Intuitionistic Fuzzy Approach to handle Imprecise Humanistic Queries in Databases

Ashu
Assistant Professor
Department of Mathematics, DAV College Jalandhar, 144008, India

ABSTRACT

L.A Zadeh introduced fuzzy set theory in 1965. Many researchers took interest in the theory and evolved many generalizations of the concept. One of them was the Intuitionistic fuzzy set introduced by K.T.Atanassov in 1986. Since its inception, the literature on IFS Theory and its application has been rapidly amounting by now to several thousands of papers. In this paper, IF approach is being applied to locate the desired person from huge database of salesmen in the company. This search is being done keeping in view the imprecise information given by the concerned person. A sample database is given and desired person is found using IFRDB model. The main stress in this paper is on similarity measures to match the query language with that of the database. The important thing is that besides considering M-Ship degree, matching of non M-Ship degree with query language has also been explained.

Keywords

Intuitionistic fuzzy sets, Intuitionistic fuzzy relational database, vague query, IFSQL

1. INTRODUCTION

Intuitionistic fuzzy set (IFS) developed by Atanassov[1-4] in 1986, as an extension of L.A.Zadeh's concept of fuzzy sets [26] given in 1965. IFS theory has a great significance as it introduced another degree of freedom into a set description. IFS based models prove to be extremely useful mainly in the situations where we face human opinions involving answers of three types: yes, no, can't say. This last type of answer leads to vagueness and IFS is a strong way to deal with vagueness. A prominent feature of IFS is that it assigns to each element, a membership (M_ship) degree and a non membership (NM_ship) degree, and thus, the IFS overpowers the concept of fuzzy set which only assigns , to each element , a M_ship degree. There are so many applications of IFS to artificial intelligence-IF expert system, IF neural networks, IF decision making, IF machine learning, IF semantic representations etc.

In addition to these, IF relational database (IFRDB) is also an upcoming approach. IFDB was introduced by Biswas,De and Roy [11]. An important feature of IF RDBMS is the processing of queries with IF predicates,e.g. predicates which correspond to natural vague terms like high, very high, medium,etc.These predicates are evaluated with IF values, which reflect on the degrees of M_ship & NM_ship of the rows in the query result which is in fact , an IFR. Although SQL is a very powerful tool, it is unable to satisfy needs for data selection based on linguistic expressions and degrees of truth. In this study, sample of a huge database is considered and IF approach is applied to select the person from the database, keeping in view, vague description.

Rest of the paper is organized in the following way. Section 2 presents preliminaries of Intuitionistic fuzzy sets and Intuitionistic fuzzy relations. Section 3 presents the related work done in the field of IFRDB; Section 4 gives the details of sample database and the query. In this section the Intuitionistic fuzzy relations are also described with the help of tables. Section 5 gives the result of the query. Finally section 6 concludes the work done.

2. BASIC DEFINITIONS [25]

2.1. IFS

An Intuitionistic fuzzy set (in short IFS) A of a non-empty set X is an object having the form A = {(x, μ_A (x), v_A(x)) : xϵX} where the functions μ_A from X into I and v_A from X into I , denote respectively the degree of membership (namely μ_A (x)) and the degree of non-membership (namely v_A(x)) of each element x ϵ X in the fuzzy set A, and $0 \le \mu_A$ (x)+ v_A(x) $\le$ 1 for all x ϵ X.

2.2. Intuitionistic fuzzy relation

Let X and Y are two sets. An Intuitionistic fuzzy relation (IFR) R from X to Y is an IFS of X × Y characterized by the membership function μ_R and non-membership function o_r. An IFR R from X to Y will be denoted by R(X →Y).
R= {< (x, y), o_r (x, y), v_R (x, y) > x$\in$ X, y$\in$Y}
Where u_r : X × Y → [0,1] and v_R(x) : X ×Y → [0,1] satisfies the condition
$0 \le \mu_R$ (x,y)+ v_R (x,y) $\le$ 1 for every (x,y) $\in$ X ×Y

2.3. Intuitionistic fuzzy Tolerance Relation

An intuitionistic fuzzy relation **R** on the Cartesian product (X × X), is called an Intuitionistic fuzzy tolerance relation if R is reflexive and symmetric.
1) Reflexive
An IF relation R (X x Y) is said to be reflexive if x1, x2 $\in$X, μ_R (x, x) = 1
2) Symmetric
An IF relation R (X x Y) is said to be symmetric if x1, x2 $\in$X
μ_R (x1, x2) = μ_R (x2, x1) and v_R (x1, x2) = v_R (x2, x1)

