

Colorations des sommets

Complément au chapitre 7 « L'homme à la cagoule »
et au chapitre 9 « L'apprentie sudokiste »

En 1852, le jeune anglais Francis Guthrie s'est demandé s'il est toujours possible de colorer une carte de géographie à l'aide de quatre couleurs, en respectant la condition que deux pays voisins ne soient pas recouverts par la même couleur. Ce n'est qu'en 1976 que deux chercheurs américains, Kenneth Appel et Wolfgang Haken, de l'Université de l'Illinois, ont pu répondre affirmativement à cette conjecture des quatre couleurs. Plus d'un siècle s'est donc écoulé entre l'énoncé et la résolution de ce problème en apparence très simple. Tout au long de ces années, les chercheurs ne sont bien sûr pas restés inactifs et de nombreuses tentatives de démonstration de la conjecture ont donné lieu à de nouveaux développements mathématiques ainsi qu'à une formulation du problème en termes de graphes. La carte à colorer a été remplacée par un graphe, chaque pays étant représenté par un sommet et deux pays voisins étant reliés par une arête.

Définition

Une « coloration des sommets » d'un graphe G est une affectation de couleurs aux sommets telle que chaque arête de G ait ses deux extrémités de couleur différente. On cherche généralement à déterminer une coloration utilisant aussi peu de couleurs que possible. Le plus petit nombre de couleurs nécessaire pour colorer les sommets d'un graphe G est appelé « le nombre chromatique » de G et est noté $\chi(G)$.

On peut aussi colorer les arêtes en évitant que des arêtes incidentes à un même sommet aient la même couleur. C'est ce que nous avons déjà défini dans le chapitre traitant de couplages.

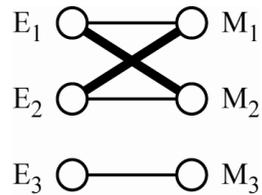
Définition

Une « coloration des arêtes » d'un graphe G est une affectation de couleurs aux arêtes telle que les arêtes ayant une extrémité en commun soient de couleur différente. On cherche généralement à déterminer une coloration utilisant aussi peu de couleurs que possible. Le plus petit nombre de couleurs nécessaires pour colorer les arêtes d'un graphe G s'appelle « l'indice chromatique » de G et est noté $q(G)$.

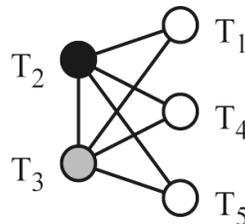
Reprenons l'exemple de coloration des arêtes vu dans un chapitre précédent. On doit réaliser 5 tâches T_1, T_2, T_3, T_4 et T_5 , chacune de ces tâches ayant une durée d'une journée. Les tâches T_1 et T_2 doivent être réalisées par l'employé E_1 , les tâches T_3 et T_4 par l'employé E_2 et la tâche T_5 par l'employé E_3 . Les tâches T_1 et T_3 requièrent la machine M_1 , les tâches T_2 et T_4 la machine M_2 et la tâche T_5 la machine M_3 . Sachant que chaque employé ne peut réaliser qu'une tâche à la fois et que chaque machine ne peut être utilisée que par un employé à la fois, combien de jours faut-il au minimum pour réaliser ces 5 tâches ?

Ce problème est équivalent à une coloration des arêtes du graphe ci-dessous à l'aide d'un nombre minimum de couleurs, de telle sorte que les arêtes qui se touchent aient des

couleurs différentes. Chaque couleur correspond à un jour. Il faut donc 2 jours. Par exemple, on peut faire les tâches T_1 , T_4 et T_5 le 1^{er} jour, et les tâches T_2 et T_3 le 2^{ème} jour.



Considérons maintenant le même problème avec une contrainte supplémentaire. Les tâches T_2 , T_3 et T_5 monopolisent l'unique engin filoguidé de l'entreprise, et cet engin ne peut pas aider à la réalisation de deux tâches simultanément. Le nombre minimum de jours nécessaires à la réalisation des 5 tâches peut être déterminé en colorant les sommets du graphe ci-dessous en un nombre minimum de couleurs, de telle sorte qu'aucune arête n'ait ses deux extrémités de même couleur. À nouveau, chaque couleur correspond à un jour. Il faut donc 3 jours. Par exemple, on peut faire les tâches T_1 , T_4 et T_5 le premier jour, la tâche T_2 le deuxième jour et la tâche T_3 le troisième jour.



