

Параграф №1

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

Технология баз данных – это технология, регламентирующая процесс создания и эксплуатации информационных систем, в которых для долговременного хранения информации используется база данных, а управляет ею 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.

Здесь следует **сделать замечание о нетрадиционном использовании термина модель в технологии БД.** В технологии БД произошла замена смысла термина модель, в ней это слово обозначает теорию моделирования, а результат моделирования есть не что иное, как БД.

2. Каковы предпосылки использования технологии БД?

- Вы **нуждаетесь в долговременном хранении информации** о некоем фрагменте реального или идеального мира (предметной области - ПрО).
- Система создается для **удовлетворения информационных потребностей** нескольких людей, в том числе тех, кто в силу разных причин не в состоянии воспринимать этот мир с помощью своих органов чувств.
- Вам необходима не бесформенная куча, извлечение информации из которой будет отнимать много времени и усилий, а **хорошо структурированное хранилище**, в котором

каждый бит информации лежит на своей полке и непротиворечиво связан с остальными

- **Моделируемый в системе мир динамичен**, и поэтому специальные пользователи должны вносить изменения в его информационное описание.
- **Большая часть пользователей будет удовлетворять свои потребности в информации** об этом мире, обращаясь к информационной системе, в которой должны обеспечиваться специальные механизмы, способствующие удобному извлечению и преобразованию информации
- **Производная информация должна получаться из первичной**, введенной непосредственно людьми, с помощью простых универсальных преобразований, и этот процесс не должен сопровождаться разработкой специальных алгоритмов.
- **Информация вводится, хранится и предъявляется преимущественно в алфавитно-цифровой форме.**

3. В чем особенности термина модель в технологии БД?

Модель данных есть теория, или инструмент моделирования, в то время как модель БД (схема БД) есть результат моделирования.

Комплекс взглядов, представлений, идей, дающий целостное представление о закономерностях и существующих связях определенной области действительности - это теория моделирования в большинстве отраслей знаний(и БСЭ), но в БД так определяется понятие "модель".

4. Каково основное назначение моделей данных?

Основное назначение МД – обеспечение процесса интерпретации (осмысления) данных и превращения их в информацию, определяют степень соответствия СУБД целям их создания.

5. Перечислите составные части любой модели данных.

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

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?

Так, например, **правила структуризации** для реляционной модели 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.

Наряду с этим реляционную схему БД составляют условия, которым должны удовлетворять данные, заполняющие тела таблиц. Иногда их называют бизнес-правилами или, чаще, **ограничениями целостности**. Последний термин связан с поддержанием БД в целостном, непротиворечивом состоянии, максимально соответствующем состоянию моделируемой предметной области. Окончательно эту проблему не решат никакие средства автоматизации. Только человек в состоянии определить, какие факты об объектах предметной области являются истинными. Но частично помочь ему в этом и призваны ограничения целостности, декларированные в схеме БД. Так, ограничение целостности, говорящее о том, что значениями атрибута Рост отношения ЧЕЛОВЕК могут быть целые числа в интервале от 50 до 300, предостережет от многих ошибок ввода. Задачей человека все же остается указание в каждом случае точного, известного ему значения из этого интервала.

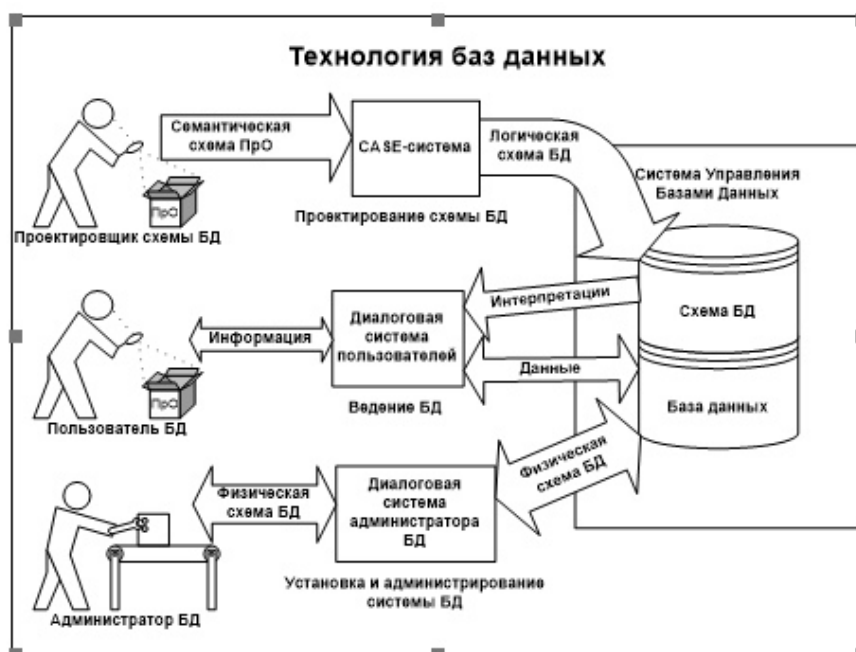
До тех пор, пока не определена как минимум структура данных, система не позволит вводить эти данные. Это происходит потому, что при создании,

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.

(Операции над данными)

Операции необходимы, чтобы первоначально наполнить схему БД (основные понятия предметной области и их взаимосвязи) **непротиворечивыми данными** и поддерживать их в дальнейшем в адекватном предметной области состоянии, а также обеспечивать получение необходимой информации.

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



На рисунке отражены самые **важные процессы технологии БД**. Прежде всего, это самый первый, самый важный и обязательный процесс – **проектирование схемы БД**. Как правило, это дело поручают профессионалам в этой технологии. От того, насколько качественной будет схема БД, существенно зависят простота и легкость осуществления всех остальных процессов и успех всей системы в целом. Также **введение БД, установка и администрирование системы БД**.

Стилизованные человечки обозначают роли, которые играют люди в системе БД.

Дуги с одной или двумя стрелками на концах олицетворяют как раз информационные потоки между людьми и программами, а также между программами. Стрелка указывает направление, в котором осуществляется передача данных.

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. Потом проектировщик разрабатывает схему БД и диалоговую систему пользователей с помощью CASE-систем.

Администратор системы осуществляет ее установку на оборудовании организации-заказчика. В дальнейшем он с помощью специальных инструментов, предоставляемых разработчиками СУБД, администрирует систему на уровне физической схемы БД – схемы хранения данных и их интерпретаций на диске. Основной обязанностью администратора БД является обеспечение эффективной бесперебойной работы БД, не взирая на возможный выход из строя аппаратуры и программ.

