CS 377: Database Systems

Homework #2 Solutions

1. **Hospital Database** (35 points): Map the ER diagram for a Hospital System Database shown below in Figure 1. The relational model should follow these guidelines:

   - Use the *smallest* possible number of relations
   - Minimize the number of attributes in the resulting relations with NULL values
   - Use the entity names in the ER-diagram as the name of the relation
   - Use the attribute names in the ER-diagram for attribute names in the relational model
   - If you augment an existing relation with an attribute x to represent a relationship R, give the attribute the name: $R.x$
   - If you define a new relation to represent a relationship, use the name of the relationship for the name of the relation
   - **Underline** the *primary key* in every relation
   - Draw an arrow from each *foreign key* to its corresponding *primary key*

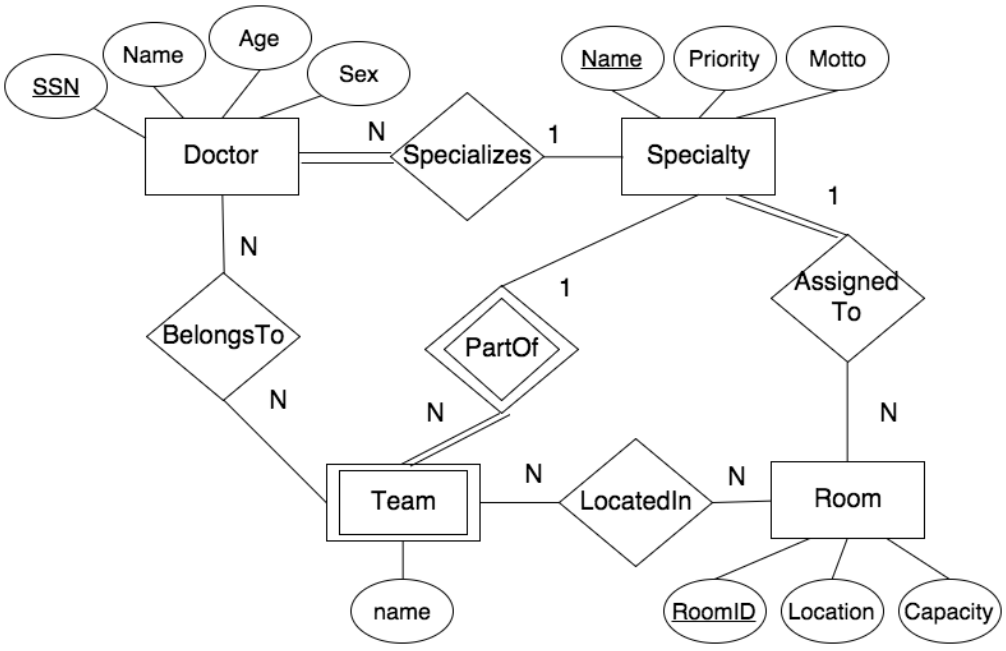

Figure 1: ER diagram for Hospital Database

(**ANSWER**) From the diagram, we convert the entities and their associated attributes to a relation:

- Doctor(SSN, Name, Age, Sex)
- Specialty(Name, Priority, Motto)

- Room(RoomID, Location, Capacity)
- Team(Specialty.Name, Name)

The next step is to take care of the relationships. We will favor expanding existing relations over creating new relations

- Specializes(Doctor, Specialty) expands the Doctor relation to a new attribute Specialty.name
- BelongsTo(Doctor, Team) introduces a new relation BelongsTo with foreign keys (Doctor.SSN and Team.Name, Specialty.Name)
- LocatedIn(Room, Team) introduces a new relation LocatedIn with foreign keys (Room.RoomID and Team.Name, Specialty.Name)
- AssignedTo(Specialty, Room), will expand the Room relation to have a Specialty.name. Note that NULL values may exist as room need not have a specialty but this is better than having a new relation.

Putting everything together yields the following relational data model.

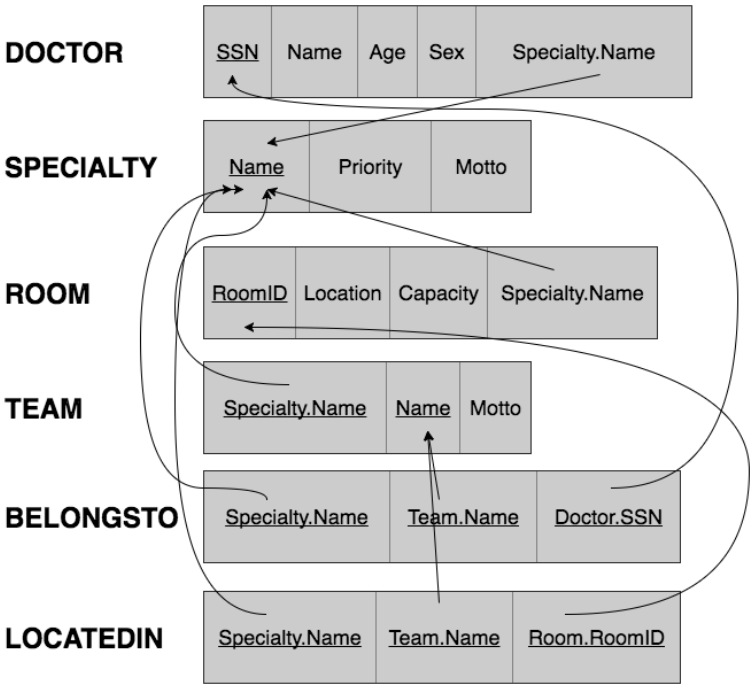

Figure 2: Relational Data Model for Hospital

2. **Library** (5+5 = 10 points): Consider the following relational schema for a library:

```
member(member_no, name, dob)
books(isbn, title, authors, publisher)
borrowed(memb_no, isbn, date)
```

