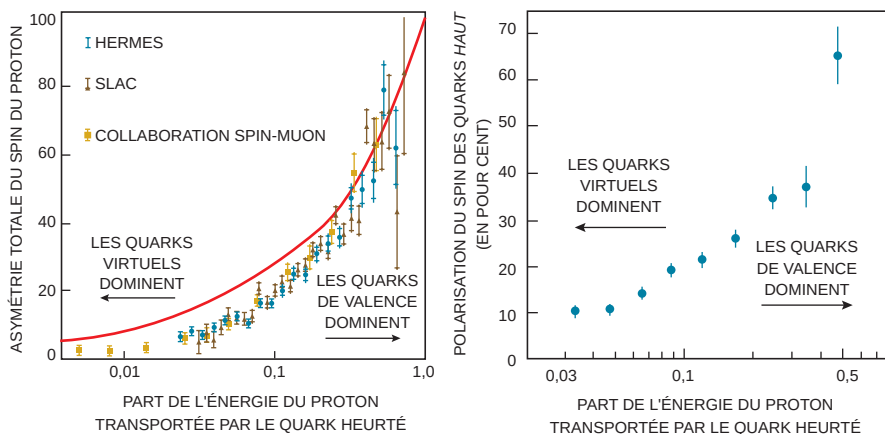




3. LA POLARISATION DU SPIN du proton est mesurée par de nombreuses expériences. L'asymétrie du spin total mesurée par le groupe HERMES au SLAC et par la Collaboration spin-muon (à gauche) montre que le spin des quarks n'apporte qu'une faible contribution au spin du proton. Des données ajustées à la courbe auraient indiqué une contribution plus importante, ce que l'on attendait au début des années 1980. La contribution des seuls quarks et antiquarks haut (à droite) a récemment été mesurée par le groupe HERMES.

neutron, qui devrait ressembler à celle du proton.

En 1988, un groupe international de physiciens, HERMES, proposa d'adapter le collisionneur d'électrons et de protons HERA, à Hambourg, afin de mesurer le spin du nucléon. Les électrons font le tour de l'anneau de HERA environ 47 000 fois par seconde, produisant un faisceau moyen 1 000 fois plus intense que celui du SLAC. Ce faisceau intense est utilisable avec des cibles gazeuses de faible densité, constituées d'hydrogène atomique, de deutérium ou d'hélium 3 purs. De telles cibles ne posent plus le problème de «dilution» du spin qui se produit avec des cibles de butanol et d'ammoniac, où de nombreuses paires de protons et de neutrons ne sont pas polarisées dans les atomes de carbone, d'azote et d'oxygène (par exemple, un seul des 15 nucléons d'un atome d'azote est polarisé).

L'obtention systématique d'un faisceau polarisé sur l'accélérateur HERA nécessita quatre ans de mise au point. Puis on a installé les détecteurs. L'équipe HERMES a commencé l'enregistrement de données à l'été 1995.

Les expériences de Hambourg et du CERN donnent des résultats très analogues (voir la figure 4), confirmant qu'un tiers seulement du spin du proton provient du spin de ses quarks de valence. De plus, on commence à identifier la contribution de chaque type de quark, en étudiant les débris de nucléons produits par les collisions. Ainsi, le groupe HERMES a récemment mesuré la relation entre le mouvement des quarks *haut* à l'intérieur du proton et leur polarisation, et il

devrait effectuer bientôt la première mesure directe des polarisations des quarks *étrange*, en analysant les collisions qui produisent un kaon négatif (particule constituée d'un quark *étrange* et d'un antiquark *haut*).

Parallèlement, les théoriciens ont amélioré les modèles : les 70 pour cent manquants dans le spin ne provoquent plus de crise. Ils semblent provenir des gluons (chaque gluon a un spin égal à un) et du mouvement orbital de tous les quarks et gluons à l'intérieur du nucléon. Mieux encore, la théorie décrit bien les mesures si l'on admet que la contribution totale des gluons est de une ou deux unités de spin et si le moment cinétique orbital est égal à moins un.

La présence de tels spins élémentaires, dans un nucléon dont le spin total est égal à seulement 1/2, était imprévue. Peut-on vérifier indépendamment ces contributions inattendues des gluons et du mouvement orbital? Personne n'a encore proposé de méthode de mesure de la contribution orbitale. Les données du groupe HERMES indiquent que les gluons contribuent effectivement au spin du nucléon, mais l'importance de cette contribution ne peut pas encore être évaluée. Des études utilisant des collisions entre protons à haute énergie devraient aussi permettre la détermination du spin des gluons. De telles expériences commenceront l'année prochaine au Collisionneur d'ions lourds relativistes, en construction à Brookhaven, aux États-Unis. Au CERN, l'expérience COMPASS mesurera aussi la contribution des gluons.

La compréhension de la structure interne des nucléons aura des retombées dans d'autres domaines de la physique des particules. La distribution du spin entre les différents éléments est intimement liée à des structures mathématiques qui apparaissent non seulement en chromodynamique quantique, mais aussi dans l'étude des interactions faibles (qui sont responsables de la désintégration de certains noyaux et participent à la fusion nucléaire dans le Soleil). En particulier, la règle de sommation établie en 1966 par James Bjorken, de l'Université Stanford, relie la dispersion des électrons autour des nucléons polarisés (un phénomène électromagnétique) et la désintégration d'un neutron (un phénomène d'interaction faible). Cette règle relie aussi la structure en quarks et en gluons des nucléons, régie par l'interaction forte, et la durée de vie du neutron. Elle est au cœur de notre description du monde de l'infiniment petit, au point que Richard Feynman déclara en 1972 que : «Sa vérification ou la démonstration de sa violation auraient un impact déterminant sur la physique des hautes énergies.»

Les expériences sur le spin vérifient la règle de sommation de Bjorken avec une précision croissante, et corroborent donc la structure mathématique de base de la chromodynamique quantique et du modèle standard de la physique des particules. Ainsi les scientifiques en apprennent-ils davantage sur les propriétés fondamentales de notre Univers, en cherchant une réponse complète à cette question, simple en apparence : qu'est-ce qui produit le spin d'un nucléon?

Klaus RITH dirige le groupe HERMES à l'Université d'Erlangen-Nuremberg. Andreas SCHÄFER dirige le groupe de théorie des hautes énergies et de chromodynamique quantique de l'Université de Regensburg.

Alan D. KRISCH, *Collisions between Spinning Protons*, in *Scientific American*, vol. 257, n° 2, pp. 42-50, août 1987.

F.E. CLOSE, M. MARTEN et C. SUTTON, *The Particle Explosion*, Oxford University Press, 1987.

Robert L. JAFFE, *Where Does the Proton Really Get Its Spin?*, in *Physics Today*, vol. 48, n° 9, pp. 24-30, septembre 1995.

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

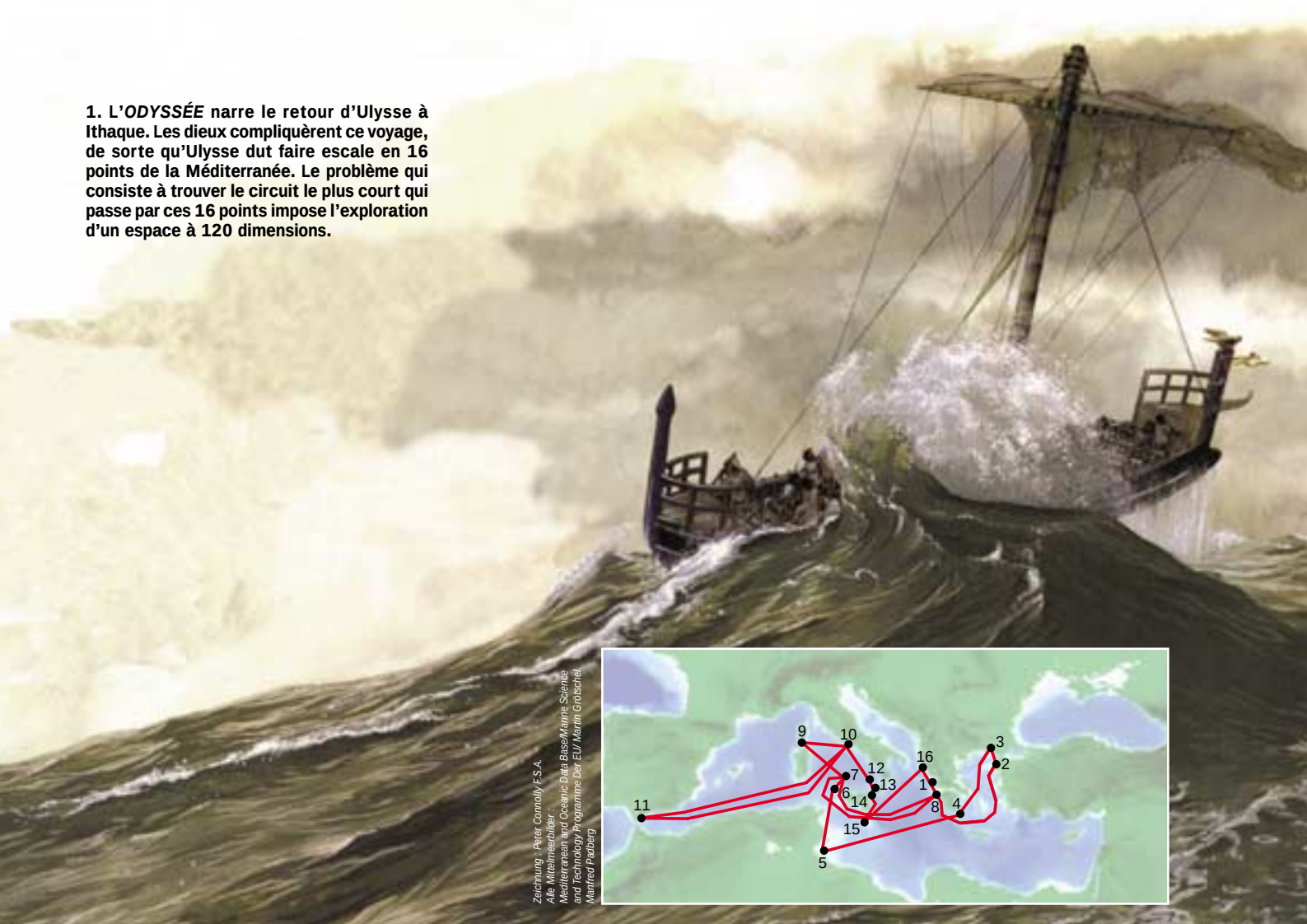
L'analyse d'un problème d'optimisation commence avec sa formalisation mathématique : on remplace chaque ville par un point, ou «sommets», et l'on relie chaque sommet du réseau par

