(RESEARCH ARTICLE)

FRDBMS for imprecise queries based on GEFRED model

Subita Kumari *

Assistant Professor, Computer Science Department, Govt. PG College for Women, MDU Rohtak, Haryana, India.

Abstract

When users work with usual software tools they have to change their many valued logical thinking (approximate reasoning) into the two-valued computer logic. Although the Structured Query Language (SQL) is a very powerful tool, it is unable to satisfy needs for data selection based on linguistic expressions and degrees of truth. In this paper, we are interested in flexible querying that is based on fuzzy set theory. Medina et al. have developed a server named fuzzy SQL, supporting flexible queries and based on a theoretic model called GEFRED. To model the flexible queries and the concept of fuzzy attributes, an extension of the SQL language named fuzzy SQL has been defined. The FRDB is supposed to have already been defined by the user. In this paper, the work of medina et al. has been extended to implement a new architecture of fuzzy DBMS based on the GEFRED model. This architecture is based on the concept of weak coupling with the DBMS SQL Server.

Keyword: FSQL; SQL; Possibility Model; Fuzzy Querying; Membership Function

1. Introduction

Databases are a very important component in computer systems. Because of their increasing number and volume, good and accurate accessibility to a database becomes even more important. Organizations work with very large data collections mainly stored in relational databases. Linguistic expressions are interesting for data extraction, analysis, dissemination and decision making. The research area of fuzziness in Data Base Management Systems (DBMS) has resulted in a number of models aimed at the representation of imperfect information in Data Bases, or at enabling non-precise queries (often called flexible queries) on conventional database schemas. However, few works have been done from a practical point of view. In this paper, we are interested to the works of Medina *et al.* who introduced the GEFRED model in 1994 and its associated language named FSQL. This language presents new concepts such as fuzzy comparators, fuzzy attributes, fuzzy constants, *etc. We* propose to implement a new architecture of the Fuzzy Relational DBMS (FRDBMS) based on the GEFRED model. This architecture is based on the weak coupling principle with the RDBMS SQL Server.

2. Basic concepts

2.1. General Terms

- Database: A classical database is a structured collection of information (records or data) stored in a computer.
- Fuzzy Database: A fuzzy database is a database which is able to deal with uncertain or incomplete information using fuzzy logic.

---

* Corresponding author: Subita Kumari

- Fuzzy Logic: Fuzzy logic is derived from fuzzy set theory by Zadeh (1965) dealing with reasoning that is approximate rather than precisely deduced from classical predicate logic. It can be thought of as the application side of fuzzy set theory dealing with well thought out real world expert values for a complex problem.
- FRDB: It is an extension of the relational database. This extension introduces fuzzy predicates under shapes of linguistic expressions that, at the time of a flexible querying, permits to have a range of answers (each one with a membership degree) in order to offer to the user all intermediate variations between the completely satisfactory answers and those completely dissatisfactory .

- FRDBMS: It is an extension of the relational DBMS in order to treat, store and interrogate imprecise data.

- FRDB Models: Two broad approaches are possibilistic model and the similarity relation based model. These models are considered in a very simple shape and consist in adding a degree, usually in the interval [0, 1], to every tuple. They allow maintaining the homogeneity of the data in DB. The main models under both approaches are Prade-Testemale, Umano-Fukami, Buckles-Petry, Zemankova-Kaendel and GEFRED of Medina *et al..*

2.2. Different Models

Under this denomination, models using the possibility theory to represent imprecision are included. The most important models in this group are Prade-Testemale model, Umano-Fukami model, and GEFRED model. The GEFRED model is an eclectic synthesis of some of the previous models. It consists of a general abstraction that allows for the use of various approaches, regardless of how different they might l

Prade-Testemale Model: Prade and Testemale published a fuzzy relational database (FRDB) model that allows the integration of what they call Incomplete or uncertain data in the possibility theory sphere. An attribute A, having a D domain, is considered. All the available knowledge about the value taken by A for an x object can be represented by a possibility distribution $\pi A(x)$ about $D \cup \{e\}$, where e is a special element denoting the case in which A is not applied to x.

