



Master2 Informatique

UE: ALGORITHMIQUE DISTRIBUÉES – 4TIN913U

Responsable : M. Gavaille

Date : 8 janvier 2025

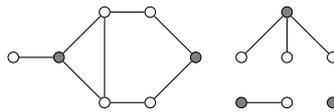
Durée : 1h30

Document: une feuille A4 recto-verso autorisée.

Ensemble dominant

Un *ensemble dominant* d'un graphe G est un ensemble S de sommets de G tel que, pour tout sommet $u \in V(G)$, soit $u \in S$ soit u a un voisin $v \in S$. Comme $S = V(G)$ constitue trivialement un ensemble dominant pour G , on essaye généralement d'en trouver un ayant peu de sommets (dans la mesure du possible).

Par exemple, deux sommets opposés d'un cycle de longueur six forment un ensemble dominant. Pour une étoile (un arbre de hauteur un), un seul sommet suffit : la racine. Notez qu'un sommet isolé dans G , c'est-à-dire de degré 0, appartient à tout ensemble dominant de G . Voir aussi les exemples ci-dessous.



On considère l'algorithme appelé FOREST, qui a pour vocation de calculer grâce à une règle locale un petit ensemble dominant, en tout cas dans certains graphes. Son principe est le suivant :

FOREST « Un sommet de G est sélectionné dans l'ensemble S si son degré est au moins deux. »

On supposera le modèle LOCAL où chaque sommet u de G ne peut utiliser que les deux primitives de communication suivantes :

- SENDALL(m) : envoie le message m à tous les voisins de u .
- RECEIVEALL() : renvoie la liste des messages reçus par u et émis à la ronde courante.

Notez bien qu'un sommet u n'a pas accès directement à la liste de ses voisins dans G . Mais il peut calculer cette liste grâce à son identifiant $ID(u)$ et aux deux primitives ci-dessus. Vous pourrez utiliser NEWROUND pour d'indiquer le démarrage d'une nouvelle ronde.

Question 1. Décrivez l'algorithme FOREST dans le modèle LOCAL et le style du cours en précisant clairement les rondes. Vous supposerez que chaque sommet u possède une variable locale SELECT $\in \{\text{TRUE}, \text{FALSE}\}$ indiquant si $u \in S$ ou pas.

Question 2. Démontrez que l'algorithme FOREST est correct, c'est-à-dire qu'il calcule bien, via les variables locales SELECT, un ensemble dominant S pour G , si G est connexe et possède au moins trois sommets.

Question 3. Comme suggéré par la question 2, expliquez pourquoi l'algorithme FOREST n'est pas correct pour certains graphes.

On dira qu'un ensemble dominant S pour G est *non trivial* si $S \neq V(G)$. Il est facile de voir que G possède un ensemble dominant non trivial si et seulement si G possède au moins une arête.

Question 4. *Modifier votre algorithme de la question 1 de façon à ce qu'il calcule, pour toute forêt G , un ensemble dominant non trivial lorsque c'est possible.*

On considère maintenant un second algorithme, appelé OUTER, dont le principe est le suivant :

OUTER « Chaque sommet détermine son degré qu'il envoie à ses voisins. Puis, un sommet est sélectionné dans l'ensemble S si son degré est au moins 4, ou bien si tous ses voisins sont de degré plus petit que 4. »

Cet algorithme est censé être efficace pour les graphes *planaires-extérieurs*, c'est-à-dire les graphes pouvant être dessinés dans le plan sans croisement d'arêtes où tous les sommets appartiennent au bord de la face extérieure. Les étoiles, arbres, forêts et tous ceux de l'exemple initial sont planaires-extérieurs. On admettra, sans le démontrer, que tout graphe planaire-extérieur possède toujours un sommet de degré au plus 2.

Question 5. *Ajoutez des règles locales supplémentaires à OUTER permettant de garantir, pour tout graphe planaire-extérieur G , le calcul d'un ensemble dominant non trivial pour G , si c'est possible. [Aide : Attention aux sommets de petits degrés (chemins, cycles).] Justifiez. Combien de rondes sont nécessaires en tout ? [NB : On ne demande pas le code précis, la description des règles et leur validité suffisent.]*

On peut montrer que l'algorithme FOREST produit, si les composantes connexes de G sont assez grandes, des ensembles dominants de taille optimale à un facteur 3 près si G est une forêt, et pour OUTER à un facteur 5 près si G est un graphe planaire-extérieur. On peut aussi démontrer qu'aucun algorithme distribué utilisant un nombre constant de rondes ne peut faire mieux en terme de facteur d'approximation pour les forêts et les planaires-extérieurs.

Soit Q_n le graphe planaire-extérieur obtenu à partir d'un chemin P_n à n sommets et en ajoutant toutes les arêtes entre sommets à distance deux dans P_n , voir l'exemple ci-après.



