

Exercise sheet for Session 2

Uncertain data management

Antoine Amarilli

November 28th, 2016

1 Exercise 1.

Consider a database schema that consists of the following tables:

- $\text{Class}(\mathbf{class}, \mathbf{date}, \mathbf{teacher}, \mathbf{room})$, indicating the planned classes
- $\text{Sick}(\mathbf{teacher}, \mathbf{date})$, indicating the dates at which a teacher is sick
- $\text{Unavail}(\mathbf{teacher}, \mathbf{date})$, indicating when a teacher is more generally unavailable
- $\text{Closed}(\mathbf{date})$, indicating dates at which the entire school is closed
- $\text{Canceled}(\mathbf{class}, \mathbf{date})$, the occurrences of classes that have to be canceled.

Remember that a *tuple-generating dependency* is a rule of the form:

$$\forall \mathbf{x} \phi(\mathbf{x}) \Rightarrow \exists \mathbf{y} \psi(\mathbf{x}, \mathbf{y})$$

where ϕ and ψ are conjunctions of atoms. An *inclusion dependency* is a tuple-generating dependency where ϕ and ψ consist of a single atom without repeated variables.

Question 1. Write tuple-generating dependencies to express the following:

- When the school is closed, all planned classes on that day are canceled
- When a teacher is sick on a day, then they are unavailable on that day
- When a teacher is unavailable on a day, all classes that they planned to give on that day are canceled

Answer.

- $\phi_1 : \forall c d t r \text{Class}(c, d, t, r) \wedge \text{Closed}(d) \Rightarrow \text{Canceled}(c, d)$
- $\phi_2 : \forall d t \text{Sick}(t, d) \Rightarrow \text{Unavail}(t, d)$
- $\phi_3 : \forall c d t r \text{Class}(c, d, t, r) \wedge \text{Unavail}(t, d) \Rightarrow \text{Canceled}(c, d)$

Question 2. Which of these tuple-generating dependencies are inclusion dependencies?

Answer. ϕ_2 is an inclusion dependency, the others are not.

Question 3. Write a conjunctive query Q that asks which classes are canceled on November 28th. (Do not assume that the Canceled table only contains classes; only return answers that occur in the Class table.) Write it in the relational calculus, and in the relational algebra.

Answer. In the relational calculus:

$$Q(c) : \exists t r \text{ Class}(c, \text{"Nov 28"}, t, r) \wedge \text{Canceled}(c, \text{"Nov 28"})$$

In the relational algebra:

$$\Pi_{\text{class}} (\sigma_{\text{date}=\text{"Nov 28"}} (\text{Class} \bowtie \text{Canceled}))$$

Question 4. Consider the database instance that contains the following facts:

- John is sick on November 28th
- The class with **class** "UDM" is taught by Antoine on November 28th in room C017
- The class with **class** "FOO" is taught by John on November 28th in room C42
- The class with **class** "UDM" is taught by Antoine on December 5th in room C47
- The school is closed on December 5th

Construct the chase of this instance by the dependencies of Question 1.

Answer. The result of the chase is as follows, with the facts added in the chase in bold (and the facts of the original instance in non-bold):

<i>Sick</i>		<i>Unavail</i>		<i>Class</i>				<i>Closed</i>	<i>Canceled</i>	
<i>teacher</i>	<i>date</i>	<i>teacher</i>	<i>date</i>	<i>id</i>	<i>date</i>	<i>teacher</i>	<i>room</i>	<i>date</i>	<i>id</i>	<i>date</i>
John	Nov 28	John	Nov 28	UDM	Nov 28	Antoine	C017	Dec 5	FOO	Nov 28
				FOO	Nov 28	John	C42		UDM	Dec 5
				UDM	Dec 5	Antoine	C47			

Question 5. Evaluate Q on the chase. What can we deduce from this?

Answer. The matches of Q on the chase are the single tuple ("FOO"). Hence, we know that $Q(\text{"FOO"})$ is entailed by the constraints and instance.

Question 6. The chase in Question 4 was finite. Would the chase by the dependencies of Question 1 be finite for any database instance? Why, or why not?

