# INFS1200/7900 Tutorial 4.1: FDs and Normal Forms – Solutions

## Introduction

**Purpose:** The purpose of this tutorial is to provide you with experience in finding both the candidate key(s) and highest normal form of a relation. To do this we will first examine the importance of identifying functional dependencies by considering various update anomalies.

**Learning Outcomes:** By the end of this tutorial you will be able to:

- Correctly identify update anomalies in a relational schema and their implications
- Identify candidate keys in a relation from a set of functional dependencies
- Justify the highest normal form of a relation

## Section A: Anomalies and Functional Dependencies

**A.1** Based on the following data, provide an example and an explanation of an insertion, deletion and modification anomaly.

| A | B | C | D | E |
|---|---|---|---|---|
| 2 | 2 | 1 | 5 | 6 |
| 2 | 3 | 1 | 5 | 4 |
| 3 | 4 | 5 | 3 | 2 |
| 3 | 5 | 5 | 1 | 3 |

Functional Dependencies:

$\{A\} \rightarrow \{C\}$

$\{B\} \rightarrow \{D, E\}$

> **Insertion Anomaly:**
>
> Insert anything into B, D and E without also inserting a value into A. Where B, D and E are values that don't already exist in the data.
>
> **Deletion Anomaly:**
>
> Deleting the tuple with B = 4 as we lose information for the functional
>
> dependency (B → {D, E}), specifically (4 → {3, 2})
>
> **Modification Anomaly:**
>
> Updating {2, 2, 1, 5, 6} to {2, 4, 1, 5, 6} creates the inconsistency of (B → {D, E}), creates the inconsistency (4 → {3, 2}) and (4 → {5, 6})

**A.2** Based on the following functional dependencies, fill in the table.

A → B

B → C

{C, D} → E

| A | B | C | D | E |
|---|---|---|---|---|
| 1 | 2 | 1 | 6 | 2 |
| 1 | 2 | 1 | 4 | 3 |
| 2 | 4 | 2 | 7 | 4 |
| 3 | 2 | 1 | 4 | 3 |

**A.3** Based on the following data, identify which of the options are *potential* functional dependencies.

| A | B | C | D | E |
|---|---|---|---|---|
| 1 | X | 1 | M | 1 |
| 2 | Y | 1 | M | 1 |
| 3 | Y | 4 | N | 3 |
| 4 | W | 2 | L | 5 |
| 5 | W | 2 | M | 1 |
| 6 | T | 5 | O | 2 |

✔ A → B

   B → A

✔ A → C

   B → C

   C → D

   C → E

✔ D → E

✔ {A, B} → C

   {B, C} → E

✔ {B, C, D} → E

**SEE NEXT PAGE FOR SECTION B**

Based on the relational schema and functional dependencies provided for the following questions, find all candidate keys.

## B.1 R [A, B, C, D, E, F]

$\{A, E\} \rightarrow \{D\}$

$\{B, C\} \rightarrow \{A\}$

$\{B\} \rightarrow \{F\}$

$\{F\} \rightarrow \{E\}$

Candidate Key(s):

{B, C}

## B.2 R [A, B, C, D, E, F]

$\{A\} \rightarrow \{B, C\}$

$\{C, D\} \rightarrow \{E\}$

$\{A, C\} \rightarrow \{E\}$

$\{B\} \rightarrow \{D\}$

$\{E\} \rightarrow \{A, B\}$

Candidate Key(s):

{B, C, F}, {C, D, F}, {A, F}, {E, F}

## B.3 R [A, B, C, D]

$\{A, B\} \rightarrow \{C, D\}$

$\{C\} \rightarrow \{A, B, D\}$

$\{D\} \rightarrow \{C\}$

Candidate Key(s):

{C}, {A, B}, {D}

## B.4 R [A, B, C, D, E, F, G, H, I, J]

$\{A, B\} \rightarrow \{C\}$

$\{A\} \rightarrow \{D, E\}$

$\{B\} \rightarrow \{F\}$

$\{F\} \rightarrow \{G, H\}$

$\{D\} \rightarrow \{I, J\}$

Candidate Key(s):

{A, B}

Based on the relational schema and functional dependencies identify the highest normal form for the following question and provide a brief justification why this holds.

## C.1 R [A, B, C, D, E, F]

{A, E} → {D}

{B, C} → {A}

{B} → {F}

{F} → {E}

Highest Normal Form:

Key is {B, C}. Hence highest normal form is 1NF because of partial dependency {B} → {F}

## C.2 R [A, B, C, D, E]

{A} → {B, C, D, E}

{B} → {A, C, D, E}

Highest Normal Form:

Keys are {A} and {B}. Highest normal form is BCNF because the LHS of all FD's are super keys.

## C.3 R [A, B, C, D, E, F, G]

{A} → {B, C, D}

{D} → {A}

{C} → {F, G}

Highest Normal Form:

Keys are {A, E} and {D, E}. Highest normal form is 1NF. {A} → {B, C, D} decomposes into {A} → {B}, {A} → {C}, and {A} → {D}. There is a partial dependency in {A} → {B} and {A} → {C} because A is a subset of the candidate key {A, E} and C and B are both non-prime attributes. Note: {D} → {A} is not a partial dependency because A is a prime attribute.

**C.4 R [A, B, C, D, E]**

{A, B} → {C, E}

{D} → {A}

{A} → {D}
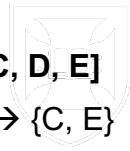
Highest Normal Form:

Keys are {A, B} and {D, B}. Highest normal form is 3NF because there are no partial or transitive dependencies. Because the LHS of the FD {A} → {D} is not a super key, the relation violates BCNF. Note: {A} → {D} is not a partial dependency because D is a prime attribute. Same reasoning applies to {D} → {A}