

Write your answers in the box below only. Do not write on the back or outside the box.

---

**Database Systems — CSci 4380**  
**Midterm Exam #1**  
**September 27, 2018**

RCS ID: \_\_\_\_\_ @rpi.edu      Name: \_\_\_\_\_

RIN # : \_\_\_\_\_

**Rules.** The exam is 110 minutes for a total of 100 points. Open book and notes. Do not use any electronic tools including your computer, phone or tablet. Work alone. You **cannot** talk to anyone in class, or share notes or thoughts.

**Question 1 (10\*3=30 points).** Write the following queries using relational algebra using the data model below. The model is described in detail in the back of the exam.

Users(email, password, name, street, state, city, zip, country, ccno)  
Houses(id, label, description, street, state, city, zip, country, price, owneremail)  
HouseAmenities(houseid, amenity)  
Trips(id, fromdate, todate, isconfirmed, totalprice, customeremail, houseid)  
Reviews(tripid, isclean, isgoodvalue, isgoodrenter, reviewtext)  
Messages(id, senderemail, receiveremail, messagetext, sentdatetime, prevmid)

- (a) Return the email and name of all users who are both renters (i.e. owners of houses) and customers (i.e. customers on trips).

**Write your answers in the box below only. Do not write on the back or outside the box.**

---

- (b) Find the id and label of all houses in Troy (city) NY (State) with price below \$100 that have both the amenity 'Helipad' and the amenity 'Batcave'.

- (c) Return the email and name of all renters who only received 5 star reviews for their houses with respect to isgoodvalue attribute.

Write your answers in the box below only. Do not write on the back or outside the box.

---

**Question 2 (12+12=24 points).** Suppose you are given the following relations to add to the data model in the appendix. Answer questions regarding each additional relation below.

- (a) Given a trip, the customer may add other names of people staying in the house with him or her as other travelers.

`OtherTravelers(tripid, firstname, lastname, email, phone)`

Other travelers on a trip are uniquely identified by their email. The same person may take multiple trips as an “other traveler”. An other traveler can have multiple phones listed in the database.

- (1) Based on the above information, list all applicable functional dependencies.
- (2) What are the key(s)?
- (3) Is this relation in BCNF? 3NF? Explain why or why not.
- (4) If it is not in BCNF, use BCNF decomposition to get relations that are in BCNF.

- (b) The following is a new relation for saving some houses in lists:

```
SavedHouses(houseid, saveddatetime, useremail, note, listname)
    houseid, useremail, listname → saveddatetime
    saveddatetime, useremail → houseid, useremail
    houseid, useremail → note
```

Answer each of the following with yes/no and write a sentence to explain your answer.

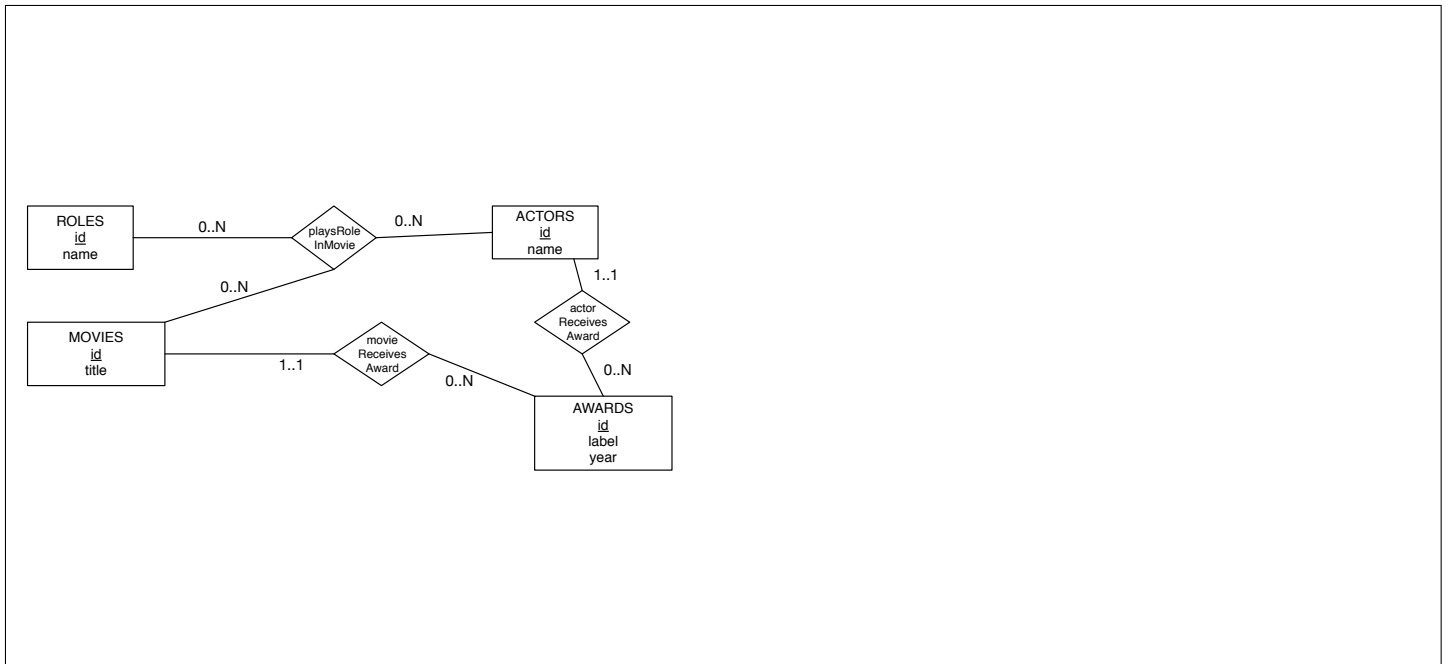
- (1) Can a user save a house in multiple lists?
- (2) Can a user save two different houses in the same datetime?
- (3) Can a user have two different notes for the same house in two different lists?
- (4) What are the keys? Is this relation in 3NF?

