Flexible Hierarchies and Fuzzy Knowledge-Based OLAP

Ermir Rogova[1]
Panagiotis Chountas[1]
Krassimir Atanassov[2]

[1] Harrow School of Computer Science, University of Westminster
[2] CLBME, Bulgarian Academy of Sciences

Flexible Hierarchies and Fuzzy Knowledge-based OLAP

Ermir Rogova
HSCS, University of Westminster
Northwick Park, HA1 3TP
London, UK
rogovae@wmin.ac.uk

Panagiotis Chountas
HSCS, University of Westminster
Northwick Park, HA1 3TP
London, UK
chountp@wmin.ac.uk

Krassimir Atanassov
CLBME – Bulgarian Academy of Sciences, B1. 105, Sofia-1113, Bulgaria
krat@argo.bas.bg

Abstract

In this paper we revise the context of "value uncertainty", as part of an OLAP based environment. A new multidimensional-cubic model named as the IF-Cube is introduced which is able to operate over data with imprecision either in the facts or in the dimensional hierarchies. These query requirements led us to introduce the concept of closure of an Intuitionistic fuzzy set over a universe that has a hierarchical structure, H-IFS. We introduce the automatic recommendation of analysis according to the concepts defined as part of a domain ontology in order to guide the decision making with the aid of H-IFS.

1. Introduction

Concepts are used to describe how the data is organized in the data sources and to map such data to the concepts described in the Domain Ontology. These concepts are used to apply extensively to business semantics described in the Domain Ontology to support the rewriting of queries, conditions and to combine OLAP features in this process. These semantics support the automatic recommendation of analysis according to the context of users' explorations and guide the decision making, a feature inexistent in current analytical tools.

Considering the H-IFS milk we try to express different ontological semantics, or kind of relations, "Figure.1" such as to what extent: Condensed whole milk is a "kind-of" whole milk? Etc.

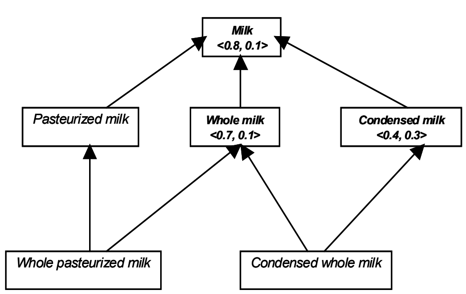

Figure 1. H-IFS 'Milk'

If we wish to summarize the sales for example of products of "Pasteurised milk" we need to take into account as well the fact that "Whole Pasteurised milk" may also be treated as "Pasteurised milk".

Previous approaches close to our work are those regarding similar questions in non-fuzzy contexts. In particular, the propagation of preference or possibility degrees in a hierarchy that we propose is in adequacy with the object model, in which a query on a given class is also addressed to the subclasses of this class. Concerning query enlargement, several works such as [1], [2] use a lattice of concepts to generalise query answers. In studies about possibilistic ontologies [3], each term of an ontology is considered as a linguistic label and has an associated fuzzy description. Fuzzy pattern matching between different ontologies is then computed using these fuzzy descriptions. This approach is related to those concerning the introduction of fuzzy attribute values in the object relational model [4].

Also, studies about fuzzy thesauri have discussed different natures of relations between concepts. Fuzzy thesauri have been considered, for instance, in [5]. However, in our context, the terms of the hierarchy and the relations between terms are not vague. These observations led us to introduce the concept of closure of an H-IFS, which is a developed form defined on the whole hierarchy. Intuitively, in the closure of a H-IFS, the "kind of, $\leq$" relation is taken into account by propagating the degree associated with an element to its sub-elements more specific elements in the hierarchy. For instance, in a query, if the user is interested in the element Milk, we consider that all kinds of Milk, Whole milk, Pasteurised milk, etc. are of interest. On the opposite, we consider that the super-elements (more general elements) of Milk in the hierarchy i.e. Milk are too general to be relevant for the user's query.

