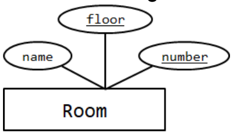
CS2102 Database Systems

Semester 1 2019/2020

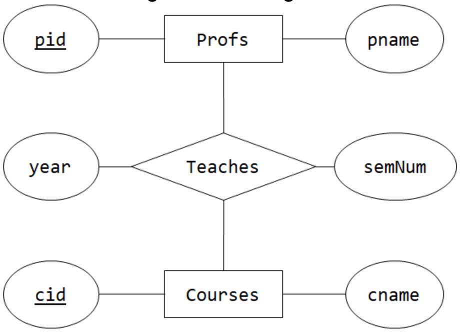
Tutorial 04

Quiz

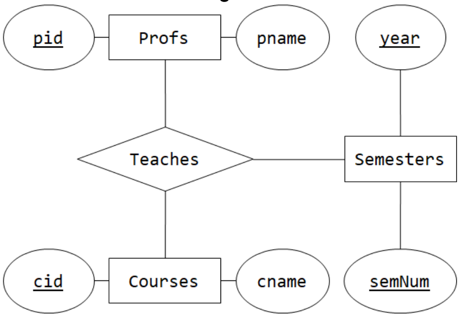
1. Consider the given ER diagram below. Write the SQL code to create the schema.



2. Consider the given ER diagram below. Write the SQL code to create the schema.



3. Consider the ER diagram below. Write the SQL code to create the schema.



4. You are creating a simple search engine. Your search engine works on a following schema `R(phrase, result)`. You are tasked to search for all the result that contains the phrase `'base'` anywhere in the phrase including the middle of the phrase.
5. Consider the following two schemas `R(a,b)` and `S(c,d)`. You are to find all the distinct pairs `(a,c)` from `R` and `S` that satisfies a certain condition `cond`. Is the `DISTINCT` keywords necessary to ensure that the result will always be a distinct pairs `(a,c)`? An example SQL query is shown below.
`SELECT _________ a,c FROM R, S WHERE cond.`
6. Consider the definition of full outer join. Which of the following SQL query is equivalent to $\pi_{a,b}(R \leftrightarrow_c S)$?
   a) `SELECT a,b FROM R FULL JOIN S ON c;`
   b) `SELECT a,b FROM R LEFT JOIN S ON c UNION SELECT a,b FROM R RIGHT JOIN S ON c;`
   c) `SELECT a,b FROM R RIGHT JOIN S ON c UNION SELECT a,b FROM R LEFT JOIN S ON c;`
   d) `SELECT a,b FROM ( R LEFT JOIN S ON c ) RIGHT JOIN S ON c;`
   e) `SELECT a,b FROM R LEFT JOIN ( S RIGHT JOIN S ON c ) ON c;`
   f) `SELECT a,b FROM R RIGHT JOIN S ON c UNION SELECT a,b FROM R LEFT JOIN S ON c UNION R JOIN S ON c;`
7. You are given a simplified map where a city is represented as a square. The city hall is located right at the center of the city. We say that the city is at coordinate `(0,0)` since your schema is `Map(x_coord, y_coord)`.
   We use the notation `(+,-)` to denote that the `x_coord` is positive value and `y_coord` is negative value. We change the sign symbol correspondingly. We say that a house is in the first quadrant (`Q1`) if its coordinate is in `(+,+)`, second quadrant (`Q2`) if its coordinate is in `(-,+)`, third quadrant (`Q3`) if its coordinate is in `(-,-)`, and the fourth quadrant (`Q4`) if its coordinate is in `(+,-)`. You are guaranteed that there will be no houses on `x_coord = 0` or `y_coord = 0` as they are the two main roads on the city.
   Write SQL query to output the quadrants of each set of coordinates in the `Map`.

Tutorial Questions *[Discussion: 8(a), 8(b), 8(c), 8(d), 9]*

Questions 8-9 uses the following database schema as Tutorial 03 Question 07. They are reproduced here for convenience.

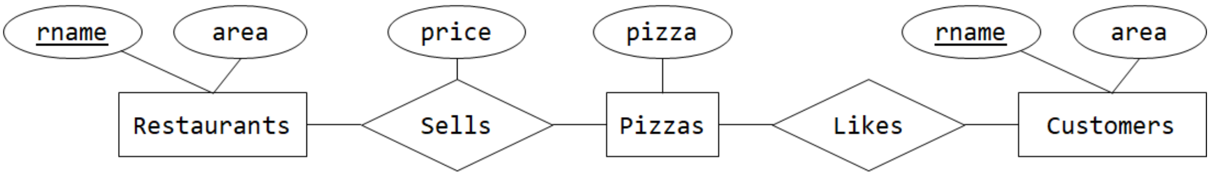
- **`Pizzas`** (<u>`pizza`</u>)
- **`Customers`** (<u>`cname`</u>, `area`)
- **`Restaurants`** (<u>`rname`</u>, `area`)
- **`Contains`** (<u>`pizza, ingredient`</u>)
- **`Sells`** (<u>`rname, pizza`</u>, `price`)
- **`Likes`** (<u>`cname, pizza`</u>)

`Pizzas` indicates all the pizzas of interest. **`Customers`** indicates the name and location of each customer. **`Restaurants`** indicates the name and location of each restaurant. **`Contains`** indicates the ingredients used in each pizza. **`Sells`** indicates the pizzas sold by restaurants and their prices. **`Likes`** indicates the pizzas that customers like.

