

Information and Database Management Systems I

(CIS 4301)

Spring 2023

Instructor: Dr. Markus Schneider

TA: Kyuseo Park

Homework 1

| | |
|----------------|--|
| Printed Name: | |
| UFID: | |
| Email Address: | |

Instructions: Please provide your answers to the questions of the following pages on separate sheets of paper. Mark to which question each answer belongs. Then scan your work into PDF 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 *Gradescope* and follow the instructions there.

Pledge (Must be signed¹ according to the UF Honor Code):

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

Student signature

¹Each 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.

Question 1 (ER Model)

[20 points]

Design an Entity-Relationship diagram that models an Internet forum and considers the requirements listed below. That means that you have to identify suitable entity sets, relationship sets, attributes, keys of entity sets (if not specified), and so on. Further, add the cardinalities (1:1, 1:m, m:1, m:n) to the relationship sets and write down your assumptions regarding the cardinalities if there could be a doubt.

Consider the following requirements:

- A **user** has an **ID** which is unique, a **name**, a **nickname**, an **email address**, a **date of birth**, an **age**, and a **phone number**. An age can be obtained in combination with the date of birth and the system date.
- A **hometown** where users were **born** has a **city name** and a **state name**. Multiple users can be associated with a hometown.
- There are two types of users: **registered users** and **unregistered users**.
- The **SSN** of a **registered user** is stored, but an unregistered user does not have it.
- A registered user **creates** a **post**. Each post belongs to one user and must have exactly one **ID**, **title**, and **content**, along with a **date**.
- Multiple posts can be associated with a registered user.
- A registered user **leaves** a comment about a post with the **text** and the **date** of the comment.
- A **comment** is associated with one post and one user. Each post can have multiple comments, and each user can comment on multiple posts.
- An unregistered user cannot create a post or leave a comment but can only **read** them (a registered user can also read).
- Both kind of users (registered and unregistered) can **like** a post. An **ID** for liking a post is stored when a user likes a post.

NOTE: Use **ERDPlus** to draw your ER diagram. This will help you become prepared for Exam 1 where we will use ERDPlus. Create an account on the ERDPlus web page to be able to store ER diagrams so that you can continue or modify them later. Note that the graphical notation is slightly different compared to Chen's notation. ERDPlus uses the *crow's foot* notation. The textbooks will help you understand and distinguish the two different notations. But ERDPlus also supports Chen's notation by using straight edges and *labels* attached to the edges. The only problem is that the text boxes containing the labels cannot be minimized with the mouse in a manner so that they closely surround the label. Unfortunately, a text box is not transparent and has a fixed minimal size that is rather large and can superimpose other components of your ER diagram. An inconvenient way to solve this problem is to export your ER diagram into the ERDPlus specific ASCII format (with the file ending .erdplus), use a text editor to change the height and width of the label that overlaps other components of your ER diagram, delete the old file or rename the old file of the ER diagram in ERDPlus, and import the modified file into ERDPlus. For labels such as *m*, *n*, or 1, the value 15 for height and width is suitable.

Question 2 (ER Model)

[20 points]

Design an Entity-Relationship diagram that models a bank management system and considers the requirements listed below. That means that you have to identify suitable entity sets, relationship sets, attributes, keys of entity sets (if not specified), and so on. Further, add the cardinalities (1:1, 1:m, m:1, m:n) to the relationship sets and write down your assumptions regarding the cardinalities if there could be a doubt.

Consider the following requirements:

- The bank has multiple branches that are identified by a branch ID.
- A branch has a name and an address.
- A bank clerk has a name, a SSN, a salary, and a position.
- Many bank clerks work for a branch.
- A customer has a name, a date of birth, an address, and a customer ID.
- Customers can have one or multiple bank accounts.
- Bank clerks help customers open a bank account.
- Multiple customers can share the same bank account. However, a loan should be held by only one customer.
- A bank account has an account number, an account type, and a balance.
- Bank clerks also offer loans to customers.
- A loan has a loan ID, an amount, and a type.
- A bank clerk can support some dependents.
- Dependents supported by a bank clerk have a name and a relationship.

NOTE: Use ERDPlus to draw your ER diagram. This will help you become prepared for Exam 1 where we will use ERDPlus. Create an account on the ERDPlus web page to be able to store ER diagrams so that you can continue or modify them later. Note that the graphical notation is slightly different compared to Chen's notation. ERDPlus uses the *crow's foot* notation. The textbooks will help you understand and distinguish the two different notations. But ERDPlus also supports Chen's notation by using straight edges and labels attached to the edges. The only problem is that the text boxes containing the labels cannot be minimized with the mouse in a manner so that they closely surround the label. Unfortunately, a text box is not transparent and has a fixed minimal size that is rather large and can superimpose other components of your ER diagram. An inconvenient way to solve this problem is to export your ER diagram into the ERDPlus specific ASCII format (with the file ending .erdplus), use a text editor to change the height and width of the label that overlaps other components of your ER diagram, delete the old file or rename the old file of the ER diagram in ERDPlus, and import the modified file into ERDPlus. For labels such as *m*, *n*, or 1, the value 15 for height and width is suitable.