2.4 Intuitionistic Fuzzy Database

As an Intuitionistic fuzzy set is a generalization of fuzzy set, Buckles and Petry defined fuzzy database [7] as a generalization of classical database. Their model is based on similarity relation for each domain of fuzzy database. A fuzzy relational database is defined as set of relations where each relation is a set of tuple. If it represents the i-th tuple it has the form (di1, di2, ………. dim). In a classical relational database each component d_{ij}, of the tuple is an element of the corresponding scalar domain Dj i.e. d_{ij} belongs to Dj. But in case of fuzzy relational database, the

element of the tuple consists of either singleton or crisp subset of the scalar domain.

1) Definition 1

An Intuitionistic fuzzy database is a set of relation where each pair of such relation R is a subset of the cross product: $2D1 \times 2D2 \times ....2Dm$ Where $2Di = P(Di) - \emptyset$ and $P(Di)$ is the power set of Di, here R is called the Intuitionistic fuzzy database relation.

2) Definition 2

Let $R \subseteq 2D1 \times 2D2 \times ....2Dm$ be an Intuitionistic fuzzy database relation. An Intuitionistic fuzzy tuple (with respect to R) is an element of R. Let $ti= (di1,di2,…,dim)$ be an intuitionistic fuzzy tuple. An interpretation of t, is a tuple = $(a_1, a_2,…,a_m)$where $a_j \in d_{ij}$ for each domain Dj.For each domain Dj if Tj be an intuitionistic fuzzy tolerance relation then the membership function is given by: $\mu_{Tj} : Dj \times Dj \rightarrow [0 ,1]$ And the non membership function is given by $v_{Tj}; Dj \times Dj \rightarrow [0 ,1]$ satisfying Atanassov's condition.

2. RELATED WORK

Atanassov introduced the concept of IFS in 1986 [1-4] Later on Intuitionistic fuzzy Logic came into existence as Atanassov & Gargov worked together in 1990. R.Biswas then entered into the field with IFR.Then Supriya K.De,Biswas & Roy [11-14] introduced IFDB following the notion of fuzzy representation of data for relational databases given by Buckles & Petry[7] .Later on Burillo & Bustince studied the properties of IF relation[8-10]. After the origin of concept, many applications came into existence. Some of them are given by E.Szmidt & J.Kacpryzck [16-17], Yaser Ahmed et.al. [25], Athar Kharal [5], etc. Many researchers worked on other aspects of IFRDB i.e. Querying [18-21].A prerequisite for the applicability of the approximate reasoning schemes is the availability of suitable IFSs for the linguistic statements involved. Constructing them is typically one of the most difficult tasks when developing an application. Therefore it is very useful to have representations of linguistic modifiers such as very and more or less at hand. These representations help in framing relations and thus concluding the required result. The concept has been used by Khatibi[23].

4. APPLICATION

Here an example is being considered to make the concept clear. Suppose a company has a large database of employees. The company has a wide distribution area and so among the list of employees there is a large number of salesmen covering various areas of state. A sample database looks like table given below:

Table 1.Employee Database

EMPLOYYEE ID	EMPLOYEE LAST NAME	POST	BASIC PAY
E1001	Stern	SALESMAN	$3000
E1002	Gounot	SALES MANAGER	$10000
E1003	Haas	PRO	$6000
E1004	Kwan	RECEPTION	$3000
E1005	Lucchessi	AGENT	$3000
E1006	Lee	COMPANY REP	$7000
E1007	Adamson	PRODUCTION INCHARGE	$10000
E1008	Thompson	CLERK	$2000

Table2. Employee Identification Database

EMPLOYYEE ID	Build	Complexion	Hair texture
E1001	Heavy	Fair	Dull
E1002	Slim	Dark	Shiny
E1003	Heavy	Very Fair	Shiny
E1004	Average	Fair	Shiny
E1005	Slim	Fair	Dull
E1006	Average	Very Fair	Average
E1007	Heavy	Average	Average
E1008	Slim	Average	Dull

