Bonnie & Clyde

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

Durée de l'épreuve: 3 heures 30 minutes

Juillet 2008

ATTENTION !

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

Important.

Sur votre table est indiqué 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 Introduction

Bonnie Parker (1910-1934) et Clyde Barrow (1909-1934) sont deux criminels américains qui ont perpétré leurs méfaits dans le sud-ouest des États-Unis pendant la grande dépression. Ils étaient spécialisés dans l'attaque à main armée de banques et on estime qu'ils ont tué douze personnes. La nouvelle génération de Bonnie & Clyde n'utilise plus d'armes à feu. En effet, suite à l'apparition des nouvelles technologies, les nouveaux criminels s'intéressent plutôt aux banques en ligne et à comment pénétrer efficacement dans leurs systèmes de sécurité. Un de ces groupes de pirates vous a contacté afin que vous les aidiez à mettre en place leurs attaques.
Les systèmes auxquels ils s'intéressent sont généralement composés d'un ensemble d'ordinateurs reliés entre eux par des liens et on distingue deux ordinateurs particuliers s et t. L'ordinateur s est celui utilisé par les administrateurs systèmes pour surveiller le système et l'ordinateur t est celui auquel s'intéressent nos hackers. Chaque lien dispose d'un certain niveau de sécurité et est donc plus ou moins difficile à contrôler. Afin de minimiser les risques d'être repérés, il est impératif de trouver l'ensemble des liens ayant le plus faible niveau de sécurité et qui leur permettra de déconnecter s et t.
Il est naturel de modéliser ce problème de la façon suivante :

Définition 1 (Réseau) *Un réseau est un triplet $G = (V, E, w)$ où V est un ensemble de nœuds, $E \subset V \times V$ est un ensemble de liens (orientés !) connectant des nœuds et $w : E \mapsto \mathbb{N}^*$ est le niveau de sécurité de chaque lien.*
Dans un réseau, on distinguera toujours deux nœuds particuliers s et t.

Si $V = \{1, \ldots, n\}$, un réseau $G = (V, E, w)$ est donc parfaitement représenté par une matrice $A^{(G)}$ telle que

$$A^{(G)}_{i,j} = \begin{cases} w(i,j) & \text{si } (i,j) \in E, \\ 0 & \text{sinon.} \end{cases}$$

On appellera sous-réseau de G, un réseau $G' = (V, E', w')$ tel que $\forall i, j \in V : A^{(G')}_{i,j} \leqslant A^{(G)}_{i,j}$.

Définition 2 (Chemin) *On appelle chemin de u à v une suite de liens (γ_0, γ_1), (γ_1, γ_2), $\ldots$, (γ_{k-1}, γ_k) telle que $\gamma_0 = u$, $\gamma_k = v$ et les γ_i sont distincts entre eux. Un tel chemin est de longueur k.*

Un chemin valide de $G = (V, E, w)$ est un chemin dont tous les liens sont dans E. Enfin, un nœud x est dit accessible dans G à partir de s s'il existe un chemin valide de s à x dans G.
Déconnecter t de s revient donc à trouver un sous-graphe G' de G tel que t n'est pas accessible dans G' à partir de s. Notre problème se ramène donc à trouver un tel G' tout en minimisant $\sum_{i,j} A^{(G)}_{i,j} - A^{(G')}_{i,j}$.

Définition 3 (Coupe) *On appelle coupe une partition (S, T) de V telle que $s \in S$ et $t \in T$. Une telle partition définit naturellement un ensemble d'arêtes permettant de déconnecter s et t (l'ensemble des $(i,j) \in E$ tels que $i \in S$ et $j \in T$). Le poids d'une coupe est défini par :*

$$c_G(S) = \sum_{\substack{(i,j) \in E \\ i \in S \text{ et } j \in T}} w(i,j)$$

Le poids d'une coupe nous indique donc exactement l'effort à fournir pour isoler T de S et a fortiori t de s. Le poids d'une coupe est donc une surestimation de l'effort nécessaire pour déconnecter t de s et on a alors :

$$\min_{G'} \sum_{i,j} A_{i,j}^{(G)} - A_{i,j}^{(G')} \leqslant \min_{S} c_G(S)$$

Néanmoins, une fois t et s déconnectés dans un sous-graphe G' de G, on peut naturellement définir une coupe induite de la façon suivante :

Définition 4 (Coupe induite) *Soit G' un sous-réseau de G tel que t n'est pas accessible dans G' à partir de s. On note alors $S(G')$ l'ensemble des nœuds accessibles à partir de s dans G'. $S(G')$ est la coupe induite par G'.*

Le poids de la coupe induite est donc inférieur à l'effort nécessaire pour obtenir G' : on a $c_G(S(G')) \leqslant \sum_{i,j} A_{i,j}^{(G)} - A_{i,j}^{(G')}$.
Notre objectif va donc se ramener à trouver une coupe de poids minimum puisqu'une telle coupe définit la meilleure façon de déconnecter t de s. L'énumération de toutes les coupes n'étant pas raisonnable, nous allons essayer de proposer des approches plus pertinentes.