Answer. The chase will always be finite, because the tuple-generating dependencies do not have existential quantifiers. Hence, the domain of the chase is always the same as that of the initial instance.

Question 7. Rewrite the query Q (in the relational calculus) to a union of conjunctive queries Q' such that, for any instance, Q' holds on the instance iff Q is entailed by the instance and the dependencies of Question 1.

Answer. We rewrite Q to the following disjuncts:

$$\begin{aligned} Q_1(c) &: \exists t r \text{ Class}(c, \text{“Nov 28”}, t, r) \wedge (\exists t' r' \text{ Class}(c, \text{“Nov 28”}, t', r') \wedge \text{Closed}(\text{“Nov 28”})) \\ Q_2(c) &: \exists t r \text{ Class}(c, \text{“Nov 28”}, t, r) \wedge (\exists t' r' \text{ Class}(c, \text{“Nov 28”}, t', r') \wedge \text{Unavail}(t', \text{“Nov 28”})) \\ Q_3(c) &: \exists t r \text{ Class}(c, \text{“Nov 28”}, t, r) \wedge (\exists t' r' \text{ Class}(c, \text{“Nov 28”}, t', r') \wedge \text{Sick}(t', \text{“Nov 28”})) \end{aligned}$$

These disjuncts can be equivalently simplified as follows:

$$\begin{aligned} Q_1(c) &: \exists t r \text{ Class}(c, \text{“Nov 28”}, t, r) \wedge \text{Closed}(\text{“Nov 28”}) \\ Q_2(c) &: \exists t r \text{ Class}(c, \text{“Nov 28”}, t, r) \wedge \text{Unavail}(t, \text{“Nov 28”}) \\ Q_3(c) &: \exists t r \text{ Class}(c, \text{“Nov 28”}, t, r) \wedge \text{Sick}(t, \text{“Nov 28”}) \end{aligned}$$

We obtain:

$$Q'(c) := Q(c) \vee Q_1(c) \vee Q_2(c) \vee Q_3(c)$$

2 Exercise 2.

Consider a database schema that consists of the following tables:

- Jedi(**jedi**), indicating the list of known Jedis
- Teach(**master**, **padawan**), indicating which Jedi trained which Jedi
- Light(**jedi**), indicating which Jedis are on the light side of the force
- Dark(**jedi**), indicating which Jedis are on the dark side of the force

Question 1. Write tuple-generating dependencies Σ that express the following:

- Anybody on the light side of the force is a Jedi
- Likewise for anybody on the dark side of the force
- If a master teaches a padawan, then both are Jedis
- Every Jedi was taught by some master
- Whenever some padawan is on the light side of the force and was taught by a master, then the master is on the dark side of the force.
- Conversely, when a padawan is on the dark side of the force, any master is on the light side of the force.

Answer.

- $\phi_1 : \forall j \text{ Light}(j) \Rightarrow \text{Jedi}(j)$
- $\phi_2 : \forall j \text{ Dark}(j) \Rightarrow \text{Jedi}(j)$
- $\phi_3 : \forall j j' \text{ Teach}(j, j') \Rightarrow \text{Jedi}(j) \wedge \text{Jedi}(j')$
- $\phi_4 : \forall j \text{ Jedi}(j) \Rightarrow \exists j' \text{ Teach}(j', j)$
- $\phi_5 : \forall j j' \text{ Teach}(j, j') \wedge \text{Light}(j') \Rightarrow \text{Dark}(j)$
- $\phi_6 : \forall j j' \text{ Teach}(j, j') \wedge \text{Dark}(j') \Rightarrow \text{Light}(j)$

Question 2. Which one of these dependencies are inclusion dependencies? Which ones can be rewritten to be inclusion dependencies?

