Introduction to Databases
Relational Calculus

Dr Paolo Guagliardo

(v26.0)

CHANGELOG

v26.0 Initial version.

Problem 1. Consider a schema consisting of a ternary relation R and a binary relation S, and let D be the database (instance) such that:

$$R^D = \{(1,1,2), (1,0,0), (1,2,2), (1,1,0), (1,2,3), (2,2,3)\}$$
$$S^D = \{(0,0), (2,1), (1,3), (1,2)\}$$

Compute the answer to each of the following queries on D under active domain semantics:

$$Q_1 = \big\{x \mid \exists y, z\ R(y,z,x) \wedge z \neq y \wedge S(x,x)\big\}$$
$$Q_2 = \big\{x, y \mid \forall z\ R(x,z,y) \rightarrow \big(S(x,z) \vee S(y,z)\big)\big\}$$
$$Q_3 = \big\{z \mid \neg\exists y\ S(y,z) \wedge S(z,y)\big\}$$

Justify your answers appropriately.

Solution. We have $\mathsf{adom}(D) = \{0,1,2,3\}$.

(Q_1) Given an assignment ν, the interpretation (D,ν) satisfies $\exists y, z\ R(y,z,x) \wedge z \neq y \wedge S(x,x)$ iff it satisfies

$$\bigvee_{a,b \in \mathsf{adom}(D)} \big(R(b,a,x) \wedge a \neq b \wedge S(x,x)\big)$$

As the only pair in S^D in which the first and second components are the same is $(0,0)$, any assignment ν that maps x to any constant other than 0 cannot (together with D) make the atom $S(x,x)$ true, and therefore it cannot satisfy the formula above. When $\nu = \{x \mapsto 0\}$, the formula is satisfied by (D,ν) because (at least) one of its disjuncts is true; indeed, $D, \nu \models R\big(1,0,\nu(x)\big) \wedge 0 \neq 1 \wedge S\big(\nu(x),\nu(x)\big)$. Thus, $Q_1(D) = \{0\}$.

(Q_2) We need to consider 16 possible assignments: $\nu = \{x \mapsto i,\ y \mapsto j\}$ for $i, j \in \mathsf{adom}(D)$. Of these, all that map x to 0 or 3, along with all that map y to 1, trivially satisfy the body of Q_2 by making the left-hand side of the implication false (no matter which value z takes), because R^D contains no triples in which the first component is 0 or 3, nor triples where the third component is 1. Thus, all of the following pairs are in $Q_2(D)$:

$$(0,0), (0,1), (0,2), (0,3), (3,0), (3,1), (3,2), (3,3), (1,1), (2,1)$$

Now, under active domain semantics, the body of the query is equivalent to

$$\bigwedge_{a \in \mathsf{adom}(D)} \neg R(x,a,y) \vee S(x,a) \vee S(y,a)$$

which is short for $\varphi_0 \wedge \varphi_1 \wedge \varphi_2 \wedge \varphi_3$, where:

$$\varphi_0 = \neg R(x, 0, y) \vee S(x, 0) \vee S(y, 0)$$
$$\varphi_1 = \neg R(x, 1, y) \vee S(x, 1) \vee S(y, 1)$$
$$\varphi_2 = \neg R(x, 2, y) \vee S(x, 2) \vee S(y, 2)$$
$$\varphi_3 = \neg R(x, 3, y) \vee S(x, 3) \vee S(y, 3)$$

We have already considered all assignments (to x and y) for which $\neg R(y, z, x)$ is false independently of the value of z. There are 6 assignments left to check:

- $\{x \mapsto 1,\ y \mapsto 0\}$ makes φ_1, and in turn the whole formula, false;
- $\{x \mapsto 1,\ y \mapsto 2\}$ makes the formula true and so it is an answer to the query;
- $\{x \mapsto 1,\ y \mapsto 3\}$ makes the formula true and so it is an answer to the query;
- $\{x \mapsto 2,\ y \mapsto 0\}$ makes the formula true and so it is an answer to the query;
- $\{x \mapsto 2,\ y \mapsto 2\}$ makes the formula true and so it is an answer to the query;
- $\{x \mapsto 2,\ y \mapsto 3\}$ makes φ_2, and in turn the whole formula, false.

Thus, in addition to the pairs identified earlier, $Q_2(D)$ will also contain $(1, 2), (1, 3), (2, 0), (2, 2)$.

