Agrégation externe de mathématique, préparation 2007
Épreuve de modélisation, option informatique

(D)

Résumé: comment reconnaître un crime parfait?

Mots-clés: graphes, coloriage de graphes, algorithme glouton.

1 Introduction

Par un beau matin de novembre, le duc de Densmore est trouvé mort suite à l'explosion d'une pièce de son château de l'île de White. Prestement diligentés, les enquêteurs constatent que le meurtre a été commis à l'aide d'une bombe placée avec précaution dans le labyrinthe du château. Or, avant son assassinat, le duc avait invité 8 femmes sur l'île. Lorsqu'elles sont interrogées, encore sous le choc de la mort du duc, elles ne peuvent se souvenir des dates précises auxquelles chacune est venue au château, mais elles se rappellent quelles autres femmes elles ont croisées au château pendant leur séjour, et chacune jure que si quiconque d'autre avait été là en même temps qu'elle, elle s'en serait aperçue, à moins qu'une personne ne se soit volontairement cachée dans le labyrinthe. Le marin qui faisait la navette vers l'île a lui aussi oublié les dates, mais il atteste n'avoir transporté chacune que pour un seul aller retour. Après leur interrogatoire, les enquêteurs établissent le rapport suivant :

- Ann dit avoir vu Betty, Cynthia, Emily, Felicia et Georgia.
- Betty dit avoir vu Ann, Cynthia et Helen.
- Cynthia dit avoir vu Ann, Betty, Diana, Emily et Helen.
- Diana dit avoir vu Cynthia et Emily.
- Emily dit avoir vu Ann, Cynthia, Diana et Felicia.
- Felicia dit avoir vu Ann et Emily.

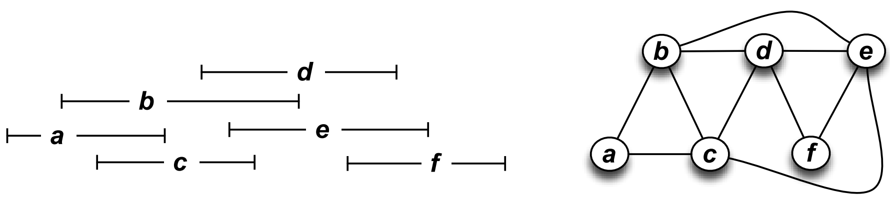

FIG. 1 – Un exemple de graphe d'intervalles

– Georgia dit avoir vu Ann et Helen.
– Helen dit avoir vu Betty, Cynthia et Georgia.

2 Les graphes d'intervalles

Pour résoudre cette énigme proposée dans l'Oulipo par le mathématicien Claude Berge, il faut faire appel à la théorie des graphes d'intervalles.

Un graphe (non orienté) est un couple $G = (V, E)$ où V est l'ensemble des sommets et E l'ensemble des arêtes de G, une arête étant une paire d'éléments de V. Le complémentaire d'un graphe $G = (V, E)$ est le graphe $\bar{G} = (V, \bar{E})$, où $\bar{E}$ est l'ensemble des paires de sommets qui ne sont pas dans E. Un *graphe d'intervalles* est un graphe non orienté $G = (V, E)$ où, à isomorphisme de graphe près, V est une famille d'intervalles $I_1, .., I_n$ de $\mathbb{R}$ et E est l'ensemble des paires d'intervalles dont l'intersection est non vide, autrement dit $E = \{\{I, I'\} : I, I' \in V,\ I \cap I' \neq \emptyset \text{ et } I \neq I'\}$. Il importe peu que les intervalles soient ouverts, fermés, ou mixtes. Tous les graphes ne sont pas des graphes d'intervalles : ainsi, le graphe cyclique de longueur 4 n'est pas un graphe d'intervalles. C'est en partant de cette remarque qu'il devient possible de trouver l'assassin du duc de Densmore. Les graphes d'intervalles interviennent dans la modélisation de nombreux problèmes d'optimisation combinatoire, et ont trouvé des applications dans les domaines les plus divers, tels l'analyse des recouvrements de gènes, la datation d'objets funéraires égyptiens, ou la modélisation de l'indifférence en sciences cognitives.

Avant d'étudier plus précisément la famille des graphes d'intervalles, rappelons quelques définitions utiles sur les graphes généraux.

