

Neutrosophic Relational Database Decomposition

MeenaArora
CSE Department
JSS Academy of Technical
Education
Noida, INDIA

Ranjit Biswas
Deptt. Of Faculty of Engg. &
Technology
Manav Rachna University
Faridabad, INDIA

Dr. U.S.Pandey
Delhi University
Delhi,
INDIA

Abstract— In this paper we present a method of decomposing a neutrosophic database relation with Neutrosophic attributes into basic relational form. Our objective is capable of manipulating incomplete as well as inconsistent information. Fuzzy relation or vague relation can only handle incomplete information. Authors are taking the Neutrosophic Relational database [8], [2] to show how imprecise data can be handled in relational schema.

Keywords— *Neutrosophic Logic; Ranking of intervals; β -value of an interval; Neutrosophic set; Neutrosophic Relation; Rank Neutrosophic INF.*

I. INTRODUCTION

In the real world there are vaguely specified data values in many applications, such as sensor information, Robotics etc. Now take an example, when we ask the opinion of an expert about certain statement, he or she may say that the possibility that the statement is true is between 0.5 and 0.7, and the statement is false is between 0.2 and 0.4, and the degree that he or she is not sure is between 0.1 and 0.3. Here is another example, suppose there are 10 voters during a voting process. In time t_1 , three vote “yes”, two vote “no” and five are undecided, using neutrosophic notation, it can be expressed as $x(0.3,0.5,0.2)$; in time t_2 , three vote “yes”, two vote “no”, two give up and three are undecided, it then can be expressed as $x(0.3,0.3,0.2)$. That is beyond the scope of the intuitionistic fuzzy set. So, the notion of neutrosophic set is more general and overcomes the fore mentioned issues. In neutrosophic set, indeterminacy is quantified explicitly and truth membership, indeterminacy-membership and falsity membership are independent. This assumption is very important in many applications such as information fusion in which we try to combine the data from different sensors. Neutrosophy was introduced by Smarandache [7]. Neutrosophic set is a powerful general formal framework which generalizes the concept of the classic set, fuzzy set, vague set etc.

The normalization process takes a relational Schema through a series of test to check up whether it satisfies a certain normal form. Consider an instance of a relation schema. In real life situation, the data available are not always precise or crisp, rather it can be in any form like it can be in natural language, any imprecise data or you can say Neutrosophic data. Consequently if, at least one data is Neutrosophic, the relation schema cannot be called to be in proper relational form. The quest to manage imprecision's is equal to major driving force in the database community is the Ultimate cause for many research areas: data mining, semi

structured data, and schema matching, nearest neighbor. Processing probabilistic data is fundamentally more complex than other data models. Some previous approaches sidestepped complexity. For example, consider an attribute SALARY (in \$) of a relation schema EMPLOYEE. If a tuple value for this attribute SALARY is precise viz. 5000, then it is a single atomic (Indivisible) value. But if a tuple value is Neutrosophic viz. “Approximately 5000”, and then it cannot be called an atomic value. Due to the importance of the need for supporting uncertain data several researchers have addressed this problem. A wide body of work deals with fuzzy modeling of uncertain data [10]

In this paper a method to decompose such relational Schemas is suggested.

II. PRELIMINARIES

Relational data model was proposed by Ted Codd's pioneering paper [5]. Since then, relational database systems have been extensively studied and a lot of commercial relational database systems are currently available [3, 4, 6]. This data model usually takes care of only well-defined and unambiguous data. However, when we talk about the imprecise data or imperfect information, it will fail to answer. But our Lay users may or may not be aware of imprecision. In order to represent and manipulate various forms of incomplete information in relational databases, several extensions of the classical relational model have been proposed [13, 9, 11, 5, 14, 15]. The vague set and vague logic proposed by Gau and Behrer provide a requisite mathematical framework for dealing within complete and imprecise information.

