Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 9, Issue. 6, June 2020, pg.120 – 131

An Appraisal of the State of Arts in Fuzzy Database Modeling

Adegoke, O. Folake[1]; Okoronkwo C. Matthew[2]; Bello Muriana[3]
[1]Department of Mathematical Sciences, Kogi State University, Anyigba, Nigeria
[2]Department of Computer Science, University of Nigeria, Nsukka
[3]Information Technology and Resource Center, kogi State University, Anyigba, Nigeria
[1] adegoke.fo@ksu.edu.ng, [2] matthew.okoronkwo@unn.edu.ng, [3] muriana.b@ksu.edu.ng

Abstract: Most real-world data are vague, imprecise and imperfect. Classical model which includes file system model, the hierarchical model and network model were lacking the ability to integrate and manipulate vague, imprecise and imperfect data. Fuzzy database model uses fuzzy sets and logic to extend the classical relational database model as a functional way of handling fuzziness. This paper appraises state of the arts in fuzzy database model that can handle uncertainty data with emphasis on the Zvieli and Chen entity relation model, Relational model of Chaudhry, Moyne and Rundensteiner, extended relational model of Yacizi and Merdan, extended entity relational model of Chen and Kerre, fuzzy object oriented database model of Ma, Zhang, Ma and Chen. The limitations and possible improvement of the models were considered.

Keywords: Imprecision, Data, Database, Uncertainty, Approach, Fuzzy and Model.

1. Introduction

In the world today, database systems are crucial for both commercial and private data management. Most times, data could be precise or imprecise depending on the type or source of the data. The imprecision of data led to the development of fuzziness in data manipulations in database systems. Often, organizations are faced with voluminous amount of data generated from non-conventional means, for instance, sensors and radio frequency identification (RFID) which is mostly not defined or imprecise.

As early as in the 1980's, Zadeh's fuzzy logic has been employed to extend various data models. The major aim of introducing fuzzy logic in database is to develop and enhance special models in other to represent and manipulate uncertain and imprecise information. This assertion has led to numerous researchers and contributions particularly with respect to the relational model or closely related forms of it.

Also, recent development in computing power has introduced the need for database in emerging application, for instance: Multimedia, GIS, CAD/CAM. This emerging application requires the manipulation and modeling of complex data with the semantic relationship. Studies has equally shown that object-oriented paradigm lends itself well in the requirement of fuzziness but classical relational database model and their extension of fuzziness do not meet the need of manipulating complex objects/data with uncertainty and imprecision. Researchers have recently concentrated on object-oriented database model in other to figure out the best ways to handle complex objects and ever-present uncertain data. A lot of merits have been achieved in the quest for this research with many models and approaches proposed by different researchers with few comprehensive reviews on the state of the arts in fuzzy database models. There is lack of elaborate appraisal or review of different approaches proposed by different researchers of handling imprecise data.

This paper is aimed to present a comprehensive appraisal of the state of the arts in database models by reviewing the proposed approaches by different researchers that have contributed significantly in the handling of imprecise data. Zvieli and Chen approach, Yazici and Merdan approach, Chen and Kerre approach, Chaudhry, Moyne, and Rundensteiner approach, Ma,Zhang, Ma, and Chen proposal are reviewed.

2. Imprecise and Uncertain Information

The following are used to describe imperfect information in database systems. They include; inconsistency, vagueness, imprecision, uncertainty and ambiguity. The management of uncertainty is crucial in database development because the information is often vague. A classification by (Amiahi Motro, 1997) states that fuzzy information is content-dependent and defined them as follows

Uncertainty: The information cannot be determined if it is true or false. For instance, Frank may be 30 years old.

Imprecision: The available information may not be specific enough to decode, for instance Mark is 30 or 32 years old (disjunction), Mark is not 30 years old (Negative), Mark is between 30 and 35 years old or Mark is over 30 years (range information), Mark is 31 ± 1 years (error margin information) or simply unknown information.

