Teaching Data Normalization: Traditional Classroom Methods versus Online Visual Methods – a Literature Review

Shiu Ram

Lecturer

School of Computing and Information Technology

Unitec New Zealand

E-mailAddress: **sram@unitec.ac.nz**

Abstract

Data normalization is one of the topics that is invariably found in prescriptions for database courses offered at tertiary institutions. Data normalization tries to minimize duplication in data storage by attempting to eliminate insert, delete and update anomalies and results in simple relations with mainly one to many relationships. According to Hoffer, Prescott and McFadden (2002, p. 189): data normalization is "the process of decomposing relations with anomalies to produce smaller, well structured relations" .

Until recently, traditional classroom methods of teaching have been used to enable students to do data normalization. However, now with the advances in software development and the power of the Internet to make available online learning resources, students have the opportunity to choose from alternative teaching-learning tools and techniques. Online visual methods provide one such tool that students can use to learn data normalization.

It will be interesting and informative to find out about online visual methods of teaching students data normalization and compare their effectiveness in facilitating learning with those of traditional classroom teaching methods.

This paper will attempt to compare the merits of online visual methods with those of traditional classroom teaching methods of teaching and learning data normalization using the information obtained from literature review.

Keywords: Anomalies, data normalisation, online visual methods.

1   Introduction

In the current information age databases form an essential component of information systems. As such, programmes at tertiary institutions that lead to majors in computer science and information systems invariably include database courses at various levels. One of the main topics included in a prescription for a database course is data normalisation. What is a database and what is data normalisation?

A database is an organised collection of related data. Data are processed to get information which is regarded as the fourth M in an organisation (the traditional three Ms being Men, Money and Machines). Information is used for making decisions which lead to actions. Correct, complete and timely information enable managers to make sound decisions, in good time, that lead to the success of their organisations and competitive advantage over their competitors.

The quality of information depends on the accuracy, integrity and consistency of the data in a database. Ideally, this can be achieved if one item of data is stored only once. In reality this may not be possible as some duplication of data may be necessary in certain database architectures, e.g. relational databases make use of tables to store data and establish relationships between the tables by common fields (columns, or attributes).

Association for Computing Machinery, Inc and Institute of Electrical and Electronic Engineers, Inc. (2001) has listed Relational Databases as one of thirteen topics that have increased in importance due to advances in technology over the past decade. This may well be an indication that data normalisation will continue to be one the topics in a database course. If data duplication in relational databases cannot be avoided, then minimizing it should be the objective that any database designer should try to achieve. Data normalisation would help the designer to achieve this objective.

The designers of database would normally be graduates majoring in computer science or information systems. It is most likely that the old graduates would have been taught how to do data normalisation in the traditional

classroom/lecture theatre teaching-learning environment. For the current students the Internet has the potential to provide an alternative environment in which the teaching is facilitated by software and the students take control of and responsibility for their learning. As this is an emerging alternative, its effectiveness is yet to be determined.

Data normalisation is a process that can be used by the designer of a database to identify and group together related data elements (attributes) and establish relationships between the groups. In relational database each group of related data is given a name (table name) and a unique identifier (primary key). Relationships between tables are formed by including the primary key of one table (parent table) as an attribute in another table (Child table). This is where the necessary duplication of data occurs.

Hoffer et al. (2002, p. 190) have depicted the process of data normalisation diagrammatically as shown below:

Figure 5-22

Steps in normalization
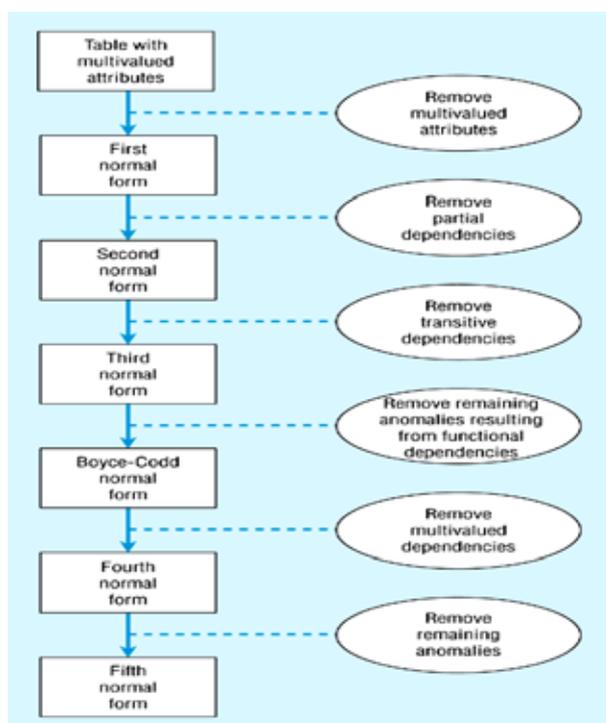
Students encountering the process of data normalisation for the first time find it difficult to understand the concepts of anomalies and dependencies and consequently are not able to normalise a given set or un-normalised data. This has been the experience so far as an alternative method of teaching this subject area has not been thoroughly tested as yet.

However, like many others Zhang, Kaschek and Kinshuk (2005) implemented a prototype to explore the potentials of knowledge management in teaching data normalisation. A literature review of all such attempts would provide some insight into the variety and quality of online systems that are currently available. An analysis of the traditional and online visual methods of teaching

data normalisation may shed some light on what could be the better alternative.

2 Traditional classroom methods of teaching data normalisation

It is a well known fact that in the traditional classroom teaching-learning environment there is face-to-face interaction between the teacher and the students. In this environment, data normalisation would be introduced to the class in the form of a lecture that would make use of teaching aids like PowerPoint slides, white boards and handouts. The teacher would try to explain the concepts to the students, work out examples and set practice exercises to be done in class and for homework. During a class session the students have the opportunity to ask questions on what they do not understand. When the students do the exercises in the class the teacher is available to help them out of their difficulties on the spot. However, when students do the home work exercises and encounter difficulties, they may have to wait until the next class to get the teacher's assistance.

2.1 Advantages of the traditional methods

Human interaction in a face-to-face situation is perhaps the best advantage of the traditional methods. This interaction can between the teacher and the whole class or an individual student. There is also the possibility that whole class can benefit from an answer that the teacher gives to an individual student's question. The interaction can continue until students are satisfied with explanations that help them to understand the subject. The teacher and the students simultaneously experience a mutual sense of satisfaction on overcoming a difficulty or solving a problem.

2.2 Disadvantages of the traditional methods

The major disadvantage of traditional methods is the time constraint imposed by the timetable – a session has a start time and an end time. This constraint may prevent prolonged discussions on a particular problem and therefore it may not be solved to the satisfaction every student in the class.

Individualized attention would help to alleviate this difficulty, but it may not be possible, especially in large classes. Also, the teacher's time is divided into other activities like preparation, marking, attending meetings, doing research, upgrading qualifications and professional development. This places a limit on the time that the teacher can give for face-to-face interactions to individuals or groups of students over and above the timetabled teaching hours.

3 Online visual methods of teaching data normalisation

According to Association for Computing Machinery, Inc and Institute of Electrical and Electronic Engineers, Inc. (2001, p. 9) "Modern networking technology enhances everyone's ability to communicate and gives people throughout the world unprecedented access to information. In most academic programs today—not

only in computer science but in other fields as well—networking technology has become an essential pedagogical tool". Tools like Visible Analyst Database Engineer and NORMIT exist today for students to learn data normalisation. Visible Analyst Database Engineer is a software that Database Administrators, Data Architects, Data Analysts and Data Modellers can use as a tool to perform in their jobs. NORMIT is a Web-Enabled Tutor for Database Normalization.

In online visual methods purpose-built software take the place of the teacher. To learn data normalisation, a student would have to run the software on appropriate hardware. Human-machine interaction takes place and the students learn at their own pace and at times that suit their convenience. As such this is a flexible earning environment (any time (24 x 7), any place (globally)) in which an individualized learning takes place.

3.1   Advantages of the online visual methods

According to Mitrovic (2002) web-enabled educational systems offer several advantages like minimising the problems of hardware and software compatibility and distributing software to individual users. As such the students can access web-enabled tutors from anywhere at anytime.