2. Principles of Intuitionistic Fuzzy Sets – Atanassov's Sets

As opposed to the classical definition [6,7] a fuzzy set given by $A' = \{< x, \mu_{A'}(x) > | x \in X\}$ where $\mu_A(x) \in [0, 1]$ is

Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)
0-7695-2874-0/07 $25.00 © 2007 IEEE

the membership function of the fuzzy set A′, an Intuitionistic fuzzy set A is given by

$$A = \{< x, \mu_A(x), v_A(x) > | x \varepsilon X\}$$

where: $\mu_A : X \rightarrow [0, 1]$ and $v_A : X \rightarrow [0, 1]$ such that $0< \mu_A(x) + v_A(x)<1$ and $\mu_A(x)\ v_A(x)\ \varepsilon\ [0, 1]$ denote a degree of membership and a degree of non-membership of x ε A, respectively. Obviously, each fuzzy set may be represented by the following Intuitionistic fuzzy set

$$A=\{<x, \mu_A'(x), (x), 1-\mu_A'(x)>| x \varepsilon X\}$$

Definition 1. Let A and B be two fuzzy sets defined on a domain X. A is included in B (denoted $A \subseteq B$) if and only if their membership functions and non-membership functions satisfy the condition: $(\forall \chi \in X)\ (\mu A(x) \leq \mu B(x)\ \&\ \nu A(x) \geq \nu B(x))$

Definition 2. Let Q and D be two Intuitionistic fuzzy sets defined on a domain X and representing, respectively, a flexible query and an ill-known datum:

- The possibility degree of matching between Q and D is defined as: $\Pi(Q / D) = \sup_{x \in X} \min (< 1-\nu_Q(x), \nu_Q(x) >, < 1-\nu_D(x), \nu_D(x) >)$
- The necessity degree of matching between Q and D is defined as: $N(Q/ D) = \inf_{x \in X} \max(<\mu_Q(x), 1-\mu_Q(x)>, <\mu_D(x), 1-\mu_D(x)>)$

3. H-IFS

The notion of hierarchical fuzzy set expresses fuzzy values in the case where these values are part of taxonomies, as for food products or bacteria for example.

The definition domains of the hierarchical fuzzy sets that we propose below are subsets of hierarchies composed of elements partially ordered by the "kind of" relation. An element l_i is more general than an element l_j (denoted $l_i \sim l_j$), if l_i is a predecessor of l_j in the partial order induced by the "kind of" relation of the hierarchy. An example of such a hierarchy is given in Figure. 1. A hierarchical fuzzy set is then defined as follows.

Definition 3. A H-IFS is an Intuitionistic fuzzy set whose definition domain is a subset of the elements of a finite hierarchy partially ordered by the "kind of" ≤ relation.

For example, the fuzzy set M defined as: {Milk<0.8,0.1>, Whole-Milk<0.7,0.1>, Condensed-Milk<0.4,0.3>} conforms to Definition-3. Their definition domains are subsets of the hierarchy given in "Figure 1".

We can note that no restriction has been imposed concerning the elements that compose the definition domain of a H-IFS. In particular, the user may associate a given <μ, ν> with an element l_i and another degree $<\mu_1, \nu_1>$ with an element l_j more specific than l_i . $<\mu, \nu> \sim <\mu_1, \nu_1>$ represents a semantic of restriction for l_j compared to l_i, whereas $<\mu_1, \nu_1> \sim <\mu, \nu>$ represents a semantic of reinforcement for l_j compared to l_i. For instance, in the following H-IFS : <1, 0>/ condensed milk + <0.5, 0.1>/Milk, the element condensed milk has a greater degree than the more general element Milk, which corresponds to a semantic of reinforcement for condensed milk compared to Milk.

4. Closure of the H-IFS

We can make two observations concerning the use of H-IFS:

- Let <1, 0>/condensed milk + <0.5, 0.1>/Milk be an expression of liking in a query. One may also assume that any kind of condensed milk (i.e whole condensed milk) interests the user with <μ, ν>→ <1,0>.
- Two different H-IFS on the same hierarchy cannot be compared using the classic comparison operations of Intuitionistic fuzzy set theory "see section 2". For example, <1, 0>/ condensed milk + <0.5, 0.1>/Milk and <1, 0>/Milk + <0.2, 0.7>/Pasteurised milk are defined on two different subsets of the hierarchy of "Figure. 1" and, thus, are not comparable.

These observations led to the closure of a H-IFS. The kind of (≤) relation is taken into account by propagating the <μ, ν> associated with an element to its sub-elements (more specific elements) in the hierarchy. If the user is interested in the element Milk, we consider that all kinds of Milk are also of interest. On the opposite, we consider that the super-elements (more general elements) of Milk in the hierarchy are too broad to be relevant for the user's query.

