Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING

João José Furtado Vasco, Colette Faucher, Eugène Chouraqui

DIAM-IUSPIM - Université d'Aix Marseille III Av. Escadrille Normandie-Niemen

F-13397 Marseille Cedex 20

Diam_ef@vmesa11.u-3mrs.fr

Abstract. *In this paper, we describe an architecture for helping frame hierarchy conception. This architecture is based on machine learning and cognitive psychological studies on categorization. Our basic assumption is that categorization should be considered as a goal-driven, context-dependent process and therefore the hierarchical organization of categories should be represented in different perspectives. The core of our architecture is a learning system of categorization that generates multi-perspective hierarchies. Concept hierarchies are, at first, generated in a probabilistic representation and after translated into a frame one.*

Key Words*: Categorisation, Frame-based Classification, Concept Formation.*

1. INTRODUCTION

Object Oriented Representations (OOR) or frame-based languages organize pieces of knowledge related to an entity in declarative structures (frames). In this context, there are many works treating the classification of an entity in a frame hierarchy ([Rechenmann 88] [Brachman 85] and [Napoli 90] to mention a few) but hardly any of these works are interested in an automatic construction of these hierarchies ([Aguirre 89] is an example). The modeling of these hierarchies requires complete acquaintance with the underlying concepts of the domain to be represented. However, from a cognitive viewpoint, it is more flexible to represent observations regarding things which *a priori* one doesn't know precisely, but which allow one to construct, incrementally and automatically, abstract representations that describe in intention these initial observations.

We propose an approach for construction of frame hierarchies that makes use of the machine learning and cognitive psychology ideas of concept formation and categorization. We have defined an architecture, called CONFORT (CONcept Formation in Object RepresenTation), for construction of categories from observations (description of specific entities; a conjunction with properties represented by attribute-value couples). This architecture can be considered a knowledge acquisition tool for helping an expert in his activity of expressing and elaborating concepts of his domain [Vasco 95a]. According to cognitive psychological studies, CONFORT is based on the assumption that categorization is a goal-driven process [Barsalou 83],[Seifert 88]. This assumption leads us to consider that concept hierarchies should be viewed from different perspectives giving rise to different hierarchical organizations according to different usage determined by the expert's categorization goals or opinions.

The core of CONFORT is FORMVIEW, a learning algorithm of incremental concept formation that we have developed using the frame-based language Objlog+ [Faucher 91]. FORMVIEW constructs multiple hierarchies of probabilistic concepts named probabilistic concept trees [Fisher 88]. These trees are a hybrid representation where cases and abstract concepts that subsume these cases are represented. From probabilistic concepts, CONFORT creates a frame-like representation. In this paper, we describe the main ideas of CONFORT focusing on the concept formation algorithm FORMVIEW and on the aspects of the transformation of *probabilistic concepts to frames.*

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

This article has the following organization. In Section 2, we define some basic notions and highlight considerations on frame-based representations and on concept formation from a machine learning perspective. In Section 3, we describe the CONFORT architecture. Section 4 describes the concept formation algorithm FORMVIEW. In Section 5 we describe the passage from probabilistic concepts to frames, and finally in Section 6 our conclusions are stated and our future researches are indicated.

2. BACKGROUND

In order to design a software architecture for helping the frame hierarchy construction, our researches are firstly focused on the concept formation problem. The main assumption, that frames represent concepts of the world, leads us to consider understanding of the human's concept formation process. Therefore, our work is fundamentally based on psychological findings about concepts and categories.

We consider a concept refer to an idea or notion by which people can understand some aspects of the world [Hampton 93]. A category is a set of entities(objects, events, actions, states, etc.) which are grouped together on the basis of some criterion of categorization. We can thus associate the notions of concept and category because in reality a concept provides a way to categorize the world into those entities that instantiate the concept, and those that do not. In other words, considering a concept a categorization criterion, we can speak of the category associated with a concept, as the set of entities that satisfies such a concept. Typically, we use the expression that a concept represents or characterizes a category. This distinction between concepts and categories is similar to the notions of intension and extension. Concepts concern to intension(information used as categorization criterion) whereas categories refer to extension(the members that satisfy the categorization criterion). Below we define more formally these notions as well as some others that will be important to the comprehension of this article.

Given

E	set of entities that will be categorized $E=\{e_1, e_2, ..., e_n \mid n \in \mathbf{N}\}$
$\mathcal{P}(\mathbf{E})$	set of parts of E, called **categories**
A	set of **attributes** describing the entities of E. Ex: A={Age, Sex, Tail}
$\upsilon(j)$	set of possible values of the attribute *j* of A, υ(Age)={young, Adult}
V	set of values of all the attributes $\in$A; $V=\{\upsilon(k) \mid k \in A\}$
O	an **observation** describing an entity e; $O=\{p_i=(j,v) \mid j \in A, v \in \upsilon(j)\} \in \mathcal{P}(A \times V)$ we call p_i a **property** of O $(1 \leq i \leq n)$
H	a strict hierarchy (disjoint categories) to establish on E : H is a finite not empty set of categories; $H \subset \mathcal{P}(E) - \{\varnothing\}$ H is an oriented acyclic graph $H=(\mathcal{P}(E), \prec, R)$ where R is the maximal element following the partial order relation $\prec$ (specialization relation).
EC	the space of representation of categories (the admissible representations according to some criterion)
$\mathbf{C_k}$	the representation of the category C_k($\in \mathcal{P}(E)$): $\mathbf{C_k} \in$ EC, called **Concept**
OB	set of experts' categorization goals experts; $OB=\{ob_1, ob_2, ..., ob_n\}$
PV	set of perspectives reflecting the categorization goals. $PV=\{pv_1, pv_2, ..., pv_n\}$. Each ob_k($\in$ OB) correspond to one pv_k ($\in$ PV) $(1 \leq k \leq n)$.
$\mathbf{H\uparrow pv}$	is a hierarchy H established on E reflecting a perspective pv ($\in$ PV)

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

2.1 Concept Formation from a Machine Learning Perspective

Machine learning systems build concepts in *intension* from exemplars of their *extension*. In this context, we concentrate on learning from observation systems or conceptual clustering [Michalsky 86][Fisher 87]. These systems recognize regularities among a set of non-preclassified entities or events and induce a concept hierarchy that organizes these observations. Generally, concept formation systems can be defined by a quadruplet(H, I, C, μ) [Thaise 91] , where :

- H is the possible space of concept hierarchies;
- I is the set of observation about entities to categorize;
- C is a partitioning of all these entities in conceptual categories C_i which are structured hierarchically($C \in H$) and optimizing a defined quality criterion;
- μ is the set of operators of construction or organization of categories which we can employ to the members of H to generate C.