Consequently, there is a genuine necessity for the different large size organizations, especially for the industries, companies having worldwide business, to develop such a system which should be able to answer the users queries posed in natural language, irrespective of the QLs and their grammar, without giving much botheration to the users. Most of these types of queries are not crisp in nature, and involve predicates with neutrosophic hedges (with concentration or dilation).

Thus, these types of queries are not strictly confined within the domains always. The corresponding predicates are not hard as in crisp predicates. Some predicates are soft and thus to answer a query a hard match is not always found from the databases by search, although the query is nice and very real, and should not be ignored or replaced according to the

business policy of the industry. To deal with uncertainties in searching match for such queries, fuzzy logic and rather vague logic [8] and Neutrosophic logic by Smarandache [7] will be the appropriate tool.

In this section, Author presents some preliminaries on the theory of Neutrosophic Logic and Neutrosophic sets (NS) which will be required for the progress of this paper.

A. Neutrosophic Logic

In the *Neutrosophic Logic* (which is a generalization of fuzzy logic, especially of intuitionistic fuzzy logic) every logical variable x is described by an ordered triple $x = (T, I, F)$, where T is the degree of truth, F is the degree of falsehood, and I the degree of indeterminacy (or neutrality, i.e. neither true nor false, but vague, unknown, imprecise), with T, I, F standard or non-standard subsets of the non-standard unit interval $]0, 1^+[$. In addition, these values may vary over time, space, hidden parameters, etc.

There is a genuine necessity to develop such a system which should be able to answer the users queries posed in natural language, without giving much botheration to the users.

Let A and B be two neutrosophic sets. One can say, by language abuse, that any element neutrosophically belongs to any set, due to the percentages of truth/indeterminacy/falsity involved, which varies between 0 and 1 or even less than 0 or greater than 1.

Thus: $x(50,20,30)$ belongs to A (which means, with a probability of 50% x is in A , with a probability of 30% x is not in A , and the rest is undecidable); or $y(0,0,100)$ belongs to A (which normally means y is not for sure in A); or $z(0,100,0)$ belongs to A (which means one does know absolutely nothing about z 's affiliation with A).

More general, $x((20-30), (40-45) \cup [50-51], \{20, 24, \text{ and } 28\})$ belongs to the set A , which means:

- With a probability in between 20-30% x is in A (one cannot find an exact approximate because of various sources used);
- with a probability of 20% or 24% or 28% x is not in A ;
- the indeterminacy related to the appurtenance of x to A is in between 40-45% or between 50-51% (limits included);

The subsets representing the appurtenance, indeterminacy, and falsity may overlap, and $n_{sup} = 30+51+28 > 100$ in this

case.

A logic in which each proposition is estimated to have the percentage of truth in a subset T , the percentage of indeterminacy in a subset I , and the percentage of falsity in a subset F , where T, I, F are defined below, is called *Neutrosophic Logic*. Constants: (T, I, F) truth-values, where T, I, F are standard or non-standard subsets of the nonstandard interval $]0, 1^+[$, where $n_{inf} = \inf T + \inf I + \inf F \geq 0$, and $n_{sup} = \sup T + \sup I + \sup F \leq 3^+$. Neutrosophic logic [7] was created by Florentin Smarandache (1995)

B. Neutrosophic Sets

Neutrosophic set is a powerful general formal framework which generalizes the concept of the classic set, fuzzy set [1], Vague set [12] etc. A neutrosophic set A defined on universe U . $x = x(T,I,F) \in A$ with T, I and F being the real standard or non-standard subsets of $]0, 1^+[$, T is the degree of truth-membership of A , I is the degree of indeterminacy membership of A and F is the degree of falsity-membership of A .

Let X be a space of points (objects), with a generic element in X denoted by x . A neutrosophic set A in X is characterized by a truth-membership function T_A , an indeterminacy-membership function I_A and a falsity-membership function F_A . $T_A(x), I_A(x)$ and $F_A(x)$ are real standard or non-standard subsets of $]0, 1^+[$. That is

$$T_A : X \rightarrow]0, 1^+[, \quad (1)$$

$$I_A : X \rightarrow]0, 1^+[, \quad (2)$$

$$F_A : X \rightarrow]0, 1^+[. \quad (3)$$

There is no restriction on the sum of $T_A(x), I_A(x)$ and $F_A(x)$ so $0 \leq \sup T_A(x) + \sup I_A(x) + \sup F_A(x) \leq 3^+$.