Пользователи выступают в **двух ролях – писателей и читателей** (аналогично ролям, сопровождаемым бумажное общение на естественном языке). **Первые** обязаны отражать в БД изменения ПрО, причем максимально оперативно и точно. **Вторые** удовлетворяют свои информационные потребности, **обращаясь к системе БД**.

8. Для чего предназначены CASE-системы?

CASE – Computer Aided Software/System Engineering – программная система для проектирования и реализации других программных систем. При наличии специализированного программного обеспечения в лице CASE-системы можно, во-первых, просто и естественно для проектировщика осуществить ввод семантической схемы, а, во-вторых, автоматически перевести ее на язык СУБД. В идеале этот перевод не должен сопровождаться потерей информации, и каждый элемент семантической схемы должен найти свое отражение в логической модели СУБД (в крайнем случае – в процедурной форме). Чем выразительней модель, тем больше особенностей ПрО в ней можно отразить. В идеале семантическая схема должна вобрать в себя все закономерности информационного описания ПрО. При этом не будет необходимости прибегать к помощи других моделей и языков, в том числе, естественного.

9. Перечислите и кратко охарактеризуйте архитектуры систем БД.

Однопользовательская архитектура (автономная архитектура) – вся система БД управляется одним компьютером, на котором расположены БД, СУБД и диалоговая система пользователя.

Двухуровневая архитектура Клиент-Сервер - состоит из множества компьютеров, объединенных в сеть. Клиенты занимаются обработкой прикладных программ, которые ведут диалоги с пользователями. Серверы занимаются обработкой БД.

Трёхуровневая архитектура Клиент-Сервер - включает в себя сервер БД, веб-сервер и браузер(клиент).

Сервер БД – несет на себе СУБД, которая выполняет SQL-операторы и осуществляет управление БД.

Веб-сервер :

- 1) принимает запросы и генерирует отклики в формате HTTP
- 2) на Веб-сервере располагается и выполняется код, реализующий

бизнес-логику приложения.

3) генерирует SQL-запросы к БД и принимает отклики на них от сервера БД.

Браузер в приложениях БД также выполняет три функции.

1) браузер является HTTP-клиентом: он генерирует запросы на страницы и на различные действия

2) содержит сценарный интерпретатор, позволяющий запускать сценарии на клиентской машине.

3) материализует представления, преобразуя HTML (или другой язык разметки) в содержимое окна на экране.

10. В чем основные различия OLTP- и OLAP-систем БД?

OLTP (система оперативной обработки транзакций) – система управления больших потоков транзакций, вносящих незначительные изменения в БД.

OLAP (инструменты оперативной аналитической обработки) – инструменты для анализа огромного объема данных.

OLTP используется для управления больших потоков транзакций, вносящих незначительные изменения в БД, а

OLAP используется для анализа огромного объема данных, представленных чаще всего через многомерные кубы.

Решение их задач на одной СУБД и БД является не эффективной.

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

ПрО представляет собой некоторый интересующий заказчика системы БД фрагмент реального или идеального мира. Первоначальные представления о ПрО возникают в голове у проектировщика БД на основании его личного знакомства с нею, а также на основании сведений, полученных у экспертов по ПрО или из других источников. Эти исключительно неформализованные представления составляют зрительные, слуховые, обонятельные, осязательные образы.

Знакомясь с различными моделями данных, проектировщик в каждом случае вооружается набором формальных понятий модели. Его задачей является перевод своего неформального восприятия ПрО на язык конкретной модели данных, используя ее формальный понятийный набор.

Известные формы – семантические модели данных, СУБД-ориентированные схемы, логические и физические схемы БД.

Семантическая схема ПрО – ее первое формальное представление. Не существует программных систем для создания БД (СУБД), поддерживающих непосредственно семантические модели, а БД – это конечная цель в технологии БД, мы не можем этим ограничиться.

Если в качестве СУБД вы выбрали одну из реляционных СУБД, вам предстоит построить реляционную схему своей ПрО, которая является одной из разновидностей **СУБД-ориентированных схем**.

Здесь следует заметить, что каждая СУБД имеет не одну, а две модели данных – **логическую и физическую**. Первая из них обращена в сторону человека (проектировщика или пользователя БД) и предлагает языковые и диалоговые инструменты общения именно для него. Собственно, вторую модель СУБД (физическую) редко кто, кроме разработчиков этой СУБД, и знает, поскольку обращена она к средствам хранения данных в оперативной и внешней памяти.

Параграф № 2.1

12. Объясните, почему человека интересуют не данные, а информация.

Данные – это факты реального мира и идеи, представленные в формализованном виде, позволяющем эксплуатировать их при помощи технических средств.

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

Сами по себе данные ничего для человека не значат, но на их основе человек может получить информацию, которая ему интересна.

Пример данных: 130

Пример информации: 130 человек сгорело заживо сегодня в Красноярске

13. Что кроме данных необходимо для получения информации?

Интерпретация и знание предметной области.

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

Системы БД относятся к классу информационных систем, интерпретации которых хранятся в программе и на диске в виде специальным образом организованных данных.

В системах БД интерпретации не фиксируются в программах, а хранятся на диске, как и сами данные. Именно их мы задаем на этапе проектирования схемы БД. 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.

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

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).

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.

Синтактика определяет правила построения БД. Семантика изучает семантические отношения между знаками и объектами реального мира (они показаны точечными дугами). И, наконец, прагматика осуществляется в голове у читателя.

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 – изучает внутренние свойства систем знаков безотносительно к интерпретации (синтаксис – правила построения знаков и знакосочетаний в рамках знаковой системы).

