

# A Markov Mesh Modelling on Uncertain Galois Lattice: Classification in Terminology

*Djamel Bouchaffra\* and Jean Guy Meunier\*\**

\* Centre of Excellence for Document Analysis and Recognition (CEDAR)  
State University of New York at Buffalo,  
Amherst, New York 14228-2567 USA

\*\* Laboratoire d'Analyse Cognitive de l'Information (LANCI)  
Université du Québec à Montréal, C.P. 8888, succ. centre-Ville,  
Montréal, Canada, H3C 3P8  
E-mails: bouchaff@pluton.lanci.uqam.ca ; meunier.jean.guy@uqam.ca

*Abstract:* Because the need to classify objects with respect to uncertain properties is increasing, it is important to seek an appropriate generalization of a Galois lattice structure. We aim to analyze an uncertain (probabilistic) Galois lattice structure. One of our interests consists to determine a complete (total) order within the objects set generated by the uncertain Galois lattice structure. Afterwards, we classify the objects according to the generation power criterion. This classification is set up using a Markov Mesh (MM) model when considering the uncertain Galois lattice as a neighborhood system. As an application, we hierarchically classify textual fragments according to their lexical properties. This operation produces a set of structured contexts (or fragments) that enables us to weight and find out the presence of different contexts of use assigned to each term. This set of contexts may aid the terminologist to capture the meaning of the terms studied.

Keywords: Galois lattice ; Uncertain Galois lattice ; Markov Mesh ; Terminological classifier.

**Summary:** Le besoin de classer des objets qui répondent *plus ou moins* à des critères faiblement définis, vagues ou imprécis se fait de plus en plus ressentir. Il nous semble important d'analyser la structure de treillis de Galois dans le cas de critères imprécis. L'un de nos intérêts est non seulement de définir un treillis de Galois dans un contexte d'imprécision (treillis incertain) mais aussi d'établir un ordre total dans l'ensemble des objets (ou des propriétés). Ensuite, nous classifions les objets par rapport à la puissance de génération. Cette classification utilise le concept de maillage propre aux champs de voisinage dont le système de voisinage est le treillis de Galois incertain. Finalement, nous appliquons le modèle pour classer hiérarchiquement des fragments textuels selon leurs propriétés lexicales. Cette opération produit un ensemble de contextes (ou fragments) ordonné nous permettant ainsi de découvrir les différents contextes d'utilisation associés à chaque terme. Ceci peut aider le terminologue à appréhender les différents sens qui sont véhiculés par les mots étudiés.

Keywords: Treillis de Galois ; Treillis de Galois incertain ; Maillage de Markov ; Classifieur terminologique.

## Introduction

One of the most interesting tasks of a taxonomist is the analysis of a natural hierarchical overlapping classification where a set of objects can be generated from several imprecise or uncertain classes. Very often, this generation process is not binary and one wishes to assign weights to the membership of an object in a class. This is encountered for example in terminology where one wants to measure and find out various uses and meanings of a word in a corpus. For instance, comparing the context of use assigned to the word *code* in the sentence "the programmer made a syntax error in the C++ code language" and the sentence "the *civil code* does not allow marriage under 14 years of age". When the corpus length reaches a hundred thousand pages, selecting contexts for words might be a big problem.

When comparing these contexts (or sets of textual fragments), one is in fact analyzing paradigmatic semantic relations of words in a textual corpus. These relations are founded on two semantic theses: the cognitive linguist says that contexts are similar if they associate the same words for the same ideas (or concepts) (Leacock, 1982). The linguistic thesis (Boutier, 1987; Boutier, F., Cova, M., Abeille, A.

We have been exploring the emergent classification strategies such as neural networks (Meunier J.G., et G. 1995) and Markovian field models (Bouchaffra D., Meunier J.G. 1995a,b). These models provide a plastic and dynamic classification. They are both dynamic because they are sensitive to new information and plastic or adaptive because they capture significant events. In this paper, we present a context mesh modelling on an uncertain Gallois lattice neighborhood and we shall apply it as an aid to linguistic classification.

