

Question de Cours

Rappelez mathématiquement le problème résolu par l'algorithme de Dijkstra. Quelle est la différence avec l'algorithme A^* ?

Matrices et graphe de Cayley

Soit $G = (V, E)$ un graphe simple à n sommets, et A sa matrice d'adjacence.

1. Rappeler la définition de A .
2. Montrer que $(A^k)_{u,v}$ est le nombre de chemins de longueur k de u à v dans G .
3. En déduire un critère matriciel de connexité.

Soit G un groupe finiment engendré par $S \subset G$. On dit que $G = \langle S \rangle$. On définit le *graphe de Cayley* $\Gamma(G, S)$ comme le graphe non orienté dont les sommets sont les éléments de G , et il y a une arête entre g_1, g_2 si et seulement si $g_1 = sg_2$ pour un certain $s \in S$.

4. Montrer que $\mathbb{Z} = \langle 1 \rangle = \langle 2, 3 \rangle$.
5. Dessiner $\Gamma(\mathbb{Z}, 1)$ et $\Gamma(\mathbb{Z}, \{2, 3\})$.
6. Dessiner $\Gamma(\mathbb{Z}^2, \{(0, 1), (1, 0)\})$.
7. On définit la distance des mots sur G relativement à S comme la distance sur $\Gamma(G, S)$. Calculer la distance entre k, l sur $\mathbb{Z}$ relativement à 1. Calculer la distance sur $\mathbb{Z}^2$ relativement à $(0, 1), (1, 0)$.
8. Donnez un algorithme pour calculer la distance des mots sur G, S . Il devra être en $\mathcal{O}(d_{G,S}(a, b) |S|)$.
9. On dit que $f : X \rightarrow Y$ est bi-Lipschitz s'il existe $K \geq 1$ tel qu'on ait toujours :

$$\frac{1}{K} d_X(x_1, x_2) \leq d_Y(f(x_1), f(x_2)) \leq K d_X(x_1, x_2).$$

Si f est surjective, on dit que c'est une équivalence bi-Lipschitz. Montrez le théorème suivant : si G est finiment engendré par S et S' alors la fonction identité $(G, d_{G,S}) \rightarrow (G, d_{G,S'})$ est une équivalence bi-Lipschitz.

Solution : Matrices et graphe de Cayley

1. La matrice d'adjacence $A \in \mathcal{M}_n(\{0, 1\})$ est définie par

$$A_{u,v} = \begin{cases} 1 & \text{si } \{u, v\} \in E, \\ 0 & \text{sinon.} \end{cases}$$

2. Par récurrence sur $k \geq 1$.

Initialisation. $(A^1)_{u,v} = A_{u,v}$ est le nombre de chemins de longueur 1 de u à v .

Hérédité. Supposons le résultat vrai au rang k . Alors

$$(A^{k+1})_{u,v} = \sum_{w \in V} (A^k)_{u,w} A_{w,v}.$$

Un chemin de longueur $k + 1$ de u à v consiste en un chemin de longueur k jusqu'à un voisin w de v , puis l'arête $\{w, v\}$. Par hypothèse de récurrence $(A^k)_{u,w}$ compte ces chemins, et $A_{w,v} = 1$ ssi $\{w, v\} \in E$: la somme donne bien le nombre de chemins de longueur $k + 1$.