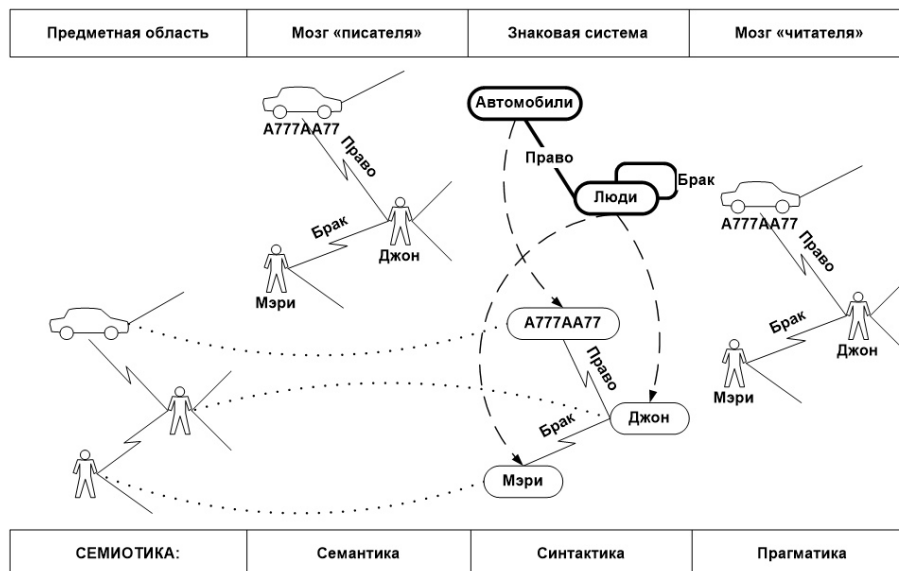
Семантика – рассматривает отношение знаков к обозначаемому (содержание знаков) или, что то же, соотношения между знаками и их интерпретациями, независимо от того, кто служит адресатом (интерпретатором).

Прагматика – изучает 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.

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.

Процесс получения информации из данных



В ПрО независимо от нашего сознания существуют некоторые объекты, обладающие характеристиками и вступающие в связи с другими объектами. На схеме в левой части показаны два объекта типа ЧЕЛОВЕК. Это супруги Джон и Мэри. Изображен также автомобиль, владельцем которого является Джон.

Соответствующие связи показаны в виде зигзагообразных дуг. Человек, чьей обязанностью является наблюдение за ПрО и отражение всех изменений в БД (назовем его писателем), так или иначе, узнал о существовании этих объектов и связей. При этом он произвел у себя в голове первичную формализацию ситуации, в частности, определил знаки для идентификации объектов и связей (на схеме зрительные образы объектов сопровождают эти знаки).

После этого писатель обратился к знаковой системе – системе БД. В ее схеме он обнаружил подходящую подсхему (она представлена графом типов в верхней части столбца Знаковая система) и создал в БД новые объекты типов ЛЮДИ и АВТОМОБИЛИ, а также связи типов БРАК и ПРАВО. Получился подграф знаков, изображенный в нижней части того же столбца. Пунктирные направленные дуги связывают знаки с их типами. Именно эти связи указывают, как интерпретировать соответствующие данные. **На этом работа писателя закончена.**

Теперь по мере необходимости любой читатель (человек, желающий получить от системы БД информацию о ПрО) **может воссоздать представления писателя о ПрО**, не прибегая к ее непосредственному исследованию. Вместо этого он обращается к системе.

В нижней части схемы показаны разделы семиотики, участвующие в описанных процессах. Синтактика определяет правила построения БД. Семантика изучает семантические отношения между знаками и объектами реального мира (они показаны точечными дугами). И, наконец, прагматика осуществляется в голове у читателя

18. В чем заключается основное назначение модели данных?

Модель данных (МД) – это интеллектуальное средство, позволяющее реализовать интерпретацию данных и таким образом

способствующее получению информации. с Громкий эпитет интеллектуальное применен в данном случае неспроста – способность к восприятию, хранению и передаче информации – одно из важнейших свойств интеллекта.

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

Как вы уже поняли, основными объектами исследований в моделировании данных являются данные и их интерпретации. Эта область человеческих знаний, как и многие другие, имеет свой элементарный объект – **атомарную единицу 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. Если иметь в виду, что к признакам объектов относятся и свойства, и характеристики, и отношения, понятно, что сколь угодно сложные Про можно представить множеством таких четверок.

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

Атомарная единица информации (АЕИ). Она определяется четверкой – <Идентификатор объекта, Наименование признака, Значение признака, [Время]>. Последний элемент четверки [Время] может быть опущен, все остальные элементы являются обязательными. **Если опустить хотя бы один из них, информация не образуется.**

Нам необходим **идентификатор**, чтобы различать объекты, **название признака**, чтобы различать признаки и **соотносить название признака с его значением для дальнейшего взаимодействия.**

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>.

21. Синтезируйте категориальную модель

Модель данных определяется двумя множествами. Первое множество - это множество правил порождения схем, а другое – множество операций над данными. В первом множестве выделяется два подмножества: правила порождения структур данных и правила порождения ограничений целостности.

Она предполагает разбиение всех объектов ПрО по категориям. Для каждой категории определяется набор признаков, значения которых характеризуют объекты данной категории. На примере этой простой модели познакомимся с основополагающими понятиями моделирования данных.

22. Дайте структурное определение модели данных.

Модель данных (МД) определяется двумя множествами G и O. G – множество правил порождения схем, O – множество операций над данными. В свою очередь во множестве G выделяются два подмножества – Gs (правила порождения структур данных) и Gc (правила порождения ограничений целостности).

23. Проиллюстрируйте компоненты модели данных на примере категориальной модели.

1-я модель данных: правила порождения допустимых структур данных.

Для категориальной модели правила множества порождения структур данных выглядят следующим образом:

- 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.

2-я модель данных: правила порождения ограничений целостности данных.

Допустимые значения признаков можно ограничить:

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

3-я модель данных: набор операций над данными.

Множество операций над данными может выглядеть для нашей модели так

- 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)**.

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

СУБД (Система управления базами данных) — комплекс специализированных программ, которые позволяют создать базу данных (БД) и манипулировать данными (вставлять, обновлять, удалять и выбирать). Система обеспечивает безопасность, надёжность хранения и целостность данных, а также предоставляет средства для администрирования БД.

Классы операций, которые необходимо реализовать в любой СУБД:

- 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 определённому условию.

(Из учебника, если понадобятся языковые и диалоговые формы:

*Управление БД на ЭВМ осуществляется специализированными программными средствами – **системами управления базами данных (СУБД)**, 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.)*

Параграф №2.2

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

Элементарные единицы данных и **интерпретаций** определяются понятиями **знак** и **тип**.

Знак – это данное, которое обозначает объект, свойства объектов или их связь.

Тип – это именованный класс подобных знаков. Имя типа определяет интерпретацию знака, который принадлежит типу. В БД они используются для структуризации, которая нужна для построения системы типов с использованием понятий, предлагаемых правилами структуризации данных этой модели.

Каждая связь конкретного типа задается в БД в виде кортежа, элементами которого являются знаки соответствующих объектов.

27. В каких частях системы БД представлены знаки и типы?(В каких частях системы БД представлены элементарные единицы данных и интерпретаций?)

