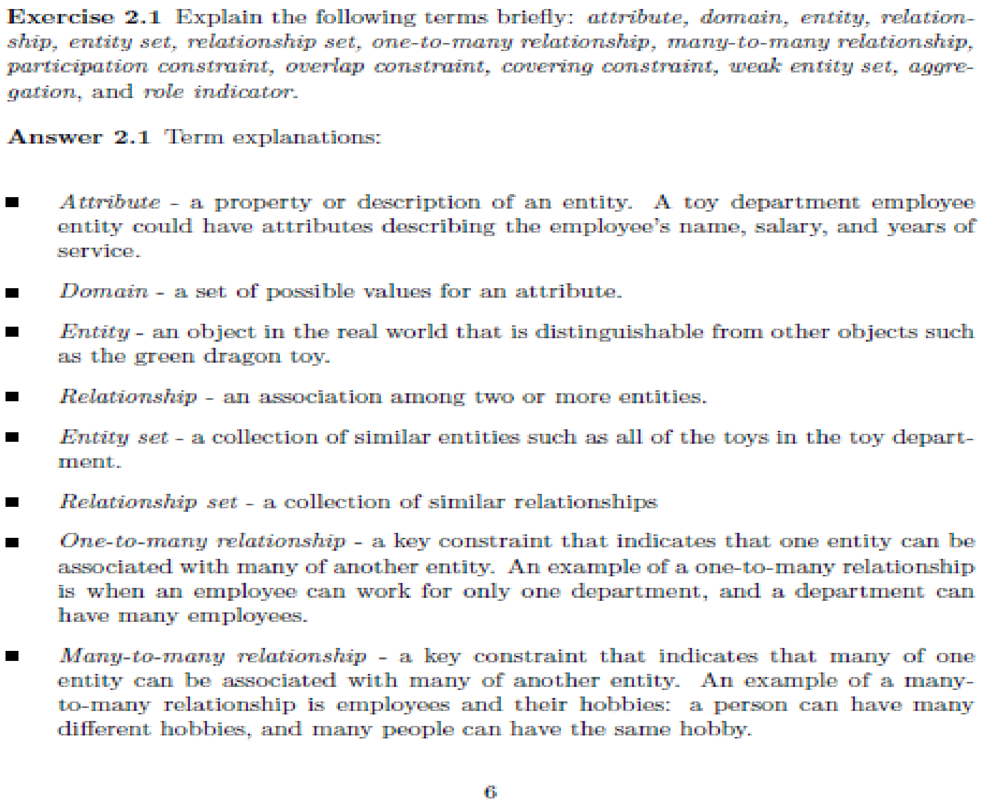
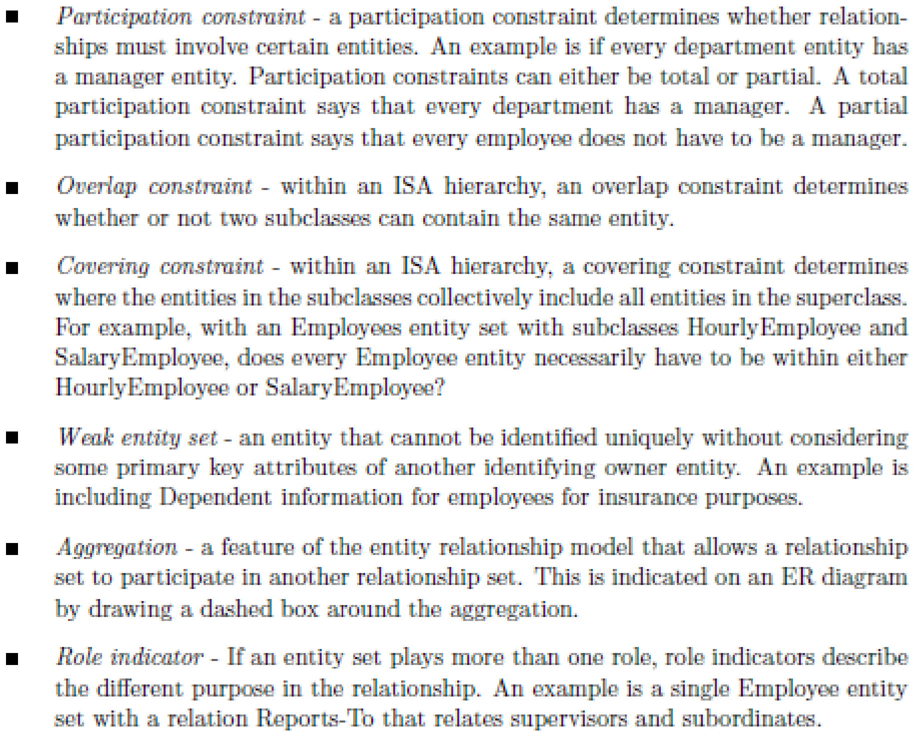


Exercise 2.1 Explain the following terms briefly: *attribute, domain, entity, relationship, entity set, relationship set, one-to-many relationship, many-to-many relationship, participation constraint, overlap constraint, covering constraint, weak entity set, aggregation*, and *role indicator*.

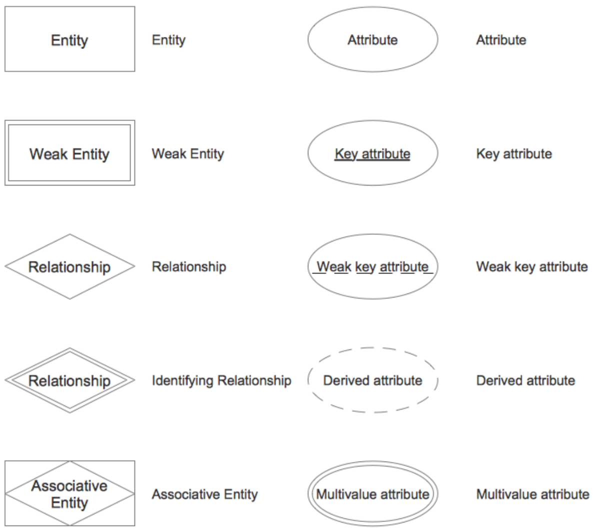
Answer 2.1 Term explanations:

- *Attribute* - a property or description of an entity. A toy department employee entity could have attributes describing the employee's name, salary, and years of service.
- *Domain* - a set of possible values for an attribute.
- *Entity* - an object in the real world that is distinguishable from other objects such as the green dragon toy.
- *Relationship* - an association among two or more entities.
- *Entity set* - a collection of similar entities such as all of the toys in the toy department.
- *Relationship set* - a collection of similar relationships
- *One-to-many relationship* - a key constraint that indicates that one entity can be associated with many of another entity. An example of a one-to-many relationship is when an employee can work for only one department, and a department can have many employees.
- *Many-to-many relationship* - a key constraint that indicates that many of one entity can be associated with many of another entity. An example of a many-to-many relationship is employees and their hobbies: a person can have many different hobbies, and many people can have the same hobby.



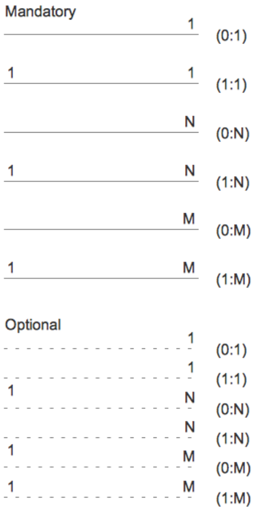
- *Participation constraint* - a participation constraint determines whether relationships must involve certain entities. An example is if every department entity has a manager entity. Participation constraints can either be total or partial. A total participation constraint says that every department has a manager. A partial participation constraint says that every employee does not have to be a manager.
- *Overlap constraint* - within an ISA hierarchy, an overlap constraint determines whether or not two subclasses can contain the same entity.
- *Covering constraint* - within an ISA hierarchy, a covering constraint determines where the entities in the subclasses collectively include all entities in the superclass. For example, with an Employees entity set with subclasses HourlyEmployee and SalaryEmployee, does every Employee entity necessarily have to be within either HourlyEmployee or SalaryEmployee?
- *Weak entity set* - an entity that cannot be identified uniquely without considering some primary key attributes of another identifying owner entity. An example is including Dependent information for employees for insurance purposes.
- *Aggregation* - a feature of the entity relationship model that allows a relationship set to participate in another relationship set. This is indicated on an ER diagram by drawing a dashed box around the aggregation.
- *Role indicator* - If an entity set plays more than one role, role indicators describe the different purpose in the relationship. An example is a single Employee entity set with a relation Reports-To that relates supervisors and subordinates.

Chen's notation



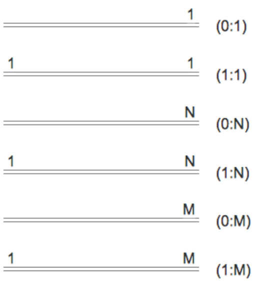
Participations

Cardinality can be shown or hidden

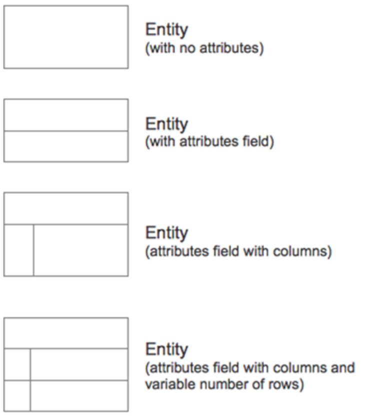


Recursive Relationship

Cardinality can be shown or hidden

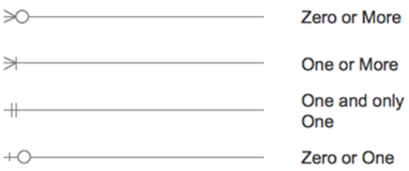


Crow's Foot notation

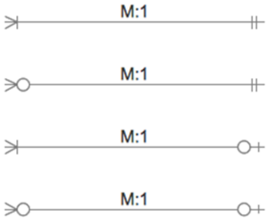

Entity
(with no attributes)
Entity
(with attributes field)
Entity
(attributes field with columns)
Entity
(attributes field with columns and variable number of rows)

Relationships

(Cardinality and Modality)



Many - to - One



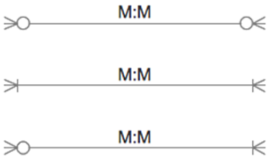
a one through many notation on one side of a relationship and a one and only one on the other

a zero through many notation on one side of a relationship and a one and only one on the other

a one through many notation on one side of a relationship and a zero or one notation on the other

a zero through many notation on one side of a relationship and a zero or one notation on the other

