

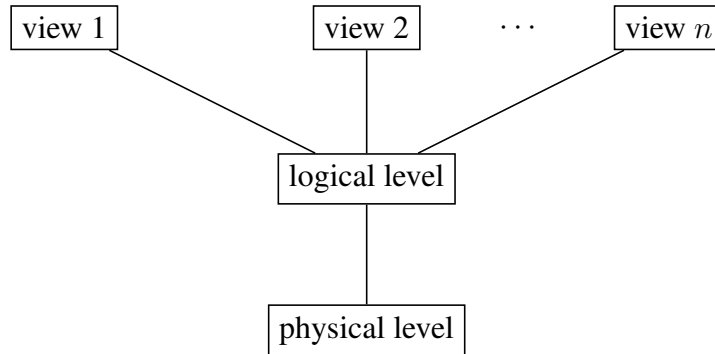
Database Management - 2018

Model Answers

1. (a) i. Multiplicity constraints include one-to-one, many-to-one and many-to-many. (3 marks)
 - ii. If R is one-to-one, then either K_2 can be included in the attributes of the translation of E_1 or K_1 can be included in the attributes of the translation of E_2 . If R is many-to-one from E_1 to E_2 , then K_2 can be included in the attributes of the translation of E_1 . Alternatively, a separate relation including K_1 and K_2 could be formed, with primary key K_1 . If R is many-to-many, then a separate relation including K_1 and K_2 should be formed, with primary key (K_1, K_2) . (6 marks)
- (b) i. NT and ND (2 marks)
 - ii. Yes. The LHS of $NT \rightarrow D$ is a superkey and the RHS of $D \rightarrow T$ is prime. (3 marks)
 - iii. No. The LHS of $D \rightarrow T$ is not a superkey. (3 marks)
 - iv. Every time a dish is ordered, its type is repeated. (2 marks)
- (c) Three from
 - i. all data is value-based - all relationships are expressed through common values.
 - ii. data must always be represented in "flat" (first normal form) relations.
 - iii. for some applications, performance is not fast enough.
 - iv. relational schemas are insufficiently flexible.

Three from: object-oriented, object-relational, and Nosql databases such as key-value stores, XML databases, document stores and graph databases. (6 marks)

2. (a) The three levels of abstraction are shown in the following diagram:



The logical level presents a conceptual model to applications and users, shielding them from the actual physical representation of information in the database, as well from changes to physical representations that might be made by a DBA.

The view (external) level presents different views of the conceptual model to different users/applications, also allowing for imposing different levels of access. The view level also shields applications from (certain) changes at the logical level, requiring only the mappings between the levels to be redefined, rather than programs and queries to be rewritten.

(8 marks)

- (b) i. Three from course \rightarrow textbook, textbook \rightarrow course, lecturer \rightarrow textbook or textbook \rightarrow lecturer.

(3 marks)

- ii. lecturer \rightarrow course and course \rightarrow lecturer

(2 marks)

- iii. Lecturers do not each only teach a single course, so it is unlikely that lecturer \rightarrow course holds. It is more likely that course \rightarrow lecturer holds (if each course is taught by a single lecturer), but this may also not be true in most institutions.

(4 marks)

- (c) The term “atomically” refers to the fact that transactions should execute to completion or else appear not to have executed at all. As an example, consider a banking application on relation Accounts(acctNo, balance) where we want to transfer £100 from account 123 to account 456. The transaction might first ensure there is at least £100 in account 123. In step 2, it might then deduct £100 from account 123. In step 3, it might add £100 to account 456: If there is a failure after step 2 and before step 3, then £100 would have been lost.

(8 marks)

3. (a) One design has all attributes in one schema: (ID, Name, Phone, Email, Spouse). Spouse should be declared as a foreign key referencing PersonID.

Another way would be to have two schemas: (ID, Name, Phone, Email) and (ID, Spouse). Both attributes ID and Spouse in this schema would be foreign keys referencing ID in the first schema.

The advantage of the second approach is not having to use NULL values for people who are unmarried.

(7 marks)

- (b) The HAVING clause is used in conjunction with the GROUP BY clause. Given a set of groups defined by GROUP BY, the HAVING clause specifies a condition which groups must satisfy in order to be included in the results.

(4 marks)

- (c) SQL specifies what restrictions must be placed on view definitions in order for them to be *updatable*.

- i. Updating a database through a view can be problematic because such an update may be ambiguous, that is, it may not be obvious what updates to apply to the tables on which the view is defined.

(2 marks)

- ii. An SQL view is said to be updatable if the query defining the view satisfies the following:

- The FROM clause has only one database relation.
- The SELECT clause contains only attribute names of the relation (no expressions, aggregates, or DISTINCT).
- Any attribute not listed in the SELECT clause can be set to null.
- There is no GROUP BY clause.

(6 marks)

- (d) i. The size of the result is estimated as $1/|R|$ if A is the primary key, and $1/i$ otherwise, where i is the number of distinct values in column A.

(2 marks)

- ii. The size of the result is estimated as $1/|R|$ since each row in S joins with at most one row in R.

(2 marks)

The assumption is that the values appearing in a column are uniformly distributed. If this is not the case, the estimation will underestimate the size of the answer.

(2 marks)

4. (a) Aliases are useful for renaming columns in the output and also for renaming tables, either to shorten the query or, for example, when two or more copies of the same table appear in a query. The syntax uses the keyword AS to introduce them.

(3 marks)

- (b) i. $BE \rightarrow BE \rightarrow BEI$ (2nd FD) $\rightarrow BEIC$ (3rd FD) $\rightarrow BEICD$ (4th FD).

(3 marks)

- ii. ABE is a key, since adding A above will give all attributes, and no subset of ABE is a key.
(3 marks)
- iii. The decomposition is not dependency-preserving, since $BE \rightarrow I$ is not preserved. The common attributes are not a key for either schema, so the decomposition is not lossless. (A, B, C, I) is not in Third Normal Form since AB is not a key and C is not prime. (C, D, E, I) is not in Third Normal Form since E is not a key and C is not prime.
(8 marks)
- (c) The query string is used to find the maximum quantity in which a given product is supplied, using an anonymous placeholder for the product value (2). The query is compiled in the second line (1). Then the value for the product is retrieved from an HTTP GET request (2), and the compiled query is executed (1), substituting the values into the query (1). The last line fetches the result (1).
(8 marks)
5. (a) The two constraints are entity integrity and referential integrity. The first requires that no attribute that is part of a primary key can contain a null value. The second requires that every non-null value that appears in a foreign key must appear in the corresponding primary key.
(6 marks)
- (b) `select distinct A from R where A in (select A from S);`
(3 marks)
- (c) i. 3-valued logic means that the system needs to work with the value UNKNOWN, in addition to the usual values of TRUE and FALSE.
(3 marks)
- ii. If the second condition is true, the expression evaluates to true. If the second condition is false, the expression evaluates to false. If the second condition is unknown, the expression evaluates to unknown.
(3 marks)
- (d) i. A weak entity type is one that doesn't have sufficient attributes of its own to form a primary key.
ii. A composite key is a key comprising more than one attribute.
iii. The Cartesian product (of two relations) is formed by combining each tuple from the one relation with every tuple from the other relation.
iv. A correlated subquery is one in which a condition in an inner query refers to an attribute from an outer query.
v. A dirty read is reading a value written by an uncommitted transaction.
(10 marks)