# CS4604 Final Exam

Dec 8, 2000

Please enter the following information:

- **Name:**

- **ID:**

GOOD LUCK and have a happy new year!!
Do not write below this line

---

| Problem | Max Score | Score |
|---|---|---|
| 1 | 30 | |
| 2 | 10 | |
| 3 | 10 | |
| 4 | 20 | |
| 5 | 10 | |
| 6 | 10 | |
| 7 | 10 | |
| 8 | 15 (XC) | |
| Total | 100 | |

1. (30 points) Short answer questions:

   (a) Give an example of an index data structure used in database systems.

   (b) What is the fundamental difference in how ODL and E/R designs treat multiple inheritance?

   (c) Which of the two normal forms 3NF and 4NF is stricter than BCNF?

   (d) Give two options of connecting a database with a programming language.

   (e) Carefully state the rule to determine if a given Datalog query is *safe*.

   (f) SQL allows only linear recursion. What does this mean?

   (g) (True/False) In the relation $R(X, Y, Z)$, if $X \rightarrow Y$, then $X \rightarrow\!\!\!\!\rightarrow Z$.

   (h) Patterns obtained by data mining typically do not have more than $1-2\%$ support. Explain why such patterns are nevertheless useful.

   (i) Under what conditions will $R \bowtie S = R \times S$?

   (j) What advice would a database tuning specialist give to a client who complains about rain leaking through the roof onto the dining table?

2. (10 points) Design an E/R diagram to model the following facts about leagues, teams, and players. League names are unique. No league has two teams with the same name. No team has two players with the same number. However, there can be players with the same number on different teams, and there can be teams with the same name in different leagues.

3. (10 points) Suppose you are given a relation $R(A, B, C, D)$ with FDs $\{AB \rightarrow C, C \rightarrow A, C \rightarrow D\}$. Assume that $R$ is decomposed into $S(A, C, D)$ and $T(B, C)$. Is this a good decomposition? Briefly explain why or why not.

4. (a) (12 points) Consider the relation `Advisor(student-name,advisor-name)`. Write a query in Datalog to find those advisors who advise exactly two students. Lets call these advisors `Two-Advisors`.

(b) (8 points) For the same relation as above, write a query in Datalog to find students whose advisors are not `Two-Advisors`. You are welcome to reuse the results of the query above.

5. (10 points) Consider the relational schema:

   ```
   HasTaken(student-ssn,course-number)
   RequiredForGraduation(course-number)
   ```

   with their obvious interpretations *i.e.,* `HasTaken` identifies the courses taken by each student and `RequiredForGraduation` lists all the courses required to graduate. For example, if `RequiredForGraduation` contains 14 courses, this means that all those 14 courses are required to graduate (you may assume that all the `course-numbers` in `RequiredForGraduation` will be different). Write a Datalog query to find the students who have not satisfied any of the requirements for graduation.

6. (10 points) Consider the relational schema:

   `OurPresidents(name,state-of-origin,number-of-terms-served)`

   with its obvious interpretation *i.e.,* `OurPresidents` identifies the names of presidents (e.g. 'Bill Clinton'), the state they are from (e.g., 'Arkansas') and the number of terms they served in office (e.g., 2). Write a query in relational algebra to find the name(s) of the president(s) who served the most number of terms.

7. (10 points) Consider the relation `Student(id,name,age,gpa)`. Write a query in SQL to find pairs of (`id`s of) students with the same name and age. A pair should be listed only once, *i.e.,* The `id`s $i$ and $j$ (if they have the same name and age) should be listed as *(i,j)* or as *(j,i)* but not as both.

8. (Extra credit question: 15 points) The *semi-join* of relations $R$ and $S$, denoted by $R \prec S$, is the set of tuples in $R$ that agree with at least one tuple of $S$ on all attributes that are common to the schemas of $R$ and $S$. Give two different expressions of relational algebra that are equivalent to $R \prec S$. *Note:* $R \prec S \neq S \prec R$. You are allowed to use only symbols from the set: $\{\cup, \cap, -, \sigma, \pi, \times, \bowtie_C, \bowtie\}$. Full credit will be given if both expressions are genuinely different, and not merely restatements of each other.