

Semantic Extensions and a Novel Approach to Conceptual Modelling

by

Somluck La-Ongsri, *B.Sc. (Applied Mathematics), M.Sc. (Information Technology)*

School of Computer Science, Engineering and Mathematics,
Faculty of Science and Engineering

February 25, 2009

A thesis presented to the
Flinders University
in total fulfilment of the requirements for the degree of
Doctor of Philosophy

Adelaide, South Australia, 2009
© (Somluck La-Ongsri, 2009)

Contents

| | |
|---|-------------|
| Contents | ii |
| List of Figures | viii |
| List of Tables | xii |
| Abstract | xiii |
| Certification | xiv |
| Acknowledgements | xv |
| 1 Introduction | 1 |
| 1.1 Background | 2 |
| 1.1.1 The Design of Conceptual Models | 2 |
| 1.1.2 A Historical Background and Perspectives of Conceptual Modelling | 5 |
| 1.2 Motivation for Research | 9 |
| 1.3 Objectives | 12 |
| 1.4 Scope | 13 |
| 1.5 Approach and Structure | 13 |
| 1.5.1 Approach of the Thesis | 13 |
| 1.5.2 Structure of the Thesis | 17 |
| 2 Revisiting the Fundamentals of Conceptual Modelling Concepts | 19 |
| 2.1 Introduction | 19 |
| 2.2 Data Models | 21 |
| 2.3 Conceptual Schema | 24 |

| | |
|---|-----------|
| <i>CONTENTS</i> | iii |
| 2.4 Semantics in Conceptual Modelling | 25 |
| 2.5 Conceptual Design as Part of Database Design | 28 |
| 2.6 Key Considerations of a Conceptual Data Model | 31 |
| 2.7 Conceptual Modelling Approaches | 33 |
| 2.7.1 The Entity-Relationship (ER) Model | 34 |
| 2.7.2 Object Role Modelling (ORM) | 38 |
| 2.7.3 Unified Modelling Language (UML) Class Diagrams | 43 |
| 2.7.4 A Comparison of ER, ORM and the UML Class Diagram | 48 |
| 2.8 Evolution of Data Models | 51 |
| 2.9 Summary | 58 |
| 3 ER Modelling Extensions: A Survey and Comparative Review | 59 |
| 3.1 Introduction | 60 |
| 3.2 Limitations of the ER Model | 63 |
| 3.3 Exploring ER Modelling Extensions | 66 |
| 3.3.1 Structural Aspect | 66 |
| 3.3.2 Data Abstraction Aspect | 70 |
| 3.3.3 Temporal Aspect | 72 |
| 3.3.4 Spatio-Temporal Aspect | 75 |
| 3.3.5 Data Warehousing Aspect | 79 |
| 3.3.6 Domain-Specific Application Aspect | 81 |
| 3.3.7 Knowledge Base Aspect | 83 |
| 3.3.8 Fuzzy Data Aspect | 85 |
| 3.3.9 XML Data Aspect | 88 |
| 3.4 A Comparative Study of ER Modelling Extensions | 90 |
| 3.5 Summary | 94 |
| 4 Mesodata in Conceptual Modelling | 96 |
| 4.1 Introduction | 97 |
| 4.2 Mesodata Approach | 101 |
| 4.2.1 What is Mesodata? | 102 |
| 4.2.2 Key Components of Mesodata | 103 |
| 4.2.3 Example Use of Mesodata | 105 |

| | | |
|----------|--|------------|
| 4.3 | The Mesodata Entity-Relationship (MDER) Model | 106 |
| 4.3.1 | MDER Data Structures | 107 |
| 4.3.2 | MDER Integrity Constraints | 111 |
| 4.3.3 | MDER Languages | 112 |
| 4.4 | The Mesodata Object Role Modelling (MDORM) Model | 112 |
| 4.4.1 | MDORM Data Structures | 114 |
| 4.4.2 | MDORM Integrity Constraints | 114 |
| 4.4.3 | MDORM Languages | 115 |
| 4.5 | Example MDER and MDORM Schemata | 116 |
| 4.6 | Summary | 116 |
| 5 | Ontology in Conceptual Modelling | 119 |
| 5.1 | Introduction | 120 |
| 5.2 | Review of Ontologies | 123 |
| 5.2.1 | What is Ontology? | 124 |
| 5.2.2 | Types of Ontology | 127 |
| 5.2.3 | Terminological Clarifications of Ontologies | 130 |
| 5.2.4 | Ontologies versus Conceptual Data Models | 134 |
| 5.3 | Ontology-Based Conceptual Data Modelling | 136 |
| 5.3.1 | Using Ontology for High-Level Conceptual Data Models | 136 |
| 5.3.2 | Ontology's Class Hierarchy Modelling | 137 |
| 5.4 | The Ontological Entity-Relationship (OntoER) Model | 141 |
| 5.4.1 | OntoER Data Structures | 142 |
| 5.4.2 | OntoER Constraints | 144 |
| 5.4.3 | OntoER Languages | 145 |
| 5.5 | The Ontological Object Role Modelling (OntoORM) Model | 145 |
| 5.5.1 | OntoORM Data Structures | 146 |
| 5.5.2 | OntoORM Integrity Constraints | 147 |
| 5.5.3 | OntoORM Languages | 148 |
| 5.6 | The Ontological Unified Modelling Language (OntoUML) Class Diagram Model | 148 |
| 5.6.1 | OntoUML Data Structures | 150 |
| 5.6.2 | OntoUML Integrity Constraints | 150 |

