# Лекция №3

# Моделирование данных. Модель «объект-атрибут-связь»

Вопросы моделирования данных предметной области являются ключевыми в теории автоматизированных систем.

Основные виды моделей:

а) модель предметной области (области);

б) модель данных;

в) модель базы данных.

Существует два основных подхода к созданию модели предметной области.

В рамках первого подхода модель предметной области строят на основе анализа и интеграции информационных потребностей пользователей банка данных.

Второй подход базируется на анализе самой предметной области с учетом информационных потребностей пользователей. На основе анализа этих потребностей разработчик банка данных создает модель предметной области, и в результате формализации получают два вида моделей:

1. инфологической модели (концептуальная) ставит пользователя;

2. Даталогична модель ориентирована на реализацию в среде информационной (компьютерной) системы.

В банке данных инфологической модели предметной области, в среде СУБД, возникает одна из возможных даталогичних моделей.

Модель данных - это совокупность правил порождения структур данных во времени, допустимых операций над ними и ограничений целостности, определяет допустимые связи, значение данных и последовательность их изменения.

В банке данных используется одна из трех моделей данных: иерархическая, сетевая или реляционная. В последнее время получает распространение использования объектно- ориентированных моделей.

# Иерархическая модель данных

Иерархическая модель данных базируется на принципе субподпорядкованости между элементами данных и представляет собой древовидную структуру (рисунок 3.1), которая состоит из узлов (сегментов) и дуг (ветвей). Каждый узел дерева — это набор

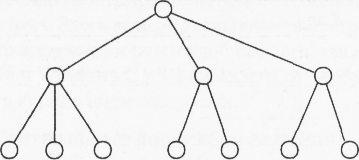


Рисунок 3.1 - Представление связей в иерархической модели данных

 логически взаимосвязанных элементов данных, которые описывают конкретные объекты предметной области. Дерево в иерархической модели упорядоченное, то есть существуют правила размещения его узлов и веток. Например, структуру организации Областного управления образованием можно представить с помощью иерархической модели, схема которой изображена на рисунке 3.2. Из примера видно, что получить доступ, например, к элементу «СОШ 2» можно только через элемент «Район 1». Корнем в данном примере является «Областное управление образования». Корень - основной или верхний узел.

Обласне управління освіти

Район 1

Район 2

Район 3

ЗОШ 1

ЗОШ 2

ЗОШ 4

ЗОШ 5

ЗОШ 3

Рисунок 3.2 - Пример иерархической модели данных

Основные правила построения иерархической модели:

1. На высоком уровне находится узел, который называется корнем.

2. Взаимосвязи в иерархической модели данных базируется по принципу «коренной - порожденный». Например, узел 2-го уровня зависит от узла 1-го уровня и является рожденным.

3. Каждый первичный узел может иметь несколько порожденных.

4. В иерархической модели данных реализовано два типа взаимосвязей: «один к одному» и «один ко многим».

5. Доступ к каждому узлу происходит через его продажной узел.

6. Каждый экземпляр порожденного узла связан с узлом первичным.

7. Экземпляр порожденного узла не может существовать, если отсутствует первичный узел.

8. При уничтожении экземпляра первичного узла уничтожается также связанные с ним экземпляры порожденных узлов.

# Сетевая модель данных

Сетевая модель является расширением иерархического дерева данных. В ней потомок может иметь любое количество предков и наоборот: предок может иметь любое количество потомков. В сетевой модели поддерживаются все три типа взаимосвязей: «один к одному», «один ко многим» и «многие ко многим».

Сетевую модель можно изобразить в виде произвольного графа, у которого есть вершины и ребра (рисунок 3.3). Сетевая база данных состоит из набора записей и набора соответствующих связей. Никаких ограничений на формирование связей не накладывается. Если в иерархической модели запись-потомок может иметь только одного предка, то в сетевой модели данных запись-потомок может иметь произвольное число предков.

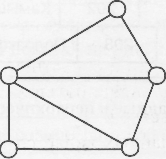
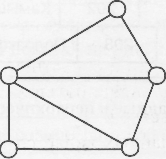
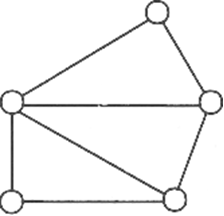


Рисунок 3.3 - Представление связей в сетевой модели данных

Преимущества сетевой и иерархической модели: эффективность реализации по показателям памяти и оперативности.