A concept formation algorithm is reduced to a hill-climbing search, in H, for a hierarchy of conceptual categories C that covers all the entities described in I and that optimizes the evaluation function f(C, I) measuring the quality criterion. The fundamental point of this research concerns the application of the μ's operator that produces an optimal value for f(C, I). On incremental concept formation, entities are treated one after another as soon as they are observed. As the general case, the classification of new entities is made by their adequacy to the existing conceptual categories.

A typical incremental concept formation system is COBWEB [Fisher 87]. It is a pioneer system influenced by research in cognitive psychology on *basic level, probabilistic concepts* and *typicality effects* [Rosch 76][Fisher 93]. In addition, it has given rise to many other successors (BRIDGER [Reich 94] and CLASSIT [Gennari 89] are examples). Table 1 shows the main lines of a procedure of a COBWEB-like system.

FUNCTION PRINCIPAL (*Root, Observation*)

1. Incorporate *Observation* in *Root*
2. Choose the best operator to employ on the partition P of the Root's direct sub-concepts, among the following:
   a) Incorporate *Observation* into a concept of P
   b) Create a new concept under *Root* to receive *Observation*
   c) Merge the 2 best concepts of P in a new concept that includes Observation
   d) Split a concept of P in its children, adding Observation to the best of these
3. If the operation 2b , read another observation
   Else return to 1 with Root = the concept of P in which Observation was inserted

Table 1. Control structure of a COBWEB-like system

2.2 Frame-based Representations

Frame-based representations have been shown to be adequate declarative knowledge representation models. They structure the world upon the frame notion which describe either a concept representing a category of entities or a concrete entity(an instance).

A frame is composed of *slots* which represent the attributes of the category's entities. Each slot has a set of *facets* to characterize them. Usually, frame-based languages have at least two kind of facets: the value facet (determine the slot value), the domain facets (determine the acceptable domain's values for the slot).

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

Frames can be linked by the specialization relation *a-kind-of* which represents a strict set inclusion relation. We can say that a frame description is the superframe's description with some additional specific knowledge. Only this specific knowledge is stored in the subframes; the specialization links allow an inheritance mechanism that finds out attributes in high-level frames. A frame instance has only the value facet and is linked to a superframe via the *is-a* link. The instance's values should satisfy the constraints defined in such a superframe such as having values that belong to the superframe's set of domain values.

3 THE CONFORT MAIN CHARACTERISTICS

CONFORT is a software architecture to help in the construction of frame hierarchies. It is based on psychological cognitive and machine learning findings on categorisation and concept formation. CONFORT supposes that goals of categorization exist (supplied by one or more experts) prior to initiation of the process. Categorization uses a scheme to weigh an observation's properties based on prior expectations of the relevance of particular properties within the task domain. Both a property's relevances and relationships are represented in a GDN (Goal Dependence Network) [Michalsky 86] according to categorization goals. Property relationships are implications between initial observed properties(typically surface properties) and those dependent on the expert domain(functional properties). Thus, an observation is represented by observer's defined properties and eventually by GDN's inferred ones. An expert should define goals, property relevance, property relationship and, in addition, he can intervene to provide feedback to concept formation process.

A goal-driven concept formation process leads us naturally to a multi-perspective representation since goals have influence on the determination of relevance for context-specific features which will favour the generation of different hierarchical organizations. For instance, to achieve the goal to buy a pet for a child, one would consider beauty and cheapness as relevant properties. As a result, animal hierarchical organization will reflect this particular situation and will probably be different from that generated from a veterinary surgeon perspective where other properties would be relevant(e.g. physiologic properties).

The core of our architecture is FORMVIEW, a hill-climbing algorithm for concept formation, that generates multiple hierarchies and uses a category quality measure that takes into account the relevance of an observation's properties and generated categories in other perspectives. Figure 1 illustrates the principal ideas of CONFORT.

In CONFORT, the construction of concept hierarchies is incremental and gradual, allowing iterative reevaluations and an expert's feedback. However, we suppose that these hierarchies should reach a certain degree of stability in order to permit their exploitation(for instance, with classification of new entities). Therefore, we have developed a phase of translation from a probabilistic representation to an abstract frame-based one. The FORMVIEW's hybrid probabilistic representation stores observations and conceptual structures that provide easy access to these observations. Reasoning at the case-level is most productive when few training observations are available and noise is not present [Fisher 89]. However, when noise increases, this strategy can be inefficient. In this situation, abstract representations are preferred since they are less sensitive to noise. That is another reason that have motivated us to develop a passage from probabilistic concepts to frame representations. This phase is executed after the incremental process of concept formation.

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

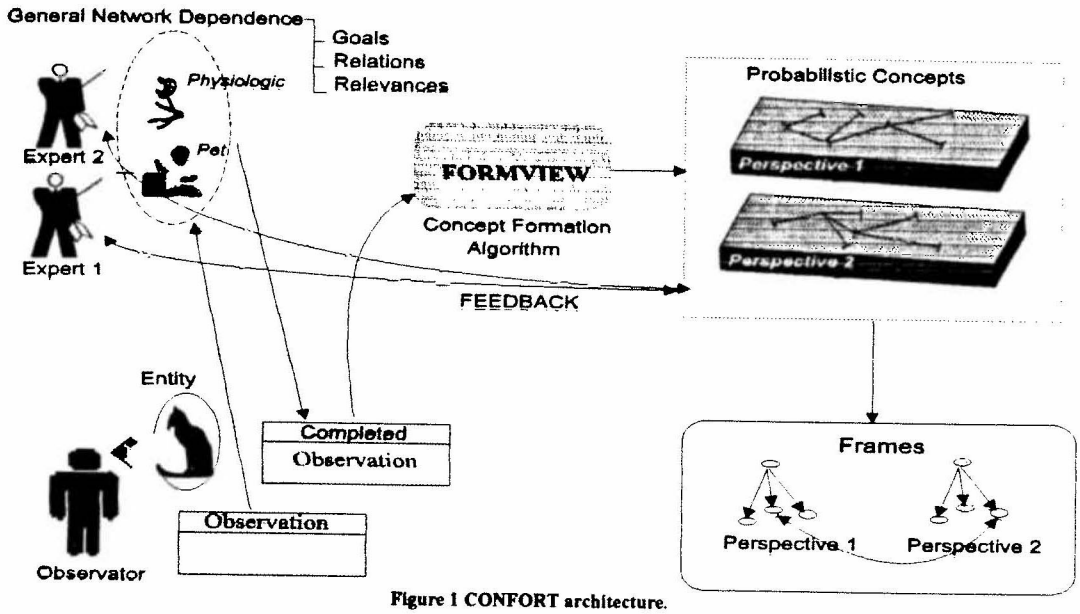

Figure 1 CONFORT architecture.