| | | |
|----------|---|------------|
| 5.6.3 | OntoUML Languages | 152 |
| 5.7 | Example of OntoER, OntoORM and OntoUML Class Diagram Schemata | 152 |
| 5.8 | Summary | 157 |
| 6 | Polymorphic Relationships in Entity-Relationship Modelling | 158 |
| 6.1 | Introduction | 159 |
| 6.2 | Review of Relationship Types in Conceptual Modelling | 161 |
| 6.2.1 | Types of Relationship | 161 |
| 6.2.2 | Constraints on Relationship Types | 163 |
| 6.3 | Motivation | 166 |
| 6.4 | Links as Overloaded Polymorphic Relationships | 169 |
| 6.4.1 | Link Type Structure: Formal Definition | 170 |
| 6.4.2 | Link Type Representation | 171 |
| 6.5 | Cardinality, Default Value and Precedence for Link Types | 173 |
| 6.5.1 | Cardinality Constraints on Link Types | 173 |
| 6.5.2 | Default Values for Link Types | 174 |
| 6.5.3 | Precedence | 176 |
| 6.6 | Summary | 177 |
| 7 | Data Modelling in Rapidly Changing Complex Environments | 179 |
| 7.1 | Introduction | 180 |
| 7.2 | Significant Issues in Driving Rapid Conceptual Modelling Techniques | 182 |
| 7.2.1 | Systems Issues | 182 |
| 7.2.2 | Data Issues | 184 |
| 7.3 | Late Binding the Conceptual Modelling | 185 |
| 7.3.1 | Deferred Schema Deployment | 185 |
| 7.3.2 | Common Conceptual Schemata | 186 |
| 7.4 | LtER Modelling | 188 |
| 7.4.1 | The LtER Schema | 188 |
| 7.4.2 | The LtER Architecture | 190 |
| 7.4.3 | The LtER Characteristics | 193 |
| 7.5 | Summary | 194 |

| | |
|---|------------|
| 8 From Conceptual Design to Logical Design for the Relational Data Model | 196 |
| 8.1 Introduction | 197 |
| 8.2 The Relational Data Model | 198 |
| 8.3 Transformation from the MDER Schema to a Relational Schema . | 201 |
| 8.4 Transformation from the MDORM Schema to a Relational Schema | 204 |
| 8.5 Transformation from the OntoER Schema to a Relational Schema | 206 |
| 8.6 Transformation from the OntoORM schema to a Relational Schema | 210 |
| 8.7 Transformation from the OntoUML Schema to a Relational Schema | 213 |
| 8.8 Summary | 216 |
| 9 Conclusions and Future Research | 219 |
| 9.1 Research Contributions | 219 |
| 9.1.1 A Survey on ER Modelling Extensions | 219 |
| 9.1.2 Accommodating Mesodata into Conceptual Modelling Methodologies | 220 |
| 9.1.3 Incorporating Ontology-based Semantics in Conceptual Modelling | 220 |
| 9.1.4 Polymorphic Relationships in ER Modelling | 221 |
| 9.1.5 Rapid Conceptual Modelling | 222 |
| 9.1.6 Data Model Mapping | 222 |
| 9.2 Future Research Directions | 223 |
| 9.2.1 Extending This Work | 223 |
| 9.2.2 New Areas to Explore: Advances in Conceptual Modelling | 225 |
| 9.3 Conclusions | 226 |
| Appendices | 228 |
| A Publications Resulting from This Thesis | 228 |
| B Sample Proposals from the CERME Framework | 229 |
| B.1 The Enhanced Entity-Relationship (EER) Model | 229 |
| B.2 The Time Extended EER (TIMEER) Model | 232 |
| B.3 The Spatio-Temporal ER (STER) Model | 236 |

