Exercise sheet for Session 1

Uncertain data management

Antoine Amarilli

November 28th, 2017

1 Exercise 1: Codd tables

Consider the following instance I:

Schedule

class	session	teacher	room
UDM	1	NULL	NULL
FOO	NULL	NULL	C42
UDM	9	NULL	NULL
FOO	NULL	NULL	C43

Lab

class	session
UDM	1
FOO	NULL

Exam

class	session
UDM	9
FOO	NULL

Question 1. Rephrase in plain English what we know about the UDM class, and what we know about the FOO class.

Answer. *The UDM class has a lab session ("1") and an exam session ("9") but we do not know where they take place or who teaches them.*

The FOO class has one lab session and one exam session, and it also has one session in C42 and one session in C43, but we do not know who teaches them, neither do we know whether they are the same as the lab or exam sessions or not.

Question 2. Write a Boolean relational algebra query that asks whether some class has a scheduled lab in the same room as the exam for that class. (By *scheduled*, we mean that the corresponding session of the class also occurs in the Schedule table.)

Answer.

$$\Pi_\emptyset \left((\Pi_{\mathbf{class},\mathbf{room}} (\mathit{Event} \bowtie \mathit{Lab})) \bowtie (\Pi_{\mathbf{class},\mathbf{room}} (\mathit{Event} \bowtie \mathit{Exam})) \right)$$

Another possibility:

$$\Pi_\emptyset \left(\mathit{Schedule} \bowtie \rho_{\mathbf{session} \mapsto \mathbf{s1}, \mathbf{teacher} \mapsto \mathbf{t1}} (\mathit{Schedule}) \bowtie \mathit{Lab} \bowtie \rho_{\mathbf{session} \mapsto \mathbf{s1}} (\mathit{Exam}) \right)$$

Question 3. Is this query *possible* on the instance? If yes, what is an example of a possible world where the query evaluates to true?

Answer. *The query is possible and an example possible world is obtained by mapping all* `NULL`*s to "a".*

Question 4. Is this query *certain* on the instance? If no, what is a counterexample possible world?

Answer. *The query is not certain and a counterexample possible world is obtained by mapping all* `NULL`*s to different values, all of which do not occur in the instance.*

Question 5. Consider the constraint that says that whenever a class has a scheduled lab then it has a scheduled exam with the same teacher (but possibly in a different room).

Is there a possible world that satisfies this constraint? Do all possible worlds satisfy this constraint?

Answer. *Use the same worlds as for Questions 3 and 4.*

Question 6. Replace three pairs of `NULL`s by named `NULL`s to obtain a v-table where the constraint is always respected.

Answer. *The following is a solution:*

Schedule

class	***session***	***teacher***	***room***
UDM	*1*	$NULL_1$	*NULL*
FOO	*NULL*	$NULL_4$	*C42*
UDM	*9*	$NULL_1$	*NULL*
FOO	$NULL_3$	$NULL_4$	*C43*

Lab

class	***session***
UDM	*1*
FOO	*NULL*

Exam

class	***session***
UDM	*9*
FOO	$NULL_3$

2 Exercise 2: v-tables and c-tables

Consider the following instance of a v-table:

Schedule

class	**session**	**teacher**	**room**
UDM	1	$\texttt{NULL}_1$	$\texttt{NULL}_2$
$\texttt{NULL}_3$	2	$\texttt{NULL}_1$	$\texttt{NULL}_4$
FOO	1	John	$\texttt{NULL}_2$

Question 1. What is this relation saying in plain English?

Answer. *There is a class UDM that has a first session, and a class FOO that has a first session in the same room. Further some class has a second session and its teacher is the same as that of the first session of the UDM class.*

Question 2. Write a query Q_1 in the relational algebra that returns the triples of a class *class* and two sessions $s_1 < s_2$ such that s_1 and s_2 are sessions of *class* that have the same teacher.

Answer.

$$\sigma_{\mathbf{s1<s2}}\left(\rho_{\mathbf{session \mapsto s2}}\left(\Pi_{\mathbf{class,session,s1}}\left(\rho_{\mathbf{room \mapsto r1, session \mapsto s1}}(\mathit{Schedule}) \bowtie \mathit{Schedule}\right)\right)\right)$$

Question 3. Evaluate Q_1 on the instance to obtain a c-table R_1.

Answer. *The result is:*

class	***s1***	***s2***	
UDM	*1*	*2*	$\texttt{NULL}_3 =$ "UDM"
FOO	*1*	*2*	$\texttt{NULL}_3 =$ "FOO" $\wedge\ \texttt{NULL}_1 =$ "John"

Question 4. Write an analogous query Q_2 that returns the triples of a class and two sessions $s_1 < s_2$ of the class that take place in the same room (but may have different teachers).

Answer. *Replace* $\mathbf{room \mapsto r1}$ *by* $\mathbf{teacher \mapsto t1}$ *in* Q_1.

Question 5. Evaluate Q_2 on the instance to obtain a c-table R_2.

Answer. *The result is:*

class	***s1***	***s2***	
UDM	*1*	*2*	$\texttt{NULL}_3 =$ "UDM" $\wedge\ \texttt{NULL}_3 = \texttt{NULL}_4$
FOO	*1*	*2*	$\texttt{NULL}_3 =$ "FOO" $\wedge\ \texttt{NULL}_3 = \texttt{NULL}_4$

Question 6. Compute a c-table representation of the union R of R_1 and R_2.

