CS 455 Principles of Database Systems

Homework 1: The Relational Data Model

Name:

Instructions: Print this assignment using single-side pages. Fill in your name above, and write in the solutions in the space provided below each question. You are allowed to use the back of each page. If you used any scratch paper to show your work, append those to the end. **Note:** It is important you use this format for gradescope.

Submission: After you've filled in the answers, scan all pages into a PDF, and submit to canvas.

Problems

1. What is the minimum and maximum number of superkeys that a relation with n attributes can have? Explain.

2. Does every relation have more superkeys than candidate keys? Explain.

3. Give an example relation schema (not just a relation instance) that has two candidate keys. You may need to list some real world assumptions about the data you are modeling.

4. The relational model of data is not the only data model in use today. For instance, the JSON model of data is widely used in emerging applications. Give a brief overview of JSON's structure, and compare it against the relational model.

5. Consider the following relations for a database that keeps track of auto sales from a used car dealership where prices can vary greatly based on the car's condition and age.

```
car(vin, model, make, price)                  options(vin, optionName, price)
sales(salespersonID, vin, date, price)        salesperson(name, email)
```

Here are some assumptions to go on:

- Dealerships often have multiple cars sharing the same make (e.g., Ford) and model (e.g., Fusion), so cars can only be identified by their vehicle identification number (`vin`).
- Certain cars comer with extra **options** (such as a sunroof and performance tires). Each option carries a different price depending on the car.
- Each **salesperson** can be uniquely identified simply by their email address.
- To record a **sale**, we store references to the salesperson, the car, date, and the price that it ultimately sold for. Note that the same car may be sold more than once, even by the same salesperson. (This happens when the car is traded-in, or is sold back to the dealership at a later time.)
- To avoid scams, the dealership has a standing policy that prohibits the same car to be sold more than once on the same day.

Below each of the following relations, (1) list all *candidate keys* and (2) list any *foreign keys* that exist. For each foreign key $A_1, ..., A_k$ that originates in relation R and points to an attribute set $B_1, ..., B_m$ in relation S, specify it using format $R(A_1, ..., A_k) \rightarrow S(B_1, ..., B_m)$.

- car

- salesperson

- options

- sale

6. Design a *schema diagram* for a company like Netflix that minimizes redundancy and NULL values. You should recall that an attribute cannot hold a list of values. Indicate all primary keys (underlined) and foreign key constraints (directed edges). You need to store data on the following:

   - A Movie has a title, description, country of origin, and a release-year. Assume that there may be two movies with the same name (e.g., Godzilla), but they are never released in the same year. You may assign each movie a unique identifier, but you shouldn't need to.
   - An Actor has a name and a biography. Assume that there may be two actors with the same name. You may assign each actor a unique identifier.
   - A User has a userid and password. The company does not allow two different users to share the same userid.
   - You need to keep track of what movies the users have seen.
   - Users can rate movies from 1 to 5 stars. Note that they need not have seen the movie to rate it! A user can only rate a movie once (though that rating can change later).
   - Users can also comment on each actor and/or movie. A comment has a date, time, and the comment itself. Unlike ratings, users can leave a multiple comments on the same movie or actor.
   - Actors play certain roles in certain movies. Actors may also play multiple roles in the same movie (see: Eddie Murphy and Mike Myers).

   Space for your schema diagram is given on the next page.

(Space for schema diagram.)