### ART classifier

A neural network such as ART (Carpenter G., Grossberg S. 1987) realizes a classification of fragments that is used to extract a set of contexts described by words. This model is stable and plastic. When a new fragment is added, ART system classifies it in the set of classes formed by the means of a competition process undertaken during the training phase. In other words, ART system does not need to reclassify already classified objects. On the ART output (fig. 1), one can see the word *code* in different contexts, it has different meanings. For instance in the class 141 *code* is used in the context of CRIMINAL CODE and in class 136 it is used in CIVIL CODE. In analysing more carefully these results, one may have the intuition that there is also some imprecise hierarchical relations between words in each class and between the classes themselves. For example in the class 136 the word *criminel* is more important than the word *paix*. Some precise hierarchy may exist between the classes 343 and 110. These questions explain our need to propose in the following sections uncertain hierarchical classification.

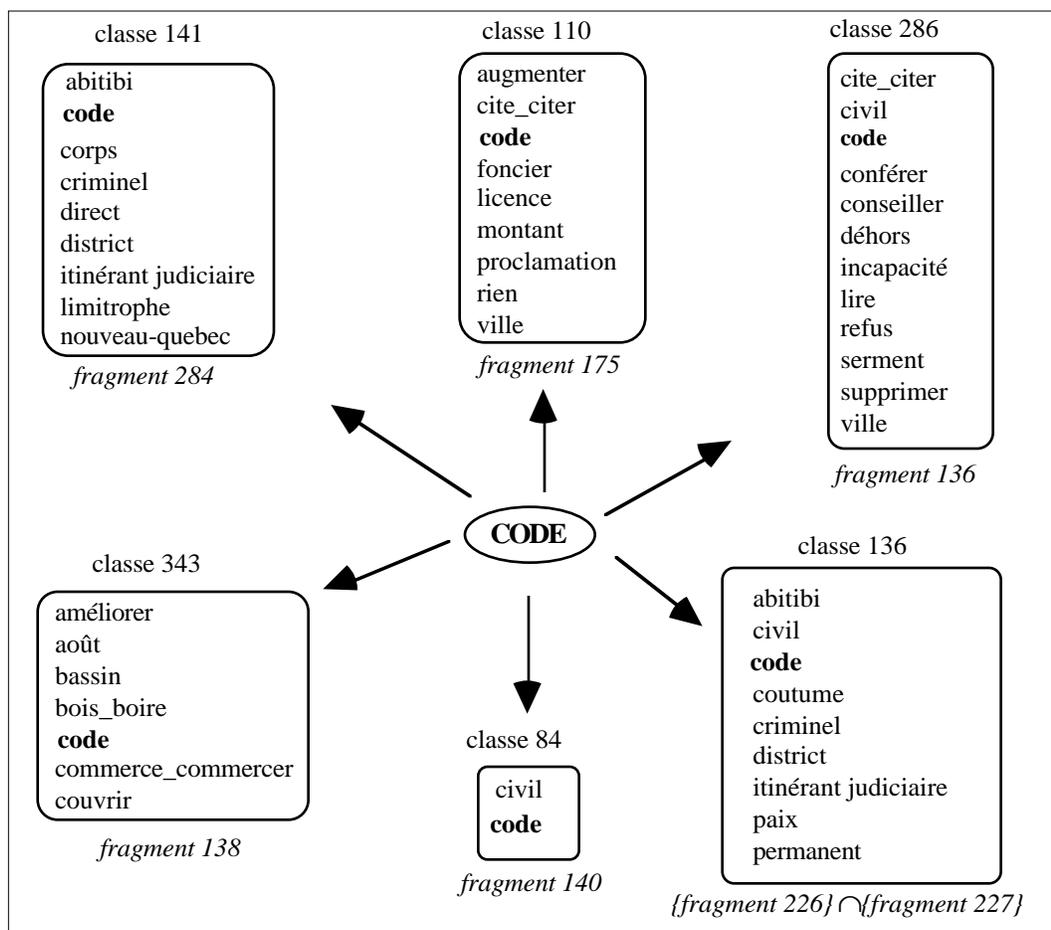


Fig.1. Different contexts of a word.

