

Тема курса:

Манипулирование данными: хранение, вставка, редактирование и удаление данных, навигация по набору данных

Тема урока:

Дизайн методом «Value-Link» (ER-метод).
Реляционная алгебра

3. МОДЕЛЬ "СУЩНОСТЬ-СВЯЗЬ" (ER-МОДЕЛЬ)

- 3.1. Модель "сущность - связь". Компоненты модели.
- 3.2. Методы абстрагирования и операции над единицами информации в ER- модели.
- 3.3. Пример построения модели.

Семантические модели предназначены для отображения смысла информации, циркулирующей в предметной области. Любая развитая модель включает в себя три компонента:

- средства описания структуры информации (структурная часть),
- описание операций над данными (манипуляционная часть) и
- описание ограничений целостности.

Одной из наиболее популярных семантических моделей является модель "сущность - связи" или ER - модель (Entity - Relationship), на использовании которой основано большинство современных подходов к проектированию баз данных. Модель была предложена **Ченом (Chen)** в **1976 г.** и представляет собой ряд графических диаграмм, включающих небольшое число разнородных компонентов. Основными компонентами ER - модели являются сущность, связь и атрибут.

3.1. Компоненты модели

ПРЕДПРИЯТИЕ
Например,
Завод ВЭМ,
АО "ПензЭнерго"

В диаграммах ER-модели сущность представляется в виде прямоугольника, содержащего имя понятия. Иногда в прямоугольнике приводят примеры понятий.

Считается, что каждое понятие полностью описывается совокупностью его свойств. Таким образом, каждое понятие должно иметь следующие характеристики:

- имя,
- объем (количество сущностей, которые можно описать данным понятием),
- содержание (совокупность свойств).

Свойство (атрибут) - это элементарная единица структуры понятия, которая служит для уточнения, идентификации, классификации, числовой характеристики или выражения состояния сущности.

Например, свойства понятия СТУДЕНТ - код студента, фамилия, имя, № зачетки, группа, адрес, возраст, пол и т.п.

Набор свойств понятия в принципе бесконечен, но при разработке ЭИС он зависит от информационных потребностей пользователя и решаемых им задач. При составлении инфологической модели следует как можно более полно описать свойства понятий, учитывая не только текущие задачи, но и возможные будущие потребности пользователей ЭИС.

Свойства понятий делятся на три типа :

- Ключевые свойства** позволяют различить сущности внутри одного понятия (ключ может состоять из нескольких свойств).
- Дифференциальные свойства** содержат смысл понятия (то, что отличает его от других понятий).
- Валентные свойства** служат для связи между разными понятиями.

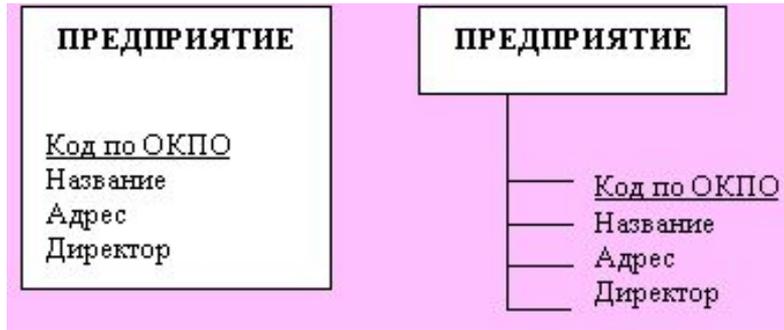
Например, "Код студента" - ключевое свойство, так как оно уникально для каждого студента;

Фамилия, Имя, Адрес - дифференциальные свойства (они отличают это понятие, например, от понятия ПРЕДПРИЯТИЕ;

Группа - валентное свойство, так как оно может применяться для связи с понятием СПЕЦИАЛЬНОСТЬ.

Для краткого описания понятия используется его схема, которая представляет собой совокупность имен свойств понятия: $P (S_1, S_2, \dots S_n)$. Ключевое свойство в схеме понятия выделяют подчеркиванием. **Например:** СТУДЕНТ (Код студента, ФИО, Адрес, Группа); СПЕЦИАЛЬНОСТЬ (Код, Название, Факультет, Группы).

