MODERN APPROACHES
an overview

- **ENTITY-RELATIONSHIP APPROACH (Chen 1976)**
  pragmatic approach, no underlying theory.
  mainly diagrams.
- **SEMANTIC HIERARCHY MODEL (Smiths' 1977)**
  introduced abstractions in data modeling:
  classification, aggregation, generalization.
  does not cover data manipulation.
- **EXTENDED RELATIONAL APPROACH (Codd 1979)**
  set theoretic extensions to the basic relational approach. contains features to capture more meaning.
- **SEMANTIC NETWORKS (Brachman 1979)**
  developments from artificial intelligence in which generalizations and inheritance play a dominant role.
- **FUNCTIONAL DATA MODEL (Shipman 1981)**
  model based on functions, with an emphasis on inheritance.
- **ACTIVE/PASSIVE COMPONENT MODELING**
  (Brodie 1982) abstractions: classification, aggregation, generalization, association. integrity control by procedures instead of structures.
- **OBJECT ORIENTED DATABASES incl. OBJECT MODELING TECHNIQUE (Rumbaugh 1991), concepts: method, encapsulation, inheritance, polymorphism.**
- **SEMANTIC DATA MODELING**
  (ter Bekke 1992) the subject of this course.
  based on data abstractions: classification, aggregation, generalization and object relativity.
  includes a complete data manipulation language.

ENTITY-RELATIONSHIP APPROACH

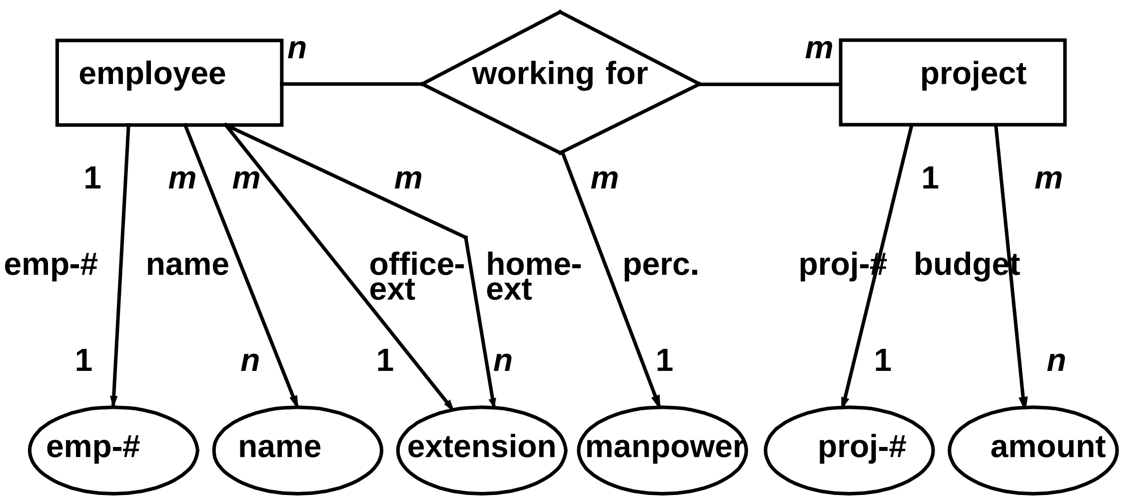
BASIC CONCEPTS:

- **ENTITIES, indicated by rectangles.**
- **RELATIONSHIPS, indicated by diamonds.**
- **ATTRIBUTES, indicated by ovals.**

GENERAL:

- **THE APPROACH IS OFTEN USED IN DESIGN TOOLS (WORKBENCHES).**
- **IMPLEMENTATION IS CARRIED OUT BY TRANSLATION INTO A RELATIONAL SCHEMA.**

EXAMPLE: about employees and their allocated projects.



ENTITY-RELATIONSHIP APPROACH
(continued)

EXERCISE: CD-CLUB.

In a club of CD (compact disc) player owners, members may borrow CDs from each other.

CDs are characterized by title, performer and owner.

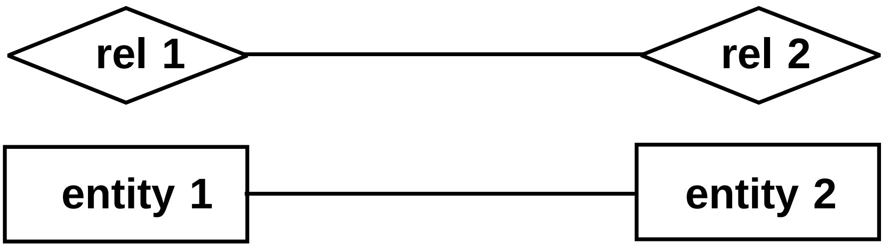
Identical CDs may be owned by more than one member.

Recorded member data are name and address.

