CS145
Bonus Activity
3^{rd} and 4^{th} Normal Forms

This activity will help you get familiar with some of the normal forms which are not covered in the lecture but you might come across them as you dive more into this field.

3rd Normal Form (3NF)

As we have seen in the lecture(Lecture 5-7 slides 104-107) on BCNF that it is not possible, in some cases, to decompose a relation into BCNF relations that have both the lossless-join and dependency-preservation properties. The solution to this is to relax our BCNF requirement slightly, in order to allow the occasional relation schema that can not be decomposed into BCNF relations without our losing the ability to check the FD’s. This relaxed condition is called "Third Normal Form".

A relation R is in 3rd Normal Form if:

If $\exists$ a nontrivial dependency A = {A1, A2, ..., An} -> B in R,
then A is a superkey for R OR B is part of some key.

You can read more about this in the section **3.5** of the textbook **Database Systems: The Complete Book by Garcia-Molina, Ullman, and Widom**

Exercise 1

Given the set of attributes below come up with the smallest number of FDs such that the relation **R(A, B, C)** is in 3NF but **not BCNF**.

3NF Decomposition Algorithm

3NFDecomp(R):

<u>let</u> K = [all attributes that are part of some key]

Find A s.t.: $\mathbf{A^+ \setminus (A \cup K) \neq \varnothing}$ and $\mathbf{A^+ \neq}$ **[all attributes]**

<u>if</u> (not found) <u>then</u> Return R

let $\mathbf{B = A^+ \setminus (A \cup K)}$, $\mathbf{C = B^C \setminus A}$

decompose R into $\mathbf{R_1(A \cup B)}$ and $\mathbf{R_2(A \cup C)}$

Return 3NFDecomp(R_1), 3NFDecomp(R_2)

Exercise 2

Give the decomposition of the relation ***R*** with attribute ***A*** and FDs ***F*** below in 3rd normal form:

A = [A, B, C, D, E]

FDs:

F1: B, C -> D

F2: D -> E

F3: E -> C

F4: E -> A

4th Normal Form (4NF)

There are occasional situations where we design a relation schema and find it is in BCNF, yet the relation has a kind of redundancy that is not related to FD's (see section 3.6.1 of the textbook). These redundancies are caused by the MVDs of the relation and can be eliminated if we also use these MVDs for decomposition. This is where we use a new normal form, called "Fourth Normal Form". In this normal form, all nontrivial MVD's are eliminated, as are all FD's that violate BCNF. As a result, the decomposed relations have neither the redundancy from FD's nor the redundancy from MVD's. The fourth normal form condition is essentially the BCNF condition, but applied to MVD's instead of FD's. Formally:

A relation R is in 4th Normal Form if:

If $\exists$ a nontrivial MVD, A = {A1, A2, ..., An}$\twoheadrightarrow$B in R,
then A is a superkey for R.

You can read more about this normal form in the section **3.6.4** of the textbook.

Exercise 3

Given the relation R with attributes A, MVDs M and FDs below, check if R is in 4th Normal Form or not:

A = (A, B, C, D, E)

FDs:

F1: A -> B
F2: B -> C

MVDs:

M1: A, B ->> C
M2: D ->> E

4NF Decomposition Algorithm

4NFDecomp(R):

Find a non trivial MVD A = {A1, A2, ..., An}$\twoheadrightarrow$B in R, s.t. A is not superkey for R

<u>if</u> (not found) <u>then</u> Return R

let **Y = (A $\cup$ B)**, **Z = A $\cup$ ([all attributes] \ (A $\cup$ B))**
decompose R into **R_1(A $\cup$ Y)** and **R_2(A $\cup$ Z)**

Return 4NFDecomp(R_1), 4NFDecomp(R_2)

Exercise 4

For the attributes A and MVDs M as defined in exercise 3, give the decomposition of the relation R into 4th normal form.