Ensuring the Drawability of Extended Euler Diagrams for up to 8 Sets

Anne Verroust, Marie-Luce Viaud

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Ensuring the Drawability of Extended Euler Diagrams for up to 8 Sets

Anne Verroust — Marie-Luce Viaud

N° 4973

octobre 2003

THÈME 3

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4973--FR+ENG

Ensuring the Drawability of Extended Euler Diagrams for up to 8 Sets

Anne Verroust*, Marie-Luce Viaud**

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Imédia

Rapport de recherche n° 4973 — octobre 2003 — 24 pages

Abstract: This paper shows by a constructive method the existence of a diagrammatic representation called extended Euler diagrams for any collection of sets $X_1, ..., X_n$, $n < 9$. These diagrams are adapted for representing sets inclusions and intersections: each set X_i and each non empty intersection of a subcollection of $X_1, ..., X_n$ is represented by a unique connected region of the plane. Starting with a description of the diagram, we define the dual graph G and reason with the properties of this graph to build a planar representation of the $X_1, ..., X_n$. These diagrams will be used to visualize the results of a complex request on any indexed video database. In fact, such a representation allows the user to perceive simultaneously the results of his query and the relevance of the database according to the query.

Key-words: Euler diagrams, Venn diagrams, hypergraphs, graph planarity, data visualization

* INRIA Rocquencourt, France
** INA, Bry sur Marne, France

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Ensuring the Drawability of Extended Euler Diagrams for up to 8 Sets

Résumé : Cet article montre de manière constructive l'existence d'une représentation sous forme de diagramme appelée diagramme d'Euler étendu de toute collection d'ensembles $X_1, ..., X_n$ pour $n < 9$. Ces diagrammes sont adaptés pour représenter les inclusions et intersections d'ensembles : chaque ensemble X_i et chaque intersection non vide d'un sous ensemble des $X_1, ..., X_n$ est représentée par une région unique du plan. À partir d'une description du diagramme, nous définissons un graphe dual G et étudions les propriétés de ce graphe pour construire une représentation planaire des $X_1, ..., X_n$.
Ces diagrammes sont destinés à être utilisés pour visualiser les résultats d'une requête complexe sur une base de données indexée de documents multimédia. Ce type de représentation permet à l'utilisateur de percevoir à la fois le résultat de sa requête et la pertinence de la base de données vis à vis de la requête.

Mots-clés : diagrammes d'Euler, diagrammes de Venn, hypergraphes, planarité de graphe, visualisation de données

1 Introduction

In the 18th century, Leonard Euler has proposed a general notation for representing sets relations, known as Euler Diagrams (cf. [Eul75]). An Euler diagram is a planar subdivision built from a collection of Jordan curves, each curve being associated to a set and each region to a non-empty intersection of sets. Euler Diagrams have been used in various applications (cf. [GKH99]). Our main goal here is to use them in a database visualization context.
Indeed, enhancing the visualization of the results of queries in databases becomes a challenging and useful task [Con94,CT95,CCLC97]. Euler diagrams gives an intuitive representation of the relationship of a collection of datasets. Moreover, if elementary queries are associated with sets, these representations generate maps of the database according to the user's viewpoint.

For example, if a user is looking for documents on writers or painters of the XIX$^{\text{th}}$ century on a given database and if he obtains the diagram of figure 1 as a result, he can infer information both on his request and on the content of the database. Indeed, the diagram shows in particular that this database will not be relevant if he looks for information on painters from the XX$^{\text{th}}$ century, because all the painters present in this database are either poets or from the XIX$^{\text{th}}$ century.

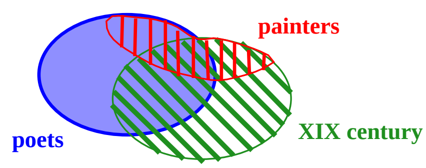

Fig. 1. A diagram associated with a query

Venn diagrams representing all the possible intersections between the initial sets [Rus01] are not adapted to our purposes. In fact, we want that each region drawn in the diagram corresponds to non empty list of selected documents. This constraint is motivated by the fact that most of the professional TV archives indexed databases are associated to a thesaurus containing exclusive indices. This characteristic implies that many combination of elementary request do not select any document. Then, we propose a graphic representation which is an extension of Euler diagrams (cf. [Eul75,LP97]).
In such an application, we must be able to represent by Euler diagrams any collection of sets and their non empty intersections. More precisely, our goal here is to show by a constructive method the existence of an extended Euler diagram for any collection of intersections of up to eight sets. We will notice that when the number of sets is greater or equal to nine, such a planar representation may not exist.
This paper is structured as follows:

- In section 2, we first define the extended Euler diagram representation for a collection of intersections Y_i of a set of sets X_i. Then, we present its dual representation, the L_connected labelled graph associated to X_i and describe the process transforming a L_connected labelled graph in an extended Euler diagram.

- In section 3, we show that any collection of intersections of up to 8 sets has a planar L_connected labelled graph representing it .

2 Diagram representation

Given $X = \{X_1, X_2, ..., X_k\}$ a collection of non empty distinct sets, we want to build a graphic representation which shows information about the sets and their intersections on a plane. Let $Y = \{Y_1, Y_2, ..., Y_{2^k}\}$ the collection of all possible intersections between the X_i. Euler diagrams [Eul75] could be used but appear to be too restrictive for our purpose. In fact, an Euler diagram consists of simple closed curves called contours associated with sets which split the plane into zones. Each set X_i is associated with an unique contour, X_i being represented by the interior of the contour.

The concrete Euler diagrams proposed in Flower and Howse's approach [FH02] are very well defined but even more restrictive. In fact, concrete Euler diagrams are Euler diagrams with few more constraints. The first constraints introduced at the curve level, make hypothesis on the set of intersections being drawn : each segment of curve delimits the interior and the exterior of exactly one set , and each intersection of curves is the crossing of exactly two contours. The introduction of "exactly" is very useful to specify formally the problem and its dual formulation with graphs, but eliminates the cases in which the set of subsets built from the intersections of the X_i does not have such properties.
We propose an extension of Euler/Venn diagrams which increases the number of set X being potentially drawable. As we will see in the next section, our diagrams can be related to the notion of vertex-planarity introduced by Johnson and Pollak [JP87] for hypergraphs.
In this paper and according to our purposes, for any collection X of up to eight elements, we want to represent by an extended Euler diagram the collection of non empty Y_i. Such diagrams are characterized by the following properties:
- An intersection point may intersect more than 2 contours,
- A curve segment may be part of more than one contour,
- Each non empty Y_i is associated with a unique zone,
- each set X_i is associated with a set of zones whose union forms a connex planar region which may contain holes.

Those properties are used to define formally the extended Euler diagrams $\mathcal{ED}$ (in section 2.1) and its dual representation, the L_connected labelled graph G (in section 2.2). Section 2.3 describes the process to transform a L_connected labelled graph G representing X in an extended Euler diagram: making the hypothesis that G is planar, we compute a drawing $D(G)$ of G and introduce special vertices in $D(G)$ to obtain a triangular mesh $D_t(G)$. Finally, from $D_t(G)$, we generate an Extended Euler diagram which is a representation of X.

2.1 Extended Euler diagrams

Definition 1. *Let L be a finite set of labels and C a set of Jordan curves in $\mathbb{R}^2$.*
We say that C is labelled *by L when each curve c of C is associated with a couple $(\lambda(c), sign(c))$ where $\lambda(c) \in L$ and $sign(c) \in \{+, -\}$.*
To each labelled curve c of C corresponds a zone $\zeta(c)$ defined by:
- if $sign(c) = +$, then $\zeta(c) = int(c)$
- if $sign(c) = -$, then $\zeta(c) = ext(c)$

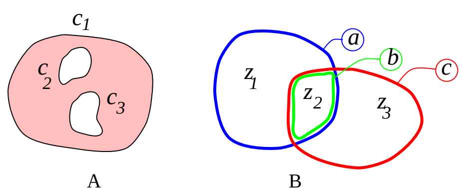

Fig. 2. A: a zone with two holes $sign(c_2) = sign(c_3) = -$; B: an extended Euler diagram with $m(z_1) = \{a\}$, $m(z_2) = \{a, b, c\}$ and $m(z_3) = \{c\}$.

Definition 2. *An* extended Euler diagram *is a triple (L, C, Z) whose components are defined as follows:*

1. *L is a finite set of labels*
2. *C is a set of Jordan curves labelled by L and verifying:*
   (a) *$\forall l \in L, \exists c \in C$, $\lambda(c) = l$ and $sign(c) = +$.*
   This curve c is unique and will be called the envelope *of l ($c_{env}(l)$).*
   (b) *if $\lambda(c) = \lambda(c')$, $c \neq c'$ and $sign(c) = sign(c')$ then c and c' do not intersect*
   (c) *if $\lambda(c) = \lambda(c')$, $c \neq c'$ and $sign(c) = +$, then $sign(c') = -$ and $c' \subset int(c)$*
3. *Z is a set of zones corresponding to the planar partition defined by the curves of C.*
   Each zone z of Z is associated to a set of labels $m(z)$ defined by
   (a) *$m(z) = \{l \in L | \forall c \in C,$ if $\lambda(c) = l$ then $z \subset \zeta(c)\}$*
   (b) *if $m(z) = m(z')$ and $m(z) \neq \emptyset$, then $z = z'$*
   We note $Z_\emptyset$ the set of zones associated to an empty set of labels.
   $Z_\emptyset$ contains at least the zone $z_\emptyset = \bigcap_{\{c | sign(c) = +\}} ext(c)$.

The set of extended Euler diagrams is noted $\mathcal{ED}$.

As a mater of fact, we have introduced Jordan curves to define zones, but those notions are equivalent. In the following, we will use rather the zones formalization.

Definition 3. *Let $X = \{X_1, X_2, ..., X_k\}$ be a set of non empty distinct subsets of $\mathcal{X}$, $Y = \{Y_1, Y_2, ..., Y_m\}$ the set of all possible non empty intersections between the X_i ($m \leq 2^k$).*
We say that the extended Euler diagram (L, C, Z) is a diagram representation *of X if and only if:*

1. *there is a bijection $\psi : L \to X; l \mapsto x$*

2. $\phi : Z \backslash Z_{\emptyset} \to Y$ *;* $z \mapsto y$ *defined by* $\phi(z) = y = \bigcap_{l \in m(z)} \psi(l)$ *is a bijection.*

Remark 1. Extended Euler diagrams can be related with Johnson and Pollak's notion of planarity for hypergraphs [JP87].

Let $H = (V, E)$ be an hypergraph and $X = \{X_1, ..., X_k\}$ be a set of non empty distinct subsets of $\mathcal{X}$ such that there are:
- a one-one map ϵ from the set of hyperedges E and $X = \{X_1, ..., X_k\}$,
- a map σ between V and the set of all possible non empty intersections between the X_i, $Y = \{Y_1, Y_2, ..., Y_m\}$
satisfying: $\forall v \in V$, v belongs to the hyperedge e of E if and only if $\sigma(v) \subseteq \epsilon(e)$.
If an extended Euler diagram (L, C, Z) is a diagram representation of X, then (L, C, Z) is a vertex-based diagram representing the hypergraph $H = (V, E)$ and H is vertex-planar according to Johnson and Pollak's definition.

Rather than working with extended Euler diagrams, we will use a dual representation: the L_connected labelled graphs defined in the next section.

2.2 L_connected labelled graphs

Definition 4. *A* labelled graph *is a triple* $G(L, V, E)$ *where:*

1. *L is a finite set of labels*
2. *V is a set of* labelled vertices*, i.e.:*
   (a) *each vertex v is labelled with a set of labels* $m(v) \subseteq L$
   (b) *two distinct vertices v and w of V have distinct sets of labels.*
3. *E is a set of edges such that:*
   (a) *each edge* $e = (v, w)$ *of E is labelled with a set of labels* $m(e) = m(v) \cap m(w)$
   (b) *if* $e \in E$ *then* $m(e) \neq \emptyset$

In the rest of the paper, $L(W)$ will be the set of labels associated to the vertices of W, i.e. $L(W) = \bigcup_{v \in W} m(v)$, where W is a set of labelled vertices.

Definition 5. *Let* $G(L, V, E)$ *be a labelled graph.*

- *Let l be a label of L. We say that* $G(L, V, E)$ *is* l_connected *if and only if the subgraph* G' *of* $G(L, V, E)$ *on the set* V' *of vertices of V having l in its set of labels is connected.*
- $G(L, V, E)$ *is said* L_connected *if and only if it is l_ connected for all l in L.*
- $G(L, V, E)$ *is said* L_complete *when E is defined by:*
  $E = \{(v, w) | v \in V, w \in V$ *and* $m(v) \cap m(w) \neq \emptyset\}$
- *A vertex v of V is said* L_connectable *to a subset W of V if and only if* $m(v) \subseteq L(W)$.

These definitions are illustrated in figure 3.

Remark 2. Given L and V, there exists only one L_complete labelled graph noted $G(L, V, E_c)$ and any L_connected labelled graph $G(L, V, E)$ is a subgraph of $G(L, V, E_c)$.

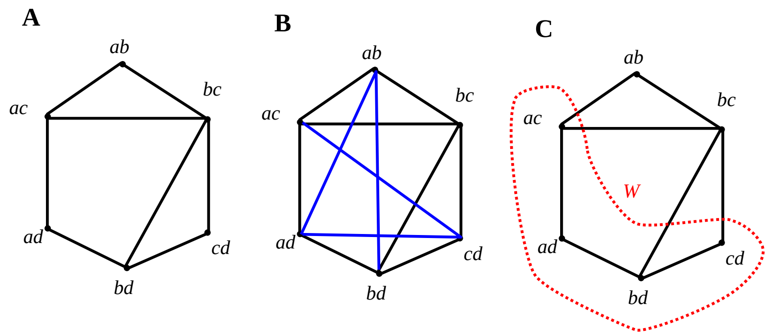

Fig. 3. Let us take $L = \{a, b, c, d\}$, $V = \{ab, ac, ad, bc, bd, cd\}$,
$E = \{(ab, ac), (ab, bc), (bc, cd), (cd, bd), (bd, ad), (ad, ac), (ac, bc), (bd, bc)\}$,
$E_c = E \cup \{(ab, bd), (ab, ad), (ac, cd), (ad, cd)\}$.
A: $G(L, V, E)$ is a L_connected labelled graph; B: $G(L, V, E_c)$ is the corresponding L_complete graph; C: ab is L_connectable to $W = \{ad, bd, cd, ac\}$.

In the following, we will note:
- $\mathcal{G}$ the set of L_connected labelled graphs
- $\mathcal{G}^{\mathcal{V}}(L, V)$ the set of L_connected labelled graphs associated to a given set of label L and a set V of labelled vertices: $\mathcal{G}^{\mathcal{V}}(L, V) = \{G(L, V, E) | G(L, V, E) \in \mathcal{G}\}$
- $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L, V)$ the set of planar graphs belonging to $\mathcal{G}^{\mathcal{V}}(L, V)$.

Definition 6. *The mapping* $dual : \mathcal{ED} \rightarrow \mathcal{G}; (L', C, Z) \mapsto G(L, V, E)$ *is defined by:* $G(L, V, E) = dual((L', C, Z))$ *if and only if*

(i) there is a one to one mapping between L' and L
(ii) there is a bijection $\delta : Z \rightarrow V; z \mapsto v$ such that $m(z) = m(\delta(z))$
(iii) $e = (v, w) \in E$ if and only if $\delta^{-1}(v)$ and $\delta^{-1}(w)$ are adjacent along a portion of curve of non null lenght in the planar partition formed by C.

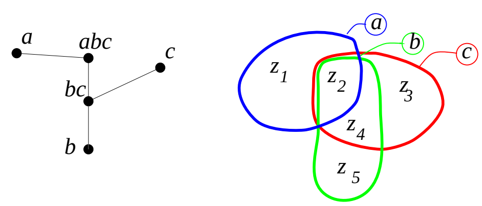

Fig. 4. An extended Euler diagram (L, C, Z) and its dual. We have $m(a) = m(z_1) = \{a\}$, $m(abc) = m(z_2) = \{a, b, c\}$, $m(c) = m(z_3) = \{c\}$, $m(bc) = m(z_4) = \{b, c\}$ and $m(b) = m(z_5) = \{b\}$.

Let $X = \{X_1, X_2, ..., X_k\}$ be a set of non empty distinct sets and $Y = \{Y_1, Y_2, ..., Y_{2^k}\}$ the set of all possible intersections between the X_i. Extended Euler diagrams and labelled graphs are introduced to built a planar representation of X. More precisely:

Definition 7. *Let* $Y' = \{Y_1, Y_2, ..., Y_m\}$ *be the subset of* Y *which elements are the non empty intersections between the* X_i ($m \leq 2^k$) *We say that the L_connected labelled graph* $G(L, V, E)$ *is a* graph representation *of* X *if and only if:*

1. *there is a bijection* $\lambda : L \rightarrow X = \{X_1, ...X_k\}$; $l \mapsto x$
2. $\chi : V \rightarrow Y'$; $v \mapsto y$ *defined by* $\chi(v) = y = \bigcap_{l \in m(z)} \lambda(l)$ *is a bijection.*

In the following, we will note $\mathcal{ED}(X)$ *the subset of* $\mathcal{ED}$ *composed of diagram representations of* X *and* $\mathcal{G}_P(X)$ *the set of planar L_connected labelled graph which are a graph representation of* X.

Remark 3. If (L,C,Z) is a diagram representation of X, then $dual((L, C, Z))$ is a graph representation of X.

We want to find a planar representation of the set Y' of non empty elements of Y such that:

- each element v of Y' corresponds to a unique zone Z_v.
- for each X_i, if $v \subseteq X_i$ and v is not the unique element of Y' included in X_i, then we can find w in Y', $w \subset X_i$, $w \neq v$ such that Z_v and Z_w are adjacent.

We will see in the next section that if we find an element of $\mathcal{G}_P(X)$ we can built such a planar representation.

