A CAI Tool for the Theory of Relation Normalization

Fangjie Xu, Huichuan Duan
School of Information Science and Engineering, Shandong Normal University, China
Fangjiexu@sdnu.edu.cn

Abstract

Teaching and learning the theory of relation normalization in a database course is a nontrivial work. Although this theory and the related concepts and algorithms have been clearly stated in authoritative texts, suitable CAIs have not been found to our best knowledge. This paper reports a CAI tool, RDBNorm, the authors designed and implemented to facilitate both teachers' classroom teaching and students' home practicing of the theory of relation normalization. RDBNorm demonstrates relation normalization associated algorithms step by step to make them easy to understand, as well as a theory tutorial presenting the key points of the theory.

Keywords: CAI tool, theory of relation normalization, closure of attributes set, function dependency, normal form.

1. Introduction

The theory of relation normalization is the core of the relational database course. Authoritative textbooks that clearly and sophisticatedly describe this theory have been elaborately written and improved edition by edition and widely accepted in computer science related teaching in universities and colleges. In China, "An Introduction to Database Systems" by Shan Wang et al [1] and "Database Systems" by Bole Shi et al [2] are two of the best preferred textbooks.

The relation normalization theory is vital in database teaching and learning, and it is also a very much difficult work for both teachers and students. This issue has been getting more and more critical as the high education is stepping from elite education toward popular education since China initiated and proceeded to the plan of enlarging the enrollment of universities and colleges. To solve this harassment, suitable CAI tools facilitating both teachers and students should be designed.

Some of the popular texts provide accompanied instruction and solution manual texts, e.g. Shan Wang [3] wrote "An Introduction to Database Systems – Learning Instruction and Problems Solution Manual" to ease the use of her text. Some teachers [4][5] have built Web sites providing tutorials in relation normalization for their course.

However, best to our knowledge, no one has reported a dedicated CAI for the teaching of relation normalization theory. Since a dedicated CAI can definitely help to improve the efficiency, effectiveness and quality of this hard work, we initiated a plan to design and implement a dedicated CAI for the teaching and learning of the theory of relation normalization. This paper reports our work on this CAI.

The remainder of the paper is organized as follows: Section 2 reports the design outlines; Section 3 presents the system design; Section 4 expresses the part of theory tutorial; Section 5 details the implementation; and finally in the last section, we give a summary of the work and tell our future plan.

2. Design outlines

For simplicity, we name "CAI tool for the theory of relation normalization" as RDBNorm.

RDBNorm aims at the key points in the theory of relation normalization, which includes two classes of concepts: FD (function dependency), and NF (normal forms, mainly 1NF, 2NF, 3NF, and BCNF), and the related algorithms to reduce a relation into different levels of NFs, which include the algorithm to search for the closure of a function dependency set, the algorithm to search for the closure of an attributes set, the algorithm to search for the minimal dependence set of a function dependence set, and algorithms to decompose a relational schema into 2NF, 3NF, and BCNF with or without loss.

All these knowledge have been well defined, so our work is to pick them from authoritative texts, of which Wang [1] and Shi [2]'s book are the most representative in Chinese.

The next crucial work in RDBNorm is choosing representative relation schema examples to demonstrate the key algorithms. Again, we choose the examples in Wang [1] and Shi [2]'s texts because their examples are not only representative, but also frequently illustrated in classroom teaching because of their popular utilization. Of course, as a CAI tool, we allow users to add their new examples as many as they like.

As a CAI tool, our first design aim of RDBNorm is presentation and demonstration, the next is tutoring and

practicing, that is, providing a tool not only to facilitate teachers' classroom teaching, but also to help students practicing system contained examples as well as newly found or designed examples.

3. System design

RDBNorm is designed to consist of 4 main modules: theory tutorial, relation schemas management, FD analyses, and NF analyses and decomposition.

In theory tutorial, we organized the key theoretical issues in a manner to ease teaching and learning. Relation schemas management provides a conventional records and table operation, i.e. addition, alteration, deletion, and presentation, for the example relations with their FDs. FD analyses contains two sub-modules: searching for the closure from an attributes set and searching for the minimal function dependence set from a given FD set. NF analyses and decomposition contains three sub-modules: 2NF analyses and decomposition, 3NF analysis and decomposition, and BCNF analyses and decomposition.