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.

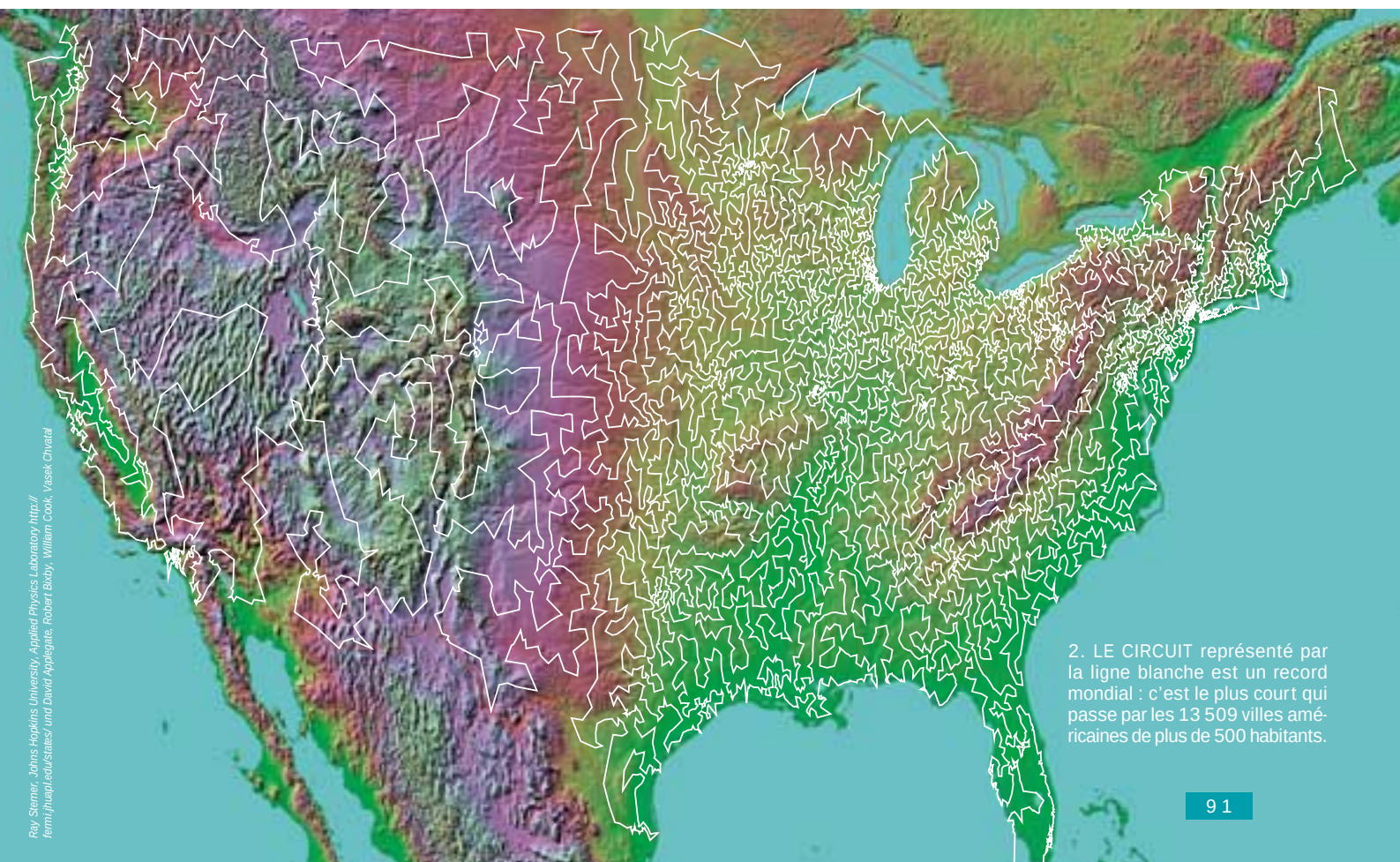
Chacune des arêtes est, dans le problème d'Ulysse, affectée d'une valeur : la distance entre les deux villes associées aux sommets. Toutefois d'autres types de valeurs sont possibles : montant de transactions, durée de fabrication, quantité de matière, etc. Un circuit n'est composé que de certaines arêtes, 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 F.S.A.
 "Alle Mittelmeerfahrer".
 "Mediterranean and Oceanic Data Base Marine Science
 and Technology Programme Der EU/ Martin Grottel,
 Manfred Patberg"



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.

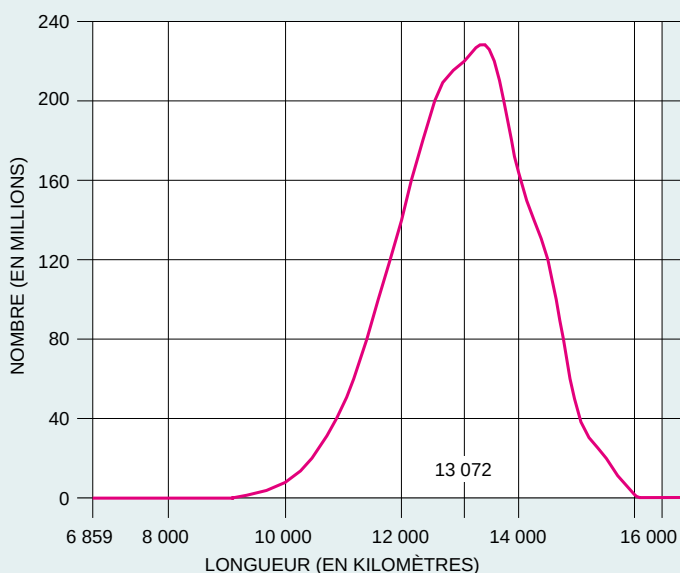
Attribuons, par exemple, des niveaux de priorité aux différentes villes à relier : 0 pour les villes peu importantes, 1 pour des villes plus importantes, 2 pour les grandes villes, etc. On impose que toute ville d'un rang i soit reliée aux autres villes de même rang, même quand i liaisons directes tombent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



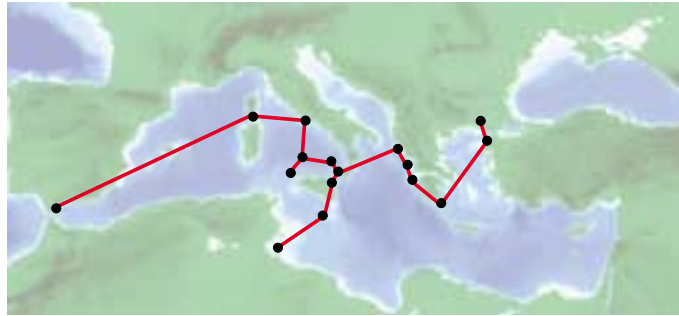
solutions possibles sans simplifier le problème, bien au contraire.

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?

