Practice Test for Språkdatabaser (VT98,980528) page 1 of 2

This test contains 10 questions each of which is worth 10 points. To pass this test you should obtain at least 60 points. You may answer in English or in Swedish. Try to give short answers.

1. a-5 Describe the three database levels of the ANSI/SPARC database architecture.

   b-5 What are the two properties of candidate keys?

2. a-5 What is the difference between a DDL and a DML?

   b-5 In the definition of some foreign key FK of a relation R we read DELETE RESTRICTED and UPDATE CASCADES. What will happen with a value of FK in R when it is deleted in the referenced relation? And what will happen when such a value is changed in the referenced relation?

3. a-5 What is the main data structure is the relational database model?

   b-5 What is the difference between the two relational algebra operators PRODUCT and JOIN?

4. a-5 What constraint does a relation have to satisfy to be in 5NF?

   b-5 What is the difference between the domain calculus and the tuple calculus?

5. a-5 Give at least two advantages of object-oriented databases compared with standard relational databases.

   b-5 Is it possible to extract the functional dependencies of a relation by inspection the values of the cells in the relation? Motivate your answer.

6. a-5 Give at least two advantages of using language databases.

   b-5 Explain the concept ”text span” that is used in the Tipster architecture.

We have two relations R and S with four attributes each. R has attributes A, B, C and D and with this relation corresponds the following irreducible set of functional dependencies: A→B, A→D, B→C and D→A. S has attributes E, F, G and H and with this relation corresponds the following irreducible set of functional dependencies: EF→G, EF→H and G→H. All other functional dependencies of these relations are either trivial or derivable from the irreducible FD sets. Both R and S are at least in 1NF.

7. a-5 Determine all candidate keys of R.

   b-5 Determine all candidate keys of S.

8. a-5 Is R in 2NF? And is it in 3NF? Motivate your answers.

   b-5 Is S in 2NF? And is it in 3NF? Motivate your answers.

9. a-5 Show that R is in BCNF. If you think that R is not in BCNF then convert it to a set of BCNF relations.

b-5 Show that S is in BCNF. If you think that S is not in BCNF then convert it to a set of BCNF relations.

The tables and attributes in the supplier-and-parts database are:

S: S#, SNAME, STATUS, CITY
P: P#, PNAME, COLOR, WEIGHT, CITY
SP: S#, P#, QTY

The names of the suppliers appear in SNAME and the names of the parts appear in PNAME.

10. a-5 Give an expression in SQL or the relational algebra or the tuple calculus for performing the following query on the suppliers-and-parts: "Get the names and weights of all red parts that are supplied by the supplier with name Moore". Mention which of the three expression types you have used.

b-5 Give an expression in SQL or the relational algebra or the tuple calculus for performing the following query on the suppliers-and-parts: "Get the cities where suppliers are located together with the number of suppliers per city". Mention which of the three expression types you have used.

Answers Practice Test Språkdatabaser VT98:980528

1. a The ANSI/SPARC database architecture contains three levels: the internal level describes how data is stored at the physical level, the external level is the level on which the database users perceive the data and the conceptual level is the logical representation level in between the first two on which the database operations work.

   b Every candidate key CK of a relation R has the following properties: 1. different tuples in R will always have different values for CK and 2. no subset of CK has property 1.

2. a A DDL or data definition language defines the structure of the data and a DML or data manipulation language is a set of commands which can be used for processing data.

   b It will not be possible to delete a value of FK from the referenced relation if the value exists in R because the definition of FK prohibits such a deletion. It is possible to update an in R existing FK value in the referenced relation. The corresponding value in R will be changed as well.

3. a The main data structure in the relational database model is the table.

   b Both PRODUCT and JOIN combine two relations into a new one but PRODUCT combines all possible row pairs while JOIN combines row pairs for which the values in shared attributes are the same.

4. a In order to be in 5NF a relation may only contain join dependencies which are implied by its candidate keys.

   b The domain calculus works with variables which represents domains while the tuple calculus works with variables which represent tuples.

5. a Object-oriented databases have as advantages that they able the database users to define their own data types and they offer the possibility of letting data types inherit structural information from each other.

   b No, it is not possible to extract the functional dependencies of a relation by inspection the values of the cells in the relation. Functional dependencies are a semantic concept. You can only obtain them if you understand what the data in the relation means. The meaning of the data does not always follow from the contents of the relation.

6. a Language databases have as advantages that they reduce the chance of inconsistencies and redundancies and they offer the possibility to examine the same data in different ways.

   b In the Tipster architecture a text span is a pair of numbers which refer to a location in a text corpus. The pair contains identification values for the start and the end of the location.

7. a The candidate keys of R are {A} and {D}.

   b The candidate key of S is {E,F}.

8. a R cannot be in 2NF since it has two candidate keys and the 2NF definition covers only relations with one candidate key. It cannot be in 3NF because all 3NF relations are also in 2NF.

   b S is in 2NF because the two nonkey attributes G and H are irreducibly dependent on the candidate key EF. However S is not in 3NF because the nonkey attribute H is transitively dependent on candidate key EF (via G).

9. a A relation is in BCNF iff every left-irreducible nontrivial FD has a candidate key on its left-hand side. In R this is not true for the FD B→C so we split R in {A,D,B} and {B,C}. Now B is a candidate key for {B,C}. The resulting set of two relations is in BCNF.

   b S is not in BCNF because of the FD G→H which does not have a candidate key in its determinant. So we split S in two relations {E,F,G} and {G,H}. These relations are both in BCNF.

10. a Here I have used a relational algebra expression:
   (((s join sp) join p) where color='red' and sname='Moore')[pname,weight]

   b Here I have used an SQL expression:
   select city,count(city) from s group by city;