2.3 From L_connected labelled graphs to extended Euler diagrams

As our purpose here is mainly to show the existence of an extended Euler diagram representing X once a planar L_connected labelled graph representing X is found, we will describe informally the process of building a extended Euler diagram from a planar L_connected labelled graph and illustrate it with figure 5.
Let $G(L, V, E)$ be an element of $\mathcal{G}_P(X)$. To have a better control on the drawing of the resulting diagram we introduce special edges and vertices during the process:

1. A straight-line drawing $D(G)$ of G is obtained (cf. [BETT99]).
2. We temporary remove the dangling edges from each internal face of $D(G)$.
3. For each internal face F of $D(G)$ which is not triangular, we insert a special vertex representing the empty set and connect it with all the vertices of F (cf. figure 5 (A)). These special vertices are introduced only to control the drawing of the diagram. Then the dangling edges are reinserted in one of the new triangular face subdividing F.
   We now have a triangulation $F_1, ..., F_n$ representing $G(L, V, E)$ and we compute a new drawing of it.
4. If a triangular face $F_i = (v_1, v_2, v_3)$ contains at least a dangling edge connected to v_i, a special vertex $v_\emptyset$ representing the empty set is inserted and edges connecting the extremities of the dangling edges and $v_\emptyset$ are created. If $v_{j,j \neq i}$ is not connected to a dangling edge then $v_\emptyset$ and v_j are connected (cf. figure 5 (C)).
   This leads to a new planar graph G' and a new drawing $D(G')$

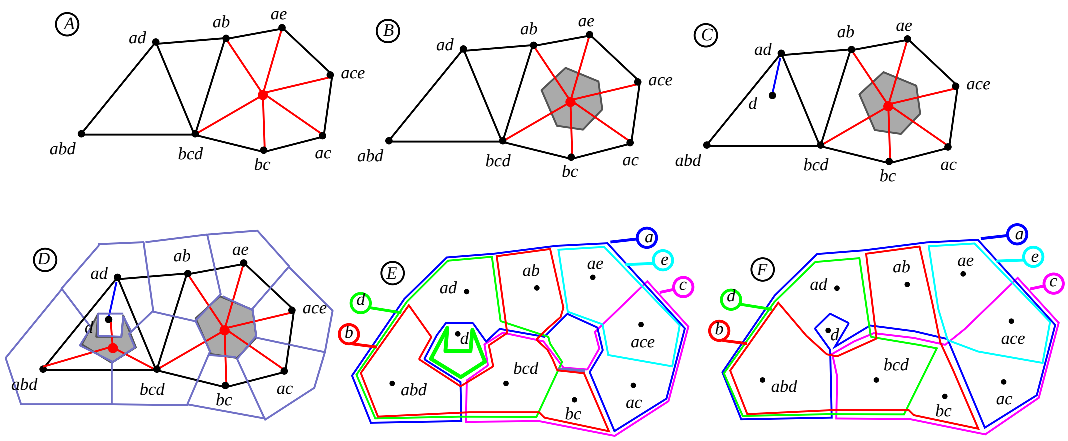

Fig. 5. The construction of an extended Euler diagram from a drawing of a planar L_connected graph ($L = \{a, b, c, d, e\}$ and $V = \{d, ad, ab, ae, ace, ac, bc, bcd, abd\}$). Each internal empty region is drawn in grey and is associated to a red vertex. A: a drawing of a L_connected labelled graph without the dangling edge (d, ad). B: the empty region associated to a non triangular face. C: the dangling edge (d, ad) is drawn. D: the graph and its associated regions. E: the extended Euler diagram with internal regions associated to an empty set of labels. F: the resulting extended Euler diagram after deformation.

5. Each vertex of G' is associated to a planar region labelled by its set of labels. The regions of the plane corresponding to an empty set of labels are associated to the new vertices (cf. figure 5 (B) and (D)). We obtain an extended Euler diagram (L, C', Z') that may contain regions associated to an empty set of labels.
6. The internal regions associated to an empty set of labels are then deformed to a region reduced to a point as in figure 5 (E). As G is L_connected, each zone assopciated to a given label is connected. We thus obtain an extended Euler diagram (L, C, Z) such that $G(L, V, E) = dual((L, C, Z))$. (cf. figure 5 (F)).

$$\begin{array}{ccccc} \boldsymbol{G(L,V,E)} & \overset{\boldsymbol{draw(1)}}{\longrightarrow} & \boldsymbol{D(G)} & \overset{extend(2,3,4)}{\longrightarrow} & D(G') \\ & \nwarrow \boldsymbol{dual} & \downarrow & & \downarrow diag(5) \\ & & \boldsymbol{(L,C,Z)} & \overset{deform(6)}{\longleftarrow} & (L,C',Z') \end{array}$$

The previous paragraph was an informal proof of the following proposition:

Proposition 1. *If there is a planar L_connected graph $G(L, V, E)$ representing X, then there is an extended Euler diagram (L, C, Z) representing X. This diagram is such that $G(L, V, E) = dual((L, C, Z))$.*

Remark 4. A dangling edge (v, w) connected to v appears when $m(w)$ is included in $m(v)$. Then to obtain the most intuitive diagram representation, the dangling edges should appear on the external face of $D(G')$, as it is done in figure 8. Such criteria should be considered as one of the visual criteria mentioned in the conclusion.

3 The planarity problem

Given a set of labels L and a set of labelled vertices V, to draw an extended Euler diagram, we need to find at least one element in $\mathcal{G}_{\mathcal{P}}^{\mathcal{V}}(L,V)$, i.e. a set of edges E such that the labelled graph $G(L,V,E)$ is planar and L_connected.
To prove the planarity of a graph, we will use the graphs K_n and $K_{n,n}$, $n \geq 2$
- K_n is the complete graph defined on n vertices: in K_n, every vertex is adjacent to every other vertex
- $K_{n,n}$ is the complete bipartite graph consisting of two disjoint vertex sets $V = \{v_1, ..., v_n\}$, $W = \{w_1 ..., w_n\}$ and the edge set $E = \{v_i, w_j | 1 \geq i, j \geq n\}$
and Kuratowski's characterization of planar graphs [Kur30]:

Theorem 1. *A graph is planar if and only if it does not contain a subdivision of K_5 or $K_{3,3}$ as a subgraph.*

We already know that, if $card(V) = 2^{card(L)}$ the diagram to draw is a Venn diagram which has a planar representation (cf. [Rus01]) for any value of $card(L)$. But this property does not hold in the general case.
In fact we have[1]:

Proposition 2. *Let k be the cardinality of L. When $k \geq 9$, there exists at least a set of labelled vertices V for which all the graphs of $\mathcal{G}^{\mathcal{V}}(L,V)$ are non planar.*

Proof. Suppose $L = \{a,b,c,d,e,f,g,h,i\}$ and
$\bigcup_{v \in V} m(v) = \{\{a,b,c\},\{d,e,f\},\{g,h,i\},\{a,d,g\},\{b,e,h\},\{c,f,i\}\}$. Then $\mathcal{G}\mathcal{V}(L,V)$ contains only one L_connected labelled graph which is a $K_{3,3}$ (cf. figure 6). □

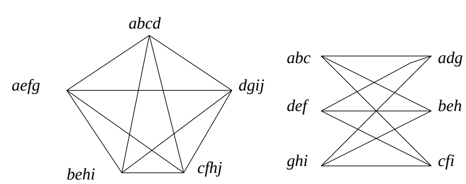

Fig. 6. Two non planar L_connected labelled graphs: on the left a K_5 and on the right a $K_{3,3}$

Let us now state some results on sets of labelled vertices and L_connected planar graphs.

Definition 8. *Let V be a set of labelled vertices and v and w two vertices of V,*
- v and w are said label_disjoint *when $m(v) \cap m(w) = \emptyset$*
- v is said label_included *in w when $m(v) \subset m(w)$*

[1] The following proposition is another version of the planarity results for Euler's Circles presented by Lemon and Pratt in [LP97].

Proposition 3. *Let $G(L,V,E)$ be a L_connected planar graph and W be a set of vertices such that $V \subset W$ and every vertex w of $W \backslash V$ is label_included in a vertex of V. Then $\mathcal{G}^{\mathcal{V}}(L,W)$ contains a planar graph.*

Proof. To prove that $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L,W)$ is not empty, we will built a graph $G(L,W,E')$ containing $G(L,V,E)$ by adding successively the vertices of $W \backslash V$ and show that the construction ensures that it is a L_connected planar graph.
Let $w \in W \backslash V$ and v a vertex of V such that w is label_included in v. Then if we add the edge $e = (v,w)$ to $G(L,V,E)$ we obtain a L_connected graph $(L, V \cup \{w\}, E \cup \{e\})$ which is still planar because the addition of e cannot contribute to add a $K_{3,3}$ or a K_5 in $(L, V \cup \{w\}, E \cup \{e\})$. By augmenting the same way the graph $G(L,V,E)$ for each $w \in W \backslash V$, we obtain at the end a L_connected planar graph on W. □

Corollary 1. *Let W be a set of labelled vertices on L and $V = \{v \in W | card(m(v)) \geq 2\}$. If $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L,V)$ is not empty, then $\mathcal{G}^{\mathcal{W}}_{\mathcal{P}}(L,W)$ is not empty.*

Proof. Let suppose that $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L,V)$ contains a L_connected planar graph $G(L,V,E)$.
$W \backslash V$ can be partitioned in two subsets:
$W_1 = \{v \in W | card(m(v)) = 1$ and v is label_included in a vertex w of $V\}$
$W_2 = \{v \in W | card(m(v)) = 1$ and v is label_disjoint of any vertex w of $V\}$
Using proposition 3, if W_1 is not empty, we built a planar L_connected labelled graph $G(L, V \cup W_1, E')$.
As the vertices of W_2 are label_disjoint of all the vertices of V, the graph $G(L, V \cup W_1 \cup W_2, E')$ is a planar L_connected labelled graph. □

Using the previous corollary, we will restrict ourselves to sets of labelled vertices V on L satisfying:
(H1) $\forall v, w \in V$, if v is label_included in w then $v = w$.
(H2) any vertex v of V has more than one label in $m(v)$.

