Database Design and Development

December 2015

Sample Exam Marking Scheme

This marking scheme has been prepared as a **guide only** to markers. This is not a set of model answers, or the exclusive answers to the questions, and there will frequently be alternative responses which will provide a valid answer. Markers are advised that, unless a question specifies that an answer be provided in a particular form, then an answer that is correct (factually or in practical terms) **must** be given the available marks.

If there is doubt as to the correctness of an answer, the relevant NCC Education materials should be the first authority.

Throughout the marking, please credit any valid alternative point.

Where markers award half marks in any part of a question, they should ensure that the total mark recorded for the question is rounded up to a whole mark.

Answer ALL questions

Marks

Question 1

Consider the following SQL statement:

```
SELECT customer.LastName, COUNT(Orders.OrderID) As NumberOfOrders

FROM orders, customers

WHERE customer.customerID = Orders.CustomerID

GROUP BY LastName

HAVING COUNT(Orders.OrderID) > 18;
```

a) What is the purpose of the COUNT function in this SQL statement? **2**

COUNT here is an aggregate function (1 mark). The purpose of this is to count the number of orders that belong to a particular customer (1 mark)

b) What is the purpose of the GROUP BY command in this SQL statement? **3**

The GROUP BY is used to groupe orders via the LastName (1 mark). This will mean that what is returned is the count of orders for each customer (1 mark) identified by their name (1 mark).

c) What is the purpose of the HAVING command in this SQL statement? **3**

This modifies the GROUP BY with a condition (1 mark). In this case, it limits the orders that will be displayed (1 mark) as those where there have been more than 18 orders for that customer (1 mark)

d) How would you modify this SQL query so that it shows the orders for customers called 'Smith' regardless of how many orders have been made? **2**

The HAVING command should be removed so that orders are shown regardless of how many (1 mark). A line should be added to the WHERE part of the statement:
AND LastName = 'Smith' (1 mark)

Total: 10 Marks

Marks

Question 2

a) Explain what a *domain* is in a relational database **and** give an example. **2**

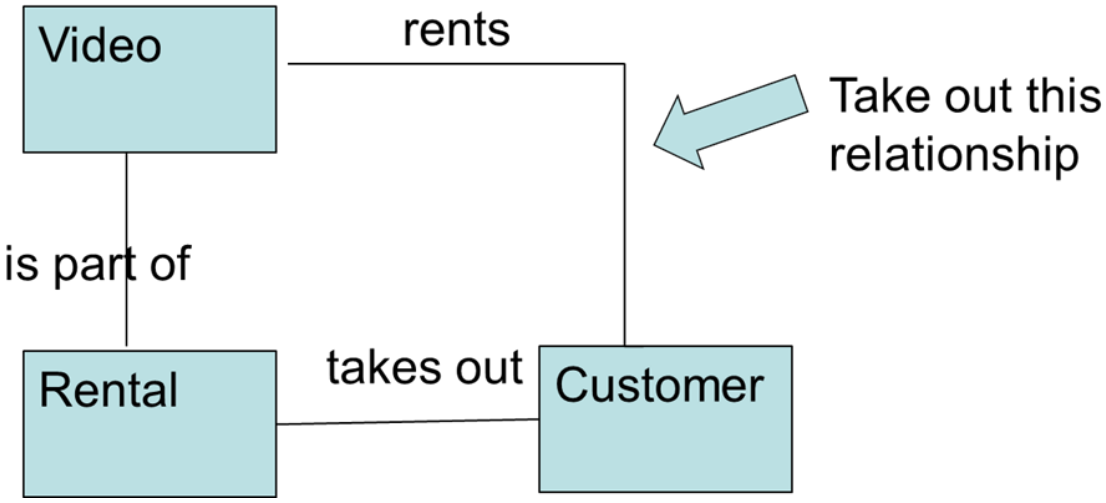
A domain specifies a set number of values that are valid for a column (1 mark). For example, the values of a column primary_colours could be defined as belonging to a domain that contained only the values: Red, Blue, Green (1 mark).

b) Show how you would use the Alter Table SQL command to add a domain on a database table. **3**

To add a constraint to make sure a quantity is above zero:
ALTER TABLE cars
ADD CHECK (colour in ('Red','Blue','Green'))
Up to 3 marks for example (1 mark for 'Alter Table', 1 mark for 'Add Check', 1 mark for the condition)

c) Explain the concept of redundancy in a relational database **and** give an example. **5**

Something is redundant if it is not needed (1 mark). It may be the case that during the process of design carried out so far that structures (entities, attributes, relationships) (1 mark for any of these) may have been created that are not really necessary because the same information can be found elsewhere (1 mark).
Example:



Here we don't need to keep the 'Rents' relationship because it is possible to connect customers with the videos they rent via the Rental entity. (2 marks).