En fait, de nombreux problèmes concrets se ramènent à la recherche d'une partition d'un ensemble d'objets en sous-ensembles ne comportant que des éléments compatibles deux à deux. Par exemple, la détermination d'un horaire scolaire est une partition du temps en périodes durant lesquelles seuls des cours pouvant avoir lieu simultanément y sont donnés. On peut modéliser ce genre de problème en termes de coloration des sommets d'un graphe : les sommets sont les objets de la partition et les arêtes représentent les incompatibilités entre les objets. Pour l'exemple de l'horaire scolaire, les sommets sont les cours devant être placés dans la grille horaire, et on relie deux cours par une arête si ces cours ne peuvent pas être donnés simultanément (parce qu'ils concernent un même enseignant, s'adressent aux mêmes étudiants, ou doivent être donnés dans une même salle).

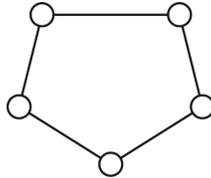
Le problème de la coloration des arêtes d'un graphe a été traité dans un chapitre précédent. Nous nous intéressons ici à la détermination du nombre chromatique. Remarquons qu'en considérant le graphe de ligne $L(G)$ d'un graphe G on a bien évidemment que colorer les arêtes de G est équivalent à colorer les sommets de $L(G)$. L'indice chromatique de G est donc égal au nombre chromatique de son graphe de ligne, ce qui s'écrit :

$$q(G) = \chi(L(G))$$

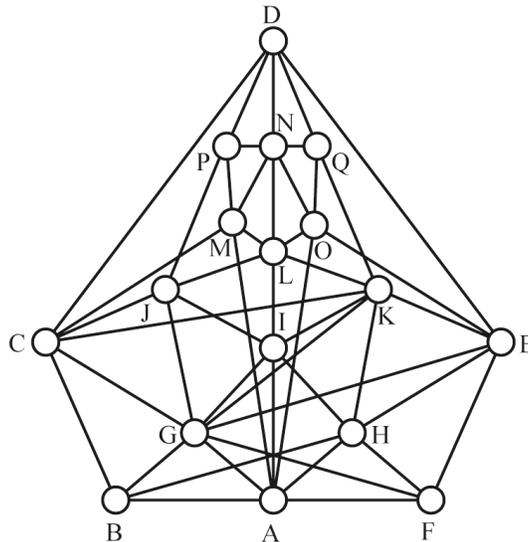
Définition

Une « clique » dans un graphe est un ensemble de sommets tous reliés 2 à 2. La taille de la plus grande clique dans un graphe G se note $\omega(G)$.

Étant donné que les sommets d'une clique doivent tous avoir des couleurs différentes, on déduit que le nombre chromatique $\chi(G)$ de G ne peut pas être inférieur à $\omega(G)$. Il peut par contre être strictement supérieur, comme c'est le cas par exemple pour le pentagone. En effet, il faut $\chi(G)=3$ couleurs pour colorer ses sommets alors que $\omega(G)=2$.



Comme autre exemple, nous pouvons donner le graphe défini par Manori pour déterminer le coupable du braquage à l'épicerie de Madame Rossier.

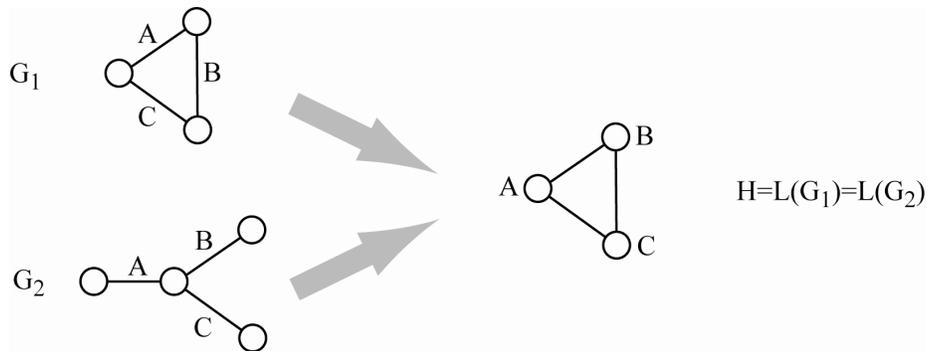


Manori a démontré que bien que ce graphe ne contienne aucune clique à 4 sommets, son nombre chromatique est supérieur à 3. En fait, on ne peut le colorer en 3 couleurs que si on ôte le sommet I.

