Technische Universität Braunschweig
Institut für Informationssysteme
http://www.ifis.cs.tu-bs.de

Wolf-Tilo Balke, Sascha Tönnies, and Benjamin Köhncke

12.01.2012

SQL Lab: Assignment 4
(due to 24. / 25.01.2012)

General Information

This week, we will transform the model of the chemical digital library from the last assignment into a relational model in a DB2 database. Therefore, we will need a SQL client (e.g. Netbeans) again. For installation instructions see last assignment.

Tasks

All of the following tasks should be performed only for the documents that were published between **01.01.2006 and 31.12.2008**. All other tables should only include data related to these documents.

Create Tables

Use SQL CREATE TABLE statements to re-create tables which can store the respective data inside your own DB2 schema (i.e. the one matching your login name, e.g. DBLAB99). **Include all statements** you used in your result sheet. Consider the following requirements:

- **Leave out the "Entities"** table and **replace this table with the "pubchem_compound"** table. Thus, the "Similarity" table will be removed and the mapping between documents and pubchem compounds has to be calculated and filled in.
- **Extend** the "WatchListDocument" and "WatchListEntity" with the result that one can rate the entry with a value between 0 and 1.
- The model may contain data types not available in DB2, thus please use **appropriate DB2 data types** were needed.
- Each created table must define a **primary key**.
- Every time information in multiple tables is related, define **foreign keys**.
- Add simple consistency constrains to your tables (i.e. UNIQUE, NOT NULL, simple CHECK conditions, …) wherever appropriate.
- **Do not import any data which is not related to any CHEMDB documents from 01.01.2006 to 31.12.2008!**

Additional Data

Create a new UserAccount entry having your group account name and the same password. This user should have been assigned to the user role. In addition add a tutor account (same password) with the admin role assigned.

Creating Views

Create SQL queries for the following questions. Use your own schema! **Include** the used **statements** and the number of resulting tuples in your result sheet.

a) Create a view, just showing all documents crawled by the "Beilstein crawler".
b) Create a view, showing each document id and the number of contained pubchem compound entries.
c) Create a materialized view containing all attributes from the document table and the respective Journal Name, Publication Year and the publisher.

Sending in Results

Print your statements on paper and hand them over to your tutor. For creating the result documents, please use a word processor and format your statements in a readable fashion (e.g. you may want to try the following online formatter: http://www.dpriver.com/pp/sqlformat.htm).

Additionally , please send your statements as plain text file (.txt) **via email to your tutor**.

If you fail to either send your solution via email or do not format them nicely on your result document, you **will be graded with -1** !

Please note, that typically there is a high load on the database server on the day before the due date. Therefore, we strongly recommend starting early working on this assignment.