

**Comprehensive Exam**  
**Databases**  
**Fall 2004**

1) (25 points)

(a) What is the purpose of developing a data model such as an entity-relationship diagram?

(b) At what level of abstraction does data modeling take place – the logical level, the physical level, or the view level (circle one)?

(c) Once an entity-relationship diagram has been completed, what is it used for? What is typically done with an entity-relationship diagram after it has been completed?

2) (25 points)

(a) In general, does normalization improve the efficiency of a database, in terms of performance (yes or no)? Be sure to explain your answer.

(b) What is the purpose of normalization? What is it intended to achieve?

(c) Many modern database management systems allow one to declare a table column that is an array, i.e., *create table student (ss integer, name string(32), dob date, **hobbylist stringArray**)*. Could such a table be in 3NF (yes or no)? In addition, what are the implications of this type of column for the normalization process?

3) (25 points) Consider the following collection of relational schemes.

employee(employee-name, street, city)

works(employee-name, company-name, salary)

company(company-name, city)

manages(employee-name, company-name, manager-name)

Note that an employee can work for more than one company in the above scheme. Give a **tuple calculus** expression for each of the following.

(a) A list of the names of those companies located in Orlando.

(b) A list of employee names for those employees who live in the same city as a company for which they work.

(c) A list of employee names for those employees who work for every company located in Orlando.

4) (25 points) Consider the following collection of relational schemes.

employee(employee-name, street, city)  
works(employee-name, company-name, salary)  
company(company-name, city)  
manages(employee-name, company-name, manager-name)

Note that this is the same as the collection of relational schemes from the previous question. Give an SQL statement for each of the following.

(a) A list of the names of those employees who work for Oracle.

(b) A list of employee names and, for each employee, the largest salary they obtain from any single company. Note that the list should only include employees who live in Orlando, and the names should appear in sorted order.

(c) A list of the names of those managers who manage an employee at a company that pays the employee a salary greater than 50000. Note that any given manager should appear at most once in the resulting list.