One of the Most Common Student Mistakes

- Find the name of the employee(s) with the lowest salary

```
SELECT    FName, LName, MIN(Salary)
FROM      Employee;
```

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

SQL Exercise

- Find the employee (SSN) who works on all project locations except Toronto.
  - Say the company has
    - P1 and P2 in Montreal,
    - P10 in Waterloo,
    - P12 in Toronto.
    - Jane (SSN = 1111111) worked on P1 and P10,
    - Peter worked on P1 and P2,
    - Kate worked on P1, P10 and P12.
  - Then
    - Jane should be included in the result while
    - Peter and Kate should not.

Review of Last Lecture

- Advanced SQL
  - View
  - Assertion
  - Trigger
- SQL exercise

Today's Plan

- Entity
- Weak entity
- Attributes
- Relationship
- Structural constraints

ENTITY-RELATIONSHIP DIAGRAMS

CHAPTER 06

Collin Roberts

University of Waterloo

Recommended Reading

- Textbook Chapter 03

Overview of Database Design

- Conceptual design
  - Describe the data informally
- Logical design
  - Mapping
  - Normalization
- Physical design
  - Storage, indexing, RAID

Benefits of ER Diagrams

- Allow us to sketch the design informally
  - Entity
  - Attribute
  - Relationship
  - Constraint
- Schemas in the ER model are entity-relationship diagrams
- DBMS independent
- There is a fairly mechanical way to convert ER diagrams to real implementations like relational databases.

Entity

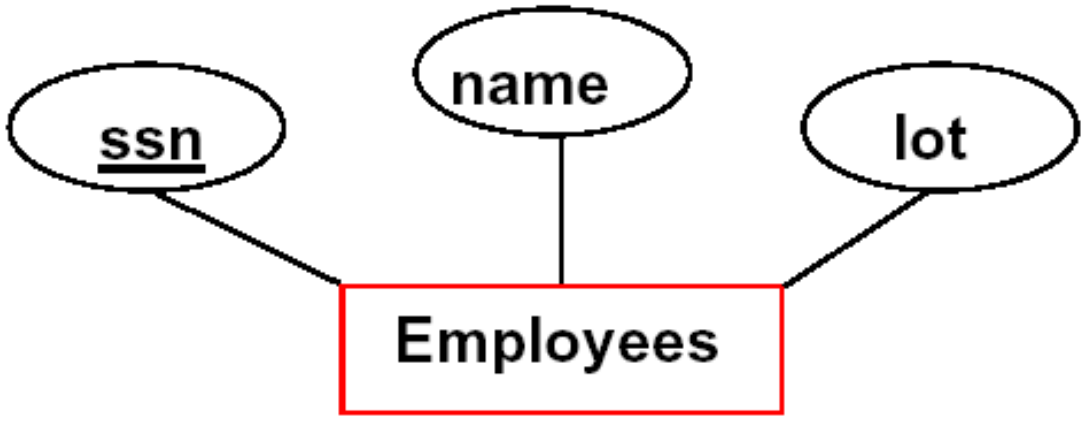
- Entity: Real-world object distinguishable from other objects
  - e.g., specific person, company, event
  - described (in DB) using a set of *attributes*.
- Entity Set: A collection of similar entities that share the same properties. E.g., all employees.
  - All entities in an entity set have the same set of attributes.
  - e.g., set of all persons, companies, events
  - Each entity set has a *key*.

Attribute

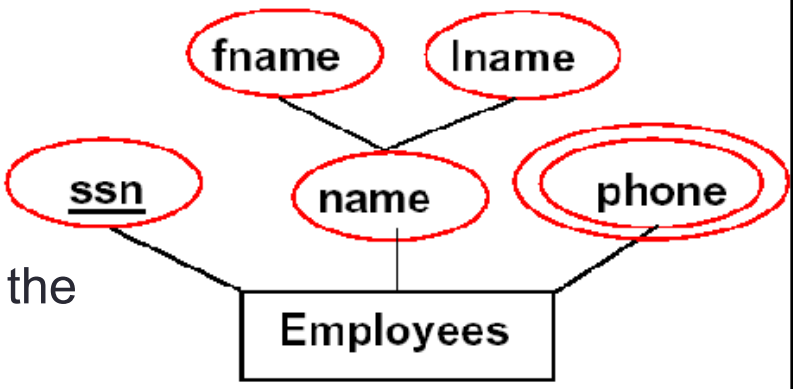
- An entity is represented by a set of attributes possessed by all members of an entity set.
- *Domain* – the set of permitted values for each attribute, e.g., 18 < age < 65

ER Diagram

- In an ER diagram, each entity set is represented by a rectangle
- Each attribute of an entity set is represented by an oval linked to the rectangle



Types of Attributes

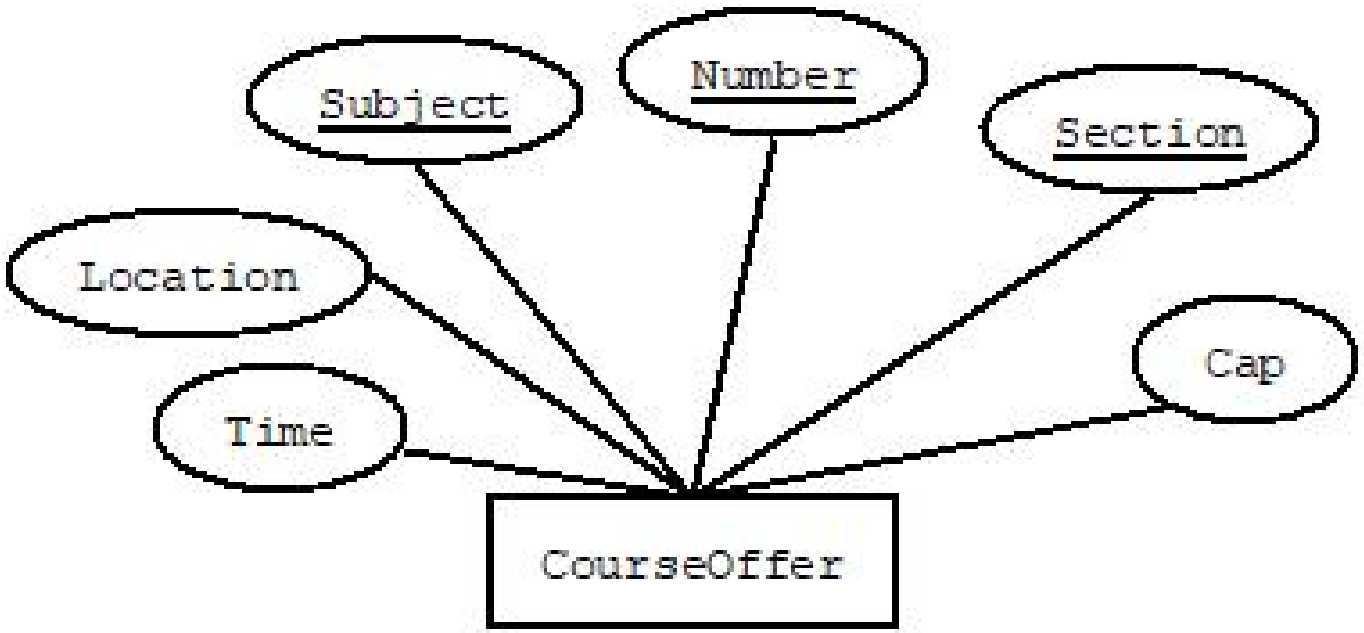


- Simple
  - Each entity has a single atomic value for the attribute
  - Example: SSN or Sex.
- Composite
  - The attribute may be composed of several components.
  - Example: Address (Apt#, House#, Street, City, State, ZipCode, Country) or Name (FirstName, MiddleName, LastName).
- Multi-valued
  - An entity may have multiple values for that attribute
  - Example: Color of a CAR or PreviousDegrees of a STUDENT. Denoted as {Color} or {PreviousDegrees}.

Keys

- A *super key* of an entity set is a set of one or more attributes whose values uniquely determine each entity.
  - Ssn, name
- A *candidate key* of an entity set is a minimal super key
  - *Ssn* is candidate key of *Employees*
  - *Did* is candidate key of *Departments*
- Although several candidate keys may exist, one of the candidate keys is selected to be the *primary key*.

