MIA 99-43: The BetaDBC LIbrary - Reference Manual and Tutoria

Table of Contents

Copyright Notice

Mjølner Informatics Report
MIA 99-43
August 1999

The BetaDBC Library

Introduction

This document describes the BetaDBC (Beta DataBase Connectivity) library for communicating with relational databases using SQL (Structured Query Language, [Date 1993]). The library implements the pattern connection used to model connections to relational databases. This pattern contains patterns for querying and manipulating relational databases. Additional support for transactions may be added by including the fragment transactions. Scrolling cursors are implemented in the fragment scrollingresultset.

The BetaDBC Library

BetaDBC Basics

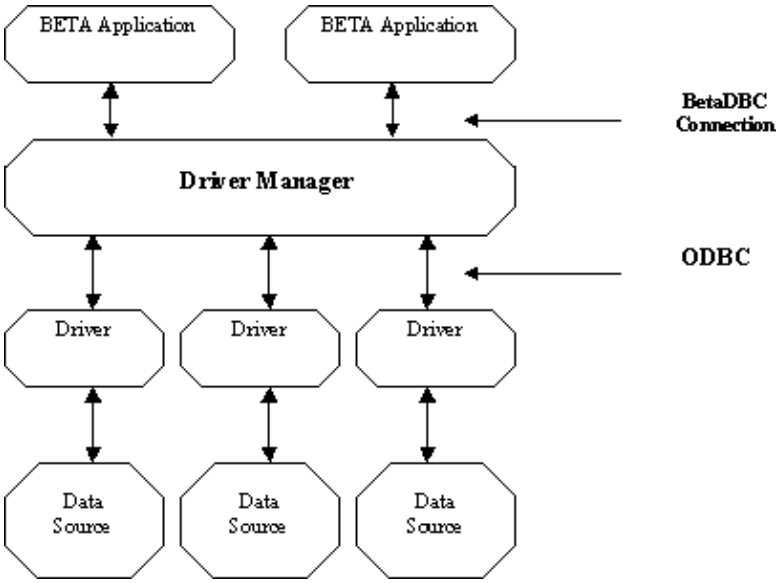
Generally speaking, there are two ways of programming to a relational database, namely using

- Embedded SQL. Using this approach SQL statements are embedded in a host language. Using a Database Management System (DBMS) specific precompiler, the application program is precompiled, transforming statements containing shared variables (variables shared by the DBMS and the application program) into DBMS library calls. After compilation the program is linked with the DBMS' runtime library.
- Call Level Interfaces (CLIs). These are libraries of functions. The functions are the native application programming interface of the DBMS or calls to it. In this way, a precompiler is not needed.

Both approaches suffer from some problems although standards have been proposed: Embedded SQL ties a program to a specific DBMS meaning that at least a new compilation will have to be made if a new DBMS is to be used. CLIs are hard to learn and often contain a lot of DBMS specific functions.

BetaDBC combines these two approaches in such a way that although shared variables may be used a precompiler is not needed and furthermore all DBMS may be treated alike. In order to achieve this BetaDBC currently builds upon and extends Open DataBase Connectivity (ODBC [Geiger, 1995])

This means that the architecture of a typical BetaDBC application may be outlined as below:



A BETA application typically connects to a DBMS via BetaDBC, calls BetaDBC functions, processes results and disconnects. The driver manager loads and unloads drivers as requested by applications, and processes function calls before sending them to a driver. The driver then processes the functions calls, submits SQL requests to data sources and returns results. Data sources encapsulate the data in form of tables that a user wants to access together with an associated operating system, DBMS, and network platform (if applicable) used to access the DBMS.

Details on how to create data sources and use BetaDBC in this environment will be given below.

