Name	**NetID**

CS411: Database Systems **Spring 2017**

Midterm, March 6

- READ THESE INSTRUCTIONS CAREFULLY BEFORE YOU START. DO NOT turn this page UNTIL we instruct you to.
- First: write your name and NetID at the top of this sheet.
- The exam lasts for 75 minutes, i.e., from 8–9.15am.
- We will not answer any questions during the exam. If you need to make any assumptions for any of the questions, please feel free to do so and then clarify the assumption in your answer.
- All questions are compulsory.
- For each question, please answer in the space provided; if you need more space, feel free to use the back side of each page. We will not provide additional space.
- The maximum score you can obtain is $18 + 20 + 10 + 25 + 27 = 100$.
- You must stop writing when time is called.
- **Cheating: No.**

Question	1	2	3	4	5	Total
Points						

Short Questions - 18 points

1. [3] Express the set intersection operator ($\cap$), between two relations R_1 and R_2, using the set difference operator ($-$).

2. [5] Consider a relation $R(A, B, C, D)$. Let the following three functional dependencies hold on R:

   - $AB \to C$
   - $BC \to D$
   - $D \to A$

   For each subset S in the following, give the closure S^+ based on the functional dependencies.

   i. BD

   ii. CD

3. [5] Consider a relation $R(A, B, C)$, and the following relational algebra expression:

$$\sigma_{Y \neq V}(\rho_{R(X,Y,Z)} R \bowtie \rho_{R(X,V,W)} R)$$

If the result of the above query is guaranteed to be empty for all instances of R, concisely provide a property of R that you can ascertain.

4. [5] Consider the ER diagram as provided in Figure 1. For each of the following queries, describe which approach of translating subclass entities would work best, and provide the appropriate reasoning.

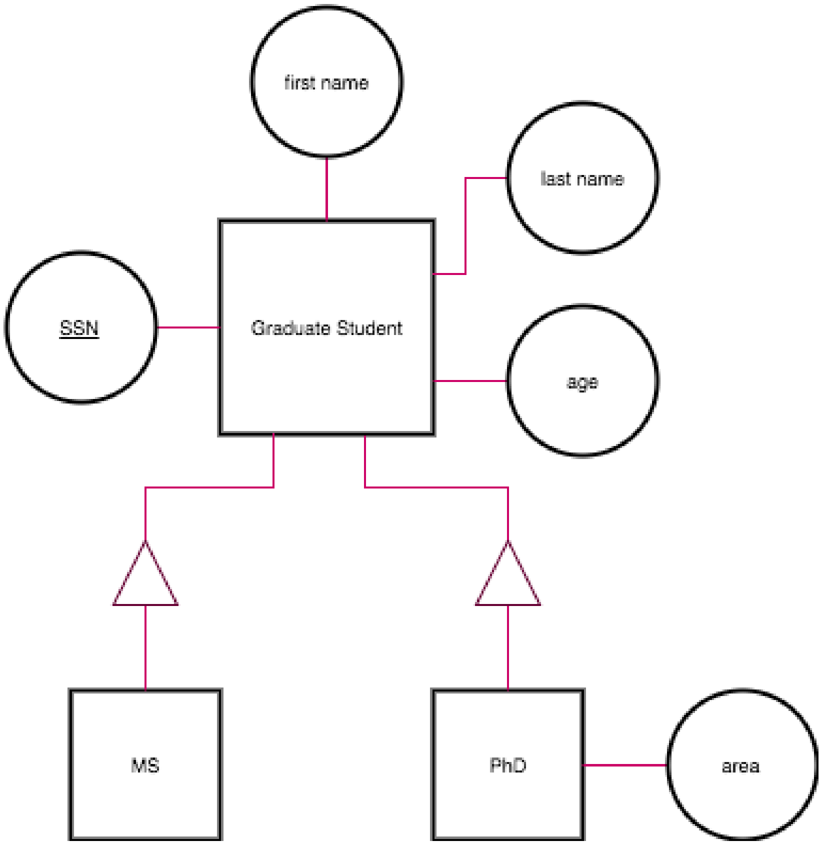

Figure 1: ER Diagram for question 4

i. Find all PhD students whose last name starts with ‘P’.

ii. Find all graduate students older than 25 years.

