International Journal of Electrical, Electronics and Computer Engineering **4**(1): 9-13(2015)
ISSN No. (Online): 2277-2626

Framework for Handling Uncertainty through Temporal Databases

Praveen Kumar Gupta*, Rahul Rishi**and Ranjit Biswas***

**School of Computer Science and Technology, Lingayas University, Faridabad, (HR) India*
***University Institute of Engg. & Tech. M.D. University, Rohtak,(HR) India*
****Department of Computer Science and Technology, Jamia Hamdard, Hamdard Nagar, (New Delhi), India*

(Corresponding author: Praveen Kumar Gupta)
(Received 04 November, 2014 Accepted 05 December, 2014)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: The present paper explores the possibility of modeling of uncertainty through temporal databases. Considering limitations of relational model and technological progress, temporal databases emerged. Adding to it, the uncertainty aspect for imitating human like thinking. Temporal models are discussed, leading to proposed conceptual temporal database framework.

Keywords: Framework, Functional Dependency, Operator, Temporal, Uncertainty

I. INTRODUCTION

The evolution of relational approach to databases is due to Codd's papers [5-7].Relational model did not address temporal dimension. As technology progressed with reference to computing and storage capabilities, existence of temporal databases became a distinct possibility.

A database which maintains past, present and future data is called a temporal database. Temporal data can be discrete or continuous. Assuming time as discrete data, for the sake of simplicity. For considering uncertainty through fuzzy data, bivalent logic is inadequate. This paper is organized in five sections. Section one is introduction. Section two describes handling uncertainty. Section three critically analyses about temporal data models. Section four describes the framework for modeling of fuzziness through temporal databases. Section five has the conclusion of the work reported here.

II. HANDLING UNCERTAINTY

The information size increased by leaps and bounds during the last century, leading to paradigmatic shift in computing and mathematics in the last century. This led to modern view of uncertainty where it is considered essential to computing, unavoidable and of great utility.

Max Black was the first to envision and consider vagueness aspect for logical analysis of data [4]. Later, Kreye [11] also worked on handling uncertainty. Theories such as probability theory, fuzzy theory, evidence theory, possibility theory etc. evolved to manage uncertainty. Evidence theory is based on dual non-additive measures: belief measures (Bel) and plausibility measures (Pl). Given a universal set X, assumed here to be finite, a belief measure is a function Bel : P(X)⟶[0,1] such that Bel(Ǿ)=0, Bel(X)=1. Associated with each Bel is a Pl, defined by equation Pl (A) = 1- Bel (A) where a set is represented by A.

Similarly, Bel (A) = 1 – Pl (A). Evidence obtained in the same context from two independent sources may be combined using Dempster's rule of combination. Possibility theory and Probability theory are recognized as special branches of Evidence theory. Possibility Theory is a special branch of Evidence Theory that deals with the bodies of evidence whose focal elements are nested. Special counterparts of belief measures and plausibility measures in possibility theory are called necessity measures and possibility measures. Since necessity measures are special belief measures and possibility measures are special plausibility measures and both measures constrain each other in a strong way. An important property of possibility theory is that every possibility measure is uniquely represented by the associated possibility distribution function. In this paper, primary focus is on one type of uncertainty, which is of interest to us and that is fuzziness and fuzzy theory that logically leads to fuzzy logic. Fuzzy logic evolved in 1965 with the publication of a paper [14] by Zadeh. The characteristic function of a crisp set assigns a value of either 1 or 0 to each individual in the universal set. This function can be generalized such that the values assigned to the elements of the universal set fall within a specified range and indicate the membership grade of these elements in the set in question. Such a function is called a membership function and the set defined by it a fuzzy set. The most commonly used range of values of membership functions is the unit interval [0, 1]. A fuzzy set can be defined mathematically by assigning to each possible individual in the universe of discourse a value representing its grade of membership in the fuzzy set. Let $U = \{u_1, u_2, \ldots\ldots\ldots\ldots u_n\}$ be a universe of discourse. A fuzzy set A in the universe of discourse U is characterized by the membership function μ_A given by $\mu_A : U \rightarrow [0, 1]$.