Umano-Fukami Model: This proposal also utilizes the possibility distributions in order to model information knowledge. In this model, if D is the discourse universe of A(x), $\pi A(x)(d)$ represents the possibility that A(x) takes the value $d \in D$. The following kind of knowledge may be modeled: unknown and applicable information, the non-applicable information (undefined), and the total ignorance (we do not know if it is applicable or non-applicable):

Unknown = $\pi A(x)(d) = 1, \forall\, d \in D$, Undefined = $\pi A(x)(d) = 0, \forall\, d \in D$

 Null = {1/Unknown, 1/Undefined}

GEFRED Model: The GEneralised model Fuzzy heart Relational Database (GEFRED) has been proposed in 1994 by Medina *et al.* .One of the major advantages of this model is that it consists of a general abstraction that allows for the use of various approaches, regardless of how different they might look. In fact, it is based on the generalized fuzzy domain and the generalized fuzzy relation, which include respectively classic domains and classic relations. It constitutes an eclectic synthesis of the various models published so far with the aim of dealing with the problem of representation and treatment of fuzzy information by using relational DB.

2.3. Fuzzy Attributes in GEFRED Model

In order to model fuzzy attributes we distinguish between two classes of fuzzy attributes-

Fuzzy Sets as Fuzzy Values: These fuzzy attributes may be classified in four data types. In all of them the values unknown, undefined, and null are included:

Fuzzy Attributes Type 1 (FTYPE1): these are attributes with "precise data", classic or crisp (traditional). However, they can have linguistic labels defined over them, which allow us to make the query conditions for these attributes more flexible.

Fuzzy Attributes Type 2 (FTYPE2): these attributes admit both crisp and fuzzy data, in the form of possibility distributions over an underlying ordered domain (fuzzy sets).

Fuzzy Attributes Type 3 (FTYPE3): they are attributes over "data of discreet non-ordered dominion with analogy" .In these attributes some labels are defined ("blond", "red", "brown", *etc.*) that are scalars with a similarity (or proximity) relationship defined over them, so that this relationship indicates to what extent each pair of labels be similar to each other.

Fuzzy Attributes Type 4 (FTYPE4): these attributes are defined in the same way as Type 3 attributes, without it being necessary for a similarity relationship to exist between the labels

Fuzzy Degrees as Fuzzy Values: The domain of these degrees can be found in the interval [0,1], although other values are also permitted, such as a possibility distribution . The meaning of these degrees is varied and depends on their use. The most important possible meanings of the degrees used by some authors are: fulfillment degree, Uncertainty degree, Possibility degree and Importance degree. The ways of using these fuzzy degrees are classified in two families: associated degrees (type 5, type 6, type 7) and non-associated degrees (type 8).

2.4. The FSQL Language

The FSQL language is an authentic extension of SQL language to model fuzzy queries.

Linguistic Labels: if an attribute is able of fuzzy treatment then linguistic labels can be defined on it. These labels will be preceded with the symbol $ to distinguish them easily. Every label has an associated trapezoidal possibility distribution (for fuzzy attributes type 1 and 2) or a scalar (for fuzzy attributes type 3 and 4). Valid statements in SQL are also valid in FSQL.

Fuzzy Comparators: besides the typical comparators (=,>, *etc.*), FSQL includes fuzzy comparators.

Function CDEG: the function CDEG (compatibility degree) may be used with an attribute in the argument to compute. It computes the fulfillment degree of the condition of the query for the attribute mentioned in its argument.

Fulfillment Thresholds: for each simple condition, a fulfillment threshold τ may be established (default is 1) with the format: <condition> THOLD τ indicating that the condition must be satisfied with minimum degree $\tau \in [0,1]$ to be considered.

Fuzzy constants: besides the typical constants (numbers, NULL…), FSQL included many constants such as fuzzy trapezoidal $[a, b, c, d], approximate values #n, $label, [n, m], UNKNOWN, UNDEFINED, etc.

