

Variational Database Management System: Theory and Application

Parisa S. Ataei

Thesis Proposal

Submitted November 3rd, 2020

Abstract

Many problems require working with data that varies in its structure and content. Likewise, many tools and techniques have been developed for dealing with variation in databases with respect to time (e.g., work on database evolution) or space (e.g., work on data integration). However, these specialized approaches neither cover all data variation needs nor provide a solution to deal with database variation both in time and space simultaneously. In this research, we propose a generic framework that considers *variation* orthogonal to relational databases. We extend the relational database theory to incorporate variation explicitly in databases and the query language: we define *variational schemas* for describing variation in the structure of a database *variational queries* for expressing variation in information needs, and *variational databases* for capturing variation in content. Although the model underlying variational database is simple, encoding variation explicitly in databases introduces complexity akin to using preprocessing directives in software. The question then arises: is this a good idea? We address this question by systematically developing two case studies that illustrate how different kinds of variation needs can be encoded and integrated in a variational database and how the corresponding information needs can be expressed as variational queries. We also design and implement a variational database management system as an abstraction layer over a traditional relational database. We demonstrate the applicability and feasibility of our approach on our two case-studies.

1 Introduction

With this initial ..., we propose the following research:

- Objective 1: Identify the kinds of variation existing in relational databases in different application domains.
- Objective 2: Design a query language and implement a database management system that accommodate at least a subset of types of variations identified in objective 1.
- Objective 3: Mechanize proofs of properties of the language and the system.
- Objective 4: Demonstrate how the proposed system can be used to manage variation in databases in different application domains.
- Objective 5: Generalize the encoding of variation and the design of the query language to cover more kinds of variation.

2 Statement of Thesis

The goal of this research is to provide a query language(s) and a database management system that explicitly account for different kinds of variation in relational databases to relieve some of programmer/DBA's labor by providing some guarantees within the query language that accounts explicitly for variation.

3 Background and Related Work

3.1 Variational Research

3.2 Forms of Variation in Relational Databases

Database Evolution:

Data Integration:

Versioning a Database:

4 Research Goals and Methods

4.1 Identify the kinds of variation existing in relational databases in different application domains

4.2 Design a query language and implement a database management system that accommodate at least a subset of identified types of variations

4.3 Mechanize proofs of properties of the language and the system

4.4 Demonstrate how the proposed system can be used to manage variation in databases in different application domains

4.5 Generalize the encoding of variation and the design of the query language to cover more kinds of variation

4.6 Summary

Figure ?? summarizes the connections between the research questions and activities. Table ?? provides the timeline for this proposal.

5 Conclusion