Definition 4. Let F be a H-IFS defined on a subset D of the elements of a hierarchy L. Its degree is denoted as <μ, ν>. The closure of F, denoted clos(F), is a H-IFS defined on the whole set of elements of L and its degree $<\mu, \nu>_{clos(F)}$ is defined as follows.

For each element **l** of L, let $S_L= \{l_1, \ldots, l_n\}$ be the set of the smallest super-elements of in D :

- **If S_L not empty, $<\mu, \nu>_{clos(F)}(S_L) = <\max_{1\leq i\leq n}(\mu(L_i)), \min_{1\leq i\leq n}(\nu(L_i))>$** else, **$<\mu, \nu>_{clos(F)}(S_L) = <0, 0>$**

In other words, the closure of a H-IFS F is built according to the following rules. For each element l_I of L:

- If l_I belongs to F, then l_I keeps the same degree in the closure of F (case where $S_L= \{ l_I \}$).
- If l_I has a unique smallest super-element l_I in F, then the degree associated with l_I is propagated to L in the closure of F, $S_L= \{ l_I \}$ with $l_I > l_I$)
- If L has several smallest super-elements $\{l_1, \ldots, l_n\}$ in F, with different degrees, a choice has to be made concerning the degree that will be associated with l_I in the closure. The proposition put forward in Definition-4, consists of choosing the maximum degree of validity μ and minimum degree of non validity v associated with $\{l_1, \ldots, l_n\}$.
- All the other elements of L, i.e., those that are more general than, or not comparable with the elements of F, are considered as non-relevant. The degree <0,0> is associated with them.

Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)
0-7695-2874-0/07 $25.00 © 2007 IEEE

Let us consider once more the H-IFS M defined as: {Milk<0.8,0.1>, Whole-Milk<0.7,0.1>, Condensed-Milk<0.4,0.3>} which is presented in "Figure.1".

The case of whole condensed milk is different: The user has associated the degree <0.8,0.1> with Milk, but has given a restriction on the more specific element whole milk (degree <0.7,0.1>). As whole condensed milk is a kind of whole milk it inherits the <μ,ν> associated with whole milk, that is <0.7, 0.1>. If the H-IFS expresses preferences in a query, the choice of the maximum allows us not to exclude any possible answers. If the H-IFS represents an ill-formulated concept, the choice of the maximum allows us to preserve all the possible values of the datum, but it also makes the datum less specific.

5. Properties of H-IFS

Two different H-IFS, defined on the same hierarchy, can have the same closure, as in the following example.

The H-IFS: Q= {Milk<1,0>, Whole-Milk<0.7,0.1>, Pasteurised-milk<1,0>, Condensed-Milk <0.4,0.3>} and R ={Milk<1,0>, Whole-Milk<0.7,0.1>, Pasteurised-milk<1,0>, Whole-Pasteurised-milk<1,0>, Condensed Milk <0.4,0.3>} have the same closure, represented in Figure 2 below.

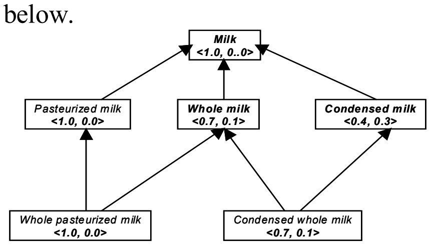

Figure. 2. Common closure of the H-IFS Q and R

Such hierarchical Intuitionistic fuzzy sets form equivalence classes with respect to their closures.

Definition 5. Two H-IFS Q and R, defined on the same hierarchy, are said to be equivalent Q≡R if and only if they have the same closure.

Property Let Q and R be two equivalent H-IFS. If $l_I \in$ dom(Q) $\cap$ dom(R), then $\langle\mu,\nu\rangle(Q.l_I) = \langle\mu,\nu\rangle(R.l_I)$

Proof According to the definition of the closure of a H-IFS F, definition 4, the closure of F preserves the degrees that are specified in F. As Q and R have the same closure (by definition of the equivalence), an element that belongs to Q and R necessarily has the same degree <μ,ν> in both. We can note that R contains the same element as Q with the same <μ,ν>, and also one more element Whole-Pasteurised-milk<1,0>. The <μ,ν> associated with this additional element is the same as in the closure of Q. **Then it can be said that the element,** Whole-Pasteurised-milk<1,0> is **derivable** in R through Q.

