

Zeitschrift: Bulletin de la Société Vaudoise des Sciences Naturelles
Band: 75 (1980-1981)
Heft: 358

Artikel: Calcul, caractérisation et identification des associations unitaires en biochronologie
Autor: Guex, Jean
DOI: <https://doi.org/10.5169/seals-277786>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 19.11.2024

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

BULLETIN N° 249 des Laboratoires de Géologie, Minéralogie, Geophysique
et du Musée géologique de l'Université de Lausanne

Calcul, caractérisation et identification des associations unitaires en biochronologie¹

PAR

JEAN GUEX²

Abstract. – A graph theoretic approach of the biochronological problem is presented. The graph representing the stratigraphic relationships between the fossil species (biostratigraphic graph) is transformed into an interval graph which can easily be interpreted in terms of a chronology by using the principle of reproducibility in the identifiability of the unitary associations.

INTRODUCTION

Divers aspects du problème posé par l'utilisation des fossiles comme outil de datation ont été abordés par l'auteur dans plusieurs travaux récents (GUEX, 1977, 1978, 1979). Ces études antérieures ne seront donc pas résumées ici : seuls les points directement liés à la présente discussion seront repris. L'objet de cette note est de présenter une formalisation mathématique des idées exposées dans ces travaux : calcul et caractérisation des associations unitaires, identification de ces associations dans les profils stratigraphiques. Ces idées peuvent être formulées de manière adéquate en utilisant certains concepts de la Théorie des Graphes. Les concepts fondamentaux de cette théorie ont été mis à la portée des non spécialistes dans plusieurs ouvrages récents : HARARY, 1969; BERGE, 1970; ROBERTS, 1976; CARRE, 1979. Les stratigraphes concernés par les problèmes théoriques de la biochronologie peuvent donc utiliser ces travaux de référence : ils y trouveront d'une part des discussions complètes sur les concepts mathématiques simples qui sont utilisés plus loin et d'autre part des exposés détaillés sur la terminologie des graphes. On commencera cette note en donnant quelques définitions indispensables à la compréhension du texte ainsi que la liste des termes biostratigraphiques introduits

¹ Travail effectué dans le cadre du projet 2.024.078 du Fonds national suisse de la recherche scientifique et publié dans le cadre du projet I.G.C.P. N° 148.

² Institut de Géologie, Palais de Rumine, Lausanne.

par l'auteur dans les travaux sus-mentionnés et qui ont une correspondance stricte avec la terminologie employée par les spécialistes des graphes.

TERMINOLOGIE ET DÉFINITIONS

A) Graphes et «graphe biostratigraphique»

Un graphe non orienté $\tilde{G} = (X, E)$ est un schéma constitué par un ensemble fini de points $X = \{x_1, \dots, x_n\}$ et par une famille $E = \{e_1, \dots, e_m\}$ de paires non ordonnées de points distincts de X . Chaque paire $e = (x, y)$ est appelée une arête du graphe et les points sont appelés les sommets du graphe.

Un graphe orienté $\vec{G} = (X, U)$ est un schéma constitué par un ensemble fini de sommets $X = \{x_1, \dots, x_n\}$ et par une famille $U = \{u_1, \dots, u_m\}$ de paires ordonnées de sommets distincts de X . Chaque paire $u = (x, y)$ est appelé un arc du graphe.

On notera G le graphe semi-orienté qui est une combinaison des deux graphes précédents.

Dans le problème étudié ici, les sommets du graphe sont les espèces fossiles.

a) Deux espèces «compatibles» (GUEX, 1977, p. 311) sont reliées par une arête dans un graphe \tilde{G} .

b) Deux espèces non compatibles dont la relation stratigraphique est connue sont reliées par un arc dans un graphe \vec{G} .

c) Deux espèces non compatibles dont la coexistence virtuelle (= déduite et non pas observée) ne peut être établie et dont la relation stratigraphique n'est pas connue ne sont reliées ni par une arête dans \tilde{G} ni par un arc dans \vec{G} .

Le graphe G montrant la totalité des informations relatives aux positions stratigraphiques des espèces (associations, exclusions avec relation stratigraphique connue ou indéterminée) sera appelé «graphe biostratigraphique» (fig. 2-f).

B) Nomenclature biostratigraphique et terminologie des graphes

- 1) Espèces compatibles: sommets voisins dans \tilde{G}
- 2) Spectre d'association d'une espèce x (GUEX, 1977, p. 311): ensemble des sommets voisins de x dans \tilde{G} (fig. 1-g)
- 3) Matrice associée: considérons un graphe $\tilde{G} = (X, E)$ et l'ensemble de ses sommets (x_1, x_2, \dots, x_n) . La matrice associée à \tilde{G} est la matrice (a_{ij}) définie par:

$$a_{ij} = \begin{cases} 1 & \text{si } (x_i, x_j) \in E \\ 0 & \text{sinon} \end{cases}$$

