

1. Понятие модели данных

Набор понятий и определений для описания данных, связей между ними и ограничений, накладываемых на данные.

2. Основные понятия сетевой модели данных: запись, набор данных

Запись – поименованная совокупность данных и/или агрегатов, не входящая в другой агрегат.

Набор – поименованная двухуровневая иерархическая структура, связывающая запись-владельца и запись-членов.

3. Обоснование требования отсутствия кортежей-дубликатов, отсутствия упорядоченности кортежей, упорядоченности атрибутов

Свойство, что тело любого отношения никогда не содержит кортежей-дубликатов, следует из определения тела отношения как множества кортежей. В классической теории множеств по определению любое множество состоит из различных элементов.

Конечно, формально свойство отсутствия упорядоченности кортежей в значении отношения также является следствием определения тела отношения как множества кортежей.

Хранить упорядоченные списки кортежей в условиях интенсивно обновляемой базы данных гораздо сложнее технически, а поддержка упорядоченности влечет за собой существенные накладные расходы.

Отсутствие требования к поддержанию порядка на множестве кортежей отношения придает СУБД дополнительную гибкость при хранении баз данных во внешней памяти и при выполнении запросов к базе данных.

Атрибуты отношений не упорядочены, поскольку по определению заголовков отношения есть множество пар <имя атрибута, имя домена>. Для ссылки на значение атрибута в кортеже отношения всегда используется имя атрибута. СУБД сама принимает решение о том, в каком физическом порядке следует хранить значения атрибутов кортежей (хотя обычно один и тот же физический порядок поддерживается для всех кортежей каждого отношения). Кроме того, это свойство облегчает выполнение операции модификации схем существующих отношений не только путем добавления новых атрибутов, но и путем удаления существующих.

4. Возможные подходы относительно обновления первичного ключа целевой сущности на которую ссылается внешний ключ

RESTRICT (ОГРАНИЧИТЬ) - не разрешать обновление, если имеется хотя бы один кортеж в дочернем отношении, ссылающийся на обновляемый кортеж.

CASCADE (КАСКАДИРОВАТЬ) - выполнить обновление и каскадно изменить значения внешних ключей во всех кортежах дочернего отношения, ссылающихся на обновляемый кортеж.

SET NULL (УСТАНОВИТЬ В NULL) - выполнить обновление и во всех кортежах дочернего отношения, ссылающихся на обновляемый кортеж, изменить значения внешних ключей на null-значение.

SET DEFAULT (УСТАНОВИТЬ ПО УМОЛЧАНИЮ) - выполнить обновление и во всех кортежах дочернего отношения, ссылающихся на обновляемый кортеж, изменить значения внешних ключей на некоторое значение, принятое по умолчанию.

IGNORE (ИГНОРИРОВАТЬ) - выполнить обновление, не обращая внимания на нарушения ссылочной целостности.

5. Достоинства и недостатки языка SQL

Достоинства:

- повсеместная распространенность;
- быстрое обучение в простых случаях;
- связывание с различными языками программирования;
- поддержка ODBC и JDBC;
- эффективность современных средств обработки запросов.

Недостатки:

- несоответствие реляционной модели данных;
- недостаточно продуманный механизм неопределенных значений;
- сложность формулировок и громоздкость.

6. Понятие обозначающей сущности. Пример

Формализует связь «многие к одному» и «одна к одной» и отличается от характеристической тем, что не зависит от обозначающей сущности. Пример: сущность «собака», связанная с обозначающей сущностью «владелец собаки» (собака может существовать без хозяина).

7. Многозначные зависимости

Пусть R - отношение, и X, Y, Z - некоторые из его атрибутов (или непересекающиеся множества атрибутов). Тогда атрибуты (множества атрибутов) Y и Z **многозначно зависят** от X (обозначается $X \twoheadrightarrow Y \mid Z$), тогда и только тогда, когда из того, что в отношении R содержатся кортежи $r_1=(x, y, z_1)$ и $r_2=(x, y_1, z)$ следует, что в отношении R содержится также и кортеж $r_3=(x, y, z)$.

8. Условия нахождения отношений в 4-ой нормальной форме

Переменная отношения R находится в четвёртой нормальной форме тогда и только тогда, когда в случае существования подмножеств A и B атрибутов этой переменной отношения R , для которых выполняется нетривиальная многозначная зависимость $A \twoheadrightarrow B$, все атрибуты переменной отношения R также функционально зависят от A .

