

ЛЕКЦИЯ №3. ПРЕДСТАВЛЕНИЕ ПРЕДМЕТНОЙ ОБЛАСТИ. МЕТОДЫ ПРЕДСТАВЛЕНИЯ ПРЕДМЕТНОЙ ОБЛАСТИ. МОДЕЛЬ СУЩНОСТЬ-СВЯЗЬ. ИНФОЛОГИЧЕСКОЕ СПИСАНИЕ ПРЕДМЕТНОЙ ОБЛАСТИ.

3.1 Инфологическое моделирование	1
3.2 Построение модели “ОБЪЕКТ –СВОЙСТВО–ОТНОШЕНИЕ”	2
3.3 Семантическое моделирование данных, ER-диаграммы.....	4
3.4 Три типа бинарных связей	6
3.5 Более сложные элементы ER-модели	7
3.6 Диаграммы сущность-связь	8

3.1 Инфологическое моделирование

3.3.1 Общие сведения об инфологическом моделировании

В базе данных отображается какая-то часть реального мира. Естественно, что полнота ее описания будет зависеть от целей создаваемой информационной системы.

Для того чтобы база данных адекватно отражала предметную область, проектировщик базы данных должен хорошо представлять себе все нюансы, присущие данной предметной области (ПО), и уметь отобразить их в базе данных. Предметная область должна быть предварительно описана. Для этого в принципе может использоваться и естественный язык, но его применение имеет много недостатков, основными из них являются громоздкость описания и неоднозначность его трактовки. Поэтому обычно для этих целей используют искусственные формализованные языковые средства. В связи с этим под *инфологической моделью* (ИЛМ) понимают описание предметной области, выполненное с использованием специальных языковых средств, не зависящих от используемых в дальнейшем программных средств.

Инфологическая модель должна строиться вне зависимости от того, будете ли вы в дальнейшем использовать какую-либо СУБД или пользоваться другими программными средствами для реализации своей информационной системы.

3.3.2 Требования, предъявляемые к инфологической модели

Основным требованием к ИЛМ, вытекающим из ее назначения, является требование адекватного отображения предметной области. ИЛМ должна быть непротиворечивой.

Несмотря на то, что реальный мир, отображаемый в ИЛМ, является по своей природе бесконечным, инфологическая модель является конечной, что обеспечивается четким ограничением предметной области. ИЛМ должна в связи с этим обладать свойством легкой расширяемости, обеспечивающим ввод новых данных без изменения ранее определенных. То же самое можно сказать и об удалении данных. В связи с большой размерностью реальных инфологических моделей должна обеспечиваться возможность композиции и декомпозиции модели.

Инфологическая модель должна легко восприниматься разными категориями пользователей. Желательно, чтобы ИЛМ строил специалист, работающий в этой предметной области, а не проектировщик систем машинной обработки данных или хотя бы проверить сделанное описание, чтобы убедиться, что специфика предметной области воспринята правильно. Инфологическая модель должна также легко и однозначно восприниматься всеми специалистами, которые в дальнейшем участвуют в процессе проектирования баз данных и программного обеспечения.

Она является ядром системы проектирования. ИЛМ содержит необходимую и достаточную информацию для дальнейшего проектирования автоматизированной системы обработки информации.

3.3.3 Компоненты инфологической модели

Инфологическая модель предметной области включает в себя ряд компонентов (рис. 1). Центральной компонентой инфологической модели является описание объектов предметной области и связей между ними (ER-модель).