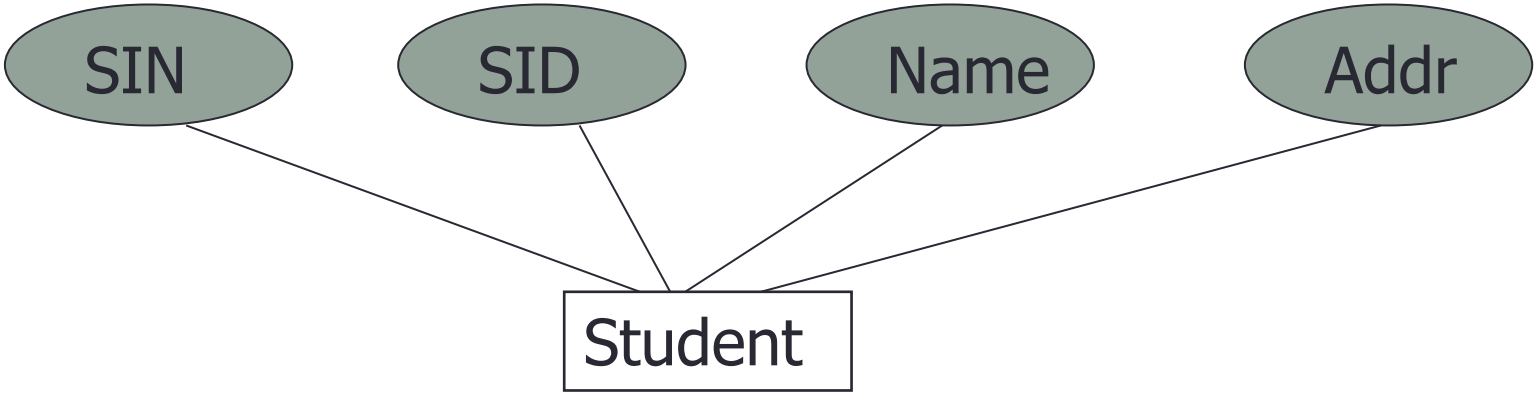
Example: Multi-attribute key



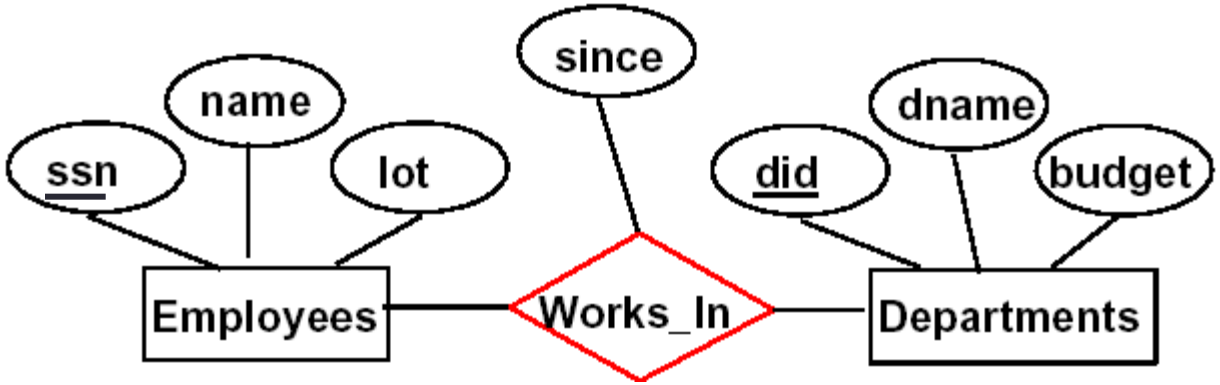
- (Location, Time) cannot be a key due to cross-list.

Exercise: SK, CK and PK

- Identify all the CK, SK and PK for the following entity:

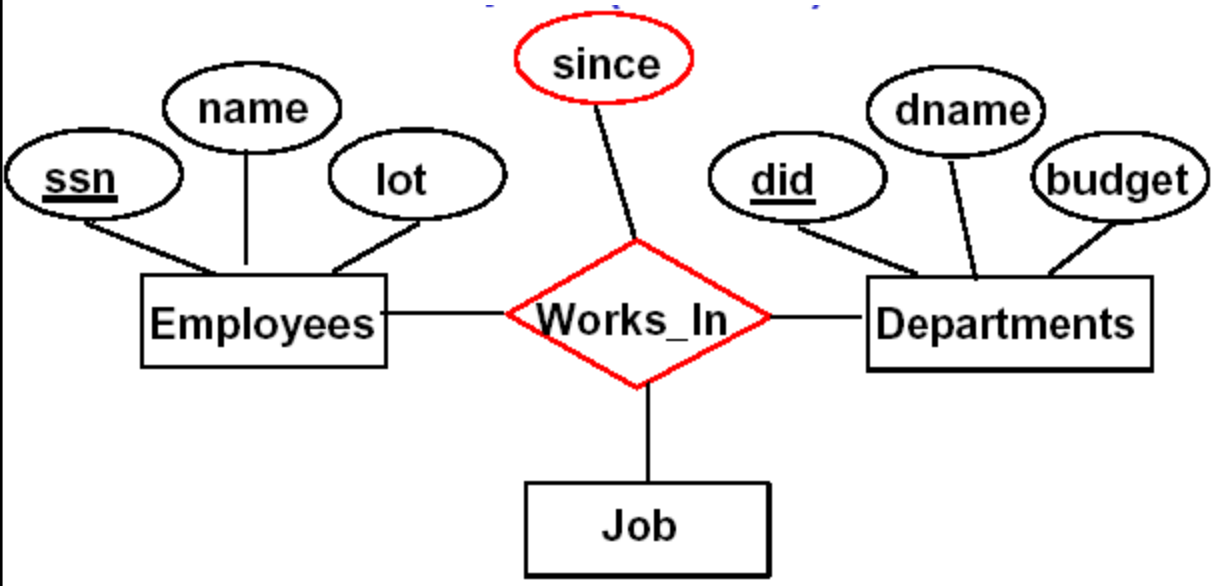


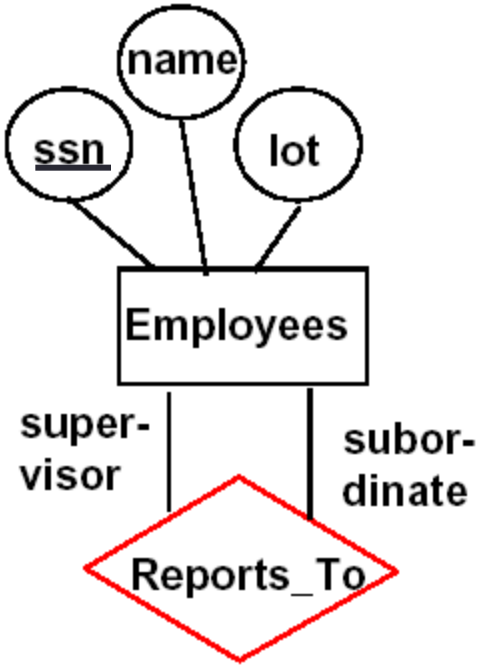
Relationship



- Relationship: Association among several entities
  - e.g., the employee John works in Pharmacy department.
- To create an instance of a relationship, we must indicate which employee and which department we want to have connected (for this relationship).
- We need the key value for an employee and the key value for a department, stored together, to represent the relationship

Relationship Property



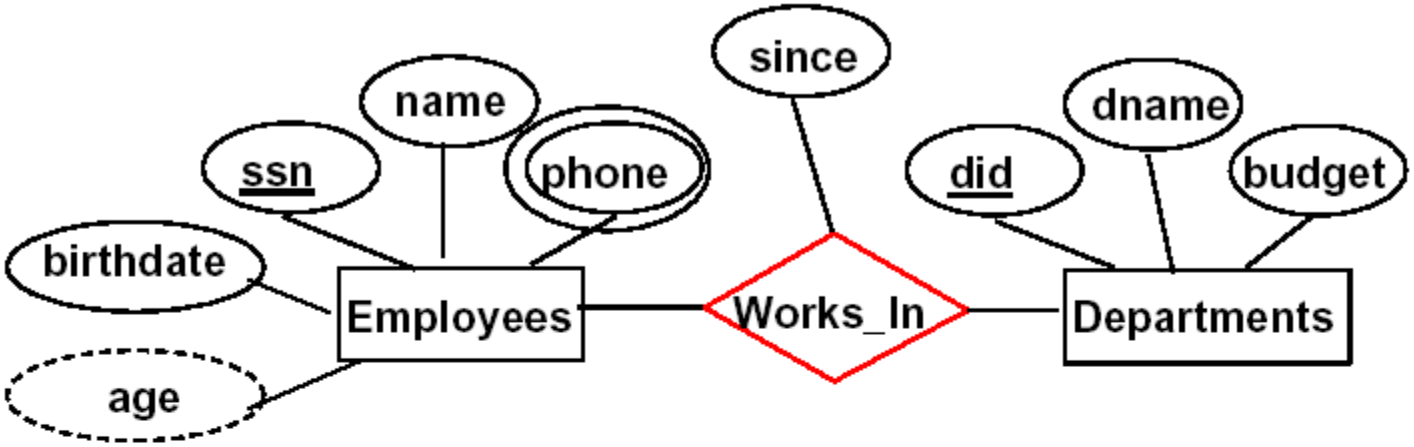


- A relationship set may have an attribute
- Same entity set can participate in different relationship sets, or in different "roles" in same set.
- Relationship degree: the number of entities involved, e.g., binary, ternary.
  - E.g., Works_in is a ternary relationship: employees of the company may have jobs at multiple depts, with different jobs at different depts.

