



UNIVERSITI PUTRA MALAYSIA

VISUAL QUERY INTERFACES FOR DATABASES

HENDRY ANHAR

FSKTM 2004 1



VISUAL QUERY INTERFACES FOR DATABASES

HENDRY ANHAR

**MASTER OF SCIENCE
UNIVERSITI PUTRA MALAYSIA
2003**



VISUAL QUERY INTERFACES FOR DATABASES

By

HENDRY ANHAR

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in
Fulfilment of the Requirements for the Degree of Master of Science**

January 2004



Ntu' yan, aal



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment
of the requirement for the degree of Master of Science

VISUAL QUERY INTERFACES FOR DATABASES

By

HENDRY ANHAR

January 2004

Chairman: Associate Professor Hj. Ali Mamat, Ph.D.

Faculty: Computer Science and Information Technology

As information technology becomes the main force behind human activities, getting information should be done with less effort. Massive programming is a task users attempt to avoid. For novice users, a few steps of simple instructions are sufficient to get information they want. The same desire is required from database users. In databases, retrieving information is the dominant activity.

In this work, a visual query interface has been developed as an effort to facilitate query formulation. Java language was used as an implementation tool. The interface consists of two main parts, database creation and query formulation. A key feature of the interface is to support users formulating queries in simple way. With this interface, a user can express his or her query by using visual supports such as tick a table for selecting the table involved in the query. The interface not only allows visual



queries but also can produce result of the queries in a visual form. The database schema, referred by users, is also shown on the screen. Although the visual supports are provided, the way a query is formulated still follows the SQL style, namely it follows SELECT-FROM-WHERE clause.

Abstrak tesis yang dikemukakan kepada senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

VISUAL QUERY INTERFACES FOR DATABASES

Oleh

HENDRY ANHAR

Januari 2004

Pengerusi: Professor Madya Hj. Ali Mamat, Ph.D.

Fakulti: Sains Komputer dan Teknologi Maklumat

Apabila teknologi maklumat menjadi kuasa utama di belakang aktiviti-aktiviti manusia, perolehan maklumat patut dilakukan dengan usaha yang sedikit. Pengaturcaraan yang kompleks merupakan tugas yang pengguna cuba hindari. Bagi pengguna novis, beberapa langkah arahan mudah memadai untuk memperoleh maklumat yang mereka ingini. Hasrat yang sama diperlukan daripada pengguna pangkalan data. Di dalam pangkalan data, perolehan maklumat merupakan aktiviti yang dominan.

Di dalam kerja ini, satu antara muka pertanyaan visual telah dibangunkan sebagai usaha untuk memudahkan perumusan pertanyaan. Atur cara Java digunakan sebagai alatan implementasi. Antara muka tersebut terdiri dari dua bahagian utama, pembinaan pangkalan data dan perumusan pertanyaan. Ciri utama antara muka ini adalah untuk menyokong pengguna merumus pertanyaan dengan mudah. Dengan

antara muka ini, seorang pengguna dapat mengungkapkan pertanyaannya dengan menggunakan sokongan visual seperti menanda jadual yang betul untuk memilih jadual yang terlibat dalam pertanyaan. Antara muka tidak hanya membenarkan pertanyaan visual tetapi juga menghasilkan jawapan pertanyaan di dalam bentuk visual. Skema pangkalan data, yang dirujuk oleh pengguna juga dipaparkan pada skrin. Meskipun sokongan visual disediakan, cara pertanyaan dirumuskan masih mengikuti corak SQL, ianya mengikuti klausa SELECT-FROM-WHERE.

ACKNOWLEDGEMENTS

I would like to thank:

Associate Professor Hj. Dr. Ali Mamat as chairman, Associate Professor Dr. Hamidah Ibrahim and Mrs. Norwati Mustapha as members of committee for their academic advice and guidance during the period of this research.

To all staffs of the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia.

Lastly, to my family, my parent Anhar Hormat and Marsian Abd. Majid, my parent in law DR. Hj. Nasrul Madin and Hj. Martini, especially to my wife Perjunianti Nasrul, and my son Mohd. Ainal Mukramin.



