* 1. Концептуальное проектирование
     1. Фундаментальные понятия

Для нормального функционирования информационной системы необходимо, чтобы концептуальная модель адекватно отображала реалии той предметной области, для которой она разрабатывается. Методологии, позволяющие эффективно отображать существующую смысловую содержательность реальности в конструкции модели, относятся к так называемым *семантическим* методологиям.

Наиболее популярной семантической моделью стала уже упоминавшаяся раннее модель "сущность — связь" (ER-модель), предложенная П. Ченом в 1976 году, которая с тех пор неоднократно усовершенствовалась самим Ченом и многими другими специалистами.

Главными элементами семантической модели данных являются *сущности, их атрибуты* и *типы связей.* Сущности часто представляют в виде *существительных,* а типы связей — в виде *глаголов.*

Семантическая модель предметной области изображается в виде диаграммы с учетом принятых обозначений для ее элементов {рис. 3.2).

* + 1. Сущности

Сущность — это то, о чем накапливается информация в информационной системе и что может быть однозначно идентифицировано.

*Сущность - тип* (в дальнейшем просто сущность) характеризуется независимым существованием и представляет множество объектов реального мира с одинаковыми свойствами. Отдельные объекты, которые входят в данный тип, называют экземплярами сущности.

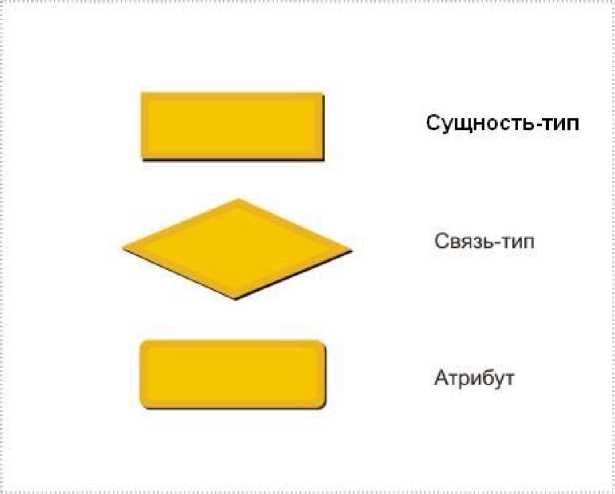


Рис. 3.2. Обозначения элементов диаграммы

Каждая сущность имеет имя и изображается на диаграммах в виде прямоугольника, а экземпляр сущности — в виде точки в прямоугольнике данной сущности (рис. 3.3).

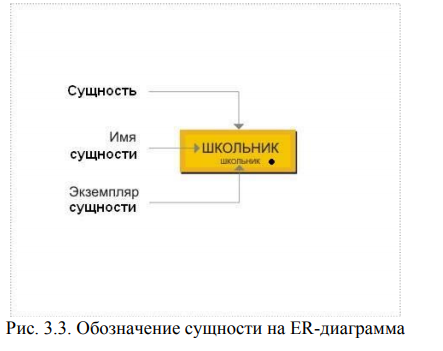


Рис. 3.3. Обозначение сущности на ER-диаграмма

* + 1. Атрибуты

*Атрибут —* это поименованная характеристика сущности, с помощью которой моделируется ее свойство. Каждой сущности присущи свои атрибуты. Например, сущность **ТОВАР** должна иметь такие атрибуты: Наименование\_товара, Индекс\_товара, Цена\_товара, Количество. На диаграммах атрибуты сущности соединяются с ней линиями (рис. 3.4).