Relationship Key

- Suppose
  - R involves E_1, …, E_n
  - Primary key of E_i is primary-key(E_i)
  - R has no attributes
  - Attribute names are all unique
- Then
  - primary-key(E_1) $\cup$ ... $\cup$ primary-key(E_n)
  - … is a superkey of R
- Why talk about keys of relationship sets?
  - To enforce cardinality constraints
  - To prepare for tabular (relational) representation

Quick Summary on ER Diagram



- **Rectangles** represent entity sets.
- **Diamonds** represent relationship sets.
- **Lines** link attributes to entity sets and entity sets to relationship sets.
- **Ellipses** represent attributes
  - **Double ellipses** represent multivalued attributes.
  - **Dashed ellipses** denote derived attributes.
- **Underline** indicates primary key attributes (will study later)

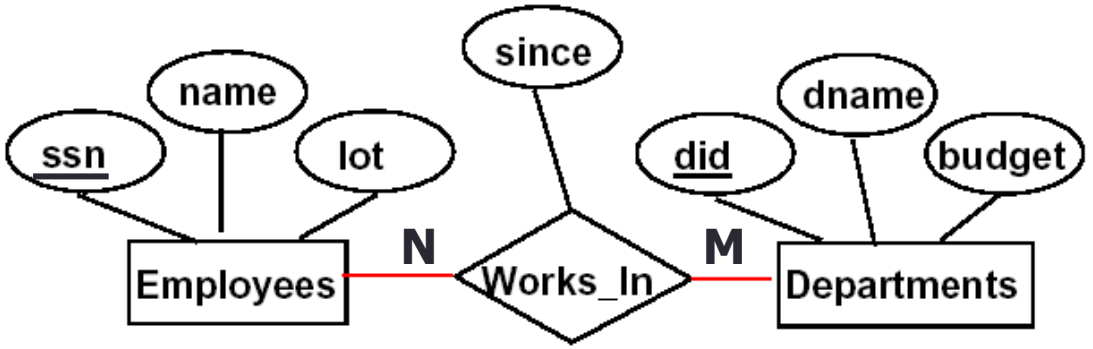
Structural Constraints

- Structural constraints on relationships:
  - Cardinality ratio (of a binary relationship): 1:1, 1:N, N:1, or M:N
    - SHOWN BY PLACING APPROPRIATE NUMBER ON THE LINK.
  - Participation constraint (on each participating entity type): total (called existence dependency) or partial.
    - SHOWN BY DOUBLE LINING THE LINK
- NOTE: These are easy to specify for Binary Relationship Types.

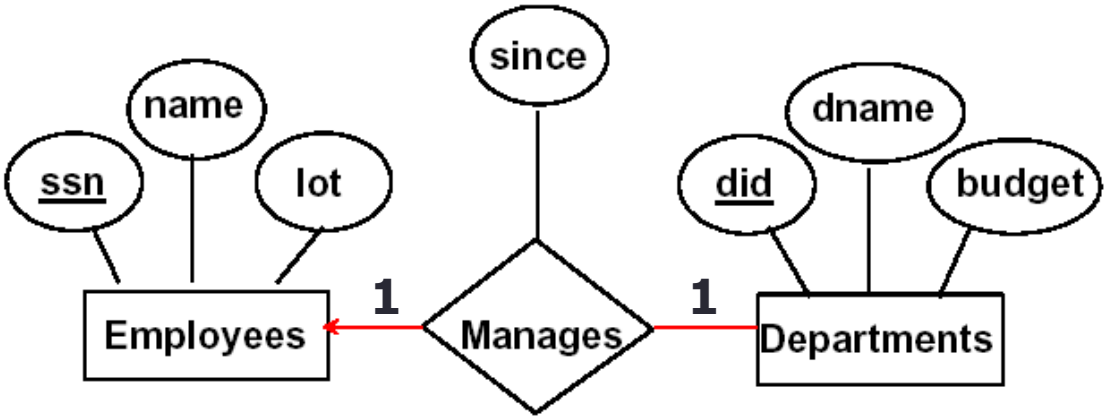
Cardinality/Key Constraint

How many times must/may an entity instance participate?

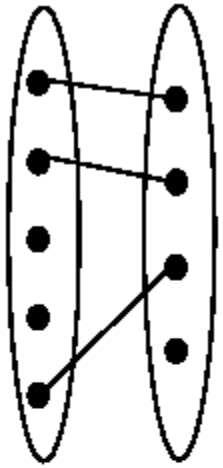
- An employee can work in many departments; a dept can have many employees.



- Each dept has at most one manager



Cardinality: 1-1

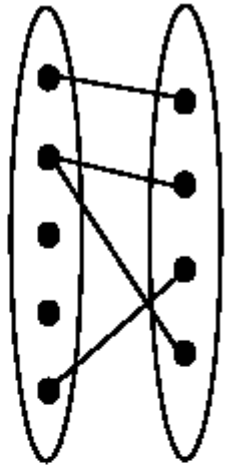


One-to-one relationship:

A customer is associated with at most one loan via the relationship *borrower*

A loan is associated with at most one customer via *borrower*

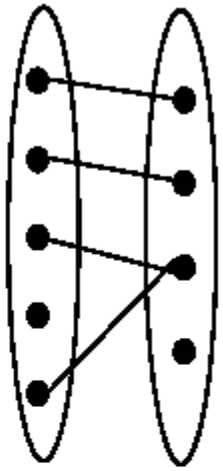
Cardinality: 1-Many



1-to Many

One-to-many relationship: a loan is associated with at most one customer via *borrower*, a customer is associated with several (including 0) loans via *borrower*

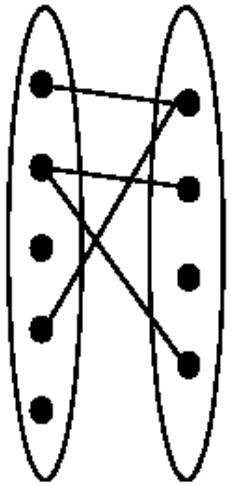
Cardinality: Many-1



Many-to-1

Many-to-one relationship: a loan is associated with several (including 0) customers via *borrower*, a customer is associated with at most one loan via *borrower*

Cardinality: Many-Many



Many-to-Many

Many-to-many relationship: A customer is associated with several (possibly 0) loans via borrower; a loan is associated with several (possibly 0) customers via borrower

Participation Constraint

- Total participation (indicated by double line): every entity in the entity set participates in at least one relationship in the relationship set
  - participation of *loan* in *borrower* is total -- every loan must have a customer associated to it via borrower
- Partial participation: some entities may not participate in any relationship in the relationship set
  - participation of *customer* in *borrower* is partial

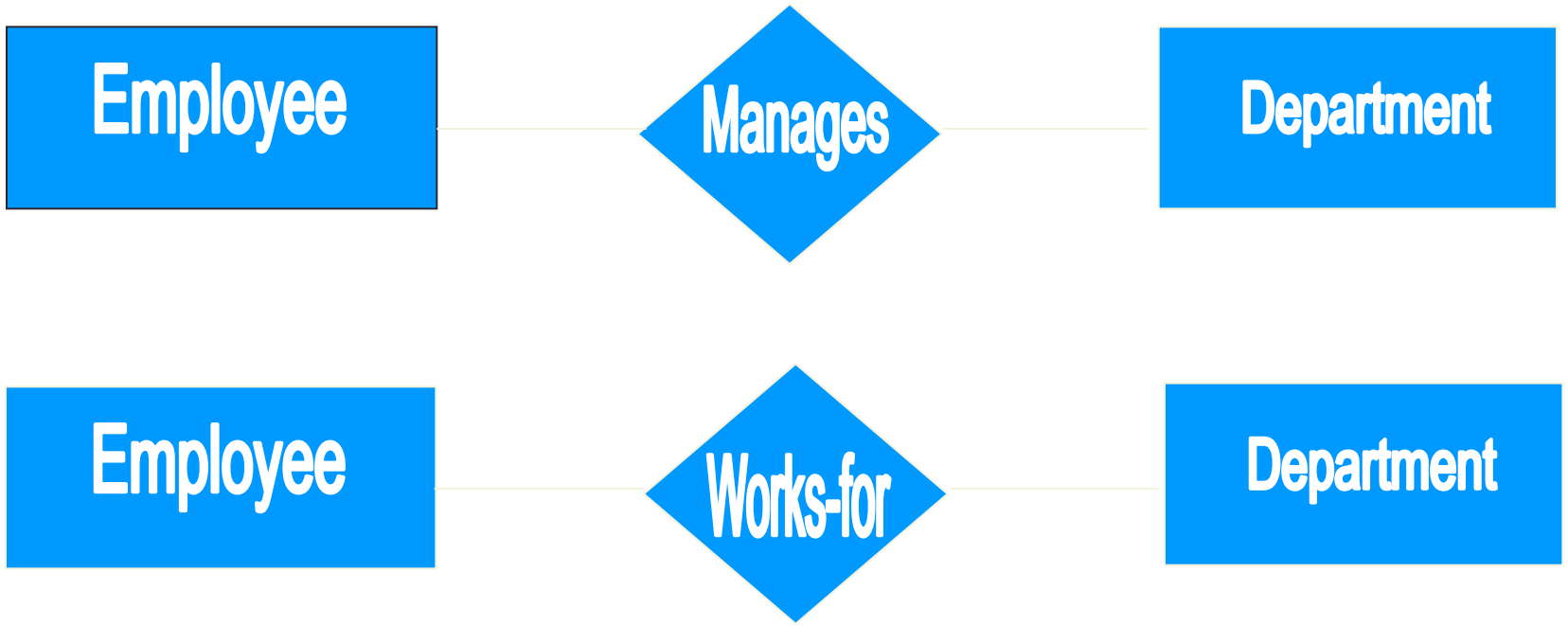
In-class Exercise #1: Structural Constraints

