

Vous avez droit à une feuille recto-verso de notes personnelles.

Ensemble dominant

On rappelle les deux algorithmes distribués vus en TP permettant de calculer un ensemble dominant S de taille proche (à un facteur constant) de l'optimal pour un graphe donné qui est soit un arbre ou une forêt (TREE), soit un graphe planaire-extérieur (OUTER).

- Algorithme TREE. Un sommet est sélectionné dans l'ensemble S si son degré est au moins deux.
- Algorithme 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.

Question 1. Décrivez l'algorithme OUTER dans le modèle LOCAL comme avec le langage et conventions du cours tout en précisant le nombre de rondes. [NB : On ne demande pas du code Java à la JBotSim! Vous pourrez utiliser NEWROUND, ID, SEND, SENDALL, RECEIVE, ...]

SOLUTION. Algorithme OUTER (2 rondes)

(code du sommet u)

NEWROUND

- SENDALL(ID(u))
- $L := \text{RECEIVEALL}()$

NEWROUND

- SENDALL(| L |)
- $D := \text{RECEIVEALL}()$
- $\text{SELECT} := (|L| \geq 4) \vee \left(\bigwedge_{d \in D} (d < 4) \right)$

Question 2. Même question pour l'algorithme TREE en corrigeant le problème des composantes connexes ayant un ou deux sommets. Précisez le nombre de rondes.

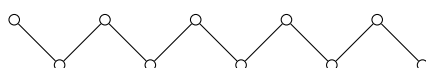
SOLUTION. Algorithme TREE (1 seule ronde)

(code du sommet u)

NEWROUND

- SENDALL(ID(u))
- $L := \text{RECEIVEALL}()$
- $\text{SELECT} := (|L| \geq 2) \vee (|L| = 0) \vee ((|L| = 1) \wedge (\text{ID}(u) < L[0]))$

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.



P_{10}



Q_{10}

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

SOLUTION. En numérotant les sommets de P_n contenu dans Q_n , $v_1, \dots, v_n$, il suffit de prendre les i sommets suivant : $v_2, v_5, v_8, \dots, v_{2+3(i-1)} = v_{3i-1}$ où i est le plus petit indice tel que $3i - 1 \geq n - 1$. (Si le dernier sommet sélectionné v_k a un indice $k < n - 1$, le sommet v_n ne sera pas dominé.) Il suit que $i \geq n/3$, et donc le nombre recherché vaut précisément $i = \lceil n/3 \rceil$.

Question 4. Montrez que pour tout n , Q_n possède un ensemble dominant d'au plus $\lceil n/5 \rceil$ sommets.

SOLUTION. De manière similaire, il suffit de prendre les i sommets suivant : $v_3, v_8, v_{13}, \dots, v_{3+5(i-1)} = v_{5i-2}$ où i est le plus petit indice tel que $5i - 2 \geq n - 2$. (Si le dernier sommet sélectionné v_k a un indice $k < n - 2$, le sommet v_n ne sera pas dominé.) Il suit que $i \geq n/5$, et donc le nombre recherché vaut précisément $i = \lceil n/5 \rceil$.

Complexité de communication

Question 5. Dans le modèle b -CONGEST, expliquez à quoi correspond le problème de la « détection d'un H » dans G , pour un graphe H fixé ? [NB : On ne demande pas sa solution.]

SOLUTION. Le problème est le suivant. En entrée on dispose d'un graphe G . Initialement les sommets connaissent leur identifiant et le graphe H (qui est fixé). Il faut trouver un algorithme distribué qui, pour chaque sommet u , calcule une sortie booléenne DETECT(u) qui est vrai si et seulement il existe une copie (au moins) de H dans G . L'algorithme doit fonctionner dans un mode synchrone où chaque message contient au plus b bits (b -CONGEST).

Question 6. 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 ?

SOLUTION. C'est un algorithme qui permet à deux participants, Alice et Bob, d'échanger des messages jusqu'à obtenir la valeur $f(x, y)$ où x, y sont respectivement les entrées d'Alice et Bob. Ils s'échangent des messages chacun à leur tour, Alice commençant, et le dernier bit doit être la réponse. Bien sûr, chacun ne connaît que son entrée et le protocole. L'objectif est de minimiser le nombre total de bits échangés.

