CS2102 Tutorial 1: Relational Model   Wk 3, Sem 2, 2020/21

1. Consider the following relation instance r of the relational schema $R(A, B, C, D)$.

A	B	C	D
0	0	0	1
2	1	2	0
1	1	2	0
0	0	1	2

   (a) Based on r, write down all the possible superkeys of R.

   (b) In addition to r, suppose that it is also known that $\{A, C\}$ is a superkey of R. Based on the given information, write down all the possible candidate keys of R. Which of these (if any) is definitely a candidate key of R?

2. Consider a relational database consisting of two relations with schema R(<u>A</u>,B) and S(<u>W</u>,X,Y,Z), where the primary keys of R and S are $\{A\}$ and $\{W\}$, respectively. Let r and s be the current instances of R and S, respectively, as shown below.

r

A	B
3	0
2	1
1	1
0	0

s

W	X	Y	Z
0	4	0	null
1	null	2	null
2	1	2	null
3	0	1	null

   Based on the current database instance, write down all the possible foreign keys in S that refer to attribute A in R.

3. Two queries Q_1 and Q_2 on a relational database with schema D are defined to be **equivalent queries** (denoted by $Q_1 \equiv Q_2$) if for every legal instance d of D, both Q_1 and Q_2 compute the same results on d.

   Consider a database with the following relational schema: $R(\underline{A}, C)$, $S(\underline{A}, D)$, and $T(\underline{X}, Y)$, with the primary key attributes underlined. Assume that all the attributes have integer domain. For each of the following pairs of queries Q_1 and Q_2, state whether or not Q_1 and Q_2 are equivalent queries.

   (a) $Q_1 = \pi_A(\sigma_{A<10}(R))$ and $Q_2 = \sigma_{A<10}(\pi_A(R))$

   (b) $Q_1 = \pi_A(\sigma_{C<10}(R))$ and $Q_2 = \sigma_{C<10}(\pi_A(R))$

   (c) $Q_1 = \pi_{D,Y}(S \times T)$ and $Q_2 = \pi_D(S) \times \pi_Y(T)$

   (d) $Q_1 = \pi_{D,Y}(S \times T)$ and $Q_2 = \pi_{D,Y}(T \times S)$

   (e) $Q_1 = (R \times \pi_D(S)) \times T$ and $Q_2 = R \times (\pi_D(S) \times T)$

   (f) $Q_1 = \pi_A(R \cup S)$ and $Q_2 = \pi_A(R) \cup \pi_A(S)$

   (g) $Q_1 = \pi_A(R - S)$ and $Q_2 = \pi_A(R) - \pi_A(S)$

4. Consider the following relational database schema discussed in class, where the primary key of each relation is underlined.

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

   Answer each of the following queries using relational algebra.

   (a) Find pizzas that Alice likes but Bob does not like.

   (b) Find all customer-restaurant pairs (C, R) where C and R are both located in the same area, and C likes some pizza that is sold by R.

   (c) Find all customer-pizza pairs (C, P) where C does not like P.

   (d) Find all customer pairs $(C1, C2)$ where $C1$ likes some pizza that $C2$ does not like.

   (e) Find all customer pairs $(C1, C2)$ where $C1 < C2$ and both customers like exactly the same pizzas (i.e., $C1$ likes P if and only if $C2$ likes P). Exclude customer pairs that do not like any pizza.

   (f) For each restaurant, find the price of the most expensive pizzas sold by that restaurant. Exclude restaurants that do not sell any pizza.

   (g) Find all customer-pizza pairs (C, P) where the pizza P is sold by some restaurant that is located in the same area as that of the customer C. Include customers whose associated set of pizzas is empty.

5. Consider the following relational algebra query expressed on the database schema in Question 4.

$$
\begin{array}{|l|}
\hline
R_1 = \pi_{pizza}(\sigma_{cname='Maggie'}(\text{Likes})) \\
R_2 = \pi_{rname}(\text{Sells}) \times R_1 \\
R_3 = \pi_{rname}(R_2 - \pi_{rname,pizza}(\text{Sells})) \\
R_4 = \pi_{rname}(\text{Sells}) - R_3 \\
R_5 = \pi_{pizza}(\sigma_{cname='Ralph'}(\text{Likes})) \\
R_6 = \pi_{rname}(\sigma_{pizza5=pizza}((\text{Sells} \times \rho_{pizza:pizza5}(R_5)))) \\
R_7 = R_4 - R_6 \\
\hline
\end{array}
$$

For each of the above relational algebra expressions R_i, write down a concise English sentence to precisely describe the information retrieved by R_i.