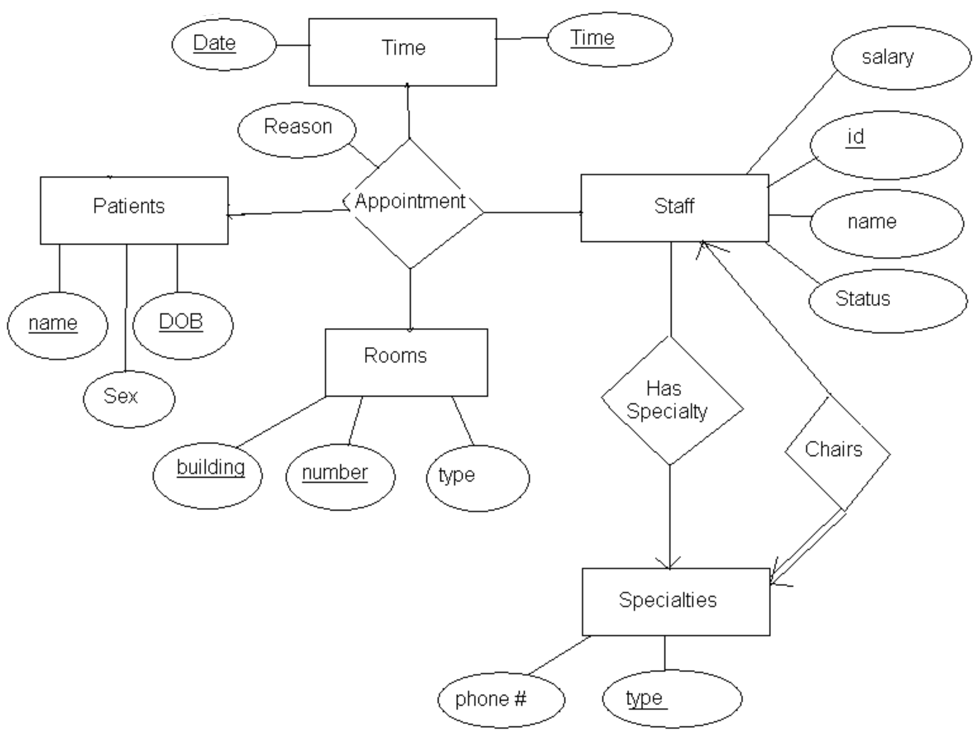
Additional Review Problems

1. BU has offered you free tuition for the spring semester on top of your usual consulting fee if you design a logical schema for a new database. Talking with the people in the administration, you gather the following information:
   - The database contains information about professors and courses.
   - Professors have a SSN, a name, and a research specialty. Each professor can be uniquely identified by her SSN.
   - Each course can be uniquely identified by a course number, and we also want to store the course's name.
   - For each course offering, we need to record the professor who teaches the course.
   - Each course is taught by exactly one professor.

   Draw a small ER diagram that captures the above information.

2. Imagine you are given the task of creating a database for a hospital. The hospital wants to keep track of doctors and staff (including their specialties), patients, rooms (when are they occupied and by whom), and appointments. The following is one possible ER diagram for this domain:



a. What constraints are specified for the Chairs relationship set?
b. What constraints are specified for the Has Specialty relationship set?

3. Translate the above ER diagram into a relational schema. Whenever possible, you should avoid giving a relationship set its own table, but instead capture it in the table for one of its connected entity sets.

   Notes:
   - One relationship set that *will* need its own table is the Appointment relationship set.
   - Because there is no information in the Time entity set except its primary key attributes, you do *not* need to create a separate table for it. Rather, its attributes can be captured in the Appointment table.
   - You may assume that a given patient has at most one appointment with a given staff member at a given date/time, and thus the room number and building do not need to be part of the primary key of Appointment.

4. Write the SQL command to create the Appointment table, making reasonable assumptions about the data types and the lengths of any string-valued attributes. Make sure to specify the primary key (which will consist of *many* attributes!) and any foreign-key constraints that make sense.

5. Write SQL commands to solve the following problems:

   a. Find the schedule and number of appointments for each doctor on November 2, 2022 (i.e., '2022-11-02'). Specifically, display the name of the doctor followed by the earliest appointment of the day for that doctor, the latest appointment time of the day, and the total number of appointments that day for each doctor who has at least one appointment on that date.

   b. Find the patient and doctor name for all appointments with a surgeon between November 2 and November 30, 2022.

   c. Find the names of all patients who had more than 3 appointments in July 2022.

   d. Find the names and salaries of all staff who make either the maximum salary or the minimum salary.

   e. Find the names of all staff members who do **not** have an appointment on November 2, 2022.

6. Write relational algebra queries to do the following:

   a. Find the patient and doctor name for all appointments with a surgeon between November 2 and November 30, 2022.

   b. Find the names of all staff members who do **not** have an appointment on November 2, 2022. For this query, you can assume that staff names are unique.

7. Assume that we're using variable-length records that begin with a header of field offsets to store rows from this relation:

    Staff(id CHAR(5), name VARCHAR(30), status VARCHAR(10),
        salary REAL, specialty_type VARCHAR(20))

    Assume that *all* of the values for a given row (including the primary key) are stored in the record, and that we're using 1-byte characters, 2-byte integer metadata values, an 8-byte double for the salary. What would the record look like for the following tuple?

    ('12345', 'Doogie Howser', 'MD', null, 'GP')

8. In the hospital domain, would you choose a B-Tree or a B+Tree to store the Staff records by their primary key? Why?

9. You have two relations, R and S:

Relation R

a	**b**	**c**
1	2	3
4	5	6
7	8	9

Relation S

c	**d**	**e**
3	4	5
7	8	9
9	0	1
5	5	5

How many rows would each of the following contain?

a. the cartesian product of R and S
b. the natural join of R and S
c. the left outer join of R and S
d. the right outer join of R and S
e. the full outer join of R and S