G. SAUCIER

Un algorithme efficace recherchant l'isomorphisme de 2 graphes

Revue française d'informatique et de recherche opérationnelle. Série rouge, tome 5, n° R3 (1971), p. 39-51

<http://www.numdam.org/item?id=M2AN_1971__5_3_39_0>

R.I.R.O.
5e année, Nº R-3, 1971, p. 39-51)

UN ALGORITHME EFFICACE RECHERCHANT L'ISOMORPHISME DE 2 GRAPHES

par G. SAUCIER [1]

Résumé. — *Cet algorithme permet de déterminer si 2 graphes G et G' sont isomorphes. Dans cette méthode, on évite l'étude indépendante et systématique des 2 graphes inexploitables pour des graphes importants, en cherchant dès le début de l'algorithme à mettre en correspondance 2 sommets de G et G'. Cette hypothèse quasi initiale permet d'utiliser des fonctions de structure plus élaborées et de conclure plus rapidement. Le gain total en nombre d'opérations, par rapport aux autres méthodes connues est très important.*

INTRODUCTION

Dans un premier temps, une hypothèse de niveau 1 est définie pour un graphe G. Cette hypothèse permet d'utiliser une représentation donnant immédiatement certains renseignements. Un algorithme d'affinement de partitions est exposé brièvement dans un second temps. Il est à noter qu'une technique similaire a été développée, indépendamment et dans le même temps, par D. G. Corneil et C. C. Gotlieb, mais que les méthodes résultantes restent très différentes. En particulier, aucune conjecture n'étant espérée dans notre méthode, un compromis efficace est fait entre l'étude exhaustive des graphes (trop longue et rendondante dans [1]) et l'utilisation d'hypothèses.

Les graphes sont supposés connexes et non marqués. L'étude des graphes non connexes nous amènera à étudier les différentes composantes connexes, l'étude des graphes marqués (sens, pondération, etc...) sera facilité par ce supplément d'information [2].

La définition de l'isomorphisme de 2 graphes est supposée connue et ne sera pas reprise ici.

(1) Université Scientifique et Médicale de Grenoble. Service de Mathématiques Appliquées.

I. HYPOTHESE DE NIVEAU 1 DANS UN GRAPHE $G(X, \Gamma)$

1.1. Définition

Soit un graphe $G(X, \Gamma)$ où X est l'ensemble des sommets, Γ la relation d'adjacence.

Une hypothèse de niveau 1 consiste à distinguer dans X un sommet S.

1.2. Structure en couches associée à un sommet S

S étant un sommet distingué dans X, nous associons à $G(X, \Gamma)$ la représentation en couches $\{ C_0, \dots C_i, \dots \}$ définie par

$$S \in C_0$$

$$x \in C_i \Leftrightarrow d(S, x) = i$$

d étant la distance au sens habituel (S, x) dans G.

Dans cette représentation, chaque sommet peut être relié à des sommets de la couche précédente, à des sommets de la même couche, à des sommets de la couche suivante.

1.3. Triplets associés aux sommets $\{ X \}$ dans une représentation en couches

A chaque sommet on associe une suite de 3 nombres (α, β, γ) où α est le nombre de liaisons avec la couche précédente, β est le nombre de liaisons avec la *même couche, γ est le nombre de liaisons avec la couche suivante.*

EXEMPLE

Soit le graphe suivant dont tous les sommets sont de degré 3 (fig. 1).

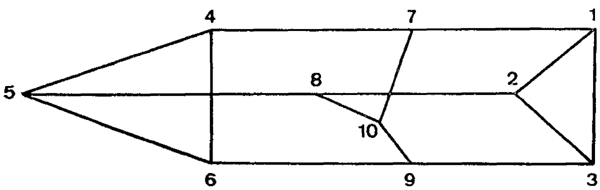

Figure 1

Si 5 est choisi comme sommet S, nous avons la structure et les triplets suivants (fig. 2).