Total: 10 Marks

Marks

Question 3

a) Explain the concept of *functional dependence* with the use of an example. **5**

Within a relation, if it is said that 'A determines B' then this means that if you know the value of 'A' then you will know the value of 'B' (1 mark). So A functionally determines B. Or B is functionally determined by A (1 mark). Note that the reverse is not true. The diagrammatic representation of an example of this is:

StudentID → StudentName ***(1 mark for an example)***

If we know the StudentID then we know the StudentName (1 mark). But the reverse is not true because we might have students who have the same name (1 mark).

Maximum of 5 marks.

b) How is functional dependence used in database development? **2**

Functional dependence is used in the process of normalisation (1 mark). It helps the developer determine which attributes belong to which entity (1 mark)

c) Explain what a deletion anomaly is in a database. **3**

An anomaly that occurs because a database is not normalised properly (1 mark). For example, a table that records students, the activities they do and the prices of those activities. If we delete the only student doing an activity then we could lose the information about that activity (2 marks for this or similar example).

Total: 10 Marks

Marks

Question 4

a) Explain what is meant by the term *superkey* in a relational database. **2**

An attribute or set of attributes that uniquely identifies a tuple (1 mark). It may contain superfluous attributes beyond those that make a tuple unique (1 mark).

b) Explain what is meant by the term *candidate key* in the relational model. **3**

Candidate key – a candidate key should be a superkey (1 mark). However, ALL the attributes of this super key must be necessary to uniquely identify it i.e. there should be no superfluous attributes that are part of the key (1 mark). It should not be the case that any subset of the attributes that go to make up this key should qualify as a superkey (1 mark).

c) Identify FIVE (5) properties that a relation must have in the relational model. **5**

- ***Relation has a name that is unique within the relational schema.***
- ***Each cell in the relation contains a single (or atomic) value.***
- ***Each attribute has a name unique within the relation.***
- ***The values of any one attribute should be drawn from the same domain.***
- ***There are no duplicate tuples (rows) within a relation.***
- ***There is no significance to the order of attributes.***
- ***There is no significance to the order of tuples.***

1 mark each for any five of these.

Total: 10 Marks

Question 5

a) What is the main purpose of conceptual database design? **5**

Conceptual design is, in a sense, a 'pure' investigation into the data needs of an organisation and/or system (1 mark). This entails the identification of entities (1 mark) and attributes etc. (1 mark) but independently of any particular model (1 mark). So the investigation here does not even assume the relational model (1 mark).

b) What are the main features of logical database design? **5**

The data is now designed in terms of the chosen model (1 mark) but not the chosen DBMS (1 mark). If this is the relational model then this involves activities like normalisation (1 mark), entity relationship modelling (1 mark) and defining attributes (1 mark).

Total: 10 Marks

Marks

Question 6

a) What is meant by *referential integrity*? Give an example. **3**

Referential integrity: any foreign key must match a candidate key (1 mark for specifically identifying the referenced key as a candidate key) in the parent table (1 mark). An example is Orders that relate to a Customer – the Customer foreign key must relate to a Customer ID that exists in that table (1 mark).

b) Explain what is meant by a *propagation constraint* **and** give an example. **3**

Propagation constraints: enforcing the rules as to what happens if data that is referenced elsewhere is altered (1 mark). For example, if there is a Customer who has a number of related Orders, what happens to these Orders if that Customer is deleted from the database? (1 mark) It also concerns updates to reference columns (candidate keys referenced as foreign keys) (1 mark).

c) Describe FOUR (4) options for propagation constraints in a database system. **4**

- ***No action, which means the record in the table with the foreign key is left as it is. (1 mark)***
- ***Cascade, which means that any change (including a delete) is replicated in the table with the foreign key in it. (1 mark)***
- ***Set Default, which means that the change in the parent table causes the record in the child table to be set to some sort of default. (1 mark)***
- ***Set Null – similar to Set Default except that the table with the foreign key has that foreign key set to Null. (1 mark)***

Total: 10 Marks

Marks

Question 7

a) Distinguish between *database transactions* and *database operations* with the use of an example. **4**

Operations are the basic units of action (1 mark) that take place with regard to database structures such as tables and columns (1 mark). They can be classified using the types listed here.
CREATE or INPUT, RETRIEVE, UPDATE, DELETE (1 mark for any example of an operation)
Transactions are made up of one or more of these operations (1 mark)

b) For a transaction in a database, identify FOUR (4) aspects that might affect performance and therefore need to be documented during physical design. **4**

- ***The tables and columns accessed and type of access (CRUD)***
- ***Columns used in any search condition***
- ***Any join conditions for queries***
- ***Expected frequency of transaction***

1 mark for each point.

c) How might the use of an index affect different types of database operations? **2**

Queries and updates (including insertions and deletions) can be affected in different ways (1 mark). Often an index that improves performance of a query can actually be slow to update (1 mark).

