Databases
TDA357/DIT621

Exercise 3

Functional Dependencies, Multivalued Dependencies and Normal Forms

1. Consider this (symbolic) relational schema: $R(a, b, c, d, e)$.

   The following functional dependencies apply to the relation:

   $a \to b$
   $b, c \to d$
   $d, e \to a$

   (a) List 3 different minimal keys of the relation. Your solution should be three sets of attributes (the keys).

   (b) Calculate the following transitive closures. Your solution should be three sets of attributes (the transitive closures), each in alphabetical order.

   $\{b, c\}^+$
   $\{a, c\}^+$
   $\{b, d, e\}^+$

   (c) Assume all the keys you identified have appropriate unique constraints in R, show a table (the contents of R) that does not respect the functional dependency $b, c \to d$. Use integer values for all the columns.

   In other words: Your solution should be a five column table (for columns a, b, c, d, and e) with at least two rows, that does not satisfy the functional dependency, but does have unique values for the three keys you found.

   (d) Decompose R into BCNF, list all the intermediate steps. Determine all keys for the resulting schema.

2. Consider the following domain:

   Each flight is identified by a flight number and a departure time. Each flight has a set of passenger IDs and a set of airport codes it lands at. Furthermore, each passenger has a set of in-flight movies they have purchased and can use on any flight.

   For the relation R(flightNo, departure, airport, passenger, movie), identify all MVDs you can find and normalize R to 4NF.

3. (a) Consider the (symbolic) table:

a	b	c	d
0	0	0	0
0	0	0	1
0	1	1	0
1	0	0	0
2	0	2	2

Identify two non-trivial functional dependencies that hold on this data and normalize it to BCNF. Provide the resulting schema, and the data in each relation of the schema as a table like the one above.

(b) Briefly explain the concept of lossless join using your result from a) as an example.

(c) Some random person claims that it's impossible to replace the '?' below so that $a \twoheadrightarrow b$ holds but neither $a \rightarrow b$ not $a \rightarrow c$ holds on the table. Either prove the random person wrong (by constructing a table) or write a short and compelling argument for why they are correct.

a	b	c
0	?	?
0	?	?
0	?	?

4. (Adapted from an exam.)

The questions below all relate to this this table with four columns (a, b, c, d) and five rows:

a	b	c	d
a0	b0	c0	d0
a0	b0	c0	d1
a1	b1	c1	d1
a2	b1	c1	d0
a0	b1	c0	d1

The values (a0, b0, ...) are symbolic, the only important thing is that a0 differs from a1 etc.

(a) Which of the following FDs hold on the table data? Your solution should be one or more of the FDs above, no motivation is required.

$a\ b \rightarrow c$
$a\ b \rightarrow d$
$c\ b \rightarrow a$
$c\ b \rightarrow c$

(b) The FD $c\ d \rightarrow a$ holds on this table. Explain why it is also a BCNF-violation.

(c) Perform one BCNF decomposition using $c\ d \rightarrow a$, and then decompose the table data into the data of both resulting tables.

Your solution should be two tables with column names and table contents (rows). You do not need to mark primary keys.

Hint 1: The natural join of the two tables should be the original tables.
Hint 2: In a correct solution, the tables have different number of rows.

(d) Find one more BCNF-violation that still holds in the tables resulting from part (c). You do not need to perform normalization or motivate your answer, just find a violation.

Your solution should be a single functional dependency.

5. Consider the following relation and functional dependencies:

$R(a, b, c, d, e)$
$a \to b$
$c \to d$
$e \to a$

Give a real-world example of attributes a, b, c, d, e that would have exactly these dependencies and none else (except of course derived ones). Add any important assumptions you make about the domain.

An attempted example would have the same format as:
Lectures (courseCode, courseTitle, date, room, teacher).

This would reasonably satisfy $a \to b$ (because course code determines the title), and $c \to d$ (because we assume that the course has at most one lecture each date). But it would not satisfy $e \to a$ because a teacher can have several courses. It might also have the unwanted dependency $a, c \to e$ if we assume that each lecture has one teacher.