CSCI 585 – Database Systems
Prof. Dennis McLeod

Sample Midterm Exam

The purpose of this sample exam is to show the style of questions the exams will contain. It is not an indication of scope of topic coverage per se. We do not provide solutions for the sample exam; you can use it as a study tool to prepare for the exams this semester, and of course ask questions regarding any issues that arise. But, the CSCI585 staff will not solve the problems for you - remember, the sample is a study tool for you.

- The duration of this exam is 80 minutes.
- Do not start the exam until you are instructed to do so.
- If you solve your exam using pencil or erasable pen, you forfeit your right to a re-grading of your exam.
- This is a closed book exam. You are only allowed one 8x11 inch sheet of paper for notes.
- You are not allowed to share notes or any other materials with other students during the exam.
- The use of electronic equipment is prohibited during the exam. All laptops, netbooks, PDAs, cell phones, etc are to be switched off.
- In case you require clarification for an exam problem, note that you are allowed to ask questions only during the first 45 minutes of the exam.
- Follow the instructions of the exam proctors at all times.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Maximum</th>
<th>Received</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem 1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Problem 2</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Problem 3</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Problem 4</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
Problem 1 (20 points)

For each concept on the left, identify the best matching phrase in the right column. Indicate the number of best matching phrase. (2 points each)

a. ________ Projection of a relational attribute
   1. Unit of information
   2. A framework for facts

b. ________ RDF
   3. Multiple inheritance
   4. Subclasses don't share members

c. ________ Transaction
   5. Universal quantification

---

d. ________ Underlying domain of a relational attribute
   6. Set of current values
   7. Set of relational (attribute, value) pairs

---

e. ________ BCNF
   8. Subject-predicate-object expression
   9. Set of possible values

---

f. ________ Tuple
   10. Foreign key

---

g. ________ Physical data independence
   11. Polymorphic type system
   12. Separation of logical and external levels

---

h. ________ Disjoint subclasses
   13. Ensures that conceptual schema does not have to change when physical storage is reorganized

---

i. ________ Conceptual schema
   14. One fact in one place

---

j. ________ Referential integrity constraint
   15. Conceptual unit of manipulation
Problem 2 (15 points)

a. Which of these constitutes a minimal set of attributes for which it holds that: no two distinct tuples in the relation can have the same values for the attributes in this set? (3 points)
   i. Candidate key
   ii. Superkey
   iii. Foreign key
   iv. Primary key

1. i only
2. iv only
3. i and iv
4. i, ii and iv
5. i, iii and iv
6. i, ii, iii and iv

b. A relation in BCNF (Boyce-Codd Normal Form): (3 points)
   1. Allows partial dependencies on a key
   2. Is an enforcement mechanism for entity integrity
   3. Has no transitive functional dependencies
   4. All of the above
   5. None of the above
   6. Both 1 and 2
   7. Both 1 and 3

   (   )

c. Rank the following from the least expressive to the most expressive: (4 points)
   EER model, OWL-DL, “pure” relational model

   (   )

d. List a minimal set of relational algebra operators needed to satisfy relational completeness. (5 points)
Problem 3 (30 points)

a. Consider the following CIOM diagram:

Describe what the members of X, Y and Z represent. List the attributes of X, Y and Z. (15 points)
b. Consider the following CIOM diagram:

Show two different ways in which this CIOM diagram can be reduced to a relational schema. Discuss briefly what assumptions or conditions are necessary for each representation to be considered a valid mapping from the CIOM diagram given above. (15 points)
Problem 4 (35 points)

Consider the following relational database conceptual schema used to keep track of patient visits in a medical clinic:

Person (SSN, Name, DoB, Gender, PlaceOfBirth)  
[SS] [S] [D] [S] [S]

Doctor (SSN, Specialty)  
[N] [S]

Patient (SSN, InsuranceID, InsuranceProvider, MedicalRecordID)  
[N] [N] [S] [N]

MedicalVisit (Date, Doctor, Patient, Diagnosis)  
[D] [N] [N] [S]

In the above, primary keys are underlined. The domain of each attribute is stated using the following convention: Numbers [N], Strings [S] and Dates [D]. The following referential integrity constraints exist:

- From Doctor.SSN to Person.SSN
- From Patient.SSN to Person.SSN
- From MedicalVisit.Doctor to Doctor.SSN
- From MedicalVisit.Patient to Patient.SSN

Answer the following questions based on the above schema.

a. In the MedicalVisit table the primary key is defined as (Date, Doctor, Patient). If the primary key were to be changed to (Doctor, Patient), how would it affect the design of the database? (5 points)
b. Let the number of tuples in |Person| = x, the number of tuples in |Doctor| = y and the number of tuples in |Patient| = z. How are the sizes of x, y and z related to one another. Briefly justify. (5 points)

c. Write the following queries in relational algebra:

i. Find the name and gender of all doctors whose specialty is “dentistry”. (5 points)

ii. Find the names of insurance providers who insure patient(s) who were born in “Los Angeles” (5 points)
d. Solve the following queries using SQL. You may express your answers in either basic SQL as presented in lectures or ANSI SQL.

i. List all specialties available in the clinic. (5 points)

ii. Find the number of patients who visited a pediatrician on October 21, 2010. A pediatrician is a doctor whose specialty is “pediatrics” (children’s health). (5 points)

iii. List the SSN and name of all infant patients. An infant is defined as someone who was born on or after October 21, 2008. (5 points)