CS2102: Database Systems

Tutorial #4: *SQL (Part 2)*

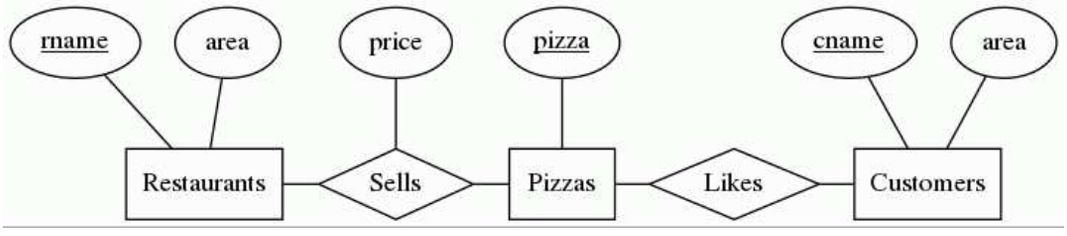
Week 6

AY 2022/23 Sem 2

1 Discussions

The following questions are to be discussed during tutorial. All answers will be released with explanation.

This tutorial discussion questions are based on the following pizza database schema. The ER diagram is shown below.



The ER diagram produces the following schemas:

Relation	Description
$Pizzas(\underline{pizza})$	All the pizzas of interest.
$Customers(\underline{cname}, area)$	The name and location of each customer.
$Restaurants(\underline{rname}, area)$	The name and location of each restaurant.
$Recipes(\underline{pizza, ingredients})$	The ingredients used in each pizza.
$Sells(\underline{rname, pizza}, price)$	Pizzas sold by restaurants and the prices.
$Likes(\underline{cname, pizza})$	Pizzas that customers like.

Additionally, we have the following foreign key constraints on the database schema:

- $(Recipes.pizza) \rightsquigarrow (Pizzas.pizza)$
- $(Sells.rname) \rightsquigarrow (Restaurants.rname)$
- $(Sells.pizza) \rightsquigarrow (Pizzas.pizza)$
- $(Likes.cname) \rightsquigarrow (Customers.cname)$
- $(Likes.pizza) \rightsquigarrow (Pizzas.pizza)$

1. **(Simple Query)** For each of the following queries, write an *equivalent* SQL query that does **not** use any subquery. Note that we do not consider set operation (*e.g., in* `Q1 UNION Q2`*, neither* `Q1` *nor* `Q2` *are considered subqueries*).

   Can say if there's aggregation in Where. always change to Having?

   (a) Query A

```sql
SELECT DISTINCT cname
FROM   Likes L
WHERE  EXISTS (
  SELECT 1
  FROM   Sells S
  WHERE  S.rname = 'Corleone Corner'
    AND  S.pizza = L.pizza
);
```

   (b) Query B

```sql
SELECT cname
FROM   Customers C
WHERE  NOT EXISTS (
  SELECT 1
  FROM   Likes L, Sells S
  WHERE  S.rname = 'Corleone Corner'
    AND  S.pizza = L.pizza
    AND  C.cname = L.cname
);
```

   (c) Query C

```sql
SELECT DISTINCT rname
FROM   Sells
WHERE  rname <> 'Corleone Corner'
  AND  price > ANY (
    SELECT price
    FROM   Sells
    WHERE  rname = 'Corleone Corner'
);
```

   (d) Query D

```sql
SELECT rname, pizza, price
FROM   Sells S
WHERE  price >= ALL (
  SELECT S2.price
  FROM   Sells S2
  WHERE  S2.rname = S.rname
    AND  S2.price IS NOT NULL
);
```

2. **(SQL Query)** Write an SQL query to answer each of the following questions on the pizza database *without using aggregate functions*. Remove duplicate records from all query results.

   (a) Find pizzas that Moe likes but is not liked by Lisa.

   (b) Find pizzas that are sold by at most one restaurant in each area; exclude pizzas that are not sold by any restaurant.

   (c) Find all tuples (A, P, P_{min}) where P is a pizza that is available in area A (*i.e.*, there is some restaurant in area A selling pizza P) and P_{min} is the *lowest* price of P in area A.

3. **(Equivalence)** Consider the query to find distinct restaurants that are located in the East area. The following are two possible SQL answers (*denoted by* Q_1 *and* Q_2) for this query.