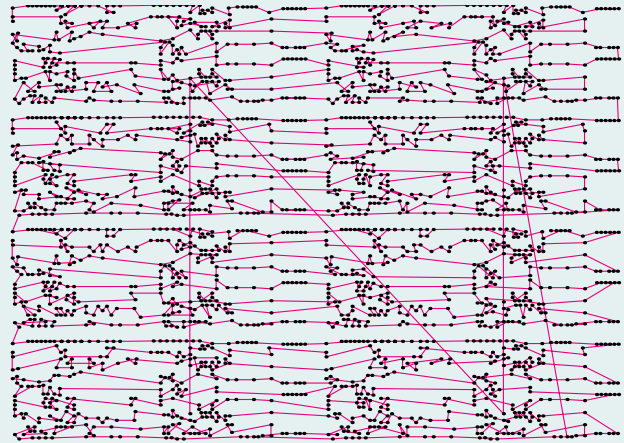
Une bonne stratégie de recherche consiste à utiliser à chaque étape les informations utilisées à l'étape précédente : c'est ainsi que sont nés les progrès des dernières années. Utilisons le voyage d'Ulysse pour examiner cette stratégie.

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

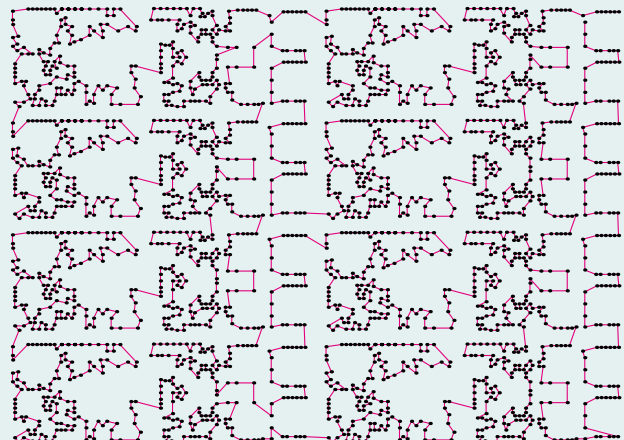


Le problème.



Une solution médiocre.

Supposons que nous utilisons 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.

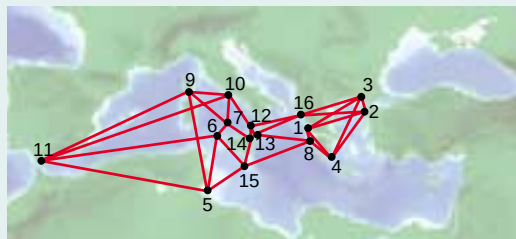
Comment détermine-t-on la borne inférieure de la fonction coût, pour un problème d'optimisation combinatoire ? L'idée fondamentale, apparemment paradoxale, consiste à simplifier le problème en le généralisant : on ajoute au problème initial un autre problème, qui possède davantage de solutions, de sorte 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.

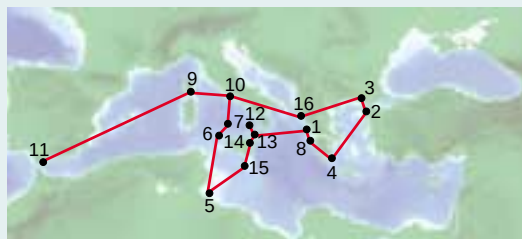
mètres. Si l'on peut dépenser plus pour assurer la fiabilité du réseau, par exemple en assignant la valeur 3 à chacune des villes, le réseau le moins cher que l'on puisse concevoir est de longueur égale à 16 112 kilomètres : avec un tel réseau, les connexions subsisteraient si trois connexions de chaque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

être aussi proche que possible de celle du problème initial. Dans le cas de l'*Odyssée*, une généralisation qui consiste à supposer toutes les distances entre les villes égales à la plus petite des distances entre villes conduit à une borne inférieure.

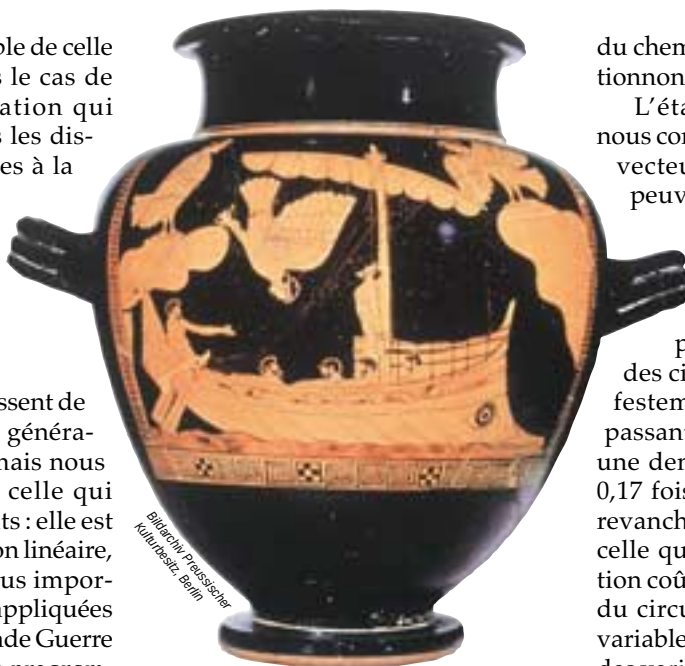
La programmation linéaire

Les mathématiciens connaissent de nombreuses techniques de généralisation (ou d'inclusion), mais nous ne considérerons ici que celle qui donne les meilleurs résultats : elle est fondée sur la programmation linéaire, l'une des techniques les plus importantes des mathématiques appliquées découvertes depuis la Seconde Guerre mondiale (voir *Gestion et programmation linéaire*, par Robert Bland, *Pour la Science*, août 1981). Son nom est trompeur, car elle n'a pas de rapport direct avec l'informatique ; c'est plutôt une optimisation de problèmes linéaires. On l'emploie quand on veut répartir de façon optimale des ressources limitées ; elle résout des problèmes de planification dans les compagnies de transport aérien, dans les raffineries de pétrole, dans les usines automobiles...

Pour inclure le problème d'Ulysse dans un problème que l'on puisse étudier par les techniques de programmation linéaire, on commence par dresser la liste des 120 liaisons directes entre les 16 villes, dans un ordre quelconque, et l'on affecte chaque liaison de la valeur 1 si Ulysse les utilise au cours de son circuit, de la valeur 0 dans le cas contraire. Ainsi chaque circuit est décrit par un ensemble de 120 chiffres binaires, ce que l'on nomme un vecteur dans un espace à 120 dimensions. Autrement dit, un circuit est associé à un point dans un espace à 120 dimensions.

Inversement, tout point de l'espace à 120 dimensions ne correspond pas à un circuit : il faut que 16 des coordonnées du point soient égales à 1, et que les 104 autres soient égales à 0.

Parmi les 653 837 184 000 circuits, comment choisir ceux dont la longueur totale est minimale ? Calculons d'abord leur longueur en multipliant chaque coordonnée (0 ou 1) par la longueur



4. QUAND ON SE CONTENTE de choisir la solution la plus directement séduisante, on ne parvient pas toujours à l'objectif fixé. Le «prévoyant Ulysse», lui, voyait loin : afin d'échapper au chant des sirènes, il demanda à ses compagnons de l'attacher au mât de son navire, tandis que ses compagnons avaient les oreilles bouchées à la cire (le vase qui représente cette scène date de 450 avant notre ère).



5. TEL LE CYCLOPE Polyphème aveuglé par Ulysse, les algorithmes de recherche de circuits n'ont pas de vue d'ensemble du problème, et ils ne progressent qu'à tâtons. Pour des raisons de gain de temps, ils ne vérifient pas, chaque fois, que chaque contrainte est respectée, et ils doivent accepter de faire parfois fausse route. C'est ainsi qu'Ulysse échappa à Polyphème en se cachant sous le ventre d'un bélier (le peintre surnommé le Maître de Sapho a réalisé cette œuvre vers 510 avant notre ère).

du chemin correspondant, puis additionnons tous les produits.

L'étape décisive vient ensuite : nous considérons les composantes du vecteur comme des variables qui peuvent prendre n'importe quelle

valeur comprise entre 0 et 1. On obtient ainsi des points de l'espace à 120 dimensions qui ne correspondent généralement pas à des circuits : Ulysse ne peut manifestement pas effectuer un circuit passant par les 16 villes s'il se rend une demi-fois d'Ithaque à Troie, ou 0,17 fois de Messine à Gibraltar. En revanche, cette généralisation est de celle que nous cherchions. La fonction coût cherchée, la longueur totale du circuit, est une fonction des 120 variables : toute perturbation de l'une des variables modifie la fonction coût d'une valeur proportionnelle à la perturbation, le facteur de proportionnalité étant égal à la longueur de l'étape correspondante.

