


AVOIDING SCHEMA CHANGE
&
A FRAMEWORK FOR NON-DISRUPTIVE SCHEMA
EVOLUTION

 University of Moratuwa, Sri Lanka.
MASTER OF BUSINESS ADMINISTRATION
www.lib.mrt.ac.lk
IN
INFORMATION TECHNOLOGY

S. S. W. W. Dissanayake
Department of Computer Science & Engineering
University of Moratuwa
December 2004

AVOIDING SCHEMA CHANGE
&
A FRAMEWORK FOR NON-DISRUPTIVE SCHEMA
EVOLUTION

By

S. S. W. W. Dissanayake



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

The Dissertation was submitted to the Department Computer Science & Engineering of the University of Moratuwa in partial fulfilment of the requirement for the Degree of Master of Business Administration.

Department of Computer Science & Engineering

University of Moratuwa

December 2004

Declaration

"I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any University to the best of my knowledge and believe it does not contain any material previously published, written or orally communicated by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, to be made available for photocopying and for interlibrary loans, and for the title and summary to be made available to outside organizations"

Signature of Candidate:.....

Date:.....



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

To the best of my knowledge, the above particulars are correct.

Supervisor:

Dr. Asoka Perera
BSc Eng (Moratuwa), MSc (Lough), PhD (Lough), CEng, MIE(SL)

Abstract

It is a well-known fact that a significant share of the cost of information systems is spent on maintenance. Some researchers estimate this share to be around 80% of the lifetime cost of information systems. A significant share of adaptive maintenance work of a business information system starts from a change in the data models it is based on. Since, almost all business information systems are based on relational databases, it is clear that relational database schema changes account for a significant share of the total maintenance cost. This is aggravated by the fact that a change in database schema has the potential of driving existing applications out of date. Modifications to such applications cost a considerable amount of time and money to business organisations.

Possible solutions to minimize the costs associated with this problem are essentially technical in nature. This dissertation explores a couple of possible solutions in relational database technology, which could help organisations to keep their system maintenance cost under check. The main solutions explored are in the area of data model patterns and schema evolution. A new branch of schema evolution called "Non Disruptive Schema Evolution" is explored as well.

Acknowledgements

I wish to thank,

- Dr. Asoka Perera, University of Moratuwa for being my supervisor and for the kind advice, guidance and support given on conducting this research project
- Dr. Ananda Jayawardena, Dr. Sanath Jayasena and Mr. Kithsiri Samarasinghe for providing guidance on conducting research projects
- The Computer Science & Engineering Department and the Management of Technology Department of the University of Moratuwa for providing the MBA in IT program
- Barry Williams, Founder and Principal Consultant of Database Answers for his support on analysing data models
- Simon Williams, Inventor of the Associative Model of Data, for the giving his time to discuss AMD related issues
- My colleagues at SriLogic who supported with their invaluable ideas

Table of Contents

DECLARATION.....	I
ABSTRACT.....	II
ACKNOWLEDGEMENTS	III
TABLE OF CONTENTS	IV
LIST OF FIGURES	VII
LIST OF TABLES	VIII
LIST OF ABBREVIATIONS	IX
EXECUTIVE SUMMARY	1
1 INTRODUCTION.....	3
1.1 BACKGROUND.....	3
1.2 OBJECTIVES	4
1.2.1 <i>Main Objective</i>	4
1.2.1.1 Scope of the main objective.....	5
1.2.2 <i>Sub Objectives</i>	5
1.2.2.1 Data model patterns to minimize schema evolutions	5
1.2.2.2 Framework for Non Disruptive Schema Evolution	7
1.3 METHOD	8
1.4 MAIN FINDINGS	10
1.5 GUIDE	11
2 LITERATURE REVIEW	13
2.1 INTRODUCTION	13
2.2 SCHEMA MODIFICATION	13
2.3 SCHEMA EVOLUTION	14
2.3.1 <i>Ontologies and Schema Evolution</i>	14

2.3.2	<i>Predictive Schema Evolution</i>	15
2.4	SCHEMA VERSIONING	15
2.4.1	<i>Partial Schema Versioning</i>	17
2.4.2	<i>Full Schema Versioning</i>	17
2.5	MANAGEMENT OF APPLICATIONS	18
2.6	DATA MODEL PATTERNS	20
2.7	CONCLUSIONS.....	23
3	DATA MODEL PATTERNS – CONCEPT EVALUATION	25
3.1	BASICS OF DATA MODEL PATTERNS.....	25
3.2	DATA MODEL PATTERNS TO ADD NEW ATTRIBUTES	27
3.3	WHY DATA MODEL PATTERNS TO ADD ATTRIBUTES?	28
4	AVOIDING SCHEMA CHANGE	30
4.1	NEW PROPERTIES DATA MODEL PATTERN	30
4.1.1	<i>The Data Model</i>	30
4.1.2	<i>A populated Data Model</i>	32
4.1.3	<i>Sample SQL for CRUD functionality</i>	33
4.1.3.1	Create	34
4.1.3.2	Retrieve	34
4.1.3.3	Update.....	35
4.1.3.4	Delete	35
4.1.4	<i>Limitations</i>	35
4.2	FLEXIBLE ITEM.....	36
4.2.1	<i>Context and Motivation</i>	36
4.2.2	<i>Introduction</i>	37
4.2.3	<i>Storing Meta data</i>	38
4.2.4	<i>Viewing Meta data</i>	41
4.2.5	<i>Storing Address data</i>	43
4.2.6	<i>Viewing Address Data</i>	46
4.2.7	<i>Storing Student Data</i>	48
4.2.8	<i>Viewing Student Data</i>	52

