CS 122A/EECS 116: Introduction to Data Management
Spring 2016, UC Irvine
Prof. Chen Li
Midterm Exam (Max. Points: 100)

Instructions:

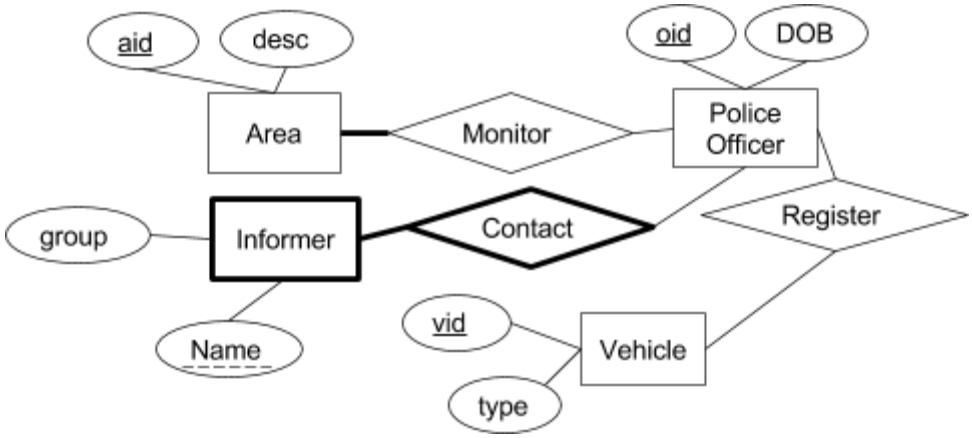
- The total time for the exam is 80 minutes. Be sure to budget your time accordingly.
- The exam is closed book and closed notes.
- Be sure to answer every part of each question.
- If you don't understand something, ask an instructor/TA for clarification.
- If you still find ambiguities in a question, note the interpretation you are taking.

NAME: Solution	**STUDENT ID: Solution**

QUESTION	POINTS	TOPIC	SCORE
1	**20**	E-R Modeling	
2	**25**	E-R to Relational	
3	**25**	Relational Algebra	
4	**20**	SQL	
5	**10**	Functional Dependencies	
TOTAL	**100**		

Question 1: E-R Modeling (20 points)

The captain of a police department wants to use a DBMS to manage the data within the bureau. The managed information is the following. Police officers monitor areas, and each area has to be monitored by at least one officer. Officers contact informers to collect information about incidents, and each informer is uniquely identified by the combination of his/her name and the officer he/she contacts. A police officer registers vehicles. The following is the E-R diagram.



(a) (7 pts) Specify cardinality constraints and participation constraints shown in the ER diagram.

	Cardinality Constraints (e.g., "M:N", "N:1", or "1:N", or "1:1")	Participation Constraints (use "Total" or "Partial" to fill in the blanks)
Monitor	M:N	Area (Total) Monitor (Partial) Police Officer
Contact	M:1	Informer (Total) Contact (Partial) Police Officer
Register	M:N	Police Officer (Partial) Register (Partial) Vehicle

(b) (6 pts) Briefly explain the meaning of the bold box surrounding "Informer" and why it is necessary.

It means "Informer" is a weak entity set. It is necessary because the entity set doesn't have enough local attributes to form a key. It needs the "oid" attribute from the "PoliceOfficer" entity set, together with its attribute "Name" to form a primary key.

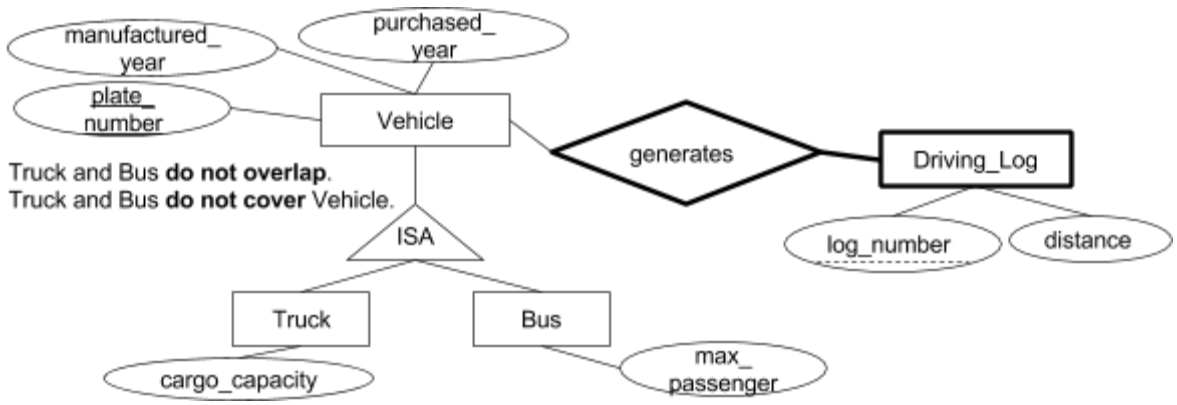
(c) (7 pts) Suppose we want to modify the ER diagram to specify the fact that a vehicle should be registered by one and only police officer. How do make the changes? Using the following space to only re-draw the parts in the original diagram with the changes.