The membership function for a fuzzy set of U takes values from the closed interval [0, 1]. A fuzzy set A is defined as the set of ordered pairs A = {(u, $\mu_A(u)$) : u ∈ U} where $\mu_A(u)$ is the grade of membership of element u in the set A. The greater the amount of $\mu_A(u)$, the greater is the truth of the statement that 'the element u belongs to set A'. For example, consider a universe U = {DOG, CAT, RAT}. A fuzzy set A of U could be A = {(DOG, 0 .7); (CAT, 0 .99); (RAT, 0.4)}. Fuzzy set theory, fuzzy sets and fuzzy logic is generalization of classical set theory, crisp sets and boolean logic, respectively. The important concepts of α-cut and strong α-cut (of a fuzzy set A is the crisp set $^{\alpha}A$ that contains all the elements of the universal set X whose membership grades in A are greater than or equal to the specified value of α) are useful for defuzzification to get equivalent crisp sets. Fuzzy functional dependency for relational databases has been defined by Hamouz and Biswas [8]. It is defined as under:

Fuzzy Functional Dependency: Let X,Y R = {A_1 , A_2 ,..., A_n}. Choose a parameter α [0, 1] and propose a fuzzy tolerance relation named as Ŗ. A fuzzy functional dependency (ffd) is said to exist if, whenever $\mathbf{t_1[X]} \xrightarrow[(\propto)\text{Ŗ}]{} \mathbf{t_2[X]}$.

Alternatively, the fuzzy functional dependency can also be represented as $t_1[X]\ E_{(\alpha)} E\ t_2\ [X]$.

The set X of attributes fuzzy-functionally determines the set Y of attributes at α-level of choice. In another terminology, the set Y of attributes is fuzzy-functionally defined by the set X of attributes at α -level of choice. It is denoted by the notation $\mathbf{X} \xrightarrow[(\propto)\text{Ŗ}]{} \mathbf{Y}$ since Ŗ is already fixed but choices on α may be set by the database designer to vary during the course of analysis. Hamouz and Biswas emphasized more on the necessity to consider integrity constraints of fuzzy nature, in database design. Fuzzy logic has become an important consideration for modeling time varying applications also.

III. TEMPORAL DATA MODELS

In this section, different temporal data models are discussed. The key concepts and features are compared of some of the important models, based on temporal and non-temporal parameters. Ariav's model [1] used tuple time stamping with time being represented by discrete time points in temporal mode. The model is conceptually simplistic but difficult to implement in efficiency and reliability terms. Ben-Zvi's time relational model (TRM) [3] has non first normal form (NFNF), as an important concept. He gave the concept of effective time and registration time, which are now known as valid time (VT) and transaction time (TT), respectively. He was the first to coin the term and notion of time-invariant key for his non first normal tuples, called tuple version sets in his terminology. The NFNF definition uses contiguous and non-contiguous time relations. A valid time relation is in time normal form (TNF) if and only if it is in snapshot BCNF and there exists no temporal dependency among its time varying attributes. Ben-Zvi differentiated between an error and a change and made both of them queriable. He recognized the need for fast access to current data. Jensen & Snodgrass model [9][10] proposed bi-temporal conceptual data model (BCDM), allowing to associate both valid and transaction times with data. [13] The domains of valid and transaction times are the finite sets D_{VT} and D_{TT}, respectively. A valid time chrononc_v is a time point belonging to D_{VT} and a transaction time chrononc_t is a time point belonging to D_{TT}. A bitemporalchrononcb = (c_t ,c_v) is an ordered pair consisting of a transaction time chronon and a valid time chronon. The schema of a bitemporal relation R, defined on the set U = {A_1,A_2,...,A_n} of non-timestamp attributes, is of the form R = (A_1,A_2,...,A_n | T), that is, it consists of n non-timestamp attributes A_1,A_2,...,A_n, with domain dom(A_i) for each i [1,n], and an implicit timestamp attribute T. The domain of T is (D_{TT}{UC})× D_{VT} , where UC is a special value that can be assumed by a transaction time chronon to express the condition "until changed". For instance, to state that a tuple valid at time c_v is current in the database, the bitemporalchronon (UC, c_v) must be assigned to the tuple timestamp. As a general rule, they associate a set of bitemporalchronons in the two-dimensional space with every tuple. An example of temporal relational database is Time DB [15]. It uses the extension approach with respect to the data structures. Time DB uses a layered approach which means it was built as a front end to a commercial DBMS that translates temporal statements into standard SQL statements. This way, it is possible to support features such as persistence, consistency, concurrency, recovery etc. without having to implement from the scratch.