**Question 3 (12 points).** You are given the following relation:  $R(A,B,C,D,E,F,G)$  with the following set of functional dependencies:

$F = \{ AB \rightarrow CE, BF \rightarrow G, C \rightarrow A \}$

- List all the keys.
- Convert this relation to 3NF using 3NF decomposition. Show your work.
- For each resulting relation, show the key and state simply whether it is in BCNF or not.

**Question 4 (6 points).** You are given the following Entity-Relationship diagram. Based on this model, list all functional dependencies that this model implies.



Write your answers in the box below only. Do not write on the back or outside the box.

---

**Question 5 (12 points).** Given the relation:  $R(A,B,C,D,E,F)$  and

$$F = \{BC \rightarrow AE, DC \rightarrow F, AF \rightarrow D\}$$

Show whether the following decomposition is lossless or not with the Chase algorithm. Show your work.

Decomposition:  $R_1(A,B,E)$ ,  $R_2(C,D,F)$ ,  $R_3(A,D,F)$ ,  $R_4(A,B,C,F)$

Write your answers in the box below only. Do not write on the back or outside the box.

---

**Question 6 (16 points).** Create an Entity-Relationship diagram for the following database, capturing all the requirements below precisely. Make sure you list all the relevant attributes, underlining the keys. For each relationship, mark the participation constraints clearly (one-to-one, one-to-many or many-to-many). If you do not find a natural key for an entity, feel free to add an id attribute.

You are creating a database for storing data for a new social media company. In this database, you store users. Each user has an email, name, password, and a phone number. Emails are unique in the database.

Each user follows some users and is followed by some users. The following relationship is not symmetric; person A may follow person B while person B does not follow A. There are a number of known locations in the database. Each location has a unique coordinates (X,Y coordinates in the form of longitude and latitude values) and a label.

Each person generates posts. Each post has a datetime, text, URL and zero or more pictures. There can be at most one post by a specific user at a given datetime. Each picture may have a label and it may also have a location associated with it from the database. In addition, each picture may have zero or more rectangles drawn on them. Each rectangle is represented by its coordinates on the image (x1,y1,x2,y2) and another tagged user. Users may like or dislike posts of users. Users may also comment on other users' posts. Each comment has a sequence id (1,2,3,...) and text.

Write your answers in the box below only. Do not write on the back or outside the box.

---

Use this page for scratch work only. Do not share your solutions or any drafts of your solutions with anyone.



# Data model to be used in Exam #1

This is a data model based on the E-R problem from Homework#3, storing information about house rentals in a system similar to Airbnb. It is simplified in various ways to make it easier to user in an exam. Please read carefully. The keys of each relation are underlined.

## **Users(email, password, name, street, state, city, zip, country, ccno)**

Stores information about all the users in the system, renters and customers including credit card number (ccno).

## **Houses(id, label, description, street, state, city, zip, country, price, owneremail)**

Each house is listed by an owner and we store the email of the owner for each house. The daily price of a house is fixed for simplicity.

## **HouseAmenities(houseid, amenity)**

This relation lists the various amenities a house has such as 'free wifi', 'garage parking', etc,

## **Trips(id, fromdate, todate, isconfirmed, totalprice, customeremail, houseid)**

Each trip has a start and end date, whether it is confirmed or not (true/false value) and a total price for the trip. Each trip is by a user (customer), so his/her email is stored. This is the person paying for the trip. Each trip for a specific house, so the id of the house is also stored. (The renter can be deduced from the owner of the house).

## **Reviews(tripid, isclean, isgoodvalue, isgoodrenter, reviewtext)**

We will only store the reviews of the houses in this database. There can only be one review for each trip, so we will use tripid as the key. Each review is for a specific house and its renter (stored in other relations) and contains 1-5 star values for isclean, isgoodvalue, isgoodrenter as well as a review text.

## **Messages(id, senderemail, receiveremail, messagetext, sentdatetime, prevmid)**

Finally, we store messages simply from one user (sender) to another user (receiver) together with the text and datetime it was sent. The attribute **prevmid** can be null if this is an original email or store the id of the message that this message was sent in response to.

Use this page for scratch work only. Do not share your solutions or any drafts of your solutions with anyone.