On connaît un procédé efficace pour chercher le minimum d'un problème linéaire quand on impose des contraintes qui sont également linéaires : c'est l'algorithme du simplexe. Considérons tous les points (ou vecteurs) de l'espace habituel dont les coordonnées sont comprises entre 0 et 1 : ils constituent un cube dont le côté a une longueur égale à 1 et dont un sommet se trouve à l'origine de l'espace. De même, dans un espace à 120 dimensions, les vecteurs dont toutes les composantes sont comprises entre 0 et 1 forment un hypercube, à 2^{120} sommets. Une contrainte exprimée par une inéquation linéaire est comme un couteau qui élimine une partie d'un fromage : on obtient un solide limité par des faces planes.

Dans l'espace à 120 dimensions, les points qui vérifient ainsi les contraintes forment un polytope (un volume limité par des hyperplans, tout comme un polyèdre est limité par des plans) convexe. Le minimum d'une fonction coût linéaire sera toujours dans un coin (ou dans plusieurs coins) du polytope. L'algorithme du simplexe examine systématiquement, mais en restant myope, les coins du polytope jusqu'à ce qu'il soit bloqué : le coin atteint est le minimum cherché (ou l'un des divers minima possibles).

Malheureusement ce minimum n'est que celui du problème généralisé, et non celui du problème initial. En général, le vecteur trouvé par la méthode du simplexe n'a pas de coordonnées entières, et l'on doit remettre en action le couteau à fromage pour transformer le polytope du problème linéaire en polytope du problème initial. Pour ce travail supplémentaire, un coup de couteau est insuffisant. Pis encore, on ne sait pas comment retailler le polytope du problème linéaire. Les programmes linéaires explicites, pour la résolution d'un circuit, sont aujourd'hui limités à 9 villes. Plus précisément, le polytope associé au problème du voyageur de commerce à 9 villes est défini par 9 équations et par 42 104 442 inéquations, avec 36 variables qui correspondent aux 36 possibilités de liaison directe entre 9 sommets. Pour des polytopes avec 8, 7 et 6 villes, on obtient respectivement 194 187, 3 437 et 100 inéquations à 28, 21 et 15 variables. Pour 10 villes, le nombre d'inéquations atteindrait au moins 51 043 900 866 (cela reste à démontrer).

Pour des problèmes encore plus complexes, on ne connaît pas, même théoriquement, le jeu d'équations qui définit le polytope, et, même si on le connaissait, il serait peu utile, car il est beaucoup trop gros. Enfin, pour des villes encore plus nombreuses, on ne pourrait même pas trouver ce jeu d'équations. Le nombre des inéquations pour le problème des 16 villes de l'Odyssée serait, en tout cas, vertigineux.

On peut trouver une inéquation linéaire qui distingue le minimum trouvé de celui du problème étudié. Quand on ajoute cette inéquation au système d'inéquations déjà trouvé, on obtient un nouveau minimum, et, si l'on a de la chance, ce point peut même appartenir au polytope réel : on dispose alors d'une solution minimale du problème initial. Cependant, en général, ce point n'a pas de coordonnées entières, ce qui impose de répéter le processus.

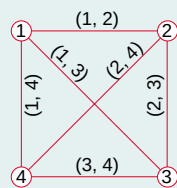
Une autre difficulté s'ajoute : généralement on ne peut pas se fonder sur le polytope des points admissibles du

problème d'optimisation, car on devrait tenir compte de toutes les contraintes, qui seraient également en nombre considérable dès que le nombre de villes est élevé. Paradoxalement on peut toutefois résoudre beaucoup de problèmes plus complexes par la méthode de la programmation linéaire. Comment cela est-il possible ?

Remplaçons à nouveau un problème par un autre, plus vaste, mais plus simple : parmi le nombre considérable de contraintes, ne retenons que certaines d'entre elles (nous découpons le morceau de fromage de manière grossière). Nous identifions d'abord un vecteur solution, c'est-à-dire le coin optimal du trop grand polytope, puis nous l'améliorons en introduisant une contrainte qui permet de couper un morceau aussi grand que possible ; on répète l'opération plusieurs fois.

Dans le cas du problème d'Ulysse, nous pourrions d'abord commencer par utiliser, comme contrainte, que chaque ville ne figure que dans deux

Les circuits et l'optimisation linéaire



On résout parfois des problèmes en les compliquant d'abord. Considérons quatre villes 1, 2, 3, 4. Le graphe complet, pour ces quatre villes, a six arêtes, repérées par les numéros des villes qu'elles relient (le plus petit numéro figure toujours en premier). On peut caractériser chaque circuit en attribuant aux arêtes qui appartiennent au circuit la valeur 1, et la valeur 0 aux autres arêtes :

CÔTÉ	(1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)
	(1,	0,	1,	1,	0,	1)
	(1,	1,	0,	0,	1,	1)
	(0,	1,	1,	1,	1,	0)

Avec ce codage, chaque circuit peut être considéré comme un vecteur d'un espace à six dimensions où deux valeurs sont possibles pour chaque dimension. De même, nous pourrions représenter l'Odyssée par un vecteur de 120 composantes, dont 16 seraient un chiffre 1, et les autres un chiffre 0. Inversement, tous les vecteurs ayant 16 composantes égales à 1 et 104 composantes égales à 0 ne sont pas des circuits. Parmi les 16 arêtes qui relient chaque ville, seulement deux peuvent être affectées de la valeur 1 : le chemin par lequel Ulysse arrive à cette ville, et le chemin par lequel il en sort. Cette condition s'exprime sous la forme d'une équation linéaire : la somme de toutes les composantes du vecteur qui appartiennent aux arêtes correspondant à une ville doit être égale à deux. Un vecteur qui remplit cette condition ne correspond pas

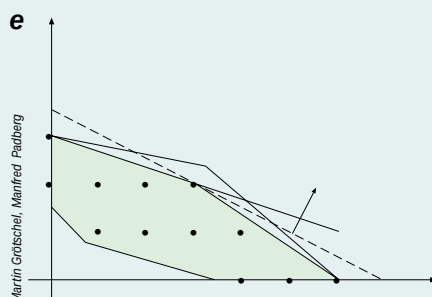
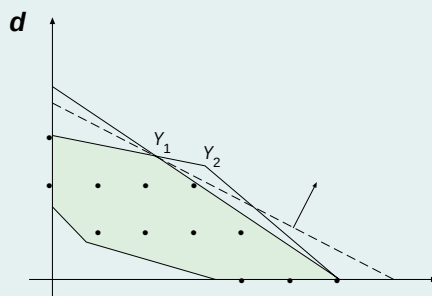
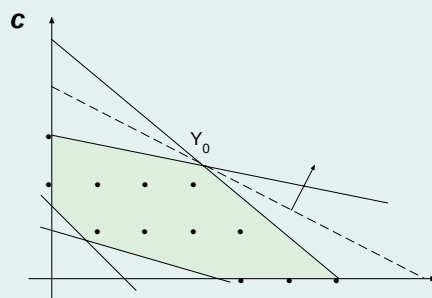
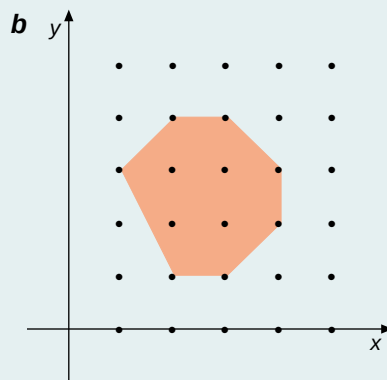
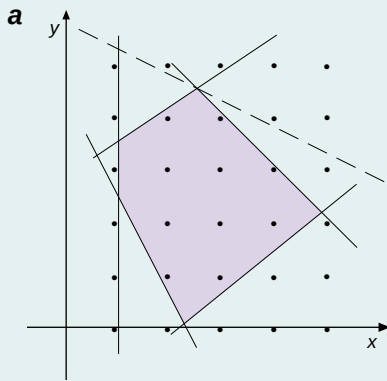


toujours à un circuit : les arêtes choisies pourraient constituer plusieurs circuits partiels (voir la figure en haut à droite) plutôt qu'un grand circuit. D'autres contraintes sont nécessaires.

Si l'on considère les sept villes du circuit partiel le plus à l'Ouest, on observe que toutes les arêtes qui relient ces villes sont affectées de la valeur 1 dans le vecteur. Au moins un des ces chiffres 1 est de trop, car sept arêtes reliant sept villes, chaque ville n'étant associée qu'à deux arêtes, constituent nécessairement un circuit fermé, ce qui exclut un circuit passant par les 16 villes. Plus généralement, on peut formuler cette condition de la façon suivante : pour chaque groupe de k villes, $k-1$ arêtes au plus entre ces villes pourront être affectées de la valeur 1. Cette condition se formalise par une inéquation linéaire : la somme de toutes les composantes associées à un groupe de k villes est nécessairement inférieure ou égale à $k-1$.

