Functional Dependencies and Normalisation

Introduction

The goal of this activity is for you to practice the technique of using functional dependencies to normalize a relational database schema and for you to take an information source and normalise from UNF to 3NF.

We will start with the universal schema, i.e., a single relation database schema, where all the attributes of interest belong to the single relation in the database schema. However, we could start from a partially normalized schema as well. In the next task given the information source you are to undergo the normalisation process to create new relations that progressively become strong to help avoid insert, update and delete anomalies.

This technique is often applied in real-world software development by a database designer when a relational database schema already exists but needs expansion, or else needs to be revised if it has been otherwise modified.

The outcome of this step is a normalized relational database schema.

Task 1

Consider the universal relation:

$$R = \{A, B, C, D, E, F, G, H, I, J\}$$

and the set of functional dependencies:

$$\begin{aligned} F = \{ \quad & AB \rightarrow C, \\ & A \rightarrow DE, \\ & B \rightarrow F, \\ & F \rightarrow GH, \\ & D \rightarrow IJ \\ & \} \end{aligned}$$

1. What is the key for *R*? Explain your answer.

   AB

   A minimal set of attributes whose closure includes all the attributes in R is a key. Since the closure of AB^+ is equal to R, one key of R is AB (and in this case, it is the only key).

2. Decompose *R* into 2NF, then 3NF relations. Explain your decisions.

a. First, we identify partial dependencies that violate 2NF. These are attributes that are functional dependent on either part of the key, A or B alone. We can calculate the closures of A^+ and B^+ to determine partially dependent attributes.

$A^+ = \{ A, D, E, I, J \}$
$B^+ = \{ B, F, G, H \}$

Hence

$A \rightarrow \{ D, E, I, J \}$
$B \rightarrow \{ F, G, H \}$

b. To normalise into 2NF, we remove the attributes that are functionally dependent on part of the key, A or B from R and place them in separate relations R1 and R2, along with the part of the key they depend on A or B, which are copied into each of these relations but also remains in the original relation, which we call R3 below.

R1 = { [A], D, E, I, J }
R2 = { [B], F, G, H }
R3 = { [A,B], C }

The keys are identified by square brackets.

c. Next, we look for transitive dependencies in R1, R2 and R3. The relations R1 has the transitive dependency:

$A \rightarrow D \rightarrow I,J$

So, we remove the transitive dependent attributes I, J from R1 into a relation R11 and copy the attribute D they are dependent on in R11. The remaining attributes are kept in a relation R12. Hence R1 is decomposed into R11 and R12 as follows:

R11 = { [D], I, J }
R12 = { [A], D, E }

The relation R2 is similarly decomposed into R21 and R22 based on the transitive dependency:

$B \rightarrow F \rightarrow G,H$

As follows:

R21 = { [F], G, H }
R22 = { [B], F }

d. The final set of relations in 3NF is:

R11, R12, R21, R22 and R3

Task 2

Consider the following information source:

Student Details

Registration No	545878
Name	Peter Smith
Address	45 Some made up place
Personal Tutor ID	556
Personal Tutor Name	Dr Nice Person
Program ID	COMPSCI_AI
Program Name	Computer Science with AI

Courses

Module Code	Module Title	Grade
COMP101	Team Project	86
COMP163	Python Programming	62
COMP132	Data Science	66
COMP152	Operating Systems	72

GPA 71.5

1. Show the information source as a table with the associated columns and data in UNF. Include a short description of the process.

UNF – Identify all fields and sample data. Identify key and underline it

Registration No	545878			
Name	Peter Smith			
Address	45 Some made up place			
Personal Tutor ID	556			
Personal Tutor Name	Dr Nice Person			
Program ID	COMPSCI_AI			
Program Name	Computer Science with AI			
Module Code	COMP101	COMP163	COMP132	COMP152
Module Title	Team Project	Python Programming	Data Science	Operating Systems
Grade	86	62	66	72

2. Convert from UNF to 1NF and show the resulting relations. Include a short description of the process.

1NF – Remove Repeating Attributes, identify new compound key

Registration No	545878
Name	Peter Smith
Address	45 Some made up place
Personal Tutor ID	556
Personal Tutor Name	Dr Nice Person
Program ID	COMPSCI_AI
Program Name	Computer Science with AI

Registration No	545878
Module Code	COMP101
Module Title	Team Project
Grade	86

3. Convert from 1NF to 2NF and show the resulting relations. Include a short description of the process.

2NF – Check Partial Dependencies, Remove any to new table with copy of determinant

Registration No	545878
Name	Peter Smith
Address	45 Some made up place
Personal Tutor ID	556
Personal Tutor Name	Dr Nice Person
Program ID	COMPSCI_AI
Program Name	Computer Science with AI

Registration No	545878
Module Code	COMP101
Grade	86

Module Code	COMP101
Module Title	Team Project

4. Convert from 2NF to 3NF and show the resulting relations. Include a short description of the process.

3NF – Check Transitive Dependencies, Remove any to new table with copy of determinant

Registration No	545878
Name	Peter Smith
Address	45 Some made up place
Personal Tutor ID	556
Program ID	COMPSCI_AI

Registration No	545878
Module Code	COMP101
Grade	86

Module Code	COMP101
Module Title	Team Project

Personal Tutor ID	556
Personal Tutor Name	Dr Nice Person

Program ID	COMPSCI_AI
Program Name	Computer Science with AI