CS 122A: Introduction to Data Management – Spring 2019

Homework 3: Relational Database Design Theory (100 points)

Due Date: **Friday, Apr 26 (5:00 PM)**

Submission

All HW assignments should be turned in with a filename that contains both your student ID and your name (e.g., 12345678_John_Doe.pdf) and must be submitted online, as a PDF file, through the associated (HW3 in this case) HW3 in **Gradescope**. See the table below for HW 3 submission opportunities. Note that after the last deadline, Saturday Apr 27th, no further HW 3 submissions will be accepted at all. That is, we will not accept assignments after that time since we will be publishing the solution at that time -- this is important so that students who want to can see the solution before the first Midterm exam. Please turn in all of your work on time! If possible, save your one dropped assignment for the end of the term when you are most likely to want/need it.

Date / Time	**Grade Implications**
Friday, Apr 26 (5:00 PM)	Full credit will be available
Saturday, Apr 27 (5:00PM)	20 points will be deducted

Submission note: please use the provided template to answer each question. Normally, your solution for each question should not EXCEED 1 page for grading purposes. Since there are four questions, four pages will be enough. However, if you think you can't answer one question in a single page, you can insert one additional blank page and use it to answer a question in two pages. In any case, please always start a new question at the beginning of a page. (The template ensures this if you don't exceed one page per question.) Also, In gradescope, mark the answer for each question so grading would be easier.

1) [25 Points] By looking at the PHLogger table:
   a) List all non-trivial functional dependencies.

   b) What is the highest normal form the PHLogger table is in currently?

   c) The external consulting experts at DBInstructor, Inc., have noticed that city and state of an address can be inferred by its postal code (zip code). What new functional dependencies would be introduced by codifying this rule?

   d) What is the highest normal form the PHLogger table is in after adding the new functional dependencies?

   e) Decompose the PHLogger table into multiple tables to the highest normal form possible.

   f) After decomposition, what is the highest normal form design that you could produce which is lossless and dependency preserving[3NF/BCNF]? Explain.

2) [25 points] Consider the following relation:

G	H	M
10	h1	m1
10	h2	m2
11	h4	m1
12	h3	m4
13	h1	m1
14	h3	m4

a) Given the current state of the database, for each one of the following functional dependencies answer a) Does this functional dependency hold in the above relation instance [Yes/No]? b) If your answer to previous question was no, explain why by listing a tuple that causes a violation.

i) $G \rightarrow H$

ii) $H \rightarrow M$

iii) $M \rightarrow H$

iv) $H \rightarrow G$

v) $M \rightarrow G$

b) List all potential candidate keys (if there are any) for the above relation.

3) [25 points] Considering the relation R(A,B,C,D,E) and the following functional dependencies, answer the questions.

FD1: AB → C
FD2: CD → E
FD3: DE → B

A. List all the candidate keys.

B. What is the highest normal form that R satisfies and why?

C. If R is not already at least in 3NF, then normalize R into 3NF and show the resulting relation(s) and their candidate keys. Your decomposition should be both join-lossless and dependency-preserving. If R is already in 3NF, just list the candidate keys of R.

D. Is your decomposition in BCNF as well?[Yes/No]. Explain.

4) [25 points] Considering the relation R(A,B,C,D,E) and the following functional dependencies, answer the questions.
FD1: A → BC
FD2: BC → AD
FD3: D → E

A. List all the candidate keys.

B. What is the highest normal form that R satisfies and why?

C. If R is not already at least in 3NF, then normalize R into 3NF and show the resulting relation(s) and their candidate keys. Your decomposition should be both join-lossless and dependency-preserving. If R is already in 3NF, just list the candidate keys of R.

D. Is your decomposition in BCNF as well?[Yes/No]. Explain.