A SURVEY ON TOOLS/SYSTEMS TO GENERATE DATABASE FROM FORM ANALYSIS

W.C. Uduwela[1] and G. Wijayarathna[2]

[1]*Department of Mathematics and Computer Science, The Open University of Sri Lanka*

[2]*Department of Industrial Management, University of Kelaniya, Sri Lanka*

INTRODUCTION

Globalization and rapid development of Information and Communication Technology (ICT) are major reasons to adapt ICT to businesses in the modern world, as it improves business competitiveness (Skoko *et al.*,2008). It has been shown that ICT helps to improve Small and Medium Enterprises' (SME) business performance too (Apulu and Latham,2009). Although SMEs are interested in adapting ICT, lack of ICT expertise and the system cost have been identified as major challenges (Harindranath *et al.*,2008). SMEs cannot develop their own ICT application either by outsourcing or purchasing off the shelf products, because it is costly. They might not have enough technical knowledge to develop their ICT application in house. Also, it is costly for them to maintain a separate ICT expertise group for this purpose (Harindranath *et al.*,2008).

Information systems (IS) play a major role in ICT applications and one of the most critical factors in any IS development is the database (DB) (Sanctum,2001). DBs need technical knowledge to develop, based on system requirements (Buchholz *et al.*,1995). Hence, it is a challenge to SMEs to develop a DB for their IS. Further, at the beginning of a software development project SMEs face difficult situations as they are not able to specify what they really need from an IS (MartÌnez and GarcÌa-Serrano, 2001). Therefore, if there is a tool/system that helps to develop DBs in a user friendly manner and cost effective way, then non-technical people will be benefited and it would be affordable for SMEs a they could develop their own IS by themselves.

Consequently a method that helps to automate the DB design process by forms may solve the problem to some extent, as forms in businesses are widely used to gather, maintain and report the data requirements of the businesses. Manual forms as well as digital forms help to gather data and keep them in a structured way. Forms are familiar, easy to read and can be understood by any end users to communicate many requirements of the system. Therefore we can consider the forms as a vital input source for DB design process (Choobineh *et al,*1988). It provides a common vocabulary and goals among end users and data processing professionals, rather than providing exhaustive requirements collection by end-users (Choobineh *et al.*,1992). Some research studies have been already carried out in order to automate the DB design process based on business forms. The purpose of this research is to study the appropriateness of form based analysis to DB design process for SMEs.

[1] Correspondences should be addressed to: W.C. Uduwela, Department of Mathematics & Computer Science, The Open University of Sri Lanka (email: wasana@ou.ac.lk)

METHODOLOGY

A comprehensive literature search was done to find existing tools/systems and approaches that generate Database Diagram (DD)/Entity Relationship Diagram (ERD) automatically centered on form analysis. Two approaches were considered for literature survey, such as DB scheme creation from legacy system DB and forms, and DB scheme creation from new forms. Analysis was carried out based on the objective of the research. It is important to evaluate the tools/systems and approaches based on the quality of the outcomes, since their quality affects both the efficiency and the effectiveness of IS (Moody, 2005). Moody found out a quality standard with a set of quality factors (completeness, flexibility, understandability, integration, correctness and implement ability) in order to evaluate the quality of data models (Arsovski *et al.*,2012 and Moody, 2005). Flexibility, understandability, integration, correctness and implementability were ignored in this analysis since they cannot be easily quantified (Moody,2003).

The response rate of authors for the evaluation of the existing tools/systems and approaches was low. Thus, completeness, which is the user requirement embedded into the data model as a set of attributes was selected as the quality factor for evaluation as there were many published research articles available to obtain an idea on completeness.

RESULT AND DISCUSSION

The outcome of the analysis of tools/systems and approaches is summarized in Table 1.