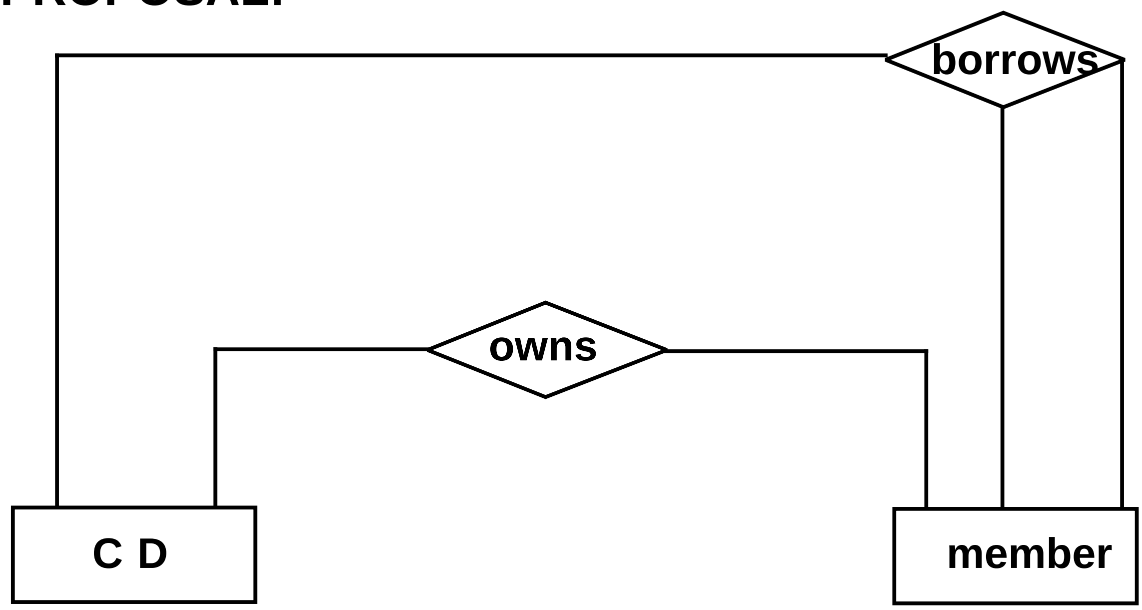
Property, borrowing and reservation are to be modeled.

E-R modeling
CD club: a proposal

INVALID ARE:



PROPOSAL:



PROBLEM:

owns

member	C D
M 1	C 1
M 1	C 2
M 2	C 3
M 3	C 4

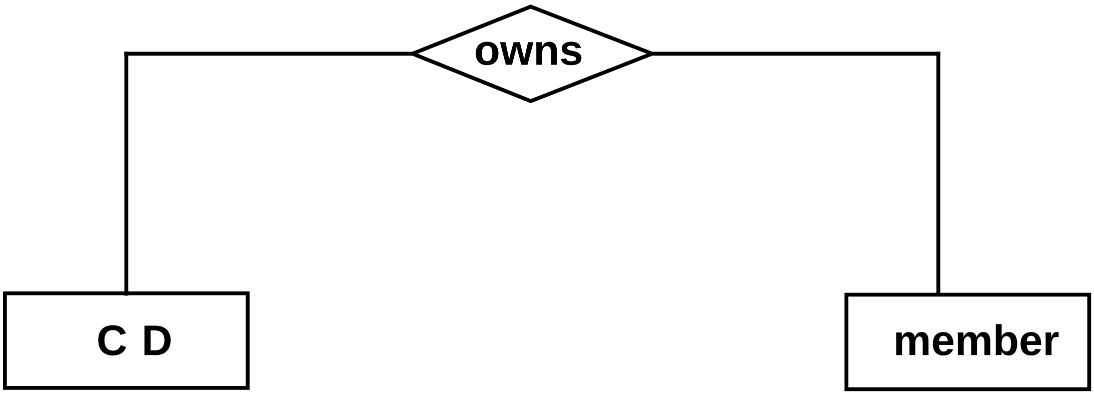
borrows

C D	from_m	to_m
C 2	M 3	M 2

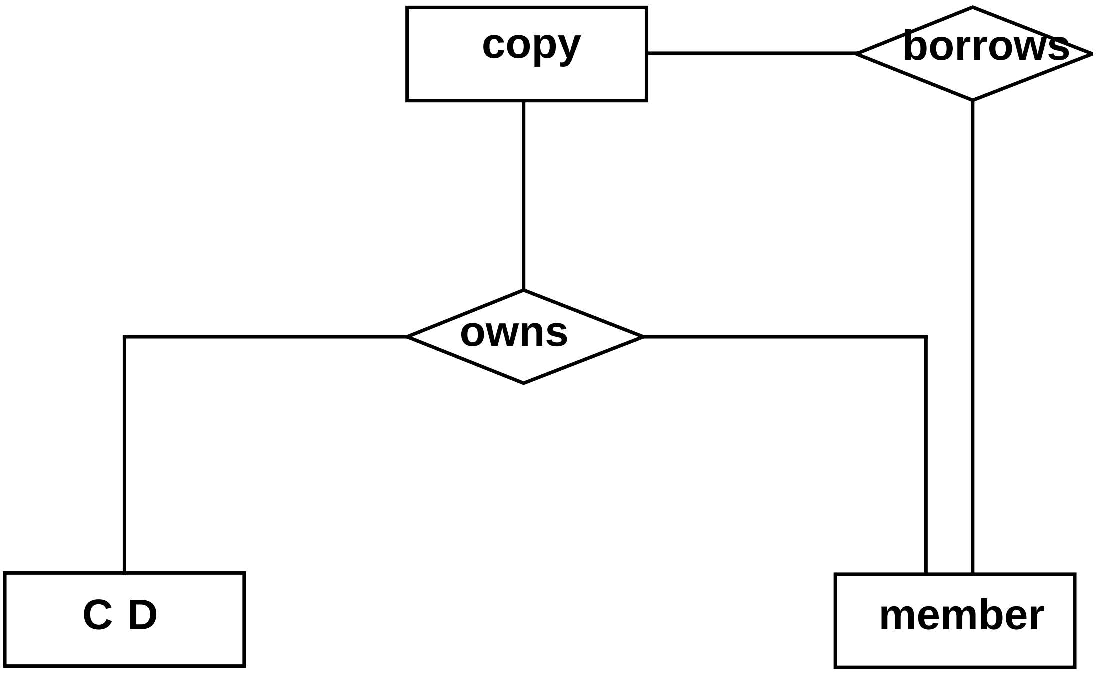
NOT
in owns

CD club: first alternative

FIRST PART:

