Journal of Computer Science 1 (3): 376-380, 2005
ISSN 1549-3636
© Science Publications, 2005

Automatic Discovery of Association Paths in Relational Databases Using Software Visualization

Haider Ali Ramadhan
Department of Computer Science, Sultan Qaboos University, P.O. Box 36, Muscat 123, Oman

Abstract: We introduce a visual framework for facilitating tasks associated with database maintenance and re-use. The prototype system embodying the framework is presented. The system utilizes various techniques and features of software visualization. The system supports visual displays of the database structure along with various implicit relationships found in it such as associations and path views. Information visualized is automatically extracted from the database schema. To assess the usefulness of the proposed framework in helping the programmers to quickly recognize path views among relations, an empirical evaluation was conducted. Results collected from the evaluation seem to support our hypothesis that the time required to manually recognize path views from the database schema is considerable and tends to increase as the depth between the relations increases. The evaluation also showed that by using our visual framework such time is negligible and tends to be static.

Keywords: Software Re-Engineering, Databases, Software Visualization, Visual Interfaces, Empirical Evaluation

BACKGROUND

Large relational databases are inherently complex. Their evolution over time may increase the difficulty of understanding them even more. Due to this evolution, their structures may degrade over time to an extent that the tasks involved in maintaining and re-engineering these databases could become very costly. Therefore, techniques that may aid the designers and engineers in the analysis of these information systems deserve special attention and research focus. The task of understanding the structure and design of a relational database system is a complex one. There are many dimensions of complexity, three important among which are:

* The overall structure of the database. This includes the set of relations, their attributes and the types of data from which the attributes may take values.
* Various links between relations. This includes the number of associations between relations and their implicit nature.
* Various direct and indirect paths among relations in the database.

To help the developers get clear understanding about the above three dimensions, we have developed a prototype system using software visualization framework [3, 6]. The system is able to visually display the structure, i.e. relations and attributes, of a relational database in a limited space. In addition, the system visually shows the implicit relationships among relations in terms of associations and path views. The visualization is organized in separate views, each dealing with one of the dimensions mentioned above. These views include the overall view, association view, detailed view and path view. All these views are automatically visualized from the specifications of the database and hence provide a general tool for visualizing any relational database. In summary, our aim is to use software visualization [2, 4, 8] features to facilitate re-engineering, maintenance and analysis of the structure and relationships found in existing relational databases.

Software Visualization [6, 8, 9, 10, 11] refers to the task of applying visual and graphical techniques to exhibit the static structure and the dynamic behavior of software systems. The main purpose of software visualization is to provide designers, engineers, programmers and users with visual aids to help them in understanding and analyzing the structure and behavior of a software. This task is achieved through abstracting low-level textual structures, i.e. code and data, into high-level visual representations, hence reducing the mapping and interpretation load.

Figure 1 shows the database composition of a small firm with 19 relations. The relations are drawn as colored dots, where the color gives the range of the attributes in a relation as per the criteria: relations with green color have attributes ranging from 1 to 5, those with blue color have attributes ranging from 6 to 10 and those with red color have attributes greater than 11. It is worth noting that these dots are randomly displayed in the view. To accommodate databases with hundreds of relations, the size of the dots is automatically reduced and the scroll bars are introduced. Generally speaking,