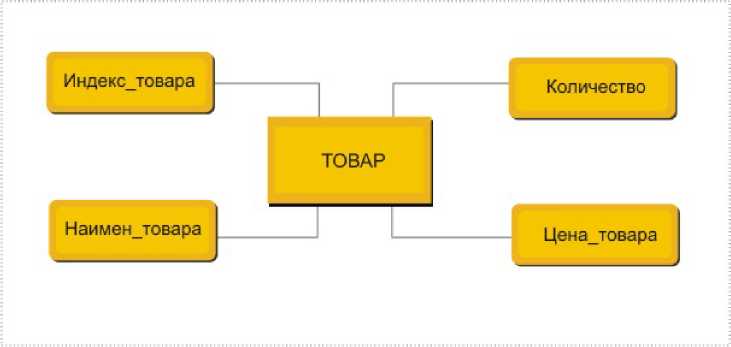


Рис. 3.4. Диаграмма представлений сущности ТОВАР и ее атрибутов

Значения каждого атрибута выбираются из соответствующего множества значений, включающего все потенциальные значения, которые могут быть присвоены атрибуту. Это множество значений называется *Доменом.*

Сущность и экземпляр сущности могут быть определены следующим образом: Сущность: **СТУДЕНТ (ФИО, Группа, Год\_рождения)**

Экземпляр сущности (Петров П.И., 93-ОА-22, 1992) .

Значения атрибутов могут часто меняться, в то время как описываемая ими сущность остается той же самой. Так, у экземпляра сущности **СТУДЕНТ** может измениться значение атрибута **ФИО,** но сама сущность останется той же.

1.1.4 Ключи

Среди атрибутов особое положение занимают такие, с помощью которых можно идентифицировать экземпляр сущности. Такие атрибуты называются *ключами.* Атрибут или несколько атрибутов, значения которых уникальным образом идентифицируют каждый экземпляр сущности, являются *потенциальным* ключом данной сущности. Потенциальных ключей может быть несколько. Например, экземпляр сущности **ФАКУЛЬТЕТ** (**Код факультета, Название\_факультета, ФИО\_декана)** может однозначно идентифицироваться любым из первых двух указанных атрибутов.

Один из потенциальных ключей может быть выбран в качестве *первичного* ключа. Обычно в качестве первичного ключа выбирается тот, который имеет наименьшую длину. Остальные потенциальные ключи называются *альтернативными.* Тот факт, что атрибут служит первичным ключом, отмечается его подчеркиванием.

Идентификацию некоторых сущностей иногда приходится осуществлять при помощи *составных* ключей, которые включают несколько атрибутов.

Например, сущность:

**ЛЕЧЕНИЕ (ФИО врача, ФИО пациента, Дата назначения, Лекарство)** ,

однозначно идентифицировать можно только составным ключом: **(ФИО врача, ФИО пациента, Дата назначения).**

1.1.5 Связи между сущностями

Две сущности могут быть связаны между собой. Подобная связь осуществляется через связь экземпляров одной сущности с экземплярами другой сущности, образуя набор экземпляров связи между двумя сущностями, который называется типом связи.

Каждому типу связи присваивается имя, которое должно представлять его функцию. Рассмотрим сущности **ПРЕПОДАВАТЕЛЬ** и **КУРС.** Между этими сущностями можно определить связь **ЧИТАЕТ,** сопоставив каждому преподавателю ту дисциплину, по которой он читает лекции, или, наоборот, каждой дисциплине — преподавателя. Связь **ЧИТАЕТ** составлена из множества пар, в каждой из которых преподаватель — из сущности **ПРЕПОДАВАТЕЛЬ**, а дисциплина — из сущности **КУРС** (рис. 3.5).