Answer. *The result (after Boolean simplifications) is:*

class	*s1*	*s2*	
UDM	*1*	*2*	$\texttt{NULL}_3 = \text{"UDM"}$
FOO	*1*	*2*	$\texttt{NULL}_3 = \text{"FOO"} \land (\texttt{NULL}_1 = \text{"John"} \lor \texttt{NULL}_3 = \texttt{NULL}_4)$

Question 7. How many rows may the possible worlds of R have?

Answer. *The possible worlds may be empty (if $\texttt{NULL}_3$ is neither "FOO" nor "UDM"), or may contain a single tuple (for instance if $\texttt{NULL}_3$ is "UDM"). The possible worlds cannot contain two different tuples, because we cannot have both $\texttt{NULL}_3 = \text{"UDM"}$ and $\texttt{NULL}_3 = \text{"FOO"}$.*

3 Exercise 3: Boolean c-tables

Consider the following instance:

Classes

session	date	teacher	room
2	Nov 28	Antoine	C017
3	Dec 5	Antoine	C47
4	Dec 12	Silviu	C47
5	Jan 9	Silviu	C47
6	Jan 16	Silviu	C47

Consider the following uncertain Boolean events:

- x_1: Room C47 collapses. All UDM classes in room C47 must be canceled.
- x_2: D&K students accept to return from vacation. If this does *not* happen, all UDM classes in January are cancelled.
- x_3: Silviu wins the lottery and escapes to the Bahamas. All of Silviu's classes must be canceled.

Question 1. Annotate the rows of the instance to make a Boolean c-table that describes the correct outcome depending on the value of the events.

Answer.

Classes

session	*date*	*teacher*	*room*	
2	*Nov 28*	*Antoine*	*C017*	
3	*Dec 5*	*Antoine*	*C47*	$\neg x_1$
4	*Dec 12*	*Silviu*	*C47*	$\neg x_1 \land \neg x_3$
5	*Jan 9*	*Silviu*	*C47*	$\neg x_1 \land x_2 \land \neg x_3$
6	*Jan 16*	*Silviu*	*C47*	$\neg x_1 \land x_2 \land \neg x_3$

Question 2. How many possible worlds does the table have?

Answer. *There are* four *possible worlds.*

Optional justification: *Indeed, we observe that sessions 5 and 6 have the same condition, tuple 2 is always present, and the following implications hold:*

- *the condition of tuple 5 implies that of tuple 4*
- *the condition of tuple 4 implies that of tuple 3*
- *the condition of tuple 3 implies that of tuple 2*

Hence, for any possible world, its set of tuples is a prefix of the initial set, so it is uniquely defined by the last tuple of the world (which may be either 2, 3, 4, or 6).

Now, we observe that these four possible worlds can actually be obtained; for instance, we can use the following valuations:

- $\{x_1 \mapsto 1, x_2 \mapsto 0, x_3 \mapsto 0\}$;
- $\{x_1 \mapsto 0, x_2 \mapsto 0, x_3 \mapsto 1\}$;
- $\{x_1 \mapsto 0, x_2 \mapsto 0, x_3 \mapsto 0\}$;
- $\{x_1 \mapsto 0, x_2 \mapsto 1, x_3 \mapsto 0\}$;

Question 3. Using only two Boolean variables x and y, create a different Boolean c-table on the same rows that describes the same set of possible outcomes.

Answer. *We number each possible world in binary. For brevity, we only write the* ***session*** *column when writing out possible worlds:*

00

session	...
2	...

01

session	...
2	...
3	...

10

session	...
2	...
3	...
4	...

11

session	...
2	...
3	...
4	...
5	...
6	...

We now write, for each tuple (again denoted by its ***session*** *value), the possible worlds where it occurs:*

2	*00*	*01*	*10*	*11*
3		*01*	*10*	*11*
4			*10*	*11*
5				*11*
6				*11*

We now introduce our two variables, x and y, which we will use to code the choice of possible worlds. Intuitively, x codes the choice of the most significant bit of the possible world number (written in binary), and y codes the choice of the least significant bit. This allows us to write the condition of each tuple (again denoted by its ***session*** *value):*

2	$\neg x \wedge \neg y$	$\vee$	$\neg x \wedge y$	$\vee$	$x \wedge \neg y$	$\vee$	$x \wedge y$
3	$\bot$	$\vee$	$\neg x \wedge y$	$\vee$	$x \wedge \neg y$	$\vee$	$x \wedge y$
4	$\bot$	$\vee$	$\bot$	$\vee$	$x \wedge \neg y$	$\vee$	$x \wedge y$
5	$\bot$	$\vee$	$\bot$	$\vee$	$\bot$	$\vee$	$x \wedge y$
6	$\bot$	$\vee$	$\bot$	$\vee$	$\bot$	$\vee$	$x \wedge y$

We can simplify these conditions, which yields:

2	$\top$
3	$x \vee y$
4	x
5	$x \wedge y$
6	$x \wedge y$

We annotate the initial table with these conditions, yielding our final answer (the $\top$, which indicates the true annotation, can be omitted):

Classes

session	***date***	***teacher***	***room***	
2	*Nov 28*	*Antoine*	*C017*	
3	*Dec 5*	*Antoine*	*C47*	$x \vee y$
4	*Dec 12*	*Silviu*	*C47*	x
5	*Jan 9*	*Silviu*	*C47*	$x \wedge y$
6	*Jan 16*	*Silviu*	*C47*	$x \wedge y$