CS411: Database Systems **Spring 2017**

Midterm 1 Solutions

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 ($-$).

   Answer: $R_1 \cap R_2 = R_1 - (R_1 - R_2)$

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

   Answer: i. $ABCD$, ii. ACD

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.

   Answer: $A \to B$

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.

   i. Find all PhD students whose last name starts with 'P'.
   ii. Find all graduate students older than 25 years.

   Answer:

   i. OO works best. ER: multiple tables to look at. Null: more data to look at.
   ii. ER works best. OO: multiple tables to look at. Null: more data to look at.

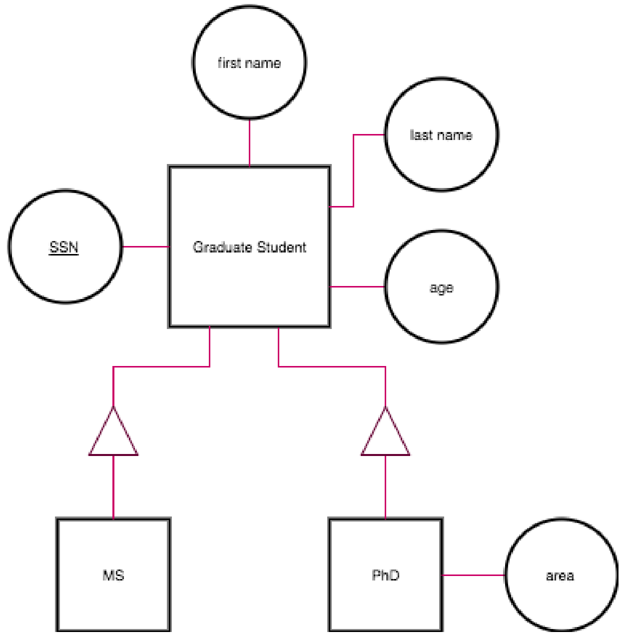

Figure 1: ER Diagram for Graduate Students

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, netID and name)
   - Instructors (such as their ID, name, email, and office)
   - Courses (such as their courseNo., 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.

   Answer: See Figure 2. Note that this is not the only correct answer.

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

   Answer: The following relational schema corresponds to Figure 2.

