CIS 631: DATA MANAGEMENT SYSTEMS DESIGN

ASSIGNMENT 1 - SOLUTIONS

Consider the following schema:

SUPPLIERS (SID: *integer*, SNAME: *string*, STREET: *string*, CITY: *string*, ZIP: *string*)
PARTS (PID: *integer*, PNAME: *string*, COLOR: *string*) CATALOG
(SID: *integer*, PID: *integer*, COST: *real*)

The primary key attributes are underlined, and the domain of each attribute is listed after the attribute name. Thus, SID is the primary key for SUPPLIERS, PID is the primary key for PARTS, and SID and PID together form the primary key for CATALOG. Attribute SID in CATALOG is a foreign key that refers to the primary key SID of SUPPLIERS and PID in CATALOG is a foreign key that refers to the primary key PID of PARTS. The CATALOG relation lists the prices charged for parts by suppliers.

State what the following query computes.

0. $\pi_{CITY}(\pi_{PID}(\sigma_{COLOR = 'red'}(PARTS)) * \sigma_{COST>100}(CATALOG) * \pi_{SID, CITY}(SUPPLIERS))$

It computes the city of the suppliers supplying a red part that costs more than $100.

Write the following queries in relational algebra.

1. Find the names of parts for which there is some supplier

$\pi_{PNAME}(PARTS * CATALOG)$

2. Find the names of parts supplied by suppliers who are at 1 Central Ave.

$\pi_{PNAME}(PARTS * CATALOG * \sigma_{STREET='1\ Central\ Ave'}(SUPPLIERS))$

3. Find the names of suppliers who supply some red part.

$\pi_{SNAME}(\pi_{SNAME, SID}(SUPPLIERS) * \pi_{SID, PID}(CATALOG) * \pi_{PID}(\sigma_{COLOR = 'red'}(PARTS)))$

4. Find the SIDs of suppliers who supply some red or green part.

$\pi_{SID}(\pi_{SID, PID}(CATALOG) * \pi_{PID}(\sigma_{COLOR = 'red'\ OR\ COLOR = 'green'}(PARTS)))$

5. Find the SID of suppliers who supply some red part or whose address is '221 Packer Street'.

T_1 $\pi_{SID}(\pi_{PID}(\sigma_{COLOR = 'red'}(PARTS)) * \pi_{SID, PID}(CATALOG))$

T_2 $\pi_{SID}(\sigma_{ADDRESS = '221 Parcker Street'}(SUPPLIERS))$

RESULT $T_1 \cup T_2$ */* We can perform this operation because T_1 and T_2 are union compatible.*

6. Find the SIDs of suppliers who supply some red part and some green part.

T_1 $\pi_{SID}(\pi_{PID}(\sigma_{COLOR = 'red'}(PARTS))* \pi_{SID, PID}(CATALOG))$

T_2 $\pi_{SID}(\pi_{PID}(\sigma_{COLOR = 'green'}(PARTS))* \pi_{SID, PID}(CATALOG))$

RESULT $T_1 \cap T_2$

7. Find the PIDs of parts that are red or are supplied by a supplier who is at the city of Newark.

T_1 $\pi_{PID}(\sigma_{COLOR='Red'}(PARTS))$

T_2 $\pi_{PID}(CATALOG * \sigma_{CITY='Newark'}(SUPPLIERS))$

Result $T_1 \cup T_2$ */* Note that we can perform this operation because T_1 and T_2 are union compatible.*

8. Find the PIDs of parts supplied by a supplier who is at the city of Newark and by a supplier who is at the city of Trenton.

T_1 $\pi_{PID}(CATALOG * \sigma_{CITY='Newark'}(SUPPLIERS))$

T_2 $\pi_{PID}(CATALOG * \sigma_{CITY='Trenton'}(SUPPLIERS))$

Result $T_1 \cap T_2$

Note that the following expression is <u>NOT a correct answer</u>.

Result $\pi_{PID}(CATALOG * \sigma_{CITY='Newark' \text{ AND } CITY='Trenton'}(SUPPLIERS))$

This expression computes the PIDs of parts supplied by a supplier who is located at both the city of Newark and the city of Trenton.

Observe also that there are no suppliers in the database located at both the city of Newark and the city of Trenton (why?)