3. Posons $B = \sum_{k=1}^{n-1} A^k$. Alors G est connexe si et seulement si $B_{u,v} > 0$ pour tous $u, v \in V$. En effet, s'il existe un chemin de u à v , il en existe un de longueur $\leq n - 1$ (sinon il contiendrait un cycle et on pourrait le raccourcir).
4. $\mathbb{Z} = \langle 1 \rangle$: tout entier n s'écrit comme n fois le générateur 1. Pour $\langle 2, 3 \rangle$: $3 - 2 = 1 \in \langle 2, 3 \rangle$, donc $\langle 2, 3 \rangle \supset \langle 1 \rangle = \mathbb{Z}$.
5. $\Gamma(\mathbb{Z}, \{1\})$ est un chemin bi-infini : $\dots - (-1) - 0 - 1 - 2 - \dots$
 $\Gamma(\mathbb{Z}, \{2, 3\})$: chaque entier n est relié à $n \pm 2$ et $n \pm 3$; le graphe est 4-régulier.
6. $\Gamma(\mathbb{Z}^2, \{(0, 1), (1, 0)\})$ est la grille carrée infinie : le sommet (a, b) est relié à $(a \pm 1, b)$ et $(a, b \pm 1)$.
7. Dans $(\mathbb{Z}, \{1\})$: $d(k, l) = |k - l|$.
 Dans $(\mathbb{Z}^2, \{(0, 1), (1, 0)\})$: $d((a, b), (a', b')) = |a - a'| + |b - b'|$ (distance de Manhattan).
8. On effectue un BFS depuis a dans $\Gamma(G, S)$. À chaque sommet courant g , on calcule ses voisins $\{sg : s \in S\}$ ($|S|$ opérations). On s'arrête dès qu'on atteint b , à profondeur $d_{G,S}(a, b)$. Chaque niveau contient au plus $|S|^{\text{profondeur}}$ sommets, mais on développe en tout au plus $d_{G,S}(a, b) \cdot |S|$ voisins avant d'atteindre b : complexité $\mathcal{O}(d_{G,S}(a, b) |S|)$.
9. Comme S et S' sont finis et engendrent tous deux G , posons

$$K = \max_{s \in S} d_{G,S'}(e, s), \quad K' = \max_{s' \in S'} d_{G,S}(e, s').$$

Tout chemin de longueur $p = d_{G,S}(x, y)$ dans $\Gamma(G, S)$ se traduit en un chemin de longueur $\leq Kp$ dans $\Gamma(G, S')$, donc $d_{G,S'}(x, y) \leq K d_{G,S}(x, y)$. Symétriquement $d_{G,S}(x, y) \leq K' d_{G,S'}(x, y)$. L'identité (surjective) est donc une équivalence bi-Lipschitz de constante $\max(K, K')$.

Question de Cours

Donnez l'algorithme de parcours en largeur sur un graphe $G = (V, E)$ à partir d'un sommet u . On précisera la structure de données choisie pour représenter le graphe.

Formule d'Euler et problème des cordes

Un graphe $G = (V, E)$ est dit *planaire* s'il peut être dessiné dans le plan de telle sorte que ses arêtes ne se croisent qu'en leurs extrémités communes. Un tel dessin divise le plan en régions appelées *faces* ; on note F leur nombre, en comptant la face extérieure non bornée.

1. Montrer la *formule d'Euler* : pour tout graphe planaire connexe,

$$|V| - |E| + F = 2.$$

2. En déduire que pour tout graphe planaire simple connexe avec $|V| \geq 3$, on a $|E| \leq 3|V| - 6$.

On place $n > 3$ points sur un cercle. On relie chaque paire par une corde, de telle sorte qu'aucun triple de cordes ne se croise en un même point intérieur.

3. Calculer le nombre de régions délimitées à l'intérieur du cercle.

Solution : Formule d'Euler et problème des cordes

1. On démontre $|V| - |E| + F = 2$ par récurrence sur $|E|$.

Initialisation. Si $|E| = 0$, la connexité impose $|V| = 1$; il y a une seule face (la face extérieure), donc $1 - 0 + 1 = 2$. ✓

Hérédité. Soit G connexe planaire à $|E| \geq 1$ arêtes.

- *Cas arbre* : $F = 1$ (pas de cycle, donc pas de face bornée) et $|E| = |V| - 1$, d'où $|V| - (|V| - 1) + 1 = 2$. ✓
- *Cas général* : G contient un cycle ; soit e une arête de ce cycle. Dans $G' = G - e$: $|V'| = |V|$, $|E'| = |E| - 1$, G' est connexe (car e était dans un cycle), et la suppression de e fusionne les deux faces qu'elle bordait, donc $F' = F - 1$. Par hypothèse de récurrence : $|V| - (|E| - 1) + (F - 1) = 2$, d'où $|V| - |E| + F = 2$.

2. Toute face est délimitée par au moins 3 arêtes, et chaque arête borde au plus 2 faces, donc $3F \leq 2|E|$. En substituant $F = 2 - |V| + |E|$:

$$3(2 - |V| + |E|) \leq 2|E| \implies |E| \leq 3|V| - 6.$$

3. Deux cordes se croisent à l'intérieur du cercle si et seulement si leurs quatre extrémités sont distinctes et alternent sur le cercle ce qui revient exactement à choisir 4 points parmi n . Par hypothèse de position générale, chaque quadruplet donne un unique point d'intersection. Le nombre de points intérieurs est donc $\binom{n}{4}$.

