Unified Modeling of Jdbc and Rdbms for Information Retrieval

Giridhary Kumar Tamuli[1], Dr Bala Buksh2

[1]*Research Scholar, School of Computer Application & Technology, Career Point University, Kota(Raj.), India*
[2]*Research Supervisor, School of Computer Application & Technology, Career Point University, Kota(Raj.),*
India

Abstract: The article refreshes the applicability of jdbc, relational database and information retrieval for creation of unified model that can used in a local computer or in a limited ecosystem. MS Access is considered for RDBMS and jdbc-odbc bridge (deprecated after Java-8) for connecting jdbc with the tables of the database.

I.   Introduction

In the era of cloud computing, big data and internet services, information gathering and retrieval have become very convenient. The information can be accessed and downloaded as per user requirements in multiple of devices like hand held mobile devices, laptops and desktops. However, information is stored in remote servers and internet communication is a must.

MS Access is a RDBMS (Relational Database management System) where tables or relations contains row or tuple and columns or attributes are used for storing information in databases (multiple tables or relations). JDBC can be used to connect to these tables of database and retrieve information as per necessity. The information is locally available in the local machine in a Client Server Model and does not require internet connection for viewing and retrieval. The model can also be applied in networking with a Server for Information Retrieval.

JDBC

JDBC stands for Java Data Base Connectivity. We will be connecting to MS Access database through Open Database Connectivity (ODBC) i.e. through jdbc-odbc bridge (deprecated after Java-8). Various drivers/ translators are also available to connect through jdbc to various proprietary databases like Oracle, Sybase, DB2, Microsoft SQL Server, MySQL, etc. Since, we are connecting to

2

MS Access database through jdbc-odbc bridge (deprecated after Java-8), the following steps are followed to acquire the connection :
1.   The relevant java packages are imported (in our case – sql package) and java class name is created.

2.   Load the JDBC database driver.
        DriverManager.registerDriver(new sun.jdbc.odbc.JdbcOdbcDriver());
3.   Create a connection
        Connection con = DriverManager.getConnection("jdbc:odbc:DSName");
4.   Create a statement
        Statement st = con.createStatement();
5.   Create a resultset, if we expect the database server to send back some data
        ResultSet rs = st.executeQuery(sql);
        while (rs.next()) {
        }
 6.   ResultSet, Statement and Connection interfaces need to be closed after execution of queries.
        rs.close();
        st.close();
        con.close();

These interfaces are executed under try and catch exception handling blocks.
        Connection.createStatement() method is used to create Statement objects, PreparedStatement objects are created using the prepareStatement() method and CallableStatement objects are created using prepareCall() method. Once executed, each object is used as the vehicle to channel SQL to the database. Statement object is

used for static SQL statements, while the PreparedStatement and the CallableStatement objects are used in conjunction with SQL containing IN or OUT parameters. **JDBC does not support getting OUT parameter as streams.**

3

While Statement object is created using a method from the Connection class, setting or retrieving the value of the parameter is part of the Statement object itself. Data from SQL queries is returned in ResultSet objects in tabular format. Each field is an unique combination of a row ID and column ID like a two-dimensional array. All data stored in a single column are an array of same data type and a row is represented as a Vector related data set that may or may not be all of the same type. Hence, the entire table can be executed as an array of Vectors.

After ResultSet object is created, the "get" method can access the data. To process a single row, we advance the cursor to that row and execute the appropriate "get" method with a column index as the argument.

RDBMS (MS ACCESS)

RDBMS (Relational Database management System) e.g. MS Access is a collection of tables or relations where tuple/row is a sequence or list of values without ordering and attribute/column is a domain of permitted atomic values. Duplicate values in a tuple are also not permitted. A referential integrity constraint is maintained between the tables through primary key and foreign key. The constraints can be implicit constraints which are inherent model-based constraints or explicit constraints which are schema-based constraints or semantic constraints which are application-based business rules known as assertions in SQL. Functional dependencies and multi-valued dependencies are another constraints in data dependencies utilized through a process called normalization. Therefore, the schema based constraints can be summarized as domain constraints, key constraints,
constraints on NULLs,
entity integrity constraints, and
referential integrity constraints.

