CMSC 23500 — Introduction to Database Systems
Discussion Session #4

April 21, 2008

In this discussion group you will be provided with the relational schema of a hospital's database, and will be asked to perform a set of queries on it using relational algebra and tuple calculus.

1 Hospital database

The schema shown in figure 1 models tha data typically handled in a hospital setting, including patients, appointments, medical procedures, etc. In particular:

- There are three types of person in the database: physicians, nurses, and patients.
- Patients have a primary care physician.
- Patients can make appointments with any physician in the hospital. An appointment may also involve a nurse (who will prep the patient before the physician examines him/her).
- A physician can prescribe medication to a patient (the same medication cannot be prescribed more than once on the same day). When a prescription is done as a result of an appointment, it is noted in the database, but there can also be prescriptions that are not linked to an appointment (e.g., getting a refill).
- Patients can stay in a hospital room for a number of days.
- A room is part of a hospital block, which is a subdivision of a hospital floor.
- As part of a stay, a patient may undergo a medical procedure, which is carried out by a physician and assisted by a nurse.
- Physicians can be trained in multiple medical procedures.
- Physicians can be affiliated with multiple hospital departments, although they will have a single primary affiliation.
- Nurses are assigned to be "on call" in the hospital's blocks.

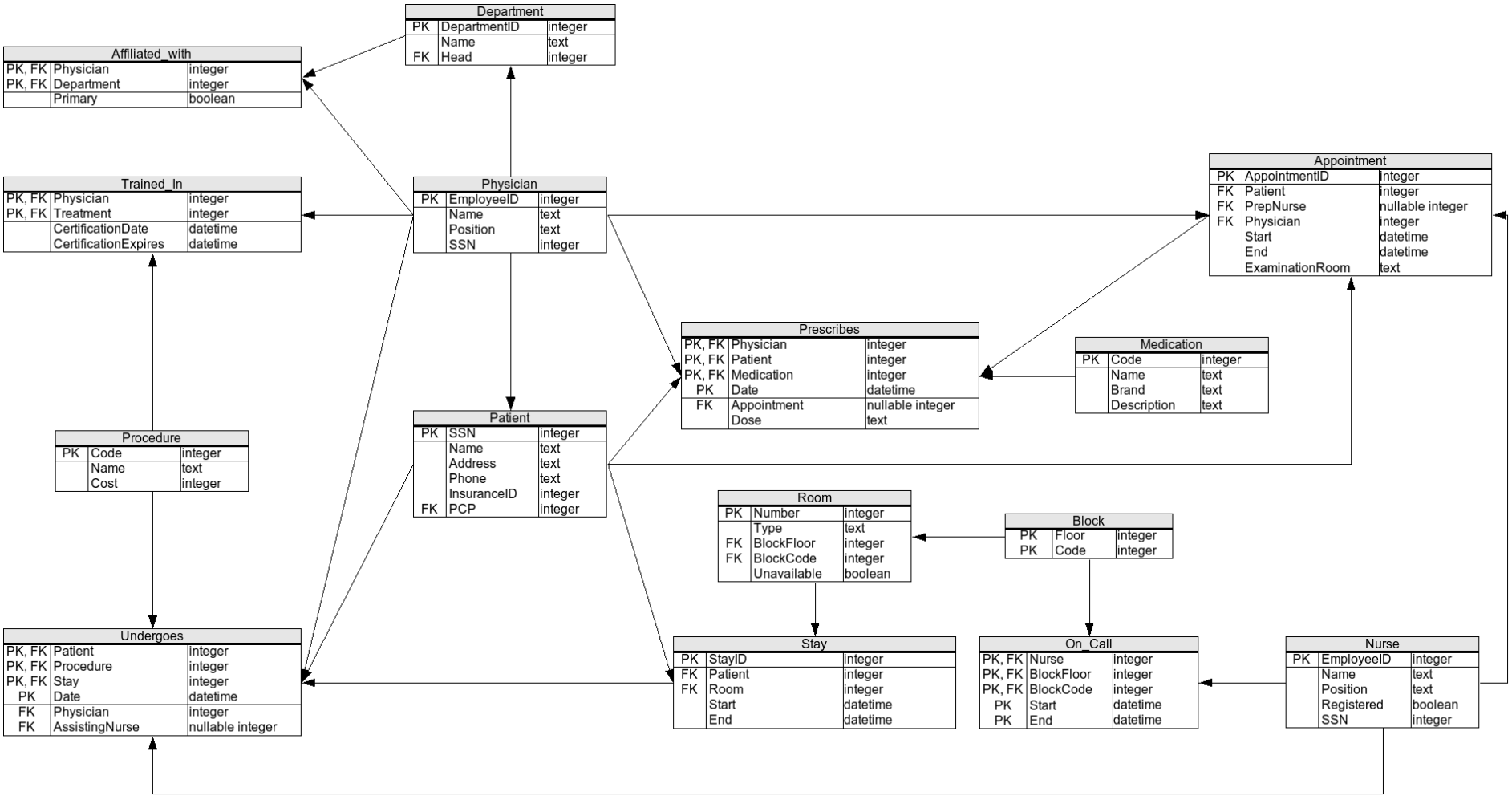

Figure 1: Relational schema for the Hospital database

2 Exercises

Assemble yourselves in groups of three. The instructor will hand each group a database query that you must write using the relational algebra. Do not discuss your query with other groups. Once you are done, write your query in the chalkboard. When every group is done, the instructor will assign you to a query (different from the one you wrote), and you will have to explain what that query is meant to retrieve. Time permitting, you will then rewrite the query using the relational tuple calculus.

3 Things to think about

Besides working on the exercises, take a look at the schema (during or after the discussion session) and consider the following:

- The AppointmentID and StayID fields in the Appointment and Stay relations are autoincrementing integer fields. Are there any other candidate keys in those relations? Why do you think it was preferable to create an artificial "ID" field in the table? Would it make sense to create such a field in the Undergoes relation?
