CS145 Midterm Examination

Monday, November 1, 1999, 11AM–12:15PM

Directions

- The exam is *open book*; any written materials may be used.
- Answer all 4 questions on the exam paper itself.
- The total number of points is 75 (i.e., one point per minute of work).
- When SQL is called for, you *must* use standard SQL2, not the Oracle version.
- You will be graded not only on correctness of queries and other database code, but on the simplicity of your answers. You may get as little as 0 points for an answer that "works" but is excessively complicated.
- Do not forget to **sign the pledge** below.

I acknowledge and accept the honor code.

____

Print your name here: ____

1	2	3	4

Problem 1: (20 points) The first question is based on the following information:

1. Manufacturers have a name, which we may assume is unique, an address, and a phone number.

2. Products have a model number and a type (e.g., "television set"). Each product is made by one manufacturer, and different manufacturers may have different products with the same model number. However, you may assume that no manufacturer would have two products with the same model number.

3. Customers are identified by their unique social security number. They have email addresses, physical addresses. Several customers may live at the same (physical) address, but we assume that no two customers have the same email.

4. An order has a unique order number, and a date. An order is placed by one customer. For each order, there are one or more products ordered, and there is a quantity for each product on the order.

Do the following:

a) In the space below, draw an entity/relationship diagram that represents the above information. Indicate keys by underlining, as usual.

b) Draw on the next page a network schema for the same information. Remember that an LRT is not required to have a key, but underline key attributes, if the LRT has a key.

c) Give a relational database schema (just the relations and attributes — not an SQL schema declaration) for the same information. Indicate keys by underlining, and do not include obviously redundant relations.

d) Suppose we had a relation `Customer(ssNo,email,addr)` as part of our relational database schema. What functional dependencies would you expect to hold?

What are the keys for this relation?

Problem 2: (15 points) A bank has the following relations, whose meanings should be apparent. Note, however, that an account can have several owners (i.e., joint accounts are allowed).

```
Accounts(acctNo, type, balance)
Owns(acctNo, custNo)
Customers(custNo, name, addr)
```

Write the following in **relational algebra**. You may write complete expressions, or break your answer into several steps by assigning intermediate relations to temporaries.

a) Find the types of accounts held by customer "Lee."

b) Produce a table giving, for each type of account, the largest balance of any account of that type. For instance, if the result consists of the two tuples (checking, 2000) and (savings, 50000), then the largest checking account has a balance of 2000, and the largest savings account has a balance of 50000, and there are no other types of accounts.

c) Find all the "dangling accounts," that is, account numbers that appear in the `Accounts` relation, but either:

1. Do not appear in the `Owns` relation, or
2. Appear in the `Owns` relation, but every customer with which they are associated fails to appear in the `Customers` relation.

Problem 3: (30 points) Consider a relation `Busses` declared in SQL2 as follows:

```
CREATE TABLE Busses (
    line CHAR(10),
    busNo INTEGER,
    stopNo INTEGER,
    arrives TIME,
    place CHAR(20)
);
```

The meaning of a tuple such as `('Greyhound', 123, 4, 14:40:00, 'Palo Alto')` is that Greyhound bus number 123 has its fourth stop at 20 minutes before 3PM, at the Palo Alto bus station. Note that this database models "long distance" busses, and Greyhound bus 123 will run only once a day. If there are other busses that follow the same route, they have different numbers. Write the following in standard SQL2:

a) The line, bus number, and stop number together are a key for `Busses` (show what you would add to the `CREATE TABLE` declaration):

b) Write a tuple-based check that says there are no "time warps"; that is, if bus number b for line a arrives at its ith stop at time t, then it arrives at its $(i+1)$st stop (if there is one) at some time later than t. You may assume times can be compared by $<$, and no busses run at midnight (i.e., for times t_1 and t_2, we have $t_1 < t_2$ if and only if t_2 is later than t_1).

c) Write an SQL2 assertion that stop numbers are consecutive integers starting at 1; i.e., if there is a stop numbered i, for some $i > 1$, then there is a stop numbered $i - 1$. *Note*: we are not asking for the more complicated constraint that the stop numbers *for each bus* are consecutive integers — just that as a whole they are consecutive.

d) Could we get the same effect as (c) with a tuple-based check? ________ Why or why not?

e) Find all the "last stops," that is, the places that have the highest stop numbers for some line/bus-number combination.

f) Find all places such that at least three busses from the same line stop at that place between noon and 1PM.

Problem 4: (10 points) Suppose we have a relation $R(a, b, c, d, e)$ with the following functional dependencies: $ab \to c$, $cd \to e$, $c \to a$, and $e \to d$.

a) What are all the keys for R?

b) Give an example of a functional dependency that is a BCNF violation for R. ______
Into what two relations does this violation tell us to decompose R?