The BetaDBC Library

The BetaDBC Interface

The BetaDBC interface is built around the concept of a connection that models a connection to a relational DBMS. The BetaDBC interface to a connection is (see below for the full interface):

```
Connection: (*)
  (# <<SLOT ConnectionLib:Attributes>>;
     declareVar: (*) ...;
     declareInteger: (*) declareVar ...;
     declareReal: (*) declareVar ...;
     declareText: (*) declareVar ...;
     declareBoolean: (*) declareVar ...;
     declareDate: (*) declareVar ...;
     declareTime: (*) declareVar ...;
     SQLStatement: (*) ...;
     directSQLStatement: (*) SQLStatement ...;
     preparedSQLStatement: (*) directSQLStatement ...;
     resultSet: (*) ...;
     open:< (*) ...;
     close:< ...;
     connectionException:< BDBCException ...;
     connectionWarning:< BDBCWarning ...;
     private: @<<SLOT ConnectionPrivate:Descriptor>>
  #);
```

Data Sources

In order to use a connection a data source must be created (See the tutorial for specifics.). To communicate with an existing data source the user must first create an instance of the connection pattern. Calling the open method on a connection:

```
open:< (*)
  (# name: ^text;
     userName: ^text;
     password: ^text;
     openConnectionException:< BDBCException (#  do INNER #);
     openConnectionWarning:< BDBCWarning (#  do INNER #)
  enter (name[],userName[],password[])
  <<SLOT ConnectionOpen:DoPart>>
  #);
```

then makes it possible to communicate with the data source. When calling open the name of the data source must be supplied whereas user name and/or password may be omitted as appropriate. Consider as an example the statement

```
('ullman97','marius',none)->sqlCon.open
```

The statement opens a connection to the data source name "ullman97" for the user "marius" without specifying a password.

Shared variables

Definition of shared variables are done via the "declare..." methods. To e.g. declare a text studioName as a shared variable named "studioName" use

```
studioName: @text;
```

```
...

do 'studioName'->declareText
   (# set::(# do value->studioName #);
      get::(# do studioName[]->value[] #)
   #)
```

SQL Statements

Data manipulation and definition is done using the SQL statement patterns. The SQL statement patterns have the following interfaces:

```
SQLStatement:(*)
  (# <<SLOT SQLStatementLib:Attributes>>;
     execute:<(*) ...;
     close:(*) ...;
     execException:< BetaDBCException ...;
     execWarning:< BetaDBCWarning ...;
     private: @<<SLOT SQLStatementPrivate:Descriptor>>;
     get:< ...;
     set:< ...
  enter set
  do INNER
  exit get
  #);

directSQLStatement:(*) SQLStatement
  (# <<SLOT DirectSQLStatementLib:Attributes>>;
     currentMarker:(*) @ ...;
     marker:(*) ...;
     b: (*) marker ...;
     c: (*) marker ...;
     d: (*) marker ...;
     f: (*) marker ...;
     i: (*) marker ...;
     s: (*) marker ...;
     t: (*) marker ...;
     execute::  ...;
     execDirectException:< BetaDBCException ...;
     execDirectWarning:< BetaDBCWarning ...;
     private: @<<SLOT DirectSQLStatementPrivate:Descriptor>>;
     set::<  ...;
     getExpanded:(*) ...
  do INNER
  #);

preparedSQLStatement:(*) ...
```

To use a statement stmt, one must first open it and then associate it with an SQL statement as in

```
stmt.open
'SELECT title, length FROM Movie WHERE studioName = \'Disney\''->stmt
```

Invoking execute on stmt will then cause the SQL statement to be executed at the database. A statement should be closed after use.

A preparedSQLStatement differs from a directSQLStatement in that a prepared statement is parsed and prepared by the data source when the statement is initialised, i.e. executing

the statement above will, if stmt is a preparedSQLStatement, cause the contents of the SQL statement to be sent to the database in order for it to be parsed and prepared for future execution. If stmt is a directSQLStatement no communication with the database will occur before calling

execute.

In this way a prepared SQL statement is a little slower to initialise than a direct SQL statement but much faster to execute. Use a preparedSQLStatement only when an SQL statement has to be executed several times.[1]

The contents of a directSQLStatement can be any SQL statement with embedded shared variables and/or markers, i.e. a directSQLStatement stmt may be initialised as in

```
'SELECT title,length FROM Movie WHERE studioName = :studioName AND year = %i'
  ->stmt
```

Here "studioName" is the name of a shared variable declared as shown above. The value of the %i marker may be set using the i pattern in directSQLStatement. Now suppose that the following statements have been executed

```
do ...; 'Disney'->studioName; 1990->stmt.i; ...
```

When executing the SQL statement the SQL contents of the statement will then conceptually be

```
SELECT title,length FROM Movie WHERE studioName = 'Disney' AND year = 1990
```

i.e., before sending an SQL statement to a database the embedded shared variables and the markers are, conceptually, substituted for their current values. After execution the contents of the statement can be changed, the markers can be reset or the statement can be closed.

describe setXXXByName

A preparedSQLStatement is used similarly to a directSQLStatement.