- The current schema has no way of enforcing a single primary affiliation (between physicians and departments). Can you see why? Although this could be enforced at the application level, or using database triggers, is there any way we could modify the schema to enforce this constraint at the schema level?
- The Trained_In relation only allows the database to keep track of the procedures the physician is *currently* certified to perform. Can you see why? How would you modify the relation so it could keep a log of all current and past certifications?
- The On_Call relation potentially allows for a nurse to be assigned to multiple blocks at the same time. Can you see why? How would you modify the relation to enforce that a nurse can be on call in only one block during a given period of time? (you can assume that, for a given nurse, the periods of time entered in the database don't overlap).

4 Problems + Solutions

1. The hospital wants to check if any physician has been performing procedures they are not certified to perform. Obtain the names of all physicians that have performed a medical procedure they have *never* been certified to perform.

$$\begin{aligned} \mathsf{TRAINED} &= \Pi_{(\text{Physician, Treatment})}\mathsf{TRAINED\_IN} \\ \mathsf{PERFORMED} &= \Pi_{(\text{Physician, Procedure})}\mathsf{UNDERGOES} \\ \mathsf{NOT\_CERTIFIED} &= \mathsf{PERFORMED} - \mathsf{TRAINED} \end{aligned}$$

2. The hospital wants a list of all appointments where a patient met with his/her primary care physician. This list should include the start and end time of the appointment, the name and employee ID of the physician, the name and SSN of the patient, and the name (if any) of the nurse that prepped the patient.

$$\begin{aligned} &\mathsf{PATIENT}'(\text{PaName}, \text{PaSSN}, \text{PaPCP}) = \Pi_{(\text{SSN,Name,PCP})}\mathsf{PATIENT} \\ &\mathsf{PHYSICIAN}'(\text{PhName}, \text{PhID}) = \Pi_{(\text{Name,EmployeeID})}\mathsf{PHYSICIAN} \\ &\mathsf{NURSE}'(\text{NuName}, \text{NuID}) = \Pi_{(\text{Name,EmployeeID})}\mathsf{NURSE} \\ &\mathsf{APPTS1} = \mathsf{PATIENT}' \bowtie_{\text{PaSSN=Patient}} \mathsf{APPOINTMENT} \\ &\mathsf{APPTS2} = \mathsf{PHYSICIAN}' \bowtie_{\text{PhID=Physician}} \mathsf{APPTS1} \\ &\mathsf{APPTS3} = \mathsf{NURSE}' \ltimes_{\text{PrepNurse=NuID}} \mathsf{APPTS2} \\ &\mathsf{APPTS4} = \sigma_{\text{PhID=PaPCP}}\mathsf{APPTS3} \\ &\mathsf{APPTS} = \Pi_{\text{PaName,PaSSN,PhID,PhName,NuName,Start,End}}\mathsf{APPTS4} \end{aligned}$$

