Assignment 1: ERD Design & DDL Conversion

University of Toronto Mississauga

Due: Thursday, February 6^{th}, 2025 at 10:00:00 PM ET

Released: Thursday, January 23^{rd}, 2025 || Version: 1.0

ER Diagram Requirements

You have been recruited by the Ministry of Health (`MoH`) as a database architect to design a database schema which accurately records and efficiently retrieves information about health care entities and their interactions in certain hospital settings. The provincial government would like you to store information about persons, hospitals and their departments, drugs, as well as medical tests.

A person may be a patient, physician, or a nurse. A person has a unique identifier, first name, last name, gender, date of birth, and an address that consists of a street, city, province, and postal code. Every person has at least one phone number; possibly more. For each telephone number, the number is recorded along with its contact type (home, work, or mobile). A phone number can only be associated to exactly one person. Each patient is assumed to have some type of health insurance, with possible categories being: public, private, and self-funded. For each physician, the `MoH` would like to record his/her medical specialty, number of years in practice, and his/her yearly salary (in CAD). For nurses, the `MoH` would like to know their yearly salary (in CAD), and their number of years in practice too.

The `MoH` has requested hospital information in the province, including the hospital name (which serves as its identifier), street address, city, and annual budget (in CAD). Every hospital has at least one medical department that provides medical services. A department is identified by its name, and the hospital in which it belongs to. The `MoH` is also interested in recording each department's annual budget. Each physician belongs to exactly one department based on his/her specialty. Each department must have at least one, possibly more, affiliated physicians. Nurses, however, may work in one or more departments, and every department must have at least one affiliated nurse.

As patients arrive to a hospital, an admissions record is created indicating the admit date, and a priority (one of: immediate, urgent, standard, or non-urgent). Patients may be admitted to many hospitals (including zero), and a hospital may admit many patients (including none). Each patient is under the care of one nurse. However, a nurse may care for one or many patients. Every patient is associated with one physician, who makes a diagnosis of the patient's condition, recording details such as the diagnosed disease, the date of diagnosis, and the prognosis (one of: excellent, good, fair, poor, or very poor).

Every physician treats at least one patient by providing a diagnosis. Patients may undergo a set of medical tests to confirm and treat a diagnosed disease. Each medical test is identified by a unique numeric identifier, the name of the test, and a fee (in CAD). Each time a patient undergoes a test, the test date, and test results (assumed to be in string format) are recorded. Physicians may provide a prescription to a patient to treat his/her illness. A prescription associates the prescribing physician to the intended patient, recording the date of the prescription. Lastly, every prescription must contain a prescribed drug, and the dosage for the intended patient. A drug is identified by its 8-digit drug code, generic name, drug category, and the unit cost. A drug may be prescribed to zero or many patients.

Tasks

1. Construct the Entity-Relationship Diagram (ERD)

Draw an Entity-Relationship Diagram (ERD) to capture the described requirements. The ERD must be electronically drawn (using Draw.io) and you will submit this in a PDF file called `a1.pdf` along with the XML file (`a1.xml`). Do not add additional attributes, relationships, or entity-sets. You must stick to the design specification above, so be careful not to assume anything outside of this context. If you wish to justify a choice made due to ambiguity, you may do so in text format in the submission PDF (on a different page from the ERD). You must use the version of Chen's Notation, as per the notational reference sheet provided in the lecture/on Quercus. Do not make up your own notation, do not use the textbook notation, do not google notation. Please ensure that your ERD is readable and that line crossing is kept to a minimum (or does not happen at all). Neither handwritten submissions nor hand-drawn illustrations will be accepted for credit.

2. Create the Data Definition Language (DDL)

Provide the corresponding PostgreSQL `CREATE TABLE` statements describing the relational schema you have designed in Task 1. You must enforce both inter- and intra-relational constraints (note: you cannot disable foreign key checks!), all relevant `CHECK` constraints, any necessary `CASCADE`, and custom data types must be included as well. You do not need to write any triggers. Please include all your statements in an executable script (`a1.ddl`) that can be run on the MCS PostgreSQL server (i.e., we will invoke the `\i` command). Scripts that do not execute on the server, when the file is run, will not be marked and a grade of 0 assigned (please see the course syllabus for more information).

Requirements and Submission

Assignments must be completed individually. Students are forbidden from collaborating with others (i.e., you cannot discuss your solutions/ideas with anyone else, however, you may seek assistance/clarification from the CSC343 Teaching Staff, but not other students, individuals, forums, external services, or external websites). Please read our "Use of Artificial Intelligence (AI)" statement in detail, as all submitted work (i.e., any graded assessments) must be the original work produced by the individual student alone – meaning you are STRICTLY PROHIBITED FROM USING GENERATIVE AI (e.g., ChatGPT). Please refer to the syllabus for additional details on assessments and the "Minimum Standards for Submitted Work".

All submissions are required to be written in LaTeX. I would strongly suggest using Overleaf as a means of editing LaTeX documents. You are not permitted to use another word processor.

All files are to be submitted using the MarkUs platform (https://markus.teach.cs.toronto.edu/utm-2025-01). You may submit as many times as you like, in fact you are encouraged to submit many times! Please ensure your answers are typed and submissions are legible/understandable.

Entity-Relationship Diagrams (ERDs) must be completed on Draw.io allowing you to extract the raw `XML` file (which is to be submitted too). The XML file should be "compressed" and include "all pages".

You must include your full name, student number, and UTORid in all files; the LaTeX file, DDL file, and in the PDF. Submit your answers to the ERD (Task 1) in one file called **`a1.pdf`**, as well as your answers to your `DDL` file called **`a1.ddl`**. Additionally, you must submit the source LaTeX file (**`a1.tex`**) and `XML` file (**`a1.xml`**).

Naming Conventions and Guidelines

Convention dictates that SQL keywords are in `UPPERCASE` while relations (i.e., entity-sets and relationships) are in `snake_case`. Additionally, relation names must be unique, so do not duplicate them!

More specifically:

1. Table Names
    - Singular and descriptive names.
    - Use underscores to separate words.
    - Examples: "student" or "course_enrollment"

2. Column Names

    - Use lowercase with underscores.
    - Prefix with a short, relevant abbreviation of the table name. Alternatively, you may use the entire table name.
    - Examples: In the "student" table: "std_name" (or "student_name"), "std_email" (or "student_email").

3. Primary Keys

    - Use the table name followed by "_id".
    - Examples: "std_id" (or "student_id"), "crs_id" (or "course_id").

4. Foreign Keys

    - Use the primary key name of the reference table.
    - Example: "student_id" in the "course_enrollment" table referring to the student table.

5. Commenting Conventions

    - Use comments to explain complex queries or logic, be concise and to the point.
    - Start with `--` for single-line comments or `/* ... */` for block comments.

It is expected that every Assignment 1 submission follows these conventions, otherwise they will be penalized.