| | | |
|----------|--|------------|
| B.4 | The Modelling of Application Data with Spatio-temporal features (MADS) | 240 |
| B.5 | The DIstributed design of SpaTIO-temporaL data (DISTIL) | 243 |
| B.6 | The starER Model | 244 |
| B.7 | The MultiDimER Model | 245 |
| B.8 | The FuzzyEER Model | 248 |
| B.9 | The ER extended for XML (EReX) Model | 253 |
| B.10 | The XSEM-ER Model | 255 |
| C | The Full Schema Mapping for MDER and OntoER models | 258 |
| C.1 | The Result of the Full Mapping of the INVENTORY MDER Schema | 258 |
| C.2 | The Result of the Full Mapping of the MEDICAL OntoER Schema | 258 |
| C.3 | Correcting the Schema: Normalisation | 259 |
| D | The Full SQL Data Definition Language (MDDL and DDL) | 263 |
| D.1 | SQL Data Definition Statements for Defining the INVENTORY Schema | 265 |
| D.2 | SQL Data Definition Statements for Defining the MEDICAL Schema | 266 |
| | Bibliography | 269 |

List of Figures

| | | |
|------|--|-----|
| 1.1 | Conceptual modelling issues during four eras. | 6 |
| 1.2 | Representation of common domain semantics | 10 |
| 1.3 | Overview of the thesis structure | 16 |
| 2.1 | Conceptual schema as the heart of an integrated system | 25 |
| 2.2 | Three main phases of database design. | 29 |
| 2.3 | An example of an ER diagram for a university database. | 35 |
| 2.4 | Cardinality ratio and participation constraint in the <code>WORKS_IN</code> and <code>MANAGES</code> relationship types. | 36 |
| 2.5 | Cardinality ratios in the ER model. | 36 |
| 2.6 | Two design options of Country | 38 |
| 2.7 | An example of an ORM diagram. | 39 |
| 2.8 | An example of a binary association with the expression of multiplicity constraints in a UML class diagram. | 44 |
| 2.9 | An example of a UML class diagram. | 46 |
| 2.10 | UML aggregation. | 47 |
| 4.1 | Examples of reference files | 99 |
| 4.2 | Mesodata layer between metadata and data | 102 |
| 4.3 | Symbols of major components in the <code>MDER</code> model. | 107 |
| 4.4 | A <code>COUNTRIES</code> mesodata entity type. | 108 |
| 4.5 | An example of a Suburb connectivity. | 109 |
| 4.6 | An example of a Suburb attribute referencing a weighted graph. | 110 |
| 4.7 | An <code>MDER</code> Example. | 111 |
| 4.8 | ORM Symbols Extension. | 113 |
| 4.9 | The <code>WGRAPH</code> value type and the mesodata mandatory role constraint. | 114 |
| 4.10 | <code>MDER</code> diagram for an example <code>INVENTORY</code> database. | 115 |

| | |
|---|-----|
| 4.11 MDORM schema for an example INVENTORY database. | 117 |
| 5.1 A portion of a medical database | 122 |
| 5.2 Different types of ontology. | 127 |
| 5.3 Ontological categories | 128 |
| 5.4 Different types of ontological specifications. | 131 |
| 5.5 Ontology spectrum | 132 |
| 5.6 Ontologies and the underlying conceptual schemata diagram. . . . | 137 |
| 5.7 An example of a representation of terms and relationships. | 139 |
| 5.8 Symbols of major components in the OntoER model. | 142 |
| 5.9 A DISEASES ontological entity type. | 143 |
| 5.10 An OntoER Example. | 144 |
| 5.11 An extension of the ORM Symbols | 146 |
| 5.12 An ontological label type | 147 |
| 5.13 A simple OntoORM diagram. | 147 |
| 5.14 A representation of ontological class types. | 150 |
| 5.15 A simple OntoUML class diagram. | 151 |
| 5.16 OntoER schema for an example MEDICAL database. | 153 |
| 5.17 A portion of the medical school location ontology. | 154 |
| 5.18 OntoORM schema for an example MEDICAL database. | 155 |
| 5.19 OntoUML class diagram for an example MEDICAL database. . . . | 156 |
| 6.1 Cardinality and participation constraints. | 164 |
| 6.2 Another representation of cardinality and participation constraints. | 165 |
| 6.3 Constraints of the ternary SUPPLY relationship. | 166 |
| 6.4 Two design diagrams. | 167 |
| 6.5 Duplication of Form and Dose attributes. | 168 |
| 6.6 Three options for the modelling of meetings. | 169 |
| 6.7 Examples of links representation. | 172 |
| 6.8 A link type. | 173 |
| 7.1 An example schema for a parts example. | 185 |
| 7.2 LtER schema. | 188 |
| 7.3 LtER architecture. | 191 |

