
Une page A4 recto-verso de notes personnelles autorisées.

Mode synchrone

Question 1. *Donnez les principales caractéristiques du modèle LOCAL, et expliquez les différences avec le modèle CONGEST et la variante b-CONGEST.*

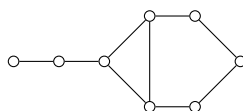
SOLUTION. [3 pts | $\sum 3.0$] Le modèle LOCAL est un modèle synchrone. Tous les processeurs possèdent un identifiant et démarrent en même temps. Les messages sont de taille illimitée. Le modèle b-CONGEST ajoute la contrainte supplémentaire que les messages sont limités à b bits. Le modèle CONGEST est identique au modèle $O(\log n)$ -CONGEST, soit lorsque $b = O(\log n)$.

On considère un algorithme distribué fonctionnant en mode synchrone, appelé DEG3, dont l'objectif est de sélectionner certains sommets d'un graphe G , et on appellera S l'ensemble de tous les sommets ainsi sélectionnés. Son principe est le suivant :

DEG3 « 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 3, ou bien si tous ses voisins sont de degré plus petit que 3. »

Question 2. *Quel est le résultat de l'application de l'algorithme DEG3 si G est le graphe ci-après ? Indiquez sur un dessin les sommets sélectionnés dans l'ensemble S en les distinguant des autres.*

SOLUTION. [3 pts | $\sum 6.0$] Cinq sommets vont être sélectionnés. Les trois sommets du triangles, ainsi que les deux sommets extrémités : le plus à gauche (de degré 1) et le plus à droite (de degré 2).



On supposera le modèle LOCAL où chaque sommet u d'un graphe 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 3. *Décrivez l'algorithme DEG3 dans le modèle LOCAL et le style du cours en précisant clairement les rondes. Vous supposerez que chaque sommet u possède une variable locale $SELECT \in \{TRUE, FALSE\}$ indiquant si $u \in S$ ou pas.*

SOLUTION. [3 pts | $\sum 9.0$]

(code du sommet u)

NEWROUND

- **SENDALL($ID(u)$)**
- **$L := RECEIVEALL()$**

NEWROUND

- SENDALL($|L|$)
- $D := \text{RECEIVEALL}()$
- SELECT := $(|L| \geq 3) \vee (\forall i \in \{1, \dots, |D|\}, D[i] < 3)$

En supposant que les indices de la liste D vont de 1 à $|D|$. Il y a deux rondes.

On rappelle qu'un *arbre* est un graphe ayant la propriété d'être connexe et sans cycle.

Question 4. Pour tout $n \geq 1$, décrivez un arbre à n sommets où lorsque DEG3 est appliqué, tous les sommets sont sélectionnés.

SOLUTION. [3 pts | Σ 12.0] Dans un graphe G où tous les sommets sont de degré < 3 , DEG3 sélectionnera tous les sommets. On peut donc considérer un chemin à n sommets.

Question 5. Pour tout $n \geq 4$, décrivez un arbre à n sommets où lorsque DEG3 est appliqué, aucun sauf un est sélectionné.

SOLUTION. [3 pts | Σ 15.0] Une étoile à n sommets, soit un sommet racine connecté à $n - 1 \geq 3$ feuilles. C'est un arbre qui ne possède qu'un seul sommet de degré ≥ 3 (la racine) qui sera donc sélectionné. Les autres (les feuilles) ont tous leurs voisins (en fait un seul, la racine) de degré ≥ 3 ne seront donc pas sélectionnés.

On dira que S est un ensemble *dominant* de G , si chaque sommet u de G est soit dans S soit voisin d'un sommet de S .

Question 6. Est-il vrai que si G est un arbre, alors l'ensemble S résultant de l'application de DEG3 sur G est dominant ? Justifiez (par un contre-exemple ou une preuve).