1. Marriage between two persons?

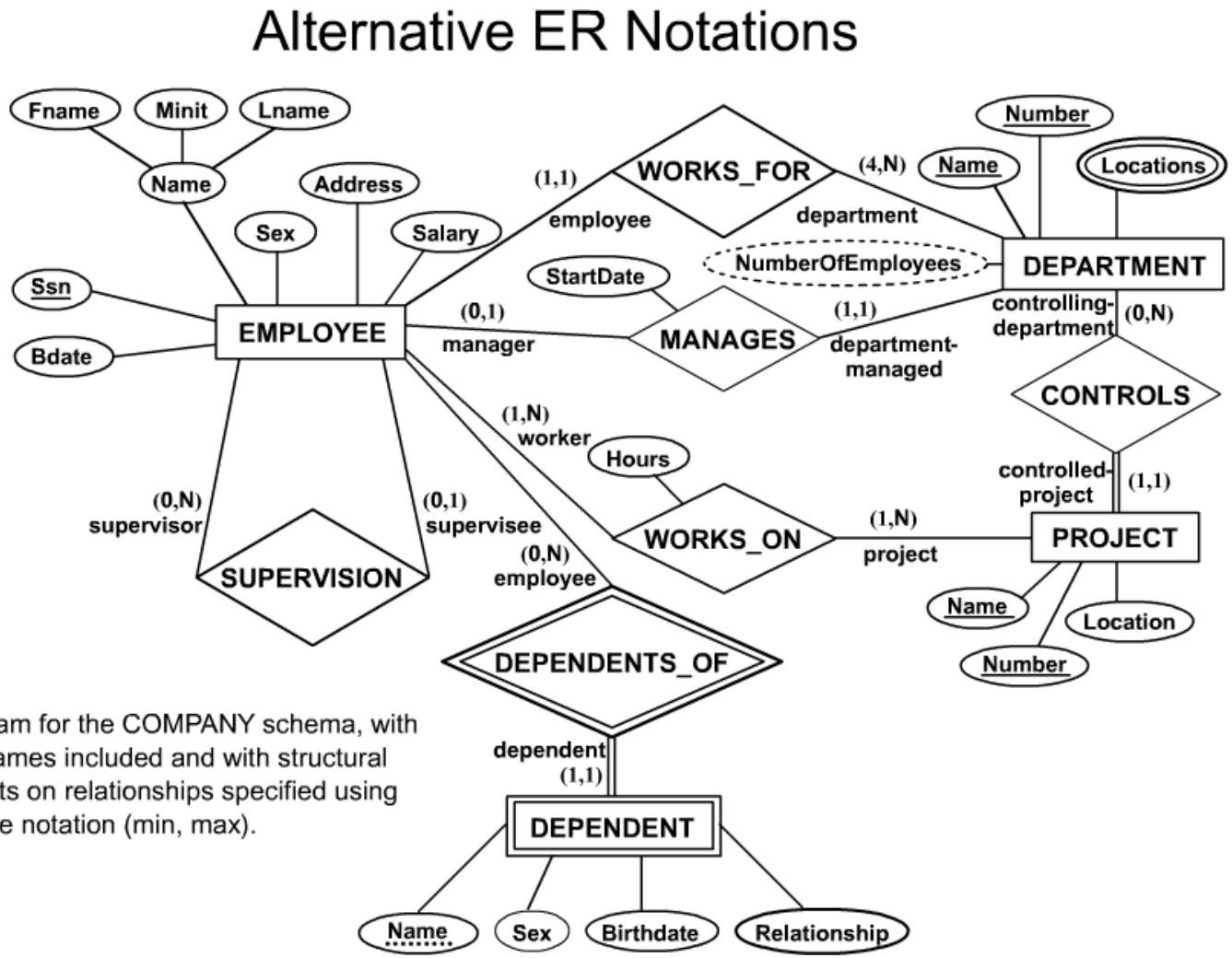
2. How about BG between two persons?

3. How about PC between adult and child?

The (min,max) notation (FYI)



COMPANY ER Schema Diagram using (min, max) notation (FYI)

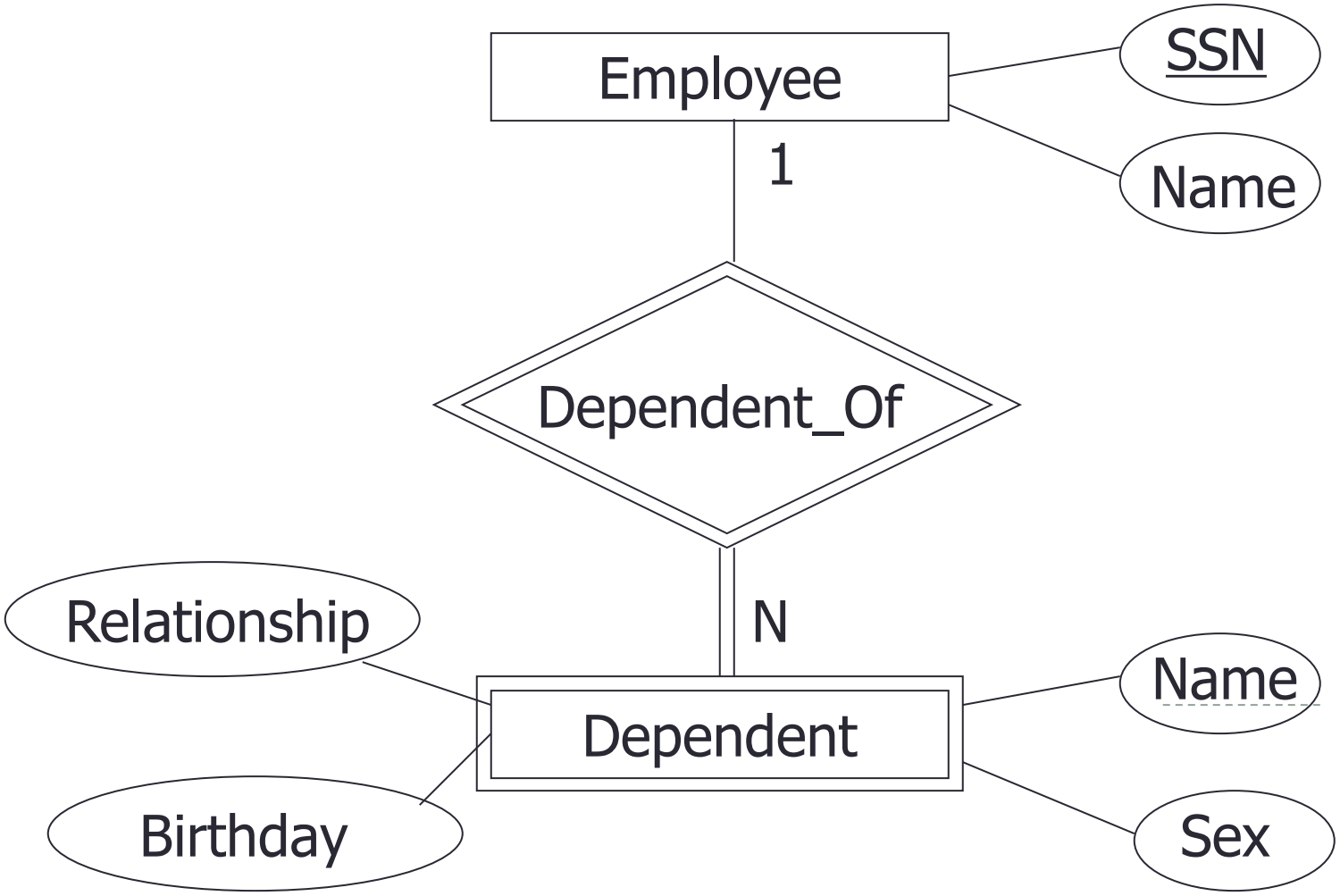


ER diagram for the COMPANY schema, with all role names included and with structural constraints on relationships specified using alternative notation (min, max).