Question 3 (Oracle)

[60 points]

Consider the following table 'Student' (the information of Students).

| STUDENTID | NAME | ADDRESS | CITY | STATE_PROV | EMAIL | DOB | MAJOR | GPA |
|-----------|--------------|--------------------|-------------|------------|-----------------------|-----------|------------------|-----|
| 1 | John Smith | 111 Main Street | Chicago | IL | johnsmith@email.com | 07-MAY-90 | Accounting | 3.5 |
| 2 | Stacy Jones | 123 Meadows Ave | Atlanta | GA | stacyjones@email.com | 17-JUL-87 | Marketing | 3.8 |
| 3 | Adam Taylor | 456 Lion Boulevard | Detroit | MI | adamtaylor@email.com | 02-FEB-85 | Computer Science | 3.2 |
| 4 | Karen White | 19 Hillside Street | Boston | MA | karenwhite@email.com | 14-JUN-82 | Psychology | 3.7 |
| 5 | Paul Johnson | 15 Summer Lane | New York | NY | pauljohnson@email.com | 09-MAR-92 | Computer Science | 3.3 |
| 6 | Jason Smith | 123 Main Street | Chicago | IL | jasonsmith@email.com | 07-MAY-90 | Accounting | 3.5 |
| 7 | Jane Smith | 222 Main Street | Chicago | IL | janesmith@email.com | 01-JAN-95 | Marketing | 3.2 |
| 8 | Bob Jones | 333 Main Street | Los Angeles | CA | bobjones@email.com | 04-NOV-92 | Computer Science | 3.9 |
| 9 | Sara White | 444 Main Street | New York | NY | sarawhite@email.com | 08-JUL-89 | Finance | 3.4 |
| 10 | Tina Gray | 555 Main Street | Boston | MA | tinagray@email.com | 05-MAR-91 | Anthropology | 3.7 |
| 11 | Jack Brown | 666 Main Street | Denver | CO | jackbrown@email.com | 10-SEP-93 | History | 3 |
| 12 | Alex Wilson | 777 Main Street | Seattle | WA | alexwilson@email.com | 12-MAY-94 | Philosophy | 3.3 |
| 13 | Manny Smith | 888 Main Street | Portland | OR | mannysmith@email.com | 02-JUL-88 | Sociology | 3.8 |

Use your CISE Oracle account to create this table and perform the operations below. Provide **SQL statements** for all operations. Show your SQL queries *and* the outputs of all results as **screenshots** in Oracle SQL Developer.

In some SQL statements we will use the logical connectives AND and OR to connect sub-conditions: <condition1> AND <condition2>, <condition1> OR <condition2>.

The ORDER BY statement followed by a comma separated list of attribute names allows one to order tuples. This statement may only appear *at the end* of a SQL statement. Sorting is performed for each attribute separately and consecutively from right to left in the attribute list. This means that attributes listed first have higher order priority than attributes following them. Adding the keywords ASC or DESC behind each attribute means that the values of the corresponding attributes are sorted in *ascending* or *descending* order respectively. Example:

```
SELECT *
FROM Student
ORDER BY Name ASC, Address DESC
```

This statement should be distinguished from the statement

```
SELECT *
FROM Student
ORDER BY Address DESC, Name ASC
```

Due to their different semantics (try it out and see!), both commands yield different table results.

Further, you will need information about the *like* command for string pattern search. Please look it up online in Oracle manuals or SQL tutorials when the command is needed.

Very few queries will require that you have an online look into SQL manuals to obtain a missing component. But in the end, these queries are not difficult.

- [10 points] Create the table 'Student', show afterwards that the table is empty at this point, then insert all records into the table, and show afterwards that and how the table is filled with tuples.
- [6 points] Find the the names and emails of all students enrolled in the Computer Science major.
- [6 points] Find the number of students enrolled in the Computer Science major or Accounting major with a GPA over 3.5.

- (d) [6 points] Find the information of students whose name has 'Smith' in it.
- (e) [6 points] Find the names, emails, GPAs, majors, and their states of all students that were born between 01/01/1990 and 01/01/1995 and have a GPA greater than 3.5.
- (f) [6 points] Find the information of students whose major is neither Marketing nor History.
- (g) [6 points] Answer the following questions.
 - (1) [3 points] Display the average GPA of all students who are from New York state.
 - (2) [3 points] Find the maximum GPA of all students in Computer Science.
- (h) [7 points] Reorganize the tuples of the table in the sense that all tuples are sorted in ascending order regarding the major. For the same major, the tuples should be sorted by GPA in descending order and then for the same GPA by name in descending order. Output the sorting attributes (only).
- (i) [7 points] Increase the GPA of students whose major is "Computer Science" by 0.1 points and show the result table.