

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.

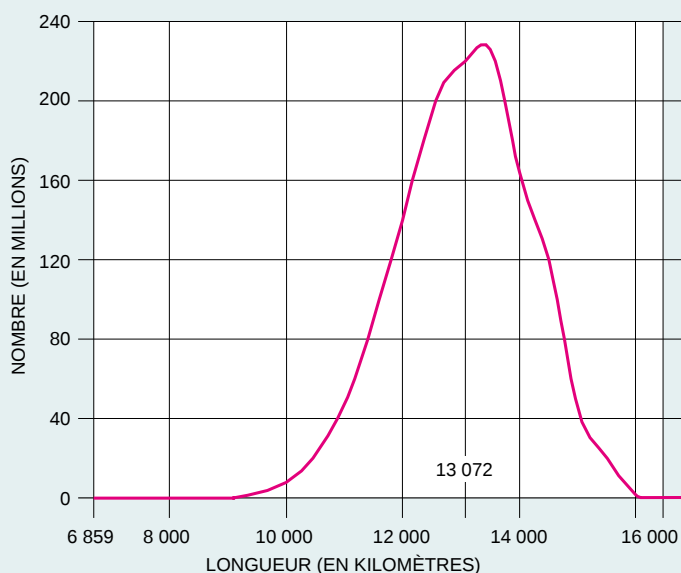
Attribuons, par exemple, des niveaux de priorité aux différentes villes à relier : 0 pour les villes peu importantes, 1 pour des villes plus importantes, 2 pour les grandes villes, etc. On impose que toute ville d'un rang i soit reliée aux autres villes de même rang, même quand i liaisons directes tombent 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.



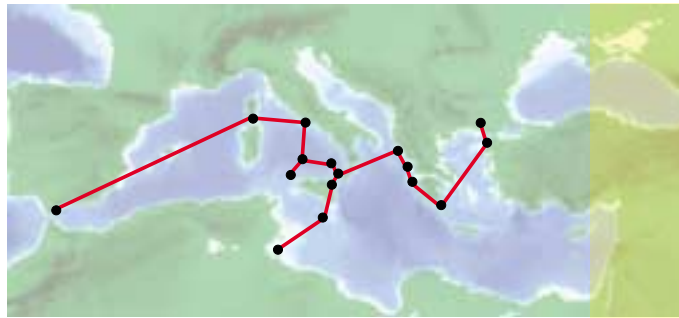
solutions possibles sans simplifier le problème, bien au contraire.

Comment trouve-t-on les solutions les meilleures? Comment savoir si une solution calculée est vraiment bonne?

Bornes supérieure et inférieure

Quand on a déterminé une fonction coût (longueur de chemin, coût, etc.) que l'on veut minimiser, l'existence d'une solution optimale est claire, car il n'existe qu'un nombre fini de solutions, de sorte qu'on doit pouvoir en trouver une pour laquelle la fonction coût prend la valeur minimale. Nous ignorons initialement cette valeur minimale, mais peut-on l'encadrer par une borne inférieure et par une borne supérieure?

On trouve facilement des bornes supérieures (chaque solution du



3. LE RÉSEAU de communications le plus économique, entre les 16 villes de l'Odyssée.

problème en constitue une), mais celles qui sont les plus intéressantes sont naturellement celles qui sont proches de la solution optimale. Des méthodes fondées sur des principes intuitifs (par exemple, joindre un sommet à son voisin le plus proche) fournissent des solutions raisonnablement proches de la solution optimale ; elles sont nommées méthodes heuristiques. On n'est pas certain qu'elles sont efficaces, mais elles le sont souvent.

L'information selon laquelle toute solution a une fonction coût supérieure à la borne inférieure est également précieuse, car, quand une solution n'est pas très éloignée de la borne inférieure, on peut l'admettre à la place de la solution optimale : une longue recherche ultérieure ne conduirait qu'à un faible gain. Comment identifier une borne inférieure? Une bonne stratégie de

