

ОБЛАСТНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
СРЕДНЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«СМОЛЕНСКИЙ АВТОТРАНСПОРТНЫЙ КОЛЛЕДЖ
имени Е.Г. Трубицына»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
по подготовке к выполнению лабораторной работы
по теме «Автоматизированное рабочее место
специалиста в Microsoft Access» по предмету
«Информационные технологии в
профессиональной деятельности»

Часть 1. Концептуальное и логическое проектирование базы данных.

Автор: Рычагов А.В.

Смоленск , 2013 г.

Концептуальное и логическое проектирование базы данных

Цель: приобрести умение анализировать предметную область (ПО) информационной системы и на основе анализа создавать концептуальную, логическую модели будущей базы данных.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

Основные понятия

Автоматизированная информационная система (АИС) – это комплекс программных и технических средств, обеспечивающих сбор, обработку и манипулирование данными.

Цель любой информационной системы (ИС) – обработка данных об объектах реального мира. Основой ИС является *база данных (БД)*. В широком смысле слова *БД* – это совокупность сведений о конкретных объектах реального мира в какой-либо ПО. В узком смысле *БД* – это поименованная, определённым образом организованная совокупность данных, отражающая состояние объектов и их отношений в рассматриваемой ПО.

Под ПО принято понимать часть реального мира, подлежащего изучению для организации управления.

Объектом называют элемент ПО, информация о котором интересует пользователя. Каждый объект описывается рядом основных свойств – *атрибутов*. *Атрибутом* называют поименованную характеристику объекта. Он показывает, какая информация об объекте интересует пользователя и должна храниться в БД. Например, ПО – высшее учебное заведение; объекты – студент, преподаватель; атрибуты – фамилия студента, его адрес проживания, группа, фамилия преподавателя, дисциплина, которую он читает, учёное звание и учёная степень.

Процесс проектирования БД состоит из трёх этапов: *концептуального, логического и физического* проектирования. Результат каждого этапа – соответствующая модель ПО, что отражает трёхуровневую архитектуру (концептуальный, внешний, внутренний уровни) любой автоматизированной ИС.

Этап концептуального проектирования

Концептуальное проектирование начинается с анализа ПО, включает анализ концептуальных требований и информационных потребностей, выявление информационных объектов (ИО) и связей между ними, построение концептуальной модели (схемы) данных.

Объединение частных представлений о содержимом БД, полученных в результате опроса пользователей, позволяет создать обобщённое неформальное описание создаваемой БД. Это описание, выполненное с использованием естественного языка, математических формул, таблиц, графиков и других средств, понятных проектировщикам ИС, называют концептуальной (инфологической) моделью данных.

Основными конструктивными элементами инфологических моделей являются *сущности, связи* между ними и их *атрибуты* (свойства).

Сущность (информационный объект) (ИО) – любой конкретный (реальный) или абстрактный объект в рассматриваемой ПО.

Связь – наблюдаемая взаимосвязь (ассоциация) между сущностями.

Для представления концептуальной модели используют различные методы и модели, например, модель “сущность” – “атрибут” – “связь” (EAR) описывает ПО на концептуальном уровне в виде EAR-диаграмм.

В них сущности помечаются прямоугольниками, ассоциации (характеры объединения сущностей) – ромбами или шестиугольниками, атрибуты – овалами, а связи между ними – рёбрами, над которыми проставляются типы связей.

Между сущностями возможны четыре типа связей: один – к одному ($1 \leftrightarrow 1$), один – ко многим ($1 \leftrightarrow \infty$), многие – к одному ($\infty \leftrightarrow 1$), многие – ко многим ($\infty \leftrightarrow \infty$).

Связь $1 \leftrightarrow 1$: в любой момент времени каждому экземпляру первого ИО соответствует 1 или 0 экземпляров другого ИО и наоборот.

Связь $1 \leftrightarrow \infty$: одному экземпляру первого ИО соответствует 0,1,2,... экземпляров второго и наоборот, каждому экземпляру второго ИО соответствует 0 или 1 экземпляр первого ИО. Аналогично определяется тип связи $\infty \leftrightarrow 1$.

Связь $\infty \leftrightarrow \infty$: одному экземпляру первого ИО соответствует 0,1,2,... экземпляров второго ИО и наоборот.

