ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology
Volume 1, Issue 6, August 2012

Normalization of Intuitionistic Fuzzy Rough Relational Databases

Chhaya Gangwal, R.N.Bhaumik

Abstract— **In this paper, a method of normalization of a relational schema with intuitionistic fuzzy rough attributes in first normal form is presented. The method is implemented by an example which shows that how the processing of imprecision and indiscernibility can be examined in relational schema using first normal form of intuitionistic fuzzy rough relational databases. Then, an algorithm of intuitionistic fuzzy rough normalization of relation schema into first normal form is constructed.**

Keywords— Intuitionistic fuzzy rough relational database, Normal forms.

1. Introduction & Preliminaries

Databases are known for the ability to store and update data in an efficient manner providing reliability and elimination of data redundancy. The relational database model has well established mechanisms built into the model for properly designing the database and maintaining data consistency. Constraints and data dependencies are used in database normalization to realize these goals and minimize such problems as update anomalies, thereby providing greater integrity maintenance.

Normalization, as first proposed by Codd [4], is the process of structuring data to minimize duplication and inconsistencies. It is a method of breaking down large tables into several smaller tables. In real life situation, the data available are not always precise or crisp, rather it can be in any form like any imprecise data or uncertain data. The imprecise, inconsistent and uncertain attributes in relational databases were discussed by many authors. In 2004, Alam et. al. [1] introduced the concept of normalization of intuitionistic fuzzy sets into first normal form and Beaubouef et.al. [3] Introduced the rough normalization in 2005. In this paper a method of normalization of a relational schema with intuitionistic fuzzy rough set in first normal form is presented.

In 1965, Zadeh [6] introduced the concept of fuzzy set. This set contains only a membership function lying between 0 and 1. But while collecting data many cases may be there where data are missing so intuitionistic fuzzy sets are required which consists of both membership value and non-membership value. Atanassov [2] introduced the concept of intuitionistic fuzzy set which is now known as IF set. Rough set theory introduced by Pawlak [5] in 1982, which is a technique for managing the uncertainty and imperfection, can analyze incomplete information effectively. Then the concept of IF rough set is introduced by coupling both IF sets and rough sets.

Manuscript received June, 2012.

Chhaya Gangwal, *PhD Scholar, Department of Mathematics, Tripura University, Suryamaninagar, Tripura-799022*

R.N.Bhaumik, *Emeritus Fellow (UGC) and Professor of Mathematics (Rtd.), Department of Mathematics, Tripura University, Suryamaninagar, Tripura-799022*

2. IF Rough Set

In this section we introduce the notion of IF rough set by combining both rough sets and IF sets. IF rough sets are the generalizations of fuzzy rough sets that give more information about uncertain or boundary region.

Definition 2.1 Let U be a universe and X, a rough set in U. An IF rough set A in U is characterized by a membership function $\mu_A : U \to [0,1]$ and a non-membership function $\upsilon_A : U \to [0,1]$ such that $\mu_A(\underline{R}\ X) = 1$, $\upsilon_A(\underline{R}\ X) = 0$
or $[\mu_A(x), \upsilon_A(x)] = [1,0]$ if $x \in (\underline{R}\ X)$ and $\mu_A(U - \overline{R}X) = 0$, $\upsilon_A(U - \overline{R}\ X) = 1$
or $[\mu_A(x), \upsilon_A(x)] = [0,1]$ if $x \in U - \overline{R}\ X$, $0 \le \mu_A(\overline{R}\ X - \underline{R}\ X) + \upsilon_A(\overline{R}\ X - \underline{R}\ X) \le 1$.

2.1.1 Example of IF Rough Sets

Let U = {Child, Pre-Teen, Teen, Youth, Teenager, Young-Adult, Adult, Senior, Elderly} be a universe. Let the equivalence relation R be defined as follows:

R^* = {[Child, Pre-Teen], [Teen, Youth, Teenager], [Young-Adult, Adult],[Senior, Elderly]}.