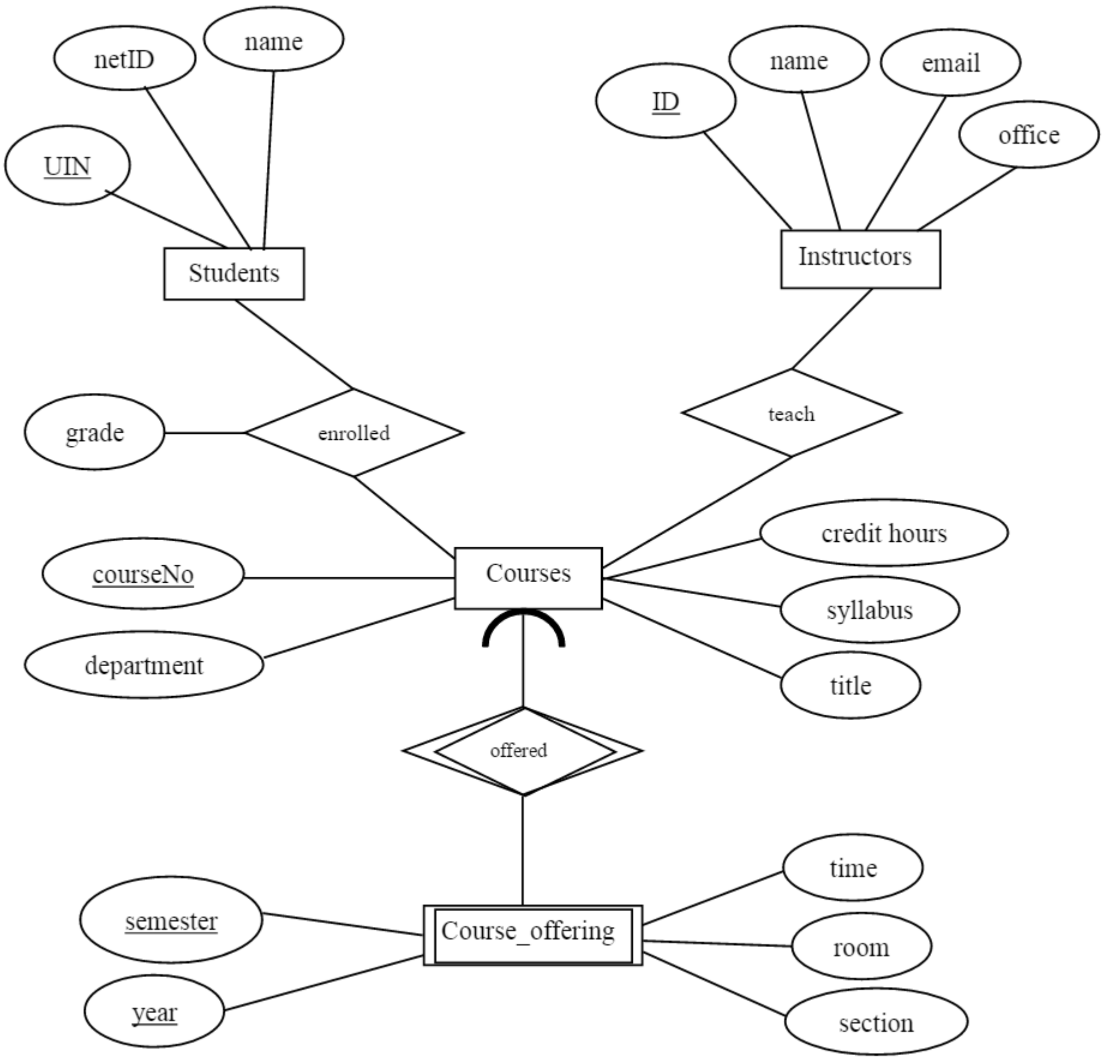

Figure 2: ER Diagram for University Database

```
Student(UIN, netID, name)
enrolled(UIN,courseNo, grade)
Courses(courseNo, department, title, syllabus, credit hours)
Instructors(ID, email, name, office)
teach(ID,courseNo)
CourseOffering(semester, year, time, room, section, courseNo).
```

Armstrong's Axioms - 10 points

The three Armstrong axioms are as follows:

- Reflexivity rule

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

- Augmentation rule:

  if $A_1A_2\dots A_n \rightarrow B_1B_2\dots B_m,$ then $A_1A_2\dots A_nC_1C_2\dots C_k \rightarrow B_1B_2\dots B_mC_1C_2\dots C_k$

- Transitivity rule:

  if $A_1A_2\dots A_n \rightarrow B_1B_2\dots B_m,$ and $B_1B_2\dots B_m \rightarrow C_1C_2\dots C_k,$ then $A_1A_2\dots A_n \rightarrow C_1C_2\dots C_k$

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

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

   Answer:

   $X \rightarrow YZ$ (given)
   $YZ \rightarrow Y$ (reflexivity rule)
   $X \rightarrow Y$ (transitivity rule)
   $WX \rightarrow WY$ (augmentation rule)
   $WY \rightarrow Z$ (given)
   $\therefore WX \rightarrow Z$ (transitivity rule)

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

   Answer:

   $X \rightarrow Y$ (given)
   $XZ \rightarrow YZ$ (augmentation rule)
   $YZ \rightarrow AB$ (given)
   $XZ \rightarrow AB$ (transitivity rule)
   $AB \rightarrow A$ (reflexivity rule)
   $XZ \rightarrow A$ (transitivity rule)
   same goes for $XZ \rightarrow B$
   $XZ \rightarrow XB$ (augmentation rule)
   $XZC \rightarrow XBC$ (augmentation rule)

$XBC \to W$ (given)
$XZC \to W$ (transitivity rule)
$XZC \to XZW$ (augmentation rule)
From above, we had $XZ \to A$
$XZW \to AW$ (augmentation rule)
$\therefore XZC \to AW$ (transitivity rule)

Alternate answer:

$X \to Y$ (given)
$XZ \to YZ$ (augmentation rule)
$YZ \to AB$ (given)
$XZ \to AB$ (transitivity rule)
$XZC \to ABCX$ (augmentation rule)
$XBC \to W$ (given)
$ABCX \to AW$ (augmentation rule)
$\therefore XZC \to AW$ (transitivity rule)

Normal Forms - 25 points

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

i. $F \to D$

ii. $DT \to F$

iii. $FN \to S$

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

   Answer: $\{F, T, N\}$ and $\{D, T, N\}$.

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

   Answer:

   Picking i first, we get: $(F, D), (T, N, F), (F, N, S)$

   Picking ii first, we get: $(D, F), (F, T), (D, T, N, S)$

   Picking iii first, we get: $(F, D), (T, N, F), (F, N, S)$

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

   Answer: First—and equivalently, third—decomposition is not dependency preserving because it violates the functional dependency $DT \to F$.

   Second decomposition is not dependency preserving because it violates the functional dependency $FN \to S$, and $DT \to F$.

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

   Answer: $(F, D), (D, T, F), (F, N, S), (F, T, N)$ or $(F, D), (D, T, F), (F, N, S), (D, T, N)$

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.

   Answer: $\Pi_{creditPoints}(\sigma_{branchId=``CU411''}((account \bowtie depositor) \bowtie borrower))$

2. [9] Let us denote the expression from part (1) as $R(A)$. Now use $R(A)$ find the maximum credit points among all the customers.
   Answer: $R - \Pi_{R(A)}(\sigma_{A_1}(\Pi_{R(A_1)} R \bowtie_{A_1 < A_2} \Pi_{R(A_2)} R))$

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

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

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

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

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

   Answer:

```
SELECT DISTINCT custId
FROM Customer
WHERE
custId IN (SELECT custId
        FROM Borrower, Depositor
        WHERE Borrower.custId = Depositor.custId)
```

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

   Answer:

```
SELECT DISTINCT custId
FROM Customer, Depositor, Account
WHERE
Customer.custId = Depositor.custId
AND
Depositor.acId = Account.acId
AND
Account.branchId = ``CU411''
```