LABORATORY MANUAL

SC2207/CZ2007: Introduction to Databases

Implementation of a Database Application

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

NANYANG TECHNOLOGICAL UNIVERSITY

1. OBJECTIVES

Upon completion of the assignment, the student should be able to:

a. Construct an entity-relationship model at a conceptual level.
b. Map the model into a schema of a relational DBMS.
c. Implement the given schema on a relational DBMS.
d. Use a database language (SQL) to retrieval data from a relational DBMS.

2. LABORATORY

This is a team-based assignment. Each team consists of **five to seven** members from your laboratory group, to be approved by your lab supervisor, to be formed during the Lab 1 session. The lab supervisor may add or remove members from your teams to ensure even spread and mix of students in each team. The final members of your team must be submitted to the lab technician during the **first laboratory session**.

There are five scheduled lab sessions for this team assignment. Laboratory sessions will start from the **third week** of the semester. For students whose lab sessions are scheduled on odd weeks, the first lab session is on Week 3. For students whose lab sessions are scheduled on even weeks, the first lab session is on Week 4.

You might need more than the scheduled five sessions for the actual implementation. You are encouraged to **start early** with your lab work (as soon as the topic is covered in the lectures).

Attendance is taken for the **first**, **third** and **fifth** lab sessions only. Attendance for the second and fourth lab sessions is not mandatory. Failing to sign-in for the first, third, or fifth lab session may result in F grade for the respective assessment.

To encourage fair contribution to the lab work, each team needs to indicate contributions from each member. The final mark of a team member may be adjusted based on the team score and individual contributions. **Appendix C is to be submitted with each submission**.

For each lab session, there is a lab supervisor and a lab technician assisting you. The lab supervisor is a professor or a teaching assistant whom you may approach for clarifications on lab work, lab report submission, graded lab reports, etc. The lab technician is a technical staff whom you may approach for lab logistics (lab attendance, SQL Server account matter, lab submission deadline, computer problem, etc.)

3. INTRODUCTION

The assignment covers the portion of the course concerning data modelling, database design and implementation from the user's viewpoint. Thus, the assignment involves modelling as well as implementation aspects of the database course.

The overall aim of the laboratory is to develop an application based on a given data model using a given database management system. This exercise will bring you through a crucial first part of the life cycle of a database application. It is assumed that the data analysis has been performed. Note that this manual provides you with more information than is required for the first laboratory session, e.g., not all constraints can be modelled in the beginning but are included at a later implementation stage. In contrast you might require additional information for an understanding of the application. Proceed by stating your assumptions in written form and / or ask your laboratory supervisor.

4. DESCRIPTION OF THE ASSIGNMENT

The description of the application is given in Appendix A and B. This includes background and general requirements of the application, conceptual information about the system and its users as well as a list of SQL queries that must be fulfilled as a minimum. Note that teamwork is required. Each team will submit one solution. **No individual submission will be accepted.**

4.1 First Laboratory Session: Creating an ER Diagram

Appendix A gives conceptual information about the project obtained after a partial system analysis was performed. Based on the description, construct a suitable ER diagram. Analyze the choice of entity sets, different types of relationships required, the usage of weak entity sets, subclasses, etc. and compare them with alternative solutions from your team members. You need to submit the following, latest **three working days** after the first laboratory session:

- A PDF document of your ER diagram. A good ER diagram is one that is self-explanatory. If you believe certain parts of your ER diagram need explanation, you can include a written description (maximum one page). Combine both the ER diagram and the written explanation (if any) as a single PDF document, labeled as follows: Lab1_XXX_TeamY.pdf, where XXX is your lab group number and Y is your team number. Marks are given for neat presentation of your ER diagram.

- Assessment for Lab 1 is based on whether the submitted ER diagram reflects correct understanding of ER diagram artefacts (entity sets, relationships, weak entities, subclasses, etc.) and whether they are used correctly and appropriately. Do note that not all information given in Appendix A can be represented in an ER diagram.

4.2 Second Laboratory Session: Finalization of the ER Diagram

There is no submission for the second laboratory session. In this lab, each team should finalize their database design based on the feedback received from their lab supervisor. Please note that the second laboratory session is a free access session; i.e., attendance is not mandatory.

4.3 Third Laboratory Session: Generation of Normalized Database Schema

In this lab, you convert the ER diagram into relational schema and ensure that the relations are at least in 3NF. Follow the general guidelines covered during the lectures and tutorials to produce suitably normalized relations. For each relation, the key(s), primary key, and functional dependencies must be specified. If a relation is generated due to the normalization of an original relation, then the normalization steps must be presented. You need to submit the following, latest **three working days** after the third laboratory session:

- A PDF document of the normalized database schema and FDs associated with each relation. Label the PDF document as: Lab3_XXX_TeamY.pdf, where XXX is your lab group number and Y is your team number. If a relation created from the ER diagram violates 3NF, then this should be highlighted along with the decomposed normalized relations. Note that for this lab, no SQL code should be submitted. Hence, the structure of your solution shall be similar to the following example:

  R1(A, B, C, D)
  Keys: AB, AD
  Primary Key: AB
  FDs: AB → CD, A → D
  The relation is in 3NF. (If relation is not in 3NF, perform the steps of the 3NF normalization.)

