

LE WAP • STATISTIQUES ET DÉCISIONS • LES LASERS • LES AMÉRINDIENS



■ POUR LA

SCIENCE

NOVEMBRE 2000

édition française de
**SCIENTIFIC
AMERICAN**

www.pourlascience.com



**Robots et
chirurgie**

**La route
des émeraudes**

Canada : \$ 8,95 / Belgique : 27 / H-B / Suisse : 11,5

M 2687 277 28 00 1



BLOC-NOTES

de Didier Noin

TRIBUNE DE LECTEURS

POINT DE VUE

Nouvelle université

par Yves Michaud

SCIENCE ET ÉCONOMIE

Un marché incomplet

par Ivar Ekeland

SCIENCE ET ASTRONOMIE

En culotte courte

par Hervé Imis

PRÉSENCE DE L'HISTOIRE

Le magnétisme grenoblois

par Jacques Baudry

PERSPECTIVES SCIENTIFIQUES

Nouveaux cousins de l'archéoptéryx

par Fabrice Lhuissier et Christophe Legendre

- Un an après... tempête d'insectes ravageurs
- Un macaque géomètre
- La sélection robotique
- Déformation professionnelle
- Vache folle : révision à la baisse des risques
- Le terrier...

LOGIQUE ET CALCUL

Mathématiques et philosophie

par Jean-Paul Delahaye

IDÉES DE PHYSIQUE

Le droit au mélange

par Bernard Cabane et Sylvie Henon

ART ET SCIENCE

La sphère et les cornes

par Adrien Douady et Eric Maudon

ANALYSES DE LIVRES

■ Un chimiste au passé simple, de Jean-Jacques

■ Coléoptères phytophages d'Europe, de G. du Châtelet

■ Biologie et physiologie animales, de B. Malpas et M. Gelsinger

■ Physiologie animale, sous la direction de Roger Eckert

■ L'invention des populations : biologie, idéologie et politique, sous

la direction d'Hervé Le Bras

Certains d'entre vous ont reçu leur numéro de Pour la Science d'octobre avec un retard très important. Nous vous prions de bien vouloir nous en excuser. Ce retard, indépendant de la volonté de nos éditeurs, est dû aux porteurs qui a perturbé l'expédition des revues à nos abonnés.

Deux encarts d'abonnement entre les pages 18 et 19, un encart broché service lecteurs et une carte d'abonnement entre les pages 98 et 99. Cédérom Infonie collé page 51.

Chaque mois, retrouvez le sommaire complet de la revue en ligne avec pour chaque article une bibliographie et un complément d'information.

www.pourlascience.com

Le WAP

Votre futur compagnon numérique 28

par Fiona Harvey

Pour créer un réseau Internet sans fil, des ingénieurs construisent des réseaux qui traitent rapidement des quantités considérables de données et des dispositifs de poche qui exploitent les ressources du réseau Internet actuel.



Les promesses et les risques du WAP 34

par Karen Bannan

Pressenti comme un standard de communication entre les téléphones cellulaires et le réseau Internet, le système WAP a déjà des concurrents.



Le saut de la troisième génération 40

par Leander Kahney

Le débit des transmissions de données augmente avec les générations de téléphones mobiles. Le déploiement des réseaux de troisième génération implique de grandes avancées techniques. Oui, mais les quelles ?



À cœur battant 46

par Cornelius Borst

Les chirurgiens opèrent aujourd'hui le cœur, mais ils n'ont plus à l'arrêter.



Comment rationaliser les décisions 52

par John Swets, Robyn Dawes et John Monahan

Des outils mathématiques d'aide à la prise de décisions améliorent les diagnostics en médecine ou dans l'industrie.



La route des émeraudes anciennes 58

par Gaston Giuliani, Michèle Heuzé et Marc Chaussidon

Grâce à une sonde ionique, on détermine aujourd'hui la provenance des émeraudes sans les détériorer.



M. Vardoulakis

Une sélection naturelle de la culture 66

par Susan Blackmore

Les comportements et les idées copiés par imitation de personne à personne – les mêmes – semblent avoir contraint les gènes humains à faire de nous ce que nous sommes aujourd'hui.



Les premiers Américains 76

par Sasha Nemecek

On découvre en Amérique des outils fabriqués par des peuples indigènes vieux de plus de 14 000 ans. Ces découvertes révèlent le mode de vie des premiers Amérindiens.



L'origine des Amérindiens 82

par Sergio Pena et Fabrício Santos

Les études génétiques du chromosome Y montrent que les premiers Amérindiens sont venus de Sibérie centrale.



Lasers à impulsions ultracourtes 86

par John-Mark Hopkins et Wilson Sibbett

Les lasers dont les impulsions ne durent que quelques milliardièmes de nanoseconde se sont imposés dans plusieurs domaines scientifiques et techniques.



Les préciosités révélées

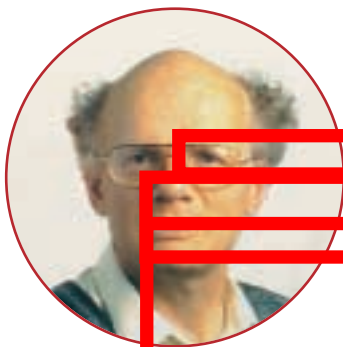
L'histoire tente de répondre à des questions simples : qui, où, quand, comment ? Ainsi les pierres précieuses, qui ornent depuis l'Antiquité les couronnes et les bijoux des familles régnantes et des riches marchands, sont sources d'interrogations : d'où viennent-elles, quand ont-elles été trouvées, comment les orfèvres les ont-ils travaillées, comment les émeraudes s'échangeaient-elles ?

La recherche historique a longtemps été confinée à l'interprétation des textes et à l'examen du style des œuvres, à des fins de datation et d'identification géographique. Il en est tout autrement aujourd'hui grâce à l'analyse physico-chimique fondée sur la diversité géologique. Les caractéristiques accessoires acquises lors de la formation des gemmes, leurs impuretés et leurs inclusions, permettent d'identifier leur provenance et de suivre leur itinéraire historique. Après avoir reconstitué les climats passés, les mouvements des roches, l'histoire de la Terre depuis sa formation, l'analyse isotopique retrace aujourd'hui les coûts de l'époque et les techniques des artisans. Tout cela grâce à des «imperfections» qui constituent l'intérêt et le principal (voir *La route des émeraudes anciennes*, pages 58 à 65). Et les analyses délicates portent sur des quantités si minimes que l'examen n'endommage en rien l'objet d'étude.

Qu'apporte la philosophie aux sciences ? Certainement un questionnement scientifique qui dépasse les seules considérations morales que les philosophes à la mode présentent aujourd'hui avec des platitudes enrubannées. La philosophie a posé les bonnes questions, les sciences, notamment la logique mathématique, ont apporté des réponses. Sur l'infini, par exemple, les mathématiques ont enrichi et renouvelé le discours philosophique ; la philosophie classique, pense le logicien, celle de Kant par exemple, est trop contrainte par sa volonté de se limiter à une vision «interne» et de bon sens. Les sciences n'ont pas ces œillères, la réalité imposant un élargissement des concepts.

La logique et la théorie de la complexité apportent des résultats troublants : les démonstrations utiles sont les plus simples ! Cela justifie-t-il la paresse du cancre ? Vous devinez la réponse, mais vous pourrez lire les détails de l'argumentation dans l'article de Jean-Paul Delahaye, *Mathématiques et Philosophie*, pages 100 à 105.

Philippe BOULANGER



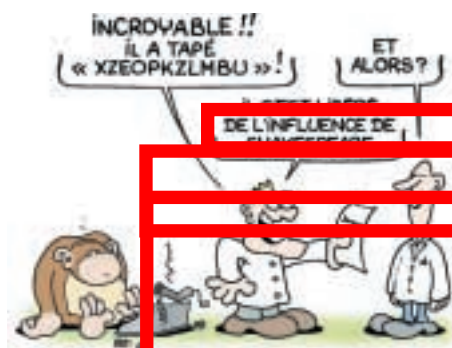
BLOC-NOTES

DIDIER NORDON

L'âne, le singe et la tortue

En science, comme chez La Fontaine, l'animal est une source de métaphores spectaculaires. Il y a toute une ménagerie. Le chat de Schrödinger, le serpent de Kekulé, le limaçon de Pascal, l'âne de Buridan, la tortue de Zénon d'Élée... Certaines métaphores sont dépourvues de nom d'inventeur. Ainsi, l'effet papillon, qui remonte à une conférence prononcée en 1972 par le météorologue Edward Lorenz. L'origine d'une autre métaphore zoologique anonyme, plus ancienne que cette dernière, commence à se perdre, celle du singe dactylographe : laissé seul avec une machine à écrire sur des touches de laquelle il frappe au hasard, jadis un singe ne tapera *Hamlet*. On peut tenir cet événement pour impossible, même si sa probabilité est, en toute rigueur, non nulle.

Selon Pierre Schramm, le singe dactylographe apparaît en 1914 dans un livre sur le hasard dû au mathématicien et homme politique Émile Borel (1871-1956). Laissons aux érudits le soin de dire si la paternité de l'image revient à Borel ou s'il a eu des prédécesseurs. L'essentiel, quant à nous, est d'avoir évoqué Borel en cette solennelle année 2000, laquelle marque – événement rare et même unique – le 129^e anniversaire de sa naissance.



Le devoir de mémoire

La prolifération des articles scientifiques pose d'énormes problèmes. Le plus grave est l'impossibilité pour un chercheur de se tenir au courant de ce qui se passe dans sa propre discipline.

Une mesure simple ferait cesser cette situation préjudiciable : que nul n'ait le droit de publier un article de physique s'il n'a pas lu et compris tous les articles de physique déjà publiés. Tous, car il en va de l'unité de sa discipline. Exigence analogue en chimie, en biologie, en mathématiques... Vu le stock impression-

nant d'articles actuellement disponibles, les chercheurs n'auront pour au moins dix ans à ne rien faire d'autre que lire. C'est un long délai, mais cela leur apprendra à ne pas avoir étudié leurs leçons. Et après ça, ils retrouveront le droit de publier. Mais la même condition restera exigée d'eux : avoir lu tous les articles de leur spécialité avant d'en ajouter un nouveau. Cela les contraindra à un rythme de publication enfin raisonnable. On pourra recommencer à penser les sciences comme étant en progrès et non comme étant une explosion, un éparpillement incontrôlé.

Tous les remèdes ont des effets secondaires imprévus et néfastes, sauf celui préconisé ici, car à coup sûr il ne sera jamais appliqué...



A qui la science profite-t-elle ?

Pour croire et proclamé que la science était sur le point de tout connaître, les savants scientifiques de la fin du XIX^e siècle (Berthelot, Kelvin...) se sont aujourd'hui la moquerie. L'histoire a montré combien leur position était naïve en même temps qu'arrogante. Champions de la modestie, nous estimons quant à nous que jamais notre savoir ne viendra à bout des mystères du monde. Et nous nous en félicitons : il restera toujours quelque chose à découvrir.

Notre position n'est ni plus ni moins fondée que celle des scientifiques. On ne peut pas plus prouver qu'il restera toujours de l'inconnu qu'on ne peut prouver le contraire. Décider de cette question est affaire non de science mais d'idéologie. Or, par un aspect au moins, la position actuelle est moins saine que celle du XIX^e siècle. À force de nous réjouir d'avoir toujours quelque chose à découvrir, nous



en arrivons à adopter une conception pour ainsi dire frivole de la science. Exemple. Lorsque des mathématiciens commentent une démonstration récente ils la jugent bonne si elle ouvre de nouveaux champs de recherches, mais stérile si elle clôt un problème sans susciter de nouvelles questions. Stérile, oui, même si le problème qu'elle résout était fameux et avait fait « sécher » longuement la profession. Qu'est-ce à dire, sinon que le fait qu'il occupe les chercheurs?

Dans le monde du XIX^e siècle, les savants du XIX^e siècle étaient plus modestes. Ils savaient que la science, contrairement à ce qu'on en dit, n'est pas mystérieuse, c'est accorder à cette dernière un statut autrement plus élevé que d'en faire un divertissement pour chercheurs.

Le lion superbe et paresseux

Devinette. Un lion marche dans la savane. Par hasard, son chemin vient à croiser celui d'une antilope. Que se passe-t-il ?

Réponse. Rien.

Ce que je dis là n'est pas pour avoir assisté à la scène, mais pour en croire un livre d'Yves Christen, *Le peuple léopard* (Michalon, 2000). En effet, l'antilope court plus vite que le lion. Le lion sait que, s'il attaque, il perdra son temps sans aucune chance d'attraper sa proie. Et l'antilope sait que le lion le sait. Elle se contente d'émettre un cri d'alarme. Le cas échéant, elle continuera de brouter sous les yeux du lion. Renoncer à son repas serait stupide et pénible – affamé, on court moins vite – alors qu'un lion repéré n'est pas un danger. Elle sait jusqu'où elle peut le laisser approcher. Un fauve ne peut débusquer une antilope bien fraîche que s'il la surprend. Selon Yves Christen, les calculs et les supputations des intentions de l'autre exigés par la chasse – de la part du prédateur comme de la part de la proie – sont une preuve éclatante de l'intelligence animale.



TRIBUNE DES LECTEURS

Vie le sucre sucré!

Didier Nodon (voir *Pour la Science*, mai 2000) discute la pertinence de la proposition «même dépourvu de saveur sucrée, le sucre restant un ingrédient de base en industrie alimentaire» (Revue Sucre info service, décembre 1999), qui, j'en conviens (et cela d'autant plus aisément que j'en suis l'insigne auteur!), frise l'aphorisme. Mais depuis quelques années, la réalité de la science a dépassé la fiction. Le vertige de cette affirmation n'est certes pas métaphysique, mais il est certainement psychophysique.

En effet, les recherches en chimioréception de la saveur sucrée ont fort logiquement recours à des composés sucrés, dont le saccharose, mais elles font aussi appel à des inhibiteurs plus ou moins spécifiques de cette saveur dans le but de modéliser les relations du type «structure/activité».

Parmi ces inhibiteurs, le sel de sodium de l'acide para-méthoxy-propanoïque, substance isolée à l'origine des graines de café vert, a fait l'objet d'un développement commercial et de brevets d'application – en association avec le sucre – dans différents produits alimentaires.

Pour ces applications, l'objectif est de maintenir une concentration de sucre importante à des fins technologiques (conservation par diminution de l'activité de l'eau, effet sur la texture, apport de masse, etc.) ou nutritionnelles (substituer des matières grasses, densité énergétique élevée pour aliments de l'effort) tout en réduisant le sucre perçu à un niveau acceptable.

Dans ce cas précis, le sucre «peru» apparemment son rôle de sucre sucré, mais sa technologie (réactions de Maillard comprises!). En d'autres termes, nous perdons un sens commun, l'efficacité même, à cautionner (un peu) le sien.

Pour finir, je vous rejoins sans retenue pour clamer à voix haute : vive le sucre sucré!

Philippe REISER, CEDUS, Paris

Darwin dans l'Histoire

L'influence de Darwin sur la science moderne (voir *Pour la Science*, septembre 2000) est indéniable. On ne doit pas pour autant oublier l'œuvre de ses prédécesseurs, car lui aussi en eut.

Ernst Mayr, à propos de la Théorie de l'évolution, fait tout commencer avec Darwin : il en a «fondé», «créé», «introduit» toutes les composantes. Toutefois, dès la fin du XVIII^e siècle et au début du XIX^e, les idées bouillonnent et l'on débat sur l'histoire de la Terre et l'origine de la Vie.

S'il était vrai que jusqu'à Darwin, tout le monde croyait à la stabilité des espèces, on peut se demander pourquoi Cuvier se donnait déjà tant de mal, dès 1830 et même dès 1800, pour combattre l'idée de la transformation des espèces, que E. Mayr considère comme la conception moderne de l'évolution et comme un des principaux «apports» de Darwin. Cuvier n'avait pas attendu Darwin (et pour cause, il était mort avant!) pour juger que cette idée était peut-être «la plus superficielle et la plus vaine de toutes celles que

nous avons jusqu'à présent» que «parmi les divers systèmes sur l'origine des êtres organisés, il n'en est pas de moins vraisemblable que celui qui en fait naître successivement les différents genres par des développements ou des métaphores graduelles», et que le soutenir aboutirait à «réduire à rien toute l'histoire naturelle, puisque son objet ne consisterait qu'en des formes variables et des types fugaces».

S'il était vrai que pour Lamarck, l'évolution a un but, pourquoi a-t-il enseigné le contraire? «C'est une véritable erreur que d'attribuer un but, une intention à la nature dans ses opérations. Elle n'en saurait avoir, puisqu'elle n'est point une intelligence, un être pensant».

S'il était vrai que Lamarck se figurait l'histoire de la vie comme une «évolution linéaire» pourquoi l'a-t-il représentée par un «tableau» (phylogénétique) «de la formation des animaux» avec de nombreuses ramifications?

S'il était vrai que le concept de l'hérédité des caractères acquis est uniquement lamarckien, on peut se demander ce qui a pu pousser Darwin à proclamer que personne n'avait rien dit avant lui en faveur de l'hérédité des caractères acquis par l'usage et le non-usage.

S'il était vrai que Darwin avait inventé de la théorie de l'évolution, comment expliquer une grande partie des naturalistes de l'époque à ce comme à son ami Lyell, vient pu s'exclamer lors de la publication de *L'origine des espèces* : «C'est du Lamarck».

Darwin certes, mais n'oublions pas que sa théorie n'est pas née du néant, mais repose sur l'œuvre de ses prédécesseurs.

Goulven LAURENT, Brest

Ce point paraît banal, pourtant passer, pour un détail, mais ça en fait son importance lorsque l'on argumente sur les crises qui ont touché l'évolution de la Vie. Il fait en effet (et par extension) référence à l'exploitation de bases de données qui utilisent parfois des termes erronés pour désigner des groupes d'animaux ou de végétaux, que l'on ne va pas forcément vérifier dans une littérature abondante et dispersée.

Lorsque C. Babin et M. Gayet parlent du «net regroupement amphibien» au cours du mas de l'évolution, ils ont raison, mais avec un argument faux – l'«extinction des labyrinthodontes» – pour deux raisons : la première est que le terme «labyrinthodonte» (littéralement, «à dent montrant une dentine plissée») pour désigner les amphibiens fossiles) est désuet. Créé en 1850, il n'est plus guère utilisé aujourd'hui, ce type de dent étant également présent chez certains poissons.

Les auteurs pourraient rétorquer qu'il en va de même pour le terme actuellement retenu de «stégocéphale» («dont la tête est composée de plaques»), valable aussi pour certains poissons et créé 18 ans plus tard. Certes, mais ce dernier a tout de même un avantage sur l'autre : il fait référence aujourd'hui à un groupe bien défini (c'est-à-dire monophylétique) comprenant tous les amphibiens fossiles (même si leur cran s'écarte pas grand chose à celui avec nos batraciens actuels), et c'est donc lui qui aurait dû être usité...

Cette remarque n'est pas, comme on pourrait le croire, une simple question de vocabulaire. On a vu (trop) souvent (même si ce n'est pas le cas ici) des courbes montrant l'aire d'extinctions «gommees» ou, au contraire, «exagérées» en fonction des noms de groupes utilisés.

La deuxième erreur est plus grave, car elle est relative à des données qui, hélas, ne sont apparemment pas mises à jour. Les stégocéphales ne se sont jamais éteints au Trias supérieur. Les tétrapodes primitifs ont certes connu un «recul» pendant cette crise (pour reprendre le terme de C. Babin et M. Gayet), mais ils l'ont passée tant bien que mal. On connaît l'effet des stégocéphales dans le Crétacé depuis les années 1960 et dans le Jurassique depuis bientôt dix ans!

J. Sébastien STEYER

Muséum national d'histoire naturelle, Paris

Les extinctions

Dans l'article de Claude Babin et Mireille Gayet (voir *Dossier Pour la Science*, juillet 2000), du reste exhaustif et embrassant toutes les hypothèses les plus récentes émises sur les diverses extinctions ayant jalonné l'évolution de la Vie, j'aimerais revenir sur un point particulier concernant l'extinction du Trias supérieur, il y a environ 215 millions d'années.

Écrivez à : tribune@pourscience.fr sur la science vos relations, vos commentaires et vos réflexions sur l'actualité scientifique et vos relations avec la science. L'adresse de *Pour la Science* sur Internet est : tribune@pourscience.fr. Découvrez d'autres contributions de lecteurs sur le site Web www.pourscience.com.

Jean-Michel Thiriet



Nouvelle université

YVES MICHAUD

L'Université de tous les savoirs connaît un succès que même les plus optimistes n'osaient envisager. Cette université d'un nouveau type fera-t-elle évoluer l'enseignement ?

L'université de tous les savoirs est cet ensemble de 366 leçons (une leçon par jour pendant toute l'année 2000) qui présente un panorama du savoir actuel. Jean-Jacques Aillagon, président de la Mission pour la célébration de l'an 2000, avait d'abord voulu fonder une «université populaire», ouverte à tous, qui mettrait en contact direct les plus grands spécialistes, savants et penseurs, et ce qu'il rappelle sans trop savoir ce que cela signifie le «grand public». Le succès a largement dépassé nos attentes : 500 personnes, en moyenne, ont fréquenté, chaque jour, l'amphithéâtre Painlevé du Conservatoire national des arts et métiers ; soit plus de 150 000 auditeurs, sans compter les 500 à 1000 personnes qui consultent les leçons sur l'Internet (www.telerama.fr).

Ce succès est plein d'enseignements sur la façon dont les sciences sont communiquées au public. Tout d'abord, nombreux sont ceux qui veulent comprendre les sciences. Certes, quelques uns de nos auditeurs viennent seulement rêver, mais notre public, qui se renouvelle beaucoup avec les thèmes abordés, est très souvent un public déjà qualifié, qui vient s'informer des avancées et des recherches, voire un public de jeunes passionnés qui viennent nourrir leur vocation scientifique.

Ainsi, malgré les efforts des journaux, revues, collections de vulgarisation, les besoins du public restent incomplètement satisfaits. Les canaux de diffusion, notamment télévisuels, sont les grands absents de la transmission du savoir, même quand ils y ont prétendument vocation. Que nous avons-nous pas entendu quand nous préparions le projet : proposer des leçons ridicules à l'époque du débat et du zapping, il fallait absolument choisir des vedettes médiatiques, il fallait faire simple, accessible à «n'importe qui». Toutes ces prévisions se sont révélées être des préjugés : on ne fait jamais de vulgarisation «pour n'importe qui» ; la plupart des personnes sont fatiguées des zappings, des «ça se discute», des «on en parle». Quant aux vedettes médiatiques, leur notoriété est relative et elle ne sert qu'à rassurer le petit monde des

médias qui imagine qu'il invite des gens connus parce que... lui-même les connaît.

J'ai été frappé, jusqu'à en être ému parfois par l'extrême disponibilité du monde scientifique et technique pour cette aventure, compris quand certains ne donnaient pas cher du projet. L'immense générosité des chercheurs et des techniciens, la nature même de leur engagement «pour la science» ou du moins qu'ils peuvent avoir une certaine liberté de leur financement de leurs recherches, d'une volonté démocratique de communiquer et de faire valoir les connaissances et leurs enjeux ? L'effort, parfois, est surtout frappant dans les sciences «dures», plus «généreux» que les sciences des sciences humaines.

L'exposition des connaissances et des démarches effectives est décalé par rapport aux programmes scolaires et universitaires. Nous avons eu la chance (parce que nous devions établir un programme sur toute une année) de procéder de manière largement transdisciplinaire, selon l'ordre dynamique de la recherche et non selon l'ordre des programmes. Pour moi qui ai longtemps enseigné l'épistémologie aux philosophes, ce contact renouvelé avec l'activité scientifique telle qu'elle se fait a bouleversé des représentations si lourdes de l'académisme qu'elles en avaient vidé la connaissance de tout intérêt. Nous avons vu constamment le rôle structurant des mathématiques, la place exceptionnelle de l'instrumentation et des montages expérimentaux, et le rôle de l'analogie.

Sous un angle différent encore, le contenu des conférences a joué un autre rôle. Il aurait été vraiment naïf pour croire que même un très grand physicien ou un très grand chimiste pourrait transmettre sa science en 30 minutes d'exposé. Ce qu'il peut faire, en revanche, c'est, en collaboration avec les autres conférenciers, fournir au public les éléments d'une réflexion contradictoire et dialectique. De ce point de vue, l'Université de tous les savoirs a parfaitement rempli son rôle de forum démocratique de réflexion sur les connaissances actuelles et sur leurs enjeux. Nous n'avons

jamais eu la prétention de dire aux auditeurs ce qu'il fallait penser de la tâche folle, des organismes génétiquement modifiés, de la pollution par les métaux lourds, du clonage thérapeutique, des thérapies géniques ou de l'interception des télécommunications par satellite, mais nous avons plusieurs fois manifesté que nos conférenciers, à eux tous, donnaient les éléments indispensables pour que chacun se fasse une opinion.

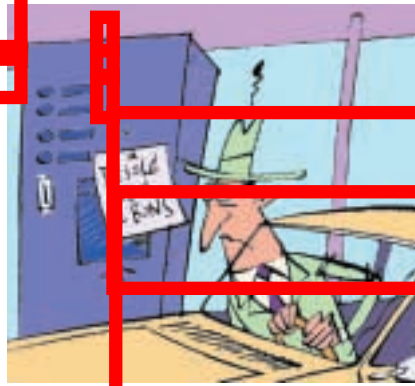
Bien d'autres leçons ont été tirées de l'expérience en cours. Tout d'abord, les meilleurs vulgarisateurs semblent être ceux qui connaissent parfaitement leur domaine et qui prennent du plaisir à y travailler : il y a une gourmandise de la connaissance. Deuxièmement ce ne sont pas les personnes qui passent le plus de temps à la télévision qui font les meilleurs conférenciers. Troisièmement, à force de «petits séminaires» et de débats convenus, de cours en petits groupes, certains de nos orateurs ne savent plus parler devant 500 personnes et ils sont soudain pris de trac. Le public, lui, ajuste très exactement ses applaudissements à la qualité de la performance.

Certaines disciplines souffrent d'un handicap par manque de culture du public : la chimie ou le droit, par exemple. En outre, il y a toujours, en France, une réticence à parler de techniques, de matériaux alors même que chacun a un avis tranché sur les dangers qu'ils font courir. Chez des économistes, on aimerait trouver l'effort théorique et généreux de pédagogie que l'on trouve chez les spécialistes de physique quantique, de supraconductivité ou de microscopie en champ proche.

L'année 2000 n'est pas terminée, et d'autres leçons suivront, mais une question me préoccupe : l'Université saura-t-elle profiter des enseignements de l'Université de tous les savoirs ?

Yves MICHAUD, professeur de philosophie à Paris 1, est responsable du projet Université de tous les savoirs (UTLS), organisé par la Mission pour la célébration de l'an 2000.

Jean-Michel Thiriet



Un marché incomplet

IVAR EKELAND

En pouvant acheter des bons d'essence à prix fixé et valables une certaine durée, tout le monde serait gagnant. Le marché ainsi complété éviterait peut-être les troubles sociaux.

Pourquoi les stations-service ne vendent-elles pas des bons de dix litres de carburant valables un mois ? Vous auriez le choix : soit faire le plein et payer l'essence (ou le gazole) au prix affiché, soit aller à la caisse et acheter, pour un prix un peu supérieur, cinq bons de dix litres qui seraient périmés un mois plus tard.

Les avantages d'un tel système sautent aux yeux. Vous êtes automobiliste et vous craignez une hausse du prix de l'essence ? Inutile d'aller remplir vos jerricans ou de stocker le carburant dans votre baignoire : achetez d'avance votre consommation du mois. Vous êtes transporteur routier, et un client vous demande une livraison de Paris à Madrid, à effectuer dans trois semaines ? Vous acceptez, mais vous lui facturez, non pas le prix du gazole à la pompe aujourd'hui, mais le prix de son de gazole, et vous vous empressiez d'acheter dès aujourd'hui la quantité de bons nécessaire.

Que de désordres évités : l'existence (de bons) précéderait l'essence. Plus de barrages sur les autoroutes ni aux abords des villes, plus de dépôts d'essence bloqués. Ces bons existeront pour diverses échéances, un, deux ou trois mois par exemple, et les transporteurs pourront conclure des contrats à effet différé en se garantissant contre les hausses des prix. Bien sûr, le bon à trois mois coûte plus cher que le bon à un mois, qui sera plus onéreux que le prix à la pompe : la différence de prix est le coût de la garantie, d'autant plus chère qu'elle s'étend loin dans le temps. Un contrat conclu aujourd'hui pour un transport qui n'aura lieu que dans dix semaines sera négocié sur la base du prix du bon à trois mois. Le moment de la livraison venu, le transporteur présentera à la pompe les bons qu'il aura achetés au moment de la conclusion du contrat, et le tour sera joué.

Le transporteur et le client sont contents, parce qu'ils sont débarrassés des mauvaises surprises : l'opération financière est bouclée à la signature du contrat, et chacun sait exactement à ce moment-là ce que cela lui coûte et ce que cela lui rapporte. Mais, m'objectera-t-on, il faut bien

que quelqu'un perde dans cette histoire ? Si ce n'est ni le transporteur ni son client, il faut donc que ce soit l'émetteur des bons ?

En bien non, et c'est toute la magie de la finance moderne. Examinons la question du point de vue de l'émetteur. Bien entendu, une simple station-service ne se livre pas à ce jeu : ce sera plutôt le rôle d'une grande compagnie pétrolière, qui tirera avantage du fait que ses bons seront valables sur tout son réseau de distribution. Pour ne pas perdre d'argent, elle même pour en gagner, elle disposera d'une panoplie de mesures.

D'abord, la fixation du prix des bons. De ce point de vue, la compagnie est dans une position très favorable par rapport aux consommateurs individuels : elle est bien mieux renseignée qu'eux sur l'évolution prévisible des prix du brut, elle est capable de faire des stocks importants quand la conjonction s'annonce mauvaise, et elle a une assise financière qui lui permet de franchir plus facilement un mauvais cap.

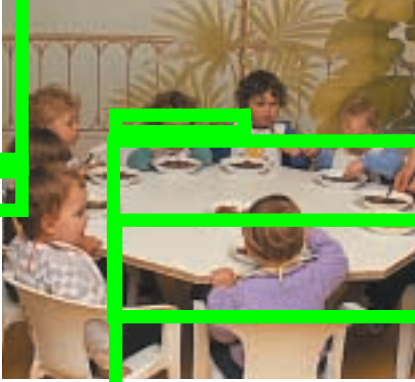
Un petit transporteur routier travaille avec des marges qui rendent insupportable une variation de cinq pour cent du prix du carburant, et il sera prêt à payer comparativement cher une assurance contre le risque, même si le coût en est très faible et surtout si c'est le client qui, en définitive, règlera la facture). Si le litre est à 10 F aujourd'hui et a une chance sur dix de passer à 10,1 F dans trois mois, (mais neuf chances sur dix de rester à 10 F), la prime d'assurance à payer pour se garantir contre ce risque est statistiquement de $(0,5 \times 0,1)$ F, soit cinq centimes, ce qui met le bon de un litre valable trois mois à 10,05 F. Les transporteurs le paieront volontiers 10,1, voire 10,2 F, surtout si une publicité bien conçue leur fait suffisamment connaître le risque.

En second lieu, la compagnie peut se prémunir contre des variations imprévues du prix du brut par des opérations financières. La flambée des prix du pétrole s'accompagne d'une flambée non moins importante du prix de certaines actions - notamment celles des compagnies pétrolières. Si donc l'on veut se prémunir contre une hausse, on peut (ce n'est qu'un

exemple de couverture parmi d'autres) constituer un portefeuille d'actions dont la valeur s'élève avec le prix du brut, et les couvrir. Si le prix monte, les consommateurs se présentent aux stations pour faire le plein, et la compagnie ne se voit donc contrainte de leur vendre du carburant à un prix inférieur au cours du jour, mais son portefeuille de couverture a pris de la valeur. Si tout est bien calculé, l'argent perdu d'un côté est regagné de l'autre, et l'opération est blanche.

Du point de vue de la théorie économique, il n'y a là rien que de très banal. Le problème du transporteur, c'est qu'il ne peut pas acheter aujourd'hui du carburant livrable dans trois mois, ce qu'un économiste exprimerait en disant que le marché n'est pas complet. Le remède à cette situation est de compléter le marché, en créant le produit manquant. C'est ce qu'on appelle un contrat à terme, et il en existe pour toutes les matières premières, mais à une échelle industrielle : le contrat à terme sur le marché du pétrole brut porte sur mille barils. Tout comme un contrat d'assurance, le rôle de ces contrats est de transférer les risques de ceux qui ne veulent pas les prendre, à ceux qui savent les gérer. Ces contrats sont complètement entrés dans les mœurs de la vie économique, et nous en utilisons tous les jours sans même nous en rendre compte. Le timbre ou le ticket de métro, par exemple, ne portent plus d'indication de prix : mais si demain le prix de la lettre ou du trajet urbain change, ils restent utilisables, et c'est dans une hausse, j'en achète d'avance. Cela signifie que ce sont en fait des contrats à terme passés avec l'Etat ou avec la RATP. Les bons de carburant rendraient le même service, et, en plus, permettraient aux compagnies de gagner davantage d'argent. C'est si commode que je reviens à ma question initiale : pourquoi personne ne se préoccupe-t-il de compléter ce marché ?

Ivar EKELAND est professeur à l'Université Paris-Dauphine.



Petit Format/A. Chaudat

En culotte courte

HERVÉ THIS

L'observation des enfants en crèche révèle l'évolution de leurs goûts alimentaires.

Les parents gastronomes font tous entendre la même antienne : leurs enfants semblent n'aimer que les féculents – pâtes, riz, pommes de terre –, les fromages les plus fades et les viandes inaptes, tel le blanc de poulet. Pourquoi les enfants commencent-ils leur vie alimentaire par ces aliments fades ? Comment les mettre sur la « voie » ? Psychologues et sociologues pourraient s'interroger sur les motivations biologiques qui poussent les parents à se désoler, mais, aujourd'hui, ce sont des « sensorialistes » qui explorent les choix alimentaires des enfants et l'évolution de ces choix.

L'expérience qui livre aujourd'hui ses premiers résultats a commencé en 1982 et s'est achevée en 1999, dans la crèche Garette de l'Hôpital de Dijon. Elle a porté sur des enfants âgés de deux à trois ans, qui, au déjeuner, choisissaient la composition de leur repas. Avec quelques contraintes, toutefois : parmi les huit aliments proposés figuraient du pain, deux entrées, une viande ou un produit carné ou un poisson, deux légumes ou féculents, deux fromages ; les plats sucrés n'étaient pas mis au menu, en raison de l'attrait présumé pour le sucré : ils étaient réservés au goûter. Les enfants pouvaient se réserver, mais ils n'étaient pas autorisés à prendre plus d'une fois le même aliment pendant le même repas. Pour les 25 enfants présents à la crèche, chaque année, les puéricultrices orchestraient et notaient ces libres choix ; les huit plats étaient sur une table, les enfants prenaient ce qu'ils voulaient, et ils pouvaient notamment ne rien manger. Au total, 420 enfants ont été suivis sur 110 repas en moyenne.