9. Правила индексирования

Число строк > 200 .

Столбцы, которые часто используются в WHERE, ORDER BY, GROUP BY, DISTINCT.

Столбцы, используемые для соединения таблиц или являющиеся внешними ключами.

Индексирование столбцов с высокой селективностью.

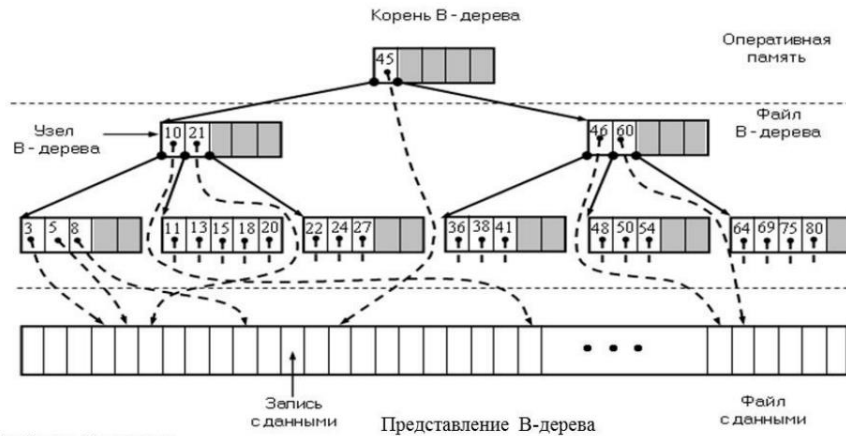
Не следует индексировать столбцы, используемые только в функциях.

Не следует индексировать часто изменяемые столбцы.

10. Способы хранения отношений

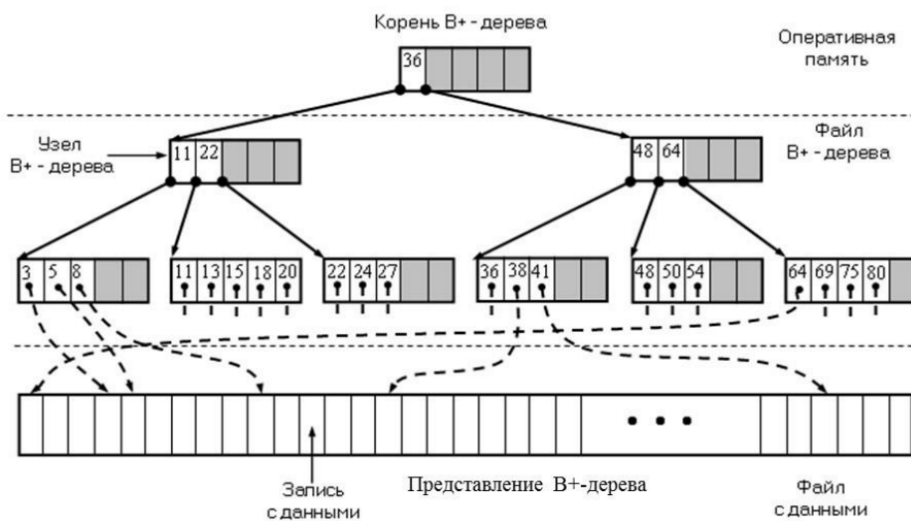
По кортежам, по столбцам.

11. Б и Б+ деревья для хранения индекса



Свойства В-дерева:

- Корень является листом, либо имеет по крайней мере двух сыновей.
 - Каждый узел, за исключением корня и листьев, имеет от $(m/2)$ до m сыновей.
 - Все пути от корня до любого листа имеют одинаковую длину.
 - Ключи в вершинах отсортированы по возрастанию.
 - ссылки на записи данных размещаются не только в листьях, но и во внутренних узлах дерева.
- +
- Нелистовые вершины содержат только ключи и ссылки на дочерние страницы;
 - листовые вершины содержат все множество ключей отношения, сами указатели на записи, плюс указатель на следующий по порядку лист;
 - ключи в листовых вершинах отсортированы по возрастанию.