2 Génération aléatoire de réseaux

Considérons la suite d'entiers (u_k) définie pour $k \geqslant 0$ par :

$$u_k = \begin{cases} \textit{votre } \mathtt{u_0} \textit{ (à reporter sur votre fiche)} & \text{si } k = 0 \\ 15\,091 \times u_{k-1} \mod 64\,007 & \text{si } k \geqslant 1 \end{cases}$$

Question 1 *Que valent :* ***a)*** u_{10} ***b)*** u_{100} ***c)*** u_{1000}.

Définition 5 *Pour $n \in \mathbb{N}^*$ et $p \in [0,1]$, on note $G_{n,p} = (V, E, w)$ le réseau à n nœuds défini par :*
– $V = \{1, \ldots, n\}$,
– $(i,j) \in E$ si et seulement si $u_{(i-1).n+j} \leqslant p \times 64006$ et $i \neq j$,
– si $(i,j) \in E$, $w(i,j) = 1 + \left\lfloor \frac{u_{(n+i).n+j} \times 400}{64006} \right\rfloor$.
On s'intéressera à $s = 1$ et $t = n$.

Question 2 *Indiquer le nombre de nœuds auxquels s est directement relié dans les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

3 Une approche gloutonne simple

Le niveau de sécurité d'un chemin de s à t est assez naturellement égal au minimum des capacités des liens qu'il emprunte puisqu'il suffit de supprimer un tel lien pour rendre le

chemin inutilisable. On étend donc naturellement w aux chemins pour définir la sécurité d'un chemin γ par

$$w(\gamma) = \min_{i=1}^{k} w(\gamma_{i-1}, \gamma_i).$$

Enfin, on associe une matrice $A^{(\gamma)}$ à un chemin γ de la façon suivante :

$$A^{(\gamma)}_{i,j} = \begin{cases} 1 & \text{si } (i,j) \in \gamma \\ 0 & \text{sinon.} \end{cases}$$

On se propose de supprimer les arêtes gênantes de sécurité minimale tant qu'il existe un chemin entre s et t.
On aura donc besoin d'une fonction trouvant un chemin entre la source s et la destination t. Pour ce faire, on parcourra récursivement le réseau en partant de s et en visitant systématiquement d'abord les nœuds d'indice les plus faibles. Afin d'arrêter la récursivité, il pourra être utile de noter au fur et à mesure du parcours les nœuds déjà visités. Le chemin obtenu par une telle procédure sera appelé chemin canonique.

Question 3 *Indiquer la longueur du chemin canonique (s'il existe) pour les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

Question 4 *Écrire une fonction qui supprime l'arête de sécurité minimale du chemin canonique (en cas d'égalité, on supprimera la première rencontrée sur le chemin de s à t) dans le réseau G. On appellera réseau résultant le réseau ainsi obtenu. Trouver le chemin canonique (s'il existe) dans le réseau résultant et donnez sa longueur pour les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

Notre procédure consiste donc à trouver un chemin canonique entre s et t et à mettre à jour le réseau résultant en supprimant la première arête de sécurité minimale rencontrée jusqu'à ce que s et t soient déconnectés.

Question 5 *Indiquer la somme des poids des arêtes ainsi supprimées pour les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

Question à développer pendant l'oral : Quelle est la complexité en temps de l'algorithme utilisé à la question précédente. Construire un exemple où cet algorithme renvoie une très mauvaise réponse (c'est-à-dire où beaucoup trop d'arêtes sont enlevées) et comparer cette réponse à la solution optimale.

Question 6 *Calculer la coupe induite par le réseau résultant final et donner son poids pour les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

4 Une approche gloutonne plus évoluée

Supprimer les arêtes une à une prend un certain temps et donne de toutes façons des solutions non optimales. On propose cette fois-ci de déconnecter s et t en supprimant des chemins entiers à chaque fois.

