Coloriages de graphes.

Épreuve pratique d'algorithmique et de programmation
Concours commun des Écoles normales supérieures

Durée de l'épreuve : 3 heures 30 minutes

Juin/Juillet 2021

ATTENTION !

N'oubliez en aucun cas de recopier votre u_0
à l'emplacement prévu sur votre fiche réponse

Important.

Il vous a été donné un numéro u_0 qui servira d'entrée à vos programmes. Les réponses attendues sont généralement courtes et doivent être données sur la fiche réponse fournie à la fin du sujet. À la fin du sujet, vous trouverez en fait deux fiches réponses. La première est un exemple des réponses attendues pour un $\widetilde{u_0}$ particulier (précisé sur cette même fiche et que nous notons avec un tilde pour éviter toute confusion !). Cette fiche est destinée à vous aider à vérifier le résultat de vos programmes en les testant avec $\widetilde{u_0}$ au lieu de u_0. Vous indiquerez vos réponses (correspondant à votre u_0) sur la seconde et vous la remettrez à l'examinateur à la fin de l'épreuve.

En ce qui concerne la partie orale de l'examen, lorsque la description d'un algorithme est demandée, vous devez présenter son fonctionnement de façon schématique, courte et précise. Vous ne devez en aucun cas recopier le code de vos procédures !
Quand on demande la complexité en temps ou en mémoire d'un algorithme en fonction d'un paramètre n, on demande l'ordre de grandeur en fonction du paramètre, par exemple : $O(n^2)$, $O(n \log n)$,...
Il est recommandé de commencer par lancer vos programmes sur de petites valeurs des paramètres et de ***tester vos programmes sur des petits exemples que vous aurez résolus préalablement à la main ou bien à l'aide de la fiche réponse type fournie en annexe***. Enfin, il est recommandé de lire l'intégralité du sujet avant de commencer afin d'effectuer les bons choix de structures de données dès le début.

1 Préliminaires

Notations et rappels

On rappelle que pour deux entiers naturels a et b, $a \mod b$ désigne le reste de la division entière de a par b, c'est à dire l'unique entier r avec $0 \leq r < b$ tel que $a = k \times b + r$ pour $k \in \mathbb{N}$.
Ce sujet ne portant que sur les graphes non orientés, chaque utilisation du mot graphe doit être entendue comme graphe non orienté. Un graphe est donc, pour ce sujet, un couple $G = \langle V, E \rangle$ où V est l'ensemble fini des sommets ou nœud du graphe et E est l'ensemble des arêtes, chaque arête étant un ensemble de sommets de cardinal 2. Notez que cette définition interdit les boucles (un sommet n'est jamais relié à lui-même par une arête).
Un coloriage d'un graphe $G = \langle V, E \rangle$ est une fonction qui associe une couleur à chaque sommet du graphe. Dans ce sujet les couleurs seront codées par des entiers et un coloriage sera donc une fonction de type $V \to \mathbb{N}$. Un coloriage d'un graphe $G = \langle V, E \rangle$ est dit valide s'il n'existe pas deux sommets de même couleur reliés par une arête. Formellement, un coloriage c est valide quand pour tout $\{a, b\} \in E \Rightarrow c(a) \neq c(b)$.
Le nombre de couleurs d'un coloriage c pour un graphe $\langle V, E \rangle$ est le nombre de couleurs différentes utilisées par c, c'est-à-dire le cardinal de l'ensemble $c(V)$. Le nombre chromatique d'un graphe est le nombre minimum de couleurs d'un coloriage valide de ce graphe.

Générateur de nombres pseudo-aléatoires

Étant donné u_0 on définit la récurrence :

$$\forall t \in \mathbb{N},\ u_{t+1} = (900\,007 \times u_t) \mod 1\,000\,000\,007$$

L'entier u_0 vous est donné, et doit être recopié sur votre fiche réponse avec vos résultats. Une fiche réponse type vous est donnée en exemple, et contient tous les résultats attendus pour une valeur de u_0 différente de la vôtre (notée $\widetilde{u_0}$). Il vous est conseillé de tester vos algorithmes avec cet $\widetilde{u_0}$ et de comparer avec la fiche de résultat fournie. Pour chaque calcul demandé, avec le bon choix d'algorithme le calcul ne devrait demander qu'au plus de l'ordre de la seconde, jamais plus d'une minute.

Question 1 *Calculer les valeurs suivantes :*

a) $u_1 \mod 1\,000$ ***b)*** $u_{42} \mod 1\,000$ ***c)*** $u_{1\,000} \mod 1\,000$ ***d)*** $u_{10^6} \mod 1\,000$

Élimination de doublons dans une liste d'entiers

Il existe différentes méthodes pour éliminer les doublons d'une liste. Dans ce sujet nous allons présenter une méthode qui peut être efficace sur les listes composées d'entiers naturels où l'on connaît une borne sur l'entier le plus grand présent dans la liste.

