Information and Database Management Systems I

(CIS 4301 UF Online)

Fall 2024

Instructor: Alexander Webber

TA: Kyuseo Park

Homework 5

Printed Name:	
UFID:	
Email Address:	

Instructions: Please provide your answers to the questions of the following pages in Word or handwritten on separate sheets of paper. Mark clearly to which question each answer belongs. Then convert or scan your work into PDF (the latter by using either a scanner or a suitable scanner app on your smartphone). Note that *only the PDF format* is allowed and that your submission must be a *single PDF file*. Finally, upload your PDF file into *Canvas* and follow the instructions there.

Note: All homework assignments are designed for a period of two, three, or even four weeks (see course deadline sheet). This means they cannot be solved in two or three hours but require a considerable amount of time and effort. Therefore, the first recommendation is to start with them as soon as they are posted. The second recommendation is to distribute the work on a homework assignment over the entire available period. The third recommendation is to submit the homework solutions *on time before the deadline*.

Pledge (Must be signed[1] according to the UF Honor Code):

On my honor, I have neither given nor received unauthorized aid in doing this assignment.

_____

Student signature

[1] Each student is obliged to print out this page, fill in the requested information in a handwritten and readable manner, make the *handwritten* signature, scan this page into PDF, and put this page as the first page of the PDF submission.

General Note

Please note that the questions in this homework assignment require for your detailed answers. This means that you have to provide a conclusive argumentation for each of your statements with sufficient detail. The only exception is the computation of the attribute closure of an attribute set. You can compute an attribute closure in your mind and directly write it down.

Question 1 (Normalization (I)) [15 points]

(a) [5 points] Consider the following table *EmployeeProjects* and determine its highest normal form. Then, normalize the table to the next higher normal form. Primary keys are underlined.

EmployeeProjects				
EmployeeID	**EmployeeName**	**ProjectID**	**ProjectName**	**ProjectDepartment**
E01	"John Doe"	P01	"Project Alpha"	"IT"
E02	"Jane Smith"	P02	"Project Beta"	"Marketing"
E01	"John Doe"	P03	"Project Gamma"	"Development"
E03	"Mike Brown"	P02	" Project Beta"	"Marketing"

(b) [5 points] Explain why the following table *BookPurchases* is not in the third normal form and normalize it.

BookPurchases			
Purchase ID	**Book Title**	**Author Name**	**Release Date**
001	"The Great Gatsby"	"F. Scott Fitzgerald"	1925
002	"To Kill a Mockingbird"	"Harper Lee"	1960
003	"The Great Gatsby"	"F. Scott Fitzgerald"	1925
004	"1984"	"George Orwell"	1949

(c) [5 points] Consider the following table *CollectionBooks* and explain if it is in the first normal form. If not, normalize the table to the first normal form.

CollectionBooks				
Book	**Author**	**Price**	**Publisher**	**Year**
"Book A"	⟨"Author X", "Author Y"⟩	20	"Publisher Z"	2020
"Book B"	⟨"Author X", "Author Z"⟩	15	"Publisher Y"	2022
"Book C"	⟨"Author Y", "Author Z"⟩	25	"Publisher Z"	2019

Question 2 (Normal Forms) [20 points]

(a) [5 points] Let $R(A, B, C, D)$ be a relation schema, and let $S = \{A \to C, B \to D, AB \to CD\}$ be a set of functional dependencies. Determine if R is in 2NF. Explain your answer. If R is not in 2NF, normalize it to 2NF.

(b) [5 points] Let $R(A, B, C, D, E)$ be a relation schema, and let $S = \{A \to B, AB \to C, C \to D, A \to E\}$ be a set of functional dependencies. Determine if R is in 2NF and 3NF. Explain your answer. If R is not in 3NF, normalize it to 3NF.

(c) [5 points] Let $R(A, B, C, D, E)$ be a relation schema, and let $F = \{AB \to CDE, CD \to ABE, E \to D\}$ be a set of functional dependencies. Determine the highest normal form of R. Check from 2NF to BCNF.

(d) [5 points] Normalize the relation schema R of the previous question to the next higher normal form.