4 FORMVIEW: CONSTRUCTING MULTI-PERSPECTIVE CONCEPT HIERARCHIES

FORMVIEW can be classified as a hill-climbing algorithm in a space of concept hierarchies (which we have defined in Section 2). Its principal characteristic is the generation of multiple hierarchies which represent different perspectives determined by categorization goals.

4.1 Knowledge Representation in FORMVIEW

An important feature of CONFORT and particularly of FORMVIEW is that they were developed in an OOR context. We integrate them into the frame-based language Objlog+[Faucher 91] permitting an automatic construction of frame hierarchies from specific entities. The details and advantages of this integration are out of the scope of this paper, and can be obtained in [Vasco 95b]. Here, we will describe FORMVIEW's components via the formal notions defined early.

4.1.1 Description of Inputs

The main FORMVIEW input is an observation describing an entity that belongs to E(cf. definitions §2). FORMVIEW uses as additional data a goal dependence network(GDN). This GDN contains:

- For each categorization goal, we define a degree of relevance(between 0 and 1) for each attribute-value pair. For instance:

  Objectif1 $\Rightarrow$[$(A_1=V_2, 0.9), (A_2=V_4. 0.3)$],

  Objectif2 $\Rightarrow$[$(A_1=V_3, 0.4), (A_3=V_5. 0.8)$]
- For each categorization goal, the existence of an attribute-value pair can determine the existence of another pair:

  $Objectif_x : (A_i=V_j \Rightarrow A_k=V_z)$

Formally:

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

Definition 1

A general dependence network consists, for each obj from OB, of two sets:

A set IP of pairs $ip=(p^{prem}, p^{inf})$ where p^{prem} (premisse property) and p^{inf} (conclusion property) $\in \mathcal{P}(A \times V)$, that is , $IP^j = \{ ip_1, ip_2, ..., ip_z \}$ and,

A set IW of doublets $pw=(p, w)$ where $p \in \mathcal{P}(A \times V)$ and $w \in [0, 1]$(p's semantic relevance), that is, $IW^j = \{ iw_1, iw_2, ..., iw_z \}$.

With $(1 \leq j \leq k)$ $(0 \leq z < card(\mathcal{P}(A \times V)^2)$.

4.1.2 Description of Concepts

At first, FORMVIEW constructs *probabilistic concepts* [Smith 81]. These concepts have the probability that an observation is classified into the category represented by the concept P(C), all possible values for their attributes and each such value having its associated *predictability* and *predictiveness* [Fisher 87]. The predictability is the conditional probability that an observation x has value v for an attribute a, given that x is a member of a category C, or $P(a=v|C)$. The predictiveness is the conditional probability that x is member of C given that x has value v for a or $P(C|a=v)$. Indeed, FORMVIEW constructs multiple hierarchies of probabilistic concepts named probabilistic concept trees [Fisher 88]. These trees are a hybrid representation where cases and abstract concepts that subsume these cases are represented.

A probabilistic concept is a conjunction of characteristics defined by a quadruplet: $(j, \upsilon(j), PD_v, PP_v)$:

Where j is an attribute from A.

$\upsilon(j)$ is a set of values of the attribute j.

PD is the set of the values of the conditional probabilities P(j=v|C) (predictability) for each value v ($v \in \upsilon(j)$).

PP is the set of the values of the conditional probabilities P(C|j=v) (predicteveness) for each value v($v \in \upsilon(j)$).

Formally:

Definition 2

Be a set of entities E defined by observations, and a hierarchy $H\uparrow pv$ built on E and representing a perspective pv. In FORMVIEW, all category of entities C of H define the probabilistic concept **CP** below:

$\mathbf{CP} = \{(j, \upsilon(j), PD, PP) | j \in A\}$

With $PD \in \mathbf{R}$

$PP \in \mathbf{R}$

we name the pair (j , v) ; $v \in \upsilon(j)$; **a property of CP**

FORMVIEW constructs multiple probabilistic concept trees which represents different points of view corresponding to different choices of categorization goals (GDN's goals).

In our work, the basic ideas on multi-perspective representations were based on the TROPES model[Marino 90]. The main feature of this model is the existence of a communication channel among hierarchies representing different perspectives. This communication is supplied with oriented links between categories called *bridges*. Two types of bridges are possible: unidirectional and bi-directional. Bi-directional bridges represent set equality relation while unidirectional one represent set inclusion relation. More precisely:

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

Definition 3

A bridge between a category C of a perspective pv ($C \in H\uparrow pv$) and another category C' in another perspective pv'($C' \in H'\uparrow pv'$), noted as bridge(C,C'), is defined as:

$$\text{bridge}(C, C') = \begin{cases} 1 & \Leftrightarrow C = C' \text{ (bi-directional bridge)} \\ 0 & \Leftrightarrow C \subset C' \text{ (unidirectional bridge from C to C')} \\ -1 & \Leftrightarrow C \not\subset C' \text{ (there is no bridge)} \end{cases}$$

When observations which are covered by a node(a concept representing a category) C are included into the set of observations which are covered by a node C' in another perspective, a bridge from C(source node) to C'(target node) is established. If the extension of C' is also included in C, a bi-directional bridge is created. Notice that both set inclusion and set equality relations accept the application of the transitivity property (horizontally, among perspectives) similar to the vertical transitivity authorized by the specialization relation in a hierarchy. In addition, the specialization relation allows FORMVIEW to establish hidden bridges between children of a bridge's source node and a bridge's target node. Figure 2 illustrates two hierarchies following two perspectives and the bridges between them.

