

БАЗЫ ДАННЫХ

2 семестр

Лекция 1

ER – диаграммы. Модель "сущность-связь".

2017

Модель «Сущность — связь»

Одной из наиболее популярных средств формализованного представления предметной области систем, ориентированных на обработку фактографической информации, является модель «сущность — связь», которая положена в основу значительного количества коммерческих CASE-продуктов, поддерживающих полный цикл разработки систем баз данных или отдельные его стадии. При этом многие из них не только поддерживают стадию концептуального проектирования предметной области разрабатываемой системы, но и позволяют осуществить на основе построенной их средствами модели стадию логического проектирования путем автоматической генерации концептуальной схемы базы данных для выбранной СУБД, например, схемы базы данных для какого-либо SQL-сервера или объектной СУБД.

Моделирование предметной области в этом случае базируется на использовании графических диаграмм, включающих сравнительно небольшое число компонентов, и самое важное — *технология построения* таких диаграмм.

Семантическую основу ER-модели составляют следующие предположения:

- та часть реального мира (совокупность взаимосвязанных объектов), сведения о которых должны быть помещены в базу данных, может быть *представлена, как совокупность сущностей*;
- каждая сущность обладает характеристическими свойствами (атрибутами), отличающими ее от других сущностей и позволяющими ее *идентифицировать*;
- сущности можно классифицировать по типам сущностей: каждый экземпляр сущности (представляющий некоторый объект) может быть отнесен к классу — *типу сущностей*, каждый экземпляр которого обладает общими для них и отличающими их от сущностей других классов свойствами;
- систематизация представления, основанная на классах, в общем случае предполагает иерархическую зависимость типов: сущность типа *A* является *подтипом* сущности *B*, если каждый экземпляр типа *A* является экземпляром сущности типа *B*;
- взаимосвязи объектов могут быть представлены как *связи— сущности*, которые служат для фиксирования (представления) взаимозависимости двух или нескольких сущностей.

Здесь следует еще раз подчеркнуть информационную природу понятия *сущность* и его соотношение с материальными или воображаемыми объектами предметной области. Любой объект предметной области обладает свойствами, часть из которых выделяется как характеристические — значимые с точки зрения прикладной задачи. При этом, например, в процессе анализа и систематизации предметной области обычно выделяются *классы* — совокупности объектов, обладающих одинаковым набором свойств, задаваемых в виде *наборов атрибутов* (значения атрибутов для объектов одного класса, естественно, могут различаться). Соответственно, на уровне представления предметной области (т. е. ее инфологической модели) объекту, рассматриваемому как понятие (объект в сознании человека), соответствует понятие *сущность*; объекту, как части материального мира (и существующему независимо от сознания человека), соответствует понятие *экземпляр сущности*; классу объектов соответствует понятие *тип сущности*.

В дальнейшем, поскольку в инфологической модели рассматриваются не отдельные экземпляры объектов, а классы, мы не будем различать соответствующие понятия этих двух уровней, т. е. будем предполагать тождественность понятий *объект* и *сущность*, *свойство объекта* и *свойство сущности*.

ER-модель, как описание предметной области, должна определить объекты и взаимосвязи между ними, т. е. установить связи следующих двух типов.

1. Связи между объектами и наборами характеристических свойств, и таким образом определить сами объекты.

2. Связи между объектами, задающие характер и функциональную природу их взаимозависимости.

Как было отмечено ранее, ER-моделирование предметной области базируется на использовании графических диаграмм, как простого (привычного), наглядного и в то же время информативного и многоаспектного способа отображения компонентов проекта. Поэтому изложение основных положений ER-модели будет иллюстрироваться материалом примера ER-диаграммы, приведенного на рис. 9.1.

