Project 1: ER Modeling & Schema Design

Due on 9/23/2016

In this project, you will apply the principles of database design that we learned in class in order to implement a relational schema for an application of your choice. You can choose an application from three different domains: *movies*, *books* and *publications*. You will start by specifying the functionalities of the application that you plan to support, then create an ER diagram to model the data, and finally translate the ER diagram to a relational schema.

1. **[10%]** Describe the functionalities supported by the application you chose. Such functionalities can be queries that can be supported, which data will be updated, and so on.

2. **[20%]** Design an ER diagram with the necessary entity sets, relations and attributes that support the functionalities you described. The ER diagram must *at the minimum* include the following:

   - 6 entity sets
   - 5 relations
   - a many-one relation
   - an ISA relationship

   Apart from the above, you should aim to specify as many constraints as possible that will be consistent with the application.

3. **[20%]** Design the relational schema that corresponds to the ER diagram. The schema will be specified as a sequence of SQL commands (for *SQLite3* DBMS) that create the relations. Include also SQL commands that populate each table with at least one tuple. The resulting schema should have the necessary key and foreign key constraints, as well as the appropriate data types.

You can choose the application from three domains:

Movies. In this domain you would be modeling the following entities: movies, actors, directors, theaters, reviews, as well as other entities of your choice. You can take a look at IMDb (`www.imdb.com`) to see how the data is modeled and what attributes you need. You can choose to support various queries, such as finding movies playing in a specific location, or finding movies sorted by average review score. You could also choose to support the addition of new reviews, movies, etc.

Books. In this domain you would be modeling the following entities: books, authors, publisher, reviews, and other entities of your choice. Goodreads (`www.goodreads.com`) is a good source of data in this form. There are many queries that you could choose to support: find a book of a specific genre, find books that have won some award, and so on. You could also choose to support a recommendation system (so you need to figure out what would be the necessary data that you need to model for this purpose).

Scientific Publications. In this domain you would be modeling the following entities: article, author, conference, journal, and other entities of your choice. You can take a look at DBLP (`dblp.uni-trier.de`) or PubMed (`www.ncbi.nlm.nih.gov/pubmed`), which are resources that collects bibliographic information from many different sources. You could support queries such as searching for a specific paper, counting the publications of an author, or finding the co-authors for a given author.

DELIVERABLES

You can download and install SQLite3 from the following link (the current version of SQLite3, which will be used for testing, is 3.14.2):

<div align="center">

`https://www.sqlite.org/download.html`

</div>

You are required to submit all the necessary material in a single zipped folder (use GZip or WinZip). Upload the zipped folder using Canvas in `Learn@UW` (Project 1). The folder will include the following deliverables:

1. A PDF document with the answers to questions 1 and 2 (the ER diagram).

2. A single plaintext file with the name `schema.sql` that includes the SQL commands for question 3.