CS411: Database Systems **Spring 2017**

Assignment 1 Solutions

Question 1. Short Questions (10 points)

Provide a short answer (4–5 sentences at most) for each of the following questions. You may use figures if necessary.

1. Consider the relation Person (Name, SSN, Age, Address, Gender). How many keys does the relation Person have? How did you arrive at your answer?

   Solution: All subsets of the attributes containing SSN would be a key. Therefore, the number of keys is $2^4 = 16$.

2. All relationships involving a weak entity set can be ignored while translating an ER diagram to a relational model. Justify or prove otherwise.

   Solution: All relationships can not be ignored. Only weak relationships can be ignored. Weak entity sets can still participate in relationships in the standard sense.

3. The expressiveness of ER models would reduce if we do not allow relationships to have attributes. Justify or prove otherwise.

   Solution: The expressiveness of ER models *wouldn't* reduce. One can have a proxy relation participating in the relationship, which has the corresponding attributes instead[1].

4. It is possible to transform a multiway relationship to multiple binary relationships without using weak entity sets. Justify or prove otherwise.

   Solution: It is not possible. Refer the Purchase example from slides.

5. The reason why we prefer to combine the relation corresponding to an entity set A, with the relation corresponding to a relationship B—where B is a many-one relationship from A to another entity set—is because we want to improve the efficiency of queries involving A.

   Solution: The claim as presented in the question is invalid. We can potentially save space upon the combination. We can not ascertain any improvement in the execution of queries involving A.

[1] described in slides

Question 2. ER Models (45 points)

Consider the following information about a database of a university.

- Departments have a department number, department name, and many research areas.
- Professors have an SSN, a name, an age, a rank and a main research area.
- Projects have a project number, a sponsor name, a starting date, an ending date and a budget.
- Graduate students have an SSN, a name, and an age. They major in a single department.
- Graduate students can either be an MS or a PhD. PhD students need to determine their speciality/main research area.
- All PhD students have a professor as an advisor.
- Each project is managed by one professor.
- Each project is worked on by one or more professors, and one or more graduate students.
- Graduate students can work on multiple projects.
- Every department has a head, who is a Professor.
- Professors can work in one or more departments. For each department they work in, there is an associated time percentage.
- Graduate students have one major department in which they are working towards their degree
- Each graduate student has another senior graduate student as a mentor.

Design and draw an ER diagram that captures the aforementioned information. Indicate the key of each entity, as well as the multiplicity of your relationships. You are free to use annotation tools such as Mac Preview or Microsoft PowerPoint to draw the ER diagrams. Please do not include scanned pictures. You may want to check out draw.io.
Note: state your assumptions clearly. Since there are many correct answers, your ER diagram will be evaluated considering your assumptions.
Solution: A possible solution is presented in Figure 1. Like stated in the question, valid asssumptions will be considered during grading for other reasonable solutions.