Pour rappel, Vizing a démontré que l'indice chromatique est toujours au plus égal au plus grand degré + 1, c'est-à-dire $\chi(G) \leq \Delta(G) + 1$. Nous pouvons faire les observations suivantes :

- un ensemble d'arêtes dans G ayant toutes une extrémité en commun correspond à une clique dans son graphe de ligne $L(G)$, ce qui montre que $\omega(L(G)) \geq \Delta(G)$;
- si $L(G)$ n'est pas un triangle, une clique dans $L(G)$ correspond à un ensemble d'arêtes toutes incidentes à un même sommet dans G , ce qui montre que $\Delta(G) \geq \omega(L(G))$.

En résumé, si $L(G)$ n'est pas un triangle, on a $\Delta(G)=\omega(L(G))$. Si $L(G)$ est un triangle, tel que déjà constaté dans un autre chapitre, et tel qu'illustré à nouveau ci-dessous, il se peut que G soit également un triangle, dans quel cas, $2=\Delta(G) < \omega(L(G))=3$.

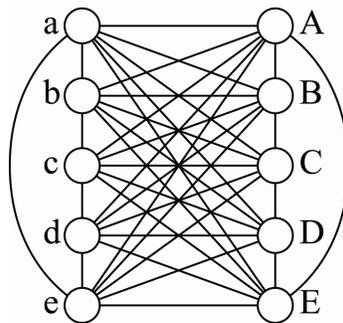


Étant donné que $q(G)=\chi(L(G))$, le résultat démontré par Vizing peut donc se réécrire comme suit :

$$\chi(L(G)) \leq \omega(L(G))+1.$$

Par le raisonnement ci-dessus, ce résultat est vrai si $L(G)$ n'est pas un triangle. Il est également vrai si $L(G)$ est un triangle puisque dans ce cas, $3=\chi(L(G)) < \omega(L(G))+1=4$.

En d'autres termes, nous venons de voir que le nombre chromatique d'un graphe de ligne n'est jamais beaucoup plus grand que la taille de sa plus grande clique. Ce résultat n'est cependant plus forcément valable pour les graphes quelconques. En effet, il a été démontré que l'écart entre $\chi(G)$ et $\omega(G)$ peut être aussi grand qu'on peut l'imaginer. Voici un exemple où $\chi(G) = \omega(G)+2$. Ce graphe a été construit comme suit. On a tout d'abord considéré deux pentagones, l'un constitué des sommets a,b,c,d,e et l'autre des sommets A,B,C,D,E. On a ensuite rajouté une arête entre tous les sommets du premier pentagone et ceux du deuxième pentagone.



Étant donné qu'il faut 3 couleurs pour colorer un pentagone, il faut 6 couleurs pour colorer ce graphe. Aussi, la plus grande clique dans un pentagone contient 2 sommets, ce qui signifie que la plus grande clique dans le graphe en entier est de taille 4. En résumé, on a $\chi(G)=6=4+2=\omega(G)+2$.

Le problème de la détermination du nombre chromatique d'un graphe G est difficile à résoudre. Les meilleurs algorithmes exacts connus à ce jour ne permettent pas de

déterminer le nombre chromatique de graphes ayant plus de 100 sommets. Ces problèmes peuvent par contre être plus faciles pour certaines familles de graphes. Par exemple, pour un graphe biparti comportant au moins une arête, on sait que $\chi(G)=2$ puisqu'on peut attribuer une couleur à l'une des parties de la partition de V , et une autre couleur à la deuxième partie. Nous avons également vu qu'il est facile de colorer un graphe d'intervalles. Pour un gros graphe quelconque, on n'a pas d'autre choix que de recourir aux heuristiques qui donnent des solutions de qualité raisonnable en des temps raisonnables. Nous présentons ici quelques unes de ces heuristiques.