For a set V of labelled vertices satisfying (H1) and (H2), we proceed as follows to show that $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L,V)$ is not empty when $card(L) < 9$:

1. we choose a subset V_0 of vertices of V among those satisfying $L(V_0) = L(V)$ and build a L_connected planar graph $G_0(L, V_0, E_0)$ on V_0 (cf sections 3.1 and 3.2).
2. we build a partition of $V = V_0 \cup V_1 \cup ... \cup V_k$, $k \leq card(V_0)$.
   This construction is described in section 3.3.
3. for each V_i, we show how to extend $G_0(L, V_0 \cup V_1 ... \cup V_{i-1}, E_{i-1})$ to obtain a L_connected planar graph $G(L, V_0 \cup V_1 ... \cup V_i, E_i)$.
   This is the subject of section 3.4.

3.1 Choice of V_0

Given V and L, let $\mathcal{T}(V)$ be the set of subsets T of V such that $L(T) = L(V)$ and $\mathcal{T}_0(V)$ be the subset of $\mathcal{T}(V)$ formed by the elements T of $\mathcal{T}(V)$ having a minimum number of vertices.

Definition 9. *Given T and T' in $\mathcal{T}_0(V)$, we rename the vertices of T and T' w.r.t. the cardinality of their associated sets of labels, i.e. $T = \{v_0, v_1 ... v_p\}$ and $T' = \{v'_0, v'_1 ... v'_p\}$ with $card(m(v_i)) \geq card(m(v_j))$ and $card(m(v'_i)) \geq card(m(v'_j))$ when $i \leq j$.*
We say that $T \geq_L T'$ if and only if:
$\exists k \leq p, \forall i < k, card(m(v_i)) = card(m(v'_i))$ *and* $card(m(v_k)) \geq card(m(v'_k))$.
$\mathcal{T}_{max}(V)$ *is the set composed by the maximal elements of* $\mathcal{T}_0(V)$ *for* $\geq_L$.

Using the definition of $\mathcal{T}_{\max}(V)$ and the hypothesis (H1) and (H2) on V we have:

Proposition 4. *Given $V_0 \in \mathcal{T}_{max}(V)$*
1. $\forall v \in V_0, \exists l \in m(v)$ *s.t.* $l \notin L(V_0 \backslash \{v\})$
2. *If $V' \in V_0$, then we can not have neither*
(a) $\exists V" \subseteq V$, *s.t.* $L(V') \subseteq L(V")$ *and* $card(V') > card(V")$
nor
(b) $\exists V" \subseteq V$, *s.t.* $L(V') \subseteq L(V")$ *and* $card(V') = card(V")$ *and* $V' <_L V"$
3. $card(V_0) < card(L)$

In the following, for a given set of labels L and a set of labelled vertices V on L satisfying (H1) and (H2),
- if W is a subset of V and $v \in W$, we note $Lu_W(v) = \{l \in m(v) | \forall w \in W, w \neq v, l \notin m(w)\}$ and $Lu(W) = \cup_{v \in W} Lu_W(v)$
- V_0 denotes an element of $\mathcal{T}_{\max}(V)$.

3.2 Construction of a L_connected planar graph on V_0 for $card(L) < 9$

Definition 10. *Given a L_connected graph $G(L, V, E)$, an edge e of E is said:*
- L_irreducible *if $G(L, V, E \backslash \{e\})$ is not L_connected.*
- totally L_irreducible *if $G(L, V, E_c \backslash \{e\})$ is not L_connected.*
A L_connected graph $G(L, V, E)$ is said L_minimal *if all the edges of E are L_irreducible.*

Let us notice that when e is totally L_irreducible, e is included in every graph of $\mathcal{G}^{\mathcal{V}}(L, V)$ and $m(e)$ contains at least one label of L which is present in only two vertices of V.

Proposition 5. *If $card(L) < 9, \mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L, V_0)$ is not empty.*

Proof. Let $G(L, V_0, E_0)$ be a L_minimal and L_connected graph of $\mathcal{G}^{\mathcal{V}}(L, V_0)$. We know by proposition 4 that $card(V_0) < card(L)$. Then, as $card(L) < 9$ we have the following cases to consider:
– $\boldsymbol{card(V_0) \leq 4}$. $\mathcal{G}^{\mathcal{V}}(L, V_0) = \mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L, V_0)$ because we need at least 5 vertices to build a non planar graph.
– $\boldsymbol{card(V_0) = 5}$. $G(L, V_0, E_0) \neq K_5$. If it was the case, K_5 would be L_minimal and any edge e of K_5 would be totally L_irreducible. Then, as K_5 has 10 edges, we must have $card(L) \geq 10$ which contradicts the hypothesis.
– $\boldsymbol{card(V_0) = 6}$. We have $card(Lu(V_0)) \geq 6$ and $card(L) \leq 8$. Then L contains one or two labels which do not belong to $Lu(V_0)$.

If $L\backslash Lu(V_0) = \{l\}$, then E_0 consists in a path joining the vertices having l in their set of labels. Then $G(L, V_0, E_0)$ is planar.

If $L\backslash Lu(V_0) = \{l, l'\}$: to build a L_connected graph on six vertices connecting two labels, we need less than 10 edges (we will see in figure 12, that only 5 edges are necessary). Then such graph cannot contain any $K_{3,3}$ or K_5.

– $\boldsymbol{card(V_0) = 7}$. As $L\backslash Lu(V_0) = \{l\}$, E_0 consists in a path joining the vertices having l in their set of labels. □

3.3 Creation of a partition of V

Definition 11. *Given V_0 an element of $\mathcal{T}_{max}(V)$ and v a vertex of $V\backslash V_0$, we call $\mathcal{W}(v, V_0)$ the set of subsets W of V_0 such that v is L_connectable to W. Let $\mathcal{W}_{\mathcal{MIN}}(v, V_0)$ be the subset of $\mathcal{W}(v, V_0)$ whose elements have the minimal number of vertices. We build a partition $\{V_0, V_1, V_2, .., V_p\}$ of V where each $V_i, i > 0$ is defined by: $V_i = \{v | \exists W \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$ and $card(W) = i\}$.*

One shall notice that hypothesis (H1) on V implies that V_1 is empty. Before extending G_0 with the V_i, we will give general results on the $V_{i,i>0}$.

Lemma 1. *If $v \in V_n$ and if $W_n = \{w_1, ...w_n\}$ is an element of $\mathcal{W}_{\mathcal{MIN}}(v, V_0)$, then $card(m(v)) \geq n$ and $m(v) = \{l_1..l_n\} \cup L_r$, with $L_r \subset L(W_n)$ and $l_i \in Lu_{W_n}(w_i)$.*

Proof. As v is in V_n, v is L_connectable to W_n. Therefore, if there was w_i in W_n such that $m(v) \cap Lu_{W_n}(w_i) = \emptyset$, then v would be L_connectable to $W_n\backslash\{w_i\}$ and $W_n \notin \mathcal{W}_{\mathcal{MIN}}(v, V_0)$. □

Lemma 2. *If $card(V_0) = card(L) - 1$ and $card(L) > 3$ then $V = V_0$.*

Proof. We know that $Lu(V_0) \geq card(V_0)$. Then, as $card(V_0) = card(L) - 1$ and V satisfies (H2), each vertex v of V_0 is such that $m(v) = Lu_{V_0}(v) \cup \{l\}$, with $l \in L\backslash Lu(V_0)$ and $card(Lu_{V_0}(v)) = 1$. If $v' \in V\backslash V_0$, $m(v')$ contains at least two labels l_1 and l_2 belonging to $Lu(V_0)$ with $l_1 \in m(v_1)$, $l_2 \in m(v_2)$ and $\{v_1, v_2\} \subset V_0$. Then, if $card(V_0) > 2$, the set of vertices $W = (V_0\backslash\{v_1, v_2\}) \cup \{v'\}$ contradicts proposition 4 (2). □

Proposition 6.
(a) If $V_{n,n\geq 2}$ is not empty then $\forall v \in V_n, \forall W_n \in \mathcal{W}_{\mathcal{MIN}}(v, V_0), n \leq \frac{card(L(W_n))}{2}$
(b) Let $v \in V_n$. If $\exists w_k \in W_n$, with $card(m(w_k) \cap Lu(W_n)) = 1$ then $n \leq \frac{card(L(W_n)) - card(W_n) + 3}{2}$

Proof. Let v be a vertex of V_n and $W_n(v, V_0) = \{w_1, ...w_n\}$ be an element of $\mathcal{W}_{\mathcal{MIN}}(v, V_0)$. We know from lemma 1 that $\{l_1..l_n\} \subseteq m(v)$ with $l_i \in Lu_{W_n}(w_i)$.
Let $Lu'(v, W_n)$ be a set of labels defined by $Lu'(v, W_n) = Lu(W_n)\backslash\{l_1..l_n\}$ and $L_r(W_n) = L(W_n)\backslash Lu(W_n)$. We can remark that:
- for $w_i \in W_n$, we cannot have simultaneously $card(L_r(W_n) \cap m(w_i)) = 0$ and