Apparently, the best advantage of this method seems to be its flexibility. It would also imply that students would take responsibility for their learning.

Some students may find it less intimidating to work with a machine than with people. The machine will not get tired of repeating something and the possibility of negative reactions that may emanate from a human teacher does not exist.

When students use this method, the teacher may be more available to attend to their problems on individual basis.

Tertiary institutions also stand to gain from online delivery of instructions. Minaei-Bidgoli, Kortemeyer & Punch (n.d., p. 1) observe that "online delivery of educational instruction provides the opportunity to bring colleges and universities new energy, students and revenues".

3.2   Disadvantages of the online visual methods

The obvious disadvantage is the absence of peers. This can lead to isolation of students in terms of socialisation. Shank (2004, p.1) believes that "existing tools for online social interaction in instructional environments are inadequate for the wide range of interactions that are needed and many social constructivist and computer supported collaborative learning (CSCL) researchers and theorists agree".

Variability in the quality of software may have implications on its effectiveness, i.e. every tool available need not necessarily be the right one. For selecting a suitable software, evaluation would need become necessary.

It may not be popular with students who learn better in face-to-face interactive environments.

Being an emerging technology tool its effectiveness is unproven.

4   Which method is better?

At this point in time it is difficult to answer this question as online visual methods have not been in use for a suitable length of time to provide evidence that they are any better than the traditional methods.

It can be argued that in future, while online visual methods of learning data normalisation may become more effective than the traditional methods, it may be to the detriment of human socialisation due to human machine interactions rather than human to human interactions.

5   Conclusions

Relational databases will continue to be an important subject area in computer science curriculum. Data normalisation is an essential topic in this subject area. To-date teaching-learning experience indicates that students find it difficult to understand the concepts of anomalies and to normalise data to obtain smaller well structured relations.

Until recently, traditional classroom methods have been used to teach students how to normalise data. Web-enabled tools are now becoming available that can take place of the teacher. As they go in use, evidence may be collected to show whether they are any better than the traditional methods of teaching data normalisation. As such this is a potential area of research and I have taken it up as a topic for my dissertation for Master of Computing qualification.

6   References

Cheap Software Megastore Ltd. (2005-2007): Visible Analyst Database Engineer Edition v7.5.4 OEM.

Accessed March 18, 2008.

http://www.cheap-software-megastore.com/index.php?target=desc&progid=1701

Hoffer, J. A., Prescott, M. B., & McFadden, F. R. (2002): *Modern database management*. (Sixth ed.). Upper Saddle River: Prentice Hall.

Minaei-Bidgoli, B., Kortemeyer, G. and Punch, W.F. (n.d.): Enhancing online learning performance: an application of data mining methods. Accessed May 14, 2008.

(http://www.lon-cpa.org/papers/Belrouz-CATE2004.pdf)

Mitrovic, A. (2002): NORMIT: A Web-Enabled Tutor for Database Normalization, *icce*, p. 1276, 2002 International Conference on Computers in Education (ICCE'02).

Shank, P. (2004): ITFORUM PAPER #81- *New Social Interaction Tools for Online Learning.* Accessed June 06, 2008.

http://itech1.coe.uga.edu/itforum/paper81/paper81.html

The Association of Computing Machinery, Inc and
The Institute of Electrical and Electronics Engineers, Inc.
(2001): *Computing   Curricula 2001Computer Science-Final Report.* Accessed June 06, 2008.

http://delivery.acm.org/10.1145/390000/384275/cc2001.pdf?key1=384275&key2=4799172121&coll=Portal&dl=ACM&CFID=31012996&CFTOKEN=69738235

UCResearch Repository: NORMIT, a Web-enabled tutor for databasenormalisation. Accessed March 18, 2008.

http://ir.canterbury.ac.nz/handle/10092/335

Zhang, L., Kaschek, R. and Kinshuk. (2005): Developing a knowledge management support system for database normalisation. Accessed May14, 2008.

(http://ieeexplore.ieee.or/apl/freeabs_all.jsp?tp=&arnumber=1508695&isnumber=32...)