4 : (1,1,1)	7 : (1,1,1)	1 : (2,1,0)
6 : (1,1,1)	9 : (1,1,1)	3 : (2,1,0)
8 : (1,0,2)	2 : (1,0,2)	
	10 : (1,2,0)	

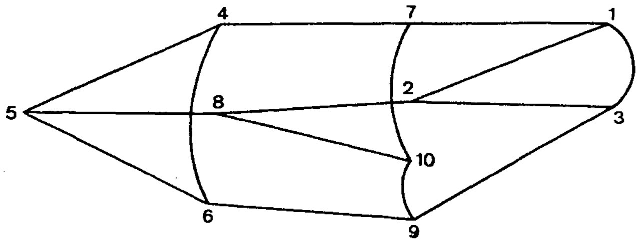

Figure 2

II. PROCEDES D'AFFINEMENT D'UNE PARTITION DE L'ENSEMBLE DES SOMMETS X

Soit un graphe $G(X, \Gamma)$ dans lequel on a distingué un sommet S. Nous avons une partition initiale des sommets en couches C_0, C_1 ... A l'intérieur d'une couche C_i, des distinctions peuvent être faites, soit à partir des triplets associés, soit en cours d'algorithme; la partition résultante sera notée P_i. Cet ensemble de partitions initiales $\{ P_i \}$ peut être affiné par étude des voisinages successifs par l'algorithme suivant :

α) Soit un ensemble de partitions $\{ P_i \}$ initiales. Numérotons les classes de chacune de ces partitions; c_0, c_1, c_2, ... A chaque classe, correspond, par exemple, un triplet défini au paragraphe précédent.

β) Associons à chaque sommet d'une couche une suite de 3 monômes (m_1, m_2, m_3). Le monôme m_1 sera constitué par la concaténation des classes de sommets de la couche précédente, adjacents à ce sommet. Le monôme m_2 traduira de même les liaisons avec les éléments d'une même couche, le monôme m_3 avec les éléments de la couche suivante.

Exemple fig. 2 *bis*.

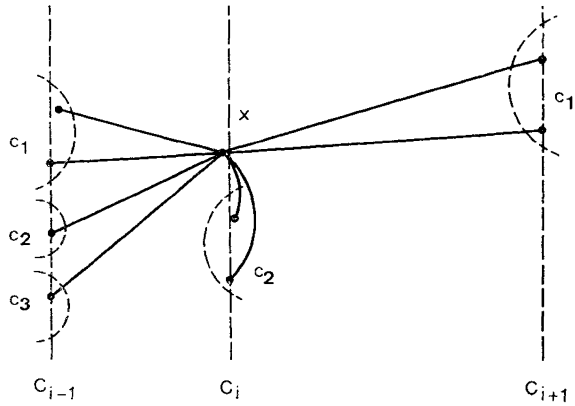

Figure 2 *bis*

à x sont associés les 3 monômes $(c_1c_1c_2c_3,\ c_2c_2,\ c_1c_1)$.

On définit ainsi une nouvelle partition $P'_j \leqslant P_j$ par la relation : 2 éléments sont dans une même classe si et seulement si les suites de 3 monômes (m_1, m_2, m_3) sont les mêmes pour ces 2 éléments.

γ) La partition relative à une couche, ayant été affinée par le processus précédent, les partitions relatives aux couches voisines sont examinées.

δ) Si aucun élément n'est plus possible, par le processus précédent, les dernières partitions trouvées constituent l'ensemble $\{ P'_i \}$.

Signification des partitions $\{ P_i' \}$

Elles traduisent les distinctions les plus fines que l'on peut faire, à partir des distinctions traduites par les partitions initiales, par étude des voisinages successifs.

EXEMPLE fig. 3.

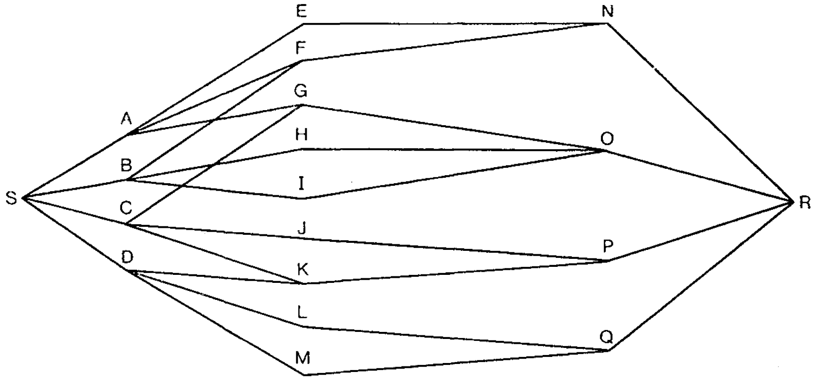

Figure 3

Les partitions $\{ P_i \}$ sont celles générées par les triplets associés.