Many-to-Many

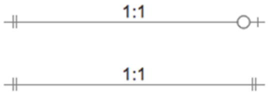


a zero through many on both sides of a relationship

a one through many on both sides of a relationship

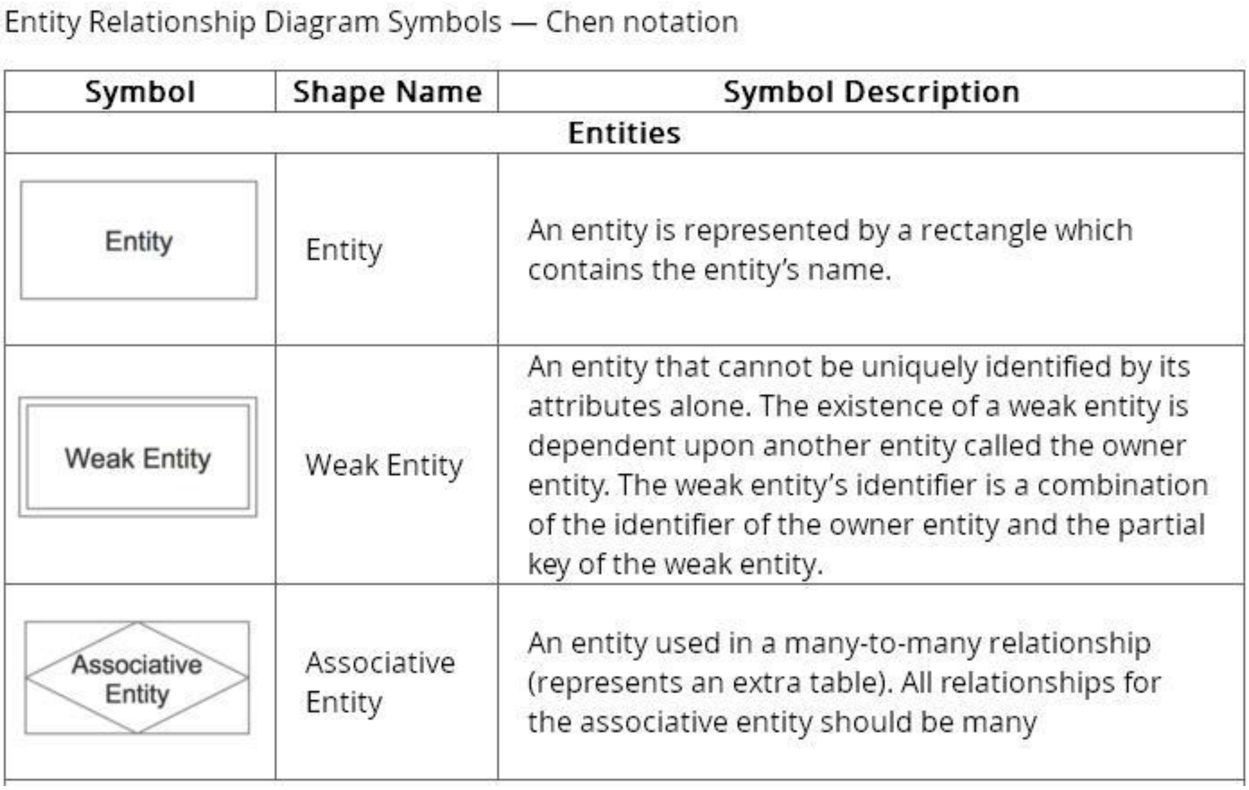
a zero through many on one side and a one through many on the other

Many-to-Many



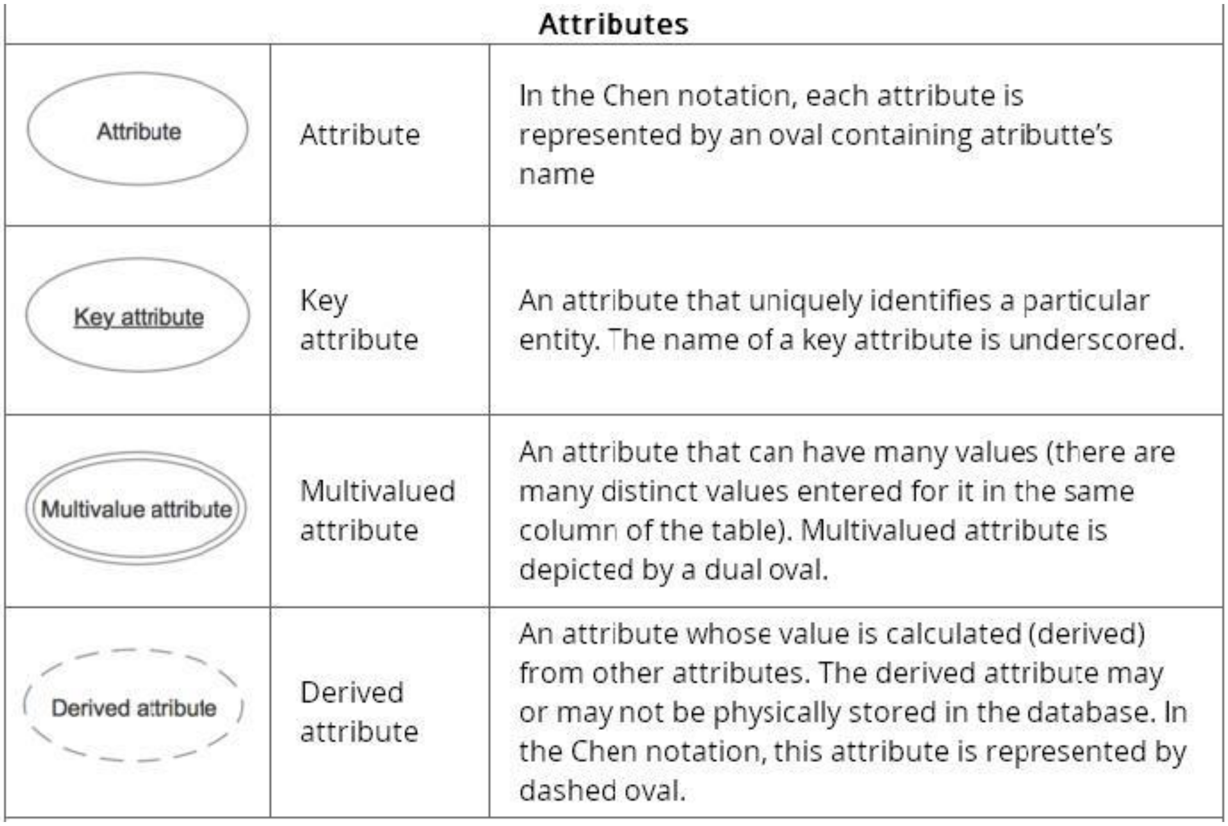
a one and only one notation on one side of a relationship and a zero or one on the other

a one and only one notation on both sides

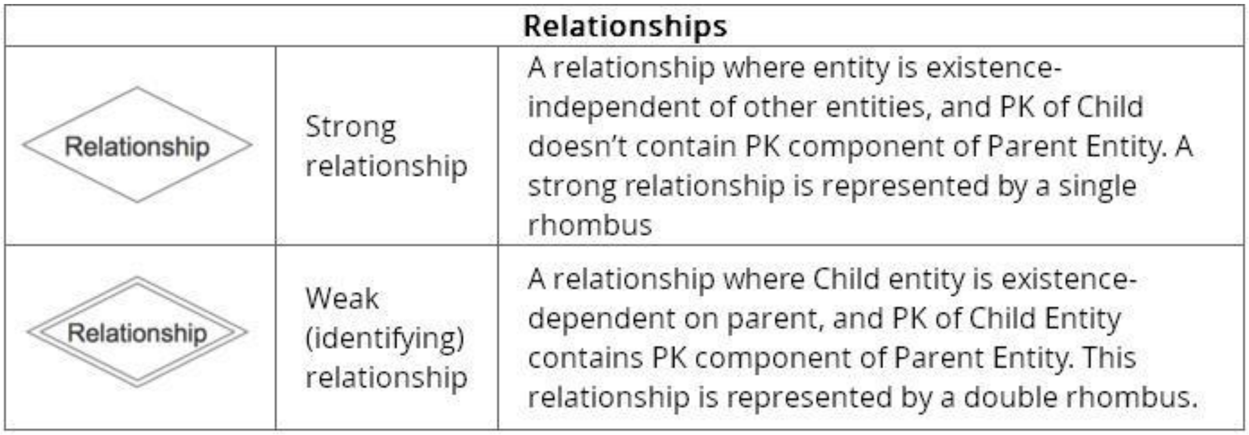


Entity Relationship Diagram Symbols — Chen notation

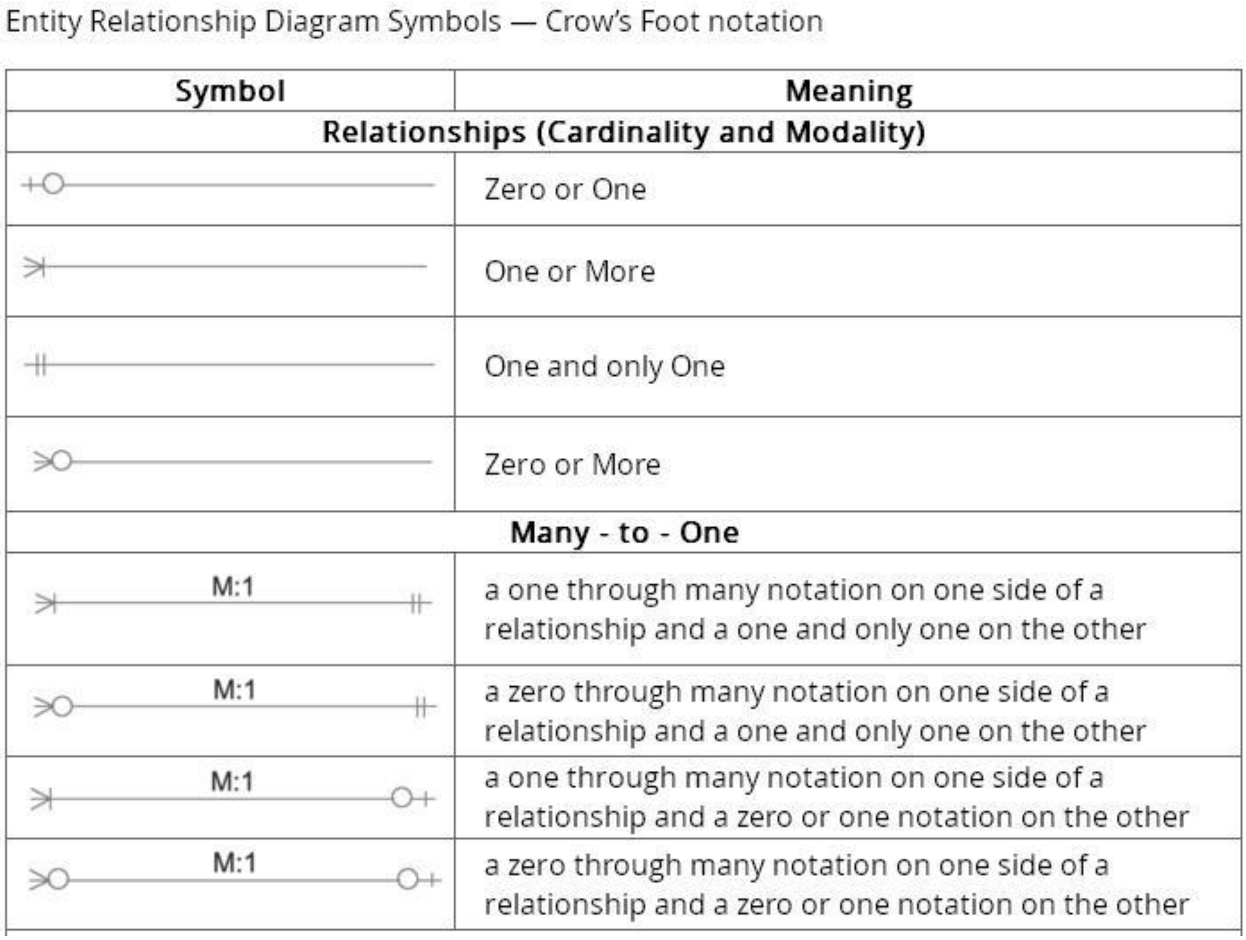
Symbol	Shape Name	Symbol Description
		Entities
Entity	Entity	An entity is represented by a rectangle which contains the entity's name.
Weak Entity	Weak Entity	An entity that cannot be uniquely identified by its attributes alone. The existence of a weak entity is dependent upon another entity called the owner entity. The weak entity's identifier is a combination of the identifier of the owner entity and the partial key of the weak entity.
Associative Entity	Associative Entity	An entity used in a many-to-many relationship (represents an extra table). All relationships for the associative entity should be many



		Attributes
Attribute	Attribute	In the Chen notation, each attribute is represented by an oval containing atributte's name
Key attribute	Key attribute	An attribute that uniquely identifies a particular entity. The name of a key attribute is underscored.
Multivalue attribute	Multivalued attribute	An attribute that can have many values (there are many distinct values entered for it in the same column of the table). Multivalued attribute is depicted by a dual oval.
Derived attribute	Derived attribute	An attribute whose value is calculated (derived) from other attributes. The derived attribute may or may not be physically stored in the database. In the Chen notation, this attribute is represented by dashed oval.