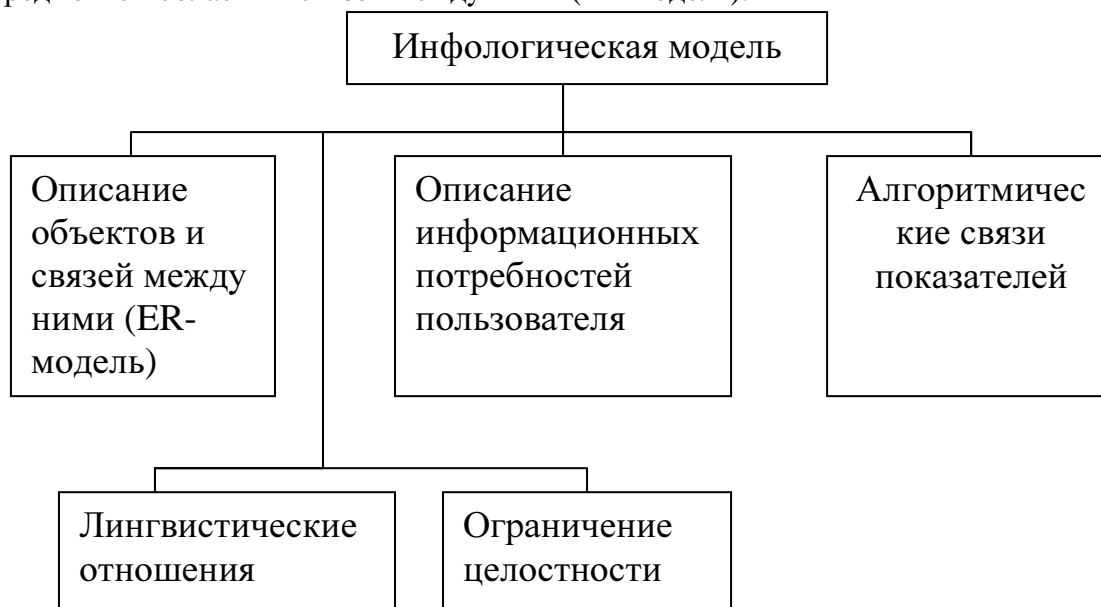


Рисунок 1 Компоненты инфологической модели

3.2 Построение модели “ОБЪЕКТ –СВОЙСТВО–ОТНОШЕНИЕ”

Для описания ИЛМ используются как языки аналитического (описательного) типа, так и графические средства в дальнейшем применяется графический способ отображения модели “объект—свойство—отношение”. В предметной области в процессе ее обследования и анализа выделяют классы объектов. Классом объектов называют совокупность объектов, обладающих одинаковым набором свойств. Например, если в качестве предметной области рассмотреть вуз, то в ней можно выделить следующие классы объектов: учащиеся, преподаватели, аудитории и т. д. Объекты могут быть реальными, как названные выше, а могут быть и абстрактными, как, например, предметы, которые изучают студенты.

При отражении в информационной системе каждый объект представляется своим идентификатором, который отличает один объект класса от другого, а каждый класс объектов представляется именем этого класса. Так, для объектов класса “ИЗУЧАЕМЫЕ ПРЕДМЕТЫ” идентификатором каждого объекта будет “НАЗВАНИЕ ПРЕДМЕТА”. Идентификатор должен быть уникальным.

Каждый объект обладает определенным набором свойств. Для объектов одного класса набор этих свойств одинаков, а их значения, естественно, могут различаться. Например, для объектов класса “СТУДЕНТ” таким набором свойств, описывающим объекты класса, может быть “ГОД РОЖДЕНИЯ”, “ПОЛ” и др.

При описании предметной области надо изобразить каждый из существующих классов объектов и набор свойств, фиксируемый для объектов данного класса.

Будем использовать для отображения объектов и их свойств следующие обозначения (рис. 2).

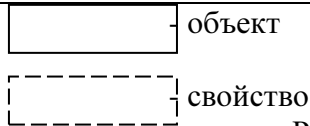


Рисунок 2 Обозначение объектов и их свойств

Каждому классу объектов в инфологической модели присваивается уникальное имя.

При построении инфологической модели желательно дать словесную интерпретацию каждой сущности, особенно если возможно неоднозначное толкование понятия.



Рисунок 3 Изображение связи «объект - свойство»

При описании предметной области надо отразить связи между объектом и характеризующими его свойствами. Это изображается просто в виде линии, соединяющей обозначение объекта и его свойств.