L'exemple est traité dans le tableau de la page suivante.

REMARQUE

— Chaque classe finale est caractérisée par une suite de 3 monômes. Il est important, lorsque l'on opèrera sur 2 graphes, que cette suite de 3 monômes soit définie biunivoquement à partir des classes des partitions initiales; ce qui permettra d'établir une correspondance biunivoque entre les classes finales à partir de la correspondance biunivoque entre les classes initiales.

Il suffira pour cela de respecter la règle suivante :

Soit une partition P s'affinant en une partition $P' < P$.

En numérotant les classes de P', on respectera

• l'ordre des classes de P : seront d'abord considérés les sommets de la première classe de P, etc...;

	COUCHE 1	COUCHE 2	COUCHE 3	COUCHE 4
Partition initiale	ABCD C_1	EHIJLM, FGK C_1, C_2	[NPQ, O] C_1, C_2	[R] C_1
Monômes associés	A : $C_1C_2C_2$ B : $C_1C_1C_2$ C : $C_1C_2C_2$ D : $C_1C_1C_2$			
Nouvelle partition	AC, BD C_1, C_2			
Monômes associés		E : C_1, C_1 H : C_2, C_2 I : C_2, C_2 J : C_1, C_1 L : C_2, C_1 M : C_2, C_1 F : C_1C_2, C_1 G : C_1C_1, C_2 K : C_1C_2, C_1		
Nouvelle partition		EJ, HI, LM, FK, G C_1, C_2, C_3, C_4, C_5		
Monômes associés	A : $C_1C_4C_5$ B : $C_2C_2C_4$ C : $C_1C_4C_5$ D : $C_3C_3C_4$		N : C_1C_4, C_1 P : C_1C_4, C_1 Q : C_3C_3, C_1 O : $C_2C_2C_5$, C_1	
Nouvelles partitions	AC, B, D C_1, C_2, C_3		NP, Q, O C_1, C_2, C_3	

• au cours de l'éclatement d'une telle classe, on respectera un ordre lexicographique des variables (c_1, c_2, ..., c_n) figurant dans la suite des 3 monômes.

EXEMPLE

La partition initiale de la 2[e] couche :

$$\underset{c_1}{\text{EHIJLM}},\ \underset{c_2}{\text{FGK}} \text{ donnera à l'étape suivante}$$

la partition en classes ordonnées somme suit :

$$\underset{c_1}{\text{EJ}},\ \underset{c_2}{\text{LM}},\ \underset{c_3}{\text{HI}},\ \underset{c_4}{\text{G}},\ \underset{c_5}{\text{FK}};$$

en effet LM est caractérisé par c_2, c_1
en effet HI est caractérisé par c_2, c_2

III. ISOMORPHISME DE 2 GRAPHES

Considérons 2 graphes G et G' dont nous recherchons l'isomorphie éventuelle. La recherche proposée se fera par les étapes suivantes.

A) Étude préalable des 2 graphes

Cette étude rapide comprendra

— le dénombrement des sommets, des arêtes, l'établissement de la partition de l'ensemble des sommets suivant la valeur de leur degré (la généralisation aux graphes marqués tiendra compte des demi-degrés, etc...) et des boucles pouvant éventuellement figurer dans le graphe. Dans les cas les plus défavorables, cette partition se réduira à une seule classe.

Si les mêmes résultats sont trouvés pour G et G' l'étape suivante est considérée (les classes se correspondent biunivoquement, chaque classe étant caractérisée par les quantités telles que : degrés, nombre de boucles).

B) Deuxième étape

Considérons une classe C, ayant un nombre minimal d'éléments, de la partition précédemment trouvée de l'ensemble des sommets de G et l'image C' de cette classe dans G'.