На ER- диаграммах свойства понятий отображаются по-разному в зависимости от разновидности модели. **Например**, на рис. 3.1а имена свойств записаны малыми буквами внутри прямоугольника, изображающего понятие, под именем понятия. На рис. 3.1б свойства вынесены за прямоугольник понятия и связаны с ним линиями, ключевое свойство выделено.



а) б)

Рис.3.1 Варианты изображения свойств понятия

Связь - это ассоциация между двумя понятиями. Она всегда является бинарной и может существовать между двумя разными понятиями или между понятием и им же самим (рекурсивная связь). В любой связи выделяются два конца, на каждом из которых указывается имя конца связи, степень конца связи (сколько сущностей данного понятия связывается), обязательность связи (т.е. любая ли сущность данного понятия должна участвовать в данной связи).

На диаграмме связь изображается в виде линии между понятиями (рис.3.2). В месте присоединения связи к понятию используется "вилка" или стрелка, если связь может относиться к нескольким сущностям этого понятия. Одноточечный вход соответствует связи, в которой может участвовать только одна сущность из данного понятия. Обязательный конец связи изображается сплошной линией, а необязательный - прерывистой линией.

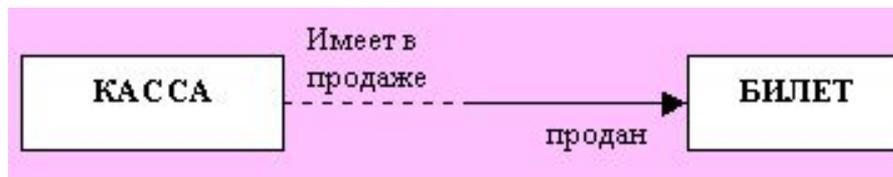


Рис.3.2. Изображение связи на ER- диаграмме.

Каждый БИЛЕТ продан одной конкретной КАССОЙ;
каждая КАССА может иметь в продаже 0, один или более билетов.

3.2. Характеристики связей и свойств

Важнейшими характеристиками связей между понятиями, а также между понятиями и их свойствами являются следующие:

- **время существования** - различают постоянные, долговременные и кратковременные связи.
Например, значение свойства РАЗМЕР_СТИПЕНДИИ для понятия СТУДЕНТ обновляется раз в семестр, а свойство НОМЕР_ЗАЧЕТКИ постоянно на протяжении обучения. Время существования влияет на то, как отображается свойство или связь в базе данных: будет ли она заложена в структуру БД, либо будет реализована алгоритмическим путем (в программе). Кратковременные, легко вычисляемые свойства и связи рекомендуется реализовывать алгоритмически.
- **избирательность** - различают необязательные, возможные, условные и обязательные связи.
Например, связь понятий СТУДЕНТ- СТИПЕНДИЯ зависит от успеваемости (условная связь), ЛИЧНОСТЬ - ИНОСТРАННЫЙ ЯЗЫК связаны необязательной связью, если предметная область - компрессорный завод и обязательной, если предметная область - лингвистический институт.
- **ассоциативность** (степень, мощность) - различают связи 1 : 1, 1 : M, N : M.

Связь 1 : 1 - это такая, при которой каждой сущности понятия А соответствует только одна сущность понятия В, **например:** СТУДЕНТ - АДРЕСНЫЕ ДАННЫЕ.

Связь 1 : М - это такая, при которой каждой сущности понятия А соответствует несколько (или 0) сущностей понятия В, **например:** ГРУППА -> СТУДЕНТ.

Связь N : М - это такая, при которой каждой сущности понятия А соответствует несколько (или 0) сущностей понятия В и наоборот, **например:** СТУДЕНТ <-> ПРЕПОДАВАТЕЛЬ.

3.3. Операции над единицами информации

Связи (отношения) между понятиями возникают в результате осмысления человеком задач реального мира. Основной механизм познания мира - это абстрагирование, которое состоит в выделении существенных и игнорировании не существенных свойств и связей. Оно позволяет разбить задачу на более простые подзадачи и упорядочивает наши знания о реальном мире.

Различают три основных **типа абстрагирования** и соответственно три вида **операций над понятиями**: агрегация, обобщение, ассоциация. Соответственно обратные логические операции над понятиями - это декомпозиция, специализация и индивидуализация.

Агрегация - это метод абстрагирования, при котором понятие - агрегат связано с другими понятиями, как целое связано с частями. **Декомпозиция** - это деление сложного понятия на компоненты.