Express the following queries in Relational Algebra:

(a) Find the name of members who have borrowed all books published by "McGraw-Hill".

(b) For each publisher, find the name and membership number of members who have borrowed more than five books of that publisher.

(**ANSWER**)

(a) Solution 1: Count each members number of books borrowed that was published by "McGraw-Hill" and join it with the total number of books published by McGraw-Hill.

$$\text{MGH\_BOOKS} = \sigma_{\text{publisher='McGraw-Hill'}}(\text{BOOKS})$$
$$\text{MEM\_BOOKS} = {}_{\text{memb\_no}}\mathcal{F}_{\text{count(isbn)}}(\text{BORROWED} * \text{MGH\_BOOKS})$$
$$\text{MGH\_COUNT} = \mathcal{F}_{\text{count(isbn)}}(\text{MGH\_BOOKS}) \qquad \text{(number of MGH books)}$$
$$\text{ANSWER} = \pi_{\text{name}}(\text{MEMBER} \bowtie_{\text{member\_no=memb\_no}} \text{MEM\_BOOKS} \bowtie_{\text{count=count}} \text{MGH\_COUNT})$$

Solution 2: Use division

$$\text{MGH\_BOOKS} = \pi_{\text{isbn}}(\sigma_{\text{publisher='McGraw-Hill'}}(\text{BOOKS}))$$
$$\text{BORROW\_AUX} = \pi_{\text{memb\_no, isbn}}(\text{BORROWED})$$
$$\text{MEM\_AUX} = \text{BORROW\_AUX} \div \text{MGH\_BOOKS}$$
$$\text{ANSWER} = \pi_{\text{name}}(\text{MEM\_AUX} * \text{MEMBER})$$

(b) We want to group on publisher and member number to find the number of books per member per publisher.

$$\text{MPB} = {}_{\text{publisher,memb\_no}}\mathcal{F}_{\text{count(isbn)}}(\text{BORROWED} * (\pi_{\text{isbn,publisher}}(\text{BOOKS})))$$
$$\text{ANSWER} = \pi_{\text{publisher,name,memb\_no}}(\text{MEMBER} \bowtie_{\text{member\_no=memb\_no}} (\sigma_{\text{count}>5}(\text{MPB})))$$

3. **Company Database Relational Algebra Queries** (10+10+10+10 = 40 points): Consider the company database relational data model discussed in class and shown below. Formulate the following queries in Relational Algebra:

   (a) Find the name of the projects in Atlanta that have been worked on at least a total of 100 person hours.

   (b) Find the name of the department(s) that pay the highest salary.

   (c) Find the fname & lname of the employee(s) who work the highest total number of hours.

   (d) Find the department(s) in which *all* employees in the department have at least one dependent.

   (**ANSWER**)

   (a)

$$\text{PH} = {}_{\text{Pname,Pno}}\mathcal{F}_{\text{sum(Hours)}}(\sigma_{\text{Plocation='Atlanta'}}(\text{WORKS\_ON} \bowtie_{\text{Pno=Pnumber}} \text{PROJECT}))$$
$$\text{ANSWER} = \pi_{\text{Pname}}(\sigma_{\text{sum}\geq 100}(\text{PH}))$$

   (b)

$$\text{MS} = \mathcal{F}_{\text{max(Salary)}}(\text{EMPLOYEE})$$
$$\text{MSE} = \text{EMPLOYEE} \bowtie_{\text{Salary=max}} (\text{MS})$$
$$\text{ANSWER} = \pi_{\text{Dname}}(\text{DEPARTMENT} \bowtie_{\text{Dnumber=Dno}} \text{MSE})$$

   (c)

$$\text{TEH(Essn, TotalHours)} = {}_{\text{Essn}}\mathcal{F}_{\text{sum(Hours)}}(\text{WORKS\_ON})$$
$$\text{MT(max)} = \mathcal{F}_{\text{max(TotalHours)}}(TEH)$$
$$\text{EWMT} = \text{TEH} \bowtie_{\text{TotalHours=max}} (\text{MT})$$
$$\text{ANSWER} = \pi_{\text{Fname,Lanme}}(\text{EMPLOYEE} \bowtie_{\text{Ssn=Essn}} \text{EWMT})$$

