B561 Assignment 4: Introduction to SQL
Solutions to Part I

C.M. Wyss

November 14, 2001

All questions for this assignment concern two relational tables: ZONES and ROADS. The SQL CREATE TABLE statements for these tables follow.

\[
\begin{align*}
\text{CREATE TABLE ZONES} & \quad \text{(ZONEID NUMBER, TYPE CHAR(1), PRIMARY KEY (ZONEID))} \\
\text{CREATE TABLE ROADS} & \quad \text{(ROADID NUMBER, SRCZONE NUMBER, ENDZONE NUMBER, DIST NUMBER, PRIMARY KEY (ROADID))}
\end{align*}
\]

The table ZONES lists the zones associated with a particular city by their ID number (a positive integer) and their type (R = residential, C = commercial, I = industrial).

The table ROADS lists the roads connecting various zones and the distance (in miles) between the zones along that road. All roads are one-way, meaning that if there is a road from zone \( i \) to zone \( j \) (i.e. \( \langle x, i, j, y \rangle \in \text{ROADS} \)), this says nothing about the existence of any roads from zone \( j \) to zone \( i \) (or lack thereof). Make no assumptions about the connectivity of zones other than the tuples provided in the ROADS relation.

Your solutions to parts I and II (below) must work on any database instance with tables ROADS and ZONES of the appropriate schemas.

To illustrate the queries below, consider the example database in figure 1.

**Part I**

1. Formulate each of the following queries in the safe domain relational calculus (sDRC). If your solution involves a complicated formula, break it up into sub-formulas and briefly describe (in English) what each signifies and how they are composed to obtain the final expression.

   (a) Find all zones of type I or R. On the sample database, this query yields the zones \{1, 2, 6, 7, 8, 9, 10\}.

   **Solution:**

   \[
   \{ \langle z \rangle \mid \exists t (\langle z, t \rangle \in \text{ZONES} \land (t = \text{'R'} \lor t = \text{'T'})) \}
   \]
(b) Find those roads that go from a zone of type I to a zone of type R, or vice-versa. On the sample database, this query yields the roads \( \{3, 7, 9\} \).

**Solution:** First, we’ll describe the property that a road begins in an ‘I’ or an ‘R’ type zone.

\[
\varphi_1(r, t) := \exists s \exists e \exists d \exists z (\langle r, s, e, d \rangle \in \text{ROADS} \land \langle z, t \rangle \in \text{ZONES} \\
\land z = s \land (t = ‘R' \lor t = ‘I')
\]

Next, we’ll describe the property that a road ends in an ‘I’ or an ‘R’ type zone.

\[
\varphi_2(r, t) := \exists s \exists e \exists d \exists z (\langle r, s, e, d \rangle \in \text{ROADS} \land \langle z, t \rangle \in \text{ZONES} \\
\land z = e \land (t = ‘R' \lor t = ‘I')
\]

Finally, we’ll put these two together to find the required roads:

\[
\{ \langle r \rangle \mid \exists t_1 \exists t_2 (\varphi_1(r, t_1) \land \varphi_2(r, t_2) \land t_1 \neq t_2) \}\.
\]

(c) Find all zones that have no roads leaving them. On the sample database, this query yields the zones \( \{3, 4, 9\} \).

**Solution:**

\[
\{ \langle z \rangle \mid \exists t (\langle z, t \rangle \in \text{ZONES}) \\
\land \neg (\exists t \forall s \exists e \exists d (\langle z, t \rangle \in \text{ZONES} \land \langle r, s, e, d \rangle \in \text{ROADS} \land z = s)) \}
\]
(d) Find all roads that do not connect zones of differing type. Thus, for example, a road \( x \) would qualify if it connects a residential zone to another residential zone. On the sample database, this query yields the roads \{1, 4, 5, 11, 12, 14, 15\}.

Solution:

\[
\{ r \mid \exists z_1 \exists z_2 \exists t_1 \exists s \exists e \exists d ( (z_1, t_1) \in ZONES \\
\land (z_2, t_2) \in ZONES \land (r, s, e, d) \in ROADS \land z_1 = s \land z_2 = 3 \land t_1 = t_2 ) \}
\]

(e) Find all zones from which a road goes to at least one of each type of zone. Thus, for example, a residential zone, \( z_r \), would qualify if there is at least one road going from it to another zone of type R, at least one road going from it to a zone of type I, and at least one road going from it to a zone of type C. On the sample database, this query yields the zones \{1, 7\}.

Solution: First, we will capture the property that there is a road going from a zone to another zone of type ‘R’.

\[\varphi_R(z_1, r) := \exists z_2 \exists t_2 \exists t_1 \exists s \exists e \exists d ( (z_1, t_1) \in ZONES \land (z_2, t_2) \in ZONES \\
\land (r, s, e, d) \in ROADS \land t_2 = 'R' \land s = z_1 \land e = z_2 \land z_1 \neq z_2 )\]

We will then capture similar properties for the ‘I’ and ‘C’ zone types:

\[\varphi_R(z_1, r) := \exists z_2 \exists t_2 \exists t_1 \exists s \exists e \exists d ( (z_1, t_1) \in ZONES \land (z_2, t_2) \in ZONES \\
\land (r, s, e, d) \in ROADS \land t_2 = 'I' \land s = z_1 \land e = z_2 \land z_1 \neq z_2 )\]

\[\varphi_R(z_1, r) := \exists z_2 \exists t_2 \exists t_1 \exists s \exists e \exists d ( (z_1, t_1) \in ZONES \land (z_2, t_2) \in ZONES \\
\land (r, s, e, d) \in ROADS \land t_2 = 'C' \land s = z_1 \land e = z_2 \land z_1 \neq z_2 )\]

Finally, we will put these together to obtain the solution:

\[\{ (z) \mid \exists r_1 \exists r_2 \exists r_3 ( \varphi_R(z, r_1) \land \varphi_I(z, r_2) \land \varphi_C(z, r_3) ) \}\].

2. The proof that the sDRC is equivalent to the RA was begun in class. Complete this proof.

Solution: See chapter 5 of Abiteboul, Hull, Vianu, Foundations of Databases.

3. Formulate the queries 1(a), 1(c), and 1(e) in the relational algebra (RA), using the translation method developed in the proof for question 2 to translate your DRC answers to question 1. Make sure you annotate your answers clearly in English, so that your grader can follow every step easily.

Solution:

(a) \(\Pi_{ZONEID}(\sigma_{TYPE = 'T' \lor TYPE = 'R'}(ZONES))\)

3
(c) \( \Pi_{\text{ZONEID}}(\text{ZONES}) = \Pi_{\text{ZONEID}}(\sigma_{\text{ZONEID}=\text{SR CZONE}}(\text{ZONES} \times \text{ROADS})) \)

(e) First, we’ll capture \( \varphi_R, \varphi_I, \) and \( \varphi_C \) with corresponding RA expressions \( E_R, E_I, \) and \( E_C. \)

\[
E_R := \Pi_{\text{ZONEID}_1} \left( \sigma_{ \text{TYPE}_2 = \text{R} } \land \text{SRCZONE} = \text{ZONEID}_1 \land \text{ENDZONE} = \text{ZONEID}_2 \land \text{ZONEID}_1 \neq \text{ZONEID}_2 \right) \left( \rho_{X \rightarrow X_1}(\text{ZONES}) \times \rho_{X \rightarrow X_2}(\text{ZONES} \times \text{ROADS}) \right)
\]

\[
E_I := \Pi_{\text{ZONEID}_1} \left( \sigma_{ \text{TYPE}_2 = \text{I} } \land \text{SRCZONE} = \text{ZONEID}_1 \land \text{ENDZONE} = \text{ZONEID}_2 \land \text{ZONEID}_1 \neq \text{ZONEID}_2 \right) \left( \rho_{X \rightarrow X_1}(\text{ZONES}) \times \rho_{X \rightarrow X_2}(\text{ZONES} \times \text{ROADS}) \right)
\]

\[
E_C := \Pi_{\text{ZONEID}_1} \left( \sigma_{ \text{TYPE}_2 = \text{C} } \land \text{SRCZONE} = \text{ZONEID}_1 \land \text{ENDZONE} = \text{ZONEID}_2 \land \text{ZONEID}_1 \neq \text{ZONEID}_2 \right) \left( \rho_{X \rightarrow X_1}(\text{ZONES}) \times \rho_{X \rightarrow X_2}(\text{ZONES} \times \text{ROADS}) \right)
\]

Our final RA expression is then \( E_R \cap E_I \cap E_C. \)