- Assessment for Lab 3 is based on whether the submitted report reflects correct understanding of keys in relations, identification of appropriate functional dependencies in each relation, how normalized relations are formed, and whether the normalizations are correctly and appropriately performed. Do note that in your final set of relations, the keys and functional dependencies in each relation may not be explicitly given in the description in Appendix A.

4.4 Fourth Laboratory Session: Implementation of the database schema

There is no submission for the fourth laboratory session. In this lab, the finalized database schema is to be implemented using SQL DDL commands. Your tables should be appropriately populated with sufficient realistic records using SQL INSERT statements so that your query solution for Appendix B results in some output records (3 to 5) for each query. Your implementation should

clearly incorporate the primary and foreign keys, data types, and any form of constraints. The lab provides MS SQL Server software for your implementation.

Please note that the fourth laboratory session is a free access session, i.e., attendance is not mandatory.

4.5 Fifth Laboratory Session: Final demonstration

In this lab, the implementation obtained from the previous laboratory session must now be extended to provide SQL query solutions for the queries in Appendix B. **At the end of the lab session**, you need to submit a single PDF document containing the followings:

- SQL DDL commands for table creation (Lab 4).
- SQL statements to solve the queries in Appendix B and additional queries. Each query should be immediately followed by the query output. Briefly explain how the output is obtained.
- A printout of all table records.
- Description of any additional effort made.

Label the PDF document as: Lab5_XXX_TeamY.pdf, where XXX is your lab group number and Y is your team number. You should prepare the PDF document in advance before coming to the lab. Some DDL commands may look like this:

```
CREATE TABLE name (
        attr1 datatype NOT NULL,
        attr2 datatype,
        ...
        PRIMARY KEY (attr1),
        FOREIGN KEY (attr3) REFERENCES name(attr1)
                ON DELETE ... ON UPDATE ...,
);
```

In addition to the PDF document, you are to capture **screen recording** of query execution as a mp4 video file. For each query in Appendix B and additional queries, first show the SQL statement, then execute the query and show the query results, all recorded as a mp4 video file. Each query video should be no more than 30 seconds and labeled as: Lab5_XXX_TeamY_Q#.mp4 where # is the query number. Zip the PDF and all mp4 files into one single ZIP file.

During the lab session, you may be given additional queries to solve. In addition, your lab supervisor may require in-person live demonstration and Q&A. All team members are to actively contribute during the demonstration session and be familiar with **all aspects** of the project. No slide presentation is required.

APPENDIX A: APPLICATION DESCRIPTION

Suppose that you are to construct a database for a university's resource management, and the requirements are as follows:

- There are four groups of people on whom a university is most dependent: stakeholders, professors, staff, and students. Some common attributes are shared by all of these persons: Person_ID (identifier), Name, Address, City, City, State, Zip, Phone, Email, and Schools.

- Each of the four groups has at least one unique attribute of its own. Stakeholders belong to a particular domain (e.g., government, funding agency, industry partners, and public); professors have Fields_of_Expertise; staff have Staff_ID, Date_Hired, and Position; students have Student_ID, Admission_Date, and Major&Minor. A particular person may belong to two or more of these groups at any time (e.g., a person can be a professor and staff).

- There are administrative staff and technical staff to help manage schools' daily operations. A technical staff is assigned to at most one laboratory with attributes Name, School, and Location. The Name and School form a joint identifier for each laboratory. The laboratories are categorized into teaching laboratories and research laboratories. Equipment in each laboratory is recorded with Name, Date_Purchased, Model_No, and ID. An equipment can only be identified given the Name and School of the laboratory.

- Students take courses taught by at least one professors on particular dates. The students are grouped into undergraduates and graduates. The undergraduates do experiments in the teaching laboratories and their attendance is taken on particular dates. Graduates who do research are assigned to at least one research laboratories and supervised by at least one professors on a particular research topic. Each professor has their own timetable for classes to teach on particular dates and time.

- Before a student enrols in a course, he/she must check that it does not schedule conflict with another course.

- Stakeholders can provide comments and suggestions for the schools' development. Each comment or suggestion on particular topic is time-stamped on a particular date.

Note that the information above may not be complete. Some aspects of the database application's details may have been omitted. It is expected that you come up with their own solutions in case of inconsistencies or missing information. However, you have to keep track of these aspects and explain your assumptions in your submitted report. Extensions to the implementation of the basic system are encouraged.

APPENDIX B: QUERIES

1. Find all stakeholders who belong to the public domain.
2. Find all stakeholders who have provided at least five comments or suggestions.
3. Find graduates who are supervised by more than one professors and assigned to more than one research laboratories.
4. Find all professors who teach more than one courses in the semester.
5. List all the equipment belonging to a particular laboratory.
6. Find all undergraduates who have not attended at least one laboratory experiments.
7. List all graduates who are doing research and taking courses in the semester.

APPENDIX C: INDIVIDUAL CONTRIBUTION FORM

Full Name	Individual Contribution to Submission 1 (Lab 1)	Percentage of Contribution	Signature

Full Name	Individual Contribution to Submission 2 (Lab 3)	Percentage of Contribution	Signature

Full Name	Individual Contribution to Submission 3 (Lab 5)	Percentage of Contribution	Signature