– On appelle *coloriage* de G une fonction c de V dans $\mathbb{N}$ telle que si $\{v_1, v_2\} \in E$, $c(v_1) \neq c(v_2)$; on appelle *nombre chromatique* d'un graphe G, noté $\chi(G)$, le plus petit entier n tel qu'il existe un coloriage c de G à valeur dans $\{1, .., n\}$. Par exemple, le graphe G de la figure 1 admet le nombre chromatique $\chi(G) = 4$, et un coloriage optimal est donné par $\{a \mapsto 1, b \mapsto 2, c \mapsto$

$3, d \mapsto 1, e \mapsto 4, f \mapsto 2\}$.

- Un sous-ensemble de sommets est appelé un *stable* si il ne contient que des sommets non reliés par des arêtes. Le *nombre stable*, noté $\alpha(G)$, est le nombre maximal de sommets dans un stable de G. Par exemple, le graphe G de la figure 1 admet le nombre stable $\alpha(G) = 2$, et $\{a, e\}$ est un stable maximal.
- Un sous-ensemble de sommets est appelé une *clique* si tous ses sommets sont deux à deux reliés par une arête. Le *nombre clique* d'un graphe G, noté $\omega(G)$, est le nombre maximal de sommets dans une clique de G. Pour le graphe G de la figure 1, $\omega(G) = 4$, avec $\{b, c, d, e\}$ comme clique maximale.
- Enfin, on appelle *nombre de couverture de cliques*, noté $k(G)$, le plus petit entier n tel qu'il existe une famille de cliques $C_1, .., C_n$ telle que $\bigcup_{i=1..n} C_i = V$. Pour le graphe G de la figure 1, $k(G) = 2$, avec $\{a, b, c\}, \{d, e, f\}$ comme couverture de clique mimimale.

On notera qu'en général on a les inégalités suivantes entre les nombres que l'on vient de définir :

$$\omega(G) \;\leq\; \chi(G) \quad \text{et} \quad \alpha(G) \;\leq\; k(G)$$

bien que sur l'exemple considéré, ces inégalités soient des égalités. Ce n'est pas un hasard, et cette double égalité s'étend à tous les graphes d'intervalles, et même à une famille de graphes beaucoup plus grande : les graphes parfaits.

Un graphe est dit parfait si son nombre chromatique est égal à son nombre clique. De manière équivalente[1], un graphe est parfait si le nombre stable est égal au nombre de couverture de cliques. La conjecture forte des graphes parfaits, prouvée en mai 2002 par Chudnovski, Robertson, Seymour et Thomas, dit qu'un graphe est parfait si et seulement si ni lui ni son complémentaire ne contiennent de cycle impair sans corde de longueur supérieure ou égale à 5. De nombreuses autres classes de graphes sont des cas particuliers de graphes parfaits, en particulier les graphes de comparaison et les graphes triangulés, que nous verrons bientôt apparaître dans la caractérisation des graphes d'intervalles.

3 Des intervalles dans tous les coins

Les graphes d'intervalles interviennent dans des problèmes de combinatoire où les objets en interaction partagent une composante physique continue, par exemple le temps, la fréquence, ou la température. Afin d'illustrer l'intérêt de ces

[1]Cette équivalence fait l'objet du théorème des graphes parfaits, conjecturé par Claude Berge au début des années soixante et prouvé en 1972 par Laszlo Lovasz.

! Attention : erreur dans la définition de graphe parfait !

Def1. Le graphe (S',A') est un sous-graphe induit du graphe (S,A) si S' est inclus dans S et si pour tout s,s' dans S', (s,s') appartient à A' ssi (s,s') appartient à A.

Def2. Un graphe (S,A) est parfait si pour tout sous-graphe induit H=(S',A'), le nombre chromatique de H est égal à son nombre de clique.

graphes en pratique, voici quelques exemples de problèmes qu'ils permettent de résoudre :

- **Le problème des frigos**. Dans un grand hopital, les réductions de financement public poussent le gestionnaire du service d'immunologie à faire des économies sur le nombre de frigos à acheter pour stocker les vaccins. A peu de chose près, il lui faut stocker les vaccins suivants :

Nom du vaccin	**Température de conservation**
Rougeole-Rubéole-Oreillons (RRO)	4..12^oC
BCG	8..15^oC
Di-Te-Per	0..20^oC
Anti-polio	2..3^oC
Anti-hépatite B	-3..6^oC
Anti-amarile	-10..10^oC
Variole	6..20^oC
Varicelle	-5..2^oC
Antihaemophilus	-2..8^oC

  Combien le gestionaire doit-il acheter de frigos, et sur quelles températures doit-il les régler ?