### hard binary relation

Let's consider a hard context  $(\mathcal{O}, \mathcal{E}, \mathcal{R}_h)$  where  $\mathcal{O}$  is the set of objects,  $\mathcal{E}$  the set of properties and  $\mathcal{R}_h$  a binary relation between  $\mathcal{O}$  and  $\mathcal{E}$ . The vocable "hard" means that there is an absence of uncertainty added in the relation between an object and a property. In other words, one has the entire association

he lattice extracted from the context  $(\mathcal{O}, \mathcal{E}, \mathcal{R}_h)$  is a set of couples  $(X, Y)$  where  $X$  and  $Y$  are two sets  $\mathcal{O}$  and  $\mathcal{E}$  respectively to the power sets  $P(\mathcal{O})$  and  $P(\mathcal{E})$ . Each element of a hard Galois lattice derived from  $(\mathcal{O}, \mathcal{E}, \mathcal{R}_h)$  must be complete with respect to  $\mathcal{R}_h$ : this means that there exists two functions  $f$  and  $g$  that:

$f(X) = \{y \in \mathcal{E} / \forall x \in X: x \mathcal{R}_h y\}$ ,  $X = g(Y) = \{x \in \mathcal{O} / \forall y \in Y: x \mathcal{R}_h y\}$ . The couple of functions  $(f, g)$  is called a Galois connection (Godin R., Missaoui R. 1994) between the two power sets  $P(\mathcal{O})$  and  $P(\mathcal{E})$ . If  $Z = X \cup Y$  and  $\mathcal{E}$  is the set of edges directed from  $X$  to  $Y$  then the couple  $(Z, \mathcal{E})$  represents a complete bipartite graph. To construct the hard Galois lattice structure, one has to define a hard partial order on the cartesian product  $P(\mathcal{O})$  and  $P(\mathcal{E})$ .

### A hard Galois lattice structure

**Definition 2.** If  $C_1 = (X_1, Y_1)$  and  $C_2 = (X_2, Y_2)$  are two elements of a hard Galois lattice, then  $C_2 \mathcal{R}_h C_1$  ( $C_2 \leq C_1$ ) iff (if and only if)  $X_2 \supset X_1$ . One can notice that:  $X_2 \supset X_1 \Leftrightarrow Y_1 \supset Y_2$ .

**Remark 1.** In the definition 2., we focus on objects generation rather than on properties generation. The definition can be given with respect to properties generation. The partial order defined on  $(\mathcal{O}, \mathcal{P}(\mathcal{E}))$  enables us to construct the graph in the sense that there is an edge from  $C_1$  to  $C_2$  if  $C_2 \mathcal{R}_h C_1$  ( $C_2 \leq C_1$ ). One can say that  $C_1$  is a parent of  $C_2$  and  $C_1$  is covered<sup>2</sup> by  $C_2$  according to objects generation.

**Lemma 2.** Let  $(\mathcal{O}, \mathcal{E}, \mathcal{R}_h)$  be a hard context, the infimum and supremum of any subset of the complete lattice  $(\mathcal{L}; \mathcal{R}_h)$  are defined as:

$$\begin{aligned} \bigcap (X_i, Y_i) &= (g.f(\bigcup X_i), \bigcap Y_i) \text{ where } i \in I (\text{finite}) \\ \bigcup (X_i, Y_i) &= (\bigcap X_i, f.g(\bigcup Y_i)). \end{aligned}$$

**cf.** see (Barbut M., Monjardet B. 1970)

Once these supremum and infimum of any subset of the lattice have been computed, one can construct the lattice structure completely. There exists many algorithms for the generation of the lattice elements (Godin J.P., 1986).

### Uncertain binary relation

In this section our aim is to analyse the Galois structure when the relation between objects and properties is uncertain. Many uncertainty measures may be used in order to express the relationship between objects and properties (Nilsson N.J., 1986 ; Zadeh L.A., 1978 ; Schaffer G., 1976). However, in this paper we focus on the probability measure over the sample space  $\mathcal{O}$ . Our aim is to choose among a family a unique Galois lattice that best (according to some sense) describes the input uncertain domain knowledge that a user has in mind. This knowledge is materialized by an input relation table.