- 4) Matrice des spectres d'association: matrice associée au graphe \tilde{G}
- 5) Ensemble d'espèces mutuellement compatibles: clique de \tilde{G} (fig. 1-f)
- 6) n-clique: clique contenant n sommets
- 7) Clique maximale: une clique est maximale si elle n'est pas contenue dans une clique plus grande
- 8) Association unitaire (GUEX, 1977, p. 311): clique maximale
- 9) Matrice d'incidence: Soit $X = \{x_1, x_2, \dots, x_n\}$ un ensemble de sommets et $F = \{S_1, S_2, \dots, S_m\}$ une famille de sous-ensembles de X . La matrice d'incidence «sommets-ensemble de sommets» est une matrice $n \times m$ dans laquelle l'entrée i, j vaut 1 si l'élément x_i appartient à l'ensemble S_j et 0 sinon
- 10) Tableau ordonné d'associations unitaires: matrice d'incidence «cliques maximales-sommets»; sous certaines conditions (cf. p. 116) une telle matrice correspond à un graphe d'intervalle (déf. 11)
- 11) Graphe d'intervalle (BERGE, 1970, p. 358): Considérons, sur une droite, une famille $J = (J_1, J_2, \dots, J_n)$ d'intervalles. On peut former un graphe dont les sommets x_1, x_2, \dots, x_n représentent respectivement les intervalles J_1, J_2, \dots, J_n , deux sommets étant reliés par une arête si et seulement si les intervalles correspondants intersectent. Un tel graphe est appelé «graphe d'intervalle».
- 12) Propriété des 1 consécutifs: se dit d'une matrice de 1 et de 0 dont les lignes peuvent être permutées de manière à ce que les 1 soient consécutifs dans toutes les colonnes (sans 0 intercalés entre eux).
- 13) Chemin: dans un graphe non orienté \tilde{G} , un chemin est une séquence alternante de sommets et d'arêtes distincts commençant par un sommet et aboutissant à un autre sommet.
- 14) Un triplet de sommets x, y et z dans \tilde{G} est appelé «astéroïdal» s'il existe des chemins C_1 entre x et y , C_2 entre x et z et C_3 entre y et z tels qu'il n'y ait pas d'arête de x à C_3 , de y à C_2 et de z à C_1 (fig. 1-a).
- 15) Cardinalité d'un ensemble X : nombre d'éléments appartenant à X .
- 16) Cycle: cf. fig. 1-b
- 17) Circuit: cf. fig. 1-c
- 18) Corde: arête reliant 2 sommets non consécutifs d'un cycle.
- 19) Chaîne: séquence d'arcs telle que chaque arc de la séquence a une extrémité en commun avec l'arc qui le précède et l'autre extrémité en commun avec l'arc qui le suit.

C) Notation

$J_{(x)}$: intervalle d'existence d'une espèce x

$J_{(x,y)}$: intervalle de coexistence du couple (x, y)

$\Gamma_{\tilde{G}}(x_i) = \{x_j\}$: ensemble des x_j voisins de x_i dans \tilde{G} (NB: $x_i \notin \Gamma_{\tilde{G}}(x_i)$)

Z_n : cycle de longueur n sans corde dans \tilde{G} (fig. 1-b)

S_3 et S_4 : circuits semi-orientés de longueur 3 ou 4 dans G (fig. 1-d/e)

$|X|$: cardinalité d'un ensemble X

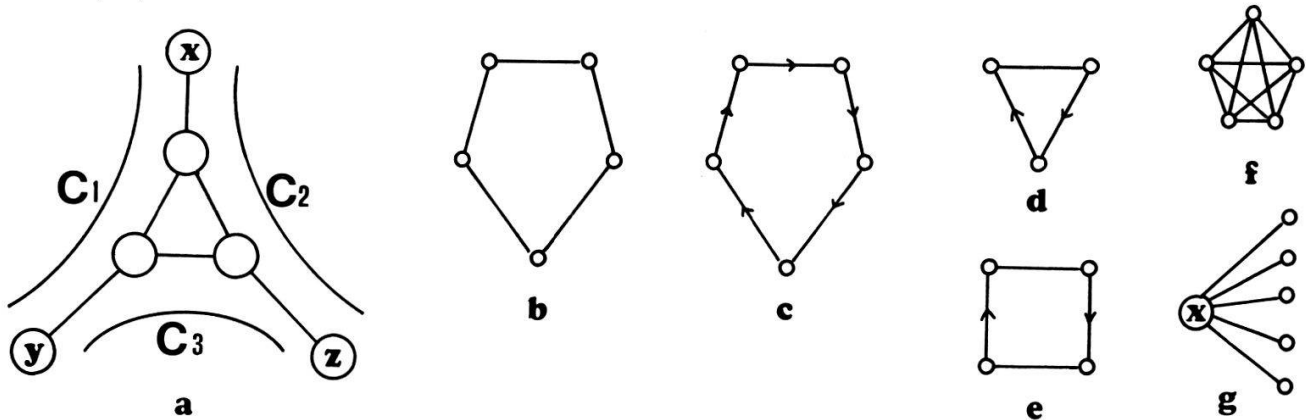


Fig. 1.

- a) Triplet astéroïdal (cf. texte)
- b) Cycle de longueur 5 sans corde
- c) Circuit de longueur 5
- d) Circuit semi-orienté S_3
- e) Circuit semi-orienté S_4
- f) Clique à 5 sommets (5-clique)
- g) Spectre d'association d'une espèce x

NATURE DU PROBLÈME

Imaginons une situation idéale dans laquelle:

- 1) la sédimentation est continue dans le temps et dans l'espace
- 2) l'enregistrement fossile des roches sédimentaires est parfait: toutes les espèces qui ont existé à un moment donné sont fossilisées partout
- 3) toutes les espèces sont ubiquistes au cours de leur durée d'existence respective.

Dans une telle situation, la séquence d'associations unitaires engendrées par les relations de coexistence entre les espèces correspondrait exactement à un *graphe d'intervalle* (définition 11). La réalité n'a bien entendu rien à voir avec une telle utopie car la nature de la distribution biogéographique des espèces fossiles (fortement limitée par les contraintes écologiques et à fortiori par le caractère fragmentaire de la documentation paléontologique) implique qu'un graphe représentant les relations de coexistence entre les espèces fossiles ne correspond pas à une chronologie même s'il satisfait les conditions qui en font un graphe d'intervalle. Cette assertion est explicitée dans la figure 2 (b à d) et le problème est discuté en détail dans GUEX, 1978.

On sait cependant qu'une séquence ordonnée d'associations unitaires peut fort bien être interprétée en termes chronologiques (principe de reproductibilité; GUEX, 1978): ceci est donc également valable pour un graphe d'intervalle.

Le premier problème qui doit être résolu lorsqu'on veut traiter de manière précise un ensemble de données biostratigraphiques éparées et fragmentaires est celui posé par la construction d'une séquence ordonnée d'associations unitaires. Dans le langage de la théorie des graphes ce problème peut être formulé ainsi: étant donné un graphe biostratigraphique, comment transformer ce graphe en un graphe d'intervalle? Cet aspect de la question va constituer l'essentiel de la discussion qui suit.