- **Le problème du CSA.** Le Conseil Supérieur de l'Audiovisuel doit attribuer de nouvelles bandes de fréquences d'émission pour la stéréophonie numérique sous-terraine (SNS). Cette technologie de pointe étant encore à l'état expérimental, les appareils capables d'émettre ne peuvent utiliser que les bandes de fréquences FM suivantes : 32..36, 24..30, 28..33, 22..26, 20..25, 30..33, 31..34, et 27..31.
  Quelles bandes de fréquences doit-on retenir pour permettre à le plus d'appareils possibles d'être utilisés, sachant que deux appareils dont les bandes de fréquences s'intersectent pleinement (pas juste sur les extrémités) sont incompatibles.

- **Le problème du wagon restaurant.** Le chef de train de l'Orient Express doit aménager le wagon restaurant avant le départ du train. Ce wagon est assez petit et doit être le moins encombré de tables possibles, mais il faut prévoir suffisemment de tables pour accueillir toutes personnes qui ont réservé :

Le baron et la baronne Von Haussplatz	19h30..20h14
Le général Cook	20h30..21h59
Les époux Steinberg	19h..19h59
La duchesse de Colombart	20h15..20h59
Le marquis de Carquamba	21h..21h59
La Vociafiore	19h15..20h29
Le colonel Ferdinand	20h..20h59

Combien de tables le chef de train doit-il prévoir ?

4 Coloriage des graphes d'intervalles

Lorsqu'on sait qu'un graphe est un graphe d'intervalles, et qu'on connait une famille d'intervalles $I_1, .., I_n$ qui l'engendre, il devient possible de résoudre efficacement certains problèmes réputés coûteux en temps de calcul.

Le problème du coloriage d'un graphe constitue un premier exemple : étant donné un graphe $G = (V, E)$, on cherche un entier n minimal et une fonction $c : V \to \{1, .., n\}$ telle que si $\{v_1, v_2\} \in E$, $c(v_1) \neq c(v_2)$. Une heuristique simple pour résoudre ce problème consiste à appliquer l'algorithme glouton suivant : tant qu'il reste reste des sommets non coloriés, en choisir un et le colorier avec le plus petit entier qui n'apparait pas dans les voisins déjà coloriés. En choisissant bien le nouveau sommet à colorier à chaque fois, cette heuristique se révelle optimale pour les graphes d'intervalles. Notons $\prec_i$ l'ordre partiel sur les intervalles ouverts défini par $I \prec_i J$ si $\inf I < \inf J$. On utilise pour l'algorithme l'ensemble $Adj(I)$ des sommets adjacents à I, autrement dit les intervalles qui intersectent I, et on construit la fonction de coloriage $c : V \to \mathbb{N}$ pendant l'exécution de l'algorithme :

Algorithme de coloriage de graphe d'intervalles

Entrée : Une famille V d'intervalles $I_1, .., I_N$.
Sortie : Une fonction de coloriage $c : V \to \mathbb{N}$.

1. pour tout $I \in V$, $c(I) \leftarrow +\infty$
2. tant qu'il existe $I \in V$ tel que $c(I) = +\infty$:
3. choisir $I \in V$ minimal pour $\prec_i$ tel que $c(I) = +\infty$
4. poser $c(I) \leftarrow \inf\,(\mathbb{N} - c(Adj(I)))$
5. fin tantque

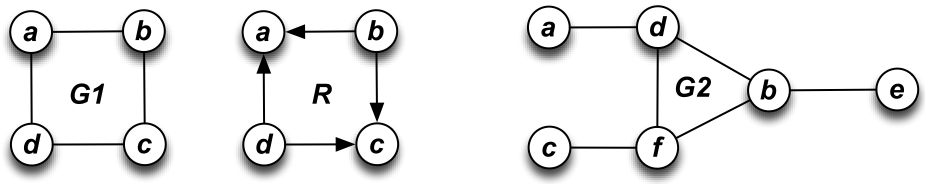

FIG. 2 – Exemples de graphes : G_1 est un graphe de comparaison (R est une orientation transitive de G), mais n'est pas triangulé, tandis que G_2 est triangulé, mais n'est pas un graphe de comparaison.

