

**B.Sc. EXAMINATION**

School of Computing

May 2012

2 hours

**Database Systems**

**(AC32004)**

This paper contains FOUR questions.

All questions carry equal marks.

Answer ALL questions.

Only calculators approved by the School for exam use may be used in this exam.

Question 1

**(a)** Should a primary key be updated? Describe the issues involved in this question, both for the relational database theory and for the practical use of an SQL DBMS.

**[10 marks]**

|  |  |
| --- | --- |
|  | It is commonly agreed that [**primary keys should be immutable**](http://programmers.stackexchange.com/questions/8320/should-a-primary-key-be-immutable) (or [as stable as possible](http://stackoverflow.com/questions/3632726/what-are-the-design-criteria-for-primary-keys) since **immutability can not be enforced in the DB**). While there is nothing that will prevent you to update a primary key (except integrity constraint), it may not be a good idea:  From a performance point of view:   * **Single** **updated key** can lead to updates of potentially lots of tables/rows. * **Unindexed foreign keys** – require a lock on the children table to ensure integrity * **Indexed foreign keys -** the **update will lead to the update of the index** (delete+insert in the index structure), this is **generally more expensive than the actual update of the base table**. * In **ORGANIZATION INDEX tables** (in other RDBMS, see clustered primary key), **the rows are physically sorted by the primary key**. A logical update will result in a physical delete+insert (more expensive)   Other considerations:   * If this key **is referenced in any external system** (application cache, another DB, export...), the reference will be broken upon update. * additionaly, **some RDBMS don't support CASCADE UPDATE**, [in particular Oracle](http://asktom.oracle.com/pls/asktom/f?p=100:11:0%3a%3a%3a%3aP11_QUESTION_ID:5773459616034). |

**(b)** Suppose that a database administrator has produced a logical database design that is to be implemented using an SQL DBMS, and that many users will require access to data in different base tables for different purposes. Only the database administrator is to define the database and determine what the users are allowed to do.

**(i)** Describe the different types of SQL statement and DBMS commands that may be required to establish the database and to make it accessible to the users.

**Data Definition Language** (DDL) statements are used to define the database structure or schema.

* **CREATE** - to create objects in the database

**Data Manipulation Language** (DML) statements are used for managing data within schema objects.

* ·         **SELECT** - retrieve data from the a database
* ·         **INSERT** - insert data into a table
* ·         **UPDATE** - updates existing data within a table
* ·         **DELETE** - deletes all records from a table, the space for the records remain

**Data Control Language (DCL) statements.**

* ·         **GRANT** - gives user's access privileges to database
* ·         **REVOKE** - withdraw access privileges given with the GRANT command

**Transaction Control** (TCL) statements are used to manage the changes made by DML statements. It allows statements to be grouped together into logical transactions.

* ·         **ROLLBACK** - restore database to original since the last COMMIT

·       **SET TRANSACTION** - Change transaction options like isolation level and what rollback segment to use

**[4 marks]**

**(ii)** Identify any kinds of statement or command you gave in (i) which may be optional and give reasons why they may be used.

**Transaction Control** (TCL) statements are used to manage the changes made by DML statements. It allows statements to be grouped together into logical transactions.

* ·         **ROLLBACK** - restore database to original since the last COMMIT

·       **SET TRANSACTION** - Change transaction options like isolation level and what rollback segment to use

Incomplete orders

**[4 marks]**

**(iii)** Some of the data to be put into the database is available in a file that combines data for two of the tables. Describe the kinds of SQL statement or DBMS command involve in storing this data by means of the import function.

Select INTO

Insert INTO

Update

**[3 marks]**

**(iv)** Users’ retrieval needs may be satisfied either via an application programme or by using SQL interactively. Give one advantage and one disadvantage of each alternative.

**[4 marks]**

Application provides a simplistic interface – no programming knowledge needed

No full control over database – restricted by what is programmed by database

Potential to damage database

Knowledge of SQL needed to use it.

Question 2

1. Select an appropriate database model (for example, relational database model or object database model) or data processing technique (for example, data mining) for each of the following scenarios, justifying each choice. In your answer you should explain which particular aspects of the model or technique make it particularly suitable for this type of scenario, with reference to: (i) the size or complexity of the data set; (ii) the nature of the queries that may need to be performed upon the data; (iii) the degree to which the data may be predominantly static or dynamic.
2. An online shop records every item bought and who the purchaser was. It operates a loyalty card scheme so it can track purchases made by address and age profile of the customers. It plans to introduce a discount system on selected products based upon factors such as the customers’ shopping habits and its own stock levels.