IV. FRAMEWORK FOR MODELING OF FUZZINESS THROUGH TEMPORAL DATABASES

In last 20 years, several proposals to deal with uncertainty or weighted data have been proposed. In an effort to Integrate uncertainty (fuzziness) with relational model, the concept of fuzzy functional dependencies has emerged and many definitions have been given in different contexts. In this paper, framework for temporal databases with fuzzy perspective, is proposed, as in Fig. 1. In the context of temporal databases, the concept of uncertainty in general and fuzzy logic in particular, is very important in the modeling process. This proposed framework intends to extend fuzzy functional dependencies already defined for relational databases to temporal databases domain. Once the equivalent crisp sets are obtained through defuzzification, the temporal dimension along with the temporal operator is applied and temporal fuzzy functional dependency can be defined to get the desired output through temporal databases. Also, such functional dependencies will be evaluated for their validity with reference to temporal-Armstrong axioms.

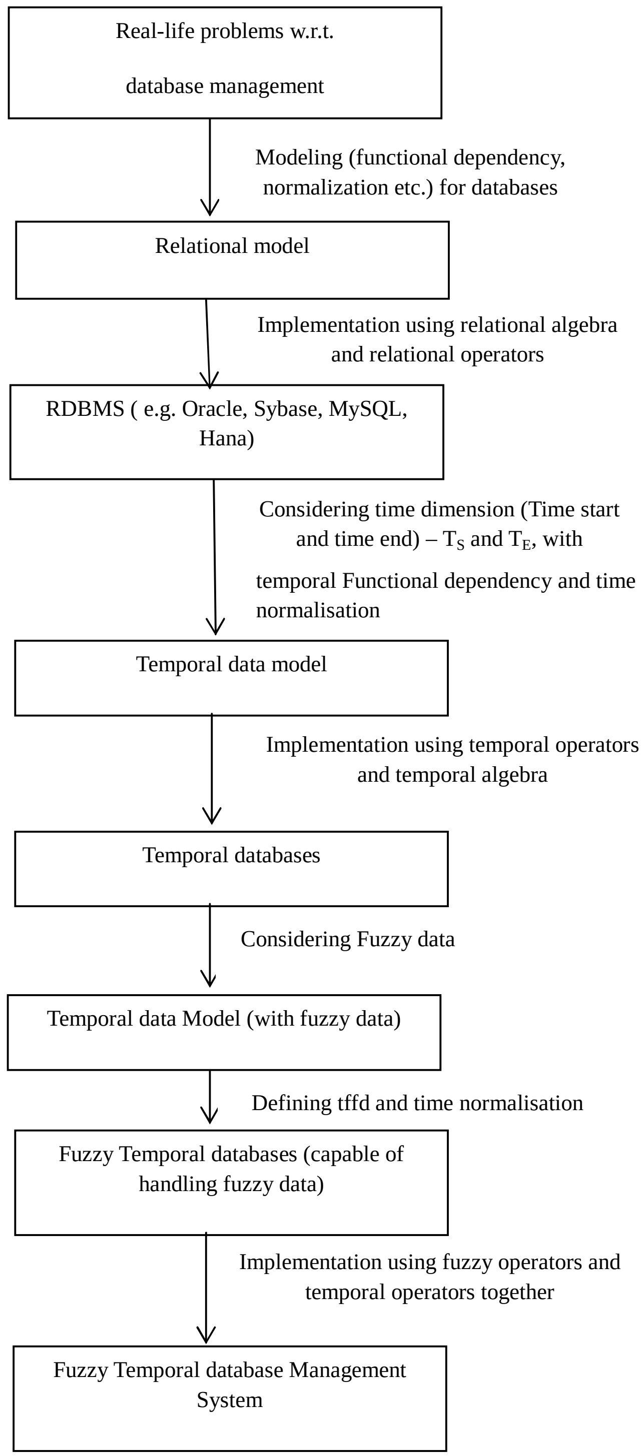

Fig. 1. Framework for modeling fuzziness through temporal databases.