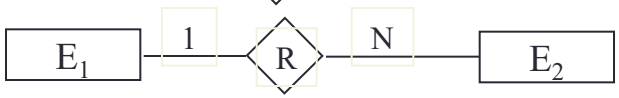
Weak Entity Types

- Weak entity: an entity that does not have a primary key
  - Identified by a double rectangle
  - Naturally, the opposite is called strong entity
- A weak entity must TOTALLY participate in an identifying relationship type with an owner or identifying entity type
  - Identified by a double diamond
- Weak entities are identified by the combination of:
  - A partial key of the weak entity type (discriminator)
    - Represented by a dash line
  - The particular entity they are related to in the identifying entity type

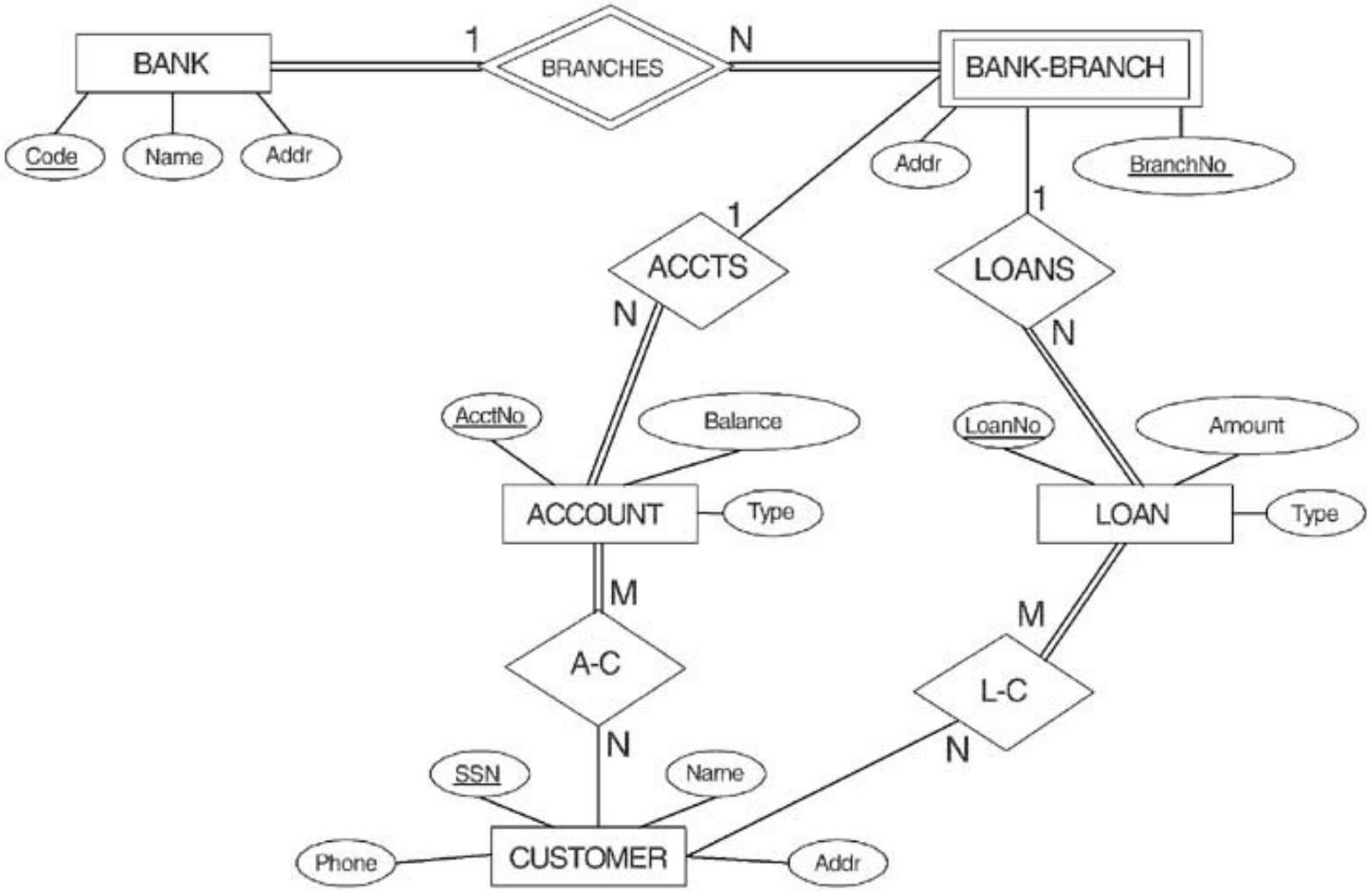
Weak Entity Type Example



SUMMARY OF ER NOTATION

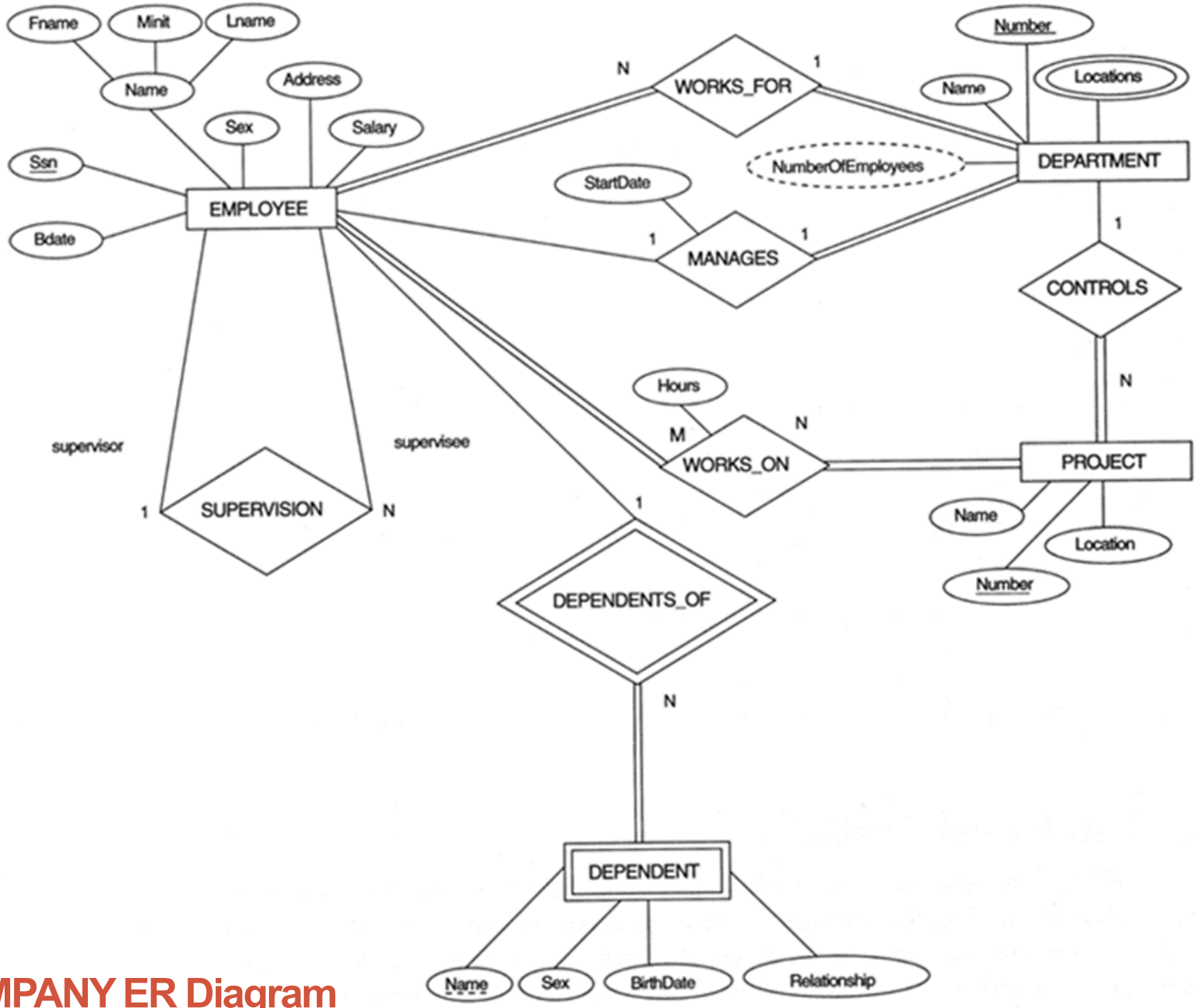
Symbol	Meaning
	ENTITY TYPE
	WEAK ENTITY TYPE
	RELATIONSHIP TYPE
	IDENTIFYING RELATIONSHIP TYPE
	ATTRIBUTE
	KEY ATTRIBUTE
	MULTIVALUED ATTRIBUTE
	COMPOSITE ATTRIBUTE
	DERIVED ATTRIBUTE
	TOTAL PARTICIPATION OF E_2 IN R
	CARDINALITY RATIO 1:N FOR E_1:E_2 IN R
	SOME USE ARROW TO REPRESENT TOTALPARTICIPATION