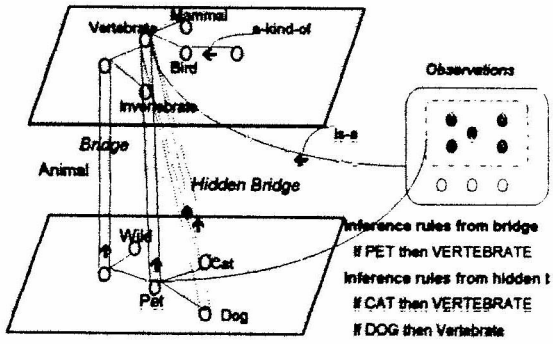

Figure 2 Multi-perspectives in CONFORT

4.2 The Utility Measure

As we have defined when we describe the main notions of the incremental concept formation algorithms, the construction of a hierarchy of concepts follows a quality criterion. FORMVIEW aims to construct such hierarchies privileging its prediction power. We will describe briefly the utility function used by FORMVIEW. This function is, like many of FORMVIEW predecessors, based on the work of Gluck and Corter on cognitive psychology[Gluck 85], who have defined a function to discover, within a hierarchical classification tree, the category more quickly remembered or the basic level category[Rosch 75]. Gluck and Corter's function, named *category utility*, allows to measure the inferential capacity of a category. They suggest that certain categories are preferred because they best facilitate predictions about new observations. Supposing observations are represented as sets of properties p_i, then Gluck and Corter's measure of category utility (CU) can be described as a trade-off between the expected number of features that can be correctly predicted about a member of a category C_k , and the proportion of the environment $P(C_k)$ to which those predictions apply: $P(C_k)E(\text{No. of correctly predicted } p_i|C_k)$.

For instance, little can be predicted about a highly general category like animals, but those properties that can be predicted (e.g. animate) apply to a large population. In contrast, many features can be predicted with near certainty about highly specific categories like robins, but these predictions are true of a relatively small population. A category of intermediate generality such as birds maximises the trade-off between the expected number of accurate predictions and the scope of their application.

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

Finally, category utility is defined as the increase in the expected number of properties that can be correctly predicted given knowledge of a category ($P(p_i|C)^2$)[1], over the expected number of correct predictions without such knowledge ($P(p_i)^2$). Gluck and Corter CU can be defined more formally as:

Definition 4

Be a set of entities E defined by observations, a hierarchy H built on E. The utility of a category C of H defining a probabilistic concept **CP** = {(j, υ(j), PD, PP)| j ∈ A} is :

$$UC(C) = \sum_k \sum_i P(C)[P(j_k=v_i|C)^2 - P(j_k=v_i)^2] \quad (a)$$

Where $v_i \in \upsilon(j)$ $(i \in N \mid 0 < i \leq \text{card}(\upsilon(j)))$, $j_k \in A$

By Bayes rule in *a*

$$UC(C) = \sum_k \sum_i P(j_k=v_i)P(j_k=v_i|C)\,P(C|j_k=v_i) - P(j_k=v_i)^2] \quad (b)$$

$P(C|j_k=v_i) \in PD$

$P(j_k=v_i|C) \in PP$

In COBWEB, Fisher changed this formula to compute the utility of a disjoint concept partition $P=\{C_1, C_2,...,C_n\}$ as the mean of the utility of each category of P. Formally:

Definition 5

The utility of a partition $P=\{ C_1, C_2,...,C_n \}$ of categories is :

$$UC(P) = \sum_k UC(C_k)/n$$

Where $(k \in N \mid 0 < k \leq n)$

Our approach requires some improvements in the formula defined above. First, we changed it to take into account the categorization goals. This was possible via the utilization of the semantic relevance of the properties defined in the GDN. In reality, we can verify that the utility of a category UC defined above (definition 2b) has a pondering factor $P(j_k=v_i)$ that represents relevance of each attribute. However, this relevance is exclusively *syntactic*(based on occurrence frequency) and we have changed it to take into consideration the semantic relevance expressed in the GDN. Therefore, the semantic utility of a category UC_S is defined as below.

Definition 7

The semantic utility of a category C defining a probabilistic concept **CP** with a set of properties $\Pi=\{ p_j \mid j \in N \}$ is :

$$UC_S(C) = \sum_j \Delta(p_j)P(C|p_j)P(p_j|C) - \sum_j P(C)P(p_j)^2$$

Where $\Delta(p_j)$ = (the semantic relevance of the property p_j + $P(p_j)$)

Similar to Fisher's COBWEB, FORMVIEW computes the utility of a partition of categories as the mean of semantic utilities of each partition category. However, it uses, during concept hierarchy construction from a particular perspective, additional information from other hierarchies which represent other perspectives. This is possible because, despite different hierarchical organization and different intensional definition, concepts between perspectives can have the same extension. These concepts are thus linked by bridges. In this

---

[1] Using a probability matching strategy, Gluck and Corter define that one can predict a property with probability $P(p_i|C_k)$ and this prediction will be correct with the same probability. Thus, E(No. of correctly predicted $p_i|C_k$) = $\sum_i P(p_i|C_k)^2$.

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

case, FORMVIEW's additional strategy is to determine a complete partition with categories of other hierarchies ; Those are target from bridges established from sources that are categories from the initial partition. For instance, given a partition P with a set of categories C_k, if a category C_i from P establishes a bridge with another category C'_i in another perspective, C'_i will make part of P, in consequence it will participate in computing the utility of this partition. Formally:

Definition 8

Be $P=\{ C_1, C_2,...,C_n \}$ a category partition named source partition. $P \supseteq H$ that represents a perspective pv ($\in$ PV) $H\uparrow pv$. Be $\mathbf{H} = \{H'_1\uparrow pv_1, H'_2\uparrow pv_2,...,H'_m\uparrow pv_m \}$ the set of hierarchies which represent others perspectives. We have $\mathbf{H}=\{ \{C'_{1k}, C'_{1k+1},..,C_{1card(H'_1\uparrow pv1)}\},\{C'_{2k}, C'_{2k+1},..,C_{2card(H'_2\uparrow pv2)}\},...,\{C'_{mk}, C'_{mk+1},..,C_{mcard(H'_m\uparrow pvm)} \}\}$. A new partition P_{com} called **complete partition**, including other categories of other existing perspectives is defined :

$$P_{com} = P \cup \{C'_{hk}\}(1 \leq h \leq m)\ (1 \leq k \leq card(H'_h\uparrow pv_h))$$

