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

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

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

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

s clés: Treillis de Galois ; Treillis de Galois incertain ; Maillage de Markov ; Classifieur inologique.

ntroduction

ne of the most interesting tasks of a taxonomist is the analysis of a natural hierarchical overlapping ification where a set of objects can be generated from several imprecise or uncertain classes. Very ı, this generation process is not binary and one wishes to assign weights to the membership of an ct 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 e word *code* in the sentence "the programmer made a syntax error in the C++ code language" and the ence "the *civil code* does not allow marriage under 14 years of age". When the corpus length reaches e hundred thousand pages, selecting contexts for words might be a big problem.
ı comparing these contexts (or sets of textual fragments), one is in fact analyzing paradigmatic semantic ions of words in a textual corpus. These relations are founded on two semantic theses: the cognitive s says that contexts are similar if they associate the same words for the same ideas (or concepts) essed (Jackendoff R. 1983). The linguistic thesis (Pottier B. 1987 ; Rastier F., Cavazza M., Abeille A.

/e have been exploring the emergent classification strategies such as neural networks (Meunier J.G., .t G. 1995) and Markovian field models (Bouchaffra D., Meunier J.G. 1995a,b). These models ide a plastic and dynamic classification. They are both dynamic because they are sentive to new mation and plastic or adaptive because they capture significant events. In this paper, we present a ‹ov mesh modelling on an uncertain Gallois lattice neighborhood and we shall apply it as an aid to inological classification .

\RT classifier

. neural network such as ART (Carpenter G., Grossberg S. 1987) realizes a classification of fragments is used to extract a set of contexts described by words. This model is stable and plastic. When a new ›rn is added, ART system classifies it in the set of classes formed by the means of a competition process ›rtaken during the training phase. In otherwords, ART system does not need to classify again already ›ified objects. On the Art output (fig. 1), one can see the word *code* in different contexts, it has different of meaning. For instance in the class 141 *code* is used in the context of CRIMINAL CODE and in ; 136 it is used in CIVIL CODE. In analysing more carefully these results, one may have the intuition there is also some imprecise hierarchical relations between words in each class and between the classes ıselves. For example in the class 136 the word *criminel* is more important than the word *paix*. Some ecise hierarchy may exist between the classes 343 and 110. Thes questions explains our need to ›re in the following sections uncertain hierarchical classification.