SCORE: _____

Question 2: E-R to Relational Translation (25 points)

In this question, when constructing an SQL DDL statement, be sure to create the columns using appropriate types. Also specify primary key(s), foreign key(s), **NOT NULL** constraints, and their DELETE policies.

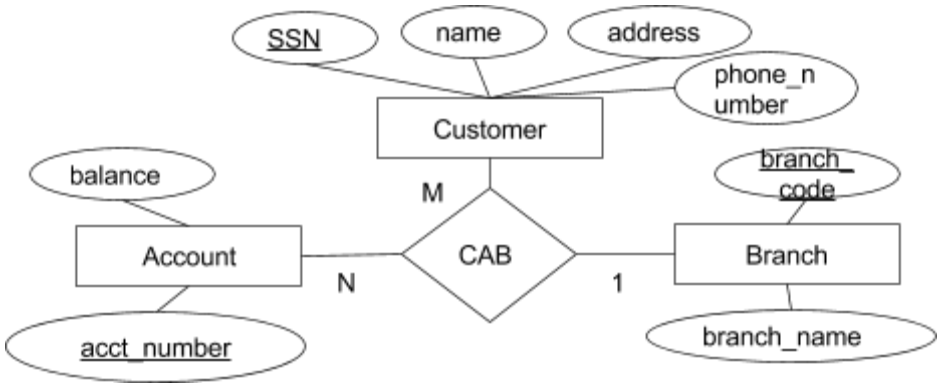
(a) (10 pts) Construct SQL DDL statements for the tables for the following E-R diagram.




```
CREATE TABLE Vehicle (
        plate_number VARCHAR(20) NOT NULL, manufactured_year INT,
        purchased_year INT, primary key plate_number);

CREATE TABLE Truck (
        plate_number VARCHAR(20) NOT NULL, cargo_capacity INT,
        PRIMARY KEY plate_number,
        FOREIGN KEY (plate_number) REFERENCES Vehicle (plate_number) ON DELETE
CASCADE);

CREATE TABLE Bus (
        plate_number VARCHAR(20) NOT NULL, max_passenger INT,
        PRIMARY KEY plate_number,
        FOREIGN KEY (plate_number) REFERENCES Vehicle (plate_number) ON DELETE
CASCADE);

CREATE TABLE Driving_Log (
        plate_number VARCHAR(20) NOT NULL, log_number INT NOT NULL,
        distance REAL,
        PRIMARY KEY (plate_number, log_number),
        FOREIGN KEY (plate_number) REFERENCES Vehicle (plate_number) ON DELETE
CASCADE);
```

SCORE: _____

Questions (b) - (d) are based on the following E-R diagram. It represents banking information about customers, bank branches, and bank accounts. It has a ternary relationship among **Customer**, **Branch**, and **Account**.



(b) (3 pts) Which of the following statements is true based on the given E-R diagram?

a. Each <**ssn**, **acct_number**> pair can identify a unique **branch_code**.
b. Each <**ssn**, **branch_code**> pair can identify a unique **acct_number**.
c. Each <**branch_code**, **acct_number**> can identify a unique **ssn**.
d. None of the above

(c) (4 pts) Construct an SQL DDL statement for a "**Customer**" table corresponding to the "**Customer**" entity set.

```
CREATE TABLE Customer (
        SSN CHAR(9) NOT NULL,  name VARCHAR(20),
        address VARCHAR(100), phone_number VARCHAR(20),
        PRIMARY KEY (SSN)
);
```

(d) (8 pts) Assume we have created **Customer**, **Account**, and **Branch** tables corresponding to the entity sets, with the underlined attributes as their primary keys respectively. Construct an SQL DDL statement for a "CAB" table corresponding to the "CAB" relationship set. Use a "DELETE" policy for the foreign key constraints so that no **CAB** records will be deleted when records in **Customer**, **Account**, and **Branch** tables are deleted.

```
CREATE TABLE CAB (
        SSN CHAR(9) NOT NULL, acct_number INT NOT NULL, code INT NOT NULL,
        PRIMARY KEY (SSN, acct_number),
        FOREIGN KEY (SSN) REFERENCES Customer(SSN) ON DELETE NO ACTION,
        FOREIGN KEY (acct_number) REFERENCES Account(acct_number) ON DELETE NO
ACTION,
        FOREIGN KEY (branch_code) REFERENCES Branch(branch_code) ON DELETE NO
ACTION
);
```

