

Рубежная работа 1

28 марта 2022 г. 17:40

Вариант 1

1. Типовые функции информационной системы

- Функции сбора и регистрации информационных ресурсов
- Функции хранения информационных ресурсов
- Функции актуализации информационных ресурсов
- Функции обработки информационных ресурсов
- Функции предоставления информационных ресурсов пользователям

2. Отличия иерархической модели данных от сетевой

- В сетевой модели осуществлена **связь** (хранение с помощью ключа, хранение связи = пара ключей + вес)
- Сетевая модель позволяет реализовать связь Many-to-many
- Сетевая модель решает **проблему** дублирования и целостности данных
- В сетевой модели нет дублирования данных

3. Типы атрибутов

- Простые - номер
- Составные - адрес, населенный пункт и тд
- Обязательные атрибуты - однозначно идентифицировать сущность
- Необязательные атрибуты
- Однозначные
- Многозначные

4. Свойства отношений

- Уникальность имени отношения в реляционной схеме (уникальное имя таблицы)
- Каждая ячейка содержит только одно неделимое значение
- Уникальность имени атрибута в пределах отношения
- Значение атрибута берется из одного и того же домена
- Каждый кортеж уникален (не может быть двух одинаковых строк)
- [OPTIONAL] Порядок следования атрибутов и кортежей не имеют значения

5. Операторы селек и их порядок

```
SELECT [DISTINCT|ALL] {*} [[ColumnExpression[AS NewName]]  
FROM TableName [As NewName].[{INNER|LEFTOUTER|FULL} JOIN Table2 [As NewName] ON condition]  
[WHERE condition]  
[GROUP BY ColumnList][HAVING condition]  
[ORDER BY ColumnList [ASC|DESC]]
```

Порядок выполнения:

1. FROM
2. [JOIN ... ON]
3. [WHERE]
4. [GROUP BY]
5. [HAVING] (если есть GROUP BY)
6. SELECT
7. [DISTINCT]
8. [ORDER BY]

6. Определение БД (по любому автору)

Базы данных

1. Совместно используемый набор логически связанных данных, предназначенный для удовлетворения информационных потребностей организации.
(Коннолли и Бегг)
2. Набор постоянно хранимых данных, используемых системами какого-либо предприятия.
(Дейт)
3. Совокупность специальным образом организованных данных, хранимых в памяти вычислительной системы и отображающих состояние объектов и их взаимосвязей в рассматриваемой предметной области.
(Хомоненко)

7. Уровни архитектуры ANSI/SPARC

Стандарт **ANSI/SPARC** — подход построения архитектуры. В ней выделяется три слоя архитектуры.

1. Внешний у-нь

2. Концептуальный у-нь
 3. Внутренний у-нь
- Внешний уровень
- Представление БД с позиции конечного пользователя (*смотрит на человека*).
1. Определяется объем и форма представления данных для принятия решения.
 2. Это не про хранение данных, это про то как принимать эффективные решения на основе данных.
- Концептуальный уровень
- Обобщающее представление БД (то как данные хранятся и какая между ними связь/*смотрит на железо*).
1. Решение о том, какие данные и как они должны храниться (*например, статус будет храниться в числах или в строках*).
 2. Какие ограничения на эти данные накладываются (*проверка регулярными выражениями*).
- Внутренний уровень
- Физическое представление БД с точки зрения конкретного инструмента.
1. Распределение данных по дисковому пространству
 2. Структуры хранения данных
 3. Как реализуется безопасность хранения данных (шифрование)
 4. Сжатие данных, оптимизация памяти
- 8. Определение тета-соединения**
- Тета соединение ($R \bowtie_F S$) - определяет новое отношение которое содержит кортежи из декартового произведения R и S, удовлетворяющие предикату F, где $F = R_{a_i} \theta S_{b_i}$, $\theta \in \{<, \leq, =, \geq, >, \dots\}$ (если =, то эквисоединение)
- 9. Определение операции "проекция"**
- Проекция ($\Pi_{a_1, a_2, a_3 \dots a_n}(R)$) - операция, которая определяет новое отношение, содержащее вертикальное подмножество исходного отношения, создаваемое по средствам извлечения значений назначенных атрибутов и исключения результирующих строк дубликатов
- 10. В чем отличие дублирования от избыточности**
- Принципиально отличие в том, что избыточность предполагает, что мы можем получить какие-то данные о кортеже на основе других кортежей, но при этом зачем-то держим повторяющиеся данные, которые, по-хорошему, можно было бы infer(я чурка) из других данных.
- А дублирование не предполагает возможности inferнуть данные из других данных.
- 11. Определение функциональной зависимости**
- Функциональная зависимость между атрибутами ($X \rightarrow Y$) - в отношении R атрибут y функционально зависит от атрибута x тогда и только тогда, когда каждому значению x соответствует в точности одно значение y, а одному y может отвечать более одного x.
- 12. Когда отношение в первой нормальной форме**
- Отношение находится в первой нормальной форме, если все его атрибуты являются простыми.
- 13. Что такое степень отношения**
- Степень** - количество атрибутов

Вариант 2

- 1. Что описывается на верхнем уровне архитектуры ANSI/SPARC**

Внешний уровень

Представление БД с позиции конечного пользователя (*смотрит на человека*).

