École polytechnique — X2013 — INF412 Fondements de l'informatique

Le problème du voyageur de commerce

Sujet proposé par David Monniaux et Bruno Salvy

On appelle *problème du voyageur de commerce* le problème informel suivant : étant donné une carte avec des villes, des chemins et des distances (ou temps de trajet) entre ces villes, fournir un itinéraire qui passe par toutes les villes, revient à la ville de départ, ne visite chaque ville qu'une fois, et qui réalise la distance (ou le temps) minimal. On explique ainsi le nom donné à ce problème : au XIX^{e} siècle, les guides à l'usage des voyageurs de commerce proposaient des itinéraires, sinon minimaux, du moins optimisés. Plus proche de nous, étant donné un ensemble de positions de trous ou de soudures à réaliser, ainsi que les temps de déplacement de la machine-outil d'un emplacement à l'autre, donner l'ordre de réalisation optimal est un problème de voyageur de commerce.

Nous nous intéresserons ici au problème formel suivant : on fixe un nombre n de sommets d'un graphe orienté, et on se donne une matrice $n \times n$ qui à chaque couple (i, j) de sommets associe soit $+\infty$, soit un entier naturel, mesurant le coût pour passer sur l'arête (i, j) ; on se donne également une borne b, le problème est de décider s'il existe un circuit (un chemin qui revient à son point de départ) passant exactement une fois par chacun des sommets et dont la somme des coûts est inférieure ou égale à b. Nous allons montrer que ce problème, noté TSP (*traveling salesman problem*), est NP-complet.

1 Mise en jambes

Question 1.1. *Montrez que TSP est dans NP.*

Un problème lié au TSP est le problème du *chemin hamiltonien* : étant donné un graphe (donné par matrice ou listes d'adjacence), et deux sommets A et B distincts dans ce graphe, dire s'il existe un chemin de A à B qui passe par tous les sommets du graphe, et ne visite chacun qu'une fois.

Question 1.2. *Montrez que le problème du chemin hamiltonien est dans NP.*

Le problème du *circuit hamiltonien* est voisin : étant donné un graphe (donné par matrice ou listes d'adjacence), dire s'il existe un circuit qui passe par tous les sommets du graphe, et ne visite chacun qu'une fois.

Question 1.3. *Montrez que le problème du circuit hamiltonien est dans NP.*

Question 1.4. *Montrez que le problème du circuit hamiltonien est un cas particulier du problème du voyageur de commerce.*

Question 1.5. *Montrez que le problème du circuit hamiltonien se réduit au problème du chemin hamiltonien.*

Question 1.6. *Montrez que le problème du chemin hamiltonien se réduit au problème du circuit hamiltonien.*

En utilisant le fait que 3SAT (dit aussi 3CNFSAT) est NP-complet, on a donc les réductions suivantes :

$$\begin{tikzcd}
& \text{3SAT} \arrow[dr, dotted, bend left] & \\
\text{voyageur de commerce} \arrow[ur, "\text{q1.1}"'] & \text{circuit hamiltonien} \arrow[l, "\text{q1.4}"'] \arrow[u, "\text{q1.3}"'] \arrow[r, bend left, "\text{q1.5}"] & \text{chemin hamiltonien} \arrow[ul, "\text{q1.2}"] \arrow[l, bend left, "\text{q1.6}"]
\end{tikzcd}$$

Il reste donc à prouver que l'on peut réduire 3SAT vers le problème du chemin hamiltonien (la flèche en pointillés).

2 Réduction 3SAT vers le problème du chemin hamiltonien

Il s'agit, à partir d'un problème 3SAT, de construire un graphe de taille polynomiale en celle du problème, qui aura un chemin hamiltonien entre deux points distingués si et seulement si ce problème 3SAT a une solution, ce chemin permettant de construire une solution du problème 3SAT, et réciproquement une solution du problème 3SAT permettant de construire un chemin hamiltonien. .

On va supposer que l'on dispose d'un graphe TRIPLETTE ayant les propriétés suivantes :

1. TRIPLETTE possède, entre autres nœuds, trois nœuds d'entrée e_1, e_2 et e_3, et trois nœuds s_1, s_2 et s_3, tous distincts. On dit que chaque nœud s_i correspond au nœud e_i.
2. Pour tout ensemble d'un (respectivement deux, trois) nœuds d'entrée, il existe un (respectivement deux, trois) chemins commençant sur ces nœuds, finissant sur les nœuds de sortie correspondant respectifs, et tels qu'à eux tous, ces chemins passent par chaque nœud une et une seule fois.
3. Pour tout ensemble d'un (respectivement deux, trois) nœuds d'entrée, il n'existe qu'un (respectivement deux, trois) chemins commençant sur ces nœuds, finissant sur des nœuds de sortie, et, dans leur ensemble, passant par chaque nœud une et une seule fois.

Soit $C_1 \wedge \cdots \wedge C_m$ le problème 3SAT, chaque C_i étant une clause formée sur les variables propositionnelles $y_1, \ldots, y_n$. On construit le graphe formé :

— Des nœuds $v_1, \ldots, v_{n+1}$. Les nœuds $v_1, \ldots, v_n$ sont respectivement associés aux variables $y_1, \ldots, y_n$.