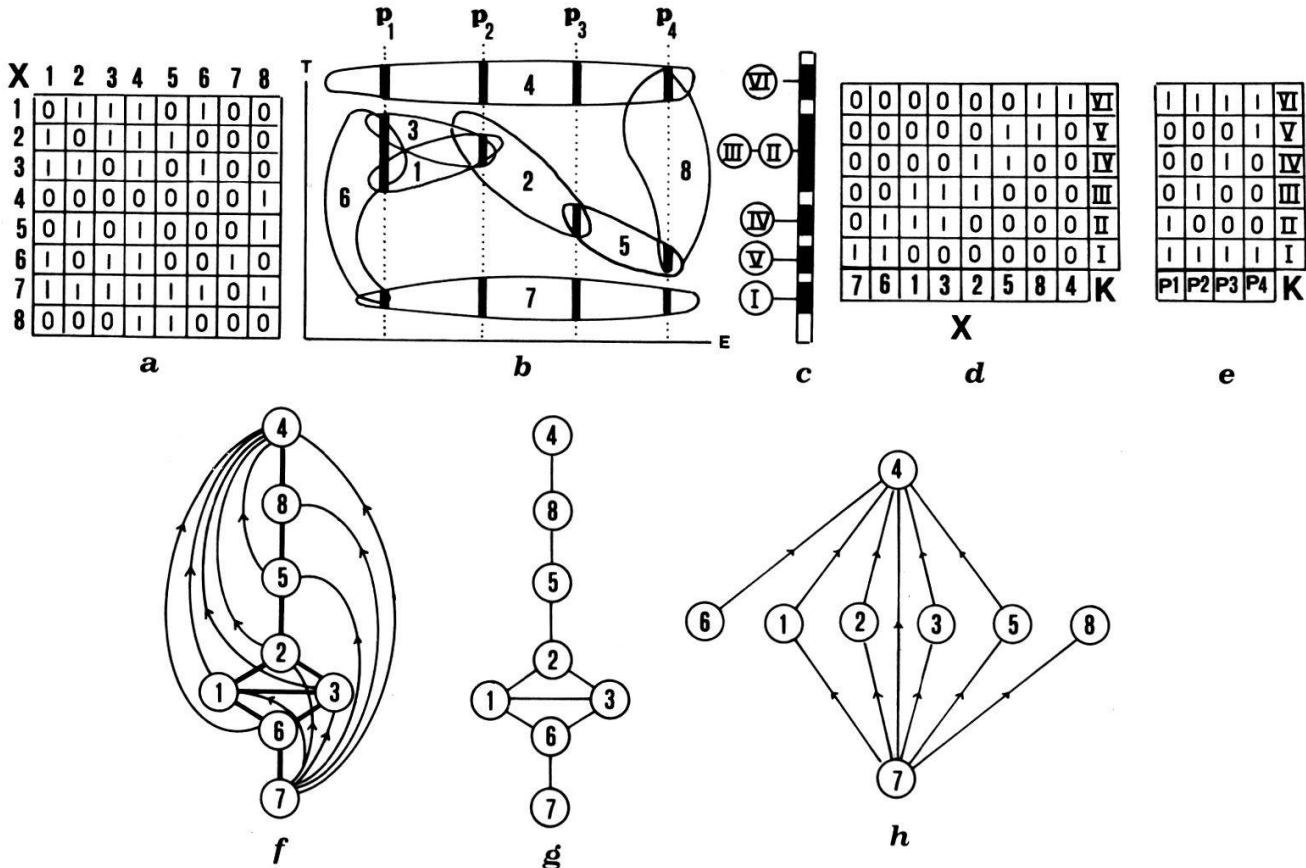


Fig. 2.

- a) Matrice A, «espèce-espèce» (définition: cf. p.116), contenant les informations biostratigraphiques tirées des profils p_1 à p_4 de la figure 2-b.
- b) Distribution spatio-temporelle de 8 espèces (1 à 8). Pointillés verticaux: tracé de 4 profils stratigraphiques (p_1 à p_4) qui traversent (dans les blocs noirs) les domaines d'existence (surfaces fermées) des 8 espèces. T = axe du temps. E = «espace».
- c) Position chronologique réelle des intervalles I à VI de la matrice d'incidence «cliques maximales-espèces» de la figure 2-d.
- d) Matrice d'incidence «cliques maximales-espèces» (K_i-x_i): cette matrice représente les relations d'association et d'exclusion entre les espèces 1 à 8 (exprimées en termes de «séquence d'associations unitaires»); elle possède la propriété des 1-consécutifs (c'est donc un graphe d'intervalle) mais la séquence ordonnée des K_i ($i = I$ à VI) ne correspond pas à l'ordre chronologique réel représenté dans la figure 2-c (d'après GUEX, 1978).
- e) Matrice «profils-cliques maximales» identifiables dans les profils 1 à 4. Cette matrice montre que seules les cliques I et VI sont identifiables de manière reproductible.
- f) Graphe G montrant les relations stratigraphiques globales (association, exclusion et superposition) observées entre les espèces 1 à 8 dans les profils p_1 à p_4 de la figure 2-b.
- g) Graphe \tilde{G} montrant les relations de coexistence observées entre les espèces 1 à 8 dans les profils p_1 à p_4 de la figure 2-b. Dans le cas particulier, ce graphe est un graphe d'intervalle.
- h) Graphe \bar{G} montrant les relations de superposition observées entre les espèces 1 à 8 dans les profils p_1 à p_4 de la figure 2-b.

NATURE DES DONNEES

Considérons un ensemble de profils stratigraphiques $P = \{p_1, \dots, p_m\}$:

Chacun des profils p_i est constitué d'un ensemble N_i de niveaux sédimentaires; le niveau j du profil i a un contenu fossile X_{ij} qui est une partie de l'ensemble des espèces $X = \{x_1, \dots, x_n\}$. Les informations sur la distribution stratigraphique locale (c'est-à-dire dans p_i) des espèces sont introduites dans une matrice $A = (a_{ij})$ ($i, j = 1, \dots, n$); on posera:

$$a_{ij} = \begin{cases} 1 & \text{si } x_i \text{ se trouve dans un niveau inférieur à celui de } x_j \text{ ou} \\ & \text{au même niveau que } x_j \\ 0 & \text{sinon} \end{cases}$$

La figure 2-a montre une matrice A qui contient les informations biostratigraphiques tirées des profils 1 à 4 de la figure 2-b. G sera le graphe dont A est la matrice d'incidence sommets-sommets (fig. 2-f):

si $a_{ij} = 1$ et $a_{ji} = 1$ on dira qu'on a une arête (i, j)
 si $a_{ij} = 1$ et $a_{ji} = 0$, on aura un arc (i, j) .

