

University of Wisconsin-Madison  
Computer Sciences Department

Database Qualifying Exam  
Fall 2003

Answer all five (5) questions (NOTE: this is different from some previous years, when you were only asked to answer 4 of 5.) Before beginning to answer a question make sure that you read it carefully. If you are confused about what the question means, state any assumptions that you have made in formulating your answer. Good luck!

Database Design

Let relation  $R$  with functional dependencies  $F$  be decomposed into  $R_1, R_2, \dots, R_n$ .

- (a) Describe a polynomial time algorithm to test whether  $F$  is preserved in the set of decomposed relations. (Just describe the algorithm. You don't have to prove that it is polynomial, or that it is correct.)
- (b) Describe an algorithm to test whether the decomposition (into  $n$  relations) is lossless-join.
- (c) Consider the best algorithm that you know for automatic decomposition into BCNF. Would you use it to automatically design a database for a complex application? Explain briefly.

Concurrency Control

This question deals with the concurrency control mechanisms discussed in Gray's Granularity of Locking paper. Assume that the hierarchy for locking is a DAG with two paths from DB to record. The first is DB File Page Record, the second is DB Index Record.

Suppose transaction  $T_1$  wants to get an SIX lock on a file  $f$ . What are all the locks it should request?

Suppose now at another time (after  $T_1$  has completed)  $T_2$  wants to get an X lock on a page  $p$ . What locks should  $T_2$  request?

Turning now to degrees of consistency, Gray states that "If each transaction in a set of transactions at least observes the degree 0 lock protocol and if transaction  $T$  observes the degree 1 (2 or 3) lock protocol then  $T$  runs at degree 1 (2 or 3) consistency."

What does this mean? In particular, your answer should explain the differences (if any) between a system in which all transactions run at level 3 and a system in which some transactions run at level 3 while others run at lower levels of consistency.

## The Cube Operator

Define the cube operator as proposed in the Gray et al. paper. Your answer should include a discussion of the ALL value.

Which would you expect to be more efficient to evaluate: a cube operator with a distributive aggregate function, or one with a holistic aggregate function? Explain your answer by sketching an algorithm for each.

## View Updates

The following questions deal with updating relational views.

Define what it means to translate an update on a view.

It is desirable that a translation of an update on a view should not change the database unnecessarily. State such a condition precisely.

Give an example of a view update for which there is no reasonable translation.

Give an example of a view update for which there is more than one reasonable translation.

## Recovery

Briefly answer the following questions:

How does the recovery manager ensure the atomicity of transactions?

How does the recovery manager ensure durability?

What is the difference between a system crash and a media failure?

Explain the WAL protocol.