Fuzzy Quantifiers: there are of two types: absolute and relative. They allow us to use expressions like "most", "almost all", "many", "very few", etc.   New Architecture of the FRDBMS

We propose the weak coupling approach with DBMS. The concept of weak coupling is shown in Figure 1.
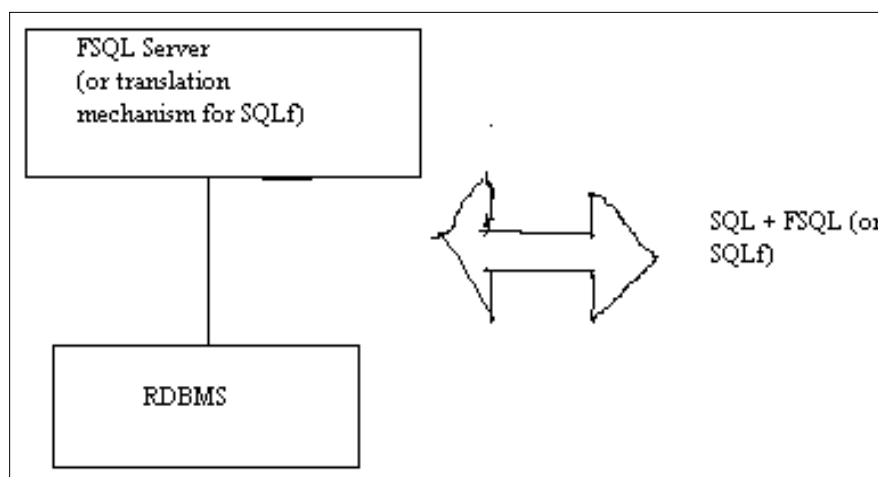

Figure 1 Weak Coupling Concept

The FRDBMS proposed respects the GEFERD model. The language of description and manipulation of the data is therefore FSQL. Seen that the FSQL language is an extension of the SQL language, a FRDBMS can model a RDB (described in SQL language) or a FRDB (described in FSQL language). The principle of this coupling is the definition of a software layer that allows the transformation of the command written by the user in FSQL language in their equivalent written in SQL. In order to implement a system which represent and manipulate "imprecise" information, Medina *et al.* have developed FIRST architecture (a fuzzy Interface for relational systems). It is built on RDBMS Client-Server architecture provided by Oracle. It extends the existing structure and adds new components to handle fuzzy information. The main important component added to this architecture is the FSQL Server which assures the translation of flexible queries written in FSQL in a comprehensible language by the DBMS (SQL).The installation of this architecture is described in Figure 2.
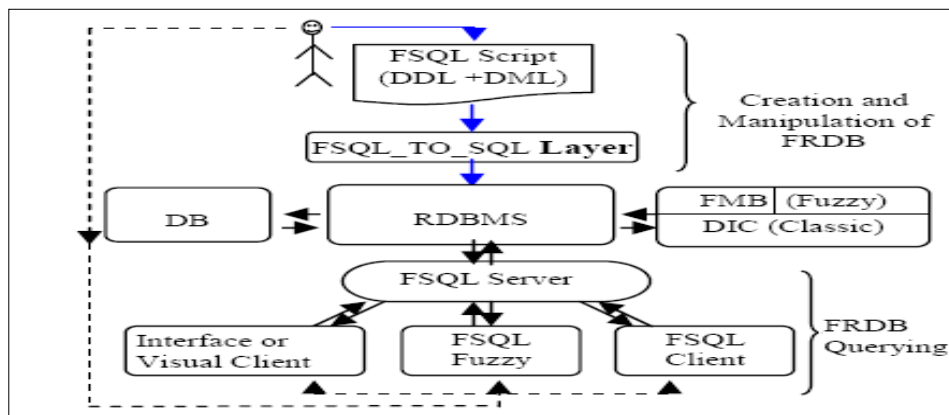

