A SYSTEMS AND PROCESS MODEL FOR DATA EXPLORATION

BY

JOHN PETER LEE

ABSTRACT OF A DISSERTATION SUBMITTED TO THE FACULTY OF THE
COMPUTER SCIENCE DEPARTMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF SCIENCE IN COMPUTER SCIENCE
UNIVERSITY OF MASSACHUSETTS LOWELL
1998

Dissertation Director: Georges G. Grinstein, Ph.D.
Professor, Department of Computer Science

Abstract

Database exploration is the process of extracting knowledge from databases using visual, analytic and database tools. Database exploration has two main modeling components: a *systems integration model* that integrates the various tools in support of data exploration tasks, and a *user-centered model* of the data exploration process itself. This thesis takes a dual approach to database exploration by (1) developing and implementing a *systems integration model* based on data exploration tasks, and (2) defining a *generalized data exploration process* model that captures the essence of data exploration sessions. The systems integration model maps between the data models and interaction styles of database and visualization *exploration domains*. The process model describes data exploration without regard for the exploration domain. Two key components of the process model are a set of exploration domain-independent *metrics* that characterize elements of the process, and a set of data exploration *interaction patterns* that characterize database exploration sessions. The generalized process model is then applied to the systems model and implementation.

Acknowledgments

There are many people who assisted me in countless ways during the course of this research. First and foremost, I wish to thank my wife, Amy, for her loving support and patience over the many years it took to produce this work. This thesis is dedicated to her. I am equally thankful for our son, Maxwell Edward Lee, our “aggressive explorer” who has put everything into perspective.

I wish to thank my thesis advisor, Georges Grinstein, who has always been encouraging and supportive of this work, and who has always inspired me to thing big. I still marvel at the breadth and depth of his intellect, creativity and compassion. I also wish to thank my thesis committee: Dan Bergeron, Haim Levkowitz and John Sieg, for their guidance in crafting this thesis. I consider myself fortunate to be able to work with such outstanding teachers.

I wish to thank the NASA Graduate Student Researchers Program for supporting three years of this effort, the financial assistance from the University of Massachusetts Lowell, and the support of past and present employers: Paul Breen (MITRE), Ed Campbell (BBN) and Lorne Grant (Spacetec IMC).

I wish to thank my sources of inspiration: Steven Munno, Wassily Kandinsky, Alexander Calder, George Chaikin, Michael Leyton, John Becker, Ernest Wantuch, Edward Abbey, SUWA and the Colorado Plateau. Finally, I wish to thank David Salzman for lighting the fire...

Table of Contents

Page

List of Tables

Page

List of Figures

Page