Tutorial 6: Basic Relational Algebra

Jin, Ziyang
34893140
f4a0b

Kim, Joon Hyung
35183128
l1m8

February 2018

Consider a database consisting of the relations, with primary keys in bold:

customer(**cid**, cname, rating, salary)
order(**cid, iid, day,** qty)
item(**iid**, iname, type, price)

customer:

cid	cname	rating	salary
1	Alan	1	1000
2	Butler	2	1000
3	Charles	3	1000
4	Dennis	4	1000
5	Frances	5	1000

order:

cid	iid	day	qty
1	1	2018-01-01	5
2	2	2018-01-02	5
3	3	2018-01-03	5
5	4	2018-01-04	5

item

iid	iname	type	price
1	xps	laptop	1500
2	mac	laptop	1500
3	civic	car	30000
4	charms	lolipop	1
5	sweet	watermelon	3

1

1 Find the details (i.e., all attributes) of the customers who have a rating higher than 6 and earn less than $125,000.

$$\sigma_{rating>6 \ \wedge \ salary<125000}(customer)$$

2 Find the names and types of items which were ordered by some customer named 'Bob' who is rated above 5.

$$\pi_{iname,type}(\sigma_{cname='Bob' \ \wedge \ rating>5}(customer) \bowtie item)$$

3 Consider the query $\pi_{iname,type}(item)$. Suppose item has 1000 tuples. Then how many tuples will the result of the above projection query contain? Explain your answer.

It depends on the actual instance of *item*. The number of tuples can vary from 1 to 1000.

(iname type) is not a key. There is also no uniqueness constraint on iname and type. So it is possible that iname and type are the same for these 1000 tuples. In this case, the resulting projection will remove duplicates and result in 1 tuple. It is also possible that iname and type are distinct from tuple to tuple, so the resulting projection will have 1000 distinct tuples. Therefore, it is possible that some of the (iname type) pairs are duplicates and some of them are distinct so the result can vary from 1 to 1000.

4 Suppose in addition to the key dependencies identified by the primary keys in the given relations, we are told that every item name uniquely determines its type, i.e., that the FD iname $\rightarrow$ type holds. Suppose there are 950 distinct item names in the item table.

1. Then how many tuples will be present in the result of the following RA^{SPJ} query. Explain you answer.

$$\pi_{iname,type}(item)$$

There will be 950 tuples in the result. Because there are 950 distinct item names in the item table. If we project (iname type) pair, none of the resulting pairs will be removed as a duplicate because each pair has a unique iname. So there will be 950 tuples in the result.

2. Would your answer change if the query instead was:

$$\pi_{iname,type,price}(item)$$

No. The resulting tuple is removed only all of the attributes in the tuple are the same as another. Since `iname`'s are distinct, the entire resulting tuples are distinct. Therefore, there is no duplicate tuple to be removed in the result. The result will have exactly 950 tuples.

5 Find the type and name of those items, whose price is at most $100 OR which were ordered by the customer with cid=123.

$$\pi_{type,iname}(\sigma_{price \leq 100}(item)) \cup \pi_{type,iname}(\sigma_{cid=123}(order) \bowtie item)$$

6 Consider the query: "Find the names of customers who did not order any laptop, i.e., did not order any item of type laptop" and the RA^{SPJ} query. Does the RA^{SPJ} query above correctly capture the English query above? Justify your answer.

$$\pi_{cname}(customer \bowtie order \bowtie \sigma_{type \neq 'laptop'}(item))$$

No. The join between customer and order only selects customers that have ordered something. If a customer has never ordered anything, the customer still qualifies the case "who did not order any laptop". However, in this join, these qualifying customers who have not ordered anything are lost. Therefore, the RA^{SPJ} query above does not correctly capture the English query above.