Definition1. (Complement) The complement of a neutrosophic set A is denoted by $c(A)$ and is defined by

$$T_{c(A)}(x) = \{1^+\} - T_A(x), \quad (4)$$

$$I_{c(A)}(x) = \{1^+\} - I_A(x), \quad (5)$$

$$F_{c(A)}(x) = \{1^+\} - F_A(x). \quad (6)$$

for all x in X .

Definition2. (Union) The union of two neutrosophic sets A and B is a neutrosophic set C , written as $C = A \cup B$, whose truth-membership, indeterminacy-membership and falsity-membership functions are related to those of A and B by

$$T_C(x) = T_A(x) + T_B(x) - T_A(x) \times T_B(x), \quad (7)$$

$$I_C(x) = I_A(x) + I_B(x) - I_A(x) \times I_B(x), \quad (8)$$

$$F_C(x) = F_A(x) + F_B(x) - F_A(x) \times F_B(x). \quad (9)$$

for all x in X .

Definition3. (Intersection) The intersection of two neutrosophic sets A and B is a neutrosophic set C , written as $C = A \cap B$, whose truth-membership, indeterminacy-membership and falsity-membership functions are related to those of A and B by

$$T_C(x) = T_A(x) \times T_B(x), \quad (10)$$

$$I_C(x) = I_A(x) \times I_B(x), \quad (11)$$

$$F_C(x) = F_A(x) \times F_B(x). \quad (12)$$

for all x in X .

C. Neutrosophic relation

In this section, we will define the Neutrosophic relation. A relation is basically set of attributes(columns) and tuples(records/rows) having atomic values i.e indivisible and consistent vales. A tuple in a neutrosophic relation is assigned a measure that will be referred to as the *truth* factor and also as the *false* factor. The interpretation of this measure is that we believe with confidence and doubt with confidence that the tuple is in the relation. The truth and false confidence factors for a tuple need not add to exactly 1. This allows for incompleteness and inconsistency to be represented.

III. CLASSICAL RELATIONAL MODEL APPROACH

A classical relational database [9] consists of a collection of relations. A relation is a table of values where each row represents a collection of related data values. In a table, each row is called a tuple, a column header is called an attribute and the table as a whole is called the relation. A relation schema $R(A_1, A_2, \dots, A_n)$ consists of a relation name R and list of attributes A_1, A_2, \dots, A_n . There are various restrictions on data in the form of constraints. Domain constraints specify that each value of an A_i must be an atomic value from the domain $\text{dom}(A_i)$ attribute. This includes restrictions on data types, on the range of values (if any), and on format of data.

IV. RANKING OF INTERVALS

Intervals are not ordered. Owing to this major weakness, there is no universal method of ranking a finite (or infinite) number of intervals. But in real life problems dealing with intervals we need to have some tactic to rank them in order to arrive at some conclusion. A method of ranking of intervals is used here in subsequent section.

Consider a decision maker (or any intelligent agent like a company manager, a factory supervisor, an intelligent robot, an intelligent network, etc.) who makes a pre-choice of a decision parameter $\beta \in [0, 1]$. The intervals are to be ranked once the decision-parameter β is fixed. But ranking may differ if the pre-choice β is renewed.

Definition 4 (β -value of an interval)

Let $J = [a; b]$ be an interval. The β -value of the interval J is a non-negative real number $J_\beta = (1 - \beta).a + \beta. b$.

Clearly, $0 \leq J_\beta \leq 1$, and $\beta = 0, J_\beta = a$ which signifies that the decision-maker is pessimistic, and also for $\beta = 1, J_\beta = b$ which signifies that the decision maker is optimistic. For $\beta = .5$ it is the arithmetic-mean to be chosen usually for a moderate decision.