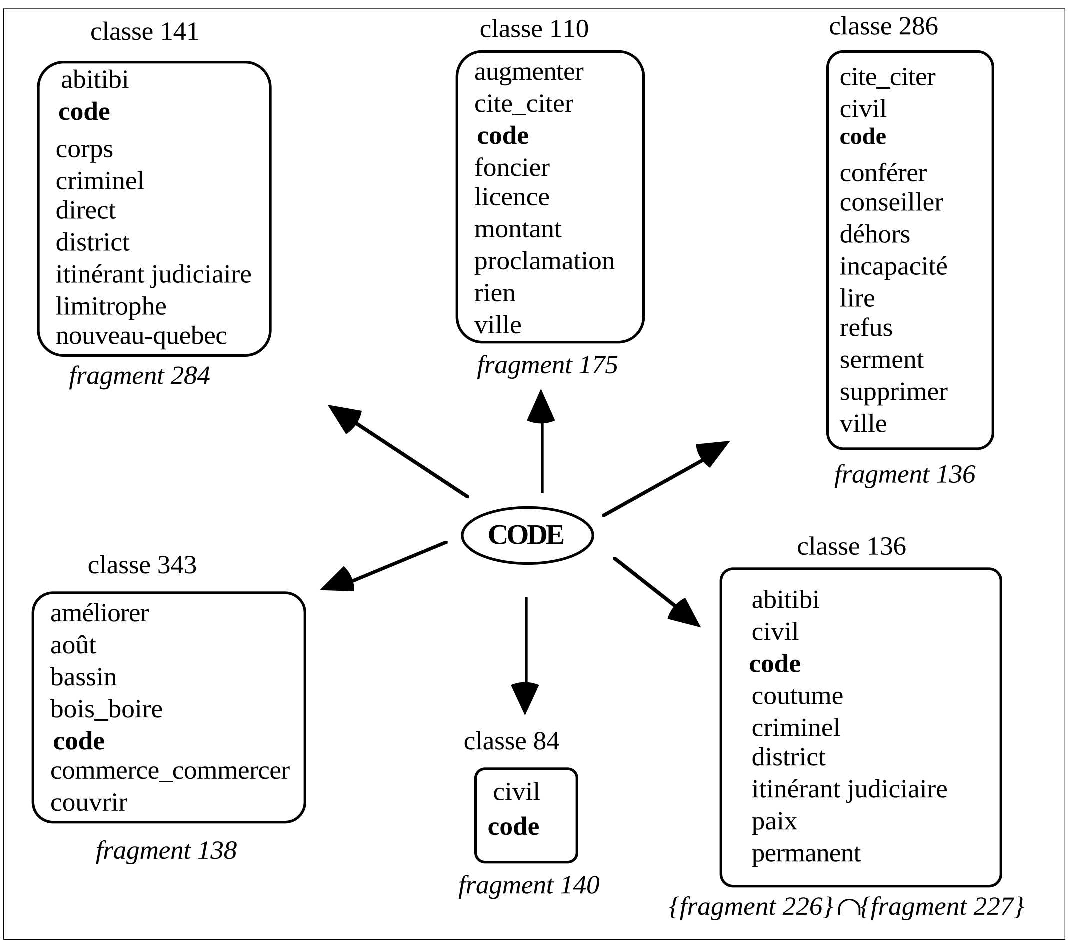


Fig.1. Different contexts of a word.

. hard binary relation

et's consider a hard context ($\mathfrak{O},\mathfrak{C},\mathfrak{R}_h$) where $\mathfrak{O}$ is the set of objects, $\mathfrak{C}$ the set of properties and $\mathfrak{R}_h$ a binary relation between $\mathfrak{O}$ and $\mathfrak{C}$. The vocable "hard" means that there is an absence of uncertainty he lattice extracted from the context $(\mathfrak{O},\mathfrak{E},\mathfrak{R}_h)$ is a set of couples (X,Y) where X and Y are two sets :h belong respectively to the power sets $P(\mathfrak{O})$ and $P(\mathfrak{E})$. Each element of a hard Galois lattice derived ı $(\mathfrak{O},\mathfrak{E},\mathfrak{R}_h)$ must be complete with respect to $\mathfrak{R}_h$: this means that there exists two functions f and g that:

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

A hard Galois lattice structure

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

ıark 1. In the definition 2., we focus on objects generation rather than on properties generation. The ː definition can be given with respect to properties generation. The partial order defined on $\mathfrak{O}),\mathfrak{P}(\mathfrak{E}))$ enables us to construct the graph in the sense that there is an edge from C_1 to C_2 if $C_2\ R_h$ $C_2 \leq C_1$). One can say that C_1 is a parent of C_2 and C_1 is covered[2] by C_2 according to objects ration.

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

$$\bigcap(X_i,Y_i) = \left(g.f(\cup X_i)\ ,\ \cap Y_i\right) \text{ where } i \in I \text{ (finite)}$$
$$\bigcup(X_i,Y_i) = \left(\cap X_i\ ,\ f.g(\cup Y_i)\right).$$

ɔf. see (Barbut M., Monjardet B. 1970)

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

ın uncertain binary relation

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

nition 3. The notation $x\mathfrak{R}_u y$ is used to outline that the binary relation between an object $x \in \mathfrak{O}$ and a erty $y \in \mathfrak{E}$ is uncertain and measured by means of a probability function over the sample space $\mathfrak{O}$. obability value μ_i is assigned to each related couple (x_i,y_i) contained in the cartesian product $(\mathfrak{O}*\mathfrak{E})$. ɔe, one can write: $\mu\{x \in \mathfrak{O}\ ;\ x\mathfrak{R}_u y\} = 1 \Leftrightarrow \{x \in \mathfrak{O};\ x\mathfrak{R}_u y\} = \{x \in \mathfrak{O}\ ;\ x\mathfrak{R}_h y\}$,

A probabilistic Galois lattice structure

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

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

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

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

$$\mu(XR_hY) = \Sigma\Sigma\ \mu(x_i \mathfrak{R}_h y_i)/p*q$$

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

$$= \left((X_1,\mu_1) ; (Y_1,\mu_2)\right) \text{ and } C_2 = \left((X_2,\mu_3) ; (Y_2,\mu_4)\right) \ C_2 \ R_u \ C_1 \Leftrightarrow X_2 = \arg\left(\max_{X_i}\{\mu\{X_1 \subset X_i\}\}\right)_{X_i} .$$

has 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 ırobability $\mu\{X_2 \supset X_1\}$ reflects what we normally mean by the certainty of the rule *if X_2 then X_1*, :e $(X_i)_{i=1,2}$ are considered as two sentences[4]. Hence, one can determine bounds to this probability.