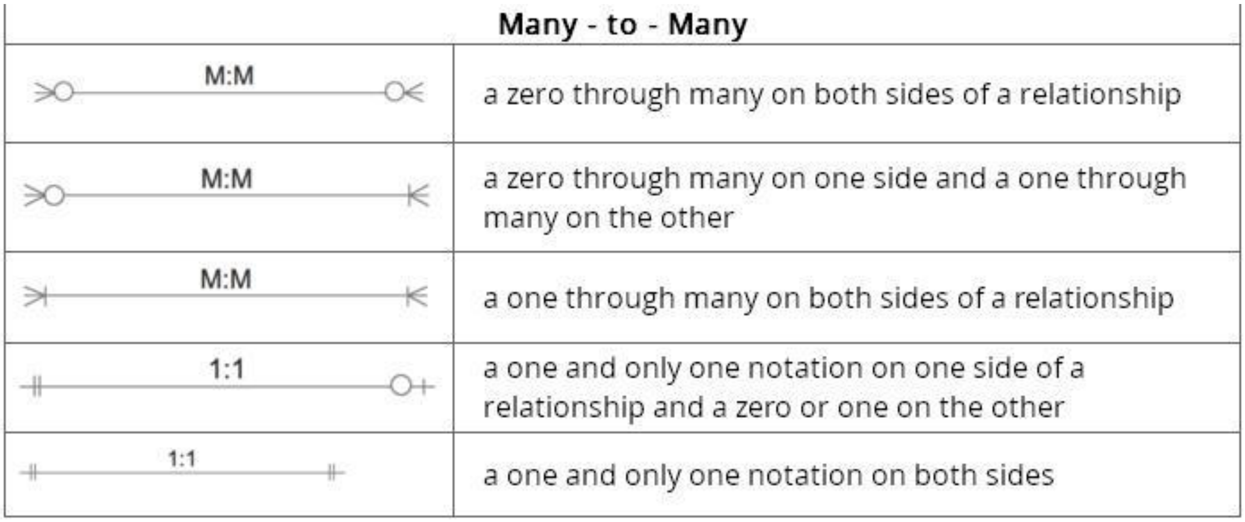


Relationships		
Relationship	Strong relationship	A relationship where entity is existence-independent of other entities, and PK of Child doesn't contain PK component of Parent Entity. A strong relationship is represented by a single rhombus
Relationship	Weak (identifying) relationship	A relationship where Child entity is existence-dependent on parent, and PK of Child Entity contains PK component of Parent Entity. This relationship is represented by a double rhombus.

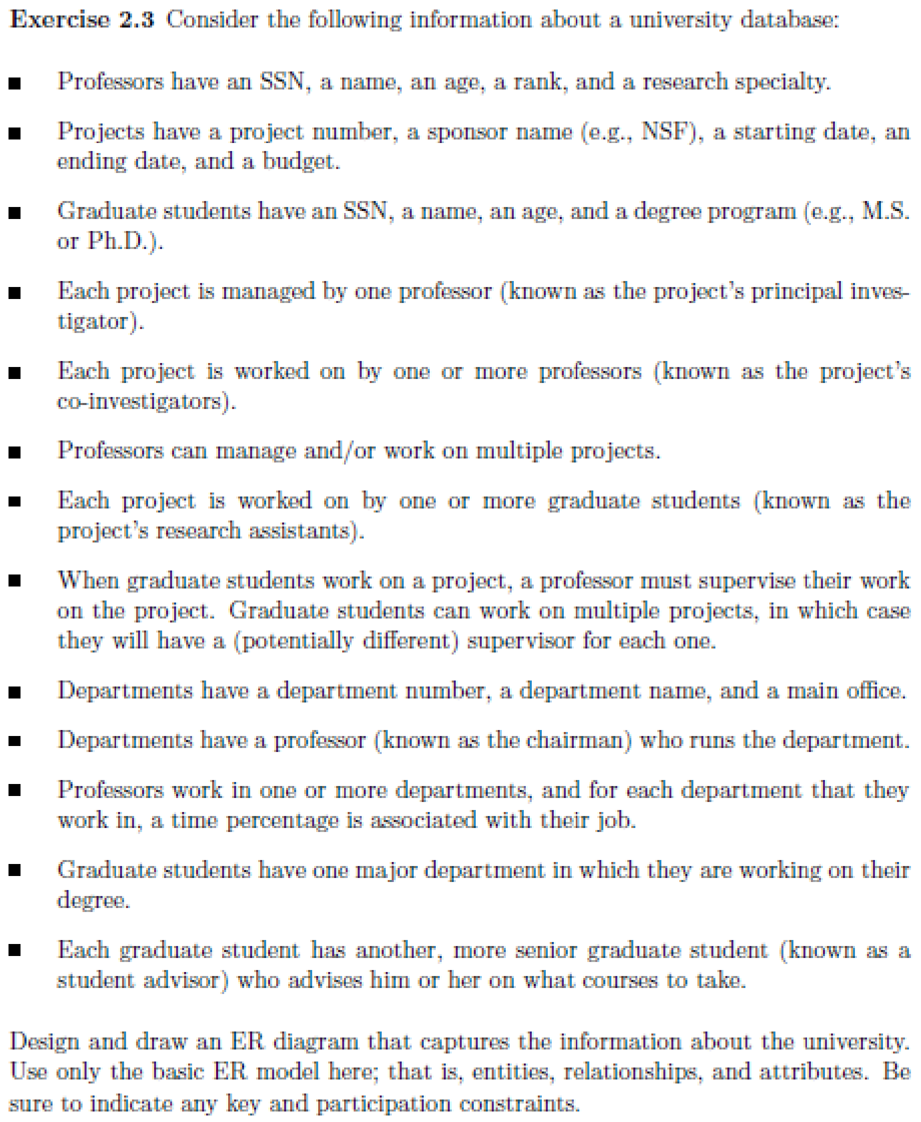


Entity Relationship Diagram Symbols — Crow's Foot notation

Symbol	Meaning
Relationships (Cardinality and Modality)	
+O———	Zero or One
>+———	One or More
++———	One and only One
>O———	Zero or More
Many - to - One	
>+——— M:1 ———++	a one through many notation on one side of a relationship and a one and only one on the other
>O——— M:1 ———++	a zero through many notation on one side of a relationship and a one and only one on the other
>+——— M:1 ———O+	a one through many notation on one side of a relationship and a zero or one notation on the other
>O——— M:1 ———O+	a zero through many notation on one side of a relationship and a zero or one notation on the other



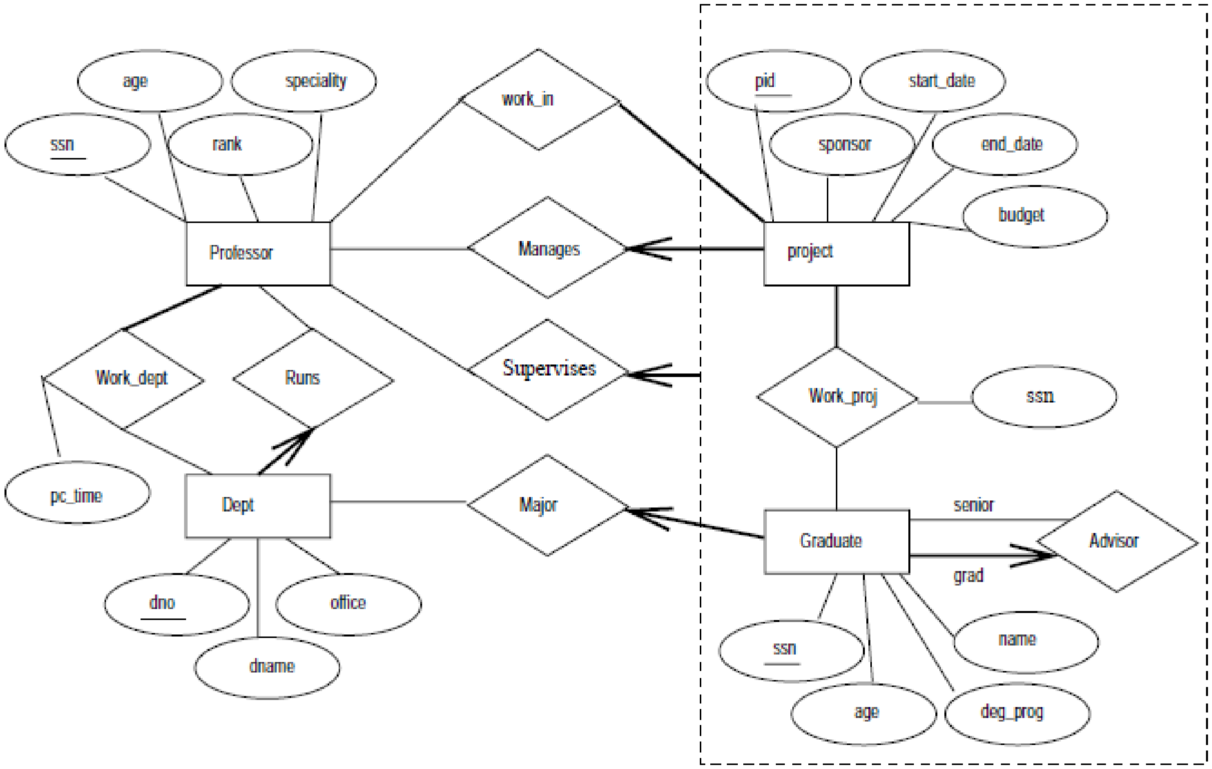
Many - to - Many	
M:M	a zero through many on both sides of a relationship
M:M	a zero through many on one side and a one through many on the other
M:M	a one through many on both sides of a relationship
1:1	a one and only one notation on one side of a relationship and a zero or one on the other
1:1	a one and only one notation on both sides



Exercise 2.3 Consider the following information about a university database:

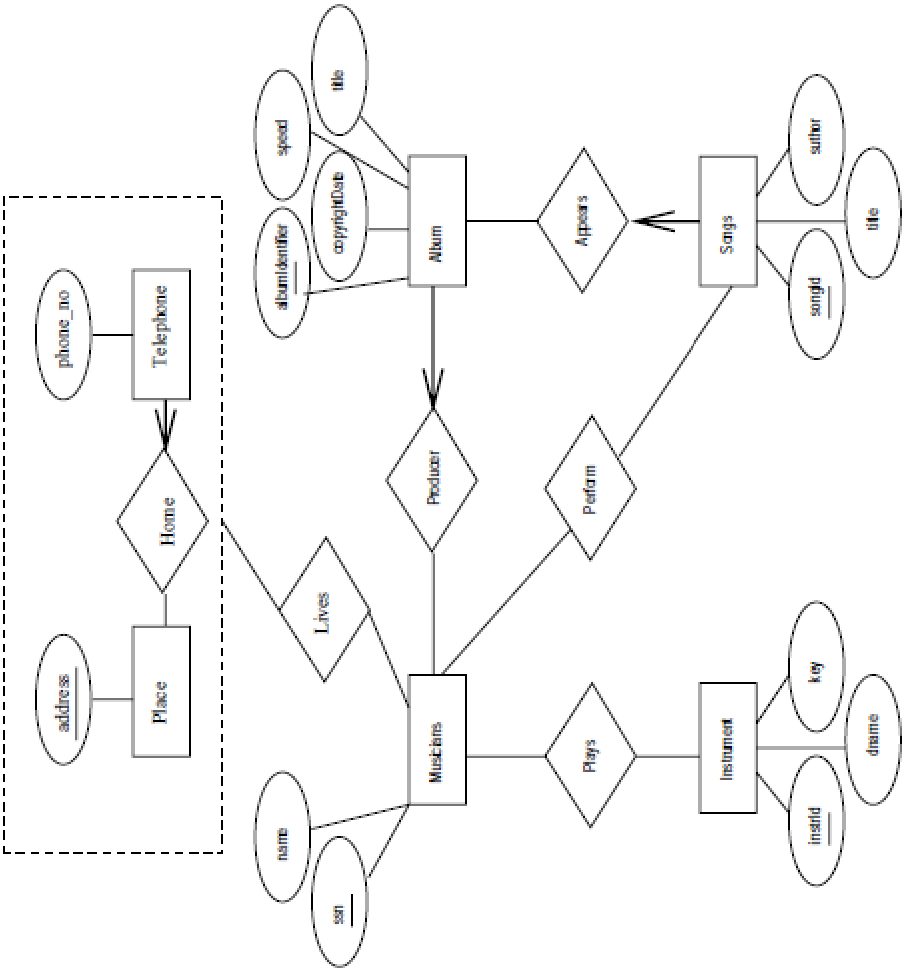
- Professors have an SSN, a name, an age, a rank, and a research specialty.
- Projects have a project number, a sponsor name (e.g., NSF), a starting date, an ending date, and a budget.
- Graduate students have an SSN, a name, an age, and a degree program (e.g., M.S. or Ph.D.).
- Each project is managed by one professor (known as the project's principal investigator).
- Each project is worked on by one or more professors (known as the project's co-investigators).
- Professors can manage and/or work on multiple projects.
- Each project is worked on by one or more graduate students (known as the project's research assistants).
- When graduate students work on a project, a professor must supervise their work on the project. Graduate students can work on multiple projects, in which case they will have a (potentially different) supervisor for each one.
- Departments have a department number, a department name, and a main office.
- Departments have a professor (known as the chairman) who runs the department.
- Professors work in one or more departments, and for each department that they work in, a time percentage is associated with their job.
- Graduate students have one major department in which they are working on their degree.
- Each graduate student has another, more senior graduate student (known as a student advisor) who advises him or her on what courses to take.

Design and draw an ER diagram that captures the information about the university. Use only the basic ER model here; that is, entities, relationships, and attributes. Be sure to indicate any key and participation constraints.



Exercise 2.5 Notown Records has decided to store information about musicians who perform on its albums (as well as other company data) in a database. The company has wisely chosen to hire you as a database designer (at your usual consulting fee of $2500/day).

- Each musician that records at Notown has an SSN, a name, an address, and a phone number. Poorly paid musicians often share the same address, and no address has more than one phone.
- Each instrument used in songs recorded at Notown has a unique identification number, a name (e.g., guitar, synthesizer, flute) and a musical key (e.g., C, B-flat, E-flat).
- Each album recorded on the Notown label has a unique identification number, a title, a copyright date, a format (e.g., CD or MC), and an album identifier.
- Each song recorded at Notown has a title and an author.
- Each musician may play several instruments, and a given instrument may be played by several musicians.
- Each album has a number of songs on it, but no song may appear on more than one album.
- Each song is performed by one or more musicians, and a musician may perform a number of songs.
- Each album has exactly one musician who acts as its producer. A musician may produce several albums, of course.

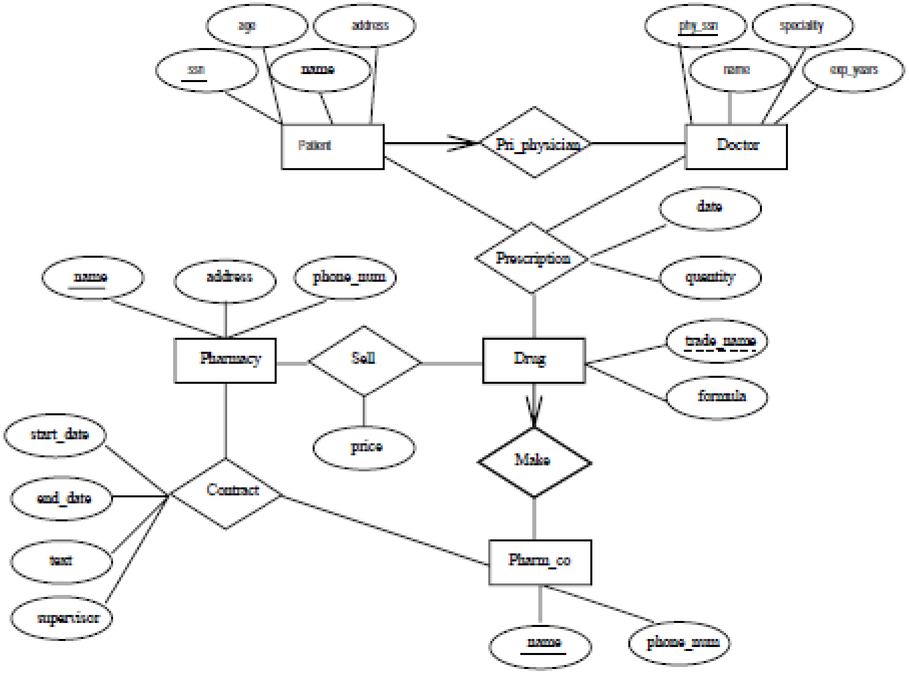


phone_no
Telephone
Home
Place
address
Lives
albumIdentifier
copyrightDate
speed
title
Album
Appears
Songs
author
title
songId
Producer
Perform
Musicians
name
ssn
Plays
Instrument
key
dname
instrId

Exercise 2.7 The Prescriptions-R-X chain of pharmacies has offered to give you a free lifetime supply of medicine if you design its database. Given the rising cost of health care, you agree. Here's the information that you gather:

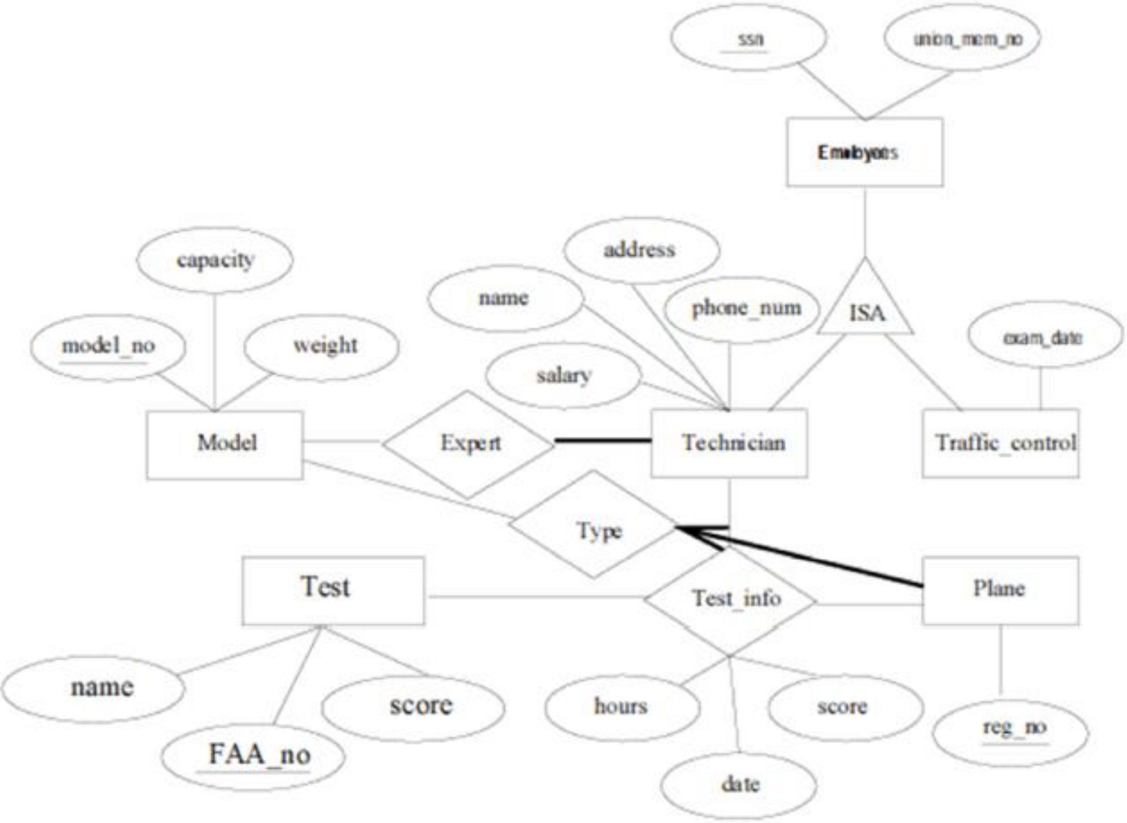
- Patients are identified by an SSN, and their names, addresses, and ages must be recorded.
- Doctors are identified by an SSN. For each doctor, the name, specialty, and years of experience must be recorded.
- Each pharmaceutical company is identified by name and has a phone number.
- For each drug, the trade name and formula must be recorded. Each drug is sold by a given pharmaceutical company, and the trade name identifies a drug uniquely from among the products of that company. If a pharmaceutical company is deleted, you need not keep track of its products any longer.
- Each pharmacy has a name, address, and phone number.
- Every patient has a primary physician. Every doctor has at least one patient.
- Each pharmacy sells several drugs and has a price for each. A drug could be sold at several pharmacies, and the price could vary from one pharmacy to another.
- Doctors prescribe drugs for patients. A doctor could prescribe one or more drugs for several patients, and a patient could obtain prescriptions from several doctors. Each prescription has a date and a quantity associated with it. You can assume that, if a doctor prescribes the same drug for the same patient more than once, only the last such prescription needs to be stored.
- Pharmaceutical companies have long-term contracts with pharmacies. A pharmaceutical company can contract with several pharmacies, and a pharmacy can contract with several pharmaceutical companies. For each contract, you have to store a start date, an end date, and the text of the contract.
- Pharmacies appoint a supervisor for each contract. There must always be a supervisor for each contract, but the contract supervisor can change over the lifetime of the contract.