Comparison of two or more intervals we will do here on the basis of β -values of them. If the value of β is renewed, the comparison results may change. The following definition will make it clear.

Definition 5(Comparing two intervals)

Let $J_1 = [a, b]$ and $J_2 = [c, d]$ be two intervals. Then for a chosen $\beta \in [0, 1]$, we define

- (i) $J_1 < J_2$, if $(J_1)_\beta < (J_2)_\beta$.
- (ii) $J_1 > J_2$, if $(J_1)_\beta > (J_2)_\beta$.
- (iii) $J_1 = J_2$, if $(J_1)_\beta = (J_2)_\beta$.

(Note: The intervals $J_1 = [a, b]$ and $J_2 = [a, b]$ are strictly equal. For the other cases of the equality " $J_1 = J_2$ ", a further internal ranking could be done on the basis of their range i.e., interval-length. If range is more, we impose that the corresponding interval is greater).

So, the relational table where the ranking is involved, we name it by Rank Neutrosophic Table and the corresponding Rank Neutrosophic Normalisation is defined in the next section.

V. RANK NEUTROSOPHIC – INF OR (RNF)

In this section we have presented a method of decomposing a relational schema with Neutrosophic attributes into basic relational form. This Method is called as Rank Neutrosophic-First Normal Form -1NF(RNF) a revision of First normal Form in Relational database.

ALGORITHM

Let us present Sequence of steps for Rank Neutrosophic normalization of relation schema into 1NF(RN) :-

- 1) Remove all the Neutrosophic-attributes from the relation.
- 2) For each Neutrosophic-attribute create one separate table with the following attributes:
 - a) All attributes in the primary key
 - b) MV (z) (membership value)
 - c) NMV (Z) (non-membership value)
- 3) For every precise value of the Neutrosophic attribute put $MV=1$ and $NMV=0$.

Thus, if there is m number of attributes in the relation schema then, after normalization there will be in total (m + 1) number of relations .In special case, when the hesitation or in deterministic parts are nil for every element of the universe of discourse the Neutrosophic number reduces to fuzzy number. In such cases, the attribute non membership value i.e. NMV (Z) will not be required in any reduced tables of 1NF.

The method of normalizing a relational schema (with Neutrosophic attributes) into 1NF is explained in this section. For the sake of simplicity, consider a relation schema R as given in Table I. with only one Neutrosophic attribute and all other three attributes being crisp. "Neutrosophic attribute" means that at least one attribute value in a relation instance is Neutrosophic.

TABLE I. RELATIONAL SCHEMA R

A_1	A_2	A_3	A_4
-------	-------	-------	-------

This relational schema R has four attributes of which say A_4 is the only Neutrosophic attribute. Consider a relation instance r of R as shown in Table II. :

TABLE II. RELATIONAL TABLE R

A ₁	A ₂	A ₃	A ₄
A ₁₁	A ₂₁	A ₃₁	A ₄₁
A ₁₂	A ₂₂	A ₃₂	A ₄₂
A ₁₃	A ₂₃	A ₃₃	A ₄₃
A ₁₄	A ₂₄	A ₃₄	A ₄₄

Suppose that A₂ is the primary key here, all the data are precise except \tilde{a} , which is an Neutrosophic number. Thus all the data except \tilde{a} is atomic. This is not in 1NF because of the non atomic data \tilde{a} .

The neutrosophic number is the Neutrosophic set of the set R of real numbers. The universe of discourse R is an infinite set. But, in this method of normalization we shall consider a finite universe of discourse, say X, whose cardinality is N.