(Q_3) Given an assignment ν, the interpretation (D, ν) satisfies $\neg\exists y\ S(y, z) \wedge S(z, y)$ iff it satisfies

$$\neg \bigvee_{a \in \mathsf{adom}(D)} \big(S(a, z) \wedge S(z, a)\big)$$

By the DeMorgan's laws, the above formula is equivalent to

$$\bigwedge_{a \in \mathsf{adom}(D)} \neg\big(S(a, z) \wedge S(z, a)\big)$$

which in turn is short for

$$\neg\underbrace{\big(S(0, z) \wedge S(z, 0)\big)}_{\varphi_0} \wedge \neg\underbrace{\big(S(1, z) \wedge S(z, 1)\big)}_{\varphi_1} \wedge \neg\underbrace{\big(S(2, z) \wedge S(z, 2)\big)}_{\varphi_2} \wedge \neg\underbrace{\big(S(3, z) \wedge S(z, 3)\big)}_{\varphi_3}$$

There are 4 assignments to check:

- $\{z \mapsto 0\}$ makes φ_0 true and, in turn, the whole formula false;
- $\{z \mapsto 1\}$ makes φ_2 true and, in turn, the whole formula false;
- $\{z \mapsto 2\}$ makes φ_1 true and, in turn, the whole formula false;
- $\{z \mapsto 3\}$ makes all φ_i's false and so it is an answer to the query.

Thus, $Q_3(D) = \{3\}$.

Problem 2. Given a schema consisting of a binary relation R and a ternary relation S, write a relational calculus query that computes the active domain.

Solution. Recall that the active domain of a database D is the set of all constants occurring in D. Thus, in RA, a query that computes the active domain (of any given database over which it is executed) must take the union of all projections on each single attributes of every relation. Since projecting over an attribute/position in RC means existentially quantifying all others, the query we are looking for is the following:

$$\Big\{x \mid \big(\exists y\ R(x,y)\big) \vee \big(\exists y\ R(y,x)\big) \vee \big(\exists y,z\ S(x,y,z)\big) \vee \big(\exists y,z\ S(y,x,z)\big) \vee \big(\exists y,z\ S(y,z,x)\big)\Big\}$$

Intuitively the above says: "return all values x that appear in the first column of R, or in the second column of R, or in the first column of S, or in the second column of S, or in the third column of S." Observe that we can use the same variables for all disjuncts (but careful: this would not work with conjunction) and write more succinctly:

$$\big\{x \mid \exists y,z\ R(x,y) \vee R(y,x) \vee S(x,y,z) \vee S(y,x,z) \vee S(y,z,x)\big\}$$

Problem 3. Consider the following schema:

Customer($\cdot$, $\cdot$, $\cdot$) of arity 3, where the first position is the customer's *ID*, the second is the customer's *Name*, and the third is the *City* where the customer lives.

Account ($\cdot$, $\cdot$, $\cdot$, $\cdot$) of arity 4, in which the first position is the account's *Number*, the second is its *Branch*, the third is the ID of the customer owning the account, and the fourth is the *Balance* on the account.

Write the following queries in relational calculus (directly, without translating from relational algebra):

(1) "ID and name of customers who own an account in a branch in their city."

(2) "ID and name of customers who do **not** own any account."

(3) "ID and name of customers who own an account with a balance which is no less than the balance of any other account."

Solution.

(1) $\{x_1, x_2 \mid \exists x_3, y_1, y_4\ \textsc{Customer}(x_1, x_2, x_3) \wedge \textsc{Account}(y_1, x_3, x_1, y_4)\}$

(2) $\{x_1, x_2 \mid \exists x_3\ \textsc{Customer}(x_1, x_2, x_3) \wedge \neg\exists y_1, y_2, y_4\ \textsc{Account}(y_1, y_2, x_1, y_4)\}$

(3) $\{x_1, x_2 \mid \exists x_3\ \textsc{Customer}(x_1, x_2, x_3) \wedge \exists y_1, y_2, y_4\ \textsc{Account}(y_1, y_2, x_1, y_4)$
$\wedge \forall z_1, z_2, z_3, z_4\ (\textsc{Account}(z_1, z_2, z_3, z_4) \rightarrow \neg(y_4 < z_4))\}$