Свойства агрегата - это совокупность свойств компонентов. Обычно составные части агрегата имеют совершенно разный набор свойств. **Например**, понятия ОБОРУДОВАНИЕ, АУДИТОРИЯ, СОТРУДНИКИ, ДИСЦИПЛИНЫ являются компонентами понятия КАФЕДРА. КАФЕДРА представляет собой понятие - агрегат.

На ER- диаграмме компоненты и агрегат изображаются прямоугольниками. Агрегат связан с компонентами связями типа "имеет / входит в" (рис. 3.3).

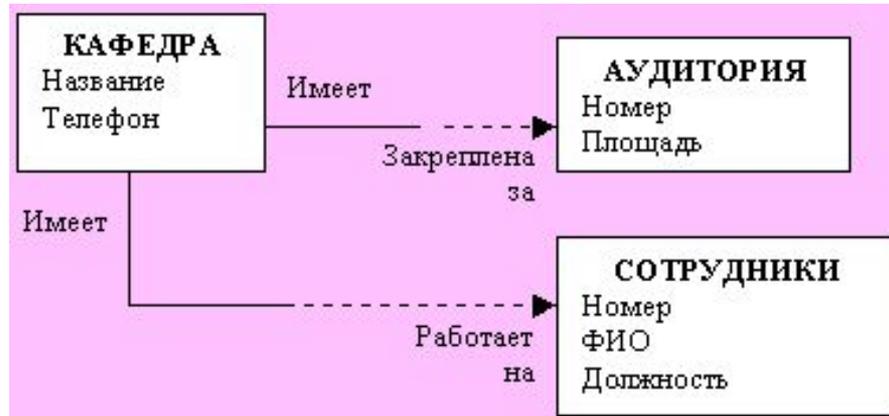


Рис.3.3 Графическое представление агрегата. Агрегат КАФЕДРА состоит из двух компонентов: АУДИТОРИЯ и СОТРУДНИКИ.

Обобщение - это метод абстрагирования, при котором обобщенное понятие связано с другими понятиями отношением "род - вид", **специализация** - операция обратная обобщению. **Например**, обобщенное понятие - УЧАЩИЕСЯ может быть разделено на категории СТУДЕНТЫ, ШКОЛЬНИКИ, АСПИРАНТЫ.

Категории обычно имеют много общих свойств, но некоторые свойства у них различны. Свойства обобщенного понятия представляют собой перечень общих свойств категорий, но могут включать и дополнительные свойства по сравнению с исходными, чаще всего это такие свойства, как Среднее значение, Сумма и т.п. На ER - диаграмме категории изображаются треугольниками, а обобщенное понятие - прямоугольником, который соединен с категориями связями типа "является видом...(is kind of..)" (рис. 3.4).