**OLAP / ROLAP**

**Size and complexity:** ROLAP is considered to be more scalable in handling large data volumes, especially models with [dimensions](http://en.wikipedia.org/wiki/Dimension_%28data_warehouse%29) with very high [cardinality](http://en.wikipedia.org/wiki/Cardinality) (i.e., millions of members).

**Nature of queries:** The data are stored in a standard [relational database](http://en.wikipedia.org/wiki/Relational_database) and can be accessed by any [SQL](http://en.wikipedia.org/wiki/SQL) reporting tool (the tool does not have to be an [OLAP](http://en.wikipedia.org/wiki/OLAP) tool).

**Change :**

**dynamic**

**[6 marks]**

1. A large car park company has a video camera system for preventing or investigating security incidents. Videos are captured 24 hours a day and archived. If a crime is reported, such as a break-in to a parked car, the car park company can access the archive to investigate the crime.

**[6 marks]**

1. A company sells a variety of teaching materials to help students whose first language is not English become familiar with common phrases and expressions. It currently has audio materials and video materials, plus a range of booklets that include images as well as textual information. It wants to transfer this to a computerised form that will allow content browsing.

**[6 marks]**

1. Name and describe THREE multimedia database types and then suggest, with a justification, the most appropriate type for holding images in a multimedia database.

**Blobs – binary large objects –** unstructured sequence of bytes

Blobs are typically [images](http://en.wikipedia.org/wiki/Image), [audio](http://en.wikipedia.org/wiki/Sound) or other [multimedia](http://en.wikipedia.org/wiki/Multimedia) objects, though sometimes binary executable code is stored as a blob.

**Clobs – character large objects –** unstructured sequence of characters

Allows the use of Like or Substring

**File reference – link to the data –** no access to the data at all

File reference is a reference to the physical address of where the data is stored in meory or in the stored device.

**[7 marks]**

Question 3

1. **(i)** Explain what the terms **measures** and **dimensions** mean in the context of a Sun model of a data warehouse that is configured for OLAP data visualisation.

**\*Measure**

**A measure is a performance indicator that is quantifiable and used to determine how well a business is operating. For example, useful measures may be Quantity Sold or Revenue.**

**\*Dimension**

**A dimension is a broad grouping of related data about a major aspect of your business. For example, you have a dimension called Products.**

**(ii)** Suggest TWO typical measures and FOUR dimensions that would be of interest to a car rental company that rents different types of cars and car accessories (eg snow tyres, GPS system, baby seat) for different seasons of the year.

Measures - Revenue, Total Rentals

Dimensions – Car, Year, Rental, Accessories

**[4 marks]**

1. Give a star schema model and a Sun model appropriate to the car rental company as described in part (ii) above.

**[8 marks]**

1. Explain why searching a MOLAP cube is significantly faster than searching a relational database that is based on the same data.

**[2 marks]**

1. Inmon defined a data warehouse as: “a **subject-oriented**, **integrated**, **time-variant** and **non-volatile** collection of data in support of management’s decision-making process”. Explain what each of these four terms means, giving an example to illustrate each answer.

**[8 marks]**

Question 4

**(a)** A large IT training company wishes to record data about the **borrowable equipment** it has, such as iPads, other tablets, portable projectors and Kinect systems. The **managers** want to be able to track who has equipment **on loan**, the **status of equipment** (damaged, on loan, on the shelf), the **borrowing behaviour** of the **training staff** (prompt or late returns, care taken of items borrowed, track record of payment of fines or replacement costs) and to send **reminder emails** to **borrowers** at appropriate intervals.

Create an Entity Relationship diagram that captures this information as a conceptual design. You should include cardinality constraints in your answer but there is no need to include attributes.

Equipment

Loans

Status

**[14 marks]**

1. What responsibility does the IT training company have towards the borrowers of its equipment, in terms of the Data Protection Act (1998)? Justify your answer by referring to the main principles of the Act as they relate to data controllers.

1. Processed fairly and lawfully.

2. Obtained for specified and lawful purposes.

3. Adequate, relevant and not excessive.

4. Accurate and up to date.

5. Not kept any longer than necessary.

6. Processed in accordance with the “data subject’s” (the individual’s) rights.

7. Securely kept.

8. Not transferred to any other country without adequate protection in situ.

**[6 marks]**

1. Explain how the IT training company might implement the database system to ensure the security of the data, in terms of precautions against accidental data loss or deliberate but inappropriate data amendment.

Transactions

Stored procedures

Back up

User Permissions

Ensuring database bridge is as far away from user as possible.

**[5 marks]**

**END OF EXAM**