Définition 6 (Flot) *On appellera flot de $G = (V, E, w)$ un ensemble de chemins Γ pondérés $(w_1, \gamma^{(1)}), .., (w_k, \gamma^{(k)})$ valides de G tel que*

$$\sum_{i=1}^{k} w_i A^{(\gamma^{(i)})} \leqslant A^{(G)}$$

Le poids d'un flot est égal à $\sum_{i=1}^{k} w_i$. On peut donc définir le réseau résiduel $G\backslash\Gamma$ de Γ dans G comme le réseau dont la matrice associée est $A^{(G)} - \sum_{i=1}^{k} w_i A^{(\gamma^{(i)})}$.

Notre objectif est de construire un flot Γ de G dont la coupe induite par le réseau résiduel est minimale.
Question à développer pendant l'oral : Soit Γ un flot de G. Montrer que $\sum_{i=1}^{k} w_i \leqslant c_G(S(G\backslash\Gamma))$
Comme précédemment, nous allons construire ce flot de façon gloutonne en cherchant un chemin canonique, en le pondérant par sa sécurité, puis en itérant sur le réseau résiduel jusqu'à ce que s et t soient déconnectés.

Question 7 *Quelle est la somme des poids du flot ainsi construit pour les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

Question 8 *Calculer la coupe induite par le réseau résiduel final et donner son poids pour les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

5 Une approche gloutonne optimale

L'approche précédente n'est pas optimale en raison de la définition du graphe résiduel. On propose d'associer la matrice suivante $F^{(\gamma)}$ à un chemin γ :

$$F_{i,j}^{(\gamma)} = \begin{cases} 1 & \text{si } (i,j) \in \gamma \\ -1 & \text{si } (j,i) \in \gamma \\ 0 & \text{sinon.} \end{cases}$$

Comme précédemment, on va chercher un ensemble de chemins pondérés $(w_1, \gamma^{(1)}), .., (w_p, \gamma^{(k)})$ valides de G connectant s à t et tel que :

$$\sum_{i=1}^{k} w_i F^{(\gamma^{(i)})} \leqslant A^{(G)}$$

Question à développer pendant l'oral : Montrer que s'il existe un ensemble Γ de chemins pondérés $(w_1, \gamma^{(1)}), .., (w_p, \gamma^{(k)})$ valides de G connectant s à t et tel que $\sum_{i=1}^{k} w_i F^{(\gamma^{(i)})} \leqslant A^{(G)}$,alors, il existe un flot $\widetilde{\Gamma}$ tel que $\sum_{i=1}^{k} w_i = \sum_{i=1}^{\widetilde{k}} \widetilde{w_i}$.

On continue donc de produire des solutions valides en calculant le graphe résiduel à l'aide des matrices $F^{(\gamma)}$.

Question 9 *Quelle est la somme des poids du flot ainsi construit pour les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

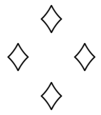
Question 10 *Calculer la coupe induite par le réseau résiduel final et donner son poids pour les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

6 Généralisation

On suppose cette fois qu'il y a plusieurs administrateurs système et donc que t doit être isolé d'un ensemble de nœuds $s_1, \ldots, s_k$.

Question 11 *En utilisant la même technique que dans la section précédente, calculer le poids de la coupe induite et permettant d'isoler le nœud $t = n$ de l'ensemble $\{1, 2, 3\}$ pour les réseaux suivants :*
a) $G_{5,1/2}$ ***b)*** $G_{10,3/10}$ ***c)*** $G_{100,1/20}$

Question à développer pendant l'oral : En supposant que l'algorithme de la section précédente est optimal, montrer qu'il reste optimal quand l'on cherche à isoler t d'un ensemble de nœuds $s_1, \ldots, s_k$.



Fiche réponse type: Bonnie & Clyde

$\widetilde{u_0}$: 4

Question 1

a) 44140

b) 53606

c) 25541

Question 2

a) 2

b) 3

c) 7

Question 3

a) 3

b) 3

c) 53

Question 4

a) Pas de chemin

b) 3

c) 54

Question 5

a) 169

b) 524

c) 26976

Question 6

a) 169

b) 524

c) 4877

Question 7

a) 169

b) 360

c) 1021

Question 8

a) 169

b) 360

c) 2359

Question 9

a) 169

b) 360

c) 1137

Question 10

a) 169

b) 360

c) 1137

Question 11

a) 278

b) 360

c) 1171