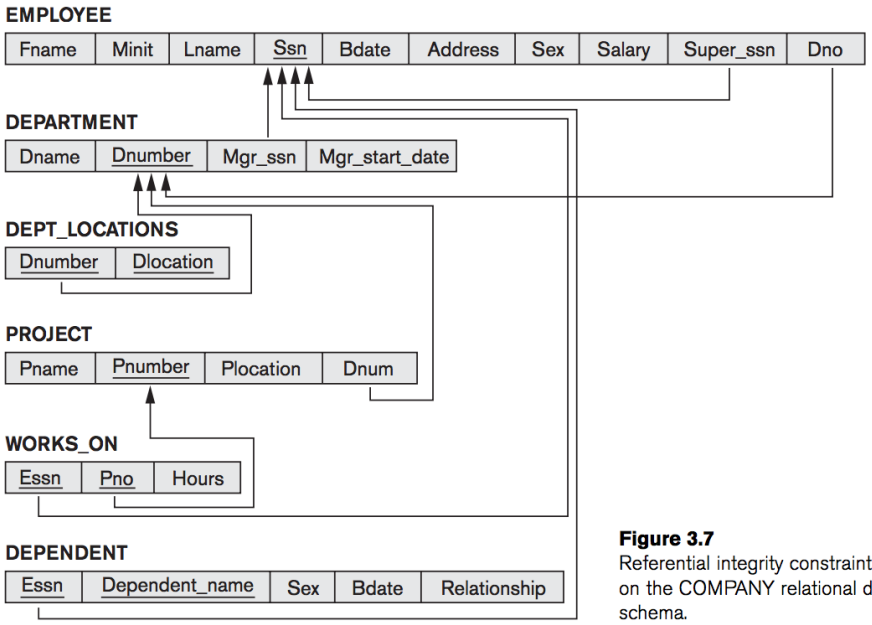

Figure 3.7
Referential integrity constraints displayed on the COMPANY relational database schema.

Figure 3: Company Database Relational Model

(d) Solution #1: Find the departments in which no employees have zero dependents

$$\text{R1} = \pi_{\text{Ssn}}(\text{EMPLOYEE} \bowtie_{\text{Ssn=Essn}} \text{DEPENDENT}) \quad \text{(employees with some dependent)}$$
$$\text{R2} = \pi_{\text{Ssn, Dno}}(\text{EMPLOYEE}) - \text{R1} \quad \text{(employees with zero dependents)}$$
$$\text{R3} = \pi_{\text{Dno}}(\text{DEPARTMENT} \bowtie_{\text{Dnumber=Dno}} \text{R2}) \quad \text{(departments with some employees that have zero dependents)}$$
$$\text{ANSWER} = \pi_{\text{Dno}}(\text{DEPARTMENT}) - \text{R3}$$

Solution # 2: Count the number of employees with dependents per employee and the number of employees per department. If they are equal, then that department has at least one dependent.

$$\text{R1} = {}_{\text{Dnumber}}\mathcal{F}_{\text{count(Ssn)}}(\text{DEPARTMENT} \bowtie_{\text{Dnumber=Dno}} \text{EMPLOYEE} \bowtie_{\text{Ssn=Essn}} \text{DEPENDENT})$$
$$\text{R2} = {}_{\text{Dnumber}}\mathcal{F}_{\text{count(Ssn)}}(\text{DEPARTMENT} \bowtie_{\text{Dnumber=Dno}} \text{EMPLOYEE})$$
$$\text{ANSWER} = \pi_{\text{Dno}}(\text{DEPARTMENT} \bowtie_{\text{Dnumber=Dnumber}} \text{R1} \bowtie_{\text{Dnumber=Dnumber} \wedge \text{count=count}} \text{R2})$$

4. **Company Database Relational Calculus Queries** (5+5+5 = 15 points): Consider the company database relational database and formulate the following queries in *Relational Calculus*:

   (a) Retrieve the names of all employees in department 5 who work more than 10 hours per week on the ProductX project.

   (b) List the names of all employees who have a dependent with the same first name as themselves.

   (c) Find the names of all employees who are directly supervised by Franklin Wong.

   (**ANSWER**)

(a)

$$\{t.\text{fname}, t.\text{lname} \mid \text{Employee}(t) \text{ AND } t.\text{dno} = 5 \text{ AND } ((\exists p)(\exists w)(\text{PROJECT}(p) \text{ AND } \text{WORKS\_ON}(w) \text{ AND } p.\text{pname} = \text{'ProductX' AND } p.\text{pnumber} = w.\text{pnum} \text{ AND } w.\text{hours} > 10 \text{ AND } w.\text{essn} = t.\text{ssn}))\}$$

(b)

$$\{t.\text{fname}, t.\text{lname} \mid \text{Employee}(t) \text{ AND } ((\exists d)(\text{DEPENDENT}(d) \text{ AND } d.\text{dependent\_name} = t.\text{fname AND } d.\text{essn} = t.\text{ssn}))\}$$

(c)

$$\{t.\text{fname}, t.\text{lname} \mid \text{Employee}(t) \text{ AND } ((\exists m)(\text{Employee}(m) \text{ AND } m.\text{fname} = \text{'Franklin' AND } m.\text{lname} = \text{'Wong' AND } t.\text{Super\_ssn} = m.\text{ssn}))\}$$