Midterm Exam #1 (Version A)
CS 122A
Spring 2018

Max. Points: 100
(Please read the instructions carefully)

Instructions:
- The total time for the exam is 50 minutes; be sure to budget your time accordingly.
- The exam is closed book and closed notes but “open cheat sheet”.
- Read each question first, in its entirety, and then carefully answer each part of the question.
- If you don’t understand something, ask one of the exam proctors for clarification.
- If you still find ambiguities in a question, note the interpretation you are taking.

<table>
<thead>
<tr>
<th>QUESTION</th>
<th>TOPIC</th>
<th>POINTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Modeling Terms</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>E-R Modeling</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>E-R to Relational Translation</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>Relational Design Theory</td>
<td>25</td>
</tr>
<tr>
<td>TOTAL</td>
<td>All</td>
<td>100</td>
</tr>
</tbody>
</table>
Question 1: Modeling Terms (10 points)

(10 pts) You’ve just been hired as a consultant for Orange County’s greenest car dealership, WaterWorld! They sell and maintain the new Aqua brand of cars based on water-powered engine technology, and they need your help to set up a database for their business. Here’s what their IT manager has to say:

- Each employee has a unique employee number, a name, and a mailing address that includes their street address, city, state, and zipcode.
- More specialized employees include mechanics and salespersons. Each mechanic has a specialty and an hourly_rate. Each salesperson has a commission rate (the percent they get of each sale).
- Cars have a unique VIN number, a model, a year, a description, a price, and a list of available exterior colors.
- Customers have a unique customer id, a name, an e-mail address, and a phone number.
- When a sale takes place, a given sale is of a particular car to a particular customer by a particular salesperson on a given date and at an agreed-upon sale price.
- When an Aqua car has a problem, a repair job will be scheduled to fix it. A repair job is associated with a particular car and the job has a cost, a description, a date, and a sequence number (e.g., this might be the 3rd repair job for the car in question).
- A given repair job will be performed by one or more mechanics, each of whom contributes some number of labor hours on their part of the repair job.

Match each of the modeling constructs in the left column below with their best-matching feature (drawn from the description above) in the right column below. Indicate your answer by writing the relevant description feature in the blank to the left of each modeling construct. (NOTE: You should use each feature ONLY ONCE, and you might find it quicker to go down the list of features looking for constructs rather than going down the list of constructs looking for features.)

** Alternative solution will be accepted ONLY if they are used once. **

<table>
<thead>
<tr>
<th>Modeling construct</th>
<th>Alternative</th>
<th>Description feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer</td>
<td>Entity</td>
<td>employee_name</td>
</tr>
<tr>
<td>Sale</td>
<td>Ternary relationship</td>
<td>sale_price</td>
</tr>
<tr>
<td><em>exterior_colors</em></td>
<td>Multivalued attribute</td>
<td>email_address</td>
</tr>
<tr>
<td><em>employee_name</em></td>
<td>Inherited attribute</td>
<td>Customer</td>
</tr>
<tr>
<td><em>email_address</em></td>
<td>Atomic attribute</td>
<td>sale_price, employee_name</td>
</tr>
<tr>
<td><em>sale_price</em></td>
<td>Relationship attribute</td>
<td>exterior_colors</td>
</tr>
<tr>
<td><em>mailing_address</em></td>
<td>Composite attribute</td>
<td>Sale</td>
</tr>
<tr>
<td>repair_sequence_number_</td>
<td>Partial key</td>
<td>VIN_number</td>
</tr>
<tr>
<td><em>VIN_number</em></td>
<td>Candidate key</td>
<td>mailing_address</td>
</tr>
<tr>
<td>VIN_number, model</td>
<td>Super key</td>
<td>VIN_number, model</td>
</tr>
</tbody>
</table>
**Question 2: E-R Modeling (30 points)**

(30 pts) Based only on the E-R model given below, indicate whether each of the statements that follow it are true (T) or false (F) by writing the appropriate letter in the space to the left of the statement.

- **T** A professor can have been the advisee of only one Ph.D. advisor.
- **F** Each textbook must contain at least one chapter.
- **F** Every person must be a student or else be a professor.
- **F** A chapter must not be about more than one topic.
- **T** A chapter must always be associated with a textbook.
- **F** A professor can have been the advisor of no more than one professor.
- **F** A professor cannot have been their own advisor.
- **T** It is possible for a person to have both a rank and a gpa.
- **T** A professor may not publish the same textbook twice using a different publisher
- **F** The number (cnum) of a chapter will uniquely identify it in the database overall.
- **F** A professor must publish at least one but possibly many textbooks.
- **F** A student must buy at least one textbook.
- **T** It is not possible to have two publishers with the same name.
- **F** A professor cannot have the same rank as the professor who advised their Ph.D.
- **F** The length of a textbook is computable from its title plus the titles of its chapters.
Question 3: E-R to Relational Translation (35 points)