$\tilde{G} = (X, E)$ sera le graphe formé par l'ensemble E des arêtes de G (fig. 2-g), alors que $\vec{G} = (X, U)$ sera formé par l'ensemble U des arcs de G (fig. 2-h).

CONSTRUCTION DU GRAPHE D'INTERVALLE

Deux théorèmes de base vont nous être utiles dans la recherche d'une solution au problème posé par la transformation d'un graphe biostratigraphique en un graphe d'intervalle:

I) FULKERSON & GROSS (1965):

«Un graphe est un graphe d'intervalle si et seulement si sa matrice d'incidence cliques maximales-sommets possède la propriété des 1 consécutifs.»

II) LEKKERKERKER & BOLAND (1962):

«Un graphe est un graphe d'intervalle si et seulement si il ne contient aucun cycle Z_n ($n \geq 4$) et si il ne contient pas de triplets astéroïdaux.»

L'idée de la méthode consiste à détruire les cycles Z_n et les triplets astéroïdaux de \tilde{G} en fonction d'une série de critères dont l'admissibilité est préalablement discutée. Son développement s'effectue en 6 étapes:

- 1) destruction des cycles de longueur 4 de \tilde{G}
- 2) détermination des cliques maximales de \tilde{G}
- 3) destruction des circuits semi-orientés S_3 et S_4 dans G
- 4) identification des cliques maximales de \tilde{G} dans les profils

- 5) construction d'un graphe orienté montrant les relations stratigraphiques entre les cliques maximales identifiables dans les profils
 6) destruction des cycles Z_n ($n > 4$) et des triplets astéroïdaux dans \tilde{G}
 La dernière étape aboutit à la construction d'un graphe d'intervalle (théorèmes I et II) que l'on sait interpréter en termes chronologiques.

1) *Destruction des cycles Z_4 (cf. Appendices)*

Les relations d'association et d'exclusion entre 4 espèces x_i, x_j, x_k, x_l peuvent engendrer un cycle Z_4 dans \tilde{G} pour deux raisons distinctes: le **remaniement** ou l'**insuffisance documentaire**.

En l'absence de critères sédimentologiques directs (resédimentation évidente, altération physique des fossiles etc....), le remaniement d'une espèce ne peut être présumé que dans la mesure où cette espèce montre une discontinuité nette dans sa distribution biochronologique: il est évident que si une espèce x se trouve dans 2 zones adjacentes A et B, il est exclu de soupçonner que x n'existe chronologiquement que dans la zone A et qu'il est remanié dans la zone B qui suit... La démonstration d'un remaniement par des critères purement biochronologiques ne peut se faire qu'avec l'aide des espèces dont les relations engendrent des associations unitaires identifiables de manière reproductible. En l'absence de preuve «physique» directe, on est donc obligé de commencer par admettre que les discontinuités dans la distribution des espèces sont dues à une insuffisance documentaire et non pas à des remaniements: l'interprétation de ces discontinuités en termes de remaniements est discutée plus loin (p.124 et Appendice II).

Trois types de Z_4 en «*situation non remaniée*» peuvent être schématisés:

Situation 1)

Si la relation stratigraphique est connue pour l'un des couples d'espèces mutuellement exclusives et impliquées dans le Z_4 , l'autre couple a alors nécessairement coexisté virtuellement. La figure 3-a démontre cette affirmation.

Situation 2)

Si la relation stratigraphique est connue pour les 2 couples d'espèces mutuellement exclusives qui sont impliquées dans le Z_4 , ces 2 couples sont nécessairement virtuellement compatibles: cette affirmation est démontrée par la figure 3-b. Le schéma ci-dessous (fig. 3-c) montre qu'une telle situation risque de se produire très souvent lorsque 2 espèces apparaissent à peu près simultanément au cours d'un intervalle de temps pendant lequel 2 autres espèces disparaissent de manière subconcomitante.

Situation 3

La relation stratigraphique est inconnue pour les 2 couples mutuellement exclusifs impliqués dans le Z_4 . Une telle situation peut également être due à une configuration relationnelle du type illustré dans la figure 3-c.

Les informations contenues dans la matrice A permettent de détecter facilement les Z_4 , de voir à quelle situation (1, 2 ou 3) on a à faire et d'en déduire les coexistences virtuelles.