Question 3 (Lossless Join Decomposition & Dependency Preservation) [20 points]

(a) [10 points] Let $R(A, B, C, D, E, F)$ be a relation schema, and let $S = \{A \to B, BD \to E, EC \to A, CF \to D, EF \to B, C \to E\}$ be a set of functional dependencies. Use the *Chase test* to determine if the following decomposition is lossless.

Decomposed relations: $R_1(A, B, E), R_2(B, D, E), R_3(C, E, F), R_4(C, D, F)$

(b) [5 points] Let $R(A, B, C, D, E, F)$ be a relation schema, and let $S = \{A \to B, BD \to E, EC \to A, CF \to D, EF \to B, C \to E\}$ be a set of functional dependencies. Use the *Nonadditive Join Test for Binary Decomposition* to determine if the following decomposition is lossless.

Decomposed relations: $R_1(A, C, E, F), R_2(B, C, D, F)$

(c) [5 points] Let $R(A, B, C, D, E, F)$ be a relation schema, and let $S = \{A \to B, BD \to E, EC \to A, CF \to D, EF \to B, C \to E\}$ be a set of functional dependencies. By applying the algorithm (predicate) *IsDependencyPreserving2* from the lecture, determine if the decomposition of R into the relation schemas $R_1(B, C, D, E)$ and $R_2(A, B, D, F)$ is dependency preserving.

Question 4 (Normalization (II)) [34 points]

To obtain more uniform student solutions, we introduce the following conventions:

- All attribute sets are written in juxtaposition notation.
- Attributes of attribute sets are listed in alphabetical order.
- FDs in sets of functional dependencies are listed in alphabetical order.
- FDs in sets of functional dependencies are processed consecutively from left to right.
- If attributes of an attribute set on either side of an FD are to be checked if they are extraneous, they are omitted from right to left. (Example: In the set $ABCD$, D is omitted first, C second, B third, and A fourth.).

(a) [12 points] Let $R(A,B,C,D,E)$ be a relation schema, and let $F = \{AD \to BE, B \to CD, BCE \to DE, DE \to AB\}$ be a set of functional dependencies. Decompose R into 3NF by using the *3NF synthesis algorithm*. Argue precisely, and explain all details.

(b) [4 points] What is the relationship between R and the relation schemas R_i of the 3NF decomposition in (a)? What is the relationship between the set F of functional dependencies and the sets F_i of functional dependencies of the 3NF decomposition in (a)? Provide a colloquial (that is, given in your own words by using the right technical terms) and a formal characterization of these relationships.

(c) [6 points] Is the 3NF decomposition from (a) also in BCNF? If so, why? If not, provide the reasons, and decompose the relation schemas of the 3NF decomposition that violate BCNF into BCNF. Argue precisely, and explain all details.

(d) [8 points] Let us assume the same relation schema $R(A,B,C,D,E)$ and the same set $F = \{AD \to BE, B \to CD, BCE \to DE, DE \to AB\}$ of functional dependencies as in (a). Decompose R into BCNF by using the *BCNF decomposition algorithm*. Argue precisely, and explain all details.

(e) [4 points] What is the relationship between R and the relation schemas R_i of the BCNF decomposition in (d)? What is the relationship between the set F of functional dependencies and the sets F_i of functional dependencies of the BCNF decomposition in (d)? Provide a colloquial (that is, given in your own words by using the right technical terms) and a formal characterization of these relationships.

Question 5 (Integrity Constraints) [11 points]

Consider the following tables:

- Employee (eID, name, email, dID, salary)
- Department (dID, name, managerID)
- Project (pID, name, startDate, endDate, dID)
- EmployeeProject (eID, pID)

(a) [3 points] Write an assertion check to ensure that every manager of a department is also an employee.

For the following two sub-questions, keep track of the number of employees who have worked on a project in a new table named *EmployeeProjectsCount*. The table has the two attributes *pID* and *numOfEmployees*.

(b) [3 points] Write a SQL statement to create the table *EmployeeProjectsCount*. Add all constraints.

(c) [5 points] Write a trigger that activates upon the deletion of a row from the table *EmployeeProject*. This trigger should decrement values of the attribute *numOfEmployees* in the table *EmployeeProjectsCount* for the associated project. If this action results in zero employees remaining for that project, the trigger should also update the *endDate* of the project to the current date.