CS145 Midterm Examination

Spring 2001, Prof. Widom

- Please read all instructions (including these) carefully.
- There are 8 problems on the exam, with a varying number of points for each problem and subproblem for a total of 75 points. *You should look through the entire exam before getting started, in order to plan your strategy.*
- The exam is closed book and closed notes, but you may refer to your three pages of prepared notes.
- Please write your solutions in the spaces provided on the exam. Make sure your solutions are neat and clearly marked. You may use the blank areas and backs of the exam pages for scratch work. Please do not use any additional scratch paper.
- *Simplicity and clarity of solutions will count.* You may get as few as 0 points for a problem if your solution is far more complicated than necessary, or if we cannot understand your solution.

NAME: ____

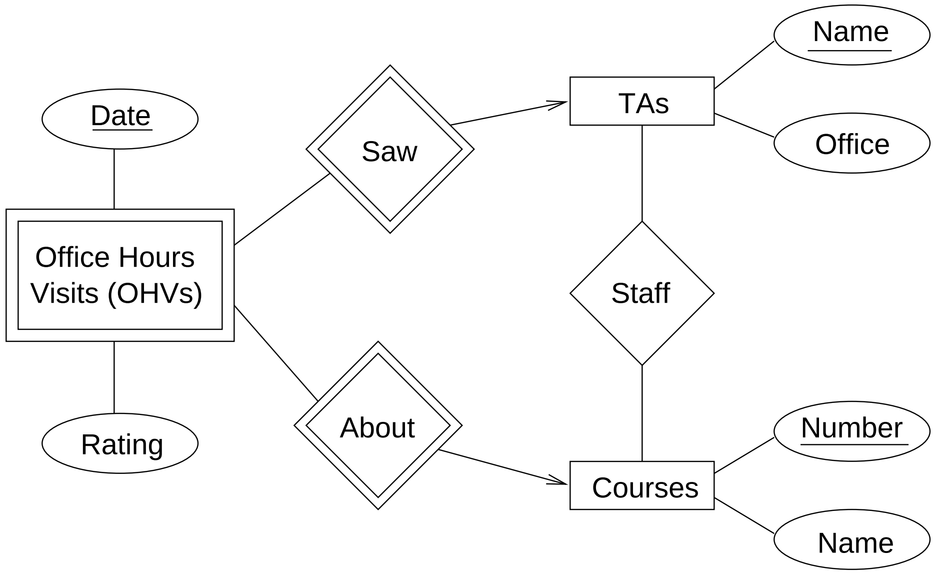
In accordance with both the letter and spirit of the Honor Code, I have neither given nor received assistance on this examination.

SIGNATURE: ____

Problem	1	2	3	4	5	6	7	8	TOTAL
Max. points	8	6	6	5	10	4	8	28	75
Points									

1. **E/R Diagrams to Relations** (8 points)

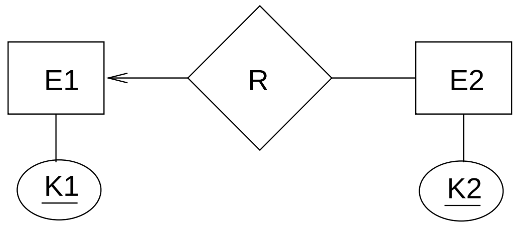
   Consider the following entity-relationship (E/R) diagram, which specifies courses, TAs for courses, and office hours visits to TAs including a rating for the quality of each visit.



   Using the algorithm described in class and in the textbook, translate the E/R diagram to a relational schema. Do not make any assumptions about the real world except those encoded in the E/R diagram. Be sure to specify (underline) keys for all relations. List your set of relations here:

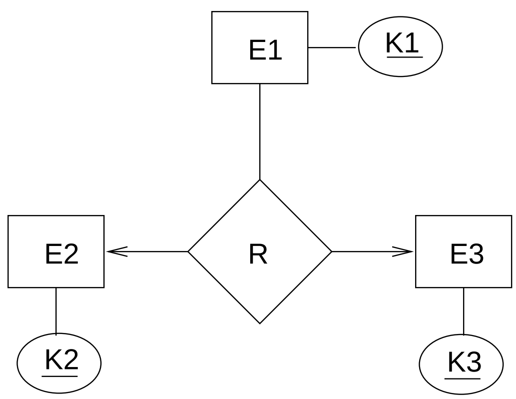
2. **E/R Diagrams and Functional Dependencies** (6 points)

(a) (2 points) Consider the following entity-relationship (E/R) diagram:



and the relation $R(K1, K2)$ for the relationship set. List all completely nontrivial functional dependencies that hold on R:

(b) (4 points) Consider the following E/R diagram:



and the relation $R(K1, K2, K3)$ for the relationship set. List all completely nontrivial functional dependencies that hold on R:

3. **Functional Dependencies** (6 points)

Consider a relation $R(A, B, C, D, E, F)$ with the following set of functional dependencies over R:

$$\begin{aligned} A &\rightarrow C \\ DE &\rightarrow F \\ B &\rightarrow D \end{aligned}$$

(a) (3 points) Based on these functional dependencies, there is one key for R.