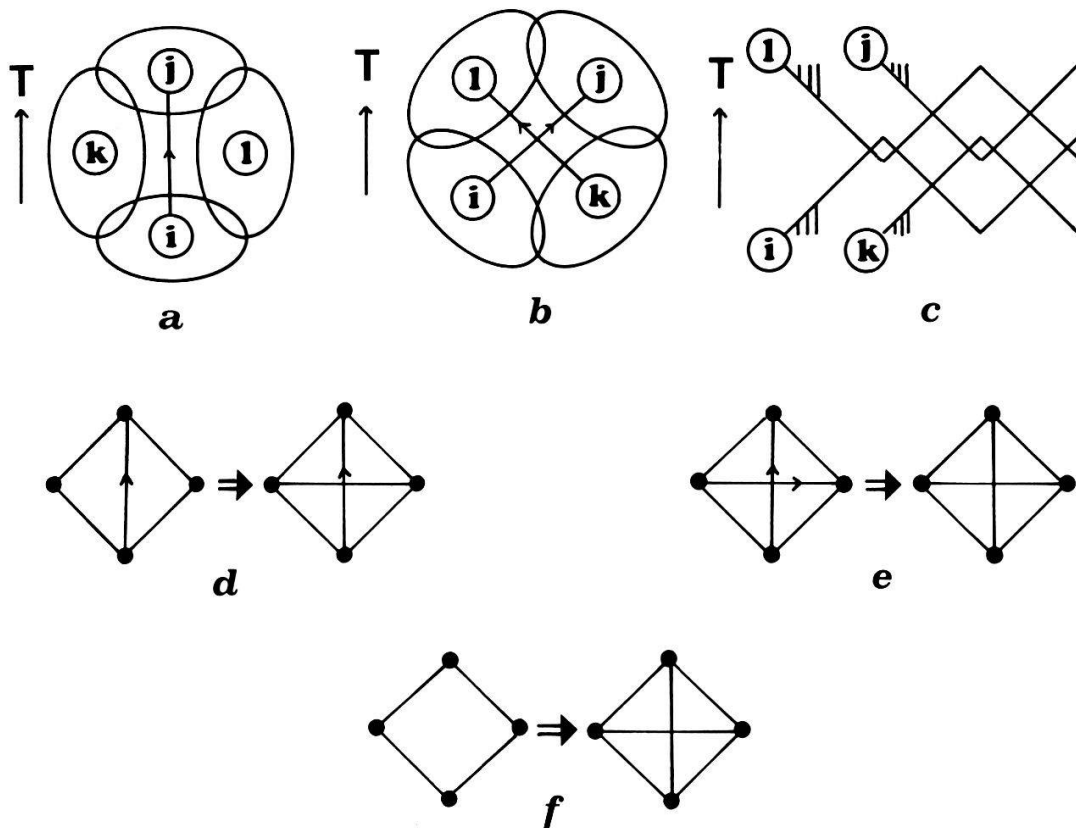


Fig. 3.

Schémas illustrant le fait qu'en «situation non remaniée», on peut détruire les cycles Z_4 par adjonction d'arêtes dans ces cycles. Les ellipses limitent les domaines d'existence des espèces i, j, k et l. T = axe du temps.

- l'arc ij indique que la relation stratigraphique entre les espèces i et j est connue.
- les arcs ij et kl indiquent que la relation stratigraphique entre les couples d'espèces ij et kl est connue.
- Configuration relationnelle hypothétique (cf. p.118) engendrant un Z_4 . Les petits traits verticaux indiquent respectivement les frontières inférieures et supérieures des domaines d'existence des espèces j et l d'une part et i et k d'autre part.
- Schéma montrant la procédure d'adjonction d'une arête dans la situation 1 (cf. p.119)
- Schéma montrant la procédure d'adjonction de 2 arêtes dans la situation 2 (cf. p.119)
- Schéma montrant la procédure facultative d'adjonction de 2 arêtes dans la situation 3 (cf. p.119)

Commençons par examiner les Z_4 qui contiennent un arc entre x_i et x_j (situations 1 et 2):

Pour chaque couple (x_i, x_j) dans \vec{G} tel que $a_{ij} = 1$ et $a_{ji} = 0$ dans la matrice A , on cherche l'ensemble des espèces communes aux 2 spectres d'association en calculant:

$$I_{ij} = \Gamma_{\vec{G}}(x_i) \cap \Gamma_{\vec{G}}(x_j)$$

Si $|I_{ij}| \leq 1$ on ne fait rien car il n'y a pas de Z_4

Si $|I_{ij}| \geq 2$, on note $\{(x_k, x_l)\}$ le(s) couple(s) d'espèce(s) qui se trouvent dans I_{ij} :

Situation 1 (fig. 3a + d): Si $a_{kl} = a_{lk} = 0$ (NB: kl et lk sont les indices ij et ji des couples d'espèces (x_k, x_l)), on ajoute une arête à Z_4 en posant $a_{kl} = a_{lk} = 1$ dans la matrice A .

Situation 2 (fig. 3b + e): Si $a_{kl} = 1$ et $a_{lk} = 0$, on remplace les 2 arcs du Z_4 considéré par 2 arêtes en posant $a_{ij} = a_{ji} = a_{kl} = a_{lk} = 1$ dans la matrice A .

Situation 3 (fig. 3c + f): Pour chaque couple (x_i, x_j) dans G , tel que $a_{ij} = a_{ji} = 0$ on calcule

$$I_{ij} = \Gamma_{\vec{G}}(x_i) \cap \Gamma_{\vec{G}}(x_j)$$

Si $I_{ij} = \emptyset$ ou si I_{ij} est une clique, on ne fait rien (pas de Z_4): sinon on ajoute 2 arêtes dans le Z_4 en posant $a_{ij} = a_{ji} = a_{kl} = a_{lk} = 1$ dans la matrice A .

Remarque: En «*situation remaniée*», l'adjonction d'arêtes à un Z_4 de type 3 (fig. 3-f) peut induire une perte sérieuse dans les informations chronologiques potentiellement contenues dans le graphe G de départ: en pratique il est donc nécessaire de comparer les résultats biochronologiques obtenus après avoir effectué la totalité des opérations décrites ici avec ceux que l'on obtient en omettant certaines de ces opérations.

2) Calcul des cliques maximales

Plusieurs algorithmes permettant de déterminer les cliques maximales d'un graphe ont été publiés dans la littérature récente: une discussion générale de ce problème est présentée par JOHNSTON (1976).

Un algorithme voisin de celui de BIERSTONE (in AUGUSTSON & MINKER, 1970) a été construit par l'auteur. On n'en donnera ici que l'idée générale.

L'algorithme procède par induction: connaissant l'ensemble des cliques maximales d'un graphe d'ordre n (= à n sommets) \vec{G}_n , on sait déterminer l'ensemble des cliques maximales de ce graphe au rang $n+1$ (= après lui avoir ajouté un sommet).

Considérons un graphe $\tilde{G}_n = (X_n, E)$ où $X_n = \{x_1, \dots, x_n\}$ est l'ensemble des sommets.

$\{x_{n+1}\}$ sera un sommet ajouté au graphe \tilde{G}_n

$\Gamma_{\tilde{G}_n}(x_{n+1})$ sera l'ensemble des voisins de $\{x_{n+1}\}$ dans X_n

$K_n = \{k_n\}$ sera l'ensemble des cliques maximales de \tilde{G}_n (= au rang n)

C_{n+1} sera un ensemble de cliques pas forcément maximales au rang $n+1$

K_{n+1} sera l'ensemble des cliques maximales de \tilde{G}_n au rang $n+1$

Il est évident que si:

- 1) $\left\{ k_n \cap \Gamma_{\tilde{G}_n}(x_{n+1}) \right\} = \emptyset$, alors k_n et $\{x_{n+1}\}$ sont 2 cliques distinctes au rang $n+1$
- 2) $\left\{ k_n \cap \Gamma_{\tilde{G}_n}(x_{n+1}) \right\} = k_n$, alors $\left\{ k_n \cup \{x_{n+1}\} \right\}$ est une clique maximale nouvelle au rang $n+1$
- 3) $\left\{ k_n \cap \Gamma_{\tilde{G}_n}(x_{n+1}) \right\} \neq \emptyset$ et $\neq k_n$, alors k_n est une clique maximale reconduite au rang $n+1$ et $\left\{ \left(k_n \cap \Gamma_{\tilde{G}_n}(x_{n+1}) \right) \cup \{x_{n+1}\} \right\}$ est une clique nouvelle au rang $n+1$.