With $C'_{hk} \supseteq C_l\ (1 \leq l \leq n)$ that is $bridge(C_l, C'_{hk}) \geq 0$

It is important to point out that target categories, which will make part of the new partition, contribute to the category utility calculation only with properties that there are not present in the concept that represents the source category. In addition, if there are several target categories with the same property, FORMVIEW selects those having the highest predictability. Formally:

Definition 9

Be $P_{com} = \{ P_{source}, C\}$ a complete partition with a source partition P_{source} and a set of categories C which P_{source}'s categories establish of bridges. Pr_{source} is the set of properties of P_{source}'s categories and Pr_{crest} the set of properties of C. We define the set of **useful properties** P_{ut} of C to compute the utility of the partition P_{source} as :

$$P_{ut} = Pr_{source} \cup \{p |\ p \in \{Pr_{crest} - Pr_{source} \} \}$$

The rational behind these two latter strategies is that a complete partition and its useful properties define a complete concept, independent from a particular perspective. Thus, new entity classification is improved because a greater quantity of properties can be induced.

4.3 The Concept Formation Process

FORMVIEW has an initial phase of "data preparation" before the start of the concept formation procedure. In this phase, following the presentation of an observation O, it searches the relationship between properties of such an observation and other ones in the GDN. This can provoke modifications in the initial observation structure due to the insertion of new attribute-value pairs(properties). More precisely:

A complete observation OC^t for a perspective t ($\in$ PV) possesses the properties of the initial observation $O=\{p_1, p_2,...,p_m\}$ plus those inferred from the GDN ($OC^t = O \cup \{p^{inf}_k\}(1 \leq k \leq card(IP^t))$.

With $(p_i, p^{inf}_k) \in IP^t\ (1 \leq i \leq m)$.

From complete observations, FORMVIEW generates several hierarchies reflecting different perspectives or points of view according to GDN's categorization goals. Each GDN categorization goal determines a perspective to consider, then a specific hierarchy. The concept formation procedure of FORMVIEW is a hill climbing search for the best partition that can be generated from the application of the operators for hierarchy organization. More

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

precisely, to categorize a new observation OC^t among the sub-concepts of a given concept(partition P_0), FORMVIEW can :

- modify the partition P_0 incorporating OC^t in one of the P_0's categories and giving rise to the partition $P_a = \{C_1, ..., C_i, ..., C_r, ..., C_k\}$; With C_i, C_r, being the two best utilities to categorize OC^t that is to say $UCs(C_i) > UCs(C_r) > UCs_z(C_z)$ $(0 < z < k, z \neq i, r)$.
- modify P_0 creating a new concept C_c for receiving OC^t, giving rise to the partition P_c ($P_c = P_0 \cup \{C_c\}$)
- modify P_0 merging the two concepts C_i et C_r in one C_{ir}, giving rise to the partition P_f ($P_f = P_0 - \{C_i, C_r\} \cup \{C_{ir}\}$)
- modify P_0 spliting the C_i in its sub_concepts C_{il} $\{C_{il} | 1 \leq l \leq q\}$, giving rise to the partition P_s ($P_s = P_0 - \{C_i\} \cup \{C_{i1}, C_{i2}, ..., C_{iq}\}$.

The choice of the partition where to integrate OC^t will be that which optimizes the category utility *UCs*. However, notice that the partition used in this computing is the complete partition for each operation, that is to say, FORMVIEW takes into consideration the possible bridges between perspectives. Thus, we have the complete partitions P^{com}_0, P^{com}_a P^{com}_c, P^{com}_f, P^{com}_s. Notice also that in this computing FORMVIEW considers only the useful properties of each complete partition.

FORMVIEW

1. From the first observation O, to find out the complete observations for each categorization goal t (OC^t)
2. Create the roots R^t for each perceptive based on the complete observations ($R^t = OC^t$)
3. For each perspective t
   - 3.1 For the next observations O
     - 3.1.1 Compute the complete observation for perspective t (OC^t)
     - 3.1.2 PrincipalLoop (R^t, OC^t)

PrincipalLoop (C, O)

1. Incorporate (C, O)
2. Compute the partitions P^{com}_a, P^{com}_c, P^{com}_f, P^{com}_s, from the partition P_0 of C's sub-concepts
3. Choose the best partition PB; PB=max($UCs(P^{com}_a)$, $UCs(P^{com}_c)$, $UCs(P^{com}_f)$, $UCs(P^{com}_s)$)
4. Choose the best category CB from PB; CB=max($UCs(C_k)$) $(1 \leq k \leq card(PB))$
5. If $PB \neq PC_c$ then replace C by CB and return to 1

Incorporate(C, O)

1. Update the predictabilities and predicteveness of the C's properties
2. Insert the new observation in the list of observations covered by C
3. EstablishBridge(C, List of observations covered by C, O)
4. If bridge(C, C') ≥ 0 (C' $\in H\uparrow$pv which did not treat O yet) then
   - 4.1 Erase bridge

EstablishBridge(C, *LstC, O*)

1. For each $H\uparrow$pv that have already treated the observation O
   - 1.1 Descend $H\uparrow$pv comparing LstC with the list of observations of H's categories (*LstCurrent*).
   - 1.2 If *LstC - LstCurrent* = 0 then
     - build bi-directional bridge between C and current category of H
   - 1.3 Else If *LstC* $\supset$ *LstCurrent*
     - build unidirectional bridge from C to current category of H
   - 1.4 Else If *LstCurrent* $\supset$ *LstC*
     - build unidirectional bridge from current category of H to C

Table 2 . FORMVIEW's control structure

Another FORMVIEM feature is the management of bridges between perspectives. This procedure is carried out when an observation is incorporated in a node of a hierarchy that represents a certain perspective. At this moment, FORMVIEW descends other hierarchies which have already treated the current observation in order to compare the set of observations which are covered by the chosen node with those covered by nodes of hierarchies developed from other perspective. Thus, it can build bridges between perspectives or even undoes

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

unidirectional bridges from the chosen node to nodes in other taxonomies that did not treat the current observation.