Примеры:

1. *Студент* $1 \leftrightarrow 1$ *Сессия*: каждый студент имеет определённый набор экзаменационных оценок в сессию. Имеется в виде ИО *Сессия* как набор оценок за текущий семестр.

2. *Стипендия* $1 \leftrightarrow \infty$ *Студент*: вид (и сумма) стипендии может многократно повторяться для различных студентов по результатам сессии.

3. *Студент* $\infty \leftrightarrow \infty$ *Преподаватель*: один студент обучается у многих преподавателей и наоборот, один преподаватель обучает многих студентов.

Концептуальная модель применяется для структурирования ПО с учётом информационных интересов пользователей ИС, она не зависит ни от программных, ни от технических решений.

Рассмотрим пример: проектирование БД ИС “Кадры”. Концептуальная модель, соответствующая ИС в виде EAR-диаграмм “сущность” – “атрибут” – “связь”, представлена на рис. 1.

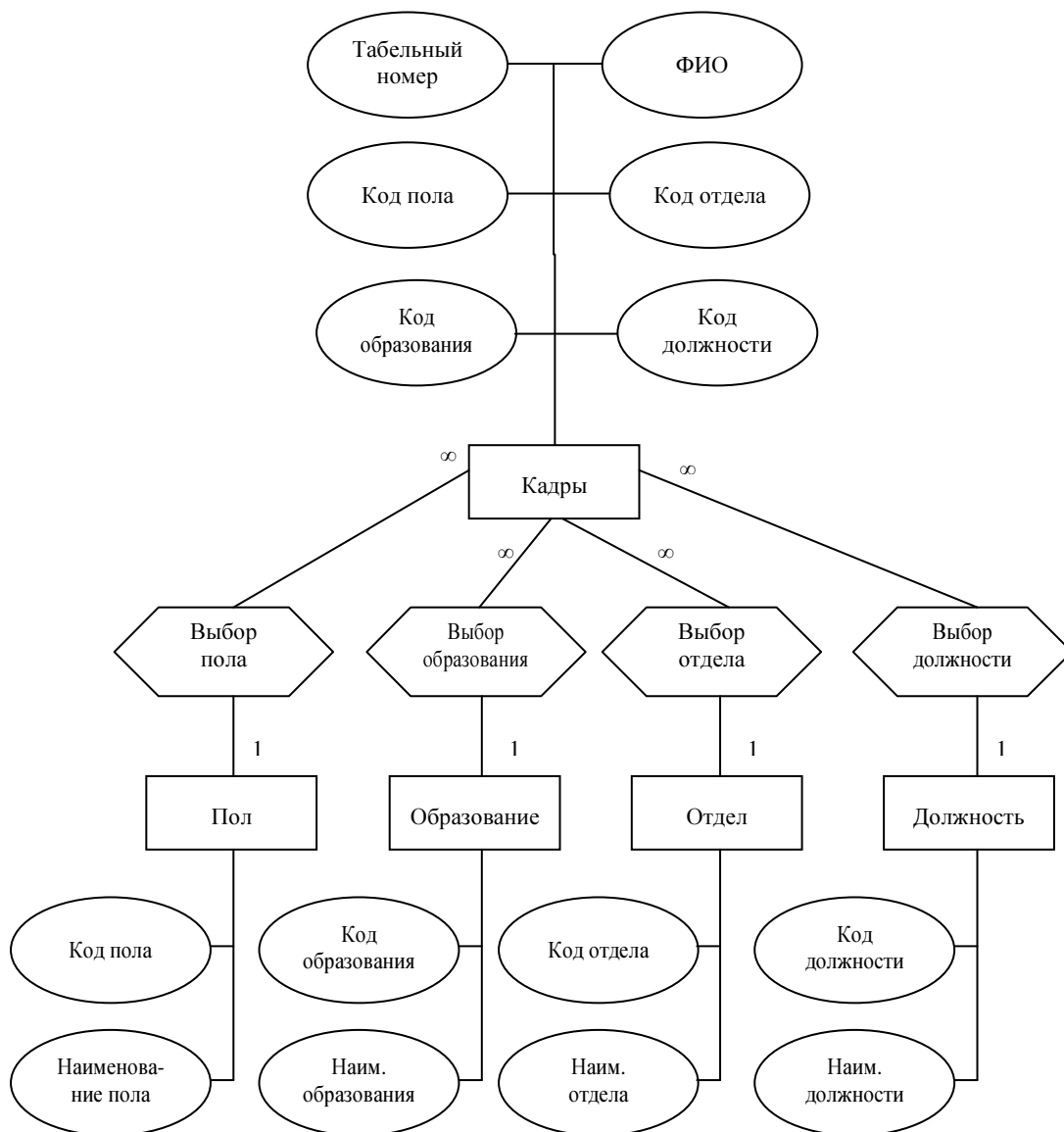


Рис 1. Концептуальная модель «Кадры»

В результате анализа ПО выделено шесть ИО (*Кадры*, *Пол*, *Образование*, *Отдел*, *Должность*), их свойства и связи.

Определим связи между сущностями:

| Название связи | Тип | Связи между сущностями |
|-------------------|-------|------------------------|
| Выбор пола | 1 ↔ ∞ | Пол, Кадры |
| Выбор образования | 1 ↔ ∞ | Образование, Кадры |
| Выбор отдела | 1 ↔ ∞ | Отдел, Кадры |
| Выбор должности | 1 ↔ ∞ | Должность, Кадры |

Итак, концептуальная модель – это описание ПО, включающее совокупность информационных объектов, их атрибутов и взаимосвязей, выявленных в результате анализа.

Этап логического проектирования

Основной задачей *логического проектирования* является разработка логической схемы (модели), ориентированной на выбранную систему управления базами данных (СУБД).

СУБД – это комплекс программных и языковых средств, предназначенных для создания, ведения и совместного применения БД любыми пользователями. СУБД осуществляет централизованное управление базой данных, обеспечивает доступ к данным, по сути, выступает в качестве интерфейса между пользователями и БД.

Одним из основных критериев выбора СУБД является оценка того, насколько эффективно внутренняя модель данных, поддерживаемая системой, способна описать концептуальную схему ПО. Существующие СУБД делятся по типам моделей данных на *реляционные, иерархические и сетевые*.