On peut donc écrire

$$C_{n+1} = \left\{ k_n, \left(k_n \cap \Gamma_{\tilde{G}_n}(x_{n+1}) \right) \cup \{x_{n+1}\} \mid k_n \in K_n \right\}$$

Notons R , l'opération qui consiste à éliminer par inclusion toutes les cliques non maximales au rang $n+1$. On peut alors écrire:

$$K_{n+1} = R(C_{n+1})$$

L'algorithme opère de manière itérative: au début, on pose

$$X_1 = \{x_1\}, K_1 = \{x_1\} \text{ et } \Gamma_{\tilde{G}_1}(x_2) = \{x_1\} \text{ ou vide.}$$

3) Suppression des circuits semi-orientés S_3 et S_4

Après avoir déterminé les cliques maximales de \tilde{G} , on effectue encore la modification suivante sur ce graphe: étant donné 2 cliques maximales K_i et K_j , soit

$$\alpha = \{x \in K_j \mid \exists \text{ arc } (y, x) \text{ avec } y \in K_i, x \notin K_i\}$$

$$\beta = \{x \in K_j \mid \exists \text{ arc } (x, y) \text{ avec } y \in K_i, x \notin K_i\}$$

Si α et β sont tous deux non vides, alors K_i est remplacé par $K_i \cup \alpha \cup \beta$ (ceci introduit de nouvelles arêtes dans \tilde{G}). En répétant ceci on détruit tous les circuits semi-orientés S_3 et S_4 de G .

Les schémas ci-dessous (fig. 4) montrent quels genres de configurations relationnelles entre les espèces induisent de tels circuits.

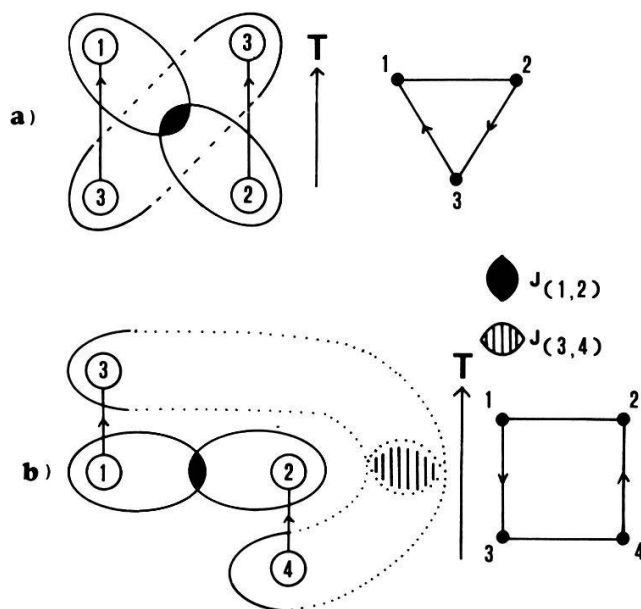


Fig. 4.

a) Exemple de configuration relationnelle engendrant un circuit semi-orienté S_3 . Les ellipses fermées sont les frontières des domaines d'existence respectifs des espèces 1 et 2 qui intersectent (surface noire). L'ellipse ouverte montre comment la distribution spatiale discontinue de l'espèce 3 engendre des relations contradictoires. T = axe du temps. Dans cet exemple $K_i = \{1, 2\}$, $K_j = \{3\}$; $\alpha = \beta = 3$ (cf. p.120)

b) Exemple de configuration relationnelle engendrant un circuit semi-orienté S_4 ; la position de l'intervalle de coexistence des espèces 3 et 4 ($J_{(3,4)}$) par rapport à $J_{(1,2)}$ est indéterminée. T = axe du temps. Dans cet exemple $K_i = \{1, 2\}$; $K_j = \{3,4\}$; $\alpha = 3$ et $\beta = 4$

4) Identification des cliques maximales dans les profils stratigraphiques

Il résulte du théorème de FULKERSON & GROSS (I) que chaque clique maximale d'un graphe d'intervalle peut être caractérisée par un sous-ensemble de sommets qui n'appartient qu'à cette clique maximale: on parlera de sommet simplicial si ce sous-ensemble se réduit à un sommet (dont tous les voisins engendrent une clique) ou une paire simpliciale si ce sous-ensemble se réduit à 2 sommets.

Chaque sommet (ou paire) simplicial(e) d'une clique maximale la caractérise entièrement puisque ce sommet ou cette paire n'apparaît dans aucune autre clique maximale.

Appelons $K = \{K_1, \dots, K_p\}$ l'ensemble des cliques maximales de \tilde{G} (modifié), numérotées pour l'instant dans un ordre quelconque.

Commençons par identifier ces cliques dans les niveaux sédimentaires des profils p_i à l'aide de leurs espèces caractéristiques (espèces et/ou paires simpliciales); pour chaque niveau n_{ij} de chaque profil p_i on calcule

$$Y_{ij} = \max. |X_{ij} \cap K_l| \\ l = 1, \dots, p$$

S'il y a une valeur unique l pour laquelle Y_{ij} est maximum, cela signifie que n_{ij} contient au moins une espèce (ou une paire d'espèces) caractéristique de K_l : on attribue donc la clique K_l au niveau n_{ij} du profil p_i et on dit que cette clique est identifiable dans le profil p_i . S'il y a plusieurs valeurs de l on ne fait rien.

Remarques:

1) L'utilisation des espèces (paires) simpliciales pour identifier les cliques maximales correspond exactement à l'opération d'identification des associations unitaires décrite par l'auteur dans des travaux antérieurs.