Question 7. À quel problème d'algorithmique distribué vu en cours la complexité de communication nous a-t-elle servi ? Citez le résultat.

SOLUTION. On s'est servi de la complexité de communication pour démontrer que dans le modèle b -CONGEST, il fallait $\Omega(\sqrt{n}/b)$ rondes dans le pire cas pour détecter un C_4 dans un graphe à n sommets.

On considère la fonction booléenne $\inf_n(x, y) \in \{0, 1\}$ qui étant donnés deux mots binaires x, y , chacun de n bits, renvoie 1 si et seulement si « $x \leq y$ », c'est-à-dire si l'entier codé en binaire par le mot x est inférieur ou égal à l'entier codé par le mot y . Par exemple, $\inf_2(01, 11) = 1$ car « $01 \leq 11$ ». En effet, $01_{\text{deux}} = 1_{\text{dix}} \leq 3_{\text{dix}} = 11_{\text{deux}}$.

Question 8. Donnez la matrice de la fonction $\inf_n$, pour $n = 2$.

SOLUTION.

$x \backslash y$	00	01	10	11
00	1	1	1	1
01	0	1	1	1
10	0	0	1	1
11	0	0	0	1

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 $f(x, y') = f(x', y) = v$.

Question 9. Donnez un ensemble discriminant F_1 pour $\inf_2$ le plus grand possible et donnez sa cardinalité. Même question pour un ensemble discriminant F_0 .

SOLUTION. Soient $F_1 = \{(00, 00), (01, 01), (10, 10), (11, 11)\}$ et $F_0 = \{(01, 00), (10, 01), (11, 10)\}$. On a $|F_1| = 4$ et $|F_0| = 3$.

Question 10. Démontrez que tout ensemble discriminant F_v , pour une fonction f et valeur v données, ne peut avoir deux entrées situées sur une même ligne ou même colonne de la matrice de f .

SOLUTION. Supposons que $(x, y), (x', y') \in F_v$ avec $x' = x$ (même ligne). Alors par définition de F_v , $f(x, y) = f(x', y') = v$ et $\bar{v} \in \{f(x, y'), f(x', y)\} = \{f(x', y'), f(x, y)\} = \{v\}$: contradiction car $v \neq \bar{v}$. On obtient la même contradiction si $y' = y$ (même colonne).

Question 11. Généralisez la question 9 à tout n , c'est-à-dire construisez des ensembles discriminants F_1 et F_0 les plus grands possibles pour $\inf_n$. Vous devez justifier les constructions ainsi que leur optimalité. [Aide : La question 10 peut-être utilisée.]

SOLUTION. De manière générale, $F_1 = \{(x, x) : x \in \{0, 1\}^n\}$ et $F_0 = \{(x + 1, x) : x \in [0, 2^n - 1[\}$. On a que F_1 est bien discriminant car F_1 correspond à une diagonale (de valeurs à 1) et les éléments inférieurs à cette diagonale sont tous égaux à 0. Les éléments de F_0 correspondent à une sous-diagonale (de valeurs à 0) où tous les éléments supérieurs sont égaux à 1.

Concernant la taille, on a $|F_1| = 2^n$ et $|F_0| = 2^n - 1$. L'optimalité de la taille se déduit de la question 10 pour $|F_1|$ puisque si $|F_v| > 2^n$, on aurait deux entrées sur la même ligne ou même colonne. Donc $|F_1|, |F_0| \leq 2^n$. De plus, $|F_0| = 2^n$ est impossible car la première ligne est rempli de valeur nulle, interdisant d'avoir 2^n valeurs à 0. Donc $|F_0| < 2^n$.

Question 12. Donnez un minorant, fonction de n , sur le nombre de scripts minimum de tout protocole de communication calculant $\inf_n$. Justifiez. En déduire la complexité de communication de $\inf_n$.