Chaque groupe d'au moins trois villes conduit à une telle condition, ce qui, pour n villes, donne $2^n - 1$ conditions (et non 2^n , car la condition d'exclusion pour un groupe est identique à celle du groupe complémentaire). Pour de grands nombres de villes, un tel nombre de conditions est bien au-delà de ce qui est praticable.

L'optimisation linéaire



La découpe des polytopes est un art difficile, car on ne sait pas où la commencer. Pour illustrer cette difficulté, examinons des exemples en deux dimensions (au lieu de trois, voire de 120). À deux dimensions, l'espace est décrit à l'aide de deux variables x et y . Une condition linéaire (exprimée par une inéquation) correspond à une droite dans le plan : tous les points situés d'un côté de la droite sont éliminés. Plusieurs conditions délimitent un polygone (dans le cas général, un polytope) qui contient les solutions acceptables du problème (a). La fonction coût est également linéaire. C'est pourquoi il existe des droites (dans le cas général, des hyperplans) sur lesquelles la fonction coût est constante. La droite en pointillés de la figure a est une telle droite : les points situés au-dessous à gauche sont tous plus chers, et les points situés à droite et au-dessus sont moins chers que ceux qui sont situés sur la droite elle-même. La solution optimale du problème (le point le moins cher du polygone) doit être un des coins de ce polygone : c'est le seul point du polygone qui rencontre la droite. Parfois l'un des côtés du polygone est parallèle à cette droite des coûts constants. En pareil cas, les deux points qui délimitent la droite, ainsi que les points intermédiaires, constituent des solutions optimales de même valeur.

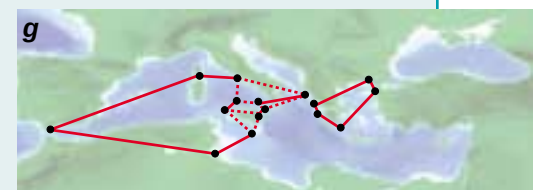
Le problème linéaire qui généralise le problème initial est ici celui qui considère tous les points de l'espace, et non seulement les points de coordonnées entières (points noirs). Dans la recherche du circuit le plus court entre les 16 étapes d'Ulysse, les valeurs entières acceptables se réduisent même à 0 et à 1. Nous ne cherchons pas ici la solution optimale au problème linéaire, car elle n'est généralement pas située sur un tel ensemble de points, mais plutôt la solution optimale parmi les points situés dans la grille de points à l'intérieur du polygone. Il nous faudrait donc rectifier ce polygone de sorte qu'il ne contienne plus que les points acceptables, ainsi que tous les points situés entre eux, délimitant une enveloppe convexe qui englobe les points acceptables (b). Une telle rectification impose l'ajout de conditions supplémentaires. Malheureusement les trouver toutes est difficile, car il peut en exister un nombre gigantesque. On adopte alors une démarche itérative.

Pour identifier un faux optimum (un optimum y_0 du problème linéaire, avec des coordonnées non entières) (c), on cherche une droite (une condition exprimée par une

inéquation) qui exclut le faux optimum tout en laissant acceptables les points qui le sont. Parmi les nombreuses droites qui possèdent cette propriété, il en existe de meilleures, c'est-à-dire des droites qui découpent un gros morceau du polygone. On résout le problème linéaire à l'aide de cette condition supplémentaire (d), on trouve un nouvel optimum y_1 qui se révèle être à nouveau un faux optimum, et l'on répète le processus jusqu'à ce que l'optimum soit situé sur un point de la grille (e).

Cet algorithme de coupe est simple dans un plan, à seulement deux dimensions (déterminer la droite de coupe la plus favorable), mais il devient très difficile à mettre en œuvre dans un espace à 120 dimensions. En 1958, à l'Université Princeton, Ralph Gomory a trouvé un algorithme qui résout le problème après un nombre fini d'étapes, mais qui se révèle trop lent pour les gros problèmes. Aujourd'hui on ne cherche plus systématiquement les hyperplans de coupe, mais on se limite à ceux qui s'adaptent le mieux aux facettes du polytope.

En considérant les conditions secondaires, la solution non admissible au problème du circuit de l'*Odyssée* décrite dans l'encadré de la page 96 reste une solution non admissible, mais qui se rapproche de la solution admissible la plus courte (la solution optimale, f) et s'améliore encore à l'étape suivante, où quelques arêtes (en pointillés) sont affectées de la valeur 1/2 (g). L'étape suivante conduit à la solution optimale (h).



chemins : un chemin qui y conduit, et un chemin qui en sort. La solution du problème linéaire correspondant conduit à une longueur minimale de circuit égale à 6 113 kilomètres, ce qui représente une borne inférieure utilisable. Une telle longueur correspond à un groupe de plusieurs circuits partiels (voir l'encadré de la page 97). Cependant, comme ce groupe n'est pas lui-même un circuit, on doit ajouter une contrainte pour éliminer le point correspondant du polytope obtenu.

Il existe de nombreuses découpes possibles ; la «meilleure» est celle qui se rapproche le plus des faces du polytope réel (ces faces sont des généralisations à plusieurs dimensions des surfaces qui limitent un polyèdre à trois dimensions). Depuis 20 ans environ, de nombreuses équipes ont étudié les découpes. Ici, nous allons ajouter une contrainte qui exclut les quatre circuits partiels de la solution provisoire trouvée. On cherche alors à résoudre le problème linéaire en tenant compte de ces inéquations, une opération où excellent les programmes modernes, qui exploitent de façon astucieuse les informations contenues dans la solution provisoire.

On obtient ainsi un vecteur solution associé à une longueur de 6 228 kilomètres, qui ne permet toujours pas à Ulysse de boucler un circuit complet (voir la figure f de l'encadré de la page 97). Poursuivons en ajoutant des contraintes qui n'avaient pas été prises en compte : on obtient la solution d'un troisième problème linéaire, qui donne une fonction finale de 6 813,5 kilomètres et un vecteur solution qui ne permet toujours pas de boucler un circuit (voir la figure g). En effet, celui-ci contient, pour certaines des 120 variables, la valeur $1/2$. Une itération supplémentaire donne une valeur de 6 859 kilomètres, et le vecteur solution qui correspond au circuit optimal (voir la figure h).

Le procédé itératif précédent est nommé algorithme de coupe. Quand on connaît les inéquations qui décrivent le polytope réel, cet algorithme se termine après un nombre fini d'étapes. En revanche, l'intuition ne nous laissait pas espérer que nous atteindrions la solution en seulement quatre étapes. Aujourd'hui on ignore les inéquations qui conduiraient à coup sûr à la résolution du problème du voyageur de



6. L'OPTIMISATION à l'aide d'arcs et de flèches : quand Ulysse revient chez lui, après 20 ans d'errance, son but n'est pas atteint. En éliminant successivement les diverses conditions restantes (c'est-à-dire en tuant les prétendants), il élimine tous les points non acceptables et atteint l'optimum.

commerce ou à celle d'autres problèmes d'optimisation combinatoire importants en pratique, mais les mathématiciens appliquent leurs résultats à des problèmes combinatoires toujours plus nombreux.

L'existence des inéquations nécessaires à la résolution du problème du voyageur de commerce ou à celle d'autres problèmes d'optimisation combinatoire n'est pas assurée : la complexité de ces problèmes va peut-être de pair avec cette inexistence. Comment faire, alors ? Nous pouvons nous contenter d'une borne inférieure, ou nous pouvons aussi appliquer un algorithme d'énumération : si l'algorithme de coupe se termine sans résoudre le problème, nous savons toutefois que certains côtés du polytope ont des coordonnées qui ne sont

ni 0 ni 1. Nous pouvons alors sélectionner l'un de ces côtés, lui assigner arbitrairement une valeur 1 (indiquant que le circuit doit passer par ce point) et tenter de résoudre alors le problème ainsi limité, et simplifié. Nous pouvons également assigner la valeur 0 (le circuit ne passe pas par ce point) et choisir parmi ces deux solutions celle qui donne le meilleur résultat.

Si l'on ne résout pas non plus ce nouveau problème, on devra effectuer une autre division, et répéter le processus, formant un arbre de solutions possibles, qu'il faudra ensuite explorer astucieusement.

L'analyse d'un tel arbre s'effectue par la méthode de branchement et coupure, une méthode perfectionnée depuis les années 1960. Ainsi les exemples que nous avons considérés ici ont été calculés à l'aide d'un programme mis au point par Manfred Padberg et Giovanni Rinaldi, de l'Institut d'analyse des systèmes informatiques de Rome (IASI). Le temps de résolution pour le problème à 16 villes n'est que de quelques millisecondes sur un ordinateur de bureau moderne, alors que l'énumération prendrait 92 heures.

Pourquoi le problème du voyageur de commerce est-il si compliqué ? Parce que sa résolution impose un très grand nombre d'inéquations. Ce nombre est-il vraiment à l'origine de sa difficulté ? En partie, certainement, car bien des problèmes d'optimisation combinatoire sont faciles à résoudre, alors que leurs polytopes exigent bien plus d'inéquations que le problème du voyageur de commerce.