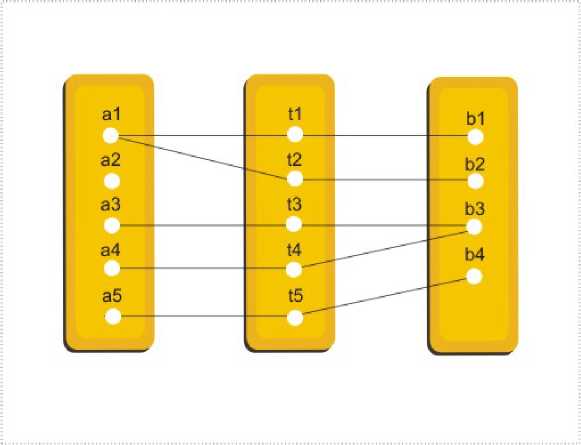


Рис. 3.5. Экземпляры типа связи ЧИТАЕТ

Полученная структура сама по себе является сущностью, состоящей из пар экземпляров, взятых из двух сущностей, связанных между собой. Сущность **ЧИТАЕТ**, полученная путем связи между сущностями **ПРЕПОДАВАТЕЛЬ** и **КУРС**, называется *составной сущностью.*

Описанная ситуация на диаграммах имеет свое графическое изображение, где тип связи обозначается в виде ромбика с указанным на нем именем связи, который соединен линиями со связываемыми

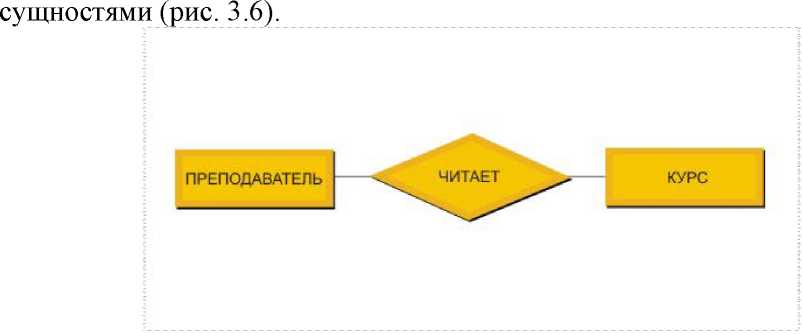


Рис. 3.6. Диаграмма типа связи **ЧИТАЕТ**

В связи могут участвовать не две, а большее количество сущностей, которые в данном случае являются участниками этой связи. Количество участников некоторой связи называется степенью связи.

В подавляющем числе случаев проектирования БД можно ограничиться рассмотрением бинарных связей.

**Мощность связи**

*Мощность* обозначает максимальное количество экземпляров одной сущности, связанных с одним экземпляром другой сущности. Например, если допустить, что у человека может быть только один супруг, то мощность связи ЖЕНАТЫ будет равна одному в каждом направлении (рис. 3.7).

Иногда помимо максимальной мощности полезно определять и минимальную мощность. В рассматриваемом примере не исключаются одинокие мужчины и женщины, поэтому минимальная мощность равна нулю в каждом направлении. Такая ситуация может быть обозначена

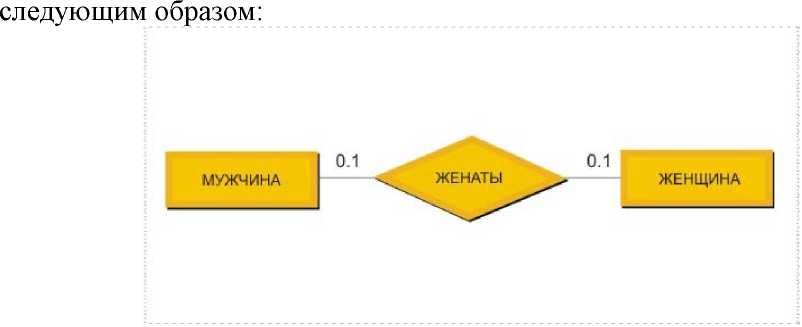


Рис. 3.7. Диаграмма связи **ЖЕНАТЫ** с указание минимальной и максимальной мощности

Некоторые связи не имеют конкретного значения максимальной мощности Например, преподаватель может читать не один курс, а, возможно, больше. Такую мощность обозначают: 1,\*, где 1 обозначает минимальную мощность. а \* обозначает "много" (существует и другой способ обозначения "много" — вместо \* ставится N). С другой стороны, если допустить, что каждый данный курс читается одним и только одним преподавателем, то мощность в обратном направлении будет 1,1.

