**Assignment1** (Chapter 1 -3 Connolly & Begg) Each question is worth 10 points each

1. Discuss the differences between DDL and DML. What operations would you typically expect to be available in each language?

Answer:

* Data definition language (DDL) is used for defining and creating a database schema or modifying an existing one. DDL cannot be used to manipulate data.
* The operations expected to be available in DDL are CREATE, ALTER, DROP, TRUNCATE, RENAME, etc.
* Data manipulation language (DML) is used for supporting the basic data manipulation operations.
* The operations expected to be available in DML are INSERT, UPDATE, DELETE, SELECT, etc., so modification, insertion, retrieval, and deletion.

1. Describe the difference between data security and data integrity.

Answer:

* Data security uses controlled access mechanisms to prevent data corruption.
* Data security makes sure that the data can only be accessed by intended users.
* Data integrity defines the quality of data. It refers to the validity of data.
* Data integrity makes sure that the data stored in the database is not altered by unauthorized parties.

1. Describe the main characteristics of the database approach and contrast it with the file-based approach.

Answer:

Main characteristics of the database approach:

* Control of data redundancy
  + File-based approach has a lot of data redundancy
  + Usually database approach stores data item in only one place
* Sharing of data
  + Database system allows multiple users to access to the data in the database at the same time
  + Database system therefore must have concurrency control
  + File-based approach does not provide data sharing with multiple users
* Multiple views of data
  + Particular users can use a subset of the database
  + Different users may have different view of the database
  + File-based approach does not have multiple views
* Insulation between program and data
  + File-based approach defines all the data files in application programs, so changing the structure means that all the programs need to be changed, too.
  + Database approach stores the data structure in the system catalog; therefore, no need to change all the programs.
* Provides meta-data
  + Database approach provides meta-data to describe the data structure and constraints
  + In file-based approach, the data definition is part of application programs.
* Improved data security
  + Database approach has better security control of the user accounts
* Data independence
  + In file-based approach, file structure is defined in the program code.
  + In database approach, the system data are separate from the programs.

1. Provide a definition for a data administrator and a database administrator. What types of interactions would these two users of the database have?

Answer:

* Data administrator definition:
  + Someone who manages the data resource, which includes database planning, development, and maintenance of standards, policies and procedures, and conceptual and logical database design.
* Database administrator definition:
  + Someone who manages the physical realization of a database system, which includes physical database design and implementation, setting security and integrity controls, monitoring system performance, and reorganizing the database, as necessary.
* Data administrator does the early stage logical database design.
* Database administrator does the later stage application and physical database design.
* Data administrator deals with the corporate data resource, which includes non-computerized data.
* Database administrator is more technical.
* DBA can assist DA by providing technical information about the DBMS, such as the system performance, security and integrity constraints, etc.

1. Name three record-based data models. Discuss the main differences between these data models.

Answer:

* Relational data model
* Network data model
* Hierarchical data model
* The relational data model is based on the concept of mathematical relations. The data in two tables are related because they have the same attributes. For the example in the book, the two tables are related because they both have branchNo.
* Compared with the relational model, the relationship is explicitly model for network data model. The records are organized as graph structures. The nodes are the records.
* The difference between hierarchical data model and network data model is that the hierarchical model allows a node to have only one parent. It can be represented as a tree.

1. What are the advantages of a relational database when compared to the file-based approach to storing data?

Answer:

* Control of data redundancy
* Data consistency
* More information from the same amount of data
* Sharing of data
* Improved data integrity
* Improved security
* Enforcement of standards
* Economy of scale
* Balance conflicting requirements
* Improved data accessibility and responsiveness
* Increased productivity
* Improved maintenance through data independence
* Increased concurrency
* Improved backup and recovery services

1. What is concurrency control and why does a DBMS need a concurrency control facility?

Answer:

* Concurrency control is to be able to manage operations on the database at the same time without the operations interfere each other.
* One of the advantage of a relational database system is data sharing.
* Multiple users can access the shared data concurrently.
* If two users are accessing the same data and one of them is changing it, then there may be interference.
* Therefore, we need concurrency control.

1. What is a transaction? Give an example of a transaction.

Answer:

* A transaction is a logical unit of work on the database.
* It can be as simple as a single command or as complicated as an entire program.
* For example, using a few lines of commands to change the salary of a staff member in the database is a transaction.

1. What is meant by the term ‘client-server architecture’ and what are the advantages of this approach? Compare the client-server architecture with two other architectures.

Answer:

* Client-server architecture is a type of network architecture.
* All the computers on the network is either a client or a server.
* Client requires some resource.
* Server provides the resource.
* The advantage of this approach includes:
  + Wider access to existing database
  + Increased performance
  + Possible reduction in hardware costs
  + Reduction in communication costs
  + Increased consistency
* Compared to teleprocessing, the client-server architecture reduces the burden on the central computer.
* Compared to file-server architecture, the client-server architecture has significantly lower communication costs.

1. What is a Transaction Processing Monitor? What advantages does a TP Monitor bring to an OLTP environment?

Answer:

* Transaction Processing (TP) Monitor is a program that controls data transfer between clients and servers in order to provide a consistent environment, particularly for online transaction processing.
* Advantages TP Monitor brings to an OLTP environment:
  + Transaction routing
    - TPM can direct transaction to specific DBMSs
  + Managing distributed transactions
    - TPM can handle transactions that need data from different DBMSs
  + Load balancing
    - TPM can direct clients to the less busy server
  + Funneling
    - TPM can establish connections between client and the DBMSs when needed to allow more users to be logged on simultaneously
  + Increased reliability
    - TPM can resubmit transactions when DBMSs fail or hold transactions when DBMSs are not available