Модель данных (МД) – формально определённая структура, которая используется для представления данных. Иерархическая МД организует данные в виде древовидной структуры, сетевая – в виде сетевой, реляционная МД – в виде таблиц (отношений).

Процесс логического проектирования состоит из следующих действий:

- выбор конкретной СУБД;
- отображение концептуальной схемы на логическую схему, получение логической МД, соответствующей внешнему уровню архитектуры любой автоматизированной ИС;
- выбор ключей;
- описание языка запросов.

При отображении концептуальной МД ПО на реляционную МД каждый прямоугольник схемы (рис. 1) – информационный объект преобразуется в таблицу (отношение).

Теория конструирования реляционных БД разработана доктором Е.Ф. Коддом в 1968 г. Сформулированные им 13 правил полностью реализованы в реляционной СУБД MS Access. Им же введены основные понятия реляционной БД, которыми являются: “тип данных”, “домен”, “атрибут”, “кортеж”, “первичный ключ”, “отношение”.

Понятие “тип данных” в реляционной МД полностью адекватно понятию типа данных в языках программирования.

Рассмотрим отношение “Должность” (рис. 2).



Рис 2. Отношение «Должность»

Домен определяется как допустимое потенциальное множество значений данного (стандартного) типа. Например, для домена “название счёта” допустимость означает, что множество значений – только те, которые представлены в плане счетов – документе как названия счетов.

Схема отношения – это именованное множество пар {имя атрибута, имя домена}. Степень или “арность” схемы отношения – мощность этого множества. Степень отношения “План счетов” равна 5.

Кортеж, соответствующий данной схеме отношения, – это множество пар {имя атрибута, значение}, которое содержит одно вхождение каждого имени атрибута, принадлежащего схеме отношения, а “значение” – допустимое значение домена данного атрибута.

Отношение – это множество кортежей, соответствующих одной схеме.

В литературе часто используют неформальную терминологию реляционных баз данных: отношение – таблица, кортеж – строка, домен – столбец.