**Definition 3.** The notation  $x \mathcal{R}_{uy}$  is used to outline that the binary relation between an object  $x \in \mathcal{O}$  and a property  $y \in \mathcal{E}$  is uncertain and measured by means of a probability function over the sample space  $\mathcal{O}$ . A probability value  $\mu_i$  is assigned to each related couple  $(x_i, y_i)$  contained in the cartesian product  $(\mathcal{O} * \mathcal{E})$ . In fact, one can write:  $\mu\{x \in \mathcal{O} ; x \mathcal{R}_{uy}\} = 1 \Leftrightarrow \{x \in \mathcal{O} ; x \mathcal{R}_{uy}\} = \{x \in \mathcal{O} ; x \mathcal{R}_h y\}$ ,

### A probabilistic Galois lattice structure

**definition 4.** The concept of complete couples can be extended and transformed into the research of an optimal configuration (X,Y) (among all possible) that maximizes the probability of their mutual relationship. The problem of determining complete couples with fixed cardinals can be written:

$$\arg(\max_{\{X,Y\}} \mu\{X \mathfrak{R}_h Y\})$$

In graph theory, this problem is equivalent to the research of complete bipartite graph  $K_{pq}$  where  $p = |X|$  and  $q = |Y|$  with maximum weight among  $C_n^p \cdot C_m^q$  possibilities ( $n = |\mathcal{O}|$  and  $m = |\mathcal{P}|$ ).

**definition 5.** The probability assigned to a configuration (X,Y) where  $|X| = p$  and  $|Y| = q$  can be defined

$$\mu(X \mathfrak{R}_h Y) = \sum \sum \mu(x_i \mathfrak{R}_h y_j) / p * q$$

The probabilistic Galois lattice  $\mathfrak{L}$  is the set of all complete couples (with respect to the sense of the definition 4.) with the following partial order:

$$C_2 = ((X_1, \mu_1) ; (Y_1, \mu_2)) \text{ and } C_1 = ((X_2, \mu_3) ; (Y_2, \mu_4)) \quad C_2 \mathbf{R}_u C_1 \Leftrightarrow X_2 = \arg(\max_{X_i} \{\mu\{X_1 \subset X_i\}\})$$

to compute  $\mu\{X_2 \supset X_1\}$  when  $\mu\{X_1\} = \mu_1$  and  $\mu\{X_2\} = \mu_2$  are known. We have supposed that the probability  $\mu\{X_2 \supset X_1\}$  reflects what we normally mean by the certainty of the rule *if  $X_2$  then  $X_1$* , where  $(X_i)_{i=1,2}$  are considered as two sentences<sup>4</sup>. Hence, one can determine bounds to this probability.

**lemma 3.** If  $\mu(X_1)$  and  $\mu(X_2)$  are the two probabilities assigned respectively to the set  $X_1$  and  $X_2$  then the probability assigned to  $\{X_2 \supset X_1\}$  is bounded, it can be written:

$$\mu(X_1) \leq \mu\{X_2 \supset X_1\} \leq \text{Min} \{1, \mu(X_1) - \mu(X_2) + 1\} \quad (1)$$

cf. see (Nilsson N.J. 1986 ; Bouchaffra D. 1993).

In order to construct the probabilistic Galois lattice, one has to choose a probability value from the relations (1). This latter probability value depends on the probability values  $\mu_1$  and  $\mu_3$  associated respectively to the sets  $X_1$  and  $X_2$ .

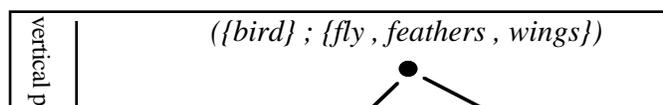
**proposition 1.** There is a nonunique ordered set which describes the inherent lattice structure defining formal groupings and probabilistic relationships among the objects and their properties.

cf. This is due to the fact that the probability assigned to  $\{X_2 \supset X_1\}$  is not unique.

A unique solution can be obtained using the entropy concept or the projection approximation method. Among all possible solutions, one selects that solution with maximum entropy (Cheeseman P., 1986).

