IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Hershey PA 17033-1117, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com
ITB7413

Chapter XI

Cooperative Query Processing via Knowledge Abstraction and Query Relaxation

Soon-Young Huh, Kae-Hyun Moon and Jin-Kyun Ahn
Korea Advanced Institute of Science and Technology, Korea

As database users adopt a query language to obtain information from a database, a more intelligent query answering system is increasingly needed that cooperates with the users to provide informative responses by understanding the intent behind a query. The effectiveness of decision support would improve significantly if the query answering system returned approximate answers rather than a null information response when there is no matching data available. Even when exact answers are found, neighboring information is still useful to users if the query is intended to explore some hypothetical information or abstract general fact. This chapter proposes an abstraction hierarchy as a framework to practically derive such approximate answers from ordinary everyday databases. It provides a knowledge abstraction database to facilitate the approximate query answering. The knowledge abstraction database specifically adopts an abstraction approach to extract semantic data relationships from the underlying database, and uses a multi-level hierarchy for coupling multiple levels of abstraction knowledge and data values. In cooperation with the underlying database, the knowledge abstraction database allows the relaxation of query conditions so that the original query scope can be broadened and thus information approximate to exact answers can be obtained. Conceptually abstract queries can also be posed to provide a less rigid query interface. A prototype system has been implemented at KAIST and is being tested with a personnel database system to demonstrate the usefulness and practicality of the knowledge abstraction database in ordinary database systems.

INTRODUCTION

Query language is used as a handy tool to obtain information from a database. However, current query processing has not been satisfactory in supporting effective decision analysis. This is largely due to exactness in nature of the relational databases and the query languages. When there is no matching data available, database users usually get a null information response. In terms of rigidness of query structure, even expert users have been frequently frustrated by the precise query specification syntax that requires them to be well-versed with the database schema for formulating correct queries. For many queries, it may be better for the database system to produce approximate answers when no exact answer is available. Usability of the database queries enhances significantly if the users are allowed to write imprecise queries, and the system understands the intent behind the queries.

Cooperative query answering aims at developing such intelligent systems that can accept less-precisely specified queries, analyze the intentions of such queries, suggest related questions or relax query conditions, and provide approximate neighborhood answers (Chu & Chen, 1994; Chu, Chen, & Lee, 1991; Cuppens & Demolombe, 1989; Minock & Chu, 1996; Motro, 1990; Prade & Testemale, 1984; Vrbsky & Liu, 1993).

One of the best applications of the cooperative query processing approach is probably the responsive online sales support in Web sites. Without intelligent help in conducting searches for a specific item on a Web site, customers are likely to be dissatisfied with their query's answer set. Unless users are very familiar with the contents of the database, they will obtain either no answer to the query or an excess of answers that might not be sorted in any usable way. This system entirely fails to provide sales support (Poo, Toh, Khoo, & Hong, 1999).

This chapter proposes a practical approach that can be easily developed for getting approximate answers from ordinary corporate databases. As a framework, it adopts an extended conceptual clustering framework, called knowledge abstraction hierarchy (KAH), capturing the value and domain abstraction and providing non-technical explanations for the approximate query answering mechanisms. Based on the KAH, a knowledge abstraction database is designed to support approximate query answering mechanisms. A prototype system has been implemented at KAIST to demonstrate the usefulness of KAH in ordinary database application systems.

This chapter is organized as follows. The second section provides backgrounds and literature review. The next section discusses the knowledge abstraction hierarchy (KAH) in terms of two abstraction perspectives: value abstraction and domain abstraction. Next, we construct a knowledge abstraction database that incorporates the KAH on a relational data model. Then, we describes the approximate query answering operations on the basis of the knowledge abstraction database. Section 6 provides the conclusion of the chapter.

BACKGROUND

A variety of approaches have been attempted along with semantic distance, fuzzy set, rule, and conceptual classification. The semantic distance approach (Ichikawa, 1986; Jain & Dubes, 1988; Motro, 1988) represents the degree of similarity between a pair of data objects by a numerical distance. This approach has advantages of ease and efficiency in developing query relaxation algorithms, since quantitative distances among data objects are easy to compute. Multiple distance metrics are also available for each attribute domain, and users are allowed to choose the direction of relaxation by mixing a selected set of distance metrics to compute a simplistic similarity value. However, this approach is limited due to the difficulty