Total 10 Marks

Marks

Question 8

a) What factors should be considered when choosing which columns to index in a database system? **5**

The columns most used in joins are chosen (1 mark). Joins are indexed because these are most often used in queries in situations where performance can be affected (1 mark) If an index can be used in this situation, the chances are that it will make the query run quicker (1 mark). Columns used for ordering can also be indexed to improve performance (1 mark). In this case, the ordering operation can use the index to do the sort rather than the rows in the database table (1 mark).

b) Describe THREE (3) factors that need to be considered when deciding whether to store derived data in a database. **3**

- ***How complicated is the calculation to be made?***
- ***Does it involve multiple table joins and so could impact performance whenever source tables are updated?***
- ***How often would a derived attribute be updated and therefore could this affect performance detrimentally?***

1 mark each for these three of points.

c) What type of documents might help a developer identify an organisation's derived data during database development? **2**

Something like an invoice (1 mark) in an organisation usually contains lots of derived data such as totals for the number of products ordered (1 mark). Other examples are acceptable.

Total 10 Marks

Marks

Question 9

a) Describe FIVE (5) advantages an organisation might gain by implementing a distributed database. **5**

- ***Emulating organisational structure.***
- ***Greater control. Data is located where it is needed and so it is easier to control.***
- ***Improved availability. Locally stored data is less likely to have access disrupted by network problems or issues with the central database.***
- ***Greater reliability.***
- ***Better performance. Where most data is retrieved from a local version (either replicated or fragmented) this could result in an increase in performance.***
- ***Easier growth. Where local sites have control over data relevant to them it is easier for them to plan and structure the growth of the database.***

1 mark each for any five of these points.

b) Describe the main aspects of the principle of transparency in a distributed database. **5**

Each site should have local autonomy (1 mark) and ensure that there is no reliance on a central site (1 mark). This is the case whether the database is replicated or fragmented (1 mark). In a fragmented case, then the sub-set of records that are stored locally must be capable of performing all necessary transactions without relying on a central database (1 mark). In the case of replication then the copy of the database should behave in all ways like the original (1 mark).

Total 10 Marks

Marks

Question 10

a) Discuss how a focus on 'subject-orientation' differentiates a data warehouse from an Online Transaction Processing System (OLTP). **4**

Database system applications in OLTP systems are usually designed around the core operations of the system (1 mark). The data warehouse should be organised instead around major subject areas (1 mark). This orientation on subject areas affects the design and implementation of the data in the data warehouse (1 mark). The major subject area of a given piece of data should form part of the unique identifier of a particular piece of data stored in a data warehouse (known as the key structure) (1 mark). Alternative answers could focus on the role of a data warehouse in providing business intelligence, transforming data into information etc.

b) Describe FOUR (4) features of an Online Analytical Processing (OLAP) tool. **4**

Consolidation: allowing aggregation of data. (1 mark)
Drilling down: this is the opposite of consolidation. It involves the breaking down of data into finer levels of detail. (1 mark)
Pivoting: being able to analyse the same data from different viewpoints. (1 mark)
Multi-dimensional data structures that have been visualised as having lots of dimensions. (1 mark).

c) Give a brief definition of data mining. **2**

Data mining is an approach that looks for patterns in large sets of data such as a data warehouse (1 Mark) The vital characteristic is that the patterns are discovered rather than just consciously sought (1 Mark)

Total 10 Marks

End of Examination Paper

Learning Outcomes matrix

Question	Learning Outcomes assessed	Marker can differentiate between varying levels of achievement
1	2	Yes
2	3	Yes
3	3	Yes
4	3	Yes
5	4	Yes
6	4	Yes
7	4	Yes
8	5	Yes
9	1	Yes
10	1	Yes

Grade descriptors

Learning Outcome	Pass	Merit	Distinction
Understand the enterprise application of database systems	Demonstrate adequate level of understanding	Demonstrate robust level of understanding	Demonstrate highly comprehensive level of understanding
Understand how to enhance the design of and further develop a database system	Demonstrate ability to perform the task	Demonstrate ability to perform the task consistently well	Demonstrate ability to perform the task to the highest standard
Be able to enhance a logical database design	Demonstrate ability to perform the task	Demonstrate ability to perform the task consistently well	Demonstrate ability to perform the task to the highest standard
Be able to develop a physical database design	Show adequate development	Show sound and appropriate development	Show innovative and highly appropriate development
Be able to enhance a database system using SQL	Demonstrate ability to perform the task	Demonstrate ability to perform the task consistently well	Demonstrate ability to perform the task to the highest standard