Database Design - 20 points

1. [10] Construct an Entity-Relationship diagram for a database maintained by the University of Illinois registrar's office. The data should include the following information:

   - Students (such as their UIN, net ID and name)
   - Instructors (such as their ID, name, email, and office)
   - Courses (such as their course number, department, title, syllabus, and credit hours)
   - Course Offering (such as semester, year, time, room, section)

   The data should also include which students are enrolled in which courses and the grade they are receiving, which courses are being taught by which instructors and information about each course offering.

   Do not forget to underline key attributes and to illustrate the mapping constraints on relationship sets (i.e., whether the relationships are many-to-many, many-to-one, one-to-many, or one-to-one). Please state any assumptions you make about the real world; for example, you may or may not choose to assume that each instructor teaches at most one course.

2. [10] Convert the above ER Diagram into a relational schema. Merge relations where appropriate. Make sure to underline key/s in each relation.

Armstrong's Axioms - 10 points

The three Armstrong axioms for your reference are as follows:

- Reflexivity rule

$$A_1A_2\dots A_n \rightarrow \text{a subset of } A_1A_2\dots A_n$$

- Augmentation rule:

$$\begin{aligned} \text{if } & A_1A_2\dots A_n \rightarrow B_1B_2\dots B_m \\ \text{then } & A_1A_2\dots A_nC_1C_2\dots C_k \rightarrow B_1B_2\dots B_mC_1C_2\dots C_k \end{aligned}$$

- Transitivity rule:

$$\begin{aligned} \text{if } & A_1A_2\dots A_n \rightarrow B_1B_2\dots B_m \\ \text{and } & B_1B_2\dots B_m \rightarrow C_1C_2\dots C_k \\ \text{then } & A_1A_2\dots A_n \rightarrow C_1C_2\dots C_k \end{aligned}$$

Derive the following rules using the three basic Armstrong's axioms:

1. [4] If $X \rightarrow YZ$ and $WY \rightarrow Z$, then $WX \rightarrow Z$

2. [6] If $X \to Y, YZ \to AB$, and $XBC \to W$, then $XZC \to AW$

Normal Forms - 25 points

Consider a relation $R(F, T, D, N, S)$ with the following functional dependencies:

i. $F \rightarrow D$

ii. $D, T \rightarrow F$

iii. $F, N \rightarrow S$

1. [5] Find two different minimal keys for this relation.

2. [8] Show two different BCNF decompositions for the relation R.

3. [4] Is any of the two decompositions above in (2) dependency preserving? If yes explain why, if no provide the dependency thats violated.

4. [8] Decompose this relation into a collection of relations that are in 3NF.

Relational Algebra and SQL - 27 points

Consider the relational schema for a banking portal which combines both the debit and credit history of all its customers (primary keys are underlined).

```
BankBranch(branchId, city)
Customer(custId, name, city)
Account(acId,branchId, creditPoints, balance)
Debt(debtId,debtAmount)
Borrower(custId, debtId)
Depositor(custId, acId)
```

Specify a relational algebra expresssion for each of the following queries.

1. [7] Find the credit points of all customers who have an account at branchId $CU411$ and have borrowed money.

2. [9] Let us denote the expression from part (1) as $R(A)$. Now use $R(A)$ to find the maximum credit points among all the customers.

Specify the following queries in SQL.

1. [6] The goal is to find the `custId` of all customers who are both borrowers and depositors. Complete the following SQL fragment by filling in the blanks.

   Note: please ONLY fill in the blanks, do not attempt to modify any part of the query fragment prior to the `WHERE` clause.

   ```
   SELECT DISTINCT custId
   FROM Customer
   WHERE

   ------------------------------------------------

   ------------------------------------------------

   ------------------------------------------------

   ------------------------------------------------
   ```

2. [5] Find the `custId` of all customers who have incurred a debt at branch $CU411$.