Suppose on February 16, the company receives a complaint about the misbehaviour of a sales man. It was found that the complaint was put up from the region Quebec and it refers to the date 1/23/2011.The complainant was not known of the name of the sales person. So when she was asked about the sales person's appearance, she put forth a vague data. She explained the culprit person as "average body build very fair complexioned and average hair texture". Now the company management has to trace out the man and to punish him accordingly. This decision has to be made using the linguistic data which is possible only with the help of Intuitionistic fuzzy tolerance relation. As we know about human imprecision of the facts, so we do take care of imprecision in this vague information. e.g. if the complainant says that the sales man was of average body build ,it may be possible that she is 80% right in her judgement and 10% wrong with remaining 10% as can't guess. So judging her perceptions about body build, one can find that the similarity between average and heavy body build is with a degree of (0.8, 0.1) and hesitation is 0.1. Similarly the similarity between very fair and fair is (0.8, 0.1) with hesitation as 0.1, etc. In this way the linguistic expression given by that lady was translated in tables following IF relations and then an SQL were

generated to locate the desired person from large database given below.

Table 3 Sales Database

Sales_date (date)	Sales_person Last name (varchar(15))	Region (varchar(15))	No.of sales (int)
1/12/2012	Lee	Ontario North	4
1/12/2012	Lee	Manitoba	1
1/23/2012	Gounot	Ontario South	2
1/23/2012	Stern	Quebec	4
1/23/2012	Lee	Quebec	5
1/23/2012	Gounot	Quebec	6
1/23/2012	Adamson	Manitoba	1
1/23/2012	Lee	Ontario North	9
1/29/2012	Stern	Quebec	2
2/16/2012	Lee	Manitoba	4
2/16/2012	Stern	Quebec	2

We can finally make intuitionistic fuzzy tolerance relations based on the description provided by her .

The domain of complexion (D1) = {very fair, fair, average, dark, very dark}.

The domain of body build (D2) = {heavy, average, and slim}.

The domain of hair texture (D3) = {dull, average, shiny}

Table 4 IF tolerance relation T_{D1}

T_{D1}	v.fair	Fair	Average	dark	v.dark
v.fair	(1,0)	(.8,.1)	(.6,.4)	(.2,.8)	(0,1)
Fair	(.8,.1)	(1,0)	(.8,.1)	(.2,.8)	(0,1)
average	(.6,.4)	(.8,.1)	(1,0)	(.8,.1)	(.2,.6)
Dark	(.2,.8)	(.2,.8)	(.8,.1)	(1,0)	(.7,.2)
v.dark	(0,1)	(0,1)	(.2,.6)	(.7,.2)	(1,0)

Table 5 IF tolerance relation T_{D2}

T_{D2}	HEAVY	AVERAGE	SLIM
HEAVY	(1,0)	(.7,.2)	(0,1)
AVERAGE	(.7,.2)	(1,0)	(.4,.4)
SLIM	(0,1)	(.4,.4)	(1,0)

Table 6 IF tolerance relation T_{D3}

T_{D3}	DULL	AVERAGE	SHINY
DULL	(1,0)	(.7,.2)	(0,1)
AVERAGE	(.7,.2)	(1,0)	(.3,.6)
SHINY	(0,1)	(.3,.6)	(1,0)

Now after the framing of tables the next task is to find the salesman which is sought for misbehaviour as described earlier by the complainant in these words: average body build very fair complexioned and average hair texture". We can translate this IF query into relational algebra form:

(Project (select (EMPLOYEES DATA)

Where BUILD=AVERAGE'

COMPLEXION='VERY FAIR'

HAIR TEXTURE='AVERAGE'

With LEVEL (BUILD) =.7

LEVEL (COMPLEXION) =.9

LEVEL (HAIR TEXTURE) =.7

With LEVEL (NAME) =0.0

LEVEL (BUILD) =.7

LEVEL (COMPLEXION) =.9

LEVEL (HAIR TEXTURE) =.7

Giving LIKELY SOUGHT AFTER SALESMAN)

5. RESULT

As per level (say α)defined in query we will check for $\mu \geq \alpha$ & $v \leq 1$-α and in that way the above SQL results in two probable cases: Employee No.E1004 or E1006 among which E1006 has more attributes tallying with given vague data. So confirming about the said date and sales in Quebec in sales data, Lee was found the culprit.

6. CONCLUSION

As the vague human perceptions were translated into mathematical tables using M_ship & NM_ship functions, the large database gave satisfactory result to the vague query. This gives importance to IF database. Actually the already existing expert systems can answer for mathematical queries only but this new approach is more users friendly and has a bright scope ahead.