Vagueness: This includes information (predicates or qualifiers) that are vague. For instance, Mark is in his early years or Mark is at the end of his youth.

Inconsistence: This present two or more information that cannot be true at the same time. For instance, Mark is 20 and 30 years old. This presents a special case of disjunction.

Ambiguity: some part of the information lacks complete semantics (or complete meaning), for instance, it is unclear whether the celebration is annually or monthly.

A number of researchers have made contributions towards the best ways to handle fuzzy based data. The next sections of the paper present a holistic appraisal of their findings and suggestions.

3. The Zvieli and Chen Approach

Zvieli and Chen presented an approach in entity relationship (ER) model with an introduction of three levels of fuzziness in ER model. The presented levels of fuzziness are as follows:

a. At the first stage, entity sets relationships and attribute sets could be fuzzy, which mean the existence of membership degree to the model. For instance, Figure 1 shows that the fuzzy entity “Company” has a 0.9 membership degree, the relationship “Recruits” has a 0.5 degree of membership and the fuzzy attribute “Accept” has a 0.7 membership degree.
b. The second level of fuzziness introduced is closely related to the occurrence of fuzzy attributes of special entities and their relationships. For instance, an “entity” average worker must be fuzzy because its instances workers are members of the entity with different membership degrees.

c. The fuzzy values of attributes of special entities and relationship are concerned at the third level. For instance, "quality" of a football player may be fuzzy (with these possibilities: bad, good, very good etc).

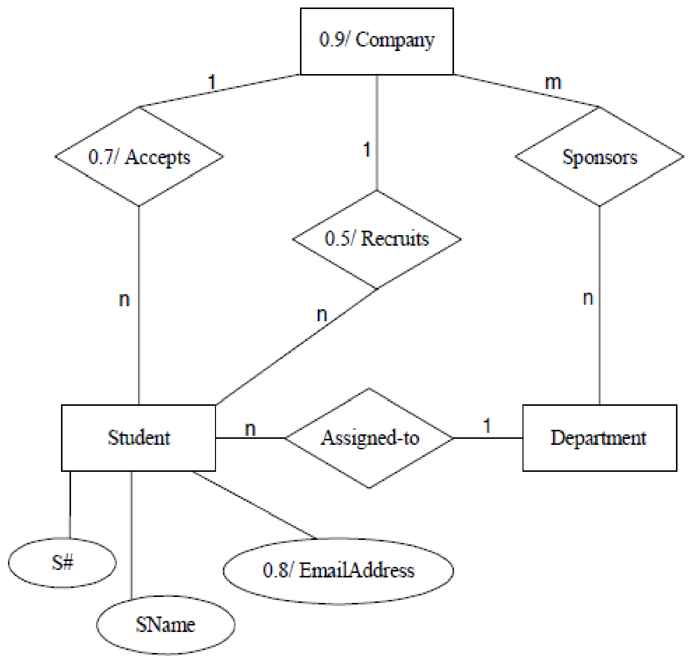

Figure 1: Membership degree of the model in certain sets (entities, relationships or attributes)

The first assertion is considered useful but toward the end, there should be consideration whether such an entity relationship or attribute "will or will not" appear during implementation.

The second assertion is equally reasonable and useful but there should be consideration of different degree meaning (membership degree, importance degree, fulfillment degree, and so on).

The third assertion is also useful and similar to specifying the data type of source attributes, with the understanding that fuzzy values belong to fuzzy data types.

4. Yazici and Merdan Approach

Yazici and Merdan proposed an extension of the IFO model, mainly for processing of imprecise and specialized data where there is existence of similarity in a label. This extension is called EXIFO, and by means of examples, there were good explanation of the implementation of the fuzzy conceptual scheme by studying the representation of uncertain attributes. In their model, three new constructors were added and by using these constructors, it is easy and possible to represent attributes that explicitly have uncertain values. As detailed in the EXIFO conceptual model, allows impression and uncertainty in database models which is based on IFO conceptual model.