Results

Executing an SQL statement stmt will yield an instance of resultSet (here rs is a reference to a resultSet):

```
stmt.execute->rs[]
```

A resultSet implements an interface to an SQL cursor in the following way

```
resultSet:(*)
  (# <<SLOT ResultSetLib:Attributes>>;
     columnCount:(*) integerValue ...;
     rowCount:(*) integerValue ...;
     column: (*)...;
     getColumn: (*)...;
     getColumnByName:(*) ...;
     result:(*)
       (# <<SLOT ResultLib:Attributes>>;
          marker:(*) ...;
          b: marker ...;
          c: marker ...;
          d: marker ...;
          f: marker ...;
          i: marker ...;
          s: marker ...;
          t: marker ...;
          private: @...
       #);
     scan:(*)
       (# current: @result;
          varNotDeclared:<(*) exception ...;
```

```
         unknownColumn:<(*) exception ...;
         pattern: ^text
       enter pattern[]
       ...
       #);
     resultSetException:< BDBCException ...;
     resultSetWarning:< BDBCWarning ...;
     private: @...
  #)
```

Given a resultSet the scan method iterates over the tuples in the resultSet. There are three distinct ways to control the scan. Firstly, one may simply execute

```
rs.scan(# ... #)
```

In the do-part of the scan one may then refer to the values of the columns in the result. This is done sequentially by referring to the markers of the current result. If e.g. rs was retrieved as shown above,

```
rs.scan(# do current.s -> putline; current.i->putint; newline #)
```

will scan over the results in the resultSet and print the values of their columns on the screen.

Also, one may enter a string when evaluating a scan pattern as in:

```
':title %i'->rs.scan(# do title->putline; current.i->putint; newline #)
```

Here title is a shared variable named "title". This statement prints the same as above but by providing an input string it is here specified that the first column of each result should be assigned to the shared variable "title" and that the second column of each result is an integer that will be fetched via the i marker. In general one may in this way specify how each column of a result should be treated.

The two ways of scanning shown above may, in some circumstances, be problematic in that they assume a specific ordering of columns in the results. Therefore, the last way of doing a scan names the columns in the resultSet, as in e.g.:

```
'length%i title:title'->rs.scan
   (# do title->putline; current.i->putint; newline #)
```

In this way the order of the columns in the result may be changed from "title, length" to "length, title" without any problems for the last way of scanning.

describe fetch

describe cursorType?

---

[1] Note that, currently, a preparedSQLStatement is implemented as a directSQLStatement.

Transactions

Many data sources support the use of transactions. In order to use this capability from BetaDBC the fragment transactions may be included. An outline of the interface of this fragment is shown below. The full interface may be seen in section 5.3.

```
ORIGIN 'betadbc';
BODY 'private/transactionsbody';
-- connectionLib: Attributes --
transactionsSupported:(*)
  (# isSupported: @boolean
  ...
  exit isSupported
  #);
autoCommitMode:(*)
  (# autoCommit: @boolean
  enter (#  enter autoCommit ... #)
  exit
    (# ...
    exit autoCommit
    #)
  #);
readUncommitted: (*) integerValue (# ... #);
readCommitted: (*) integerValue (# ... #);
repeatableRead: (*) integerValue (# ... #);
serializable:(*) integerValue (# ... #);
transactionLevelSupported:(*) booleanValue
  (# level: (*) @integer
  enter level
  ...
  #);
transactionLevel:(*)
  (# level: (*) @integer
  enter (#  enter level ... #)
  exit
    (# ...
    exit level
    #)
  #);
commit:(*) (#  ... #);
rollBack:(*) (#  ... #)
```

The scope of a transaction is a whole connection including all statements allocated in it. The default is that every execution of an SQL statement starts a new transaction and automatically commits the effects of this statement after the statement has completed. This auto commit mode may be changed to manual commit mode by evaluating

```
false->sqlCon.autoCommitMode
```

where sqlCon is an instance of a connection. Note that this is only meaningful if transactionsSupported evaluates to true. In manual commit mode a series of database manipulations may be committed by executing commit. Equivalently a series of database manipulations may be aborted by executing rollBack.

