CIS 550: Database and Information Systems

Homework 3: Relational DB Design

This homework is about relational database design, from ER diagrams to relational schemas and their normalization.

Submission: Create **one** .pdf file with your answers, with the answer to each question on a separate page(s), and submit using Gradescope. Be sure to label your responses as numbered here.

Question 1 (24 points)

Model the data that supports this application using an ER diagram for the elements below. Use the notation from class (e.g. 0..1, 1..1, 0..*, 1..* annotations on edges; double boxes on weak entity sets, the defining relationship set, etc.). Create an electronic version of each diagram using draw.io, Powerpoint, Word, Visio, OpenOffice Draw, OmniGrae, Google Drive Document or any other software that you are familiar with, and export it to PDF format. **Do not turn in handwritten diagrams as they are hard to read!**

Consider the following Zoom-like scenario. Note that this may vary slightly from the way Zoom is set up, so use the scenario as described.

- An account has an account number, person name, email, language, and type. An account is identified by its account number.

- A meeting plan is identified by an id, and is associated with a unique account (the host). A meeting plan can be scheduled or unscheduled.
    - An unscheduled meeting has a title. (Think of this as your personal Zoom room.)
    - A scheduled meeting has a title, optional description, time, duration, start date, end date, and recurrence interval. The recurrence interval can be “daily”, “weekly”, “every T/Th”, etc. (Think, for example, of Zoom office hours that are planned for the semester.)

- A meeting is an instance of a meeting plan, and is identified by a date and time in addition to the meeting plan id. Not every meeting plan must have a meeting instance, but every meeting must be an instance of a meeting plan.

- A meeting may be recorded zero or more times. A recording has an id (which identifies the recording), video content and chat text.

A. (20 points)

Draw an ER diagram for this scenario.

Note:
- Both entity and relationship sets can have attributes
- Primary keys should be determined from the descriptions, and should be shown in the ER diagram by underlining the attribute(s)

- Cardinality constraints should be indicated according to the descriptions
- Feel free to give the entity and relationship sets any names that are reasonable.

B. (4 points)

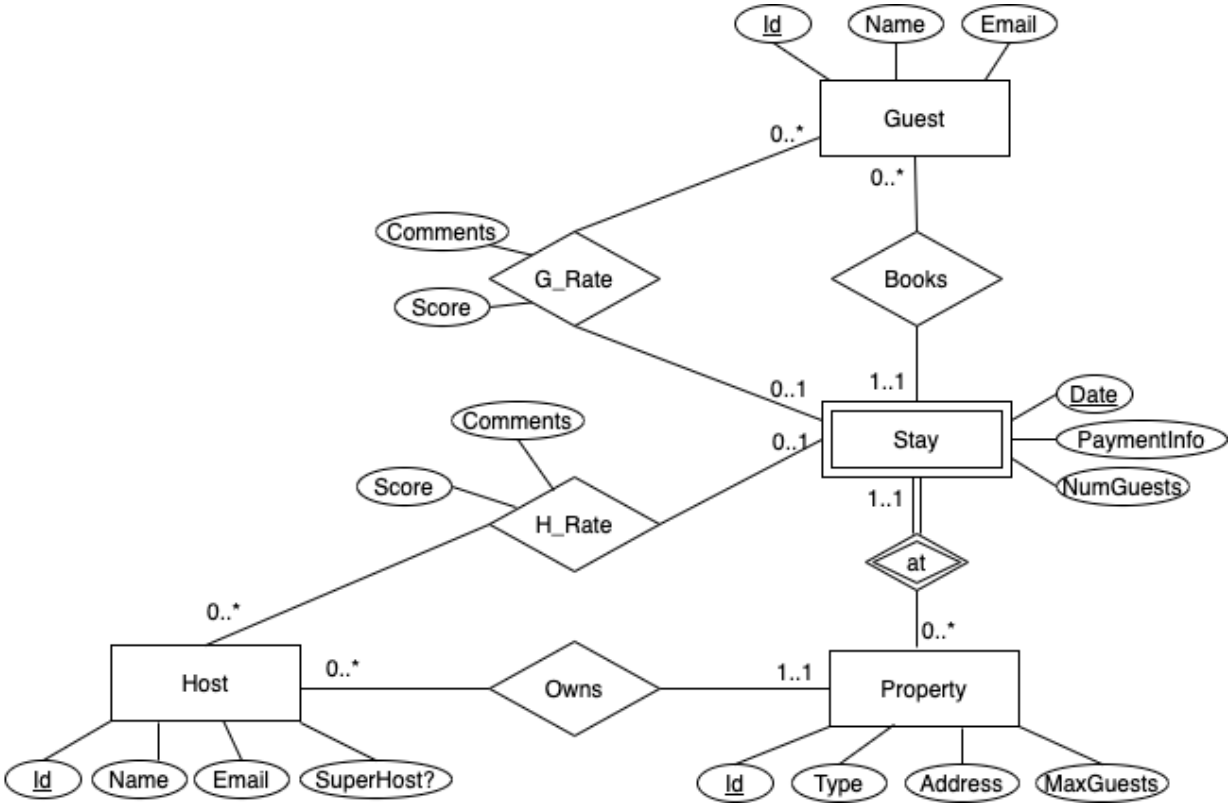
Suppose that you wish to enforce the constraint that the date and time of a scheduled meeting instance are "correct". That is, a meeting instance is identified by a date and time (in addition to the meeting plan id), and a scheduled meeting has a set of potential meeting instance dates and times that can be determined by the start date, interval, end date and time. The date and time of a scheduled meeting instance is correct if it is one of the potential meeting instance dates.

