Database Systems
(2013/07/05)

[1]

Consider a fictitious DBMS, called F-DBMS that uses a relational approach and supports a simple SQL-like language. The overview of this fictitious DBMS is as follows:

> F-DBMS uses a basic file system (such as the one in UNIX or Linux) to store relations. There is one file for each relation. For example, there is a file named R for relation R. The file for a relation has one line for each tuple. Values of attributes of a tuple are stored as character strings, separated by two consecutive special marker characters ## that are assumed not to appear in the data. The DB schema is stored in a special file named `schemas`. For each relation, the file `schemas` has a line beginning with that relation name, in which attribute names alternate with types. The two consecutive characters ## separate these elements.

To process a SQL-like query such as `SELECT A1, A2 FROM R WHERE <Condition>`,
(i) F-DBMS reads file `schemas` to find out the schema of relation `R` in the `FROM` clause (ii) Check that the query is semantically valid for `R` (e.g., check that the name of the relation in the FROM clause exists in file `Schemas`, names of attributes in the `SELECT` clause are correct, etc), (iii) Read each line of the file named `R`, check `<Condition>` is satisfied, and output the values of attributes `A1` and `A2` in the line as a result tuple, if the condition is true.
Assume that grouping/ordering related clauses and insert/delete/update statements in SQL are also supported in some simple ways such as above. Assume also that embedding SQL statements in some programming language is also possible. Except for supporting SQL statements as described above, no other mechanism is provided in this DB system.

Describe possible problems (or inconvenience) of this fictitious DBMS for various database applications.

(20 points)

[2]

(a) A set of relational operators is complete if any other relational operations can be expressed in terms of operators from this set. What are these operators? (5 points)

(b) Briefly describe the converting procedure from the ER-diagram to the relational schema discussed in the class. You are required to describe (i) the basic idea together with how to reduce the number of tables, and (ii) how to handle the weak entity set. (15 points)

[3] Consider a database for the scheduling of classrooms for mid-term exams. For your convenience, use the following four entity sets.

- EXAMS with attributes `exam-id` and `time-for-exam`.
- COURSES with attributes `c-name`, `c-number` and `dept-name`.
- SECTIONS with attributes `s-number` and `no-of-enrollments`. This entity set is weak, which depends on COURSES. (Note that a course consists of one or more sections.)
- ROOMS with attributes `r-number`, `capacity` and `building`.

There can be multiple sections for a course. However, there can be only one exam for each section of a course. Multiple rooms can be assigned to one exam when we cannot find a room whose capacity is enough for a certain exam.

Design an ER-diagram for the above database. For each entity set, indicate a key appropriately. You also have to indicate appropriate mapping cardinalities in the relationships between entity sets.

(20 points)

[4]

(a) Show that the following two inference rules for functional dependencies follow from Armstrong's axioms. In your proof, you have to indicate which Armstrong's axioms are used. You can assume that sets of attributes in the left side and in the right side are disjoint.

(i) If $X \rightarrow A$ and $X \rightarrow B$ hold, then $X \rightarrow AB$ also holds.

(ii) If $X \rightarrow Y$ and $WY \rightarrow Z$ hold, then $WX \rightarrow Z$ also holds

(10 points)

(b) Consider a relation R (A, B, C, D, U, V, W) with the set of functional dependencies

$\{AB \rightarrow DV, AD \rightarrow CW, DW \rightarrow UV, A \rightarrow D, D \rightarrow A\}$

Find all the keys for R. (5 points)

Explain how you can be sure that you found all the keys for R. (5 points)

[5] Consider the Employee database below.

Employee (<u>e-name</u>, e-street, e-city)

Company (<u>c-name</u>, c-city)

Works (<u>e-name</u>, <u>c-name</u>, salary)

Manages (<u>e-name</u>, m-name)

Write the following queries in SQL.

Find the companies, which are located in New York, whose employees earn a higher salary on average than the average salary at 'City Bank'. (20 points)