3. The hospital wants to check if any physician has been performing procedures with an expired certification. Obtain the names of all physicians that have performed a medical procedure that they are certified to perform, *but* such that the procedure was done at a date (Undergoes.Date) after the physician's certification expired (Trained_In.CertificationExpires).

$$\begin{aligned} \mathsf{UNDERGOES1} &= \mathsf{UNDERGOES} \bowtie_{\text{Physician=EmployeeID}} \mathsf{PHYSICIAN} \\ \mathsf{UNDERGOES2} &= \Pi_{(\text{EmployeeID,Name,Procedure,Date})}\mathsf{UNDERGOES1} \\ \mathsf{UNDERGOES3} &= \mathsf{UNDERGOES2} \bowtie_{\text{EmployeeID=Physician,Procedure=Treatment}} \mathsf{TRAINED\_IN} \\ \mathsf{UNDERGOES4} &= \sigma_{\text{Date>CertificationDate}}\mathsf{UNDERGOES3} \\ \mathsf{NOT\_CERTIFIED} &= \Pi_{(\text{Name})}\mathsf{UNDERGOES4} \end{aligned}$$

4. The hospital has recently discovered that room 1232 is the deadliest room in the hospital. More patients die in that room than can be attributed to chance alone.

Produce the names of all the nurses who have ever been on call for that room.

$$\begin{aligned}
\mathsf{ROOM1} &= \mathsf{ROOM} \bowtie_{\mathrm{BlockFloor=Floor,BlockCode=Code}} \mathsf{BLOCK} \\
\mathsf{ROOM2} &= \Pi_{(\mathrm{Floor,Code})}(\sigma_{(\mathrm{Number=1232})}\mathsf{ROOM1} \\
\mathsf{ROOM3} &= \mathsf{ROOM2} \bowtie_{\mathrm{BlockFloor=Floor,BlockCode=Code}} \mathsf{ON_CALL} \\
\mathsf{ROOM4} &= \mathsf{ROOM3} \bowtie_{\mathrm{Nurse=EmployeeID}} \mathsf{NURSE} \\
\mathsf{DEADLY_NURSE} &= \Pi_{(\mathrm{Name})}\mathsf{ROOM4}
\end{aligned}$$

5. The hospital suspects that some physicians are unfairly favoring the pharmaceutical brand SNAFU over other brands. Produce a list with the names of physicians who have prescribed SNAFU medications, including the number of such prescriptions (see section 6.4.2 of the book for aggregate functions).

$$\begin{aligned}
&\mathsf{PHYSICIAN}'(\mathrm{PhName, PhID}) = \Pi_{(\mathrm{Name,EmployeeID})}\mathsf{PHYSICIAN} \\
&\mathsf{PRESCRIBES1} = \mathsf{PRESCRIBES} \bowtie_{\mathrm{PhID=Physician}} \mathsf{PHYSICIAN}' \\
&\mathsf{PRESCRIBES2} = \mathsf{PRESCRIBES1}' \bowtie_{\mathrm{Medication=Code}} \mathsf{MEDICATION} \\
&\mathsf{PRESCRIBES3} = \Pi_{(\mathrm{PhID,Medication})}(\sigma_{\mathrm{Brand=\text{“}SNAFU\text{”}}}\mathsf{PRESCRIBES2}) \\
&\mathsf{SNAFU} =_{\mathrm{PhID}} \mathcal{F}_{\mathsf{COUNT(Medication)}}
\end{aligned}$$