SOLUTION. [3 pts | Σ 18.0] C'est vrai non seulement pour tout arbre, mais pour tout graphe G . Soit X l'ensemble des sommets de G qui ne sont pas sélectionnés par l'algorithme DEG3. Les sommets de X doivent donc être de degré < 3 et avoir au moins un voisin de degré ≥ 3 . Soit $u \in X$. Il a donc un voisin v de degré ≥ 3 . Donc $v \in S$. Il suit que u possède un voisin dans S . On a donc montré que soit un sommet est dans S , soit il a un voisin dans S . C'est la définition d'un ensemble dominant pour G .

Complexité de communication

Question 7. 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. [3 pts | Σ 21.0] 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.

On considère la fonction booléenne $\text{odd}_n(x, y) \in \{0, 1\}$ 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. Par exemple, $\text{odd}_8(11011101, 01101110) = 1$ car $11011101 \wedge 01101110 = 01001100$ qui contient un nombre impair de 1. On a aussi $\text{odd}_8(11011110, 01101001) = 0$. Notez que $\text{odd}_n(2^i, 2^j) = 1$ si et seulement si $i = j$. On notera $M(n)$ la matrice de la fonction odd_n .

Question 8. Donnez les matrices $M(1)$ et $M(2)$.

SOLUTION. [3 pts | $\sum 24.0$]

$$M(1) = \begin{array}{c|cc} & y & 0 & 1 \\ \hline x & & & \\ \hline 0 & & 0 & 0 \\ 1 & & 0 & 1 \end{array} \quad M(2) = \begin{array}{c|cccc} & y & 00 & 01 & 10 & 11 \\ \hline x & & & & & \\ \hline 00 & & 0 & 0 & 0 & 0 \\ 01 & & 0 & 1 & 0 & 1 \\ 10 & & 0 & 0 & 1 & 1 \\ 11 & & 0 & 1 & 1 & 0 \end{array}$$

Question 9. Démontrez que, pour tout $n \geq 1$, $M(n+1) = \begin{bmatrix} M(n) & M(n) \\ M(n) & \overline{M}(n) \end{bmatrix}$ où $\overline{M}(n)$ représente la matrice $M(n)$ où tous les bits ont été complétés. En déduire que $M(n+1)$ possède une diagonale de zéros.

SOLUTION. [3 pts | $\sum 27.0$] Les quatre sous-matrices de $M(n+1)$ correspondent aux mots $x', y' \in \{0, 1\}^{n+1}$ de la forme : $(0x, 0y)$, $(0x, 1y)$, $(1x, 0y)$, $(1x, 1y)$ avec $x, y \in \{0, 1\}^n$. Il est clair que la parité des trois premières sous-matrices ne peut changer par rapport à $M(n)$ puisque figure toujours un 0 dans comme premier bit de x' ou de y' . Donc les trois premières sous-matrices valent $M(n)$. Pour la dernière, la parité est inversée puisque x' et y' commence tous les deux par 1, et donc elle vaut $\overline{M}(n)$.

Il est facile de vérifier par induction que $M(n+1)$ possède donc une diagonale montante de 0. C'est effectivement vrai pour $n = 1$. Puis, dans $M(n+1)$ on a deux matrices $M(n)$ alignées sur la même diagonale, ce qui forme donc bien une diagonale de 0 dans $M(n+1)$.

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 $f(x, y') = f(x', y) = v$. On notera $CC(f)$ la complexité de communication de f .

Question 10. 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 $CC(f) \geq k + 1$. [Aide : Utilisez les liens établis dans le cours entre $CC(f)$, le nombre de scripts et la cardinalité des ensembles discriminants pour f .]

SOLUTION. [3 pts | $\sum 30.0$] Par définition de F_v , la matrice de f doit contenir au moins une entrée (x_0, y_0) avec $f(x_0, y_0) = \bar{v}$, sinon F_v ne serait pas discriminant de cardinalité $2^k \geq 2$. Donc cette entrée (x_0, y_0) forme un ensemble discriminant $\{(x_0, y_0)\}$ pour \bar{v} . D'après le cours, $CC(f) \geq \log_2 \sigma(f)$, le nombre de script minimum pour f , et $\sigma(f) \geq |F_v| + |\{(x_0, y_0)\}| = 2^k + 1$. Et donc $CC(f) \geq \log_2(2^k + 1) > \log_2(2^k) = k$. Comme $CC(f)$ et k sont des entiers, cela implique que $CC(f) \geq k + 1$.

