Ответы на вопросы по УП11

1)База данных - совместно используемый набор логически связанных данных, предназначенный для удовлетворения информационных потребностей организации.

Банк данных - это система, которая хранит и обрабатывает информацию. Он может включать в себя базы данных, системы управления базами данных, инструменты для анализа данных и другие компоненты.

Сходства между БД и Банком данных:

* Оба представляют собой структуры данных, предназначенные для хранения информации.
* Они могут быть использованы для различных целей, таких как хранение данных, управление информацией и анализ данных.
* Они оба могут быть структурированы и организованы для облегчения поиска и извлечения информации.

Различия между БД и Банком данных:

* Банк данных обычно состоит из нескольких баз данных, в то время как база данных может состоять из одной или нескольких таблиц.
* Банк данных может иметь централизованное управление, в то время как базы данных могут управляться разными организациями.
* Банк данных может хранить различные типы данных, такие как текст, изображения, аудио и видео, в то время как база данных обычно хранит один тип данных.

2) Система управления базами данных (СУБД) - программное обеспечение, которое позволяет создавать, изменять и управлять базами данных. СУБД выполняют следующие функции:

– Создание и изменение структуры баз данных;

– Управление доступом к данным;

– Обеспечение целостности данных;

– Оптимизация запросов к базе данных;

Требования к СУБД:

Надежность: СУБД должна обеспечивать сохранность данных в случае сбоев в системе;

Масштабируемость: возможность масштабирования системы для работы с большими объемами данных;

Производительность: СУБД должна обеспечивать высокую скорость обработки запросов;

Безопасность: защита данных от несанкционированного доступа;

Гибкость: возможность адаптации СУБД к различным условиям работы;

Классификация СУБД:

Реляционные СУБД (MS SQL Server, Oracle, MySQL, PostgreSQL);

Иерархические СУБД (IBM DB2, Microsoft SQL Server, Sybase);

Сетевые СУБД (Oracle, Microsoft SQL Server);

Объектно-ориентированные СУБД (ObjectDB, Versant Object Database);

В NoSQL СУБД (MongoDB, Cassandra, Redis).

3) Архитектура СУБД состоит из следующих компонентов:

Ядро СУБД - отвечает за управление данными, выполнение запросов и обеспечение безопасности данных;

Менеджер транзакций - обеспечивает атомарность, согласованность и изоляцию транзакций;

Процессор запросов - обрабатывает запросы пользователей и генерирует необходимый SQL-код для выполнения запросов;

Подсистема хранения данных - управляет физической структурой данных на диске;

Администратор баз данных - предоставляет инструменты для создания, изменения и удаления баз данных.

4) Настольные СУБД предназначены для работы на отдельных рабочих станциях. Они обычно имеют простой интерфейс и предназначены для небольших объемов данных. Примеры настольных СУБД включают Microsoft Access, SQLite и Oracle Database Lite.

В архитектуре файл/сервер обработка запросов происходит следующим образом: клиент отправляет запрос на сервер, который затем передает его базе данных. База данных обрабатывает запрос и возвращает результаты клиенту. Недостатком этой архитектуры является то, что она может быть неэффективной при большом количестве клиентов, так как каждый клиент должен отправлять запрос на сервер и ждать ответа. Кроме того, безопасность в этой архитектуре может быть проблемой, так как все данные хранятся на одном сервере, который может быть атакован.

5) Система данных - это набор данных, организованных таким образом, чтобы их можно было легко найти, извлечь и использовать. Системы данных могут включать в себя базы данных, файлы, таблицы и другие структуры, которые используются для хранения и управления данными.

6) Многомерная СУБД - это тип СУБД, который используется для хранения и анализа многомерных данных. Многомерные данные представляют собой набор измерений, каждое из которых имеет свои атрибуты.

Измерение - это одно из направлений многомерных данных, которое представляет собой набор атрибутов, определяющих характеристики объекта в данном измерении.

Ячейка - это значение в многомерной модели данных, которое получается путем пересечения значений по всем измерениям.

Срез - это подмножество многомерной модели данных, полученное путем фиксации значений по некоторым измерениям.

7) Предметная область - это совокупность объектов и явлений, которые описываются в базе данных.