$card(Lu_{W_n}(w_i)\backslash\{l_i\}) = 0$, since $card(m(w_i)) \geq 2$.
- $L(W_n) = \{l_1, ..., l_n\} \cup Lu'(v, W_n) \cup L_r(W_n)$.
As $\{l_1, ..., l_n\}, Lu'(v, W_n)$ and $L_r(W_n)$ are disjoint, we have
$card(L(W_n)) = n + card(Lu'(v, W_n)) + card(L_r(W_n))$.
Then, to prove proposition 6 we consider two cases:
-1- if $\forall w_i \in W_n, Lu_{W_n}(w_i)\backslash\{l_i\} \neq \emptyset$ then $card(L(W_n)) \geq n + card(Lu'(W_n)) \geq 2n$
-2- if $\exists w_k \in W_n$, s.t. $Lu_{W_n}(w_k)\backslash\{l_k\} = \emptyset$.
Let $T = \{v, W_n\backslash\{w_k\}\}$. We have $L(T) = L(W_n)$ and $card(T) = card(W_n)$. As W_n is a subset of V_0 then by proposition 4, we must have: $card(m(w_k)) \geq card(m(v)) \geq n$. Thus, from the hypothesis on w_k, $m(w_k)$ must contain at least $(n-1)$ labels distinct from $\{l_1, ..., l_n\}$.
Let $P = \{v, w_k\}$. Either $L(P) = L(W_n)$ and $n = 2$, or $L(P) \subset L(W_n)$, and, as $card(L(P)) \geq 2n-1$, we must have (a): $card(L(W_n)) \geq 2n$.
W_n cannot contain more that one vertex w such that $m(w) \subseteq L(P)$, to satisfy the minimality hypothesis on W_n. Thus, at least $card(W_n) - 2$ vertices of W_n contain one label or more not in L(P). Then we have :
$card(L(W_n)) \geq card(W_n) - 2 + card(L(P)) \geq card(W_n) - 2 + (2n-1)$. □

Using proposition 6 we have:

Corollary 2. *If* $2n > card(L)$ *then* V_n *is empty.*

3.4 Construction of a L_connected planar graph when $card(L) < 9$

Let $G_0 = (L, V_0, E_0)$ a planar L_connected graph. To extend G_0 with the vertices of V_i with $i > 1$, we will use the following remark:

Remark 5. Let V a set of labelled vertices, W a subset of V and $G(L, W, E)$ a L_connected labelled graph. If a vertex v of $V\backslash W$ is L_connectable to a subset W' of W, then the set of edges $E' = \{(v, w') | w' \in W'\}$ is such that $G(L, W \cup \{v\}, E \cup E')$ is a L_connected labelled graph.

Properties on the insertion of vertices of V_2

Lemma 3. *Let* $G_0(L, V_0, E_0)$ *be a L_connected planar graph and* v *an element of* V_2*. If* $\mathcal{W}_{\mathcal{MIN}}(v, V_0)$ *contains a set of vertices* $W_2 = \{w_1, w_2\}$ *such that* (w_1, w_2) *is an edge of* E_0*, then* v *can be inserted in* $G_0(L, V_0, E_0)$ *while keeping planarity and L_connectivity.*

Proof. The insertion of v creates a path $p(v) = \{(w_1, v), (v, w_2)\}$, parallel to the edge (w_1, w_2) of E_0. If $G_0(L, V_0, E_0)$ is planar then $G(L, V_0 \cup \{v\}, E_0 \cup \{p(v)\})$ is planar. □

Corollary 3. *Let* $G_0(L, V_0, E_0)$ *be a L_connected planar graph. If* $card(V_0) \leq 4$ *then all the vertices of* V_2 *can be inserted in* $G_0(L, V_0, E_0)$ *while keeping planarity and L_connectivity.*

Proof. If $card(V_0) \leq 4$, the complete graph $K_{card(V_0)}$ on V_0 is planar and L_connected. Thus, using lemma 3, all the vertices v of V_2 can be inserted in it without breaking the planarity and the L_connectivity. As $G_0(L, V_0, E_0)$ is a subgraph of $K_{card(V_0)}$, its extension is also a subgraph of the extension of $K_{card(V_0)}$.
Thus, if $E' = \{p(v) | v \in V_2$ and $p(v) = \{(w_1, v), (v, w_2)\}$ with $\{w_1, w_2\} \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)\}$, the graph $G(L, V_0 \cup V_2, E_0 \cup E')$ is planar and L_connected. □

Corollary 4. *If V is such that $card(V_0) = 2$, then $\mathcal{G}_{\mathcal{P}}^{\mathcal{V}}(L, V)$ is not empty for any L.*

Properties on the insertion of vertices of V_3

Lemma 4. *Let $G_0(L, V_0, E_0)$ be a L_connected planar graph. If V_3 contains a vertex v such that there is $W_3 = \{w_1, w_2, w_3\} \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$ with $card(L(W_3)) < 7$, then all the vertices v' of V_3 such that $W_3 \in \mathcal{W}_{\mathcal{MIN}}(v', V_0)$ can be inserted in G_0 while keeping the planarity and the L_connectivity on G_0.*

Proof. Let v and v' be two distinct vertices of V_3 associated to W_3. We first insert v in G_3 by adding three edges $(v, w_1), (v, w_2), (v, w_3)$.
-1- Suppose that $\forall w_i \in W_3$, $Lu_{W_3}(w_i) \backslash m(v) \neq \emptyset$. Then $\forall w_i \in W_3, card(m(w_i)) = 2$ and $m(w_i) \in Lu(W_3)$. To satisfy the hypothesis on W_3, we must have $card(L(\{v, v'\})) < 5$. As $card(m(v)) \geq 3$ and $card(m(v') \geq 3$ and v and v' are distinct, $m(v')$ has exactly one label l' not in $m(v)$. Let $w_i \in W_3$ s.t. $l' \in m(w_i)$. Then v' is L_connectable to $\{v, w_i\}$ and can be inserted in G_0 by creating a path (v, v', w_i) parallel to the edge (v, w_i) (cf. figure 7 A). This operation leads to a planar L_connected graph.

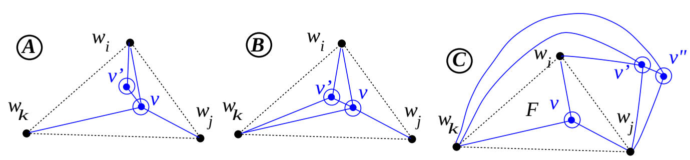

Fig. 7. Insertion of v and v' vertices of V_3. A: v' has one label l' not in $m(v)$ and belonging to $m(w_i)$. B: v and v' have one label in common belonging to $m(w_j)$. C: v and v' are label_disjoint and v'' and v' have at least one label in common.

-2- Suppose that $\exists w_j \in W_3$, s.t. $Lu_{W_3}(w_j) \backslash m(v) = \emptyset$. Then v and v' have the label of $Lu_{W_3}(w_j) \cap m(v)$ in common. Thus v' is L_connectable to $\{v\} \cup (W_3 \backslash \{w_j\})$ and v' can be inserted on the edge (w_j, v) and connected to the vertices w_i of W_3 with $i \neq j$ by adding the edges (v', w_i) (cf. figure 7 B). The resulting graph is planar and L_connected. □

Using a similar reasoning than for the corollary 3, we have:

Corollary 5. *Let $G_0(L, V_0, E_0)$ be a L_connected planar graph. If $card(V_0) \leq 4$ then all the vertices v of V_3 such that there is $W_3 = \{w_1, w_2, w_3\} \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$ with $card(L(W_3)) < 7$ can be inserted in $G_0(L, V_0, E_0)$ while keeping planarity and L_connectivity.*

Properties on the existence of V_4

Lemma 5. *Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) < 9$ and $card(V_0) \neq 4$, then V_n is empty for $n \geq 4$.*

Proof.
- if $n > 4$, using proposition 6, V_n is empty.
- if $n = 4$ and $card(V_0) < 4$, V_4 is empty by definition.
- if $n = 4$ and $card(V_0) > 4$. Suppose that $\exists v \in V_4$ and $W_4 \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$. To satisfy the hypothesis on the cardinality of V_0, we must have $card(L(W_4)) \leq 7$. Using proposition 6, we deduce that V_n is empty for $n \geq 3$. □

Extending G_0 when $card(L) < 9$

Lemma 6. *Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(V_0) = 3$ and $card(L) < 9$, then $\mathcal{G}_{\mathcal{P}}^{\mathcal{V}}(L, V)$ is not empty.*

Proof. Let $G_0(L, V_0, E_0)$ a planar L_connected graph. As $card(V_0) = 3$, $V = V_0 \cup V_2 \cup V_3$.

- From corollary 3, we can insert in G_0 the vertices of V_2 without breaking the planarity and the L_connectivity.

- From corollary 5, when $card(L) < 7$, we can insert in G_0 any vertex of V_3 without breaking the planarity and the L_connectivity.

- If $card(L) \geq 7$, then for any $v \in V_3$, $W_3 = V_0 = \{w_1, w_2, w_3\} \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$. We know by lemma 1 that v is such that $m(v) \cap Lu_{V_0}(w_j) \neq \emptyset$ for $j = 1, 2, 3$.

