Assignment I

Normalization (paper lab)

The purpose of this exercise is to get understanding of the different normal forms and of the problems that can be prevented by normalization. After the lab, the student should be able to determine if a database schema conforms to the Boyce-Codd normal form (BCNF). The students should be familiar with modeling real world scenarios in terms of ER model and translate an ER diagram into a relational database schema.

1 Background reading

- Elmasri/Navathe: chapter 3, 10, sections 7.1, 11.1 and 11.2.
- Padron-McCarthy/Risch: chapter 2, 5, 6, and 11.

2 The scenario – A Library Database

The city library of Uppsala has a database to keep track of its books, the people who borrow books, and which books have been borrowed by who. The tables, with the data, are included below in sections 4 and 5. (For well-known reasons, there are very few books in the library.)
Unfortunately, the design of the database is not very good. Your mission is to analyze the problems with the design, and suggest a better one.
The output of your work should be a report that addresses all the faults you find regarding normalization along with a description of why they are problematic. Furthermore, you should develop an alternative design that is in BCNF.

3 Exercises

1) Specify all functional dependencies in the library scenario. Then state in which normal form (1NF, 2NF, 3NF or BCNF) each of the existing tables is, and why. Assume that there is one telephone number per address.
2) For each table that doesn't fulfill the requirements for BCNF, explain the problems that this lack of normalization has and their potential consequences. Give some examples.
3) Design a new database for the library, where all the tables fulfill BCNF without losing any information. Use the top-down approach for the relational database design by starting with an ER-diagram for the database and mapping the diagram to relational tables.

Hints:
- Consider dates as atomic (there is a type 'date' in SQL);
- For more details on ER-to-Relational mapping see E/N chapter 7.1.

4 The schema of the existing database

There are three tables:

- A table called **BOOK**, which contains data about the books. It has the attributes **TitleNr** (a number that this library assigns), **ISBN**, **CopyNr** (which is used to separate different copies of the same book), **Title**, **PublYear**, **Author**, and **AuthorNat**. The primary key consists of **TitleNr**, **CopyNr** and **Author**. An alternative key is formed by **ISBN**, **CopyNr** and **Author**.
- A table called **CUSTOMER**, which contains data about the persons who can borrow books. It has the attributes **CustomerNr** (a unique number identifying a person, assigned by the library), **PersonNr** (which is a unique number identifying a person, assigned by the Swedish state), **Name**, **Address**, **Tel**, and **NrBooks** (the number of books that this person has borrowed at the moment). **CustomerNr** is the primary key. **PersonNr** is an alternative key.
- A table called **LOAN**, where the loans are stored. It has the attributes **TitleNr**, **CopyNr**, **CustomerNr**, **Date** (which is the date when the book was borrowed), and **BorrowerName** (which is the name of the customer who borrowed the book). The primary key consists of **TitleNr** and **CopyNr**.

5 The contents of the existing database

The tables look like this:

BOOK

TitleNr	ISBN	CopyNr	Title	PublYear	Author	AuthorNat
1	0071148108	1	Database	1997	Silberschatz	USA
1	0071148108	1	Database	1997	Korth	USA
1	0071148108	1	Database	1997	Sudarshan	India
2	0805317538	1	Fundamentals	1994	Elmasri	USA
2	0805317538	1	Fundamentals	1994	Navathe	USA
2	0805317538	2	Fundamentals	1994	Elmasri	USA
2	0805317538	2	Fundamentals	1994	Navathe	USA
3	0198642253	1	Mord	1995	Guillou	Sweden
3	0198642253	2	Mord	1995	Guillou	Sweden
4	3411021764	1	Våld	1998	Guillou	Sweden

CUSTOMER

CustomerNr	PersonNr	Name	Address	Tel	NrBooks
1	6312111658	McCarthy	Vägen 7	282677	1
2	4403149901	Larsson	Gatan 6	146000	1
3	4010229910	Ericsson	Gatan 8	241000	1
4	4501129921	Schwarzenegger	Vägen 3	174590	0

LOAN

TitleNr	CopyNr	CustomerNr	Date	BorrowerName
1	1	3	7/1 98	Ericsson
3	2	1	1/9 98	McCarthy
2	1	2	7/1 98	Larsson

6 Handing in solutions

Your report should contain your answers to the questions in part 3. The database design in exercise 3.3 should be presented as both ER diagram and the corresponding relational tables in format **table_name**(*column1*, *column2*,…). Motivate why your relations are in BCNF by specifying all functional dependencies.