Chapter 4
Processing of Queries with Fuzzy Similarity Domains

Soraya O. Carrasquel
Universidad Simón Bolívar, Venezuela

Rosseline Rodríguez
Universidad Simón Bolívar, Venezuela

Ricardo R. Monascal
Universidad Simón Bolívar, Venezuela

Leonid Tineo
Universidad Simón Bolívar, Venezuela

ABSTRACT

There are some data models and query languages based on the application of fuzzy set theory. Their goal is to provide more flexible DBMS that allow the expression of user preferences in querying as well as imprecision in data. In this sense, the FuzzyEER data model proposes four kinds of fuzzy attributes. One of them, named Type 3, consists of a set of labels provided of a similarity relation. An extension of SQL, named FSQL, allows the expression and use of fuzzy attributes. Nevertheless, FSQL does not allow using fuzzy attributes in some clauses based on data ordering, due to semantics problem. This chapter presents a solution for this problem in case of Type 3 fuzzy attributes. Main contribution consists in how to process queries involving such attributes by means of an extension to an existing RDBMS. Formal semantics, grammar, catalogue definition and translation schemas are contained in this chapter.

INTRODUCTION

Traditional databases only handle precise data and conditions. In many scenarios this is not quite fit for representing the actual information needs of users. In the real world, there might be partial or imprecise information about the values of some data. The user may also have selection criteria that are not precise.

The incorporation of some of these concepts in the modeling and manipulation of databases gave birth to different fuzzy relational model proposals (Buckles & Petri, 1982; Fukami, Umano, Muzimoto & Tanaka, 1979). These were later generalized to an extended model for fuzzy relational databases, known as GEFRED (Generalized model of Fuzzy Relational Databases) (Medina, Pons, & Vila, 1994).

From the GEFRED model, the Enhanced Entity-Relationship model (EER) and the Structured Query Language (SQL) language were formally extended into the FuzzyEER model (*Encyclopedia of Information Science and Technology*, 2014) and the FSQL language (Galindo, Urrutia, & Piattini, 2006),

that is an extension of SQL that allows flexible conditions in queries. The EER is an ER model that allows for generalizations/specializations and categories. The FuzzyEER model is an extension of the EER model for the creation of conceptual schemas with fuzzy notation and semantics. This extension includes elements of different conceptual model proposals for fuzzy databases, such as fuzzy attributes, fuzzy entities, fuzzy relations, and fuzzy specializations, among others.

In order to incorporate concepts from the FuzzyEER model (Galindo et al., 2006) in a relational DMBS, previous work includes a relational implementation schema called FIRST (Fuzzy Interface for Relational SysTem) (Medina, Pons, & Vila, 1995). The most recent version of this schema is called FIRST-2. The metadata that defines the fuzzy elements of FuzzyEER in FIRST-2 are stored in a relational catalog called FMB (Fuzzy Metaknowledge Base).

For representing fuzzy data, FuzzyEER and FSQL (Galindo et al., 2006) define for types of fuzzy attributes:

Type 1: Attributes with precise data values, provided with linguistic labels, interpreted as fuzzy numbers with the purpose of being used in fuzzy conditionals.
Type 2: Attributes with fuzzy data values, represented as fuzzy numbers. These attributes are possibility distributions on an ordered domain.
Type 3: Attributes with values in a domain formed by labels provided with a similarity relations between such labels. It can additionally have possibility distributions.
Type 4: Similar to Type 3, without the similarity relationship.

FSQL forbids using fuzzy attributes in ordering and grouping criteria, except for Type 1, that are actually precise values. This is unsatisfying. The representation of fuzzy attributes is very limited and only allows a certain set of values. FSQL allows any fuzzy binary relationship in place of a similarity relationship, as it does not have a well-defined semantics for them. These deficiencies must be addressed and corrected. The focus of this chapter is in Type 3 fuzzy attributes, restricted to the case of a set of labels with a provided similarity relationship.

We extend SQL in order to integrate this attributes type in the database queries. The main contribution of this chapter is to describe the translation schemes that allow translate instructions of extended SQL to instructions of native SQL. Our proposal consists into implement a mildly coupled architecture where the translation schemes are incorporated into a mediator layer which extends a Relational Database Management System (RDBMS).

BACKGROUND

Fuzzy Set theory was proposed by Zadeh (1965) as a way of representing imprecision and uncertainty. His initial motivation was control system applications, but with time they were beginning to be used in prediction and optimization, pattern recognition and expert systems. This theory provides a mathematic on formal computational framework for representing the notions vague or imprecise nature.

In classical set theory the membership of an element to a set is rigid, given by an indicator function whose codomain is $\{0, 1\}$, where 0 represents exclusion and 1 inclusion. Fuzzy sets are characterized by a membership function (μ_F) whose codomain is in the real interval $[0, 1]$. As the membership degree of an element approaches 1, this element is possibly (or certainly) more included in the set. Thus, $\mu_F(x)=0$

Related Content

Efficient Retrieval Technique for Microarray Gene Expression
J. Jacinth Salome (2012). *International Journal of Information Retrieval Research (pp. 43-51).*
www.irma-international.org/article/efficient-retrieval-technique-microarray-gene/74783

On the Design and Implementation of Interactive XML Applications
Jeff Brown, Rebecca Brown, Chris Veladoand Ron Vetter (2013). *Information Retrieval Methods for Multidisciplinary Applications (pp. 19-30).*
www.irma-international.org/chapter/design-implementation-interactive-xml-applications/75898

A Soft-Decision Histogram from the HSV Color Space for Video Shot Detection
Shamik Sural, M. Mohanand A. K. Malumdar (2005). *Video Data Management and Information Retrieval (pp. 237-253).*
www.irma-international.org/chapter/soft-decision-histogram-hsv-color/30768

Real Time License Plate Recognition from Video Streams using Deep Learning
Saquib Nadeem Hashmi, Kaushtubh Kumar, Siddhant Khandelwal, Dravit Lochanand Sangeeta Mittal (2019). *International Journal of Information Retrieval Research (pp. 65-87).*
www.irma-international.org/article/real-time-license-plate-recognition-from-video-streams-using-deep-learning/217484

Clustering Web Information Sources
Athena Vakali, Geroge Pallisand Lefteris Angelis (2008). *Personalized Information Retrieval and Access: Concepts, Methods and Practices (pp. 98-117).*
www.irma-international.org/chapter/clustering-web-information-sources/28070