Let X = {Child, Pre-Teen, Youth, Young-Adult} be a subset of universe U.

We can define X in terms of its lower and upper approximations:

$\underline{R}X$ = {Child, Pre-Teen}, and

$\overline{R}X$ = {Child, Pre-Teen, Teen, Youth, Teenager, Young-Adult, Adult}.

The membership and non-membership functions $\mu_A : U \to [0,1]$ and $\upsilon_A : U \to [0,1]$ on a set A are defined as follows:

μ_A (Child) = 1, μ_A (Pre-Teen) = 1 and ν_A (Child) = 0, ν_A (Pre-Teen) = 0

μ_A (Young-Adult) = 0, μ_A (Adult) = 0, μ_A (Senior) = 0, μ_A (Elderly) = 0,

ν_A (Young -Adult) = 1, ν_A (Adult) = 1, ν_A (Senior) = 1, ν_A (Elderly) = 1,
μ_A (Teen) = 0.3, μ_A (Teenager) = 0.4, μ_A (Young) = 0.5,
ν_A (Teen) = 0.5, ν_A (Teenager) = 0.4, ν_A (Young) = 0.3.

Such a set A defined in U on rough set X is called IF rough set in which IF values are represented as [μ_A (Child), $1-\nu_A$ (Child)] = [1,0], [μ_A (Teenager), $1-\nu_A$ (Teenager)] = [0.4,0.3] etc.

3. IF Rough Relational Database Model

In the proposed model, a relation schema R consists a set of attributes $A_1, A_2, ..., A_n$. The domain of an attribute A_i is denoted by domain (A_i). A relation r of the relation schema R is thus a set of tuples r. A tuple t_i takes the form $(d_{i_1}, d_{i_2}, ..., d_{i_m}, d_{i_\mu}, d_{i_\nu})$ where d_{ij} is a domain value of a particular domain set D_j and $d_{i\mu} \in D_\mu$ where D_μ is the interval $I = [0,1]$, the domain for IF membership value, and D_ν is the interval $[0,1]$, the domain IF non-membership values. In the ordinary relational database $d_{ij} \in D_j$. In the IF rough relational database except for the membership and non-membership values $d_{ij} \subseteq D_j$ and d_{ij} is not restricted to be a singleton, $d_{ij} \neq \phi$. Let $P(D_i)$ denote any non-null member of the power set of D_i.

Definition 3.1: An IF rough relation R is a subset of the set cross product $P(D_1) \times P(D_2) \times ......P(D_m) \times D_\mu \times D_\nu$, where D_μ and D_ν are the domain for IF membership and non-membership value respectively.

Definition3.2: Tuples $t_i = (d_{i1}, d_{i2}, ..., d_{in}, d_{i\mu}, d_{i\nu})$ and $t_k = (d_{k1}, d_{k2}, ..., d_{kn}, d_{k\mu}, d_{k\nu})$ are redundant if $[d_{ij}] = [d_{kj}]$ for all $j = 1, ...,..n.$

4. IF Rough First Normal Form

In this section the method of normalizing a relational schema into First Normal Form of IF rough relational databases is presented. First Normal Form of IF rough relational databases means the intersection of tuples and attributes contains one and only one value.

4.1 Algorithm

Here, the sequence of steps for IF rough normalization of relation schema, mainly on first normal form in IF rough set is presented as follows:-

1) Remove all the indiscernible attributes and repeating attributes from the relation.
2) Entering appropriate data in the empty tuples. And placing repeating data along with a
   copy of the attribute in a relation
3) For each indiscernible attribute create one separate table with the following attributes:
a) All attributes in the primary key
b) Put membership value (MV) of the attributes
c) Put non-membership value (NMV) of the attributes
4) For every precise value of the attribute, put MV = 1 and NMV= 0.

If there is m number of attributes in the relation schema, then after normalization, there will be $(m+1)$ number of relations. Now, the method of IF rough normalization of relation schema into first normal form is explained. A relation schema R with indiscernibility relation in attribute domain and all other crisp attributes are considered. By "indiscernibility relation" means that at least one attribute value in a relation instance is indiscernible.