Soit S un sommet de C;

$*$ choisissons un sommet S' dans C'; nous cherchons si S' peut être l'image de S par un isomorphisme cherché. Pour cela, étudions les structures en couches associées à S et S' dans G et G'.

α) *Ces structures en couches doivent avoir :*

· même nombre de couches,

· même nombre d'éléments dans une couche,

· même partition initiale des éléments d'une couche suivant la valeur du triplet associé. On numérotera de la même façon les classes correspondantes de G et G' ayant même valeur de ce triplet.

Si ces conditions ne sont pas remplies, retour à $*$ (nouveau choix de S').

β) *Les partitions initiales sont affinées* à l'aide de l'algorithme énoncé. S'il n'y a pas correspondance biunivoque entre les partitions obtenues à chaque étape (même ensemble de monômes caractéristiques pour les classes obtenues), retour à $*$ (choix de S'). Supposons cette correspondance possible. Si les classes ainsi obtenues se réduisent à un seul sommet, l'isomorphie cherchée est établie. Dans le cas contraire, on passe à γ.

γ) *Hypothèse d'ordre supérieur : identification progressive des classes de plus d'un élément.*

On considère dans la 1re couche (puis 2e, etc...) une classe de plus d'un élément dans G, et la classe image dans G'.

** Soient x et x' 2 éléments de ces 2 classes, que nous essayons de mettre en correspondance.

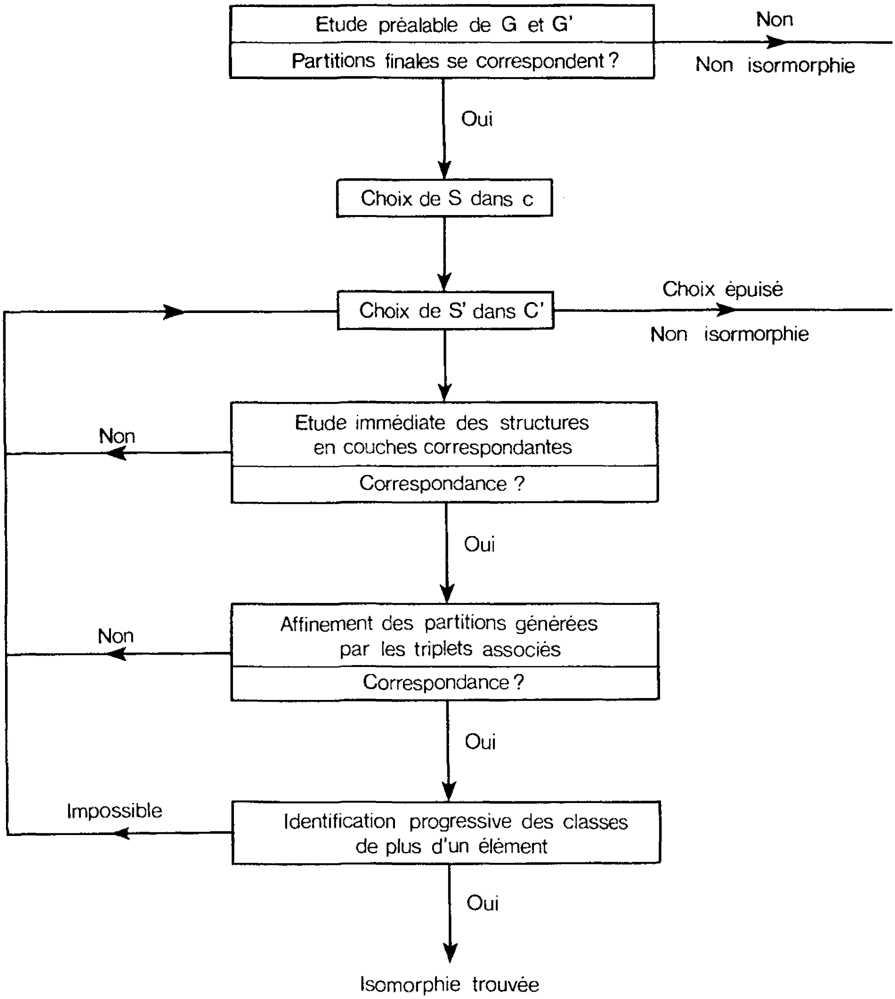

Figure 4

Ce choix correspond à un affinement des partitions dans G et G' (x et x' constituent une classe unique dans G et G' numérotée de la même façon dans G et G'). Ces partitions sont affinées par l'algorithme énoncé. Si les classes finales se correspondent biunivoquement, nous avons :

— soit résultat final cherché si ces classes se réduisent à un seul élément,

— soit passage à l'hypothèse suivante (2 éléments sont mis en correspondance dans 2 classes correspondantes).