Связь между объектом и его свойством может быть различной. Объект может обладать только одним значением какого-то свойства. Например, каждый человек может иметь только одну дату рождения. Назовем такие свойства единичными. Для других свойств возможно существование одновременно нескольких значений у одного объекта. Пусть, например, при описании “СОТРУДНИКА” фиксируется в качестве его свойства “ИНОСТРАННЫЙ ЯЗЫК”, которым он владеет. Так как сотрудник может знать несколько иностранных языков, то такое свойство будем называть множественным. При изображении связи между объектом и его свойствами для единичных свойств будем использовать одинарную стрелку, а для множественных свойств — двойную.

Кроме того, некоторые свойства являются постоянными, их значение не может измениться с течением времени. Назовем такие свойства статическими, а те свойства, значение которых может изменяться со временем, будем называть динамическими.

Другой характеристикой связи между объектом и его свойством является признак того, присутствует ли это свойство у всех объектов данного класса либо отсутствует у некоторых объектов. Например, для отдельных служащих может иметь место свойство “УЧЕНАЯ СТЕПЕНЬ”, а другие объекты этого класса могут не обладать, указанным свойством. Назовем такие свойства условными.

В инфологической модели отображаются не отдельные экземпляры объектов, а классы объектов. Когда в ИЛМ изображено обозначение объекта, то ясно, что речь идет о классе объектов, обладающих описанными свойствами. Поэтому в инфологическую модель в большинстве случаев можно в явном виде не вводить еще и обозначение для класса объектов. Явное изображение класса объектов необходимо только в том случае, если в ПО для данного класса объектов фиксируются не только характеристики, относящиеся к отдельным объектам этого класса, но и какие-то интегральные характеристики, относящиеся ко всему классу в целом. Например, если для класса объектов “СОТРУДНИКИ ПРЕДПРИЯТИЯ” фиксируется не только возраст каждого из сотрудников, но и средний возраст всех сотрудников, то в инфологической модели необходимо отразить не только объект “СОТРУДНИК”, но и класс объектов “СОТРУДНИКИ”. Для отображения класса объектов можно использовать какое-то специфическое обозначение или такое же, которое используется для объектов.

3.3 Семантическое моделирование данных, ER-диаграммы

Широкое распространение реляционных СУБД и их использование в самых разнообразных приложениях показывает, что реляционная модель данных достаточна для моделирования предметных областей. Однако проектирование реляционной базы данных в терминах отношений на основе кратко рассмотренного нами механизма нормализации часто представляет собой очень сложный и неудобный для проектировщика процесс.

При этом проявляется ограниченность реляционной модели данных в следующих аспектах:

- Модель не предоставляет достаточных средств для представления смысла данных. Семантика реальной предметной области должна независимым от модели способом представляться в голове проектировщика. В частности, это относится к упоминавшейся нами проблеме представления ограничений целостности.
- Для многих приложений трудно моделировать предметную область на основе плоских таблиц. В ряде случаев на самой начальной стадии проектирования проектировщику приходится производить насилие над собой, чтобы описать предметную область в виде одной (возможно, даже ненормализованной) таблицы.
- Хотя весь процесс проектирования происходит на основе учета зависимостей, реляционная модель не предоставляет каких-либо средств для представления этих зависимостей.
- Несмотря на то, что процесс проектирования начинается с выделения некоторых существенных для приложения объектов предметной области ("сущностей") и выявления связей между этими сущностями, реляционная модель данных не предлагает какого-либо аппарата для разделения сущностей и связей.

На использовании разновидностей ER-модели основано большинство современных подходов к проектированию баз данных (главным образом, реляционных). Модель была предложена Ченом (Chen) в 1976 г. Моделирование предметной области базируется на использовании графических диаграмм, включающих небольшое число разнородных компонентов. В связи с наглядностью представления концептуальных схем баз данных ER-модели получили широкое распространение в системах CASE, поддерживающих автоматизированное проектирование реляционных баз данных. Среди множества разновидностей ER-моделей одна из наиболее развитых применяется в системе CASE

ЛЕКЦИЯ №3 Представление предметной области. Методы представления предметной области. Модель сущность-связь. Инфологическое списание предметной области. фирмы ORACLE. Ее мы и рассмотрим. Более точно, мы сосредоточимся на структурной части этой модели.