Aujourd'hui, Sophie Nicklaus et Sylvie Issanchou du Laboratoire de recherches sur les arômes de la Station INRA de Dijon, en collaboration avec Vincent Boggio, de la Faculté de médecine de Dijon, dépouillent les résultats. Les questions étudiées sont nombreuses : quels aliments ont-ils été consommés ? Les enfants ont-ils des aversions générales pour certains plats ? Comment interpréter les différences de choix entre les plats ? Comment expliquer les variations des choix ? Par des particularités individuelles des enfants ? À ces questions, le dépouillement des données

donne des confirmations et des informations nouvelles.

Tout d'abord, comme le pressentent les parents, les enfants choisissent surtout des féculents et des viandes ; les fromages choisis sont ceux qui ont peu de goût et une texture molle, et rares sont les amateurs de roquefort dès l'âge de deux ans. La culotte courte, peut-être, mais pas en couche. Le pain « signe des temps » : les féculents sont les favoris, mais pas aux pâtes au riz ou aux frites. Et les plats les plus consommés sont les frites, les chips, les lasagnes, les quiches, le poisson pané, le riz, la purée, le jambon, le bifteck. Pas de surprise, mais une confirmation quantifiée.

Les choix des viandes, notamment, ont étonné les chercheurs : les enfants font peu de différence entre les viandes, choisissant presque indistinctement le rôti de porc, le dinde, le gigot ou les abats. En revanche, pour les légumes (moins souvent choisis que les viandes), les choix sont tranchés : le type de légumes et le type de préparation culinaire déterminent les choix des enfants. Par exemple, l'épinard, qui a une mauvaise réputation parmi les enfants, est pourtant le légume le plus consommé... quand il est nappé de sauce blanche. Les légumes tels que les endives, le chou (cru ou cuit), les tomates ou les haricots verts attirent peu les gastronomes de crèche : apparemment les produits de texture dure et fibreuse (difficiles à mâcher) sont rejetés, de même que ceux dont l'amertume est prononcée.

Pourquoi ces choix des enfants ? Les sensorialistes ont cherché à mettre en relation ces choix avec la valeur nutritionnelle, et ils ont découvert que plus les aliments ont une valeur énergétique notable plus ils sont choisis (à l'exception des fromages). Est-ce en fait dans cette observation une influence encore forte de l'inné, chez les enfants ? On sait que les nourrissons à qui l'on fait goûter des solutions salées, sucrées, acides ou amères font des grimaces de dégoût pour l'acide ou l'amer, et une mimique de plaisir quand ils testent du sucré : ils sont analogues aux singes frugivores, qui associent la saveur sucrée à la présence de sucres,

molécules à haute densité d'énergie, et l'amer aux végétaux qui contiennent des alcaloïdes toxiques. C'est progressivement que l'enfant apprend, plus tard, par conditionnement, à associer le goût à la satisfaction de son alimentation : le conditionnement est celui qui apparaît, par exemple, quand le bébé associe le lait à la satisfaction.

La conditionnement préalable de molécules énergétiques et non sucrées, telles les graisses ; la culture, elle, explique que des aliments même tout jeunes se délectent de mets au goût puissant, que les explications biologiques feraient rejeter.

La deuxième synthèse importante concerne la forte variabilité des choix pour les produits à goût puissant. Aujourd'hui, aucune des variables analysées n'explique cette variabilité : ni le sexe, ni l'allaitement maternel, ni le rang de l'enfant dans la famille, ni son indice de masse corporelle (une mesure de la masse grasseuse, que l'on obtient en divisant le poids par la taille au carré)... On ignore donc encore comment s'expliquent les variations importantes existant d'un enfant à l'autre.

Les choix alimentaires des jeunes enfants avant été très peu étudiés, il n'est pas étonnant que ces premières études aient de nombreuses observations non expliquées. De surcroît, il faudra encore attendre longtemps avant de pouvoir contrôler les choix des enfants étudiés aux préférences des adultes que ces enfants deviendront. C'est à ce prix que l'on comprendra les mécanismes de l'apprentissage alimentaire au cours de l'enfance, et que l'on découvrira l'importance des expositions précoces dans les choix ultérieurs, supposée depuis longtemps, mais peu démontrée. L'industrie s'intéresse évidemment aux résultats de ces études : pour offrir des produits qui plairont plus longtemps, elle veut comprendre la dynamique des préférences alimentaires des tout-petits.

Prochain rendez-vous France Info et Pour la Science le 29 novembre 2000, avec la chronique Info Sciences de Marie-Odile Monchicourt.



Le magnétisme grenoblois

PIERRE AVERBUCH

Il y a 50 ans, Louis Néel expliquait le comportement des corps magnétiques utilisés aujourd'hui tant dans les dynamos de vélos que dans les disques durs des ordinateurs.

Aujourd'hui, Grenoble est renommée, parmi les physiciens, pour ses laboratoires et ses grands équipements, le synchrotron européen, par exemple. Ce centre européen de la physique est né avec l'arrivée, en 1940, de Louis Néel (ci-dessus), qui s'est intéressé au magnétisme. Puis il a contribué à y développer la physique des solides, la physique nucléaire et de nombreuses autres recherches.

À la fin du XIX^e siècle, on distingue les corps paramagnétiques des corps ferromagnétiques. Les premiers s'aimantent sous l'influence d'un champ extérieur, mais perdent leur aimantation dès que ce champ disparaît. C'est le cas de l'aluminium, de l'oxygène, du sodium, du platine, etc. Les seconds conservent l'aimantation qu'ils ont acquise dans un champ magnétique, même après la suppression de ce dernier, devenant ainsi des aimants permanents. Au-delà d'une certaine température (découverte par Pierre Curie et nommée aujourd'hui température de Curie), les substances ferromagnétiques deviennent paramagnétiques. C'est le cas du fer, du nickel, du cobalt, des alliages contenant ces éléments et de la magnétite ou pierre d'aimant, que Platon connaissait déjà : «Non seulement la pierre attire des anneaux de fer, mais elle leur communique ce pouvoir, qui permet à ces anneaux d'en attirer d'autres.» Les corps ferromagnétiques sont utilisés depuis longtemps pour fabriquer des boussoles, mais aussi de nombreux objets plus modernes, telles les bandes magnétiques : les particules d'oxyde de fer placées sur le support en plastique sont aimantées lors de l'enregistrement et conservent les informations grâce à leur orientation. De même, les bits des disques durs d'ordinateurs, qui codent l'information sous forme binaire, sont des petits aimants dont les champs sont orientés dans un sens ou dans l'autre.

Au début du XX^e siècle, Paul Langevin explique le paramagnétisme par le

magnétisme électronique à l'échelle atomique : les électrons ont des propriétés magnétiques parce qu'ils tournent autour du noyau. Les corps paramagnétiques sont ceux dont les éléments ont un moment magnétique intrinsèque. Le moment magnétique de chaque atome est orienté au hasard à l'échelle macroscopique, le corps n'est pas aimanté. En revanche, dans un champ magnétique, les moments atomiques s'alignent comme l'aiguille d'une boussole sur le champ terrestre. L'agitation thermique s'oppose à cet ordre. Paul Langevin explique ainsi la diminution du paramagnétisme avec la température.

Peu après, Pierre Weiss élabore une phénoménologie du ferromagnétisme, toujours en vigueur aujourd'hui. Elle est fondée sur l'existence d'un couplage entre les moments magnétiques intrinsèques des atomes. Weiss invente le terme de «champ moléculaire», proportionnel à l'aimantation, pour désigner le phénomène responsable de ce couplage. Il suppose que les substances ferromagnétiques sont formées d'un grand nombre de domaines microscopiques (les domaines de Weiss), où les moments magnétiques sont parallèles. Chaque domaine est un petit aimant, mais leurs champs se compensent, en raison des orientations diverses des domaines. Lorsque l'on place un corps ferromagnétique dans un champ extérieur, l'aimantation des domaines s'aligne sur ce champ et le corps s'aimante.

L. Néel arrive à l'Université de Strasbourg en 1928. Jeune agrégé sortant de l'École normale supérieure, il y est nommé assistant de physique. Il prépare sa thèse sur le comportement des corps ferromagnétiques au voisinage de la température de Curie. Parmi les travaux de recherche qu'il poursuit ensuite, l'un des plus remarquables est la découverte de l'antiferromagnétisme.

Avec l'avènement de la mécanique quantique, on avait compris que l'électron, en plus de son mouvement orbital autour du noyau atomique, a un moment ciné-

tique intrinsèque (le spin), auquel est associé un moment magnétique. Selon le principe d'exclusion de Pauli, deux électrons dont les spins sont parallèles ne peuvent se rapprocher trop l'un de l'autre. L'énergie de répulsion électrostatique entre deux électrons est donc même forte quand leurs spins sont parallèles que quand ils sont tête-à-tête. En 1926, Werner Heisenberg comprend que les moments magnétiques s'alignent dans un même domaine de Weiss, à cause de ce «couplage d'échange».

PLUSIEURS MAGNÉTISMES

En 1931, la communauté des physiciens est convaincue que le couplage d'échange favorise le parallélisme des spins, dit positif, et le seul possible. Toutefois, L. Néel propose une interaction d'échange négative pour expliquer que certains corps, tel le manganèse et ses oxydes, ont un paramagnétisme indépendant de la température. Pour comprendre son raisonnement, imaginons que les atomes des corps antiferromagnétiques occupent les cases d'un grand échiquier. Supposons que les interactions d'échange couplent entre eux les atomes placés sur des cases ayant une lettre en commun (les atomes qui n'ont qu'un sommet en commun ne sont pas couplés). Si le couplage d'échange est positif, les spins tendent à être parallèles et le corps est ferromagnétique, tel que Weiss l'a décrit. Si le couplage est négatif, l'énergie est minimale quand les spins des atomes qui occupent les cases blanches sont orientés dans la direction opposée à celle des spins des atomes qui occupent les cases noires. On a alors un corps antiferromagnétique, tel que L. Néel l'a décrit.

Les atomes des solides étant organisés en réseau cristallin, l'aimantation tend à s'aligner selon certains axes dits de facile aimantation. Revenons à notre échiquier : supposons que l'axe de facile

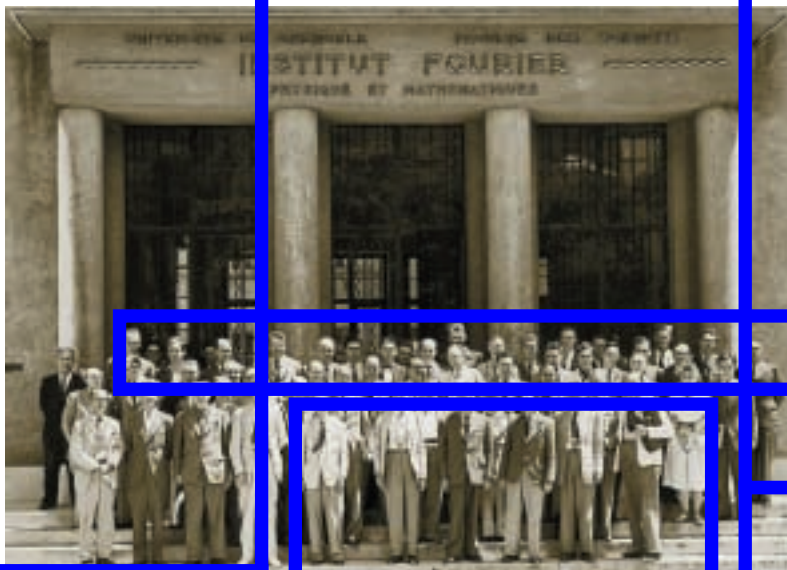
aimantation lui soit perpendiculaire. Soumettons-le à un champ magnétique parallèle à cette direction : le couplage est positif avec les atomes des cases noires et négatif avec les atomes des cases blanches (ou inversement). Au total, le couplage est nul et aucun magnétisme global n'apparaît. En revanche, si le champ est parallèle au plan de l'échiquier, il fait pencher tous les spins dans sa direction : un paramagnétisme apparaît, qui semble indépendant de la température.

Bien que ce modèle s'appuie sur des expériences réalisées avec des cristaux, seule une poignée de physiciens s'y rallient. La communauté scientifique attendra plus de dix ans pour reconnaître la validité de ce modèle, quand la diffraction des neutrons confirmera l'orientation des moments magnétiques atomiques.

Pendant la guerre, L. Néel, professeur à Strasbourg, est mobilisé. En le charge de trouver un moyen de protéger les navires de la flotte britannique, il découvre les lois de l'hystérésis magnétique en champ faible, connues depuis le XIX^e siècle, pour appliquer aux navires un champ magnétique qui détruit l'aimantation globale des coques métalliques de sorte qu'elles n'attirent plus les mines dévastatrices.

L'ARRIVÉE À GRENOBLE

Après l'annexion de l'Alsace par l'Allemagne, L. Néel cherche un nouveau laboratoire d'accueil. À Grenoble, le doyen de la Faculté des sciences avait obtenu du Front populaire la construction d'un bâtiment pour les physiciens et les mathématiciens. L. Néel y obtient sa mutation, mais les lois raciales du régime de Vichy interdisent à son élève, Louis Weil, d'être universitaire. Bien qu'il soit professeur d'université, L. Néel tisse alors des liens étroits avec les industriels et obtient que la société Ugine embauche Weil. En retour, cette société va prendre le marché des aimants destinés, à l'époque, aux dynamos des véhicules, grâce aux travaux de Weil sur les poudres métalliques. Ce dernier utilise les résultats des recherches que L. Néel a entreprises pour mieux comprendre les phénomènes d'hystérésis qu'il a utilisés lorsqu'il a désaimanté les bateaux. L. Néel explique comment l'état d'un système magnétique dépend des diffé-



Le premier grand rassemblement international sur le magnétisme eut lieu juste après la Seconde Guerre mondiale, à Grenoble. Louis Néel, qui en fut le moteur, est le quatrième à partir de la droite, au premier rang.

rentes configurations dans lesquelles il s'est trouvée. Il comprend que l'hystérésis est due au fait que, dans les champs faibles, les parois entre domaines de Weiss se déplacent et que, dans les champs forts, la structure de la décomposition en domaines est modifiée. Dans le cas de grains magnétiques assez petits, il n'y a plus d'un seul domaine : la séparation de domaines abaisse l'énergie du matériau, mais ce gain de stabilité est inférieur au coût énergétique de la création d'une paroi. En utilisant des poudres métalliques qu'elle oriente puis qu'elle compacte, la société Ugine fabrique des aimants dont l'aimantation est plus stable que s'ils étaient d'un matériau massif.

En 1944, L. Néel devient directeur d'un laboratoire de l'Institut électrotechnique de Grenoble dédié à des essais pour les industriels. Puis, juste après la guerre, il crée le Laboratoire d'électrostatique et de physique du métal. À cette époque, les besoins de l'électronique imposent l'étude du magnétisme des ferrites, oxydes mixtes de deux ou trois métaux, dont l'un au moins est aimanté. On s'aperçoit que la variation de leur aimantation avec la température diffère de celle des corps ferromagnétiques. Pour expliquer ce comportement, L. Néel généralise alors son modèle d'antiferromagnétisme au cas où les atomes couplés négativement sont de natures différentes. Dans un matériau antiferromagnétique, les atomes des cases noires

et ceux des cases blanches sont identiques chimiquement et ont des moments magnétiques de même amplitude, qui s'annule en l'absence de champ extérieur. Dans les ferrites et dans les corps ferromagnétiques en général, les atomes des cases noires et ceux des cases blanches sont d'espèces chimiques dif-

férentes : les deux sous-réseaux ne se compensent plus, et un magnétisme macroscopique apparaît.

Les travaux de L. Néel ont également servi en géophysique : certaines ferrites naturelles évoluent lentement sur des échelles de temps géologiques. Les laves qui en contiennent s'aimantent en refroidissant. Grâce à ces boussoles fossiles, on reconstitue l'histoire du champ magnétique terrestre dont la direction varie au cours du temps. On doit alors tenir compte de la dérive des continents. Les travaux de L. Néel ont ainsi

indirectement permis d'utiliser la « mémoire magnétique » des roches pour reconstituer l'histoire de la dérive des continents.

Pour poursuivre des recherches en magnétisme, les physiciens ont besoin d'instruments de plus en plus variés : Weil construit des liquéfacteurs d'hydrogène, puis d'hélium pour atteindre les basses températures nécessaires à l'étude du magnétisme des ferrites. Une équipe de chercheurs spécialistes de la résonance magnétique s'implante à Grenoble pour étudier les effets magnétiques dynamiques. L. Néel cherche un réacteur nucléaire comme source de neutrons (diffraction par la matière et sensibles aux orientations des moments magnétiques, ils permettent de visualiser l'antiferromagnétisme). Il contacte le Commissariat à l'énergie atomique au moment où ce dernier cherche à créer un troisième centre en province. À la fin de l'année 1955, L. Néel devient directeur du Centre d'étude nucléaire de Grenoble. Les capacités de recherche de la ville doublent avec ces nouveaux laboratoires. Il s'y développe une synergie entre la recherche universitaire et l'industrie. En 1970, L. Néel reçoit le prix Nobel de physique qui couronne sa carrière, consacrée au magnétisme. Grenoble lui doit en partie son statut de centre scientifique européen et de pôle de l'industrie électronique.

Pierre AVERBUCH est physicien honoraire au Centre de recherche sur les très basses températures, à Grenoble.

Pierre AVERBUCH, Les débuts du magnétisme grenoblois, in Bulletin de la Société française de physique, n° 125, p. 16, juillet 2000.

■ POUR LA SCIENCE

VOTRE ABONNEMENT 100 % NUMÉRIQUE

1 an au prix de 48 €

~~soit 4 € seulement par numéro au lieu de 5,20 €~~

Pour consulter et télécharger
en toute liberté sur www.pourlascience.fr
les 12 prochains numéros
dès leur parution.



Oui, je m'abonne
à Pour la Science 100% numérique
1 an (12 numéros) pour 48 €

Je préfère découvrir toutes
les offres d'abonnement disponibles
en cliquant ici.

Nouveaux cousins de l'archéoptéryx

La découverte de divers quasi volatiles à plumes et à poils confirme que les oiseaux descendent des dinosaures... par l'archéoptéryx.

Qu'est-ce qu'un oiseau ? Les paléontologues sont perplexes après la découverte de plusieurs petits dinosaures du Crétacé inférieur (-144 à -66,4 millions d'années) dont le squelette se distingue à peine de celui d'un volatile. Ces fossiles ne remettent pas en cause l'importance de l'archéoptéryx, car le plus ancien oiseau connu vivait au Jurassique supérieur, il y a 150 millions d'années, bien avant les animaux trouvés récemment. Toutefois, la découverte de plusieurs dinosaures à la posture d'oiseau, à plumes ou à poils, intrigue les paléontologues, car elle indique comment le vol et les plumes sont apparus.

La trouvaille la plus récente est celle du *Bambiraptor feinbergi*, un petit prédateur coureur du groupe des théropodes, les dinosaures bipèdes dont le plus fameux est le tyranosaure. Il y a quelques mois, des amateurs de fossiles du Montana ont découvert le *Bambiraptor feinbergi* dans une couche d'alluvions fluviales datant d'environ 80 millions d'années, c'est-à-dire du Crétacé. Ils ont mis au jour environ 90 pour cent d'un squelette. Cette exceptionnelle complétude permet de reconstituer l'articulation de la tête et celles des membres.

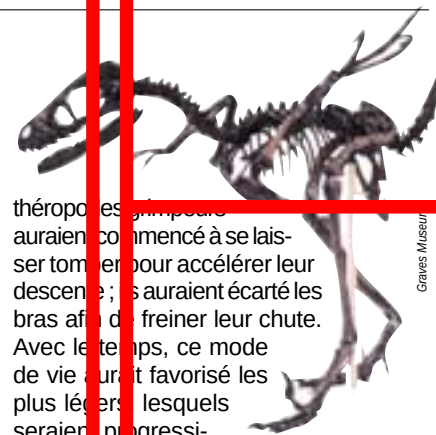


Bambiraptor feinbergi (ci-dessus, une représentation d'artiste et, en haut à droite, le squelette reconstitué) avait des ailes, mais aussi des griffes de chasseur.

Les proportions du squelette, notamment les longueurs des bras et des doigts, rapprochent *Bambiraptor feinbergi* de l'archéoptéryx. De plus, il possède plusieurs caractéristiques aviaires : une sorte de bréchet, une arête osseuse saillante sur le sternum de la plupart des oiseaux, et une articulation de l'épaule orientée vers le côté, qui facilitait les battements d'ailes. Le petit prédateur, qui mesure un mètre environ, est pourtant un membre de la famille des dromaeosaures, un sous-groupe des théropodes auquel appartenait le Velociraptor, de sinistre réputation depuis le film *Jurassic Park*. Comme son cousin, *Bambiraptor feinbergi* avait une «patte de tueur». Cette extrémité est une caractéristique du sous-groupe des dromaeosaures : elle est constituée de trois doigts griffus, dont le deuxième, très mobile, est armé d'une griffe en forme de faucille, longue et taillée (le poignard). C'est pour quoi, les paléontologues pensent que *Bambiraptor feinbergi* est bien un dinosaure, proche des oiseaux.

Plusieurs fossiles, qui évoquent des oiseaux, ont aussi été trouvés en Chine. Le plus intéressant, parce que le plus ancien, est celui de *Caudipteryx*. Il s'agit d'une espèce du groupe des oviraptorosaures, d'autres dinosaures bipèdes typiques du Crétacé, ainsi nommés parce qu'on les soupçonne d'avoir consommé les œufs des autres espèces. Le *Caudipteryx* gambadait en Chine il y a 124 millions d'années. Les bras et la queue couverts de plumes, doté d'une tête munie d'un bec (denté), il ressemble tellement à un oiseau qu'il pourrait fort bien... en être un ! S'agit-il d'un dinosaure ou d'une forme d'oiseau que son mode de vie terrestre a fini par priver du don de voler, comme c'est le cas, par exemple, des autruches modernes ? Si des oiseaux incapables de voler ont existé seulement 10 à 15 millions d'années après l'archéoptéryx, c'est que l'évolution des volatiles primitifs a été bien plus compliquée et tortueuse qu'on ne le présumait.

D'autres fossiles chinois renforcent cette impression. Certains semblent avoir une queue couverte d'ailes, d'autres des poils qui pourraient avoir été des prémices de plumes. Un autre petit théropode du Crétacé chinois trouvé tout récemment, a un doigt particulièrement long, qui est une adaptation manifeste à l'escalade dans des arbres. Même si ce théropode est postérieur à l'archéoptéryx, il remet au goût du jour une explication séduisante de l'évolution des dinosaures vers l'oiseau : des



Graves Museum

théropodes grimpeurs auraient commencé à se laisser tomber pour accélérer leur descente ; ils auraient écarté les bras afin de freiner leur chute. Avec le temps, ce mode de vie aurait favorisé les plus légers, lesquels seraient progressivement devenus capables de planer.

Naturellement, ces fossiles ne supplantent en intérêt l'archéoptéryx, car les caractéristiques aviaires sont moins nettes. Plus ancien que tous les quasi volatiles trouvés récemment, «l'antique porteur d'ailes» avait déjà un plumage organisé comme celui des oiseaux contemporains ; la structure de ses plumes était aussi «moderne» que celle que nous connaissons. Il est indubitable que l'archéoptéryx volait, et que ce sont ses ancêtres qui ont accompli la «révolution de la plume». Quand ? Probablement au Jurassique inférieur ou moyen, soit entre 208 et 163 millions d'années avant notre ère. Pour le savoir, il faudrait trouver d'autres restes de quasi oiseaux plus vieux que l'archéoptéryx. C'est malheureusement peu probable, car sur toute la planète, les fossiles datant du début du Jurassique sont rarissimes.

Peter WELLNHOFER



Peter Wellnhofer

Ces poils, présents sur le petit prédateur grimpeur chinois *Sinornithosaurus millenii* étaient-ils des précurseurs de plumes ?

Planètes volantes

Des astrophysiciens ont découvert des objets célestes ressemblant à des planètes ... mais qu'en sont pas.

À l'instar de notre planète Terre qui tourne autour du Soleil, vous pensiez naïvement qu'une planète est un corps massif en orbite autour d'une étoile. Ce n'est plus tout à fait exact depuis que l'équipe de Maria Rosa Osondo, de l'Institut d'astrophysique de Tenerife, a découvert des corps qui ne sont pas assez gros pour être des étoiles, mais qui ne sont pas non plus en orbite autour d'une étoile.

En pointant les télescopes optiques des îles Canaries vers le nuage d'Orion, à plus de 1 000 années-lumière de la Terre, les astrophysiciens ont exploré les émissions infrarouges de cette zone. Après un temps d'exposition de plusieurs heures, les images ont révélé la présence de 18 corps inhabituels, séparés de quelques années-lumière. De quoi s'agit-il ?

Le télescope Keck, de l'Observatoire de Hawaï, a enregistré le spectre de trois de ces objets : il révèle que la température de ces corps est d'environ 2 000 degrés. Par ailleurs, l'éclat apparent de ces objets et l'âge de la zone où ils ont été observés montrent qu'ils ont environ cinq millions d'années. À partir de ces deux paramètres – âge et température –, les astrophysiciens disposent de modèles de prédiction de leur masse : elle serait comprise entre cinq et 15 fois la masse de Jupiter (c'est-à-dire entre 950 et 2 850 fois la masse de la Terre).

Les objets observés ne seraient donc pas assez massifs pour être le siège d'une fusion nucléaire telle qu'elle règne au cœur des étoiles. Les naines brunes sont les plus petits objets connus abritant une telle fusion : elles ont une masse environ égale à 75 fois celle de Jupiter. Ainsi, les 18 corps observés ne sont pas des étoiles : seraient-ce des planètes ? Cependant, elles ne gravitent pas autour d'étoiles ! Les astrophysiciens se perdent en conjectures.

Selon une première hypothèse, ces planètes auraient été éjectées d'une ancienne orbite qu'elles suivaient autour d'une étoile. Ce phénomène se produit lorsque deux planètes sur des orbites proches se croisent : l'une d'elles peut être éjectée. Une autre hypothèse stipule que ces objets résulteraient d'une agglomération de matière, dans le nuage d'Orion.

Dans un futur proche, les astronomes chercheront dans divers amas suffisamment jeunes pour y détecter d'autres corps du même type que ceux du nuage d'Orion. Ils explorent également le voisinage du Système solaire où de tels

corps trop froids pour émettre encore dans l'infrarouge, devront être détectés dans le domaine des ondes millimétriques, domaine pour lequel est construit, au Chili, le grand interféromètre ALMA (Atacama Large Millimeter Array), soit Grand Interféromètre millimétrique (Atacama).

Ces objets de recherche de planètes extra-solaires seront complétés en 2004 par le satellite *Corot* dont le Centre national d'études spatiales a récemment décidé le lancement.

En attendant, Planète ? Non-planète ? le vocabulaire des astrophysiciens s'enrichira peut-être d'un nouveau mot.

BRÈVES

Yvon de la fatigue

Comment évaluer la baisse de vigilance après une journée longue et d'intense activité ? Une vingtaine de personnes ont consommé régulièrement de petites quantités d'alcool et subi des tests visant à quantifier leurs réactions. De même, 20 personnes ont été soumises à des privations de sommeil et ont subi les mêmes tests. Au bout de 18 heures de veille, la baisse de vigilance était équivalente à celle d'une personne ayant une alcoolémie de 0,5 gramme par litre, la limite au-delà de laquelle on ne doit plus conduire. Plus la veille se prolonge, plus les résultats se dégradent, comme si l'« alcoolémie » continuait à augmenter. Pourra-t-on faire des expériences pour évaluer la validité de l'expression ivre de colère ?

Des Romains... très américains

Les archéologues ont précisé les habitudes alimentaires des habitants de Magontiacum (Mayence). Ils ont fouillé une aire de culte domestique de la ville antique. Un lieu grouillant de vie, où les citoyens venaient faire des vœux et des offrandes. Les archéologues ont retrouvé des noyaux de dattes, des pignons et des restes de fruits, dans les cendres des foyers. La découverte de restes de fruits méditerranéens confirme que les militaires romains importaient un bon quart de ce qu'ils consommaient. Comme les Français, les Américains et tous les occupants successifs de cette cité qui, depuis 2 000 ans, est une ville de garnison.

Les femmes et la recherche

Selon une enquête ministérielle, les femmes sont deux fois moins nombreuses dans le secteur de la recherche et du développement privé que dans le même secteur public contrairement aux hommes qui sont quasiment aussi nombreux dans les deux secteurs. Dans la recherche publique, les femmes représentaient 38 pour cent en 1997 (30 pour cent au CNRS). Dans le secteur privé, leur proportion ne dépassait guère 17 pour cent en 1997. Comme dans la recherche publique, elles sont davantage présentes dans certains secteurs, tels que la pharmacie et les industries agroalimentaires ou textiles. Et pourtant, les femmes chercheurs sont mieux représentées, dans les entreprises, que les femmes ingénieurs !



Dans le nuage d'Orion (près de l'étoile brillante, à droite), 18 corps célestes ont été découverts, dont trois (encadrés) ont été analysés. Ces objets ne sont ni des étoiles ni des planètes, car, d'une part, ils ne sont pas assez massifs, d'autre part, ils ne tournent autour d'aucune étoile.

Le mystère du méthane

Pourquoi ne détecte-t-on pas de méthane à proximité des immenses réservoirs de méthane qui dorment au fond des océans? Deux solutions: ou bien il ne s'en échappe pas, ou bien il est dégradé. Quel est le coupable? On sait que des archéobactéries méthanogènes, présentes dans des réservoirs dépourvus d'oxygène produisent du méthane et que d'autres bactéries se nourrissent de ce gaz, mais, jusqu'à présent, on ne connaissait aucun micro-organisme capable de détruire le méthane. Une équipe allemande vient d'identifier des bactéries anaérobies qui oxydent le méthane en l'absence d'oxygène; elles vivent en symbiose avec d'autres bactéries qui réduisent les sulfates, fournissant ainsi l'oxygène nécessaire à la transformation du méthane en dioxyde de carbone et en hydrogène.

Des nouvelles d'Ötzi

Ötzi, l'homme de 5 300 ans, découvert en 1991 dans un glacier alpin, n'est peut-être pas mort de froid et d'épuisement comme on l'a d'abord pensé. D'après de récentes analyses, il aurait souffert d'arthrite et de troubles coronaires, et aurait eu une quarantaine d'années, âge quasi canonique



pour grimper les montagnes. Un paléopathologiste britannique pense qu'il ne serait pas mort seul à cet endroit, mais qu'il y aurait été enterré par des compagnons. De nouveaux échantillons d'ADN et de tissus vont être analysés; on devrait bientôt mieux comprendre comment il est mort et d'où il venait.

Manteaux identiques

D'après les informations transmises par le télescope spatial Hubble, Io, l'un des quatre principaux satellites de Jupiter, aurait un manteau identique à celui de la Terre. Io a une intense activité volcanique et Hubble a enregistré des spectres des éruptions volcaniques du volcan le plus actif; on en a déduit la température et la pression à l'intérieur du volcan. Les calculs ont montré que le manteau d'Io est dépourvu de fer. On suppose qu'au début de la vie d'Io, de l'eau, disparue aujourd'hui, aurait réagi avec le fer, de l'hydrogène aurait été libéré et le fer oxydé. Au cours d'un tel mécanisme, le manteau se serait appauvri en fer tandis que le cœur se serait enrichi, comme sur Terre.

La perception sonore

De l'influence de l'orthographe sur la perception auditive.

Lorsque vous entendez la phrase «Le blond a fait un stage de billes en plomb dans une ville», seule l'absence de sens profond vous apparaît, alors que le temps de reconnaissance de chaque mot vous importe peu. Il en va autrement pour les psycholinguistes qui étudient la perception de la parole. L'équipe de Ludovic Ferrand, du Laboratoire de psychologie expérimentale du CNRS, a montré que l'orthographe influence la perception auditive et sur le temps de reconnaissance d'un mot parlé.

Dans les années 1970, les premiers travaux sur la lecture silencieuse ont révélé que pendant celle-ci, on code non seulement l'information visuelle – la forme des lettres –, et orthographique – les lettres –, mais aussi l'information phonologique, c'est-à-dire une information du même type que celle d'un mot entendu, dans une sorte de vocalisation interne perçue par le lecteur. Comme une petite voix pégée dans le cerveau.

Depuis 1975, on a tout d'abord étudié le rôle du codage de l'information sonore dans la lecture silencieuse. Par exemple, les mots *ville* et *bille* ont la même terminaison, mais le groupe de lettres commun ne se prononce différemment. L'une des deux prononciations est dite régulière (*ville*) car elle est plus fréquente dans la langue, l'autre (*bille*) est irrégulière. Des volontaires à qui l'on a présenté sur un écran l'orthographe des groupes de lettres, devaient distinguer parmi ceux-ci les mots et les non-mots (des groupes de lettres qui n'ont aucun sens, tel *tohle*). Les résultats ont montré que temps de reconnaissance d'un mot régulier écrit, tel *ville*, est plus court que celui d'un mot irrégulier, tel *bille*. Sans un codage de l'information de type sonore, le temps de reconnaissance de tous les mots serait identique.

Un tel codage dans une lecture silencieuse s'explique par l'existence d'un dictionnaire mental, un stockage sous forme sonore des mots entendus avant que l'on n'ait appris à lire et à écrire. Ce codage phonologique renforce l'efficacité du codage visuel et du codage orthographique dans la lecture: nous lisons environ cinq mots par seconde avec très peu d'erreurs.

Puisque l'information sonore aide à la lecture silencieuse, L. Ferrand et ses collègues ont voulu savoir si la visualisation des mots et, par conséquent, l'orthographe, améliore la perception audi-

tive. Ils ont répondu par l'affirmative avec une méthode similaire aux expériences des années 1970, des volontaires entendaient des groupes de lettres dans un casque et identifiaient s'il s'agissait d'un mot ou d'un non-mot, tandis que les expérimentateurs mesuraient le temps de la décision (celle-là s'effectue par comparaison des sons entendus avec le dictionnaire mental).