Si les classes finales ne se correspondent pas, retour à l'hypothèse précédente : nouveau choix de x et x' ou retour à 2 classes précédemment considérées pour lesquelles les hypothèses possibles n'auront pas été épuisées.

Si les hypothèses de l'étape γ sont épuisées sans avoir trouvée une correspondance biunivoque cherchée, retour au choix de S' (*).

D'où l'organigramme suivant fig. 4.

EXEMPLE (Steen [3]).

Soient les 2 graphes suivants dont nous recherchons l'isomorphie fig. 5.

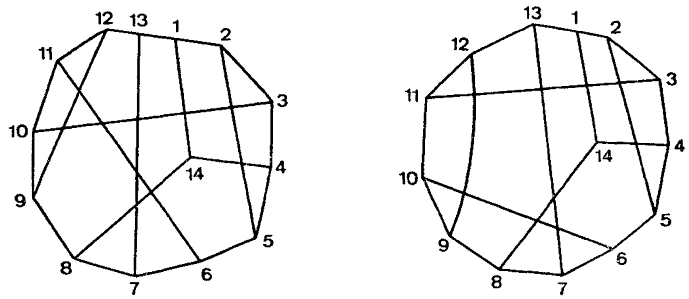

Figure 5

G et G' sont homogènes, chaque sommet étant de degré 3.

Choisissons le sommet 1 de G comme sommet S.

La représentation en couches associée est fig. 6 :

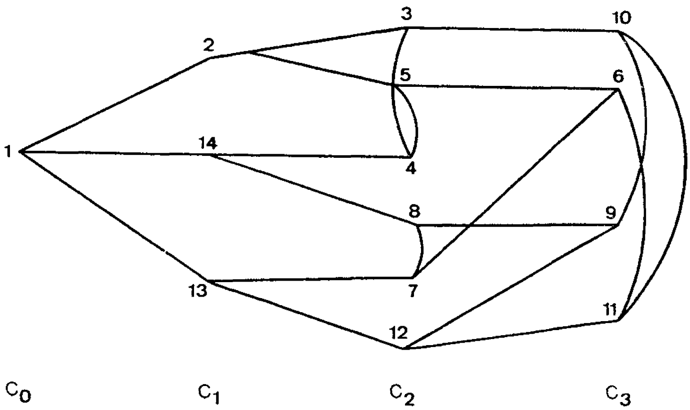

Figure 6

la partition des sommets suivant la valeur du triplet :

$$\begin{array}{cccccc} (2,14,13) & (3,5,7,8) & (4) & (12) & (10,11) & (6,9) \\ (1,0,2) & (1,1,1) & (1,0,2) & (1,2,0) & (1,2,0) & (2,1,0) \end{array}$$

Choix de S'

· Le sommet 1 de G' considéré comme sommet S' donne une représentation en couches comprenant 5 couches.

· Les sommets 2, 3, 4, 5 de G' donnent une couche C_2 comprenant 5 éléments (au lieu de 6) dans la structure en couches correspondante.

· Le sommet 6 donne le même nombre d'éléments par couches mais une répartition différente suivant la valeur du triplet dans la 2[e] couche.

· Le sommet 7 de G' donne une 2[e] couche à 5 éléments.

· Le sommet 8 donne une représentation en couches avec même nombre d'éléments par couches et même partition initiale (voir fig. 7).