Les *méthodes constructives* de coloration des sommets d'un graphe parcourent séquentiellement les sommets, en attribuant à chaque sommet la plus petite couleur possible (les couleurs sont numérotées). La qualité de la solution ainsi obtenue dépend fortement de l'ordre dans lequel les sommets sont considérés. Celui-ci peut être choisi a priori ou alors construit de manière dynamique. Il est important de remarquer que tout graphe possède un ordre des sommets à partir duquel la méthode constructive produit une coloration en un nombre minimum de couleurs. Pour s'en convaincre, il suffit de considérer une coloration du graphe G en $\chi(G)$ couleurs et d'ordonner les sommets de G en prenant d'abord les sommets de couleur 1, puis ceux de couleur 2, et ainsi de suite.

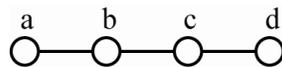
Propriété

Une méthode constructive n'utilisera jamais plus de $\Delta(G)+1$ couleurs.

Preuve

Quand on colore un sommet, un maximum de $\Delta(G)$ couleurs peuvent déjà apparaître dans son voisinage. □

Il est intéressant de noter qu'un algorithme de coloration séquentielle peut se tromper sur le graphe suivant qui ne contient que 4 sommets.



En effet, si on colore ce graphe avec l'ordre a,d,b,c, on donnera la couleur 1 à a et d, la couleur 2 à b et la couleur 3 à c. Ce graphe peut pourtant facilement être coloré à l'aide uniquement de 2 couleurs (couleur 1 pour a et c et couleur 2 pour b et d). Ce graphe est noté P_4 de l'anglais (Path on 4 vertices) car c'est une chaîne sur 4 sommets.

Il a été démontré (par Chvátal) que si un graphe ne contient aucun quadruplet de sommets a,b,c,d ayant la structure d'un P_4 (c'est-à-dire que a est relié à b mais pas à c et d, b est relié aussi à c mais pas à d. et c est également relié à d), alors l'algorithme séquentiel produira une coloration optimale en $\chi(G)$ couleurs, quel que soit l'ordre utilisé.

Les méthodes de coloration constructives les plus répandues peuvent être décrites comme suit. Tout d'abord, parmi les méthodes constructives basées sur des ordres statiques, les plus connues sont les suivantes :

RANDOM

Les sommets sont triés de manière aléatoire.

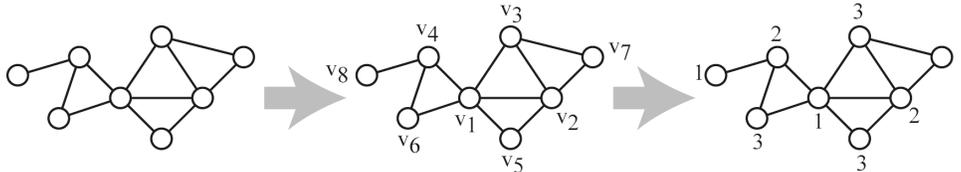
LF (Largest First)

Les sommets sont triés par ordre de degré non croissant.

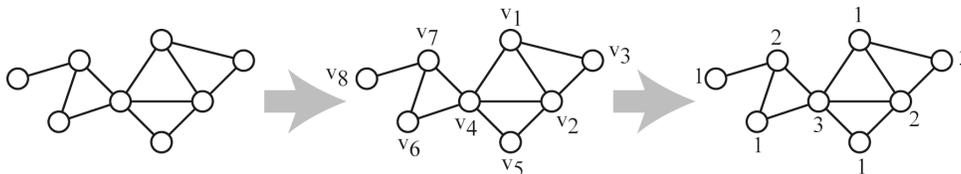
SL (Smallest Last)

L'ordre $v_1, v_2, \dots, v_{|V|}$ des sommets est tel que v_i a le plus petit degré dans le graphe ne contenant plus que les sommets v_1, v_2, \dots, v_i .

Illustration



Coloration avec l'algorithme LF



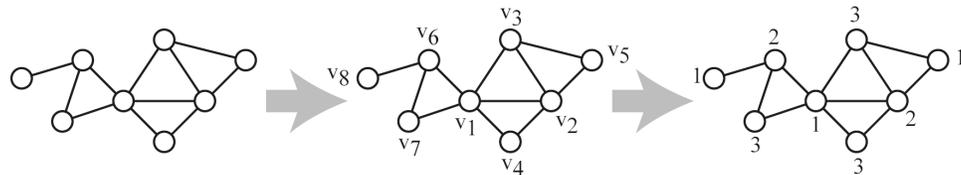
Coloration avec l'algorithme SL

Étant donnée une coloration partielle de G , définissons le degré de saturation $DSAT(v)$ d'un sommet v dans G comme le nombre de couleurs différentes déjà affectées aux sommets voisins de v . Étant donné un sous-ensemble W de sommets dans G , $DEG_W(v)$ mesure le nombre de sommets w dans W qui sont voisins de v . Dans ce qui suit, le sous-ensemble A de sommets correspond à l'ensemble des sommets qui ne sont pas encore colorés. Parmi les méthodes constructives basées sur des ordres dynamiques, la plus utilisée est la suivante :

