет один и только один  $y \in E_2$ .

Особенность заключается в том, что требуется обязательное наличие единственного образа для каждого аргумента (отступление от правил-многозначные функции)

48. Укажите основные отличия математического отображения от семантически значимого отображения

**Семантически значимое** отображение (англ. mapping) (в дальнейшем просто – отображение) – это понятие, определяющее некоторый закон

предметной области, по которому каждому объекту моделируемого мира может быть поставлен в соответствие (а может быть и нет) один или более объектов.

Термин **семантически значимый** предполагает **значимость** с точки зрения семантики изучающей семантические отношения, которые образуются **между объектами и знаками**, представляющими эти объекты в знаковой системе. Благодаря этому эпитету мы дистанцируемся от абстрактных (в частности, математических) отображений и функций. В математическом отображении образы и их прообразы не меняются со временем. В течение времени эти экземпляры могут возникать и исчезать, у них могут меняться образы.

49. Дайте определения, характеризующие роли элементов множеств, участвующих в отображении.

Экземпляр отображения (или экземпляр-отображения) – соответствия одному объекту одного другого объекта.

Отображением объекта – все экземпляры одного и того же отображения, соответствующие одному объекту-прообразу.

Образ объекта  $x$  – объект  $y$ , поставленный в соответствие объекту  $x$  одним из экземпляров отображения.

Прообраз объекта  $y$  – объект  $x$ , которому соответствует  $y$ .

Область образов объекта - совокупность образов порождаемая всем отображением объекта  $x$ .

В каждом конкретном случае эта совокупность:

- может быть пустой (отображение неприменимо к объекту  $x$ ) или
- содержать один объект или
- содержать много объектов.

50. Почему в моделировании данных приходится рассматривать реальные и потенциальные ООО и ОЗО? Дайте им определения.

Семантика отображений стабильна и в любой момент времени отражает тот закон который справедлив для всех возможных состояний их экземпляров. По этой причине наряду с динамичными РООО и РОЗО для отображений необходимо рассматривать потенциальные области определения и значений отображений.

**Реальной областью определения отображения (РООО)** называется совокупность всех объектов имеющих образы при отображении

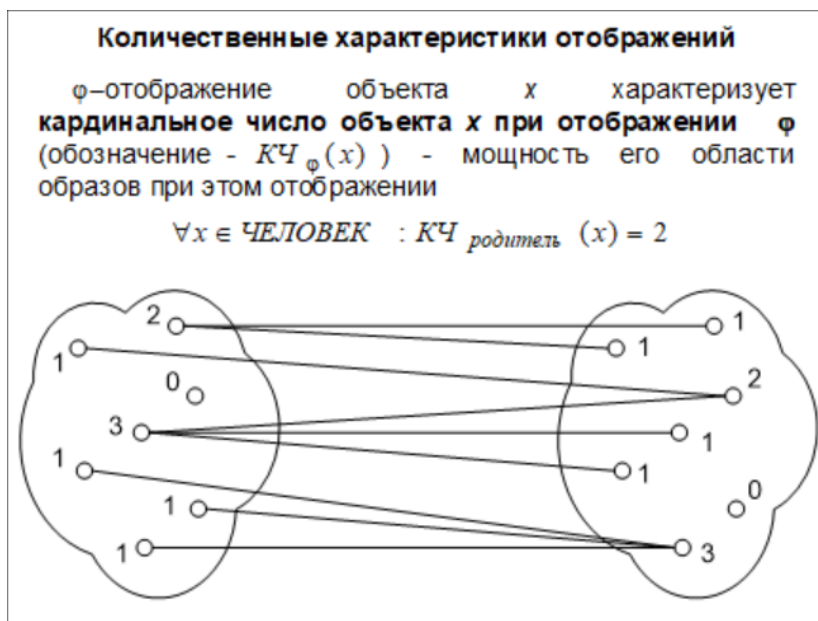
**Реальной областью значений отображения (РОЗО)** называется совокупность всех образов этих объектов

Область определения отображения (ООО) (англ. domain) – это совокупность объектов которые в силу смысла закона могли иметь в прошлом или имеют в настоящем или смогут иметь в будущем хотя бы один -образ.