The researchers introduced the use of fuzzy-valued attributes, incomplete-valued attributes. In the first instance the true data could belong to a specific set or subsets of certain values. For example, a set of colors may be the domain of this attributes {orange, red, yellow, blue} or a subset {orange and yellow} which present a relationship between the colors. In their second presentation, the data value is unknown. For example, the domain of the attributes could be a set of years between 2000 and 2012. In the third instance, the main or true data value available but not expressed precisely: for example, it could present a domain of attributes with check whether a certain phone number exist, each of the presented attribute has a definition and a graphical representation. In their research, they proposed a high-level primitive semantics which are related to each with the logic operators OR, XOR, or AND to model fuzzy entity type. The main contribution of this approach is the use of an extended NF relation (Non-First Normal Form) for transformation of a conceptual design into a logical design.

The strategy aimed at analyzing the attributes that compose the conceptual model in order to establish NF model. It focused on conceptual modeling approach for the representation of complex uncertain information, by introducing the use of an object-oriented databases paradigm and an algorithm for transforming a conceptual schema specification of the EXIFO model into a logical schema of the fuzzy object- oriented database model (FOOD). EXIFO tries to preserve the acquired strength of semantic approaches while aliening concepts from the object-oriented paradigm and fuzziness by adding new constructors.

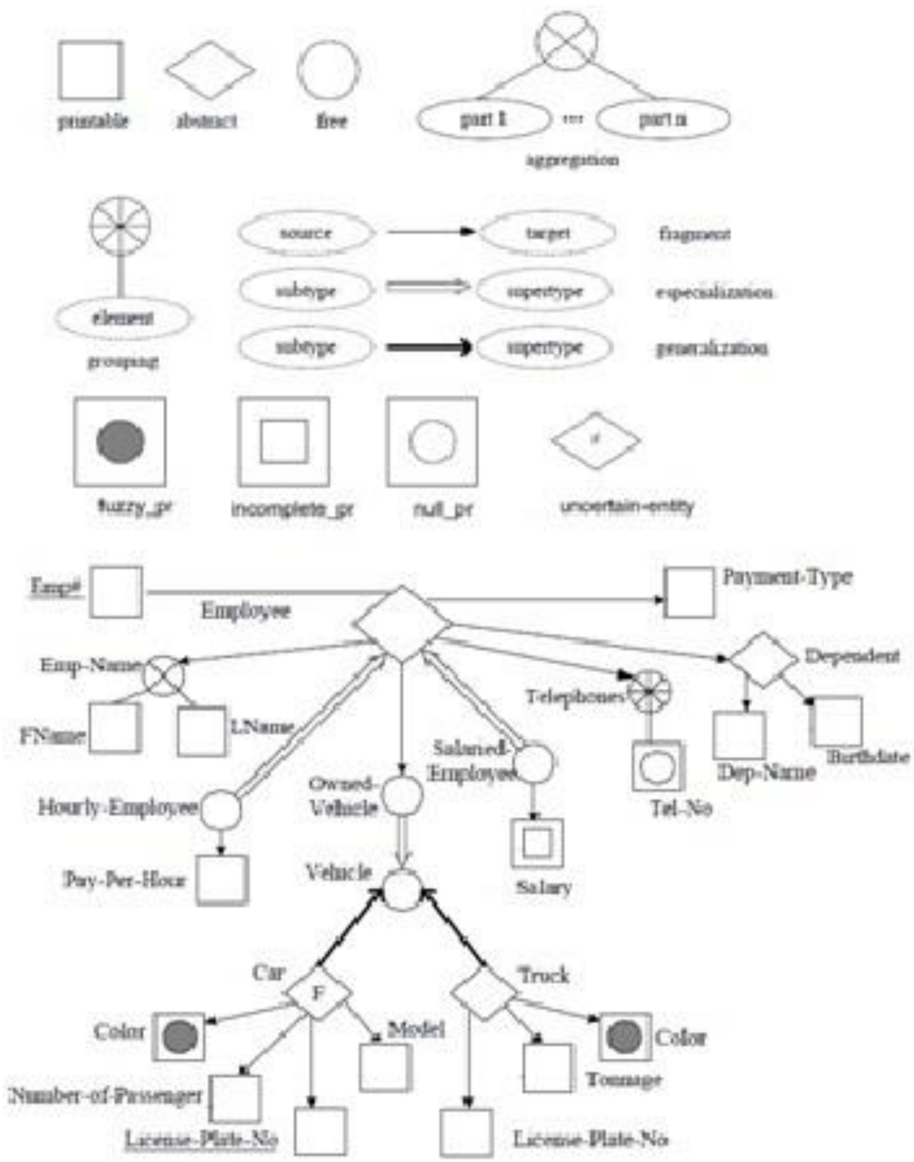

Figure 2: Fuzzy EXIFO model proposed by Yazici and Merdan. Notation and example of employee-vehicle

5. The Chen and Kerre Approach