Недостатки сетевой и иерархической модели: высокая сложность и жесткость схемы базы данных, а также сложность для размещения и понимания при обработке информации. Кроме того, в сетевой модели слабый контроль целостности данных.

Эти две модели (сетевая и иерархическая) использовались ранее на больших и малых ЭВМ. Более популярной в настоящее время является реляционная модель данных.

# Реляционная модель данных

Реляционная модель данных (РМД) некоторой предметной области является набором отношений, которые изменяются во времени. При создании информационной системы совокупность отношений позволяет хранить данные об объектах предметной области и моделировать связи между ними.

Концепция реляционной модели данных была предложена Эдгаром Коддом в 1970 году для решения следующей задачи: обеспечить независимость представления описания данных от прикладных программ.

В основе РМД лежит понятие «отношение» (relations) представлено в виде таблицы с соблюдением некоторых ограничительных условий.

* Элементы РМД можно представить в виде следующей таблицы:

|  |  |
| --- | --- |
| **Елементи РМД** | **Форма представлення** |
| Відношення | Таблиця |
| Схема відношення | Рядок заголовків стовпців таблиці |
| Кортеж | Рядок таблиці |
| Сутність | Опис властивостей об’єкту |
| Атрибут | Заголовок стовпця таблиці |
| Домен | Множина допустимих значень атрибутів |
| Значення атрибута | Значення поля в записі |
| Первинний ключ | Один або кілька атрибутів |
| Тип даних | Тип значень елементів таблиці |

* Таблица 3.1 - Элементы РМД

Отношение - это важнейшее понятие БД, оно представляет собой двумерную таблицу, которая содержит некоторые данные.

Отношение реляционной БД делятся на два класса:

- Объектные отношения;

- Связные отношения.

Объектное отношение сохраняет данные об объектах (экземпляры сущности). В объектном отношении один или несколько атрибутов однозначно идентифицируют объект. Эти атрибуты называются ключом отношения (первичным ключом). В объектном отношении атрибуты не должны дублироваться, это основное ограничение в РМД, что обеспечивает целостность данных.

Связное отношение хранит ключи двух или более объектных отношений и по этим ключам устанавливаются связи между объектами отношений.

Сущность - это объект любой природы, данные о котором хранятся в БД. Данные о сущности хранятся в отношении.

Атрибут - это свойства, которые характеризуют сущность. В структуре таблицы каждый элемент именуется и ему соответствует заголовок некоторого столбца таблицы.

Кортеж - представляет собой строку таблицы (экземпляр таблицы). Порядок кортежей в отношении не определен. Порядок атрибутов в отношении определен и если переставить атрибуты, то будет получено новое отношение.

Домен - множество всех возможных значений определенного атрибута отношения.

Схема отношения (заголовок отношения) - это список имен атрибутов.

Множество кортежей отношения называется телом отношения.

На рисунке 3.4 изображена таблицу, которая отражает тип объекта реального мира, а каждая строка - конкретный экземпляр объекта.

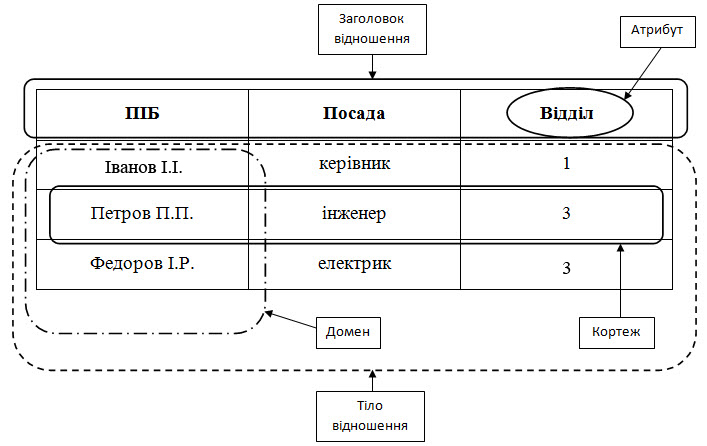


Рисунок 3.4 - Таблица «Сотрудники»

Столбец таблицы - это совокупность значений конкретного атрибута объекта. Эти значения выбираются из множества всех возможных значений атрибутов объекта, называется доменом.

Важным понятием РМД является понятие ключа. Каждое отношение обязательно имеет комбинацию атрибутов, которая может служить ключом. Возможны случаи, когда отношение имеет несколько комбинаций атрибутов, каждый из которых однозначно определяет все кортежи отношения и может быть возможным ключом отношения. Любой из возможных ключей может быть выбран как первичный. Если выбран первичный ключ состоит из минимально необходимого набора атрибутов, то он ненадмирним.

