CS143: Homework #4

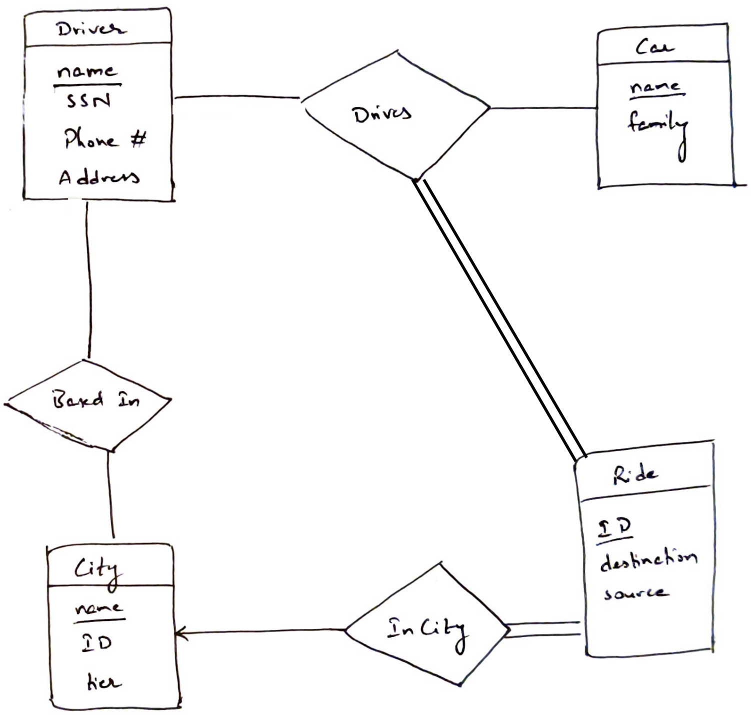
Entity-Relationship Model

1. Nevertaxi Inc. has decided to store information about drivers who drive in the cities Nevertaxi is based in (as well as other company data) in a database. The following information describes the situation that the Nevertaxi database must model.

   - Each driver that drives at Nevertaxi has a name, an address, and a phone number. Poorly paid drivers often share the same address.
   - Each car that is used in rides using Nevertaxi has a name (e.g., Thomahawk-387, BMW-fleet-23) and a family (e.g., sportscar, hatchback, sedan).
   - Each city Nevertaxi is based in has a name, tier and ID.
   - Each ride using NeverTaxi has a ID, source, destination, distance and an associated driver.
   - Each driver can drive multiple cars and each car can be driven by multiple drivers.
   - Each ride is associated with one city and a city can have multiple rides associated with it.
   - A ride is associated with only one driver, but a driver can be associated with multiple rides.
   - A city can have multiple drivers associated with it and a driver can be associated with multiple cities.

   (a) Draw an ER diagram for Nevertaxi. Be sure to indicate all key and cardinality constraints and any assumptions that you make. Identify any constraints that you are unable to capture in the ER diagram and briefly explain why you could not express them.

   ANSWER:



It is not easy to include cardinality constraint for a tertiary relationship, so the relationship set ‘‘Drives’’ does not include the constraint that every ride is associated with exactly one driver and one car.

The meaning of source and destination of a ride is not clear. Assuming that they indicate the exact pick up and drop off locations, they must be associated with a particular city, but the current design does not reflect this.

(b) Convert your ER diagram into relational schema and write a ‘CREATE TABLE‘ statement for each relation. Be sure to indicate the primary keys of the relations if they exist.

ANSWER:

CREATE TABLE Driver(name VARCHAR(10) PRIMARY KEY, SSN INT, phone_num int, address VARCHAR(50));

CREATE TABLE Car(name VARCHAR(10) PRIMARY KEY, family VARCHAR(20));

CREATE TABLE Ride(ID int PRIMARY KEY, source VARCHAR(10), destination VARCHAR(10));

CREATE TABLE City(name VARCHAR(10) PRIMARY KEY, ID INT, tier INT);

CREATE TABLE Drives(ID INT, driver_name VARCHAR(10), car_name VARCHAR(10), PRIMARY KEY(ID, driver_name, car_name));

CREATE TABLE BasedIn(driver_name VARCHAR(10), city_name VARCHAR(10), PRIMARY KEY(driver_name, city_name));

CREATE TABLE InCity(ID INT PRIMARY KEY, city_name VARCHAR(10));

Note that from our semantic understanding of the domain, we may conclude that the ID of Drives table is key, but the information captured in ER diagram does not guarantee this fact.

2. You are to design a database that maintains information for producing a weekly television guide for a given region (such as Northern California). The data should include information about television

shows, television networks, cities, channels, show times, etc. For starters, you may make the following assumptions:

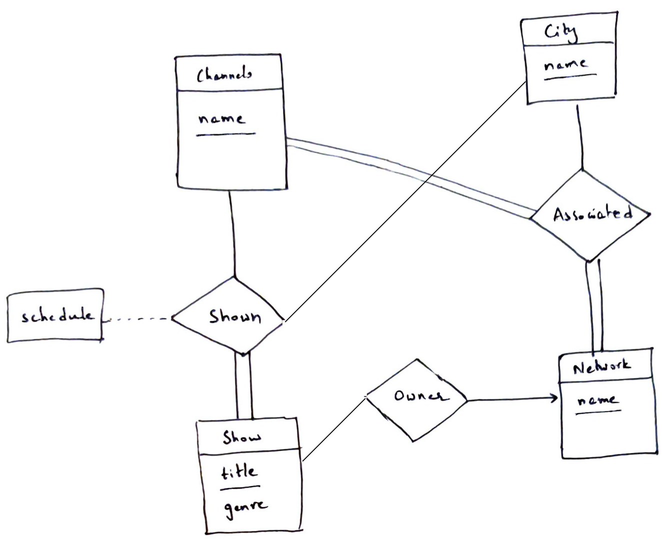
- A given channel in a given city is associated with one network.
- A given show is either owned by a network (and shown on a channel associated with that network) or is a local show and may be shown on any channel.
- Not all shows are shown in all cities, and the days and times for a given show may differ from city to city.
- You may ignore cable channels, which generally are not city-dependent.