Фундаментальные (базовые) свойства отношений

Отсутствие кортежей-дубликатов. Из этого свойства вытекает наличие у каждого отношения *первичного ключа* (ПК) – набора атрибутов, значения которых однозначно определяют кортеж отношения. Понятие ПК является исключительно важным в контексте с понятием *целостности* БД.

Целостность базы данных – свойство, определяемое способностью СУБД защищать компоненты и связи БД от искажений в результате некорректных операций или сбоев и отказов технических средств.

Отсутствие упорядоченности кортежей. Отсутствие требования к поддержанию порядка на множестве кортежей отношения даёт дополнительную гибкость СУБД при хранении БД во внешней памяти и при выполнении запросов к ней.

Атомарность значений атрибутов. Значения всех атрибутов являются атомарными. Это следует из определения домена как потенциального множества значений одного простого типа данных, т.е. среди значений домена не могут содержаться множества (т.е. одна ячейка таблицы – одно значение).

Реляционная модель данных (РМД) – совокупность связанных таблиц. Реляционную модель можно условно разделить на три части, описывающие разные аспекты реляционного подхода: *структурную, манипуляционную и целостную.*

В *структурной* части модели фиксируется, что единственной структурой данных, используемой в реляционных БД, является нормализованное N-арное отношение (таблица).

В *манипуляционной* части модели утверждаются два фундаментальных механизма манипулирования реляционной БД – реляционная алгебра и реляционное исчисление. Основной функцией манипуляционной части реляционной модели является обеспечение меры реляционности любого конкретного языка БД.

В *целостной* части РМД фиксируются два базовых требования целостности, которые должны поддерживаться в любой реляционной СУБД: *целостность сущностей* и *целостность по ссылкам.*

Требование *целостности сущностей* означает, что любой кортеж любого отношения отличим от любого другого кортежа *этого* отношения, другими словами, любое отношение должно иметь *первичный ключ*. Это требование автоматически удовлетворяется, если в системе не нарушаются базовые свойства отношений.

Требование *целостности по ссылкам* иначе называют требованием *внешнего ключа.*

Напомним, *первичный ключ* отношения (таблицы) – один или несколько атрибутов (полей), значения которых однозначно определяют кортеж (запись) таблицы. Связь между отношениями (таблицами) осуществляется путём включения общих полей. Поле *первичного ключа* одной (говорят, главной) таблицы присутствует в качестве обычного поля в связанной (подчинённой) таблице, его и называют *внешним ключом* по отношению к главной таблице, например, поле *Код должности* в таблице *Должности* – *первичный ключ*, такое же поле *Код должности* в таблице *Кадры* – *внешний ключ* (рис. 3). Требование *целостности по ссылкам* означает, что для каждого значения *внешнего ключа* подчиненного отношения должен быть кортеж с таким же значением *первичного ключа* главного отношения, либо значение *внешнего ключа* должно быть неопределённым (ни на что не указывать). Ограничения целостности должны поддерживаться СУБД.

Для соблюдения *целостности сущностей* достаточно гарантировать отсутствие в любом отношении кортежей с одним и тем же значением первичного ключа.

В обеспечении *целостности по ссылкам* существуют три подхода.

Первый подход заключается в том, что *запрещается* производить удаление кортежа, на который существуют ссылки, т.е. сначала нужно либо удалить ссылающиеся кортежи, либо изменить значения из внешнего ключа должным образом.

При *втором* подходе при *удалении* кортежа, на который имеются ссылки, во всех ссылающихся кортежах значение внешнего ключа автоматически становится неопределённым.

Третий подход (каскадное удаление) состоит в том, что при *удалении* кортежа из отношения, на которое есть ссылка, автоматически удаляются все ссылающиеся кортежи.

В современных СУБД (в том числе MS Access) обычно можно выбрать способ поддержания целостности по ссылкам для каждой отдельной ситуации (связи) определения внешнего ключа.

Нормализация базы данных

При проектировании реляционных БД большое внимание уделяется нормализации таблиц. В процессе нормализации обеспечивается защита целостности данных путём устранения их дублирования. В результате исходная таблица разбивается на две или более связанных таблиц, которые могут быть “собраны” вместе с помощью операции объединения.