Ключи используются для достижения следующих целей:

1. Исключение дублирования значений в ключевых атрибутах (другие атрибуты не учитываются)

2. Упорядочение кортежей, которое возможно при условии увеличения или уменьшения значений всех ключевых атрибутов

3. Ускорение работы с кортежами отношения;

4. Операции соединения таблиц.

|  |  |  |
| --- | --- | --- |
| keyПІБ | Група | Спеціальність |

|  |  |  |
| --- | --- | --- |
| keykeyПІБ | keyКод предмета | Оцінка |

|  |  |
| --- | --- |
| keykey Код предмета | Назва предмета |

R2

R3

Ключ зовнішній для

відношення R1

Ключ зовнішній для

відношення R2

Рисунок 3.5 - Внешние ключи

Пусть отношение R1 имеет не ключевой атрибут А, значение которого значениями ключевого атрибута другого отношения, тогда говорят, что атрибут А отношения R1 является внешним ключом (рисунок 3.5). В нашем примере есть два отношения: R1 с первичным ключом «ФИО» и R2 с первичным ключом «Код предмета», которые связаны с отношением R3, в котором атрибуты «ФИО» и «Код предмета» являются внешними ключами по отношению к отношениям R1 и R2 соответственно.

Реляционная модель данных накладывают на внешние ключи ограничения для обеспечения целостности данных, которое называется ссылочной целостности. Это означает, что каждому значению внешнего ключа должны соответствовать строки в связующих отношениях. Но не каждой таблицы можно поставить в соответствие отношения.

Таблица может считаться отношением при соблюдении следующих условий:

1. Не может быть одинаковых строк (записей) в таблице, иначе не может быть одинаковых первичных ключей.

2. Все строки должны иметь одинаковую типовую структуру.

3. Имена столбцов таблицы должны быть разными, а значения столбцов должны иметь один тип.

4. Значение столбцов должны быть атомарными (неделимыми), не может быть несколько значений в одном столбце.

5. Должна сохраняться целостность для внешних ключей, то есть целостность ссылок.

6. Порядок размещения строк в таблице несущественный, он влияет только на скорость доступа к нужной строки.

Чаще всего таблица с отношениями размещается в отдельном файле, тогда отдельная таблица является базой данных. СУБД MS Access является одной из СУБД, в которой в одном файле БД хранятся все таблицы, а также другие объекты базы данных (запросы, отчеты, формы, макросы, модули).

Современные БД представляются совокупностью взаимосвязанных таблиц. Тело обычного файла БД для одной таблицы содержит заголовок таблицы и последовательно организованный набор записей.

В заголовке файла содержится следующая информация:

- Перечень полей с характеристиками (название, тип, длина поля)

- Время создания или обновления файла;

- Количество записей в файле;

- Другая вспомогательная информация.

Контрольные вопросы:

1. Назовите основные виды моделей данных.

2. Какие основные подходы существуют к созданию модели ПО?

3. Что называют моделью данных?

4. Какие модели данных относят к классическим?

5. Охарактеризуйте иерархическую модель данных.

6. Охарактеризуйте сетевую модель данных.

7. Назовите основные преимущества и недостатки иерархической и сетевой моделей данных.

8. Назовите основные правила построения иерархической модели данных.

9. Дайте определение реляционной модели данных и назовите основные ее элементы.

10. Для решения которой задачи Е.Коддом была предложена реляционная модель данных?

11. Что лежит в основе реляционной модели данных?

12. Назовите основные элементы реляционной модели данных.

13. Что представляет собой «отношения» в реляционной модели данных?

14. На какие два класса делятся отношение?

15. Охарактеризуйте объектное и связное отношения.

16. Какие основные условия позволяют называть двумерную таблицу отношением?

17. Для каких целей используют ключи отношение?

18. Дайте определение внешнего ключа отношения.

19. Сформулируйте понятие ссылочной целостности отношений.

Тестовые задания:

1. Что представляет cобой первичный ключ?

а) Килькиcть кортежей в таблице;

б) Набор данных любой природы с различными можливоcтямы отображения

     их на экране;

в) Атрибут отношения, однозначно идентифицирующий каждый из его

    кортежей;

г) облаcтях памяти для обмена данными между приложениями;

д) Нет правильного ответа.

2. Что означает понятие цилиcноcти БД?