Обобщение - это абстракция, которая позволяет соотнести множество знаков или типов с одним общим типом в соответствии с отношением есть некоторый (англ. is a). В зависимости от вида объектов обобщения рассматривают два частных случая обобщения – классификацию и генерализацию. Классификация – это обобщение знаков до типа (отношение экземпляр-класс), а генерализация – это обобщение типов до типа (отношение подкласс-суперкласс).

Обобщенный тип обладает всеми признаками, общими для базовых знаков или типов. Иначе говоря, все признаки обобщенного типа могут быть унаследованы базовыми типами.

Агрегация - это абстракция, посредством которой объект конструируется из других, базовых объектов в соответствии с отношением есть часть (англ. part of). Агрегация возможна как на уровне знаков (сложный знак собирается из более простых знаков), так и на уровне типов (сложный тип собирается из более простых типов). Более того, агрегация на уровне типов предполагает множество агрегаций на уровне знаков.

Добавлено примечание ([1]): Вопрос редактируется

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

Абстракция - метод исследования, основанный на том, что при изучении какого-то явления не учитываются его несущественные стороны и признаки; это позволяет упрощать картину изучаемого явления и рассматривать его в чистом виде.

Абстракция предполагает, что несущественные частные детали должны быть опущены, а внимание должно быть сосредоточено на важных общих особенностях.

Абстракция используется для придания множествам объектов некоторой семантики.

В моделировании данных абстракция используется для образования категорий данных. Кроме того, абстракция позволяет на основе одних категорий построить другие, более общие.

В технологии БД выделяют два вида абстракции: обобщение и агрегацию.

- **Обобщение** – это абстракция, которая позволяет соотнести множество знаков или типов с одним общим типом с отношением *есть некоторый*. В зависимости от вида объектов обобщения рассматривают два частных случая обобщения – классификацию и генерализацию. Можно использовать для строительства многоуровневых иерархий.
 - Классификация – это обобщение знаков до типа (отношение экземпляр-класс).
 - Генерализация – это обобщение типов до типа (отношение подкласс-суперкласс). Все признаки обобщенного типа могут быть унаследованы базовыми типами.
- **Агрегация** – это абстракция, при которой объект конструируется из других, базовых объектов с отношением *есть часть*. Агрегация возможна на уровне знаков (сложный знак собирается из более простых знаков) и на уровне типов (сложный тип собирается из более простых типов). Более того, агрегация на уровне типов предполагает множество агрегаций на уровне знаков. Так же как и обобщение, агрегацию можно использовать для строительства многоуровневых иерархий.

Агрегация и обобщение применяются взаимодополняющим образом и выражают структурные и классификационные аспекты типизации. Структуру типа можно представить как агрегат базовых типов, а сам агрегат будет объектом обобщения. Классификация типов может быть выражена иерархией обобщения, а структура – иерархией агрегации.

29. Поясните, как абстракции используются для интерпретации данных (знаков).

Абстракции используются в двух видах.

Обобщение - соотношение множества знаков или типов с одним общим типом в соответствии с отношением есть некоторый (англ. is a). В свою очередь обобщение делится на: классификацию – это обобщение знаков до типа (отношение экземпляр-класс), и генерализацию – это обобщение типов до типа (отношение подкласс-суперкласс).

Агрегация – это абстракция, посредством которой объект конструируется из других, базовых объектов в соответствии с отношением есть часть (англ. part of). Агрегация возможна как на уровне знаков (сложный знак собирается из более простых знаков), так и на уровне типов (сложный тип собирается из более простых типов). Более того, агрегация на уровне типов предполагает множество агрегаций на уровне знаков.

30. Перечислите традиционно используемые в моделях данных формы данных.

Множество - это собрание правильно идентифицированных объектов, удовлетворяющих условию принадлежности.

Множество – это комплекс, все элементы которого находятся в одной позиции.

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 .

В отличие от множества элементы комплекса распределены по позициям, определяемым целыми положительными числами. Причем в каждой позиции может быть любое число различных элементов.

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

Домен - множества, элементы которых синтаксически однородны, например, множество целых чисел от 60 до 70, строки букв длиной до 10 символов и т.д.

Атрибуты - именованные домены, представляющие семантически значимые объекты. Атрибуты являются интерпретацией объектов реального мира и их характеристик.

31. Укажите свойства множеств.

Чтобы определить множество, нужно указать характерное свойство элементов, то есть свойство, которым будут обладать все элементы этого множества.

СВОЙСТВА МНОЖЕСТВА:

- Количество элементов во множестве не фиксировано (то есть количество может быть сколько угодно, даже бесконечным)
- Элементы множества не подчиняются порядку расположения
- Повторяющиеся элементы во множестве не имеют смысла.

Так же множество-такая форма данных, которая определяется своим интерфейсом (способна добавлять, исключать элементы из множества, определять принадлежность к множеству)

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

Главные свойства множества:

1. Интенционал множества- определяет свойства, общие для всех элементов множества (определяет принадлежность элементов к множеству, условие принадлежности в математике)

2. Экстенционал множества- определяет актуальную реализацию множества путем явного указания его элементов. (например: элемент-человек, множество-класс людей)

Почему в моделировании данных уместно говорить о нескольких реализациях множества?

При моделировании в БД динамических 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 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.

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

Комплекс - это собрание правильно идентифицированных объектов, удовлетворяющих условию i -принадлежности.

Множество - это **комплекс**, все элементы которого находятся в одной, предположим первой позиции

Кортеж – это **комплекс**, который имеет по одному элементу в каждой позиции от 1 до n

Комплекс не является структурным понятием ни одной из моделей данных, в отличие от его специализаций.

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

В отличие от элементов домена, элементы атрибута обладают не только синтаксической однородностью, но и семантической (в их условии принадлежности фигурируют особенности конкретной предметной области). Например, элементами домена могут быть даты, а элементами атрибута даты рождения взрослого населения России.

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 . Бинарное отношение задается множеством упорядоченных пар $\langle s_1, s_2 \rangle$, которые удовлетворяют некоторому критерию.

Логическое отношение - 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 данных - Пусть задано множество из n типов или доменов $T_i (i=1, \dots, n)$, причем все они необязательно должны быть различными. Тогда r будет **отношением**,

определенным на этих типах, если оно состоит из двух частей: **заголовка и тела** (заголовок еще иногда называют схемой или интенсиналом отношения, а тело – экстенсиналом отношения), где:

- заголовок – это множество из n атрибутов вида $A_i:T_i$; здесь A_i – имена атрибутов отношения r , а T_i – соответствующие имена типов;

- тело – это множество из m кортежей t ; здесь t является множеством компонентов вида A_i , в которых v_i – значение типа T_i , т.е. значение атрибута A_i в кортеже t .