Le tableau estDoublon est supposé déjà alloué à la taille $N_{max} = 100\,000$
La liste l ne contient que des entiers dans l'intervalle $[0, N-1]$ avec $N_{max} \geq N$

```
fonction ÉlimineDoublon(l) :
    pour i ∈ l faire
        estDoublon[i] ← Faux
    fin
    listeSansDoublon ← liste vide
    pour i ∈ l faire
        si non estDoublon[i] alors
            Ajoute i à listeSansDoublon
            estDoublon[i] ← Vrai
        fin
    fin
    retourner listeSansDoublon
```

Question à développer pendant l'oral 1 *Expliquer le fonctionnement de la fonction* ÉlimineDoublon*, justifier de sa correction et donner sa complexité en fonction de la taille de la liste l (on ne prendra pas en compte le temps d'allocation du tableau* estDoublon*).*

Question 2 *On note $L(M) = \{u_i \mod M \mid 0 \leq i < M\}$. Calculer le cardinal de l'ensemble $L(M)$ pour les valeurs de M suivantes :*

a) $M = 10$ ***b)*** $M = 5\,432$ ***c)*** $M = 98\,765$

Générateur de graphes pseudo-aléatoires

Étant donnés deux nombres N et M on définit le graphe $\mathcal{G}(N, M)$ comme le graphe qui a pour sommets l'ensemble $V = \{0 \dots N-1\}$ et pour arêtes l'ensemble $E = \{\{l_i, g_i\} \mid 0 \leq i < M\}$ où l_i et g_i sont définis de la façon suivante :

$$\begin{aligned} l_i &= u_{2i} \mod N \\ g_i &= (l_i + 1 + (u_{2i+1} \mod (N-1))) \mod N \end{aligned}$$

Comme il est possible d'avoir une paire d'entiers (i, j) (avec $i \neq j$) telle que $l_i = l_j$ et $g_i = g_j$ ou alors $l_i = g_j$ et $g_i = l_j$, il est possible que $\mathcal{G}(N, M)$ contienne un nombre d'arêtes inférieur strictement à M. Pour construire le graphe, il est recommandé de construire dans un premier temps la liste des voisins de chaque sommet avec éventuellement des doublons, puis d'utiliser la fonction ÉlimineDoublon pour les éliminer.

Question 3 *Calculer le nombre de voisins du sommet 0 dans les graphes $\mathcal{G}(N, M)$ pour les valeurs de N et M suivantes :*

a) $N = 5$, $M = 10$ ***b)*** $N = 100$, $M = 300$

c) $N = 1\,234$, $M = 123\,456$ ***d)*** $N = 98\,765$, $M = 234\,567$

Question 4 *Calculer le nombre total d'arêtes des graphes $\mathcal{G}(N, M)$ pour les valeurs de N et M suivantes :*

a) $N = 5$, $M = 10$ ***b)*** $N = 100$, $M = 300$

c) $N = 1\,234$, $M = 123\,456$ ***d)*** $N = 98\,765$, $M = 234\,567$

Question à développer pendant l'oral 2 *Expliquer comment le graphe $\mathcal{G}(N, M)$ est généré et quelle est la complexité de cette construction en fonction de N et M. Expliquer aussi la structure de données choisie pour stocker le graphe.*

2 Une première heuristique pour approximer le nombre chromatique

Un « recoloriage » d'un sommet s dans un coloriage consiste à remplacer la couleur du sommet s par le plus petit entier naturel qui n'est pas utilisé par un des voisins de s.
Un premier algorithme, que l'on nomme COLORIER1 pour colorier les graphes consiste simplement à partir du coloriage où tous les sommets sont coloriés à 0 puis à recolorier chacun des sommets, en commençant par le sommet 0, puis le sommet 1, puis le sommet 2, etc., jusqu'au sommet $N - 1$.