ER DIAGRAM FOR A BANK DATABASE



In-class Exercise #2

- In the following Company ER diagram, find out
  - Example of weak entity?
    - Identify the PK for this weak entity
  - Example of identifying relationship?
  - What is the primary key for DEPENDENT?
  - How many super keys for DEPARTMENT?
  - What is the primary key for PROJECT?
  - How many super keys for PROJECT?
  - Modify the diagram so that an employee must have a supervisor.
  - Explain the semantics (implied by the structural constraints) on the WORKS_ON relationship



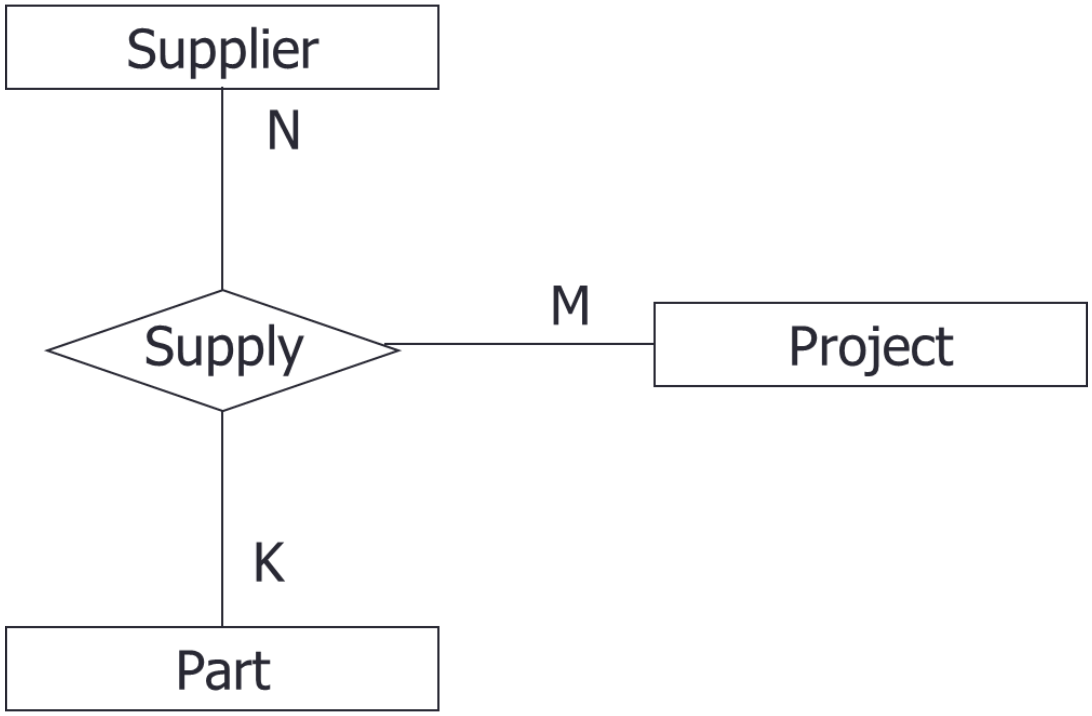
COMPANY ER Diagram

In-class Exercise #3

- Draw an ER diagram to represent:
  - A customer can borrow at most 1 loan and a loan has at least one borrower.
    - What is the cardinality for the relationship?
    - What is the participation for Loan and the participation for Customer?

Relationships of Higher Degree

- Relationship types of degree 2 are called binary
- Relationship types of degree 3 are called ternary and of degree n are called n-ary
  - AKA multiway relationship



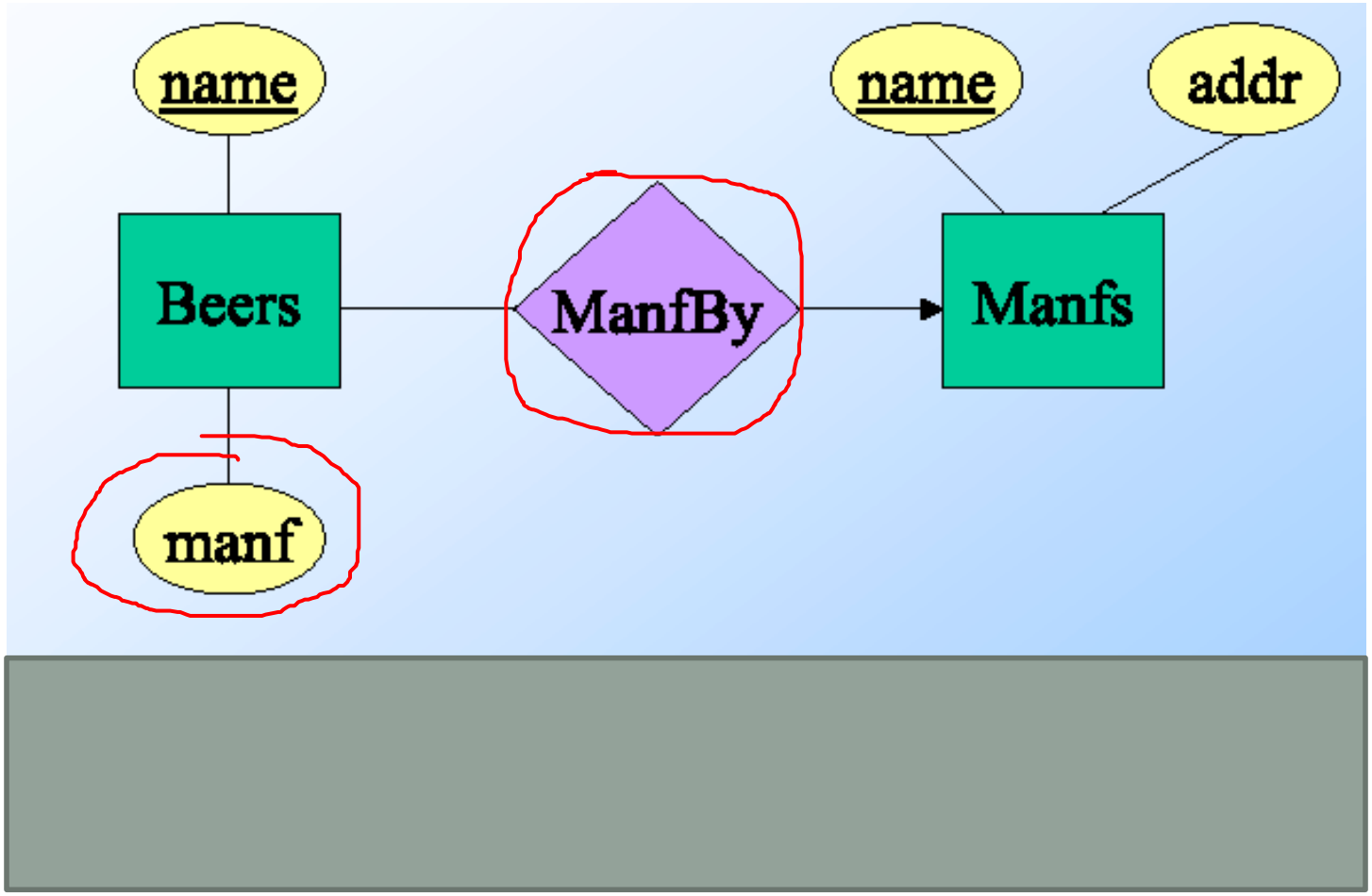
Design Choice

- How many attributes?
  - Avoid redundancy
- Attribute or Entity?
  - Don’t use an entity when an attribute will do
- Entity or relationship?
  - Difficult to answer
  - Depend on which one makes common sense
- Binary relationship or n-ary?
  - Use binary as often as possible
  - Binary is simple but might not be adequate

Avoid Redundancy

- Redundancy occurs when we say the same thing in two different ways.
- Redundancy wastes space and (more importantly) encourages inconsistency.
  - The two instances of the same fact may become inconsistent if we change one and forget to change the other, related version.

Good or Bad?