4. On construit un graphe planaire G comme suit :

- **Sommets** : les n points du cercle et les $\binom{n}{4}$ intersections intérieures, soit $|V| = n + \binom{n}{4}$.
- **Arêtes** : chaque corde est découpée en segments par ses intersections. La corde reliant deux points est croisée par $\binom{n-2}{2}$ autres cordes (choix des 2 autres extrémités parmi les $n - 2$ restants), donc divisée en $\binom{n-2}{2} + 1$ segments. On ajoute aussi les n arcs du pourtour du cercle, d'où

$$|E| = n + \binom{n}{2} \left(\binom{n-2}{2} + 1 \right) = n + \binom{n}{2} + 3\binom{n}{4},$$

en utilisant $\binom{n}{2}\binom{n-2}{2} = 3\binom{n}{4}$.

G est connexe et planaire. Par la formule d'Euler :

$$F = 2 - |V| + |E| = 2 - \left(n + \binom{n}{4} \right) + \left(n + \binom{n}{2} + 3\binom{n}{4} \right) = 2 + \binom{n}{2} + 2\binom{n}{4}.$$

Le nombre de régions intérieures est $F - 1$ (on retire la face extérieure) :

$$R(n) = 1 + \binom{n}{2} + \binom{n}{4}.$$

Vérification : $R(4) = 1 + 6 + 1 = 8$ ✓

Question de Cours

Rappelez la définition d'un arbre, en tant que graphe. Comparez à la définition inductive d'un arbre binaire.

Composantes fortement connexes

Un *graphe orienté* $G = (V, A)$ est un ensemble de sommets V et d'arcs $A \subseteq V \times V$. On note $u \rightarrow v$ l'arc de u vers v . On dit que v est *accessible* depuis u s'il existe un chemin orienté de u à v , c'est-à-dire une suite $u = v_0 \rightarrow v_1 \rightarrow \dots \rightarrow v_k = v$. On note $u \rightsquigarrow v$ cette relation.

1. Montrer que la relation $u \sim v$ définie par " $u \rightsquigarrow v$ et $v \rightsquigarrow u$ " est une relation d'équivalence sur V .

Les classes d'équivalence de $\sim$ sont appelées *composantes fortement connexes* (CFC) de G . On note $\text{CFC}(G)$ l'ensemble de ces classes. On définit le *graphe des CFC* $\mathcal{G} = (\text{CFC}(G), \mathcal{A})$ où il y a un arc $C_1 \rightarrow C_2$ dans $\mathcal{A}$ si $C_1 \neq C_2$ et il existe $u \in C_1, v \in C_2$ tels que $u \rightarrow v \in A$.

2. Montrer que $\mathcal{G}$ est un DAG (graphe orienté acyclique).
3. On définit le *graphe transposé* $G^T = (V, A^T)$ où $A^T = \{v \rightarrow u \mid u \rightarrow v \in A\}$. Montrer que G et G^T ont les mêmes CFC.
4. On suppose que l'on dispose d'un algorithme $\text{ACCESSIBLE}(G, u)$ qui retourne, en $\mathcal{O}(|V| + |A|)$, l'ensemble des sommets accessibles depuis u dans G . En déduire un algorithme calculant la CFC d'un sommet u donné, en $\mathcal{O}(|V| + |A|)$.
5. En déduire un algorithme calculant toutes les CFC de G , en $\mathcal{O}(|V|(|V| + |A|))$.

On cherche maintenant un algorithme plus efficace. On numérote les sommets dans l'ordre de *fin* d'un DFS sur G : le sommet dont le DFS termine en dernier reçoit le numéro $|V|$, etc. On appelle *puits* une CFC sans arc sortant dans $\mathcal{G}$.

6. Montrer que la CFC du sommet de plus grand numéro de fin est une *source* (CFC sans arc entrant dans $\mathcal{G}$) de $\mathcal{G}$.
7. En déduire que cette même CFC est un *puits* dans le graphe des CFC de G^T .
8. Décrire l'algorithme de Kosaraju :
 - (a) Effectuer un DFS sur G et enregistrer les sommets par ordre décroissant de fin.
 - (b) Effectuer les DFS sur G^T dans cet ordre, chaque DFS ne visitant que les sommets non encore visités ; chaque arbre de DFS est une CFC.