Ici, les mots n'étaient plus orthographiques, mais phonologiques. Par exemple, le son *on* s'écrit *omb*, *on*, *ont*, *ong*, *or*, etc. Parmi ces différentes formes, *ond* est une écriture dominante, nommée régulière, tandis que *omb*, *or* ou *ont* sont la plus rare, nommée irrégulière. Par ailleurs, certains mots, tel *stage*, sont complètement réguliers, car le son *age* ne s'écrit que d'une seule façon. Les résultats montrent que l'on reconnaît plus rapidement les mots réguliers, tel *blond* que les mots irréguliers, tel *plomb*. Les mots complètement réguliers sont reconnus encore plus rapidement.

Ainsi, l'orthographe influence sur la perception de la parole. L'imagerie médicale a validé cette hypothèse en mettant en évidence l'interaction des aires auditives et visuelles du cerveau. Ce résultat a été confirmé par une étude de 1998 de J. Morais de l'Université libre de Bruxelles. Des personnes ne sachant ni lire ni écrire et des volontaires sachant lire et écrire devaient détecter, dans des mots parlés, des sons tel *b* dans *table*; seuls les seconds en sont capables. Selon J. Morais, dès que vous apprenez à lire et à écrire, l'alphabète modifie dans votre système de perception de la parole.

Ainsi, l'écriture et la lecture, systèmes culturels par excellence, ont une influence sur un système naturel, la parole. La prochaine fois que vous entendez «Le blond a fait un stage de billes en plomb dans une ville», ne haussez plus les épaules dédaigneusement, cette phrase est riche d'enseignements.



Un an après... tempête d'insectes ravageurs

Après la tempête qui a ravagé la France en décembre 1999, une nouvelle menace plane sur les forêts : les insectes.

Aujourd'hui, les insectes dévoreurs de bois, les xylophages règnent en maîtres dans les forêts françaises depuis les tempêtes de décembre 1999 ; ils ne cessent d'accroître leur territoire et leur appétit : après avoir pullulé dans les arbres morts et couchés, certains de ces insectes vont attaquer les arbres vivants. Après les catastrophes forestières, dont on observe directement les conséquences, un deuxième danger, invisible celui-là, menace : la pullulation de ces insectes ravageurs.

Selon Jean-François Abgrall, ingénieur de recherche au Cemagref, à Grenoble, l'importance des dernières tempêtes, avec 140 millions de mètres cubes de bois abattus, laisse présager un risque phytosanitaire, une catastrophe aussi étendue et dommageable que la tempête elle-même. La pullulation des xylophages augmenterait de 30 à 50 pour cent, le volume de résineux détruits par la tempête. D'après de précédentes études réalisées lors de tempêtes dans d'autres régions, aussi bien en France qu'en Europe, la pullulation culmine entre deux et quatre ans après la tempête ; il faut près de dix ans, lorsque les conditions climatiques sont favorables au développement des insectes, pour que la situation se normalise.

Les données collectées sont d'excellents outils pour l'estimation des risques et l'établissement des stratégies de gestion à mettre en œuvre. Après la tempête de 1982, qui avait mis au sol plus de 10 millions de mètres cubes de bois (principalement des épicéas), dans le Massif Central et dans les Alpes du Nord, 200 millions de typographes, les insectes ravageurs les plus dangereux, ont été éliminés dans les pessières (les forêts d'épicéas) des Alpes du Nord. Pour ce faire, on a utilisé des pièges contenant des phéromones, des substances volatiles qui attirent les partenaires sexuels. Néanmoins, plus de 200 000 mètres cubes d'épicéas dans cette région qui ont été détruits par ces insectes durant les six années qui suivirent la tempête de 1982.

La tempête de décembre 1999 a mis à terre 14 fois plus d'arbres que celle de 1982, détruisant une plus grande diversité d'espèces sur des régions plus

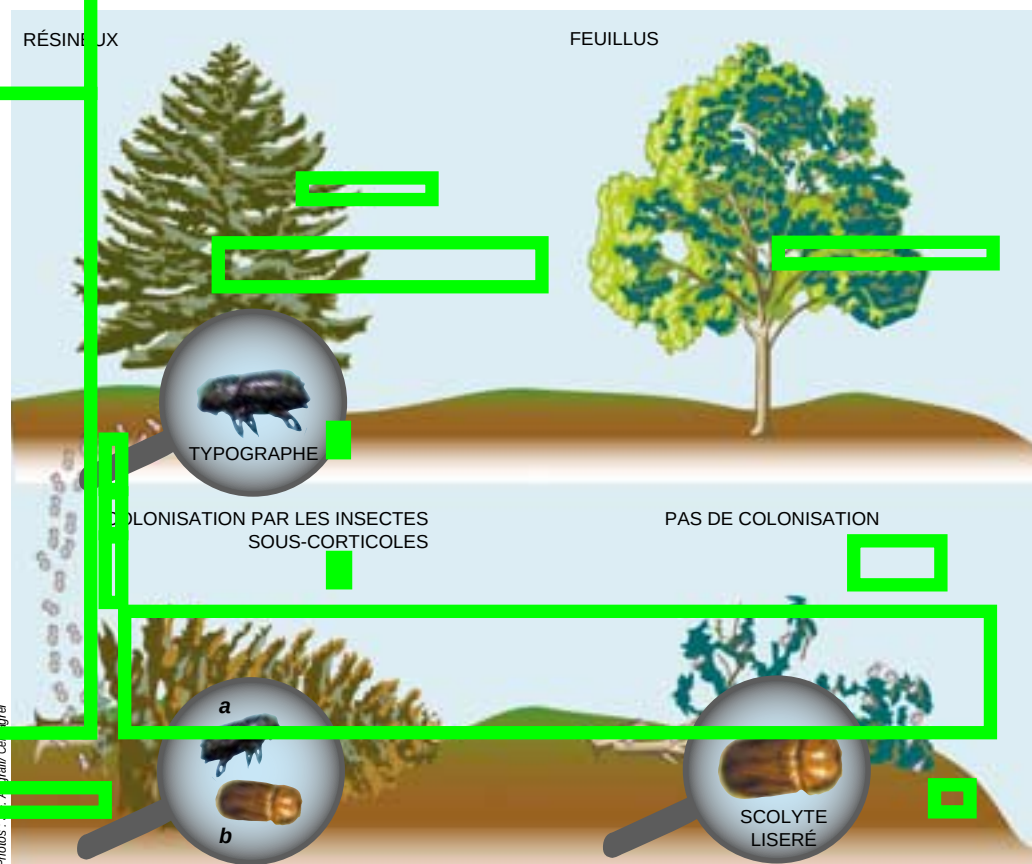
étendues. Indépendamment des conditions climatiques qui jouent un rôle important dans la prolifération, on imagine les dégâts que pourraient commettre au cours des dix prochaines années. Les participants à ce «festin» sont d'une part les insectes ravageurs de bois de cœur qui creusent de profondes galeries dans les feuillus et dans les résineux allant jusqu'au cœur de l'arbre, et, d'autre part, les sous-corticoles qui pullulent dans les résineux et qui creusent des galeries de ponte entre l'écorce et le bois ; ils provoquent ainsi la mort par destruction des tissus conducteurs de sève.

Les arbres eux-mêmes luttent contre ces prédateurs : ils produisent une résine qui imprègne leurs tissus et ils créent des zones de tissus nécrosés autour des galeries de ponte des insectes. Malheureusement, ces mécanismes consomment beaucoup d'énergie et contribuent à l'affaiblissement de l'arbre, conduisant à sa mort lorsque les mécanismes de défense sont débordés par l'afflux d'insectes. Les sous-corticoles colonisent les arbres restés sur pied de façon massive : ils émettent des phéromones spécifiques qui attirent les renforts. Au cours de l'année 2000, deux vagues d'attaque sur

les arbres gisants ont été observées, l'une fin mars par les agents de piqûre, l'autre en avril par les sous-corticoles. Ces attaques ont rendu les arbres gisants inutilisables en menuiserie.

Dès le début de l'année, les sous-corticoles se sont multipliées dans les troncs gisants, constituant ainsi un «stock» d'adultes hivernants. Le risque de pullulation vient des insectes sous-corticoles qui essaieront au cours du printemps prochain : quittant le bois mort, ils se déplaceront en masse pour attaquer les arbres restés sur pied. Les premiers problèmes phytosanitaires apparaîtront au printemps 2001.

L'extrication des arbres gisants hors de la forêt est le moyen de lutte le plus efficace contre la propagation des insectes. Cependant, deux générations de typographes se sont déjà développées dans les bois gisants. Comment gagnera-t-on cette course de vitesse avec les xylophages ? En organisant dès le printemps 2001 un abattage rapide, coordonné et efficace des bois sur pied «contaminés» ? En utilisant des phéromones qui attirent les insectes pour les piéger ? La solution se fait attendre, et pourtant la survie de la forêt épargnée par la tempête en dépend.



Les insectes sous-corticoles (a, le typographe) qui se sont développés tout au long de l'année sur les résineux gisants attaquent les résineux sur pied. Les ravageurs de bois de cœur (b, le scolyte liséré) ne colonisent pas les feuillus restés sur pied, ils creusent des galeries dans les feuillus et dans les résineux gisants, entraînant ainsi une grande dévalorisation de leur valeur marchande.

Paysages perceptifs

Des observations tomographiques du cerveau révélant que les aires cérébrales stimulées changent selon l'émotion ressentie. On a demandé à des personnes dont on enregistrerait une image cérébrale par tomographie par émission de positons de se remémorer des épisodes tristes, heureux ou angoissants qu'ils avaient vécus. Chaque émotion aurait son propre paysage perceptif. On expliquerait ainsi les variations de nos ressentiments. Simultanément, on enregistrerait aussi le rythme cardiaque des personnes qui subissaient la tomographie. On a constaté que la modification du rythme cardiaque précède l'activation de l'aire cérébrale concernée par l'émotion : la manifestation biologique de l'émotion précède sa manifestation affective.

Balles nerveuses!

Les joueurs de tennis servent leur première balle avec une telle puissance que les services gagnants sont de plus en plus nombreux et nuisent au suspens de certains matchs. Afin d'y remédier, la Fédération internationale de tennis propose l'utilisation de balles dites "lentes". D'après les différents tests et simulations, de telles balles de même masse que les balles classiques, mais dont le volume est augmenté de six pour cent, laissent 16 millisecondes de plus au joueur pour se préparer à recevoir la balle! Les scores sont les mêmes, mais les échanges sont un peu plus longs.

La douleur du nouveau-né

Contrairement à une croyance populaire, les nouveau-nés ressentent la douleur. Qui plus est, des sensations douloureuses qui résultent, par exemple, de soins intensifs ou d'explorations fonctionnelles diverses après un accouchement prématuré influent sur le comportement face à la douleur, plusieurs années après la

naissance. Des études sur l'animal ont montré que les réactions physiologiques et comportementales sont perturbées par une

leur répétée ou prolongée. Des enfants de 18 mois, nés avant terme, semblent moins sensibles à la douleur que des enfants du même âge nés à terme. Plus ces enfants ont subi de gestes traumatiques, moins ils semblent sensibles à la douleur. Des douleurs répétées influeraient sur le comportement de l'enfant, voire même de l'adulte.

Un macaque géomètre

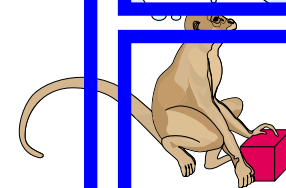
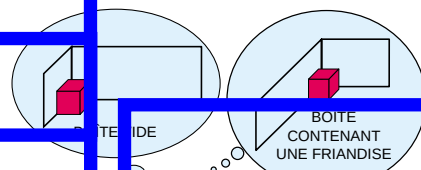
Les singes recourent à un codage de la géométrie de leur environnement pour s'orienter.

Beaucoup de nos actions quotidiennes, banales, requièrent des représentations mentales de l'environnement proche ou lointain, grâce auxquelles nous localisons précisément un objet ou un lieu à atteindre, même dissimulés à la vue, ou trop éloignés. La première étape de la construction d'une représentation spatiale, lorsque nous sommes confrontés à un environnement nouveau, consiste à sélectionner des informations, puis à les organiser. Des études récentes ont mis en évidence que l'un des modes d'organisation de l'espace repose sur le codage de la géométrie de certains de ses éléments. Intuitivement, la géométrie évoque un univers de formes, les mathématiques, et un ensemble de notions abstraites dont on peut penser qu'elles ne sont accessibles qu'aux êtres humains, et en particulier à ceux qui les ont acquises à l'école.

Toutefois, dans notre laboratoire du Centre de recherche en neurosciences cognitives du CNRS, à Marseille, nous avons montré que les singes codent également la géométrie de l'espace, et, de plus, qu'ils sont meilleurs que des enfants âgés de six ans, car ils déjouent les pièges de ce type de codage! Dans notre expérience, un singe macaque est enfermé dans une cage transparente en plexiglas, placée dans une pièce rectangulaire totalement vide à l'exception de quatre boîtes identiques posées aux quatre coins.

Dans sa cage, au milieu de la pièce, le singe observe un expérimentateur qui dissimule une friandise dans l'une de ces boîtes. Puis on recouvre la cage d'un tissu noir et on la tourne plusieurs fois sur elle-même afin de désorienter le macaque par rapport à l'emplacement de la cible. Le tissu est ensuite retiré à distance, puis la cage est ouverte également à distance, et l'animal cherche alors la friandise. Les seules informations dont il dispose sont celles fournies par la géométrie de la pièce.

Le coin où est cachée la récompense a le grand côté du rectangle à gauche et le petit à droite. Cependant, dans un rectangle, deux angles sont de ce type, géométriquement équivalents. Avec ces seules informations, le singe



Un macaque se repère dans l'espace : il sait retrouver la boîte, placée au coin d'une pièce rectangulaire où l'on a caché une friandise.

recherche la friandise dans l'une ou l'autre des deux boîtes placées dans ces deux coins, et néglige les deux autres. Il utilise la géométrie, mais n'a aucun moyen de différencier le coin où se trouve la récompense. Ainsi le singe, à l'instar du rat ou des êtres humains, utilise la forme de l'environnement pour s'orienter.

Dans une seconde expérience, nous avons ajouté sur l'une des largeurs de la pièce un grand panneau de couleur. Après avoir à nouveau repéré la boîte qui contenait la friandise, puis avoir été désorienté, le macaque se dirige sans hésitation vers la friandise cachée. Banal, diriez-vous... Oui! Toutefois, selon les travaux de Linda Hermer et Elisabeth Spelke de l'Institut de technologie du Massachusetts, l'enfant jusqu'à deux ans, lui, ignore totalement cette précieuse indication : de même que le rat, il continue à rechercher la récompense dans les deux angles géométriquement équivalents. L'adulte utilise facilement les deux types d'informations, géométriques et locales (le panneau coloré).

Ainsi l'adulte et le macaque codent simultanément et sans ambiguïté la géométrie de l'environnement et les informations locales. Les psychologues américains avaient avancé que le langage joue un rôle fondamental dans le codage des informations locales, mais nos travaux sur le singe montrent que leur théorie n'est pas la seule hypothèse possible, puisque les singes ne dépendent pas du langage. Ils ont peut-être développé une autre façon de prendre en compte les caractéristiques locales de leur environnement en raison de nécessités vitales. Par exemple, il est important pour eux de reconnaître, par sa couleur, un fruit qui est mûr. Pour autant, les macaques connaissent-ils le théorème de Pythagore?

Catherine THINUS-BLANC et Stéphane GOUTEUX, Centre de recherche en neurosciences cognitives, Marseille

La sélection robotique

Par un procédé analogue à la sélection naturelle, des robots marcheurs sont créés ex-nihilo.

À Prague, une légende veut qu'un vieux rabbin ait fabriqué un être artificiel en argile, le Golem, afin de soulager les habitants de la ville de leurs corvées quotidiennes. Les roboticiens ne font pas autre chose, sauf qu'ils remplacent l'argile par du métal ou du plastique, et comptent plus sur des moteurs électriques pour mouvoir leurs créations que sur des extraits de la Cabbale. Toutefois, les roboticiens ajoutent aujourd'hui un ingrédient jusque-là réservé à Dame Nature : la sélection naturelle.

Hod Lipson et Jordan Pollack, de l'Université Brandeis, ont «construit» sur ordinateur des robots constitués d'éléments passifs (des barres rigides et des articulations) et d'éléments actifs (des barres rétractables). Ces robots, commandés par un réseau de neurones, ont pour tâche de se déplacer en ligne droite. Ces robots assemblés (au cours d'une simulation), ont été soumis à des contraintes «évolutives» artificielles, par un procédé qui s'apparente à la sélection naturelle. À partir d'une population de robots virtuels, le

procédé de sélection choisit les plus performants, c'est-à-dire ceux qui parcourent la plus grande distance en un temps donné.

Puis, au hasard, le «programme de sélection» ajoute, retire ou modifie une structure passive de l'architecture, ou des neurones du système de commande des «individus» sélectionnés. Lorsqu'apparaît un neurone qui relie le système de commande à une barre rigide, celle-ci devient alors automatiquement une barre rétractable. Ce processus se répète jusqu'à ce qu'un neurone n'est pas relié à une barre rétractable, et plusieurs générations sont requises avant d'observer le moindre mouvement. On passe d'une «génération» de robots à la suivante lorsque le programme informatique a, premièrement, modifié au hasard les robots virtuels puis, deuxièmement, éliminé les moins performants.

D'abord modélisée sur un ordinateur, une population initiale de 200 machines a peu à peu évolué et fait apparaître des robots capables de se déplacer.

Les simulations effectuées sur les modèles retenus après 300 à 600 générations ont été fabriquées, grâce à une «imprimante en trois dimensions» qui a façonné les éléments des différents robots en plastique. L'intervention humaine est réduite à son minimum : elle consistait à monter les éléments fabriqués et à installer des moteurs électriques au niveau des barres rétractables.

Chacune des machines finales était composée en moyenne d'une vingtaine de barres. Parmi elles, les roboticiens ont eu la surprise d'obtenir des spécimens symétriques alors qu'aucune instruction n'avait imposé, ni proposé cette solution : ils concluent que la marche en ligne droite symétrique est naturellement «sélectionnée». De plus, les «vainqueurs» de cette compétition étaient robustes, c'est-à-dire que de petites modifications, telle la longueur des barres, n'influent pas sur leur capacité à marcher.

Au final, une structure en forme de tétraèdre avec une barre centrale parcourt plus de trois centimètres par cycle de mouvements. Par ailleurs, une structure en forme de flèche dotée de trois «bras» rétractables a fait preuve d'inventivité en utilisant un déphasage synchronisé de mouvements : tandis que la barre centrale se rétracte, les deux extérieures s'allongent et font glisser l'ensemble.

Ces robots autonomes qui se construisent eux-mêmes et évoluent font rêver les ingénieurs des agences spatiales qui y voient des engins d'exploration idéaux. Devant de telles applications, oublions la fin de la légende du Golem, selon laquelle la créature s'est rebellée contre les êtres humains qui l'avaient créée.

Déformation professionnelle

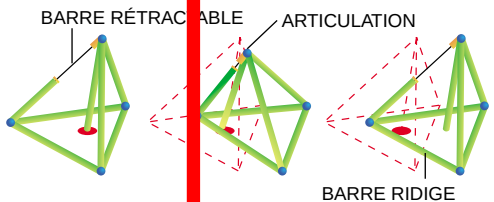
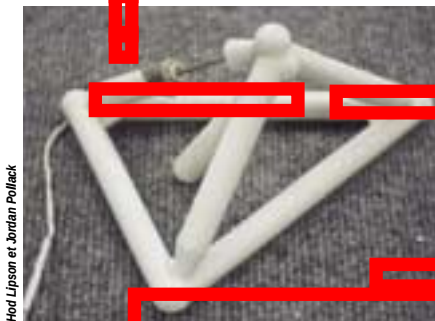
Les chanteurs d'opéra présentent des déformations de la colonne vertébrale qui résultent de la pratique intensive du chant.

Des anomalies de la forme et de la courbure du rachis cervical des chanteurs d'opéra avaient été constatées, il y a une quinzaine d'années, par Nicole Scotto di Carlo, de l'unité CNRS-ESA 6057, à l'Université de Provence. Toutefois, les données n'étaient pas suffisantes pour que l'on réponde à la question : ces anomalies sont-elles d'origine pathologique ou résultent-elles d'une déformation inhérente à la pratique du chant ? Afin de déterminer la cause de ces anomalies, N. Scotto di Carlo a entrepris une étude radiologique systématique du rachis cervical de plusieurs chanteurs professionnels : les déformations du rachis cervical observées résultent bien de la pratique intensive du chant.

Pour cette étude, 12 chanteurs professionnels (qui avaient entre 15 et 25 ans de pratique vocale), 12 chanteurs débutants (5 ans de pratique vocale) et 12 personnes qui ne chantent pas et constituent le groupe témoin ont été sélectionnés. Aucun ne présentait de pathologie rachidienne ni de trouble de la statique, aucun ne pratiquait d'activité sportive mettant la tête en mouvement. Des radiographies de la tête et du cou ont été réalisées sur des chanteurs au repos et pendant qu'ils émettaient des voyelles en voix parlée et en voix chantée dans les registres grave, médium et aigu. Les sujets témoins ont été radiographiés au repos et en cours de bâillement, quand l'ouverture de leur bouche est proche de celle des chanteurs pendant l'émission de l'aigu.

L'analyse des radiographies a révélé des anomalies dans la forme du rachis cervical pour tous les chanteurs professionnels : plus de la moitié présente une rectitude cervicale, c'est-à-dire un rachis cervical rectiligne (voir la figure 1c), et 42 pour cent ont une hypocyphose, c'est-à-dire une légère courbure à concavité antérieure (voir la figure 2a). Seulement 16 pour cent des chanteurs débutants ont une rectitude cervicale.

De surcroît, cette étude a révélé l'existence de postures spécifiques en fonction des registres. Chez les chanteurs professionnels, lors du passage du grave à l'aigu, l'ouverture buccale augmente, la tête se relève, le rachis cervical recule et la courbure cervicale s'inverse (voir la



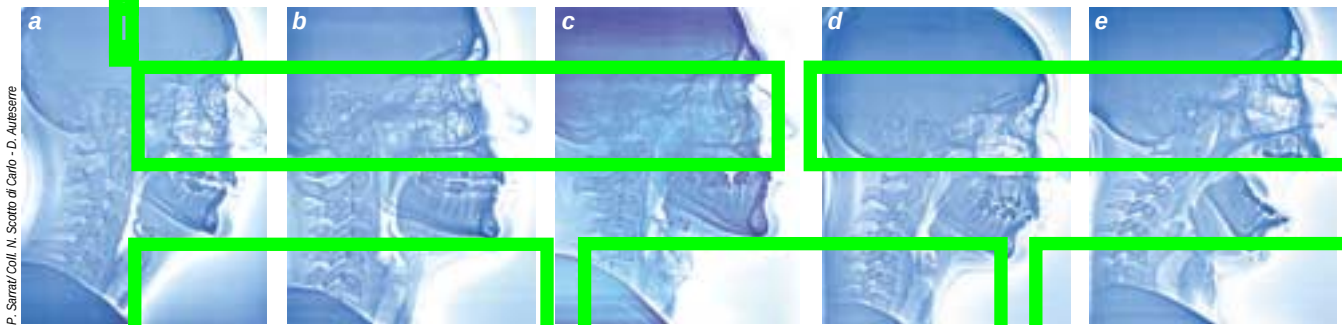
L'évolution d'un robot sélectionné artificiellement appliqué à des robots dont la tâche est de marcher en ligne droite aboutit, après 300 à 600 générations, à des «spécimens» qu'une «imprimante en trois dimensions» a fabriqués. Certains ont la forme d'un tétraèdre (au milieu) muni d'une tige centrale qui sert d'appui lorsqu'une barre se rétracte (le principe de la «marche du robot» a été schématisé en bas). Ce robot avance de deux à trois centimètres par cycle de mouvements et chaque cycle dure quelques secondes.

figure 1e). En revanche, chez les chanteurs débutants, les modifications sont moins prononcées et, dans l'aigu, la courbure du rachis cervical ne s'inverse pas. Lors du passage du grave à l'aigu, le relèvement de la tête et le recul du rachis cervical sont imposés l'un et l'autre par l'ouverture buccale. Chez les chanteurs professionnels, le recul du rachis cervical dans l'aigu s'accompagne d'une inversion de courbure laquelle dégage le pharynx qui se dilate alors suffisamment pour permettre la bascule antérieure

du larynx indispensable à l'émission de l'aigu.

Ainsi, la mobilité du rachis cervical et les déformations qu'il subit pendant le chant semblent être la cause des anomalies de la courbure observée chez tous les chanteurs examinés. Ces anomalies n'étant associées à aucun trouble fonctionnel, on admet que la posture observée chez les artistes lyriques au repos résulte des déformations cervicales subies tout au long des années de pratique intensive du chant.

Dans le chant, les déplacements antéro-postérieurs du rachis peuvent atteindre cinq à six centimètres (à la septième vertèbre cervicale). Cette extrême mobilité expliquerait que toute lésion de cette région du rachis (y compris un simple torticolis) qui réduirait sa mobilité perturberait la voix chantée et notamment l'émission de l'aigu. Pour ne pas mettre leur voix en danger, les chanteurs professionnels doivent éviter toute manipulation ou intervention chirurgicale visant à corriger les «anomalies» de leur rachis cervical.



P. Sarraf/ Coll. N. Scritto di Carlo - D. Aulesserre

La comparaison de radiographies faites sur des personnes qui ne chantent pas (a) et sur des chanteurs d'opéra professionnels a révélé plusieurs anomalies dans la courbure du rachis cervical. Ainsi, on observe dans la moitié des cas une rectitude cervicale (b), parfois une hypolordose (c), c'est-à-dire une légère courbure à convexité postérieure de la colonne vertébrale, ou encore une hypociphose, c'est-à-dire une légère courbure à convexité antérieure (d). Chez certains chanteurs, pendant l'émission de l'aigu, on constate une cyphose prononcée (e).

Vache folle : révision à la baisse des risques

Nouvelles estimations du nombre des personnes qui mourront de la maladie transmise par l'agent de la vache folle.

Depuis son déclenchement en Angleterre au début des années 1990, l'épidémie de la vache folle a tué 180 000 bovins et provoqué l'apparition d'une nouvelle maladie humaine. Cette dernière ressemble en tout point à la maladie de Creutzfeldt-Jakob, décrite au début du xx^e siècle par Hans Creutzfeldt et Alfons Jakob. Maladie rare, cette dernière touche une personne sur un million. Dans huit pour cent des cas, elle apparaît sous une forme dite sporadique, parce qu'elle survient «au hasard» chez des personnes, qui ont en moyenne 65 ans. En revanche, la nouvelle forme de la maladie de Creutzfeldt-Jakob, frappe surtout les jeunes : plus de la moitié des malades ont moins de 35 ans. Aujourd'hui, les biostatisticiens Roy Anderson et ses collègues de l'Université d'Oxford, révisent à la baisse les conclusions, inquiétantes, qu'ils tiraient dans leurs estimations précédentes. Tenant compte des cas dénombrés jus-

qu'en 1999, ils estiment qu'il y aura vraisemblablement entre 63 et 136 000 victimes en tout, certes, mais c'est moins que les 500 000 victimes britanniques de la maladie de Creutzfeldt-Jakob évoquées dans le passé.

Malgré les nouvelles prévisions, le bilan du premier semestre de l'année 2000 est déjà de 24 décès, dont cinq dans un même village du Leicestershire. Toutefois, selon Annick Alperovitch, qui dirige l'Unité de génétique et d'épidémiologie des maladies neuro-dégénératives à l'Hôpital de la Pitié-Salpêtrière, le nombre de cas augmente parce que l'épidémie ne fait que commencer. Comme la durée moyenne d'incubation de la maladie de Creutzfeldt-Jakob se mesure en dizaines d'années, le nombre annuel de victimes risque encore d'augmenter pendant quelques années.

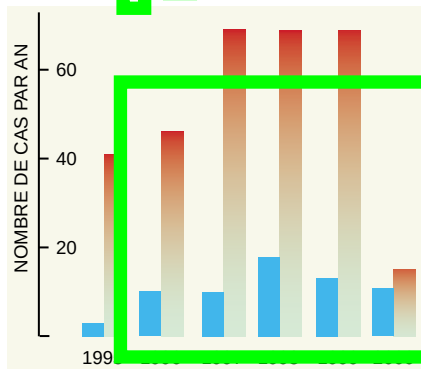
Pour calculer le nombre probable de victimes de la maladie, R. Anderson et ses collègues évaluent le risque, pour une personne, de tomber malade à un instant et à un âge donné. Ce risque dépend de trois facteurs principaux : la susceptibilité génétique des individus à la maladie, le risque qu'un bovin contamine une ou plusieurs personnes, et la répartition statistique des durées d'incubation dans la population. Le facteur génétique est le mieux connu. On sait que toutes les personnes qui ont été atteintes par cette nouvelle maladie ont une caractéristique

génétique commune, que partagent environ 40 pour cent des britanniques. Dans leur étude, les biostatisticiens d'Oxford considèrent donc que seuls les individus porteurs de cette caractéristique sont menacés par la nouvelle maladie de Creutzfeldt-Jakob.

On évalue moins bien la capacité de contamination d'un bovin malade introduit dans la chaîne alimentaire. Les biostatisticiens anglais supposent que, si une vache malade est consommée, elle ne contamine que deux personnes au plus. Inspiré par les dernières observations, cette hypothèse est une révision optimiste des précédentes estimations.

Enfin, on ignore totalement la répartition des durées d'incubation dans la population britannique, mais on peut la simuler à l'aide de lois statistiques, qui toutes ont pour paramètre principal la durée moyenne d'incubation. Ainsi armée d'une méthode de calcul du risque de tomber malade, l'équipe de R. Anderson en a fait varier tous les paramètres ; elle étudie ainsi des millions de scénarios, ne retenant que ceux qui conduisent à des prévisions reproduisant les statistiques de début d'épidémie.

Plus de cinq millions de configurations de paramètres ont été envisagées. Comment mettre de l'ordre ? Les biostatisticiens ont défini quatre niveaux de mortalité hypothétiques pour l'année 2000, puis pour la période 2001-2002. Il ont ensuite réparti leurs prévisions



Le nombre de victimes britanniques de la nouvelle forme de la maladie de Creutzfeldt-Jakob (en bleu) reste notablement inférieur à celui des victimes de la forme «classique», dite sporadique, de cette maladie (en rouge), qui fait chaque année et dans tous les pays industrialisés, quelques dizaines de victimes, généralement des personnes âgées. Pour l'année 2000, les mortalités moyennes correspondent aux chiffres enregistrés à la fin du mois d'août.

suivant ces quatre niveaux, et les ont triées en fonction des durées moyennes d'incubation. Finalement, ils ont isolé, dans chaque catégorie, la prévision la

plus pessimiste et la plus optimiste. Ils ont obtenu en effet un tableau qui indique dans quelle fenêtre se trouvera le nombre de décès futurs en fonction du nombre de cas qui se déclareront en 2000, puis en 2001-2002. Les limites basse (63) et haute de l'épidémie (136 000) sont les valeurs extrêmes ; ainsi, la prévision la plus défavorable – 136 000 victimes – a été trouvée pour une durée moyenne d'incubation supérieure à 60 ans.

Selon A. Alperovitch, la durée moyenne d'incubation est certainement inférieure à 60 ans ; elle serait plutôt de l'ordre de 20 ans. Avec une telle valeur de la durée moyenne d'incubation, le nombre probable de victimes britanniques du nouveau variant de la maladie de Creutzfeldt-Jakob atteindrait 3 000. Soit en moyenne 100 cas par an d'ici à 2030. Nous sommes encore loin de ce chiffre puisque le nombre de décès était égal à 14 en 1999 et pourrait être de 30 en 2000.

Le terrier du poulpe

Le poulpe cache ses provisions de coquillages au fond de l'abri qu'il creuse dans le sable.

La moitié des prises mondiales de poulpe s'effectue au large des côtes d'Afrique de l'Ouest et du Nord. Cette pêche s'est développée dans les années 1960 au large du Sahara, puis dans les eaux sénégalaises, après un spectaculaire accroissement de l'abondance des poulpes en 1986. Selon les années, la prise de poulpe varie de 2 000 à plus de 15 000 tonnes au Sénégal. Pourquoi un tel écart ? Afin d'évaluer la pérennité de cette ressource, une étude récente a précisé les caractéristiques biologiques de l'espèce *Octopus Vulgaris* : le poulpe est casanier et le renouvellement annuel des populations met l'espèce hors de danger.

Alain Claverivière et ses collègues, de l'Institut de recherche pour le développement, et leurs collègues africains ont étudié des poulpes vivant en bassin d'étude et d'autres vivant en pleine mer. Parmi ces derniers, ils en ont marqué 6 000, dont 1 200 ont été repêchés. Au Sénégal, les poulpes vivent sur tous les types de fonds, mais préfèrent les sédiments de sable fin. Le poulpe creuse des abris dans ces fonds meubles pour se protéger des prédateurs, dont les principaux sont les raies, les daurades, les murènes et les mérus. Ces terriers, dont la profondeur et le diamètre varient avec la taille du poulpe, sont de deux types. Dans le premier (une che-

minée cylindrique et lisse) le poulpe en danger s'abrite en se retournant pour présenter ses tentacules qui cachent tous les débris de coquilles. Le second type, d'une forme plus évasée, est rempli de vieilles coquilles sous lesquelles le poulpe s'enfonce rapidement. Le poulpe creuse lui-même son terrier qu'il ne quitte que pour s'alimenter et se reproduire. Cet habitat particulier expliquerait l'abondance de poulpes sur les fonds meubles de l'Afrique de l'Ouest.

On a découvert que les poulpes font des provisions de coquillages vivants, à l'intérieur ou à la périphérie de leurs terriers. Le suivi des poulpes marqués a révélé une durée de vie de l'ordre d'une année. Au terme de leur croissance, les femelles atteignent au maximum un poids de cinq kilogrammes et les mâles de six à huit kilogrammes. La mort des femelles survient peu après la éclosion des œufs qu'elles couvent pendant deux mois environ. Puis, elles jeûnent, maigrissent et meurent d'épuisement. Les mâles maigrissent comme les femelles qu'ils ont fécondées, et meurent presque en même temps qu'elles.

Des explosions démographiques surviennent lorsque les larves trouvent un environnement favorable à leur développement, tel que les fonds de sable riches en éléments nutritifs déclenchées par les algues. Ce phénomène, observé en 1999 au Sénégal, a précédé un énorme accroissement des populations de poulpes l'été suivant. Malgré la surpêche de 1999, le renouvellement rapide des poulpes dans les eaux sénégalaises les protège de la surexploitation. En revanche, leurs prédateurs se raréfient.

Votre futur compagnon numérique

FIONA HARVEY

Pour créer un réseau Internet sans fil, des ingénieurs construisent des réseaux qui traitent rapidement des quantités considérables de données et des dispositifs de poche qui exploitent les ressources du réseau Internet actuel.