I certify that an Examination Committee met on 16th January 2004 to conduct the final examination of Hendry Anhar on his Master of Science thesis entitled “Visual Query Interfaces for Databases” in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Hj. Md. Nasir Sulaiman, Ph.D.

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Hj. Ali Mamat, Ph.D.

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Hamidah Ibrahim, Ph.D.

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Norwati Mustapha

Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

GULAM RUSUL RAHMAT ALI, Ph.D.

Professor/Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:



This thesis submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Master of Science. The members of the Supervisory Committee are as follows:

Hj. Ali Mamat, Ph.D.

Associate Professor

Faculty of Computer Science and Information Technology

Universiti Putra Malaysia

(Chairman)

Hamidah Ibrahim, Ph.D.

Associate Professor

Faculty of Computer Science and Information Technology

Universiti Putra Malaysia

(Member)

Norwati Mustapha

Faculty of Computer Science and Information Technology

Universiti Putra Malaysia

(Member)

AINI IDERIS, Ph.D.

Professor/Dean

School of Graduate Studies

Universiti Putra Malaysia

Date:



DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UPM or other institutions.

HENDRY ANHAR

Date:



TABLE OF CONTENTS

	Page
DEDICATION	ii
ABSTRACT	iii
ABSTRAK	v
ACKNOWLEDGEMENTS	vii
APPROVAL	viii
DECLARATION	x
LIST OF FIGURES	xiii
CHAPTER	
1 INTRODUCTION	1
1.1 Introduction to Database Access	1
1.2 Problem Statements	3
1.3 Research Objective	5
1.4 Research Scope	5
1.5 Thesis Organization	6
2 LITERATURE REVIEW	8
2.1 Visual Languages	8
2.2 Visual Representation	10
2.2.1 Form Based Representation	11
2.2.2 Diagram Based Representation	16
2.2.3 Icon Based Representation	22
2.2.4 Hybrid Method	24
2.2.5 Browsing Strategy	26
3 INTERFACE DESIGN	28
3.1 Visual Supports	28
3.2 The Screen	29
3.2.1 Screen for Creating a Database	29
3.2.2 Screen for Formulating Queries	32
3.3 Scope of Query	35
3.3.1 Single Table Queries	35
3.3.2 Multi Table Queries	36
3.4 Visual Output	37
4 SYSTEM IMPLEMENTATION	38
4.1 Introduction	38
4.2 Database and Schemas Creation	38
4.2.1 Data Structures	39
4.2.1.1 Relation Schema	39
4.2.1.2 Data Types of Attributes	41



4.2.1.3	Actual Data	42
4.3	Query Processing	43
4.3.1	Query Transformation	44
4.3.1.1	Internal Representation of a Query	46
4.4	Query Execution	58
4.4.1	Identifying a Query	58
4.4.2	Query Connected to Database	60
4.4.3	Query Evaluation	62
5	USING AND TESTING THE SYSTEM	65
5.1	Introduction	65
5.2	Creating Database	66
5.3	Entering Data	67
5.4	Schema Representation	69
5.5	Query Formulation	71
5.5.1	Selecting an Attribute	71
5.5.2	Setting Condition of a Query	72
5.5.3	Examples of Query Formulation	74
5.6	Visual Output	79
6	CONCLUSION AND FURTHER WORK	83
6.1	Results and Discussion	83
6.2	Conclusion	84
6.3	Further Work	85
	REFERENCES	86
	BIODATA OF THE AUTHOR	90



LIST OF FIGURES

Figure	Page
2.1 QBE Interface	11
2.2 Table Representation in Microsoft Access	13
2.3 QBE Window in Microsoft Access	14
2.4 A Window for Performing a Query in Microsoft Access	15
2.5 Window for Representing Query Result in Microsoft Access	15
2.6 Formulating Query in QBD	17
2.7 Database Schemas Representation in SUPER	19
2.8 Working Schemas Representation in SUPER	19
2.9 Selecting Windows Representation in SUPER	20
2.10 Result Window Representation in SUPER	20
2.11 Iconic Representation of Objects in ICONICBROWSER	22
2.12 Operation Applicable to VCR Object Type	23
2.13 Setting the Condition of a Query	23
2.14 An Example of the Use of the SKI System	25
2.15 Example of Use of the LID System	27
3.1 Interface for Creating Database	30
3.2 Interface for Producing the Actual Data	31
3.3 Interface for Formulating Queries	32
3.4 A Format for a Query's Result	37