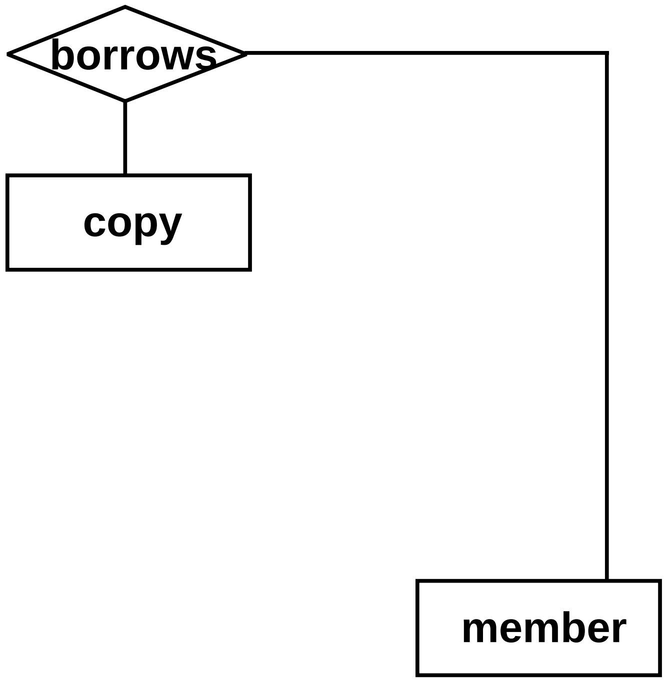


SECOND PART:

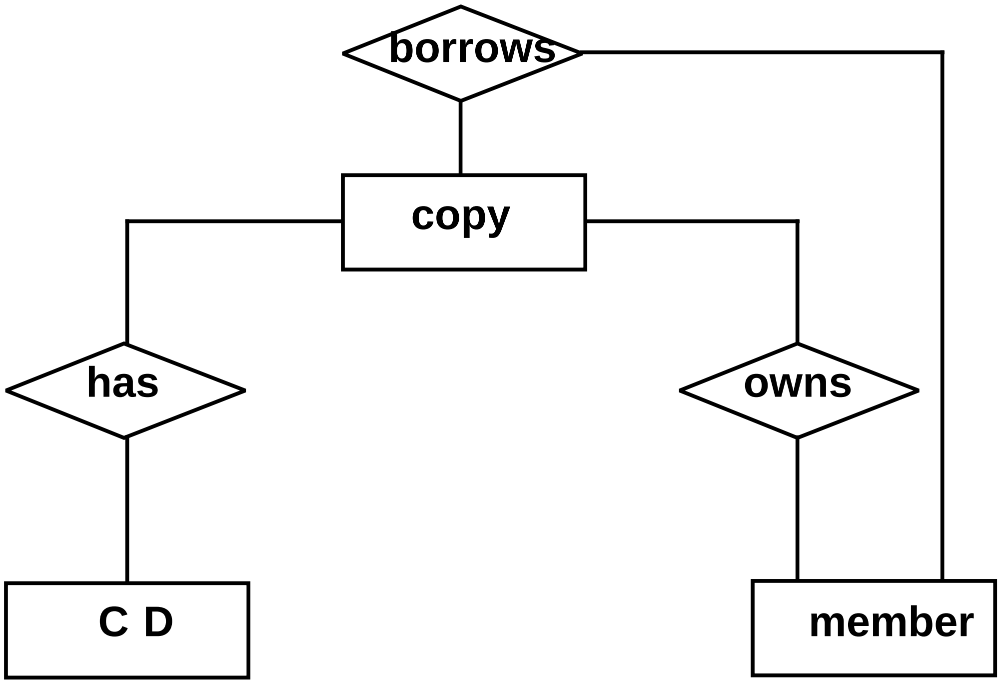


E-R modeling
CD club: second alternative

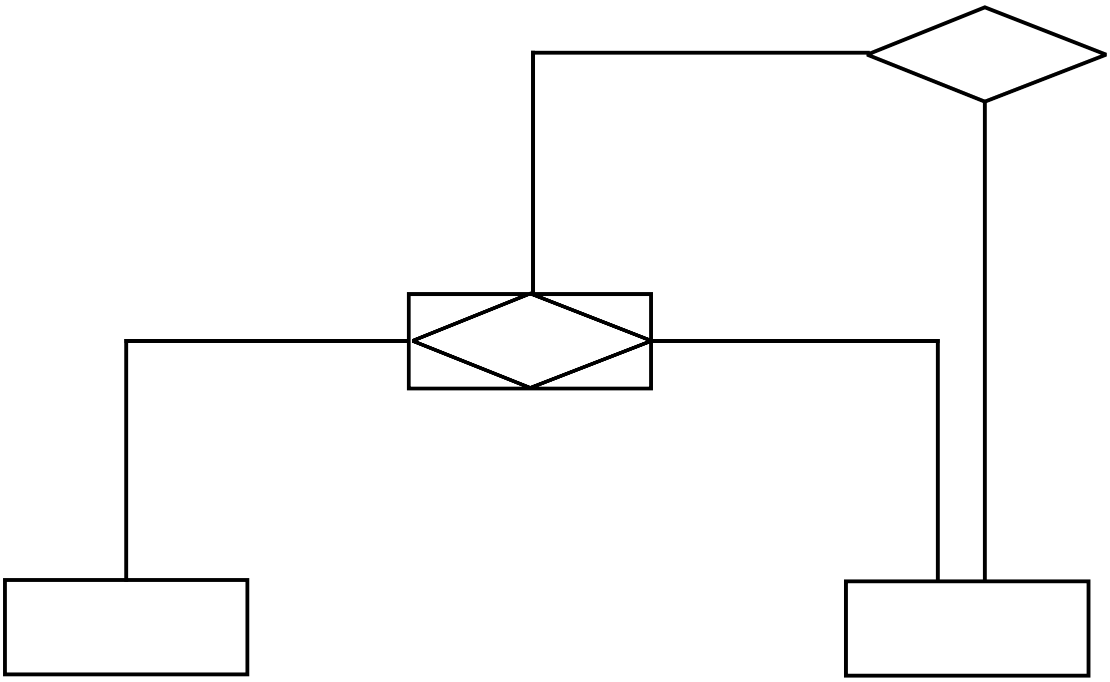
FIRST PART:



SECOND PART:



A BETTER SOLUTION WOULD BE:



BUT:

THIS SOLUTION DEMONSTRATES THE INADEQUACY OF ENTITY AND RELATIONSHIP AS BASIC CONCEPTS.

THESE ARE INTERPRETATIONS RELATIVE TO A USER VIEW.

ENTITY-RELATIONSHIP APPROACH
(continued)

PROBLEMS WITH VIEW INTEGRATION:

- **LACK OF MODEL ORTHOGONALITY**
- **PLURALISM OF PERCEPTIONS**
- **LACK OF DESIGN RELIABILITY**

METHOD FOR VIEW INTEGRATION:

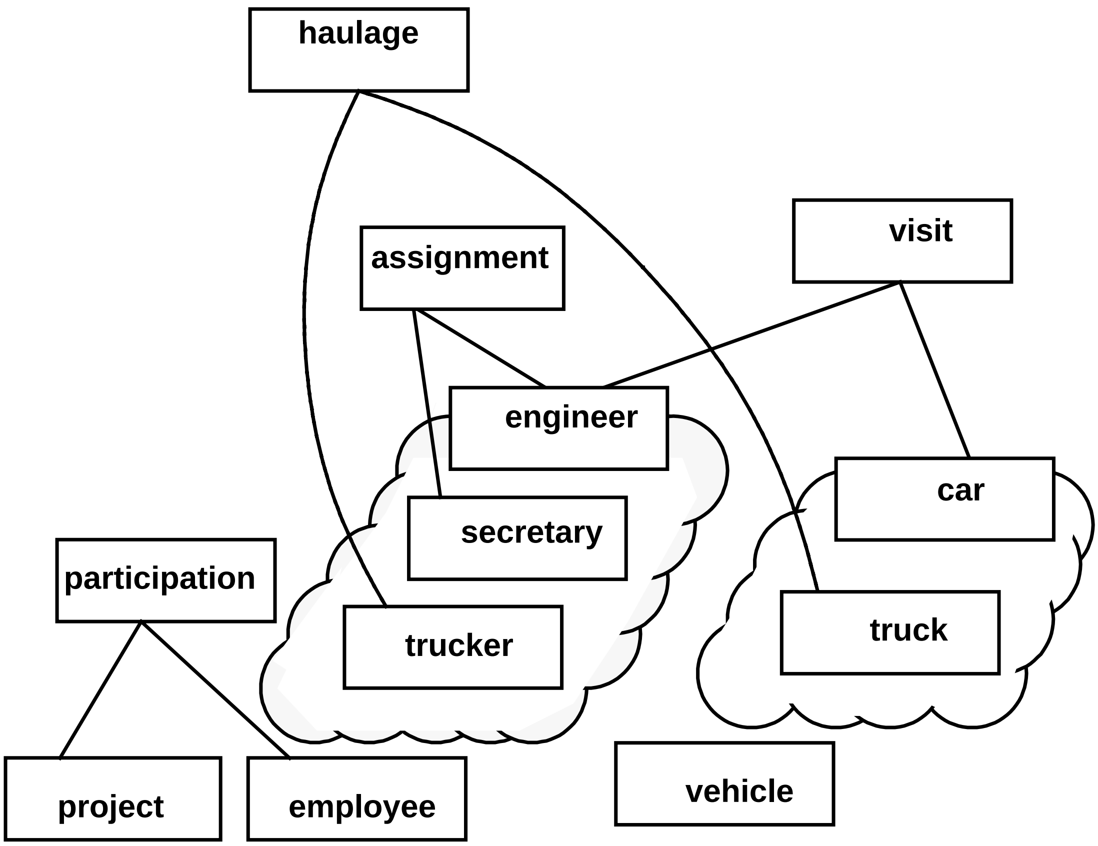
- **ANALYSIS OF DIFFERENCES IN NAMING (SYNONYMS AND HOMONYMS)**
- **ANALYSIS OF DIFFERENCES IN MODELING**
  (ATTRIBUTE, ENTITY)
  (ENTITY, RELATIONSHIP)
- **ANALYSIS OF RESTRICTIONS AND OPERATIONS (INHERENT, STATIC, DYNAMIC)**