Fig. 1 shows the function structure of RDBNorm CAI Tool.

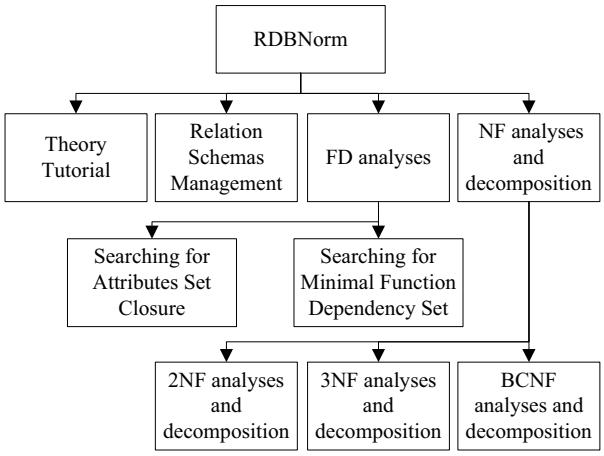

Fig. 1 RDBNorm modules

4. Theory tutorial

We have decomposed the theory of relation normalization into 18 items: (1) Function dependency; (2) Full function dependency and transitive function dependency; (3) Key (candidate key, primary key, and external key) and 1NF; (4) 2NF and its decomposition algorithm; (5) 3NF and its decomposition algorithm; (6) BCNF and its decomposition algorithm; (7) Logical implication and Armstrong axiomatic system; (8) The five inferential rules of Armstrong axiomatic system; (9) The closure of function dependency set; (10) The closure of an attributes set and the generating algorithm; (11) The minimal dependence set of a function dependence set and the generating algorithm; (12) The decomposition of relational schema and lossless decomposition; (13) The algorithm for discriminating lossless decomposition; (14) The sufficient and necessary conditions for lossless decomposition and the decomposition for preserving function dependency; (15) The decomposition algorithm for preserving function dependence in 3NF; (16) The lossless decomposition algorithm for preserving function dependence in 3NF; (17) The lossless decomposition algorithm for BCNF; (18) Universal relations.

For the ease of classroom presentation, theory tutorial is designed into tabs, each tab contains one theory item. Fig. 2 shows the interface.

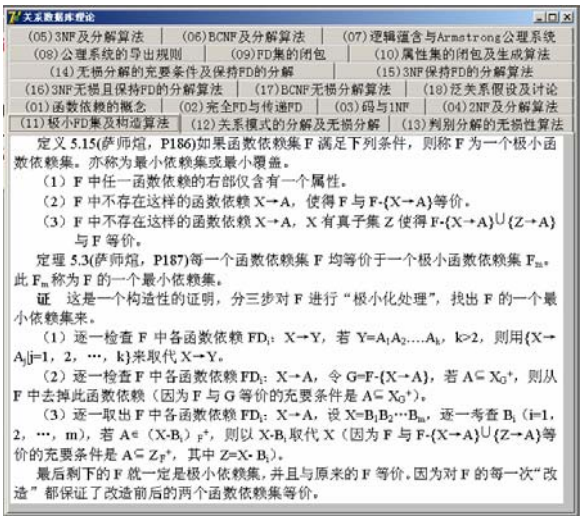

Fig. 2 Theory tutorial interface

5. Database design

In order to store and manage the many and ever increasing numbers of relation schemas to be taught and analyzed, a database RDBNorm is designed. RDBNorm consists of 3 tables: *r_gxms*, *r_FD*, and *r_Key*.

r_gxms table is introduced to store the relation schemas, which is defined as:

r_gxms(xh, mc, attr, mca, ms, ly_zz, ly_ym, ly_wz)

where *xh* is the key, *mc* and *mca* are the relation name and description, *attr* is the attributes list. Other fields describe the source of the relation.

r_Key table is designed to store the candidate keys of a relation schema, which is defined as:

r_Key(xh, gxmsxh, key)

where *gxmsxh* is a foreign key associating *r_gxms.xh*, and *key* is a list of attributes constituting the key.

r_FD table is used to describe the function dependencies, which is defined as:

r_FD(xh, gxmsxh, FDx, FDy)

where *gxmsxh* is a foreign key associating *r_gxms.xh*, *FDx* and *FDy* represent the left and right side of a function dependency respectively.