We take a vertex v of V_3 and insert v in G_0 by adding three edges $(v, w_1), (v, w_2), (v, w_3)$ inside the face $F = (w_1, w_2, w_3)$. This leads to a planar L_connected graph.
If V_3 contains another vertex v' then we have two cases to consider:
-1- there is $j \in \{1, 2, 3\}$ such that $m(v) \cap Lu_{V_0}(w_j) = m(v') \cap Lu_{V_0}(w_j)$. Then v' is L_connectable to $\{v\} \cup (W_3 \backslash \{w_j\})$ and v' is inserted on the edge (w_j, v) and connected to the vertices w_i of W_3 with $i \neq j$ by adding the edges (v', w_i) (cf. case (B) of figure 7).
-2- for any $j \in \{1, 2, 3\}$ $m(v) \cap Lu_{V_0}(w_j)$ and $m(v') \cap Lu_{V_0}(w_j)$ are distinct. In this case three edges $(v', w_1), (v', w_2)$ and (v', w_3) are added outside the face $F = (w_1, w_2, w_3)$ (cf. case (C) of figure 7).
Suppose that V_3 contains another vertex v''. As $card(L) < 9$, we can not have **-2-** for v, v' and v'' (for any $j \in \{1, 2, 3\}$ $m(v) \cap Lu_{V_0}(w_j)$, $m(v') \cap Lu_{V_0}(w_j)$ and $m(v'') \cap Lu_{V_0}(w_j)$ are distinct). Then, as the hypothesis of **-2-** is satisfied for v and v', the vertex v'' will be inserted either inside F or outside F and connected to v or v' as in **-1-**.
Thus all the vertices of V_3 can be inserted in G_0 and the resulting graph is planar and L_connected. □

Lemma 7. *Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) < 8$ and $card(V_0) = 4$ then $\mathcal{G}_{\mathcal{P}}^{\mathcal{V}}(L, V)$ is not empty.*

Proof. Let $G_0(L, V_0, E_0)$ be a planar L_connected graph.
From proposition 6, V_n is empty when $n > 3$. Thus $V = V_0 \cup V_2 \cup V_3$.
- From corollary 3 we can insert in G_0 the vertices of V_2 without breaking the planarity and the L_connectivity.
- If $card(L) < 8$ and $card(V_0) = 4$, any $v \in V_3$ with $W_3 \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$ is such that $card(L(W_3)) < 7$. Then from corollary 5, we can insert in G_0 any vertex of V_3. □

By a similar reasoning we prove the following lemmas[2].

Lemma 8. *Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) = 8$ and $card(V_0) = 4$ then $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L, V)$ is not empty.*

Lemma 9. *Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) < 8$ and $card(V_0) = 5$ then $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L, V)$ is not empty.*

Lemma 10. *Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) = 8$ and $card(V_0) = 5$, $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L, V)$ is not empty.*

Lemma 11. *Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) < 9$ and $card(V_0) = 6$ then $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L, V)$ is not empty*

Using lemma 6 to 11 we obtain our main result:

Theorem 2. *When $card(L) < 9$ then for any set V of labelled vertices on L, $\mathcal{G}^{\mathcal{V}}_{\mathcal{P}}(L, V)$ is not empty.*

Then, using proposition 1, we have:

Corollary 6. *For any set of non empty distinct sets $X = \{X_1, .., X_k\}$ such that $k < 9$ there is an extended Euler diagram representing X.*

Interpreting this result with Johnson and Pollak's definition [JP87] of hypergraph planarity, remark 1 and the previous lemma, we obtain the following result on hypergraphs:

Corollary 7. *Any hypergraph having at most eight hyperedges is vertex-planar.*

Remark 6. Let us consider the set of vertices V of figure 5.
$V_0 = \{abd, ace\}$, $V_1 = \{ab, ac, ad, ae, bc, d\}$ and $V_2 = \{bcd\}$. $G(L, V_0, E)$ has only one edge (abd, ace) and the L_connected labelled graph built by inserting successively the vertices of V_2 and V_1 leads to an extended Euler diagram where each label correspond to a connected region without hole in it (cf. figure 8).

[2] to improve the readability of the paper, the proofs appear in annex

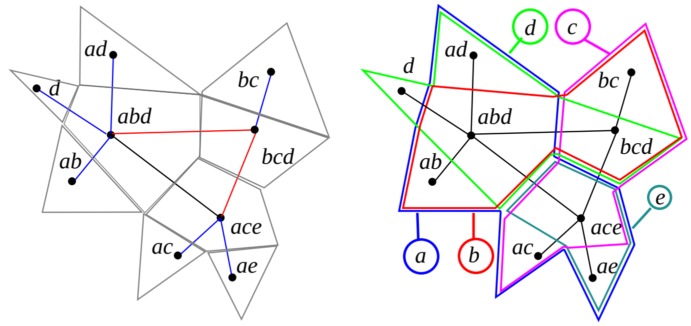

Fig. 8. The L_connected labelled graph built by computing the V_i and the corresponding extended Euler diagram

3.5 Conclusion

We have shown that there exists a planar L_connected graph for any collection of intersections between up to eight sets $\{X_1, ...X_k\}$. This planar L_connected graph can be used to build an extended Euler diagram representing $\{X_1, ...X_k\}$.
Interpreting our work using Johnson and Pollak's notion of planarity [JP87] we have shown in this paper that any hypergraph having at most eight hyperedges is vertex-planar.
We are currently working on the algorithm to produce the planar graph and the extended Euler diagram.
However, to reach the purposes described in the introduction, i.e. to create a semantically structured map of the results of a complex query, we have to address a few more task. Indeed, for most of the collections of intersections, there exists many planar graphs satisfying the constraint of L_connectivity, and the graph built from the proof may not be the most adapted to our purposes. Then at this graph level, we may have to introduce some graphical criterion to provide the user the most readable diagram. Moreover, we still have to find the best embedding according to visibility and usability criterion.

References

Bad90. A. Baddeley. *Human Memory, Theory and Practice.* Lawrence Erlbaum Associated Ltd. Publishers, U.K., 1990.

BETT99. G. Di Battista, P. Eades, R. Tamassia, and I.G. Tollis. *Graph Drawing: algorithms for the visualization of graphs.* Prentice Hall, 1999.

CCLC97. T. Catarci, M.F. Costabile, S. Levialdi, and C.Batini. Visual query systems for databases: a survey. *Journal of Visual Languages and Computing*, 8:215–260, 1997.

Con94. M. Consens. *Creating and filtering structural data visualizations using Hygraph patterns.* Tech. Rep. CSRI– 302, Toronto, Canada, February 1994.

CT95. T. Catarci and L. Tarantino. A hypergraph-based framework for visual interaction with databases. *Journal of Visual Languages and Computing*, 6(2):135–166, 1995.

Eul75. L. Euler. *Lettres à une Princesse d'Allemagne sur divers sujets de physique et de philosophie*, volume T. 2, letters 102-108. Berne, Société Typographique, 1775.

FH02. J. Flower and J. Howse. Generating Euler diagrams. In *Diagrams 2002*, pages 61–75, Callaway Gardens, Georgia, USA, April 2002.

GKH99. J. Gil, S. Kent and J. Howse. Formalizing Spider diagrams. In *IEEE Symposium on Visual Languages*, pages 130–137, Tokyo, September 1999.

JP87. D.S. Johnson and H. O. Pollak. Hypergraph planarity and the complexity of drawing Venn diagrams. *Journal of Graph Theory*, 11(3):309–325, 1987.

Kur30. K. Kuratowski. Sur le problème des courbes gauches en topologie. *Fundamenta Mathematicae*, 15:271–283, 1930.

LP97. O. Lemon and I. Pratt. Spatial logic and the complexity of diagrammatic reasoning. *Machine Graphics and Vision (Special Issue on Diagrammatic Reasoning)*, 6(1):89–109, 1997.

Rus01. F. Ruskey. A survey of Venn diagrams. *The electronic journal of combinatorics*, March 2001.

4 Annex

Lemma 8. Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) = 8$ and $card(V_0) = 4$ then $\mathcal{G}_{\mathcal{P}}^{\mathcal{V}}(L, V)$ is not empty.

Proof. $V = V_0 \cup V_2 \cup V_3 \cup V_4$.
-1- Suppose $v \in V_4$ and $W_4 \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$. Then, as $4 = n > \frac{card(L(W_4)) - card(W_4) + 3}{2}$, using proposition 6 (b) we have: $\forall w_i \in V_0$, $card(Lu_{V_0}(w_i)) > 1$. As $card(L) = 8$, $\forall w_i \in V_0$, $m(w_i) = Lu_{V_0}(w_i)$ and G_0 does not contain any edge.
We insert v in G_0 and the graph $G(L, V_0 \cup \{v\}, E \cup \{(v, w_1), (v, w_2), (v, w_3), (v, w_4)\})$ is planar and L_connected.
Using proposition 4 (2), we have: any subset W of V containing v and such that $L(W) = L(V_0)$ must have more than four vertices. Thus, any vertex v' belonging to V_i with $i \leq 4$ is such that $m(v')$ has exactly one label l' not in $m(v)$. Let $w_i \in V_0$ s.t. $l' \in m(w_i)$. The insertion of v' in G_0 is done by creating a path (v, v', w_i) parallel to the edge (v, w_i) (cf. left graph of figure 9). This operation leads to a planar L_connected graph.