Question 5 *Dans le coloriage renvoyé par* COLORIER1 *sur les graphes $\mathcal{G}(N, M)$ quelle est la couleur du sommet $\lfloor N/2 \rfloor$ et quel est le nombre de couleurs utilisées pour pour les valeurs de N et M suivantes :*

a) $N = 5$, $M = 10$ ***b)*** $N = 100$, $M = 300$

c) $N = 1\,234$, $M = 123\,456$ ***d)*** $N = 98\,765$, $M = 234\,567$

Question à développer pendant l'oral 3 *Justifier que* COLORIER1 *renvoie toujours un coloriage valide. Quelle est la complexité de votre programme pour calculer* COLORIER1 *(on ne prendra pas en compte le temps de construction du graphe) ?*

Question à développer pendant l'oral 4 *Exhiber un exemple de graphe où le nombre de couleurs du coloriage renvoyé par* COLORIER1 *est supérieur strictement au nombre chromatique. Le nombre de couleurs utilisées par* COLORIER1 *reste-t-il, même dans le pire cas, proche du nombre chromatique ?*

3 Colorier avec deux couleurs

Tester si un graphe est coloriable avec deux couleurs est un problème bien plus simple que de tester si un graphe est coloriable avec K couleurs pour $K \geq 3$. Nous allons donc commencer par nous intéresser au cas bicolore.
Pour le cas bicolore on remarquera que dans un coloriage, la validité d'un coloriage correspond au fait qu'un nœud est toujours de la couleur qui n'est pas celle de ses voisins. On peut donc générer un premier coloriage en appliquant cette propriété récursivement puis vérifier que le coloriage obtenu est bien correct.

On note $P(N, M, K)$ comme la chaîne de caractère $b_0 b_1 b_2 \dots b_9$ où $b_i = 1$ quand $\mathcal{G}(N + i, M + i)$ est coloriable avec K couleurs et $b_i = 0$ sinon.

Question 6 *Calculer $P(N, M, 2)$ pour les valeurs de N et M suivantes :*

a) $N = 10,\ M = 6$ ***b)*** $N = 1234,\ M = 543$ ***c)*** $N = 98\,765,\ M = 45\,678$

Question à développer pendant l'oral 5 *Décrire l'algorithme utilisé pour tester si un graphe est coloriable à deux couleurs. Donner et justifier sa complexité (qui ne prend pas en compte le temps de construire le graphe).*

4 Une seconde heuristique pour approximer le nombre chromatique

Étant donné un graphe G, on dit qu'il est naïvement coloriable avec K couleurs si G est vide ou s'il existe un sommet n de degré strictement inférieur à K tel que G' soit naïvement coloriable avec K couleurs où G' est le graphe G où l'on a retiré le sommet n et ses arêtes adjacentes.

Question à développer pendant l'oral 6 *Justifier qu'un graphe naïvement coloriable avec K couleurs est un graphe qui est coloriable avec K couleurs. Montrer que si G est un graphe naïvement coloriable alors n'importe quel G' obtenu en enlevant des arêtes et des sommets de G est naïvement coloriable avec K couleurs.*

Question 7 *Calculer le plus petit K tel que $\mathcal{G}(N, M)$ est naïvement coloriable avec K couleurs pour les valeurs de N et M suivantes :*

a) $N = 10,\ M = 32$ ***b)*** $N = 123,\ M = 3\,456$

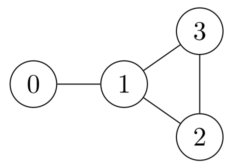
c) $N = 4\,321,\ M = 100\,000$ ***d)*** $N = 54\,321,\ M = 423\,000$

Question à développer pendant l'oral 7 *Décrire l'implémentation de l'algorithme utilisé. Donner et justifier sa complexité.*

5 Une troisième heuristique pour approximer le nombre chromatique

Nous introduisons maintenant un algorithme, Colorier2, pour colorier les graphes. L'idée de cet algorithme est de partir d'un coloriage invalide et de corriger progressivement ses erreurs.
Plus précisément, l'algorithme Colorier2 travaille récursivement, prenant en paramètre un graphe G et un coloriage C. Si C est valide, l'algorithme s'arrête sinon l'algorithme recolorie le sommet qui a le plus de voisins de la même couleur que lui (rappel : on recolorie toujours avec la plus petite couleur non utilisée par un voisin). En cas d'égalité entre plusieurs sommets qui ont le même nombre de voisins de même couleur, l'algorithme recolorie le plus petit sommet (c'est-à-dire le sommet codé par le plus petit nombre).

Par exemple sur le graphe ci-dessous, si tous les sommets sont initialement coloriés à 0, alors COLORIER2 recoloriera d'abord le sommet 1 (qui a trois voisins de couleur 0) en la couleur 1. À ce moment, seuls 2 et 3 sont voisins de même couleur, on recolorie donc le sommet 2 (le sommet 3 a aussi un voisin de même couleur que lui, mais $3 > 2$) en la couleur 2 (la couleur 0 est prise par le sommet 3 et la couleur 1 par le sommet 1). L'algorithme s'arrêtera ensuite.



