

## End-user Enable Database Design and Development Automation

Uduwela W.C. and Wijayarathna G.

Department of Mathematics and Computer Science, Faculty of Natural Sciences, Open

University of Sri Lanka

Faculty of Computing and Technology, University of Kelaniya, Sri Lanka

Email: wasana@ou.ac.lk, gamini@kln.ac.lk

### Abstract

Information System (IS) is a combination of software, hardware, and network components working together to collect, process, create, and distribute data to do the business operations. It consists with “update forms” to collect data, “reports” to distribute data, and “databases” to store data. IS plays a major role in many businesses, because it improves the business competitiveness. Although SMEs are interested to adopt IS, they are often suffered by other factors: time, underline cost, and availability of ICT experts. Hence, the ideal solution for them is to automate the process of IS design and development without ICT expertise for an affordable cost. The software tools are available on the Web to generate the “update forms” and “reports” automatically for a given database model. However, there is no approach to generate the databases of IS automatically.

Relational database model (RDBM) is the most commonly used database model in IS due to its advantages than the other data models. The reason of the advantages of the model is its design, but it is not a natural way of representing data. The model is a collection of data that is organized into multiple tables/relations linked to one another using key fields. These links represent the associations between relations. Typically, tables/relations represent entities in the domain. A table/relation has column/s and rows where column/s represent the attributes of the entity and rows represent the records (data). Each row in a table should have a key to identify that row uniquely. Designers should have to identify these elements from the given data requirements in the process of the RDBM design, which is difficult for non-technical people. The process of design of RDBM has few steps: *collect the set of data requirements, develop the conceptual model, develop the logical model, and convert it to the physical model*. Though there are approaches to automate some steps of the process of design and development of RDBM, they also request the technical support. Thus, it is required to develop a mechanism to automate the database design and development process by overcoming the difficulties in the automation approaches of RDBM, so that non-technical end-users will be able to develop their databases by themselves. Hence, a comprehensive literature survey was conducted to analyze the feasibilities and difficulties of the automation of the process of RDBM design and development.

Uduwela, W. et al., the author says that the “form” is the best way to collect data requirements of the database model for its automation, because form is in semi structured way than the natural

language (the most common way to present the data requirements is natural language) and it is very closer to the underline database.

Approaches were available to automate the development of the conceptual model based on the given data requirements. This is the most critical step in the RDBM design process, because it needs to identify the elements of the model (entities, attributes of them, relationship among the entities, keys and the cardinalities). Form based approaches were analyzed using the data available in the literature to recognize the places where the user intervention is needed. The analysis says that all approaches need user support and it is needed to make the corrections of the outcome, because the elements are not consistent among business domains; it differs from domain to domain and with the same domain also. Further, they demand user support to make the initial input according to the data requirements (set of forms) to identify the elements of the conceptual model.

The next step of the process is developing the logical model based on the conceptual model. The outcome of the logical model should be a normalized database to eliminate the data insertion, updating and deletion anomalies by reducing its data redundancies. Data redundancies often caused by Functional Dependencies (FD) that are a set of constraints between two sets of attributes in a relation. The database can be normalized by removing undesirable FDs (remove partial dependencies and transitive dependencies). We could not identify any approach that generates normalize database diagram automatically from the data requirements directly. Existing approaches request the FDs to generate the normalized RDBM. Designers' high perception power and skills are needed to identify the correct FDs, because it also depends on the domain which is a problem for the automation. FDs can be found by doing data mining, but it also generates an incorrect set of FDs if there are insufficient data combinations. Developing the physical model from the logical model is straightforward and relational database management systems help to automate it.

According to the analysis, it can be concluded that the existing approaches on conceptual model development cannot develop accurate models as they has to develop distinct models for each problem. Normalization approaches also cannot be automated as FDs also vary among business domains and with the same domain also. These concludes that there should a database model that can be designed and developed by end-users without any expert knowledge. The proposed model should not be either domain specific or problem specific. It would be better if the approach could convert the data requirements to the database model directly without intermediate steps like in the DBM design process. Further, it would be better the proposed model could be run on the existing database management systems too.

**Keywords:** *Database, Forms, conceptual model, Functional Dependencies, Normalization, ICT, SME*