Let us suppose that X: {x₁, x₂,... x_n} and the Neutrosophic number \tilde{a} , proposed by a database expert is a NS (Neutrosophic Set) given by:

$$\tilde{a} = \{ (x_i, \mu_i, V_i) : X_i \in X, I= 1, 2, 3, \dots, N\}.$$

Then the Table II can be replaced by the following table III :

TABLE III. THE RELATIONAL INSTANCE R

A ₁	A ₂	A ₃	A ₄
A ₁₁	A ₂₁	A ₃₁	A ₄₁
A ₁₂	A ₂₂	A ₃ ₂	{(X _i , μ _i , V _i), (X ₁ , μ ₁ , V ₁), ... (X _n , μ _n , V _N)}
A ₁₃	A ₂₃	A ₃₃	A ₄₃
A ₁₄	A ₂₄	A ₃₄	A ₄₄

Now remove all the Neutrosophic attributes (here A₄ only), from Table III. Replace Table III by the following two tables Table IV. and Table V. :

TABLE IV. THE RELATION R₁

FNAME	FID	YEARLY SALARY (in thousands of dollars)
Juhi	F001	4563
Manoj	F002	6789
Geeta	F003	Approximately 56
Arun	F004	8987

TABLE V. THE RELATION R₂

A ₂	A ₄	MV (A ₄)	NMV (A ₄)
A ₂₁	A ₄₁	1	0
A ₂₂	X ₁	μ ₁	V ₁
A ₂₂	X ₂	μ ₂	V ₂
A ₂₂	X ₃	μ ₃	V ₃
....
A ₂₂	X _n	μ _n	V _n
A ₂₃	A ₄₃	1	0
A ₂₄	A ₄₄	1	0

In Table V. we have all the attributes of the primary-key of r(here only one attribute A₂), the Neutrosophic attribute A₄ and two new attributes which are Membership _Value(A₄) or MV(A₄)and Non Membership _Value(A₄) or NMV(A₄). Corresponding to all precise values of A₄, the MV(A₄) value is put 1 and the NMV(A₄) value is 0.

Now we see that the relation schema is in 1NF. Such a method of normalization is called Neutrosophic normalization and the normal form is called Neutrosophic 1NF or 1NF (N).

VI. EXAMPLE WITH HYPOTHETICAL DATA

We study the method here by an example with hypothetical data.

Consider an example with hypothetical data. Consider a relation schema FACULTY as shown in Table VI. below whose primary key is FID and the attribute YEARLYSALARY is a Neutrosophic attribute.

TABLE VI. THE RELATION SCHEMA- FACULTY

FNAME	FID	YEARLY SALARY (in thousands of dollars)
-------	-----	---

Consider a relation instance of this relation schema given by the following Table VII.

TABLE VII. THE RELATIONAL TABLE - FACULTY

A ₁	A ₂	A ₃
A ₁₁	A ₂₁	A ₃₁
A ₁₂	A ₂₂	A ₃₂
A ₁₃	A ₂₃	A ₃₃
A ₁₄	A ₂₄	A ₃₄

VII. CONCLUSION

In this instance FNAME and FID are crisp attribute whereas YEARLY-SALARY is neutrosophic attribute, all the attribute values for FNAME are atomic; all the attribute values for the attribute FID are atomic. But all the attribute values for the attribute Yearly Salary are not atomic. The data “approximately 56” is an Neutrosophic number $\tilde{56}$. Suppose that for this relation, a database expert proposes the Neutrosophic number $\tilde{56}$ as an NS given by $\tilde{56} = \{(55,.8,.1),(56,.9,.03),(56.5,.7,.10)\}$.

Therefore Table VII could be replaced by the following Table VIII.

TABLE VIII. RELATIONAL TABLE WITH NEUTROSOPHIC ATTRIBUTES.

FNAME	FID	YEARLY SALARY (in thousand of dollars)
Juhi	F001	4563
Manoj	F002	6789
Geeta	F003	{(55,.8,.1),(56,.9,.03),(56.5,.7,.10)}
Arun	F004	8987

Now remove the Neutrosophic attribute YEARLY SALARY (YS) for this instance and divide it into two relations given as in Table IX and Table X.