1. Draw an ER diagram that captures the preceding information. Identify any constraints not captured by the ER diagram.
2. How would your design change if each drug must be sold at a fixed price by all pharmacies?
3. How would your design change if the design requirements change as follows: If a doctor prescribes the same drug for the same patient more than once, several such prescriptions may have to be stored.



Computer Sciences Department frequent fliers have been complaining to Dane County Airport officials about the poor organization at the airport. As a result, the officials decided that all information related to the airport should be organized using a DBMS, and you have been hired to design the database. Your first task is to organize the information about all the airplanes stationed and maintained at the airport. The relevant information is as follows:

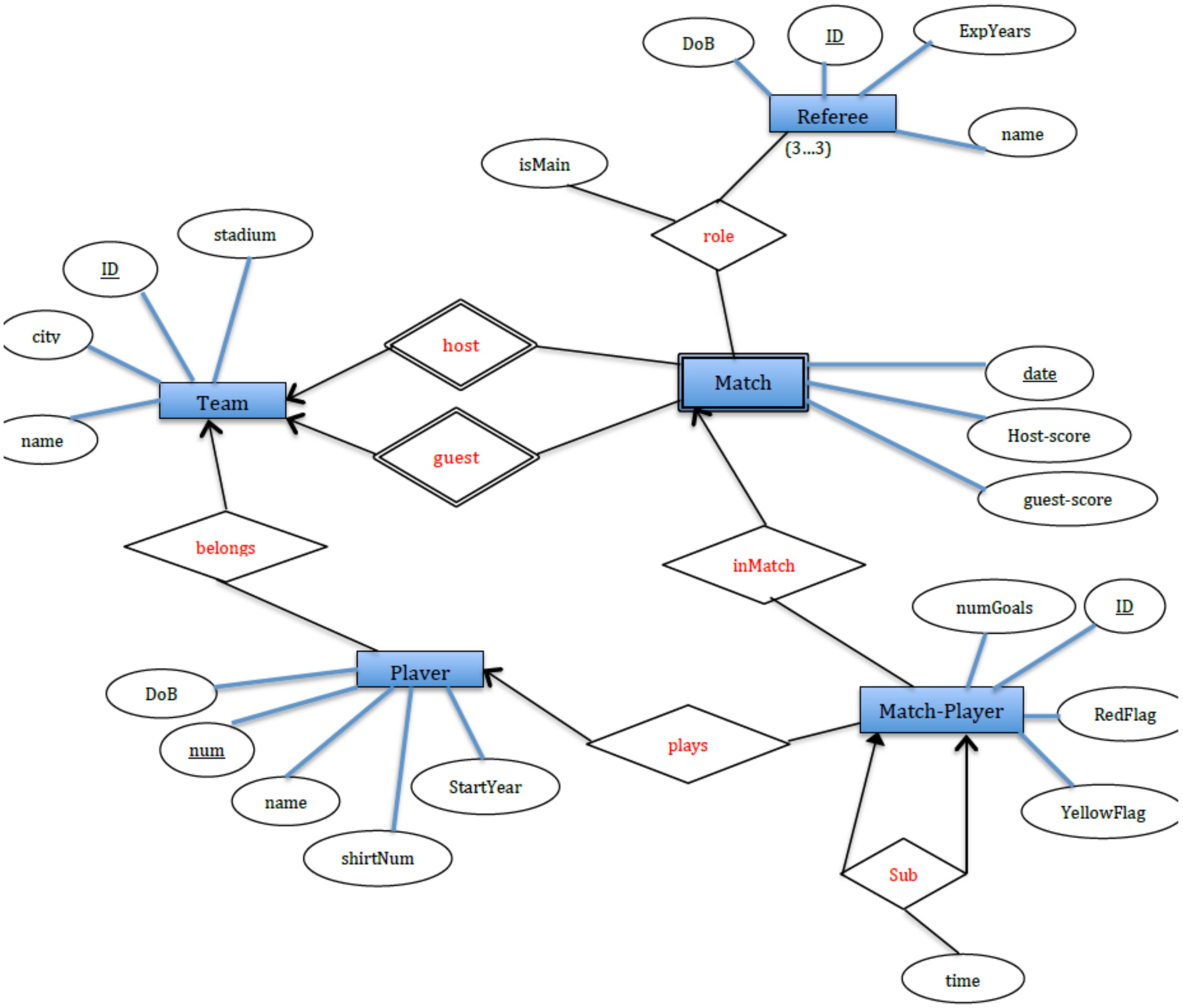
- Every airplane has a registration number, and each airplane is of a specific model.
- The airport accommodates a number of airplane models, and each model is identified by a model number (e.g., DC-10) and has a capacity and a weight.
- A number of technicians work at the airport. You need to store the name, SSN, address, phone number, and salary of each technician.
- Each technician is an expert on one or more plane model(s), and his or her expertise may overlap with that of other technicians. This information about technicians must also be recorded.
- Traffic controllers must have an annual medical examination. For each traffic controller, you must store the date of the most recent exam.
- All airport employees (including technicians) belong to a union. You must store the union membership number of each employee. You can assume that each employee is uniquely identified by a social security number.
- The airport has a number of tests that are used periodically to ensure that airplanes are still airworthy. ach test has a Federal Aviation Administration (FAA) test number, a name, and a maximum possible score.
- The FAA requires the airport to keep track of each time a given airplane is tested by a given technician using a given test. For each testing event, the information needed is the date, the number of hours the technician spent doing the test, and the score the airplane received on the test.



Assume we have the following application that models soccer teams, the games they play, and the players in each team. In the design, we want to capture the following:

- We have a set of teams, each team has an ID (unique identifier), name, main stadium, and to which city this team belongs.
- Each team has many players, and each player belongs to one team. Each player has a number (unique identifier), name, DoB, start year, and shirt number that he uses.
- Teams play matches, in each match there is a host team and a guest team. The match takes place in the stadium of the host team.
- For each match we need to keep track of the following:
    - The date on which the game is played
    - The final result of the match
    - The players participated in the match. For each player, how many goals he scored, whether or not he took yellow card, and whether or not he took red card.
    - During the match, one player may substitute another player. We want to capture this substitution and the time at which it took place.
- Each match has exactly three referees. For each referee we have an ID (unique identifier), name, DoB, years of experience. One referee is the main referee and the other two are assistant referee.

Design an ER diagram to capture the above requirements. State any assumptions you have that affects your design (use the back of the page if needed). Make sure cardinalities and primary keys are clear.



Assumptions:

1- In Match-Player entity set, we added a unique identifier for each record ID.
2- The final result in Match entity set is captured using two attributes Host-score and guest-score
3- The attribute 'isMain' in relationship 'role' is true if the referee is the main referee in the match, otherwise, it will be false.

Map the ERD in Question 1 to create the relational model corresponding to the described application. Basically, list the CREATE TABLE statements with the attribute names, and appropriate data types. Also make sure to have the primary keys and foreign keys clearly defined (use the back of the page if needed).

```
Create Table Team (
 ID: int Primary Key,
City: varchar(100),
Name: varchar(100),
Stadium: varchar(100));
```

```
Create Table Player (
 num: int Primary Key,
DoB: date,
Name: varchar(100),
StartYear: int,
ShirtNum: int,
TeamID: int Foreign Key References Team(ID));
```

```
Create Table Referee (
 ID: int Primary Key,
DoB: date,
Name: varchar(100),
ExpYear: int);
```

```
Create Table Match (
 HostID: int Foreign Key References Team(ID),
GuestID: int Foreign Key References Team(ID),
Date: date,
Host-score: int,
Guest-score: int,
Primary Key (HostID, GuestID, Date));
```

```
Create Table RefereeRole (
 HostID: int,
GuestID: int,
Date: date,
RefID: int Foreign Key References Referee(ID),
isMain: Boolean,
Foreign Key (HostID, GuestID, Date) References Match (HostID, GuestID, Date),
Primary Key (HostID, GuestID, Date, RefID));
```

```
Create Table Match-Player (
 ID: int Primary Key,
PlayerNum: int Foreign Key References Player(num),
MatchDate: date,
HostID: int,
GuestID: int,
numGoals: int,
redFlag: Boolean,
yellowFlag: Boolean,
subID: int Foreign Key References Match-Player(ID),
subTime: int,
Foreign Key (HostID, GuestID, MatchDate) References Match (HostID, GuestID, Date));
```

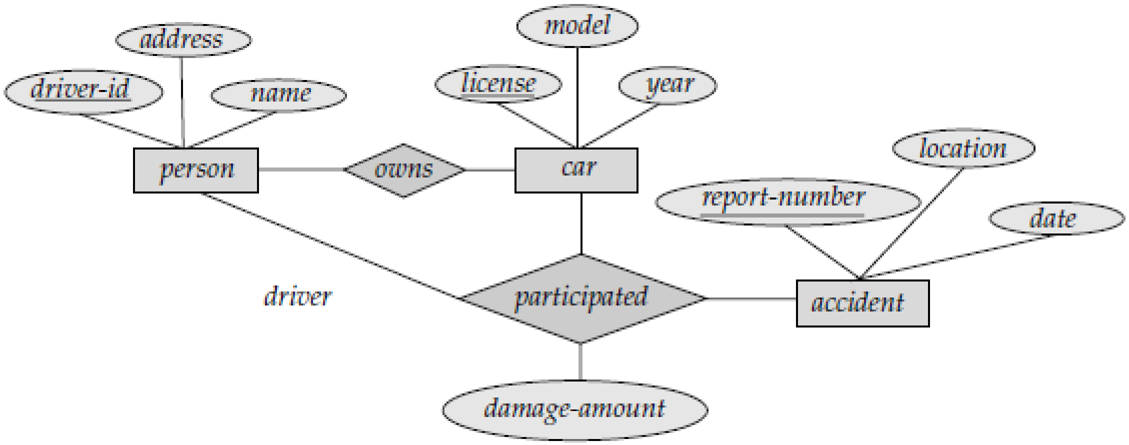


Figure 2.1 E-R diagram for a Car-insurance company.

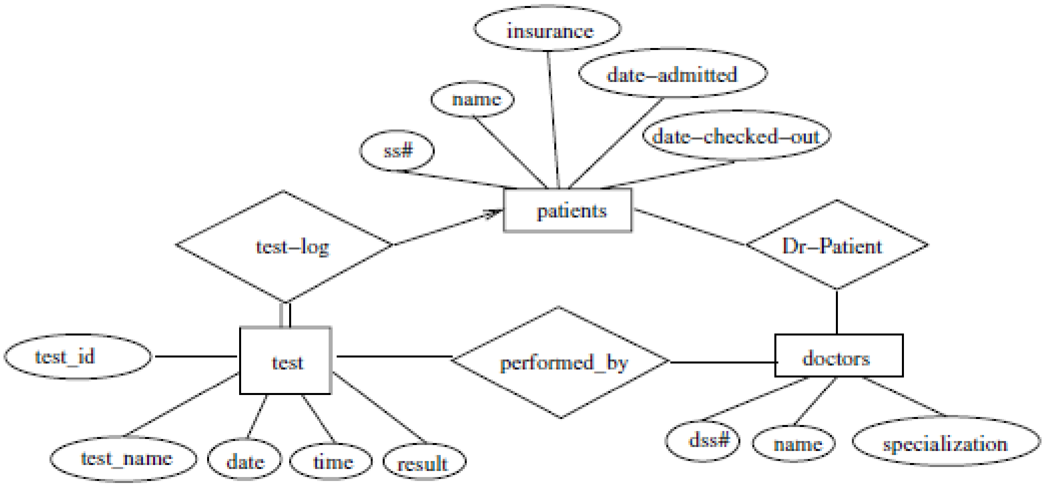


Figure 2.2 E-R diagram for a hospital.