Попытаемся сопоставить все приведенные ранее определения. Человека, прежде всего, интересует истинность или ложность того или иного факта, поэтому ему ближе и понятней определение логики. В информационных системах (и в базах данных, в частности) принято хранить только истинные факты и предполагать выполнение гипотезы о замкнутости мира: если кортеж не содержится в теле отношения, то соответствующее ему утверждение ложно. Понятно, что в таком случае **математическое отношение и логическое отношение определения эквивалентны. Отношение в моделировании данных отличается от математического отношения отсутствием порядка в элементах кортежей**, и, таким образом, одно отношение в смысле определения **Отношение в моделировании данных соответствует сразу нескольким отношениям, в смысле определения математического отношения**. Именно поэтому отношение в моделировании данных используется в базах данных.

37. В каком виде задаются в БД интерпретации данных?

Интерпретации данных задаются в виде таблиц.

38. Каким формам данных приписываются интерпретации?

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)?

Интерпретация как тип объектов

- Множество используется для представления имен типов
- Кортёж представляет множество значений, отношений, характеризующий данный объект

Интерпретация как тип связей между типами объектов

- Кортёж, элементы которого являются знаки соответствующих объектов, задает связь конкретного типа

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

Формы данных в виде графа



Пример домена



Формы данных в виде таблицы



Добавлено примечание ([2]): @all требуется проверка на полноту ответа

41. Что ассоциируется с понятиями “интенционал БД” и “экстенционал БД”?

Интенционал БД

- это дефиниционная БД (или дефиниционное множество). Он определяет свойства, общие для всех элементов всех реализаций БД или множества (в математике его называют условием принадлежности)
- представляет собой схему БД, включающую определения структур данных и ограничения целостности.

Экстенционал БД

- является репрезентативной БД (или репрезентативным множеством). Он определяет актуальную реализацию БД (или множества) путем явного указания его элементов
-

Параграф №2.3

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

ОЦ - логическое условие, предназначенное для реализации данного множества, атрибута, отношения или нескольких отношений либо в истину, либо в ложь.

ОЦ вводятся в модели данных в целях повышения ее семантичности и расширения возможностей поддержания целостности данных.

43. Дайте определения понятий, связанных с процессом верификации ОЦ

С процессом верификации декларативных ОЦ в СУБД связан ряд определений.

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

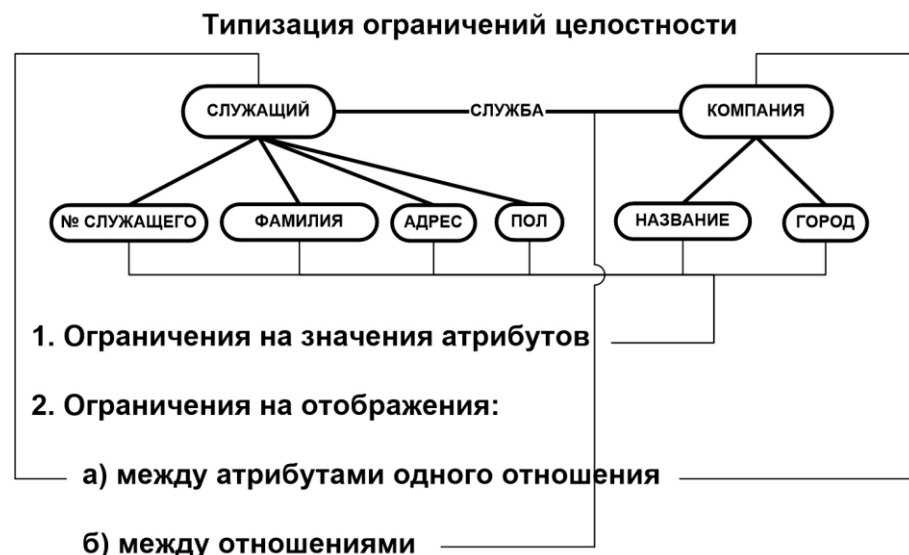
- 1) **правильно построено**, если оно соответствует синтаксическим правилам задания ОЦ;
- 2) **удовлетворено состоянием** БД DBS_k , если оно истинно для него;
- 3) **удовлетворяемо**, если существует некоторое состояние БД DBS_k , удовлетворяющее C_i ;
- 4) **недостоверно**, если никакое состояние БД не удовлетворяет C_i ;
- 5) **логическим следствием** C_1, \dots, C_n (т.е. избыточным ограничением), если C_i удовлетворяют все состояния БД, которые удовлетворяют C_1, \dots, C_n ;
- 6) **эквивалентно** C_j , если C_i и C_j – 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.

Эффективные алгоритмы проверки этих ОЦ реализованы в СУБД и не требуют дополнительного программирования. В случае недостаточности выразительных возможностей декларативных ОЦ, предусмотренных в модели данных, прибегают к процедурным средствам определения ОЦ. Многие СУБД предоставляют средства для программных проверок требуемых свойств данных на языке сервера (в виде триггеров). Как правило, они предусматривают достаточно мощный язык программирования, на котором можно алгоритмически выразить любые необходимые ОЦ. По степени 43 эффективности программы проверок данных, реализованные на сервере БД, хотя и уступают возможностям аппарата верификации декларативных ОЦ, все-таки превосходят программные проверки данных на стороне клиентского приложения. К последним надо прибегать в последнюю очередь

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. областью действия ОЦ первого типа являются отдельные атрибуты.
- 2) Ограничения на отображения:
 - а) между атрибутами одного отношения, Областью действия ОЦ в первом случае является одно отношение, как агрегат атрибутов.
 - б) между отношениями. ОЦ второго подтипа могут охватывать несколько отношений (чаще всего – два),

На схеме показаны области действия каждого из типов ОЦ.

Так областью действия ОЦ первого типа являются отдельные атрибуты. Встречаются случаи, когда они действуют на несколько атрибутов одного отношения.