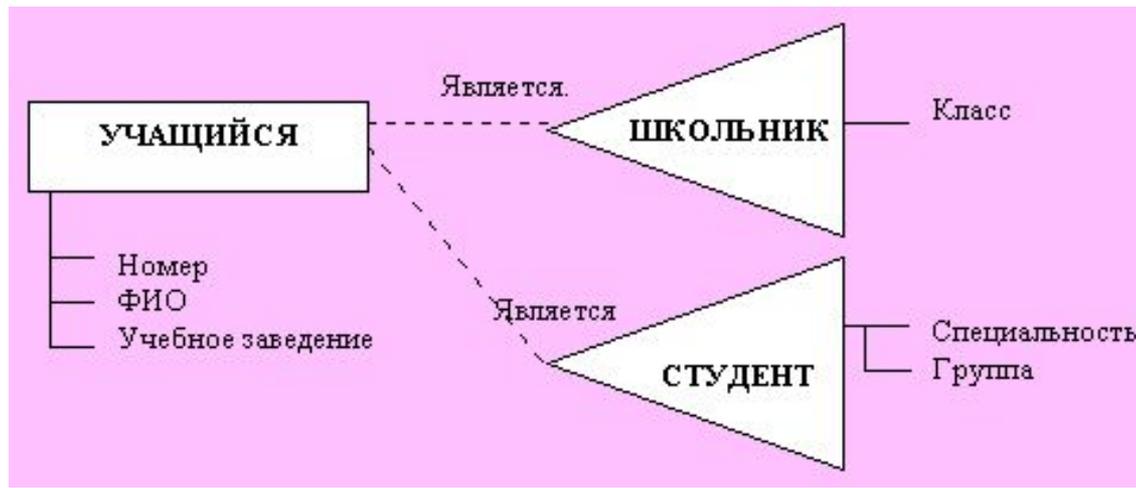


Рис. 3.4. Графическое представление обобщенного понятия

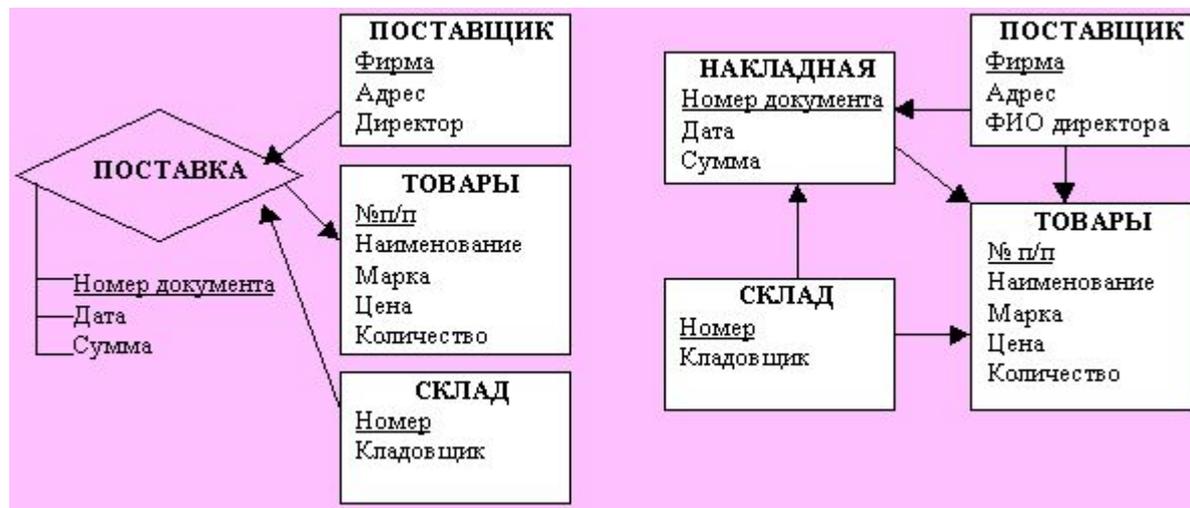
Ассоциация - это такая логическая операция, которая устанавливает связи между конкретными сущностями разных понятий.

Индивидуализация - это независимое рассмотрение связанных ассоциацией понятий. Обычно ассоциативные связи возникают в ходе реальных экономических процессов (бизнес - процессов).

Например - понятия **СТУДЕНТ**, **ПРЕПОДАВАТЕЛЬ**, **ДИСЦИПЛИНА** ассоциируются с понятием **ЧТЕНИЕ ЛЕКЦИЙ** или **СДАЧА ЭКЗАМЕНА**. Понятия **ТОВАР**, **МАГАЗИН**, **ПРОДАВЕЦ** связаны отношением **ПРОДАЖИ**. Чаще всего свойства ассоциации соответствуют реквизитам экономического документа, в котором отражается результат соответствующего бизнес - процесса, **например**, свойствами понятия **ПРОДАЖИ** могут быть такие реквизиты товарного и кассового чека, как дата продажи, номер документа, сумма.

На ER-диаграмме ассоциация может отображаться по-разному (рис. 3.5):

- В виде понятия, которое изображается ромбом, который соединен линиями со связанными понятиями.
- В виде множества бинарных связей между простыми понятиями.



а) б)

Рис. 3.5. Графическое представление ассоциации

Вариант А применяется обычно в двух случаях: если ассоциация соответствует процессу, который сопровождается составлением документов, реквизиты которых не отображены в виде свойств других понятий, либо если процесс соединяет более двух понятий.

ER - модели широко применяются как при ручном, так и при автоматизированном проектировании баз данных. При этом в некоторых системах, например, в системе IDEF, отсутствуют сложные понятия (агрегаты, ассоциации, обобщенные понятия), они заменяются связями. В разных системах применяются различные обозначения мощности связей и классов членства (обязательности связи) и т. д.