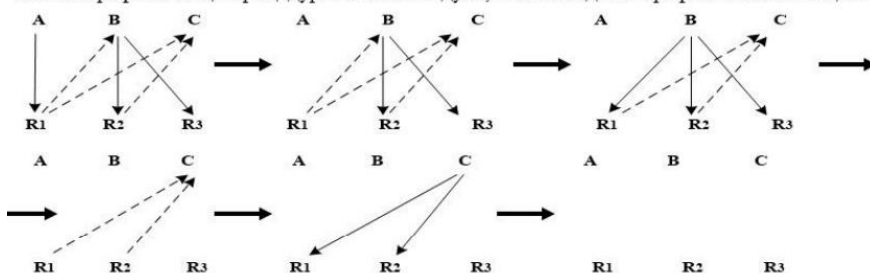


12. Распознавание тупиковых ситуаций при обработке транзакций

Алгоритм редукции графа:

1. Из графа ожидания удаляются все дуги, исходящие из вершин-транзакций, в которые не входят дуги из вершин-объектов (моделирование ситуации успешного завершения транзакций, не ожидающих удовлетворения захватов).
2. Для тех вершин-объектов, для которых не осталось входящих дуг, но существуют исходящие, ориентация одной из исходящих дуг изменяется на противоположную (моделирование удовлетворение захватов).

Если в графе в конце процедуры остались дуги, то в исходном графе имел место цикл.



13. Механизм событий

Механизм событий позволяет прикладной программе и серверу БД уведомить другие программы о некотором событии и тем самым синхронизировать их выполнение.

Create dbevent <имя события>. Вызов события raise dbevent <имя события>.

Программа регистрируется как получатель события register dbevent.

get dbevent имя процедуры(Имя события=Dbeventname, ее параметры).

В этой процедуре проверяется название события. Если оно известно, то что-то делается.

14. ETL процесс

Процесс управления хранилищами данных, состоящий из:

Извлечение данных из внешних источников

Преобразование данных:

- Обобщение данных

- Перевод значения

- Создание полей

- Очистка данных

Загрузка данных в хранилище данных

15. Наследование в объектно-реляционной модели данных

Наследование представляет собой возможность одному объекту приобретать свойства другого объекта, например, если одна таблица определяется как подтаблица другой, то она наследует поведение супертаблицы (ограничения, триггеры, индексы, функции и пр.). Аналогично строчный тип, определенный как подтип базового типа, наследует поля данных и поведение супертипа.

Create table t2 of type NAME under t3. Если нужно работать только с супертипом пишется ...Only(супертип)

16. Отличия задачи классификации от задачи кластеризации

Перечень групп в кластеризации определяется в процессе работы алгоритма.

В классификации перечень групп уже задан изначально.

17. Мандатная модель безопасности

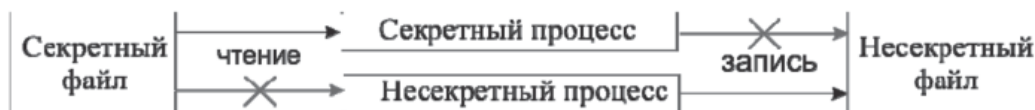
Пусть в системе имеются данные (файлы) двух видов: *секретные* и *несекретные*, а пользователи этой системы также относятся к двум категориям: с уровнем допуска к несекретным данным (несекретные) и с уровнем допуска к секретным данным (секретные).

1. Свойство простой безопасности: несекретный пользователь (или процесс, запущенный от его имени) не может читать данные из секретного файла.

2. Пользователь с уровнем доступа к секретным данным не может записывать данные в несекретный файл. Если пользователь с уровнем доступа к секретным данным скопирует эти

данные в обычный файл (по ошибке или злему умыслу), они станут доступны любому «несекретному» пользователю. Кроме того, в системе могут быть установлены ограничения на операции с секретными файлами (например, запрет копировать эти файлы на другой компьютер, отправлять их по электронной почте и т. д.). Второе правило безопасности гарантирует, что эти файлы (или данные в них) никогда не станут несекретными и не «обойдут» эти ограничения. Таким образом, вирус, например, не сможет похитить конфиденциальные данные.

Общее правило звучит так: пользователи могут читать только документы, уровень секретности которых не превышает их допуска, и не могут создавать документы ниже уровня своего допуска. То есть теоретически пользователи могут создавать документы, прочесть которые они не имеют права.



18. Принцип многоэкземплярности в модели Белла-Лападула

В рамках одного отношения может существовать множество кортежей с одним и тем же значением первичного ключа.