**Показатель кардинальности**

Для того чтобы указать количество возможных связей для каждого экземпляра участвующего в связи сущности, используют *показатель кардинальности.*

Для бинарных связей показатель кардинальности может иметь следующие значения:

*"оДин к оДному"* (1 : 1), *"оДин ко многим"(1* : NJ*, "многие ко многим"* (М : N).

Если максимальная мощность связи в обоих направлениях равна одному, мы называем ее связью *"оДин к оДному"* (1 : 1).

Например, на факультете может быть один декан, и обратно, один и тот же декан может руководить только одним факультетом, что может быть обозначено следующим образом:

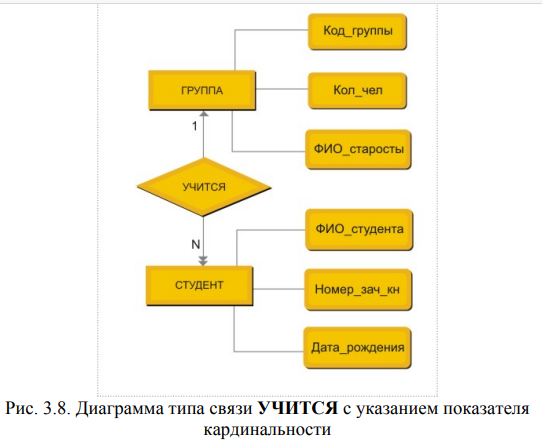
**ФАКУЛЬТЕТ <-----> ДЕКАН.**

Если максимальная мощность в одном направлении равна одному, а в другом — многим, то связь называется *"один ко многим"* (1 : N).

Например, в группе учится много студентов, но каждый студент учится только в одной группе:

**ГРУППА** < **>> СТУДЕНТ**.

Концептуальная модель подобной связи приведена на рис. 3.8.



На диаграмме использовано два способа обозначения вида бинарной связи: символическая (со стороны сущности **ГРУППА** выход связи помечен символом 1, а со стороны сущности **СТУДЕНТ —** символом N) и стрелками (в направлении, где максимальная мощность равна многим, проставлена двойная стрелка, а со стороны, где она равна единице — одинарная). Реально при построении диаграмм выбирают один из них.

И наконец, если максимальная мощность в обоих направлениях равна многим, то такая связь относится к типу *"многие ко многим"* (M : N). Например, преподаватель работает в разных группах, и в одной и той же группе работают различные преподаватели:

**ПРЕПОДАВАТЕЛЬ <<----->> ГРУППА**.

Связь между сущностями осуществляется посредством атрибутов. Например, рассмотрим две сущности:

Сущность: **СТУДЕНТ**

Атрибуты: **Номер зачетной книжки**

**ФИО студента**

Сущность: **ГРУППА**

Атрибуты: **Код группы**

**Количество студентов ФИО старосты**

Для их связи в число атрибутов сущности **СТУДЕНТ** необходимо добавить код группы, в которой он учится, и значение которого будет использовано для связи экземпляра одной сущности с экземпляром другой сущности.

1.1.6 Супертип и подтип

Для удовлетворения новых требований, выдвигаемых все более усложняющимися приложениями, в семантическое моделирование были введены дополнительные концепции, расширяющие его возможности. Дополнительные концепции базируются на таких понятиях, как супертип и подтип, а также используют процесс наследования атрибутов.

*Супертип —* это сущность, включающая разные подтипы, которые необходимо представить в модели данных.

*ПоДтип —* это сущность, являющаяся членом супертипа, но выполняющая отдельную роль в нем.

Супертип может иметь несколько разных подтипов. Так, например, подтипы: **АССИСТЕНТ, СТАРШИЙ ПРЕПОДАВАТЕЛЬ, ДОЦЕНТ, ПРОФЕССОР** являются членами супертипа **ПРЕПОДАВАТЕЛЬ.** Это означает, что каждый экземпляр подтипа является в то же время и экземпляром супертипа. Связь между супертипом и подтипом относится к типу "один к одному".