Definition 6. Let F be a H-IFS with dom(F) = $\{l_1, \ldots, l_n\}$, and F_{-k} the H-IFS resulting from the restriction of F to the domain dom(F) \ $\{l_k\}$. l_k is deducible in F if

$$\langle\mu, \nu\rangle clos_{(F\text{-}k)} (l_k) = \langle\mu, \nu\rangle clos_{(F)} (l_k)$$

As a first intuition, it can be said that removing a **derivable** element from a H-IFS allows one to eliminate redundant information. But, an element being **derivable** in F does not necessarily mean that removing it from F will have no consequence on the closure. For instance, if the element Pasteurised milk is **derivable** in Q, according to **Definition 6**, removing Pasteurised-milk <1,0> from Q would not modify the degree of Pasteurised milk itself in the resulting closure, but it could modify the degree of its sub-element Whole-pasteurised-milk. Thus, Pasteurised-milk <1,0> can not be derived or removed.

Definition 7. In a given equivalence class (that is, for a given closure C), a H-IFS is said to be minimal if its closure is C and if none of the elements of its domain is **derivable.** For instance, the H-IFS S_1 and S_2 are minimal (none of their elements is derivable). They cannot be reduced further. S_1= Milk<1,0> S_2= {Milk<1,0>, Whole-Milk<0.7,0.1>, Whole-Pasteurised-milk<1,0>, Condensed-Milk <0.4, 0.3>}.

In the next section, a complementary solution is proposed when it comes to lack of answers to a query, i.e. when the user wants to retrieve complementary answers close to his initial query as part of an OLAP environment. New models have appeared to manage incomplete datacube [8], imprecision in the facts and the definition of fact using different levels in the dimensions [9]. Nevertheless, these models continue to use inflexible hierarchies thus making it difficult to merge reconcilable data from different sources that arise due to different perceptions-views about a particular modelling reality.

6.1 Overview of the Cube Model VS Semantics of the IF-Cube

According to [10] a cube structure is defined as a 4-tuple, <D, M, A, f> where the four components indicate the characteristics of the cube. These characteristics are:

a set of n dimensions D = $\{d_1, d_2, \ldots, d_n\}$ where each d_i is a dimension name, extracted from a domain $dom_{dim(i)}$. A set of k measures M = $\{m_1, m_2, \ldots, m_k\}$ where each m_i is a measure name, extracted from a domain $dom_{measure(i)}$. The set of dimension names and measures names are disjoint; i.e., D $\cap$ M = 0. A set of t attributes A = $\{a_1, a_2, \ldots, a_t\}$ where each a_i is an attribute name, extracted from a domain $dom_{attr(i)}$.

In contrast, an **IF-Cube** is an abstract structure that serves as the foundation for the multidimensional data cube model. Cube C is defined as a five-tuple *(D, l, F, O, H)* where:

- *D* is a set of dimensions
- *l* is a set of levels $l_1, \ldots, l_n,$

Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)
0-7695-2874-0/07 $25.00 © 2007 **IEEE**

- A dimension $d_i = (l \leq O, l_{\perp}, l_{\top})$ $dom(d_i)$ where $l = l_i$ $i=1...n$. l_i is a set of values and $l_i \cap l_j = \{\}$, $\leq O$ is a partial order between the elements of l.
  To identify the level l of a dimension, as part of a hierarchy we use dl. $l_{\perp}$: base level $l_{\top}$: top level for each pair of levels l_i and l_j we have the relation: $\mu_{ij} : l_i \times l_j \rightarrow [0,1]$ $\nu_{ij} : l_i \times l_j \rightarrow [0,1]$ $0 < \mu_{ij} + \nu_{ij} < 1$
- F is a set of fact instances with schema $F = \{<x, \mu_F(x), \nu_F(x)> | x \in X\}$, where $x=<att_1, ..., att_n>$ is an ordered tuple belonging to a given universe X, $\mu_F(x)$ and $\nu_F(x)$ are the degree of membership and non-membership of x in the fact table F respectively.
- H is an object type history that corresponds to a cubic structure(l, F, O, H') which allows us to trace back the evolution of a cubic structure after performing a set of operators i.e. aggregation.

The example below provides a sample imprecise cube (D, l, F, O, H) *i.e. sales* and a conceptual non-rigid hierarchy product with reference to milk consisting of $l_i, ..., l_n$ levels with respective levels of membership and non membership $< \mu_{ij}\ \nu_{ij}, >$.

