DB

Stanford University Computer Science Department
2000 Comprehensive Exam in Databases

- The exam is *open book and notes.*
- Answer all 6 questions on the exam paper itself, in the space provided for each question.
- The total number of points is 60; questions may have differing point values.
- You have 60 minutes to complete the exam (i.e., one minute per point). It is suggested you review the entire exam first, in order to plan your strategy.
- *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.

Provide your magic number here: ______________

1	2	3	4
5	6		

A. The Honor Code is an undertaking of the students, individually and collectively:

(1) that they will not give or receive aid in examinations; that they will not give or receive unpermitted aid in class work, in the preparation of reports, or in any other work that is to be used by the instructor as the basis of grading;

(2) that they will do their share and take an active part in seeing to it that others as well as themselves uphold the spirit and letter of the Honor Code.

B. The faculty on its part manifests its confidence in the honor of its students by refraining from proctoring examinations and from taking unusual and unreasonable precautions to prevent the forms of dishonesty mentioned above. The faculty will also avoid, as far as practicable, academic procedures that create temptations to violate the Honor Code.

C. While the faculty alone has the right and obligation to set academic requirements, the students and faculty will work together to establish optimal conditions for honorable academic work.

By writing my "magic number" below, I certify that I acknowledge and accept the Honor Code.

Problem 1: (10 points) Suppose relation $R(A, B, C, D, E)$ has functional dependencies

$A \rightarrow B$
$BC \rightarrow D$
$BE \rightarrow C$
$AD \rightarrow E$
$CE \rightarrow A$

a) What are all the keys of R?

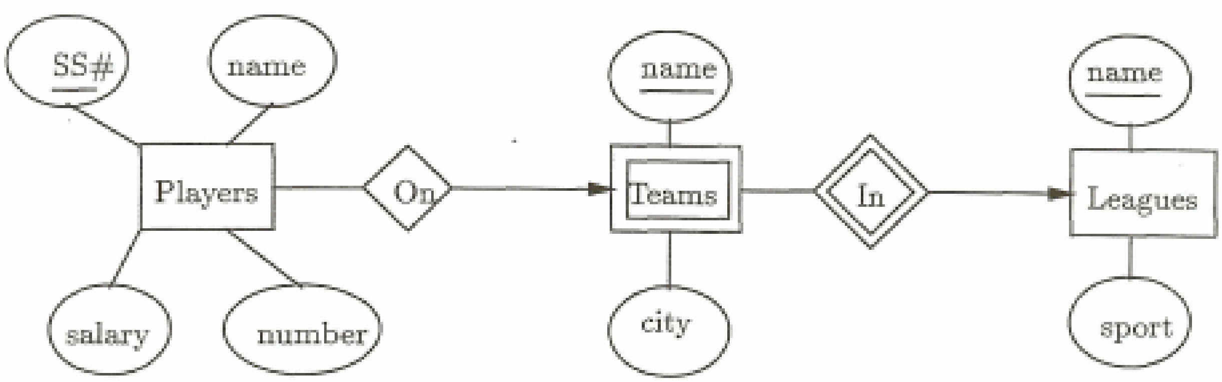
---

b) Of the five given FD's, which violate the BCNF condition?

---

c) Of the five given FD's, which violate the 3NF condition?

Problem 2: (10 points) Below is an entity-relationship diagram:



a) Choose a relational database schema that represents this E/R diagram as faithfully as possible. Do not use a relation if its information is sure to be contained in some other relation of your schema.

---

b) Suppose that we delete the SS# attribute from *Players*. Exploit the fact that a team will not give the same number to two players in order to find a similar E/R diagram with a suitable key for *Players*. Draw your diagram below:

c) Convert your E/R diagram from (b) to an appropriate relational database schema.

Problem 3: (10 points) A database system, containing of two objects A and B executes three transactions: T_1, T_2 and T_3. Initially A has a value of 10, and B a value of 20.

T_1: First writes a value of 30 for A; then changes B to 40. Finally, T_1 commits. T_1 runs with isolation level `SERIALIZABLE`.

T_2: Starts by changing A to 50, then modifies B to 60. At this point T_2 does a rollback, and all changes are undone. T_2 runs with isolation level `SERIALIZABLE`.

T_3: Is a read-only transaction that first reads A and then reads B. The isolation level of T_3 is discussed below.

We do *not* know in what order these three transactions execute.

a) Assume that T_3 runs with isolation level `SERIALIZABLE`. What are all the possible A, B values that T_3 can read? Give each answer as a pair $[X, Y]$, where X is the A value read, and Y is the B value read by T_3.

---

b) Assume that T_3 runs with isolation level READ COMMITTED. What additional A, B values can T_3 read? [Do not list pairs given in part (a).]

---

c) Assume that T_3 runs with isolation level READ UNCOMMITTED. What additional A, B values can T_3 read? [Do not list pairs given in parts (a) or (b).]

Problem 4: (5 points) Suppose we have a relation with schema $ABCD$, the functional dependency $A \rightarrow B$ and the multivalued dependency $A \rightarrow\rightarrow C$. Give five other nontrivial multivalued dependencies that R must satisfy (i.e., MVD's that follow logically from the two given dependencies). Note that for a MVD to be nontrivial, its left and right sides must have no attribute in common, and there must be some attribute that is in neither the left nor right side.

Problem 5: (20 points) Consider the following relation:

```
Advised(Advisor, Student, Year)
```

A tuple (A, S, Y) in `Advised` specifies that advisor A advised student S who graduated in year Y. Assume that `Student` is a key for this relation.

a) Consider the following SQL query, which finds all advisors who advised a student who graduated in the same year that Hector Garcia-Molina or Jennifer Widom graduated.

```
SELECT Advisor
FROM Advised
WHERE Year IN
    (SELECT Year FROM Advised
    WHERE Student = 'Hector Garcia-Molina'
        OR Student = 'Jennifer Widom')
```

Write an SQL query, *without any subqueries* and *without the keyword* DISTINCT, that always produces the same set of tuples as the above query.

---

b) Are there are any circumstances in which your "equivalent" query can produce an answer different from that of the query above? Explain, if so.

c) Using SQL3 recursion, write a query that finds all "descendants" of Jeff Ullman, i.e., all students whose advisor was Jeff Ullman, or whose advisor's advisor was Jeff Ullman, or whose advisor's advisor's advisor was Jeff Ullman, and so on.

d) Write a SQL query that finds the advisor(s) with the longest advising span, i.e., with the longest period from their earliest advisee to their latest advisee.

Problem 6: (5 points) Two relations $A(x)$ and $B(y)$ each contain integers (i.e., their tuples have one component, which is an integer). Give `CREATE TABLE` statements for A and B with `CHECK` clauses sufficient to assure that $A \cap B$ will always be empty.