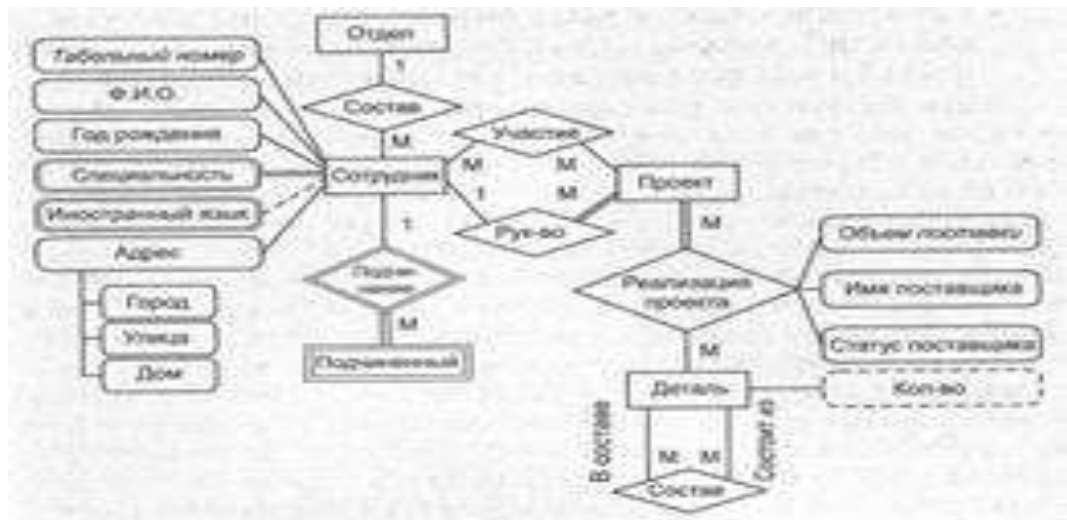


Рис. 12.1 Пример ER-модели

Сущность. Сущность, с помощью которой моделируется класс однотипных объектов, определяется в как «предмет, который может быть четко идентифицирован». Так же как каждый объект уникально характеризуется набором значений свойств, сущность должна *определяться* таким *набором атрибутов*, который позволял бы различать отдельные экземпляры сущности. Каждый экземпляр сущности должен быть отличим от любого другого экземпляра той же сущности (это требование аналогично требованию отсутствия кортежей-дубликатов в реляционных таблицах). Например, для однозначной идентификации каждого экземпляра сущности «Сотрудник» вводится атрибут «Табельный номер», который вследствие своей природы будет всегда иметь уникальное значение в рамках предприятия. То есть, уникальным идентификатором сущности может являться атрибут, комбинация атрибутов, комбинация связей или комбинация связей и атрибутов, однозначно отличающая любой экземпляр сущности от других экземпляров сущности того же типа.

Сущность имеет *имя*, уникальное в пределах модели. При этом *имя сущности* — это *имя типа*, а не некоторого конкретного экземпляра.

Сущности подразделяются на *сильные* и *слабые*. Сущность является слабой, если ее существование зависит от другой сущности сильной по отношению к ней. Например, сущность «Подчиненный» является слабой по отношению к сущности «Сотрудник»: если будет удалена запись, соответствующая некоторому сотруднику, имеющему подчиненных, то сведения о подчинении также должны быть удалены.

Свойства. Природа свойства, как *характер связи* свойства с сущностью (объектом), может быть различной. Рассмотрим основные виды свойств.

Свойство может быть *множественным* или *единичным* — т. е. атрибут, задающий свойство, может одновременно иметь несколько значений или, соответственно, только одно. Например, сотрудник может иметь несколько специальностей, но единственное значение — «Табельный номер».

Свойство может быть *простым* (не подлежащим дальнейшему делению с точки зрения прикладных задач) или *составным* — если его значение составляется из значений простых свойств. Например, свойство «Год рождения» является простым, а свойство «Адрес» — составным, так как включает значения простых свойств «Город», «Улица», «Дом».

