ITFOOD: Indexing Technique for Fuzzy Object Oriented Database.

Priyanka J. Pursani, Prof. A. B. Raut

Abstract: The Indexing Technique for Fuzzy Object Oriented Database Model is the extension towards database models to support complex data and to handle fuzziness in the complex database along with indexing to speed up the evaluation of fuzzy queries. In today's era as traditional commercial market is changing rapidly towards specialized market where need of internet is increasing. So to satisfy the requirements of this specialized market Object Oriented Database (OOD) has been developed which would soon become the primary database technology as Relational database were not designed to handle the type of multimedia data frequently found on the web. The data present on the web is highly uncertain. So to handle the uncertainty on the web, Fuzzy Object Oriented Database (FOOD) model is required. Due In this paper FOOD index is proposed which deals with various kinds of fuzziness and provides indexing technique based on R tree indexing which supports various fuzzy queries.

Index Terms: Object Oriented database, Fuzzy Set theory, Fuzzy queries, Fuzzy Indexing, R tree indexing.

I. Introduction:

Among various database models, OOD has been on the focus of researchers as it supports a semantic modelling of data with abstraction mechanisms more powerful than traditional data models. At the same time, they offer tools to manage hierarchies of structured objects in an efficient way through the inheritance relation [1]. The object-oriented database model, which provides powerful data modeling features, has gained more popularity in recent years [2], [3].

Most of the existing database models are designed under the assumptions that the data/information stored is precise and queries are crisp. In fact, these assumptions are often not valid for many of the next generation information systems since they may involve complex information with uncertainty. In general, data/ information in databases may be uncertain for the following reasons:

- A decision in many knowledge-intensive applications usually involves various forms of uncertainty.

ISSN: 2278 – 1323

- Integrating data from various sources is not usually a crisp process, while unifying various heterogeneous data into an integrated form, due to possible semantic differences (and other reasons), sometimes forcing data to be completely crisp may result in falsity and useless information.
- Information in some non-traditional applications is inherently both complex and uncertain, i.e., representing subjective opinions and judgments concerning medical diagnosis, economic forecasting, or personal evaluation.
- In natural languages, numerous linguistic terms with modifiers (e.g., "very," "more or less," etc.) and quantifiers (e.g., "many," "few," "most," etc.) are used when conveying vague information.

Handling uncertainty in databases was first proposed on relational-based database models [4], [5], [6].

The indexing of Fuzzy object oriented databases has not received much attention yet. Index structures allow fast access to data by content. The current crisp index structures [7] developed for OODBs are inappropriate to represent and efficiently access fuzzy data for the Fuzzy Object Oriented Database model i.e. FOOD model. Fuzzy querying allows one to express vague predicates represented by fuzzy sets. Conventional index structures cannot be used directly since fuzzy predicates may not refer to the entry values of the index. Therefore, an efficient indexing mechanism for the Fuzzy Object Oriented Database model is needed to allow fast access to the objects with crisp or fuzzy values. In order to support exact, range, and fuzzy queries efficiently, a index structure that can use both crisp and fuzzy attributes as organizing attributes of objects should be used. Therefore, in this proposed work a new index structure, dealing with different kinds of fuzziness in the Fuzzy Object Oriented Database model i.e. FOOD model is introduced .The Index structure handles various types of flexible queries including crisp, range, and fuzzy queries.

In the literature very few indexing techniques are available for fuzzy Object Oriented Databases. The indexing techniques defined in the framework of object-oriented data models are either *structural* [8], [9], [12], [13], or *behavioral* [11]. Structural indexing is based on object attributes and can be classified into techniques supporting nested predicates, such as the ones presented in [4] and techniques supporting queries issued against an inheritance hierarchy [9]. On the other hand, behavioral indexing aims at efficient execution of queries containing method invocations. It is based on precomputing or caching the results of a method and storing them in an index. The major difficulty in this approach is the

ISSN: 2278 – 1323

detection of changes invalidating the results of a method.

The fuzzy indexing structure proposed in [10] uses one index per fuzzy predicate tied to an attribute. This indexing structure only deals with homogeneous domains and assumes that the underlying relations are crisp.

II. ITFOOD Model:

The Architecture of the proposed model for Indexing Technique for Fuzzy Object Oriented Database is given in Fig 1.