Question 8 *Dans le coloriage renvoyé par* COLORIER2 *sur les graphes* $\mathcal{G}(N, M)$ *en partant du coloriage où tous les sommets sont coloriés à 0, quelle est la couleur du sommet* $\lfloor N/2 \rfloor$ *et quel est le nombre de couleurs utilisées pour les valeurs de* N*,* M *suivantes :*

a) $N = 5$, $M = 10$

b) $N = 100$, $M = 300$

c) $N = 1\,234$, $M = 123\,456$

d) $N = 98\,765$, $M = 234\,567$

Question à développer pendant l'oral 8 *Décrire l'implémentation de l'algorithme de* COLORIER2 *utilisé. Donner et justifier sa complexité.*

6 Vers un algorithme pour colorier de façon optimale les graphes

Déterminer si un graphe est coloriable avec K couleurs pour $K \geq 3$ est un problème difficile pour lequel les seuls algorithmes connus ont une complexité exponentielle en le nombre de sommets. Un tel algorithme exponentiel consiste à générer tous les coloriages récursivement : tout d'abord on teste toutes les possibilités pour colorier un premier sommet, puis un second, puis un troisième, etc. À chaque fois que l'on a un coloriage complet on peut alors vérifier s'il est correct.

Bien que cette méthode soit valide, elle passe un temps important à générer des coloriages nécessairement invalides. En effet, rien ne sert de continuer à générer les coloriages récursivement si l'on a déjà deux sommets voisins coloriés avec la même couleur. Une seconde idée pour gagner du temps consiste à éviter de tester des coloriages symétriques, par exemple on peut décider que le premier sommet ne sera jamais colorié qu'avec la couleur 0 et donc que pour le second sommet, il suffit de tester la couleur 0 et la couleur 1, etc. Plus généralement, si on a un coloriage partiel avec ℓ couleurs, il suffit de tester les couleurs 0 à ℓ pour le sommet courant. D'autres idées pour réduire le temps de calcul sont possibles, on peut par exemple jouer sur l'ordre dans lequel on colorie les sommets (en coloriant à chaque étape un sommet contraint par ses voisins plutôt que de d'abord colorier le sommet 0 puis le sommet 1, etc.).

Question 9 *Calculer $P(N, M, K)$ pour les valeurs de N, M et K suivantes :*

a) $N = 5$, $M = 18$, $K = 3$

b) $N = 15$, $M = 40$, $K = 3$

c) $N = 35$, $M = 85$, $K = 3$

d) $N = 45$, $M = 100$, $K = 3$

e) $N = 55$, $M = 250$, $K = 4$

f) $N = 65$, $M = 825$, $K = 7$

g) $N = 75$, $M = 925$, $K = 7$

N'hésitez pas, pour cette question, à laisser tourner vos algorithmes quelques minutes mais attention, avec un algorithme trop naïf, il est impossible de réussir les cas les plus durs même en laissant tourner le calcul des heures (l'algorithme ayant une complexité exponentielle).
Pour obtenir tous les points, il faudra donc trouver et implémenter diverses optimisations qui réduisent le temps de calcul. Il est fortement conseillé de commencer par un algorithme naïf et d'implémenter petit à petit les optimisations en testant à chaque fois quels sont les cas résolubles rapidement.

Question à développer pendant l'oral 9 *Expliquer les idées (implémentées ou non) qui ont été explorées pour tester plus rapidement si les graphes sont coloriables avec K couleurs.*

Fiche réponse type : Coloriages de graphes.

$\widetilde{u_0}$: 54

Question 1

a) 378

b) 706

c) 573

d) 94

Question 2

a) 7

b) 3463

c) 62416

Question 3

a) 2

b) 3

c) 172

d) 7

Question 4

a) 6

b) 288

c) 114064

d) 234562

Question 5

a) 1 / 3

b) 3 / 7

c) 27 / 51

d) 3 / 8

Question 6

a) 1110001100

b) 1111101011

c) 0010111111

Question 7

a) 5

b) 38

c) 36

d) 12

Question 8

a) 0 / 2

b) 2 / 5

c) 41 / 46

d) 4 / 6

Question 9

a) 0000101101

b) 1010101111

c) 0100011000

d) 0110111110

e) 0000001011

f) 0000000111

g) 0000110111

Fiche réponse : Coloriages de graphes.

Nom, prénom, u_0 : ....

Question 1

a)

b)

c)

d)

Question 2

a)

b)

c)

Question 3

a)

b)

c)

d)

Question 4

a)

b)

c)

d)

Question 5

a)

b)

c)

d)

Question 6

a)

b)

c)

Question 7

a)

b)

c)

d)

Question 8

a)

b)

c)

d)

Question 9

a)

b)

c)

d)

e)

f)

g)