Другой класс ограничений, задаваемых на отображения между атрибутами и отношениями, является вторым типом. Сюда попадают и ключи, и ограничения ссылочной целостности, и обязательность значения атрибута, и многое другое. Более того, через понятие отображения возможно выражение многих таких ОЦ, которые до сих пор выражались либо на естественном языке, либо на языке логики предикатов первого порядка, либо на языке программирования. (например, ограничить число служащих, которые работают в определенном отделе)

Выделяют два подтипа ОЦ 2-го типа. В первом случае рассматриваются отображения между атрибутами одного отношения, во втором случае – между отношениями. **Областью действия ОЦ в первом случае является одно отношение, как агрегат атрибутов. ОЦ второго подтипа могут охватывать несколько отношений (чаще всего – два), поскольку они отражают закономерности взаимосвязей между кортежами этих отношений.**

46. Перечислите виды ОЦ 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 есть закон соответствия, сопоставляющий каждому элементу множества A некоторый (единственный) элемент множества B . $\phi: A \rightarrow B$ означает, что задано отображение A в B , называемое ϕ .

Традиции, сложившиеся в математике, требуют обязательного наличия единственного образа для каждого аргумента, но существуют отступления от этого правила, в частности, при рассмотрении многозначных функций.

У Кофмана самый общий случай называется соответствием. При этом не накладывается никаких ограничений – у конкретного аргумента (прообраза) образа может не быть вовсе или образов может быть произвольное число. Специализацией соответствия является отображение, требующее наличия

хотя бы одного образа у каждого прообраза, специализацией отображения является функция, определение которой в точности соответствует первому определению отображения

48. Укажите основные отличия математического отображения от семантически значимого отображения

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

Традиции, сложившиеся в математике, требуют обязательного наличия единственного образа для каждого аргумента, но существуют отступления от этого правила, в частности, при рассмотрении многозначных функций.

Семантически значимое 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.

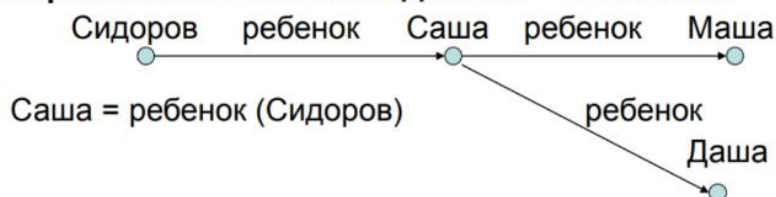
Законы, действующие в мире объектов, представляются в знаковой системе в виде отображений, прообразами и образами которых являются знаки-объекты, знаки-связи и знаки-значения. Термин семантически значимый предполагает значимость с точки зрения семантики, изучающей семантические отношения, которые образуются между объектами и знаками, представляющими эти объекты в знаковой системе. Благодаря этому эпитету мы дистанцируемся от абстрактных (в частности, математических) отображений и функций.

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

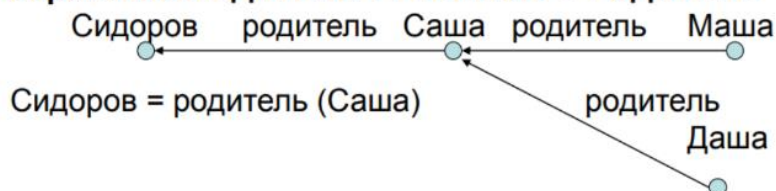
Отношение РОДИТЕЛЬ-РЕБЕНОК

Родитель	Ребенок
Сидоров	Саша
Саша	Маша
Саша	Даша

Отображение РЕБЕНОК: РОДИТЕЛЬ -> РЕБЕНОК



Отображение РОДИТЕЛЬ : РЕБЕНОК -> РОДИТЕЛЬ



Дополнение

Один из экземпляров отображения РЕБЕНОК показан в текстовой форме (Саша = ребенок (Сидоров)). Для этого использована традиционная для функции нотация – $\langle \text{образ} \rangle = \langle \text{имя отображения} \rangle (\langle \text{прообраз} \rangle)$.

Ниже представлены экземпляры обратного отображению РЕБЕНОК отображения РОДИТЕЛЬ. Они передают ту же самую информацию, только сменилась ориентация отображения, и объекты поменялись ролями (образы стали прообразами и наоборот).

*Это связано с тем, что отображение РЕБЕНОК эквивалентно инверсии отображения РОДИТЕЛЬ, и оба этих отображения полностью определяются бинарным отношением РОДИТЕЛЬ-РЕБЕНОК, представленным в табличной форме в верхней части слайда

Каждый факт соответствия φ одному объекту одного другого объекта представляет экземпляр отображения φ (или экземпляр φ -отображения); все экземпляры одного и того же отображения φ , соответствующие одному объекту-прообразу, будем называть отображением φ этого объекта (или φ -отображением объекта). Рассмотрим объект x и отображение φ . Будем называть образом (англ. image) объекта x при отображении φ (или кратко, φ -образом объекта x) объект y , поставленный в соответствие объекту x одним из экземпляров отображения φ (краткая символьная запись экземпляра – $y = \varphi(x)$). Объект x в таком случае будем рассматривать как прообраз (англ. preimage) объекта y при отображении φ (или кратко, φ -прообразом объекта y). Все отображение φ объекта x порождает совокупность образов, которую будем называть областью φ -образов объекта x (обозначение – $\{\varphi(x)\}$). В каждом конкретном случае эта совокупность: - может быть пустой (отображение φ неприменимо к объекту x) или - содержать один объект или - содержать много объектов

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

Совокупность всех объектов, имеющих образы при отображении φ называется **реальной областью определения отображения (РОО) φ** , а совокупность всех φ -образов этих объектов – **реальной областью значений отображения (РОЗО) φ** .

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

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

Область значений отображения φ (ОЗО) – это совокупность объектов, которые в силу смысла закона **могли быть** в прошлом или

являются в настоящем или смогут **стать** в будущем ф-образами некоторых объектов.

Например, если отображению ДОЧЬ придать юридический смысл, то ООД 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?

Количественные характеристики отображений j -отображение объекта x характеризует кардинальное число (англ. cardinal number) объекта x при отображении j (обозначение – $KЧ_j(x)$) – мощность его области образов при этом отображении. Значение кардинального числа элемента равно количеству инцидентных ему (выходящих из него) дуг экземпляров отображения. Для большинства отображений разброс значений кардинальных чисел его прообразов достаточно велик. Но встречаются и такие семантически значимые отображения, при которых кардинальные числа всех прообразов совпадают. (пример, у любого человека только два родителя)

Минимальное кардинальное число (МинКЧ) отображения j – это наименьшее из кардинальных чисел $KЧ_j(x)$ объектов x , являющихся экземплярами ООД j

Максимальное кардинальное число (МаксКЧ) отображения j – это наибольшее из кардинальных чисел $KЧ_j(x)$ объектов x , являющихся экземплярами ООД j

52. Перечислите и укажите характеристики введенных типов отображений.