Answer. ϕ_1 , ϕ_2 , and ϕ_4 are inclusion dependencies. ϕ_3 is not an inclusion dependency, but can be rewritten to two inclusion dependencies:

$$\begin{aligned}\phi_3^1 &: \forall j j' \text{ Teach}(j, j') \Rightarrow \text{Jedi}(j) \\ \phi_3^2 &: \forall j j' \text{ Teach}(j, j') \Rightarrow \text{Jedi}(j')\end{aligned}$$

Question 3. Consider the instance I where the Jedis are Obi-wan (light side) and Anakin (dark side), and the first taught the second. Is the chase of this instance by Σ finite? Why?

Is the chase accurate with respect to the Star Wars movies?

Answer. Let us show that the chase is infinite. We show by induction that, at any step of the chase, there is a fact $F = \text{Teach}(a, b)$ such that $a \neq b$ and F is the only *Teach* fact where a occurs:

- The claim is true on the initial instance, as witnessed by the following fact: *Teach*(“Obi-wan”, “Anakin”).
- Assume that the claim is true after n chase rounds, and let us show that it is still true after $n + 1$ chase rounds. Consider the witnessing fact *Teach*(a, b) after n chase rounds. If we do not create any *Teach* fact in the round where a occurs, then there is nothing to show. If we do, it must have been because we applied ϕ_4 to the fact *Jedi*(a) (which must then hold). But then, the fact that we create is *Teach*(z, a), where z is a null value that occurs only in that fact and at that position. This new fact witnesses that the claim is still true after $n + 1$ chase rounds.

Assuming now by way of contradiction that the chase terminates after a finite number n of rounds, considering the witnessing fact *Teach*(a, b) that the result of the chase must contain, as the result of the chase satisfies ϕ_3 , we know that *Jedi*(a) holds we know by the above reasoning that ϕ_4 applies to it. Hence, the result of the chase does not satisfy ϕ_4 , a contradiction.

The chase is not accurate with respect to Star Wars, because Obi-Wan’s only known master (Qui-Gon) is on the light side of the force, yet the chase asserts that Obi-Wan has a master on the dark side of the force.

Question 4. Is there an instance I whose chase by Σ is finite?

Answer. Letting I be the empty instance, the chase of I by Σ is I itself, which is finite.

Question 5. Write a conjunctive query Q that asks whether a dark Jedi trained a dark Jedi. Write it both in the relational algebra and in the relational calculus.

Answer. In the relational calculus:

$$Q() : \exists j j' \text{ Teach}(j, j') \wedge \text{Dark}(j) \wedge \text{Dark}(j')$$

In the relational algebra:

$$\Pi_{\emptyset} ((\rho_{\text{jedi} \rightarrow \text{master}}(\text{Dark}) \times \rho_{\text{jedi} \rightarrow \text{padawan}}(\text{Dark})) \bowtie \text{Teach})$$

Question 6. Is Q entailed by the instance I and tuple-generating dependencies Σ ? Why (not)?

Answer. Q is not entailed by I and Σ .

Indeed, we can show by induction that the chase is an infinite line graph of the form $\text{Teach}(\text{"Obi-Wan"}, \text{"Anakin"}), \text{Teach}(z_1, \text{"Obi-Wan"}), \text{Teach}(z_2, z_1), \text{Teach}(z_3, z_2), \text{etc.}$, and that $\text{Light}(a)$ holds iff a is "Obi-Wan" or is z_i with even i , and conversely $\text{Dark}(a)$ holds iff a is "Anakin" or is z_i with odd i . Hence, the query Q does not have a match in the chase of I by Σ . This implies that Q is not entailed by I and Σ .

Question 7. Is there an instance I' where Q does not hold, but such that I' entails Q under Σ ?

Answer. The instance I' consisting of the following facts satisfies the conditions:

- $\text{Teach}(a, b)$
- $\text{Dark}(b)$
- $\text{Light}(b)$

Question 8. Write a formula in first-order logic that asserts that a Jedi either follows the light side or the dark side, but not both. Can this be expressed as a tuple-generating dependency?

Answer. A formula to express this is:

$$\phi : \forall x (\text{Dark}(x) \vee \text{Light}(x)) \wedge \neg(\text{Dark}(x) \wedge \text{Light}(x))$$

We cannot express ϕ as a tuple-generating dependency.