Affinement des partitions initiales

Graphe G

	P_0 : (2, 13, 14) C_1	(3, 5, 8, 7) (4) (12) C_1 C_2 C_3	(10, 11) (6, 9) C_1 C_2
	2 : $\varnothing$, $\varnothing$, C_1C_1 13 : $\varnothing$, $\varnothing$, C_1C_3 14 : $\varnothing$, $\varnothing$, C_1C_2		
Nouvelle partition	(2) (14) (13) C_1 C_2 C_3		
		3 : C_1, C_2, C_1 5 : C_1, C_2, C_2 7 : C_3, C_1, C_2 8 : C_2, C_1, C_2 4 : C_2, C_1C_1, $\varnothing$ 12 : C_3, $\varnothing$, C_1C_2	
Nouvelle partition		(3) (5) (8) (7) (4) (12) C_1 C_2 C_3 C_4 C_5 C_6	
			10 : C_1, C_1C_2, $\varnothing$ 11 : C_6, C_1C_2, $\varnothing$ 6 : C_2C_4, C_1, $\varnothing$ 9 : C_3C_6, C_1, $\varnothing$
Nouvelle partition			(10) (11) (6) (9) C_1 C_3 C_3 C_4

L'ordre lexicographique défini précédemment a été respecté dans l'écriture des nouvelles classes.

Graphe G' fig. 7.

(les classes initiales sont numérotées comme dans G suivant la valeur du triplet).

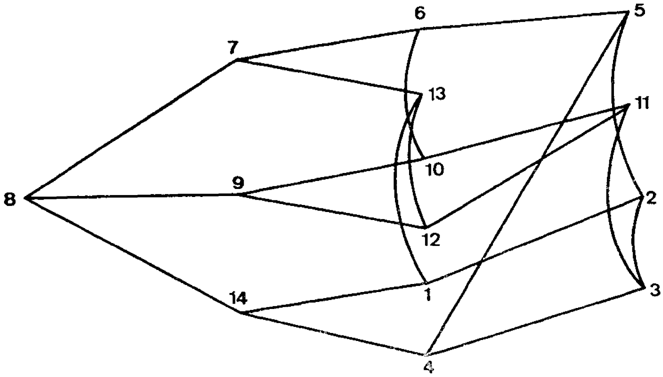

Figure 7

(7, 9, 14) C_1	(6, 10, 12, 1) (13) (4) C_1 C_2 C_3	(2, 3) (5, 11) C_1 C_2
7 : C_1C_2, $\varnothing$ 9 : C_1C_1, $\varnothing$ 14 : C_1C_3, $\varnothing$		
(9) (7) (14) C_1 C_2 C_3		
	6 : C_2, C_1, C_2 10 : C_1, C_1, C_2 12 : C_1, C_2, C_2 1 : C_3, C_2, C_1	

les suites de monômes associées à ces 4 sommets ne sont pas les mêmes que dans G; correspondance impossible : retour au choix de S'.

• Les sommets 9, 10, 11, 12 donnent un nombre de sommets en couches C_2 égal à 5.

• Le sommet 13 donne un même nombre de sommets par couche mais une partition initiale suivant la valeur du triplet différent.

• Sommet 14 de G' même nombre d'éléments par couche, même partition initiale fig. 8.

Affinement de cette partition

Nous nous plaçons donc comme dans Steen dans le cas le plus défavorable où le sommet S' satisfaisant est resté en dernier lieu. Les procédés de comptage sur les structures en couches correspondantes et la partition initiale ne laissent subsister que 2 sommets (8 et 14) de G'. Le sommet 8 est éliminé par l'algorithme d'affinement. Cet algorithme donne la solution pour 14.

(1, 4, 8) C_1	(13, 3, 5, 7) (2) (9) C_1 C_2 C_3	(11, 10) (6, 12) C_1 C_2
1 : C_1C_2 4 : C_1C_1 8 : C_1C_3 (4) (1) (8) C_1 C_2 C_3		
	13 : C_2, C_1, C_2 3 : C_1, C_2, C_1 5 : C_1, C_2, C_2 7 : C_3, C_1, C_2 2 : C_2, C_1C_1, $\varnothing$ 9 : C_3, $\varnothing$, C_1C_2	
	(3) (5) (13) (7) (2) (9) C_1 C_2 C_3 C_4 C_5 C_6	
		11 : C_1, C_1C_2, $\varnothing$ 10 : C_6, C_1C_2, $\varnothing$ 6 : C_2C_4, C_1, $\varnothing$ 12 : C_3C_6, C_1, $\varnothing$
		(11) (10) (6) (12) C_1 C_2 C_3 C_4