Основными понятиями ER-модели являются сущность, связь и атрибут.

Сущность – это реальный или представляемый объект, информация о котором должна сохраняться и быть доступна. В диаграммах ER-модели сущность представляется в виде прямоугольника, содержащего имя сущности. При этом имя сущности - это имя типа, а не некоторого конкретного экземпляра этого типа. Для большей выразительности и лучшего понимания имя сущности может сопровождаться примерами конкретных объектов этого типа.

Ниже изображена сущность АЭРОПОРТ с примерными объектами Шереметьево и Хитроу:



Рисунок 4 Пример сущности.

Каждый экземпляр сущности должен быть отличим от любого другого экземпляра той же сущности (это требование в некотором роде аналогично требованию отсутствия кортежей-дубликатов в реляционных таблицах).

Связь – это графически изображаемая ассоциация, устанавливаемая между двумя сущностями. Эта ассоциация всегда является бинарной и может существовать между двумя разными сущностями или между сущностью и ей же самой (рекурсивная связь). В любой связи выделяются два конца (в соответствии с существующей парой связываемых сущностей), на каждом из которых указывается имя конца связи, степень конца связи (сколько экземпляров данной сущности связывается), обязательность связи (т.е. любой ли экземпляр данной сущности должен участвовать в данной связи).

Связь представляется в виде линии, связывающей две сущности или ведущей от сущности к ней же самой. При этом в месте "стыковки" связи с сущностью используются трехточечный вход в прямоугольник сущности, если для этой сущности в связи могут использоваться много (many) экземпляров сущности, и одноточечный вход, если в связи может участвовать только один экземпляр сущности. Обязательный конец связи изображается сплошной линией, а необязательный - прерывистой линией.

Как и сущность, связь – это типовое понятие, все экземпляры обеих пар связываемых сущностей подчиняются правилам связывания.

В изображенном ниже примере связь между сущностями БИЛЕТ и ПАССАЖИР связывает билеты и пассажиров. При том конец сущности с именем "для" позволяет связывать с одним пассажиром более одного билета, причем каждый билет должен быть связан с каким-либо пассажиром. Конец сущности с именем "имеет" означает, что каждый билет может принадлежать только одному пассажиру, причем пассажир не обязан иметь хотя бы один билет.

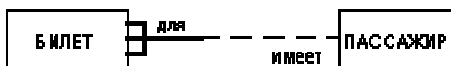


Рисунок 5 пример связи

Лаконичной устной трактовкой изображенной диаграммы является следующая:

- Каждый БИЛЕТ предназначен для одного и только одного ПАССАЖИРА;
- Каждый ПАССАЖИР может иметь один или более БИЛЕТОВ.

На следующем примере изображена рекурсивная связь, связывающая сущность ЧЕЛОВЕК с ней же самой. Конец связи с именем "сын" определяет тот факт, что у одного отца может быть более чем один сын. Конец связи с именем "отец" означает, что не у каждого человека могут быть сыновья.



Рисунок 6 Реккурсивная связь

Лаконичной устной трактовкой изображенной диаграммы является следующая:

- Каждый ЧЕЛОВЕК является сыном одного и только одного ЧЕЛОВЕКА;
- Каждый ЧЕЛОВЕК может являться отцом для одного или более ЛЮДЕЙ ("ЧЕЛОВЕКОВ").

Атрибутом сущности является любая деталь, которая служит для уточнения, идентификации, классификации, числовой характеристики или выражения состояния сущности. Имена атрибутов заносятся в прямоугольник, изображающий сущность, под именем сущности и изображаются малыми буквами, возможно, с примерами.

