

L'Odyssée abrégée

MARTIN GRÖTSCHEL • MANFRED PADBERG

La recherche du circuit le plus court entre plusieurs villes est le prototype d'une classe importante de problèmes d'optimisation. Parfois cette recherche est si difficile qu'on se contente d'une solution approchée. De nouveaux algorithmes donnent souvent la meilleure solution.

Ulysse, le légendaire roi d'Ithaque, erra pendant 20 ans avant de serrer dans ses bras sa chère Pénélope. Les historiens ont identifié les 16 étapes de l'*Odyssée* (voir la figure 1), et les mathématiciens ont calculé que le périple aurait été plus court si les dieux ne s'en étaient pas mêlés... et si Ulysse avait bénéficié des résultats récents en théorie de l'optimisation.

Le monde moderne a ses Ulysses : les voyageurs de commerce et les hommes d'affaires, qui vont de ville en ville. Imaginons que l'un d'entre eux doive visiter les 16 villes où Ulysse est passé. L'ordre de parcours des villes lui importe peu, mais il se préoccupe de la longueur totale du circuit : le temps, c'est de l'argent. Ce problème est nommé problème du voyageur de commerce. De formulation anodine, il devient si difficile, quand le nombre de villes est notable, que les mathématiciens en cherchent encore une solution. Mieux encore, le problème du voyageur de commerce est un exemple d'une classe de problèmes qui se posent chaque fois qu'on doit identifier une combinaison optimale parmi de nombreuses possibilités. Par exemple, on le rencontre à propos de chargements de camions, de bateaux ou d'avions, de gestion du personnel, d'utilisation de machines, de détermination d'horaires de sociétés, de réalisation de circuits intégrés, d'investissements, de configuration de réseaux de communication...

Entre les 16 villes de l'*Odyssée*, le circuit le plus court mesure 6 859 kilomètres (voir l'encadré de la page 92), et Ulysse n'aurait parcouru que 9 913 kilomètres si, ayant emprunté les ailes de Dédale, il avait volé en ligne droite entre les étapes imposées par les dieux. Comment avons-nous déterminé le

circuit le plus court, parmi les 653 837 184 000 possibilités? Nous avons successivement calculé la longueur de tous les circuits, à l'aide d'un ordinateur de bureau qui a fonctionné pendant 92 heures.

Un tel examen constitue la méthode la plus simple de résolution de problèmes combinatoires, quand le nombre de possibilités est fini. Toutefois, cette utilisation de la puissance de calcul brute n'est possible que dans de rares cas, et elle devient rédhibitoire pour des problèmes à peine plus compliqués que celui d'Ulysse. Nous avons choisi cet exemple, parce qu'il est à la limite des possibilités des ordinateurs actuels, quand ceux-ci font l'examen systématique des combinaisons. Pour 25 villes, en revanche, même les ordinateurs les plus puissants du monde seraient incapables d'examiner toutes les possibilités en un temps raisonnable.

En pratique, les problèmes d'optimisation sont souvent bien plus difficiles que celui d'Ulysse. Des avancées théoriques récentes permettent aujourd'hui de les résoudre. Ainsi, au cours de l'été 1998, David Applegate, Robert Bixby et William Cook, de l'Université de Houston, et Vasek Chvatal, à New Brunswick, ont utilisé les techniques que nous examinerons plus loin pour déterminer le circuit le plus court entre plus de 13 000 villes des États-Unis (voir la figure 2), battant de loin leur propre record, de 7 397 villes seulement.

La modélisation et les graphes

L'analyse d'un problème d'optimisation commence avec sa formalisation mathématique : on remplace chaque ville par un point, ou «sommets», et l'on relie chaque sommet du réseau par

un segment (une «arête»). Quand toutes les paires de sommets sont reliées par une arête, le graphe formé par les sommets et par les arêtes est dit complet. Un graphe complet à 16 sommets comporte exactement 120 arêtes.