Les monômes caractéristiques sont les mêmes d'où l'isomorphie cherchée :

G	1	2	14	13	3	5	8	7	4	12	10	11	6	9
G'	14	4	1	8	3	5	13	7	2	9	11	10	6	12

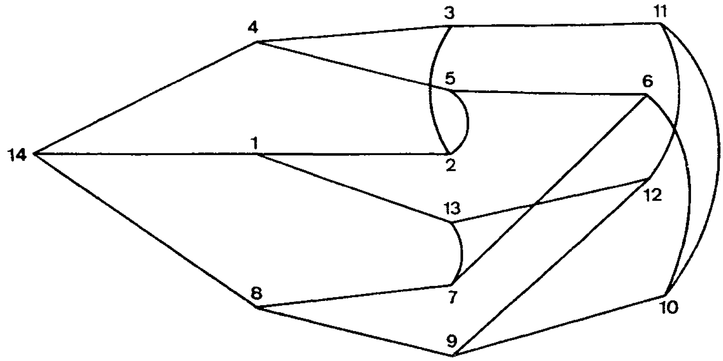

Figure 8

IV. COMPARAISON AVEC LES AUTRES METHODES

La méthode proposée évite une étude longue et systématique menée parallèlement sur les 2 graphes. Le premier graphe est étudié par rapport à un seul sommet privilégié (un seul algorithme d'affinement utile). Dans le 2e graphe, les sommets sont essayés successivement et donc, dans le cas le plus défavorable, exigent n études successives (si n est le nombre de sommets). Or, ces n études ne seront pas forcément n algorithmes d'affinement; l'algorithme d'affinement étant la partie, la plus longue et la plus onéreuse de ces méthodes. Un sommet peut être éliminé dès la construction de la structure en couches correspondante (comptage du nombre de couches et d'éléments dans une couche). Le cas de $(n + 1)$ algorithmes d'affinement constitue donc le cas le plus défavorable, très peu probable en pratique. Or, dans la méthode la plus performante (1), l'établissement du graphe quotient terminal exige à elle seule les $2n$ algorithmes d'affinement pour l'ensemble des 2 graphes. L'utilisation d'hypothèses est soit la même dans les 2 méthodes, soit diminuée dans notre méthode par suite de l'hypothèse initiale.

Exemple précédent

Dans l'exemple précédent, nous nous sommes placés dans le cas le plus défavorable où le « bon » sommet de G' est essayé en dernier lieu et nous sommes amenés à élaborer :

- 10 représentations en couches (solution éliminée après comptage du nombre de couches ou de sommets dans une couche),
- 2 représentations en couches et établissement des partitions initiales suivant les degrés des sommets,
- 2 algorithmes d'affinement.

Nous sommes donc très loin des 28 études complètes nécessitées par la méthode de Corneil et Gotlieb [1].

Il faut donc remarquer que l'élaboration d'une structure en couches avec les informations correspondantes est une étape intermédiaire judicieuse pour l'étude d'un graphe, par rapport à un sommet, et qu'il est lourd et inutile d'effectuer systématiquement les affinements jusqu'à stationnarité pour chaque sommet.

V. CONCLUSION

En conclusion, en présence de 2 graphes à comparer, il apparaît fondamental de ne pas rechercher une étude systématique et indépendante des 2 graphes mais de commencer le plus rapidement possible, les étapes de comparaison.

BIBLIOGRAPHIE

[1] D. G. Corneil et C. C. Gotlieb, *An efficient algorithm for graph isomorphism*, J.A.C.M., vol. 17, n° 1, january 1970, pp. 51-64.

[2] G. Saucier, *Codage des automates asynchrones* Thèse de doctorat ès sciences mathématiques. Université de Grenoble, 16-11-1970.

[3] J. P. Steen, *Algorithme de recherche d'un isomorphisme entre 2 graphes*. Thèse 3e cycle Mathématiques Appliquées, Lille, 26-2-1968.

[4] S. H. Unger, *A heuristic program for testing pairs of directed line graphs for isomorphism*. Comm. A.C.M., 7, 1 (janv. 1964), pp. 26-34.