Assignment #2

CS 348 - Fall 2022

Due : 11:59 p.m., Thur, Oct 27, 2022

Appeal Deadline: One week after return

(Total weight 10%)

Submission Instruction

This assignment will be submitted through Crowdmark. See the website for more detailed instructions. In particular, do not forget to submit one file per question to make the lives of TAs easier.

This part consists of 5 questions on SQL (lectures 3-6), E-R model (lecture 7), E-R model to relational mapping (lecture 8), and design theory (lectures 9-10), and a bonus question on the optional topic "recursion" in lecture 6. The first 5 questions has a total of 100 marks (10% of the final grade), and the bonus question is of 10 marks (a bonus of 1% to the final grade).

Question 1.

[20 marks in total]

a) (5 points) Suppose we have three relations $r(A, B), s(B, C)$, and $t(B, D)$, with all attributes declared as **not null**. Are there instances of r, s and t such that the result of

$$r \textbf{ natural left outer join } (s \textbf{ natural left outer join } t).$$

has a null value for A but a non-null value for D? If yes, given an instance of relations r, s, and t; if not, please explain why not using no more than 5 sentences. Note that **natural left outer join** between s and t refers to a left outer join between s and t based on the common columns in the two tables being joined.

b) (5 points) Consider the following schema:
`Manufacturer` (`maker`, `name`, `location`)
`Product` (`maker`, `model`, `type`)
`PC` (`model`, `speed`, `ram`, `hd`, `price`)
`Laptop` (`model`, `speed`, `ram`, `hd`, `screen`, `price`)
`Printer` (`model`, `color`, `type`, `price`)
Write an SQL trigger to carry out the following action: after **deletion** of rows in the `Product` table, for each maker in the `Manufacturer` table, check if the maker produces any other products, and if the maker does not, removes this maker from the `Manufacturer` table.

c) (5 points) SQL allows a foreign-key dependency to refer to the same relation, as in the following example:

create table `manager`
(employee_ID **char(20)**,
manager_ID **char(20)**,
primary key employee_ID,
foreign key (manager_ID) **references** `manager`(employee_ID) **on delete cascade**)

Here, employee_ID is a key to the table `manager`, meaning that each employee has at most one manager. The foreign-key clause requires that every manager also be an employee. First explain what is **delete cascade**? Then consider the following table instance, explain exactly what happens when the 4th row (4,5) in the relation `manager` shown in Table 1 is deleted, where the first column is employee_ID and the second column is manager_ID.

employee_ID	manager_ID
1	2
2	4
3	4
4	5
5	7
6	7
7	7

Table 1: Relation `manager`

d) (5 points) Using Armstrong's axioms, prove that the following is true: if $\mathcal{X} \longrightarrow \mathcal{Y}$ and $\mathcal{X} \longrightarrow \mathcal{Z}$ then $\mathcal{X} \longrightarrow \mathcal{YZ}$. As before $\mathcal{X}$, $\mathcal{Y}$, etc. are sets of attributes and $\mathcal{XZ}$ implies the union of $\mathcal{X}$ $\mathcal{Z}$. Clearly indicate at each step of your derivation, which of Armstrong's axioms you are using.

Question 2.

[20 marks in total] Assume you are developing an information system for a public library that will use a relational database system to keep track of all the books, as well as readers' loaning and returning records. An initial analysis phase of the project has resulted in the following (informal) description of the relevant data for the system.

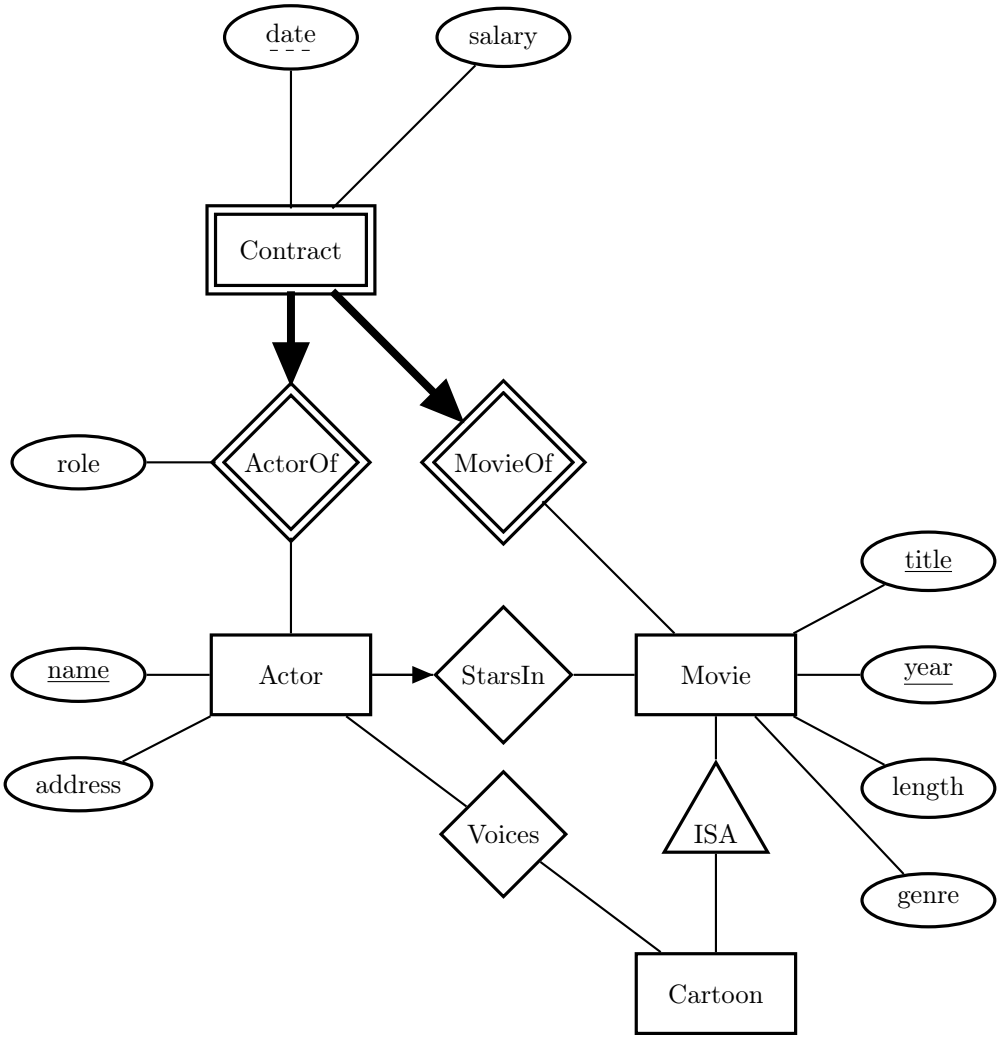
- A book has one or several authors, a year, a publisher, and is identified by ISSN. A book can have multiple copies, each copy is associated with a copyID. A book can be categorized into one category.
- A category has a name and a general description.
- A registered reader has a registration ID, name, and phone number. A reader is identified by the registration ID.
- When a reader borrows a book, the date of the loan as well as the due date are assigned to this book.
- A reader can borrow at most 10 books from the library.
- A reader has to pay penalty fines upon failing to return a book on or before its due date. Each penalty bill has a ID, and a value. The fines is calculated based on the number of days that the book was kept past the due date.