Four transaction isolation levels (as defined in the SQL standard) are available, namely readUncommitted, readCommitted, repeatableRead, and serializable. Whether a transaction isolation level, such as serializable, is supported in a given connection may be checked by evaluating e.g.

```
serializable->sqlCon.transactionLevelSupported
```

The BetaDBC Library

Scrollable Cursors

Describe scrollable cursors. Problem: they do not work on Access = no testing done so far

The BetaDBC Library

Tutorial

The sample programs shown in the tutorial may be found in the tutorial directory accompanying BetaDBC. The examples use a database schema and examples from [Ullman, 1997]. It supposes that the reader is familiar with basic SQL and focuses on teaching the essentials of using BetaDBC. All examples take (up to) three command line arguments: a data source, a user name, and a password. If the data source used permits it the user name and/or password may be omitted.

Creating a data source

In order to use BetaDBC a suitable data source must be created. Currently, data sources are created outside BetaDBC. Different procedures must be followed depending on the operating system used. This will be changed in a future release of BetaDBC.

Creating a data source on Windows 95 and NT

In the 'Start' menu choose 'Settings' and 'Control Panel'. Start the ODBC (32 Bit) application from the 'Control Panel'. Press 'Add...' and choose the database driver that you want to use, then press 'Finish'. Follow the driver specific instructions.

Creating a data source on Unix

rewrite to match clemen's new setup

You will need a .odbc.ini file in your home directory. Create such a file if it does not exist.

Suppose a driver for the PostgreSQL database is located in "/users/kursus/dprog2/RDB/lib/libcliPG.so". If you want to connect to the database "marius" on the database server "delirium" insert the following in your .odbc.ini file ('ReadOnly = 0' tells the driver that you want to manipulate the "marius" database)

```
[ullman97]
# data source containing the 'marius' db on delirium
# this database contains the movie tables from ullman97
Driver = /users/kursus/dprog2/RDB/lib/libcliPG.so
Database = marius
Servername = delirium
ReadOnly = 0
```

This defines a data source named "ullman97", that you may use BetaDBC to connect to. Each data source at least specifies which ODBC. In the example "ReadOnly" is a driver specific attribute of the data source "ullman97". See the documentation for the drivers used for definitions of driver specific attributes.

In the following it is assumed that a suitable data source named "ullman97" has been created.

Creating a database

Check this section with the new setup

The next step will then be to create and insert values into a database. Let's use the following sample database schema

```
Movie(title, year, length, inColor, studioName, producerCNo)
StarsIn(movieTitle, movieYear, starName)
MovieStar(name,address,gender,birthdate)
MovieExec(name, address, certNo, netWorth)
Studio(name, address, presCNo)
```

A series of SQL statements creating the database schema may be found in createmoviedbtables.txt. The text files used in this section are shown in section 6.

Create the tables corresponding to this schema by running the executesqlfile program also found in the tutorials directory:

```
[postgres@delirium tutorial]$ ./executesqlfile createmovie.txt ullman97 marius
```

How the executesqlfile program is implemented will be discussed later. You may now insert some values in the database by running

```
[postgres@delirium tutorial]$ ./executesqlfile moviedbtables.txt ullman97 marius
```

The tables may later be deleted by running

```
[postgres@delirium tutorial]$ ./executesqlfile dropmovies.txt ullman97 marius
```

Querying and retrieving from the database

This section will introduce the basics of BetaDBC: connecting to data sources, executing simple queries and retrieving the results.