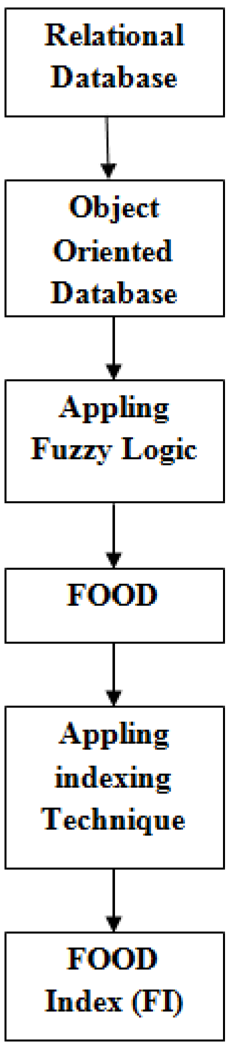

Fig.1. Architecture of ITFOOD model

A. FOOD Model:

The FOOD model is similarity-based [14]. For fuzzy attributes, fuzzy domains and similarity matrices are defined. Similarity matrices are used to represent vague relationships within the fuzzy attributes. The domain, dom, is the set of values that the attribute may take, irrespective of the class it falls into. The range of an attribute, rng, is the set of allowed values that a member of a class, i.e., an object, may take for an attribute. A range for each attribute of the class is defined as a subset of a fuzzy domain. The range definition for attribute ai of class C is represented by the notation, $rng_c(a_i)$, where

$a_i \in$ Attr(C)={$a_1,a_2,.....,a_n$}

Attr(C) refers to the attributes of class C. Similar objects are grouped together to form a class. An object belongs to a class with a degree of membership. Fuzziness may occur at three different levels in the FOOD model; the attribute level, the object/class level, and the class/ superclass level. The degree of membership of the attribute is given manually by the user.Similarity Matrix of fuzzy valued attribute colour can be given as shown in table I.

ISSN: 2278 – 1323

	Red	Orange	White
Red	1	0.6	0
Orange	0.6	1	0.1
White	0	0.1	1

TABLE I: Similarity Matrix for Fuzzy valued attribute Colour.

Attribute Definitions of FOOD Model: A fuzzy object can have attributes in the form of a set of values from a fuzzy domain. The similarities between the elements in that fuzzy domain are represented using a similarity matrix. For instance, consider a class, BOOK, with an attribute defined in the domain, Colour = {red, orange, white}, with the similarity matrix shown in Table 1. According to the table, the degree of similarity between the elements "red" and "orange" is 0.6.

Every class has a range definition for each of the fuzzy attributes with the corresponding relevance values indicating the importance of that attribute in the definition of that class. Range values show ideal values of a class. However, an attribute of a class can take any value from a domain without considering the range values. Object/class MDs are calculated from the similarity of object values to the range values using formula (1). Similarly, class/superclass MDs are calculated from the similarity of range values of two classes. In this model, semantics is associated with range definitions to permit a more precise definition of a class. When an attribute is multi-valued, semantics defines the relationships among these values. There are three semantics used in the FOOD model: AND, OR, and XOR. AND semantics requires objects to have all the values given in the range definition, OR semantics requires objects to have at least one of the values given in the range definition, and finally, XOR semantics requires objects to have exclusively one of the values given in the range definition. The details of the attribute representations and the other features of the FOOD model are described in [15].

Yazici et.al has well explained the relationship between class/object and superclass/subclass [15],[16].

B. FOOD Index:

Indexing in OODBS is a lot more complicated than in RBDS. One difference between objects and relational tuples is that objects are not flat. Therefore one should be able to index on instance variables that are nested several levels deep in an object to be indexed.

Indexing for OODBS is first proposed for the GemStone data model. It is a generalization of an indexing technique for path expressions [17].

FOOD index is developed using R tree:

- R tree is height balanced tree similar to B tree.

ISSN: 2278 – 1323

- Each node in the R-tree represents a rectangular region in space that encloses a group of uncertain objects [19].
- R tree is well suited for secondary memory. It makes R tree applicable for large database where index is too large to fit in main memory [18].The basic structure of R tree is shown in fig. 2.

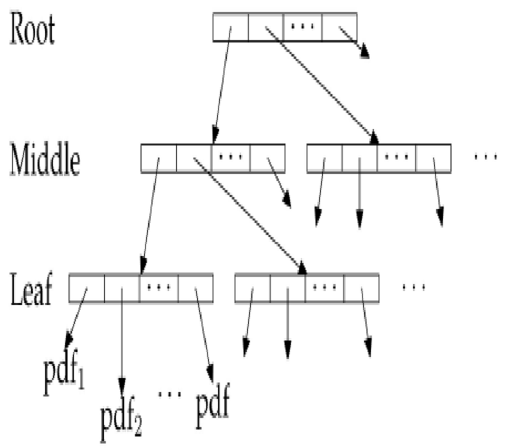

Fig . 2. Structure of R Tree

C. Conclusion:

In this work, we introduced the FOOD index (FI) which can be implemented to enhance the performance of fuzzy queries in Object Oriented Database. It can be used for inheritance hierarchy and to deal with fuzziness in OOD model.

