

Comprehensive Exam
Databases
Fall 2005

1) Recall that SQL provides the special value *null* for representing missing or unknown values. As a result, SQL makes use of a three-valued logic for evaluating logical expressions. The three values used by the logic include *true*, *false* and *unknown*.

(a) (14 points) For this question you are required to fill in the blanks below showing how SQL evaluates terms in a logical expression.

true and unknown _____

false and unknown _____

unknown and unknown _____

true or unknown _____

false or unknown _____

unknown or unknown _____

not unknown _____

(b) (1 point) Based on the above, it is possible that a logical expression in a selection predicate might evaluate to a final value of *unknown*. In such a case what is the final *unknown* value converted to?

(c) (10 points) Give two reasons why the use of nulls in a relational database is considered controversial.

2) (25 points) Consider the following relational schemes. Note that attributes forming the primary key for each relation have been underlined.

student(<u>student-name</u> , street, city)	-- Basic student information.
offering(<u>department</u> , <u>number</u> , <u>section</u> , time, location, population)	-- Courses currently offered; for CSE5260 department is "CSE" and number is 5260.
titles(<u>department</u> , <u>number</u> , title)	-- Course titles; "CSE5260" is "Database Systems"
enrollment(<u>student-name</u> , <u>department</u> , <u>number</u> , <u>section</u>)	-- Indicates which students are enrolled in which courses.

Give a relational algebra expression for each of the following.

(a) The department, number, section and title for every course section that has more than 25 students enrolled, i.e., population > 25.

<u>department</u>	<u>number</u>	<u>section</u>	<u>title</u>
CSE	5260	1	Database Systems
MTH	5100	3	Discrete Math
PSY	4260	2	Abnormal Psychology
PSY	4260	3	Abnormal Psychology
CHM	2035	3	Introduction to Chemistry
:			

(b) A list of student names along with the titles of those courses they are enrolled in, and the location where the student takes each course.

<u>student-name</u>	<u>title</u>	<u>location</u>
John Smith	Database Systems	ECE 127
Bob Jones	Discrete Math	Crawford 220
Mary Brown	Abnormal Psychology	Link 105
:		

(c) A list of the names of those cities that have at least two student residents. For this question you are required to use the aggregate operator.

(d) The same query as in part (d). However, in this case your answer must not use the aggregate operator.

(3) Suppose an interface is developed for a system accessing a student database in an academic environment. This interface is used by staff in the registrar's office, and performs such functions as adding, deleting, and updating student information. For example, one button on the interface labeled *delete student*, will delete all information from the database for a given student. When clicked on, this button will result in a tuple being deleted from a student table, several tuples being deleted from an enrollments table, plus another from an advisors table, etc. In other words, clicking this button on the interface will result in several different SQL delete statements being executed, each on a different table. Unfortunately, the computer system that supports the database frequently crashes, and often times this occurs right in the middle of executing the above queries.

(a) (8 points) How might the above scenario result in an inconsistent database?

(b) (9 points) What SQL mechanism should be used to ensure that an inconsistent database does not occur?

(c) (8 points) Explain how the mechanism in (b) works. In other words, how does it prevent an inconsistent database from occurring?

4) (25 points) Give a *formal* definition for each of the following.

(a) (3 points) First normal form

(b) (4 points) Second normal form

(c) (4 points) Third normal form

(d) (4 points) BCNF

(e) (5 points) Dependency preserving decomposition

(f) (5 points) Loss-less join decomposition.