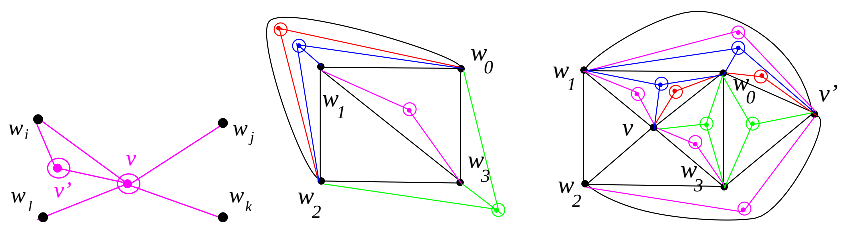

Fig. 9. $card(L) = 8$ and $card(V_0) = 4$. On the left: the insertion of elements of V_4 consists in creating four pink edges for the first element v and two pink edges for the next one. On the middle: the graph K_4 is replaced by the rightmost graph and the insertion of the vertices of V_2 (resp. V_3) are transformed (the correspondence is given by the colors).

-2- Suppose V_4 is empty. The complete graph K_4 is planar and contains G_0.
Using corollaries 3 and 5, we know that any element of V_2 and V_3 can be inserted in K_4 while keeping planarity when every vertex v of V_3 is L_connectable to a set of vertices W_3 satisfying $L(W_3) < 7$.
Let us now suppose that V contains two vertices of V_3, v and v', such that $W_3 = W_3(v, V_0) = W_3(v', V_0)$ satisfies $card(L(W_3)) = 7$. We note w_0 the vertex of $V_0 \backslash W_3$ and $l_0 = Lu_{V_0}(w_0)$.
We have two cases to consider:
A. $\forall v, v' \in V_3$ associated to W_3, $\exists w_j \in W_3$ with $Lu_{W_3}(w_j) \cap m(v) = Lu_{W_3}(w_j) \cap m(v')$. Then all the vertices of V_3 associated to W_3 are inserted in G_0 as in figure 7 (B).
B. $\exists v, v' \in V_3$ and $W_3 = \{w_1, w_2, w_3\} \subset V_0$ such that $Lu_{W_3}(w_j) \cap m(v) \neq Lu_{W_3}(w_j) \cap m(v')$ for $j = 1, 2, 3$. Then we have:

- $\boldsymbol{card(L(\{v, v'\}) = 6}$ **and** $\boldsymbol{card(m(v)) = card(m(v')) = 3}$:
  if $card(L(\{v, v'\}) = 7$, $L(\{v, v'\}) = L(W_3)$ which contradicts proposition 4 2.
  Let us note $m(v) = \{l_1, l_2, l_3\}$ and $m(v') = \{l'_1, l'_2, l'_3\}$.

- $\boldsymbol{\forall w_j \in W_3, card(Lu_{W_3}(w_j)) = 2}$ and $m(w_j) = \{l_j, l'_j, l_c\}$ with $\{l_c\} = L(W_3)\backslash Lu(W_3)$: suppose $\exists w_i \in W_3$ s.t. $card(Lu_{W_3}(w_i)) = 3$, then, as $card(L(W_3)) = 7$, $\forall w_j \in W_3\backslash\{w_i\}$, $card(m(w_j)) = 2$. Thus $T = \{v, v', w_i\}$ is such that $L(T) = 7$ and $T >_L W_3$ which contradicts proposition 4 2.
- $\boldsymbol{m(w_0) \subseteq \{l_0, l_i, l'_j\}, 1 \leq i \neq j \leq 3}$, using proposition 4 and the fact that v and v' belong to V_3.

We suppose without loss of generality that $m(w_0) \cap m(w_2) = \emptyset$, and consider the middle graph of figure 9. We cannot insert v and v' in K_4 by connecting v and v' with the vertices of W_3 without breaking the planarity of the resulting graph. We then replace K_4 by the rightmost graph G' of figure 9, deleting the edges (w_0, w_2) and (w_1, w_3) and connecting w_0 with v and v'. We now have to show that the vertices of V_2 associated to $\{w_1, w_3\}$ and to $\{w_0, w_2\}$ can be inserted between two adjacent vertices of G' and that the vertices of V_3 associated to the faces (w_0, w_2, w_3), (w_0, w_1, w_2) and (w_2, w_1, w_3) can be inserted inside faces of G'.

Let w be a vertex to be inserted in G' ($w \in V\backslash(V_0 \cup \{v, v'\})$). We have:

- $\boldsymbol{card(m(w)) < 4}$, otherwise, we can find a subset T of V such that $L(T) = L(V_0)$ and $T >_L V_0$:
  If $card(m(w)) \geq 4$ and $m(w)$ contains l_0, we take $T = \{w, w_1, w_2, w_3\}$.
  If $card(m(w)) \geq 4$ and $m(w)$ does not contain l_0 then we can find w_i and w_j in W_3 such that $T = \{w, w_i, w_j, w_0\}$ satisfies $L(T) = L(V_0)$ and $T >_L V_0$.
- $\boldsymbol{m(w)}$ cannot contain both $\boldsymbol{l_0}$ and $\boldsymbol{l_c}$ or both $\boldsymbol{l_i}$ and $\boldsymbol{l'_i}$, $\boldsymbol{i > 0}$, otherwise , we can find a subset T of V such that $L(T) = L(V_0)$ and $T >_L V_0$, which contradicts proposition 4.

Then, $m(w)$ contains one or two labels taken from $\{l_i, l'_j\}$ and at most one label from $\{l_0, l_c, l_k, l'_k\}$ with $1 \leq i \neq j \neq k \leq 3$, and:
(a) If $m(w) \subset L(W_3)$, then, w is L_connectable either to $\{v, w_2\}$, to $\{v, w_1\}$, to $\{v, w_3\}$, to $\{v', w_1\}$ or to $\{v', w_2\}$.
(b) If $m(w) = \{l_0, l_i, l'_j\}$ with l_i or $l'_j \in m(w_0)$ then $w \in V_2$ and w is L_connectable to $\{w_0, v\}$ or to $\{w_0, v'\}$
(c) If $m(w) = \{l_0, l_i, l'_j\}$ with $\{l_i, l'_j\} \cap m(w_0) = \emptyset$ then $w \in V_3$ and w can be inserted in one of the triangular faces of G' adjacent to w_0.
Thus the insertion of any $w \in V\backslash(V_0 \cup \{v, v'\})$ in G' is possible as in figure 9 and the resulting graph is planar and L_connected. □

Lemma 9. Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) < 8$ and $card(V_0) = 5$ then $\mathcal{G}_{\mathcal{P}}^{\mathcal{V}}(L, V)$ is not empty.

Proof. Let $G_0(L, V_0, E_0)$ be a L_connected planar graph. By lemma 5, $V = V_0 \cup V_2 \cup V_3$. V_3 is empty: indeed, if $\exists v \in V_3$ then for $W_3 \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$, we must have $card(L(W_3)) < 6$ because $card(V_0) = 5$ and $card(L(W_3)) \geq 6$ by proposition 6 (a).
Then we have to show that any vertex of V_2 can be inserted in G_0 while keeping the L_connectivity and the planarity.
As $card(Lu(V_0)) \geq 5$ and $card(L) \leq 7$, L contains either one label l or two labels l and l' which are not in $Lu(V_0)$. We consider three cases:

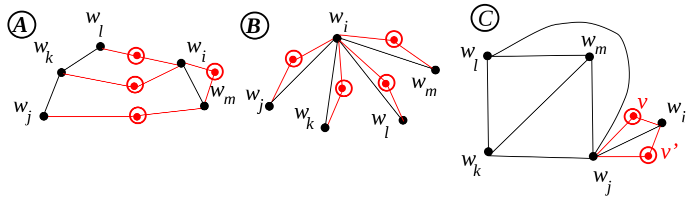

Fig. 10. $card(L) < 8$ and $card(V_0) = 5$. the three cases. The insertion of any vertex of V_2 consists in creating one of the red paths.

A. $L = Lu(V_0) \cup \{l, l'\}$ and no vertex of V_0 has both l and l' in its set of label. Then V_0 contains two label_disjoint vertices w_i and w_j. As l and l' play the same role, we may consider without loss of generality that w_j, w_k, w_l have l in their sets of labels and that w_i and w_m have l' in their sets of labels. Then the graph G_0 having the set of edges $E_0 = \{(w_j, w_k), (w_k, w_l), (w_i, w_m)\}$ is L_connected and planar. Any vertex v of V_2 has a set of labels of the form $m(v) = \{l, l_{v,v=i,m}\}$, $m(v) = \{l', l_{v,v=j,k,l}\}$ or $m(v) = \{l_i, l_m\}$ and can be inserted in G_0 by adding one of the red paths of figure 10 A.
B. $L = Lu(V_0) \cup \{l, l'\}$ and a vertex w_i of V_0 contains l and l' in its set of labels. Then the graph G_0 having the set of edges $E_0 = \{(w_i, w)|w \neq w_i and w \in V_0\}$ is L_connected and planar. Any vertex v of V_2 can be inserted in G_0 by adding one of the red paths of figure 10 B.
C. $L = Lu(V_0) \cup \{l\}$. There is one vertex w_i of V_0 such that $Lu_{V_0}(w_i) = \{l_{i1}, l_{i2}\}$ and the other vertices w_j of V_0 are such that $m(w_j) = \{l_j, l\}$ with $Lu_{V_0}(w_j) = \{l_j\}$.
Then $G_0(L, V_0, E_0)$ is a subgraph of the graph formed by a complete graph K_4 on the vertices of $V_0 \backslash \{w_i\}$ and by an edge (w_i, w_j) (cf. figure 10 C.). If $m(w_i)$ contains l, then V_2 is empty. In the other case, V_2 can be composed of two vertices v and v' with $m(v) = \{l_{i1}, l\}$ and $m(v') = \{l_{i2}, l\}$. The insertion of these vertices in G_0 can be made by adding the paths (w_j, v, w_i) and (w_j, v', w_i). □

Lemma 10. Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) = 8$ and $card(V_0) = 5$, $\mathcal{G}_{\mathcal{P}}^{\mathcal{V}}(L, V)$ is not empty.