Question 11. 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.

SOLUTION. [3 pts | $\sum 33.0$] De manière générale, c'est faux. Pour odd_2 par exemple, on peut considérer $F_0 = \{(11, 00), (10, 01)\}$ et $F'_0 = \{(01, 10), (00, 11)\}$. Pourtant, les entrées $\{(11, 00), (00, 11)\} \subset F_v \cup F'_v$ ne violent pas la propriété du rectangle.

Question 12. Donnez un ensemble discriminant F_0 pour odd_2 le plus grand possible, sans démontrer l'optimalité.

SOLUTION. [3 pts | $\sum 36.0$] Il y a plusieurs solutions possibles. Par exemple, sur la diagonale montante de zéros, $F_0 = \{(11, 00), (10, 01), (01, 10)\}$. Attention ! on ne pouvait pas mettre le 4e élément diagonal.

Question 13. De manière générale, démontrez que pour tout $n \geq 1$, odd_n possède un ensemble discriminant de cardinalité au moins n . Justifiez le caractère discriminant de cet ensemble. [Aide : Pensez aux entrées puissances de deux.]

SOLUTION. [3 pts | $\sum 39.0$] On prend $F_1 = \{(2^i, 2^i) : i \in \{0, \dots, n-1\}\}$, c'est-à-dire des entrées situées sur la diagonale descendante composée d'un seul bit. Clairement, $\text{odd}_n(2^i, 2^j) = 1$ si et seulement si $i = j$. Dit autrement, F_1 est discriminant pour la valeur 1. De plus $|F_1| = n$ comme demandé.

Parenthèse : Pouvait-on faire plus grand que n ? On aurait pu, par exemple, considérer un ensemble maximum I_3 de triplets d'indices $\{i, j, k\}$ deux à deux disjoints (en fait s'intersectant sur un nombre impaires de positions : 0 ou 2 donc ici). En prenant $F_1 = \{(x, x) : x = 2^i + 2^j + 2^k, \{i, j, k\} \in I_3\}$, on avait également que F_1 était discriminant pour la valeur 1, puisque ces éléments diagonaux ont $\text{odd}_n = 1$ et les autres $\text{odd}_n = 0$. Et de même pour I_5, I_7, \dots , l'exemple demandé était pour I_1 . Seulement le calcul de I_3 n'est pas facile. C'est la taille d'une clique maximum du graphe de Kneser $K(n, 3)$ où les sommets sont les sous-ensembles de $\{1, \dots, n\}$ ayant $3 = k$ éléments, et sont adjacents si leur intersection est vide. Il se trouve qu'il est connu que cette taille est $\lfloor n/k \rfloor$, ce qui est maximal quand $k = 1 \dots$ Donc pas intéressant.

Question 14. Démontrez que $\text{CC}(\text{odd}_{2^k}) \geq k + 1$ pour tout entier $k \geq 1$.

SOLUTION. [3 pts | Σ 42.0] D'après les questions précédentes 13 et 10, la fonction possède un ensemble discriminant de cardinalité 2^k . Et donc $\text{CC}(\text{odd}_{2^k}) \geq k + 1$.

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		1	1	0	0
01		1	1	0	0
10		0	0	1	1
11		0	0	1	1

Question 15. Montrez que tout protocole de communication pour f_0 nécessite au moins 4 scripts.

SOLUTION. [3 pts | Σ 45.0] On peut construire deux ensembles discriminants F_0 et F_1 chacun de taille 2 (en considérant les deux 0 et deux 1 diagonaux centraux de la matrice). D'après le cours, le nombre de scripts doit être au moins de $|F_0| + |F_1| = 4$.

Question 16. Donnez un protocole de communication pour f_0 de coût 2. [Son principe suffit.]

SOLUTION. [3 pts | Σ 48.0] 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, et en déduit $f(x, y)$ qu'il retransmet à Alice.