By introducing Fuzziness in OOD, it helps to enhance the speed to access the objects in related groups. Along with it we apply indexing on FOOD model which increases the speed to access the objects. Hence we can say the performance of Object Oriented Database can be increased by implementing ITFOOD.

References:

[1]. [Pasi & Yager, 1999] , Gabriella Pasi and Ronald R. Yager, "Calculating Attribute Values Using Inheritance Structures in Fuzzy Object-Oriented Data Models" , *IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews,* Vol. 29, No. 4, pp. 556-565, November 1999

[2]. [Bancilhon et al.,1992]F. Bancilhon, C. Delobel, and P. Kanellakis, "Building an Object- Oriented Database System" *Morgan Kaufmann Publishers, Inc.*,1992.

[3]. [Harrington,2000] J.L. Harrington " Object-Oriented Database Design" Morgan Kaufmann Publishers, Inc., 2000.

[4]. [Bosc & Pivert, 1995] P. Bosc and O. Pivert, "SQLf: A Relational Database Language for Fuzzy Quering," *IEEE Transactions on Fuzzy Systems*, Vol. 3, No. 1, pp. 1-17, 1995.

[5]. [Petry, 1996] F.E. Petry "Fuzzy Databases, Principles and Applications". *Kluwer Academic Publishers*, 1996.

[6] [Takahashi, 1993] Y. Takahashi, "Fuzzy Database Query Languages and Their Relational Completeness Theorem,"

ISSN: 2278 – 1323

IEEE Transactions on Knowledge and Data Engineering, Vol. 5, No. 1, pp. 122-125, 1993.

[7]. [Bertino & Guglielmina, 1993] E. Bertino and C. Guglielmina, "Path-index: An approach to the efficient execution of object-oriented queries," *IEEE Transactions on Knowledge and Data Engineering.*, vol. 10, no. 1, pp. 1–27, Feb. 1993.

[8]. [Bertino, 1991] E. Bertino, "An indexing technique for object-oriented databases," in *Proceedings of 7th Inernational Conference on Data Engineering* , pp. 160–170, 1991.

[9]. [Bertino & Kim, 1989] E. Bertino and W. Kim, "Indexing techniques for queries on nested objects ,"*IEEE Transactions on Knowledge and Data Engineering*, Vol. 1, No. 2, pp. 196–214, 1989.

[10].[Bosc & Pivert,1992] P. Bosc and O. Pivert, "Fuzzy querying in conventional databases," in *Fuzzy Logic for Management of Uncertainty*, L. A. Zadeh and J.Kacprzyk, Eds. New York: Wiley, 1992, pp. 645–671.

[11]. [Bertino & Kim, 1989] E. Bertino and W. Kim, "Indexing techniques for queries on nested objects," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 1, No. 2, pp. 196–214, 1989.

[12]. [Kemper & Moerkotte, 1991] A. Kemper and G. Moerkotte, "Access support in object bases," in the Proceedings of ACM SIGMOD, 1991.

[13]. [Kim et. al. ,1989] W. Kim, K. C. Kim, and A.Dale, "Indexing techniques for object-oriented databases," *in Object-Oriented Concepts, Databases, and Applications, MA: Addison-Wesley*, 1989.

[14]. [Zadeh, 1971] L.A. Zadeh, "Similarity Relations and Fuzzy Orderings," published in the Journal of Information Sciences, Vol. 3, No. 2, pp. 177-200, 1971.

[15]. [Koyuncu & Yazici, 2003] M. Koyuncu and A. Yazici, "IFOOD: An intelligent fuzzy object-oriented database architecture," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 15, No. 5, pp. 1137–1154, 2003.

[16].[Yazici et.al. ,2008] A. Yazici, Cagri Ince and Murat Koyunce "FOOD Index: A Multidimensional Index Structure for Similarity-Based Fuzzy Object Oriented Database Models", *IEEE Transactions on Fuzzy Systems,* Vol. 16, No. 4, 2008.

[17].[Lee] M. Lee, "Object-Oriented Database Theory: An Introduction & Indexing in OODBS" *Database Hall of Fame, TU Muenchen*.

[18]. [Gavrilla , 1994] D.M.Gavrila, "R Tree indexing", 1994.

ISSN: 2278 – 1323

[19]. [Kao et.al. , 2010] B. Kao et al., B. Kao, S. Lee, F. Lee, D. Cheung & W. Ho "Clustering Uncertain Data using Voronoi Diagrams and R-Tree Index*",* *IEEE Transactions On Knowledge and Data Engineering,* Vol. 22, No. 9, pp. 1219 – 1233,2010

ISSN: 2278 – 1323