Объект - это элемент предметной области, который имеет определенные свойства и характеристики.

Атрибут (элемент данных) - это характеристика или свойство объекта, которое может быть измерено или описано.

Запись - это экземпляр объекта, то есть конкретный объект с определенными значениями атрибутов.

Кортеж - это упорядоченный набор значений атрибутов, который представляет собой описание объекта.

8) Проектирование базы данных - это процесс создания структуры базы данных, которая будет удовлетворять требованиям пользователей и обеспечивать эффективное хранение и обработку данных. Проектирование БД включает в себя определение структуры таблиц, установление связей между таблицами, определение необходимых индексов и ограничений.

К базам данных предъявляется ряд требований, таких как :

Нормализация данных позволяет избежать дублирования информации и обеспечивает целостность данных.

Производительность базы данных должна быть высокой, чтобы обеспечить быстрое выполнение запросов пользователей.

Безопасность данных обеспечивается путем шифрования данных, использования механизмов аутентификации и авторизации пользователей, а также обеспечения физической безопасности серверов и баз данных.

9) Этапы жизненного цикла базы данных включают в себя следующие этапы:

– Анализ требований

– Проектирование базы данных

– Разработка структуры таблиц

– Создание схемы базы данных

– Загрузка данных

– Тестирование и отладка

– Развертывание базы данных

– Поддержка и обновление базы данных

10) Модель “сущность-связь” (ER-модель) предназначена для представления и анализа структуры данных в базе данных. Она позволяет определить объекты (сущности) и связи между ними, а также свойства каждого объекта. На ER-диаграмме сущности изображаются в виде прямоугольников, а связи - в виде линий между ними. Типы связей могут быть разными: один-к-одному, один-ко-многим, многие-ко-многим. Класс принадлежности сущности определяет, является ли данная сущность обязательной или необязательной.

11) Связь 1:1 означает, что каждой записи в одной таблице соответствует ровно одна запись в другой таблице. Для преобразования такой связи в реляционные таблицу, необходимо создать две таблицы, и связать их через общий атрибут или набор атрибутов.

Связь 1:М означает, что одной записи в первой таблице соответствует несколько записей во второй таблице. В этом случае, для преобразования в реляционную таблицу, нужно создать таблицу с атрибутами первой связи и таблицу с атрибутами второй связи, а затем установить связь между ними через внешний ключ.

М:N означает много-ко-многим связь, когда одной записи в одной таблице может соответствовать несколько записей в другой таблице, и наоборот. Для преобразования в реляционные таблицы, необходимо создать три таблицы: две с атрибутами связей и третью таблицу, которая будет связывать первые две через два внешних ключа.

12) Существует три основных типа связей между таблицами:

Один к одному (1:1) - это связь между двумя таблицами, при которой каждой записи в одной таблице соответствует только одна запись в другой. Например, таблица “Сотрудники” и таблица “Контакты”, где каждый сотрудник имеет только один контакт.

Один ко многим (1:М) - это связь, при которой одной записи в одной таблице соответствует много записей в другой. Например, в таблице “Сотрудники” может быть много записей о проектах, над которыми они работают.

Многие ко многим (М:N) - это связь, при которой много записей из одной таблицы связаны с многими записями из другой таблицы. Например, если у нас есть таблица “Проекты” и таблица “Сотрудники”, то каждый проект может иметь несколько сотрудников, и каждый сотрудник может работать над несколькими проектами. В этом случае необходимо создать третью таблицу для связи между ними.

13) Computer-Aided Design (CAD) - автоматизированная система, предназначенная для решения задач проектирования. Информационная согласованность в САПР - это точность и надежность данных, используемых в процессе проектирования. Она гарантирует, что вся информация является точной, полной и непротиворечивой на протяжении всего процесса проектирования, от первоначальной концепции до конечного продукта.

14) Независимость данных - это принцип, согласно которому данные должны быть отделены от процессов их обработки. Это означает, что данные должны храниться и обрабатываться таким образом, чтобы они могли быть использованы различными приложениями и системами без необходимости изменения их структуры или формата.

15) Концептуальный уровень - это уровень, на котором определяются основные понятия и объекты предметной области, а также связи между ними. На этом уровне определяются сущности, их атрибуты и отношения между ними, которые затем используются для построения логической модели данных.