Chacune des arêtes est, dans le problème d'Ulysse, affectée d'une valeur : la distance entre les deux villes associées aux sommets. Toutefois d'autres types de valeurs sont possibles : montant de transactions, durée de fabrication, quantité de matière, etc. Un circuit n'est composé que de certaines arêtes, et un circuit complet qui passe par toutes les arêtes une seule fois est un circuit hamiltonien, ainsi nommé d'après le mathématicien irlandais William Hamilton (1805-1865). La longueur du circuit est la somme des longueurs des arêtes empruntées.

Pourquoi la recherche du plus court circuit hamiltonien est-elle si difficile? Parce que le nombre de circuits possibles devient considérable quand le nombre de villes augmente? Non : le nombre de solutions possibles n'est pas une bonne mesure du degré de la complexité des problèmes combinatoires. En effet, imaginons que l'on veuille créer un réseau de télécommunications entre des villes, en faisant éventuellement passer les lignes par des villes intermédiaires, et supposons que le coût de liaison soit proportionnel à la longueur des lignes empruntées par les communications. La solution la plus onéreuse est celle qui correspond au graphe complet, où chaque ville est directement reliée à toutes les autres, mais il existe des solutions moins coûteuses : on élimine progressivement du graphe l'arête la plus longue que l'on puisse ôter sans dissocier le réseau en deux moitiés disjointes, puis on répète l'opération avec le graphe

1. L'ODYSSÉE narre le retour d'Ulysse à Ithaque. Les dieux compliquèrent ce voyage, de sorte qu'Ulysse dut faire escale en 16 points de la Méditerranée. Le problème qui consiste à trouver le circuit le plus court qui passe par ces 16 points impose l'exploration d'un espace à 120 dimensions.

Zeichnung: Peter Connolly / ESA
Alle Mittelmeerländer
Mediterranean and Oceanic Data Base Marine Science
and Technology Programme der EU / Martin Griesel
/ Manfred Patberg



Ray Stemer, Johns Hopkins University, Applied Physics Laboratory <http://fermi.jhuapl.edu/states/> und David Applegate, Robert Bibo, William Cook, Vasek Chvatal

2. LE CIRCUIT représenté par la ligne blanche est un record mondial : c'est le plus court qui passe par les 13 509 villes américaines de plus de 500 habitants.

obtenu. On obtient ainsi un réseau de longueur égale à 4 540 kilomètres, pour les 16 villes d'Ulysse (voir la figure 3).

Cet algorithme est «myope» parce qu'il choisit, à tout moment, la plus avantageuse des diverses possibilités, sans chercher si le choix effectué à une étape donnée ne compromet pas l'accès à une solution qui serait finalement meilleure. Souvent on fait mieux fait d'accepter un inconvénient temporaire quand on escompte un gain supérieur, mais, dans ce cas, la myopie n'est pas gênante : on peut démontrer que l'algorithme conduit toujours à une solution optimale.

Combien de possibilités se présentent-elles, dans ce problème ? Pour n villes, ces solutions sont des graphes qui ont exactement $n - 1$ arêtes sans

circuit fermé passant par une partie des arêtes. De tels graphes sont des arbres. Il y a exactement un siècle, Arthur Cayley (1821-1895) a démontré que, pour n villes, il existe exactement n^{n-2} arbres. Ainsi, pour 16 villes, le nombre d'arbres est égal à 72 057 594 037 927 936, ce qui est bien supérieur au nombre de circuits possibles entre les différentes villes visitées par Ulysse. Pourtant, un ordinateur de bureau trouve en quelques millisecondes la solution optimale au problème des arbres.

Si le nombre de solutions possibles ne fait pas la difficulté du problème du voyageur de commerce, pourquoi ce problème est-il si compliqué ? Nous ne le savons pas. La théorie de la complexité aide à comprendre

la différence entre les problèmes faciles, tels ceux des arbres, et les problèmes difficiles, tel celui du voyageur de commerce, mais cette théorie montre aussi que nombre de problèmes combinatoires d'importance pratique sont complexes.