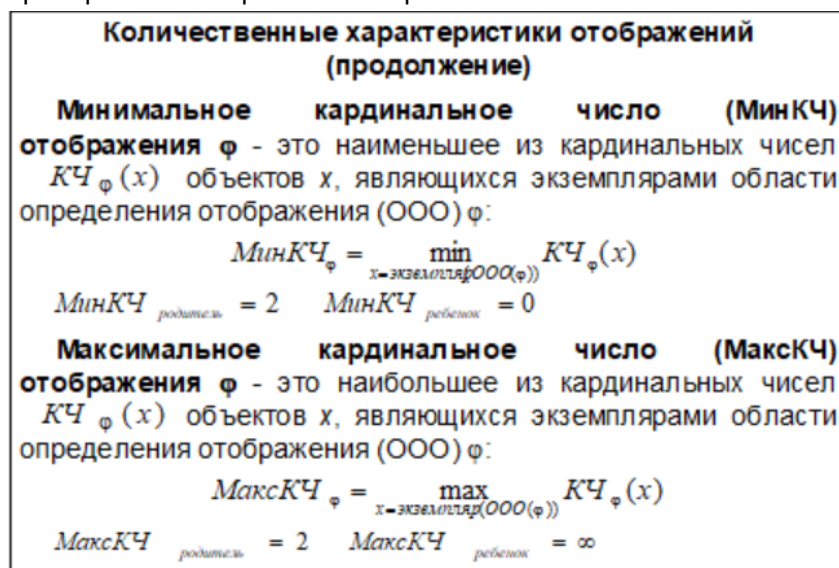
Область значений отображения (ОЗО) (англ. codomain) – это совокупность объектов которые в силу смысла закона могли быть в прошлом или являются в настоящем или смогут стать в будущем -образами некоторых объектов.

Например если отображению ДОЧЬ придать юридический смысл то ООО : ДОЧЬ – совокупность всех людей а ОЗО ДОЧЬ – совокупность всех женщин (ДОЧЬ: ЧЕЛОВЕК -> ЖЕНЩИНА).

51. Дайте определения КЧ МинКЧ и МаксКЧ. К каким элементам относятся эти характеристики?



Кардинальное число (КЧ)-мощность области образов объекта при отображении. Носит экстенциональный характер поскольку касается отдельных объектов являющихся прообразами конкретного отображения.



52. Перечислите и укажите характеристики типов отображений.

**Определение 2.3.7.** Отображение  $\varphi$  называется:

- **неограниченным**, если  $\text{МинКЧ}_{\varphi} = 0$ ,  $\text{МаксКЧ}_{\varphi} = \infty$ ;
- **полностью определенным**, если  $\text{МинКЧ}_{\varphi} \neq 0$ ;
- **функциональным**, если  $\text{МаксКЧ}_{\varphi} = 1$ , в том числе:
  - **частичным функциональным**, если  $\text{МинКЧ}_{\varphi} = 0$ ,  $\text{МаксКЧ}_{\varphi} = 1$ ;
  - **полным функциональным**, если  $\text{МинКЧ}_{\varphi} = 1$ ,  $\text{МаксКЧ}_{\varphi} = 1$ .

53. Поясните фразу бинарное отношение множеств определяет два отображения между ними .

Есть определение отображения в моделировании данных:  
«Бинарное отношение  $R$  множеств  $S_1$  и  $S_2$  определяет два отображения  $R: S_1 \rightarrow S_2$  и  $R^{-1}: S_2 \rightarrow S_1$ , каждое из которых является обратным по отношению к другому». Другими словами, отображение  $R$  эквивалентно инверсии отображения самого себя. При таком раскладе объекты меняются ролями (образы становятся прообразами, и наоборот). Если один из атрибутов отношения использовать как  $POOO$ , а второй – как  $POZO$ , каждый кортеж будет определять один экземпляр одного отображения. Если поменять атрибуты местами, те же кортежи определят экземпляры обратного отображения.

Бинарное отношение определяющее пару функциональных отображений есть отношение один к одному (1:1) бинарное отношение с одним функциональным а другим нефункциональным отображениями есть отношение один ко многим (1:M) или многие к одному (M:1) и бинарное отношение определяющее пару нефункциональных отображений есть отношение многие ко многим (M:N).

54. Какими символическими нотациями можно описать бинарные отношения? Сравните их информативность на примере различных типов бинарных отношений.

**Первая нотация обеспечивает исчерпывающую информацию(характеризует МаксКЧ и МинКЧ)** как о бинарном отношении так и об обоих определяемых им отображениях –  $R(S_1(0 \infty) : S_2(0 1))$ .  $R(S_1(0 \infty) : S_2(0 1))$  – бинарное отношение  $R$  определено на множествах  $S_1$  и  $S_2$ .

$S_1(0 \infty)$  означает что МинКЧ и МаксКЧ отображения  $S_2 \rightarrow S_1$  равны 0 и  $\infty$  соответственно – любой элемент  $S_2$  может быть связан минимум с 0 и максимум с  $\infty$  элементов  $S_1$ .

