

Database Management - 2017

Model Answers

1. (a) i. The cardinality of the PaysFor relationship type is many-to-one. (1 mark)
 - ii. Premium is a weak entity type. This means that it doesn't have sufficient attributes of its own to act as a primary key. (3 marks)
 - iii. The relation schema would be Premium(PolicyNumber, DueDate) with primary key (PolicyNumber, DueDate). There would be a foreign key constraint for PolicyNumber in Premium, referencing the primary key of Policy. (5 marks)
 - (b) i. {Track, Artist} is a key. Using the closure algorithm, we can add TrackTitle from the first FD, then Album from the third, and finally AlbumTitle from the second. Neither Track nor Artist determines all attributes. (4 marks)
 - ii. One BCNF decomposition is as follows. The LHS of the first FD (Track) is not a superkey for {Track, TrackTitle, Album, AlbumTitle, Artist}, so decompose into {Track, TrackTitle} and {Track, Album, AlbumTitle, Artist}. The LHS of the second FD (Album) is not a superkey for {Track, Album, AlbumTitle, Artist}, so decompose into {Album, AlbumTitle} and {Track, Album, Artist}. The LHS of the third FD (Track) is not a superkey for {Track, Album, Artist}, so decompose into {Track, Album} and {Track, Artist}. The final 4 relation schemas are therefore {Track, TrackTitle}, {Album, AlbumTitle}, {Track, Album} and {Track, Artist}. (7 marks)
 - (c) The GROUP BY clause is used to group rows together which have the same value(s) for some attribute(s). Then typically an aggregate function is applied to each group of rows, e.g. to find the total number of parts supplied by each supplier. (5 marks)
2. (a) i. It returns (person, item) pairs such that the person visits a store which sells the item. (3 marks)

- ii. It returns (person, item) pairs such that the person likes the item and also visits a store which sells it. (3 marks)
- iii. No, it would not always return the same answer. This query would return (person, item) pairs such that the person visits a store which sells the item and some person likes the item. (4 marks)
- (b) Null values are used to represent unknown or inapplicable values. Testing for null values in SQL requires the use of IS NULL, rather than using equals. The presence of null values forces the query processor to adopt 3-valued logic. Null values complicate the processing of aggregation functions. (5 marks)
- (c) The default action disallows the update or deletion of the referenced value. The CASCADE action propagates the deletion or update to the foreign key value. The SET NULL action sets the corresponding foreign key value to null. (5 marks)
- (d) Boyce-Codd Normal Form (BCNF) is stronger than Third Normal Form (3NF) and eliminates more update anomalies (redundancies). However, a dependency preserving 3NF decomposition can always be found, which is not always the case for BCNF. (5 marks)
3. (a) i. For the closure of CD , we start with CD . From the third FD, we can add A . From the first FD, we can add B , to give $ABCD$. (2 marks)
- ii. BCE is a superkey, since the second FD allows us to add D using the closure algorithm. From the third FD, we can add A , and from the fourth, we can add G , yielding all attributes. None of BC , BE or CE is a key, so BCE is a key. (4 marks)
- (b) In embedded SQL, a preprocessor turns SQL statements into procedure calls that fit with the surrounding host-language code. All embedded SQL statements begin with EXEC SQL, so the preprocessor can find them easily. In SQL, the shared variables must be preceded by a colon. In the host language, shared variables behave like any other variable. Shared variables are declared as in the following example:
- ```
EXEC SQL BEGIN DECLARE SECTION;
 char theBar[21], theBeer[21]; float thePrice;
EXEC SQL END DECLARE SECTION;
```
- and used as follows:
- ```
EXEC SQL SELECT price INTO :thePrice
    FROM Sells
```

```
WHERE bar = :theBar
AND beer = :theBeer;
```

(7 marks)

- (c) i. The transitive FD is $\text{Employee} \rightarrow \text{Manager}$. “Transitive” means that the FD can be inferred from the other two FDs “via” the Department attribute. (2 marks)
- ii. Transitive FDs cause redundancy which in turn can lead to update anomalies. (2 marks)
- iii. A Third Normal Form design would be (Employee, Department) and (Department, Manager). (2 marks)
- (d) Object-relational database systems provide features such as user-defined types, type inheritance, table inheritance, array types, multiset types, object identity, reference types and methods (any 6). (6 marks)
4. (a) i. SQL views are useful to make it easier for users/programmers to query a database by providing them with appropriate sets of attributes from underlying tables. They also provide a means of restricting classes of users to subsets of the database. (2 marks)
- ii. The logical layer and the physical layer. (2 marks)
- iii. A view is defined by means of a query, so the data “in” the view is not materialised (stored). (2 marks)
- (b) A dirty read operation is when one transaction reads data written by another, uncommitted transaction. This means that if that transaction aborts, the data read will be incorrect. This is disallowed by the isolation levels READ COMMITTED, REPEATABLE READ and SERIALIZABLE. (7 marks)
- (c) i. Three from $\text{film} \rightarrow \text{director}$, $\text{film} \rightarrow \text{year}$, $\text{year} \rightarrow \text{film}$, $\text{director} \rightarrow \text{film}$ or director , $\text{year} \rightarrow \text{film}$. (3 marks)
- ii. $\text{director} \rightarrow \text{year}$, and $\text{year} \rightarrow \text{director}$ (or film , $\text{year} \rightarrow \text{director}$) (2 marks)
- iii. A director does not necessarily only make films in one year, so $\text{director} \rightarrow \text{year}$ does not hold. The FD $\text{year} \rightarrow \text{director}$ also does not hold because every film made in a year does not have the same director. The FD film , $\text{year} \rightarrow \text{director}$ probably holds because it is unlikely that two different versions of the same film would be made in the same year (or that a film has multiple directors).

(4 marks)

- (d) They return rows appearing (i) in either table, (ii) in both tables, and (iii) in table A but not table B, respectively.

(3 marks)

5. (a) Any four of the following. It saves programmer time by providing a declarative query language, e.g. SQL. It saves programmer time by automatically checking constraints. It saves maintenance time by ensuring data independence. It provides correct concurrent access to the database for multiple, simultaneous users. It provides automatic recovery from failure. It provides security to ensure appropriate access to data.

(8 marks)

- (b) i. An error will be generated if the same primary key value already exists, or if the Price attribute is set as NOT NULL, or if there is a foreign key constraint and no matching primary key value.

(3 marks)

- ii. An error will be generated if the same primary key value already exists, or if a CHECK constraint requires a positive price value, e.g., , or if there is a foreign key constraint and no matching primary key value.

(3 marks)

- iii. An error will be generated because the primary key cannot take a null value.

(1 mark)

- (c) i. A left outer join joins two tables, including every row from the left table and padding these with null values if there is no matching row from the right table.

- ii. An ISA relationship type is used in the ER model to represent a generalisation/specialisation hierarchy.

- iii. A dependency preserving decomposition is one in which all the original functional dependencies can be inferred from those represented in the decomposed relations.

- iv. An atomic transaction is one which either runs to completion or whose effects on the database are undone.

- v. A noSQL database is one which allows for a more flexible data model, usually without a schema being required.

(10 marks)