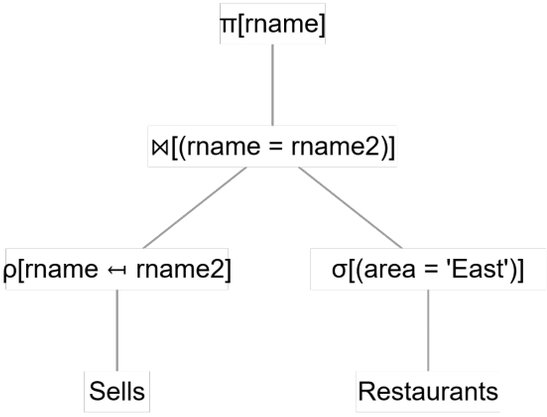
   Q_1 Query 1

```sql
SELECT DISTINCT S.rname
FROM   Sells S JOIN Restaurants R
  ON   S.rname = R.rname AND R.area = 'East';
```

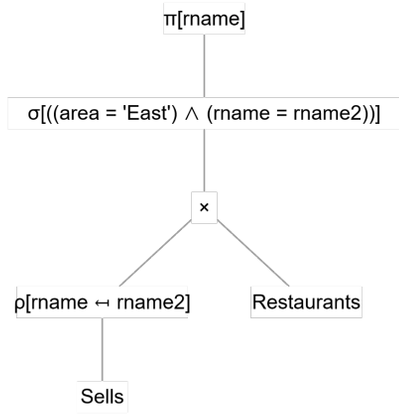
Q_2 Query 2

```
SELECT DISTINCT S.rname
FROM   Sells S, Restaurants R
WHERE  S.rname = R.rname
  AND R.area = 'East';
```

The semantics of these two SQL queries are defined by the relational algebra expressions shown below. Discuss whether Q_1 and Q_2 are equivalent queries.



(a) Q_1



(b) Q_2

4. **(Equivalence)** Consider the query to find distinct restaurants that are located in the East area or restaurants that sell some pizza that Lisa likes, where the restaurants that do not sell any pizza are to be excluded. The following are two possible SQL answers (*denoted by* Q_1 *and* Q_2) for this query.

Q_1 Query 1

$R \bowtie \phi = \phi$.

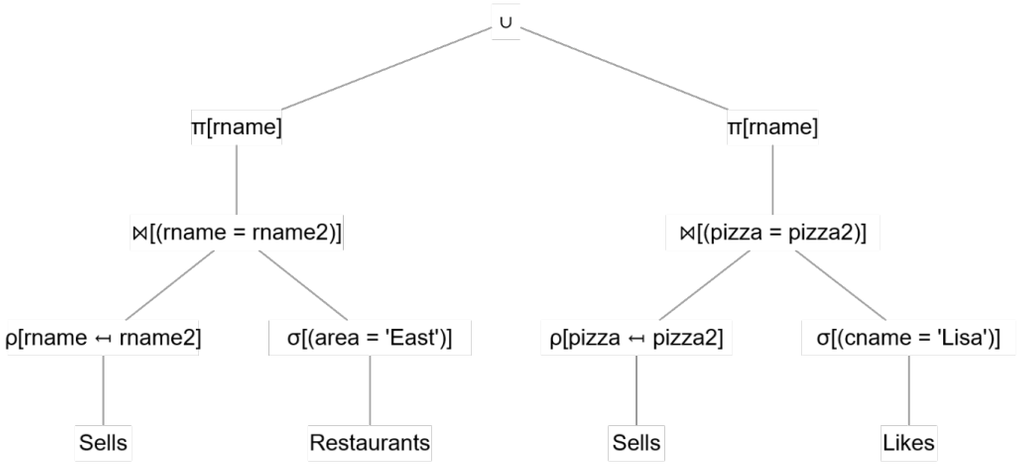
```
SELECT DISTINCT S.rname
FROM   Sells S JOIN Restaurants R
  ON   S.rname = R.rname AND R.area = 'East'
UNION
SELECT DISTINCT S.rname
FROM   Sells S JOIN Likes L
  ON   S.pizza = L.pizza AND L.cname = 'Lisa';
```

Q_2 Query 2

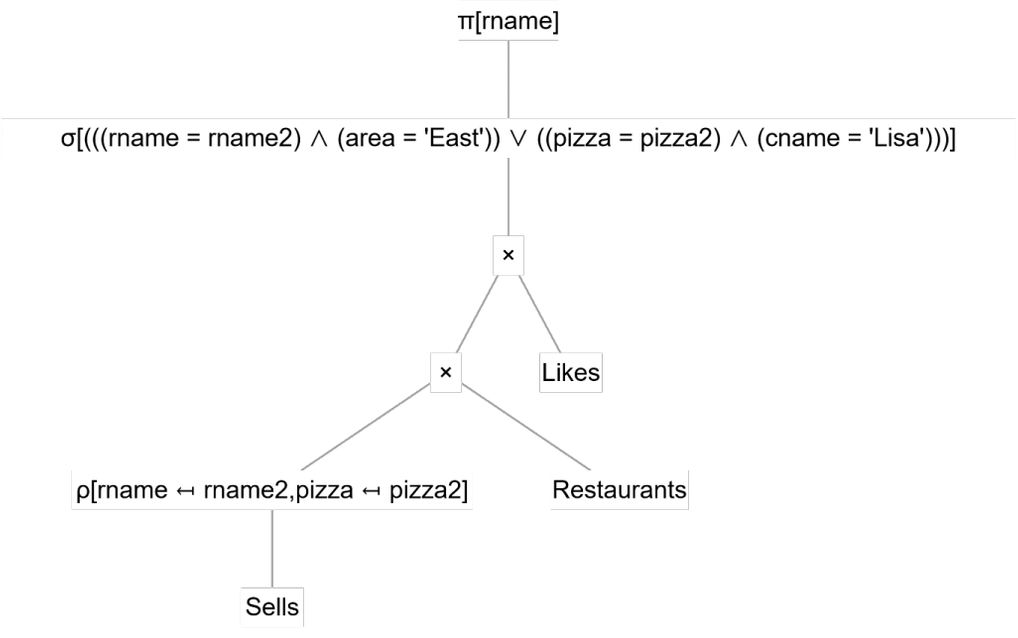
```
SELECT DISTINCT S.rname
FROM   Sells S, Restaurants R, Likes L
WHERE  (S.rname = R.rname AND R.area = 'East')
  OR   (S.pizza = L.pizza AND L.cname = 'Lisa');
```

