Algorithmes pour la construction de treillis de Galois généraux

Richard EMILION[a], Gérard Lambert[b], Gérard LEVY[c]

[a]UFR SEGMI, Modalx, Université Paris X

[b,c]CERIA, Université Paris IX-Dauphine

Résumé. Nous proposons un nouvel algorithme rapide de construction des treillis de Galois et de leur diagramme de Hasse, pour des variables très générales. Nous généralisons aussi la construction par ordre lexicographique.

Algorithms for general Galois lattice building

Abstract. *We propose a new fast algorithm for building general Galois lattices and drawing their Hasse diagramm when the basic objects are described by very general variables. We also generalize the lexicographic order building.*

Abridged English Version

The best known Galois lattice (GL) is the binary GL defined from a relation between a set $\mathcal{I}$ of n objects and a set $\mathcal{J}$ of p properties : each node represents a subset of objects with their common properties. Letting $d_{ij} = 1$ if the object $i \in \mathcal{I}$ has property $j \in \mathcal{J}$ and $d_{ij} = 0$ otherwise, we see that $d_i = (d_{ij})_{j=1,\dots,p}$ lies in the binary lattice $\{0,1\}^p$. In our present work we propose a fast algorithm which lists the nodes of a GL and draw its Hasse diagram for descriptions d_i lying in any arbitrary lattice $\mathcal{F}$. We therefore can deal with real vectors, sets, functions, fuzzy sets, cumulative histograms, probability cumulative distribution functions, etc.

Given two lattices $< \mathcal{E}, \leq, \vee, \wedge >$ and $< \mathcal{F}, \leq, \vee, \wedge >$, a *Galois connection* (GC) between $\mathcal{E}$ and $\mathcal{F}$ is a pair of mappings (f,g) such that $f : \mathcal{E} \longrightarrow \mathcal{F}$ and $g : \mathcal{F} \longrightarrow \mathcal{E}$ are decreasing, $h = gof : \mathcal{E} \longrightarrow \mathcal{E}$ and $k = fog : \mathcal{F} \longrightarrow \mathcal{F}$ are extensive. Let $I_h = \{x \in \mathcal{E} : h(x) = x\}$ be the set of *closed* (or invariant) elements of $\mathcal{E}$. The *Galois lattice* (GL) defined by a GC (f,g) is the set of nodes $\{(x, f(x)), x \in I_h\}$. Let $\mathcal{E} = \mathcal{P}(\mathcal{I})$, be the power set of $\mathcal{I}$. Let the descriptions $d_i \in \mathcal{F}$ be given for any $i \in \mathcal{I}$. Then (*see* [3]) :

The unique maximal GC (f,g) *between* $\mathcal{E}$ *and* $\mathcal{F}$*, such that* $f(\{i\}) = d_i$*, is given by* $f(A) = \wedge_{a \in A} d_a$ *for any* $A \in \mathcal{E}$ *and* $g(z) = \{i \in \mathcal{I} : z \leq d_i\}$ *for any* $z \in \mathcal{F}$.

DICHOTOMIC SEARCH ALGORITHM

Proposition 1

Let X_0 *and* $K \neq \emptyset$ *be two disjoint subsets of* $\mathcal{I}$. *Let* $i_0 \in K$.

a) *We have* $h(X_0 \cup \{i_0\}) = X_0 \cup A$ *where* $A = \{i \in \mathcal{I} \backslash X_0 : f(X_0) \wedge f(i_0) \leq f(i)\}$.

b) *If* $A \subseteq K$ *then* $X_0 \cup A$ *is the smallest closed set containing* X_0 *and* i_0 *and contained within* $X_0 \cup K$.

c) *If a closed set contains* X_0 *and does not contain* i_0*, then it also does not contain any element of the set* $R = \{i \in K : f(X_0) \wedge f(i) \leq f(i_0)\}$.

Let **Closed**(X_0, K) denote a procedure which lists all the closed sets of $\mathcal{I}$ containing X_0 strictly and contained within $X_0 \cup K$. Proposition 1 then leads to

Procedure **Closed**(X_0, K) ;
var i_0 : *element of* $\mathcal{I}$; z, z_0 : *elements of* $\mathcal{F}$; X, A, R : *subsets of* $\mathcal{I}$;
begin
  $z_0 := f(X_0)$;
  if $K \neq \emptyset$ ***then***
    begin
      choose an element i_0 *of* K; $z := z_0 \wedge f(i_0)$; $A := \{i \in \mathcal{I} \backslash X_0 : z \leq f(i)\}$;
      if $A \subseteq K$ ***then***
        begin $X := X_0 \cup A$; *insert node* (X, z) *in* GL; **Closed1**(X, $K \backslash A$) ; ***end*** ;
      $R := \{i \in K : z_0 \wedge f(i) \leq f(i_0)\}$; **Closed1**($X_0$, $K \backslash R$) ;
    end ;
end ;

Closed($\emptyset, \mathcal{I}$) will then lists all the closed sets. Our implementation is iterative and envolves a linear compact data structure through which operations on sets are performed by just exchanging pairs of integers.

HASSE DIAGRAM

For any $X_0 \subsetneq \mathcal{I}$, let $z_0 := f(X_0)$ and let δ be the mapping defined by

$$\delta : \mathcal{I} \backslash X_0 \rightarrow \mathcal{F},\ \delta(i) = z_0 \wedge f(i) \text{ for any } i \in \mathcal{I} \backslash X_0.$$

Let $z_1^*, ..., z_k^*$ denote the maximal elements of the set $\delta(\mathcal{I} \backslash X_0)$. The following result generalizes that of Bordat obtained for binary GLs (*see* [2]) :

Proposition 2
The smallest closed sets which strictly contain X_0 *are the sets* $X_j = X_0 \cup \delta^{-1}(z_j^*)$.

A GENERALIZATION OF GANTER'S ALGORITHM

Let $\mathcal{F} = \mathcal{F}_1 \times ... \times \mathcal{F}_j \times ... \mathcal{F}_p$, where $\mathcal{F}_j$ is a totally ordered finite alphabet represented, without loss of generality, as a set of integers $\mathcal{F}_j = \{0, 1, ..., M_j\}$, $0 < 1 < ... < M_j$.

Let $\preceq$ denote the *lexicographic order* on $\mathcal{F}$. For any $a = (a_1, ..., a_p) \in \mathcal{F}$ such that $a \prec M = (M_1, ..., M_p)$, let $a^+ \in \mathcal{F}$ be the successor of a w.r.t. $\preceq$. Define the transition index of a, say $i^+(a)$, as the greatest $j \in \{1, ..., p\}$ such that $a_j < M_j$. If $t = i^+(a)$, let $a^* = (a_1^*, ..., a_p^*) \in \mathcal{F}$ be defined as $a^* = (a_1, ..., a_{t-1}, M_t, ..., M_p)$. The well-known and fast Ganter's algorithm is a particular case of the following when $\mathcal{F}_j = \{0, 1\}$ (*see* [6]).

Proposition 3
Let $a \in \mathcal{F}$ *such that* $a \prec M$, *then* $y = k(a^+)$ *is the least closed element greater than* a *(w.r.t.* $\preceq$*) iff* $a \preceq y \preceq a^*$ *and thus iff* $\forall j < i^+(a)\ y_j = a_j$.