En 2000, en voyage d'affaires à New York, vous avez quelques heures pour flâner dans la Cinquième Avenue. Soudain, votre ordinateur portable vibre ; vous sortez de votre poche l'appareil gros comme une calculatrice actuelle et, sur l'écran vidéo, vous découvrez que les actions de la Société *Captrôme* perdent de la valeur. Vous avez créé cette entreprise de nez électronique en 2000, mais, fin 2004, votre partenaire

financier a voulu qu'elle entre au second marché. Dans le microphone de votre portable, vous criez : « Appelle Viviane ! » Devant vous, des passants se retournent, tandis que votre secrétaire vous révèle que votre ancien partenaire financier – avec qui vous êtes fâché – vient de placer sur le marché 30 pour cent des actions de la société, qu'il possédait encore. Debout sur le trottoir, vous criez : « Accès finances personnelles ! » Votre compagnon numérique se connecte immédiatement à votre site boursier favori. Pendant que les passants vous regardent, vous achetez la moitié des actions tout juste mises sur le marché.

Vous devez rentrer d'urgence à Paris pour concrétiser l'achat. Vous demandez au compagnon numérique de se connecter sur le service de réservation d'une compagnie aérienne et vous achetez un billet. À l'aide de l'appareil, vous informez aussi Viviane et votre femme du changement de programme. Enfin, vous demandez à votre portable de vous indiquer le bar le plus proche. Après cette alerte, un remue-ménage s'impose.

Cet scénario paraît fantaisiste, mais la plupart des grandes entreprises de télécommunications dépensent des milliards afin de le rendre plausible. Pourquoi ? Parce qu'elles ont observé le succès considérable des téléphones cellulaires : tous ces utilisateurs de téléphones sans fil pourraient devenir des consomma-

teurs à temps complet. En France, plus de 24 millions de personnes possèdent déjà un portable, soit 37 pour cent de la population. Dans plusieurs pays occidentaux, une personne sur deux est équipée. En tout, 117 millions d'Européens utilisent la téléphonie mobile. C'est en Finlande que la frénésie de la téléphonie mobile culmine : 71 pour cent des personnes utilisent un *kännykkä*. Outre Atlantique, 34 pour cent des Américains téléphonent quotidiennement de leur *mobile phone*. Au Japon, 45 pour cent des Nippons abusent du *Keitai Denwa*, 58 pour cent des Hongkongais font de même, etc. Tous ces abonnés utilisent déjà des services que les opérateurs ont ajoutés à la simple communication. Toutefois, l'exercice est souvent inefficace : le paiement sécurisé reste difficile, le recours à un opérateur est impératif, les temps d'attente sont longs, etc. Aujourd'hui, les utilisateurs de téléphone portable n'utilisent des prestations par téléphone qu'en cas de nécessité. La situation changerait si les accès étaient automatiques et simples.

Toutefois, ils ne le deviendront que si les données arrivent assez rapidement dans les portables. Les services pourront alors y être proposés sur une multitude de minipages. Pour de nombreux experts commerciaux, l'évolution ultérieure du marché est assurée : obnubilés par l'idée de gagner du temps, la « ménagère de moins de 50 ans » et beaucoup d'autres



À gauche et à droite : avec l'aimable autorisation de Lernout & Hauspie et Nokia

Les gadgets électroniques de demain auront probablement une tout autre apparence que ceux d'aujourd'hui. La Société belge *Lernout & Hauspie* a mis au point un prototype (à gauche) d'appareil portable, que les utilisateurs commandent à la voix. Ils lui dictent aussi leurs courriers électroniques. La Société finlandaise *Nokia*, spécialiste de téléphonie, a développé un appareil capable d'afficher des images et des films vidéo (à droite).

consommateurs feront leurs achats sur leurs terminaux portables. Des entreprises se chargeront alors de les aider à bien utiliser le temps ainsi sauvé...

Les bases de la prochaine révolution de la téléphonie sont en construction. L'Europe est plus avancée que les États-Unis et l'Asie. En Finlande, par exemple, les abonnés utilisent déjà leur téléphone pour envoyer des messages à leurs amis, pour payer des factures, pour connaître l'état de la circulation routière et pour payer leurs consommations dans un café d'Helsinki : ils composent seulement les numéros appropriés. L'avance européenne résulte d'une meilleure harmonisation des techniques. Le GSM est adopté dans tous les pays d'Europe. Au Japon, tous les adolescents sont fous de l'option *iMode*, qui assure la connexion au réseau Internet. Plus de 12 millions de Japonais l'utilisent. Les sociétés américaines ont été plus lentes à offrir ce genre de services parce qu'aucune norme standard n'est répandue aux États-Unis. Pourtant, les opérateurs américains n'ont pas hésité à lancer des campagnes publicitaires agressives afin de convaincre leurs clients d'acheter des téléphones cellulaires reliés au réseau Internet. Parallèlement, des fabricants de matériels informatiques, notamment les constructeurs d'ordinateurs de poche, exploitent le nouveau marché.

Le *Palmer V II*, par exemple, accède au réseau Internet par les airs, grâce à une mini-antenne. Le *Jornada* intègre un modem, lequel peut être branché sur un portable, etc.

Aujourd'hui, les communications effectuées par ces appareils sont terriblement lentes : le débit est inférieur à seulement dix kilobits par seconde en moyenne, soit une transmission cinq fois plus lente qu'avec un micro-ordinateur connecté par un fil. Avec ces débits, on peut effectuer des tâches simples, tel l'achèvement d'un message, mais la navigation sur Internet est pénible. Le téléchargement de films vidéo ou de musique est impossible ! À cette faiblesse technique s'ajoute celle de l'infrastructure. Aujourd'hui, les réseaux ne transmettent des données par les airs que dans les grandes agglomérations.

Les progrès techniques et les investissements considérables devraient pallier ces défauts. Les premiers téléphones cellulaires étaient des appareils analogiques : ils transmettaient un signal électrique proportionnel à l'intensité sonore de la voix. Les téléphones de deuxième génération – ou « téléphones 2G » – sont tous numériques. Les séquences de 0 et de 1 qu'ils envoient ou reçoivent codent la voix, et ils peuvent en principe coder n'importe quel type d'informations. Les sociétés de télé-

communications mettent aujourd'hui au point des réseaux sans fil de troisième génération (3G), qui transmettent rapidement les données. Les opérateurs annoncent des débits de deux mégabits par seconde : on pourra télécharger des morceaux de musique et des films vidéo.

Le gouvernement britannique a déjà réservé certaines bandes du spectre électromagnétique à la téléphonie mobile de troisième génération : la mise aux enchères des licences correspondantes lui a rapporté presque 250 milliards de francs. En Allemagne, où les mêmes fréquences viennent d'être mise aux enchères, les montants atteints devraient être comparables. En France, le gouvernement a décidé de vendre les licences à un prix fixe aux sociétés qu'il choisira. Le gouvernement américain a des projets similaires. En Europe, le Royaume-Uni devrait être le premier pays équipé de réseaux 3G. De grands opérateurs, tels *British Telecommunications*, *Vodafone* et *France Telecom*, promettent que leurs réseaux 3G seront opérationnels dès 2002. Les services de données à haut débit ne seront peut-être pas actifs aussitôt, et les populations rurales seront équipées bien après les citadins : en 2010, semble-t-il.

Impatients de capter le marché, plusieurs opérateurs s'apprêtent à proposer des solutions intermédiaires : les



«réseaux 5G». Aux États-Unis, par exemple, la Société *Sprint PCS* projette de pousser la vitesse de pointe de ses applications Internet à 144 kilobits par seconde grâce à la «commutation de paquets». Jusqu'à présent, presque tous les réseaux de téléphonie mobile utilisaient la commutation de circuits : un canal de communication, c'est-à-dire une bande spécifique de fréquence, est réservé à chaque utilisateur qui téléphone ; quand tous les canaux sont occupés, téléphoner devient impossible. La commutation par paquets évite cette saturation : le signal est découpé en paquets de données qui sont étiquetés avec l'adresse de destination. Puis tous ces paquets sont envoyés au réseau, qui les achemine par les chemins les plus disponibles. À la réception, un ordinateur les met dans le bon ordre, puis les livre au circuit d'écoute ou à l'écran. Les

débits maximaux de la commutation par paquets sont bien supérieurs à ceux de la commutation de circuits. La commutation par paquets est comparable à la gestion ordinaire du trafic routier : tous les véhicules sont libres d'emprunter n'importe quelle chaussée, n'importe quand. En revanche, la commutation de circuits se compare à la gestion extraordinaire du trafic mise en place par la gendarmerie lors de la visite d'un président étranger ; le véhicule de ce dernier emprunte une partie du réseau libérée pour lui. Pendant ce temps, le reste des véhicules n'a pas le droit de rouler sur l'itinéraire du président. Si nous étions tous des présidents, la commutation par paquets serait bien moins intense !

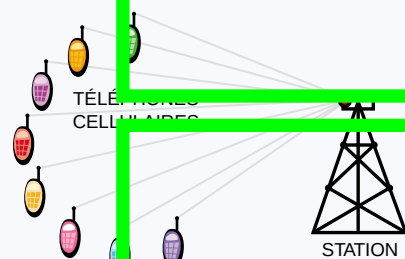
La commutation par paquets a cependant un inconvénient majeur : la

transmission est moins fiable, la collecte et le rassemblement des paquets de données prend du temps. Dans la rue Sleidan, à Strasbourg, quand il pleut, ou devant le supermarché de Suippes, en Champagne, quand le réseau est encombré, le paquet n° 735 n'atteint parfois pas le récepteur ou arrive trop tard. Les systèmes de transmission de données attendent jusqu'à l'arrivée du paquet qui manque et demande une retransmission s'il tarde. Les systèmes spécialement conçus sur la voix, en revanche, n'attendent pas longtemps ; à la fin d'un court délai de latence, ils réassemblent le signal avec des vides. Les mots sont alors fléchés, ce qui rend la conversation plutôt fatigante. Les opérateurs ont donc cherché à développer des méthodes pour minimiser le problème. La suppression des «paquets de silence» en est une : environ 30 pour

La télécommunication sans fil

De nombreux acronymes sont utilisés par l'industrie des télécommunications sans fil. Ils forment trois catégories qui cor-

respondent à trois stades successifs du développement des réseaux aériens :



Réseau cellulaire : pour assurer une bonne couverture d'un grand territoire, les opérateurs de réseaux téléphoniques le divisent en un grand nombre de zones plus petites nommées cellules. Chacune est équipée d'une station de base,

AMPS : est l'acronyme de *Advanced Mobile Phone System*, soit «système perfectionné de téléphonie portable», désigne une fréquence de réseau cellulaire analogique. Ce réseau utilise la bande voisine de 800 mégahertz. Il sera bientôt remplacé par les réseaux FDMA.



avec une forêt d'antennes qui émettent vers tous les «téléphones cellulaires» présents dans la cellule.

Bande passante : dans les télécommunications, c'est la fréquence utilisée pour la transmission d'un signal modulé en fréquence (l'antenne émet une onde de fréquence déterminée, et le signal est codé par des variations de la fréquence autour de l'onde «porteuse»). En physique, c'est aussi la gamme de fréquences qu'un système électronique transmet (un filtre, un amplificateur, un réseau, etc.). La bande passante se mesure en nombre de cycles par seconde, c'est-à-dire en hertz.

Débit de transmission : c'est le nombre de bits transmis par un système électronique par seconde.

Première génération : les systèmes analogiques (début des années 1980 jusqu'à aujourd'hui)

Encore largement utilisés, les téléphones cellulaires apparus au début des années 1980 captent et émettent des signaux analogiques. Ils transforment les sons en variations électriques dont la fréquence dépend de l'intensité sonore, puis ils émettent ce signal électrique par un canal réservé à la conversation. À l'autre bout de cette ligne, le signal est transformé en signal sonore par le petit haut-parleur de l'appareil.

FDMA : est l'acronyme de *Frequency Division Multiple Access*, soit «accès multiples par partage de la fréquence», désigne une façon d'allouer automatiquement

des fréquences aux utilisateurs, de sorte qu'un grand nombre de personnes puissent téléphoner par l'intermédiaire d'une même station. Un réseau local AMPS utilise, par exemple, 832 canaux différents, placés à 30 kilohertz les uns des autres sur l'échelle des fréquences. Le système FDMA peut être employé en même temps que les systèmes CDMA ou TDMA.

GRÂCE AU SYSTÈME FDMA, un grand nombre de téléphones cellulaires peuvent envoyer simultanément leurs signaux à une même station (à gauche). Comme le montre la figure ci-dessus, chacun des utilisateurs se voit assigner une fréquence différente. Dans la figure, chacune de ces fréquences est symbolisée par une bande de couleur différente. Les bandes de couleur sont pas utilisés en permanence.

Deuxième génération : les premiers systèmes numériques (début des années 1980 jusqu'à aujourd'hui)

CDMA : cet acronyme de *Code Division Multiple Access*, soit «accès multiples par partage de code», désigne une technique de partage d'une même fréquence par les utilisateurs. Avec ce système, le signal

cent d'une conversation téléphonique est formé de silences ; pourquoi les transmettre ?

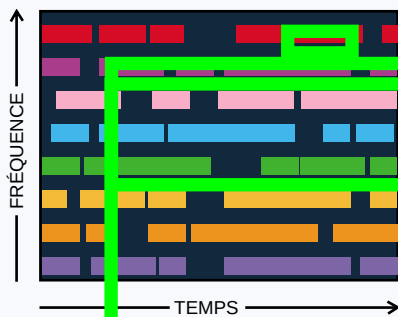
Les supertéléphones arrivent

Si un téléphone reçoit des données, de la musique, des images, pourquoi ressemblerait-il aux téléphones actuels, et si l'on veut perfectionner l'ergonomie des «compagnons numériques», quelle apparence doit-on donner à ces appareils ? Pour les opérateurs, la question est essentielle. Quand les réseaux de troisième génération seront construits, les clients ne les emploieront que si les terminaux sont bien pensés. Doit-on conserver la forme du mini-ordinateur ? Ou celle du téléphone ? De l'agenda électronique ? Apparemment, chaque public préfère un appareil spécifique : l'ado-

lescent japonais devra avoir son terminal en forme de cœur fluorescent, l'ingénieur allemand son solide appareil de ceinture, l'universitaire français son bureau de poche, le boursier sa minisalle des marchés, etc. Ces machines devront être polyvalentes, adaptables, éventuellement modifiables en fonction de la mode. Aussi les fabricants explorent-ils les modèles et des formes variés. Les principales sociétés, telles *Nokia*, *Ericsson* et *Motorola*, créent des «supertéléphones 3G» qui ne ressembleront à aucun des téléphones cellulaires actuels. Ils auront des écrans couleur de haute résolution, de plusieurs dizaines de centimètres carrés, pour la présentation de graphiques et de vidéos. Certains seront équipés d'un clavier et d'une souris miniaturisée, mais la plupart s'utiliseront à l'aide d'écrans sensibles et de styles, à la façon des machines *Palm*

ou d'autres ordinateurs de poche actuels.

Proches des baladeurs par le format, ces terminaux du futur devraient en assurer la fonction. Le téléchargement de musique est une application étudiée par plusieurs sociétés. Elle se fera sans doute sous le format déjà bien connu MP3, ou sous le format qui le remplacera. Robert Madge, fondateur de la Société *Madge Networks*, pense que ces applications seront les premières utilisées. Aujourd'hui, ceux qui ont déjà téléchargé de la musique sur Internet savent que de la mise au point reste à faire avant d'arriver aux visions de R. Madge. Cependant, plusieurs solutions sont envisagées : par exemple, l'utilisateur pourrait télécharger les morceaux musicaux commandes pendant la nuit. Comprimés, ces fichiers seraient conservés sur des cartes mémoires, telle la *Memory Stick* de la Société Sony.



LE SYSTÈME CDMA divise le signal en «jetons de communication» (petits paquets en couleurs dans la figure), qui sont ensuite transmis par les nombreux canaux d'une certaine bande de fréquence, puis réassemblés.

temps», désigne une technique de partage du temps de communication par les utilisateurs. Les téléphones de ces derniers disposent de façon périodique d'une certaine durée de communication dans le canal (fréquence) qui leur a été attribué. Les appareils reçoivent tous les signaux, mais ils peuvent ne conserver que ceux qui leur sont destinés, parce qu'ils «savent quand écouter».

GSM : cet acronyme de *Global System for Mobile Communication*, soit «système global de téléphones mobiles», désigne le système de télécommunication sans fil utilisé en Europe et au Japon. Il garantit la compatibilité des réseaux ; ainsi, un «téléphone GSM» français peut être employé en Allemagne, et réciproquement. La norme GSM est fondée sur le système de division du temps de communication TDMA. Les systèmes GSM utilisent la bande de fréquence des 900 mégahertz.

PCS : cet acronyme de *Personal Communication Service*, soit «service de communication personnelle», désigne un type de réseau sans fil nord-américain, qui utilise la bande de fréquence de 1 900 mégahertz. La variante *Sprint PCS*, par exemple, utilise le système CDMA d'étiquetage par code de morceaux de signal (jetons).

téléphonique est divisé en un grand nombre de «jetons», tous étiquetés avec le code de l'appareil portable dont ils proviennent. Le système CDMA peut être employé en même temps que les systèmes FDMA ou TDMA.

TDMA : cet acronyme de *Time Division Multiple Access*, soit «accès multiples par partage du

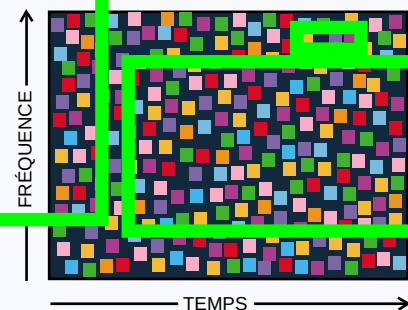
Troisième génération : les systèmes numériques à large bande (devaient apparaître au début des années 2000)

Les entreprises de télécommunications veulent modifier bientôt leurs réseaux afin d'atteindre des débits de dix kilobits par seconde.

EDGE : cet acronyme de *Enhanced Data Rates for Global Evolution*, soit «taux de données améliorées pour l'évolution globale», a été composé parce qu'il suggère l'idée d'être «en pointe». Il désigne une forme rendue plus rapide du système TDMA, c'est-à-dire de partage des temps de communication. Le taux de transmission du système EDGE devrait dépasser 384 kilobits par seconde. Le système EDGE va être employé pour reconditionner les systèmes GSM en Europe et AT&T en Amérique du Nord.

GPRS : cet acronyme de *General Packet Radio Service*, soit «service général de paquets radio», désigne l'un des améliorations prévues du système GSM. Elle consiste à introduire la commutation par paquets. Au lieu d'envoyer des données à l'aide de circuits dédiés, un système par commutation de paquets divise l'information en paquets susceptibles d'être transmis par n'importe quel des canaux disponibles.

W-CDMA : cet acronyme de *Wideband Code Division Multiple Access*, soit «accès multiples à large bande par partage de code», désigne une forme améliorée de système CDMA. Le taux de transmission peut y culminer à deux mégabits par seconde. W-CDMA répartit les paquets d'informations qu'il prépare (jetons de communication) sur une bande de fréquence beaucoup plus large que ne le fait CDMA.



GRÂCE AU SYSTÈME TDMA, un grand nombre de téléphones cellulaires peuvent se partager de façon efficace le temps de communication sur un même canal (fréquence). Chaque canal est mis à disposition d'un utilisateur pendant une durée fixe, à intervalles réguliers (bandes de couleurs dans la figure ci-dessus).

Bluetooth, ou comment se débarrasser des câbles

Enfin un moyen crédible de se débarrasser des câbles. Harald «Dent bleue» est le roi qui unifia la Norvège au Danemark, il y a 1 000 ans. Le nom de l'impitoyable Viking fait aujourd'hui l'unité chez tous ceux qui attendent avec impatience d'être délivrés des câbles d'ordinateurs et autres attrape-poussières qui encombreront inutilement nos bureaux. Ce réseau radio, *Bluetooth*, de faible portée, fut initialement inventé pour la transmission entre les chaînes stéréo et les casques sans fils. Très rapidement, il a été adopté pour des applications électroniques variées.

Même si les premiers appareils dotés de ce système arrivent à peine sur le marché, plus de 1 000 constructeurs électroniques ont déjà adopté *Bluetooth*, qui est, de facto, devenu la norme pour la communication sans fils. On le trouve dans des ordinateurs portables, des téléphones cellulaires, des jouets, des réfrigérateurs, etc.

Un réseau radio *Bluetooth* a une portée maximale d'environ dix mètres. Il fonctionne par micro-ondes, c'est-à-dire dans la bande de fréquence ISM réservée aux applications industrielles, médicales et scientifiques (2,4 à 2,5 gigahertz). Cette bande de fréquence a l'avantage d'être utilisable sans autorisation de l'administration. Afin d'éviter les interférences avec d'autres types

d'appareils qui opèrent dans la même bande, le réseau *Bluetooth* varie sa fréquence 1 600 fois par seconde.

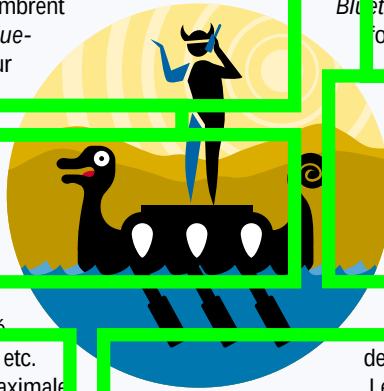
Avec *Bluetooth*, les câbles sont inutiles, et l'installation de réseaux domestiques ou de bureau sera beaucoup plus facile.

Bluetooth facilitera aussi l'échange automatique d'informations entre appareils proches. Les hommes d'affaires pourraient ainsi échanger leurs coordonnées en rapprochant, puis en activant leurs

organiseurs. Ces derniers appareils deviendraient aussi beaucoup plus faciles à accorder avec le système d'information de l'entreprise. À peine leurs propriétaires présents dans les locaux de l'entreprise, une procédure automatique les «synchroniserait».

Une autre application serait l'identification permanente de personnes à l'intérieur des usines des bureaux ou des villes...

Les constructeurs électroniques estiment que *Bluetooth* sera très vite omniprésent. La vitesse avec laquelle l'industrie l'adoptera dépendra toutefois de la vitesse de renouvellement des appareils portables. Rapide pour les téléphones, les agendas électroniques (organiseurs), etc., il est moins pour les réfrigérateurs ou les machines à laver.



X-Plane

Les constructeurs de téléphones pensent aussi à tous ceux qui préfèrent l'écrit : les jeux ou la lecture d'informations sont difficiles sur un écran minuscule. Aussi, les constructeurs électroniques imaginent des appareils munis d'écrans plus grands, comme les livres électroniques qui apparaissent en France depuis un ou deux ans. Ces sortes d'adolescents ont la surface d'une page et une mémoire assez vaste pour stocker des dizaines d'ouvrages. Des constructeurs qui envisagent un développement du téléchargement de films ont mis au point des systèmes équipés de minuscules caméras vidéo ou de projecteurs. Ils permettront soit la réalisation des films dans toutes les circonstances, soit une projection après un téléchargement.

Parler, écouter et regarder son écran

Avec tous les prototypes, l'utilisateur pourra simultanément parler, écouter et regarder son écran. Certains ont un écran séparé du corps de l'appareil ou sont dotés de microcasques qui associent des écouteurs et un microphone. Toutefois, les ingénieurs voudraient supprimer les fils. Aussi de nombreux prototypes transmettent-ils les signaux entre le microcasque et le corps de l'appareil à l'aide d'ondes électromagnétiques de faible puissance. Un réseau

radio à faible distance, nommé *Bluetooth* a été mis au point à cette fin ; il assurera également la transmission des données entre le terminal et des appareils situés à proximité : la caisse enregistreuse d'un supermarché, par exemple. Ainsi, le compagnon numérique servira aussi de portefeuille électronique.

Toutefois, ces machines risquent de coûter cher (la mise au point d'un prototype de supertéléphone coûte plus de 18 millions de francs), et quelques spécialistes du marketing prédisent qu'elles seront réservées à un public limité. «La plupart des individus n'auront pas besoin de l'animation vidéo intégrée continue», observe Lars Godell de la Société *Forrester Research*, qui pense que le compagnon numérique courant sera un «téléphone intelligent» simple et proche des téléphones cellulaires actuels ; son écran sera sans doute un peu plus grand, et son clavier amélioré facilitera la saisie des données. Ces téléphones intelligents assureront la téléphonie, l'envoi et la réception de courriers électroniques, l'accès au réseau Internet et, peut-être, l'écoute de musique téléchargée.

Pourquoi les spécialistes du téléphone seraient-ils les seuls à concevoir des supertéléphones ? Ces appareils transformeront des informations numériques en voix, données, jeux, musique, etc. Forts de leurs savoir-faire, les spécialistes du traitement ou de la création

de chacun de ces types d'information veulent aussi créer leurs types de compagnons numériques. Ainsi, les sociétés informatiques cherchent à relier à Internet leurs ordinateurs de poche et agendas électroniques ; grâce à des transmissions radio, ces systèmes deviendront de petits bureaux électroniques transportables partout. Une connexion automatique régulière avec le micro-ordinateur du bureau rafraîchira les données utilisées dans «le bureau de poche». Certains fabricants envisagent des interfaces vocales pour ces appareils, mais le surcoût réservera ces systèmes aux hommes et aux femmes d'affaires. De même, les fabricants de consoles de jeux vidéo veulent aussi que leurs clients puissent jouer partout. Selon L. Godell, les Sociétés *Nintendo* et *Sony* prévoient de doter leurs consoles d'accès au réseau Internet d'ici à deux ans. Même intérêt chez les fabricants de baladeurs : ils préparent des modèles qui téléchargeraient des fichiers MP3 à partir du réseau Internet.

Plus généralement, tous les constructeurs de systèmes électroniques portables veulent transmettre des données au réseau Internet aux terminaux sans fil. Certains projets sont très originaux. Par exemple, la Société *Swatch* s'est associée avec la Société *Hewlett Packard* pour créer une «montre Internet» : elle aura l'aspect d'une montre normale, mais son possesseur

pourra télécharger des messages électroniques de la musique ou des résultats sportifs grâce à un minuscule émetteur-récepteur.

La guerre des logiciels

Tous ces matériels portables sont de petits ordinateurs, qui sont animés par un système d'exploitation, c'est-à-dire un programme qui traduit les instructions reçues dans le langage du microprocesseur. Comment adapter les systèmes d'exploitation actuels à la gestion des échanges sans fil avec le réseau Internet? Aucun des systèmes actuels (les équivalents pour les systèmes portables de Windows 97, MacOS...) ne semble vraiment avantage. Le système d'exploitation mis au point par *Palm*, par exemple, est utilisé sur les machines *Visor* de la Société *Handspring*. Le système d'exploitation *EPOC*, créé par le fabricant britannique d'ordinateurs de poche *Psiion*, est déjà bien implanté : il est employé sur les téléphones cellulaires avec accès au réseau Internet des Sociétés *Nokia* et *Ericsson*. Ces systèmes, mis au point par des spécialistes du matériel portable, sont aujourd'hui concurrencés par la Société *Microsoft* : en 1996, le géant du logiciel a mis au point son propre système d'exploitation pour systèmes portables, *Windows CE*. Ce système a été adopté par *Hewlett Packard*, *Compaq* et *Casio*, pour leurs ordinateurs de poche. La Société *Microsoft* fabrique aussi des logiciels pour les «téléphones intelligents». Elle s'est associée au spécialiste des téléphones suédois *Ericsson* afin de créer, à l'aide de ces logiciels, du matériel portable facile à personnaliser : on adaptera ainsi les machines aux diverses professions (médecins, avocats, agents commerciaux...) ou aux diverses catégories de consommateurs (adultes, enfants...).

Beaucoup d'appareils portables seront probablement équipés pour reconnaître la parole. Toutefois, les avancées dans ce domaine sont lentes. La qualité des logiciels de dictée s'est améliorée (certains d'entre eux transcrivent avec précision entre 95 et 99 pour cent des mots qu'on leur dicte), mais, cette application mise à part, les tâches linguistiques les plus simples sont à peine à la portée des ordinateurs les plus puissants. Les performances insuffisantes limitent les applications. Pourtant, de nombreuses équipes se penchent sur le problème. Les Sociétés *Microsoft*, *IBM* et *Philips* étudient notamment des

systèmes d'exploitation qui seraient commandés par la voix. La Société belge *Lernout & Hauspie*, spécialisée dans la reconnaissance vocale, vient de se renforcer en rachetant ses concurrents : *Dragon* et *Voice phone* ; elle a créé le système *NAK*, une forme de commande par la voix qui fonctionne à partir d'un nombre limité de commandes vocales. Le nom *NAK* est inspiré du mot *nakulu*, qui signifie «écho» en hawaïen. Le système ne fonctionne que sur les appareils portables construits autour du microprocesseur *StrongARM* de la Société *Intel* (une puce puissante et qui consomme peu d'énergie). Avec le système *NAK*, les appareils peuvent exécuter les programmes de synthèse vocale pour lire les textes des courriers électroniques ; ils peuvent également prendre des réponses sous la dictée, puis les envoyer après réception d'une commande («envoyez le courrier électronique»). La Société *Lernout & Hauspie* annonce que *NAK* aura un vocabulaire de plus de 30 000 mots, un large répertoire de commandes et la capacité de comprendre plusieurs langues ; les matériels *NAK* seraient disponibles dès 2001 et coûteraient environ 30 000 francs.

Au système où apparaissent de nouvelles fonctions pour appareils portables, peut-être aurons-nous bientôt un compagnon numérique muni d'une tige, qui nous grattera obligeamment le dos pendant que nous téléphonerons? De quelles fonctions exactement sera doté le «compagnon numérique du peuple», celui qui sera courant dans dix ans? La question est évidemment difficile, car il y a de nombreuses incertitudes subsistant. Quel sera le succès des services qui se préparent pour l'Internet aérien? À quelle vitesse se feront les retours sur les investissements énormes consentis en ce moment? La bande passante de l'Internet aérien augmentera-t-elle assez vite pour rendre son utilisation supportable? Les opérateurs et les constructeurs de matériels ont leurs hypothèses, mais ils ont surtout besoin que nous acceptions ces futurs compagnons numériques. Aussi émettent-ils rigoureusement un message de leur cru : «Les téléphones portables actuels seront bientôt aussi dépassés que les monocles.»

Fiona HARVEY est journaliste spécialiste de l'industrie des télécommunications.

Les promesses et les risques du WAP

KAREN BANNAN

Pressenti comme un standard de communication entre les téléphones cellulaires et le réseau Internet, le système WAP a déjà des concurrents.

Nom bre d'industriels sont convaincus que l'Internet mobile commencera sous une forme dénommée «WAP» : grâce à cet ensemble de règles techniques, les pages Internet sont transmises sous une forme allégée et affichées sur les téléphones portables. Les systèmes aux spécifications WAP ont des performances que beaucoup jugent insuffisantes. Pourtant, le WAP est pressenti comme la norme des échanges de données entre le réseau Internet et les réseaux cellulaires. D'un autre côté, l'arrivée d'un concurrent japonais compromet son avenir.

WAP : les trois lettres sont l'acronyme de *Wireless Application Protocol*, soit «protocole pour les applications sans fil». Ce protocole décrit la transmission de données vers les appareils portables. Les premiers utilisateurs de ces services ont vite constaté ses limites : les pages se chargent lentement, elles sont affichées sur des écrans minuscules, la liaison est interrompue constamment, etc. Le rôle du WAP est d'améliorer les échanges : il

standardise la transmission, l'affichage et la préparation des données sur les réseaux cellulaires.

Plusieurs grands opérateurs tentent aujourd'hui d'imposer cette norme, mais sa pénétration est lente. Le public européen est le premier à qui des systèmes WAP ont été proposés. Il les adopte, mais sans empressement. En France, par exemple, sur 24 millions d'abonnés de la téléphonie mobile, moins de 100 000 ont aujourd'hui un appareil qui exploite le WAP. Pourtant, toutes les grandes entreprises françaises de télécommunication — *France Telecom, SFR, Bouygues Telecom* — proposent déjà des téléphones et des services WAP. Aux États-Unis, où le protocole fut mis au point, seuls quelques opérateurs, tels *PCS* ou *Nex-tel*, envisagent bientôt des services WAP.

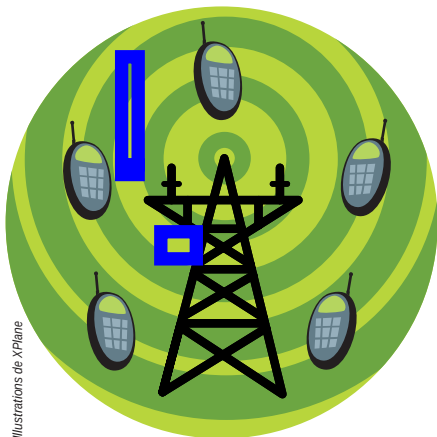
Si toutes ces grandes entreprises préparent des systèmes à la norme WAP, c'est parce qu'elles espèrent que l'Internet mobile, dont le WAP est la toute première forme, se développera rapidement. Elles veulent être des pionnières de cette technique, afin de renforcer leurs positions sur l'énorme marché mondial de la téléphonie mobile. Chaque fois que s'ouvre un marché, les clients initiaux sont essentiels, car c'est avec eux que vient l'expérience nécessaire à la conquête des parts de marché ultérieures. C'est pourquoi les grands opérateurs cherchent aujourd'hui à intéresser leurs abonnés aux services WAP.

Le protocole WAP est issu des efforts faits par la Société *Phone.com* pour faire communiquer les réseaux cellulaires avec le réseau Internet. Fondée au milieu des années 1990, cette société a vite constaté que les systèmes de bord des appareils portables sont incapables



de manipuler des pages HTML (*Hyper-Text Markup Language*), composées de textes, d'images et de «liens» ; en cliquant dessus, les internautes téléchargent immédiatement les pages Internet pertinentes en rapport avec la portion de page marquée. Les pages HTML contiennent aussi des sons, des images vidéo, des boutons, etc. : autant de parties de la page qui renferment souvent, aussi, des liens. Ce type de zone à haute densité d'information n'est téléchargeable que si le réseau transmet assez rapidement, ce qui n'est pas le cas des réseaux cellulaires. Les appareils portables ont d'autres difficultés avec les pages HTML : elles intègrent souvent des animations, lesquelles mobilisent de la puissance de calcul, qui est souvent limitée.