DSATUR

L'ordre des sommets est construit en choisissant à chaque étape le sommet v de A qui maximise $DSAT(v)$. Si deux sommets v et w ont le même degré de saturation, v précède w pour autant que $DEG_A(v) \geq DEG_A(w)$.

Illustration



Coloration avec l'algorithme DSATUR

Le tableau ci-dessous donne quelques indications sur les performances de ces quatre heuristiques lorsqu'on les applique à des graphes générés aléatoirement. Nous notons $G_{n,d}$ un graphe aléatoire formé de n sommets et dans lequel chaque arête a une probabilité d d'exister, indépendamment des autres arêtes. La probabilité d est appelée la *densité* du graphe. Tous les résultats du tableau ci-dessous sont des moyennes calculées sur dix

graphes aléatoires. Nous donnons pour chaque heuristique le nombre moyen de couleurs utilisées pour colorer le graphe $G_{n,d}$ considéré.

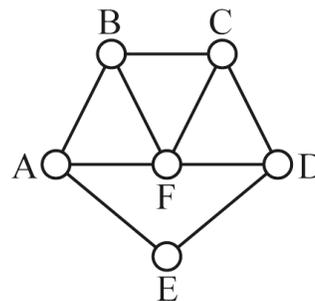
n	d	RANDOM	LF	SL	DSATUR
100	0.4	17.60	16.10	16.65	14.70
	0.5	21.35	20.15	20.50	18.45
	0.6	26.20	24.50	24.85	22.60
300	0.4	39.00	37.10	38.10	34.30
	0.5	48.20	46.00	46.80	43.30
	0.6	61.00	58.20	58.90	53.70
500	0.4	57.20	56.00	56.80	51.00
	0.5	72.40	69.20	71.80	65.00
	0.6	91.40	87.60	89.60	82.80
1000	0.4	99.50	96.50	97.00	92.00
	0.5	126.50	122.00	124.50	116.00
	0.6	160.50	155.00	157.00	146.50

Pour des graphes aléatoires d'au plus 1000 sommets, toutes ces méthodes ont des temps de calcul de l'ordre de la seconde. Notons qu'il existe des algorithmes beaucoup plus efficaces, mais plus lourds à implanter. Par exemple, pour le graphe à 1000 sommets et densité 0.5, les résultats du tableau ci-dessus se situent entre 116 et 127 couleurs alors que les meilleurs algorithmes connus à ce jour parviennent à déterminer des colorations de ces mêmes graphes en 82 couleurs (mais avec quelques heures de calcul).

Nous terminons ce chapitre en reparlant de Manori qui a utilisé la coloration des sommets d'un graphe pour trouver la solution au premier sudoku qui lui a été soumis par la jeune Lei. En fait, tout problème de sudoku revient à colorer un graphe à 81 sommets en 9 couleurs avec quelques sommets ayant déjà une couleur imposée. Ce graphe contient un sommet par case et deux sommets sont reliés par une arête s'ils correspondent à des cases se trouvant sur la même ligne, la même colonne ou le même carré 3x3 de la grille.

Pour la grille proposée par Lei, Manori ne considère que 6 cases particulières représentées ci-dessous.

				5	7		6
	3		1	A	B	9	F
							C
	1	9	5	E		4	D
				3			7
				6			
				7			1
							3



Manori remarque que ces 6 cases ne peuvent recevoir que les chiffres 2, 4 et 8. Il doit donc colorer ce graphe en 3 couleurs, l'une représentant le 2, une autre pour le 4 et une dernière pour le 8. Il fait ensuite remarquer que la seule solution possible attribue une même couleur à A et C, une 2^{ème} couleur à B et D, et finalement une 3^{ème} à E et F. Il suffit alors d'observer que les cases D et E ne peuvent pas recevoir le chiffre 4 pour conclure que la couleur représentant le chiffre 4 ne peut être attribuée qu'aux cases A et C.