What is it? $\boxed{}$

(b) (3 points) Add to the above set of functional dependencies the dependency $A \rightarrow B$. Now suppose we want A to be a key. Name one more functional dependency that, if added to the set, makes A a key. As an additional restriction, the new functional dependency must have only one attribute on the left-hand side and only one attribute on the right-hand side.

Write the dependency here: $\boxed{}$

4. **Functional Dependencies and Relational Algebra** (5 points)

Consider a relation $R(A, B, C)$. Write a relational algebra expression that always returns empty if and only if the functional dependency $A \rightarrow B$ holds on R. Mark your answer clearly, and remember that you will be graded on simplicity as well as correctness.

5. **BCNF and 4NF** (10 points)

On this problem you will get 2 points for each correct answer, -1 points for each incorrect answer, and 0 points for each answer left blank.

We use the following two relational schemas:

Schema 1: $R(A, B, C, D)$

Schema 2: $R_1(A, B, C)$, $R_2(B, D)$

(a) Consider Schema 1 and suppose that the only functional dependencies that hold on the relations in this schema are $A \to B$, $C \to D$, and all dependencies that follow from these. Is Schema 1 in Boyce-Codd Normal Form (BCNF)?

Circle one: YES NO

(b) Consider Schema 2 and suppose that the only functional dependencies that hold on the relations in this schema are $A \to B$, $A \to C$, $B \to A$, $A \to D$, and all dependencies that follow from these. Is Schema 2 in BCNF?

Circle one: YES NO

(c) Suppose we omit dependency $A \to D$ from part (b). Is Schema 2 in BCNF?

Circle one: YES NO

(d) Consider Schema 1 and suppose that the only functional and multivalued dependencies that hold on the relations in this schema are $A \to BC$, $B \to D$, $B \twoheadrightarrow CD$, and all dependencies that follow from these. Is Schema 1 in Fourth Normal Form (4NF)?

Circle one: YES NO

(e) Consider Schema 2 and suppose that the only functional and multivalued dependencies that hold on the relations in this schema are $A \to BD$, $D \to C$, $C \twoheadrightarrow AB$, $B \twoheadrightarrow D$, and all dependencies that follow from these. Is Schema 2 in 4NF?

Circle one: YES NO

6. **Relational Algebra and SQL** (4 points)

On this problem you will get 2 points for each correct answer, -1 points for each incorrect answer, and 0 points for each answer left blank.

Consider a relation $R(A, B, C)$ and the following two statements:

Statement 1: A is a key for R.

Statement 2: Functional dependency $A \rightarrow BC$ holds on R.

(a) Are these two statements equivalent in relational algebra?

Circle one: YES NO

(b) Are these two statements equivalent in SQL?

Circle one: YES NO

7. **SQL** (8 points)

Consider a relation `Scores(ID,name,score)` for recording student scores on this midterm. Attribute `ID` is a key. You are to write a query in SQL that returns the frequency distribution of scores, in descending order by score. For example, the output might look like:

score	number
74	6
73	2
70	1
69	15
...	...

The first column records the score on the exam, and the second column records the number of students who received that score. Do not include rows for scores where no students received that score. Delineate your answer clearly, and remember that you will be graded on simplicity as well as correctness.

8. **Query Equivalences** (28 points)

On this problem you will get 2 points for each correct answer, -1 points for each incorrect answer, and 0 points for each answer left blank.

Each row of the following table shows two queries. In the blank third column of the table write "YES" if the two queries are equivalent, and "NO" if they are not equivalent. Remember that two queries are equivalent if they always return exactly the same answer on all databases.

All queries refer to a schema containing two relations:

$R(A, B)$ where A is a key and B is a key
$S(A, B)$ where A is a key

You may assume that the relations do not contain NULL values, but do not make any other assumptions about the relations.

Query 1	Query 2	Equiv.?
$\pi_A(R - S)$	$\pi_A(R) - \pi_A(S)$	
$\pi_B(R - S)$	$\pi_B(R) - \pi_B(S)$	
$\pi_B(R \cup S)$	$\pi_B(R) \cup \pi_B(S)$	
$\sigma_{R.A=5}(R)$	$\sigma_{R.A=5}(\pi_{R.A,R.B}(R \bowtie S))$	
$\pi_{R.A}(\sigma_{R.B=S.B}(R \times S))$	`select R.A from R,S where R.B=S.B`	
$\pi_{R.B}(\sigma_{R.A=S.A}(R \times S))$	`select R.B from R,S where R.A=S.A`	
$\pi_B(R) -$ $\pi_{R1.B}(\sigma_{R1.B \geq R2.B}(\rho_{R1}(R) \times \rho_{R2}(R)))$	`select min(B) from R`	
$\pi_B(R) -$ $\pi_{R1.B}(\sigma_{R1.B < R2.B}(\rho_{R1}(R) \times \rho_{R2}(R)))$	`select max(B) from R`	
`select R.B from R,S where R.A=S.A`	`select B from R` `where A in (select A from S)`	
`select B from R` `where A not in (select A from S)`	`select R.B from R,S where R.A<>S.A`	
`select A from R`	`select A from R group by A`	
`select B from S`	`select B from S group by B`	
`select B from S group by B`	`select distinct B from S`	

`(select B from S)`
`except all`
`(select B from R)`
	`select B from S`
	`where B not in (select B from R)`	