Martin GRÖTSCHEL est professeur de mathématiques à Berlin et Manfred PADBERG est professeur à l'Université de New York.

E. BRADFORD, *Ulysses Found*, Harcourt, Brace & World, New York, 1963.

E. L. LAWLER et al. (Hrsg.), *The Traveling Salesman Problem. A Guided Tour of Combinatorial Optimization*, Wiley, Chichester, 1985.

Mechthild STOER, *Design of Survivable Networks*, in *Mathematics*, vol. 1531, Springer, Berlin, 1992.

Manfred PADBERG, *Linear Optimization and Extensions*, Springer, Berlin, 1995.

Michael JÜNGER, Gerhard REINELT et Giovanni RINALDI, *The Traveling Salesman Problem*, in *Annotated Bibliographies in Combinatorial Optimization*, Mauro Dell'Amico et al. (Hrsg), Wiley, Chichester, pp. 199-221, 1997.

David APPLEGATE et al., *On the Solution of Traveling Salesman Problems*, in *Documenta Mathematica*, Numéro spécial, *Proceedings of the ICM 98*, pp. 645-656, 1998 ; <http://www.mathematik.uni-bielefeld.de/documenta/xvol-icm/ICM.html>.

Rainer BURKARD et al., *Well-Solvable Special Cases of the Traveling Salesman Problem : A Survey*, in *SIAM Review*, vol. 40, pp. 496-546, 1998.



L'art du pavage

IAN STEWART

Comment couvrir le plan avec un pavé pentagonal.

Les mathématiques et les arts se rencontrent souvent, grâce à la symétrie. La symétrie absolue manque de variété, mais elle peut constituer la base d'œuvres artistiques remarquables : les papiers peints, les tissus, les carrelages, les jardins ou les mosaïques ont souvent une beauté captivante. Pensons aux toiles de Victor Vasarely, aux superbes kimonos du Musée de Tokyo ou à l'Alhambra, à Grenade.

Bien que les bases mathématiques de la symétrie des pavages soient connues depuis longtemps, les artistes continuent d'innover. L'artiste britannique Rosemary Grazebrook, par exemple, a récemment exploré un système de pavage très simple et pratique, différent des pavés rectangulaires usuels.

La définition mathématique de la symétrie est simple, mais subtile : une symétrie d'une figure est une transformation qui laisse la figure globalement inchangée. Par exemple, une rotation de 90 degrés dans le sens des aiguilles d'une montre transforme un carré en un carré identique : cette rotation est une symétrie du carré. Une figure peut avoir plusieurs symétries, dont l'ensemble, muni de la loi de composition des transformations, est une structure mathématique de groupe (la composition est associative, admet un élément neutre, et tout élément possède un inverse).

Les pavages, maintenant, sont de plusieurs types. Celui qui a été le plus étudié est celui où les pavés forment un réseau régulier à deux dimensions. Curieusement les mathématiques des pavages furent d'abord explorées dans le cas difficile de l'espace à trois dimensions, celui des cristaux de sel, de sucre ou de tartre, et c'est seulement plus tard que les mathématiciens ont étudié l'espace à deux dimensions. En 1891, le cristallographe

russe E. Fedorov démontra que, pour les pavages du plan, il existe 17 groupes de symétries (voir la figure 2).

Ce résultat s'applique bien sûr aux papiers peints ou aux tissus, mais pourquoi, alors, les magasins d'ameublement proposent-ils tant d'échantillons différents ? Pourquoi existe-t-il tant de types de carreaux de salle de bain que les choix conjugaux deviennent souvent des scènes de ménage ? Dix-sept ne suffisent-ils pas ? Dans la plupart des cas, les différences tiennent à des caractéristiques telles que la couleur, la texture ou la nature des éléments de motifs ; la symétrie, elle, reste limitée. Par exemple, un carrelage de salle de bain composé de carreaux carrés ornés de l'image d'un canard a les mêmes symétries qu'un carrelage composé de carreaux carrés qui porteraient l'image d'une algue (à moins que le motif dessiné n'ait lui-même des symétries, qui s'ajoutent alors à celles des carreaux).

Nous n'examinerons pas ici certains pavages qui ont peu de symétries, tels les remarquables inventions du mathématicien Roger Penrose : ces pavages recouvrent entièrement le plan, mais ils ne sont pas périodiques. Les pavages que nous considérerons ici sont fondés sur une unité de base, une figure qui est répétée indéfiniment dans deux directions distinctes. Imaginons par exemple un carrelage composé de carreaux carrés tels qu'on en trouve dans tant de salles de bain. Sur un mur de salle de bain infini, le carreau carré se répète indéfiniment. Choisissons un carreau du carrelage : si l'on se déplace d'un nombre entier de côtés, verticalement ou horizontalement, on retombe sur un autre carreau identique au premier ; de même, on retrouve un carreau si l'on se déplace dans des directions intermédiaires, à condition que les déplacements verticaux

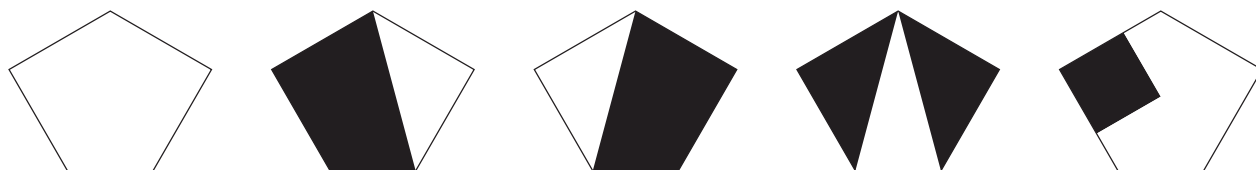
ou horizontaux soient égaux à un nombre entier de côtés.

Les droites verticales et horizontales qui joignent les centres des carreaux forment ce que l'on nomme un réseau. Les papiers peints ou les tissus ont une symétrie de réseau, car ils sont généralement réalisés à partir de grands rouleaux de papier ou de tissu, le long desquels une machine à tambour imprime un dessin. Lorsque le papier est collé sur un mur, ou quand des morceaux de tissus sont cousus ensemble, on raccorde généralement les bandes à leurs bords. On peut aussi décaler légèrement les bandes latéralement ou verticalement ; le réseau n'est alors plus carré.

Cette structure de réseau ne s'impose pas pour les pavages, dont les pavés sont des unités indépendantes, mais elle reste pratique. Le réseau carré, par exemple, possède des symétries de réflexion autour de droites verticales, horizontales ou de direction diagonale et passant par le centre d'une maille ou par un sommet du réseau, ainsi que de rotations d'angle égal à 90 degrés. Un nid d'abeilles, dont les cellules sont des hexagones réguliers, est aussi un réseau, mais avec des symétries différentes, telles des rotations de 60 degrés.