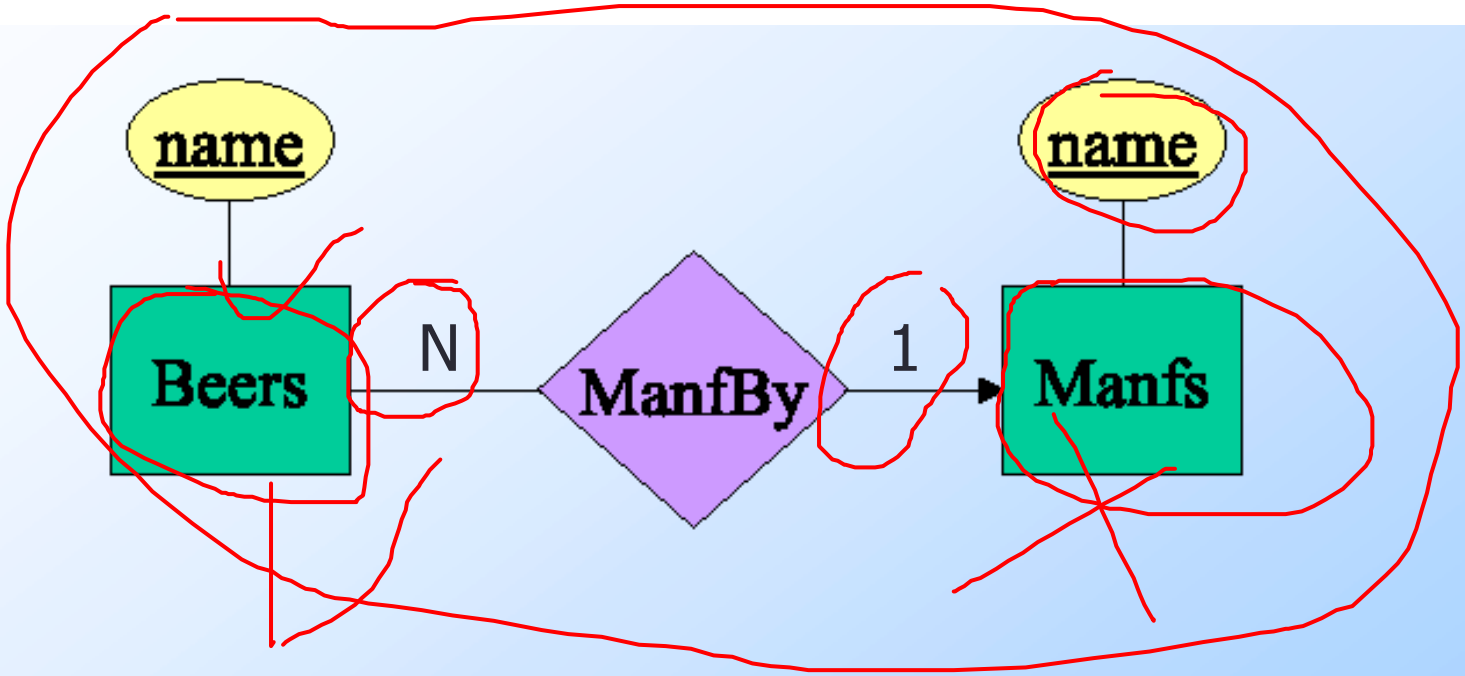


Attribute vs Entity

- An entity set should satisfy at least one of the following conditions:
  - It is more than the name of something; it has at least one nonkey attribute.

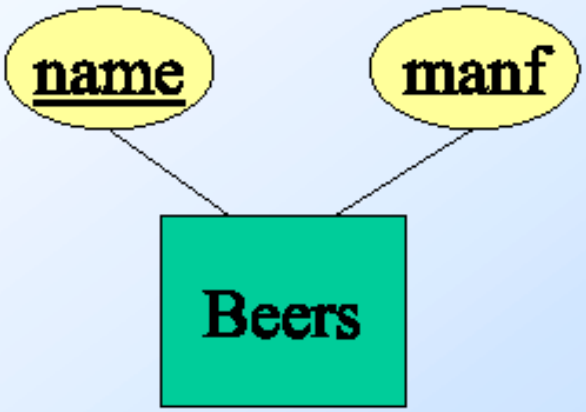
    or
  - It is the "many" in a many-one or many-many relationship.

Good or Bad?



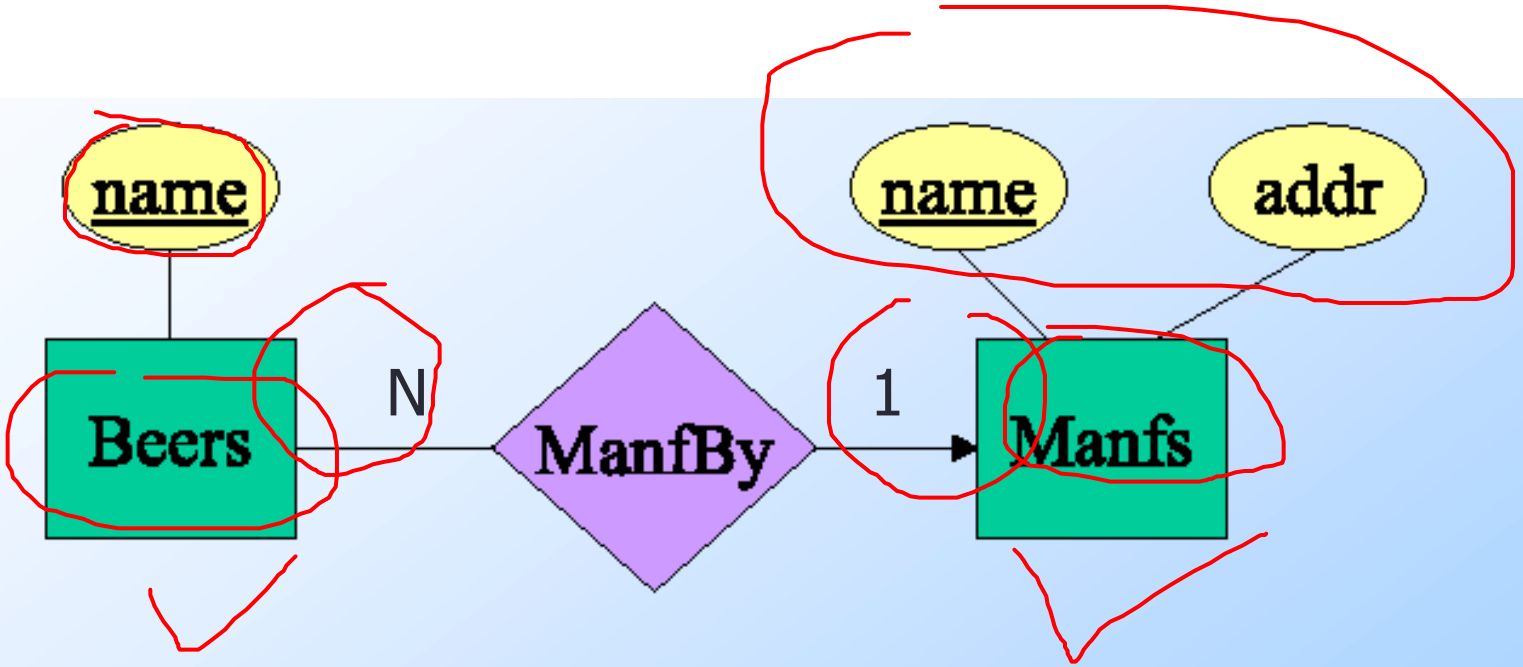
Since the manufacturer is nothing but a name, and is not at the "many" end of any relationship, it should not be an entity set.

Good or Bad



There is no need to make the manufacturer an entity set, because we record nothing about manufacturers besides their name.