The two researchers focused on the extension of EER concepts, for instance (super-class, subclass, generalization, specialization, category, and shared subclass) without including graphical representations. The main idea is that if E1 is a superclass of E2 and $e \in E2$, then $E1(e) \leq E2(e)$, where E1(e) and E2(e) are the membership functions of e to E1 and E2, respectively. They discuss three kinds of challenges in fuzzy relationships and discovered the following constrains: The inheritance constraint, which means that a subclass instance inherits all relationship instances in which it has participated as a super class entity, the total inclusion constraint for entity E is defined when for any instance in E, $\exists\ \alpha$ such that $\alpha > 0$, where α is one membership degree in the fuzzy relationship, and the cardinality constraints 1:1, 1:N, and N:M are also studied with fuzzy relationships.

The fuzzy ER model proposes a model generated by M = ($\boldsymbol{E}$, $\boldsymbol{R}$, $\boldsymbol{A}$) expressed by E as entity type, R as interrelation type, and A as an attributes, also adding label types that generate at the first level, L1(M) = E, R, A_E, A_R and proposes four set types, with a corresponding graphic notation and where μ_X is the membership function to the set X (one Entity, one Relationship or one Attribute) and D_E is the domain of $\boldsymbol{E}$ composed of all possible entity types concerned:

$E = \{\mu_E(E)/E\colon E \in D_E\ and\ \mu_E\ (E) \in [0,1]\}$

$R = \{\mu_R(R)/R\colon R$ is a relationship type involving entity types in $D_E\ and\ \mu_R\ (E) \in [0,1]\}$

$A_E = \{\mu_{AR}(A)/A\colon\ A$ is an attribute type of entity type $E\ and\ \mu_{AE}(A) \in\ [0,1]\}$

$A_R = \{\mu_{AR}(B)/B\colon\ B\ A$ is an attribute type of relationship type $R\ and\ \mu_{AR}(B) \in\ [0,1]\}$

The participation constraint is modeled setting that an entity E λ- participates in R if for every e of E, there exists an f in F such that $\mu_R(e,f) >= \lambda$. The cardinality constraint is shown in Figure 3 where N and M are fuzzy sets. The concept of fuzzy quantifier is not used in this approach, at the second level, for each entity type E and relationship type R, the sets of their values can be fuzzy sets, reflecting possible partial belonging of the corresponding values to their types. The third level of fuzzy extensions concerns attributes and their values. For each attribute type A, any of its values can be a fuzzy set.