4.1	Database Creation Module	39
4.2	The Structure of the Relation Schema	40
4.3	An Example of Database Schema	40
4.4	The Representation of Database Schema	41
4.5	The Structure of Data Type	41
4.6	An Instance of Data Type	42
4.7	The Structure of the Actual Data	42
4.8	An Instance of Actual Data	43
4.9	Query Processor Module	43
4.10	The Structure of QG-1	44
4.11	An Instance of QG-1	45
4.12	The Structure of QG-2	48
4.13	The Structure of Seven Categories Block of Space	50
4.14	An Instance of the QG-1 and QG-2	58
4.15	Transferring Internal Query from Disk into Main Memory.	59
4.16	Structure of the Query	60
4.17	Arrangement of Tables Value	60
4.18	Matching queryArray and Data	61
4.19	An Instance of Query and Table	62
5.1	Interface to Create a Database	66
5.2	Database Schemas	67
5.3	Interface for Listing the Existing Tables	68
5.4	Interface to Formulate Data	69



5.5	Representation of Database Schema	70
5.6	The Process of Formulating the query5-1	75
5.7	The Process of Formulating the query5-2	77
5.8	Output of query5-1	80
5.9	Output of query5-2	81
5.10	Representation of an Image	81



CHAPTER 1

INTRODUCTION

1.1 Introduction to Database Access

An important issue in database (Atzeni and Antonellis, 1992, Dan and Richard, 1995, Thomas et al., 1999) is how to retrieve information from stored databases. This issue usually relates to database access techniques and classes of users. Since not all levels of users have high knowledge in databases, they need a lot of time to learn any aspects of database access.

To retrieve the information from a database storage, a user usually uses a traditional language, such as SQL. This kind of language is hard to use by non-programmers, such as novice users, who have limited expertise in Database Management Systems (DBMS).

As a solution, visual representation paradigms are introduced. The visual representation provides visual tools to assist users to interact with the system. This system is known as Visual Query System, while the language used is known as Visual Query Language.

In Catarci et. al. (1997), more than 40 systems have been proposed with different characteristics and advantages based on visual representation paradigms, such as forms,



icons, diagrams, and combination of them. These systems are aimed to facilitate users to interact with Database Management System, especially to do tasks in accessing and browsing the database.

Most of the systems include abilities to show database elements in separate windows, such as in Graphical Query Language (GQL) (Annamaria et. al., 1992). This system represents database schema and database instances in visual representation paradigms. The entities are represented as circles and their associated attributes are represented as ovals, the attributes' names appear inside the ovals, and relationship between entities are denoted by lines. With this Graphical Query Language, users feel at ease to use the system. Users can also easily perform a query, since they are required to know database construction such as table, attribute, and relationship between tables to perform a query.

Querying the database system is supported by the capability to browse the database itself. With this capability, it is possible to know the condition of data values in physical storage. In IconicBrowser (Catarci et. al., 1997), the system uses a 'list-of-values' to represent attribute's values to browse the database. Then users can choose the condition of data to be retrieved by clicking the range of values in the 'list-of-values'. Obviously, users can easily determine what kinds of data are shown in the output.



1.2 Problem Statements

In this research, the problem focuses on users who have limited knowledge in database system. This problem appears when a user tries to retrieve information from a database storage. Another problem appears when a user tries to formulate a query because a user has to know the structure of the data in the database. This means that users should know the database structure before formulating a query to improve information quality.

Many users, as a matter of fact, do not have sufficient skills on how to write program codes or do not have enough time to learn them. In principle, when a user begins to type a programming code, he/she is expected to understand database concepts and query syntax, as well as the underlying data model. This is a very difficult task which needs strong skills and leads to many errors.