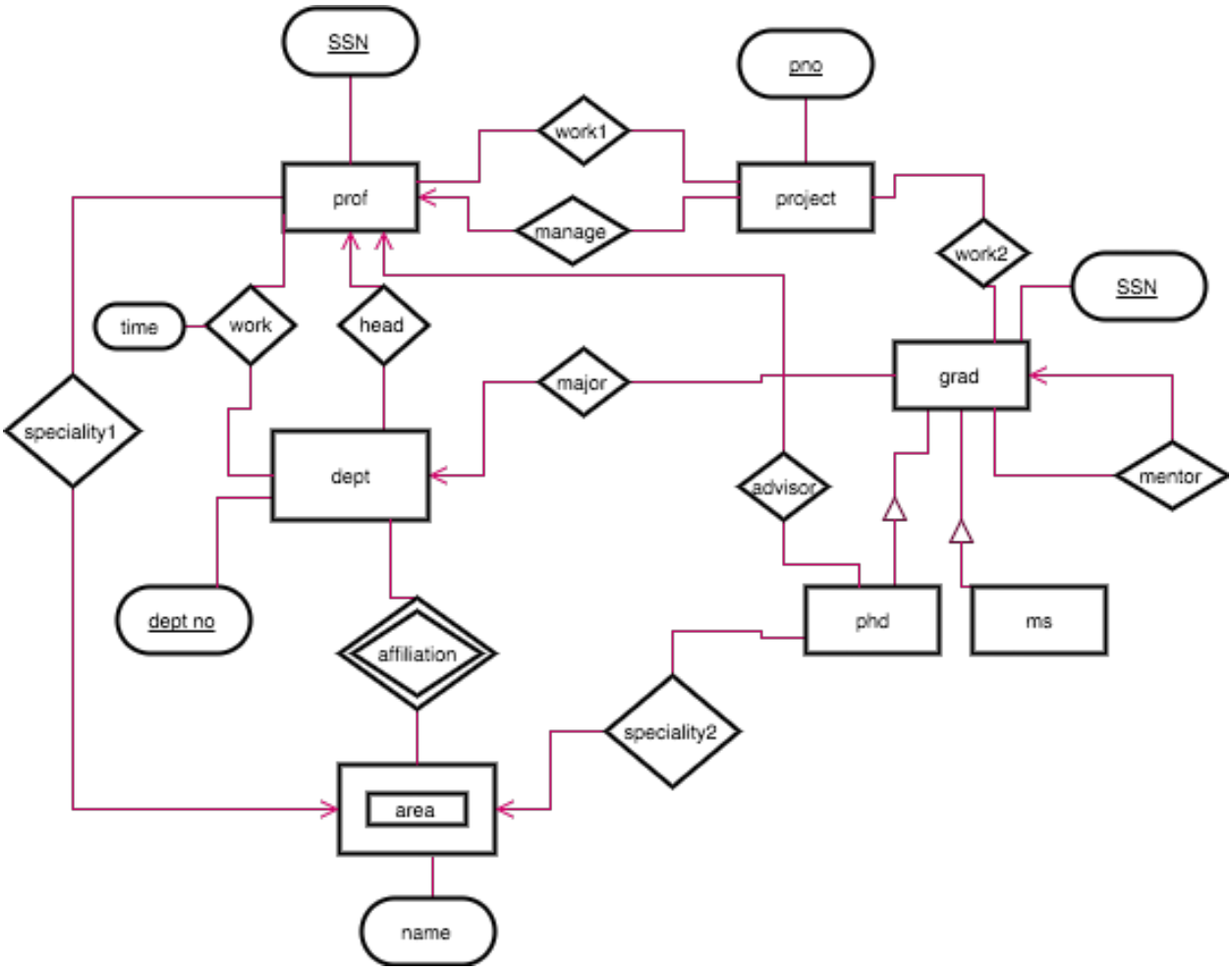

Figure 1: ER for Q2. $\rightarrow$ depicts the *exactly one* integrity constraint

Question 3. Relational Model (45 points)

1. Convert the ER model from the previous question to a relational model.

   Solution: The relational model is as follows:

   ```
   Department(Number, Name)
   ResearchArea(DepartmentNumber, Name)
   Professor(SSN, Name, Age, Rank)
   Speciality1(DepartmentNumber, ResearchAreaName, ProfessorSSN)
   Project(Number, Sponsor, StartDate, EndDate, Budget)
   Work1(ProfessorSSN, ProjectNumber)
   Manage(ProfessorSSN, ProjectNumber)
   GradStudent(SSN, Name, Age)
   Major(DepartmentNumber, GradStudentSSN)
   MS(GradStudentSSN)
   PhD(GradStudentSSN)
   Speciality2(GradStudentSSN, DepartmentNumber, ResearchAreaName)
   Advisor(GradStudentSSN, ProfessorSSN)
   Work2(GradStudentSSN, ProjectNumber)
   Head(DepartmentNumber, ProfessorSSN)
   Work(ProfessorSSN, DepartmentNumber, Time)
   Mentor(MenteeGradStudentSSN, MentorGradStudentSSN)
   ```

2. Which approach did you use to convert the subclass entity set? Show us alternative schema designs.

   Solution:

   (a) ER approach

   ```
   GradStudent(SSN, Name, Age)
   MS(GradStudentSSN)
   PhD(GradStudentSSN)
   ```

   (b) Object Oriented Approach

   ```
   GradStudentMS(SSN, Name, Age)
   GradStudentPhD(SSN, Name, Age)
   ```

   (c) Null Approach

   ```
   GradStudent(SSN, Name, Age, IsPhD)
   ```

3. Compare between all the designs you came up with in part 2. Talk about scenarios when each alternative would be a better choice in comparison to the rest.

   Solution: NULL approach won't work well when a query needs only one of the two sub-classes. A typicaly example is when we need the SSNs of all MS students. That being said NULL approach minimizes the number of tables which is useful.

The object oriented approach minimizes space and avoids redundancy. The ER approach doesn't seem suitable for the setting in this question due to the data duplication. For more details, refer section 4.6.4 from the textbook.