4.2.9	<i>Supporting the new Attribute</i>	53
4.2.10	<i>Limitations</i>	54
4.2.11	<i>Other uses of “Flexible Item”</i>	54
5	NON DISRUPTIVE SCHEMA EVOLUTION	56
5.1	FRAMEWORK FOR RESEARCH IN NDSE	57
5.2	RENAMING AN ATTRIBUTE AS AN NDSE	60
5.2.1	<i>Use of Views</i>	60
5.2.2	<i>Use of a Property File</i>	60
6	CONCLUSIONS & FUTURE WORK	61
6.1	CONCLUSIONS	61
6.1.1	<i>Data Model Patterns can be used to avoid schema changes</i>	61
6.1.2	<i>Non Disruptive Schema Evolution is a promising prospect</i>	62
6.2	FUTURE RESEARCH	62
6.2.1	<i>Documentation of New Data Model Patterns</i>	62
6.2.2	<i>Research on new applications of “Flexible Item” data model pattern</i>	63
6.2.3	<i>Research on Non Disruptive Schema Evolution</i>	63
	REFERENCES	64
	APPENDIX-A	67
	INTERVIEWS AND DISCUSSIONS	67
	INTERVIEW & DISCUSSION FRAMEWORK	67
	GENERAL CONCLUSIONS	68

List of Figures

FIGURE 4.1 CLASS VIEW OF NEW PROPERTIES DATA MODEL PATTERN.....	30
FIGURE 4.2 DATA MODEL FOR NEW PROPERTIES	31
FIGURE 4.3 DATA MODEL TO STORE META DATA.....	38
FIGURE 4.4 DATA MODEL TO STORE DATA.....	43
FIGURE 5.1 FRAMEWORK FOR NDSE RESEARCH.....	59



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

List of Tables

TABLE 4-1 STUDENT TABLE.....	32
TABLE 4-2 ATTRIBUTE TABLE.....	33
TABLE 4-3 STUDENT__ATTRIBUTE TABLE.....	33
TABLE 4-4 RELATION TABLE.....	39
TABLE 4-5 ATTRIBUTE TABLE.....	40
TABLE 4-6 RELATION_ATTRIBUTE TABLE.....	41
TABLE 4-7 VIEWING META DATA	42
TABLE 4-8 DATA TABLE.....	44
TABLE 4-9 TUPLE TABLE.....	45
TABLE 4-10 DATA_INTEGER TABLE.....	45
TABLE 4-11 DATA_VARCHAR TABLE	46
TABLE 4-12 VIEWING ALL ADDRESSES.....	47
TABLE 4-13 DATA TABLE.....	48
TABLE 4-14 TUPLE	50
TABLE 4-15 DATA_VARCHAR TABLE	50
TABLE 4-16 DATA_DATA TABLE.....	51
TABLE 4-17 DATA_INTEGER TABLE.....	51
TABLE 4-18 VIEWING STUDENT RECORDS.....	53

List of Abbreviations

CRUD Create, Retrieve, Update, Delete

DBA Data Base Administrator

DBMS Data Base Management System

DDL Data Definition Language

ER Entity Relationship

ERP Enterprise Resource Planning

GUI Graphical User Interface

IT Information Technology

NDSE Non Disruptive Schema Evolution

OR Object-Relational

RDBMS Relational DataBase Management System

SDLC System Development Life Cycle

SQL Structured Query Language

XML eXtensible Markup Language



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Executive Summary

Business organisations spend a considerable amount of time and money upgrading applications when the schema of a central relational database changes. Currently there is no widely accepted, commercially viable solution for this problem other than living with it. The objective of this research project was to investigate alternative, low cost strategies available to business organisations to deal with the problems caused by schema changes in relational databases.

The possibility of using data model patterns (can be thought of as the equivalent of design patterns in Object Oriented software design), to avoid or minimize schema changes was investigated. It was successfully demonstrated that an already published data model pattern ("New Properties") could facilitate adding a new property to an entity without requiring a change in the database schema. Further the model pattern "Flexible Item" was identified and documented to achieve the same objective while addressing certain limitations of the former.

The practical implication of this is that, if in the business analysis of a new system, if a particular entity is identified as being very critical to the system and having a possibility of needing more attributes in the future, one of the two data model patterns discussed in this dissertation could be selected for use.

Further, a conceptual framework for future research on and area that we call "Non Disruptive Schema Evolution"

was identified and provided. This can be thought of as an extension of traditional schema evolution, with the added feature that, a particular change to the schema will not render any application invalid.

In conclusion, it is evident that there are simple, low cost and formal techniques of addressing the problem of schema changes in relational database, than what the contemporary thinking suggest.



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk