On Imprecision Intuitionistic Fuzzy Sets & OLAP – The Case for KNOLAP

Ermir Rogova and Panagiotis Chountas

Harrow School of Computer Science, Data & Knowledge Management Group,
University of Westminster, Watford Road, Northwick Park, HA5 3TP, London, UK
{rogovae,chountp}@wmin.ac.uk

Abstract. Traditional data repositories are typically focused on the storage and querying of crisp-precise domains of data. As a result, current commercial data repositories have no facilities for either storing or querying imprecise-approximate data. However, when considering scientific data (i.e. medical data, sensor data etc) value uncertainty is inherited to scientific measurements. In this paper we revise the context of "value uncertainty", and examine common models related to value uncertainty as part of the OLAP model. We present our approach for extending the OLAP model to include treatment of value uncertainty as part of a multidimensional model inhabited by flexible date and non-rigid hierarchical structures of organisation.

1 Introduction

In this paper we introduce the semantics of the Intuitionistic Fuzzy cubic representation in contrast to the basic multidimensional-cubic structures. The basic cubic operators are extended and enhanced with the aid of Intuitionistic Fuzzy Logic [1], [2].

Since the emergence of the OLAP technology [3] different proposals have been made to give support to different types of data and application purposes. One of this is to extend the relational model (ROLAP) to support the structures and operations typical of OLAP. Further approaches [4], [5] are based on extended relational systems to represent data-cubes and operate over them. The other approach is to develop new models using a multidimensional view of the data [6].

Nowadays, information and knowledge-based systems need to manage imprecision in the data and more flexible structures are needed to represent the analysis domain. New models have appeared to manage incomplete datacube [7], imprecision in the facts and the definition of fact using different levels in the dimensions [8].

Nevertheless, these models continue to use inflexible hierarchies thus making it difficult to merge reconcilable data from different sources with some incompatibilities in their schemata. These incompatibilities arise due to different perceptions-views about a particular modelling reality.

In addressing the problem of representing flexible hierarchies we propose a new multidimensional model that is able to treat with imprecision over conceptual hierarchies based on Intuitionistic Fuzzy logic. The use of conceptual hierarchies enables us to:

O. Castillo et al. (Eds.): Theor. Adv. and Appl. of Fuzzy Logic, ASC 42, pp. 11–20, 2007.
springerlink.com

- define the structures of a dimension in a more perceptive way to the final user, thus allowing a more perceptive use of the system.
- query information from different sources or even use information or preferences given by experts to improve the description of hierarchies, thereby getting more knowledgeable query results. We outline a unique way for incorporating “kind of” relations, or conceptual imprecise hierarchies as part of a Knowledge based multidimensional analysis (KNOLAP).

2 Semantics of the IF-Cube in Contrast to Crisp Cube

In this section we review the semantics of Multidimensional modeling and Intuitionistic Fuzzy Logic and based on these we propose a unique concept named as Intuitionistic Fuzzy Cube (IF-Cube). The IF-Cube is the basis for the representation of flexible hierarchies and thus flexible facts.

2.1 Principles of Intuitionistic Fuzzy Logic

Each element of an Intuitionistic fuzzy [1], [2] set has degrees of membership or truth (μ) and non-membership or falsity (ν), which don’t sum up to 1.0 thus leaving a degree of hesitation margin (π).

As opposed to the classical definition of a fuzzy set given by $A' = \{< x, \mu_{A'}(x) > |x \in X\}$ where $\mu_A(x) \in [0, 1]$ is the membership function of the fuzzy set A′, an intuitionistic fuzzy set A is given by:

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

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 $\mu_A(x)$ $v_A(x) \in [0, 1]$ denote a degree of membership and a degree of non-membership of $x \in A$, respectively.

Obviously, each fuzzy set may be represented by the following Intuitionistic fuzzy set

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

For each intuitionistic fuzzy set in X, we will call $\pi_A (x) = 1 - \mu_A(x) - v_A(x)$ an intuitionistic fuzzy index (or a hesitation margin) of $x \in A$ which expresses a lack of knowledge of whether x belongs to A or not. For each $x \in A$ $0<\pi_A (x)<1$.

2.2 Overview of the Cube Model

A logical model that influences the database design and the query engines is the *multidimensional-cubic* view of data in the warehouse. In a multidimensional data model, there is a set of *numeric measures* that are the objects of analysis. Examples of such measures are sales, budget, etc. Each of the numeric measures depends on a set of *dimensions,* which provide the context for the measure. The attributes of a dimension may be related via a hierarchy of relationships. In the above example, the product name is related to its category and the industry attribute through a hierarchical relationship, see “Fig.1”.

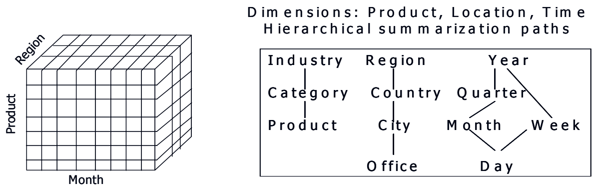

Fig. 1. "Cube "Sales" – Rigid Hierarchies for Product, Location Time Dimensions"

According to [6] 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 $dom_{attr(i)}$. A one-to-many mapping $f : D \rightarrow A$, i.e. there exists, corresponding to each dimension, a set of attributes.

2.3 Semantics of the IF-Cube

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 $l_1, \ldots, l_n,$
- A dimension $d_i = (l \le O, l_\perp, l_\top)$ $dom(d_i)$ where $l = l_i$ $i=1 \ldots n$.
  l_i is a set of values and $l_i \cap l_j = \{\}$,
  $\le O$ 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
  $\mu_{ij} : l_i \times l_j \rightarrow [0,1]$ $\quad \nu_{ij} : l_i \times l_j \rightarrow [0,1]$ $\quad 0 < \mu_{ij} + \nu_{ij} < 1$
- *F* is a set of fact instances with schema $F = \{<x, \mu_F(x), \nu_F(x)> \mid x \in X\}$, where $x = <att_1, \ldots, att_n>$ 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 $l_i, \ldots, l_n$ levels with respective levels of membership and non membership $< \mu_{ij}\ \nu_{ij}, >$.

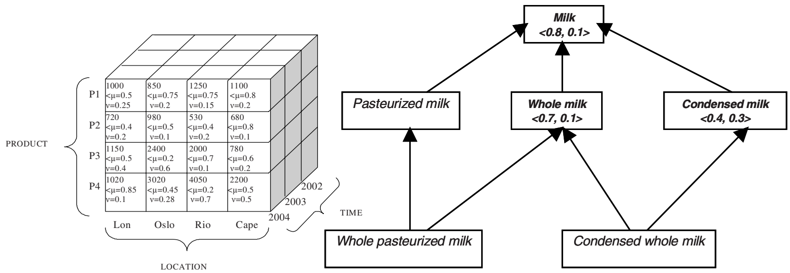

Fig. 2. "Imprecise Cube 'Sales' – Conceptual – Ontological, IF Hierarchy 'Milk' "

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.
- offer a unique feature such as keeping track of the history when we move between different levels of a hierarchical order.

In the next section, the fundamental cubic operators are defined and explained with the aid of examples. The examples make use of cubic slices commonly known as fact tables. Each operator is presented in the following format: the operator's name, symbol, textual description, input, output, mathematical description and an example of the operator.

3 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 (θ_1, …, θ_n) associated with the aid of logical operators *p (* i.e. ∧, ∨, etc.) . The set of possible facts (cubic instances) that satisfy the θ should carry a degree of membership μ and non-membership v expressed as

$$F = \{<x, min(\mu_F(x), \mu(\theta(x))), max(v_F(x), v(\theta(x)))> \mid x \in X\} \quad (1)$$

This guaranties a resulting cube populated with fact instances that satisfy the predicate (θ) either completely or to some degree of certainty.

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

Example: Find the sales amount of 1000 with membership of greater than 0.4 and non membership of less than 0.3 for all products in all cities during 2004

$$\boldsymbol{\Sigma_{(amount>1000 \,\wedge\, (\mu>0.4 \,\wedge\, v<0.3) \,\wedge\, year=2004\,)}(Sales)=C_{Result}}$$

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

$$D_o = D_1 \cup D_2,\quad l_o = l_1 \cup l_2,\ O_o = O_1 \cup O_2\quad H_o = H_1 \cup H_2,\ F_o = F_1 \text{ X } F_2$$

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

Mathematical notation: $C_{i1} \otimes C_{i2} = C_o$
Example: Consider the two cubes we want to relate, C_{i1}: C_{Sales} and C_{i2}: $C_{Discounts}$. $C_{Discounts}$ has the same dimensions as C_{Sales} except the measure amount is not sale but is a discount. In that case the cubic product would be:

$$\mathbf{C_{Sales'}} \otimes \mathbf{C_{Discounts}} = \mathbf{C_{Result}}$$

ProdID	StoreID	Amount	<μ, ν>
P1	S1	10	.7, .2
P2	S2	15	.5, .5

⊗

ProdID	StoreID	Discount	<μ, ν>
P2	S1	2	.5, .5
P3	S3	5	.3, .3

=

S.ProdID	S.StoreID	S.Amount	D.ProdID	D.StoreID	D.Discount	<μ, ν>
P1	S1	10	P2	S1	2	.5, .5
P1	S1	10	P3	S3	5	.3, .3
P2	S2	15	P2	S1	2	.5, .5
P2	S2	15	P3	S3	5	.3, .5

Fig. 3. Fact-Sales', Fact-Discounts and Fact-Result

Join (Θ): It can be expressed using Cubic Product operator. $C_{i1} = (D_1, l_1, F_1, O_1, H_1))$ and $C_{i2} = (D_2, l_2, F_2, O_2, H_2)$ are candidates to join if $D_1 \cap D_2 \neq 0$,
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: $C_o = (D_o, l_o, F_o, O_o, H_o)$

Mathematical notation: $C_{i1}\ \Theta\ C_{i2} = \sigma_p(C_{i1} \otimes C_{i2})$

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: $C_{i1} \cup C_{i2} = C_o$
Example: Consider the two cubes we want to relate, C_{i1}: C_{Sales_North} and C_{i2}: C_{Sales_South}, in that case the union of these two cubes would be:

$$\mathbf{C_{Sales\_North}} \cup \mathbf{C_{Sales\_South}} = \mathbf{C_{Result}}$$

ProdID	StoreID	Amount	<μ, ν>
P1	S1	10	.7, .2
P2	S2	15	.5, .5

∪

ProdID	StoreID	Amount	<μ, ν>
P1	S1	10	.5, .5
P3	S3	5	.3, .3

=

S.ProdID	S.StoreID	S.Amount	<μ, ν>
P1	S1	10	.7, .2
P2	S2	15	.5, .5
P3	S3	5	.3, .3

Fig. 4. Fact-Sales_North, Fact-Sales_South and Fact-Result

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 $D_o = D_1 = D_2$, $l_o = l_1 = l_2$, $O_o = O_1 = O_2$, $H_o = H_1 = H_2$,
$F_o = F_1 \cap F_2 = \{ < x, \min(\mu F_1(x), \mu F_2(x)), \max(\nu F_1(x), \nu F_2(x)) > \mid x \in X \}$
Mathematical notation: $C_{i1} - C_{i2} = C_o$
Example: Consider the two cubes we want to relate, C_{i1}: C_{Sales_North} and C_{i2}: C_{Sales_South}, in that case the difference between North and South sale cubes would be:

$$\mathbf{C_{Sales\_North} - C_{Sales\_South} = C_{Result}}$$

ProdID	StoreID	Amount	<μ, ν>
P1	S1	10	.7, .2
P2	S2	15	.5, .5

-

ProdID	StoreID	Amount	<μ, ν>
P1	S1	10	.5, .5
P3	S3	5	.3, .3

=

S.ProdID	S.StoreID	S.Amount	<μ, ν>
P1	S1	10	.5, .5
P2	S2	15	.5, .5

Fig. 5. Fact-Sales_North, Fact-Sales_South and Fact-Result

3.1 Extended Operators

Aggregation (A): An aggregation operator A is a function $A(G)$ where $G = \{<x, \mu_F(x), \nu_F(x)> \mid x \in X\}$ where $x = <att_1, \ldots, att_n>$ is an ordered tuple belonging to a given universe X, $\{att_1, \ldots, att_n\}$ 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 $\{<x', \mu_F(x'), \nu_F(x')> \mid x' \in X\}$. 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: $C_o = (D, l_o, F_o, O_o, H_o)$

The definition of the extended group operators allows us to define the extended group operators ***Roll up (Δ), and Roll Down (Ω).***

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

$C_o = (D_o, l_o, F_o, O, H_o)$.
Input: $C_i = (D_i, l_i, F_i, O, H_i)$
Output: $C_o = (D_o, l_o, F_o, O, H_o)$

An object of type history is a recursive structure $H = \begin{cases} \omega & \textit{is the initial state of the cube} \\ (l, D, A, H') & \text{is the state of the cube after performing an operation on the cube} \end{cases}$

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 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').*

Input: C_i=*(D, l, F, O, H)*
Output: C_o=*(D',l', F', O, H')* where *F'* → *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 operator points to the need of defining the IF extensions for traditional group operators [9], 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 $(att_1, \ldots, att_n)$, let att_{n-1} defined on the domain $U=\{u_1, \ldots, u_n\}$. The fact F consists of fact instances F_i *with* $1 \le i \le m$. 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 $F_i[att_{n-1}] = \{<\mu_i(u_{ki}), \nu_i(u_{ki})>/ u_{ki} \mid 1 \le k_i \le n\}$. The IF_{sum} of the attribute att_{n-1} of the fact table F is defined by:

$$IF_{SUM}((att_{n-1})(F)) =$$

$$\{<u>/ y \mid ((u = \min_{i=1}^{m} (\mu_i(u_{ki}), \nu_i(u_{ki})) \wedge (y = \sum_{ki=k1}^{km} u_{ki}) (\forall_{k1, \ldots km} : 1 \le k1, \ldots km \le n))\}$$

Example: $IF_{SUM}((Amount)(ProdID))$
={<.8,.1>/10}+{(<.4,.2>/11),(<.3,.2>/12)}+{(<.5,.3>/13),(<.5,.1>/12)}
={(<.8∧.4,.1∧.2>/10+11),(<.8∧.3, .1∧.2>/10+12)}+{<.5, .3>/13, <.5, .1>/12}
={(<.4, .2>/21), (<.3, .2>/22)} + {<.5, .3>/13, <.5, .1>/12}
={(<.4∧.5,.2∧.3>/21+13),(<.4∧.5,.2∧.1>/21+12),(<.3∧.5,.2∧.3>/22+13),(<.3∧.5, .2∧.1>/22+12) = {(<.4, .3>/34), (<.4, .2>/33), (<.3, .3>/35), (<.3, .2>/34)}
={(<.3, .3>/34), (<.4, .2>/33), (<.3, .3>/35)}

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:

$$IF_{AVG}((att_{n-1})(F) = IF_{SUM}((att_{n-1})(F)) / COUNT((att_{n-1})(F))$$

Example: $IF_{AVG}((Amount)(ProdID))$
= $IF_{SUM}((Amount)(ProdID)) / COUNT((Amount)(ProdID))$
= {(<.3, .3>/34), (<.4, .2>/33), (<.3, .3>/35)} / 3
= {(<.3, .3>/11.33), (<.4, .2>/11), (<.3, .3>/11.66)}

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:

$$IF_{MAX}((att_{n-1})(F)) =$$

$$\{<u>y \mid ((u = \min_{i=1}^{m} (\mu_i(u_{ki}), \nu_i(u_{ki})) \wedge (y = \max_{i=1}^{m} (\mu_i(u_{ki}), \nu_i(u_{ki})))(\forall_{k1, \ldots km} : 1 \le k1, \ldots km \le n))\}$$

Example: $IF_{MAX}((Amount)(ProdID))$
IF_{MAX}={<.8,.1>/10},{(<.4,.2>/11),(<.3,.2>/12)},{(<.5,.3>/13),(<.5, 0.1>/120)}
={(<.8∧.4,.1∧.2>/max(10,11)),(<.8∧.3,.1∧.2>/max(10,12)},{<.5,.3>/13,<.5, .1>/12}
={(<.4,.2>/11),(<.3,.2>/12)},{<.5,.3>/13, <.5, .1>/12} ={(<.4∧.5,.2∧.3>/max(11,13)),
(<.4∧.5,.2∧.1>/max(11,12)),(<.3∧.5,.2∧.3>/max(12,13)),(<.3∧.5,.2∧.1>/max(12,12))
={(<.4,.3>/13),(<.4,.2>/12),(<.3,.3>/13), (<.3,.2>/12)} = **{(<.3,.3>/13),(<.3,.2>/12)}**

IF_{MIN} **:** The IF_{MIN} aggregate, like its standard counterpart, is only defined for numeric domains. Given a fact F defined on the schema X $(att_1, \ldots, att_n)$, let att_{n-1} defined on the domain $U=\{u_1, \ldots, u_n\}$. The fact F consists of fact instances f_i *with* $1 \le i \le m$. 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 $f_i[att_{n-1}] = \{<\mu_i(u_{ki}), v_i(u_{ki})>/ u_{ki} \mid 1 \le k_i \le n\}$. The IF_{sum} of the attribute att_{n-1} of the fact table F is defined by:

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

We can observe that the IF_{MIN} is extended in the same manner as IF_{MAX} aggregate except for replacing the symbol **max** in the IF_{MAX} definition with **min**. Once we have defined our Intuitionistic Fuzzy multidimensional model and have defined the IF cubic-algebra, the concept of knowledge based OLAP is introduced. Ideally, in a Knowledge based OLAP environment for summarizing purposes it is desirable to use Intuitionistic Fuzzy hierarchies like milk see “Fig.2” instead of rigid hierarchies like Product in “Fig.1”.

4 The Case for Knowledge Based OLAP-KNOLAP

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 definitions are used to apply more extensively the business semantics described in the Domain Ontology, to support the rewrite 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 in order to guide the decision making, feature inexistent in current analytical tools.

With respect to the Intuitionistic Fuzzy hierarchy milk, we try to express different ontological semantics, or “kind of” relations such as to what extent:

- Condensed whole milk is a “kind-of” Whole milk?
- Condensed whole milk is a “kind-of” Condensed milk?
- Pasteurised whole milk is a “kind-of” Whole milk?
- Pasteurised whole milk is a “kind-of” Pasteurised milk?
- Pasteurised milk is ”kind-of” milk? Etc.

It is obvious from the above examples that if we wish to summarise 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” when applying i.e. the ***IF_{SUM}.***

These observations led us to introduce the concept of closure of an Intuitionistic fuzzy set over a universe that has a hierarchical structure, which is a developed form defined on the whole hierarchy. Intuitively, in the closure of this Intuitionistic fuzzy set, the "kind of" 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, Pasteurized 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.

Let us consider the Intuitionistic fuzzy set M defined as: {Milk<0.8,0.1>, Whole-Milk<0.7,0.1>, Condensed-Milk<0.4,0.3>}} which is presented in "Fig.6". Then the next step is to calculate the $<\mu, \nu>$ values for "Pasteurized milk", "Whole Pasteurized milk " and "Condensed whole milk."

- If the hierarchical IF structure expresses preferences in a query, the choice of the maximum values for μ and minimum value ν from the pairs of values $<\mu, \nu>$ from the parent elements to the sub elements allows us not to exclude any possible answer (high possibility necessity degrees). In real cases, the lack of answers to a query generally makes this choice preferable, because it consists of widening the query answer rather than restricting it.
- If the hierarchical IF represents an ill-known concept, the choice of the maximum value for μ and minimum value ν allows us to preserve all the possible values, but it also makes the answer less specific. In a way, it also participates in enlarging the query, as a less specific datum may share more common values with the query (the possibility degree of matching can thus be higher, although the necessity degree can decrease).

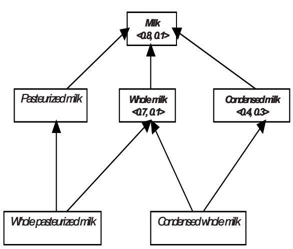

Fig. 6. "IF Hierarchy 'Milk' "

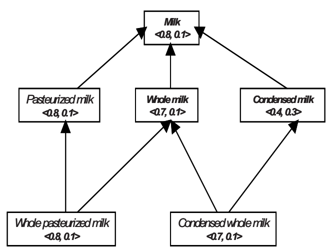

Fig. 7. "Fully weighted Hierarchy 'Milk' "

"Fig.7" is a fully weighted Hierarchy after applying the maximum values for μ and minimum value ν from the pairs of values $<\mu, \nu>$ from the parent elements to the sub elements, i.e. from (whole-milk, condensed-milk) to (condensed-whole-milk), from (milk) to (pasteurized milk), and from (whole-milk, pasteurized milk) to (pasteurized-whole-milk).

The complete study of the hierarchical IF requires the formal definition of the IF hierarchical closure. We will further need to formally define the containment of an IF hierarchical set to another.

Furthermore if one wishes to consider multiple versions of evolving IF hierarchies, the similarity between different versions of IF hierarchical set in the geometrical framework introduced by [10], [11] needs to be examined as well.

5 Conclusions

In this paper 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. Classical models imposed a rigid structure that made the models present difficulties when merging information from different but still reconcilable sources. We introduce the automatic recommendation of analysis according to the context of users' explorations in order to guide the decision making with the aid of Intuitionistic fuzzy set over a universe that has a hierarchical structure and the corresponding hierarchies.

These features are inexistent in current OLAP tools. Furthermore we notice that our IF cube can be used for the representation of Intuitionistic fuzzy linguistic terms.

There is a need to formally define the closure of Intuitionistic fuzzy set over a universe that has a hierarchical structure as well the containment between different versions of these sets. We also need to study the impact of imprecision with respect to star and snowflake data warehouse conceptual structures. Finally, a graphical way needs to be developed to represent the results of the operations in order to get a more intuitive way to read the information obtained.

References

[1] K., Atanassov (1999). *Intuitionistic Fuzzy Sets*, Springer-Verlag, Heidelberg

[2] K., Atanassov *Intuitionistic Fuzzy Sets,* Fuzzy Sets and Systems, 20, 87–96, (1986)

[3] R. Kimball, *The Data Warehouse Toolkit.* New York: John Wiley & Sons, 1996.

[4] S., Chaudhuri U. Dayal V., Ganti. Database Technology for Decision Support Systems. In: Computer, Vol. 34, p. 48-55, 2001

[5] M., Jarke et al. Fundamentals of data warehouses. Springer, London, 2002

[6] 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

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

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

[9] E. Rundensteiner L. Bic. Aggregates in posibilistic databases, VLDB'89, pp. 287-295, 1989

[10] E., Szmidt & J. Kacprzyk Distances between intuitionistic fuzzy sets, Fuzzy Sets and Systems, Vol. 114, No.3,505–518, 2000

[11] E., Szmidt & J. Kacprzyk A new concept of similarity measure for Intuitionistic fuzzy sets and its use in group decision making LNAI 3558, pp272–282, Springer, 2005