

Лабораторная работа 1

ПРОЕКТИРОВАНИЕ СХЕМЫ БД

Цель работы состоит в разработке схемы реляционной базы данных (БД) для информационной системы.

Теоретические сведения

Проектирование программного обеспечения любой информационной системы включает следующие этапы:

- 1) анализ предметной области и выработка требований к системе;
- 2) логическое представление реальности;
- 3) идентификация и проектирование видимых сущностей;
- 4) разработка абстрактных (поддерживаемых) сущностей и концептуальной схемы базы данных;
- 5) проектирование структуры приложения.

Предметная область – это множество всех предметов или объектов некоторой части реального мира, свойства которых и отношения между которыми рассматриваются в научной теории или в практической деятельности человека.

Для сбора, хранения, поиска и выдачи информации о предметной области и ее объектов в настоящее время в информационных системах широко используются *базы данных*.

Анализ предметной области начинается с выделения сущностей и определения их свойств или атрибутов.

Выделенные сущности могут быть двух типов: видимые сущности и поддерживаемые сущности. *Видимые сущности* представляют собой объекты предметной области, которые может распознать человек. *Поддерживаемые сущности или абстрактные сущности* разрабатываются проектировщиком базы данных для физической поддержки общей логической модели.

Объекты реального мира, помимо непосредственных, прямых связей, имеют друг с другом иные, более сложные причинно-следственные связи. Эти связи и процессы должны каким-то образом отражаться в базе данных, если мы имеем в виду не статичное хранилище, а информационную модель части реального мира.

Иными словами, в базе, помимо собственно данных и непосредственных связей между ними, должны храниться знания о данных, а она сама должна адекватно отражать процессы, происходящие в реальном мире. Следовательно, необходимо иметь средства хранения и управления такой информацией.

Данные требования выливаются в решение следующих задач:

- а) необходимо, чтобы база данных в любой момент времени правильно отражала состояние предметной области – данные должны быть взаимно непротиворечивыми;
- б) база данных должна отражать правила предметной области, законы, по которым она функционирует (business rules);
- в) необходим постоянный контроль состояния базы данных, отслеживание всех изменений, и адекватная реакция на них;
- г) необходимо, чтобы возникновение некоторой ситуации в базе данных четко и оперативно влияло на ход выполнения прикладной программы, поэтому многие программы требуют оперативного оповещения о всех происходящих в базе изменениях.

По способу установления связей между сущностями различают три модели:

1) *Иерархическая модель данных* представляет собой дерево, вершинами которого являются сущности, а дугами – отношения между ними. Достоинством этой модели данных является высокая эффективность. Недостатками являются отсутствие строгой математической основы; неоднородность структуры; неполнота представления, так как не любую предметную область можно представить деревом; асимметрия отношений между сущностями и сложность реорганизации структуры данных.

2) *Сетевая модель данных* – это модель в виде графа произвольного вида, т.е. в виде совокупности поименных узлов, связанных произвольным образом поименными дугами.

Узлами являются сущности, а *дугами* – разнообразные отношения между этими сущностями.

Достоинствами сетевой модели являются адекватность представления предметной области и равноправие между данными (симметрия).

Недостатки – это сложность, неоднородность структуры, отсутствие строгой математической основы, невысокая эффективность и сложная процедура реорганизации такой модели.

3) *Реляционная модель данных* – совокупность n -арных отношений, каждая из которых представляет соответствующую сущность предметной области. Имеется строгая математическая основа – *реляционная алгебра (исчисление)*. Данные в отношениях (атрибуты) являются равноправными, и любые отношения легко реорганизовывать. Модель данных имеет много достоинств: простота, однородность, полнота представления предметной области.

Недостатком такой модели является низкая эффективность вследствие разобщенности семантически взаимосвязанных данных.

На физическом уровне отношения представляются таблицами, атрибуты – столбцами, а наборы значений атрибутов (кортежи) – строками.