As a solution, many authors have propose several tools that can support users to interact with a system without any experience on queries and databases. Tools such as windows, icons, menus, and pointers, are used as the basic idea of visual interface. Visual interface (Brossier et. al., 1995, Deepa and Tosiyasu, Dolin, 1999, Kamran and Mark, 1993) allows users to traverse the database naturally and pick, view, and update data. Users also can query the database without having to learn and write syntax, and semantic of query is expressed as a visual query, i.e. by drawing.

Although the above mentioned visual interface systems are already available, they have some weaknesses, and not suitable for expressing queries. For example, the system does not display the schema of a database and expressing a query is not straightforward.

Even the existing commercial system such as Microsoft Access has some deficiencies in this context. The deficiencies are:

1. Microsoft Access separates the interface for representing the overall database substance with the interface for formulating query. In formulating a query, user will need to activate another interface when he want to see the structure of a database.
2. To formulate a query (Query By Example), users are required to follow steps which have been stated. This could complicated when a query consists of several clauses (where clause).

Based on those limitations, it would be better to have a system, which is simpler and more efficient in expressing a query. The system should provide users with the ability to browse the database while expressing a query and users only need to follow simple steps to interrogate a database. Consequently this research will develop a system which is capable to overcome the problems mentioned above.

1.3 Research Objective

The objective of the research is to provide a system that allows users to express database queries using visual supports. The system is based on a relational data model and a database may include an image although it is not stored as part of the database.

1.4 Research Scope

This research concentrates on the design and implementation of a visual query interface to assist users in manipulating databases. For this purpose, a system, named Visual Query Interface (VQI) is designed and implemented based on visual approach. The design of the system focuses on users who have limited programming skills. Also, the system uses visual tools to perform a query.

Since the VQI system only supports alphanumeric data type and does not support very complex ones, such as graphics or sound bytes, we propose to implement relational data model to the system. Based on the model, users can use all of the database's attributes to build queries.

The VQI system is not a very complex system. Not all types of queries are supported by the system. The limitations are aimed to help author to detect errors easily when the system cannot process query or produce a wrong result of the query. The VQI system



works well on a single processor. The system is developed using Java Programming (Gary and Cay, 1996, Merlin et. al., 1999).

The implemented database is limited to seven tables, and each table is limited to nine attributes. When a user builds up more than seven tables, for example eight tables, the system only accesses the first seven tables and ignores the last one. Also, if a user uses more than nine attributes on a table, for example ten attributes, the system accesses only the first nine attributes and ignores the last one. The limitations will assist author to improve the system. When any error occur in the system, the author can easily determine which parts of the programming code that need to be changed.

1.5 Thesis Organization

The thesis is organized into 5 chapters. Chapter 1 presents the introduction, research objective, problem overview, research scope, and thesis organization.

Chapter 2 presents literature review. It starts with the description of visual language, visual query interfaces, including visual query approaches such as form-based and diagram-based representations.

Chapter 3 focuses on interface design. The layout of screens with visual support facilities is presented. The chapter also describes the query scope supported by the system.



Chapter 4 presents the system implementation. In this chapter, we describe the internal system architectures and query processing steps.

Chapter 5 focuses on using and testing the system. It describes how to use the system, create database and formulate a query using visual supports and it also concentrates on the system testing.

Chapter 6 presents the conclusion of the research. It also presents further work on the research.



CHAPTER 2

LITERATURE REVIEW

This chapter starts with a discussion on visual languages and visual query interfaces. The paradigms of visual presentations that are used as the basis for the visual query interfaces are discussed.

2.1 Visual Languages

Visual programming language (Andries, 1996, Bimbo et. al., 1992, Burnett and Baker, 1994, Chang, 1998, Margaret et. al., 1994, Raymon, 1991) is the way of using graphical technique in computer programming where the syntax of programming codes are denoted by symbols such as circles, diamonds and lines. Examples of visual programming language include Visual Basic Programming, Visual FoxPro, Oracle, etc.

A subclass (aimed at extracting information from database) of the more general class of visual language is visual query language. A visual query language (Catarci et. al., 1996, Dionisio and Cardenas, 1996, Kent and Robert, 1992, Oria and Tamer, 1998, Tiziana et. al., 1993, Yuksel, 1995) is a programming language based on the use of visual representation to depict the domain of interest and express the related requests.