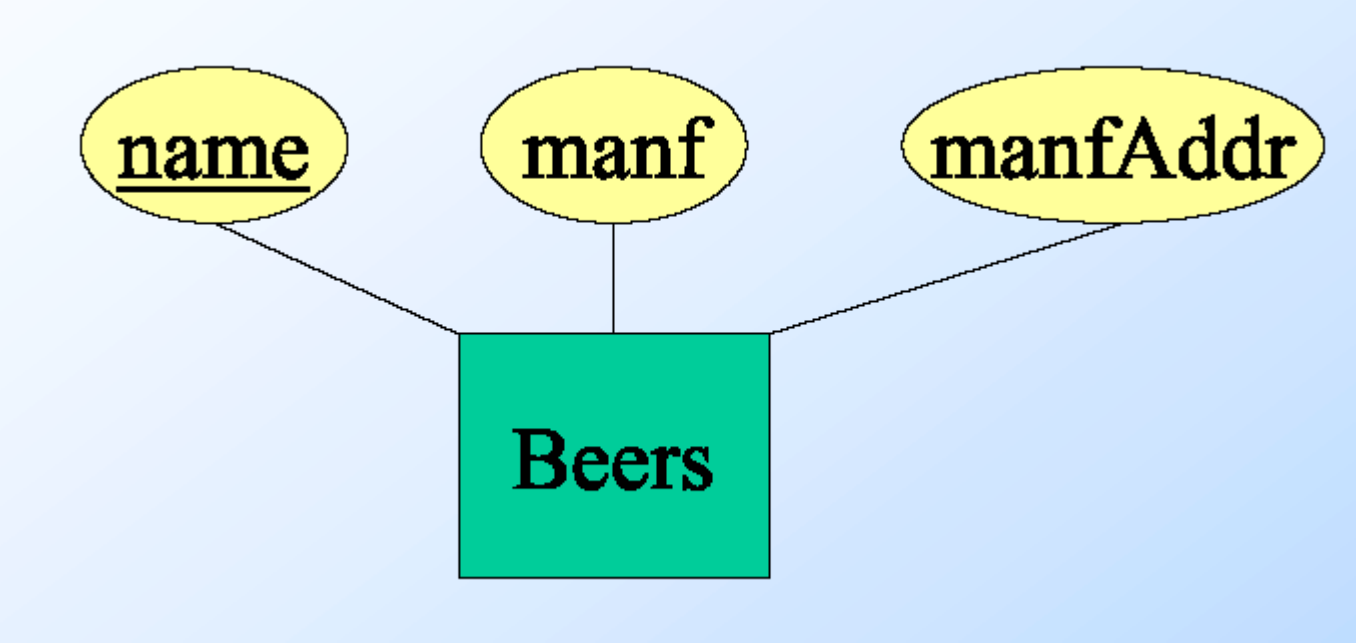
Good or Bad?



- *Manfs* deserves to be an entity set because of the nonkey attribute *addr*.
- *Beers* deserves to be an entity set because it is the "many" of the many-one relationship *ManfBy*.

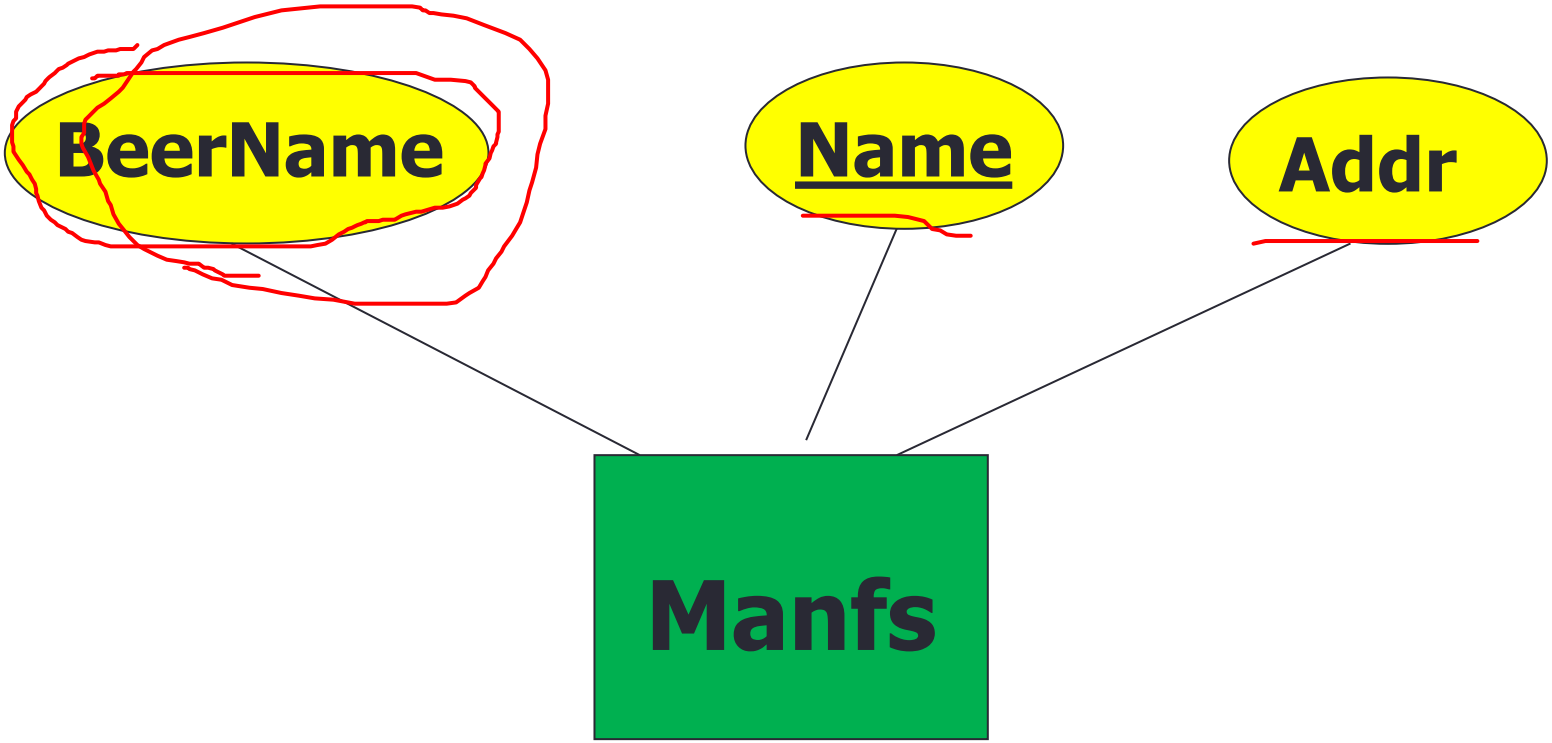
Entity or Relationship?

- Good or bad? Why?



Entity or Relationship?

- Good or bad? Why?



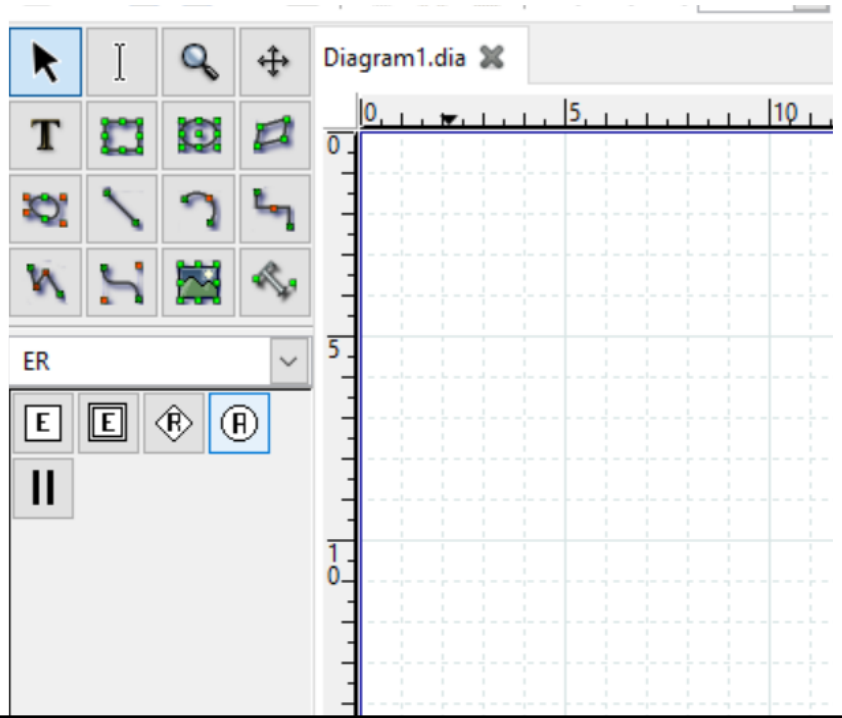
ERD Exercise

- Employees: Each employee has an employee number, a name, an address, and belongs to one department.
- Departments: Each department has a name and a department code.
- Projects: Each project has a project number. Some number of employees and a manager (also an employee) are associated with each project. For each employee it is necessary to record the amount of time he/she has worked on the project. It is possible for an employee to work on more than one project but each project only has one manager. Each project uses some parts that are explained below.
- Parts: Each part has a part number and a description. In addition, it is necessary to represent the quantity and types of subcomponents (other parts), which make up each part as well as the parts of which it is a component.
- Warehouses: Each warehouse has a number, an address, and contains an inventory of parts. The quantity of each in stock in each warehouse must be recorded.
- Suppliers: Each supplier has a number, a name, and an address. It is necessary to record the parts that a supplier can potentially supply as well as the parts that he/she is currently supplying. It must also be possible to identify for each project, which supplier is supplying which parts and in what quantities.

Free Diagraming Tool

- Dia: http://dia-installer.de/download/index.html.en

- If you want to draw ER diagram in Dia, make sure to choose ER from the drop down menu.



Review of Today's Lecture

- Entity
- Weak entity
- Attributes
- Relationship
- Structural constraints