Таблица должна иметь уникальное в пределах базы данных имя, оно называет сущность предметной области, а каждая строка конкретный объект; столбец – совокупность значений конкретного атрибута рассматриваемых объектов. Эти значения выбираются из множества допустимых значений –

домена. Каждый столбец имеет имя, уникальное в пределах таблицы. В отличие от полей строки не имеют имен, количество строк в таблице логически не ограничено. Любая таблица должна иметь один или несколько столбцов, значение которых однозначно идентифицируют каждую ее строку. Столбец или их комбинация, обладающая таким свойством, называется *первичным ключом*. Таким образом, он должен обладать свойством уникальности, другим его свойством должна быть минимальность, когда ни один из входящих в ключ столбцов не может быть исключен из него без нарушения свойства уникальности.

Для поддержания связей между таблицами используются *внешние ключи*, когда в данной таблице используются поля, являющиеся первичными ключами для другой таблицы.

Основные требования при физической реализации реляционной базы данных таковы:

- а) каждая таблица должна иметь уникальное в базе данных имя и состоять из однотипных строк;
- б) каждая таблица должна состоять из фиксированного числа столбцов и простых (не составных) значений в каждом столбце;
- в) ни в какой момент времени в таблице не должно быть двух строк, дублирующих друг друга; строки должны отличаться хотя бы одним значением, чтобы была возможность однозначно идентифицировать любую строку таблицы;
- г) каждому столбцу должно быть присвоено уникальное имя;
- д) полное информационное содержание базы данных должно быть представлено в виде явных значений самих данных и только таким образом, а не использованием указателей или ссылок;
- е) при обработке данных должно быть обеспечено свободное обращение к любой строке и столбцу.

Для этих целей используется формальный аппарат ограничений на формирование таблиц. Этот аппарат называется *нормализацией таблиц*. Имеется несколько уровней нормализации:

- а) таблица находится в первой нормальной форме, когда она не содержит повторяющихся полей и составных значений полей;
- б) таблица находится во второй нормальной форме, когда она удовлетворяет требованиям первой нормальной формы и все ее поля, не входящие в первичный ключ, связаны полной функциональной зависимостью с первичным ключом;
- в) таблица находится в третьей нормальной форме, когда она удовлетворяет требованиям второй нормальной формы и ни одно из ее неключевых полей функционально не зависит от любого другого неключевого поля. Это объясняет независимые изменения любого неключевого поля.

Пример выполнения лабораторной работы

В качестве примера рассмотрим проектирование упрощенной базы данных для информационной системы по обслуживанию хранилищ, музеев, выставок и аукционов.

Выделим базовые сущности рассматриваемой предметной области и опишем их атрибуты.

Авторы (Authors) – это создатели произведений живописи, графики или фотографий. Атрибутами сущности являются:

- а) код автора (**AuthorId**);
- б) фамилия автора (**Lastname**);
- в) имя автора (**Firstname**);
- г) отчество автора (**Middlename**);
- д) дата рождения автора (**DateOfBirth**);
- е) дата смерти автора (**DateOfDeath**);
- ж) код страны проживания автора (**Country**).

Произведения (Artworks) – это картины на бытовые и исторические темы, пейзажи, портреты, натюрморты, панно, панорамы, настенные росписи, гравюры, литографии, эстампы и художественные фотографии. Атрибутами сущности являются:

- а) код произведения (**ArtworkId**);
- б) название произведения (**Title**);
- в) жанр (**Genre**);
- г) средства создания (**Tools**);
- д) код автора (**AuthorId**);
- е) дата создания (**CreatDate**);
- ж) цена (**Price**);
- з) отдел хранения (**DepId**).

Сотрудники галереи или музея (Employees). Атрибутами сущности являются:

- а) код сотрудника (**EmployeeId**);
- б) фамилия сотрудника (**Lastname**);
- в) имя сотрудника (**Firstname**);
- г) отчество сотрудника (**Middlename**);
- д) должность (**Position**);
- е) оклад (**Salary**);
- ж) дата начала работы (**BeginDate**);
- з) дата окончания работы (**EndDate**);
- и) отдел (**DepId**).

Отделы галереи или музея (Departments) – отделы, в которых хранятся произведения и работают сотрудники. Атрибутами сущности являются:

- а) код отдела (**DepId**);
- б) название отдела (**Name**).

Реляционная модель для анализируемой предметной области представлена на рис. 1.