The scope of discussion in this proposed framework is limited to only fuzziness (and excludes ambiguity). In case of tuples with fuzzy components, the membership function of the fuzzy subset is interpreted as a possibility distribution. The membership function is a graphical representation of the magnitude of participation of each input. It associates a weight with each of the inputs that are processed, define functional overlap between inputs, and ultimately determines an output response. The rules use the input membership values as weighting factors to determine their influence on the fuzzy output sets of final output conclusion. There are different membership functions associated with each input and output response [12]. Operation of a fuzzy system has 3 steps i.e. fuzzification, inference, and finally defuzzification.

Fuzzification: We consider fuzzy numbers and applying to them the concepts of fuzzy set theory. For each input and output variable selected, we define two or more membership functions(Fig.II) and is illustrated as under:

If we take x like a variable and poor, acceptable and good as triangular membership functions respectively (Fig. II), the membership function Poor is to be defined by three points (x_1, x_2, x_3), membership function. Acceptable is defined by three points (x_2, x_3, x_4) and membership function. Good is defined by (x_3, x_4, x_5) membership function. Assuming that all the membership functions are triangular,

$$y^{poor}{}_{(}x, x_1, x_2, x_{3)} = \max\left(min\left(\frac{x-x_1}{x_2-x_1}, \frac{x_3-x}{x_3-x_2}\right),\ 0\right)$$

$$y^{acceptable}{}_{(}x, x_2, x_3, x_{4)} = \max\left(min\left(\frac{x-x_2}{x_3-x_2}, \frac{x_4-x}{x_4-x_3}\right),\ 0\right)$$

$$y^{good}{}_{(}x, x_3, x_4, x_{5)} = \max\left(min\left(\frac{x-x_3}{x_4-x_3}, \frac{x_5-x}{x_5-x_4}\right),\ 0\right)$$

Any of the values will belong to at least one membership function with a certain degree of membership.

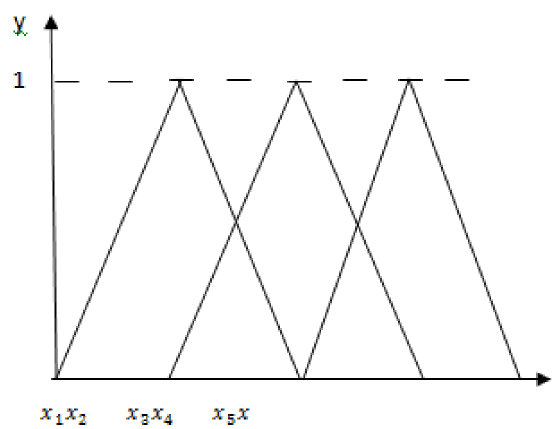

Fig. 2. Example of Membership Function (MF) for given input.

Inference and Defuzzification: In many situations, for a system whose output is fuzzy, it is easier to take a crisp decision if the output is represented as a single scalar quantity. This conversion of a fuzzy set to single crisp value is called defuzzification.

Once the functions are inferred, scaled, and combined, they are defuzzified into a crisp output which drives the system. There are many defuzzification methods e.g. center of area method, center of maxima method and mean of maxima method [12].

However, only ordinary fuzzy sets are being considered and such sets being operated by standard fuzzy operations (i.e. complement, t-norm and t-conorm operations) within this proposed framework. Any Fuzzy model application can be built up based on three phases mentioned above, with reference to the database parameters considered.