Afin de pallier ces défauts, les ingénieurs de la Société *Phone.com* ont d'abord créé HDML, ou *Handled Device Markup Language* : ce langage est aux réseaux cellulaires ce que HTML est à l'Internet. Les pages créées avec le langage HDML se transmettent plus facilement par les réseaux aériens ; surtout,



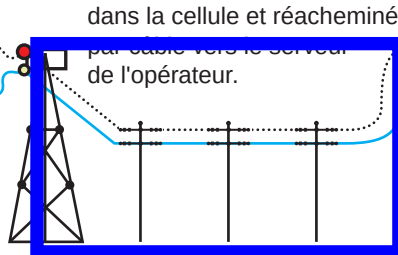
Illustrations de X-Plane

Comment un système WAP connecte votre téléphone au réseau Internet ?

- 1 Un utilisateur tape une adresse de site à l'écran de son terminal portable. Le micronavigateur de bord envoie une requête (un signal numérique) par les airs.
- 3 Ce serveur est relié à l'Internet et dispose d'un portail WAP. Ce logiciel cherche automatiquement la page HTML demandée.



- 2 Le signal est reçu par la tour de la station de base présente dans la cellule et réacheminé par câble vers le serveur de l'opérateur.



- 5 La page WML est transmise vers la station de base puis vers le téléphone mobile à l'origine de la requête. Le micronavigateur de bord la reçoit, puis l'affiche sur le petit écran du terminal portable.

- 4 Un programme de conversion transforme la page HTML en une page WML, un format qui convient mieux à la transmission par les airs et à l'affichage sur un écran de téléphone mobile. Puis le portail prépare la page pour la transmission à travers les airs.

elles s'affichent sur les mini-écrans des terminaux portables.

Rapidement, plusieurs opérateurs américains ont incorporé les logiciels de la Société *Phone.com* à leurs téléphones mobiles. Forte de ce succès, *Phone.com* s'est associée avec les trois grands fabricants de téléphones, *Motorola*, *Nokia* et *Ericsson*, afin de concevoir le langage WML (*Wireless Markup Language*) proche de HTML, mais standardisé. En juin 1997, *Motorola*, *Nokia*, *Ericsson* et *Phone.com* ont créé ensemble la fondation *WAP Forum*, qui rassemble tous les industriels intéressés par la définition de la norme WAP. Cette fondation compte aujourd'hui plus de 530 membres.

Un système WAP exploite les ressources du réseau Internet en les allégeant. Quand le propriétaire d'un téléphone portable lance le téléchargement d'une page sur le petit navigateur de l'appareil, l'appel est d'abord transmis à travers les airs vers la station de la cellule, puis acheminé vers le serveur WAP de l'opérateur. Sur ce dernier se trouve une « passerelle » vers le réseau Internet, c'est-à-dire un programme qui cherche automatiquement la page HTML demandée par l'utilisateur et la convertit au format WML. Au cours de

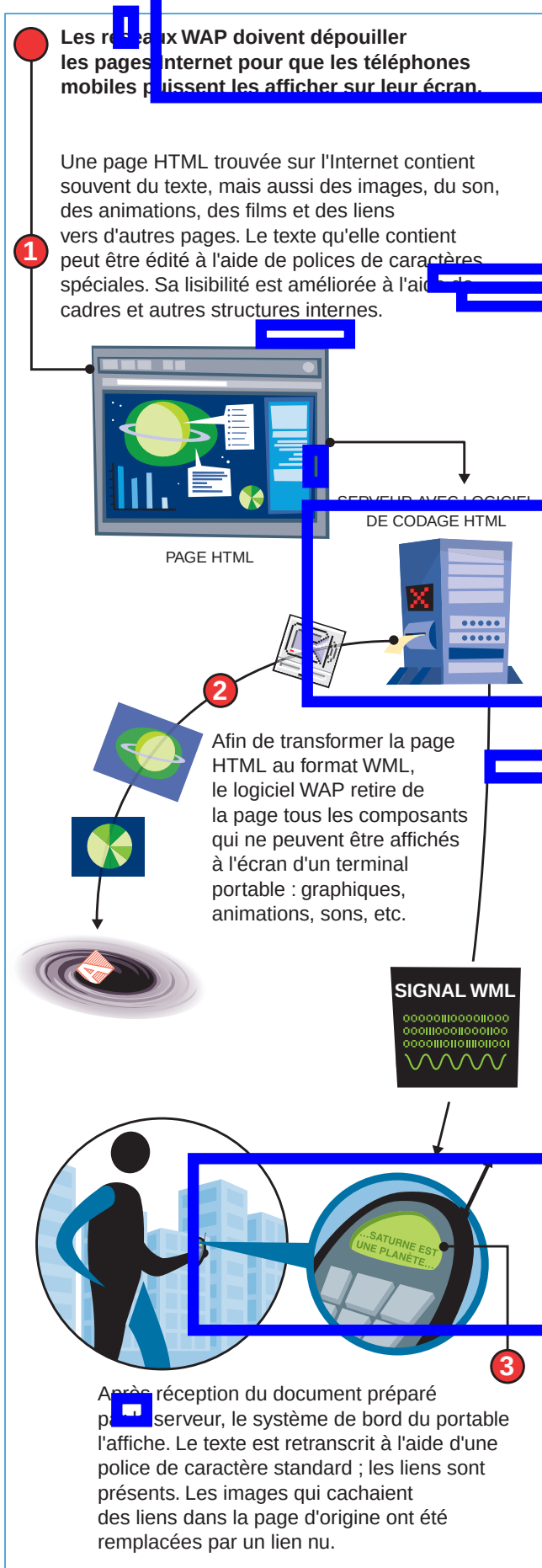
cette opération, toutes les images, les graphiques, les polices spéciales de caractères, par exemple, sont supprimés. C'est donc une page expurgée de toutes les zones à haute densité d'information qui s'affiche sur l'écran de l'appareil portable demandeur. Dans la plupart des cas, l'essentiel s'y trouve car il se concentre en général dans le texte de la page HTML.

L'opération est toutefois complexe. Ainsi, on ne doit pas transmettre sur le réseau cellulaire les images qui cachent des liens ! On doit récrire toutes les pages qui contiennent de telles images, afin de remplacer les liens graphiques – ceux qui se cachent sous les images – par des liens hypertextes. Autre problème : la transformation du document HTML en langage WML rend illisibles les pages dont la compréhension est indissociable du contenu graphique. Aussi, de nombreux éditeurs de pages Internet ont choisi de doubler leur sites HTML par un site WML. C'est, par exemple, le cas du site d'aide à la navigation *MapQuest*, ou de *MSNBC.com* (le portail de *Microsoft* de services Internet), ou encore de *Go2On-line.com* (site de recherche d'adresses Internet). Plus de 5 000 autres sites adaptés au WAP sont référencés sur le site

www.climania.com. Les services proposés sur ces sites sont de nombreux. Le propriétaire d'un terminal peut, par exemple, y consulter les résultats sportifs, les horaires des compagnies aériennes ou encore les dernières promotions du libraire en ligne *Amazon.com*.

Les faiblesses du WAP

Comparée à celle de l'Internet, l'offre WAP est maigre. C'est pourquoi de nombreux sceptiques prédisent l'échec du protocole. Les militants de la liberté sur le réseau Internet sont parmi les plus bruyants : pour eux, les spécifications du WAP limitent de façon inacceptable l'accès au réseau. À les écouter, la navigation sur un téléphone WAP est frustrante et onéreuse. Ils refusent donc l'idée que de nouveaux services Internet s'établissent sous cette forme. En outre, ils n'aiment pas que les opérateurs imposent aux utilisateurs de passer par leur site avant d'accéder à l'Internet. Le WAP a aussi des détracteurs dans l'industrie des télécommunications. Ces derniers jugent ses performances misérables : « Ce protocole a été conçu pour les réseaux de téléphones cellulaires à bas débit



des années 1990, c'est-à-dire pour des téléphones sans capacités graphiques, dotés d'écrans de deux à quatre lignes de texte», estime Rich Luhr, un consultant de la Société Hirschel Shostech.

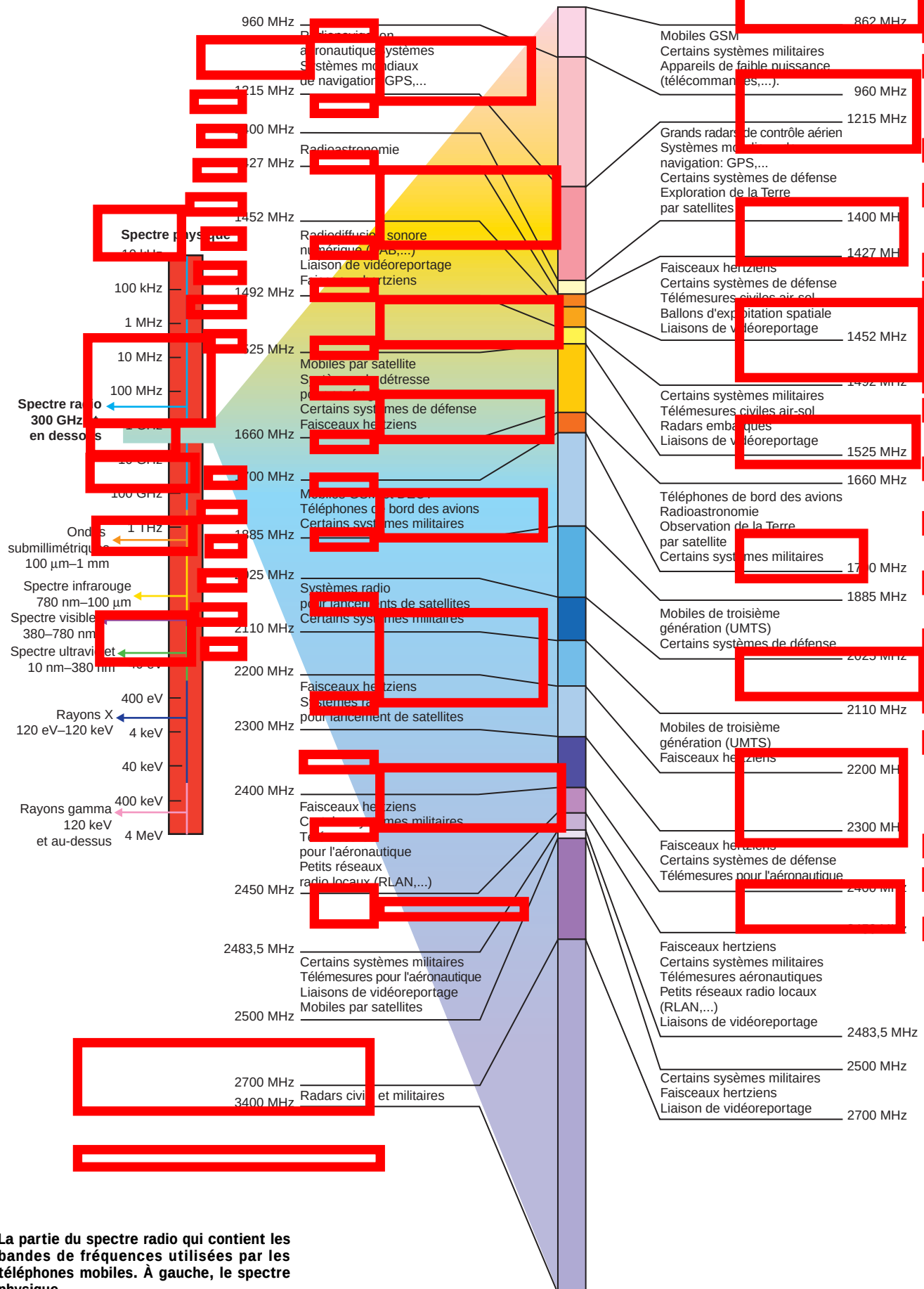
À ces critiques, les industriels engagés dans le WAP opposent que ce protocole pourra être révisé quand des réseaux cellulaires plus performants seront opérationnels (voir *Le sort de la troisième génération*, page 28). Ce moment pourrait être plus lointain que ce que pensent les détracteurs du WAP : les investissements à faire pour établir des réseaux 3G sont si énormes, que certains industriels se contenteront peut-être de perfectionner les réseaux actuels (2G) ; dans ce but, ils feront évoluer leurs réseaux GSM et autres DECT vers des réseaux GPRS ou EDGE. Dans ce cas de figure, le WAP persisterait bien plus longtemps que prévu. Aux passionnés de l'Internet, les industriels rétorquent que les terminaux WAP ne sont pas faits pour naviguer sur le réseau, mais pour y «picorer». Selon Skip Speaks, directeur de la division téléphonie cellulaire de la Société Ericsson, le WAP est conçu pour fournir un contenu spécifique d'une façon adaptée aux appareils sans fil, mais certainement pas pour de longs surfs sur le réseau. La navigation sur Internet est comme la restauration par un buffet à volonté : grâce à un riche choix de plats, chacun compose son menu. En revanche, le WAP ressemble plutôt au service de chambre d'un hôtel ; chaque client consulte un menu, commande, puis attend la livraison. Le service est payant et restreint, mais il est quand même utile !

Afin de pallier la pauvreté des services, le WAP Forum veut inciter les entreprises à éditer plus de pages WAP. La difficulté de la programmation en WML limite cette édition : même les créateurs de sites rompus à HTML la trouvent difficile, d'autant plus qu'il n'existe pas d'outils de programmation pour la faciliter. Selon S. Goldman, le président du WAP Forum, de tels environnements apparaîtront quand le WAP se démocratisera. Discours hypocrite de partisan ? Malgré l'activisme du WAP Forum, de nombreux créateurs de sites restent circonspects ; étant donné les investissements en temps et en énergie nécessaires, ces entreprises craignent de s'approprier un protocole voué à la désuétude.

Toutefois, si le WAP s'impose, elles auront eu tort. Ce protocole a l'avantage considérable d'être «ouvert» : chacune de ses spécifications est adoptée après discussion entre les membres compétents du WAP Forum. Publiée sur le site de la fondation, elle est accessible à tous, développeurs comme opérateurs. Pour ces derniers, cette normalisation est un avantage de taille, car elle élargit la concurrence. Plutôt que d'avoir toujours affaire à la même maison, ils pourront, à terme, se procurer des logiciels chez d'autres fournisseurs que Phone.com. L'apparition d'une norme assainit aussi la concurrence entre développeurs et entre créateurs de sites.

Un autre aspect du WAP attire les critiques : sa sécurisation réputée déficiente. Selon les détracteurs du protocole, les informations véhiculées par les réseaux WAP risquent d'être interceptées. Pourquoi ? Parce que ces réseaux comportent un maillon faible : le serveur. Toutes les informations envoyées par un utilisateur y sont décryptées, puis elles transitent par la passerelle vers l'Internet. Même si ces opérations ne durent qu'une fraction de seconde, des pirates compétents ont le temps de récupérer des informations. Sur le réseau Internet, la sécurisation

L'utilisation des fréquences radio en Europe



La partie du spectre radio qui contient les bandes de fréquences utilisées par les téléphones mobiles. À gauche, le spectre physique.



L'homme et l'animal

Claude Combes
Christophe Guittton

«160 pages brillantes
d'érudition, de profondeur
et d'ouverture d'esprit»

Philippe Dorchies
professeur de parasitologie,
La Dépêche vétérinaire

«Une promenade plaisante
claire et documentée, guidée
conjointement par un biologiste
et un vétérinaire.»

Eulélia

«Un livre riche et intéressant
alliant philosophie et sciences
grand public publié par deux
scientifiques de l'université
de Perpignan»
Midi Libre

Code P012 • 95 F
160 pages

Voir bon de commande
page 98

BELIN • POUR LA SCIENCE

des données importantes – identification, numéros de cartes de crédit, de compte, etc. – est assurée à l'aide du protocole SSL (*Secure Sockets Layer protocol*, soit «protocole de sécurisation»). Une série de spécifications comparable vient d'être définie pour le WAP : le WTLS (*Wireless Transport Layer Security*, soit «couche de sécurité pour le transport sans fil»). Les règles du WTLS spécifient comment crypter les données transmises à travers les airs pendant leur transit entre le téléphone cellulaire et l'opérateur de réseau. Comparé au SSL, le protocole WTLS a l'avantage d'exiger moins de ressources et de mémoire. Pour S. Goldman, le problème de la sécurité sur les systèmes WAP est aujourd'hui minime.

Le concurrent du WAP

Pendant, le WAP est menacé par un concurrent apparu brusquement dans le Soleil levant : le plus grand des opérateurs japonais remporte aujourd'hui un succès fulgurant grâce à l'*i-mode*, un système de navigation sur l'Internet et pour téléphones mobiles. En août 2000, plus de 12 millions de Japonais étaient déjà équipés d'un téléphone doté de l'*i-mode*. Deux mois avant, ils n'étaient que dix millions ! Exactement comme un système WAP, l'*i-mode* de la Société NTT DoCoMo sert à télécharger des informations, à envoyer des courriers électroniques, à jouer, etc. L'*i-mode* n'est pas une norme industrielle, c'est une «solution propriétaire», dont les caractéristiques techniques ne sont maîtrisées que par la Société NTT DoCoMo.

Le succès appelant le succès, l'opérateur japonais vient de convaincre la Société America on Line Japon (AOL Japon) de conclure une alliance stratégique : l'opérateur japonais veut profiter du savoir-faire d'AOL dans les services en ligne. Si la coopération avec le premier des fournisseurs d'accès Internet du monde est profitable au Japon, nul doute qu'elle sera promptement étendue au reste de la planète. L'alliance serait alors très puissante. L'opérateur japonais a conscience de cet enjeu et cherche à mettre toutes les chances de son côté. Il vient de s'implanter dans les deux grands marchés, l'Europe et les États-Unis, où il n'est pas encore présent. Apparemment, NTT DoCoMo ne veut pas devenir opérateur lui-même (il n'a pas acheté de licences en Europe),

mais cherche des alliances avec des sociétés déjà présentes. Aussi le groupe japonais a-t-il créé, il y a moins d'un an, une filiale aux États-Unis ; il a aussi installé un centre de recherches dans la Silicon Valley. De plus, NTT DoCoMo arrive aujourd'hui en Europe : le groupe vient de créer sa filiale européenne à Londres, investissant 30 millions de francs. Il investit aussi 50 millions de francs à Munich, pour y fonder un autre centre de recherches. Manifestement, la stratégie fonctionne : NTT DoCoMo vient de passer un premier accord commercial avec un gros opérateur européen, la Société hollandaise KPN Mobile Net, qui proposera des services fondés sur la technique japonaise. Si NTT DoCoMo séduit d'autres opérateurs et une partie du public européen, l'influence du WAP en sera limitée.

D'autant que l'adoption prochaine du protocole n'est pas garantie aux États-Unis ! Par souci d'économie, certains des opérateurs américains, qui exploient déjà l'HDML, décideront peut-être d'en rester à ce précurseur du WAP (également dû à *Phone.com*) en attendant l'avènement de réseaux cellulaires plus performants (de troisième génération). Cette décision leur éviterait d'avoir à changer les téléphones portables dotés de services au format HDML, qu'ils ont déjà distribués à leurs clients.

L'avenir du WAP dépendra de l'attitude des opérateurs et, surtout, de celle des utilisateurs de téléphones mobiles. Aujourd'hui, ces derniers sont frustrés par les performances limitées et le coût des services WAP. Aussi, nombre d'observateurs estiment que les téléphones WAP ne se démocratiseront que le jour où les opérateurs factureront les services de données au même prix que les communications vocales. Autre obstacle à l'avènement du WAP : les entreprises de télécommunications contrôlent dans une large mesure les téléphones. L'Internet ne s'est pas développé tant qu'il a été coûteux et restreint. Il s'est développé le jour où chacun a créé son site en toute liberté et à peu de frais.

Hélas, le WAP est aujourd'hui dominé par ses créateurs.

Karen BANNAN est journaliste spécialisée dans l'industrie informatique.

Le saut de la troisième génération

LEANDER KAHNEY

Le débit des transmissions de données augmente avec les générations de téléphones mobiles. Le déploiement des réseaux de troisième génération implique de grandes avancées techniques. Oui, mais lesquelles ?

«Avec la «3G», vous allez voir, entendre, lire...» Les opérateurs et les constructeurs promettent que cette «3G», la troisième génération de téléphonie, sera merveilleuse. Le débit des transmissions des données sur les réseaux 3G atteindrait 2,4 mégabits par seconde, alors qu'elle est inférieure à 10 kilobits par seconde sur les réseaux mobiles d'aujourd'hui. Oui, 2,4 mégabits par seconde, soit deux fois plus qu'avec les réseaux câblés les plus rapides d'aujourd'hui ! Ces débits considérables changeront notre vie, nous disent les promoteurs de la «3G», car nos compagnons numériques pourront enfin nous procurer des images, des films, de la musique, du courrier. Bref, toutes les informations du monde qui transitent par le réseau Internet. Quand cette révolution aura-t-elle lieu ? Tout de suite ou presque, disent les opérateurs : au Japon, elle a déjà com-

mencé ; elle battra son plein chez nous en 2002, puis en 2003 aux États-Unis.

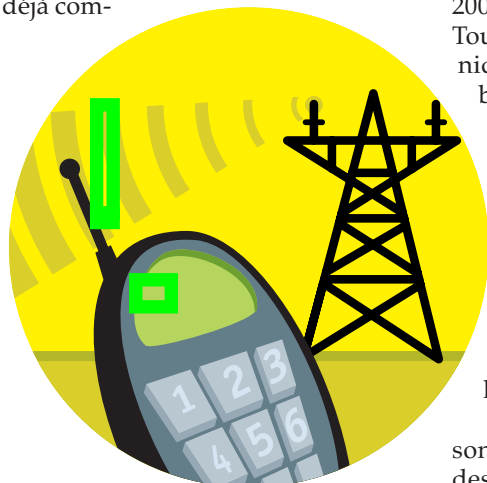
Les signes 3G ou UMTS, pour *Universal Mobile Telecommunications System* («système universel de télécommunication mobile») sont équivalents : ils désignent l'ensemble des solutions techniques envisagées pour conférer de hauts débits

aux réseaux cellulaires actuels. La plupart des «systèmes 3G» sont des versions perfectionnées des systèmes de transmission actuels, différentes selon les industriels et selon les pays. Cette situation confuse se traduit par une saut à l'alphabet, sauce américaine : CDMA, GSM, CDMA, CDMA-2000, W-CDMA, GPRS, EDGE, UWM, etc. Toutefois, trois de ces solutions techniques remplaceront probablement bientôt les systèmes actuels : EDGE, CDMA-2000 et W-CDMA. Pour résumer, en Europe et en Asie, le système GSM sera remplacé par le système W-CDMA. En Amérique du Nord, où plusieurs types de réseaux «2G» coexistent, les réseaux CDMA vont migrer vers le CDMA-2000, tandis que les réseaux TDMA évolueront vers EDGE.

Aujourd'hui, ces systèmes n'en sont qu'au stade des prototypes ou des tests. Chacun a ses forces et ses

faiblesses. EDGE a l'avantage considérable de pouvoir être mis en œuvre après une mise à niveau limitée des infrastructures. Toutefois, son débit maximal est seulement de 384 kilobits par seconde, ce qui est bien inférieur aux deux mégabits par seconde atteints par les systèmes de type W-CDMA ou CDMA-2000. Ce dernier doit son débit remarquable à sa forme de commutation, laquelle répartit de multiples paquets d'informations sur un large spectre de fréquences. Le système W-CDMA est un descendant du CDMA, un système qui attribue à chaque liaison téléphonique un code particulier. Avec un tel principe, de multiples signaux radio partagent sans problèmes une bande large de fréquences. Chaque téléphone capte tous les signaux, mais il ne retient que ceux qui sont étiquetés avec le bon code. Dans d'autres variantes du CDMA, le système saute sans cesse d'une fréquence à l'autre suivant une séquence fixe.

Les systèmes à large bande de type W-CDMA semblent s'imposer. Certes, l'emploi de codes implique une gestion lourde, susceptible de limiter les performances. Toutefois, ce type de réseau semble faire le meilleur usage d'une large bande de fréquences. En outre, les réseaux W-CDMA supportent bien le bruit de fond et sont peu sensibles aux interférences. Alors que les canaux du réseau CDMA sont larges de 1,25 mégahertz dans les



Illustrations de XPlane

d'une station serait beaucoup mieux utilisée si les antennes émettaient selon la position des destinataires de leurs signaux. Le système de la Société *ArrayComm* utilise un circuit de traitement du signal qui a été créé pour l'armée américaine à des fins d'espionnage des communications radio. Une fois qu'il a traité le signal émis par l'appareil portable d'un utilisateur, ce circuit intégré calcule sa position avec précision, puis en déduit des instructions adéquates pour les antennes. Ces dernières émettent alors dans la bonne direction. L'opération se répète à intervalles réguliers pour tous les utilisateurs présents dans la cellule. Utiliser une même radiofréquence pour écouter ou émettre vers plusieurs utilisateurs est possible. Comme les émissions dépendent de la position des téléphones, les interférences entre les différents signaux sont réduites. Cet emploi multiple des mêmes fréquences permet d'offrir des hauts débits à chacun des utilisateurs. Au final, le spectre de l'opérateur est utilisé de façon très efficace.

Pour M. Cooper, le système est facile à installer. Les antennes sont déjà en place, et nombre de stations de base disposent de la puissance de calcul nécessaire pour remplacer les microcontrôleurs cités précédemment. Le système a toutefois un défaut : le débit de transmission dépend de la vitesse de l'utilisateur ; plus un utilisateur se déplace vite, plus le débit des informations qu'il reçoit est limité. Aujourd'hui, le système d'*ArrayComm* suit efficacement les piétons, mais pas les voitures. La Société *ArrayComm* veut lancer bientôt les premiers essais à grande échelle. L'entreprise s'est associée à Sony pour transmettre avec cette méthode des vidéos, de la

musique et des jeux à l'intérieur de San Diego.

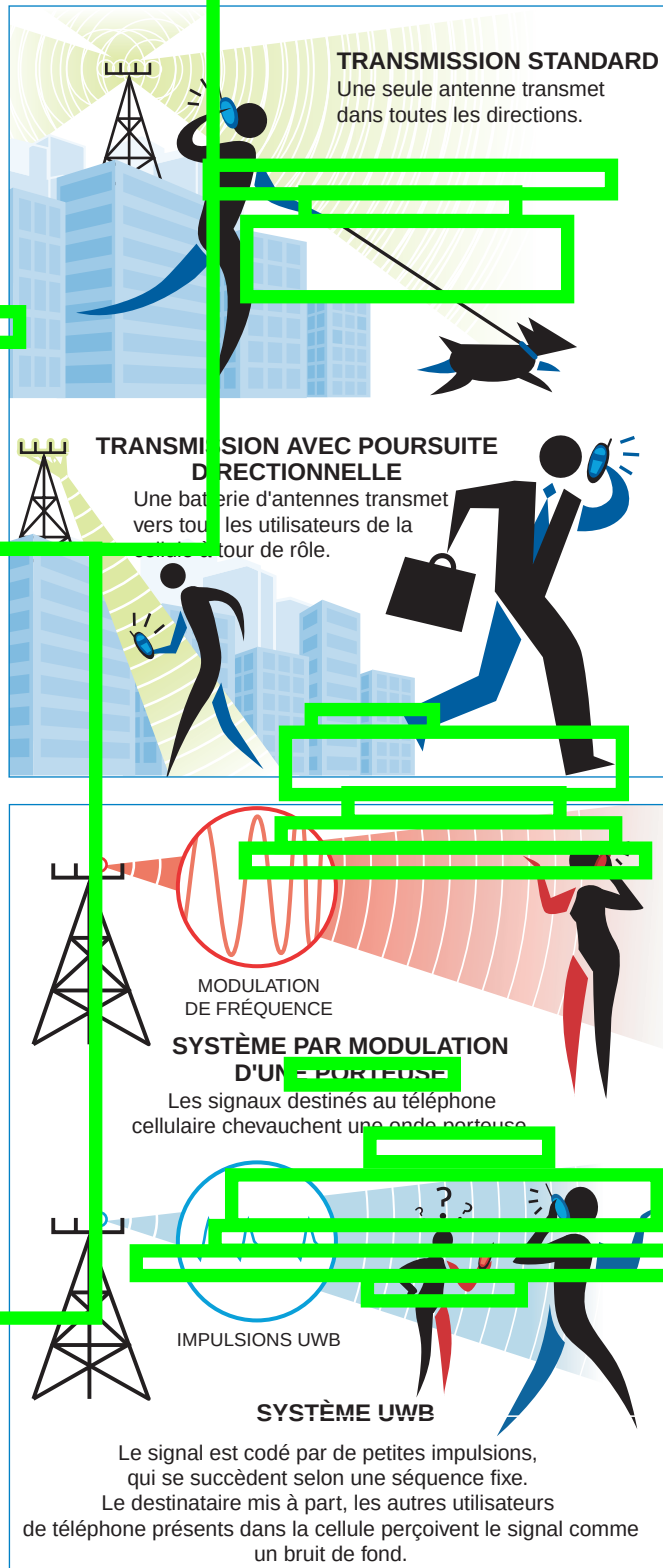
D'un côté, l'inventeur Larry Fullerton prépare un autre système, également concurrent de W-CDMA et de EDGE. Depuis deux décennies, il perfectionne une technique d'émission radio inventée dans les années 1960 : une sorte de morse ultrarapide pour

stations d'émission. Un commutateur commande l'envoi d'impulsions électromagnétiques ultrabrèves dans l'antenne. Ces dernières durent moins d'un milliardième de seconde et peuvent être produites jusqu'à 40 millions de fois par seconde. Elles sont espacées dans le temps de façon à former un « motif temporel », lequel code l'information. Seuls les récepteurs qui connaissent les motifs décodent les signaux.

Ce système UWB (pour *Ultra Wide Band*, soit « bande ultralarge ») utilise pleinement la quasi-totalité du spectre : les impulsions transportent en effet de l'énergie entre un et trois gigahertz. Comme différents transmetteurs UWB utilisent différents motifs, la transmission simultanée d'un grand nombre de signaux est possible sans interférence. Pour un transmetteur, le signal d'un autre appareil passe en effet pour du bruit de fond. Pour les récepteurs radio classiques, le signal UWB ressemble à des parasites aléatoires qui peuvent être filtrés tant que leur puissance reste faible.

L. Fullerton est aujourd'hui le directeur technique de la Société *Time Domain*, qui commercialisera UWB, avec l'intention d'en pousser très haut les performances : un gigabit par seconde ? Quelques autres entreprises, telle que *Multispectral Solutions*, comptent employer l'UWB. Jusqu'au printemps 2000, l'Autorité américaine de régulation des télécommunications n'avait accordé que des permissions limitées pour expérimenter le système UWB. En mai, elle a toutefois autorisé les essais à grande échelle.

Le système UWB a un grand éventail d'applications potentielles : systèmes radar personnels anticollisions, systèmes imageurs capables de voir à travers les murs... S'imposera-t-il



pour les communications sans fil à large bande? Seulement si une difficulté majeure est réglée : afin d'éviter les interférences avec les signaux des autres émetteurs radio, la puissance des systèmes UWB est limitée à moins de 50 millièmes de watt ; sa portée est donc réduite à quelques mètres. C'est pourquoi la Société *Time Domain* prévoit que les systèmes UWB seront d'abord limités aux réseaux locaux, à l'intérieur des bâtiments. Ils y joueront le rôle d'une sorte de réseau Bluetooth élargi (voir l'encadré de la page 2), et c'est seulement plus tard qu'ils seront employés sur des territoires plus vastes.

Multiplexage radio

Une troisième approche, le multiplexage, pourrait bouleverser l'organisation des réseaux «3G». Le multiplexage est employé dans les télécommunications par fil depuis les années 1950 pour transmettre plusieurs signaux par le même canal. Il est aujourd'hui utilisé couramment dans la transmission par fibres optiques. En radio, plusieurs des

général d'une bande toutes comprises dans la bande passante, on envoie jusqu'à plus de 100 signaux sur une même fibre sans provoquer d'interférences. Si ce principe n'a pas encore été appliqué aux communications radio, c'est parce que les circuits de traitement du signal nécessaires étaient insuffisants. Le temps et la recombinaison rapide de signaux entremêlés impose des circuits intégrés qui traitent les signaux très vite.

De tels circuits apparaissent. La Société canadienne *Wi-LAN* a breveté une méthode de multiplexage nommée W-OFDM, acronyme anglais d'une expression qui signifie «multiplexage orthogonal à large bande par répartition en fréquences».

Selon Hatim Zaghloul, directeur général de la Société *Wi-LAN*, la méthode W-OFDM permet des débits de transmission élevés sur une partie limitée du spectre des fréquences radio. Elle utilise qu'environ dix pour cent de la bande «ISM», réservée aux applications industrielles, scientifiques et médicales, laquelle peut être utilisée sans autorisation préalable. Ces fré-

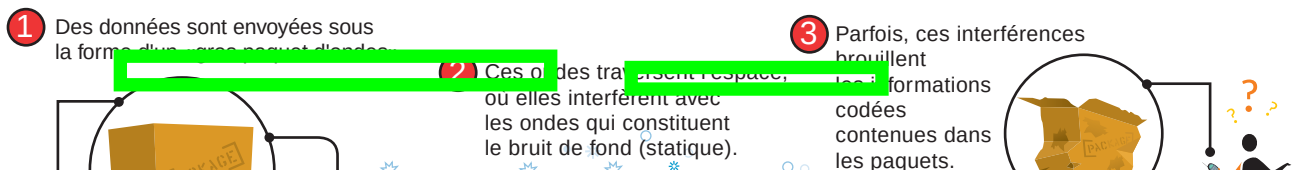
quences sont régulièrement espacées, qui transmettent chacun un mégabit par seconde. L'émetteur utilise en parallèle chacun des canaux, où il transmet à une vitesse plus réduite que s'il n'utilisait qu'un seul canal. Cette lenteur se révèle être un avantage, car la transmission est alors moins sensible au bruit de fond et aux réflexions parasites. En les subdivisant, les systèmes qui multiplexent utilisent le spectre plus efficacement.

