

AN ABSTRACT OF THE DISSERTATION OF

Parisa Ataei for the degree of Doctor of Philosophy in Computer Science presented on June ?, 2021.

Title: Theory and Implementation of a Variational Database Management System

Abstract approved: _____

Eric Walkingshaw

Databases used in the same context that are intended to satisfy the same information need share commonalities while varying in some aspects as well. For example, databases used to store information for a software system are intended to satisfy the client’s information need and they share some parts of the schema, however, they still vary in some parts of the schema and most of the content. That is, they are *variants* of a conceptual hypothetical database that captures all the *variation* among them. This pattern appears repeatedly in databases. Instances of variation occurring in databases are: schema evolution, database integration, database versioning, data extraction, and software development either using software product lines approaches or not. While there are specialized approaches to some instances of variation in databases there are no generic solutions to manage variation in databases.

In this thesis, we answer the question: “can we abstract out this repeating pattern of variation appearing in databases and bring the hypothetical database that captures all the variation among a number of databases to life?” To this end, we formalize the database that captures all variation of a number of databases, called a *variational database*, and a query language that allows interaction with the said database, called *variational relational algebra*. [maybe two languages, another: *variational SQL*, depends on my timing!] We implement these concepts in *Variational Database Management System (VDBMS)*, demonstrate the feasibility of our concepts by developing two real-world use cases, and examine the performance of VDBMS over our two use cases.

©Copyright by Parisa Ataei
June ?, 2021
All Rights Reserved

Theory and Implementation of a Variational Database Management System

by

Parisa Ataei

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June ?, 2021
Commencement June 2021

Doctor of Philosophy dissertation of Parisa Ataei presented on June ?, 2021.

APPROVED:

Major Professor, representing Computer Science

Director of the School of Electrical Engineering and Computer Science

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Parisa Ataei, Author

ACKNOWLEDGEMENTS

[Eric. Committee. jeff. abu. parents. friends.]

TABLE OF CONTENTS

	<u>Page</u>
1 Introduction	1
1.1 Motivation and Impact	1
1.1.1 Motivating Example	1
1.2 Contributions and Outline of this Thesis	1
2 Background	2
2.1 The Relational Database	2
2.2 Relational Algebra	2
2.3 The Formula Choice Calculus	2
3 The Variational Database Framework	3
3.1 Variational Needs in a Relational Database	3
3.2 Variation Space in a Variational Database Framework	3
3.3 Variational Set	3
3.3.1 Variational Set Configuration	3
3.4 Variational Schema	3
3.4.1 Variational Schema Configuration	3
3.5 Variational Table	3
3.5.1 Variational Table Configuration	4
3.6 Variational Database	4
3.6.1 Variational Database Configuration	4
3.7 Properties of a Variational Database Framework	4
4 The Variational Query Language	5
4.1 Variational Relational Algebra	5
4.1.1 VRA Configuration	5
4.1.2 VRA Semantics	5
4.1.3 VRA Type System	5
4.1.4 VRA Variation-Minimization Rules	5
4.2 Variational Query Language Properties	5
5 Variational Database Management System (VDBMS)	6
5.1 Implemented Approaches	6

TABLE OF CONTENTS (Continued)

	<u>Page</u>
5.2 Experiments	6
6 Related Work	7
6.1 Instances of Variation in Databases	7
6.2 Instances of Database Variation Resulted from Software Development . . .	7
6.3 Variational Research	7
7 Conclusion	8
Bibliography	8
Appendices	9
A Variational Database Usecases	10

LIST OF ALGORITHMS

Algorithm

Page

Chapter 1: Introduction

[points: - def variation in db and is everywhere - instances but context specific solution don't suffice - thus a generic framework that addresses problems]

[- dimension, behaviour - instances but context specific solution don't suffice - example of instances - well-studied: schema evolution - partial: SPL - new instance either out of combination of other or new - generic framework to instantiate for each instance and address all variational needs.]

1.1 Motivation and Impact

[motivation]

1.1.1 Motivating Example

[combination of instances, behaviours, and dimensions]

1.2 Contributions and Outline of this Thesis

[contribution]

Chapter 2: Background

[background]

2.1 The Relational Database

[relational database]

2.2 Relational Algebra

[relational algebra]

2.3 The Formula Choice Calculus

[formula choice calculus]

Chapter 3: The Variational Database Framework

[needs. must have configuration.]

3.1 Variational Needs in a Relational Database

[needs and examples of them.]

3.2 Variation Space in a Variational Database Framework

[fexp. evaluation.]

3.3 Variational Set

[vset]

3.3.1 Variational Set Configuration

[vset configuration.]

3.4 Variational Schema

[vsch]

3.4.1 Variational Schema Configuration

[vsch configuration.]

3.5 Variational Table

[vtab]

3.5.1 Variational Table Configuration

[vtab configuration]

3.6 Variational Database

[vdb]

3.6.1 Variational Database Configuration

[vdb configuration]

3.7 Properties of a Variational Database Framework

[well-formed vdb properties.context-specific properties.]

[show that they hold for vdb.]

Chapter 4: The Variational Query Language

[vql]

4.1 Variational Relational Algebra

[vra]

4.1.1 VRA Configuration

[vra configuration]

4.1.2 VRA Semantics

[vra semantics]

4.1.3 VRA Type System

[type sys]

4.1.4 VRA Variation-Minimization Rules

[rules]

4.2 Variational Query Language Properties

[prop. show for vra.]

Chapter 5: Variational Database Management System (VDBMS)

[vdbms]

5.1 Implemented Approaches

[apps]

5.2 Experiments

[exp.]

Chapter 6: Related Work

[related work! have to work on this!]

6.1 Instances of Variation in Databases

[schema evolution. database versioning. data integration. data provenance.]

6.2 Instances of Database Variation Resulted from Software Development

[SPL. data model. query.]

6.3 Variational Research

[blah]

Chapter 7: Conclusion

[conclusion]

APPENDICES

Appendix A: Variational Database Usecases

A.1 Variation in Space

[enron email usecase]

A.2 Variation in Time

[employee evolution usecase]