Consider the simple SQL statement

```
SELECT *
FROM Movie
WHERE studioName = 'Disney' AND year = 1990;
```

An application that uses BetaDBC, executes the above query and retrieves the result may look like:

```
ORIGIN '../betadbc';
-- program: Descriptor --
(#
   sqlCon: @connection;
   stmt: @sqlCon.directSqlStatement;
   rs: ^sqlCon.resultSet
do
   (2->arguments,3->arguments,4->arguments)->sqlCon.open;
   'SELECT title, length FROM Movie WHERE studioName  = \'Disney\' AND year = 1990'
     ->stmt.open;
   stmt.execute->rs[];
   rs.scan
     (#
     do
        'title: '->puttext;
        current.s->putline;
        'length: '->puttext;
        current.i->putint;
        newline
```

```
     #);
   stmt.close
   sqlCon.close
#)
```

The program starts out by declaring a connection, a directSQLStatement belonging to that connection and a resultSet belonging to that connection. The connection is used in order to connect to a data source in the first line of the program:

```
(2->arguments,3->arguments,4->arguments)->sqlCon.open;
```

Connections open method takes as arguments a name of the connection, a username and a password. Thus an invocation of the program like

```
[postgres@delirium tutorial]$ ./simple ullman97 marius foobar
```

means that the first statement will be an attempt to connect the user "marius" with password "foobar" to the data source named "ullman97". If this succeeds

```
'SELECT title, length FROM Movie WHERE studioName = ''Disney'' AND year = 1990'
     ->stmt.open;
```

will open the directSQLStatement "stmt" and set its content to the query we want to execute. Executing the query yields a resultSet holding a cursor for the result

```
stmt.execute->rs[];
```

The resultSet may then be scanned. During the scan 'current' will hold a reference to a tuple in the resultSet. The values of this result may then be accessed consecutively by using the marker attributes

```
 rs.scan
     (#
     do
        'title: '->puttext;
        current.s->putline;
        'length: '->puttext;
        current.i->putint;
        newline
     #);
```

Finally, in order to free resources the directSQLStatement and the connection are closed.

Executesqlfile

The simple scheme presented in the last section can now be used for implementing the executesqlfile program. The executeLoop shows how to reuse an SQLStatement by simply replacing it's textual contents

```
executeTxt[]->sqlCon.stmt
```

Embedded SQL

Using shared variables makes it possible to use the values of BETA objects in place of a concrete value in SQL statements. Using BETA and BetaDBC does not include using a preprocessor, which makes it necessary to declare shared variables imperatively, as in

```
   sqlCon:@connection
   studioName:@text;
do ...;
```

```
   'studioName'
     ->sqlCon.declareText
       (# set:: (#  do value->studioName #);
          get:: (#  do studioName[]->value[] #)
       #);
   ...
```

Here a shared text variable named "studioName" is declared. The 'set' pattern is final bound to describe how the shared variable's value is to be set. 'get' is final bound to describe how the value of the shared variable is to be fetched.

Then, using embedded SQL syntax, one may use shared variables in SQL statements:

```
   stmt:@sqlCon.directSQLStatement;
do ...;
   'INSERT INTO Studio(name, address) VALUES (:studioName, :studioAddr)'
     ->stmt.open;
   ...
```

This means that when executing stmt, ":studioName" and ":studioAddr" will (conceptually) be replaced by the values of the BETA text variables "studioName" and "studioAddr", and the resulting SQL statement will then be executed.

Using BetaDBC it is possible to declare most commonly used objects as shared variables (i.e., boolean, integer, real, text, date and time). The figure below shows a full program that will execute the statement above. "stmt.getExpanded" returns in a text how the SQL statement would look if it was executed at that point.

Embedded SQL

Include use of "fetch"

Suppose that we are executing a statement that return a result t statement. Embedded SQL can then also be used to fetch results directly into shared variables. In BetaDBC this is done through the use of the scan pattern. Suppose we are executing

```
'SELECT MovieExec.name, netWorth FROM Studio, MovieExec WHERE presCNo = certNo AND Studio.name
```