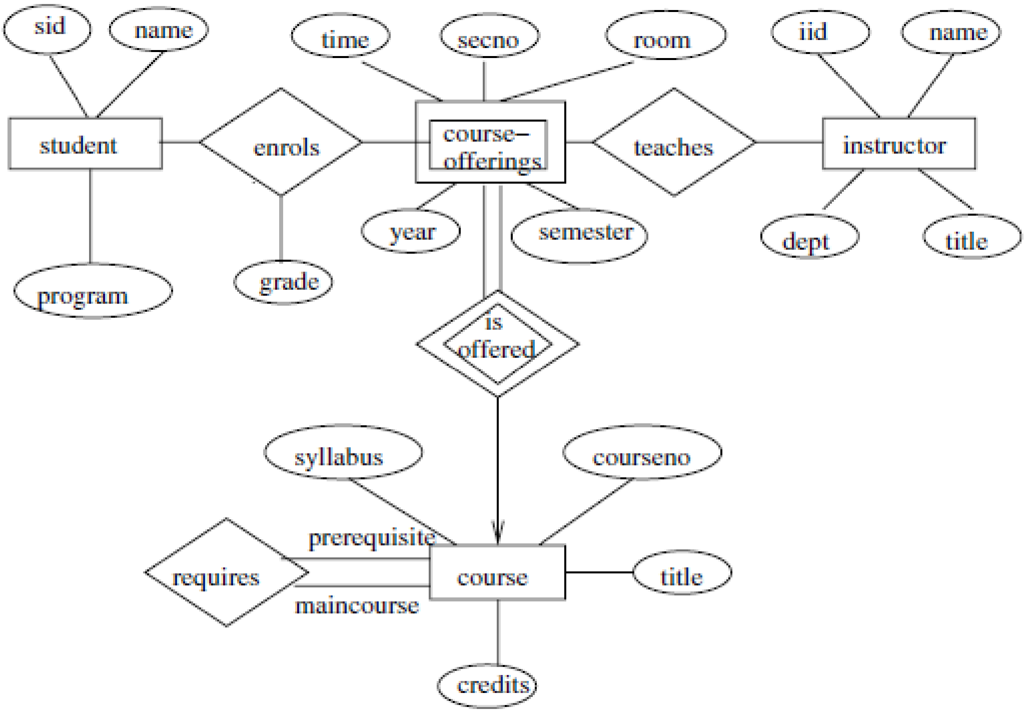


Figure 2.3 E-R diagram for a university.

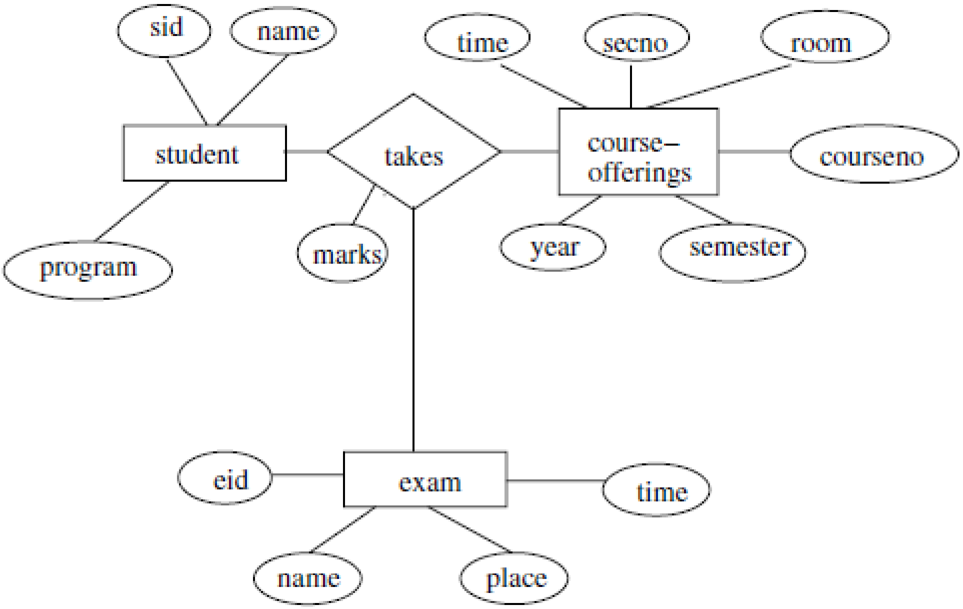


Figure 2.4 E-R diagram for marks database.

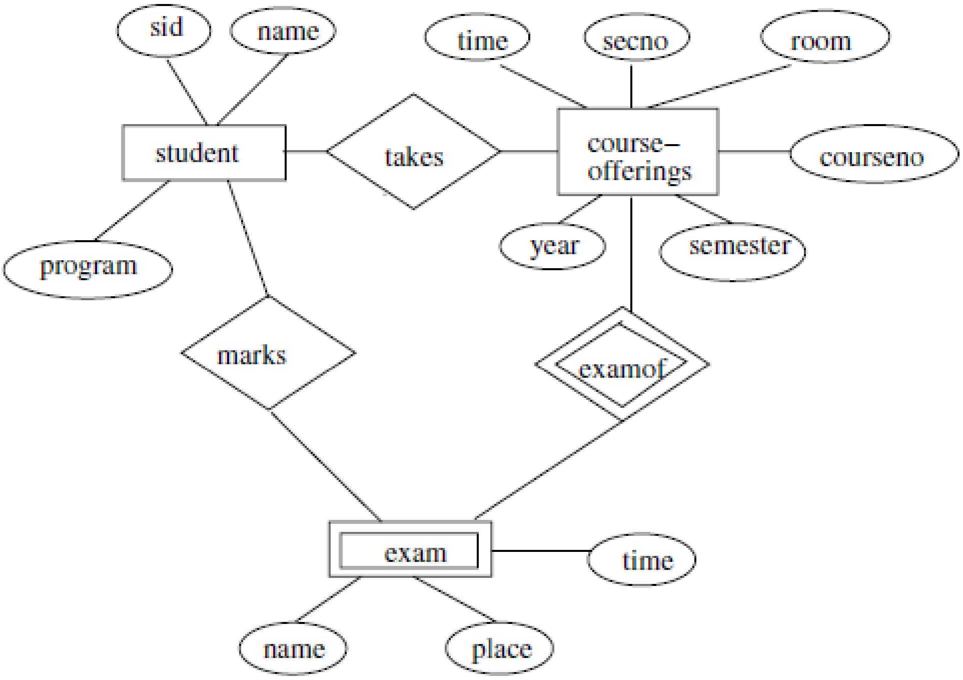


Figure 2.5 Another E-R diagram for marks database.

Explain the difference between a weak and a strong entity set.

Answer: A strong entity set has a primary key. All tuples in the set are distinguishable by that key. A weak entity set has no primary key unless attributes of the strong entity set on which it depends are included. Tuples in a weak entity set are partitioned according to their relationship with tuples in a strong entity

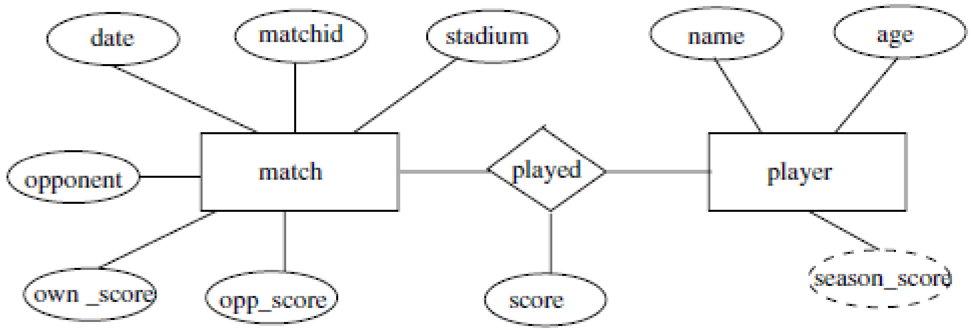


Figure 2.6 E-R diagram for favourite team statistics.

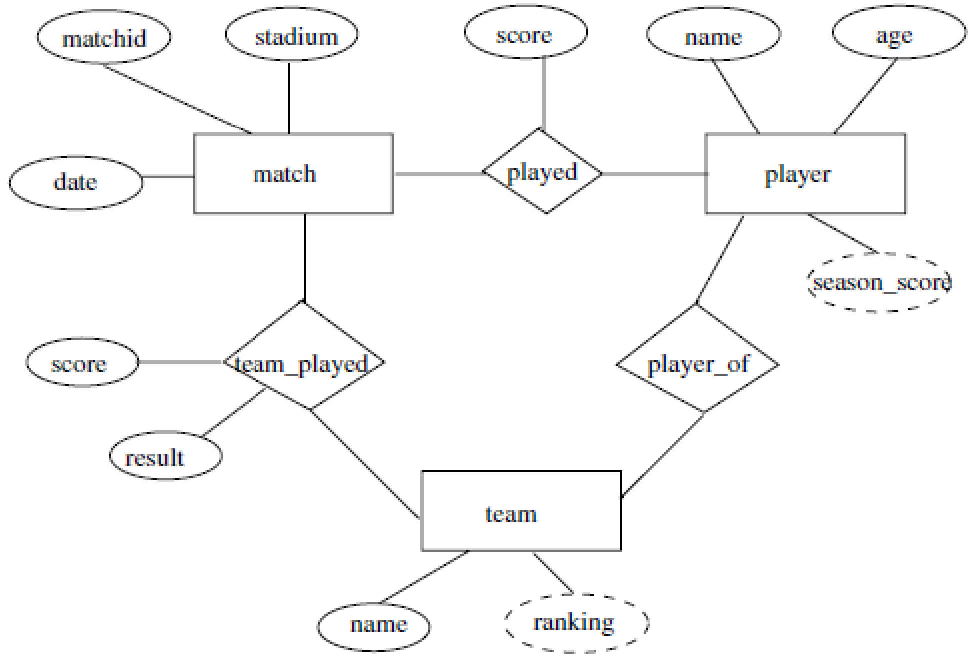


Figure 2.7 E-R diagram for all teams statistics.

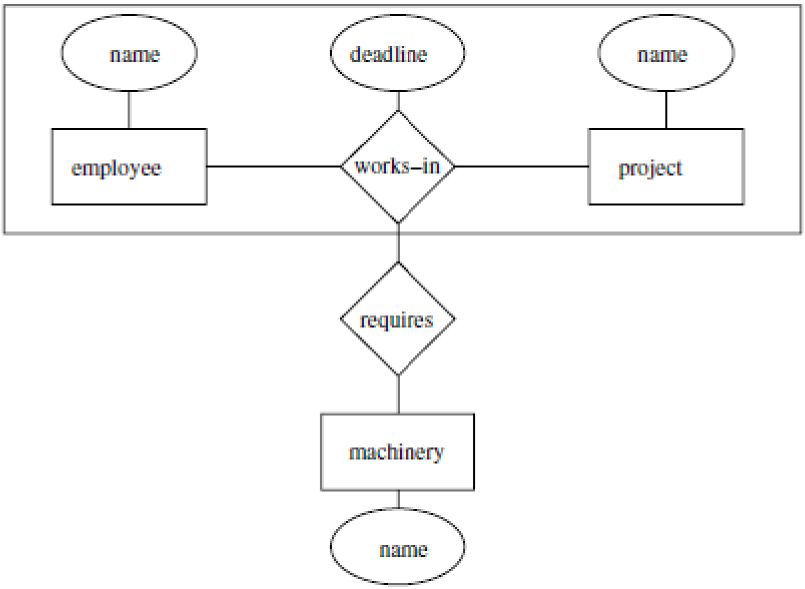


Figure 2.8 E-R diagram Example 1 of aggregation.