The defined IF OLAP Cube and the proposed OLAP operators allow us to: accommodate imprecise facts, utilise *conceptual hierarchies* used for aggregation purposes in the cases of roll-up and roll down operations and offer a unique feature such as keeping track of the history when we move between different levels of a hierarchical order

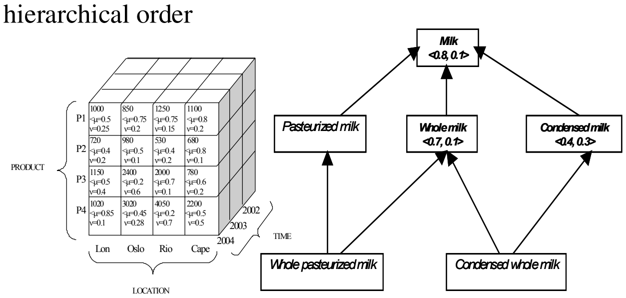

Figure. 3. "Imprecise Cube 'Sales' – Conceptual – Ontological, IF Hierarchy 'Milk' "

In the next section, the fundamental cubic operators are defined and explained.

7. Cubic operators

Selection (Σ): The selection operator selects a set of fact-instances from a cubic structure that satisfy a predicate (θ). A predicate (θ) involves a set of atomic predicates ($\theta_1, \ldots, \theta_n$) associated with the aid of logical operators p (i.e. $\wedge$, $\vee$, etc.) . The set of possible facts (cubic instances) that satisfy the θ should carry a degree of membership μ and non-membership ν expressed as

$F = \{<x, min(\mu_F(x), \mu(\theta(x))), max(\nu_F(x), \nu(\theta(x)))> | x \in X\}$.

Input: $C_i = (D, l, F, O, H)$ and the predicate θ

Output: $C_o = (D, l, F_o, O, H)$ where $F_o \subseteq F$ and $F_o = \{f \mid (f \in F) \wedge (f \text{ satisfies } \theta)$

Mathematical notation: $\sum_{\theta}(C_i) = C_o$

Cubic Product (⊗): This is a binary operator $C_{i1} \otimes C_{i2}$. It is used to relate two cubes C_{i1} and C_{i2} assuming that $D_1 \subseteq D_2$ and O_1 , O_2 are reconcilable partial orders. Thus, l_1, l_2 could lead to l_o being a ragged hierarchy.

Input: $C_{i1} = (D_1, l_1, F_1, O_1, H_1)$ and $C_{i2} = (D_2, l_2, F_2, O_2, H_2)$

Output: $C_o = (D_o, l_o, F_o, O_o, H_o)$ where $D_o = D_1 \cup D_2$, $l_o = l_1 \cup l_2$, $O_o = O_1 \cup O_2$ $H_o = H_1 \cup H_2$, $F_o = F_1 \text{ X } F_2$

$F_o = \{<<x, y>, min(\mu_{f1}(x), \mu_{f2}(y)), max(\nu_{f1}(x), \nu_{f2}(y),)> | <x, y> \in X \times Y\}$

Mathematical notation: $C_{i1} \otimes C_{i2} = C_o$

Union (∪): The union operator is a binary operator that finds the union of two cubes. C_{i1} and C_{i2} have to be union compatible. The operator also coalesces the value-equivalent facts using the minimum membership and maximum non-membership.

Input: $C_{i1} = (D_1, l_1, F_1, O_1, H_1)$ and $C_{i2} = (D_2, l_2, F_2, O_2, H_2)$

Output: $C_o = (D_o, l_o, F_o, O_o, H_o)$ where $D_o = D_1 = D_2$, $l_o = l_1 = l_2$, $O_o = O_1 = O_2$, $H_o = H_1 = H_2$, $F_o = F_1 \cup F_2 = \{ < x, max(\mu F_1(x), \mu F_2(x)), min(\nu F_1(x), \nu F_2(x)) > | x \in X\}$

Mathematical notation: $C_{i1} \cup C_{i2} = C_o$

***Difference (-):*.** The difference operator removes the portion of the cube C_{i1} that is common to both cubes. C_{i1} and C_{i2} have to be union compatible.

Input: $C_{i1} = (D_1, l_1, F_1, O_1, H_1)$ and $C_{i2} = (D_2, l_2, F_2, O_2, H_2)$

Output: $C_o = (D_o, l_o, F_o, O_o, H_o)$ where $D_o = D_1 = D_2$, $l_o = l_1 = l_2$, $O_o = O_1 = O_2$, $H_o = H_1 = H_2$, $F_o = F_1 \cap F_2 = \{ < x, min(\mu F_1(x), \mu F_2(x)), max(\nu F_1(x), \nu F_2(x)) > | x \in X \}$

Mathematical notation: $C_{i1} - C_{i2} = C_o$

Aggregation (A): An aggregation operator A is a function $A(G)$ where $G = \{<x, \mu_F(x), \nu_F(x)> | x \in X\}$ where $x=<att_1, ..., att_n>$ is an ordered tuple belonging to a given universe X, $\{att_1, ..., att_n\}$ is the set of attributes of the elements of X, $\mu_F(x)$ and $\nu_F(x)$ are the degree of membership and non-membership of x. The result is a bag of the type $\{<x', \mu_F(x'), \nu_F(x')> | x' \in X\}$. To this extent, the bag is a group of elements that can be duplicated and each one has a degree of μ and ν.

Input: $C_i = (D, l, F, O, H)$ and the function $A(G)$

Output: $C_o = (D, l_o, F_o, O_o, H_o)$

Roll up (Δ): The result of applying Roll up over dimension d_i at level dl_r using the aggregation operator A over a datacube $C_i = (D_i, l_i, F_i, O, H_i)$ is another datacube $C_o = (D_o, l_o, F_o, O, H_o)$.

An object of type history is a recursive structure H where ω is the initial state of the cube and (l, D, A, H') is the state of the cube after performing an operation on the cube. The structured history of the datacube allows us to keep all the information when applying *Roll up* and get it all back when *Roll Down* is performed. To be able to apply the operation of *Roll Up* we need to make use of the IF_{SUM} aggregation operator.

Roll Down (Ω): This operator performs the opposite function of the *Roll Up* operator. It is used to roll down

Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)
0-7695-2874-0/07 $25.00 © 2007 IEEE

from the higher levels of the hierarchy with a greater degree of generalization, to the leaves with the greater degree of precision. The result of applying *Roll Down* over a datacube C_i = *(D, l, F, O, H)* having H=*(l', D', A', H')* is another datacube C_o= *(D', l', F', O, H'),* where *F' is a set of fact instances defined by operator A.*

To this extent, the *Roll Down* operative makes use of the recursive history structure previously created after performing the *Roll Up* operator. The definition of aggregation operators points to the need of defining the IF extensions for traditional group operators [11], [12] such as *SUM, AVG, MIN and MAX.* Based on the standard group operators, we provide their IF extensions and meaning.

***IF_{SUM}* :** The IF_{sum} aggregate, like its standard counterpart, is only defined for numeric domains. Given a fact F defined on the schema X $(att_1, ..., att_n)$, let att_{n-1} defined on the domain $U=\{u_1, ..., u_n\}$. The fact F consists of fact instances F_i *with* $1 \leq i \leq m$. The fact instances F_i are assumed to take Intuitionistic Fuzzy values for the attribute att_{n-1} for $i = 1$ *to* m we have - $F_i[att_{n-1}] = \{<\mu_i(u_{ki}), \nu_i(u_{ki})>/ u_{ki} \mid 1 \leq k_i \leq n\}$. The IF_{sum} of the attribute att_{n-1} of the fact table F is defined by:

$$IF_{SUM}((att_{n-1})(F)) = \{<u>/ y \mid ((u= \min_{i=1}^{m} (\mu_i(u_{ki}), \nu_i(u_{ki})) \wedge (y = \sum_{ki=k1}^{km} u_{ki}) (\forall_{k1, ...km} : 1 \leq k1, ...km \leq n))\}$$

***IF_{AVG}* :** The IF_{AVG} aggregate, like its standard counterpart, is only defined for numeric domains. This aggregate makes use of the IF_{SUM} that was discussed previously and the standard *COUNT.* The IF_{AVG} can be defined as: $IF_{AVG}((att_{n-1})(F) = IF_{SUM}((att_{n-1})(F)) / COUNT((att_{n-1})(F))$

***IF_{MAX}*:** The IF_{MAX} aggregate, like its standard counterpart, is only defined for numeric domains. The IF_{sum} of the attribute att_{n-1} of the fact table F is defined by:

$$IF_{MAX}((att_{n-1})(F)) = \{<u>y|((u= \min_{i=1}^{m} (\mu_i(u_{ki}),\nu_i(u_{ki})) \wedge (y= \max_{i=1}^{m} (\mu_i(u_{ki}),\nu_i(u_{ki})))(\forall_{k1,...km} :1\leq k1,...km\leq n))\}$$

***IF_{MIN}*:** Given a fact F defined on the schema X $(att_1, ..., att_n)$, let att_{n-1} defined on the domain $U=\{u_1, ..., u_n\}$. The fact F consists of fact instances f_i *with* $1 \leq i \leq m$. The fact instances f_i are assumed to take Intuitionistic Fuzzy values for the attribute att_{n-1} for $i = 1$ *to* m we have $f_i[att_{n-1}] = \{<\mu_i(u_{ki}), \nu_i(u_{ki})>/ u_{ki} \mid 1 \leq k_i \leq n\}$. The IF_{sum} of the attribute att_{n-1} of the fact table F is defined by:

$$IF_{MIN}((att_{n-1})(F)) = \{<u>/ y|((u= \min_{i=1}^{m} (\mu_i(u_{ki}),\nu_i(u_{ki})) \wedge (y= \min_{i=1}^{m} (\mu_i(u_{ki}),\nu_i(u_{ki})))(\forall_{k1,...km} :1\leq k1,...km\leq n))\}$$

8. Conclusions

Whereas in classic fuzzy sets, all the elements are on the same level and are associated with a degree explicitly defined, this is not necessarily the case in H-IFS because several levels of detail exist in the hierarchy, and the hierarchical links between the elements have to be taken into account. The hierarchical links are defined by the "kind of, ≤" relation. H-IFS that have the same closure define equivalence classes, called minimal H-IFS.

We have presented a new multidimensional-cubic model named as the IF-Cube. The main contribution of this new model is that is able to operate over data with imprecision in the facts and the summarisation hierarchies. These features are inexistent in current OLAP tools. Furthermore we notice that our IF cube can be used for the representation of linguistic terms.

9. References

[1] S. Rice, J. F. Roddick: Lattice-Structured Domains, Imperfect Data and Inductive Queries. pp. 664-674, LNCS DEXA 2000

[2] P. Chountas, I. Petrounias: Virtual Integration of Temporal and Conflicting Information. pp. 243-248, IEEE CS Press, IDEAS 2001

[3] "Evaluation of Term-Based Queries Using Possibilistic Ontologies," Soft Computing for Information Retrieval on the Web, E. Herrera-Viedma, G. Pasi, and F. Crestani, eds. Springer-Verlag,2005.

[4] M. Koyuncu, A. Yazici: IFOOD: An Intelligent Fuzzy Object-Oriented Database Architecture. IEEE Trans. Knowl. Data Eng. 15(5): pp. 1137-1154, 2003

[5] S. Miyamoto and K. Nakayama, "Fuzzy Information Retrieval Based on a Fuzzy Pseudothesaurus," IEEE Trans. Systems, Man and Cybernetics, vol. 16, no. 2, pp. 278-282, 1986.

[6] K., Atanassov Intuitionistic Fuzzy Sets, Springer-Verlag, Heidelberg, 1999

[7] K., Atanassov Intuitionistic Fuzzy Sets, Fuzzy Sets and Systems, pp 20, 87–96, 1986

[8] C. Dyreson, Information retrieval from an incomplete data cube, VLDB, Morgan Kaufman Publishers, pp. 532-543, 1996.

[9] T. Pedersen, C. Jensen, and C. Dyreson, A foundation for capturing and querying complex multidimensional data, Information Systems, vol. 26, pp. 383-483, 2001.

[10] H. Thomas & A., Datta, A Conceptual Model and Algebra for On-Line Analytical Processing in Decision Support Databases. Information Systems Research 12: 83-102, 2001

[11] D., Dubois et al. Handling Incomplete or Uncertain Data and Vague Queries in Database Applications, Plenum Press, 1988

[12] H., Prade, Annotated bibliography on fuzzy information processing. Readings on Fuzzy Sets in Intelligent Systems, H. Prade, D. Dubois, and R. Yager, Eds. Morgan Kaufmann Publishers Inc., 1993

Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)
0-7695-2874-0/07 $25.00 © 2007 IEEE