| | | |
|------|--|-----|
| 7.4 | Two design choices. | 193 |
| 8.1 | An example of a relation <code>PROPERTY_FOR_RENT</code> | 199 |
| 8.2 | A referential integrity constraint. | 200 |
| 8.3 | MDER schema for an example <code>INVENTORY</code> database. | 202 |
| 8.4 | Mapping the mesodata entity types. | 203 |
| 8.5 | MDORM schema for an example <code>INVENTORY</code> database. | 205 |
| 8.6 | OntoER schema for an example <code>MEDICAL</code> database. | 207 |
| 8.7 | Corresponding source relational schemata for the ontological entity types. | 208 |
| 8.8 | OntoORM schema for an example <code>MEDICAL</code> database. | 211 |
| 8.9 | Mapping an OMR constraint in <code>OntoORM</code> | 213 |
| 8.10 | OntoUML schema for an example <code>MEDICAL</code> database. | 214 |
| B.1 | Subclasses and specialisation in an EER diagram. | 230 |
| B.2 | The temporal entity type <code>employee</code> | 233 |
| B.3 | The temporal relationship type <code>WORKS_FOR</code> | 233 |
| B.4 | The temporal attribute <code>Salary</code> | 234 |
| B.5 | Representation of snapshot participation constraints in <code>TIMEER</code> | 234 |
| B.6 | Representation of lifespan participation constraints in <code>TIMEER</code> | 235 |
| B.7 | The temporal relationship type <code>WORKS_FOR</code> with participation constraints | 235 |
| B.8 | Entity type capturing existence time. | 237 |
| B.9 | Spatial entity types. | 237 |
| B.10 | A Spatio-temporal entity type with valid time support. | 238 |
| B.11 | An attribute with valid time support. | 238 |
| B.12 | A spatial attribute. | 238 |
| B.13 | A spatio-temporal attribute in <code>STER</code> | 239 |
| B.14 | A spatio-temporal relationship type. | 239 |
| B.15 | An example of a multi-association relationship type. | 241 |
| B.16 | An example of a starER diagram. | 244 |
| B.17 | Examples of specialisation, aggregation and membership relationship types. | 245 |
| B.18 | Notations for a multidimensional model. | 246 |

| | | |
|------|--|-----|
| B.19 | A conceptual multidimensional schema. | 247 |
| B.20 | Entity employee with fuzzy attributes. | 249 |
| B.21 | Possibility distributions for the fuzzy attribute. | 250 |
| B.22 | Fuzzy entity with a membership degree. | 250 |
| B.23 | An example of a fuzzy relationship. | 251 |
| B.24 | An example of a fuzzy participation constraint. | 251 |
| B.25 | An example of a fuzzy cardinality constraint. | 252 |
| B.26 | An example of fuzzy (min,max) notation. | 252 |
| B.27 | An example of an EReX schema. | 253 |
| B.28 | XSEM-ER data node types and cluster types. | 256 |
| C.1 | Result of mapping the INVENTORY MDER schema. | 259 |
| C.2 | Result of mapping the MEDICAL OntoER schema. | 260 |
| C.3 | Functional dependencies on schemata. | 261 |
| C.4 | Normalisation into 3NF. | 261 |
| C.5 | A normalised schema of the MEDICAL OntoER schema. | 262 |

List of Tables

| | | |
|-----|---|-----|
| 2.1 | Examples of Multiplicities. | 45 |
| 2.2 | Equivalent data structures in ER, ORM and the UML class diagram. | 49 |
| 2.3 | Equivalent integrity constraints in ER, ORM and the UML class diagram. | 50 |
| 2.4 | Equivalent languages in ER, ORM and the UML class diagram. | 50 |
| 2.5 | The data models in ten historical eras. | 51 |
| 3.1 | Overview presentation of the CERME survey framework. | 62 |
| 3.2 | A comparison of the main CERME proposals. | 92 |
| 3.3 | Resources available for each CERME aspect. | 93 |
| 4.1 | Examples of mesodata types and their operators. | 103 |
| 4.2 | Examples of mesodata types and the structure of source relations. | 104 |
| 4.3 | A relation zipcoderel. | 105 |
| 5.1 | Corresponding use of terms between semantic domain relationships and common domain structures | 140 |
| B.1 | Assigning temporal aspects to ER constructs. | 236 |