All temporal databases and its applications need to include the following aspects which are quite significant for future work i.e. – temporal SQL and query optimization, testing for consistency of databases and distributed computing, heterogeneous databases and data migration issues, improvements in data-mining through grid computing and cloud computing. Exception handling in temporal databases with fuzzy data (e.g. only vice chancellor of the university has the discretion to over-rule this condition of allowing to appear for exam, under special circumstances) will fall under the category of complex issues to be considered for future work.

V. CONCLUSION

In this paper, the constraints of relational databases were discussed and analyzed both in terms of temporal perspective and fuzzy aspects. Since the concept of integrity constraints is central to databases and to address the critical issues of handling uncertainty using fuzzy logic, it leads us to consider the need to propose the framework for handling uncertainty through temporal databases. Computing cost is quite high in temporal databases which can be significantly reduced if uncertainty is included using fuzzy logic within the design of temporal database framework. Key constraints and normalization in temporal databases need further consideration. The framework for modeling fuzziness through temporal databases will include defining temporal fuzzy functional dependencies in temporal databases which will help in consistent and valid database schema design and to make useful applications in real world, keeping in view the advantage of latest storage technologies and processing powers. Some of the potential areas of application of this proposed framework may be logistics sector like railways reservation system, financial sector like stock market operations and e-governance sector like citizen healthcare system.

The capability of fuzzy sets to express gradual transitions from membership to non-membership and vice-versa has a broad utility, especially when it is considered in temporal databases. The role of evidence theory in general and fuzzy set theory in particular may be quite useful in modeling of such highly complex systems with human like thinking.

REFERENCES

[1]. G. Ariav, "A Temporally Oriented Data Model", *ACM Transactions on Database Systems*, Volume **11**, No.4, 1986, pp. 499-527.

[2]. W.W. Armstrong, "Dependency structure of database relationships", *Proceedings of the IFIP 74*, 1974, pp. 580-583.

[3]. B. Zvi, "The Time Relational Model", *Ph. D. Thesis*, Computer Science Department, UCLA, 1982.

[4]. M. Black, "Vagueness: An exercise in logical analysis", *Philosophy of Science*, Volume **4**, 1937, pp. 427–455.

[5]. E.F. Codd, "A relational model of data for large shared data banks", *Communications of the ACM*, 1970, Vol. **13**, No. 6, pp. 377-387.

[6]. Codd, E. F. and Rustin R., "Further Normalisation of the Database relational model", *Database Systems*, Englewood Cliffs, NJ: Prentice- Hall, 1972, pp. 33-64.

[7]. E.F. Codd, "Recent investigation in relational database systems", *Proceedings of IFIP 74*, 1974, pp. 1017-1021.

[8]. S. Hamouz and R. Biswas, "Fuzzy functional dependencies in Relational Databases", *International Journal of Computer Cognition*, Vol. **4**, No. 1, 2006, pp. 39-43.

[9]. C.S. Jensen and R. T. Snodgrass, "Proposal for a Data Model for the Temporal Structured Query Language", Department of Computer Science, University of Arizona, *TEMPIS technical report number* **37**, 1992.

[10].C. S. Jensen and R.T. Snodgrass, "Temporal Data Management", *IEEE Transactions on Knowledge and Data Engineering*, Volume **11**, No. 1, 1999, pp. 36-44.

[11]. M.E. Kreye, Y.M. Goh and L.B. Newnes, "Manifestation of uncertainty- A classification", *Proceedings of International conference on Engineering design*, Technical University of Denmark, 2011.

[12]. S.S. Singh, R. Sayal and P.V. Rao, "Analysis and usage of fuzzy logic for database queries" *International journal of computer applications*, Volume **16**, No. 3, pp. 19-26.

[13]. R. T. Snodgrass, "The Temporal Query Language TQuel." *ACM Transactions on Database Systems*, Volume **12**, No. 2, 1987, pp. 247- 298.

[14]. L.A. Zadeh, "Fuzzy sets", *Information and Control*, Volume **8**, 1965, pp. 338-353.

[15]. Steiner A. "A time relational DBMS: Time DB", *Ph. D. Thesis*, 1998, Chapter 4, pages 53- 79.