

CMSC 508 - Semester project

Problem Definition and Database Design

John Leonard
August 21, 2022

Objective

The objective of the first phase of the final project is to design a database that contributes to a solution to a real-world problem selected by your student team.

Teams are composed of TWO members unless explicitly allowed by the instructor.

Students may choose a problem related to their own personal interests and background if you like. **All project proposals MUST HAVE my approval.** Some examples may be online bike store management, DNA database sequencing, music/movies rating system, sports/players database, collaborative networks, etc.

The main aim is designing a database that satisfies the information and functional requisites extracted from the analysis of the problem statement, in order to demonstrate your expertise and proficiency on the use of database design methodologies to create an accurate and faithful database. The design must reflect all expected information, scenarios, and functionality.

Deliverables

- Problem statement: a 2-3-page document with the definition of the real-world problem, providing contextualization, scope of the system, and listing all expected information and functionality. This should include 20 distinct queries that could be asked of the database written in plain English statements. The problem statement must be a Markdown document stored in this repository.

Note: distinct queries are those that are entirely different. Counterexample: “Display a list of student last names in alphabetical order” is not distinct from “Display a list of student names and email addresses”. However, “Display a list of student last names in alphabetical order” is distinct from “Who are the five students with the highest GPA?”

- Entity-relationship diagram (ERD). Created using proper software tools and consistent notation. Note: your design should consist of at least 4 major entities. This may be uploaded as an image (.png preferred).
- Translation of the ERD into the relational model as relational schemas. **Listing of relations, their attributes, types, domain, and constraints, identification of the primary key and foreign keys.**
- Identification of the Functional Dependencies that exist for each relation, **based on the real-world domain of the database.**
- Normalization of the relations according to BCNF/4NF. A description, in a text format, of the process of decomposing the tables extracted from the ERD translation into relations satisfying BCNF/4NF **that includes a justification of the normalization claim.** Decomposition via BCNF/4NF must be lossless.
- Sample data for each relation that could be used to populate the database to illustrate the behavior and interrelationships among data.

Important! You must verify that the solution can reflect all data scenarios and information requisites identified in the problem statement.