TABLE IX. FACULTY-1 RELATION

FNAME	FID
Juhi	F001
Manoj	F002
Geeta	F003
Arun	F004

TABLE X. FACULTY-2 RELATION

FID	YS	MV(Y S)	NMV(YS)
F001	4563	1	0
F002	6789	1	0
F003	55	.8	.1
F003	56	.9	.03
F003	56.5	.7	.10
F004	8987	1	0

Clearly, it is now in 1NF, called by 1NF (RN). For FACULTY-1, the Primary Key is FID, but for the newly created FACULTY-2 the Primary Key is {FID, YS}.

In the above section we have presented a method of normalization of a relational schema with Neutrosophic attribute in 1NF (RN). In recent years neutrosophic algebraic structures have been investigated (see for instance [16], while the neutrosophic framework has found practical applications in a variety of different fields, such as relational database systems. We have implemented the method of normalization by an example given in section IV which proves that how the imprecise data can be handle in relational schema using First Normal Form of Rank Neutrosophic databases. Thereby we claim that the algorithm suggested in section V is totally a new concept which can easily handle the neutrosophic attributes of First normal Form.

REFERENCES

- [1] Atanassov, K., Intuitionistic Fuzzy Sets : Theory and Applications, Physica-Verlag (2000), New- York.1955.
- [2] Biswas,R ., Intuitionistic fuzzy relations Bull. Sous. Ens.Flous.Appl. (BUSEFA L) 70 ,1997,22-29.
- [3] Chiang D., Chow L. R. and Hsien N,“Fuzzy information in extended fuzzy relational databases”, Fuzzy Sets and Systems 92, pp.1-10.
- [4] Codd, E.F., A relational model of data for large shared data banks, /Communications of the ACM, Vol.13 (6), p.377-387, June 1970 .
- [5] E. F. Codd, Extending the Database Relational Model to Capture More Meaning, ACM Trans. Database Systems , 4(4):397-434,Dec. 1979.
- [6] Elmasri and Navathe, Fundamentals of Database Systems, Addison-Wesley,New York,third edition,2000.
- [7] F. Smarandache (2002a), A Unifying Field in Logics: Neutrosophic Logic, in Multiple-Valued Logic / An International Journal , Vol.8, No.3,385-438,2002.
- [8] Gau, W.L. and Buehrer, D. J., Vague sets, IEEE Transactions on Systems, Man and Cybernetics, Vol.23 (1993) 610-614
- [9] J. Biskup, A Foundation of Codd’sRelationalMaybe-operations, ACM Trans. Database Systems, 8, 4:608-636, Dec. 1983.
- [10] J. Galindo, A. Urrutia, and M. Piattini, “Fuzzy Databases: Modeling, Design, and Implementation”. Idea Group Publishing, 2006.
- [11] M. L. Brodie, J. Mylopoulous, and J. W.Schmidt, On the Development of Data Models, On Conceptual Modeling , 19-47, 1984
- [12] Meena Arora and Ranjit Biswas, Deployment of Neutrosophic technology to retrieve answers for queries posed in natural language, in 3rd International Conference on Computer Science and Information Technology ICCSIT 2010, IEEE catalog Number CFP1057E-art, volume No. 3, ISBN: 978-1-4244-5540-9, Pages : 435-439, 2010.
- [13] S. Parsons, Current Approaches to Handling Imperfect Information in Data and Knowledge Bases, IEEE Trans, Knowledge and Data Engineering , 3:353- 372, 1996.
- [14] W. Lipski, On Semantic Issues Connected with Incomplete Information Databases, ACM Trans. Database Systems, 4, 3:262-296, Sept.1979.
- [15] W. Lipski, On Databases with Incomplete Information, Journal of the Association for Computing Machinery , 28:41-70, 1981.
- W.B. Kandasamy, Smarandache Neutrosophic Algebraic Structures, Hexis, Phoenix, 2006.
- [16] W.B. Kandasamy, Smarandache Neutrosophic Algebraic Structures, Hexis, Phoenix, 2006.