601.315 Databases, Spring 2022
Project Phase B: ER Diagram and Relation Schemas

Due: Fri, 18 Mar at 11pm. Use of late days *is NOT* permitted.

Part 0: Updates Since Phase A. In response to your Phase A submission, you may have received comments via Gradescope suggesting a shift in topic direction, the creation of a richer domain, or the focusing of your work to a smaller area. Please review your comments and indicate any changes you've made in response. If no modifications were needed, simply note that as your response for this part.

Part 1: E-R Diagram. The next step in your course project is to create an entity-relationship (E-R) diagram for your selected topic domain. Your diagram should include entities, attributes and relationships, along with relationship participation constraints (total or partial, or in some cases numeric constraints). The diagram might also include derived attributes, composite attributes, multi-valued attributes, weak entity sets, and specialization. Consider alternate designs as you go, and aim to select one that will minimize redundancy. As you draft your diagram, consider the questions you submitted in part A. Ensure that your design allows you to answer these questions. Several rounds of diagram revision may be needed in order to come up with a clean design. Your final version may be legibly hand-drawn, or diagrammed using software.

Part 2: Draft Schemas. Next, convert your finalized E-R diagram into a set of relation schemas. Refer back to class slides and videos which describe the steps in the conversion process. Give a list of schemas resulting from your design, indicating primary key attributes in each relation by underlining them. Include a short paragraph to describe any decisions you made in the conversion process that do not seem straightforward.

Part 3: Finalized Schemas. Once your draft schema is completed, consider whether there are ways to clean up your design. For example, if the natural primary key of a relation consists of multiple attributes, do you instead want to introduce a unique identifier for each entity in the relation (one that is determined by the database using auto-numbering upon insertion of each tuple) to use as a primary key? Take time to list the set F of functional dependencies warranted in your topic domain. Are your relations all in third normal form (3NF) relative to your set F? If not, document the steps of your decomposition process so that your final design contains relations that are in 3NF. Are there any other modifications to your first draft set of schemas that seem necessary? For this part, write a brief discussion of your work to revise the schemas you laid out in Part 2 above. Be sure to include your set F of functional dependencies and any steps taken to normalize your relations, and to clearly label your final set of relation schemas, again indicating primary key attributes using underlining. Finally, also give a numbered list of constraints other than primary key constraints that you'll want to enforce in the database. These could be foreign key constraints, domain constraints, participation constraints, etc.

Deliverable. Create a PDF file containing your answers to all parts listed above. Make sure that the response for each part is clearly labeled as Part 0, Part 1, etc. The top of the document must list the names and JHED IDs of each partner. Submit your work as a PDF via Gradescope by the deadline listed above for Phase B. One partner will submit the work as a team submission upload, and will indicate all partner names. Therefore, only one partner should submit.