Proof. Any vertex of V_0 has at least two labels in its set of labels. Then, as $card(V_0) = 5$ and $card(L) = 8$, we have $card(Lu(V_0)) \leq 7$ and in any subset W of V_0 containing three vertices there is at least one vertex w_i such that $card(Lu_{V_0}(w_i)) = 1$. Moreover, using the results of proposition 4 (a), any $v \in V_3$ is such that $card(L(W_3)) = 6$ for $W_3 \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$ and $m(v)$ cannot contain only labels from $Lu(V_0)$. We have three cases to consider:
A. $L = Lu(V_0) \cup \{l\}$. We know that if $v \in V_3$ then $l \in m(v)$ and $l \in Lu(W_3)$. Thus there are 2 vertices w_0 and w_1 in W_3 whose sets of labels do not contain l. Figure 11 (A) shows the unique case satisfying such constraints. Any other vertex v' of V_3 has l in common with v and can be inserted in W as shown in figure 4 (B).
Let $W' = V_0 \backslash \{w_0, w_1\}$. If $w_i \in W'$ then $m(w_i) = \{lu(w_i), l\}$. To L_connect l, we need to add 2 edges between the 3 vertices of W'. Moreover, V cannot contain a vertex $v \in V_2$ L_connectable to $W_2 \subset W'$ otherwise proposition 4 (a) would not be satisfied. Then there are two vertices of V_0 which are not connected neither by an edge of G_0 nor by a path of V_2 or V_3. Thus the resulting graph is planar.

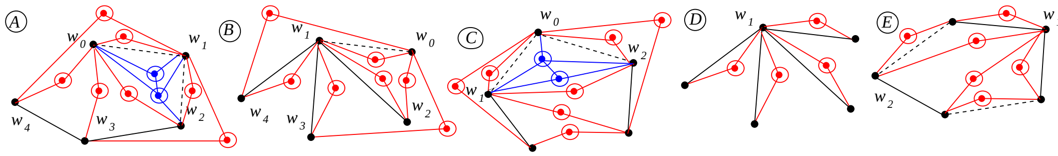

Fig. 11. $card(L) = 8$, $card(V_0) = 5$. A: $card(Lu(V_0)) = 7$. When $m(w_0)$ and $m(w_1)$ do not contain the common label l, the elements of V_3 can be inserted using the blue paths. The insertion of any element of V_2 consists in creating one of the red paths. B and C: $card(Lu(V_0)) = 6$. When the dashed line exists, $m(w_0)$ contains either l or l'. Otherwise, $m(w_0) \cap \{l, l'\} = \emptyset$. The blue (resp. red) paths correspond to the insertion of vertices of V_3 (resp. V_2). B: $m(w_1)$ contains l and l'; C: $m(w_1)$ contains l' and $m(w_2)$ contains l. D and E: $card(Lu(V_0)) = 5$. D: $m(w_1)$ contains $\{l, l', l''\}$; E: each $m(w_i)$ contains at most two labels among l, l' and l''. $m(w_1)$ contains labels l and l' and $m(w_2)$ contains label l''.

B. $L = Lu(V_0) \cup \{l, l'\}$.There is only one vertex w_0 of V_0 such that $card(m(w_0) \cap Lu(V_0)) = 2$. If V_3 contains a vertex v then $W_3 = \{w_0, w_i, w_j\} \in \mathcal{W}_{\mathcal{MIN}}(v, V_0)$ and we cannot have $Lu_{V_0}(w_k) \subset m(v)$ with $k = i$ or $k = j$. Otherwise, i.e. if $Lu_{V_0}(w_i) \subset m(v)$ then , either $card(m(w_i)) = 2$ and $T = \{w_0, v, w_j\}$ would be such that $T >_L W_3$ or $m(w_i) = \{Lu_{V_0}(w_i), l, l'\}$ and v is not in V_3. Thus $m(v) = \{lu(w_0), l, l'\}$ and V_3 contains at most two vertices which can be inserted in the same face because their associated sets of labels contain $\{l, l'\}$ (cf. figure 11 (C)) and the insertion of vertices of V_3 in G_0 leads to a L_connected planar graph.
Let us consider now the insertion of the vertices of V_2. A vertex v belonging to V_2 is such that $m(v) = \{l_1, l_2\}$ with either $l_1 \in Lu(V_0)$ and $l_2 \in \{l, l'\}$ or $l_1 \in Lu_{V_0}(w_0)$ and $l_2 \in Lu(V_0) \backslash Lu_{V_0}(w_0)$.
- If there is a vertex w_1 in V_0 such that $m(w_1)$ contains l and l', we have the configuration of figure 11 (B): a vertex v of V_2 is inserted in $G_0(L, V_0, E_0)$ by adding either a path connecting w_0 and another vertex of V_0 or a path connecting w_1 and another vertex of V_0.
- If no vertex of V_0 contains the two labels l and l' in its set of label, l (resp. l') belongs to at least two sets of labels associated to vertices of V_0. Let w_1 and w_2 be two vertices of V_0 such that $l' \in m(w_1)$ and $l \in m(w_2)$, we have the configuration of figure 11 (C) for the insertion of the vertices of V_2 in $G_0(L, V_0, E_0)$.
In all the cases the resulting graph is planar.
C. $L = Lu(V_0) \cup \{l, l', l''\}$. The labels l, l' and l'' appear in at least two vertices of V_0 and any vertex of V_0 has at least one of these labels in its set of labels. Then, as $card(V_0) = 5$, there is a vertex w_1 in V_0 which contains at least two of the three common labels in its set of labels. Thus, to satisfy the hypothesis on V_0, V_3 must be empty.
- If a vertex w_1 of V_0 contains these three labels in its set of labels. Then $G_0(L, V_0, E_0)$ where $E_0 = \{(w_1, w) | w \in V_0 \backslash \{w_1\}\}$ is planar and L_connected (cf. figure 11 (D)) and the elements of V_2 are inserted by connecting w_1 and another vertex of V_0.
- If all the vertices of V_0 contain at most two of the three common labels in their sets of labels, we have the configuration of figure 11 (E) and a way to insert the vertices of V_2 in $G_0(L, V_0, E_0)$. □

Lemma 11. Let V a set of labelled vertices on L and $V_0 \in \mathcal{T}_{max}(V)$. If $card(L) < 9$ and $card(V_0) = 6$ then $\mathcal{G}_{\mathcal{P}}^{\mathcal{V}}(L, V)$ is not empty.

Proof. As $card(V_0) = 6$, we must have $card(L) \geq 7$. Let $G_0(L, V_0, E_0)$ be a L_connected planar graph.
When $card(L) = 7$ then by lemma 2 we have $V = V_0$ and $\mathcal{G}_{\mathcal{P}}^{\mathcal{V}}(L, V)$ is not empty.
When $card(L) = 8$, using the same reasoning than in the previous lemma, V_3 cannot contain any vertex. We know also that any element v of V_2 is such that $m(v) = \{l\} \cup \{l_i\}$ where $l_i \in Lu(V_0)$. We have two cases to consider:
A. $card(Lu(V_0)) = 7$ and a label l belongs to the sets of labels of at least seven vertices of V_0. If l is present in all the $m(w)$ for $w \in V_0$, V_2 is empty. Otherwise, we are in the case of figure 12 (A). Let $w_1 \in V_0$ with $l \notin m(w_1)$. To satisfy the hypothesis on V_0, an element v of V_0 must be such that $m(v) \cap m(w_1) \neq \emptyset$ and can be inserted by adding a path between w_1 and another vertex of V_0.
B. $card(Lu(V_0)) = 6$ and each vertex of V_0 contains either l or l' in its set of labels. If V_0 contains a vertex w_1 such that the two labels l and l' belong to $m(w_2)$, then we have the configuration of figure 12 (D). In the other cases, we are in case (B) or (C) of figure 12. $\square$

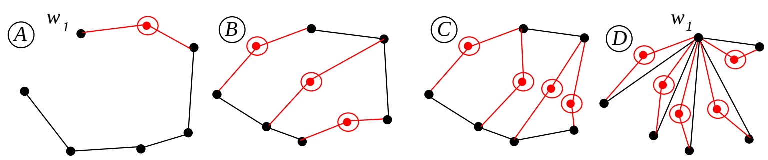

Fig. 12. $card(V_0) = 6$ and $card(L) = 8$. The black edges represent $G(L, V_0, E)$ and the insertion of vertices of V_2 are made using the red paths. A: $card(Lu(V_0)) = 7$ and $m(w_1) \cap \{l\} = \emptyset$; B: $card(Lu(V_0)) = 6$, three vertices of V_0 contain l in their sets of labels and the three others contain l' in their sets of labels; C: $card(Lu(V_0)) = 6$, four vertices contain l in their sets of labels the two others contain l' in their sets of label; D: $card(Lu(V_0)) = 6$ and $m(w_1)$ contains $\{l, l'\}$.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399