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

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

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

ı order to construct the probabilistic Galois lattice, one has to choose a probability value from the ıations (1). This latter probability value depends on the probability values μ_1 and μ_3 associated ectively to the sets X_1 and X_2.

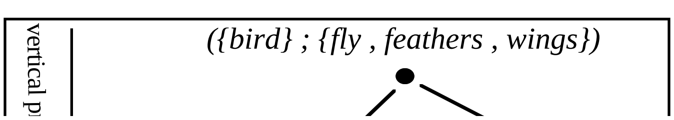
***ɔosition 1.** There is a nonunique ordered set which describes the inherent lattice structure defining ral groupings and probabilistic relationships among the objects and their properties.*

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

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

A complete order within the objects set

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



ɔosition 2. *The partial order ® is a complete order and the set (X,®) is said to be completely and well red.*

of. This order is complete because in the rational numbers contained in the set [0..1] a natural order ɔe induced, this order is *complete*. It is a *well order* since any subset A (A ≠ ∅) contained in X has a *mal element.*

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

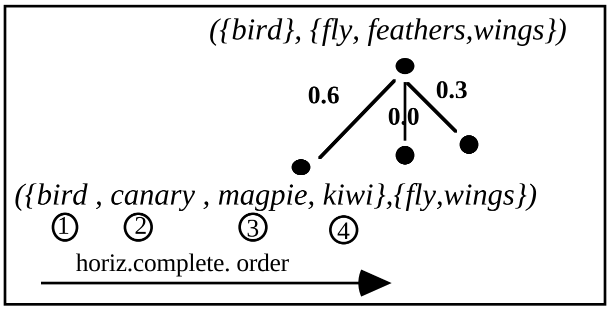


Fig. 3. A horizontal complete order in a probabilistic Galois-lattice.

. A relation measure within the objects set

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

$$\mu\{x \text{ in } X \,;\, x\mathcal{R}_u Y\} = \Sigma\alpha_j\,\mu\{x\mathcal{R}_u y_j\}$$

:e α_i are weights that may privilege some specific properties.

'he generation power criterion and the MM concept

'he Markov Mesh concept which is a subclass of Markovian Random Fields class enables us to classify ɔbjects with respect to the criterion *power of generation.* The unique lattice determined previously ɛsents 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.

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

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

:e | . | is the cardinal symbol and $succ^1$ stands the first successors of the node i.

nition 8. γ is A Markov Mesh with respect to a lattice neighborhood system L if :
ɔ$\{\gamma = \omega\} > 0\ \forall\omega \in \Omega$ and Prob$\{\gamma_{si} / \gamma_{sj}\,,\, s_j \in A_i\}$ = Prob$\{\gamma_{si} / \gamma_{sj}\,,\, s_j \in B_{si}$ and $B_{si} \subset A_{si}\}$.

\pplication to terminology

ı 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 d MUGAL is divided into three main modules. The first one selects words (according to various ria, e.g., functionnal words, happax, probabilistic distribution, discrimination indices) from the

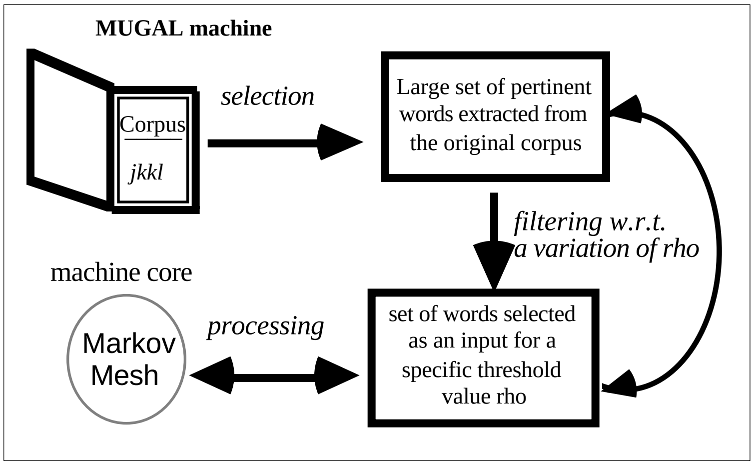


Fig 4. The MUGAL system archtecture.

:onclusion

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

REFERENCES

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