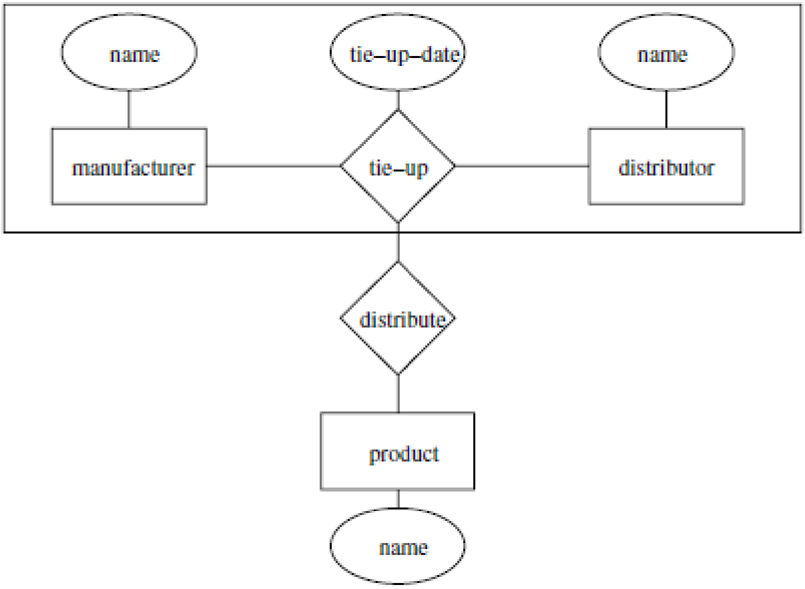


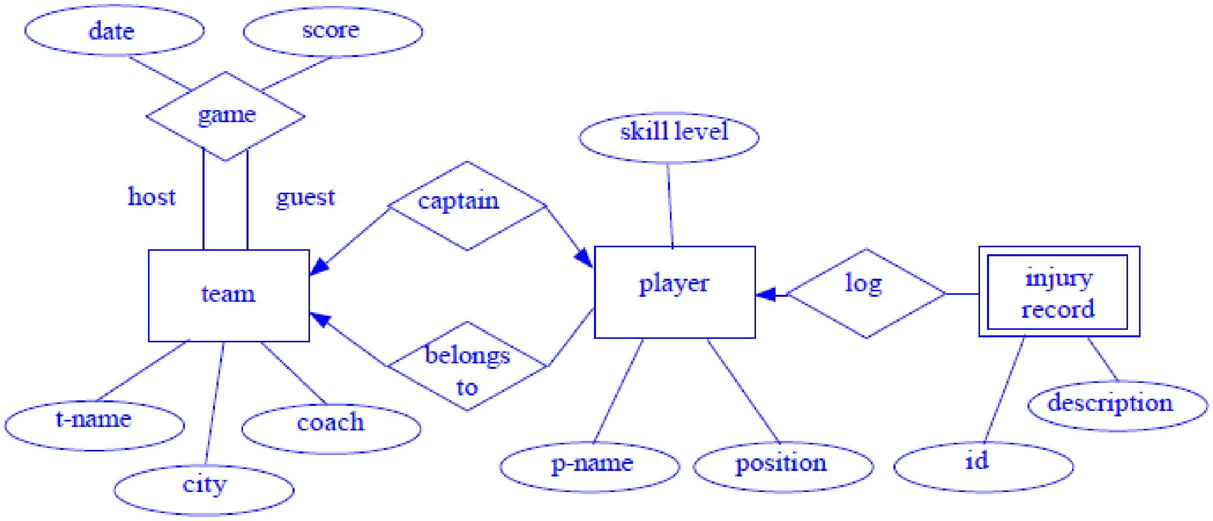
Figure 2.9 E-R diagram Example 2 of aggregation.

Suppose you are given the following requirements for a simple database for the National Hockey League (NHL):

- the NHL has many teams,
- each team has a name, a city, a coach, a captain, and a set of players,
- each player belongs to only one team,
- each player has a name, a position (such as *left wing* or *goalie*), a skill level, and a set of injury records,
- a team captain is also a player,
- a game is played between two teams (referred to as host_team and guest_team) and has a date (such as *May 11th, 1999*) and a score (such as *4 to 2*).

Construct a clean and concise ER diagram for the NHL database using the Chen notation as in your textbook. List your assumptions and clearly indicate the cardinality mappings as well as any role indicators in your ER diagram.

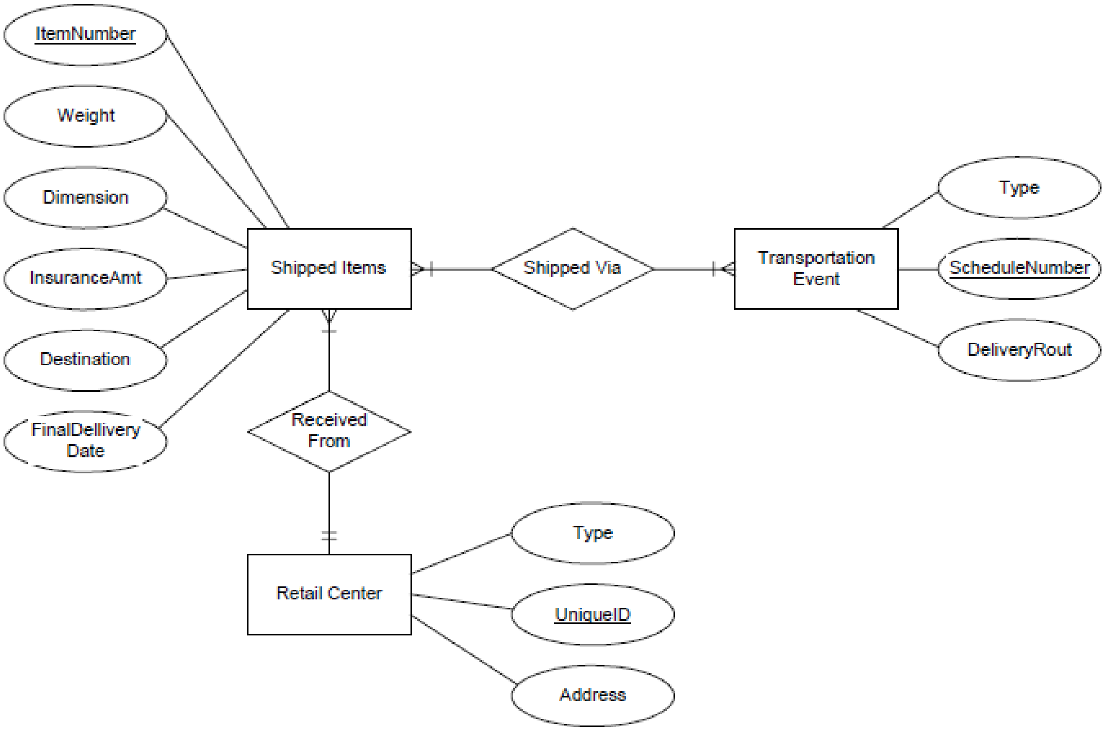
Here is one sample solution. Note that other diagrams are possible depending on assumptions.



UPS prides itself on having up-to-date information on the processing and current location of each shipped item. To do this, UPS relies on a company-wide information system. Shipped items are the heart of the UPS product tracking information system. Shipped items can be characterized by item number (unique), weight, dimensions, insurance amount, destination, and final delivery date. Shipped items are received into the UPS system at a single retail center. Retail centers are characterized by their type, uniqueID, and address. Shipped items make their way to their destination via one or more standard UPS transportation events (i.e., flights, truck deliveries). These transportation events are characterized by a unique scheduleNumber, a type (e.g, flight, truck), and a deliveryRoute.

Please create an Entity Relationship diagram that captures this information about the UPS system. Be certain to indicate identifiers and cardinality constraints.

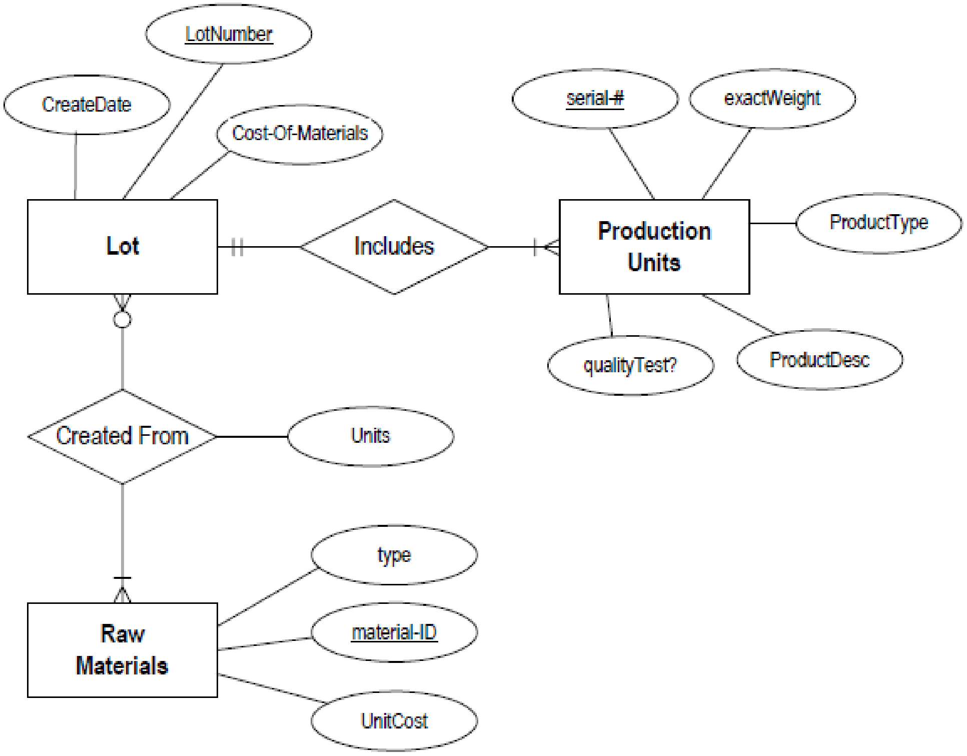
Solutions:



Grading:

Entities correctly identified: 5
Attributes correctly identified: 5
Primary keys correctly identified: 5
Relationships and cardinality correctly identified: 10

Production tracking is important in many manufacturing environments (e.g., the pharmaceuticals industry, children's toys, etc.). The following ER diagram captures important information in the tracking of production. Specifically, the ER diagram captures relationships between production lots (or batches), individual production units, and raw materials.



1. We need a database to keep track of our CD sets with opera recordings (we have several hundred). There is a lot of information about recordings that is interesting, but we limit ourselves to a simplified version of the database.