Abstract

Conceptual modelling presented as a framework for database design is a discipline of great importance in many areas in computer science that seeks to represent real-world phenomenon using semantic primitives. To date, traditional (static) conceptual models have been successfully used and extended to deal with the semantics of relatively stable real world applications. However, the capturing of semantics is a seemingly endless task as it involves various dimensions and categories.

It is argued in this thesis that the incorporation of complex domain structures in conceptual modelling to represent the semantic domains of an attribute and the relationships within a concept in ontologies would provide more expressive and richer semantics. Additionally, it is argued that basic relationships in the entity-relationship model may need to be modified or extended to handle a broad spectrum of situations that arise from differing perspectives of the real world. Furthermore, it is argued that a conceptual model should allow rapid and simultaneous storage of data and data modelling as unexpected and sudden events require data to be modelled rapidly. This thesis begins with an extensive review of the field of conceptual modelling and an exploration of the concepts of meso-data, ontologies and semantic relationships in conceptual modelling as well as various aspects of extensions to the entity-relationship model.

Using these foundations, a classification of ER modelling extensions (**CERME**) framework is introduced that forms the basis of common aspects and comparative criteria which can be used to categorise and compare various proposals. In addition, the Mesodata Entity-Relationship (**MDER**) model, Mesodata Object Role Model (**MDORM**), Ontological Entity-Relationship (**OntoER**) model, Ontological Object Role Modelling (**OntoORM**) and Ontological Unified Modeling Language (**OntoUML**) class diagrams are presented that allow advanced semantics to be associated with the domains of an attribute. It is also demonstrated that these proposed models can be mapped into the commonly accepted standard relational model. Furthermore, for some of the modelling issues that are not easily accommodated into the ER model, this thesis introduces a new relationship construct, **polymorphic relationships**, to handle this situation. To this end, a novel approach to conceptual modelling, the **LitER** model, is presented that incorporates the previously proposed concepts of mesodata, ontologies and polymorphic relationships into the model which allows data to be modelled rapidly.

Certification

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed

Dated

Somluck La-Ongsri

Acknowledgements

My PhD journey and this thesis would not have been accomplished without the support of other colleagues and friends. I would like to express my gratitude to those who have helped me along my journey:

— my supervisor, Professor John F. Roddick, for his guidance on the ideas and concepts that provided me with inspiration for my thesis and for his support and encouragement throughout my candidature;

— my teacher, Associate Professor Suphamit Chittayasothorn, for his teaching, guidance and support that provided me with inspiration for my academic career and PhD research and for his constructive feedback on my work;

— my sponsor, Sukhothai Thammathirat Open University, Thailand, for providing the scholarship that gave me the opportunity to pursue a PhD abroad;

— Flinders fellows: Carl Mooney for his assistance with \LaTeX and proofreading, Denise de Vries and Paul Gardner-Stephen for their helpful suggestions and proofreading, Murk Bottema for his assistance, helpful suggestions and reading parts of the drafts, Paul Calder who always made himself available to students and for reading parts of the drafts, Rino Calaycay who is friendly and always ready to help, Neville Williams for his warmth and friendliness, Martin Luerssen for providing his \LaTeX sources and lastly, Edi Winarko, May Zhao, Kenny Ma, Huan-Min Shen, Richard Leibbrandt, Anna Shillabeer and Ping Liang for sharing their experiences;

— Flinders loop bus driver, David Petch, for his caring and generosity; Security officers for assisting me to get home safely after long nights of study; The Document Services Unit for delivering research materials; The Staff Development and Training Unit for their useful training courses and programs;

— my best friend, Steppy, for his continual support in reviewing my manuscript and in bringing me happiness in my last tough year; my close friends (PTong, PPai, Synraha's family and relatives, Dao and PJuk) for their wonderful friendship and support; and finally

— my family for providing the support that only a family can and for their understanding, encouragement and confidence in me; my heartfelt thanks goes to Mum and Dad who will be proud to see me finally achieve my PhD and my supportive sister and brother, PPen and PPat.