Figure 2 FRDMS Architecture

3. Conclusion and Future Scope

Fuzzy relational data bases have been extensively studied in a theoretical level. The majority of these works used the fuzzy sets formalism to model the linguistic terms as "moderate", "means" and to value the predicates including such terms. Medina et al. have developed a server named fuzzy SQL, supporting flexible queries and based on a theoretic model called GEFRED. This server has been programmed in PL/SQL language under Oracle database management systems. To model the flexible queries and the concept of fuzzy attributes, an extension of the SQL language named fuzzy SQL has been defined. The FSQL language extends the SQL language, to support the flexible queries, with many fuzzy concepts. The FRDB is supposed has already been defined by the user. In this proposal, we extend the work of medina et al. to implement a software layer which will convert FSQL queries to SQL queries. This architecture is based on the concept of weak coupling with the DBMS. This will facilitate the user a powerful and easy to use data mining tool which allows him to query data from databases by using linguistic expressions in order to improve the quality of selection process. The proposed architecture of FRDBMS based on the GEFRED model makes use of weak coupling concept with the DBMS. As a future work a new architecture supporting the concept of strong coupling with DBMS can be developed. As futures perspectives of this work, we also mention the automatic mapping of existing relational DB to FRDB. This point is theoretically done but not implemented yet, so we think that it will contribute to make easier the use of the FRDB in real applications.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1]   Touzi, A. G., & Hassine, M. A. B. (2009). New architecture of fuzzy database management systems. *Int. Arab J. Inf. Technol.*, *6*(3), 213-220.

[2]   Bosc, P., & Pivert, O. (1995). SQLf: A Relational Database Languageb5n6for Fuzzy Querying. *Computer Journal of IEEE Transactions on Fuzzy Systems*, *3*(1), 80-89.

[3] Bosc, P., & Pivert, O. (2000). SQLf query functionality on top of a regular relational database management system. *Knowledge Management in Fuzzy Databases*, 171-190.

[4] Galindo, J., Medina, J. M., Pons, O., & Cubero, J. C. (1998). A server for fuzzy SQL queries. In *Flexible Query Answering Systems: Third International Conference, FQAS'98 Roskilde, Denmark, May 13–15, 1998 Proceedings 3* (pp. 164-174). Springer Berlin Heidelberg.

[5] Galindo, J., Aranda, M. C., Caro, J. L., Guevara, A., & Aguayo, A. (2002). Applying fuzzy databases and FSQL to the management of rural accommodation. *Tourism Management*, *23*(6), 623-629.

[6] Galindo, J. (2005). New characteristics in FSQL, a fuzzy SQL for fuzzy databases. *WSEAS Transactions on Information Science and Applications*, *2*(2), 161-169.

[7] Galindo, J., Urrutia, A., & Piattini, M. (Eds.). (2006). *Fuzzy databases: Modeling, design, and implementation*. IGI Global.

[8] Grissa, A., Ben Hassine, A., & Ounelli, H. (2006). Extended_FSQL_Server: A Server for the Description and the Manipulation of FRDB. In *Proceedings of the 4th International Multi Conference on Computer Science and Information Technology (CSIT)* (pp. 454-464).

[9] Mama, R., & Machkour, M. (2021). Fuzzy querying with SQL: Fuzzy view-based approach. *Journal of Intelligent & Fuzzy Systems*, *40*(5), 9937-9948.

[10] Rachid, M., & Mustapha, M. (2022). A Comparative Study of Existing Fuzzy Query Systems of Database. In *Distributed Sensing and Intelligent Systems: Proceedings of ICDSIS 2020* (pp. 465-476). Cham: Springer International Publishing.

[11] Mama, R., & Machkour, M. (2019, July). A study on fuzzy interrogation systems of database. In *2019 International Conference of Computer Science and Renewable Energies (ICCSRE)* (pp. 1-6). IEEE.

[12] Mama, R., & Machkour, M. (2019, October). A study of fuzzy query systems for relational databases. In *Proceedings of the 4th International Conference on Smart City Applications* (pp. 1-5).

[13] Mama, R., & Machkour, M. (2020). Fuzzy questions for relational systems. In *Innovations in Smart Cities Applications Edition 3: The Proceedings of the 4th International Conference on Smart City Applications 4* (pp. 104-114). Springer International Publishing.