### A complete order within the objects set

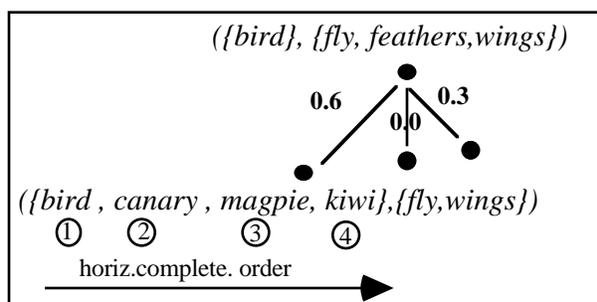
Because of the absence of a partial order within the set of objects (or properties) in the hard Galois structure [Fig. 2.], we are investigating the possibility of ordering the elements contained in X and Y. The probabilistic Galois structure enables to define a partial order in the set of objects or in the set of properties. It may be very helpful in terminology when one wants to approach the synonymy concept which is one of the paths leading to semantics.



**Definition 2.** The partial order  $\leq$  is a complete order and the set  $(X, \leq)$  is said to be completely and well ordered.

**Proof.** This order is complete because in the rational numbers contained in the set  $[0..1]$  a natural order is induced, this order is *complete*. It is a *well order* since any subset  $A$  ( $A \neq \emptyset$ ) contained in  $X$  has a *minimal element*.

Figure 3 indicates the total order defined on the set  $\{\text{bird, canary, magpie, kiwi}\}$  at level 2. The objects of the lattice are ordered with respect to the set of properties: *it is a horizontal complete order*.



**Fig. 3.** A horizontal complete order in a probabilistic Galois-lattice.

**Definition 3.** A relation measure within the objects set

Many measures can be performed in order to express the uncertain relationship within the objects set with respect to the properties set  $Y$  ( $|Y| = q$ ). The following measure is being experimented:

$$\mu\{x \text{ in } X ; x \mathfrak{R}_u Y\} = \sum \alpha_j \mu\{x \mathfrak{R}_u y_j\}$$

where  $\alpha_j$  are weights that may privilege some specific properties.

**The generation power criterion and the MM concept**

The Markov Mesh concept which is a subclass of Markovian Random Fields class enables us to classify objects with respect to the criterion *power of generation*. The unique lattice determined previously presents the neighborhood system (Besag J., 1974 ; Bouchaffra D., Meunier J.G., 1995 (a,b) ; Rushin R.L., Kusuoka S., 1993). In order to analyse and approach the hyponymy and hyperonymy inological concepts, we defined the following generation power.

**Definition 7.** The generation power assigned to a node  $s_i$  of a Galois lattice can be written as:

$$\gamma_{s_i} = \sum_{j \in \text{succ}^1(i)} \mu_{s_i s_j} \cdot \left( \frac{|Y_i| - |Y_j|}{|X_j| - |X_i|} \right),$$

where  $|\cdot|$  is the cardinal symbol and  $\text{succ}^1$  stands the first successors of the node  $i$ .

**Definition 8.**  $\gamma$  is a Markov Mesh with respect to a lattice neighborhood system  $L$  if :

$$\mu\{\gamma = \omega\} > 0 \quad \forall \omega \in \Omega \quad \text{and} \quad \text{Prob}\{\gamma_{s_i} / \gamma_{s_j}, s_j \in A_i\} = \text{Prob}\{\gamma_{s_i} / \gamma_{s_j}, s_j \in B_{s_i} \text{ and } B_{s_i} \subset A_{s_i}\}.$$

**Application to terminology**

In this study the uncertain binary relation is the relative cooccurrence frequency of a word in a fragment. An object is a fragment and the properties are specific words describing this fragment. Our system MUGAL is divided into three main modules. The first one selects words (according to various criteria, e.g., functional words, hapax, probabilistic distribution, discrimination indices) from the