The basic FORMVIEW's control structure is summarized in Table 2. There we can see the functions *FORMVIEW* and *PrincipalLoop*. We can also see the *Incorporate* and *EstablishBridge* functions which are responsible for management of bridges.

4.4 An Example of Multi-hierarchies Generation

We now consider an example of categorization in FORMVIEW. We have defined two perspectives. In order to facilitate comparisons one of these perspectives(perspective 1) defines a COBWEB-view, that is, we do not define anything in the GDN. Perspective 2 possesses a statement defining the relevance of the property *FoodType=Packaged* and its implication to *Character=Smart*. The tiny GDN and the observations used in this example[Martin 94] are described in Table 3.

Name	Found	FoodType	Mobility	Covering	Legs	Reproduction	Appears
Finch	Inside	Packaged	Flies	Feathers	Two	Prodeggs	Plain
Angelfish	Outside	Fresh	Swims	Scales	Zero	Prodeggs	Pretty
Macaw	Inside	Packaged	Flies	Feathers	Two	Fertilize	Plain
Hamster	Inside	Packaged	Walks	Hair	Four	Fertilize	Plain
Leopard	Outside	Fresh	Walks	Hair	Four	Prodeggs	Pretty
GoldFish	Inside	Packaged	Swims	Scales	Zero	Prodeggs	Pretty
Guppy	Inside	Packaged	Swims	Scales	Zero	Fertilizes	Plain
Pigeon	Outside	Fresh	Flies	Feathers	Two	Fertilize	Plain

GDN Objective 2 FoodType=Packaged ==> Character=Smart

Foodtype=Packaged's relevance = 1

Table 3 Animal observations and a tiny GDN

After 3 observations the generated hierarchies are the same ones. However, the fourth observation(*Hamster*) leads to a different hierarchy organization since from the perspective 2(objective 2), FORMVIEW generates a hierarchy that cluster *Hamster*, *Finch* and *Macaw* because they share the *FoodType=packaged* and *Character=Smart* properties(Figure 3a). These properties, defined in the GDN with high relevance, accentuate the relation between observations in which they participate. Figure 3b shows the hierarchies generated from all observations of Table 3. There, we can clearly realize the two different clusters which reflect the importance of the GDN's specified properties.

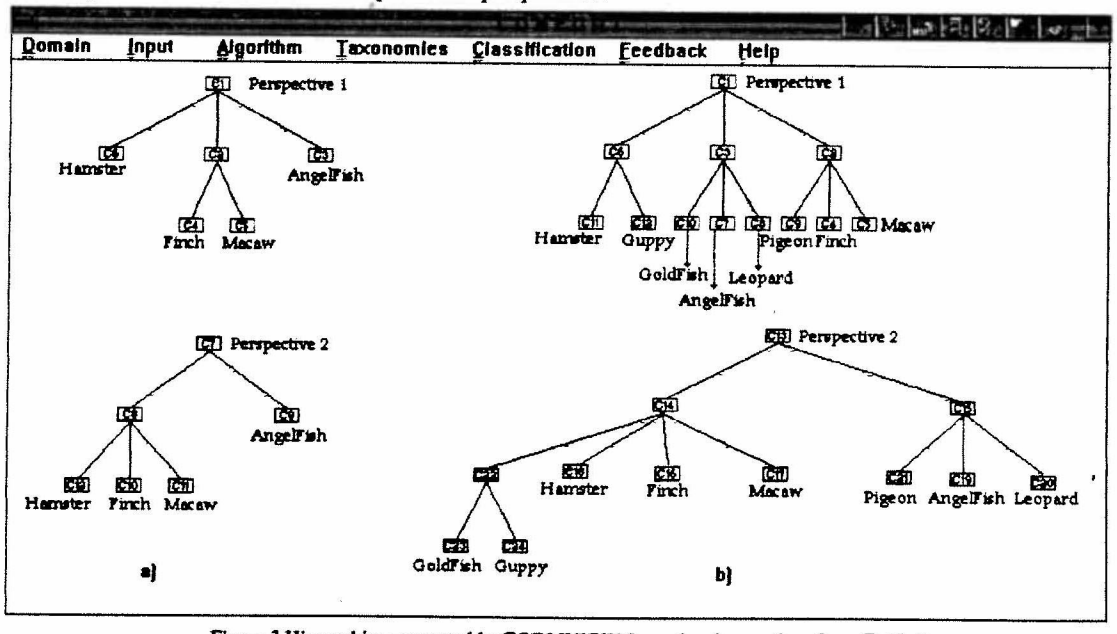

Figure 3 Hierarchies generated by FORMVIEW from the observations from Table 3

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

5 FROM PROBABILISTIC CONCEPTS TO FRAMES

As we have mentioned a probabilistic concept tree has information about frequency of both a concept's properties and observations. In the passage from probabilistic concepts to an abstract representation, we keep this information in order to give more power to the representation as well as providing useful heuristics for classification of future instances. More precisely, the frequencies of occurrence of properties and their probabilities are used for creation of descriptive facets which define default values, and sufficient and necessary properties. The passage from probabilistic concepts to frames can be seen in two dimensions. The horizontal dimension consists of the definition of properties(slots and descriptive facets) which will compose the frame. The vertical dimension consists of the definition of the levels of the frame hierarchies, that is, what frames should be maintained in the hierarchy and what do not. Our heuristics to define vertical and horizontal dimensions are based on psychological findings that account for the probabilistic character of concepts [Smith 81],[Rosch 76],[Fisher 88].

5.1 The vertical dimension

The vertical dimension consists of searching in the probabilistic concept hierarchies, for concepts which do not have a *large importance*(in the sense *predictive power*) and, therefore, do not justify their existence. Thus, in a concept hierarchy, if there is a level having a concept that contains the predictive power less or equal than its parent level, it does not have an importance that justifies its passage from a probabilistic concept to a frame.

