T-202-GAG1: Project 5 – Solution

# Introduction

This is a solution to Project 5 in T-202-GAG1 Databases. Two SQL scripts to create and fill the tables resulting from the normalization process accompany this solution, as well as two SQL scripts to check for FDs and MVDs.

# Checking for FDs and MVDs

It is possible to write relatively simple queries to check for the possibility of functional dependencies and multi-valued dependencies. It should be noted however, that although in this project it is assumed that dependencies holding in the data hold in real life, this is not the case in general, as potential dependencies discovered in the data must be checked against the database requirements.

## Checking for FDs

A query such as the following can be used to check for a functional dependency. This particular query checks for the FD A 🡪 B in R1 (which is known to hold, as A is the key).

SELECT IF (COUNT(\*)=0, 'R1: A --> B', 'R1: A -/> B') AS FD  
FROM (SELECT A  
 FROM R1  
 GROUP BY A  
 HAVING COUNT(DISTINCT B) > 1  
) X;

A program was created to generate queries to examine all possible FDs for the five relations.

## Checking for MVDs

A query such as the following can be used to check for a multi-valued dependency. This particular query checks for the MVDs A 🡪🡪 B and A 🡪🡪 C in the table R2.

SELECT IF (COUNT(\*) = 0, 'MAYBE MVD', 'NO MVD') AS MVD   
FROM (SELECT A  
 FROM R2  
 GROUP BY A  
 HAVING (COUNT(DISTINCT B) \* COUNT(DISTINCT C)   
 <> COUNT(\*))  
) X;

Note that such queries should only be run against relations in either 3NF or BCND that have keys of three or more columns. Also note that similar queries should be run to look for all possible MVDs, in this case by running three different queries using A, B, and C as the GROUP BY column.

# Normalization of R1

## Analysis

Primary key: A

Determined FDs: B 🡪 C, D 🡪 E

Minimal cover: A 🡪 BD, B 🡪 C, D 🡪 E

Other keys: None

Normal form: 2NF

Decomposition: R1ABD, R1BC, R1DE

## Table: R1BC

Columns: B, C

Key: B

FDs: B 🡪 C

Normal forms:

* Since all FDs are key FDs, the table is in BCNF
* Since the key has a single column, the table is in 4NF
* Since all keys are single column, the table is in 5NF

## Table: R1DE

Columns: D, E

Key: D

FDs: D 🡪 E

Normal forms:

* Since all FDs are key FDs, the table is in BCNF
* Since the key has a single column, the table is in 4NF
* Since all keys are single column, the table is in 5NF

## Table: R1ABD

Columns: A, B, D

Key: A

FDs: A 🡪 BD

Normal forms:

* Since all FDs are key FDs, the table is in BCNF
* Since the key has a single column, the table is in 4NF
* Since all keys are single column, the table is in 5NF

# Normalization of R2

## Analysis

Primary key: ABC

Determined FDs: A 🡪 D

Minimal cover: ABC 🡪 E, A 🡪 D

Other keys: None

Normal form: 1NF

Decomposition: R2ABCE, R2AD

## Table: R2AD

Columns: A, D

Key: A

FDs: A 🡪 D

Normal forms:

* Since all FDs are key FDs, the table is in BCNF
* Since the key has a single column, the table is in 4NF
* Since all keys are single column, the table is in 5NF

## Table: R2ABCE

Columns: A, B, C, E

Key: ABC

Other FDs: None

Normal forms:

* Since all FDs are key FDs, the table is in BCNF
* Since the key has multiple columns, MVDs were sought using the queries above, but none were found and t˙he table should therefore be in 4NF
* The table may or may not be in 5NF

# Normalization of R3

## Analysis

Primary key: AB

Determined FDs: D 🡪 A, D 🡪 E

Minimal cover: AB 🡪 CD, D 🡪 A, D 🡪 E

Other keys: BD

Normal form: 2NF

Decomposition: R3ABCD, R3DE

## Table: R3DE

Columns: D, E

Key: D

FDs: D 🡪 E

Normal forms:

* Since all FDs are key FDs, the table is in BCNF
* Since the key has a single column, the table is in 4NF
* Since all keys are single column, the table is in 5NF

Note: This table should NOT contain the A column, as that would lead to redundancy in the collection of tables.

## Table: R3ABCD

Columns: A, B, C, D

Keys: AB, BD

Other FDs: D 🡪 A

Normal forms:

* Since one FD has 3NF, the overall table is in 3NF
* Since the key only has two columns, MVDs were not sought and the table should be in 4NF
* The table may or may not be in 5NF

# Normalization of R4

## Analysis

Primary key: ABD

Determined FDs: A 🡪 C, A 🡪 E, E 🡪 C

Minimal cover: ABD 🡪 ABD, A 🡪 E, E 🡪 C

Other keys: None

Normal form: 1NF

Decomposition: R4ABD, R4EC, R4AE

Note: This is the decomposition due to FDs. As MVDs were identified (see below) further decomposition was needed.

## Table: R4EC

Columns: E, C

Key: E

FDs: E 🡪 C

Normal forms:

* Since all FDs are key FDs, the table is in BCNF
* Since the key has a single column, the table is in 4NF
* Since all keys are single column, the table is in 5NF

## Table: R4AE

Columns: A, E

Key: A

FDs: A 🡪 E

Normal forms:

* Since all FDs are key FDs, the table is in BCNF
* Since the key has a single column, the table is in 4NF
* Since all keys are single column, the table is in 5NF

## Table: R4ABD

Columns: A, B, D

Key: ABD

Other FDs: None

Normal forms:

* Since there are no non-key attributes, the table is in BCNF
* Since the key has three columns, MVDs were sought resulting in the following MVDs: A 🡪🡪 B, A 🡪🡪 D; the table is therefore NOT in 4NF and must be decomposed
* The table may or may not be in 5NF

Decomposition: R4AB, R4AD

## Table: R4AB

Columns: A, B

Key: AB

FDs: None

Normal forms:

* Since there are no non-key attributes, the table is in BCNF
* Since the table is too small for MVDs, the table is in 4NF
* Since the table is too small for JDs, the table is in 5NF

## Table: R4AD

Columns: A, D

Key: AD

FDs: None

Normal forms:

* Since there are no non-key attributes, the table is in BCNF
* Since the table is too small for MVDs, the table is in 4NF
* Since the table is too small for JDs, the table is in 5NF

# Normalization of R5

## Analysis

Primary key: ABC

Determined FDs: E 🡪 A, D 🡪 B

Minimal cover: ABC🡪 DE, E 🡪 A, D 🡪 B

Other keys: ACD, BCE, CDE

Normal form: 3NF

Decomposition: No decomposition is needed