SEMANTIC HIERARCHY MODEL

ABSTRACTIONS:

- CLASSIFICATION
- AGGREGATION
- GENERALIZATION

EXAMPLE:



EXTENDED RELATIONAL MODEL

CONCEPTS:

- SURROGATES
- E AND P RELATIONS
- SUBTYPES AND SUPERTYPES

S:

S ¢
alpha
beta

P:

P ¢
gamma
delta
epsilon

SP:

SP ¢
zeta
eta
delta
iota
kappa
lambda

SSNT:

S ¢	sup-No	name	town
alpha	1023	Smith	Leeds
beta	1024	Johnson	Liverpool

SPSPDQ:

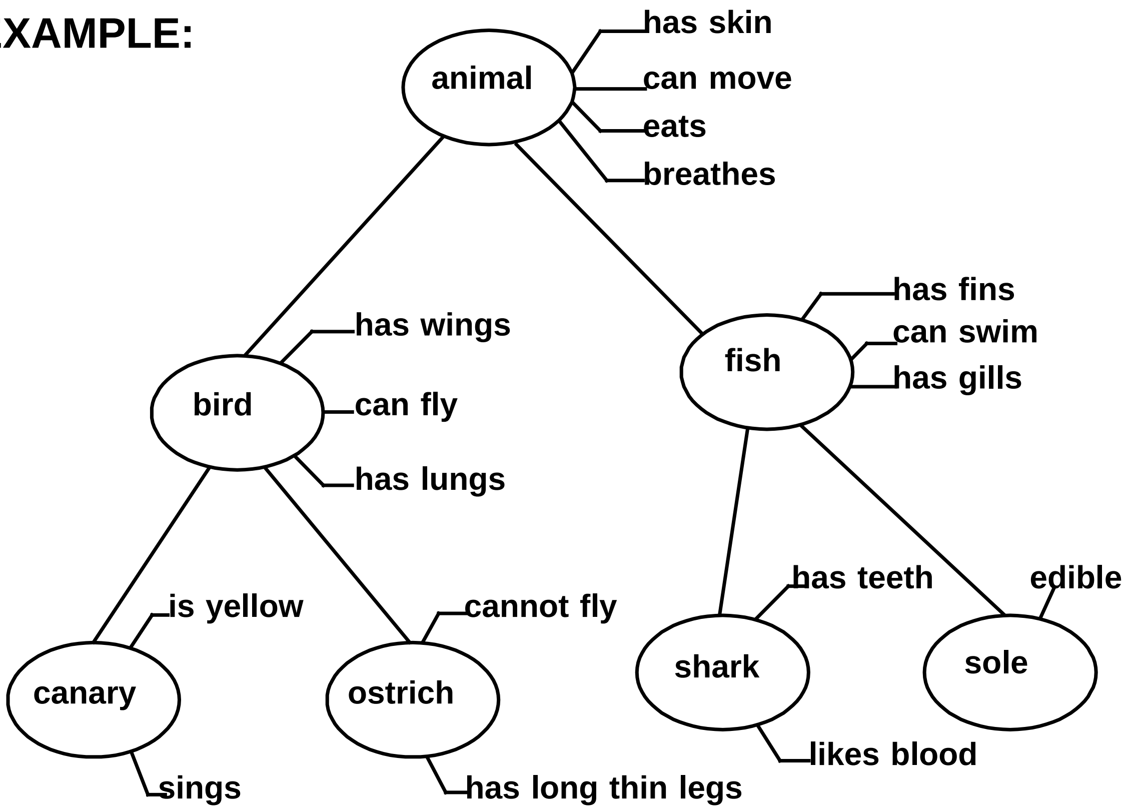
SP ¢	S ¢	P ¢	date	quantity
zeta	alpha	gamma	900401	30
eta	beta	delta	900403	50
delta	alpha	gamma	900407	10

SEMANTIC NETWORKS

INHERITANCE LINKS:

- GENERIC/GENERIC RELATIONSHIPS
  EXAMPLE: A DOG IS A MAMMAL.
- GENERIC/INDIVIDUAL RELATIONSHIPS
  EXAMPLE: JOHN IS A BACHELOR.
- IGNORANCE OF INHERITANCE
  EXAMPLE: BIRDS CAN FLY, BUT AN OSTRICH CANNOT FLY.

EXAMPLE:



FUNCTIONAL DATA MODEL

- **DATA DEFINITION: FUNCTIONS.**

DECLARE Study(Student) => Faculty
Function: Each Student corresponds to 1 Faculty

DECLARE Supervisor(Faculty) => Dean

DEFINE Supervisor(Student) =>
Supervisor(Faculty(Student))
Supervisor *of the* Faculty *of the* Student.

DECLARE Student(   ) =>> ENTITY
Definition of an entity.

DECLARE Name(Student) => STRING
Definition of a property.