2) Les associations virtuelles déduites au cours des étapes 1 et 3 ci-dessus restent sans influence sur la procédure d'identification car les éventuelles paires simpliciales engendrées par ces manipulations n'ont (par définition) pas été observées dans les profils stratigraphiques: une conséquence de ceci peut être que certaines cliques maximales ne seront éventuellement identifiables nulle part: ces cliques peuvent être éliminées.

5) Construction du graphe G_k

Les informations sur les relations stratigraphiques locales (dans p_i) entre les cliques maximales identifiables dans les profils p_i peuvent être introduites dans une matrice $B = (b_{ij})$; on posera

$$b_{ij} = \begin{cases} 1 & \text{si } K_i \text{ se trouve au-dessous de } K_j \text{ dans un profil } p_i \\ 0 & \text{sinon} \end{cases}$$

Dans le graphe orienté $G_k = (K, U_k)$ associé à B , les cliques maximales qui présentent le plus grand nombre de contradictions dans leurs relations de position stratigraphique observée ($b_{ij} = b_{ji} = 1$) peuvent être éliminées (p. ex. les cliques emballées dans les surfaces pointillées de la fig. 5-a): G_k n'a ainsi plus de circuits. Certaines cliques maximales dont les relations sont indéterminées ($b_{ij} = b_{ji} = 0$) peuvent être trivialement réunies (fig. 5-a: surfaces blanches).

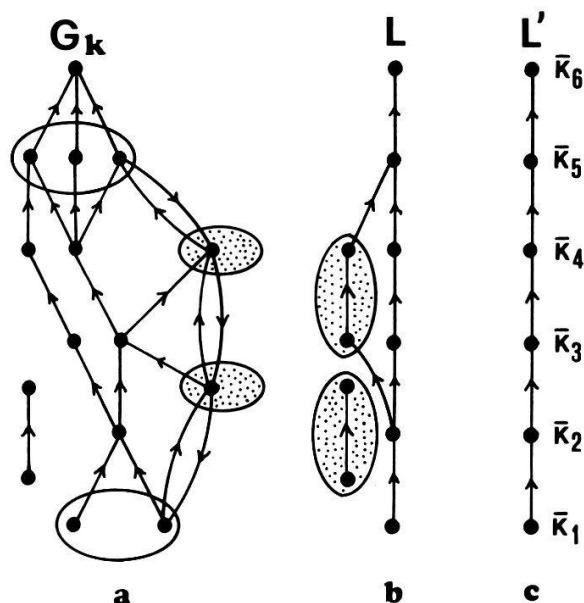


Fig. 5.

Grphe orienté associé à une matrice B (cf. p.122) hypothétique, montrant les relations stratigraphiques entre des cliques maximales identifiables dans des profils imaginaires.

a) Les 2 cliques maximales emballées dans les surfaces pointillées sont celles qui montrent le plus grand nombre de relations contradictoires (arcs opposés: $b_{ij} = b_{ji} = 1$): ces cliques sont éliminées.

Les cliques emballées dans les surfaces blanches sont réunies.

b) Les cliques maximales emballées dans les surfaces pointillées n'appartiennent pas à L .

c) Séquence ordonnée de cliques maximales obtenue après avoir achevé les opérations décrites dans le texte.

6) Construction du graphe d'intervalle

a) Recherche de la plus longue chaîne de cliques de G_k

Appelons L la plus longue chaîne de cliques maximales de G_k (fig. 5-b: les cliques se trouvant dans les surfaces hachurées n'appartiennent pas à L) et rebaptisons ces cliques \bar{K}_p , $p = 1, \dots, p_L$ (s'il y a plusieurs L possibles on choisit celle qui contient les cliques identifiables dans le plus grand nombre de p_i). Le graphe réduit G_k donné par la séquence L n'engendre pas forcément un graphe d'intervalle par suppression de l'orientation: il peut encore contenir des cycles Z_n et des triplets astéroïdaux: ces configurations interdites induisent des discontinuités dans la distribution des 1 dans les colonnes de la matrice d'incidence «cliques maximales-espèces» (cf. fig. 1-f). Ces discontinuités peuvent être supprimées¹: on remplace simplement par des 1 les 0 qui sont intercalés entre des 1 dans les colonnes de cette matrice. On obtient ainsi un graphe d'intervalle dans lequel l'ordre obtenu sur l'ensemble des cliques est établi uniquement à l'aide des cliques maximales identifiables de manière reproductible dans les p_i (celles-ci appartiennent nécessairement à L).

¹ NB: ceci ne modifie pas l'ordre qui existe sur l'ensemble des cliques.

b) Réintégration d'une partie des cliques éliminées

Appelons K^* l'ensemble des cliques qui ne sont pas dans L (fig. 5 a/b: surfaces hachurées). Pour chaque $\bar{K}_k \in K^*$ on calcule

$$W_{pk} = \max_{p = 1, \dots, p_L} |\bar{K}_k \cap \bar{K}_p|$$

S'il y a une valeur unique p pour laquelle W_{pk} est maximum on remplace \bar{K}_p par $\bar{K}_p \cup \bar{K}_k$, sauf si cela réintroduit une relation contradictoire, auquel cas on élimine \bar{K}_k définitivement. S'il y a plusieurs valeurs de p , on élimine également \bar{K}_k . On obtient finalement une séquence L' de \bar{K}_p (fig. 5-c) qui est ordonnée de manière univoque; cette séquence est transformée en un graphe d'intervalle selon la procédure décrite plus haut: celle-ci ne préserve pas forcément la maximalité des cliques: si cela s'avère nécessaire, on peut éliminer les cliques qui ne sont plus maximales.

INTERPRETATION CHRONOLOGIQUE

1) *Etablissement des biochronozones*

Le graphe d'intervalle (= tableau ordonné des associations unitaires) étant construit, on peut maintenant l'utiliser pour l'interprétation chronologique.

On a déjà vu (GUEX, 1979, p. 179) comment les associations unitaires identifiables de manière reproductible servaient de point de départ à la définition de biochronozones. La manière précise selon laquelle les associations unitaires sont groupées pour constituer une zone (choix des limites) ne peut être qu'empirique. Dans une optique informatique on pourrait cependant dire que le groupement des associations unitaires non reproductibles autour de celles qui le sont (= squelette chronologique) relève d'une optimisation du contrôle superpositionnel entre les biochronozones dont on veut définir les limites.