— De m copies de TRIPLETTE, chacune correspondant à une clause C_i, avec les 3 entrées et sorties respectives étiquetées par les variables de la clause.

— Pour chaque variable y_i, on répertorie les clauses $C_{i_1}, \ldots, C_{i_p}$ dans laquelle elle intervient positivement, avec $i_1 < \cdots < i_p$; on connecte v_i à l'entrée correspondante à y_i dans C_{i_1}, la sortie correspondante de C_{i_1} à l'entrée correspondante de C_{i_2}, etc., et la sortie de C_{i_p} est connectée à v_{i+1}. On appelle ces branchements le «sous-graphe positif associé à y_i».

— On opère exactement de la même manière pour les clauses où elle intervient négativement. On appelle ces branchements le «sous-graphe négatif associé à y_i».

Question 2.1. *Montrez qu'à tout chemin hamiltonien de v_1 à v_{n+1} on peut associer une valuation de $y_1, \ldots, y_n$ satisfaisant $C_1 \wedge \cdots \wedge C_m$.*

Question 2.2. *Montrez qu'à toute valuation de $y_1, \ldots, y_n$ satisfaisant $C_1 \wedge \cdots \wedge C_m$ on peut associer un chemin hamiltonien.*

Question 2.3. *Donnez un graphe TRIPLETTE vérifiant les 3 conditions.*

Ceci achève la preuve que le problème du chemin hamiltonien, et donc ceux du circuit hamiltonien et du voyageur de commerce, sont NP-complets. Le problème du voyageur de commerce reste NP-complet si les distances indiquées respectent l'inégalité triangulaire, et le reste encore si elles sont des distances euclidiennes dans le plan.

3 Solutions approchées

Dans le cours, un algorithme en temps polynomial a été présenté pour trouver un chemin presque optimal (à un facteur 2 de l'optimal) lorsque la matrice des coûts est symétrique et vérifie l'inégalité triangulaire : si A est la matrice, $A_{ij} + A_{jk} \geq A_{ik}$ pour tout (i, j, k). Toujours avec cette hypothèse d'inégalité triangulaire, mais maintenant sans la symétrie de la matrice, les meilleurs algorithmes connus s'approchent de l'optimum à un facteur $\alpha \log n$. L'idée du plus classique d'entre eux remonte à 1982 et mène à un facteur $\log_2 n$ de l'optimum. Ce résultat n'a été amélioré que très récemment, en 2008, où une solution à un facteur $0,999 \log_2 n$ de l'optimum a été élaborée. Cette section repose sur le résultat plus simple de 1982.

Question 3.1. *Montrer que sans l'hypothèse d'inégalité triangulaire, il n'existe pas d'algorithme pour trouver une solution approchée à un facteur constant en complexité polynomiale sauf si P=NP. [Indication : prendre un problème de circuit Hamiltonien et rendre le graphe complet en y rajoutant toutes les arêtes manquantes avec un coût $c > 1$.]*

À partir de maintenant, nous considérons que la matrice de coût vérifie l'inégalité triangulaire.

Question 3.2. *Montrer que tout circuit est au plus à un facteur n de l'optimal.*

On admettra qu'il est possible de trouver en complexité polynomiale un recouvrement de tous les sommets du graphe par une union de circuits (chaque sommet du graphe appartient à un et un seul de ces circuits), la somme des coûts de cette union étant optimale. (L'algorithme correspondant s'appelle "la méthode hongroise").

Question 3.3. *Montrer que la somme de ces coûts est au plus celle du circuit optimal, et que le nombre de circuits est au plus $n/2$.*

Question 3.4. *En choisissant un sommet dans chacun de ces circuits et en considérant récursivement le sous-graphe induit par ces sommets, montrer que l'on aboutit à une solution à un facteur $\log_2 n$ de l'optimal en complexité polynomiale.*

Références

[1] David L. Applegate, Robert E. Bixby, Vasek Chvátal, and William J. Cook. *The traveling salesman problem : a computational study.* Princeton university press, 2007. ISBN 9780691129938.

[2] Markus Bläser. A new approximation algorithm for the asymmetric TSP with triangle inequality. *ACM Trans. Algorithms*, 4(4) :Art. 47, 15, 2008. ISSN 1549-6325. doi: 10.1145/1383369.1383378. URL `http://dx.doi.org/10.1145/1383369.1383378`.

[3] A. M. Frieze, G. Galbiati, and F. Maffioli. On the worst-case performance of some algorithms for the asymmetric traveling salesman problem. *Networks*, 12(1) :23–39, 1982. ISSN 0028-3045. doi: 10.1002/net.3230120103. URL `http://dx.doi.org/10.1002/net.3230120103`.

[4] Michael Machtey and Paul Young. *An introduction to the general theory of algorithms.* Computer science library, Theory of computation series. North-Holland, New York, 1978. ISBN 044400226X.

[5] Sartaj Sahni and Teofilo Gonzalez. *P*-complete approximation problems. *J. Assoc. Comput. Mach.*, 23(3) :555–565, 1976. ISSN 0004-5411.