Table I: Relation schema R

A_1	A_2	A_3	A_4	A_5	A_6

This relational schema R has six attributes of which A_2 is the only attribute with indiscernibility relation. Consider a relation instance r of R given by:

Table II: Relational Table R

A_1	A_2	A_3	A_4	A_5	A_6
a_{11}	a_{21}	a_{13}	a_{14}	a_{15}	a_{16}
a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}
a_{31}	a_{23}	a_{33}	a_{34}	a_{35}	a_{36}

Suppose that A_1 is the primary key here, all the data are precise except a_{22} which is in indiscernibility relation. Thus all the data except a_{22} is atomic. This is not in first normal form because of the non-atomic data.

Let us suppose that $X = \{x_1, x_2, ..., x_n\}$ be the finite universe of discourse and indiscernible attributes a_{22} is given by $a_{22} = \{(x_i, \mu_i, \nu_i) : x_i \in X_i, i = 1,2,....,n\}$.

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

Table III: The Relation Instance r

A_1	A_2	A_3	A_4	A_5	A_6
a_{11}	a_{21}	a_{13}	a_{14}	a_{15}	a_{16}
a_{21}	$\{(x_1, \mu_1, \nu_1), (x_2, \mu_2, \nu_2),.. .., (x_n, \mu_n, \nu_n)\}$	a_{23}	a_{24}	a_{25}	a_{26}
a_{31}	a_{23}	a_{33}	a_{34}	a_{35}	a_{36}

Now remove all the indiscernibility relation from Table III. Replace Table III by the following two tables:

Table IV: The Relation r_1

A_1	A_3	A_4	A_5	A_6
a_{11}	a_{13}	a_{14}	A_{15}	a_{16}
a_{21}	a_{23}	a_{24}	A_{25}	a_{26}
a_{31}	a_{33}	a_{34}	A_{35}	a_{36}

In Table V, we have all the attributes of the primary key of r, the indiscernible attribute A_2 and two new attributes are membership value (A_2) and non-membership value (A_2). Corresponding to all precise value of A_2, the membership value is put 1 and non-membership value is 0.

Table V: The Relation r_2

A_1	A_2	$MV(A_2)$	$NMV(A_2)$
a_{11}	a_{21}	1	0
a_{21}	x_1	μ_1	ν_1
a_{21}	x_2	μ_2	ν_2
…..	…..	……	……
a_{21}	x_n	μ_n	ν_n
A_{31}	a_{23}	1	0

So that the relation schema is in 1NF. Such a method of normalization is called IF rough normalization and the normal form is called IF rough 1NF.

4.2 Illustration

In this section we take a relation schema Patient. This relational schema has an attribute set like {Patient identification number (PID#), Patient name (PNm), Age, Telephone number (Tel), Patient address (Addr.), Department (Dept), Visit date (VisDt)}. Here the Patient ID represents the attributes' role as a primary key and the attributes Age is indiscernible attributes.

Table VI: The Relation Schema Patient

PID#	PNm	Age	Tel	Addr.	Dept	VisDt.

A relation instance r of the relational schema R as shown in TableVII:

Table VII: The Relational Table- Patient

PID#	PNm	Age	Tel	Addr.	Dept	VisDt.
P001	Paul	Adult	222….	Durga Nagar	Ortho	08.03.05
P002	Jadu	Senior	222….	Indra Nagar	Medicine	05.04.06
P003	Rohit	Adult	222….	Krishna Nagar	Eye	04.02.06
P004	Bob	{Adult, Teen}	222….	Shamali Bazar	Surgery	01.04.05
P005	Tom	{Adult, Senior}	222….	Rani Bazar	ENT	01.07.05
P006	John	Adult	222….	Krishna nagar	Medicine	.2.05.06

In this relational instance Patient name, Patient ID, Telephone number, Addr., Department and Visit date are crisp attributes Here, the indiscernibility lies in the Age attribute since there are some patients who are not able to tell the exact Age. So that all the attributes values for PID#, PNm, Tel, Addr., Dept and VisDt. are atomic but all the attribute values for the attribute Age and are not atomic. The data {Adult, Teen} and {Adult, Senior} of Age Attributes are indiscernible. So that to handle impreciseness and indiscernibility for this relation, we propose IF rough set theory.