$S_2(0 1)$  означает что МинКЧ и МаксКЧ отображения  $S_1 \rightarrow S_2$  равны 0 и 1 соответственно – любой элемент  $S_1$  может быть связан минимум с 0 и максимум с 1 элементом  $S_2$ .

**Определение 2.3.8.** Говорят что бинарное отношение определяющее пару функциональных отображений

Вторая нотация: характеризует только МаксКЧ

1:1 (“один-к-одному”) – оба отображения функциональны

1:M (“один-ко-многим”) – одно отображение функционально второе – нет

M:N (“многие-ко-многим”) – оба отображения не функциональны

55. Укажите отличие сложного отображения от простого отображения. Приведите примеры сложных отображений укажите их характеристики и тип.

Простые отображения - каждый экземпляр имеет один объект-прообраз и один объект-образ.

Сложные отображения - каждый экземпляр имеет более одного объекта-прообраза и/или объекта-образа.

Пример: Про ближайших родственников. Единственный тип объектов – ЧЕЛОВЕК и три отношения. Два из этих отношений бинарны (простое) (БРАК и РОДИТЕЛЬ-РЕБЕНОК) а одно тернарно (сложное) (ОТЕЦ-МАТЬ-РЕБЕНОК).



56. Дайте три определения возможного ключа.

Возможный ключ (потенциальный ключ ключ-кандидат) -

- ❖ если атрибут (группа атрибутов) отношения функционально определяет все другие атрибуты этого отношения (функционально определяет - т.е. отображения между возможным ключом и любым атрибутом отношения функциональны).
- ❖ если атрибут (группа атрибутов) отношения функционально определяет отношение этот атрибут отношения. Т.е. отображение между возможным ключом и самим отношением функционально. Другими словами возможный ключ уникально идентифицирует кортежи отношения.
- ❖ Если атрибут (группа атрибутов) отношения не имеет значений-дубликатов в кортежах отношения

57. Что означает определенность значения атрибута ?

В каждом кортеже отношения в атрибуте должно лежать какое-то значение

Обязательный атрибут - в каждом кортеже отношений в этом атрибуте должно лежать какое-то значение (NOT NULL)



Между отношением и некоторым его атрибутом отображение полностью определено (а точнее оно – полное функциональное).

58. Какие особенности каких отображений скрываются в определении возможного ключа и обязательности значений атрибута?

Если между первичным ключом отношения и некоторым другим атрибутом этого же отношения отображения полностью определено (полное функциональное) то этот атрибут не может иметь неопределенных значения в кортежах отношений

Т.е. между первичным ключом и отношением существует взаимнооднозначное соответствие (каждому значению первичного ключа соответствует один и только один кортеж и наоборот)

Первое определение: Если атрибут отношения функционально определяет все другие атрибуты этого отношения, этот атрибут называется возможным ключом (потенциальным ключом).

Слова «функционально определяет» означают, что отображения между возможным ключом и любым атрибутом отношения функциональны.

Второе определение: если атрибут отношения функционально определяет отношение, этот атрибут называется возможным ключом. Т.е. отображение между возможным ключом и самим отношением функционально.

Третье определение: Если атрибут отношения не имеет значений-дубликатов в кортежах отношения, этот атрибут называется возможным ключом. **В отличие от двух предыдущих определений мы анализируем не общий закон ПрО, а конкретный прецедент, представленный экстенсионалом отношения.**

## Глава 2.4

59. Дайте определение расширенного состояния БД.

Перевод БД из состояния DBS  $i$  в состояние DBS  $i+1$  (или в неопределенное состояние) причем изменение состояния происходит или при изменении реализации данных или при изменении управляющих элементов. Сообщения об ошибках также относятся к управляющим элементам входящим в расширенное состояние БД.

60. Что такое операция над данными?

Процесс перевода БД из одного состояния в другое или в неопределенное состояние причем изменение состояния происходит или при изменении реализации данных или при изменении управляющих элементов.



Каждое состояние БД должно соответствовать ее схеме т.е. выполнение операций не должно приводить к нарушению свойств данных отраженных в схеме.

Важный момент определения операций в моделировании данных - обеспечение достаточной простоты их усвоения пользователем.

61. Какие два компонента можно выделить в любой операции над данными?