The semantics of these two SQL queries are defined by the relational algebra expressions shown below. Discuss whether Q_1 and Q_2 are equivalent queries.

NOT. Make Likes empty.


∪
π[rname]
π[rname]
⋈[(rname = rname2)]
⋈[(pizza = pizza2)]
ρ[rname ↦ rname2]
σ[(area = 'East')]
ρ[pizza ↦ pizza2]
σ[(cname = 'Lisa')]
Sells
Restaurants
Sells
Likes

(a) Q_1


π[rname]
σ[(((rname = rname2) ∧ (area = 'East')) ∨ ((pizza = pizza2) ∧ (cname = 'Lisa')))]
×
×
Likes
ρ[rname ↦ rname2,pizza ↦ pizza2]
Restaurants
Sells

(b) Q_2

2 Challenge

The answers to the following questions is given without explanation. Please discuss them on Canvas.

1. **(SQL Query)** Write an SQL query to answer each of the following questions on the pizza database *without using aggregate functions*. Remove duplicate records from all query results.

   (a) Find all tuples (A, P, P_{min}, P_{max}) where P is a pizza that is available in area A (*i.e.*, there is some restaurant in area A selling pizza P), P_{min} is the *lowest* price of P in area A and P_{max} is the *highest* price of P in area A.

2. **(Update)** Consider *again* the following relational schema discussed in Tutorial 2 Question 2 (*Discussions*).

```
CREATE TABLE Offices (
  office_id    INTEGER,
  building     TEXT NOT NULL,
  level        INTEGER NOT NULL,
  room_number  INTEGER NOT NULL,
  area         INTEGER,
  PRIMARY KEY (office_id),
  UNIQUE (building, level, room_number)
);

CREATE TABLE Employees (
  emp_id      INTEGER,
  name        TEXT NOT NULL,
  office_id   INTEGER NOT NULL,
  manager_id  INTEGER,
  PRIMARY KEY (emp_id),
  FOREIGN KEY (office_id) REFERENCES Offices (office_id)
    ON UPDATE CASCADE,
  FOREIGN KEY (manager_id) REFERENCES Employees (emp_id)
    ON UPDATE CASCADE
);
```

   Suppose that the office with `office_id = 123` needs to be renovated. Write an SQL statement to reassign the employees located in this office to another temporary office located at room number 11 on level 5 at the building named *Tower1*.

   Hint: You can use subquery in an update statement.

3. **(Backward Reasoning)** You are given the following schema:

   - `Students(matric, sname)`
   - `Workings(pid, matric, since)`
   - `Projects(pid, pname)`
   - `Categories(pid, cname)`

   You are also given the following foreign key:

   - (`Workings.matric`) $\rightsquigarrow$ (`Students.matric`)
   - (`Workings.pid`) $\rightsquigarrow$ (`Projects.pid`)
   - (`Categories.pid`) $\rightsquigarrow$ (`Projects.pid`)

   Consider the following SQL query:

```
SELECT *
FROM   Students NATURAL JOIN Workings
               NATURAL JOIN Projects
               NATURAL JOIN Categories;
```

Say it produces the following result:

matric	sname	pid	since	cname
A0001	AA	P01	2002	CA
A0001	AA	P01	2002	CB
A0001	AA	P02	2004	CB
A0002	BB	P01	2003	CA
A0002	BB	P01	2003	CB
A0003	CC	P03	2004	CA
A0003	CC	P03	2004	CC
A0003	CC	P03	2004	CD
A0004	AA	P03	2004	CA
A0004	AA	P03	2004	CC
A0004	AA	P03	2004	CD

Now consider another the query to find all pair of distinct projects' pid (*i.e.*, `(p1, p2)`) such that the two projects have exactly the same set of categories[1]. It produces the following result:

pid	pid
P01	P04
P04	P01
P03	P05
P05	P03

Simply note that P01 and P04 have exactly the same set of categories. Similarly, P03 and P05 have exactly the same set of categories. What is a possible result of the following SQL query? Show your answer using only P01, P03, P04, and P05?

```
SELECT pid FROM Categories
EXCEPT ALL
SELECT DISTINCT pid FROM Categories WHERE cname <> 'CA';
```

[1] This involves some constructs that you will only learn in Lecture 06.