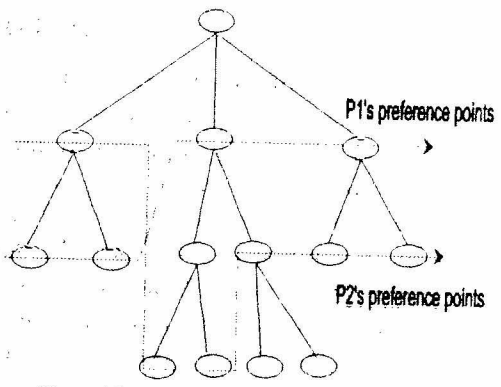

Figure 4 Preference points for two properties in a hypothetical tree

To determine the better prediction level and which probabilistic concepts should be transformed in frames, we use a strategy based on Fisher's work [Fisher 89] on simplification of probabilistic concept hierarchies. He has defined the notion of past preferences which define that in a probabilistic concept hierarchy, we can identify preference points for each attribute of a probabilistic concept. A preference point determines, for an attribute, the level of the hierarchy where this attribute can have its value better predict. In other words, we can predict, with reliable force, a value for this attribute in this level of the hierarchy, without descend more specific levels in the hierarchy. Figure 4 illustrates this strategy [Fisher 95]. The P1's preference points determine the hierarchy level where P1's value can be predicted reliably.

5.1.1 The definition of preference points

The basic assumption for determine preference points is that prediction of a missing attribute should occur at a node(probabilistic concept) that historically has facilitated the greatest number of correct predictions[Fisher 89]. The determination of this node is reached via counts which are updated during the categorization process. Since the categorization process consists in a descending research in the probabilistic concept hierarchy, every observation attribute should be compared with those of the probabilistic concept in each hierarchy level; If the observation's value is equal to the node's most frequent value then such a value would have been correctly predicted at this point. For each attribute and node, a count is maintained of the number of times the attribute would be correctly predicted during

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

the categorization. Another count is used to keep the number of times the attribute would be predicted at the node's children. The preference point is one where the node's attribute count is greater than its children's counts.

We can extend the preference point notion to the preference concept. In a preference concept, all its attributes are preference points. This means that the preference concept's children are not important in the sense that attribute prediction can be realized in a superior node. Therefore, CONFORT does not translate concept preference's children to frames.

5.2 The horizontal dimension

The horizontal dimension consists of defining the frame's properties. We use descriptive facets to improve its characterization, giving an additional power to this kind of representation. In fact, the probabilistic representation is characterized by the storing of frequencies of attributes, values and concepts. Therefore, our idea is to profit from this information to generate descriptive facets which will be a useful heuristic for classifications of the next observations.

The properties represented in frames are defined by slots and facets. The slots represent the probabilistic concept's attributes. The facets characterize these slots determining each slot's value and domain as well as specific facets to define the sufficient or necessary nature of a slot and its default values.

5.2.1 The definition of sufficient properties

Sufficient properties are those that have predicteveness equal to 1. For instance, if the predictiveness of a property *p(*with attribute *a* and value *v)* for a category C (i.e. P(C|*p*)) is 1 then existence of *p* is sufficient to classify another observation in C. In other words, if CONFORT has identified that all collected observations that have property p are covered by C, then, as a result, CONFORT creates a facet for the attribute *a* defining the sufficiency of *v* for the classification of other instances. The sufficiency of properties can be verified for a set of properties P={ p1,p2,p3,..pn }. In this case, predicteveness should be computed from P(C|p1,p2,p3,..,pn)

5.2.2 The definition of necessary properties

Necessary properties are those that have predictability equal to 1. For instance, if the predictability of a property *p(*with attribute *a* and value *v)* for a category C (i.e. P(p|C)) is 1 then existence of *p* is necessary to classify another observation in C. In other words, being informed that all the observations covered by C always have the property p, CONFORT creates a facet for the attribute *a* defining the necessity of *v* for the classification of other instances. Similar to the definiton of sufficient properties, necessary properties can be verified for a set of properties P={ p1,p2,p3,..p_n } and, therefore, predictabilities should be computed from P(p1,p2,p3,..p_n|C).

5.2.3 The definition of default values

The determination of default values is another feature of our model. In fact, few works have been developed on default value assignment. Specifying a default value to a frame slot means that such value is generally true to this slot and consequently can be inherited by the associated sub-frames and instances. In CONFORT, when probabilistic concepts are translated to frames, default values are defined from attribute-value's predictability and predicteveness. The basic idea is to consider default values as those having predictability and predicteveness greater than a user defined contextual threshold.

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, F. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

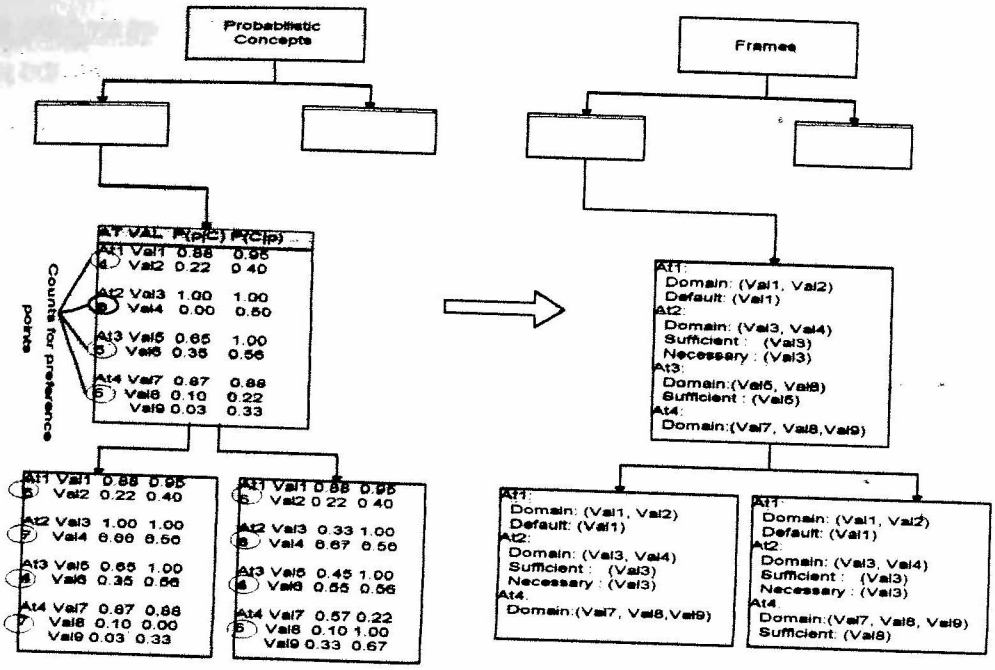

Figure 5 Passage from probabilistic concept to frame

Figure 5 exemplifies the passage from probabilistic concepts to frames. We can see that attribute *At3* was not transformed to the frame representation because preference points of its super-frame were better. Necessary and sufficient properties are defined from probabilistic concept's predictabilities(P(p|C)) and predictiveness(P(C|p)), respectively. The *At1* attribute's *Val1* was determined a default value since it has predictability and predictiveness greater than 0.85 (a user defined parameter).