16) Логический уровень - на этом уровне происходит преобразование концепций и объектов предметной области в структуры данных, которые могут быть использованы в конкретной СУБД. На логическом уровне определяются типы данных, ограничения целостности, триггеры и процедуры для работы с данными.

17) Физический уровень - это уровень, на котором происходит отображение логических структур данных в структуры, поддерживаемые конкретной СУБД. На этом уровне определяется размещение данных на носителях, методы доступа к данным, оптимизации запросов и т.д.

18) На уровне пользователя информация представляется в виде таблиц, форм, отчетов, графиков и других элементов интерфейса, которые позволяют пользователю работать с данными.

19) Иерархическая модель данных представляет собой структуру, в которой данные организованы в виде древовидной структуры. Пример: каждый сотрудник находится на определенном уровне иерархии и имеет подчиненных.

20) Сетевая модель данных позволяет представить данные в виде сети, где каждый узел может иметь несколько связей с другими узлами. Пример: где пользователи связаны друг с другом через отношения дружбы, подписки.

21) Реляционная модель данных состоит из отношений, которые представляют собой множества объектов. Каждый объект имеет набор атрибутов, которые определяют его свойства. Домен - это множество возможных значений атрибута. Атрибут - это свойство объекта, которое может быть определено на домене.

22) Иерархическая и сетевая модели данных отличаются от реляционной модели тем, что они используют структуры данных, отличные от отношений. В реляционной модели данные связаны между собой через внешние ключи, которые указывают на соответствующие записи в других отношениях.

23) Кортеж - это строка в таблице, которая состоит из значений атрибутов. Кортеж представляет собой набор значений, соответствующих одной строке в таблице.

Отношение - это таблица в реляционной базе данных. Отношение состоит из набора атрибутов и множества кортежей, которые соответствуют этим атрибутам.

24) Отношения в реляционной модели обладают рядом свойств и ограничений:

Атомарность: каждый атрибут должен иметь только одно значение для каждого кортежа.

Отсутствие дубликатов: кортежи не должны содержать повторяющихся групп значений.

Упорядоченность: кортежи должны быть упорядочены, но порядок атрибутов не имеет значения.

Кортежам соответствуют значения атрибутов, но не наоборот.

Значения атрибутов должны быть из одного домена.

Отношения должны быть нормализованы, то есть разбиты на несколько отношений для устранения избыточности данных.

Отношения не должны иметь циклических зависимостей.

Каждая пара атрибутов должна быть связана первичным ключом.

Внешние ключи должны быть уникальными и не должны дублироваться.

25) Поле — это столбец в таблице, который содержит значения атрибутов.

Запись — это строка в таблице, содержащая значения атрибутов для конкретного объекта.

Ключ — это атрибут или группа атрибутов, которые однозначно идентифицируют запись в таблице. Различают первичный и внешние ключи.

Первичный ключ — это атрибут или набор атрибутов, который однозначно определяет запись в таблице и не содержит нулевых значений.

Внешний ключ — это атрибут или набор атрибутов в одной таблице, которые соответствуют первичному ключу в другой таблице и используются для установления связи между таблицами.

26) Реляционная целостность — это набор правил и ограничений, которые гарантируют корректность и согласованность данных в реляционной базе данных. Она включает в себя два основных аспекта:

Целостность таблиц: это правило гарантирует, что данные, хранящиеся в таблицах, соответствуют определенным правилам. Например, в таблице не может быть пустых значений для первичных ключей или нарушение ограничений уникальности.

Ссылочная целостность: это правило обеспечивает согласованность между связанными таблицами, гарантируя, что значения внешних ключей всегда соответствуют значениям первичных ключей в других таблицах. Это предотвращает появление несогласованных данных, таких как наличие записи о товаре, который не существует в таблице товаров.

27) Постреляционная модель расширяет возможности реляционной модели, позволяя работать с нетипичными данными, такими как массивы, повторяющиеся группы и сложные типы данных. Она также снимает некоторые ограничения, такие как обязательное наличие первичного ключа и обязательное соблюдение ссылочной целостности. Однако постреляционные системы могут быть менее производительными и менее стандартизированными, чем реляционные системы.