Использование понятий супертипа и подтипов позволяет при моделировании выделить для подтипа свои собственные атрибуты и атрибуты, наследуемые им от супертипа.

На диаграмме (рис. 3.9) подтипы соединяются линиями с кружком, который в свою очередь соединяется с супертипом. На каждой линии, идущей от подтипа, располагается U-образный символ, который обозначает направление включения. Верхняя часть **U** "открывается" в сторону супертипа. Внутри кружка располагается буква **D**, если подтипы не пересекаются, и буква О — для пересекающихся подтипов. В последнем случае экземпляр супертипа может быть членом сразу нескольких подтипов. Изображенная на диаграмме ситуация исключает пересечение подтипов, поэтому в кружок помещен символ **D**.

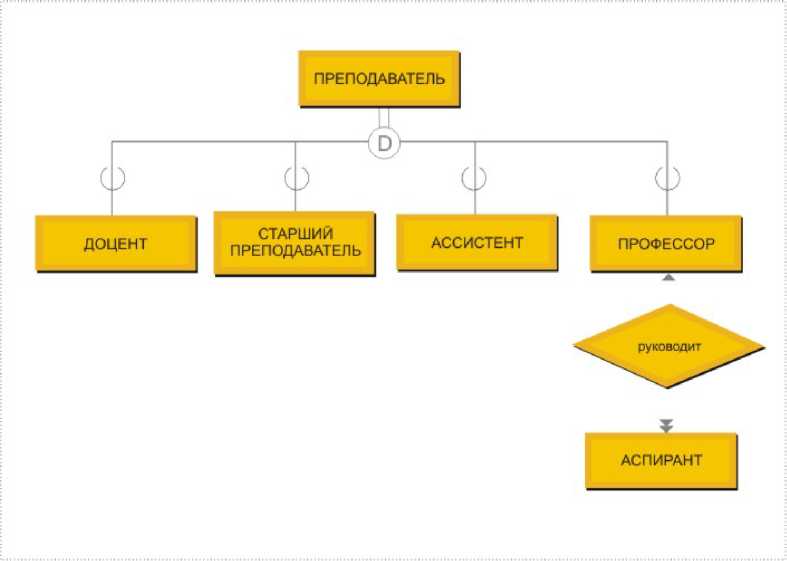


Рис. 3.9. Диаграмма с использованием понятий супертип и подтип

1.2 Пример моделирования локальной ПрО

С помощью рассмотренных выше понятий могут быть получены ER- модели для большинства схем баз данных в традиционных административно-управленческих приложениях. Если ПрО обширная, то построение ее концептуальной модели будет протекать более успешно, если эту ПрО разбить на несколько локальных предметных областей. Объем локальной ПрО выбирается таким образом, чтобы в нее входило не более 6-7 сущностей. Как ранее упоминалось, отправными элементами для построения ER-модели локальной ПрО очень часто являются используемые в организации документы.

Предположим, что определена локальная ПрО: поставка товаров на склад. Пусть используемая форма поставки имеет вид, как на рис. 3.10.

Покажем, как, используя приведенную форму, можно построить концептуальную модель этой небольшой локальной предметной области.