6 CONCLUSION

We defined a software architecture for the incremental construction of concept hierarchies. Our objective was to design a tool to assist in the conception of frame hierarchies while maintaining the nature of a knowledge acquisition tool. This architecture has a learning algorithm which generates multiple probabilistic concept hierarchies representing different perspectives. At last, we showed the transformation of the probabilistic representation into a frame-based one.

ACKNOWLEDGMENTS

The first author's work was supported by the Brazilian Research Council(CNPQ) grant 200522/93-0, University of Fortaleza(UNIFOR) and Data Processing of Ceara (SEPROCE). Comments by Phil Smith helped improve the style and understandability of the paper.

REFERENCES

[Aguirre 89] Aguirre, J.L.: *Construction automatique de taxonomies à partir d'exemples dans un modèle de connaissances par objets*. Thèse de doctorat. Laboratoire ARTEMIS/IMAG, Grenoble, 1989.

[Barsalou 83] Barsalou, L.W.: *Ad Hoc Categories*. Memory and Cognition,11(3),1983.

[Brachman 85] Brachman, R.J., Schmolze, J.G.: *An Overview of the KL-ONE Knowledge Representation System*. Cognitive Science, 9(2): 171-216, 1985.

[Faucher 91] Faucher, C.: *Elaboration d'un langage extensible fondé sur les schemas le langage objlog+*. Thèse de doctorat, Université de Droit, d'Economie et des Sciences d'Aix-Marseille, 1991.

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

[Fisher 87] Fisher, D.H.: *Knowledge Acquisition via Incremental Conceptual Learning*. Machine Learning, vol 2, numero 2, 1987.

[Fisher 88] Fisher, D. H. *A Computational Account of Basic Level and Typicality Effects*. Proc. of Seventh National Conference on Artificial Intelligence, 1988.

[Fisher 89] Fisher, D. H.: *Noise-Tolerant Conceptual Clustering*. IJCAI, 1989.

[Fisher 93] Fisher, D., Yoo, J.: *Categorization, Concept Learning, and Problem-Solving: A Unifying View*. The Psychology of Learning and Motivation. Vol 29. 1993.

[Fisher 95] Fisher, D.: *Iterative Optimization and Simplification of Hierarchical Clusterings. Technical Report CS-95-01. Depto of Computer Science, Vanderbilt University, Nashville, TN, 1995.* .

[Gennari 89] Gennari, J.H, Langley, P., Fisher, D.: *Models of Incremental Concept Formation*. Artificial Intelligence, 40, 1989.

[Gluck 85] Gluck, M. A., Corter, J.E.: *Information, uncertainty, and the utility of categories*. Proceedings of the Seventh Annual Conference of the Cognitive Science Society. Irvine, CA, Lawrence Erlbaum Associates, 1985.

[Hampton 93] Hampton, J. Dubois, D.: Psychological Models of Concepts: Introduction. In Categories and Concepts: Theoretical Views and Inductive Data Analysis. Academic Press, 1993.

[Marino 90] Marino, O., Rechenmann, F., Uvietta, P.: Multiple Perspectives and Classification Mechanism in Object-Oriented Representation. Cognitiva 90, 1990.

[Martin 94] Martin, J., Bilman, D.: Acquiring and Combining Overlapping Concepts. Machine Learning, 16, 121-155, 1994.

[Michalsky 86] Michalsky, R., Carbonnel, J., Mitchell, T.: *Machine Learning, An Intelligence Approach*. Vol II. Morgan Kaufmann, CA. 1986.

[Minsky 75] Minsky, M. *A framework for representing knowledge*. The psychology of Computational Vision, P.H.Winston(ed), McGrawHill, pp.156-189, 1975.

[Napoli 90] Napoli, A., Ducournau, R., Laurenço, C.: An Approach to Object-Oriented Classification. Proc. 1st ASIS SIG/CR Classification Research Workshop, Toronto, 1990.

[Rechenman 88] Rechenman, F.: *SHIRKA, un système de gestion de bases de connaissances centrées objet*. Manuel d'utilisation, INRIA/ARTEMIS, 1988.

[Reich 94] Reich, Y.: *Macro and Micro Perspectives of Multistrategy Learning*. In Michalsky and Tecuci(Eds), Machine Learning: A Multistrategy Approach. Vol. IV. Morgan Kauffmann, 1994.

[Rosch 75] Rosch, E., Mervis, C.: *Family Resemblances: studies in the internal structure of categories*. Cognitivie Psychology 7, 1975.

[Seifert 89] Seifert, C.: *A Retrieval Model Using Feature Selection*. Proc. of the Sixth International Workshop on Machine Learning. Morgan Kauffmann. 1989.

[Smith 81] Smith, E.E, Medin, D.L.: *Categories and Concepts*. Library of Congress Cataloging in Publication Data. Cognitive Science series 4, 1981.

[Thaise 91] Thaise: *L'approche logique de l'intelligence artificiel*. Tome 4: De l'apprentissage artificiel aux frontières de l'IA. Chapitre 1: Apprentissage Artificiel, 1991.

[Vasco 95a] Vasco, J.J.F, Faucher, C., Chouraqui, E.: *Incremental Concept Formation in an Object-Oriented Language*. Scandinavian Conference in Artificial Intelligence, Trondheim, 1995.

[Vasco 95b] Vasco, J.J.F., Faucher, C., Chouraqui, E.: Knowledge Acquisition Based on Concept Formation *Using a Multi-Perspective Representation. Florida Artificial Intelligence Research Symposium FLAIRS/95*. 1995.

ISSN: 2324-9773

Furtado-Vasco, J., Faucher, C., Chouraqui, E., Chouraqui, E., & Chouraqui, E. (1995). CONSTRUCTION OF FRAME HIERARCHIES USING MACHINE LEARNING. 6th ASIS SIG/CR Classification Research Workshop, 37-52. doi:10.7152/acro.v6i1.12660

ISSN: 2324-9773