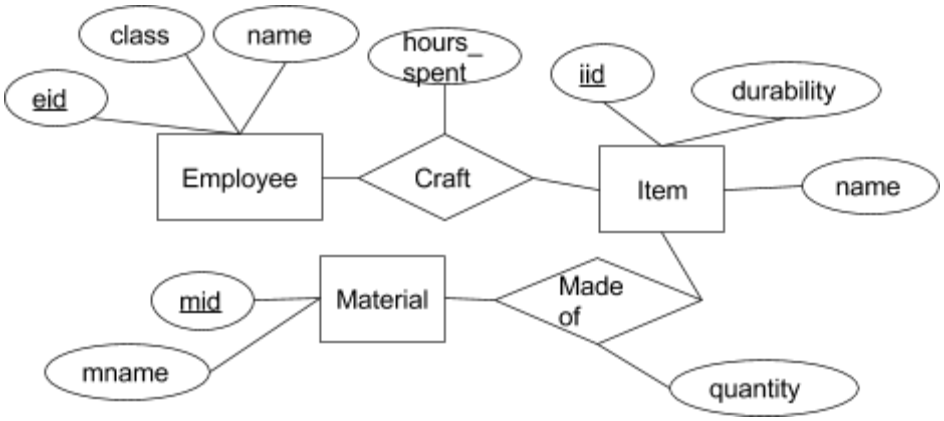
SCORE: _____

Question 3: Relational Algebra (25 points)

You are the owner of a blacksmith shop on an island called "Azeroth". There are several employees working for you, and each of them can craft items based on recipes. Each recipe requires certain materials. You have a database with the following schemas:

1. **Employee(**eid:number, name:string, class:string)
2. **Item(**iid:number, name:string, durability:number)
3. **Craft**(eid:number, iid:number, hours_spent:number}
4. **Material**(mid:number, mname:string}
5. **Madeof**(iid:number, mid:number, quantity:number)

The following is the corresponding ER diagram:



Write the following queries using relational algebra.

(a) (5 pts) Find iids of items who have an "hours_spent" less than 50.
π rid σ hour_spent < 50 (Craft)

(b) (6 pts) Find eids of employees who can craft all the items.
π eid, iid Craft ÷ π iid Item

(c) (7 pts) Find names of materials required by items crafted by an employee from the class 'Paladin' or 'Priest'.
π mname ((π iid((π eid σ Class = 'Paladin' ∨ Class = 'Priest' Employee) ⋈ Craft) ⋈ π iid, mid Madeof) ⋈ Materials)

(d) (7 pts) Find iids of items that can be crafted by an employee named 'Thall' but cannot be crafted by an employee named 'Uther'.
π iid (π eid σ Name = 'Thall' Employee ⋈ Craft) - π iid (π eid σ Name = 'Uther' Employee ⋈ Craft)

SCORE: _____

Question 4: SQL (20 points)

You are interviewed by a shoe company for a data science expert position. A database has been already implemented and populated, with the following schema:

```
Shoe ( sid INTEGER,
       brand:VARCHAR(20),
       size INTEGER,
       color VARCHAR(20),
       price DECIMAL(6,2),
       PRIMARY KEY sid);

Customer ( cid INTEGER,
           name VARCHAR(30),
           shoe_size INTEGER,
           gender VARCHAR(6),
           PRIMARY KEY cid);

Purchase ( cid INTEGER,
           sid INTEGER,
           purchase_date DATE,
           PRIMARY KEY (CID, SID),
           FOREIGN KEY (CID) REFERENCES Customer (cid),
           FOREIGN KEY (SID) REFERENCES Shoe (sid));
```

(a) (5 pts) List the ids and genders of customers who purchased shoes on "October-20-2015".

```
SELECT C.cid, C.gender
FROM Customer C, Purchase P
WHERE C.cid = P.cid AND P.purchase_date = '2015-10-20';
```

(b) (6 pts) List names of female customers (without duplicates) that have ***not*** purchased a "Nike" shoe.

```
SELECT DISTINCT C.name
FROM Customer C
WHERE C.gender = 'female' AND C.cid NOT IN (
                    SELECT P.cid
                    FROM Purchase P, Shoe S
                    WHERE S.sid = P.sid AND S.brand = 'Nike');
```

SCORE: _____

(c) (4 pts) For each brand, list the average price of shoes with this brand.

```
SELECT S.brand, AVG(S.price)
FROM Shoe S
GROUP BY S.brand;
```

(d) (5 pts) List ids of customers who purchased at least one shoe from every brand.

```
SELECT C.cid
FROM Customer C
WHERE NOT EXISTS ( SELECT DISTINCT S1.brand
                   FROM Shoe S1
                   EXCEPT
                   SELECT DISTINCT S2.brand
                   FROM Purchase P, Shoe S2
                   WHERE P.cid = C.cid AND P.sid = S2.sid);
```

SCORE: _____

Question 5: Functional Dependencies and Keys (10 pts)

There is a relation **R**(A, B, C, D, E) with five attributes and the following two functional dependencies:

- AB → CDE
- CD → ABE

Suppose AB is a primary key. For each set of attributes in the following table, specify YES/NO in each column.

Attributes	Candidate Key?	Super Key?	Primary Key?
AB	Yes	Yes	Yes
ABC	No	Yes	No
CD	Yes	Yes	No
CDE	No	Yes	No

SCORE: _____