9. Find the PIDs of parts supplied by each and every supplier.

$\pi_{PID, SID}(CATALOG) \div \pi_{SID}(SUPPLIER)$

10. Find the PIDs of parts supplied by each and every supplier who supplies at least one part.

$\pi_{SID, PID}(CATALOG) \div \pi_{SID}(CATALOG)$

11. Find the PIDs of parts supplied by each and every supplier who is at the city of Newark or at the city of Trenton (equivalently: find the PIDs of parts supplied by each and every supplier who is at the city of Newark and by each and every supplier who is at the city of Trenton).

$\pi_{SID, PID}(CATALOG) \div \pi_{SID}(\sigma_{CITY='Newark' \text{ OR } CITY='Trenton'}(SUPPLIERS))$

Equivalently:

T_1 $\pi_{SID, PID}(CATALOG) \div \pi_{SID}(\sigma_{CITY='Newark'}(SUPPLIERS))$

T_2 $\pi_{SID, PID}(CATALOG) \div \pi_{SID}(\sigma_{CITY='Trenton'}(SUPPLIERS))$

Result $T_1 \cap T_2$

12. Find the PIDs of parts supplied by each and every supplier who is at the city of Newark or by each and every supplier who is at the city of Trenton.

T_1 $\pi_{SID, PID}(CATALOG) \div \pi_{SID}(\sigma_{CITY='Newark'}(SUPPLIERS))$

T_2 $\pi_{SID, PID}(CATALOG) \div \pi_{SID}(\sigma_{CITY='Trenton'}(SUPPLIERS))$

Result $T_1 \cup T_2$

13. Which one of the queries 11 and 12 is more restrictive (if any)?

Query 12 is more general (11 is more restrictive). For instance, a PID of a part supplied by every supplier of Trenton but not by every supplier of Newark appears in the answer of 12 but not in the answer of 11.

14. Find pairs of PIDs such that the part with the first PID is sold at a higher price by a supplier than the part with the second PID.

$\pi_{PID, PID1}(\sigma_{COST>COST1}(CATALOG * \rho_{PID \rightarrow PID1, COST \rightarrow COST1}(CATALOG)))$

15. Find the SIDs of suppliers who supply at least two different parts (you are not allowed to use a grouping/aggregation operation for this query).

$\pi_{SID}(\sigma_{PID \neq PID1}(\pi_{SID, PID}(CATALOG) * \rho_{PID \rightarrow PID1}(\pi_{SID, PID}(CATALOG))))$

16. Find the SIDs of suppliers who supply at least two different parts (you have to use a grouping/aggregation operation for this query).

T_1 $_{SID}F_{count(PID)}$ (CATALOG)

RESULT π_{SID} ($\sigma_{count(PID) \geq 2}$ (T_1))

17. For every part supplied by a supplier who is at the city of Newark, print the PID and the SID and the name of the suppliers who sell it at the highest price.

T_1 π_{PID} (CATALOG * $\sigma_{CITY='Newark'}$ (SUPPLIERS))

T_2 $_{PID}F_{MAX(COST)}$ (T_1 * CATALOG)

T_3 $\rho_{MAX(COST) \rightarrow COST}$ (T_2)

T_4 $\pi_{PID, SID}$ (T_3 * CATALOG) /* Notice that relations T_3 and CATALOG have two common attributes: COST and PID.*/

RESULT $\pi_{PID, SID, SNAME}$ (T_4 * $\pi_{SID, SNAME}$ (SUPPLIERS))

18. For every part, find its PID, its PNAME and the number of suppliers who sell it.

$\pi_{PID, PNAME, COUNT(SID)}$ (($_{PID}F_{COUNT(SID)}$ CATALOG) * PARTS)

19. List the PID, PNAME and average cost of all parts.

$_{PID, PNAME}F_{avg(COST)}$ (CATALOG * PARTS)

20. Find the average cost of red parts.

$F_{avg(COST)}$ ($\sigma_{COLOR='Red'}$(PARTS) * CATALOG)

21. Find the average cost of parts supplied by suppliers named 'Yosemite Sham'.

$F_{avg(COST)}$ (CATALOG * π_{SID} ($\sigma_{SNAME = 'Yoshemite Sham'}$ (SUPPLIERS)))