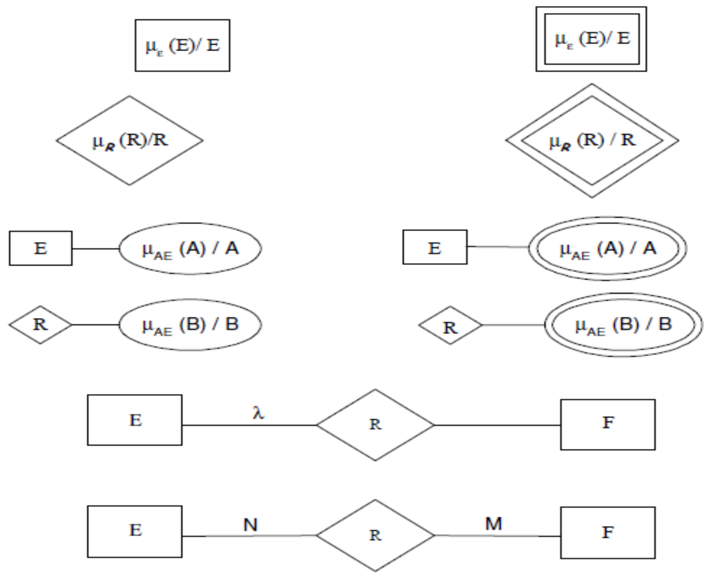

Figure 3: The fuzzy ER notation proposed by Chen and Kerre

6. The Chaudhry, Moyne, and Rundensteiner Approach

Chaudhry, Moyne, and Rundensteiner proposed a model for designing Fuzzy Relational Databases (FRDB) in line with the extension of the ER model of taking special interest in converting crisp databases into fuzzy ones. The way to achieve it is to define *n* linguistic labels as *n* fuzzy sets over the universe of an attribute. Then, each tuple in the crisp entity is changed to up-to *n* fuzzy tuples in a new entity (or *n* values in the same tuple). Each fuzzy tuple (or value) do not store the crisp value but a linguistic label and a grade of membership, giving the degree to which, the corresponding crisp entity belongs to the new entity, the crisp entity and the new fuzzy entity are mapped to separate tables.

Their ER model includes fuzzy relationships as relationships with at least one feature, namely, the membership grade. They proposed FERM, a design methodology for mapping a fuzzy ER data model to a crisp relational database in four steps (constructing a fuzzy ER data model, transforming it to relational tables, normalization, and ensuring correct interpretation of the fuzzy relational operators). They also present the application of FERM to build a prototype of a fuzzy database for a tactful control system for a semiconductor manufacturing process in an expanded model presented in their 1994 paper. Focusing on their proposal for the control processes, in each process, imprecise values are experimented and associated to linguistic labels, and every value involves a process called “DB Fuzzifier construct,”

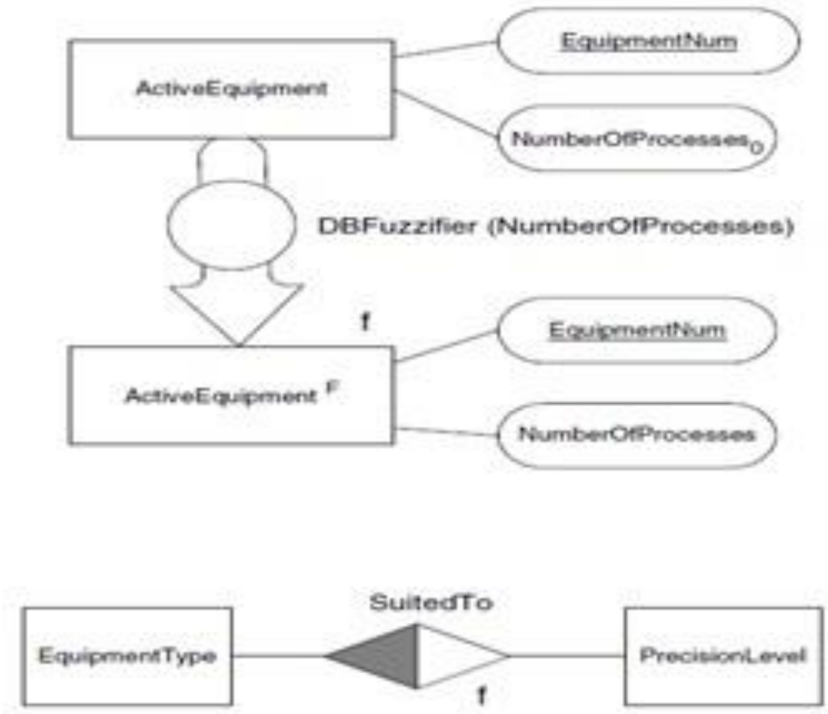

Figure 4: The model proposed by Chaudhry, Moyne, and Rundensteiner. An instance of DB Fuzzifier translational and Fuzzy Interrelation