Parfois, une modification minime de l'énoncé d'un problème change complètement son type. Supposons, par exemple, que l'on se préoccupe de la fiabilité d'un réseau de télécommunications : on veut que le réseau continue de fonctionner malgré les incendies, les séismes ou les coups de pelleuse malencontreux, qui détruiraient certaines lignes. Le réseau ne peut alors avoir la structure arborescente précédente, car la suppression d'une arête dissocierait l'arbre en deux arbres qui ne communiqueraient plus (ce type de considérations a parfois été négligé : le remplacement des câbles métalliques par des fibres optiques, aux États-Unis, coûtait si cher que les opérateurs ont cherché des solutions minimales... qui ont conduit à l'effondrement catastrophique du réseau téléphonique de Chicago, en 1998).

Plus les circuits sont nombreux, entre les paires de sommets d'un réseau, plus le réseau est fiable ; le réseau le plus fiable serait celui qui correspond au graphe complet, mais il serait trop onéreux. Pour décider du type de redondance à assurer, les ingénieurs doivent étudier la fiabilité en fonction des coûts de réalisation et de maintenance. En pratique, certaines liaisons sont plus importantes que d'autres, et des priorités doivent être prises en compte. On formalise facilement ces nouvelles données, mais la complexité des problèmes augmente alors parfois.

Attribuons, par exemple, des niveaux de priorité aux différentes villes à relier : 0 pour les villes peu importantes, 1 pour des villes plus importantes, 2 pour les grandes villes, etc. On impose que toute ville d'un rang i soit reliée aux autres villes de même rang, même quand i liaisons directes tombent en panne. On cherche alors le moins cher de tous les réseaux qui satisfont cette condition (voir l'encadré de la page 94). Les problèmes réels comportent en outre une foule de conditions secondaires : un câble ou un relais ne peuvent traiter simultanément qu'un nombre limité de conversations, ce qui diminue le nombre de

Tous les chemins mènent à Ithaque

Combien de circuits passent-ils par n villes ? Calculons ce nombre par récurrence. Supposons que nous disposions de la liste des circuits passant par $n - 1$ villes. Nous obtiendrions la liste pour n villes en ajoutant à chaque circuit à $n - 1$ villes un crochet par la n -ième ville ; dans chaque cas, $n - 1$ crochets sont possibles. De ce fait, le nombre de circuits possibles passant par n villes est égal au produit de $(n - 1)$ par le nombre de circuits passant par $n - 1$ villes (par ce procédé de construction, les circuits ne sont comptés qu'une seule fois). Pour trois villes, il n'existe qu'un seul circuit (quand on ne distingue pas un circuit du circuit inverse). Pour quatre villes, trois circuits sont possibles ; pour cinq villes, il en existe 3×4 ; et, plus généralement, il

existe $3 \times 4 \times 5 \times \dots \times (n - 1)$, soit $(n - 1)!/2$ circuits qui passent par n villes (dans cette formule, le point d'interrogation désigne la fonction factorielle). De ce fait, entre les 16 villes de l'Odyssée, le nombre de circuits possibles est égal à 653 837 184 000.

Nous avons calculé la longueur de tous ces circuits et obtenu, pour les plus courts, les longueurs suivantes : 6 859, 6 865, 6 870, 6 875, 6 889, 6 890, 6 895, 6 898, 6 900 kilomètres. Le circuit le plus long mesurait 16 434 kilomètres. Le graphique ci-dessous indique le nombre de circuits en fonction de leur longueur. La médiane est à 13 072 kilomètres : la moitié des circuits sont plus longs que cette valeur. La longueur de circuit la plus fréquente, 13 359 kilomètres, est obtenue 229 millions de fois.