Justifier sa correction et montrer que sa complexité est $\mathcal{O}(|V| + |A|)$.

Solution : Composantes fortement connexes

1. *Réflexivité.* Le chemin vide donne $u \rightsquigarrow u$, donc $u \sim u$.
Symétrie. Si $u \sim v$ alors $u \rightsquigarrow v$ et $v \rightsquigarrow u$, ce qui donne immédiatement $v \sim u$.
Transitivité. Si $u \sim v$ et $v \sim w$, on a $u \rightsquigarrow v \rightsquigarrow w$ donc $u \rightsquigarrow w$, et $w \rightsquigarrow v \rightsquigarrow u$ donc $w \rightsquigarrow u$: ainsi $u \sim w$.
2. Supposons qu'il existe un cycle $C_1 \rightarrow C_2 \rightarrow \dots \rightarrow C_k \rightarrow C_1$ dans $\mathcal{G}$. Alors il existe $u_1 \in C_1$ accessible depuis un sommet de C_k , lui-même accessible depuis un sommet de C_{k-1} , ..., accessible depuis u_1 . Tout sommet de C_2 est donc mutuellement accessible avec tout sommet de C_1 , contredisant $C_1 \neq C_2$. $\mathcal{G}$ est donc acyclique.
3. Soit C une CFC de G et $u, v \in C$. Par définition, $u \rightsquigarrow v$ et $v \rightsquigarrow u$ dans G , ce qui signifie exactement $v \rightsquigarrow u$ et $u \rightsquigarrow v$ dans G^T (les chemins sont renversés). Donc u et v sont dans la même CFC de G^T . La bijection $C \mapsto C$ est bien une égalité des partitions.
4. $\text{CFC}(G, u)$:
 (a) $R \leftarrow \text{ACCESSIBLE}(G, u)$ sommets atteignables depuis u .
 (b) $R' \leftarrow \text{ACCESSIBLE}(G^T, u)$ sommets depuis lesquels u est atteignable.
 (c) Retourner $R \cap R'$.
 Correction : $v \in R \cap R'$ ssi $u \rightsquigarrow v$ et $v \rightsquigarrow u$, i.e. $v \sim u$. Complexité : deux parcours en $\mathcal{O}(|V| + |A|)$.
5. On itère $\text{CFC}(G, u)$ sur tous les sommets non encore affectés à une CFC. Chaque appel est en $\mathcal{O}(|V| + |A|)$ et il y a au plus $|V|$ appels : complexité totale $\mathcal{O}(|V|(|V| + |A|))$.
6. Soit u le sommet de plus grand numéro de fin dans le DFS de G . Supposons qu'il existe un arc $C' \rightarrow C$ dans $\mathcal{G}$ avec $C = [u]_{\sim}$. Alors il existe $v \in C'$ tel que le DFS a visité v avant de terminer u (sinon u serait accessible depuis v et ils seraient dans la même CFC). La propriété du DFS montre que la fin de v est antérieure à celle de u : contradiction avec la maximalité de u . Donc C est une source de $\mathcal{G}$.
7. Dans G^T , les arcs sont renversés : une source de $\mathcal{G}$ devient un puits du graphe des CFC de G^T . Par (3), les CFC sont les mêmes, donc la CFC de u est un puits dans le graphe des CFC de G^T .
8. *Correction.* Par (6) et (7), le DFS sur G^T lancé depuis u (sommet de plus grand fin) ne peut atteindre que des sommets de $[u]_{\sim}$ (car $[u]_{\sim}$ est un puits dans le graphe des CFC de G^T). Il calcule donc exactement $[u]_{\sim}$. En supprimant ces sommets et en répétant avec le prochain sommet non visité, on obtient les CFC une par une dans l'ordre topologique inverse de $\mathcal{G}$.
Complexité. Le premier DFS sur G est en $\mathcal{O}(|V| + |A|)$. L'ensemble des DFS sur G^T visite chaque sommet et chaque arc au plus une fois : $\mathcal{O}(|V| + |A|)$. Complexité totale : $\mathcal{O}(|V| + |A|)$.