7. ACKNOWLEDGMENT

8. REFERENCES

[1] Atanassov K.T: On the temporal IFS theory Proc. Of 9th international conf.IPMU 2002,Vol.iii Annecy France,July 1-5,2002,1833-1837

[2] Atanassov K.T.: Open Problems in IFS theory, Proc. Of 6th joint conf. on inf. Sciences, Research Triangle Park(North Carolina,USA)March 8-13,2002,113-116

[3] Atanassov K.T.; IFS Past, Present & Future, CLBME-Bulgarian academy of sciences

[4] Atanassov K.T.; IFS , Theory and applications,Physica Verlag-Studies in fuzziness & soft computing ISBN 978-3-7908-2463-6

[5] Athar Kharal ; Homeopathic drug selection using IFS:www.Icidistributors.com

[6] Binyamin Yussof,et.al., A new similarity measures on IFS, World Academy of Sciences , Engg. & Technology 78 2011.

[7] B.P Buckles, F.E Petry "A Fuzzy Model for Relational Databases" in Fuzzy Sets and Systems, 7:213-226, 1982.

[8] P. Burillo, H. Bustince, "Intuitionistic fuzzy relations (Part I)", Mathware Soft Computing. 2 (1995) 5–38.

[9] Burillo, P. and Bustince, H., "Intuitionistic fuzzy relations (Part I)", Mathware Soft Computing. 2 (1995) 5–38.

[10] Bustince, H. and Burillo, P., "Structures on Intuitionistic fuzzy relations",Fuzzy Sets and Systems, 78 (1996) 293–303.

[11] De, S.K., Biswas, R. and Roy, A.R., "On Intuitionistic fuzzy Databases, Proceedings of the second International Conference on IFS", Sofia, 3 -4 October (1998) NIFS (4) 34-356.

[12] De, S.K., Biwas, R. and Roy, A.R., "An application of Intuitionistic fuzzy sets in medical diagnosis", Fuzzy Sets and Systems 117 (2001) 209–213.

[13] De, S.K., Biswas, R. and Roy, A.R., "On Intuitionistic fuzzy sets", NIFS 3(4) (1997) 14-20.

[14] De, S.K, Biswas, R. and Roy, A.R., "Multicriteria based decision making using Intuitionistic fuzzy set theory", Journal of Fuzzy Math, 6(4) (1998)837-842.

[15] Deval Popat ,Hema Sharda, David Taniar :Classification of fuzzy data in DBMS, KES 2004:691-697

[16] E,Szmidt & Kacpryzk , IFS in some medical applications, I Int. Conference on IFSs , Sofia , Sept. 2001 NIFS 7 (2001) , 58-64

[17] E,Szmidt & Jim F. Baldwin , Assigning the parameters for IFS, Banska Bystrica 22.9.2005.NIFS 11 (2005), 6 , 1-12

[18] Grzegorzewski, Mrowka ; Flexible querying via IFS , International Journal of intelligent systems vol. 22 Issue 6 Pages 587-597, June 2007.

[19] Kolev B. IF relational databases.In Proceedings of 7th international conference on IFS, 23-24 Aug 2003, Sofia, Bulgaria, 109-113.

[20] Mirosla Hudec, An approach to fuzzy database querying , analysis & realisation,UDC 004.4'2,DOI:10.2298/csis 09021274

[21] MP Singh, R.Tiwari, M.Mahajan and Diksha Rani, An architecture for handling fuzzy queries in data warehouses, Contemporary computing,II int,conf. IC3, 2009 , Noida , 240-249 , Springer Verlag Berlin Heidelberg 2009.

[22] Sathi Mukherjee & Kajla Basu , Solving IF assignment problem by using similarity measures & score functions, I.journal of pure & applied Sc. Technologies,2(1)(2011),pp 1-18

[23] Vahid Khatibi, Gholam Ali; A fuzzy-evidential hybrid inference engine for coronary heart disease risk assessment, Expert *Systems with Applications, 2010. 37(12): p. 8536-8542*

[24] Xin Zhang,et.al. Method for aggregating triangular fuzzy IF information & its application to decision making Baltic journal of sustainability 2010 16(2):280-290

[25] Yaser Ahmed et.al. Applying IF approach to reduce search domain in an accidental case IJACSA, Vol.1, No. 4, Oct.2010

[26] Zadeh L.A., (1965), Fuzzy Sets, Inf. & Control, 8,338-353.