Отобразим концептуальную модель ПО на логическую схему, ориентируясь на СУБД MS Access, получим фрагмент логической модели БД “Кадры” (рис. 3).

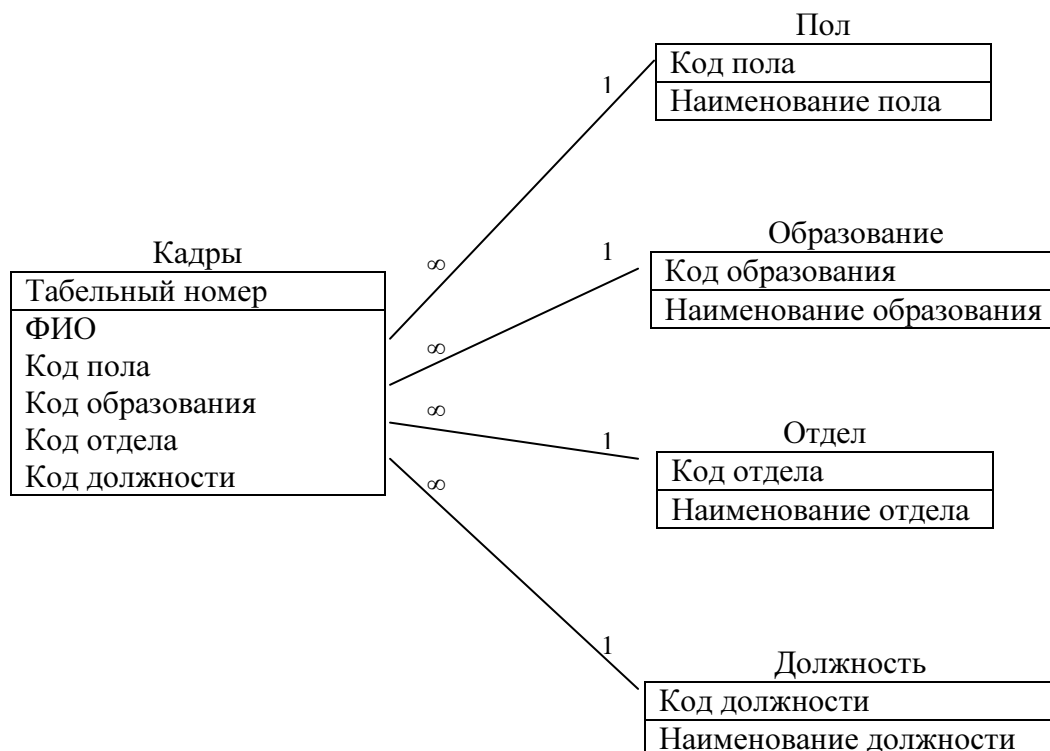


Рис 3. Логическая модель базы данных

Следующий этап – физическое проектирование: логическая модель данных отображается на физическую схему, в результате получается физическая модель, определяющая размещение данных, методы доступа и технику индексирования. Физическая модель соответствует внутреннему уровню архитектуры любой АИС. В современных СУБД (в том числе MS Access) процесс физического проектирования БД осуществляется автоматизированно средствами самой СУБД.

ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ

1. Понятие предметной области базы данных. Объект, атрибут ПО.
2. В чем заключается этап концептуального проектирования базы данных.
3. Что такое инфологическая модель, ее основные конструктивные элементы.
4. В чем заключается этап логического проектирования база данных.
5. Что такое процесс нормализации базы данных, для чего он нужен?

ЛИТЕРАТУРА

1. Михеева Е.В. Информационные технологии в профессиональной деятельности: учеб. пособие для студ. учреждений сред. проф. образования / Е.В. Михеева. – 10-е изд., испр. – М.: Издательский центр «Академия», 2012. – 384 с.
2. Михеева Е.В. Практикум по информационным технологиям в профессиональной деятельности: учеб. пособие для студ. учреждений сред. проф. образования / Е.В. Михеева. – 11-е изд., испр. – М.: Издательский центр «Академия» 2012. – 256 с.
3. Филимонова Е.В. Информационные технологии в профессиональной деятельности: учебник. / Е.В. Филимонова. – Изд-е 2-е, доп. и перераб. – Ростов н/Д: Феникс, 2008. – 381, [1] с. – (СПО).