7. Ma, Zhang and Chen Approach

These researchers proposed a model with three levels of the work of Zvieli and Chen and integrate into the Fuzzy Extended Entity-Relationship model (FEER model) a way of managing multifaceted objects in the real world at a conceptual level, associating an importance degree of each of the components (attributes, entities, etc.) to the scheme. However, their definitions of generalization, specialization, category, and aggregation impose very restrictive conditions. They also provide an approach to mapping an FEER model to a Fuzzy Object-Oriented Database scheme (FOODB). Figure 5 shows the following: single-valued attribute type, multi-valued attribute type, disjunctive fuzzy attribute type, conjunctive fuzzy attribute type, null attribute type, open or null attribute type, disjunctive imprecise attribute type, conjunctive imprecise attribute type, entity with grade of membership, relationship with grade of membership, and attribute with grade of membership. In the approach, the researchers introduce an extended object-oriented database model to handle imperfect as well as complex objects. They extend some major notions in object-oriented databases such as objects, classes, objects-classes relationships, subclass/super-class, and multiple inheritances.

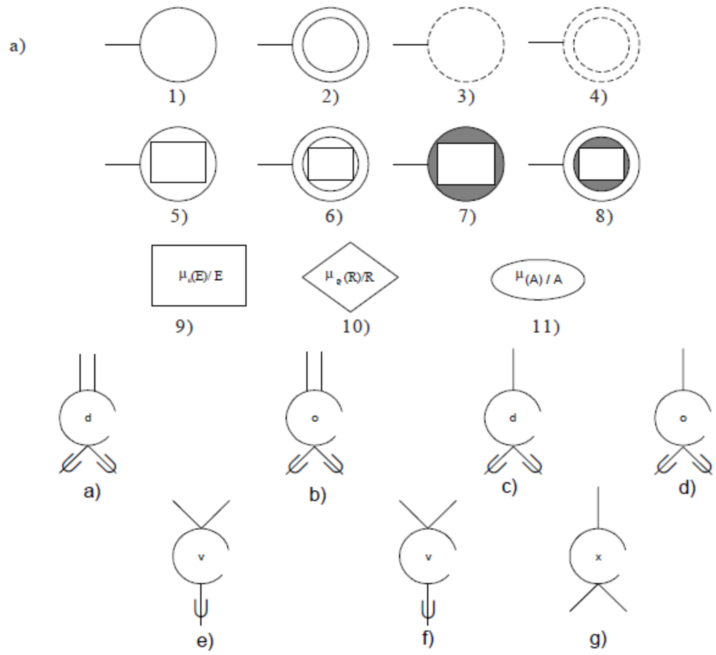

Figure 5: FEER notation by Ma, Zhang, fuzzy attributes, entities, and interrelations/specialization, aggregation and fuzzy categories