В заключение приведем рекомендуемый порядок построения ER-модели:

1. В каждом внешнем представлении выявить понятия и их свойства, исходя из анализа экономических документов. Их должно быть не более 6-7 в каждом представлении.
2. Обозначить понятия именами, которые должны быть краткими, отражать смысл и быть привычными для пользователя.
3. Выбрать ключевое свойство (свойства). Если такого нет, то вводят условное свойство "Порядковый №".

4. Выявить связи между понятиями и сущностями.

- a. Найти общие свойства понятий (обобщение);
- b. Рассмотреть все пары понятий и определить ассоциативные связи;
- c. Определить степень каждой связи.
- d. Записать формулы расчета, связывающие вычисляемые свойства.

5. Объединить модели, построенные для разных внешних представлений. Объединение ведут по 2-4 модели за один шаг, получая более сложные объекты и устраняя противоречия (например, разные наименования одних свойств и одинаковые - для разных понятий).

Задача построения модели ПО осложняется тем обстоятельством, что деление информации на свойства, понятия и связи условно, и один и тот же элемент информации можно описать по-разному. **Например**, информацию о преподавателе можно рассматривать как понятие, а можно и как свойство понятия ДИСЦИПЛИНА; любое ассоциированное понятие можно рассматривать как связь между несколькими понятиями. Принятие конкретного решения зависит от специфики предметной области и применяемых средств автоматизации проектирования.

3.4. Пример построения ER- модели

В качестве предметной области рассмотрим видеотеку (см. раздел 2. 2). Детализируя диаграмму потоков данных, для каждого из 4-х основных процессов мы получили 4 хранилища данных: Члены видеотеки, Фильмы, Аренда и Поставщики. В ER- модели им будут соответствовать 4 одноименных понятия: Определим структуру каждого понятия, принимая во внимание реквизиты документов, которые обрабатываются в видеотеке: учетные карточки членов видеотеки, карточки учета фильмов, справочник поставщиков фильмов и журнал учета аренды фильмов. Свойства понятий сведем в таблицу:

Понятие	Свойства	Понятие	Свойства
Члены видеотеки	ФИО Адрес Телефон Интересы	Фильмы	№ по каталогу Название Тип (тематика) Страна-производитель Год выпуска Анонс Цена Количество лент
Поставщики	Фирма Адрес Телефон Расчетный счет К кому обращаться	Аренда	№ п/п Дата выдачи Отметка о возврате Фильм ФИО клиента

Рассмотрим ассоциативные связи между понятиями, которые соответствуют информационным процессам.

А. В процессе учета членов видеотеки при регистрации нового члена необходимо убедиться, не был ли он записан раньше, не является ли должником. При исключении из членов следует проверить, вернул ли клиент все арендованные фильмы. Данный процесс использует информацию из понятий Члены и Аренда (рис. 3. 6а). Линия со стрелкой отображает связь между понятиями типа "один ко многим", например, каждому члену видеотеки может соответствовать много записей в журнале аренды, но каждой записи в журнале соответствует только один конкретный клиент. При регистрации каждого нового клиента нужно проверить все карточки о членах видеотеки, поэтому на диаграмме показана рекурсивная связь.

В. При выдаче фильмов в аренду регистратор должен проверить, является ли клиент членом видеотеки, не числятся ли за ним несданные фильмы, затем в журнал аренды заносятся сведения об аренде фильма, а количество имеющихся лент уменьшается на 1. При возврате фильмов следует в журнале внести пометку о возврате, а количество имеющихся в наличии лент данного фильма увеличить на 1. Кроме того, регистратор может предложить клиентам справки о имеющихся в наличии фильмах. Соответствующая диаграмма показана на рис. 3.6б.

С. Процесс "Учет поставщиков" соединяет понятия Фильмы и Поставщики (рис. 3.6в) с помощью ассоциированного понятия "Поставки".

Д. Управление фильмотекой заключается в периодическом предоставлении руководству отчетов об имеющихся фильмах и о ходе их аренды, поэтому данный процесс соединяет понятия "Фильмы" и "Аренда" (рис. 3.6г). Содержание отчетов отображает простое понятие "Отчеты начальству".