Table VIII: The Relational Instance r

PID#	PNm	Age	Addr.	Tel	Dept	VisDt.
P001	Paul	Adult	Durga Nagar	222….	Ortho	08.03.05
P002	Jadu	Senior	Indra Nagar	222….	Medicine	05.04.06
P003	Rohit	Adult	Krishna Nagar	222….	Eye	04.02.06
P004	Bob	{(Adult,0.8,0.1), (Teen,0.7,0.2)}	Shamali Bazar	222….	Surgery	01.04.05
P005	Tom	{(Adult,0.7,0.1),(Senior,0.9,.03)}	Rani Bazar	222….	ENT	01.07.05
P006	John	Adult	Krishna Nagar	222….	Medicine	.2.05.06

In order to find IF rough normalization of relation schema into first normal form (1NF) we need the following steps:

Step 1: Create Patient-1 Relation: In this step, first we will remove the indiscernible attributes from the Relational Table and create a Patient-1 Relation. For Patient-1 Relation, the Primary Key is PID#.

Table IX: Patient-1 Relation

PID#	PNm	Addr.	Tel	Dept	VisDt.
P001	Paul	Durga Nagar	222….	Ortho	08.03.05
P002	Jadu	Indra Nagar	222….	Medicine	05.04.06
P003	Rohit	Krishna Nagar	222….	Eye	04.02.06
P004	Bob	Shamali Bazar	222….	Surgery	01.04.05
P005	Tom	Rani Bazar	222….	ENT	01.07.05
P006	John	Krishna Nagar	222….	Medicine	.2.05.06

Step 2: Create Patient-2 Relation: Now we shall create Patient-2 Relation. For this newly created relation the Primary

Key is {PID#, Age}. In this relation, imprecision and uncertainty can be handle through indiscernibility and IF membership and non-membership value. For every precise value of the attribute Age, we shall put membership value = 1 and non-membership value = 0.

Table X: Patient-2 Relation

PID#	Age	MV	NMV
P001	Adult	1	0
P002	Senior	1	0
P003	Adult	1	0
P004	Adult	0.8	0.1
P004	Teen	0.7	0.2
P005	Adult	0.7	0.1
P005	Senior	0.9	0.03
P006	Adult	1	0

Clearly, Table IX and Table X now in first normal form called by 1NF in IF rough set theory.

5. Conclusion

In this paper we have presented a method of normalization of a relational schema with intuitionistic fuzzy rough in first normal form. We have implemented the method by an example which proves that how the imprecise and indiscernible data can be handle in relational schema using First Normal Form of intuitionistic fuzzy rough databases. The notion of indiscernibility from rough set theory coupled with the membership and non-membership values from intuitionistic fuzzy set to represent uncertain information in a manner that maintains the degree of uncertainty for each tuples of the database is also presented.

References:

1. Alam, M.A. et.al., “Normalization of intuitionistic fuzzy relational databases”, NIFS 10(1), 1-6, 2004.

2. Atanassov, K., “Intuitionistic fuzzy sets”, Fuzzy Sets and Systems, 20, 87–96, 1986.

3. Beaubouef, T. et.al., “Normalization in a Rough Relational Database,” Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing Lecture Notes in Computer Science, 3641/2005, 275-282, 2005.

4. Codd, F. E, “A relational model of data for large shared data banks”, Comm. ACM., 13, 377-387, 1970.

5. Pawlak, Z., “Rough sets”, Internat. J. Comput. Inform. Sci., 11, 341-356, 1982.

6. Zadeh, L.A., “Fuzzy sets”, Information Control, l8, 338-353, 1965.