3.4 Три типа бинарных связей

На рис. 7 показаны три типа бинарных связей. В связи 1:1 («один к одному»)

одиночный экземпляр сущности одного типа связан с одиночным экземпляром сущности другого типа. На рис. 7, а связь СЛУЖЕБНЫЙ_АВТОМОБИЛЬ связывает одиночную сущность класса СОТРУДНИК с одиночной сущностью класса АВТОМОБИЛЬ. В соответствии с этой диаграммой, ни за одним сотрудником не закреплено более одного автомобиля, и ни один автомобиль не закреплен более чем за одним сотрудником.

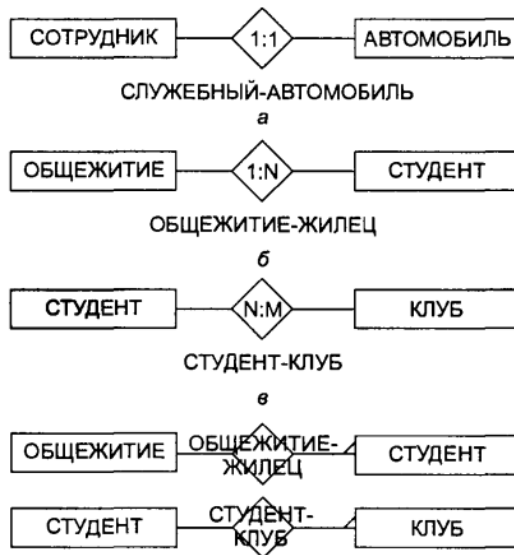


Рисунок 7 Три типа бинарных связей.

На рис. 7, б изображен второй тип связи, 1:N («один к N» или «один ко многим»), В этой связи, которая называется ОБЩЕЖИТИЕ-ЖИЛЕЦ, единичный экземпляр сущности класса ОБЩЕЖИТИЕ связан со многими экземплярами сущности класса СТУДЕНТ. В соответствии с этим рисунком, в общежитии проживает много студентов, но каждый студент живет только в одном общежитии.

Позиция, в которой стоят 1 и N, имеет значение. Единица стоит на той стороне связи, где располагается ОБЩЕЖИТИЕ, а N стоит на той стороне связи, где располагается СТУДЕНТ. Если бы 1 и N располагались наоборот, и связь записывалась бы как N:1, получилось бы, что в общежитии живет один студент, причем каждый студент живет в нескольких общежитиях. Это, разумеется, не так. На рис.7, в показан третий тип бинарной связи, N:M (читается «N к M» или «многие ко многим»). Эта связь называется СТУДЕНТ-КЛУБ, и она связывает экземпляры сущностей класса СТУДЕНТ с экземплярами сущностей класса КЛУБ. Один студент может быть членом нескольких клубов, а в одном клубе может состоять много студентов.

Числа внутри ромба, символизирующего связь, обозначают максимальное количество сущностей на каждой стороне связи. Эти ограничения называются максимальными кардинальными числами, а совокупность из двух таких ограничений для обеих сторон связи называется максимальной кардинальностью (maximum cardinality) связи. Например, о связи, изображенной на рис. 3.3, б, говорят, что она обладает максимальной кардинальностью 1:N. Кардинальные числа могут иметь и другие значения, а не только 1 и N. Например, связь между сущностями БАСКЕТБОЛЬНАЯ КОМАНДА и ИГРОК может иметь кардинальность 1:5, что говорит нам о том, что в баскетбольной команде может быть не более пяти игроков.

Связи трех типов, представленных на рис. 7, называются иногда связями типа «ИМЕЕТ», или связями обладания (HAS-A relationships). Такой термин используется потому, что одна сущность имеет (has) связь с другой сущностью. Например: сотрудник имеет автомобиль, студент имеет общежитие, клуб имеет студентов.

