VeraSQL: An educational client for relational databases

André Greubel[1], Theresa Rudolph[1] and Martin Hennecke[1]

Abstract: In this paper, we present VeraSQL, an educational database client specifically designed to support the transition between spreadsheets and relational databases as taught in many German school curricula. Most importantly, VeraSQL can execute SQL queries on CSV-files. VeraSQL is written in Java, open source, and can be used and downloaded free at www.verasql.de.

Keywords: SQL, query languages, educational database client, database curriculum

1   Overview

In German school curricula, lessons on databases and SQL follow lessons on spreadsheets and corresponding software like Microsoft Excel or Libre Office Calc. While learning about these topics, learners often contact ideas and methods, which can also be utilized for databases. However, the current set of professional and educational tools commonly used in classrooms neither supports this transition nor highlights similarities between these topics.

Educational software usually only offers an *educational SQL environment*, rather than a full *database client*. Most importantly, it is not easily possible to select a custom data source (e.g., data from spreadsheets) in most environments. For example, it is not possible to change the data source in BlocklySQL [Pö19] at all. Other solutions like DBSnap++ [Si] can import a spreadsheet as data source, but are unable to easily build and execute SQL-queries. Due to this lack of suitable educational software, some teachers and schoolbooks (e.g,. [Br07]) even opt to use professional database management software instead.

Our educational tool, VeraSQL[2], was explicitly designed to close this gap by:

- Offering a visual user interface for SQL-queries. This interface, like block-based environments, enables the creation of queries with drag-and-drop.

- Offering a custom SQL-compiler able to execute a given SQL-query step-by-step on CSV-spreadsheets.

---

[1] University of Wuerzburg, Institute for Computer Science, Mathematik West (30), 97074 Wuerzburg
{andre.greubel, martin.hennecke}@uni-wuerzburg.de, theresa.rudolph@stud-mail.uni-wuerzburg.de
[2] Vera is a German Acronym for "visual creation of queries" (german: [V]isualle [Er]stellung von [A]bfragen)

- Simplifying presentation of exercises by showing a given query simultaneously with its result.

2   VeraSQL

In VeraSQL, queries are displayed on the left and can be executed and stored to / loaded from an SQL-file with the left buttons in the bottom control area. Clicking on the shape of an expression in the bottom control area adds a condition to the end of the corresponding expression in the query display area on the left.

The right buttons in the QueryControlArea allow for the selection of a data source. VeraSQL can either be connected to a database server or use local data sources.

The top button connects to a database server. In this case, queries are sent to the server and the result is displayed in the MainArea.

The lowest button in the right row loads a single CSV-file as a table into VeraSQL. Similarly, the middle button loads every CSV-file in a directory as a table into VeraSQL. In these cases, VeraSQL automatically enters *local database mode*. In this mode, a custom SQL-compiler is used to locally execute the query on the selected spreadsheet-tables (instead of sending it to a database server).

This local database mode is one of the key features of VeraSQL and currently responsible for roughly 40% of the 13.000 nonempty lines of Java-Code.

The local compiler responsible for the execution of the SQL-query is based on a simplified custom grammar in ANTLR. Most of the time, the compiler follows SQL-compliant behaviour. However, as CSV-files do not specify data formats, VeraSQL automatically infers them. If every entry in a column can be parsed to a `BigDecimal`, the column is represented that way. Otherwise, the column type defaults to a `java.lang.String`. Because of this, we recommend formatting dates and times using ISO 8601 in the user data to prevent non-standard compliant sorting.

A more detailed technical description of VeraSQL (including limitations and technical details about the compiler), as well as VeraSQL itself, is available at www.verasql.de.

Bibliography

[Br07]   Brichzin, P. et al.: Informatik. Oldenbourg, München, 2007.

[Pö19]   Pöhner, N. et al.: BlocklySQL: A new block-based editor for SQL: Proceedings of the 14th Workshop in Primary and Secondary Computing Education, 2019; pp. 1–2.

[Si18]   Silva, Yasin N. et al.: DBSnap++: Proceedings of 23rd Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE'18); pp. 170–175.