- **DATA MANIPULATION: NESTED LOOPS.**

```
FOR EACH Student
   SUCH THAT Supervisor(Faculty(Student)) = ’Watson’
   PRINT Name(Student)
```

- **CONSTRAINTS**

DEFINE CONSTRAINT OwnDeptHead(Department) => Dept(Head(Department)

ACTIVE AND PASSIVE COMPONENT MODELING

AGGREGATION, GENERALIZATION AND ASSOCIATION

```
employee = OBJECT
    AGGREGATE OF
        salary-number;
        employee-name: ESSENTIAL;
        salary: ESSENTIAL;
            PRIMARY KEY salary-number;
    GENERIC OF
        secretary: EXCLUSIVE;
        manager: EXCLUSIVE;
    END OBJECT;

secretary = OBJECT
    AGGREGATE OF
    .
    .
    END OBJECT;

manager = OBJECT
    AGGREGATE OF
    .
    .
    END OBJECT;

control = OBJECT
    AGGREGATE OF
        name;
        PRIMARY KEY name;
        ASSOCIATION OF employee;
        MANUAL MEMBERSHIP;
    END OBJECT;
```

SEMANTIC DATA MODELING
OBJECT TYPE ORIENTED DATA MODELING

" A given object in the real world may well be regarded as:

- **an entity by some people, as**
- **a property by others, and as**
- **an association by still others**
- **etc.**

It is a goal of semantic modeling to support such flexibility of interpretation ".

LEVELS OF ABSTRACTION IN A DBMS

• **EXTERNAL LEVEL**	
(an interpretation)	**VIEW INDEPENDENCE**
• **CONCEPTUAL LEVEL**	
(common model)	**DATA INDEPENDENCE**
• **INTERNAL LEVEL**	
(an implementation)	

SEMANTIC MODELING IS CONCERNED WITH A CONCEPTUAL MODEL THAT ALLOWS VIEW INDEPENDENCE.

REQUIREMENTS FOR SEMANTIC MODELS

BASIC CONCEPTS IN:

- **RELATIONAL MODEL**
  domain
  relation
  attribute
  .
  .

- **ENTITY-RELATIONSHIP MODEL**
  entity
  relationship
  value
  .
  .

BASIC CONCEPTS HAVE A FIXED INTERPRETATION.
(e.g. a domain can not be interpreted as a relation, etc.).

CONCLUSION:

DATA MODELS BASED ON MORE THAN ONE BASIC CONCEPT ARE NOT SEMANTIC !

SEMANTIC ABSTRACTIONS

- **CLASSIFICATION**

  resulting in the concept of type.

- **AGGREGATION**

  collection of different types into a new type.

- **GENERALIZATION**

  intersection of types from different aggregations of types.

- **SPECIALIZATION**

  union of types from different aggregations of types.

CLASSIFICATION

TYPE: BASIC CONCEPT

EXAMPLES:

name
peter
john
.

height
1.76
1.82
.
.

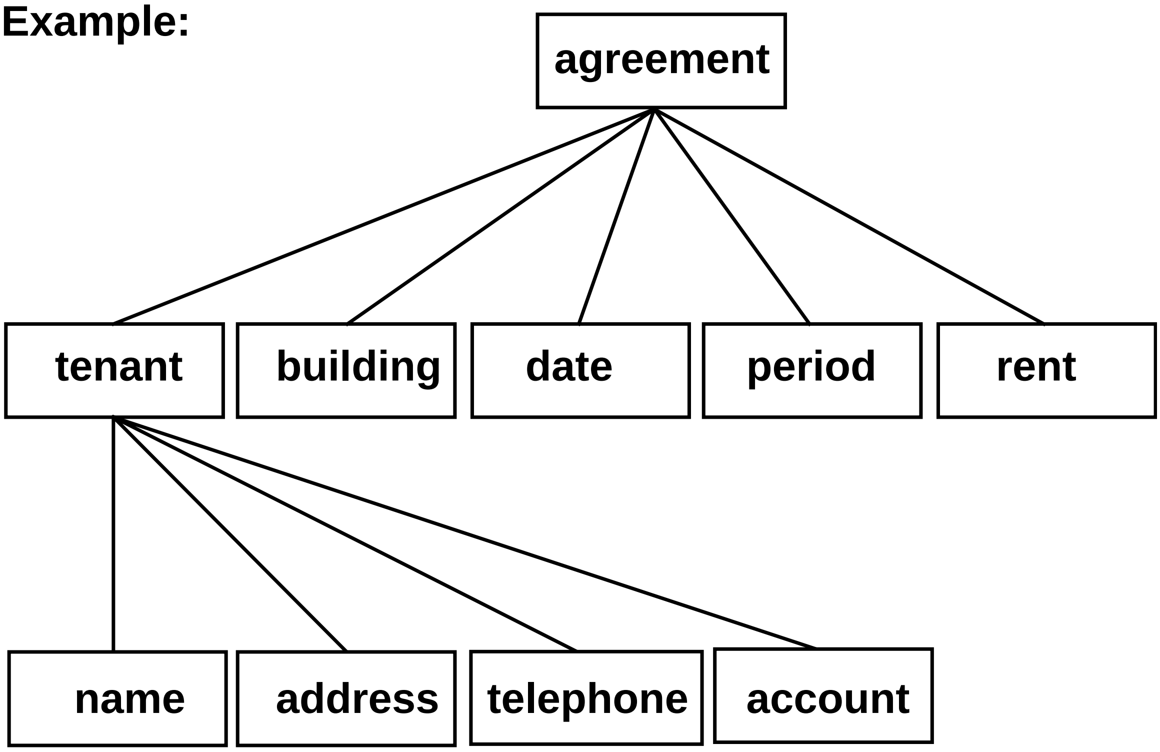
NOTE:

property is not defined by its extension

(i.e. values are illustrations of the concept).