Связи. Кроме связей между объектом и его свойствами, инфологическая модель отражает связи между объектами разных классов. *Связь* определяется как «ассоциация, объединяющая несколько сущностей». Эта ассоциация всегда может существовать между разными сущностями или между сущностью и ею же самой (рекурсивная связь).

Как и сущность, связь является *типовым* понятием, т. е. все экземпляры связываемых сущностей подчиняются правилам связывания типов. Принципиальность различия типов связей между типами и экземплярами иллюстрируется ER-диаграммами для типов и экземпляров, представленными на рис. 12.2.



Рис. 12.2 Пример ER-диаграмм для типов и экземпляров сущностей

Сущности, объединяемые связью, называются *участниками*. *Степень связи* определяется количеством участников связи.

Если каждый экземпляр сущности участвует, по крайней мере, в одном экземпляре связи, то такое участие этой сущности называется *полным* (или *обязательным*); в противном случае — *неполным* (или *необязательным*).

Количественный характер участия экземпляров сущностей (один или многие) задается *типом связи* (или *мощностью связи*). Возможны следующие типы: «один к одному» (1:1), «один ко многим» (1:М), «многие к одному» (М:1), «многие ко многим» (М:М).

Следует отметить, что инструмент связей — это средство представления *сложных объектов*, каждый из которых может рассматриваться как множество некоторым образом взаимосвязанных *простых объектов*. Деление на простые и сложные объекты, также как и характер взаимосвязи, является условным и определяется особенностями анализа предметной области, т. е. в конце концов — характером использования данных о предметах в решаемых прикладных задачах. При этом с точки зрения, например, конструктора, ДЕТАЛЬ является сложным объектом, а с точки зрения поставщика — простым.

Среди многих разновидностей взаимосвязей наиболее частыми являются такие отношения иерархического типа, как «часть — целое», «род — вид».

Отношение «часть — целое» используются для представления *составных объектов*. Например, МАШИНЫ состоят из УЗЛОВ, УЗЛЫ состоят из ДЕТАЛЕЙ. Здесь возможны как отношения «*один ко многим*», так и «*многие ко многим*».

Отношение «род — вид» — для представления *обобщенных объектов*. Например, СОТРУДНИКИ подразделяются по профессии на КОНСТРУКТОРОВ, ПРОГРАММИСТОВ, РАБОЧИХ; ПРОГРАММИСТЫ — на ПРИКЛАДНЫХ ПРОГРАММИСТОВ и СИСТЕМНЫХ ПРОГРАММИСТОВ. Иерархические отношения, и в частности — «родо-видовые», обычно используются как основа классификации объектов по наборам характеристических признаков. Причем «видовые» объекты *наследуют* свойства «родовых».

Другой широко используемой разновидностью взаимосвязи является агрегирование — объединение простых объектов в сложный по принципу их принадлежности *агрегату* или их совместного участия в некотором процессе. Агрегирование, рассматриваемое здесь как более общий случай иерархических отношений, объединяет объекты разной природы с единственным общим свойством «совместное участие». Агрегированные объекты именуются обычно отглагольными существительными, например, «*Состав*»: ПОДРАЗДЕЛЕНИЕ *состоит из* СОТРУДНИКОВ; «*Поставка*»: ПОСТАВЩИК *поставляет* ДЕТАЛИ.

Супертипы и подтипы. Сущность может быть расщеплена на два или более взаимоисключающих *подтипов*, каждый из которых включает общие атрибуты и/или связи. Эти общие атрибуты и/или связи явно определяются один раз на более высоком уровне. В подтипах могут определяться собственные атрибуты и/или связи. В принципе выделение подтипов может продолжаться на более низких уровнях, но в большинстве случаев оказывается достаточно двух-трех уровней.

Сущность, на основе которой определяются подтипы, называется *супертипом*. Подтипы должны образовывать полное множество, т. е. любой экземпляр супертипа должен относиться к некоторому подтипу. Иногда для полноты множества надо определять дополнительный подтип, например, **ПРОЧИЕ**.