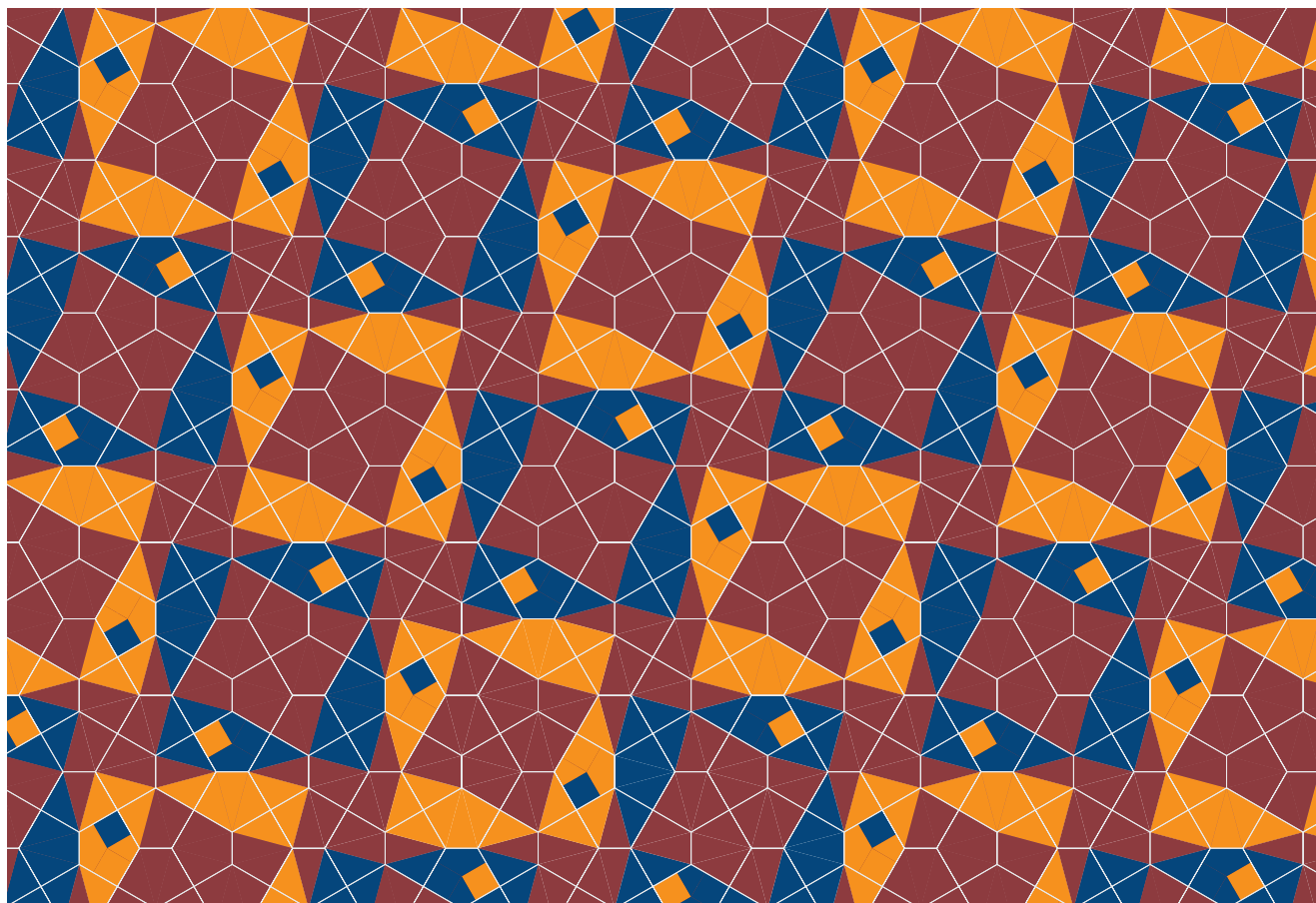
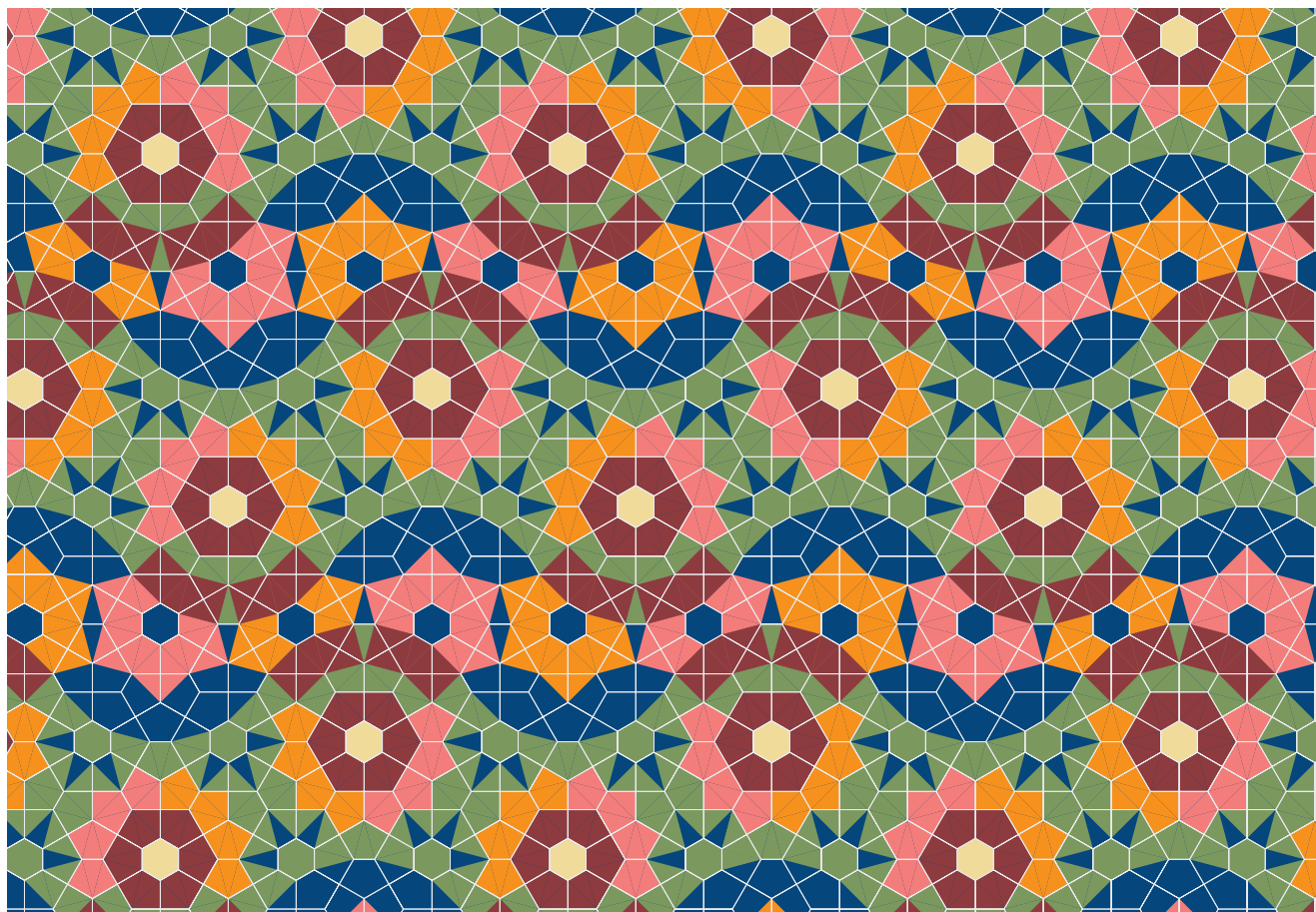
R. Grazebrook a découvert que certains pavés pentagonaux peuvent engendrer une foule de réseaux différents. Les pavés qu'elle utilise ont deux angles égaux à 90 degrés et trois angles égaux à 120 degrés, de sorte qu'on peut les associer par quatre ou par trois (voir la figure 1).

On peut juxtaposer quatre des pavés pentagonaux de R. Grazebrook pour former un hexagone plus large que haut, avec lequel on peut paver le plan, et quand on utilise à la fois les pentagones



Des pavés pentagonaux colorés comme ci-dessus peuvent former un réseau quand on les associe à des hexagones réguliers (ci-contre en haut) ou quand on les juxtapose (ci-contre en bas).

Brian Christie



Brian Christie avec l'autorisation de Rosemary Graebrook



Algorithmes et Optimisation

Que vous soyez...

... affairiste notoire, politicien corrompu, sportif dopé, escroc rutilant...
ou tout simplement prudent,

Cryptez vos données sensibles!

Belzebuth et la Société de Calcul Mathématique, SA, présentent

Sieve II le diabolique

La référence ultime en matière de cryptographie logicielle

• **simple d'installation** : Windows 95, 98, NT,

• **simple d'emploi** : mot de passe de 4 caractères,

• **inviolable** : Sieve II le diabolique utilise un procédé nouveau, le cryptage dynamique récursif, conçu par la SCM. Un prix de 10 000F est offert depuis deux ans à la première personne qui parviendra à décrypter le texte (écrit en français banal) qui accompagne le logiciel.

• **inconditionnellement garanti à vie**. Sieve II le diabolique crypte absolument tout : données (tout format), sons, images, exécutables, etc.

Livraison en recommandé, sous pli discret.

Prix de vente unitaire HT : 1 500F,
TTC : 1 809F.

(Pour quantités importantes, nous consulter)

BON DE COMMANDE

Je commande exemplaires du logiciel de cryptage Sieve II le diabolique, au prix de 1 809 F l'unité (1 500F HT).

