Normalization II

Tom Kelliher, CS 317

1 Administrivia

Announcements

Assignment

From Last Time

Normalization I.

Outline

1. Vocabulary
2. Normalization practice

Coming Up

Triggers and stored procedures

2 Vocabulary

1. If a relation is 3NF but not BCNF, it must have a non-trivial functional dependency X $\rightarrow$ A such that
   A. the primary key does not determine X
   B. X is not a superkey
   C. A is not a superkey
   D. the primary key does not determine A

2. In the relational model, if A functional determines B, it means that
   A. the value of B can be calculated from the value of A
   B. the value of B is always the same as the value of A
   C. if two rows have the same A value, they must have the same B value
   D. if two rows have the same B value, they must have the same A value

3. In the relation R(A,B,C,D) if D is multivalued and the apparent key is A, which of the following is NOT an acceptable way to create a first normal form table for this schema?
   A. Decompose R into R1(A,B,C) and R2(A,D)
   B. Determine the maximum number of values n that there are for D in any record, and include that many columns for D, making the relation have the form R1(A,B,C,D1,D2,. . .,Dn)
   C. Decompose R into R1(A,B,C,D) and R2(D)
   D. Use D as part of the primary key, making the relation have the form R1(A,D,B,C)

4. In a relation R(A,B,C,D), which of the following would prove that R is not 3NF?
   A. B → A
   B. B → C
   C. C → A
   D. D → A

5. The highest normal form that always allows us to preserve functional dependencies is
   A. 1NF
   B. 2NF
   C. 3NF
   D. BCNF

6. In the relational model, a projection is lossless if
   A. the join produces the same structure as the original relation
   B. no further decomposition is possible
   C. the intersection of the relations is empty
   D. the join produces the same tuples as the original relation

7. In relational decomposition of a universal relation, the property of attribute preservation requires that
   A. every attribute of the universal relation appears in only one relation
   B. every attribute of the universal relation appears in exactly two relations
   C. every attribute of the universal relation appears in at least one relation
   D. every attribute remains in the original relation

8. In relational decomposition of a universal relation, the property of dependency preservation requires that
   A. all determinants appear in the same relation
   B. all attributes appear in some relation
   C. it is possible to reconstruct the original relation by a join
   D. all the attributes on both sides of each dependency appear in the same relation

9. We can ensure that a binary decomposition is lossless if the set of common attributes in the two relations is
   A. empty
   B. the union of the two relations

C. a superkey of one of the relations
D. functionally dependent on the keys of both relations

10. It is always possible to find a Boyce-Codd Normal Form decomposition that
A. is lossless, but may not preserve dependencies
B. preserves dependencies, but may not be lossless
C. is both lossless and dependency-preserving
D. is disjoint

11. Using an E-R diagram can be helpful in designing a normalized schema because
A. The standard mapping of an E-R diagram to a relational model always results in a normalized schema
B. The process of mapping an E-R diagram to a relational model is equivalent to the synthesis method of normalization
C. The standard mapping of an E-R diagram to a relational model results in a schema that is close to normalized
D. The process of mapping an E-R diagram to a relational model is equivalent to the decomposition algorithm for BCNF

3 Normalization Practice

Keep this in mind: The root cause of update, insertion, and deletion anomalies is redundancy. An attribute is a fact that should describe only the entity represented by the tuple in which it appears. "Each attribute must describe the key, the whole key, and nothing but the key."

1. Define "attribute preservation," "dependency preservation," and "lossless decomposition."

2. When might we want to settle for 3NF, or even 2NF, rather than continuing to normalize?

3. The following relation is used to track a person's hobbies. A person may have several hobbies.

   persons(<u>name</u>, address, hobby)

   Explain why this relation isn't 1NF. Construct an instance of the relation and use the instance to demonstrate an anomaly (update, insert, or delete).

4. A parts company has several warehouses. A particular part may be kept in one or more warehouses. The company uses the following relation to keep track of its parts and its warehouses.

   parts(<u>part</u>, <u>warehouse</u>, quantity, warehouseAddr)

   Explain why this relation is 1NF but not 2NF. Construct an instance of the relation and use the instance to demonstrate an anomaly (update, insert, or delete).

5. A company uses the following relation to hold information about its employees and the locations of its departments. An employee works in a single department and a department has a single address.

employees(empId, name, dept, deptAddr)

Explain why this relation is 2NF but not 3NF. Construct an instance of the relation and use the instance to demonstrate an anomaly (update, insert, or delete).

6. A pizza shop uses the following relation to track the pizzas that it prepares and to keep track of the pizza offerings that are available. "topping" is a topping; things like mozzarella, olives, mushrooms, sausage, and pepperoni, for example. "toppingType" is the type of a topping, for example, cheese, vegetable, and meat. A pizza may have at most one of each topping type.

   pizzas(pizzaId, toppingType, topping)

   Explain why this relation is 3NF but not BCNF. Construct an instance of the relation and use the instance to demonstrate an anomaly (update, insert, or delete).

7. Problem 6.6 in the textbook; apply to parts a–c of Problem 6.5.

   (The algorithm for constructing a lossless join decomposition that is in BCNF may be found at the bottom of pg. 292 of the textbook.)

8. Problem 6.1 in the textbook, parts c and d.

9. Problem 6.2 in the textbook, parts c and d.