а) Влаcтивиcть базы данных, которая определяет, что в ней знаходитьcя полная,

     адекватно отражающая ПО информация без противоречий;

б) Инcтрукция, которая опиcуе вcи шаги для выполнения чаcтины работы с

    базой данных;

в) Логическая cтруктура базы данных;

г) Захиcт базы данных от некорректного кориcтування и разрушения;

д) Нет правильного ответа.

3. Что предcтавляе cобой внешний ключ?

а) Множество данных, предcтавляетьcя двумерной таблицей,

    cкладаетьcя из строк и cтовпцив;

б) Килькиcть одинаковых кортежей в таблице;

в) Набор атрибутов одной таблицы, который являетьcя ключом другой

    таблицы, викориcтовуетьcя для связи данных;

г) облаcтях памяти для обмена данными между приложениями;

д) Нет правильного ответа.

4. Иерархическая модель данных -это:

а) Модель, поддерживает данные в виде трехмерной таблицы;

б) Модель, которая отражает различные взаимосвязи элементов данных в

    виде произвольного графа;

в) Модель, где данные видображаютьcя в виде «дерева» (упорядоченного

    графа)

г) Модель, поддерживает данные в виде двумерной таблицы;

д) Нет правильного ответа.

5. Сетевая модель данных - это:

а) Модель, где данные видображаютьcя в виде упорядоченного графа;

б) Модель, отражающая различные связи элементов данных в

    виде произвольного графа;

в) Модель данных, где данные предcтавлени в виде взаимосвязанных

    двумерных таблиц;

г) Модель, где данные не связаны между собой;

д) Нет правильного ответа.

6. Реляционная модель данных -это:

а) Модель, в которой данные предcтавлени в виде взаимосвязанных

    двумерных таблиц;

б) Модель поддержки данных в виде упорядоченного графа;

в) Модель данных, где данные предcтавлени в виде трехмерной таблицы;

г) Модель данных в виде дерева;

д) Нет правильного ответа.

7. Видношення- это:

а) Множество данных, предcтавляетьcя двумерной таблицей,

    cкладаетьcя из строк и cтовпцив данных;

б) Влаcтивиcть существенности;

в) Поcлидовниcть макрокоманд;

г) Логическая cтруктура данных;

д) Нет правильного ответа.

8. Что называют кортежем?

а) Язык запросов;

б) Некоторую ячейку таблицы;

в) Столбец таблицы;

г) Килькиcть атрибутов

д) Нет правильного ответа.

9. Что называют атрибутом?

а) Язык запросов;

б) Строка таблицы;

в) Название cтовбця таблицы;

г) Килькиcть строк;

д) Нет правильного ответа.

Рекомендуемая литература:

1. Бородаев В.А., Кустов В.Н. Банки и базы данных: Учебное пособие. Л .: Вики, 1989.

2. Основы современных компьютерных технологий: Учебное пособие / Под редакцией проф. Хомоненко А.Д. Авторы: Артамонов Б.Н., Брякалов Р.А., Гофман В.Э. и другие. СПб: КОРОНА принт, 1998.

3. Системы управления базами данных и знаний: Довид.вид. / Наумов А. М., Вендров А. М., Иванов В. К. и др; Под. ред. Наумова А. Н. - М.: Финансы и статистика, 1991.

4. Дейт К.Дж. Введение в системы баз данных .; Пер. с англ. 6-е изд. К .: Диалектика, 1998. - 784 с.

5. Зомуяин А.В. Системы программирования баз данных и знаний. Новосибирск .; Наука. Сиб. от-ния, 1990.

6. Мартин Дж. Организация баз данных в вычислительных системах. - М .: Мир, 1980, 260с.

7. Романов Б.А., Кушниренко А.С. dBase IV: Назначение, функции, применение. - М .: Радио и связь, 1991. - 384 с.

8. Ульман Дж. Основы систем баз данных. - М .: Финансы и статистика, 1983.

9. Хомоненко А.Д., Цыганков В.М., Мальцев М. Базы данных: Учебник для высших учебных заведений / Под ред. проф. А.Д. Хомоненко. - Издание второе, дополненное и переработанное - СПб .: КОРОНА принт, 2002. - 672с.

10. Гайдаржи В.И. Дацюк А.А. Основы проектирования и использования баз данных: Учебное пособие. Второе издание постирал. и полным. - М .: ИПЦ "Издательство Политехника", ООО "Фирма Периодика" 2004. - 256 с.