4

A database schema diagram e.g. entity-relationship diagram along with primary and foreign keys, represents various constraints applied to a database.
A query language (procedural or non-procedural) is used to extract data from the database which also results in a relation. Hence, various relational operations like selection, join operation, cartesian product operation, set operations (union, intersection and difference), etc. are applied to query results.

After ascertaining the constraints in relation schemas, the relations are subjected to the process of normalization based upon functional dependencies up to third normal form (3NF) or to Boyce-Codd normal form (BCNF) - a stronger 3NF. Normalization is carried out to
(1) minimize redundancy, and
(2) minimize the insertion, deletion and update anomalies. It is, however, not required to normalize to the highest possible normal form. For performance reasons, relations can be in 2NF also which will but result in various anomalies.

4NF and 5NF have also been defined based upon multivalued and join dependencies. The practicality of 4NF and 5NF are not present due to inherent constraints.

Information Retrieval

Although web search engines are the dominant ways of finding information, information retrieval is the process of retrieving documents in response to query by a user. The information can be in the form of written texts, abstracts, documents, books, web pages, e-mails, messages, digital library, etc. Also, can be images, audio recordings, video strips, and maps. Mostly the informations are in the form of unstructured or semi-structured data. However, this information need not be in machine readable form as in Relational Databases.

5

In Information Retrieval systems, retrieval and browsing are the two form of interaction of an user and are characterized by types of users, types of data, type of information need and scale of repository. Most Information Retrieval Systems process the text collections to create indexes and operate upon an inverted index which comprises of vocabulary and document information.

There are two type of information retrieval process – statistical and semantic. Documents are broken down into text or word, counted, weighted, and measured for relevance during in statistical process which can be Boolean, vector space, and probabilistic. **Semantic** process use knowledge-based techniques including statistical analysis for retrieval process.

Unified Modeling (JDBC, MS Access and Information Retrieval)

Although there is limitations with MS Access in terms of limited number of users and quantity of rows and columns in a table, the database can be effectively utilized in a small ecosystem or in a local machine, The tables along with attributes in the database can be manually created considering the various constraints and data types or the tables can be created through jdbc. Data can then be entered into the table as tuple. Multiple tables can be created and relationship between the tables are created with Primary Key (PK) and Foreign Key (FK). Any information entered in the table can be retrieved as in Information Retrieval.

DSN and database name need to be updated with the system before connection between jdbc and MS Access, although jdbc-odbc bridge (deprecated after Java-8) is used in the connection. A Graphical User Interface (GUI) can also be created in applet (deprecated), frame window (limited capability) or swing application to insert, update and delete operation in the table of database or to exit the program as a whole.

6

Conclusion

The unified modeling in a local or networked environment provides an isolated and secure setup to focus and concentrate on the task at hand. The environment is free of digital marketing ads and pop-up windows which can be very distracting at times. Recommender systems and digital marketing have their own advantages and benefits.

Hence, in the present-day cloud computing environment, these unified modeling can be used to operate in a safe and secure ecosystem, free from the concern of hacking and other hazards. The administrator can securely monitor the safety and security in this type of setup.

References

[1].   Silberschatz Abraham, Korth Henry F, Sudarshan S. Database System Concepts. Sixth Edition. India: McGraw Hill Education (India) Private Ltd 2013.

[2].   Elmasri Ramez, Navathe Shamkant B. Fundamentals of Database Systems. Seventh Edition. India: Pearson India Education Services Pvt Ltd 2017.

[3].   Schildt Herbert. Java: The Complete Reference. Seventh Edition. India: Tata McGraw Hill Education Private Ltd 2011.

[4].   Kurniawan Budi. Java for the Web with Servlets, JSP, and EJB: A developer's guide to J2EE Solutions. First Edition. New Riders Publishing 2012.

[5].   Siple Matthew. The Complete Guide to Java Database Programming. First Edition. India: Tata McGraw Hill 1998.

[6].   Reese George. Database Programming with JDBC and Java. Second Edition. India: O'Reilly Shroff Publishers & Distributors Pvt Ltd 2000.

[7].   John V. Peterson. Absolute Beginner's Guide to Databases. India: Techmedia

[8].   Craig Eddy, Timothy Buchanan. Sams Tech Yourself Microsoft Access 2000 in 24 Hours. India: Techmedia.