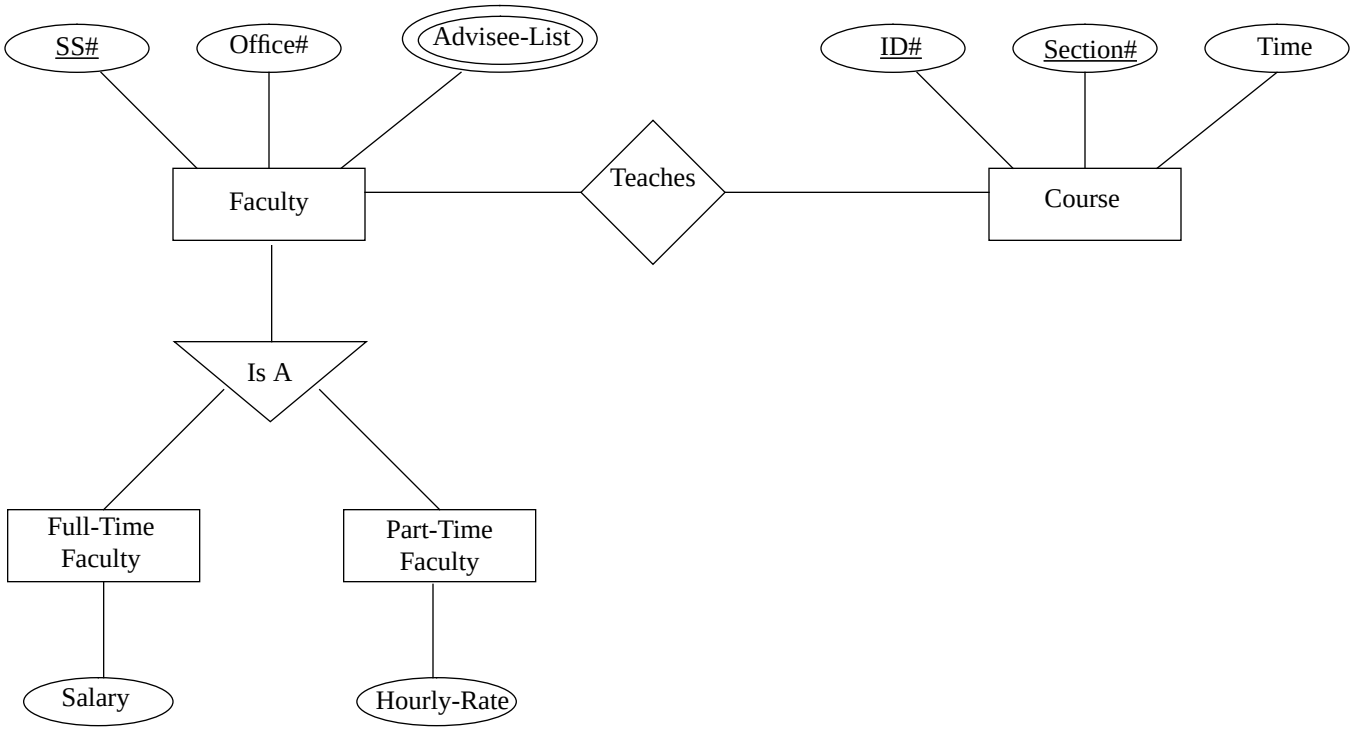
Comprehensive Exam
Databases
Spring 2000

1) (20 points) Consider the following ER diagram for representing information associated with faculty and courses at some university (note that Advisee-List is a multi-valued attribute).



Give a collection of relational schemes, i.e, tables, for the above ER diagram. For each relational scheme be sure to specify which attributes form the primary key.

2) (20 points) Suppose that a logical model is to be developed for the information associated with a medical clinic. In this clinic patients are allowed to make appointments with doctors in advance, or they are allowed to walk-in without an appointment. In either case, clinic requirements dictate that the date and time they are admitted, as well as the time they are checked out be tracked. Note that there are no over-night stays, but there can be multiple visits in the same day by the same patient. Also of great concern to the clinic is the policy number for each patients medical insurance. It is assumed that patients will have only one medical insurance policy, members of the same family may be listed on the same policy, and patients without insurance are *not* allowed to visit the clinic. Each patient is assigned to exactly one doctor the first time they are admitted to the clinic, and this doctor is the one who sees the patient on all future visits.

A log record is maintained for each medical test performed at the clinic, which contains the social security number of the patient on which the test was performed, the type of test, the date the test was performed, and the result of the test. Although multiple tests can be given to a patient on the same day, no two tests of the same type will be given to the same patient on the same day.

Finally, each doctor has a unique office number.

Note that the following entity sets could be identified from the above description:

- patient(ss#, name, policy#)
- visitation(ss#, date-admitted, time-admitted, time-checked-out)
- doctor(ss#, name, office#)
- log(ss#, test-type, date, result)

(a) What would be a candidate key for the *patient* entity set?

(b) What would be a candidate key for the *visitation* entity set?

(c) What would be a candidate key for the *log* entity set?

(d) List all of the possible super-keys for the *doctor* entity set.

3) (15 points) What is the purpose of normalization?

4) (10 points) List and explain two (2) advantages of 3NF over BCNF.

5) (15 points) Consider the following set F of functional dependencies for the relational scheme R=(A,B,C,D,E,F).

A=>B
A=>D
C=>EA
C=>B

Now consider the decomposition of R into R1=(A,B,F) and R2=(A,C,D,E). Is this decomposition dependency preserving (yes or no)? If so, then explain why, and if not, then explain why not.

6) (20 points) Consider the following collection of relation schemes.

branch (branch-name, branch-city, assets)
customer (customer-name, customer-street, customer-city)
account (branch-name, account-number, balance)
loan (branch-name, loan-number, amount)
depositor (customer-name, account-number)
borrower (customer-name, loan-number)

Give an SQL statement for each of the following.

(a) A list of the names of those customers who have at least one account and at least one loan.

(b) A list of the names of those branches that have an average account balance of more than 1200.