As an example, suppose that a scheduled meeting had a start date of 07/01/2021, interval of 1 week, end date of 07/20/2021, and time of 1pm. This gives a set of potential meeting dates and times of (07/01/2021, 1pm), (07/08/2021, 1pm), and (07/15/2021, 1pm). A "correct" instance of this scheduled meeting would be (07/01/2021, 1pm) and an "incorrect" one would be (07/05/2021, 12pm).

Does your ER diagram capture the fact that scheduled meeting instances should be "correct", e.g. through keys or cardinality constraints? If not, how can this be guaranteed?

Question 2 (12 points)

The following ER diagram models an AirBNB database, in which hosts own properties, guests stay at properties, and both hosts and guests can rate stays.



Translate the ER diagram to the relational model by specifying the resulting relations, their attributes, keys, and foreign keys. You may modify the attribute names in foreign keys as appropriate, and should specify when the foreign key must exist using NOT NULL. Use the abbreviated syntax from lectures rather than SQL DDL. Once you have identified the primary key(s), you do not need to specify a NOT NULL constraint for these keys, as we assume all keys must be non-null. However, you will need to specify this constraint for any other attributes you believe cannot be null.

For example, if every department must have a manager (who is an employee) but not every employee has to have a home department, you would write:

Employee(ssn, name, title, home-dept)
    home-dept FOREIGN KEY REFERENCES Department(id)

Department(id, name, manager)
    manager FOREIGN KEY REFERENCES Employee(ssn)
    manager NOT NULL

Question 3 (12 points)

For each of the following facts, say whether or not they can be represented by the ER diagram in the previous question. You should not assume any constraints other than those represented by the key or cardinality constraints shown in the diagram. Briefly justify your answer.

As an example, the answer to whether "The host with id=1 does not own any property" can be represented is "*Yes, the participation of host in the owns relationship is 0..* and therefore is not mandatory.*"

A. (4 points)

The host with id 5 owns the property with id 1 and the property with id 2.

B. (4 points)

The stay at the property with id 10 on 9/15/19 was booked by two different guests, the guest with id 5 and the guest with id 6.

C. (4 points)

The guest with id 4 added a review for the stay at the property with id 11 on 9/10/19, but did not book the stay.

(continued overleaf)

Question 4 (10 points)

Consider the following instance of a relation R(A, B, C, D):

A	B	C	D
a1	b2	c2	d1
a2	b2	c3	d1
a1	b2	c2	d2

A. (2 points)

Does this instance satisfy AB → C?

B. (2 points)

Does this instance satisfy A → C?

C. (2 points)
Does this instance satisfy B → C?

D. (4 points)

Can the following be proven using Armstrong's Axioms? Explain your answer.
If AB → C and A → C , then B → C

Question 5 (10 points)

Consider the following schema R and set of functional dependencies F:

R(A,B,C,D), F= {AB → C , BC → D , CD → A , AD → B}

A. (4 points)

State all candidate keys. Justify your answer.
Remember that a candidate key is a set of attributes that functionally determines all attributes in the schema, and that is minimal in the sense that no proper subset of those attributes is also a candidate key.

B. (3 points)

Is R in BCNF? Justify your answer.

C. (3 points)

Is R in 3NF? Justify your answer.

(continued overleaf)

Question 6 (16 points)

Consider the following schema R and set of functional dependencies F:
R(A,B,C,D,E), F= {A → B , BC → E , ED → A}

A. (4 points)

State all candidate keys. Justify your answer.

B. (3 points)

Is R in 3NF? Justify your answer.

C. (6 points)

You are given the following decomposition of R into BCNF:

R1(A, B)
R2(A, C, E)
R3(A, C, D)

What is the projection of F over each R1, R2 and R3 (F1, F2 and F3, respectively)? (**Hint:** Be careful to look in F+ for the dependencies that hold in each Ri, not just in F.)

D. (3 points)

Does the decomposition in part C preserve dependencies? Justify your answer.

Question 7 (16 points)

Consider the following schema R and set of functional dependencies F:

R(A,B,C,D,E), F= {A → B , A → C , C → D}

A. (4 points)

State all candidate keys. Justify your answer.

B. (6 points)

Give a decomposition of R into BCNF. Explain how you arrive at it, and justify whether or not your decomposition preserves dependencies.

C. (6 points)

Give a decomposition of R into 3NF. Explain how you arrive at it.