RDBNorm database is implemented in Microsoft SQL Server 2000. Fig. 3 shows its diagram, in which keys and foreign keys are clearly presented.

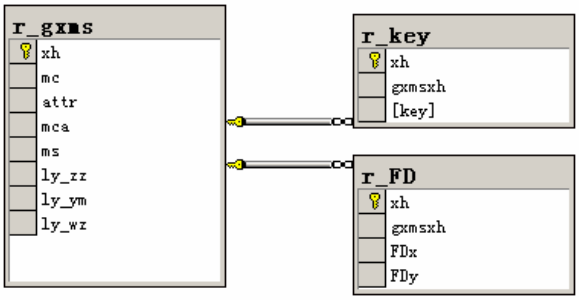

Fig. 3 RDBNorm database design

6. The implementation of the algorithms in relation normalization

As CAI software, RDBNorm provides demonstrative show of the detailed steps of the algorithms. To achieve this spirit, we have designed a suitable interface scheme for all the algorithms. In the scheme, a window is divided into three parts, a toolbar providing devices for users to operate, a frame for the description of the algorithm helping teachers to elucidate the ideas, and a frame for step by step demonstration of the execution of the algorithm.

6.1 An example algorithm implementation -- searching for the closure of an attributes set

All the algorithms in RDBNorm have been implemented in terms of the design scheme we have just told. In this section we illustrate one of them, the algorithm for searching for the closure of an attributes set, as an example. Fig. 4 shows the interface. Of course, we must use Chinese in a class of Chinese students.

As illustrated in Fig. 3, the process of the algorithm running is presented in a tree structure, which is a much friendly user interface design.

6.2 The algorithms implemented

In the current version of RDBNorm, we have implemented the following algorithms besides the one we have just delineated in the above section: the algorithm to search for the minimal dependence set of a function dependence set, and algorithms to decompose a relational schema into 2NF, 3NF, and BCNF with or without loss.

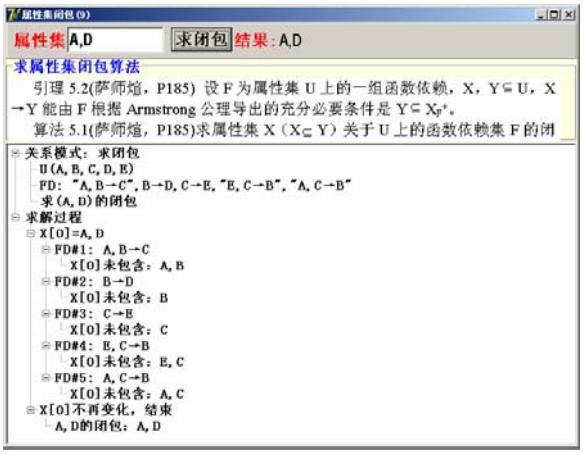

Fig. 4 Interface for searching for the closure of an attributes set algorithm

All these algorithms are implemented in an interface style shown in section 6.1.

7. Summary

By step by step demonstrating of the key algorithms, such as the algorithm to search for the closure of a function dependency set, the algorithm to search for the closure of an attributes set, the algorithm to search for the minimal dependence set of a function dependence set, and algorithms to decompose a relational schema into 2NF, 3NF or BCNF, RDBNorm largely facilitates relation normalization teaching and learning in a database course. Our classroom practice does show that RDBNorm provides expected help for both classroom teaching and home learning. Our future work includes adding more sophisticated algorithm, improving the display interface, supplying more examples, and providing more functionalities for students.

References

[1] Shan Wang and Shixuan Sa, "An Introduction to Database Systems", 4th Ed., Higher Education Press in China, Beijing, China, 2006.

[2] Bole Shi, Baokang Ding, and Wei Wang, "Database Systems", 2nd Ed., Higher Education Press in China, Beijing, China, 2003.

[3] Shan Wang, "An Introduction to Database Systems – Learning Instruction and Problems Solution Manual", 3rd Ed., Higher Education Press in China, Beijing, China, 2003.

[4] Jan Lemeire, Relational Database tutorial, http:// parallel.vub.ac.be /documentation/database/tutor/, 2003.

[5] Osmar Zaiane, Relational Database Design, http: //www.cs.sfu.ca /CC/354/zaiane/material /notes/Chapter7/ node1.html, 1998.