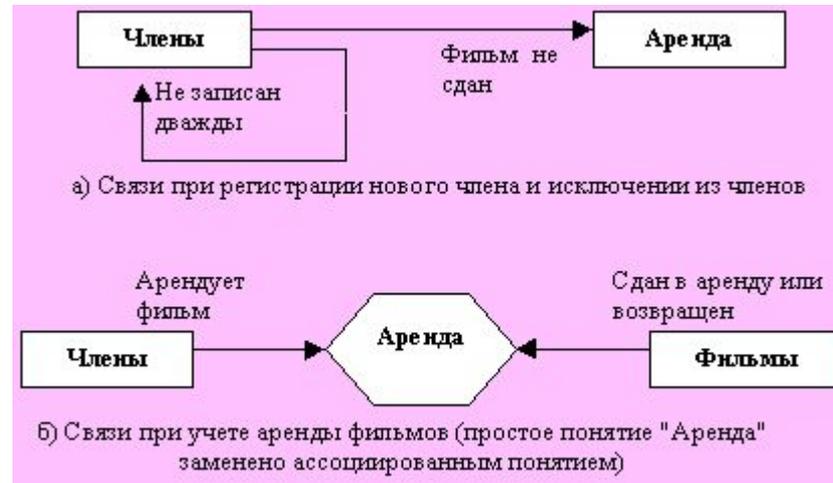




Рис.3.6. Построение фрагментов ER-модели при анализе информационных процессов

Обратите внимание, что ассоциированные понятия, как правило, имеют временное свойство (Дата поставки, Дата регистрации и т.п.) и ключевое свойство типа "Номер документа", так как они соответствуют информационным процессам, протекающим во времени.

В заключение объединим модели, полученные при анализе разных внешних представлений, устранив двусмысленности, дублирование атрибутов и рекурсивную связь (рис.3.7).

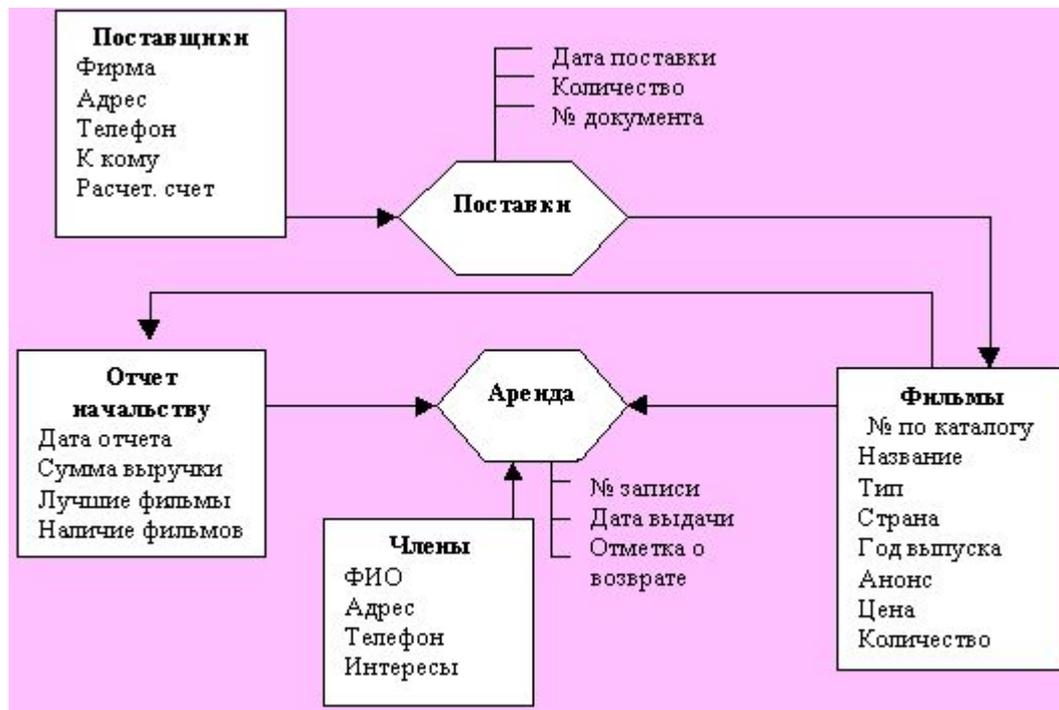


Рис.3.7. Объединенная диаграмма ER-модели предметной области "Видеотека"

Можно построить другой вариант ER- диаграммы, в которой ассоциированные понятия заменены непосредственными бинарными связями между объектами. Таким образом варианты модели могут различаться. Следует выбрать тот вариант, который наиболее просто, понятно и полно отображает необходимую информацию о предметной области.