Then,

```
':presName :presNetWorth'
     ->(stmt.execute).scan
       (# ... #)
```

will cause the first column in each result tuple to be assigned to the shared integer variable "presName", and the second column to "presNetWorth". The full code is shown below:

```
ORIGIN '../betadbc';
INCLUDE '~beta/basiclib/formatio';
-- program: Descriptor --
(# sqlCon: @connection;
   stmt: @sqlCon.directSqlStatement;
   studioName,presName: @text;
   presNetWorth: @integer
do 'studioName'
     ->sqlCon.declareText
       (# set::  (#  do value->studioName #);
          get::  (#  do studioName[]->value[] #)
       #);
   'presName'
     ->sqlCon.declareText
```

```
      (# set::  (#  do value->presName #);
         get::  (#  do presName[]->value[] #)
      #);
   'presNetWorth'
     ->sqlCon.declareInteger
      (# set::  (#  do value->presNetWorth #);
         get::  (#  do presNetWorth->value #)
      #);
   (2->arguments,3->arguments,4->arguments)->sqlCon.open;
   'Input a studio name: '->puttext;
   getLine->studioName.puttext;
   'SELECT MovieExec.name, netWorth FROM Studio, MovieExec WHERE presCNo = certNo AND Studio.na
     ->stmt.open;
   stmt.getExpanded->putline;
   ':presName :presNetWorth'
     ->(stmt.execute).scan
       (#
       do 'The net worth of the president %s \nof %s is %i $\n'
            ->putFormat
              (#  do presName[]->s; studioName[]->s; presNetWorth->i #)
       #);
   stmt.close;
   sqlCon.close
#)
```

An ad hoc query evaluator

We now have most of the building blocks to create an ad hoc query evaluator, i.e. a program that connects to a data source and in a loop prompts for SQL statements that are to be executed on this data source. The following implements such a program.

As long as the user inputs anything but an empty line this input is sent to the data source as an SQL statement:

```
getline->stmt;
(if (stmt).empty then leave L if);
stmt.execute->res[];
```

If successful, the result is examined. First the column information of the resultSet is extracted:

```
(for j: res.columnCount repeat
  '%s: %s\t'->putFormat
              (#
              do (j->res.getColumn).name[]->s;
                 (j->res.getColumn).dataTypeName[]->s
              #)
for);
```

Each resultSet has a columnCount yielding the number of columns in the resultSet. For each column, information such as name and dataTypeName (a DBMS specific datatype name) may be retrieved.

If the columnCount is non-zero, the results are fetched:

```
(if res.columnCount > 0 then
    res.scan
      (#
      do (for i: res.columnCount repeat
              (if (i->res.getColumn).DataType##
               // text## then
                  current.s->puttext
               // integerObject## then
                  current.i->putint
```

```
                // realObject## then
                   current.f->putreal
                // booleanObject## then
                   (if current.b then
                       'true'->puttext;
                    else
                       'false'->puttext
                   if)
                // time## then
                   current.t->puttime
                else
                   'Unknown data type!!!'->puttext
               if);
               '\t'->puttext
         for);
         newline
      #)
   else
      'DML/DDL statement executed successfully!'->putLine
if)
```

Again the information about columns in the resultSet is used. By evaluating `(i->res.getColumn).DataType##`

the BETA pattern corresponding to the SQL datatype in column i is found.

Text files for the tutorial

createmoviedbtables.txt

```
CREATE TABLE MovieStar (
    name CHAR(30),
    address VARCHAR(255),
    gender CHAR(1),
    birthdate INTEGER
);

CREATE TABLE Movie (
    title VARCHAR(255),
    year INTEGER,
    length INTEGER,
    inColor BIT,
    studioName CHAR(50),
    producerCNo INTEGER
);

CREATE TABLE StarsIn (
    movieTitle VARCHAR(255),
    movieYear INTEGER,
    starName CHAR(30)
);

CREATE TABLE MovieExec (
    name CHAR(30),
    address VARCHAR(255),
    certNo INTEGER,
    netWorth INTEGER
);

CREATE TABLE Studio (
    name CHAR(50),
    address VARCHAR(255),
    presCNo INTEGER
);
```

deletemoviedbtables.txt

```
DROP TABLE MovieStar;

DROP TABLE Movie;

DROP TABLE StarsIn;

DROP TABLE MovieExec;

DROP TABLE Studio;
```

populatemoviedbtables.txt

The BetaDBC Library

References

[Date 1993] Date, C.J. and Darwen, H., A Guider to the SQL Standard, Addison Wesley, Reading, MA, 1993.

[Geiger 1995] Geiger, K. Inside ODBC, Microsoft Press, 1995.

[Ullman 1997] Ullman, J.D., Widom, J., A First Course in Database Systems, Prentice Hall International, 1997.