AGGREGATION
The collection of certain types into a new type.

Example:



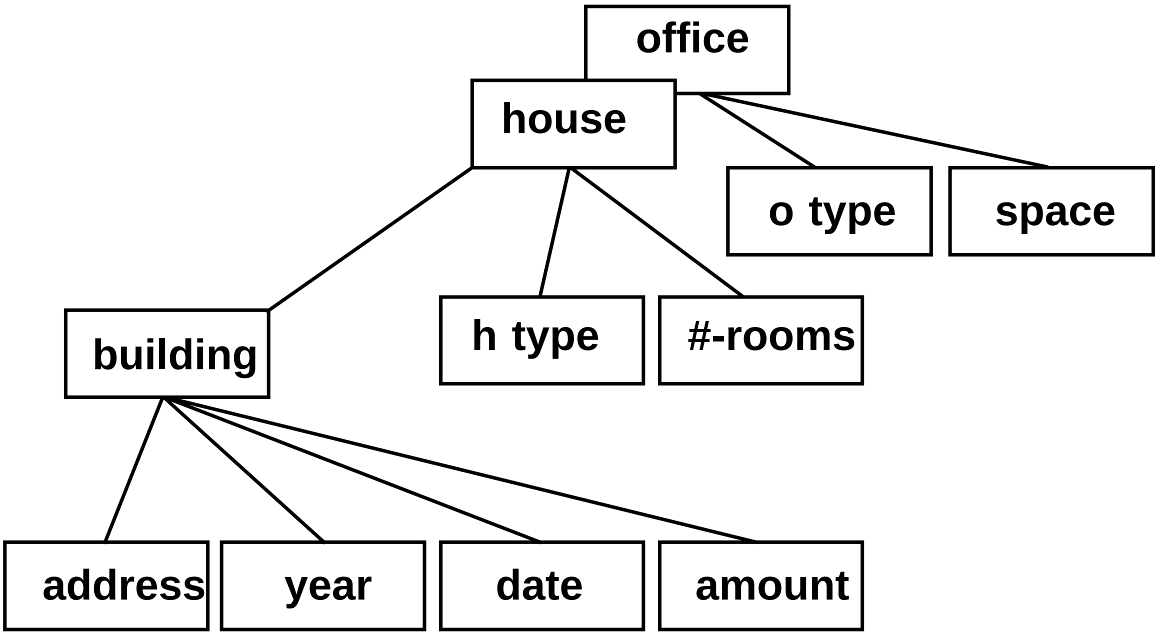
VIEW INDEPENDENCE:

- **tenant can be considered as a type but also as an attribute.**
- **building is an attribute, but can later be considered as a type.**

RELATED TO THE VERB "TO HAVE".

GENERALIZATION
The intersection of types from different types.

Example:



PROPERTY:

- **INHERITANCE OF ATTRIBUTES.**

RELATED TO THE VERB "TO BE".

TYPE DEFINITIONS

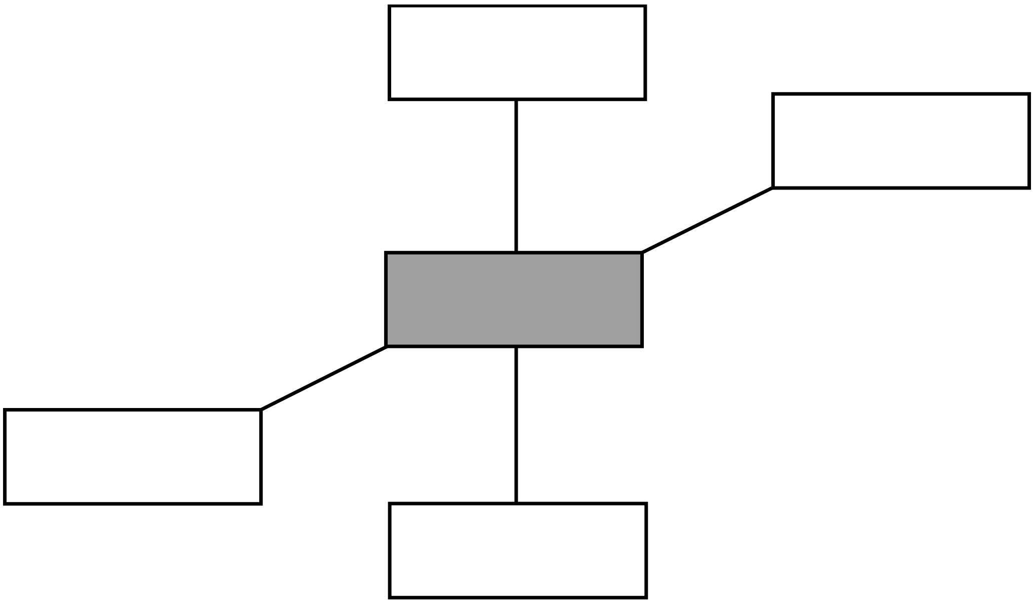
- **AGGREGATION**

type **tenant**	**=**	**name, address, telephone, account**
type **agreement**	**=**	**tenant, building, date, period, rent**

- **GENERALIZATION**

type **building**	**=**	**address, year, date, amount**
type **house**	**=**	**[building], house type, number of rooms**
type **office**	**=**	**[building], office type, space**

VIEW INDEPENDENCE



OBJECT CAN BE INTERPRETED AS:

- **TYPE**
- **ATTRIBUTE**
- **RELATIONSHIP**
- **GENERALIZATION**
- **SPECIALIZATION**
-
-

THIS PROPERTY IS ALSO KNOWN AS:

OBJECT RELATIVITY.