 1. Определяется объем и форма представления данных для принятия решения.
 2. Это не про хранение данных, это про то как принимать эффективные решения на основе данных.
- 2. Дайте определение СУБД**

СУБД – программное обеспечение, с помощью которого пользователи могут определять, создавать, и поддерживать БД, а также осуществлять к ней контролирующийся доступ.

СУБД – комплекс языковых и программных средств, предназначенный для создания, ведения и использования БД многими пользователями.
- 3. Перечислите типы связей в модели сущность-связь**

Связь 1 к 1: первичный ключ одного из отношений является внешним ключом

Связь 1 к n: значения не ключевого поля берутся из потенциального ключа другого отношения

Связь n к n: может быть реализован только через таблицу связку, содержащую пары потенциальных ключей.

4. В чем отличие реляционной и постреляционной моделей данных

- **Реляционная модель** предполагает неделимость данных, хранящихся в полях записей таблиц.
- В **ПРМД** полем данных является агрегат (*совокупность множества полей данных, имеющая уникальную идентификацию*)

5. Что такое кардинальность отношения в реляционной модели данных

Кардинальность — это количество кортежей, которое содержит отношение.

6. Дайте определение операции "Выборка" в реляционной алгебре

Выборка ($\sigma_{\text{предикат}}(R)$) - операция, которая определяет результирующее отношение, которое содержит только те кортежи из исходного, которые удовлетворяют заданному условию - предикату.

7. Дайте определение операции "Естественное соединение" в реляционной алгебре

Естественное соединение ($R \bowtie S$) - соединение по эквивалентности двух отношений, выполненное по всем общим атрибутам, из результатов которого исключается по одному экземпляру каждого общего атрибута

8. Дайте определение аномалии модификации

Модификации - изменение значения одной записи повлечет за собой просмотр всей таблицы и изменение некоторых других записей

9. Операторы селек и их порядок

```
SELECT [DISTINCT|ALL] {*|[ColumnExpression[AS NewName]]}
FROM TableName [As NewName].[{INNER|LEFTOUTER|FULL} JOIN Table2 [As NewName] ON condition]
[WHERE condition]
[GROUP BY ColumnList][HAVING condition]
[ORDER BY ColumnList [ASC|DESC]]
```

Порядок выполнения:

1. FROM
2. [JOIN ... ON]
3. [WHERE]
4. [GROUP BY]
5. [HAVING] (если есть GROUP BY)
6. SELECT
7. [DISTINCT]
8. [ORDER BY]

10. Какие условия должны быть, чтобы отношение находилось во второй нормальной форме

Отношение находится во второй нормальной форме, если оно находится в первой нормальной форме и каждый не ключевой атрибут функционально полно зависит от первичного ключа.

11. Дайте определение понятию "суперключ"

Суперключ - атрибут или множество атрибутов, единственным образом инициализирующих кортеж

12. Что такое частичная функциональная зависимость

Частичная функциональная зависимость - зависимость не ключевого атрибута от части составного потенциального ключа.

Вариант 3

1. Дайте определение понятию "Данные"

Данные — это поддающиеся многократной интерпретации представления информации в формализованном виде, пригодные для передачи, интерпретации и обработки

2. Что такое полная функциональная зависимость

Полная функциональная зависимость - не ключевой атрибут полностью зависит от составного ключа.

3. Концептуальная структура ANSI/SPARC

Концептуальный уровень

Обобщающее представление БД (то как данные хранятся и какая между ними связь/смотрит на железо).

1. Решение о том, какие данные и как они должны храниться (*например, статус будет храниться в числах или в строках*).
2. Какие ограничения на эти данные накладываются (*проверка регулярными выражениями*).

4. В чем отличие между концептуальной и логической модели данных

Логическая модель данных определяет **способ** организации данных ака иерархическая, сетевая, реляционная. Концептуальная же модель описывает предметную область в виде взаимосвязанных объектов, не вдаваясь в детали реализации.

5. Определение домена в реляционной системе

Домен - множество допустимых значений атрибута

6. Что такое левое-внешнее соединение в реляционной алгебре

Левое внешнее соединение ($R \rhd S$) - типа соединения, при котором результирующее отношение включает также кортежи отношения R, не имеющие совпадающих значений в общих столбцах отношения S.

7. Определение объединения в реляционной алгебре

Объединение ($R \cup S$) - определяет новое отношение, которое включает все кортежи, содержащиеся только в R, все кортежи, содержащиеся только в S, и кортежи, содержащиеся одновременно в R и S за исключением дубликатов.