(35 pts) Translate the following E-R schema into an appropriate set of SQL tables. As usual, avoid using more tables than necessary, and be sure that your translated design – expressed as CREATE TABLE statements in SQL – includes any/all appropriate (i) primary keys, (ii) unique keys, (iii) NOT NULL constraints, (iv) FOREIGN KEY constraints, and (v) ON DELETE options. We’ve started the answer for you below, providing the full answer for USER and a part of the answer for PICTURE. Finish filling in the details of PICTURE on the left, adding any missing attributes and/or constraints, and then finish the the job by adding any additional CREATE TABLE statements on the right below. Note: Try to fit your answer on this page. Only if space on this page is not enough you can use the blank page at the end of this exam.

```
CREATE TABLE USER (
    userid VARCHAR(20),
    name VARCHAR(40),
    age INT,
    PRIMARY KEY(userid)
);

CREATE TABLE PICTURE (
    picno VARCHAR(20),
    caption VARCHAR(100),
    date DATETIME,
    thumbnail VARBINARY(2000),
    imgfilename VARCHAR(100),
    userid VARCHAR(20),
    post_date DATETIME,
    PRIMARY KEY(userid, picno),
    FOREIGN KEY (userid) REFERENCES USER(userid) ON DELETE CASCADE
);

Create Table Tagged_in(
    tagger_id VARCHAR(20),
    pic_userid VARCHAR(20),
    picno VARCHAR(20),
);  
```

SQL DDL syntax reminder:
```
CREATE TABLE FooTable (
    id INTEGER,
    name VARCHAR(40),
    price DECIMAL(9,2),
    FOREIGN KEY (id) REFERENCES BarTable (id)
);  
```
Primary Key(tagger_id, pic_userid, picno),
Foreign Key(pic_userid, picno) References PICTURE(userid, picno),
Foreign Key(tagger_id) References USER(userid)
);

Create Table Friend(
    liker   VARCHAR(20),
    likee   VARCHAR(20),
    Primary Key(liker, likee),
    Foreign Key(liker) References User(userid),
    Foreign Key(likee) References User(userid)
);

Question 4: Relational Design Theory (25 points)

Answer each of the following questions about relational DB design and functional dependencies:

(5 pts) Consider a relation $R$ ($a$, $b$, $c$, $d$, $e$) with FDs $a \rightarrow b$, $b \rightarrow c$, $c \rightarrow d$, $c \rightarrow e$.

a. Compute the attribute closure ($b^+$) of the attribute $b$: $b, c, d, e$

b. Circle the highest normal form that $R$ satisfies: 1NF 2NF 3NF BCNF

(10 pts) Assuming that a (possibly different) table $R$ ($a$, $b$, $c$, $d$, $e$) currently contains the following data values, answer the following questions about $R$’s set of functional dependencies.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Joseph</td>
<td>Joe</td>
<td>30</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Susan</td>
<td>Sue</td>
<td>20</td>
<td>y</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Fred</td>
<td>Frodo</td>
<td>100</td>
<td>w</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Susan</td>
<td>Sue</td>
<td>30</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Christina</td>
<td>Chrissie</td>
<td>100</td>
<td>w</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Jacob</td>
<td>Jake</td>
<td>50</td>
<td>w</td>
<td></td>
</tr>
</tbody>
</table>

a. Is it possible that this is the same table $R$ described above? YES NO

b. Does the functional dependency $e \rightarrow c$ currently hold in $R$? YES NO

c. Does the functional dependency $b \rightarrow c$ currently hold in $R$? YES NO

d. Does the functional dependency $c \rightarrow b$ currently hold in $R$? YES NO

e. Does the functional dependency $(a, b) \rightarrow d$ currently hold in $R$? YES NO
(10 pts) Consider a (possibly still different!) table \( R(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}) \) with FDs \((\mathbf{b}, \mathbf{a}) \rightarrow \mathbf{c}, \ \mathbf{e} \rightarrow \mathbf{d}\).

a. List the candidate key(s) for \( R \): \textit{abe}

b. Circle the highest normal form that \( R \) satisfies: \( 1\text{NF} \ 2\text{NF} \ 3\text{NF} \ \text{BCNF} \)

c. Normalize \( R \) into a lossless-join BCNF design \((R_1, R_2, \ldots )\):
\[
\begin{align*}
& R_1(\mathbf{a}, \mathbf{b}, \mathbf{c}) \\
& R_2(\mathbf{e}, \mathbf{d}) \\
& R_3(\mathbf{a}, \mathbf{b}, \mathbf{e})
\end{align*}
\]

d. Is your BCNF design dependency-preserving? \textit{YES} \ NO