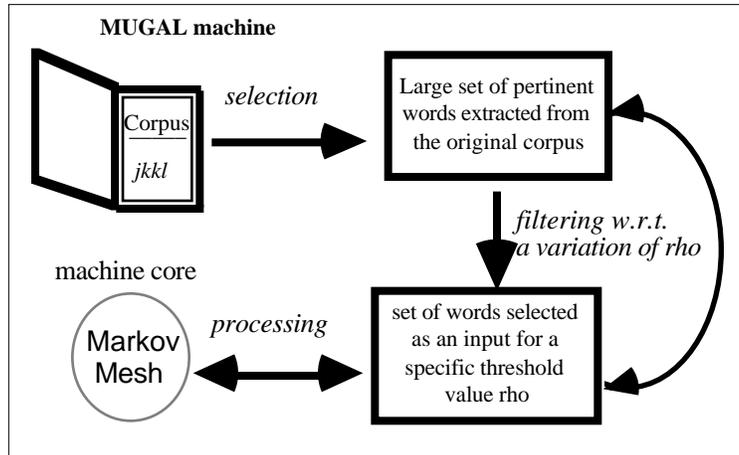


Fig 4. The MUGAL system architecture.

## Conclusion

We have attempted to present a generalization of the Galois lattice structure using the probability measure. It expresses the uncertainty between objects and properties. The total order proposed within the objects (or properties) seems to be very necessary and natural when dealing with uncertainty. The unique Galois lattice classifying fragments according to their lexical properties helps terminologists to find out different meanings of words. This lattice is inherent to the Markov Mesh module which analyses hyponymy and hyperonymy concepts.

## REFERENCES

- Butt M., Monjardet B. (1970). *Ordre et Classification*. Algèbre et Combinatoire. Tome II, Hachette, Paris.
- Decri, J.P., collab. (1973). *La Taxinomie*. Vol. I ; l'Analyse des correspondances. Vol. II, Dunod, Paris.
- de G J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *J. Roy. Statist. Soc.*, Series A, vol. 36, p. 192-236, .
- de G J.P. (1986). Calcul pratique du treillis de Galois d'une correspondance. *Mathématique et Sciences Humaines* 96 p. 31-47.
- Deffra D., Meunier J.G., (1995a). A Markov Random Field Approach to Automatic Information Retrieval. *Proceedings of the Third International Conference on Document Analysis and Recognition, ICDAR'95*, August 14-16, Montréal, CA, published by IEEE Society.
- Deffra D., Meunier J.G., (1995b). A Thematic Knowledge Extraction Modelling through a Markovian Random Field Approach. *6th International DEXA Conference and Workshop on Database and Expert Systems Applications, DEXA'95*, Sept 19-22, London, UK.
- Deffra D. (1993). Theory and Algorithms for analysing the consistent Region in Probabilistic Logic. *An International Journal of Computers & Mathematics with applications*, vol. 25, No. 3, February, edit. Ervin Y. Zeman, Published by Pergamon Press.
- Deffra D., Grossberg S. (1987). A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine. *Computer Vision, Graphics and Image processing*, 37, p. 54-115.
- Deffra D., Zeman P. (1986). Probabilistic vs. Fuzzy Reasoning. In *Uncertainty in AI*. L.N. Kanal and J.F. Lemmer, Elsevier Science publishers, New York, N.Y., p. 85-102.
- Deffra D., Hanks P. (1990). Word association norms, mutual information and lexicography. *Computational Linguistics*, Vol. 16, p. 22-29.
- Deffra D., Kusuoka S. (1993). Statistical Mechanics and Fractals. *Lecture notes in mathematics*, Springer Verlag, .
- Deffra D., Missaoui R. (1994). An Incremental Concept Formation Approach for Learning from Databases. *Artificial Intelligence* 133 pp. 387-419, Elsevier.
- Deffra D. (1983). *Semantics and Cognition*. Cambridge, MIT Press.
- Deffra D., Salem A. (1994). *Statistique Textuelle*. Dunod, Paris.
- Deffra D. (1992). Discrimination through the regularized nearest cluster method. In *Computational Statistics*, Eds. Y. Ge, Y. Whittaker J., Eds, Physica Verlag, Heidelberg, p. 103-118.
- Deffra D., Meunier J.G., Nault G. (1995). ART and computer text analysis. Rapport interne de recherche LANCI.
- Deffra D., N.J. (1986). Probabilistic Logic. *Artificial Intelligence* 28, p. 71-87, Elsevier Science Publishers B.V., North-Holland.