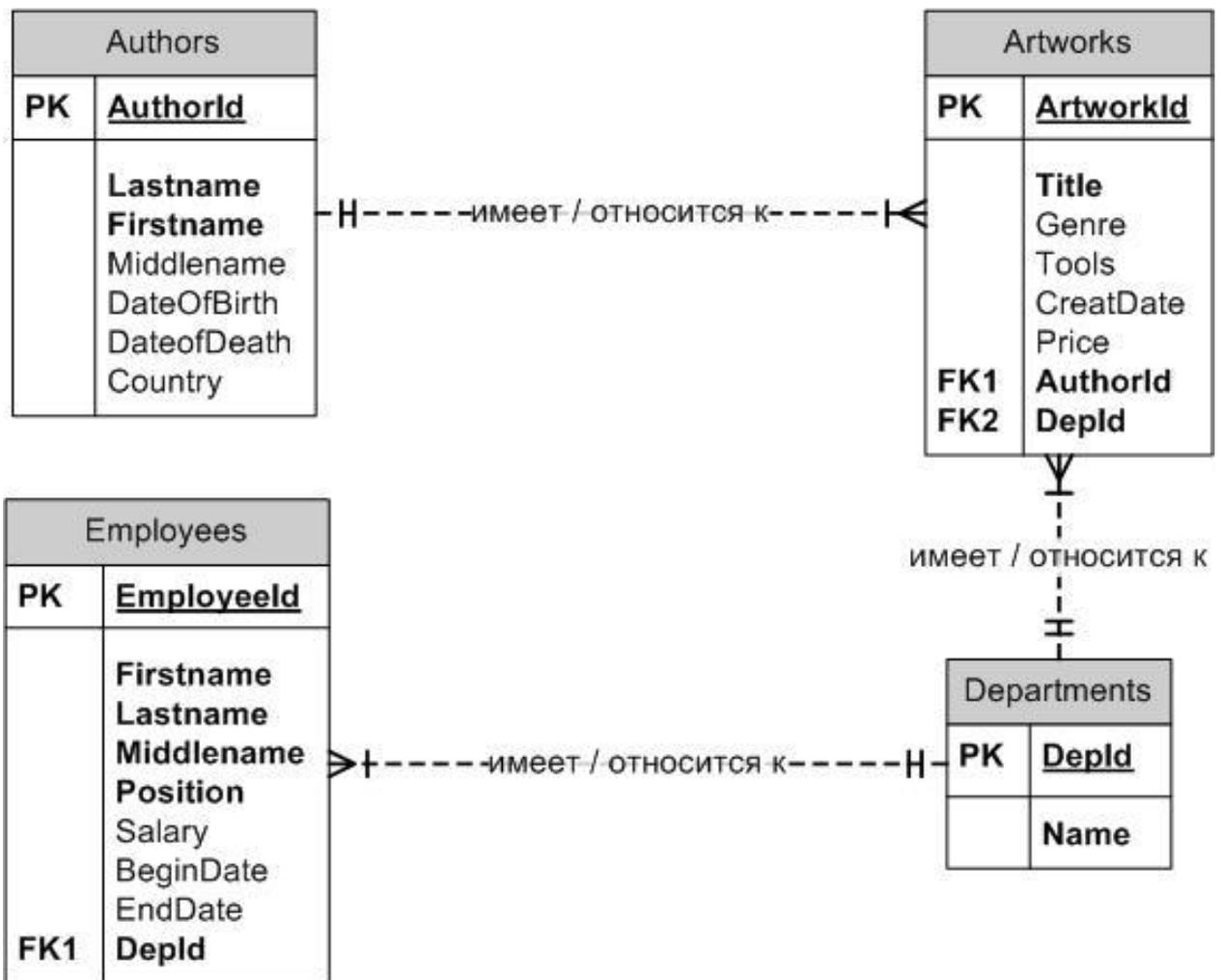


Рисунок 1 – Диаграмма «сущность-связь» реляционной базы данных

Задание

- 1) Разработать схему реляционной БД для задания из Приложения (задание выбирается согласно номеру студента по журналу).
- 2) Показать, что разработанная схема нормализована.