Avec un de ces systèmes, la Société *Wi-LAN* a atteint un débit de 32 mégabits par seconde. Lors d'essais menés au début de l'année 2000, un système W-OFDM a transmis un film vidéo vers une voiture qui roulait à la vitesse de 110 kilomètres par heure. Les nouvelles versions de W-OFDM, en préparation, devraient atteindre un débit de 155 mégabits par seconde avant la fin de l'année 2001. La Société *Wi-LAN* étudie notamment des processeurs de signaux assez puissants pour transmettre simultanément selon plus de canaux. La méthode de multiplexage W-OFDM devrait être déployée dans des systèmes sans fil mais fixes, à la fin de l'année 2000 et dans

Comparaison

transmission standard de signal/transmission par multiplexage

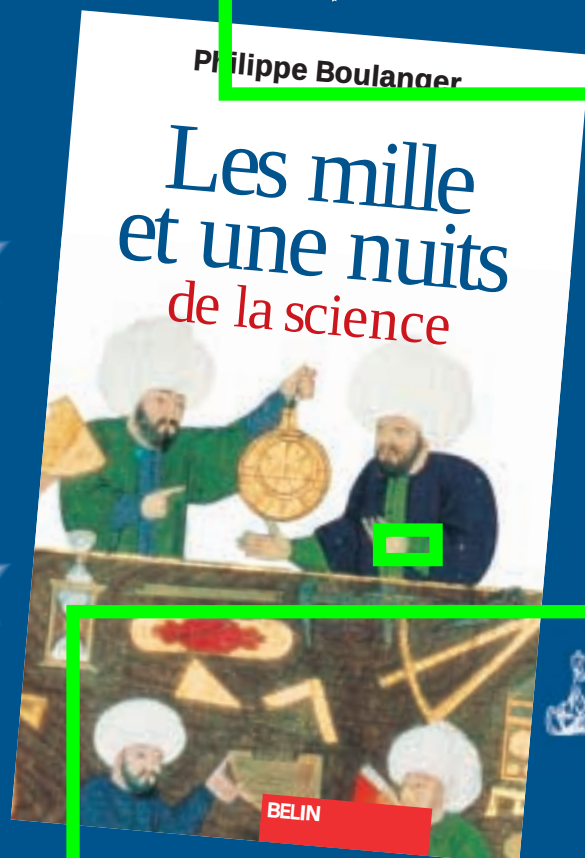
TRANSMISSION STANDARD



TRANSMISSION PAR MULTIPLEXAGE



Entrez dans l'ambiance orientale
des mille et une nuits et laissez-vous
charmer par l'imagination
de Shéhérazade contant la science
et les tribulations des savants



«Dans ce livre, Philippe Boulanger,
Directeur de la revue *Le Point*,
réussit le tour de force de vulgariser des
concepts difficiles en charmant son lecteur
par une subtile danse... des neurones».

Hervé Ponchelet
Le Point

Code 2445 • 240 pages • Prix 100 F
Voir bon de commande p. 98

BELIN • POUR LA SCIENCE

les systèmes mobiles en 2003. En revanche, cette méthode imposerait une restructuration profonde des réseaux cellulaires actuels. Sera-t-elle une solution pour la quatrième génération de téléphones mobiles?

Là est l'essentiel du problème posé par l'adoption d'une solution technique «3G» : chaque choix détermine les suivants, parce qu'il implique d'importants investissements dans les infrastructures. Vu le grand nombre de solutions techniques à l'étude, l'avènement d'une forme de télécommunication sans fil à haut débit dépendra surtout des ressources que les opérateurs ont ou sauront trouver (dans le porte-monnaie des utilisateurs). Quel que soit le choix technique qu'ils devront faire des compromis entre la largeur de bande disponible et le nombre d'utilisateurs qui la partageront : «L'essentiel du problème est dans le plan d'exploitation, pas dans la technique, estime Craig Mathias, de la Société d'études de marchés *MobileInsights*.

L'arrivée des systèmes 3G et leurs performances ne dépendent que des opérateurs. S'ils veulent fournir des réseaux à hauts débits de transmission de données, ils y arriveront. Toutefois, seules les communications vocales leur permettent aujourd'hui de gagner de l'argent.» D'où les nombreuses possibilités, qui nous vantent les performances d'infrastructures qui n'existent pas, et n'existeront peut-être que si leur démarrage se révèle assez rentable...

Selon M. Cooper, l'industrie des télécommunication est dominée par de grandes entreprises, trop souvent en situation de monopoles. Ces grands opérateurs pensent naturellement adopter un système unique pour tous les réseaux qu'ils exploitent, mais des réseaux plus variés et mieux adaptés aux divers usages seraient préférables.

Les réseaux nationaux de communications vocales coexisteront sans doute avec des réseaux à haut débit de données. Certains seront chers et à haut débit, d'autres bon marché et à débit modeste. Enfin, les performances des réseaux sans fil devraient égaler un jour celles des réseaux câblés.

Leander KAHNEY est rédacteur au journal californien *Wired News*.

À cœur battant

CORNELIUS BORST

Les chirurgiens opèrent aujourd'hui le cœur sans avoir à l'arrêter.

On doit généralement opérer les personnes âgées qui, au moindre mouvement, souffrent de douleurs dans la poitrine, qui mettent une heure à se lever et à s'habiller parce qu'elles se fatiguent vite : leurs douleurs résultent souvent de l'obstruction des artères coronaires, les vaisseaux qui alimentent le cœur en oxygène. Ces vaisseaux, issus d'une ramification de l'aorte, sont solidaires du cœur. L'opération est le pontage coronarien.

Celui-ci est banal : chaque année, dans le monde, environ 800 000 personnes en subissent un. Toutefois, il fait courir des risques aux patients, sur-

tout quand ils ont des troubles pulmonaires, des maladies rénales ou des troubles de la coagulation sanguine. Techniquement, le chirurgien prélève des vaisseaux dans la poitrine ou dans la jambe du malade et il les greffe en parallèle du vaisseau endommagé, au-delà des rétrécissements. L'opération impose une ouverture de la poitrine (les chirurgiens parlent de « scier » la cage thoracique, car il coupent le sternum avec une scie et ouvrent la cage thoracique). Pendant environ une heure, ils arrêtent le cœur, car ils ne peuvent suturer correctement un vaisseau du cœur pendant que celui-ci bat.

Pendant cet arrêt du cœur, le malade est relié à une machine, nommée cœur-poumon, qui assure la circulation sanguine, dite extra-corporelle, et fournit de l'oxygène aux tissus jusqu'à ce que le cœur redémarre. Cette machine complexe a été introduite vers 1960 dans les blocs opératoires modernes de chirurgie cardiaque. C'est cette circulation artificielle déviée dans une machine cœur-poumon qui provoque parfois de graves complications, notamment chez les personnes âgées ou affaiblies. Ces complications expliquent les longs séjours postopératoires – jusqu'à huit jours – et, souvent, des convalescences de deux à trois mois. De surcroît, certaines personnes supportent mal l'ouverture de la cage thoracique et, au cours de leur convalescence, elles deviennent plus sensibles à certaines infections, telles les pneumonies. Comment éviter ces conséquences des opérations ?

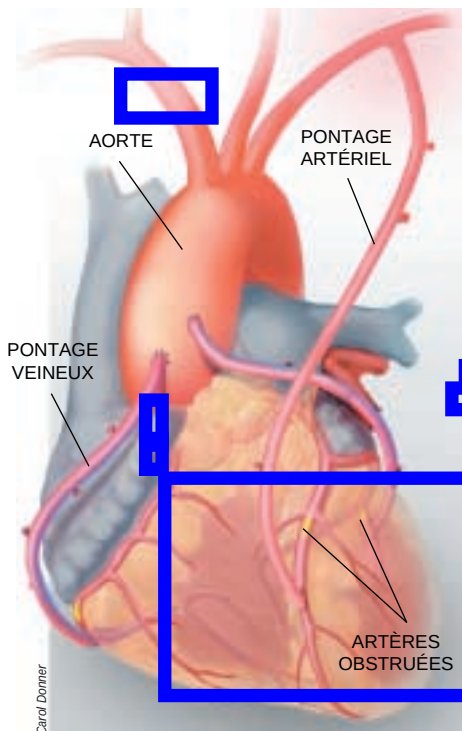
Vers le milieu des années 1990, les chirurgiens ont mis au point deux techniques chirurgicales pour faciliter les pontages coronariens. Avec d'autres médecins, nous avons été de ceux qui ont cherché à supprimer la machine cœur-poumon afin d'opérer un cœur battant. D'autres équipes ont testé la chirurgie cardiaque par endoscopie, une opération dont les incisions sont à peine plus grandes qu'un trou de serrure.

Opérer un cœur qui bat

Les douleurs de poitrine qui imposent les pontages coronariens sont dues à l'athérosclérose, c'est-à-dire au durcissement des artères causé par des substances, tel le cholestérol, qui s'accumulent parfois sur les parois des artères dont elles rétrécissent le diamètre. La maladie évolue progressivement et, vers la cinquantaine, une petite douleur dans la poitrine apparaît à l'effort, car le rétrécissement coronarien limite le débit de sang, qui devient insuffisant lors des activités soutenues. La vie de ces malades est alors menacée : en l'absence de traitement, ils risquent un infarctus.

Des millions de personnes dans le monde sont handicapées par ces douleurs dans la poitrine. Certains facteurs génétiques favorisent l'athérosclérose, mais une alimentation déséquilibrée ou un mode de vie stressant en sont plus souvent responsables. Nous examinerons ici l'amélioration des traitements des maladies coronariennes, mais nous prôtons avec vigueur une alimentation saine, la pratique de l'exercice physique et l'absence de tabac.

Lorsqu'un diagnostic d'athérosclérose est posé, le médecin prescrit des médicaments ou opte pour l'angioplastie, c'est-à-dire l'insertion d'un petit ballon oblong dans l'artère obstruée : le gonflement du ballon étire les parois élastiques et dilate le vaisseau. Le cardiologue place aussi parfois dans le vaisseau une minuscule structure métallique en forme de ressort, qui le maintient ouvert. Cependant, quand les traitements précédents sont trop risqués ou lorsque le cardiologue prévoit que l'artère se rétrécira de nouveau après l'angioplastie, il doit envisager un pontage pour rétablir un débit sanguin adéquat jusqu'au cœur. Lors de ces pontages coronariens, on doit généralement greffer deux à cinq vaisseaux sur les artères du cœur. Pour chaque greffe, les chirurgiens



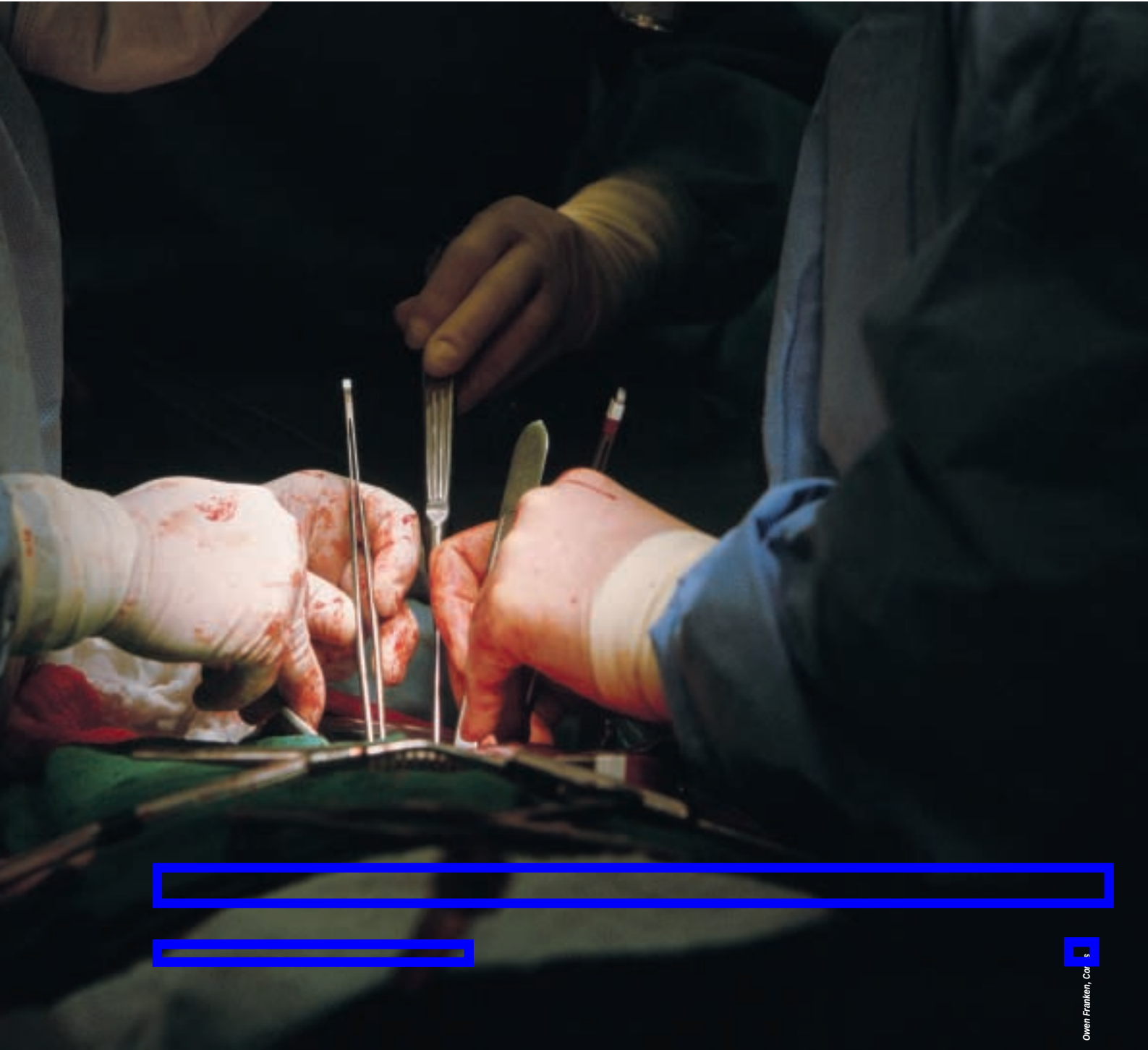
1. LORS D'UN PONTAGE SUR LE CŒUR, on relie par des vaisseaux des grosses artères, telle l'aorte, aux artères coronaires en aval des rétrécissements, afin que le sang contourne les zones obstruées. Les chirurgiens utilisent soit des greffons artériels (des artères déviées à proximité du cœur), soit des greffons veineux (des segments de veines prélevés dans la jambe).

effectuent une suture entre un vaisseau donneur et l'artère coronaire bouchée : chaque greffe dure quelques minutes.

Nous l'avons vu, la machine cœur-poumon est l'une des principales sources des complications de la chirurgie cardiaque. Pour mettre en place ce dispositif, le médecin le connecte, d'une part à l'oreillette droite du cœur (qui reçoit le sang veineux) et, d'autre

part, à l'aorte, qui véhicule le sang après son oxygénation. Puis il introduit dans l'artère coronaire une solution paralysante qui stoppe les battements cardiaques. Ces manipulations décrochent parfois des fragments de plaques athéromateuses dans l'aorte. Si ces fragments sont emportés dans le cerveau, ils provoquent une attaque cérébrale en obs-

truant des capillaires. En outre, la machine cœur-poumon perturbe les défenses naturelles de l'organisme : les opérations provoquent parfois des fièvres, des lésions d'organe et des hémorragies. Après l'opération, le malade requiert quelquefois une assistance respiratoire et, lorsque le cœur bat de nouveau, ses contractions sont parfois anormales (arythmie). Enfin,



2. LES PONTAGES CORONARIENS CLASSIQUES imposent de scier la cage thoracique : on arrête le cœur et on relie le malade à une machine cœur-poumon qui fait circuler le sang et l'oxygène dans l'organisme. Une telle opération est très éprouvante.

les malades souffrent parfois d'hypotension, d'une diminution du débit sanguin et de la production d'urine. Dans quelques cas rares, on doit remplacer la machine cœur-poumon débranchée par une pompe mécanique.

Les risques sont aujourd'hui mieux connus. La probabilité d'un décès peu après un pontage coronarien augmente avec l'âge. En France, par exemple, ce risque passe de un à trois pour cent entre 20 et 50 ans à 10 pour cent entre 81 et 90 ans. Un malade sur quatre manifeste au moins une complication opératoire. En 1998, parmi les malades de plus de 60 ans, trois à cinq pour cent d'entre eux sont morts à l'hôpital. Souvent, les malades ne reprennent des activités normales qu'après deux à trois mois au moins, parfois à cause de pertes de mémoire et d'attention, ainsi que d'une grande faiblesse physique ou d'un état dépressif. Depuis 15 ans, nous étudions des solutions pour mieux traiter les maladies coronariennes. Avec mes collègues, nous avons mis au point un dispositif mécanique qui immobilise uniquement la zone du vaisseau obstrué, et non le cœur tout entier.

Une vieuvre sur le cœur

Le premier système chirurgical à cœur battant a été présenté en mars 1993 : Richard S. Stava, qui était médecin de l'armée américaine, a exposé le prototype d'un robot que des chirurgiens commanderaient à distance pour une chirurgie d'urgence sur les champs de bataille. Parallèlement, nous avons cherché à opérer sur un cœur actif sans équipement complexe.

Au printemps 1994, avec Erik Jansen, cardi-chirurgien à Utrecht, nous avons reproduit sur des cochons un acte de chirurgie sur cœur battant qui avait été proposé par Federico Benetti, du Centre de chirurgie cardiovasculaire de Buenos Aires, et par Enio Buffolo, de l'École de médecine de São Paulo.

Afin de ponter l'artère coronaire, les deux médecins Sud-américains immobilisaient une petite région de la surface du cœur par des points de suture de stabilisation pratiqués sur le tissu adjacent au site du pontage et par une pression exercée par une grosse pince chirurgicale maintenue manuellement. Ainsi, seuls quelques centi-

mètres carrés du cœur étaient immobilisés et la fonction de pompage était à peine diminuée. Cependant, malgré l'élégance du procédé, F. Benetti et E. Buffolo ont fait peu d'émules dans l'immédiat.

En mai 1994, à Utrecht, lors d'une opération expérimentale sur un porc, j'étais chargé de tenir la pince. Malheureusement, la région du cœur où nous voulions effectuer le pontage n'était pas complètement immobile. Pourquoi ne pas remplacer les points de suture et la main humaine par un dispositif mécanique rigide ? Après quelques semaines, E. Jansen a utilisé avec succès un prototype de stabilisateur cardiaque construit par la Société Rik Moesvelt Beck.

Peu après, avec Paul Gründeman, nous avons inventé un stabilisateur cardiaque nommé Octopus, qui immobilise n'importe quelle petite zone de la surface d'un cœur battant (voir la figure 4). Le nom provient des coupelles de succion qui fixent l'instrument sur le cœur. En septembre 1995, nous avons utilisé pour la première fois l'Octopus pour un pontage sur un malade humain ; depuis, plus de 50 000 personnes dans le monde ont été opérées avec un Octopus. Parmi les 400 malades qui ont été opérés à Utrecht, le taux de mortalité est nul au cours de l'opération ou pendant les 30 jours suivants.

D'autres chirurgiens ont mis au point d'autres dispositifs de stabilisation mécanique différents (par exemple, par pression) et, aujourd'hui, les cardio-chirurgiens disposent de 13 sortes d'appareils.

Des études sur les 200 000 malades qui ont subi un pontage coronarien sur cœur battant avec l'aide d'un stabilisateur mécanique montrent qu'ils ont eu moins de complications au cours de l'intervention chirurgicale, que l'opération a nécessité moins de transfusions sanguines, que les patients sont restés moins longtemps sous respiration artificielle ou en unité de soins intensifs et qu'ils ont quitté l'hôpital et repris leurs activités normales plus rapidement que les malades opérés selon les techniques classiques. De plus, le coût global de l'opération est réduit d'environ un tiers.

En 1999, près de dix pour cent des opérations coronariennes dans le monde se sont faites sans machine cœur-poumon, et nous espérons que les stabilisateurs mécaniques seront utilisés dans plus de 50 pour cent des



Chris Timmers, Utrecht University Medical Center

3. LES MACHINES CŒUR-POUMON assurent l'irrigation sanguine et l'oxygénation des tissus de l'organisme lorsque le cœur est arrêté, au cours des opérations classiques.

opération de pontage en 2005. Dans des hôpitaux démunis de machines cœur-poumon, tels ceux des pays en développement, ces systèmes permettent une chirurgie coronarienne, qui serait autrement inaccessible. Si les machines cœur-poumon sont superflues, l'ouverture de la cage thoracique l'est-elle également ?

F. Benetti a encore amélioré la chirurgie sur cœur battant en réduisant la taille de l'incision à huit centimètres seulement, entre les côtes du côté gauche de la poitrine, pour des malades qui ne requièrent qu'un seul pontage de l'ar-

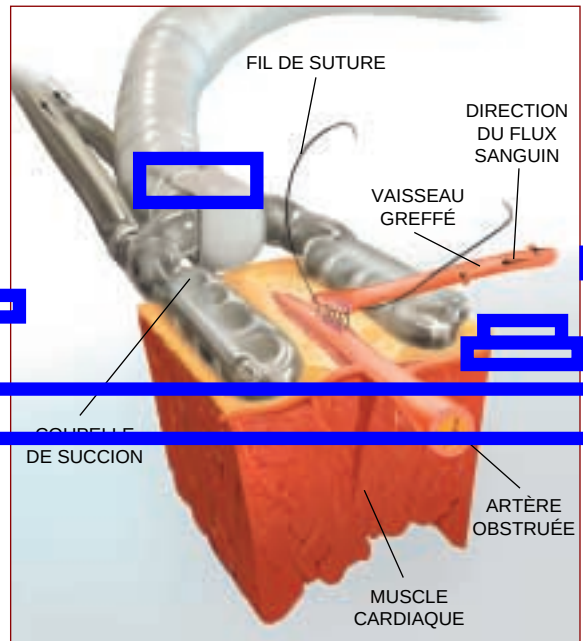
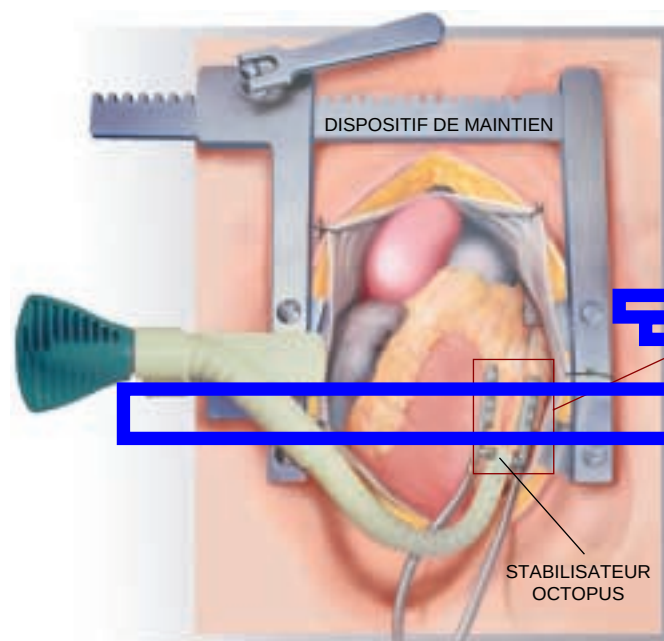
rière coronarienne principale, sur le devant du cœur. Le chirurgien écarte encore les côtes adjacentes, mais le procédé est moins réhudiciable que l'ouverture de la cage thoracique tout entière.

La chirurgie par un trou de serrure

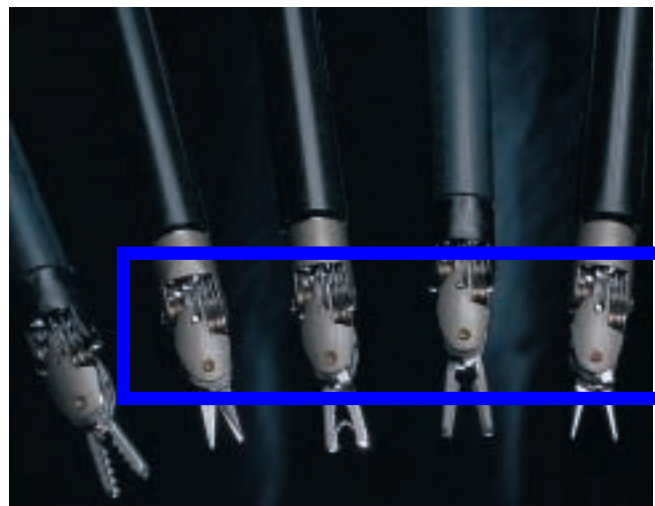
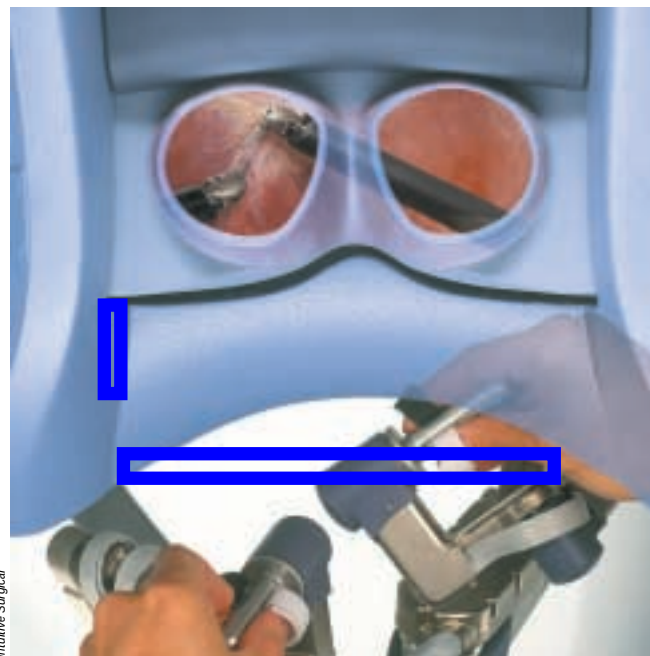
Les chirurgiens ont évidemment été très intéressés par ce progrès. En novembre 1994, lors d'un colloque à Rome, Valavanur Subramanian, de New York, a projeté un film de présentation de cette technique : celle-ci

s'est développée en Europe. Depuis le premier colloque international sur la chirurgie coronarienne à invasion minimale, à Utrecht, en septembre 1995, plusieurs milliers de personnes ont été opérées ainsi.

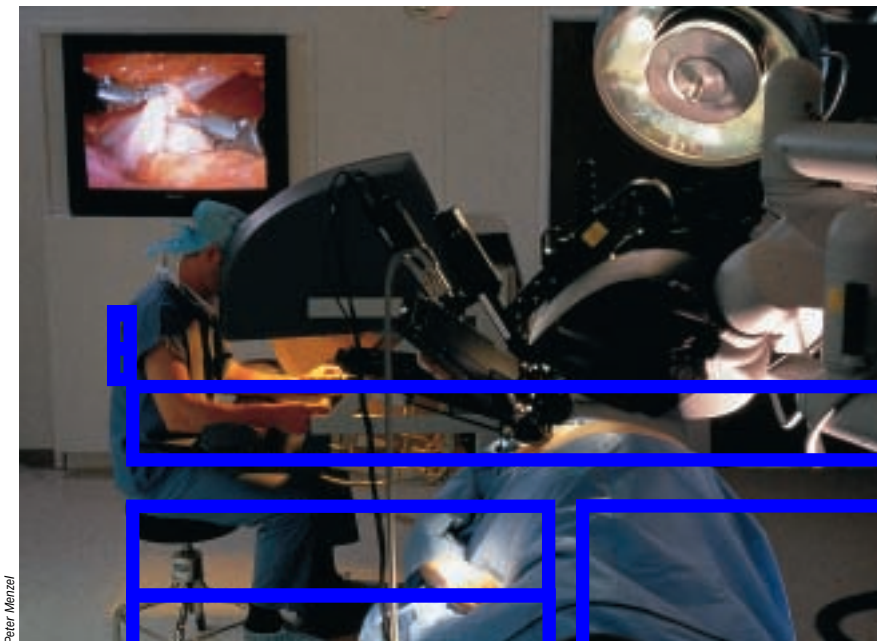
Cependant, la chirurgie sur cœur battant ne supplante pas totalement les pontages par chirurgie classique pour de nombreux malades, l'opération classique reste la meilleure solution. Par exemple lorsqu'un pontage à l'arrière du cœur est nécessaire, la chirurgie sur cœur battant est souvent délicate, car le chirurgien doit extraire le cœur partiellement



4. LE STABILISATEUR CARDIAQUE OCTOPUS immobilise par succion (à droite) une petite région du cœur battant pendant que les chirurgiens suturent le greffon de pontage. L'Octopus est fixé au dispositif métallique utilisé pour écarter la paroi thoracique (à gauche).



5. LE CHIRURGIEN VOIT aujourd'hui l'intérieur de la cavité thoracique grâce à une caméra insérée par une petite ouverture entre les côtes. Assisté par un robot, le pontage coronarien est effectué par de minuscules instruments (à droite, en taille réelle). Les mouvements de la main du chirurgien sur la console d'un ordinateur sont transmis aux instruments chirurgicaux (à gauche).



Peter Menzel

6. PENDANT UN PONTAGE CORONARIEN avec un robot, le chirurgien (à gauche) est assis à quelque distance du malade. Les autres membres de l'équipe chirurgicale observent le déroulement de l'opération sur un écran (en haut, à gauche).

de la cage thoracique : cette manœuvre déforme l'organe et diminue la quantité de sang qu'il pompe.

Pourtant, ces dernières années, des chirurgiens ont découvert des mesures simples qui évitent ces accidents. Mon collègue P. Gründeman a montré qu'une inclinaison de 15 à 20 degrés vers le bas de la table d'opération empêche une chute excessive de la pression sanguine. À l'hôpital de Recife, Ricardo Lima utilise le péricarde (la membrane qui entoure le cœur) pour soulever partiellement l'organe hors de la cage thoracique et accéder à l'envers du cœur sans compromettre la pression sanguine. Il a été suivi par de nombreux confrères.

D'autres solutions sont étudiées. Ainsi, en chirurgie abdominale, des opérations, telle l'ablation de la vésicule biliaire, s'effectuent par des incisions de la taille d'un trou de serrure, grâce à la chirurgie endoscopique : les chirurgiens insèrent un tube rigide relié à une caméra vidéo miniature (l'endoscope) par une incision et les instruments chirurgicaux par deux autres incisions : l'image de l'endoscope guide les mouvements du chirurgien. Pourquoi n'opère-t-on pas le cœur ainsi ?

À Stanford, des médecins ont mis au point des procédés de chirurgie cardiaque par endoscopie dans un thorax fermé et sur cœur arrêté, tandis que le malade est relié à une machine cœur-poumon.

La connexion à une machine cœur-poumon, ainsi que l'arrêt du cœur sans ouverture de la cage thoracique requièrent divers tubes et des cathéters manipulés à partir de l'aîne. Cependant, l'opération n'est pas toujours bien supportée. De plus, la suture d'un pontage est très exigeante, mais la palette des instruments actuels de la chirurgie endoscopique est restreinte et la marge de manœuvre dans une cage thoracique fermée est étroite : les essais d'opération cardiaque par endoscopie ont été abandonnés après trois tentatives. Apparemment, on doit effectuer des ouvertures plus grandes (de six à huit centimètres) pour obtenir des sutures fiables des greffons aux artères coronaires : au milieu de l'année 2000, plus de 6 000 malades avaient subi une telle opération des coronaires.

Un robot chirurgien

Les chirurgiens opéreront-ils un jour un cœur battant dans la cage thoracique fermée ? Les insuffisances de l'endoscopie classique seront peut-être compensées par des robots qui manipuleraient à distance des instruments chirurgicaux. Avec un tel équipement, les médecins visualisent en trois dimensions l'intérieur de la cavité thoracique, et les actions de l'opérateur sur une console seraient transmises aux instruments chirurgicaux dans le

thorax, sans tremblement et avec plus de précision.

Alain Carpentier et Didier Loulmet, de l'hôpital Broussais, à Paris, puis Friedrich Mohr, Volkmar Falk et Anno Diegeler, de l'Université de Leipzig, ont été les pionniers d'une telle chirurgie coronarienne à thorax fermé, avec cependant l'assistance d'une machine cœur-poumon. En 1998, ils ont combiné le procédé mis au point à l'Université de Stanford et le système de chirurgie endoscopique robotisé, nommé da Vinci.

En septembre 1999, à l'Université de l'Ontario, Douglas Boyd a effectué la première opération assistée par ordinateur sur cœur battant et dans une cage thoracique close. Toutefois, l'opération a duré presque une journée. Au milieu de l'année 2000, des chirurgiens de Munich, de Leipzig, de Dresde et de Londres ont réduit le temps d'opération à moins de cinq heures pour des pontages coronariens simples, sur cœur battant et cage thoracique fermée ; les opérations ont été satisfaisantes.

Les techniques robotiques s'imposent progressivement. Bientôt les chirurgiens devront s'entraîner à la pratique des opérations coronariennes par endoscopie, de la même manière que les pilotes s'exercent dans des simulateurs de vol. Ensuite, d'autres innovations faciliteront sans doute la chirurgie des maladies cardiaques coronariennes, par exemple, grâce à un système à pression en cours de mise au point, les chirurgiens fixeront rapidement un pontage sans point de suture.

Les pontages coronariens seront peut-être un jour « dépassés » ; en attendant, la prévention des maladies cardiaques coronariennes et le perfectionnement constant des techniques restent des priorités.

Cornelius BORST est professeur de cardiologie expérimentale au centre médical de l'Université d'Utrecht, aux Pays-Bas.

Cornelius BORST et Paul F. GRÜNDEMANN, *Minimally Invasive Coronary Artery Bypass Grafting : An Experimental Perspective*, in *Circulation*, vol. 99, n° 11, pp. 1400-1408, 23 mars 1999.

Minimally Invasive Cardiac Surgery, sous la direction de Robbin G. Cohen et al., Quality Medical Publishing, 1999.

Minimal Access Cardiothoracic Surgery, sous la direction d'Anthony P.C. Yim et al., W.B. Saunders Company, 2000.

Comment rationaliser les décisions

JOHN SWETS • ROBYN DAWES • JOHN MONAHAN

Des outils mathématiques d'aide à la prise de décisions améliorent les diagnostics en médecine ou dans l'industrie.

Une opacité sur une mammographie est parfois le signe d'une tumeur. Doit-on l'opérer? Un compte qui étudie les dossiers de mise en liberté conditionnelle est alerté par le comportement violent d'un détenu : doit-on, malgré tout, le libérer? Dans un aéroport, un technicien observe une image anormale lors d'une inspection aux ultrasons d'une structure d'avion. Est-elle le signe d'une fissure inquiétante?

Souvent, les décisions sont difficiles à prendre car les données sont incomplètes ou ambiguës. En médecine, en sécurité publique, dans l'industrie, dans les administrations, dans l'armée, au gouvernement, les enjeux sont parfois considérables, car des vies humaines sont en jeu.

Toutefois, les méthodes d'aide à la décision sont presque les mêmes, quels que soient les domaines. Des améliorations méthodologiques dans un domaine profitent donc généralement aux autres domaines. Ces méthodes, dont deux au moins sont validées, sont malheureusement méconnues ou inutilisées. L'une d'entre elles améliore la «pertinence» d'une décision : elle augmente la probabilité que la décision prise soit la bonne. L'autre améliore l'«utilité» d'une prise de décision ; elle évite les alertes inutiles. Ces méthodes sont statistiques et concernent donc un grand nombre d'individus ; au niveau individuel, elles influencent le choix du traitement thérapeutique, mais n'établissent pas une prédiction de l'évolution de la maladie d'une personne.