28) Объектно-ориентированный подход к моделированию данных основывается на понятиях объектов, классов, методов, наследования, инкапсуляции и полиморфизма.

Объект - это основной элемент объектно-ориентированной модели, представляющий собой экземпляр класса и обладающий свойствами (атрибутами) и методами.

Класс - это шаблон, по которому создаются объекты. Он определяет структуру и поведение объектов.

Метод - это процедура или функция, связанная с классом и определяющая поведение объектов этого класса.

Наследование - это механизм, позволяющий создавать новые классы на основе существующих. Новый класс наследует свойства и методы родительского класса, и может добавлять свои собственные.

Инкапсуляция - это принцип, согласно которому данные и методы, связанные с объектом, скрываются от внешнего мира. Это обеспечивает защиту данных и упрощает изменение реализации.

Полиморфизм - это возможность использовать один и тот же метод для операций над объектами разных классов. Это позволяет писать более универсальный и гибкий код.

29) Объектно-реляционные модели данных представляют собой компромисс между реляционными и объектно-ориентированными моделями. Они позволяют хранить объекты в реляционных таблицах, но при этом сохраняют некоторые объектно-ориентированные свойства, такие как наследование и полиморфизм.

30) Многомерная модель данных предназначена для анализа и представления данных в виде многомерных кубов.

Измерение - это одна из сторон многомерного куба, представляющая собой категорию данных. Например, измерение “время” может иметь значения “год”, “месяц”, “день”.

Ячейка - это пересечение нескольких измерений, которое содержит значение факта. Например, ячейка “продажи за год” будет содержать количество проданных товаров за определенный год.

Факт - это числовое значение, которое хранится в ячейке. Факты обычно представляют собой количественные показатели, такие как суммы, количества, проценты и т.д.

31) Ключ отношения - это атрибут или набор атрибутов, которые однозначно идентифицируют кортеж в таблице.

Первичный ключ отношения - это особый тип ключа, который уникально идентифицирует каждый кортеж в отношении. Это означает, что для каждого кортежа существует только один набор значений первичного ключа. В реляционной модели данных первичный ключ обязателен для каждой таблицы.

32) Альтернативный ключ отношения - это ключ, который, как и первичный ключ, уникально идентифицирует кортеж, но не является первичным ключом. Альтернативные ключи могут быть полезны, если первичный ключ не подходит для конкретной задачи или если требуется дополнительная степень защиты от дублирования данных.

33) В реляционной модели данных отношения связаны между собой с помощью внешних ключей. Это гарантирует, что данные в связанных таблицах остаются согласованными и не содержат противоречий.

34) Основные принципы реляционной алгебры включают в себя:

– Атомарность: операции выполняются над одним отношением за раз.

– Совместимость типов: аргументы и результаты операций должны иметь совместимые типы.

– Ассоциативность: порядок выполнения операций не имеет значения, если аргументы совместимы.

– Коммутативность: порядок аргументов не имеет значения.

– Дистрибутивность: операция может быть распределена между другими операциями.

Базовые теоретико-множественные операции включают в себя объединение, пересечение, разность и декартово произведение. Специальные теоретико-множественные операции включают выбор, проекцию, соединение и деление.

35) Операция проекции - это процесс выбора определенных атрибутов из отношения для создания нового отношения.

Операция селекции - это процесс выбора кортежей из отношения, удовлетворяющих определенным условиям.

Операция объединения - это процесс объединения двух или более отношений, имеющих одинаковые атрибуты, в одно отношение.

36) Операция пересечения: результатом операции пересечения двух множеств является множество, состоящее из всех элементов, которые присутствуют в обоих множествах.

Операция разности: результатом операции разности двух множеств является множество, состоящее из элементов первого множества, которые не входят во второе множество.

Операция соединения: результатом операции соединения двух или более множеств является единое множество, включающее все элементы исходных множеств.

37) Операция декартово произведение: результатом декартового произведения двух множеств есть множество всех возможных сочетаний этих множеств.

Основные объекты базы данных:

Таблица - основная структурная единица базы данных, состоящая из строк и столбцов.

Строка (запись) - набор данных, описывающих конкретный объект или ситуацию.