Please feel free to make additional assumptions about the real world in your design, as long as the assumptions are reasonably realistic and are stated clearly as part of your solution.

Specify an entity-relationship diagram for your database. Don't forget to underline key attributes and include arrowheads and double lines.

Note that this question is fairly open-ended and there is no single right answer, but some designs are better than others.

ANSWER:




```
In this design, we do not include the constraint that when a show is owned by a network,
it can be shown only by the channels associated with the network.  Instead of this design,
we may also create two subclasses, IndependentShow and NetworkShow to differentiate the
two types of shows.
```

3. This problem is based on an E/R design for a database used in a tech company shown in Figure 1. This database stores information about programmers. Each programmer has a name, which uniquely identifies the programmer. A programmer may in fact be a team leader who in turn leads a team of programmers. For example, Elaine leads a team consisting of Michael and Bryan. Bryan works on project C. Michael works on project A and in turn leads a team consisting of Jane and David who work on project A and B respectively. Each team leader is also associated with the name of the team he leads.

   Convert the E/R diagram to relations. For the translation of subclasses, assume that we generate one table per each subclass, instead of creating one gigantic table for the ISA relationship.

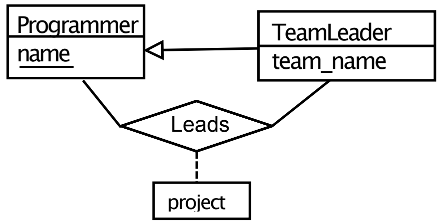


Figure 1: E/R diagram for a tech company

ANSWER:

```
Programmer(name)
TeamLeader(name, team)
Leads(TeamLeader.name, Programmer.name, project)
```

Relational Design Theory

1. Suppose that we decompose the schema $R(A, B, C, D, E, F)$ into (A, B, C, F) and (A, D, E). When the following set of functional dependencies hold, is the decomposition lossless?
   $A \to BC, CD \to E, B \to D, E \to A$
   Explain your answer.

   ANSWER:
   (A, B, C, F) `INTERSECT` $(A, D, E) = A$, `and` A `is a key for` (A, D, E), `so the decomposition is lossless.`

2. List non-trivial functional dependencies satisfied by the following relation. You do not need to find all functional dependencies. It is enough to identify a set of functional dependencies that imply all functional dependencies that is satisfied by the relation.

A	B	C
a_1	b_1	c_2
a_1	b_1	c_2
a_2	b_1	c_1
a_2	b_1	c_3

   ANSWER:
   $A \to B$
   $C \to A$

3. Assume *Student* and *Class* entity sets that we have used in the class. The *Student* and *Class* sets are connected by Take relationship set. We now convert the Take relationship set into a table **Take(sid, dept, cnum)** using our standard translation algorithm, where sid is the key for a student and (dept, cnum) is the key for a class.
   Explain how functional dependencies can be used to indicate the following:

(a) A one-to-one relationship exists between entity sets *Student* and *Class*.
ANSWER:
$sid \rightarrow dept, cnum$
$dept, cnum \rightarrow sid$

(b) A many-to-one relationship exists between entity sets *Student* and *Class*.
ANSWER:
$sid \rightarrow dept, cnum$

4. Assume the following set of functional dependencies hold for the relation R(A, B, C, D, E):
$A \rightarrow BC$, $CD \rightarrow E$, $B \rightarrow D$, $E \rightarrow A$

(a) Is A a key for R? Explain your answer.
ANSWER:
`Yes,` A `is a key.` $A+ = ABCDE$.

(b) Is BC a key for R? Explain your answer.
ANSWER:
`Yes,` BC `is a key.` $BC+ = ABCDE$.

5. Assume the following set of functional dependencies hold for the relation *R(A, B, C, D, E, F)*:
$A \rightarrow BC$, $C \rightarrow E$, $B \rightarrow D$
Is it in **BCNF**? Explain your answer. If it is not, normalize it into a set of relations in **BCNF**.
ANSWER:
`It is not in BCNF.`
`The key is` AF, `so A` $\rightarrow$ `BC, C` $\rightarrow$ `E and B` $\rightarrow$ `D all violate BCNF.`
$R(A, B, C, D, E, F) \Longrightarrow R1(A, B, C, D, F) and R2(C, E) using C \rightarrow E$
$R1(A, B, C, D, F) \Longrightarrow R3(A, B, C, F) and R4(B, D) using B \rightarrow D$
$R3(A, B, C, F) \Longrightarrow R5(A, F) and R6(A, B, C) using A \rightarrow BC$

`The final BCNF tables are:`
$R2(C, E)$
$R4(B, D)$
$R5(A, F)$
$R6(A, B, C)$