CS 564 PS2

September 29, 2017

Instructions / Notes:

- Using the IPython version of this problem set is **strongly recommended**, however you can use only this PDF to do the assignment, or replicate the functionality of the IPython version by using this PDF + your own SQLite interface
- Note that the problems reference tables in the SQLite database (in the PS2.db file) however solution queries must be for any general table of the specified format, and so use of the actual database provided is *not necessary*
- See Piazza for submission instructions
- Have fun!

1 Problem 1

For Parts 1 & 2 of this problem you will need to provide a single SQL query which will check whether a certain condition holds on the **hospital** table in the provided database. You need to evaluate any requested conditions in the following way: **your query should return an empty result if and only if the condition holds on the instance.** If the condition *doesn't hold*, your query should return something non-empty, but it doesn't matter what this is.

Note our language here: the conditions that we specify cannot be proved to hold **in general** without knowing the externally-defined functional dependencies; so what we mean is, *check whether they* **are not violated** *for the provided instance.*

You may assume that there are no 'NULL' values in the tables.

1.1 Part (a)

Is $\{provider\}$ a **superkey** for relation $Hospital$?

1.2 Part (b)

Does $\{Zip\} \rightarrow \{City, State\}$ hold for relation $Hospital$?

1.3 Part (c)

A **multivalued dependency (MVD)** is defined as follows: let R be a schema i.e. a set of attributes, and consider two *sets of attributes* $X \subseteq R$ and $Y \subseteq R$. We say that a multivalued dependency (MVD), written:

$$X \twoheadrightarrow Y$$

holds on R if whenever there are two tuples t_1, t_2 such that $t_1[A] = t_2[A]$, there also exists a third tuple t_3 such that:

- $t_3[A] = t_1[A] = t_2[A]$
- $t_3[B] = t_1[B]$
- $t_3[R \setminus B] = t_2[R \setminus B]$

Note that $R \setminus B$ is all the attributes in R that are not in B, and that t_3 need not be distinct from t_1 or t_2. Note especially that an MVD holds on an entire *relation*, meaning that any two tuples (in any order) in the relation should satisfy the above conditions if the MVD holds. **See the end of the lecture 8 slides for more on MVDs!**

For this part of this problem you will need to provide a single SQL query which will check whether a MVD holds for relation *courses*.

Write your query to check if the MVD $\{course\} \twoheadrightarrow \{book\}$ holds for a relation *courses*.

2 Problem 2

Consider a relation $S(A, B, C, D, E, F)$ with the following functional dependencies:

- $\{A\} \rightarrow \{D\}$
- $\{A\} \rightarrow \{E\}$
- $\{D\} \rightarrow \{C\}$
- $\{D\} \rightarrow \{F\}$

In each part of this problem, we will examine different properties the provided schema (i.e., the provided relation with the above functional dependencies).

To answer **yes**, provide python code that assigns the variable

```
answer
```

```
answer = True
explanation = "Lise is correct because all keys are superkeys."
```

To answer **no**, provide python code that assigns the variable

```
answer
```

```
answer = False
explanation = "D is not a superkey because its closure is {D,C,F}."
```

2.1 Part (a)

CS564 student Jeff claims that if A, B is a superkey. Is Jeff correct?

2.2 Part (b)

Jeff also claims that the decomposition ABC, CDE, EFA is lossless. Is Jeff correct?

2.3 Part (c)

CS564 Maria claims that the decomposition ABC, CDE, EFA is dependency preserving. Is Maria correct?

2.4 Part (d)

Now consider a relation $R(X, Y, Z)$ with some list of functional dependencies $f_1, f_2, \ldots, f_n$. Now suppose that K is a **key** for this relation, given these functional dependencies.

CS564 student Liam claims that if we add any new functional dependency $f_{n+1} = U \rightarrow V$ to our list of functional dependencies, then K will still be a key for R given $f_1, f_2, \ldots, f_{n+1}$. Is Liam correct? If yes, explain why. If no, provide a counter-example in your explanation.

3 Bonus Problem

Prove the *transitivity rule for MVDs*: If $A \twoheadrightarrow B$ and $B \twoheadrightarrow C \implies A \twoheadrightarrow C \setminus B$, using only the basic definition of an MVD; and where A, B, C are *sets of* attributes such that $A \cup B \cup C \subseteq R$, where R is the full set of attributes.