EXERCISE: CD-CLUB.

In a club of CD (compact disc) player owners, members may borrow CDs from each other.

CDs are characterized by title, performer and owner.

Identical CDs may be owned by more than one member.

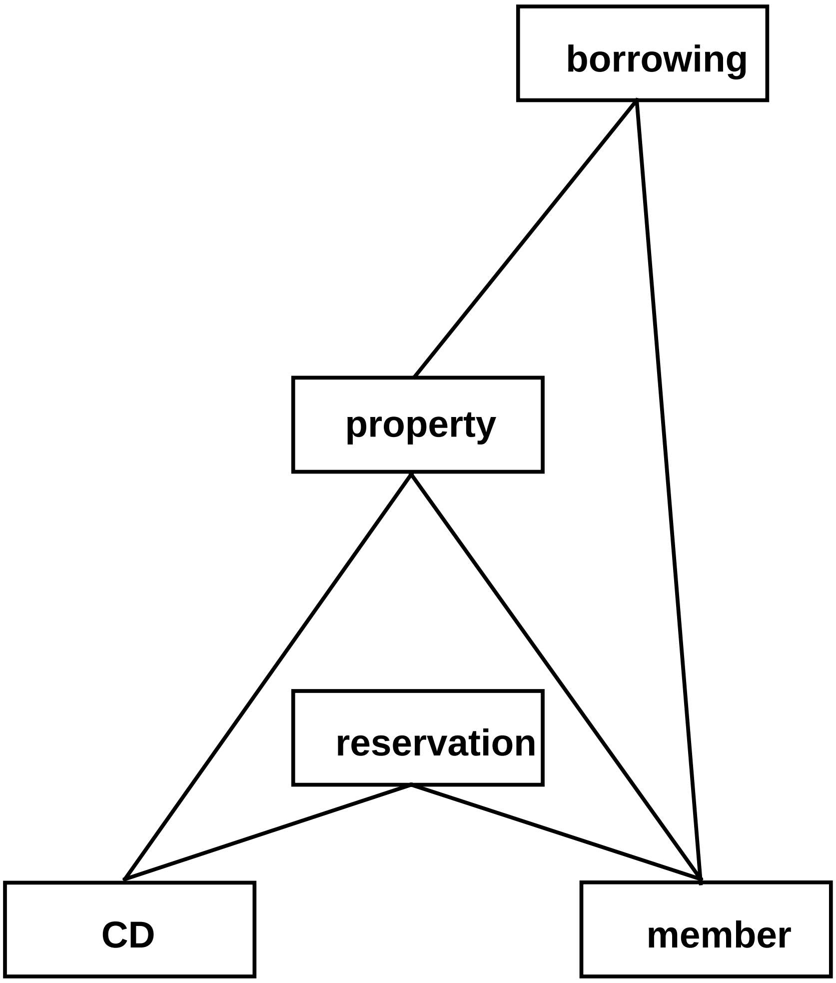
Recorded member data are name and address.

Property, borrowing and reservation are to be modeled.

SEMANTIC DATA MODELING
CD club: the unique solution

type **CD**	**= title, performer**
type **member**	**= name, address**
type **reservation**	**= CD, member, period**
type **property**	**= CD, member**
type **borrowing**	**= property, member, from_date, to_date.**

The abstraction hierarchy:



DATA MODELING
Exercises E-R approach

EXERCISE 1

Employees are defined in an enterprise schema by the attributes name, address, town, job, salary and department. Provide the E-R schema for the entity employee.

EXERCISE 2

The E-R schema needs extending because job must also be regarded as an entity with the attributes job description and maximum salary. Which attributes now define the entity employee?

EXERCISE 3

A few constraints are added to the above E-R schema:

- **a** **The number of employees per department is limited by the maximum maxstaff. Indicate the constraint in the schema.**
- **b** **The total salary budget per department is limited by the maximum totsal. Can this constraint be applied to the E-R schema as well? Explain briefly.**

MODERN APPROACHES
evaluation

E-R APPROACH

- **two separate concepts (entity and relationship) interfere with the exchange of interpretations.**
- **cardinality constraints cannot be used for transitive relationships.**
- **constraints for other set functions do not exist.**
- **generalizations are introduced on an ad hoc basis.**

DATABASE ABSTRACTIONS:

- **aggregation and generalization viewed from the classical relational model; are not formally defined.**

RM/T MODEL

- **overloaded with concepts.**
- **subtype and supertype are not formally defined.**

SEMANTIC NETWORKS

- **exclusively based on is-a connections with different interpretations.**

FUNCTIONAL DATA MODEL:

- **lacks certain important consistency constraints (such as unique identifications).**
- **interesting operator to derive information without using a relational join.**

ACM/PCM:

- **integrity control by procedures instead of structures.**

COMPARISON

- **RELATIONAL:**

  data are represented in tables.

  ADVANTAGE:

  simple data model, query language, wide range of hardware.

  DISADVANTAGE:

  limited query facilities, pitfalls, limited integrity control, low performance.

- **SEMANTIC:**

  data represented in tables, emphasis on abstractions.

  ADVANTAGE:

  structuring capabilities, query language, reliability, flexibility, high performance.