Классификация отображений в моделировании данных



Неограниченное отображение - Отображение называется неограниченным, если любой прообраз может или вообще не иметь образа, или иметь произвольное количество образов.

Полностью определенное отображение - Пусть для отображения $S1 \rightarrow S2$ имеем пару кардинальных чисел $(1, \infty)$, т.е. предполагается, что каждый объект $S1$ отображен, по крайней мере, в один объект $S2$.

Функциональное отображение - если максимальное кардинальное число равно единице (в это случае отображение есть функция) Представлению функционального отображения в форме графа свойственно не более одной дуги, выходящей из вершины-прообраза.

Частично функциональное отображение - Если МинКЧ функционального отображения равно 0. В таком случае любой прообраз может либо не иметь образа, либо иметь не более одного образа. Соответственно, в графе такого отображения могут быть либо висячие вершины-прообразы, либо вершины-прообразы, которым инцидентна одна дуга экземпляра отображения.

Полное функциональное отображение - . Если МинКЧ функционального отображения равно 1. В таком случае каждый прообраз должен иметь один и только один образ. Если посмотреть

на граф экстенционала такого отображения, то нетрудно заметить, что из каждого элемента $ООО$ выходит ровно одна дуга экземпляра отображения.

Дополнительно:

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

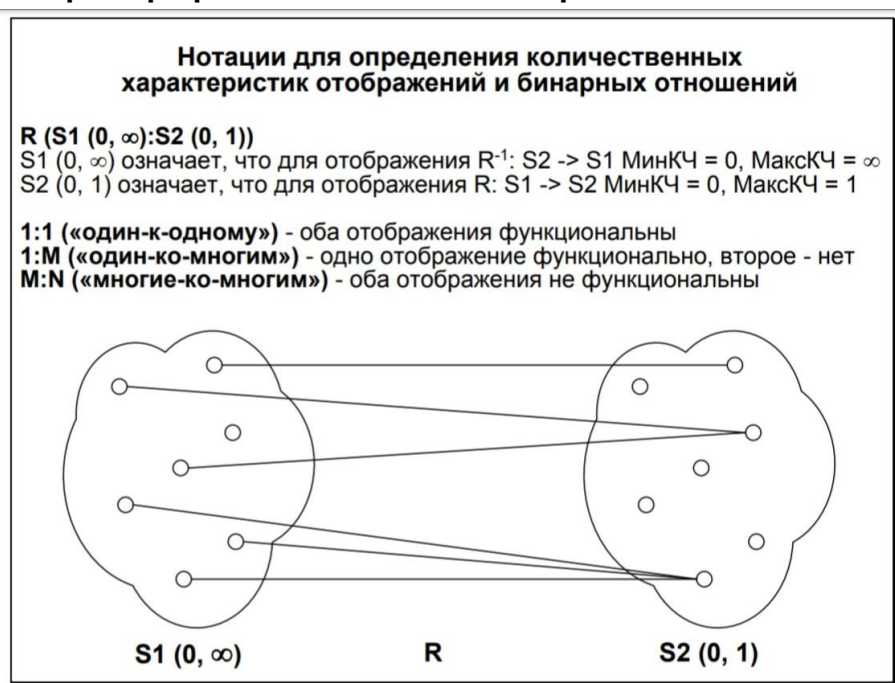
53. Поясните фразу бинарное отношение множеств определяет два отображения между ними.

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.

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



Первая нотация обеспечивает исчерпывающую информацию, как о бинарном отношении, так и обоих определяемых им отображениях – $R(S1 (0, \infty) : S2 (0, 1))$. Она говорит, прежде всего, о том, что бинарное отношение R определено на множествах S1 и S2. S1 (0, ∞) означает, что МинКЧ и МаксКЧ отображения $S2 \rightarrow S1$ равны 0 и ∞ соответственно. Иными словами, любой элемент S2 может быть связан минимум с 0 и максимум с ∞ элементов

$S1.S2(0,1)$ означает, что МинКЧ и МаксКЧ отображения $S1 \rightarrow S2$ равны 0 и 1 соответственно. Иными словами, любой элемент $S1$ может быть связан минимум с 0 и максимум с 1 элементом $S2$.

Поскольку в БД имеют дело с ПрО, в которых и число объектов, и число связей, и число значений характеристики конечно, знак ∞ в данном случае означает отсутствие ярко выраженной верхней границы МаксКЧ. И, если оно для текущего экстенционала БД равно i , то не исключено такое изменение состояния БД, что оно станет равным $i+1$.

Вторая нотация, хоть и используется чаще, менее информативна, поскольку характеризует лишь МаксКЧ отображений.

Определение. Говорят, что 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)**.

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

Сложные отображения - такие отображения у которых прообраз или образ или и то и другое одновременно являются кортежами других отображений

В **сложных отображениях** каждый экземпляр которых имеет более одного объекта-прообраза и/или объекта-образа, а в **простых отображениях**, у каждого экземпляра которых есть один объект-прообраз и один объект-образ.

Примером является отношение:

Ребенок: Мать x Отец -> Ребенок (к кортежу мама, папа ставится соответствие Ребенка)

Классическим примером такой ситуации в моделировании данных является поставка некоторым поставщиком детали для конкретного проекта. Каждый факт такой поставки

представляет собой тернарную связь между конкретными поставщиком, деталью и проектом.

Соответствие понятий

Классификация Кофмана	Моделирование данных
Соответствие	Отображение
Отображение	Полностью определенное отображение
Функция	Полное функциональное отображение

Классификация отображений в моделировании данных



Роj
 Реi
 Муж : женщина -> мужчина
 Жена : Мужчина -> Женщина

56. Дайте три определения возможного ключа.

Если атрибут (группа атрибутов) отношения...

- 1) функционально определяет все другие атрибуты этого отношения
- 2) функционально определяет отношение (Т.е. отображение между возможным ключом и самим отношением функционально. Другими словами, возможный ключ уникально идентифицирует кортежи отношения)
- 3) не имеет значений-дубликатов в кортежах отношения

...этот атрибут (группа атрибутов) называется возможным ключом (потенциальным ключом, ключом-кандидатом) отношения.

57. Что означает определенность значения атрибута?

Определенность значения атрибута есть отсутствие неопределенности значения в этом атрибуте

Неопределенные значения (часто в моделях данных они задаются псевдопеременной или константой NULL) используются в качестве значений атрибутов любых типов в тех случаях, когда:

- либо значение атрибута для кортежа пока не известно,
- либо данная характеристика не применима к конкретному объекту, и атрибут, соответствующий этой характеристике, никогда не примет какого-то определенного значения для кортежа, представляющего этот объект.