The following are all the foreign key constraints on the database schema:

- `Contains.pizza` is a foreign key that refers to `Pizzas.pizza`
- `Sells.rname` is a foreign key that refers to `Restaurants.name`
- `Sells.pizza` is a foreign key that refers to `Pizzas.pizza`
- `Likes.cname` is a foreign key that refers to `Customers.pizza`
- `Likes.pizza` is a foreign key that refers to `Pizzas.pizza`

Furthermore, we provide the simplified ER diagram corresponding to the schema above minus `Contains`.



8. Answer each of the following queries using SQL. For parts (a) to (e), remove duplicate records from all query results.
   a) Find all pizzas that 'Alice' likes but 'Bob' does not like.
   b) For each customer, find the pizzas sold by restaurants that are located in the same area as the customer's area. Exclude customers whose associated set of pizzas is empty.
   c) For each customer, find the pizzas sold by restaurants that are located in the same area as the customer's area. Include customers whose associated set of pizzas is empty.
   d) Tutorial 03 Question 07 supposes the existence of relation `Dislikes` (<u>`cname, pizza`</u>). The relation indicates the pizzas that customers do not like. The database still also satisfies the following constraint: for every customer $c \in \pi_{\text{cname}}(\text{Customers})$ and for every pizza $p \in \pi_{\text{pizza}}(\text{Contains})$, either $(c, p) \in \text{Likes}$ or $(c, p) \in \text{Dislikes}$ (*in other words, if a customer do not like a pizza, the customer automatically dislike the pizza*). Create the `Dislikes` relation as a `VIEW`.

9. Consider the following query. For each customer, find the pizzas sold by restaurants with a price under \$20 where the restaurants is located in the same area as the customer's area. _Exclude_ customers whose associated set of pizzas is empty.
   The following are two possible SQL answers (denoted by Q1 and Q2) for the above query.
   Q1:
```
SELECT DISTINCT C.cname, S.pizza
FROM   ( Customers C JOIN Restaurants R ON C.area = R.area )
       JOIN Sells S ON R.rname = S.rname;
WHERE  price < 20;
```
   Q2:
```
SELECT DISTINCT C.cname, S.pizza
FROM   ( Customers C JOIN Restaurants R on C.area = R.area )
       JOIN Sells S ON R.rname = S.rname AND price < 20;
```
   Observe that Q1 and Q2 are similar except that the predicate on price appears as a selection predicate in Q1 but appears as part of the second join predicate in Q2. The semantics of these two SQL queries are defined by the relational algebra expressions[1] shown below.

10. The Prescriptions-R-X chain of pharmacies has offered to give you a free lifetime supply of medicine if you design its database. Given the rising cost of health care, you agree. Here's the information that you gather:
    - Patients are identified by an SSN, and their names, addresses, and ages must be recorded.
    - Doctors are identifier by an SSN. For each doctor, the name, specialty, and years of experience must be recorded.
    - Each pharmaceutical company is identified by name and has a phone number. For each drug, the trade name, and formula must be recorded. Each drug is sold by a given pharmaceutical company, and the trade name identifies a drug uniquely from among the products of that company. If a pharmaceutical company is deleted, you need not keep track of its products any longer.
    - Each pharmacy has a name, address, and phone number.
    - Every patient has a primary physician. Every doctor is the primary physician of at least one patient.
    - Each pharmacy sells several drugs and has a price for each. A drug could be sold at several pharmacies, and the price could vary from one pharmacy to another.
    - Doctors prescribe drugs for patients. A doctor could prescribe one or more drugs for several patients, and a patient could obtain prescriptions from several doctors. Each prescription has a date and a quantity associated with it. You can assume that, if a doctor prescribes the same drug for the same patient more than once, only the last such prescription needs to be stored.
    - There is exactly one contract between a pharmacy and a pharmaceutical company if and only if that pharmacy sells some drug that is made by that pharmaceutical company. For each contract, you have to store a start date, an end date, and the text of the contract.

    a) Consider the ER diagram shown on the next page for Prescriptions-R-X. What are the constraints that are not captured by this design? Modify the ER design to capture as many of the constraints as possible.
    b) Translate your ER design in part (a) into a relation schema using SQL (assume reasonable data types for the domain constraints). Your solution should capture as many of the application's constraints as possible. Identify any constraints that are not captured by your relational schema.
    c) How would your design in part (a) change if each drug must be sold at a fixed price by all pharmacies?
    d) How would your design in part (a) change if the design requirements change as follows: If a doctor prescribes the same drug for the same patient more than once, several such prescriptions may have to be stored.

[1] For convenience, we abuse the relational algebra notation to use `Customers C` to denote $\rho_C(\text{Customers})$

e) Suppose that pharmacies appoint a supervisor for each contract. There must always be a supervisor for each contract, but the contract supervisor can change over the lifetime of the contract. Supervisors are identified by an SSN, and their start dates must be recorded. Modify your ER design in part (a) to capture this additional requirement.

f) Translate your ER design in part (a) into a relational schema. Identify any constraints that are not captured by your relational schema.