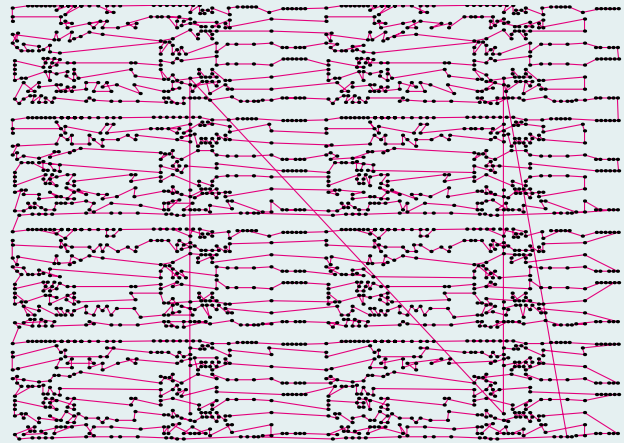
recherche consiste à utiliser à chaque étape les informations utilisées à l'étape précédente : c'est ainsi que sont nés les progrès des dernières années. Utilisons le voyage d'Ulysse pour examiner cette stratégie.

La recherche myope, qui ne tient compte que du gain directement obtenu à chaque étape, est utile pour la recherche d'une borne supérieure : à partir de chacune des villes, on se

Le long voyage d'une perceuse

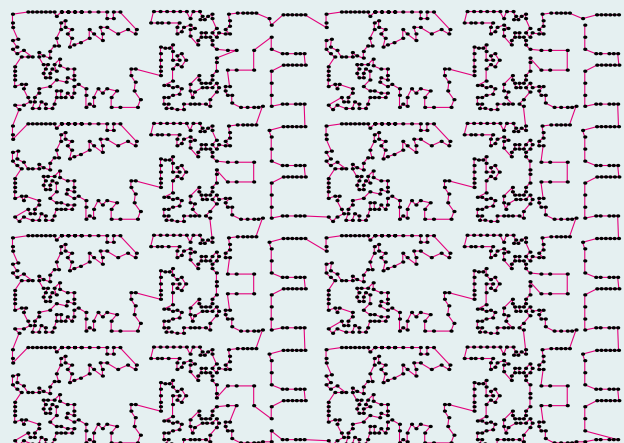


Le problème.



Une solution médiocre.

Supposons que nous utilisions une perceuse ou un faisceau laser pour percer 2 392 trous dans un circuit imprimé. Tout comme Ulysse, la perceuse devra atteindre chaque trou (l'équivalent d'une ville) une fois seulement. Si le temps de perçage est constant, la seule façon d'accélérer la réalisation du circuit consiste à réduire le parcours de la perceuse. La solution représentée en haut à droite, par exemple, est médiocre : sa longueur est égale à 718 876 unités, contre 378 032 unités pour la solution optimale (en bas à droite). Une station de travail met à peine une heure pour trouver cette dernière solution si elle exécute un algorithme efficace. Pour ce problème, on a admis que le temps de parcours de la perceuse était proportionnel à la distance entre deux trous.



La solution optimale.

Martin Grötschel, Manfred Padberg

rend dans la plus proche des villes voisines encore non visitées ; puis, quand toutes les villes ont été visitées, on revient au point de départ.

Quand on applique cette stratégie à des villes quelconques, à l'aide de papier et de crayon, on fait rapidement deux observations : le résultat dépend beaucoup du point de départ, et la solution obtenue est facile à améliorer. Ainsi, lorsque deux arêtes se croisent, on gagnera à remplacer la partie de circuit composée des quatre sommets liés par ces deux arêtes par un circuit où les arêtes sont des bords opposés du quadrilatère. Parfois on gagne même à remplacer les deux côtés du quadrilatère par les autres côtés, et d'autres améliorations de ce type sont possibles.

Pour perfectionner l'algorithme de recherche du circuit le plus court, on doit ainsi corriger les effets néfastes de la myopie. Si l'on s'obstinaît à améliorer les solutions en ne considérant que les solutions voisines, on risquerait de rester piégé dans un minimum local, c'est-à-dire une solution qui n'est pas particulièrement

bonne, mais que les règles d'évolution ne permettent pas d'améliorer (de même, si on laisse rouler une bille sur une tôle déformée, elle ira se loger dans un creux, mais, si elle n'a pas assez d'énergie pour en ressortir, elle risque d'y rester même si le creux n'est pas le point le plus bas de la tôle). On devra donc utiliser des perturbations qui conduiront assez loin de la solution identifiée à chaque étape. De telles méthodes ont souvent des noms amusants, tels que méthode de Monte-Carlo, recuit simulé, évolution, algorithme du déluge...