SOLUTION. C'est au moins $|F_1| + |F_0| = 2^n + 2^n - 1$ à cause de la propriété des rectangles. On remarque que $|F_1| + |F_0| > 2^n$ car pour $n \geq 1$, $2^n - 1 \geq 1$. La complexité de communication est au moins le logarithme en base deux du nombre minimum de scripts, ce qui fait $\geq \log_2(|F_1| + |F_0|) > \log_2(2^n) = n$, soit $n + 1$ bits. Ce n'était pas demandé, mais on note que ce minorant est optimal puisque la complexité de communication de toute fonction booléenne (sur des entrées de n bits) est au plus $n + 1$.

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 :

x \ y				
	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 13. Montrez que tout protocole de communication pour f_0 nécessite au moins 4 scripts.

SOLUTION. On peut construire deux ensembles discriminants F_0 et F_1 chacun de taille 2. D'après le cours, le nombre de scripts doit être au moins de $|F_0| + |F_1| = 4$.

Question 14. Donnez un protocole de communication pour f_0 de coût 2. [Son principe suffit.]

SOLUTION. Alice indique à Bob si son entrée x commence par un 0 ou un 1. Puis Bob, avec y , détermine le numéro de colonne de la matrice de f_0 , et en déduit la valeur de $f_0(x, y)$ qu'il retransmet à Alice.

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 15. Proposer un algorithme non distribué qui, étant donnée la matrice de f , permet de calculer l'entier k tel que f est k -similaire. Donnez sa complexité en fonction de n . [Son principe suffit.]

SOLUTION. On trie les L lignes de M en temps $O(T \cdot L \log L)$ où T est le temps de comparaison de deux lignes. Puis, un simple balayage permet de calculer k avec un temps supplémentaire négligeable $O(L)$. On a $L = T = O(2^n)$, et donc la complexité recherchée est $O(n2^{2n})$.

Question 16. 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 13 et 14.]

SOLUTION. Il suffit à Alice de transmettre le numéro i de ligne x (parmi les k distinctes de la matrice) et Bob en déduit avec son entrée y la valeur b qu'il transmet à Alice. Cela suffit car ils connaissent tous deux la matrice f . Dans ce protocole, les scripts sont donc de la forme (i, b) où $i \in \{1, \dots, k\}$ et $b \in \{0, 1\}$, ce qui fait $2k$ scripts différents en tout. Le coût est aussi de $\lceil \log_2(k) \rceil + 1$ bits, puisqu'il suffit de $\lceil \log_2(k) \rceil$ bits pour décrire i et de 1 bit pour b .

Question 17. Pour tout entiers k et n , $1 < k \leq n$, trouver une fonction booléenne sur les mots de n bits qui soit k -similaire et de complexité de communication au moins $\log_2(k) + 1$. [La description de sa matrice suffit pour décrire la fonction, mais vous devez argumenter pour sa complexité de communication.]

SOLUTION. On part de la matrice identité de dimension $k \times k$. On prolonge chacune des k lignes avec $2^n - k$ zéros (par exemple), et on recopie (par exemple) la première ligne $2^n - k$ fois de façon à obtenir une matrice de 2^n lignes et 2^n colonnes. Clairement, il existe un ensemble discriminant F_1 de taille k (diagonale de 1 de l'identité) et F_0 de taille $k - 1$ (diagonale de 0 dessous la diagonale de 1 précédente). Il suit que le nombre de scripts est au moins $2k - 1$, ce qui fait, d'après le cours, une complexité de communication d'au moins $\lceil \log_2(2k - 1) \rceil$ bits, ce qui est toujours au moins $\log_2(k) + 1$.

Pour ce dernier argument, le seul moyen d'avoir $\lceil \log_2(2k - 1) \rceil \neq \lceil \log_2(2k) \rceil$, est que $2k - 1$ soit une puissance de deux et $2k$ juste au dessus. Cela n'est jamais possible si $k > 1$. Il suit que $\lceil \log_2(2k - 1) \rceil = \lceil \log_2(2k) \rceil = \lceil \log_2(k) \rceil + 1$, ce qui est toujours au moins $\log_2(k) + 1$ comme requis.