Table 1: Summary of the tools/systems and approaches

Tools and Target User	Approach Used	User Intervention	Completeness of the Tool
(Choobineh, *et al*,1992), (Choobineh, *et al*,1988) DB Designers and end users	Consist of two systems • Form Definition System: to create forms to collect end users requirements. • Expert DB Design System: to produce ERD based on form analysis. Rules are divided in to six phases (form selection, entity identification, attribute attachment, relationships identification, cardinality identification and consistency checking.	Yes. Except the rules in first phase of the Expert DB Design System, rules in other phases used in conjunction with a designer dialog. (Assume users are experienced DB designers).	For completeness, the approach should be combined into other sources in addition to forms such as natural language description.
(Lukovi,*et al*,2007), (Pavicevic, *et al*,2006), (Mogin, and Luković), (Pavicevic, *et al*,2005), (Mogin,*et al*,1994). DB Designer and end users	• conceptual modeling of a DB schema; Form type concept was used (tree structure over the instances of component types). Derived by generalization and introducing certain structuring rules into screen forms. Designer creates initial set of attributes and constraints of the form type. • automated design of relational DB sub schemas in the 3rd normal form • checking the consistency of constraints embedded into a DB schema and a set of sub schemas • automated integration of sub schemas into a relational DB schema	Yes. Using the screen form designers need to specify form types of various structures. Then the modeling process is raised to the level, which is closer to the users without an advanced knowledge of the DB design.	This tool can be even used for complex systems and it is capable of producing an integrated DB schema of a high quality in a reasonably short time. The tool is practically used to generate DB schemas.

(Veronica. *et al*,1989) DB designer	Two modes exist: • Expert mode: requirements are state directly • Novice mode: requirements are inferred from examples and purposeful dialogue. Rules and heuristics are used to make inferences from examples.	Yes. Need DB designer to review the requirements collected.	Results are in the expected way.
(Mfourga, 1997) DB Designer	From form-based interfaces of legacy systems. • Form Analysis. Two types of analysis used; Static Analysis; Identifies structural components and their relationships based on logical and physical aspects of form & Dynamic Analysis; Discovers constraints among components: cardinality constraints, functional dependencies, and existence dependencies. • Extracting ERD. entity derivation, relationship derivation, attribute attachment, cardinality determination, conceptual normalization, and schema integration	This approach has not been yet automated. The overall process needs user interaction, especially for form analysis.	They recommend that this approach to automate and can supplement existing DB reverse engineering techniques where forms constitute important uses of the DB
(Shu,*et al*,1983) DB Designers	To collect relevant information needed for DB design. Two groups of data identify based on forms, such as data and integrity constraints data and anticipated process which is use, modify or produce data. Form headings and hierarchical structure of forms used.	It is manual process. It can be understood by both specialist and non-specialist. Outcome (DB design specification) is formal, &machine manipulated.	50 cases were used to applied the theory in manually and the feedback was encouraged them.
(Wu,*et al*,2004) Decision Suppoer System developers	• Form Analysis: Given a set of business forms, decompose them into structure and joined data. • Heading structure design; to find relationships. • Meta-template design; heading structures serve as basis for designing meta-templates and DBs.	Not automated the process of DB design though all relevant data is there.	Feedback was encouraged them to recommend he approach to automate with further studies.

Most of the tools/systems and approaches supported DB designers to ease their task (Choobineh *et al.*,1992 and Lukovi *et al.*,2007). Both DB designers and users were assisted to collect correct requirements accurately through form analysis (FA) (Shu *et al.*,1983 and Pavicevic *et al.*,2005). Some tools/systems and approaches among the above aid to generate DB for novice users with minimum DB knowledge (Veronica *et al.*,1989). Completeness of outcomes were in the expected way excluding a tool which was proposed to combine Natural Language with FA to improve its completeness (Choobineh *et al.*,1992).

The majority used form features such as Form Type, Form Instance, Form Schema, Form Template, Form Title, etc. to develop the base of tools/systems and approaches. Form Type is a collection of form field. Form schema is associated constraints. Form Template is medium dependent representation. Form instance is a collection of value for form fields (Choobineh *et al.*,1988 and Choobineh *et al.*,1992). A majority applied rules/heuristics on analyzed form features to identify the components of ERD/DD.