Question 6. *Montrez que pour tout n , P_n possède un ensemble dominant d'au plus $\lceil n/3 \rceil$ sommets.*

Question 7. *Montrez que pour tout n , Q_n possède un ensemble dominant d'au plus $\lceil n/5 \rceil$ sommets.*

Question 8. *En supposant n assez grand, combien de sommets de P_n sont sélectionnés au pire par l'algorithme FOREST modifié comme à la question 4 ? Même question pour l'algorithme OUTER modifié comme à la question 5 pour le graphe Q_n ? Justifiez.*

Complexité de communication

Question 9. *Expliquez ce qu'est un protocole de communication qui calcule une fonction booléenne f ? Quelles en sont les principales règles ? Quel est l'objectif recherché ?*

Question 10. *À quel problème d'algorithmique distribué vu en cours la complexité de communication nous a-t-elle servi ? Citez le résultat.*

On considère la fonction booléenne $\text{pand}_n(x, y) \in \{0, 1\}$ qui étant donnés deux mots binaires x, y , chacun de n bits, renvoie 1 si et seulement si le mot $x \wedge y$ (ET-bit-à-bit) contient un nombre impair de 1. Dit autrement $\text{pand}_n(x, y)$ renvoie la parité du nombre de 1 de $x \wedge y$.

Par exemple, $\text{pand}_6(110111, 011011) = 1$ car $110111 \wedge 011011 = 010011$ qui contient un nombre impair de 1. On a aussi $\text{pand}_6(110111, 011010) = 0$. Notez que $\text{pand}_n(2^i, 2^j) = 1$ si et seulement si $i = j$. On notera $A(n)$ la matrice de la fonction pand_n .

Question 11. *Donnez les matrices $A(1)$ et $A(2)$.*

Question 12. *Démontrez que, pour tout $n \geq 1$, $A(n+1) = \begin{bmatrix} A(n) & A(n) \\ A(n) & \bar{A}(n) \end{bmatrix}$ où $\bar{A}(n)$ représente la matrice $A(n)$ où tous les bits ont été complétés. En déduire que $A(n+1)$ possède une diagonale de zéros.*

On rappelle qu'un ensemble discriminant pour une fonction booléenne f et une valeur $v \in \{0, 1\}$ est un ensemble F_v d'entrées (x, y) pour f tel que, pour tout $(x, y) \neq (x', y') \in F_v$, $f(x, y) = f(x', y') = v$ et $\bar{v} \in \{f(x, y'), f(x', y)\}$, où $\bar{v} = 1 - v$ est la valeur complémentaire de v . Cela revient donc à dire qu'on ne peut pas avoir $f(x, y) = f(x', y) = v$. On notera $\text{CC}(f)$ la complexité de communication de f .

Question 13. *Démontrez que si f possède un ensemble discriminant F_v tel que $|F_v| = 2^k$, pour un certain entier $k \geq 1$ et une valeur $v \in \{0, 1\}$, alors $\text{CC}(f) \geq k + 1$. [Aide : Utilisez les liens établis dans le cours entre $\text{CC}(f)$, le nombre de scripts et la cardinalité des ensembles discriminants pour f .]*

Question 14. *Soient F_v et F'_v deux ensembles discriminants pour une fonction f et la valeur v . Est-il vrai que $F_v \cup F'_v$ est aussi un ensemble discriminant pour f et la valeur v ? Justifiez.*

Question 15. *Donnez un ensemble discriminant F_0 pour pand_2 le plus grand possible, sans démontrer l'optimalité.*

Question 16. *De manière générale, démontrez que pour tout $n \geq 1$, pand_n possède un ensemble discriminant de cardinalité au moins n . Justifiez le caractère discriminant de cet ensemble. [Aide : Pensez aux entrées puissances de deux.]*

Question 17. *Démontrez que $\text{CC}(\text{pand}_{2^k}) \geq k + 1$ pour tout entier $k \geq 1$.*

Dans la suite on ne considèrera que des protocoles de communications sans-préfixe et sans mot vide. Soit f_0 la fonction booléenne définie sur les mots binaires de 2 bits et dont la matrice est la suivante :

	y	00	01	10	11
x		00	01	10	11
	00	1	1	0	0
	01	1	1	0	0
	10	0	0	1	1
	11	0	0	1	1

Question 18. *Montrez que tout protocole de communication pour f_0 nécessite au moins 4 scripts.*

Question 19. *Donnez un protocole de communication pour f_0 de coût 2. [Son principe suffit.]*

On dira qu'une fonction booléenne f sur des mots de n bits est k -similaire si la matrice de f comporte exactement k lignes différentes (parmi les 2^n possibles). Par exemple, la fonction f_0 ci-dessus est 2-similaire.

Question 20. *Proposez un protocole de communication pour une fonction f qui est k -similaire. Précisez, en fonction de k (éventuellement de n), le nombre de scripts de votre protocole, ainsi que son coût. [Aide : Retrouver les résultats des questions 18 et 19.]*