2) *Interprétation des discontinuités dans la distribution zonale des espèces*

On a dit plus haut qu'en l'absence de critères physiques directs, le remaniement d'une espèce ne pouvait être supposé que dans la mesure où celle-ci montrait une discontinuité nette dans sa distribution biochronologique. En d'autres termes le remaniement de cette espèce ne peut être diagnostiqué qu'après l'attribution de chaque ensemble de niveaux de chaque profil stratigraphique aux différentes zones établies. Une fois ces corrélations effectuées, on peut construire un diagramme qui montre le contenu spécifique de chacune des zones: ce contenu est l'union des contenus fossilifères (X_{ij}) des niveaux attribuables à une zone donnée. Si certaines espèces montrent une distribution discontinue dans ce diagramme, on peut éventuellement supposer qu'elles ont été remaniées.

APPENDICES

I) Les graphes extrêmes

Une raison purement technique de chercher d'abord à éliminer les Z_4 est la suivante: MOON & MOSER (1965) ont démontré que le nombre maximum de cliques maximales d'un graphe à n sommets pouvait être égal à $3^{n/3}$.

De tels graphes sont appelés «graphes extrêmes» ou «graphes de MOON & MOSER». Il est évident que la recherche des cliques maximales d'un graphe biostratigraphique est sans intérêt pour nous si le nombre de ces cliques devient trop grand: la première chose à faire consiste donc à isoler et à traiter les structures susceptibles d'être à l'origine d'une telle situation.

On laissera au lecteur le soin de vérifier qu'un graphe extrême contenant $3^{n/3}$ cliques maximales devient une n -clique si l'on détruit tous ses Z_4 par adjonction d'arêtes: cette opération «empêche» un graphe quelconque (comme ceux que nous étudions) d'être localement extrême (*i.e* de contenir des sous-graphes qui sont une imbrication de Z_4).

II) Remarque générale sur la destruction des cycles et des circuits

Les lecteurs se demanderont peut-être s'il ne serait pas plus simple d'utiliser la matrice A pour chercher directement un ordre sur l'ensemble des espèces et d'en déduire ultérieurement la séquence d'associations unitaires résultant de cet ordre. La raison pour laquelle nous ne procédons pas ainsi est simple: Les cycles ou les circuits du graphe G (engendrés par les espèces dont la distribution est discontinue) pourraient aisément être détruits en retirant certaines espèces de ce graphe. Une telle réduction de G se justifierait si l'on savait a priori quelles espèces ont une distribution discontinue *dans le temps*: or ça n'est évidemment pas le cas puisque c'est précisément l'une des choses que l'on cherche à établir. Exemple: considérons 2 espèces A et C pour lesquelles la relation stratigraphique est indéterminée et admettons qu'il existe une espèce B située au-dessus de A dans un profil et au-dessous de C dans un autre profil. Il est évidemment faux de conclure que C est plus récent que A si ces espèces sont impliquées dans un cycle ou dans un circuit. Il est également clair que l'élimination de l'une des espèces du cycle ou du circuit n'aide en rien à la mise en évidence de la discontinuité dans la distribution chronologique de l'une des espèces. Il n'existe d'autre part aucun critère objectivement valable pour choisir l'espèce qui devrait être éliminée: retirer les espèces impliquées dans des cycles ou dans des circuits en fonction de leur rareté géographique reviendrait à conserver uniquement celles qui ont statistiquement le plus de chance d'avoir subi des remaniements. En conséquence il nous a donc paru préférable de mettre en évidence le plus possible d'associations virtuelles déductibles à l'aide des configurations

relationnelles entre les espèces et entre les cliques maximales. Cette procédure présente l'avantage de prendre en considération la totalité des espèces dans un problème donné.

BIBLIOGRAPHIE

- AUGUSTSON, J.G. & MINKER, J. 1970. – An analysis of some graph theoretical cluster techniques. *J. Ass. Comput. Mach.* 17.
- BERGE, C. 1970. – Graphes et hypergraphes. Dunod éd.
- CARRE, B. 1979. – Graphs and networks. Clarendon Press, Oxford.
- FULKERSON, D.R. & GROSS, O.A. 1965. – Incidence Matrices and Interval Graphs. *Pacif. J. Math.* 15.
- GUEX, J. 1977. – Une nouvelle méthode d'analyse biochronologique. *Bull. Lab. Géol. Univ. Lausanne* 224 et *Bull. Soc. Vaud. Sc. Nat.* 351.
- 1978. – Influence du confinement géographique des espèces fossiles sur l'élaboration d'échelles biochronologiques et sur les corrélations. *Bull. Lab. Géol. Univ. Lausanne* 230 et *Bull. Soc. Vaud. Sc. Nat.* 354.
- 1979. – Terminologie et méthodes de la biostratigraphie moderne: commentaires critiques et propositions. *Bull. Lab. Géol. Univ. Lausanne* 234 et *Bull. Soc. Vaud. Sc. Nat.* 355.
- 1980. – Datations paléontologiques et graphes d'intervalle. In «*Regards sur la Théorie des Graphes*», *Colloque de Cerisy*. Presses Polytechniques Romandes, HANSEN & DE WERRA Éd.
- HARARY, F. 1969. – Graph Theory. Addison-Wesley Publ., Reading Mass.
- JOHNSTON, H.C. 1976. – Cliques of a graph: variations on the Bron-Kerbosch algorithm. *Int. J. Comput. & Inf. Sci.* 5.
- LEKKERKERKER, C.B. & BOLAND, J.C. 1962. – Representation of a Finite Graph by a Set of Intervals on the Real Line. *Fund. Math.* 51.
- MOON, J.W. & MOSER, 1965. – On cliques in graphs. *Israel J. Math.* 3.
- ROBERTS, F. 1976. – Discrete mathematical models. Prentice-Hall, Englewood Cliffs, N.J.

REMERCIEMENTS

L'auteur a eu d'utiles discussions avec les Professeurs HENRI MASSON, FRANÇOIS REUSE et DOMINIQUE DE WERRA (Lausanne) et avec le D^r JACQUES ROUX (Orléans): ces collègues trouveront ici l'expression de sa reconnaissance.

Manuscrit reçu le 22 octobre 1980.