NOM :

PRÉNOM :

ADRESSE DE LIVRAISON :

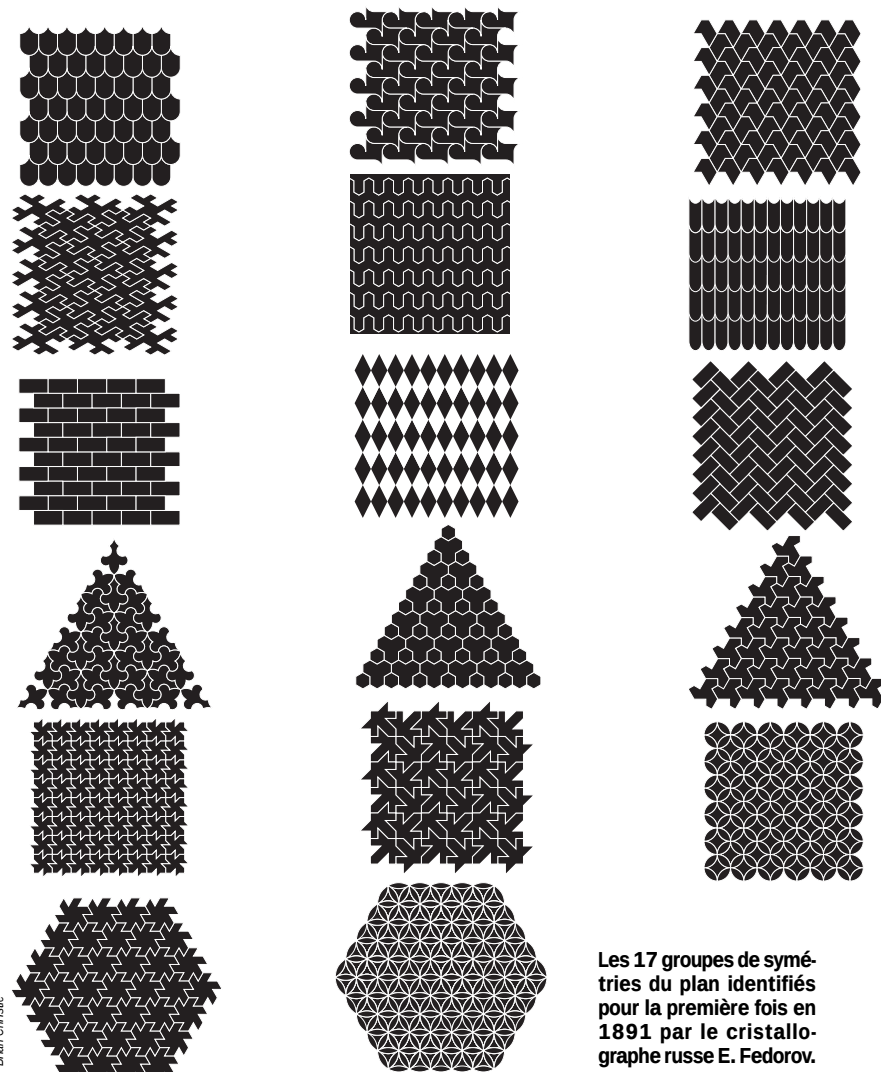
Bon à retourner, accompagné du règlement (chèque à l'ordre de SCM) à : Société de Calcul Mathématique, SA, 111, Faubourg Saint Honoré, 75 008 Paris.

de R. Grazebrook et des hexagones réguliers, on peut obtenir les 17 classes de symétries sauf une (Laquelle? Et comment obtient-on les autres? Je vous laisse le chercher).

R. Grazebrook doit l'idée initiale de ses pavages à Martin Gardner, qui fit les articles de la rubrique «Jeux mathématiques» de *Pour la Science* pendant des années. Elle étudiait l'art de l'Alhambra quand un article de Martin Gardner lui révéla les relations entre les motifs islamiques et les pavages du graveur Maurits Escher. Elle introduisit alors deux procédés distincts pour colorier ses pavés pentagonaux : soit en divisant les pavés en trois triangles (la série dite «Pentland»); soit en les divisant en quatre, (deux carés, un quadrilatère en forme de cerf-volant et un pentagone plus petit, de la série «Penthouse»). On peut naturellement colorier les pavés pentagonaux de R. Grazebrook d'autres manières, mais ces deux séries seulement engendrent une surprenante variété de motifs.

Réactions

L'article sur les lancers de pièces et les dés (*Abrogeons la loi des moyennes*, juin 1998, p. 122) a intéressé Tom Guldbrandsen, de Lyngby, au Danemark. Supposons que nous lancions un grand nombre de fois un dé, et que nous notions le nombre de 1, de 2, ..., de 6 obtenus. Quelle est la probabilité que les six lancers soient en nombreux égaux, quand vient le n -ième lancer? T. Guldbrandsen a d'abord observé que l'égalité ne peut se produire qu'après 6, 12, 18... lancers : le nombre de lancers doit être un multiple de 6. Il a trouvé une formule qui donne, en fonction de n , la probabilité que les six totaux soient égaux. Comme cette égalité peut être vérifiée plusieurs fois, il a obtenu comme probabilité : 0,021903735824.



Les 17 groupes de symétries du plan identifiés pour la première fois en 1891 par le cristallographe russe E. Fedorov.

Brigitte



La fin de la lignée humaine?

JACQUES NINIO

Les mutations délétères semblent si nombreuses que la lignée humaine devrait être éteinte depuis longtemps. Où est l'erreur?

La science-fiction prête aux mutants des capacités extraordinaires. Pourtant, nous sommes tous des mutants, mais dans quelle proportion? Lorsqu'un enfant naît, il hérite d'un mélange de gènes de son père et de sa mère. En raison des erreurs de reproduction, quelques gènes de l'enfant diffèrent légèrement de ceux de ses parents. Ces erreurs, les mutations, sont des modifications localisées des séquences d'ADN, la molécule en double hélice qui constitue les gènes. Certaines mutations, nocives, finissent par être éliminées. D'autres sont bénéfiques et permettent une meilleure adaptation de l'espèce (voir la figure 1). Le plus souvent, elles sont sans effet.

Comment évaluer la fréquence des mutations? Une première méthode consiste à recenser les individus porteurs d'une anomalie génétique absente chez les parents. Les statistiques portent sur les anomalies les plus fréquentes : l'achondroplasie, l'hémophilie A ou la dystrophie musculaire de Duchenne. Pour ces anomalies, le taux de mutations est compris entre un et 5×10^{-5} par gène et par génération. Toutefois, on ne peut pas généraliser ces fréquences à l'ensemble des mutations du génome, car les anomalies génétiques demeurent anecdotiques : elles concernent peu d'individus et des portions très courtes de certains gènes (une dizaine de nucléotides sur le millier que comporte le gène), de sorte que l'incertitude statistique est élevée.

Une évaluation plus sûre du taux de mutation est fondée sur le séquençage systématique d'un même gène chez un grand nombre d'individus et chez leurs parents. On recense chez les enfants les mutations dans un gène, puis on extrapole à l'ensemble du génome (voir la figure 2). Selon ces études, les biologistes estiment le taux M de mutations produites d'une génération sur l'autre à 10^{-8} par nucléotide et par génération, soit soixante mutations au total par

individu puisque le matériel génétique d'une cellule humaine comporte 23 paires de chromosomes, soit six milliards de nucléotides.

MUTATIONS TRANSITOIRES ET MUTATIONS FIXÉES

Les biologistes des populations évaluent différemment le taux de mutation. D'une part, la datation du plus jeune fossile connu commun à deux lignées indique quand ces lignées, deux espèces de singes par exemple, se sont séparées. D'autre part, on compare les séquences d'ADN actuelles de ces deux espèces. On déduit de ces mesures le taux moyen de mutations effectivement fixées au cours de l'évolution, qui se sont répandues dans une espèce. Ainsi, chez l'homme, en moyenne, deux mutations sont fixées par milliard de nucléotides et par an, ce qui, si l'on suppose

qu'une génération dure 20 ans, correspond à un taux extrapolé E_x de 4×10^{-8} mutations fixées par nucléotide et par génération.

De ce taux de mutations fixées dans l'espèce, comment déduit-on M , le taux qui nous intéresse, c'est-à-dire le taux de mutations produites «à la source» chez les individus? Le raisonnement nécessite l'examen du taux F de mutations fixées par un mécanisme purement aléatoire, en négligeant les effets de la sélection naturelle.

D'abord, on admet que, sur le long terme, le taux E de mutations effectivement fixées ne peut être supérieur au taux F de mutations fixées sans sélection naturelle : l'évolution trie et élimine, mais n'accroît pas la variabilité.

Chez les bactéries, la variabilité est faible, et la durée entre générations est courte. La sélection agit de manière féroce en éliminant toutes les mutations qui produiraient un désavantage, fût-il minime (E est bien inférieur à F). Chez l'homme, la variabilité est relativement grande, la durée entre les générations est longue, et quand au bout de quelques siècles ou quelques millénaires, la sélection a retenu une mutation avantageuse dans un gène, dix autres se sont fixées dans le même gène, mutations qui sont neutres ou légèrement délétères : la sélection naturelle n'a pas beaucoup joué et le taux de mutations effectivement fixées au cours de l'évolution est proche du taux sans sélection naturelle (E est proche de F).

Par ailleurs, les biologistes des populations supposent que le taux F des mutations qui se seraient fixées sans sélection est égal au taux M de mutations produites. C'est un résultat intouchable de la théorie de l'évolution, au cœur de la génétique des populations actuelle. Cette égalité théorique relie les données mesurées sur les espèces (comparaison des séquences) aux données sur la production des



1. Les mutations, dues à des modifications de l'ADN, sont la source de diversité entre les espèces, par exemple entre le premier mouton domestique d'il y a 8 000 ans, dont le mouflon de Corse (en haut) est le descendant direct, et le mouton à laine mérinos (en bas), race apparue au Moyen Âge.