Столбец - атрибут, определяющий свойство объекта или ситуации.

Индекс - структура данных, которая ускоряет поиск и сортировку данных в таблице.

Первичный ключ - столбец или группа столбцов, однозначно идентифицирующих каждую строку таблицы.

Внешний ключ - столбец в одной таблице, который ссылается на первичный ключ другой таблицы.

Транзакция - группа операций, которые либо все успешно выполняются, либо все откатываются в случае ошибки.

Триггер - это программа, которая автоматически выполняется при возникновении определенного события в базе данных.

38) Создание базы данных включает в себя несколько этапов:

– Определение целей и задач базы данных;

– Определение структуры данных, необходимых для решения поставленных задач;

– Разработка схемы базы данных и моделей отношений между таблицами;

– Проектирование таблиц базы данных с учетом ограничений целостности и бизнес-правил;

– Создание таблиц базы данных с использованием соответствующего программного обеспечения

Правила создания таблиц баз данных:

– Каждая таблица должна иметь первичный ключ;

– Каждый столбец должен иметь уникальное имя;

– Значения в столбце не должны дублироваться;

– Столбцы, которые используются для связи между таблицами, должны иметь уникальные имена.

39) Строковые данные (String) - представляют собой последовательности символов, которые могут быть различной длины.

Числовые данные (Numeric) - включают в себя различные типы чисел, такие как целые числа (Integer), десятичные числа (Decimal) и числа с плавающей запятой (Float).

Дата и время (Date and Time) - используются для хранения дат и времени. Включают в себя типы данных Date, Time и DateTime.

Логические данные (Boolean) - могут принимать только два значения: True или False.

40) – Имя атрибута

– Тип данных

– Обязательность

– Значение по умолчанию

– Длина и точность

Взаимосвязь между таблицами базы данных может быть определена через внешние ключи. Внешние ключи - это столбцы в одной таблице, которые соответствуют первичным ключам в другой таблице. Это позволяет установить связь между двумя таблицами и обеспечивает целостность данных.

41) Сущность - это объект реального мира, информацию о котором мы хотим хранить в базе данных.

Атрибут - это свойство сущности, которое имеет значение.

Идентификатор сущности - это атрибут или комбинация атрибутов, которые уникально идентифицируют каждую сущность в таблице.

42) “Один к одному” (1:1) - одна запись в одной таблице связана с одной записью в другой таблице;

– “Один ко многим” (1:N) - одна запись в первой таблице связана со многими записями во второй таблице, но каждая запись во второй таблице связана только с одной записью в первой таблице;

– “Многие ко многим” (N:M) - одна запись в каждой таблице связана с множеством записей в другой таблице, и наоборот. В этом случае необходимо создать дополнительную таблицу для связывания двух исходных таблиц.

43) Нормализация таблиц - это процесс организации данных в таблицах таким образом, чтобы исключить дублирование, избыточность и обеспечить целостность данных. Цель нормализации - уменьшение избыточности данных, упрощение их обновления и обеспечение целостности.

Первая нормальная форма (1NF):

– Все атрибуты атомарные, то есть они не могут быть разделены на более мелкие части.

– Все неключевые атрибуты зависят от первичного ключа.

Вторая нормальная форма (2NF):

– Таблица находится в 1NF.

– Ни один неключевой атрибут не зависит от другого неключевого атрибута, а зависит только от первичного ключа.

Третья нормальная форма (3NF):

-Таблица находится во 2NF.

-Ни один неключевой атрибут не зависит функционально от любого другого неключевого атрибута.

44) Преобразование логической модели в физическую включает в себя создание таблиц, индексов, триггеров и других объектов, которые необходимы для работы с данными на физическом уровне.

45) Функциональная зависимость - это отношение между атрибутами таблицы, при котором значение одного атрибута однозначно определяется значением другого атрибута .

Вторая нормальная форма (2НФ) требует, чтобы все атрибуты, не входящие в состав первичного ключа, были функционально зависимыми от ключа. То есть, для каждого атрибута должно быть выполнено условие: атрибут должен зависеть от ключа, но не от других атрибутов. Если это условие не выполняется, то таблица не находится во второй нормальной форме и требуется ее нормализация - разделение на две или более таблиц.