La décision n'impose pas toujours la mathématisation des problèmes. Dans certains domaines, telles la médecine

ou la météorologie, où de nombreux facteurs interviennent, les outils statistiques ne sont qu'une aide à la décision, les examens complémentaires ou les données nouvelles facilitant les décisions. En revanche, dans d'autres domaines, les analyses statistiques se sont souvent révélées plus pertinentes que les jugements subjectifs, même formulés par des professionnels expérimentés.

Nous examinerons ici les décisions limitées à deux éventualités. Une tumeur est-elle présente, oui ou non? Le délinquant peut-il être libéré sans risque de récidive, oui ou non? L'aile de l'avion est-elle défectueuse, oui ou non? Les nombreux problèmes font apparaître des choix entre plus de deux possibilités, mais les décisions oui/non prédominent. Toutefois, on est généralement confronté à quatre solutions (et non deux) : les «vrais positifs» (une opacité se révèle être une tumeur) et les «vrais négatifs» (l'absence d'opacité correspond à une absence de tumeur), ou la décision prise est dans l'ambiguïté ; les «faux positifs» (il n'y a pas de tumeur, malgré une opacité) et les «faux négatifs» (il y a une tumeur sans opacité), pour lesquelles la décision prise risque de ne pas être la bonne.

Les outils du métier

Si les tests diagnostiques donnaient les réponses attendues, les statistiques d'aide à la décision seraient évidemment inutiles. Toutefois, les résultats bruts des tests doivent généralement être interprétés. Par exemple, les ophtalmologistes mesurent la pression intraoculaire pour dépister les glau-

comes, des compressions du nerf optique et un accroissement de la dureté du globe, qui conduisent parfois à la cécité. Une pression très faible indique un œil sain, tandis qu'une pression élevée signale un glaucome... mais les valeurs intermédiaires?

Les informations statistiques éclairent partiellement la situation. Pour simplifier, admettons que la pression soit la seule mesure diagnostique disponible pour le dépistage du glaucome. Supposons qu'une pression inférieure à 10 (avec des unités arbitraires) corresponde toujours à un œil sain, que les pressions supérieures à 40 témoignent toujours d'un glaucome ; les mesures comprises entre 10 et 40 s'observent aussi bien sur des yeux sains que sur des yeux malades.

Pour élaborer des outils diagnostiques fiables, on doit d'abord identifier une grande population d'individus dont on mesure la pression intraoculaire, puis déterminer quels individus ont un glaucome au cours d'une période donnée. Pour chaque valeur de la pression, on détermine alors la probabilité qu'une personne ait un glaucome. Ces probabilités, et d'autres considérations présentées plus loin, définissent un point de partage rationnel, ou seuil diagnostique : les scores supérieurs ou égaux à ce seuil conduisent à un diagnostic positif (le patient souffre d'un glaucome), les scores inférieurs à un diagnostic négatif (le patient ne présente pas de glaucome).

Évidemment, un seul test diagnostique procure moins d'informations qu'une combinaison de tests. Pour accroître la pertinence d'un diagnostic, le décideur a intérêt à effectuer plusieurs examens qui fournissent

chacun une information en donnant un poids plus important aux mesures les plus prédictives de la pathologie étudiée. Des algorithmes mathématiques qui définissent les meilleurs tests et qui calculent le risque de glaucome d'après l'analyse des résultats combinés sont des méthodes statistiques prédictives.

Les données totalement objectives, telle la mesure de la pression, ne sont pas les seules utiles pour l'amélioration de la pertinence des méthodes statistiques prédictives ; des impressions subjectives peuvent également être quantifiées. On les objective en réalisant, par exemple, une liste de critères perceptibles (comme l'irrégularité d'un grain de beauté éventuellement malin), classés sur une échelle de 1 à 5.

On estime la pertinence globale des méthodes statistiques prédictives en étudiant les courbes ROC (de l'an-

glais *Receiver Operating Characteristic*, qui signifie «courbes des caractéristiques de l'efficacité»). Ces courbes ont été utilisées pour la première fois au cours de la Seconde Guerre mondiale quand on a étudié la fiabilité des radars : on voulait quantifier leur capacité à distinguer les interférences aléatoires (le bruit de fond) du signal indiquant réellement la présence d'avions ennemis.

Les analyses des courbes ROC donnent une évaluation de la proportion d'individus atteints (et non atteints) en fonction d'une valeur seuil choisie : quelle proportion d'individus réellement atteints par la maladie étudiée seront dépistés (vrais positifs)? Et quel pourcentage d'individus sains seront par erreur considérés comme malades (faux positifs, ou fausses alarmes)?

Pour chaque seuil choisi, on détermine la proportion relative de vrais

positifs par rapport aux faux positifs. On obtient ainsi un graphe dont la courbure indique la pertinence : plus elle est forte, plus la prédiction est pertinente, car plus le rapport des vrais positifs sur les faux positifs est élevé.

Les vrais positifs et les faux positifs ne sont pas les seuls résultats possibles. Un diagnostic fondé sur des seuils donne aussi des vrais négatifs (des individus correctement estimés sains) et des faux négatifs (des individus malades non dépistés). Ces résultats complémentaires des précédents sont ignorés lors de la construction des courbes ROC. Par exemple, une proportion de vrais positifs égale à 80 pour cent correspond à une proportion de faux négatifs égale à 20 pour cent, ce qui représente une erreur considérable (malade sur cinq n'est pas dépisté).

Comme les méthodes diagnostiques qui distinguent parfaitement les



1. LES QUESTIONS auxquelles un expert doit répondre par oui ou par non abondent, en médecine comme dans de nombreux autres domaines. Des méthodes statistiques confirmées, mais sous-utilisées, augmentent les chances de faire le bon choix.

individus malades des individus sains sont rares, chaque organisme doit décider de l'importance qu'il accorde à la détection de tous les vrais positifs ou de la majorité d'entre eux. Ce choix a des conséquences, car à un plus grand nombre de vrais positifs correspond également un plus grand nombre de faux positifs. Un seuil doit être adapté à chaque situation.

Revenons à l'exemple du glaucome. Les médecins n'utilisant que la pression intraoculaire trouveront la plupart des cas de glaucome s'ils choisissent un seuil diagnostique très «laxiste», égal à 10, par exemple (en supposant que l'échantillon de population du test ait montré qu'une personne sur deux atteintes d'un glaucome avaient une pression supérieure à ce seuil). Avec un tel choix, de nombreuses personnes saines seront déclarées malades ou faux positifs, et

inutilement inquiétées et traitées. Pour réduire ce type d'erreurs, les médecins peuvent choisir un seuil diagnostique assez «sévère», égal à 35 par exemple (très peu de personnes saines de l'échantillon avaient une pression aussi élevée), mais on dépisterait alors moins de la moitié de tous les individus malades ; les nombreux faux négatifs seraient alors privés de traitement.

On choisit le seuil en fonction des proportions de faux négatifs et de faux positifs qu'il engendre (ces proportions dépendent de la prévalence de la maladie dans la population étudiée). Les statisticiens ont identifié des principes et mis au point des outils mathématiques pour déterminer le seuil optimal. Par exemple, on préfère un seuil «laxiste» quand une maladie est fréquente dans une population ou quand il est très important de détec-

ter tous les cas avérés ; en revanche, le seuil doit être «sévère» quand la maladie est rare ou que le coût du traitement est élevé (en dehors de toute considération éthique).

Bien que les méthodes statistiques prédictives et les courbes ROC soient peu utilisées par les médecins, les exemples de leur intérêt abondent. L'une des meilleures illustrations provient de la psychiatrie.

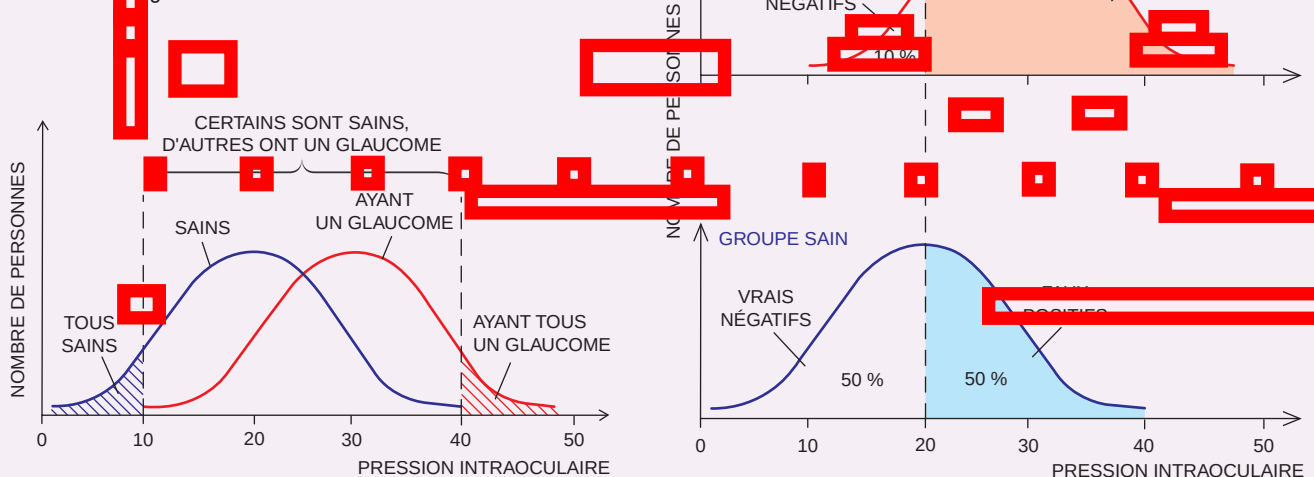
Régulièrement, les psychiatres et les psychologues doivent évaluer si des individus incarcérés ou perturbés risquent de devenir violents. Les personnes qui menacent la sécurité d'autrui doivent être identifiées et traitées, autant pour leur propre bien que pour celui de leur entourage. En revanche, le médecin n'a pas le droit d'intervenir dans la vie de personnes qui n'ont pas besoin de soins. Contre toute attente, en 1993, l'étude la plus

Améliorer la prise de décision en quatre étapes

Comment les médecins s'assurent-ils que le test diagnostique qu'ils utilisent est le plus pertinent, et qu'il distingue au mieux les individus malades de ceux qui sont sains ? La construction de courbes ROC est l'outil principal de mesure du degré de pertinence. L'étude d'un exemple permet de mieux comprendre son élaboration. Suivons les étapes d'une analyse qui évalue la pertinence du diagnostic du glaucome à partir d'une mesure de la pression intraoculaire des patients.

DEUXIÈME ÉTAPE : pour chaque valeur de la pression, on détermine la probabilité qu'un diagnostic positif soit correct pour les patients. Cette probabilité est déterminée à partir de la fraction de l'échantillon pour laquelle la prédiction est exacte. L'axe vertical de la courbe représente 100 pour cent de chaque population. Si le seuil de pression intraoculaire choisi, pour la détermination des malades, est de 20, alors 90 pour cent des personnes ayant réellement un glaucome sont correctement diagnostiquées (vrais positifs) ; en revanche, 50 pour cent de personnes saines sont déclarées atteintes (faux positifs).

PREMIÈRE ÉTAPE : à partir d'un grand nombre de personnes dont on connaît la pression intraoculaire et la présence, ou non, d'un glaucome, on sépare les sujets sains et ceux qui sont atteints d'un glaucome. On marque sur un graphique le nombre d'individus pour chaque valeur de la pression intraoculaire. Le graphique de cette population montre que les pressions comprises entre 10 et 40 ne révèlent pas de façon fiable les personnes qui souffrent d'un glaucome.



complexe d'évaluation clinique par les méthodes statistiques a présenté un étonnant manque de pertinence : le diagnostic établi par des médecins qui examinaient des personnes amenées au service des urgences d'un grand hôpital psychiatrique était faux dans un cas sur deux, comme si on l'avait fait au hasard ; autrement dit, les psychiatres des urgences ne pouvaient absolument pas prévoir qui, parmi les malades, commettrait un acte violent au cours des six mois suivant le diagnostic.

Après cette étude, les biostatisticiens ont mis au point des méthodes statistiques prédictives pour évaluer le risque de violence. Le «guide d'évaluation du risque de violence», l'une des méthodes les plus étudiées, mesure douze variables, parmi lesquelles une évaluation du risque de développer une déficience mentale et l'inadapta-

tion à l'école primaire. Lors d'une vérification des prévisions établies par cette méthode, les délinquants sortis d'un hôpital psychiatrique étaient répartis en deux groupes : ceux pour lesquels le risque de récurrence violente était «élevé», ceux pour lesquels le risque était «faible». Cinquante-cinq pour cent des personnes du groupe à risque élevé ont récidivé, et seulement 19 pour cent l'ont fait dans le groupe à risque faible. La pertinence du guide était ainsi bien supérieure à celle d'un tirage au sort.

Une méthode statistique prédictive plus récente s'est révélée encore meilleure pour la même prédiction. Toutefois, les méthodes statistiques restent controversées : les médecins doivent-ils les utiliser comme avis ou bien doivent-ils se conformer strictement aux indications des méthodes statistiques prédictives?

Les méthodes statistiques prédictives se révèlent également très utiles pour le dépistage du cancer du sein. Lors des examens classiques, les radiologues interprètent les mammographies de manière subjective.

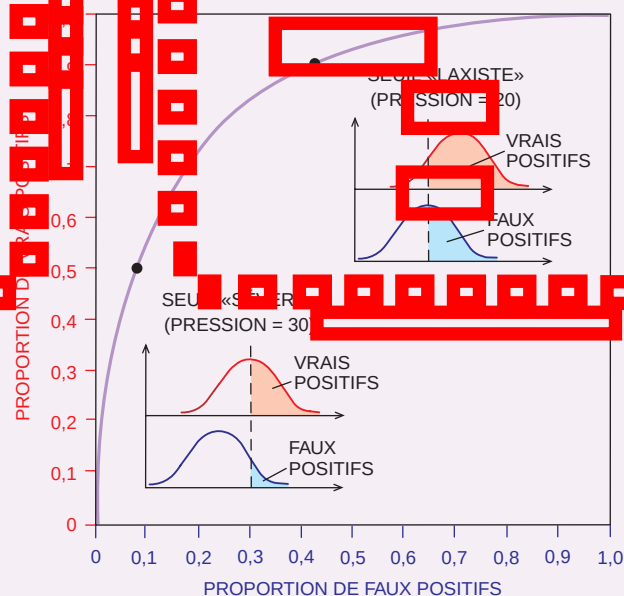
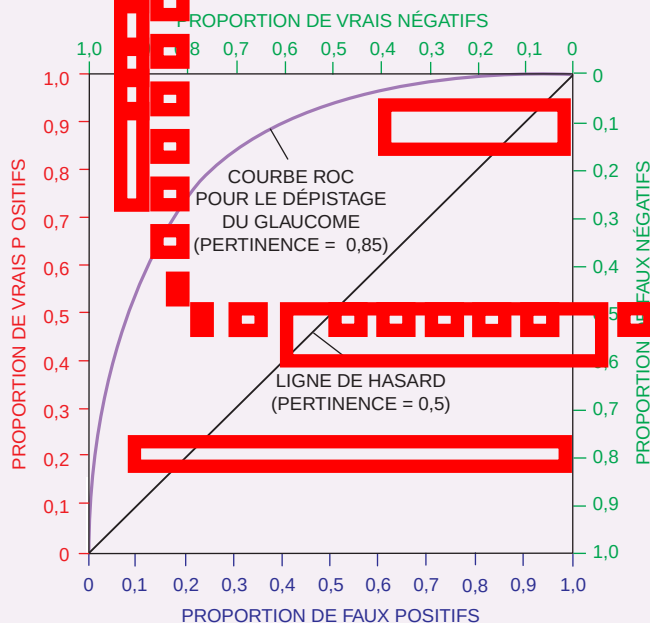
Le dépistage du cancer

Un protocole d'examen établi par des radiologues spécialisés, qui associe des méthodes statistiques prédictives, conduit à une interprétation plus pertinente : le protocole invite d'abord le radiologue à examiner des critères qualitatifs (tels que la régularité du contour de la masse), identifiés par des spécialistes du sein ; puis on calcule un risque de cancer du sein, dont on tient compte pour poser le diagnostic. Les informations apportées par l'analyse statistique sont une aide réelle : les radiologues non spécialisés ayant utilisé cette

TROISIÈME ÉTAPE : on construit la courbe ROC en reportant, pour chaque valeur seuil, la proportion de vrais positifs en fonction de celle des faux positifs. Si l'on obtient une droite, on doit conclure que le test diagnostique a 50 pour cent de chances de conduire au bon diagnostic : un tirage à pile ou face serait aussi efficace. Plus la courbe s'incurve, plus le test est pertinent (le rapport des vrais positifs sur les faux positifs augmente). La pertinence est l'aire de la surface sous la courbe ; elle est nulle si le test n'est pas pertinent et maximale si le dépistage du cancer est peu prédictif.

QUATRIÈME ÉTAPE : lorsque la pertinence atteinte est acceptable, on doit choisir un seuil pour un diagnostic définitif. Le choix du seuil doit fournir une proportion de vrais positifs élevée sans entraîner une pro-

portion inacceptable de faux positifs. Chaque point de la courbe représente un seuil particulier, allant du plus «sévère», en bas à gauche, au plus «laxiste», en haut à droite. Les seuils «sévéres» (*cartouche inférieure*) limitent le nombre de faux positifs au prix de nombreuses personnes malades non diagnostiquées (forte proportion de faux négatifs, c'est-à-dire faible proportion de vrais positifs). Les seuils «laxistes» (*cartouche supérieure*) augmentent le nombre de personnes malades diagnostiquées comme étant malades). Le seuil optimal pour une population donnée dépend de facteurs tels que la gravité de la maladie diagnostiquée, la prévalence de la maladie dans la population, les dommages potentiels pour ceux qui sont diagnostiqués, ainsi que les coûts économiques, sociaux ou émotionnels pour les faux positifs.



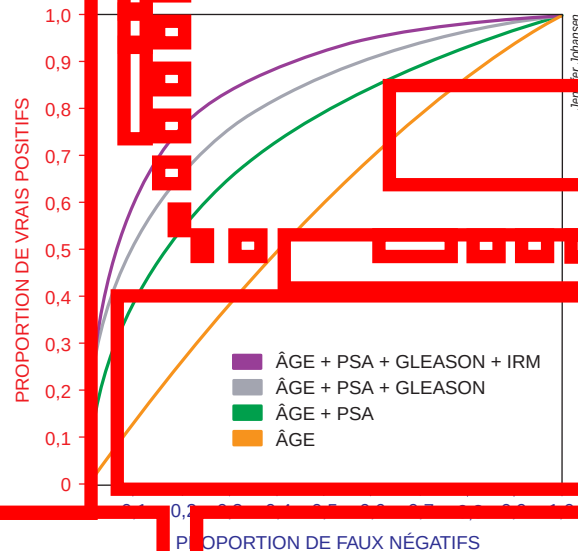
méthode ont une interprétation nettement plus pertinente, comparable en précision à celle des spécialistes.

Des méthodes statistiques prédictives sont également utilisées par les médecins pour le traitement du cancer de la prostate. Une de ces méthodes donne d'excellents résultats. Lorsqu'un homme est déclaré cliniquement atteint d'un cancer de la prostate (sur la base d'un examen médical, d'une biopsie simple et d'examen non invasifs), la question du traitement le plus adapté s'impose. Ni la prostatectomie ni la radiothérapie conformationnelle (radiation concentrée sur la prostate) n'élimineront la tumeur si elle a envahi les tissus environnants, ou si elle a formé des métastases. Les médecins cherchent donc à déterminer le stade évolutif de la tumeur avant de choisir un traitement.

Malheureusement, des tumeurs qui semblent initialement confinées à la prostate se révèlent souvent être à des stades bien plus évolués. Pendant des années, les médecins avaient peu de moyens fiables pour prédire chez quels patients l'envahissement de la tumeur dépassait la glande. Aujourd'hui, ils utilisent un tableau admis par la communauté médicale, qui indique le stade d'une tumeur en fonction des caractéristiques de la tumeur humaine et de la présence de métastase.

Les chercheurs qui ont élaboré ces tableaux savaient que trois indices donnaient chacun une valeur prédictive indépendante : le « stade clinique de la tumeur » (déterminé par des examens non invasifs, de taille et d'étendue) ; la concentration sanguine en antigène prostatique spécifique (PSA, une protéine libérée par les cellules de la prostate) et le score de Gleason (un indicateur de l'agressivité de la tumeur, établi sur les analyses histologiques d'une biopsie). Les biostatisticiens ont élaboré une méthode statistique prédictive qui intègre la plupart des combinaisons de résultats pour ces trois variables et qui calcule la probabilité que le diagnostic initial de « cancer localisé » soit exact. Les probabilités sont finalement présentées sous la forme d'un tableau facile d'utilisation.

L'ensemble des exemples proposés a montré que la prise de décision repose sur une valeur de seuil critique. Le pro-



2. LES COURBES ROC comparent la pertinence de diverses mesures pour l'évaluation du stade évolutif d'un cancer de la prostate. Les variables étudiées sont : l'âge des patients, la concentration sanguine en PSA (antigène prostatique spécifique), l'agressivité de la tumeur (représentée par le score de Gleason) et l'aspect de la tumeur lors d'un examen d'imagerie par résonance magnétique (IRM). Le modèle prenant en compte les quatre variables s'est révélé être le meilleur (courbe du haut).

blème principal des analyses statistiques est la difficulté de choisir des seuils diagnostiques rationnels, comme pour les tests de dépistage du virus du SIDA (VIH).

Le dépistage du VIH commence avec un test simple, qui détecte la présence d'anticorps anti-VIH, molécules produites par l'organisme lorsque le système immunitaire commence à réagir à l'intrusion du virus. Quand le résultat est positif (à partir d'un certain taux d'anticorps), les laboratoires pratiquent un autre test plus complexe. On limite ainsi le nombre de faux positifs. Les tests sur les anticorps sont souvent problématiques, car ils ne reposent pas tous sur les mêmes valeurs seuils et n'ont pas tous la même pertinence. Des seuils variés seraient justifiés si chaque test était destiné à une population distincte, mais ce n'est pas le cas.

Les seuils sont également gênants d'un autre point de vue. Initialement mis au point pour distinguer les dons de sang sains des échantillons contaminés, on les a conservés pour l'identification des personnes infectées par le virus. Or, si l'on n'hésite pas à jeter 500 millilitres de sang sain déclaré – à tort – positif (faux positif), il est beaucoup plus grave d'inquiéter une personne saine en l'envoyant subir d'autres tests de dépistage du VIH parce que le premier test aura été faussement positif. Pis encore, les seuils initiaux

ont été utilisés sans distinction pour tous les donneurs de sang : ceux qui ont un faible risque d'infection, comme ceux qui ont un risque élevé. Lors de l'évaluation des individus à risque élevé, les seuils devraient être plus « laxistes » que pour les populations à risques faibles (afin d'augmenter la probabilité de découverte), même si ce choix augmente le nombre de faux positifs.

Ces dernières années, les tests de confirmation sont devenus plus précis, tandis que des traitements contre l'infection virale prolongent la vie et améliorent l'état de santé des malades. Les diagnostics de faux positifs sont plus rares aujourd'hui, et les personnes séropositives sont mieux dépistées qu'autrefois. Aujourd'hui, la question n'est plus celle du seuil de séropositivité, mais reste celle du dépistage des personnes contaminées et qui l'ignorent.

De la météorologie à l'aéronautique

Les médecins ne sont pas les seuls à utiliser les méthodes statistiques prédictives : les météorologues ont adopté cet outil pour leurs prévisions il y a plus de 15 ans. *Météo France*, par exemple, publie régulièrement en données météorologiques des programmes statistiques qui estiment le risque de tornades, d'ouragans, de fortes pluies, etc. *Météo France* transmet ces prédictions objectives aux météorologues régionaux, qui les modifient en fonction de nouvelles informations ou de facteurs que les programmes informatiques n'auraient pas correctement pris en compte.

Enfin, d'autres professions ont également adopté les méthodes statistiques d'aide à la décision. Par exemple, les commissions universitaires américaines cherchent ainsi à mieux évaluer les candidats qui postulent pour des études en troisième cycle. Ils utilisent des méthodes prédictives pour prévoir les notes des étudiants à partir des notes obtenues au cours du cursus universitaire et des résultats aux tests d'aptitude pour l'entrée en troisième cycle. Les étudiants dont les notes sont supérieures à une valeur seuil préétablie sont généralement reçus, et ceux dont

les notes sont inférieures à une valeur seuil inférieure sont généralement refusés. Le comité évalue ensuite, plus subjectivement, les qualifications des postulants qui n'ont pas été admis par les méthodes statistiques prédictives. En France, les dossiers ne sont pas sélectionnés de cette façon statistique, mais par des évaluations... peut-être plus subjectives.

Les méthodes objectives d'établissement des seuils ont aussi été sous-utilisées dans l'industrie aéronautique. Lors de l'entretien des avions, on doit rechercher les défaillances graves, mais rares, telles que des fissures dans les ailes. Quand un diagnostic d'aile fissurée n'est pas fait, les conséquences sont catastrophiques : de nombreux passagers meurent si l'avion s'écrase. En revanche, une décision faussement positive met inutilement l'avion hors service, ce qui perturbe le service et coûte cher aux compagnies. À première vue, les bénéfices et les coûts sont en faveur d'un seuil «laxiste», les vies n'ayant pas de prix. Cependant, comme ces fissures sont rares, un seuil «laxiste» entraînerait un nombre incalculable de faux positifs. Aujourd'hui, aucune technique statistique ne résout le dilemme.

Les utilisateurs d'alarmes de cockpits (comme les compagnies aériennes ou l'armée) n'ont pas non plus réussi à choisir quels seuils de décision adopter. En vol, les alarmes se déclenchent en de multiples circonstances : lorsque les appareils de mesure enregistrent qu'un autre avion est trop proche, quand l'avion se rapproche trop du sol, quand un moteur est en panne ou quand un vent violent balaie la piste d'atterrissage. Toutefois, ces alarmes sonnent de façon intempestive, en raison du manque de précision des détecteurs ou du laxisme de leurs seuils. Aussi les pilotes hésitent-ils à tenir compte de ces alarmes trop fréquentes ; les interventions systématiques requerraient d'ailleurs de perturber les vols. Toutefois, en raison du grand nombre de fausses alarmes, les pilotes risquent de finir par les ignorer ou par réagir trop lentement à une véritable urgence.

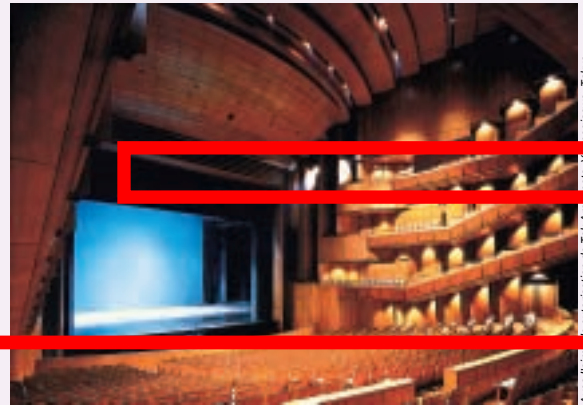
Les méthodes statistiques prédictives augmentent la pertinence des diagnostics, et les formules pour établir des seuils de partage amélioreraient l'utilité de ces choix. Ces outils ont encore d'autres avantages : en standardisant les analyses sur lesquelles se fondent les diagnostics, les méthodes

Une question de goût

Les architectes et les amateurs de vin ont trouvé deux applications amusantes des méthodes statistiques prédictives. La «méthode architecturale», qui s'applique à des salles d'opéra, fut mise au point grâce à des chefs d'orchestre qui avaient noté la qualité sonore de 23 salles. Puis des ingénieurs acousticiens mesurèrent physiquement différentes propriétés acoustiques (telles que le délai entre les sons directs et les sons réfléchis, la diffusion des ondes sonores provoquée par les irrégularités du plafond et des murs) de chacune de ces 23 salles. Les analyses statistiques ont ensuite révélé quelles sont les caractéristiques des meilleures salles d'opéra. On a alors amélioré la conception de nouvelles salles.

En œnologie, une méthode statistique prévoit la qualité future des Bordeaux rouges (qualité mesurée par leur prix de vente aux enchères) à partir d'un test sur des échantillons de vins encore jeunes et imbuables. Auparavant, les experts prévoient la qualité future en sentant et en goûtant le nouveau produit. Il y a dix ans, les œnologues ont constaté que les années marquées par un hiver humide, des mois d'août et de septembre secs et une période de croissance chaude produisent l'excellent vin.

Ils ont alors formulé l'«équation bordelaise», une méthode statistique prédictive qui combine les conditions météorologiques et les années de vieillissement pour définir la probabilité qu'un vin soit excellent dans plusieurs années. Cette équation explique 83 pour cent des divergences de prix des vins rouges de Bordeaux à maturité, mis aux enchères. Pourtant cette équation, suscita une vive réaction des œnologues, qui la décrièrent violemment dans un article.



Le nouvel opéra de Tokyo, dont les qualités acoustiques ont été validées par des méthodes statistiques de prévision du son.

statistiques facilitent l'identification des paramètres importants. Elles fournissent aux décideurs des arguments faciles à communiquer et qui leur permettent de préciser de vagues impressions. De surcroît, elles facilitent la formation des débutants.

Ces outils sont parfois accueillis avec réticence, notamment quand les experts craignent qu'ils ne les remplacent. Les médecins également veulent garder la maîtrise de leurs diagnostics et de leurs recommandations, et donner un compte rendu argumenté de leur décision. Les résultats d'une méthode statistique prédictive sont parfois difficiles à prendre en compte, surtout quand la logique qui sous-tend l'analyse n'est pas évidente.

Nous comprenons ces préoccupations, mais les bénéfices apportés par les outils statistiques justifient certainement que les personnes amenées à prendre des décisions, qui ont la vie d'autrui entre leurs mains, y réfléchissent.

John A. SWETS est spécialiste de la «politique» de santé américaine et radiologue à Boston. Robyn DAWES est professeur de sciences sociales à l'Université de Carnegie Mellon. John MONAHAN dirige l'association de recherche Loi-santé mentale de la Fondation MacArthur.

A.W. MARTIN et al., *Combination of Prostate-Specific Antigen, Clinical Stage and Gleason Score to Predict Pathological Stage of Localized Prostate Cancer*, in *Journal of the American Medical Association*, vol. 277, n° 18, pp. 1445-1451, 14 mai 1997.

Kenneth A. FREEDBERG et Jeffrey H. SAMET, *Think HIV : Why Physicians Should Lower their Threshold for HIV Testing*, in *Archives of Internal Medicine*, vol. 159, n° 17, pp. 1994-2003, 27 septembre 1999.

John A. SWETS, Robyn M. DAWES et John MONAHAN, *Psychological Science Can Improve Diagnostic Decisions*, in *Psychological Science in the Public Interest* (supplément à *Psychological Science*), vol. 1, n° 1, pp. 1-26, mai 2000.

La route des émeraudes anciennes

GASTON GIULIANI • MICHÈLE HEUZÉ • MARC CHAUSSIDON

Grâce à une sonde ionique, on détermine aujourd'hui la provenance des émeraudes sans les détériorer. Les géologues et gemmologues ont analysé des pierres historiques et reconstitué la route des émeraudes anciennes.

Les premières émeraudes ont-elles été ramassées en Égypte, en Bactriane (la région recouvrant le Nord de l'Afghanistan, le Sud du Turkménistan, de l'Ouzbékistan et du Tadjikistan actuels) ou en Inde du Sud ? Aujourd'hui, les géologues aident les historiens à reconstituer la chronologie de la découverte des gisements d'émeraudes.

En occident, les premiers objets décorés d'émeraudes apparaissent au IV^e siècle avant notre ère, dans l'Empire grec, sous le règne d'Alexandre le Grand. Puis, l'un de ses généraux fonde la dynastie ptolémaïque en Égypte, où les mines d'émeraudes commencent à être exploitées. Bien qu'elles restent assez rares, ces gemmes apportent une touche de vert à la bijouterie, qui les associe avec des perles et d'autres pierres, tels les grenats. Des bracelets allient les émeraudes à des motifs égyptiens, tels les fleurs de lotus et les serpents, et des motifs grecs, tels les cygnes et les couronnes de pampres. À partir du II^e siècle avant notre ère, cette gemme est à la mode dans l'Empire romain : les gisements d'Égypte, qui vient d'être conquise, sont exploités de façon intensive. Les émeraudes, désormais abondantes, s'égrènent le long des colliers, seules ou avec des perles. Elle forment des grappes ou des mosaïques sur les boucles d'oreilles, ornent les bracelets et les pendentifs en or, travaillés selon une technique d'ajourage en dentelle (*l'opus interasile*) très prisée des Romaines.

Dans l'Antiquité, on ne taille pas les émeraudes en facettes. On polit

les angles des cristaux pour renforcer leur éclat et améliorer leur transparence. Après la division de l'Empire romain au V^e siècle, l'orfèvrerie byzantine est richement ornée de gemmes et notamment d'émeraudes, au contraire du reste de l'Europe, où les grenats apportés par les hordes barbares sont omniprésents. Au Moyen Âge, les émeraudes qui rehaussent l'orfèvrerie religieuse de touches de couleur sont souvent des réemplois de pierres de bijoux antiques. Plus, à la Renaissance, les émeraudes colombiennes envahissent l'Europe : les Espagnols et les Portugais en font le commerce dans le monde entier, jusqu'en Inde, où les maharajas les apprécient.

Les émeraudes de Bactriane

Les recherches sur l'histoire des gemmes nécessitent autant de connaissances minéralogiques qu'historiques. Deux géologues (Gaston Giuliani et Marc Chaussidon) et une historienne du bijou (Michèle Heuzé) ont collaboré pour retrouver la route des émeraudes anciennes. Depuis 1998, les géologues disposent d'une sonde ionique à Vandœuvre, près de Nancy (voir l'encadré de la page 64) : à plusieurs reprises, cette sonde a confirmé l'origine d'émeraudes historiques, mais, plus intéressant encore, elle a permis de reconstruire les pans oubliés de l'histoire de ces gemmes et de faire tomber certaines légendes.