La programmation de ces méthodes est simple quand les problèmes sont de complexité moyenne. Aussi les problèmes d'optimisation ont-ils été souvent examinés par des amateurs, qui n'ont pas toujours obtenu de bons résultats. Pour étudier des problèmes du voyageur de commerce avec des dizaines de milliers de sommets (un ordre de grandeur courant, en pratique), on doit structurer les données et mettre en œuvre des techniques informatiques raffinées, afin d'obtenir des solutions en un temps acceptable.

Simultanément les heuristiques sont restées très utiles pour la résolution de problèmes pratiques.

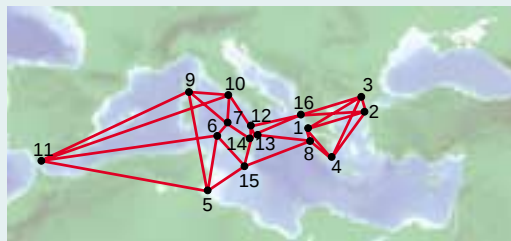
Comment détermine-t-on la borne inférieure de la fonction coût, pour un problème d'optimisation combinatoire ? L'idée fondamentale, apparemment paradoxale, consiste à simplifier le problème en le généralisant : on ajoute au problème initial un autre problème, qui possède davantage de solutions, de sorte que chaque solution du problème initial soit également une solution du problème élargi. On peut, par exemple, rendre moins strictes les conditions secondaires du problème initial, voire les supprimer.

Une solution minimale du problème élargi se caractérise par un coût inférieur à celui de toute autre solution et, par conséquent, inférieur à celui des solutions du problème initial. C'est exactement ce que l'on cherche : une borne inférieure pour le problème initial.

Évidemment le problème élargi doit être plus simple que le problème initial, et sa solution optimale doit

La fiabilité des réseaux

Supposons que nous devons relier les 16 villes étapes d'Ulysse par un réseau de télécommunications qui résiste aux défaillances de ses lignes. Pour des raisons historiques, nous décidons d'attribuer une priorité 1 à Ithaque (1), Troie (2), Djerba (5), Circée (10) et Messine (13) ; les autres villes auraient un niveau de priorité inférieur (0). Autrement dit, les liaisons entre les villes prioritaires devraient être préservées même en cas de panne des connexions directes entre ces villes. Si les coûts de construction sont proportionnels à la longueur des lignes, le réseau le moins coûteux a une longueur de 5 842 kilomètres (*ci-dessous*). Aujourd'hui les concepteurs de réseau cherchent empiriquement un équilibre raisonnable entre les coûts de réalisation et la fiabilité en modifiant les para-



Toutes les villes sont de rang 3.

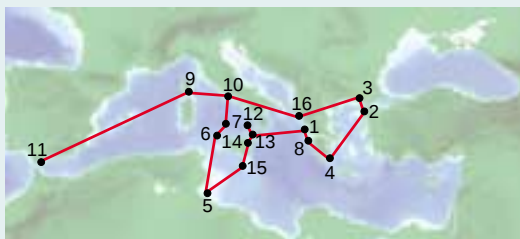
mètres. Si l'on peut dépenser plus pour assurer la fiabilité du réseau, par exemple en assignant la valeur 3 à chacune des villes, le réseau le moins cher que l'on puisse concevoir est de longueur égale à 16 112 kilomètres : avec un tel réseau, les connexions subsisteraient si trois connexions de chaque ville tombaient simultanément en panne. Si le central d'une ville était en panne (par suite d'un incendie du

central, par exemple), les autres villes resteraient interconnectées. Le réseau ne serait interrompu que si les villes 8 (Zakinthos) et 16 (Corfou) tombaient simultanément en panne. Dans les problèmes réels, tel celui du réseau téléphonique de l'Est des États-Unis, la solution

la plus économique serait celle de la figure en bas à droite, où les villes d'importance 0 sont représentés par un cercle, et les villes prioritaires par un carré. Un tel réseau comporte 116 villes, mais le coût de la liaison entre deux villes n'est pas toujours proportionnel à la distance qui les sépare : le franchissement des fleuves, des gorges ou des montagnes coûte très cher.



Martin Grötschel, Clyde Monma et Mechthild Stoer



Les villes 1, 2, 5, 10 et 13 ont une priorité 1.

Réseau téléphonique de l'Est des États-Unis.