Подтип наследует свойства и связи супертипа. Например, тип сущности **ПРОГРАММИСТ** является подтипом сущности **СОТРУДНИК**. Программисты обладают всеми свойствами сотрудников и участвуют во всех связях, однако обратные утверждения неверны.

Тип сущности, его подтипы, подтипы этих подтипов и т. д. образуют *иерархию типов сущности*, пример которой приведен на рис. 12.3



Схема 12.3 Пример иерархии типов сущностей.

ER-диаграмма

Как отмечалось ранее, одна из основных целей семантического моделирования состоит в том, чтобы результаты анализа предметной области были отражены в достаточно простом, наглядном, но в то же время формализованном и достаточно информативном виде.

В этом смысле ER-диаграмма является очень удачным решением. В ней сочетаются функциональный и информационный подходы, что позволяет представлять как совокупность выполняемых функций, так и отношения между элементами системы, задаваемые структурами данных. При этом графическая форма позволяет отобразить в компактном виде (за счет наглядных условных обозначений) типологию и свойства сущностей и связей, а формализмы, положенные в основу ER-диаграмм, позволяют использовать на следующем шаге проектирования логической структуры базы данных строгий аппарат нормализации.

Сущности. Каждый тип сущности в ER-диаграммах представляется в виде прямоугольника, содержащего имя сущности. В качестве имени обычно используются существительные (или обороты существительного) в единственном числе. Для отражения сущностей слабых типов используются прямоугольники, стороны которых рисуются двойными линиями. Например, в рассматриваемой далее ER-диаграмме, приведенной на рис. 5.4, ПОДЧИНЕННЫЙ — сущность слабого типа.

Свойства. Свойства служат для уточнения, идентификации, характеристики или выражения состояния сущности или связи. Свойства отображаются в виде эллипсов, содержащих имя свойства. Эллипс соединяется с соответствующей сущностью или связью линией.

Имена ключевых свойств подчеркиваются, например, свойство «Табельный номер» сущности СОТРУДНИК.

Контур эллипса рисуется двойной линией, если свойство многозначное, например, свойство «Специальность» сущности СОТРУДНИК.

Контур эллипса рисуется штриховой линией, если свойство производное, например, свойство «Кол-во» сущности ПОСТАВЩИК.

Эллипс соединяется пунктирной линией, если свойство условное, например, свойство «Иностранный язык» сущности СОТРУДНИК.

Если свойство составное, то составляющие его свойства отображаются другими эллипсами, соединенными с эллипсом составного, например, свойство «Адрес» сущности **СОТРУДНИК** состоит из простых свойств «Город», «Улица», «Дом».

Связи. Связь — это графически изображаемая ассоциация, устанавливаемая между сущностями. Каждый тип связи на ER-диаграмме отображается в виде ромба с именем связи внутри. В качестве имени обычно используются отглагольные существительные.

Стороны ромба рисуют двойными линиями, если это связь сущности слабого типа с сущностью, от которой она зависит. Например, связь «Подчинение», связывающая сущность слабого типа **ПОДЧИНЕННЫЙ** с сущностью **СОТРУДНИК**, от которой она зависит.

Участники связи соединены со связью линиями. Двойная линия обозначает полное участие сущности в связи с данной стороны. Например, связь «Подчинение» со стороны сущности **ПОДЧИНЕННЫЙ**.

Связь может быть модифицирована указанием роли. Например, для рекурсивной связи «Состав» указаны роли: «Деталь *состоит из ...*» и «Деталь *входит в состав ...*».

Тип связи указывается индексами «1» или «М» над соответствующей линией. Например, связь «Руководство» имеет тип «один ко многим»: один сотрудник может руководить многими проектами; связь «Участие» имеет тип «многие ко многим»: один сотрудник может участвовать во многих проектах, и в проекте могут участвовать многие сотрудники.