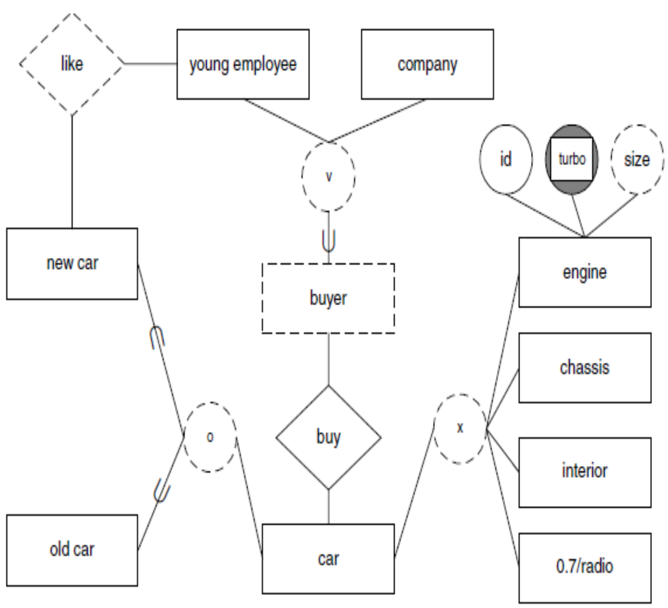

Figure 6: An instance of Ma, Zhang, Ma and Chen notation for a car assembly company case

8. Conclusion

The storage and retrieval of information in today's modern society has become a daily occurrence and it is applicable in virtually every facet of life. This has made the management of information to be very essential for effective and efficient service delivery to the users. The incorporation of logic in the management of database models has been a crucial area in database research because of the importance in data management especially in intensive applications where fuzzy data is indispensable. Over the years, fuzzy database model is mainly investigated with respect to the popular relational database model, However, classical relational database model and the extension of fuzziness do not meet the current need of modeling complex data with imprecision and uncertainty. Hence, there is need for further extension of models to be able to accommodate complex fuzzy data and meet contemporary challenges. This paper elaborates on the issue of the state of the art in fuzzy database modeling in which different approaches by researchers were discussed in detail.

References

1. Amiahi Motro and Philippe Smets (1997). uncertainty management in information systems, department of information and software systems engineering George Mason University, Fairfax, VA 22030, USA
2. Chaudhry, N., Moyne, J., & Rundensteiner, E. A design methodology for databases with uncertain data. In Proceedings of the Seventh International Working Conference on Scientific and Statistical Database Management, Charlottesville, VA pp. 32-41, 1994
3. Fujishiro, I. The design of a graph-oriented schema for the management of individualized fuzzy data. Japanese Journal of Fuzzy Theory and Systems, *3*(1): 1-14, 1991.
4. Umano, M., &Fukami, S. Fuzzy relational algebra for possibilitydistribution- fuzzy-relation model of fuzzy data. Journal of IntelligentInformation Systems, *3*, 7-28, 1994.
5. Vandenberghe, R. M. An extended entity-relationship model for fuzzy databases based on fuzzy truth values. In Proceedings of the Fourth International Fuzzy Systems Association World Congress, Brussels pp. 280-283, 1991.
6. Van Gyseghem, N., & de Caluwe, R. The UFO model: Dealing with imperfect information. In Fuzzy and uncertain object oriented databases, concepts and models pp: 123-185, 1997.
7. Vert, G., Morris, A., Stock, M., & Jankowski, P. Extending entity relationship modeling notation to manage fuzzy datasets. Proceedings of the Eighth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Spain pp: 1131-1138, 2000.
8. Yazici, A., Buckles, B. P., & Petry, F. E. handling complex and uncertain information in the ExIFO and NF2 data models. IEEE Transactions on Fuzzy Systems, pp: 659-675, 1999.

9. Yazici, A., & Cinar, A. Conceptual modeling for the design of fuzzy object-oriented database. In O. Pons, M. A. Vila, & J. Kacprzyk (Eds.), Knowledge management in fuzzy databases (pp. 12-35). Physica- Verlag, 2000.
10. Yazici, A., & Koyuncu, M. Fuzzy object-oriented database modeling coupled with fuzzy. Fuzzy Sets and Systems, 89: 1-26, 1997.
11. Yazici, A., & Merdan, O. Extending IFO data model for uncertain information. In Proceedings of the Fourth International Conferenceon Information Processing and Management of uncertainty (IPMU'96, Vol. III), Spain pp. 1283-1282, 1996.