Так вот иногда законы ПрО требуют обязательного наличия определенного значения атрибута в каждом кортеже. В частности, одной из особенностей первичного ключа является то, что в каждом кортеже отношения он должен иметь определенное уникальное значение, иначе все его значения не смогут быть знаками объектов.

Определение: Если между отношением и некоторым его атрибутом отображение полностью определено (а точнее, оно – полное функциональное), этот атрибут **не может иметь неопределенных значений** в кортежах отношения.

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

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.**

Параграф №2.4

59. Дайте определение расширенного состояния БД.

Это такое состояние БД DBSi+1, в котором используются такие дополнительные объекты в операциях, как индикаторы текущих и другие управляющие элементы, которые в строгом смысле не относятся к объектам БД, но они связаны с ней и могут изменяться в результате выполнения операций. Совместно с конкретной реализацией данных они определяют состояние БД. Текущее состояние БД определяется совокупностью значений хранящихся данных, а также значением индикатора текущей, указывающим на строку, к которой осуществлялось последнее обращение.

Выполнение операции дать следующую строку не приведет к изменению реализации БД, но состояние БД изменится, так как изменится значение индикатора текущей.

60. Что такое операция над данными?

Операции над данными – динамические свойства модели данных, выражающиеся множеством операций, которые определяют допустимые действия над некоторой реализацией, выражаемые средствами языка манипулирования данными, переводят БД из состояния DBSi в состояние DBSi+1 (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. Это означает, что выполнение операций не должно приводить к нарушению свойств данных, отраженных в схеме.

Одним из важных моментов определения операций в моделировании данных является обеспечение достаточной простоты их усвоения пользователем.

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

Источник данных и приемник данных

62. Какие действия над данными предусмотрены в технологии БД?

Операции над данными переводит БД из одного состояния в другое

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)

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

А) логической позиции в БД – **селекция по текущей**;

Б) значений данных – **селекция по данным**;

В) связей между данными – селекция по связям.

Данные можно селектировать, базируясь на их логических позициях в множествах, отношениях и т.п. Такой порядок может быть использован для селекции по логической позиции, например, для селекции первого (FIRST), последнего (LAST), следующего (NEXT), предыдущего (PREVIOUS) или n-го элемента. Назовем этот тип селекции селекцией по текущей.

Упорядочение может быть фиксированным или произвольным, если порядок не задан явным образом.

В языке данных селекция посредством текущей принимает форму операций манипулирования текущими.

A) *Данные можно селектировать, базируясь на их логических позициях в множествах и отношениях. Хотя модель может и не поддерживать понятие упорядоченности данных, компьютерная реализация в конечном итоге определяет некоторый порядок. Такой порядок может быть использован для селекции по логической позиции, например, для селекции первого (FIRST), последнего (LAST), следующего (NEXT), предыдущего (PREVIOUS) или n-го элемента. Назовем этот тип селекции селекцией по текущей.*

В языке данных селекция посредством текущей принимает форму операций манипулирования текущими. Мы можем установить текущую в следующую логическую позицию: SET CURRENCY TO NEXT СТУДЕНТ.

Другим вариантом является манипулирование текущими, подразумеваемое другим действием, например: GET NEXT СТУДЕНТ.

Здесь действие выборки подразумевает переустановку текущей.

Б) *Селекция данных может быть задана значениями, относящимися к атрибутам. Например, можно произвести селекцию всех служащих со значением атрибута Зарплата, меньшим, чем 10000. Этот тип селекции называется селекцией по данным. В языке данных селекция по данным обычно принимает форму условия на значения атрибутов, которое задает критерий селекции. Этот критерий может определять как простые, так и сложные условия отбора данных. Простое условие определяется на одном атрибуте и одном значении атрибута и обычно имеет вид: <имя атрибута> <операция сравнения> <значение>.*

В) *Селекция данных может производиться не только по логической позиции и значению, но и в соответствии с логическими связями между данными. Например, при наличии связи типа СОБСТВЕННОСТЬ между типами объектов СТУДЕНТ и АВТОМОБИЛЬ можно осуществлять селекцию всех студентов, владеющих автомобилями, или всех автомобилей, принадлежащих студентам. Селекция такого рода называется селекцией по связям. Пример на выдачу информации о студентах, владеющих*

автомобилями: GET СТУДЕНТ WHERE СТУДЕНТ.Фамилия = АВТОМОБИЛЬ.Фамилия владельца

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

Операции обычно задаются в терминах селекции и действия. Действие определяет характер операции, а селекция – критерий отбора данных, над которыми должно быть произведено действие.

Операции над данными, выражаемые средствами языка манипулирования данными, переводят БД из состояния DBSi в состояние DBSi+1 (или в неопределенное состояние), причем изменение состояния происходит или при изменении реализации данных, или при изменении управляющих элементов.

65. Что такое процедура БД?

Процедура базы данных представляет собой последовательность операций, выполняемых при определенных условиях. Общность этих условий обеспечивает возможность автоматического, без вмешательства пользователя, инициирования процедур. Действия, выполняемые процедурой, могут быть весьма разнообразны и не ограничиваться частью БД. Процедура посредством информации состояния уведомляет систему и пользователя об успешном или неуспешном завершении.

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?

Функции агрегирования – вычисление значений, которые непосредственно не хранятся в БД, например, вычисление сумм, подсчет числа экземпляров, определение минимума и максимума.

Виртуальный атрибут - вычисление значения атрибута, например, вычисление возраста человека по заданной текущей дате и дате рождения, что выполняется по запросу о возрасте.

Триггеры целостности - контроль целостности БД. Процедуры БД для проверки и/или поддержки ОЦ

Триггеры безопасности – обеспечение контроля над доступом, который может быть связан не только с парольной защитой, но и с анализом данных и проверкой полномочий и аутентичности пользователя. (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.)

Операторы доступа – процедуры БД, которые могут использоваться также для расширения языка данных операциями, первоначально в нем не предусмотренными. (например, процедура удаления кортежей-дубликатов, позволяющая модифицировать спецификационную операцию, в результате которой образуется отношение)

Триггеры, запускаемые включением, удалением или обновлением - процедуры БД, которые инициируются операциями модификации одних данных и выполняют косвенные модификации других данных.

Процедуры администратора БД - Процедуры БД, которые могут применяться для сбора статистики или для реализации каких-либо других функций администрирования БД. Они не изменяют данные, но могут модифицировать состояние БД.