**Действие** определяет характер операции.

**Селекция** – критерий отбора данных над которыми должно быть произведено действие.

Действия:

- Установка индикаторов текущих элементов (SET CURRENCY)
- Включение или добавление в БД новых элементов (INSERT)
- Обновление или модификация существующих элементов в БД (UPDATE)
- Удаление существующих элементов из БД (DELETE)
- Выборка информации из БД (SELECT)

Способы селекции данных:

- Селекция по указателям текущих (GET NEXT СТУДЕНТ)
- Селекция по данным (GET СТУДЕНТ WHERE Пол = 'ж')
- Селекция по связям (GET СТУДЕНТ WHERE СТУДЕНТ. Фамилия = АВТОМОБИЛЬ. Фамилия владельца)
- Комбинированный способ (GET NEXT СТУДЕНТ WHERE Пол = 'ж' AND СТУДЕНТ. Фамилия = АВТОМОБИЛЬ. Фамилия владельца)

62. Какие действия над данными предусмотрены в технологии БД?

- 1) **установка** текущих – типичные ключевые слова – SET CURRENCY FIND;
- 2) **включение** (добавление новых элементов данных в БД) – INSERT ADD;
- 3) **обновление** (модификация существующих элементов данных в БД) – UPDATE MODIFY;
- 4) **удаление** (исключение элементов данных из БД) – DELETE REMOVE;
- 5) **выборка** (получение данных из БД) – SELECT GET.

63. Какими способами можно селектировать данные для выполнения тех или иных действий с ними?

Селекция может осуществляться посредством:

- 1) логической позиции в БД – селекция по текущей;
- 2) значений данных – селекция по данным;
- 3) связей между данными – селекция по связям.

64. Какие два класса операций и языков манипулирования данными выделяются в технологии БД? Каковы их особенности?

**Навигационные** - результат всегда представлен **единственным элементом** (в зависимости от модели это могут быть значение атрибута кортеж сущность связь и т.д.) полученным **при прохождении по логическому пути** (т.е. при навигации) в структуре БД.

**Спецификационные** - операции могут специфицировать как бы **новую под схему** определяемую на схеме БД которой в общем случае **соответствует множество элементов** существующих в БД.

Язык все операции которого являются навигационными называется навигационным языком. Язык все операции которого являются спецификационными называется спецификационным языком (SQL). В нем отсутствуют все характерные для навигационных языков черты.

**Особенности** запросов на **навигационном языке**:

- они всегда представляют собой программу;
- для их формулировки требуется язык программирования с условными конструкциями и конструкциями циклов;
- в них обязательно используется селекция по текущей;
- для них необходимо действие установки текущих.

65. Что такое процедура БД?

**Процедура БД** - последовательность операций выполняемых при определенных условиях.

**Процедуры БД** специфицируются в схеме и в общем случае состоят **из трех** разделов: **условия действия и уведомления**.

В разделе условия специфицируется проверка достоверности предпосылок действия.

Назначение раздела действий следует из названия.

В разделе уведомления указывается что именно должна передавать процедура при возврате управления и/или задается условие возврата.

66. Какие виды процедур БД вам известны?

1. **функции агрегирования** – вычисление значений которые непосредственно не хранятся в БД например вычисление сумм подсчет числа экземпляров определение минимума и максимума. Во многих моделях функции агрегирования являются конструкциями языка данных.
2. **Виртуальный атрибут**– вычисление значения атрибута например вычисление возраста человека по заданной текущей дате и дате рождения что выполняется по запросу о возрасте
3. **Триггеры целостности**– контроль целостности БД. Верификация ограничений рассмотренных в предыдущем параграфе может осуществляться автоматически (декларативный способ определения ОЦ) или с помощью процедур БД. Служат для проверки и/или поддержки ОЦ
4. **Триггеры безопасности** – обеспечение контроля над доступом к данным и проверки полномочий и аутентичности пользователей. Например процедура БД может получать в качестве исходных данных идентификационные параметры пользователя время дату вид действия и вырабатывать сообщения для пользователя журнала или лица ответственного за соблюдение правил доступа.

5. **Операторы доступа** - расширения языка данных операциями первоначально в нем не предусмотренными. Например операция сортировки может быть определена как процедура БД предшествующая выдаче результата спецификационной операции. Другой пример – процедура удаления кортежей- дубликатов позволяющая модифицировать спецификационную операцию в результате которой образуется отношение.
6. **Триггеры запускаемые включением удалением или обновлением (прикладные)** - процедуры БД которые инициируются операциями модификации одних данных и выполняют косвенные модификации других данных. Если тип связей между отношениями является полным (хотя бы одно из определяемых им отображений полностью определено) то удаление кортежа может привести к запуску процедуры удаляющей все кортежи связанные с удаляемым.
7. **Процедуры администратора БД (Триггеры БД запускаемые другими событиями в БД (системные триггеры)** - сбор статистики или для реализации каких-либо других функций администрирования БД. Они не изменяют данные но могут модифицировать состояние БД.