5 Deux caractérisations des graphes d'intervalles

Revennons à la mort mystérieuse du duc de Densmore. Pour comprendre comment les témoignages recueillis permettent de trouver la coupable, il faut apprendre à reconnaitre les graphes qui sont des graphes d'intervalles. Une première caractérisation peut être obtenue grâce au résultat suivant :

Théorème 1 *Un graphe G est un graphe d'intervalles si et seulement si il existe une énumération $C_1, .., C_n$ des cliques maximales (pour l'inclusion) de G, telle que pour tout sommet v, les indices k tels que $v \in C_k$ sont consécutifs.*

Cette propriété permet de définir un algorithme de reconnaissance des graphes d'intervalles en temps linéaire, consistant à chercher une permutation permettant de satisfaire un ensemble de contraintes de contiguïtés. L'outil permettant de représenter un ensemble de permutations avec contraintes de contiguïté, appelé arbre PG, est toutefois relativement long à détailler. Pour notre problème, on cherche par ailleurs une caractérisation plus visuelle.

Soit $G = (V, E)$ un graphe. Un cycle de longueur n est une suite de sommets $v_0, .., v_n$ telle que $v_0 = v_n$ et $\{v_i v_{i+1}\} \in E$ pour tout i ; une corde d'un cycle $v_0, .., v_n$ est une arête $\{v_i v_j\}$ telle que $|i - j| > 1$. Un graphe est dit *triangulé* si tout cycle de longueur $n \geq 4$ admet une corde.

Une *orientation transitive* d'un graphe $G = (V, E)$ est une relation transitive irréflexive $R \subset V^2$ telle que pour tout $v_1, v_2 \in V$, $\{v_1, v_2\} \in E$ ssi $v_1 R v_2$ ou $v_2 R v_1$. Un graphe $G = (V, E)$ est un *graphe de comparaison* si il admet au moins une orientation transitive.

Théorème 2 *Un graphe G est un graphe d'intervalles si et seulement si G est un graphe triangulé et le complémentaire $\bar{G}$ est un graphe de comparaison.*

6 Questions

En dehors de la question de programmation (question 2) , les questions sont facultatives et indicatives.

1. Qui a tué le duc de Densmore ? (indication : lorsqu'on retire un suspect du graphe des rencontres, si le graphe obtenu reste un graphe qui n'est pas un graphe d'intervalles, alors c'est que ce suspect est innocent).
2. Proposer une structure de donnée adaptée pour représenter un graphe d'intervalles dont une représentation sous forme de famille d'intervalles est connue. Implémenter de manière efficace l'algorithme de coloriage de graphes d'intervalles et illustrer cet algorithme sur une application bien choisie citée dans le texte.
3. Quelle est la complexité de cet algorithme ?
4. Prouver l'algorithme de coloriage.
5. Dans quelle mesure les problèmes d'optimisation cités à la section 3 portent-ils sur des graphes d'intervalles ? Quels problèmes d'optimisation "classiques" soulèvent-ils ? Ces problèmes peuvent-ils être résolus efficacement en général ?
6. Proposer un algorithme efficace pour construire une clique maximum d'un graphe d'intervalles dont on connaît une représentation sous forme de famille d'intervalles.
7. Proposer un algorithme efficace pour construire un stable maximum (un ensemble de sommets indépendants) d'un graphe d'intervalles dont on connaît une représentation sous forme d'intervalles. On pourra chercher à quelle condition l'intervalle dont l'extrémité droite est la plus à gauche appartient à un stable maximum.
8. Montrer que pour les graphes d'intervalles, le nombre chromatique est égal au nombre clique ($\chi(G) = \omega(G)$ d'après les notations du texte).
9. Montrer que le complémentaire d'un graphe parfait est parfait.
10. Démontrer le théorème 1.
11. Démontrer le théorème 2 (indication : on pourra considérer, pour une orientation transitive R de $\bar{G}$, la relation binaire $\tilde{R}$ sur l'ensemble des cliques maximales de G, définie par $C_1 \tilde{R} C_2$ si il existe $v_1 \in C_1, v_2 \in C_2$ tels que $v_1 R v_2$, et montrer que si G est un graphe triangulé, $\tilde{R}$ est un ordre total sur les cliques maximales).