8. Определение 3 нормальной формы

Отношение находится в третьей нормальной форме, если оно находится во второй нормальной форме и ...

- все не ключевые атрибуты взаимно независимы и полностью зависят от первичного ключа.
- ни один не ключевой атрибут не находится в транзитивной функциональной зависимости от первичного ключа.

9. Алгоритм соединения таблиц вложенными циклами

Для каждой строки [r] из [Ведущая таблица]

Для каждой строки [s] из [Ведомая таблица]

Если УдовлетворяетУсловию ([r],[s],[Условие соединения])

Вывести ([r],[s]);

10. Плюсы и минусы иерархической модели данных

Плюсы

- Простота восприятия логической структуры базы данных человеком
- Высокое быстродействие при транзакционной обработке

Минусы

- Требуется много памяти для хранения (*производительность*)
- Сложно контролировать целостность данных (*безопасность*)
- Дублирование данных
- Скорость операций записи
- Огромные трудности при реорганизации структуры (иерархии)
- Невозможна связь Many-to-many

11. Определение потенциального ключа

Потенциальный ключ - суперключ, который не содержит подмножество, содержащее суперключ

Вариант 4

1. Определение СУБД

СУБД – программное обеспечение, с помощью которого пользователи могут определять, создавать, и поддерживать БД, а также осуществлять к ней контролируемый доступ.

СУБД – комплекс языковых и программных средств, предназначенный для создания, ведения и использования БД многими пользователями.

2. Определение внешнего ключа

Внешний ключ - атрибут или множество атрибутов внутри отношения, которые соответствуют потенциальному ключу некоторого может быть того же самого отношения

3. Определение полусоединения

Полусоединение ($R \rhd_F S$) - определяет те кортежи из R, которые входят в соединение R и S.

4. Определение кортежа в реляционных моделях

Кортеж - отдельная строка

5. Что такое транзитивная функциональная зависимость

Транзитивная функциональная зависимость - существует функциональная зависимость $z(x)$, если существует такое множество атрибутов y такое, что есть функциональная зависимость $x \rightarrow y$ и есть функциональная зависимость $y \rightarrow z$.

$X \rightarrow Z,$

$\exists Y: X \rightarrow Y$

$Y \rightarrow Z$

6. Внутренняя структура ANSI/SPARC

Внутренний уровень

Физическое представление БД с точки зрения конкретного инструмента.

1. Распределение данных по дисковому пространству
2. Структуры хранения данных

3. Как реализуется безопасность хранения данных (шифрование)
4. Сжатие данных, оптимизация памяти

7. Определение разности

Разность ($R - S$) - отношение, которое состоит из кортежей, имеющих в отношении R, но отсутствующих в отношении S.

Операция разность требует совместимость по объединению.

Определения, не вошедшие в варианты

1. **Сущность** - объект, который может быть идентифицирован неким способом, отличающим его от других объектов.
2. **Атрибут** - свойство сущности
3. **Поле данных** - неделимая, уникально адресуемая единица данных
4. **Сегмент** - совокупность полей данных, тоже имеющих свою уникальность
5. **Схема отношений** - строка заголовков и столбцов.
6. **Атрибут** - отдельный столбец
7. **Поле** - пересечение картежа и атрибута
8. **Первичный ключ** - один из потенциальных ключей, который выбран для уникальной идентификации кортежей данного отношения.
9. **Целостность:**
 - Сущностная** - в отношении ни один атрибут первичного класса не содержит NULL значение
 - Ссылочная** - если в отношении существует внешний ключ, то его значение должно соответствовать существующему значению потенциального ключа другого отношения
10. **Пересечение ($R \cap S$)** - отношение, которое объединяет кортежи, присутствующие как в отношении R, так и в отношении S.
11. **Декартовое произведение ($R \times S$)** - определяет новое отношение. Которое является результатом конкатенации каждого кортежа из отношения R с каждым кортежем из отношения S.
12. **Нормализация** - преобразование отношений к виду, отвечающему одной из нормальных форм.
13. **Аномалии.** Есть три вида:
 - Модификации** - изменение значения одной записи повлечет за собой просмотр всей таблицы и изменение некоторых других записей. *(было)*
 - Удаления** - при удалении записи может пропасть и другая информация.
 - Добавления** - информацию в таблицу нельзя поместить пока она неполная или требуется дополнительный просмотр
14. Отношение находится в **нормальной форме Бойса-Кодда**, если детерминанты всех зависимостей являются потенциальными ключами
15. Отношение находится в **четвертой нормальной форме**, если оно находится в нормальной форме Бойса-Кодда и не содержит многозначных зависимостей.
16. **Денормализация** - искусственная декомпозиция отношений для уменьшения количества соединений.