CMPSC380
Principles of Database Systems, Fall 2014

Final Project: Advanced Topics in Data Management

Purpose: DeeBee: Implementation of A Relational Database Management System

Pledge:

Hand in: Friday, December 12th, 2014

Abstract: DeeBee is a small relational database management system implemented for educational purposes. It implements a subset of the structured query language, enough to support simple database operations; and is designed for modularity, so that additional advanced database features can be added in the future.

1 Introduction

Relational database management systems (RDBMSs) are everywhere. The relational model is a model of data storage in which data is stored in *tuples*, or rows, which are grouped together to form *relations*, or tables [2–4]. The relational model is perhaps the most popular models of data storage currently in use, with Silberschatz, Korth, and Sudarshan calling it "[t]he primary data model for commercial data-processing applications" [4, page 39].

A majority of modern relational database management systems, from the SQLite embedded database in every Android phone and iPhone [5] to the MySQL databases used in many web applications [1], implement the *Structured Query Language*, or SQL. SQL is a *query language*; domain-specific declarative programming language that is used by database administrators, users, and application software to interact with a database [4].

In order to learn more about how such SQL databases function 'under the hood', I have implemented my own small RDBMS, called DeeBee. DeeBee implements a small subset of the SQL, chosen to be expressive enough to allow basic database operations to be performed but minimal enough to allow DeeBee to be implemented within the constraints of the Computer Science 380 Final Project. Implementing DeeBee has yielded many insights into the challenges, techniques, and patterns involved in the design and implementation of an RDBMS.

At the time of writing, the DeeBee codebase comprises almost 1700 lines of Scala source code and over 500 lines of comments. While this is small compared to many 'real-world' systems, it still represents a significant undertaking for a single individual over a 24-day period. Therefore, I have made use of a number of software engineering tools and practices in order to best maximize my productivity. While these techniques are not the focus of this assignment, I will touch on them briefly as well.

2 DeeBee

References

[1] Dale Dougherty. LAMP: The Open Source Web Platform - O'Reilly Media, 2001. `http://www.onlamp.com/pub/a/onlamp/2001/01/25/lamp.html`.

[2] Hector Garcia-Molina, Jeffrey D Ullman, and Jennifer Widom. *Database System Implementation*, volume 654. Prentice Hall Upper Saddle River, NJ:, 2000.

[3] Jan L Harrington. *Relational database design and implementation: clearly explained*. Morgan Kaufmann, 2009.

[4] A. Silberschatz, H. Korth, and S. Sudarshan. *Database System Concepts*. McGraw-Hill Education, 2010.

[5] SQLite. Famous Uses of SQLite, 2014. `https://www.sqlite.org/famous.html`.