   Operas have titles and a year when they were first performed. We may ignore operas with the same title by different composers. We also assume that an opera has only one composer. Data about the composers include name, nationality, year of birth, and year of death. The same data is recorded for the librettists (text writers).

   The roles in the operas are sung by singers with different voice types (Rodolfo in La Bohème is a tenor, Marcello is a baritone, etc.). A role name may be used in several operas (Leonora in Il trovatore and in La forza del destino, for instance), and this we must take into account.

   Each opera may have been recorded several times. A recording is on a record label, was made a specific year, and has a conductor. Conductors are not so important, so we only record their names. A recording is released in a CD box with a unique number. We assume that each recording is released only once. In a recording, the different roles are sung by singers. We record the name, nationality, year of birth, and year of death also for singers.

   a) Develop an E/R model that describes the database.
   b) Translate the E/R model into a relational model. All relations must be in BCNF (prove, or at least motivate carefully, that this is the case). Indicate primary keys and foreign keys in all relations.
   c) Using your relations in question b, write an SQL statement that gives the CD numbers and opera titles of the recordings where Herbert von Karajan was the conductor and Mirella Freni appeared in one of the roles.

2. The transport company that Krusty Kookies Sweden AB uses for pallet transports has a database with the following relations:

   Drivers(driverId, name, address)
   Trucks(truckId, palletCapacity)
   Deliveries(deliveryId, truckId, driverId, deliveryDate, deliveryAddress, nbrPallets)

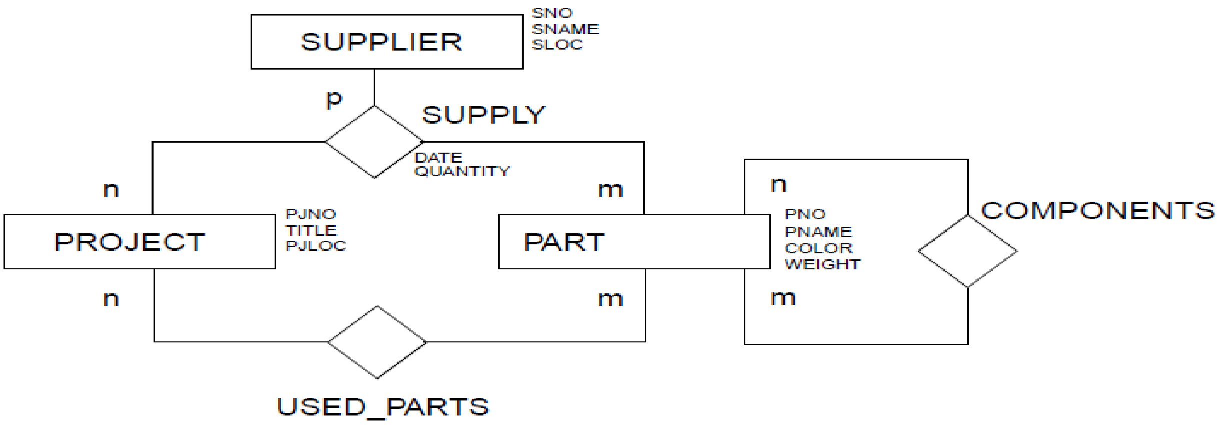
   The attribute names are, hopefully, self explanatory. In Deliveries, truckId and driverId are foreign keys.

   a) Draw an E/R diagram that describes the database.

   Write SQL statements for the following tasks (you may create and use views, if you wish):

   b) Create the table Deliveries. Invent suitable types for the attributes and indicate reasonable integrity constraints.
   c) List the names and addresses of all drivers, alphabetically by name.
   d) Print the average pallet capacity of the trucks.
   e) List all data about the deliveries where the full capacity of the truck wasn't used.
   f) Print the name of the driver who has made most deliveries (names, if there is more than one).
   g) Print the names of the drivers who haven't made any delivery.

3. Design a relational database schema corresponding to the given ER diagram. Determine primary and foreign keys.



4. A calendar program that allows users to browse each other's calendars and to book common appointments shall be developed. The program has a database which keeps track of the users and their calendars.

You use the calendar to store data concerning appointments. An appointment starts and ends at a given time on a given day and is described by a text. You may specify that you wish to be reminded of an appointment. Reminders are of different kinds: a signal in the computer's loudspeaker, a pop-up window with the description of the meeting, or an e-mail containing the description. You may, for each reminder, specify how long before the appointment that you wish to be reminded.

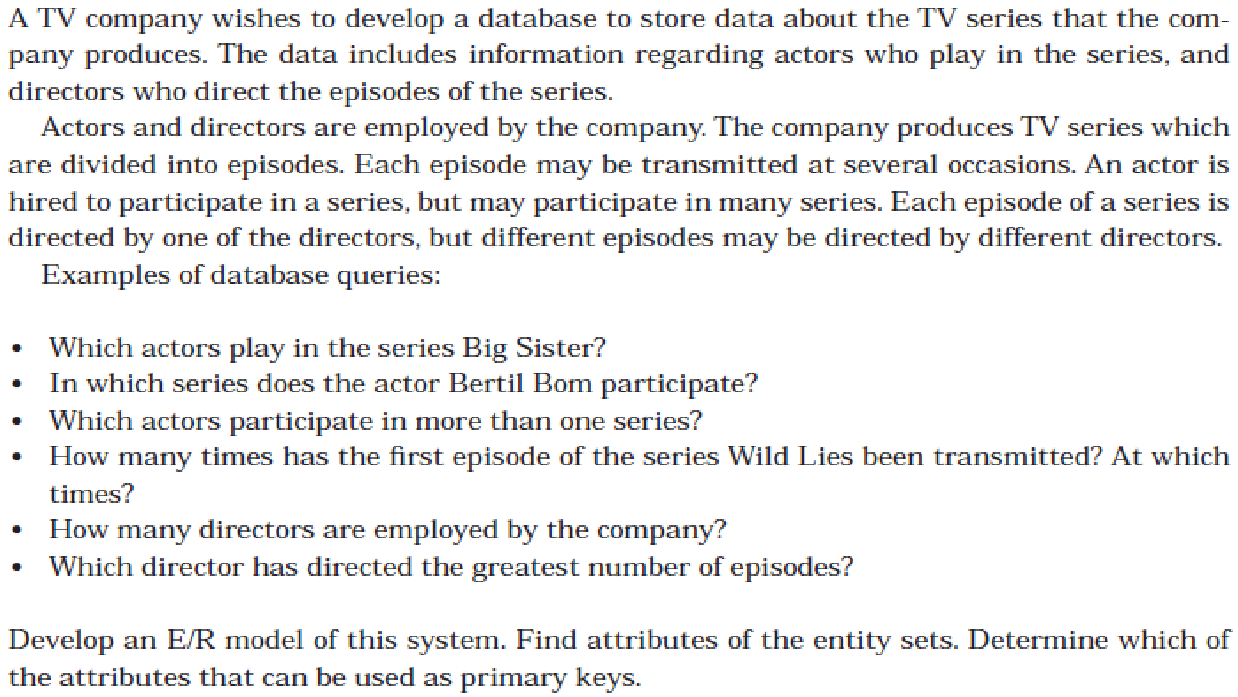
Develop an E/R model of the database.

5. Develop an E/R model of a database that is to be used in the following banking system:

A bank offers different kinds of accounts to its customers, e.g., checking accounts, savings accounts, and stock fund accounts. The different accounts have different interest rates and other conditions, such as withdrawal fees, number of withdrawals without a fee, etc. It must be possible for the bank to easily change interest rates and other conditions.

A stock fund account is linked to one or more stock funds. The stock fund shares are bought at rates that vary over time. Money that is deposited in a stock fund account is used to buy shares in the stock fund(s) that the account owner has chosen to link to the account. If more than one stock fund is linked to an account the money is divided between the funds according to percentages determined by the account owner.

6.



A TV company wishes to develop a database to store data about the TV series that the company produces. The data includes information regarding actors who play in the series, and directors who direct the episodes of the series.

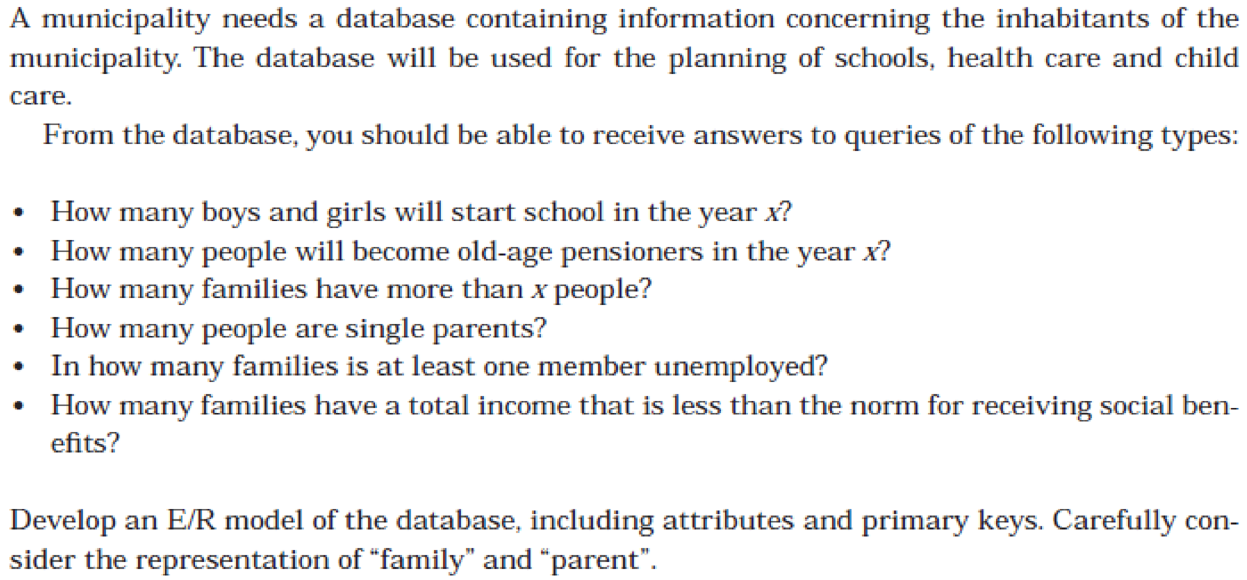
Actors and directors are employed by the company. The company produces TV series which are divided into episodes. Each episode may be transmitted at several occasions. An actor is hired to participate in a series, but may participate in many series. Each episode of a series is directed by one of the directors, but different episodes may be directed by different directors.

Examples of database queries:

- Which actors play in the series Big Sister?
- In which series does the actor Bertil Bom participate?
- Which actors participate in more than one series?
- How many times has the first episode of the series Wild Lies been transmitted? At which times?
- How many directors are employed by the company?
- Which director has directed the greatest number of episodes?

Develop an E/R model of this system. Find attributes of the entity sets. Determine which of the attributes that can be used as primary keys.

7.



A municipality needs a database containing information concerning the inhabitants of the municipality. The database will be used for the planning of schools, health care and child care.

From the database, you should be able to receive answers to queries of the following types:

- How many boys and girls will start school in the year *x*?
- How many people will become old-age pensioners in the year *x*?
- How many families have more than *x* people?
- How many people are single parents?
- In how many families is at least one member unemployed?
- How many families have a total income that is less than the norm for receiving social benefits?

Develop an E/R model of the database, including attributes and primary keys. Carefully consider the representation of "family" and "parent".