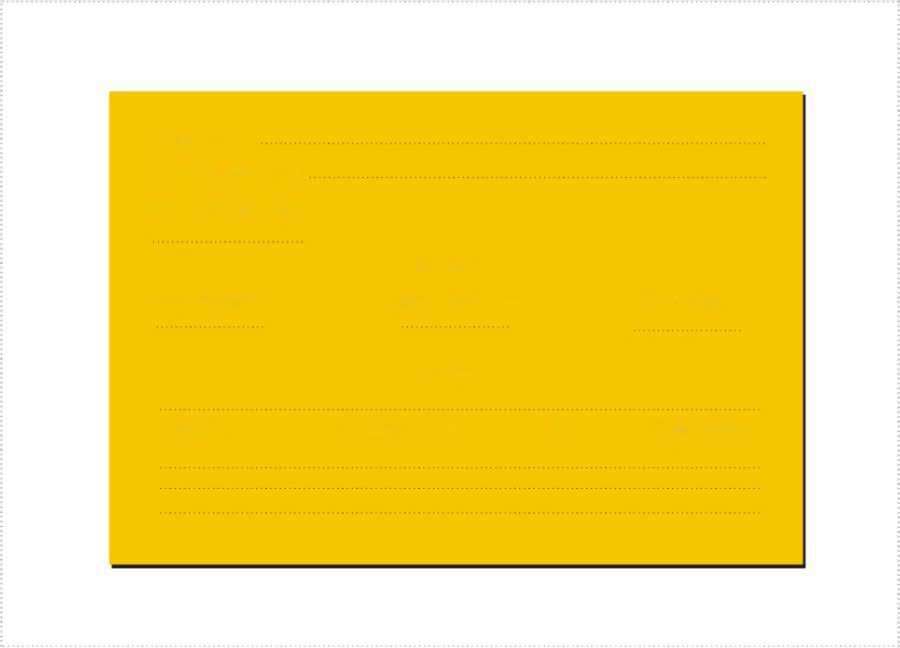


Рис. 3.10. Форма поставки

Поставщик

Адрес поставщика

Индекс поставщика

Поставка

Дата поставки Индекс поставки № склада

Товар

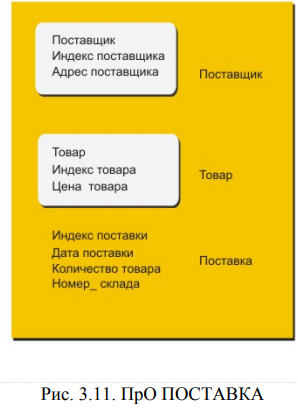
Индекс Наименование Цена Количество

Итак, анализируемая форма содержит следующую информацию: **Поставщик, Индекс поставщика, Адрес поставщика, Товар. Индекс товара, Цена товара. Количество товара, Поставка, Индекс поставщика, Дата поставки** и **Номер склада.**

Выделим две сущности: **ПОСТАВЩИК** и **ТОВАР** (рис. 3.11).

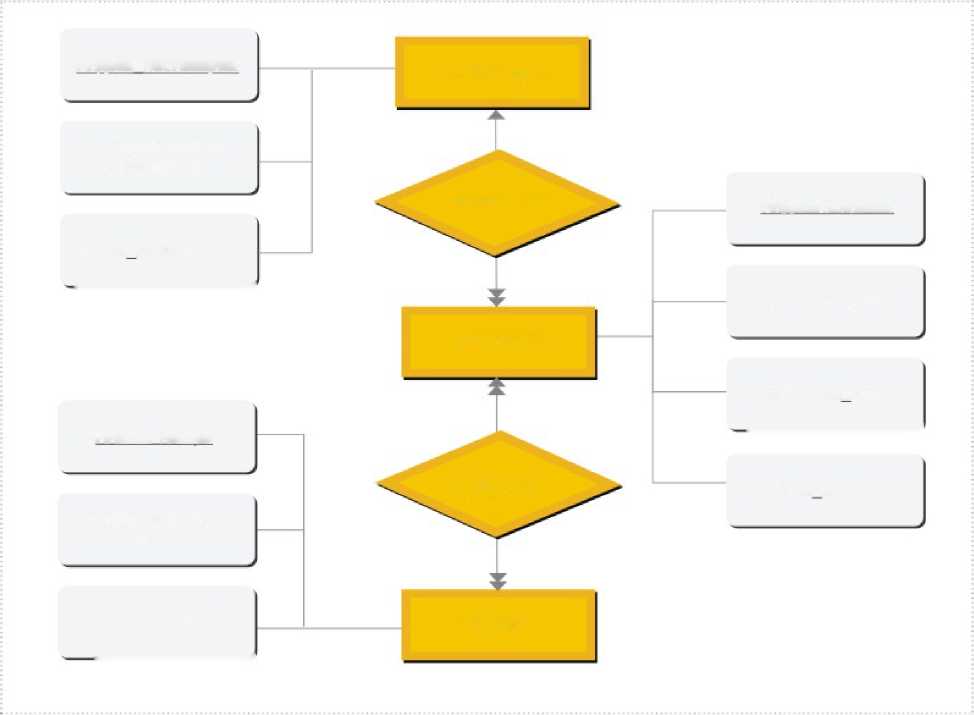
Оставшиеся атрибуты характеризуют сущность **— ПОСТАВКА.** Сформируем ее и установим определенные типы бинарных связей между тремя сущностями, исходя из следующих рассуждений: один и тот же поставщик может осуществить ряд поставок, но каждая поставка осуществляется только одним поставщиком.