Draw an ER diagram that represents the database described above. In addition, add comments *below the diagram* to explain any requirements *not* captured by your ER diagram.
Put down your solution on the *next* page.

Write down your solution for Q2 on this page (20 points).

Question 3.

[20 marks in total] Convert the ER diagram given below into a set of relational DDL commands (e.g., create table) that define a relational schema. The commands should include primary and foreign key constraints and assertions where appropriate. Include comments to clarify any unusual translation decisions, and indicate any parts of the ER specification that are not enforced by your relational schema.



Write down your solution for Q3 on this page (20 points).

Question 4.

[15 marks in total]

Consider the following collection of relations and dependencies. Assume that each relation is obtained through decomposition from a relation with attributes $ABCDEFGHI$ and that all the known dependencies over relation $ABCDEFGHI$ are listed for each question. (The questions are independent of each other, obviously, since the given dependencies over $ABCDEFGHI$ are different.) For each (sub)relation: (a) state a candidate key based on the given set of dependencies; (b) State whether the relation is in BCNF or 3NF or neither, and explain why; (c) If it is not in BCNF, decompose it into a collection of BCNF relations.

a. (5 points) $R_1(A, C, B, D, E)$, $\{A \rightarrow B,\ C \rightarrow D\}$

b. (5 points) $R_2(A, D, G, H)$, $\{D \rightarrow G,\ G \rightarrow H\}$

c. (5 points) $R_3(A, I, C, B)$, $\{A \rightarrow I,\ I \rightarrow A,\ B \rightarrow H\}$ Note that H is not included in R_3, but we still have $B \rightarrow H$ given. This is intentional :)

Question 5.

[25 marks in total] Given the relaton $R = ABCDEF$, with the set of dependencies

$$F = \{AB \rightarrow CE,\ ACD \rightarrow B,\ BE \rightarrow CF,\ AC \rightarrow DE,\ BF \rightarrow A\},$$

answer each of the following questions. **Note**: When solving each task you may use the information already obtained from the tasks that preceded it. In particular you may choose to use the minimal cover for solving questions b to d.

a. (5 points) What is the attribute closure of of AB? Please show how you derive its attribute closure.

b. (5 points) List all candidate keys for R (as groups of attributes separated by comma, e.g. A, BC).

c. (5 points) Derive a minimal cover for F. List the set of dependencies obtained in each step of the minimal cover algorithm.

d. (5 points) Determine if a decomposition $R_1 = ABC,\ R_2 = ACD,\ R_3 = ACEF$ preserves the dependencies $AB \rightarrow E$ and $BE \rightarrow F$. Explain your answer.

e. (5 points) Given $R_2 = ABCDE$ and $F_2 = \{A \rightarrow B,\ BC \rightarrow E,\ ED \rightarrow A\}$, derive a `BCNF` lossless-join decomposition of R_2. Show each step of the algorithm.

Question 6.

(Bonus 10 points) Consider a `manager` relation with the same schema as Q1(c).

a. (5 points) Write a recursive SQL query that outputs the ids of all the employees are working under employee id 7 [5 points]. For example, given the database instance shown in Q1(c), the list of employees working under id 7 is (1,2,3,4,5,6,7).

b. (3 points) Use this example to explain why your query works.

c. (2 points) State whether your recursive query is linear or non-linear and why.

Note that this question requires additional reading/learning of the optional materials on the topic "Recursion" from Lecture 6. This topic won't be tested in the exams.