3.5 Более сложные элементы ER-модели

Мы остановились только на самых основных и наиболее очевидных понятиях ER-модели данных. К числу более сложных элементов модели относятся следующие:

- Подтипы и супертипы сущностей. Как в языках программирования с развитыми типовыми системами (например, в языках объектно-ориентированного программирования), вводится возможность наследования типа сущности, исходя из одного или нескольких супертипов. Интересные нюансы связаны с необходимостью графического изображения этого механизма.
- Связи "many-to-many". Иногда бывает необходимо связывать сущности таким образом, что с обоих концов связи могут присутствовать несколько экземпляров сущности (например, все члены кооператива сообща владеют имуществом кооператива). Для этого вводится разновидность связи "многие-ко-многими".
- Уточняемые степени связи. Иногда бывает полезно определить возможное количество экземпляров сущности, участвующих в данной связи (например, служащему разрешается участвовать не более, чем в трех проектах одновременно). Для выражения этого семантического ограничения разрешается указывать на конце связи ее максимальную или обязательную степень.
- Каскадные удаления экземпляров сущностей. Некоторые связи бывают настолько сильными (конечно, в случае связи "один-ко-многим"), что при удалении опорного экземпляра сущности (соответствующего концу связи "один") нужно удалить и все экземпляры сущности, соответствующие концу связи "многие". Соответствующее требование "каскадного удаления" можно сформулировать при определении сущности.
- Домены. Как и в случае реляционной модели данных бывает полезна возможность определения потенциально допустимого множества значений атрибута сущности (домена).

3.6 Диаграммы сущность-связь

Схемы, изображенные на рис. 3.3, называются диаграммами «сущность—связь», или ER-диаграммами (entity-relationship diagrams, ER-diagrams). Такие диаграммы стандартизированы, но не слишком жестко. В соответствии с этим стандартом, классы сущностей обозначаются прямоугольниками, связи обозначаются ромбами, а максимальное кардинальное число каждой связи указывается внутри ромба¹. Имя сущности указывается внутри прямоугольника, а имя связи указывается рядом с ромбом.

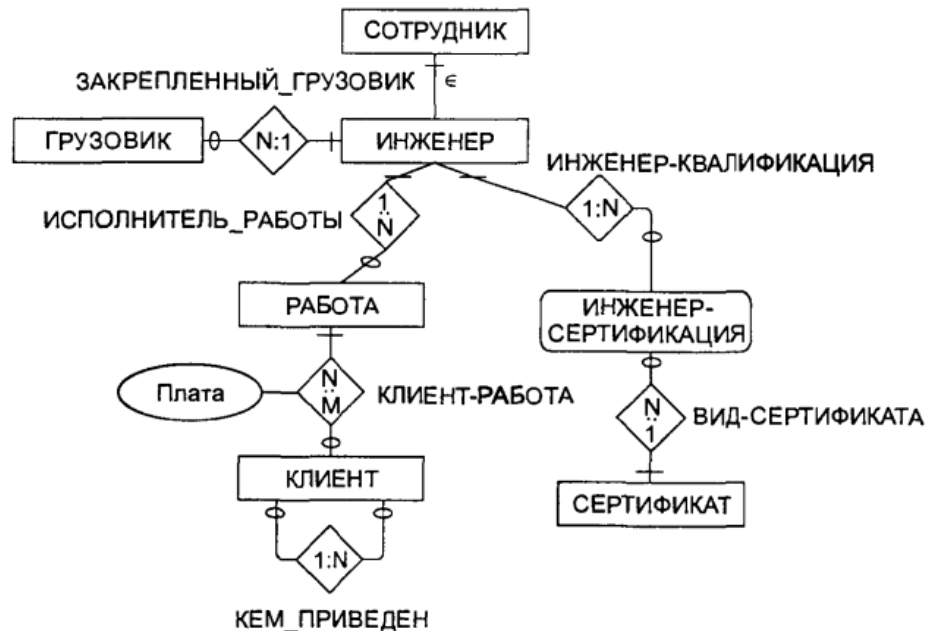


Рисунок 8 Пример диаграммы «сущность-связь»