Мощность связи между сущностями **ПОСТАВКА** и **ТОВАР** должна быть установлена M:N, так как каждая поставка может содержать несколько товаров, и один и тот же товар может содержаться в нескольких поставках. Исходя из вышесказанного, диаграмма модели предметной области **ПОСТАВКА** примет такой вид, как на рис. 3.12.



Атрибуты **Индекс поставщика, Индекс поставки** и **Индекс товара** были введены дли однозначной идентификации экземпляров рассматриваемых сущностей, так как ни один из остальных атрибутов не подходит на эту роль. Как уже упоминалось, такие идентификационные атрибуты называются первичными ключами.

При построении концептуальной модели следует избегать избыточности информации. После того, как выделены сущности, ключи, определяют и удаляют имеющиеся избыточные связи. Большое внимание уделяется анализу атрибутов. Забегая вперед, следует указать на то, что в хорошо спроектированной БД должно соблюдаться правило: среди атрибутов сущности должна наблюдаться зависимость описательного атрибута от ключевого, но не должна существовать зависимость между описательными атрибутами.



Индекс поставщика

поставляет

включает

Наименование поставщика

Адреспоставщика

*'*

Индекс товара

Наименование товара

Цена\_ товара

<

Индекс поставки

Количествотовара

>

Номсклада

Дата\_поставки

ТОВАР

ПОСТАВКА

ПОСТАВЩИК

Рис. 3.12. Диаграмма модели предметной области ПОСТАВКА

Завершающим этапом построения концептуальной модели исследуемой ПрО является спецификация всех сущностей, входящих в модель. Для данного примера результаты этого шага должны быть сведены к следующему:

1. Спецификация сущностей:

**ПОСТАВЩИК: Индекс поставщика—** идентификационный атрибут **Адрес\_поставщика** — описательный атрибут **Наименование\_поставщика** — описательный атрибут **ПОСТАВКА: Индекс поставки** — идентификационный атрибут **Количество\_товара—** описательный атрибут **Дата\_поставки** — описательный атрибут **Номер\_склада —** описательный атрибут

**ТОВАР: Индекс товара** — идентификационный атрибут **Наименование\_**товара — описательный атрибут **Цена\_товара —** описательный атрибут

1. Спецификация типов связей:

**ПОСТАВЛЯЕТ**: связь **ПОСТАВЩИК** <-—> **ПОСТАВКА** 1:N

**ВКЛЮЧАЕТ:** связь **ПОСТАВКА <** > **ТОВАР** M:N

1. Спецификация атрибутов:

**Индекс поставщика**: символьный, 6 символов

**Адрес\_поставщика**: символьный, 50 символов

**Цена\_товара**: денежный

Сформированные спецификации заносятся в словарь данных.

После создания моделей каждой выделенной предметной области производится объединение локальных концептуальных моделей в одну общую, как правило, довольно сложную схему.

Ниже представленны различные примеры ERR диаграммы изучите их..