CONCLUSIONS

It has been identified that all existing tools or systems need user intervention to generate ERD. The user must have a technical knowledge to generate ERD/DD and improve accuracy of outcomes. Thus, these tools/systems are not affordable for SMEs, to enhance IT adaption for their business.

According to the analysis based on conclusions given by the relevant authors of the tools/systems and approaches were achieved in the satisfactory level of the completeness of the generated ERD/DD as some of them use their tools for real application too. In addition, most of the tools are working properly by producing the correct ERD/DD, even for the complex DB requirements. Positive feedbacks of the outcomes emphasize that FA is a good approach for the DB design process. Forms use in the business help to gather system requirements easily and it is easy to read and understand by any user since it is a methodological way. Finally, it could be concluded that FA is one of the best approaches to develop a tool that is user friendly (used even by non-technical people) without any cost to generate ERD/DD.

REFERENCES

Anatomy of a Web Application: Security Considerations, Sanctum Inc, 2001

Apulu I., and Latham, A., (2009)., Journal TMC Academic, vol. 4, pp.64-80.

Arsovski, Z., Petrovic, D.R., Milanovic, I., Rankovic, V.,Kalinic, Z., (2012). Int. J. for Quality research; 2012, vol 6, no. 1, pp.47-53

Buchholz, E., Cyriaks, A. H., Düsterhöft, H. Mehlan and B. Thalheim, (1995)., Proc. 1st Int. Workshop on Applicat.of Natural Language to DBs

Choobineh, J., Mannino, M. V., Nunamaker J.F., J.R., and Konsynski B.R. (1988), IEEE Transactions on Software Engineering, vol 14, no. 2, pp.242-253.

Choobineh, J., Mannino, M. V., and Tseng, V. P. (1992). Communications of the ACM, Vol. 35, no. 2, pp.108-120

Harindranath, G., Dyerson, R., and Barnes, D., (2008), European Conf. Inform. Syst., pp.889-900.

Lukov,i I., Mogin, P., Pavic, J., and Ristic, S. (2007). Software practice and Experinece, vol. 37, pp.1621–1656

MartÌnezP., and GarcÌa-Serrano, A., (2001), Proc. the 5th Int. Conf. London, UK, pp.276-287.

Mfourga, N. (1997). WCRE '97 Proc. of the 4th Working Conf. on Reverse Engineering (WCRE '97) . Washington,DC,USA: IEEE Computer Society Washington, DC, USA ©1997, pp.184-193

Mogin, P., and Luković, I.A Prototyping CASE Tool, *http://homepages.mcs.vuw.ac.nz/~pmogin/ISATA_95E.pdf*

Mogin, P., Lukovic, I., and Karadzic, Z. (1994)., Proc. of the Int. Conf. on Technical Informatics, vol. 5, Timisoara, Romania, 16-19 November 1994. 'Politehnica' University of Timisoara: Timisoara, Romania, pp.49-58

Moody, D., (2005). J. Data & Knowledge Eng. - Special issue: Quality in conceptual modeling, Vol. 55, pp.243-276.

Moody, D. L,(2003). Proc. 11th European Conf. Inform. Syst, Naples, Italy, pp.16-21.

Pavicevic, J., Lukovic, I., Mogin, P., and Govedarica, M. (2006). ICSOFT 2, INSTICC Press, pp.157-160.

Pavicevic, J., Lukovic, I., Mogin, P., and Ristic, S. (2005). XIII Scientific Conf. on Industrial Systems IS'05, Herceg Novi, Montenegro, pp.321-330.

Shu, N.C., (1983), Int. conference on Management of Data, New York, USA, pp.161-172.

Skoko, H., Ceric, A., and Huang, C., (2008), Int. J. Bus. Research, Vol. 8, No. 4, pp.161-165.

Veronica. P.,Tseng and Michael. V., Mannino, (1989), J. of Management Information Systems, vol. 6, no. 2, pp.51-75.

Wu,J., Doonga, H., Leeb, C., Hsiac, T., and Liang, T., (2004, vol. 36, no. 3, January, pp.313–335