Au IV^e siècle avant notre ère, le philosophe grec Théophraste rapporte les récits d'Alexandre le Grand, qui

fait consigner tout ce qu'il voit, notamment les us et coutumes des pays qu'il traverse lors de ses conquêtes (son empire s'étend jusqu'en Bactriane). Alexandre doit cette soif de connaissance à son précepteur, Aristote, dont Théophraste, le disciple, écrit : «L'émeraudes qui servait à orner des vases d'or était portée en bague, taillée de forme concave pour que la lumière puisse y rassembler ses rayons. [...] Les émeraudes dont on se sert pour orner les coupes et autres vases d'or viennent de Bactriane vers le désert ; on y va à cheval pour les chercher au temps des vents étésiens ou des vents d'Est annuels : on les voit dans ce temps-là parce que les sables sont agités avec violence par ces vents. Cependant, les pierres que l'on y trouve ne sont pas grosses» (les vents étésiens sont des vents violents, tels ceux qui soufflent en Méditerranée orientale pendant l'été). En l'absence de preuves tangibles de l'existence des gisements de Bactriane, les historiens ne croyaient plus aux vieux écrits de Théophraste, jusqu'à ce que l'on analyse, avec la sonde ionique, une boucle d'oreille gallo-romaine retrouvée dans l'Ain, à Miribel, et conservée au Muséum national d'histoire naturelle de Paris.

Les écrits de Théophraste sont cités en référence jusqu'au XIII^e siècle, mais

1. LES ÉMERAUDES DE CE TRÉSOR ROMAIN du III^e siècle, découvert à Eauze, dans le Gers, proviendraient des gisements égyptiens, exploités de façon intensive dans l'Antiquité (selon les analyses destructives menées à la fin des années 1980 au laboratoire de géochimie de l'Université Pierre et Marie Curie).

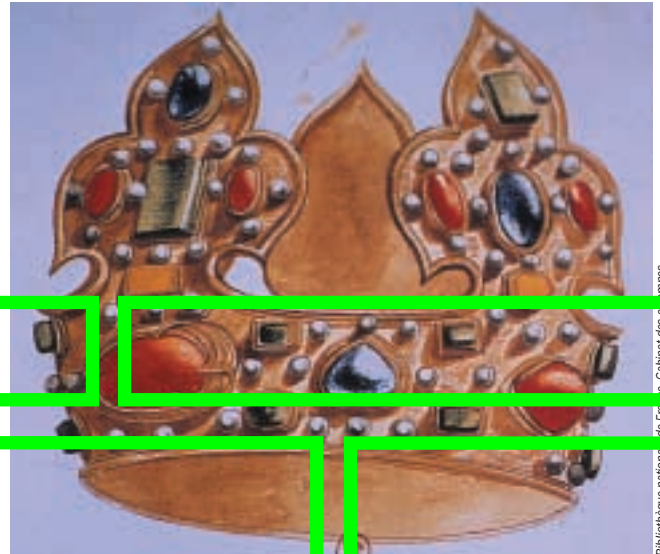
au XVII^e siècle, Jean-Baptiste Tavernier, un grand courtier amateur de gemmes voyageant en Asie, ne trouve aucune trace d'émeraude et en nie toute existence à l'Est : «C'est une erreur ancienne. [...] Jamais l'Orient

n'en a produit, ni dans la terre ferme ni dans les îles et qu'en ayant fait une exacte perquisition dans tous mes voyages, personne n'a jamais su marquer aucun lien de l'Asie où elle se trouve.» Dans les années 1960, on

pense toujours que les émeraudes de Bactriane sont un mythe.

Or, dès 1958, un berger découvre au Pakistan des gisements qui seront exploités à partir des années 1960. Entre 1970 et 1980, on découvre





2. CETTE ÉMERAUDE, EXCEPTIONNELLE PAR SA TAILLE (3 centimètres sur 2,4 centimètres), est d'origine autrichienne. Il s'agit peut-être de la «table» d'émeraude figurée sur l'aquarelle de droite, au centre du lys frontal de la Sainte Couronne (aujourd'hui perdue). Si tel est le

cas, l'exploitation du gisement autrichien d'Habachtal remonterait au XII^e ou au XIII^e siècle. Cependant, le polissage de la pierre ressemble à celui que les Romains pratiquaient. Selon cette dernière observation, l'exploitation du gisement autrichien remonterait à l'Antiquité.

d'autres gisements en Afghanistan. On commence à soupçonner que Théophraste avait raison : la provenance pakistanaise de l'émeraude fixée sur la boucle d'oreille gallo-romaine du Muséum, que la sonde ionique a révélée, confirme la véracité de ses propos.

Cependant, aucun texte historique n'indique de quand date le ramassage de cette émeraude et la découverte du gisement pakistanaise. Dans l'Antiquité, on a peut-être ramassé des émeraudes en Afghanistan (si les vents étésiens dont parle Théophraste sont ceux de la vallée de Panjshir) à l'époque grecque, puis au Pakistan (d'où vient la boucle d'oreille analysée) à l'époque romaine. Cependant, en 329 avant notre ère, Alexandre le Grand se rend près de l'emplacement actuel des mines pakistanaises, en passant à quelques kilomètres de celui des afghanes. Dans cette région, il ne rapporte l'existence que d'un seul gisement, probablement le gisement pakistanaise, qui, selon cette hypothèse, aurait été connu dès l'époque grecque.

La seule certitude sur le gisement de Bactriane est sa nature alluvionnaire : les pierres que l'on ramassait dans le sable y avaient été déposées par le vent et les cours d'eau. Ce gisement était probablement de maigre rendement : en Afghanistan et au Pakistan, les archéologues ont retrouvé des milliers de perles minérales, par exemple en agate, mais aucune en émeraude.

Le jérôme vert de la boucle d'oreille gallo-romaine de Miribel a beaucoup voyagé. Dès le IV^e millénaire avant notre ère, la Bactriane, grâce à son or, à son argent mais, plus encore, à ses lapis-lazuli et ses grenats, est le point de départ d'une route commerciale vers les cités-États d'Irak, puis d'Égypte. La région joua un rôle primordial dans les relations entre l'Inde, le monde des steppes et l'Asie centrale. D'après l'historien romain Pline l'ancien, les caravanes empruntant la route de la soie pour aller vers la Chine ou revenir vers la Méditerranée mettent sept jours pour traverser la Bactriane. Lorsque les Parthes envahissent la région, elles passent par la mer Noire.

Les pierres transitent par les centres commerciaux majeurs, Begram, en Afghanistan, et Taxila, au Pakistan, avant de rejoindre l'Indus, l'océan Indien, le golfe Persique ou la mer Rouge. Est-ce ce long périple qu'a suivi la petite émeraude de la boucle d'oreille du Muséum ? Peut-être. À moins que les ambassadeurs des rois de Bactriane, qui se rendaient fréquemment en Europe à l'époque romaine, ne l'aient apportée dans leurs bagages.

L'autre lointaine provenance des émeraudes dans l'Antiquité est la Scythie. Pline affirme que l'on y trouve les plus belles. Où est ce gisement ? Au I^{er} siècle, la Scythie est un terme générique que les Romains emploient pour désigner les régions lointaines des barbares. Plus tard, à Byzance, on appelle Scythes tous ceux qui ne sont pas des

Grecs. Ainsi, le gisement scythe serait soit en Asie mineure, soit dans l'Oural. Pline n'emploie pas le terme de Germanie, ce qu'il aurait dû faire si le gisement était en Autriche. Toutefois, il n'est pas géographe. Dans l'Oural, dont le gisement n'est en exploitation que depuis 1830, les bijoux retrouvés lors des fouilles archéologiques et datant de l'époque romaine ne comportent aucune émeraude. La question reste ouverte, en attendant qu'un autre gisement soit découvert ou que les futures analyses permettent de répondre à cette question.

Les mines d'Égypte : seize siècles d'exploitation

L'origine pakistanaise de la boucle d'oreille gallo-romaine accrédite les récits de Théophraste, qui ne fait pas référence aux émeraudes d'Égypte. Ainsi, elles n'ont pas encore été découvertes : les datations qui font remonter l'exploitation des gisements égyptiens avant le passage d'Alexandre le Grand (parfois même 13 siècles avant notre ère) ne sont pas fondées. L'extraction commence au plus tard au III^e siècle avant notre ère, sous le règne de Ptolémée IV Philopator (de -221 à -203). Les nombreux bijoux découverts dans les ruines d'Herculanium et de Pompéi témoignent de l'abondance des émeraudes.

Les lapidaires arabes citent quatre variétés d'émeraudes égyptiennes, dont les plus belles sont parfois, mais rarement, gravées des portraits de souverains ptolémaïques : en Scythie, sur les

centaines de gemmes gravées que l'on a retrouvées dans la région de la mer Noire, seules deux sont des émeraudes.

Jusqu'au XIII^e siècle, les émeraudes égyptiennes alimentent l'Europe, le Proche-Orient et l'Inde. Puis les mines d'Égypte s'épuisent et les pillages des anciens tombeaux de la région d'Alexandrie constituent un nouveau type de gisement! En Europe, cette pénurie conduit les orfèvres à réemployer des émeraudes antiques. L'emplacement des mines d'Égypte est peu à peu oublié. L'explorateur français Frédéric Caillaud les redécouvre en 1816, en Haute Égypte, sur la mer Rouge. Au XIX^e et XX^e siècles, on envisage de les réexploiter, mais elles ne sont pas rentables.

L'énigme du gisement autrichien

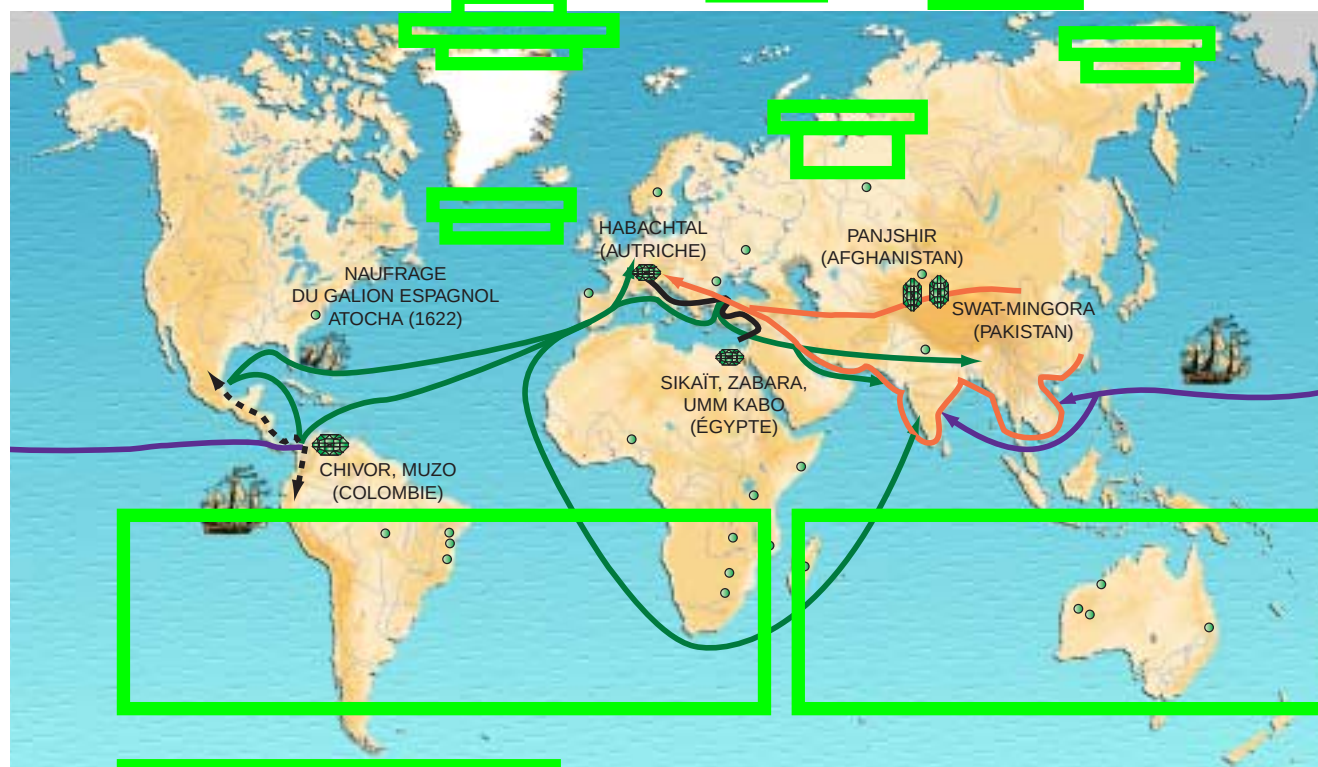
Les orfèvres du Moyen Âge ne connaissent-ils pas le gisement d'Habachtal? On a parfois affirmé, sans aucune preuve, que les Romains le connaissent déjà. La seule émeraude fixée sur un bijou romain (une boucle

d'oreille) est retrouvée en Autriche, de surcroît dans la région d'Habachtal, est vraisemblablement égyptienne, d'après les restes de la roche mère dans laquelle elle a cristallisé (voir l'encadre des pages 62 et 63). L'archevêque de Salzbourg aurait exploité la mine au Moyen Âge. Cependant, le premier document qui la mentionne (une lettre de la veuve de l'archiduc Ferdinand Charles d'Autriche à son frère), ne daterait que de 1622 et le deuxième (une chronique minière) de 1727.

L'analyse de l'émeraude dite de la Sainte Couronne, conservée au Muséum national d'histoire naturelle (voir la figure 2), apporte un nouvel éclairage au débat. Les résultats sont compatibles avec trois provenances : l'Autriche, le Brésil et le Zimbabwe. Ces deux dernières sont exclues du fait de la découverte des gisements au XX^e siècle, postérieurement à la datation présumée de l'usage de la gemme. Celle-ci aurait orné le centre du lys frontal de la Sainte Couronne (voir la figure 2), datée du XII^e ou du XIII^e siècle. Cette pierre, exceptionnelle par sa taille (elle mesure 3 centimètres de longueur

et 2,4 centimètres de large) et par son poids (51,6 carats), a échappé aux ventes évolutives. Sa provenance autrichienne ne bouleverse les données historiques, puisqu'elle prouverait l'existence d'une exploitation dès le Moyen Âge. De plus, son polissage, d'une qualité exceptionnelle, semble avoir été réalisé selon des techniques romaines employées du II^e au IV^e siècle : il laisse en effet voir une antériorité de son ramassage.

Avant que les conquistadores espagnols arrivent en Colombie, au XVI^e siècle, les autochtones exploitaient déjà les gisements d'émeraudes depuis au moins six siècles, comme l'attestent leurs bijoux. D'après leurs légendes, ils les auraient même découverts avant notre ère. Les Espagnols font le commerce des émeraudes qu'ils vont chercher chez les Incas, les Mayas, les Toltèques et les Aztèques (en Équateur, au Pérou et en Amérique centrale). Ils pillent les trésors des souverains locaux. Notamment, l'Espagnol Gonzalo Jimenez de Quesada en aurait rapporté près de 7 000 de son expédition en Colombie, en



3. LA ROUTE COMMERCIALE DES ÉMERAUDES AVANT LE XVIII^e siècle, seuls cinq gisements ont été exploités à des époques diverses (représentés par des cristaux verts). La route de la soie (en rouge) a longtemps été utilisée pour commercialiser les émeraudes d'Orient vers l'Europe. Les Grecs et les Romains ont rapporté les émeraudes égyptiennes en Europe, à partir du III^e siècle avant notre ère (en suivant la route en noir). Avant l'arrivée des conquistadores en Amérique du Sud, les autochtones y effectuaient déjà le

commerce des émeraudes entre le Pérou, l'Équateur et l'Amérique centrale (le long du trajet en pointillés noirs). Puis, les Espagnols en ont rapportées des milliers en Europe, d'où elles repartaient à bord des galions portugais vers les comptoirs indiens de Goa et de Diu (routes en vert). À partir de 1571, les Espagnols sont directement allés en Inde, en partant de Colombie (route maritime en violet). Depuis, de nombreux autres gisements ont été découverts dans le monde entier (points verts).

La formation des émeraudes

L'émeraude est la variété verte du béryl, un minéral constitué de silicium, d'oxygène, de béryllium et d'aluminium (de formule chimique $\text{Be}_3\text{Al}_2\text{Si}_6\text{O}_{18}$), qui cristallise dans le système hexagonal. Sa couleur verte est due à d'infimes quantités de chrome, et parfois de vanadium, qui remplacent des atomes d'aluminium. Les autres variétés de béryl colorées contiennent des substituants différents : la couleur bleu-vert de l'aigue-marine est due à la présence de fer, et celle de la bixbite, rouge, résulte de la présence de manganèse et de césium dans le réseau cristallin.

L'émeraude est un minéral rare, car le chrome (et le vanadium) et le béryllium ont des comportements géochimiques opposés : le chrome se concentre dans les roches issues du manteau (l'enveloppe de la Terre située sous la croûte continentale ou océanique), tandis que le béryllium a une affinité pour la croûte continentale et se concentre dans les magmas granitiques. La formation d'une émeraude résulte d'un concours de circonstances exceptionnelles qui réunit le chrome (et le vanadium) et le béryllium dans un même lieu. On distingue quatre types de gisements d'émeraudes.

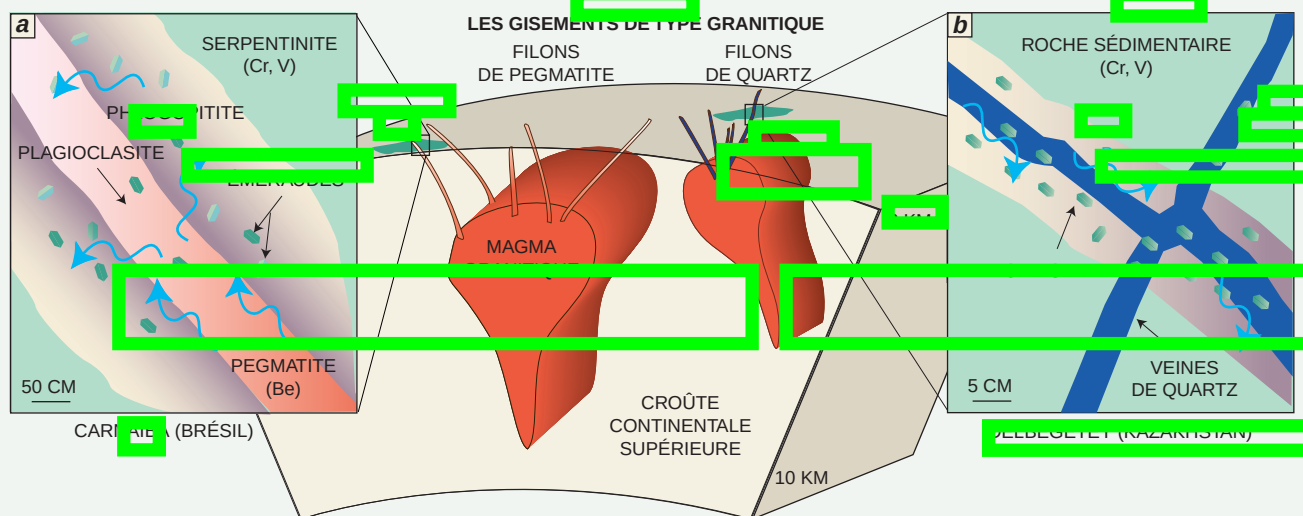
Les gisements de type granitique (voir la figure), tel le gisement de Carnaíba, au Brésil (voir l'encart a), résultent de la rencontre entre une serpentinite, issue de la transformation d'une roche du manteau pauvre en silice (basique), mais riche en magnésium, en fer, en chrome et en vanadium, et un magma granitique riche en silice (acide) et en alumine. En remontant vers la surface terrestre, le magma fissure les roches basiques, notamment la serpentinite, et y injecte des filons de pegmatite. À mesure que le magma refroidit, ses divers constituants se sollicitent les uns après les autres, formant une série de roches magmatiques. La pegmatite est le « bout » de cette chaîne : ses constituants (du quartz, du feldspath et du mica, dont les cristaux atteignent parfois des tailles de l'ordre du mètre) sont les derniers à se solidifier. Le refroidissement du magma libère un fluide hydrothermal constitué d'eau, de dioxyde de carbone et de sels, dont la température est comprise entre 400 et 600 °C. Ce fluide en

traverse les zones de contact entre les filons de pegmatite et la serpentinite, dont il solubilise la plupart des constituants (notamment le béryllium de la pegmatite, et le chrome et le vanadium de la serpentinite). Il transforme ces roches respectivement en plagioclase (constituée de feldspath blanc riche en sodium) et en phlogopite (un mica, c'est-à-dire une roche à structure feuilletée, contenant de la phlogopite : un mica magnésien de couleur marron). Ainsi, le chrome provenant de la dissolution de la chromite contenue dans la serpentinite et le béryllium de la pegmatite, sont concentrés dans la phlogopite et dans la plagioclase, où les émeraudes cristallisent.

Ces émeraudes englobent des minéraux de leur matrice, notamment de la phlogopite, dans des cavités intracristallines : les inclusions solides leur font perdre de leur transparence. Des petites gouttes de fluide hydrothermal sont également piégées dans les émeraudes, sous forme d'inclusions fluides.

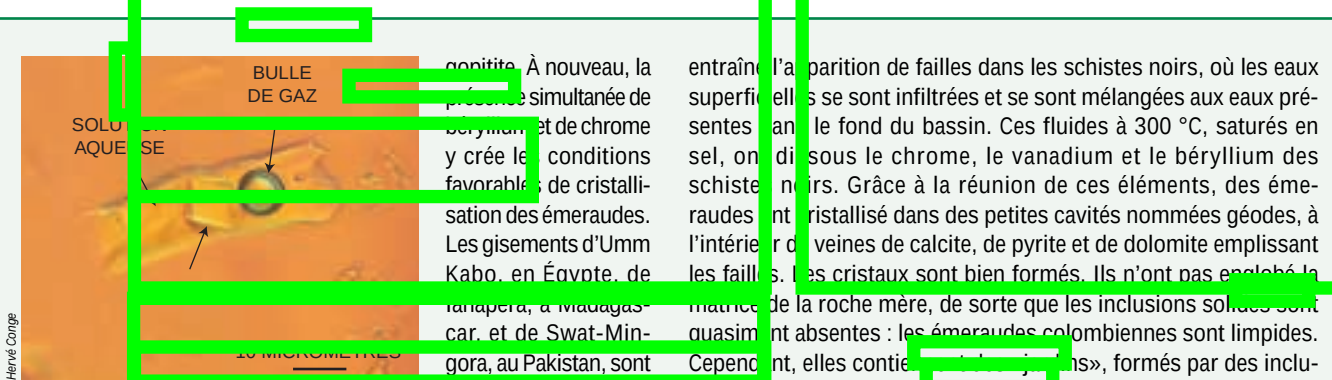
Les gisements de type granitique sont les plus courants. On en trouve notamment au Brésil, en Zambie, au Zimbabwe, en Australie et en Russie. Le gisement de Delbenev, au Kazakhstan (voir l'encart b), est une variante du type précédent, car l'émeraude est contenue dans des filons de quartz et des greisens (des roches constituées de quartz et de muscovite, un mica blanc) issus du magma granitique, qui a fissuré une roche sédimentaire. Le béryllium est apporté par les fluides hydrothermaux issus du granite, et le chrome probablement par les roches sédimentaires.

Les gisements associés à des accidents tectoniques profonds, tel le gisement de Santa Terezinha de Goiás, au Brésil (voir l'encart c), sont situés aux zones de failles profondes et de cisaillements dans la croûte terrestre. Ces accidents tectoniques permettent la remontée, dans les parties plus superficielles de la croûte, de fluides métamorphiques (qui circulent dans les parties inférieures de la croûte). Ces fluides, entre 400 et 600 °C, sont constitués de gaz carbonique, d'eau et d'éléments légers, notamment de béryllium. Au cours de leur remontée vers la surface, ils circulent à travers des roches basiques, notamment des schistes à talc, qui se transforment, là aussi, en phlo-



Dans les différents types de gisements, les roches «mères» contenant du chrome (en vert) sont transformées en roches «filles» (en violet) par le fluide hydrothermal qui y circule (flèches bleues). Celui-

ci s'est enrichi en béryllium (Be), qu'il a solubilisé en traversant la pegmatite (a) ou des roches plus profondes (non représentées), en chrome (Cr) provenant de la dissolution de la chromite contenue



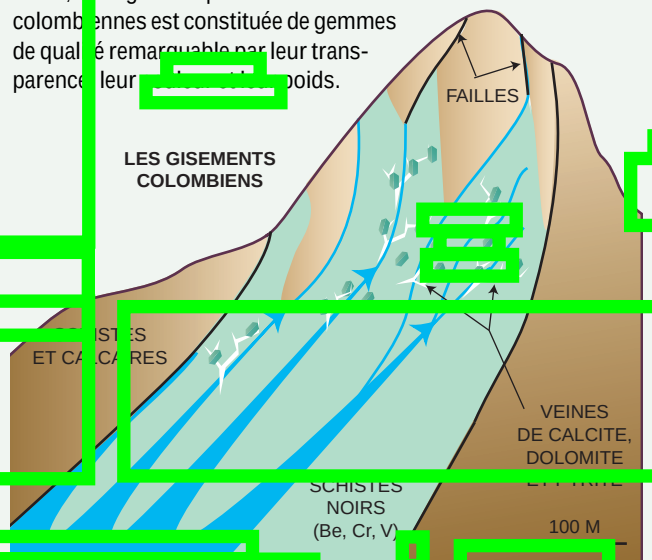
Les inclusions triphasées (cube de sel, solution aqueuse, gaz) sont caractéristiques des émeraudes colombiennes.

quartz associées soit à de la phlogopite, soit à des assemblages de dolomite et de quartz (voir l'encart d).

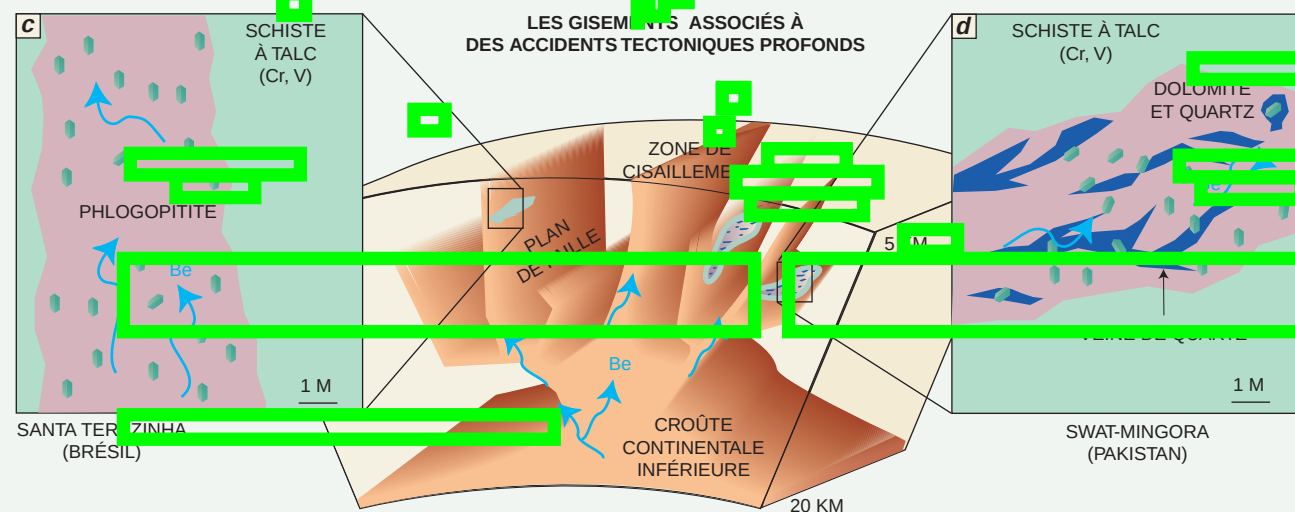
Les gisements issus du métamorphisme régional, tel celui d'Habachtal, en Autriche, résultent de la transformation chimique des roches à l'état solide à cause d'une élévation de température et de pression liée à leur enfouissement dans la croûte terrestre) et des déformations tectoniques qu'elles subissent. Les émeraudes cristallisent dans des phlogopites qui se sont formées au contact de roches basiques (les serpentinites) et de roches volcaniques acides (les schistes à muscovite et grenat). Ces dernières sont enrichies respectivement en chrome (et en vanadium) et en béryllium, nécessaires à la formation des émeraudes. Ces éléments ont été solubilisés par des fluides entre 500 et 550°C, composés d'eau et de gaz carbonique. Les gisements de Sikait et de Zabara, en Égypte, de Poona, en Australie, et de Gravelotte, en Afrique du Sud, appartiennent à ce type particulier de gisement.

Dans les gisements colombiens, les roches porteuses du béryllium et du chrome sont des roches sédimentaires. Ces éléments, issus du démantèlement de roches de surface, dont certaines contenaient du chrome et d'autres du béryllium, se sont déposés dans les sédiments du bassin de la cordillère orientale, les schistes noirs à grain très fin, principalement constitués d'argile et riches en matières carbonées. Les épisodes de déformation du bassin ont

entraîné l'apparition de failles dans les schistes noirs, où les eaux superficielles se sont infiltrées et se sont mélangées aux eaux présentes dans le fond du bassin. Ces fluides à 300 °C, saturés en sel, ont dissous le chrome, le vanadium et le béryllium des schistes noirs. Grâce à la réunion de ces éléments, des émeraudes ont cristallisé dans des petites cavités nommées géodes, à l'intérieur de veines de calcite, de pyrite et de dolomite emplissant les failles. Les cristaux sont bien formés. Ils n'ont pas englobé la matrice de la roche mère, de sorte que les inclusions solides sont quasiment absentes : les émeraudes colombiennes sont limpides. Cependant, elles contiennent des « inclusions fluides », formées par des inclusions fluides, où coexistent un cube de sel, une phase vapeur, composée d'azote et de gaz carbonique, et une solution aqueuse. Ainsi, une grande partie des émeraudes colombiennes est constituée de gemmes de qualité remarquable par leur transparence et leur poids.



Coupe géologique du gisement colombien de Coscuez qui s'est formé il y a 38 millions d'années. À cette époque, cinq kilomètres de sédiments le recouvraient. Des fluides salés, qui circulaient dans les failles du bassin sédimentaire ont solubilisé le béryllium (Be), le chrome (Cr) et le vanadium (V) des schistes noirs. Ces éléments ont permis la cristallisation des émeraudes dans les veines de calcite, de dolomite et de pyrite, qui se sont formées aux extrémités des failles.



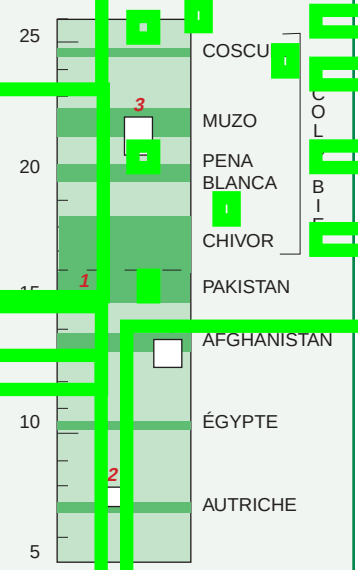
dans des roches basiques (serpentinite, schistes à talc) et en vanadium (V). L'émeraude cristallise dans les roches «filles» où ces deux éléments sont réunis : généralement des phlogopites (c),

mais aussi parfois des veines de quartz ou des assemblages de quartz et de dolomite (d) et de quartz et de micas blancs, les greisens (b).

La carte d'identité des émeraudes

Pour identifier le gisement d'où provient une émeraude, on détermine sa composition isotopique en oxygène, un élément qui constitue 65 pour cent des atomes du réseau cristallin. Les émeraudes contiennent trois isotopes stables de l'oxygène, de masse atomique 16, 17 et 18. Le noyau de l'oxygène 16 comprend huit protons et huit neutrons, tandis que les noyaux de l'oxygène 17 et 18 comportent respectivement un et deux neutrons de plus. Ces isotopes ont des concentrations distinctes : à cause de leur différence de masse, ils ne sont pas incorporés dans les mêmes proportions lors de la cristallisation à partir du fluide hydrothermal. La composition isotopique des émeraudes d'un gisement donné dépend des conditions dans lesquelles il s'est formé, notamment de la température de cristallisation et de la composition isotopique des roches « mères » partiellement dissoutes par le fluide hydrothermal (voir l'encadré des pages 62 et 63). Par exemple, les valeurs élevées du rapport des concentrations de l'oxygène 18 et de l'oxygène 16 (noté $^{18}\text{O}/^{16}\text{O}$) dans les schistes noirs de Colombie sont en partie responsables de la valeur élevée de ce rapport dans les émeraudes du gisement.

Nous avons mesuré le rapport $^{18}\text{O}/^{16}\text{O}$ dans des émeraudes issues de 62 gisements distincts, en utilisant la technique classique d'« extraction de l'oxygène » des silicates. Les échantillons, de quelques milligrammes, sont mis en contact avec un oxydant puissant à 650 °C, qui les attaque chimiquement et en extrait l'oxygène, sous forme de gaz. Celui-ci est ensuite transformé en dioxyde de carbone et analysé dans un spectromètre de masse. Cette étude a montré que chaque gisement d'émeraude était caractérisé par un intervalle spécifique de valeurs du rapport $^{18}\text{O}/^{16}\text{O}$. Ce rapport constitue la « carte d'identité » géographique d'une émeraude.



Depuis la fin de l'année 1998, nous avons mis au point l'analyse par sonde ionique de la composition isotopique de l'oxygène dans les minéraux de manière non destructive et avec une grande précision. Cet instrument bombarde la surface des émeraudes d'ions césium sur une aire de quelques micromètres carrés. Le faisceau d'ions accélérés arrache de très faibles quantités d'oxygène (quelques couches atomiques). La composition isotopique de l'oxygène est analysée par le spectromètre de masse de la sonde ionique. Les cratères d'impact, d'environ 10 micromètres de diamètre et 0,05 micromètre de profondeur, sont invisibles à l'œil nu, même avec une loupe. Cette technique étant quasiment non destructive, elle a été employée sur des émeraudes historiques (appartenant à des musées ou des collections privées), pour déterminer leur provenance.

L'analyse de quatre émeraudes historiques par la sonde ionique indique leur provenance : celle de la boucle d'orille gallo-romaine de Miribel (1), celle de la Sainte-Croix (2), celle d'un galleon espagnol coulé en 1622 (3) et celle du trésor indien du Nizam d'Hydrabad (4). Leurs rapports isotopiques sont portés en ordonnée ainsi que ceux des gisements connus avant le XVIII^e siècle.

1537, quand il fonde Bogota, l'actuelle capitale. En 1521, lors du sac de Tenochtitlan, l'ancienne capitale aztèque dont les ruines sont aujourd'hui enfouies sous la ville de Mexico, Hernan Cortés arrache à l'empereur Moctezuma un bloc de calcaire où sont enchâssées dix émeraudes d'un vert profond de 2,5 à 5 centimètres de hauteur : cette pièce, conservée au musée de Vienne en Autriche, est l'une des plus célèbres de la minéralogie.

Les conquistadores ne découvrent les premiers gisements d'émeraudes que 20 ans après leur arrivée : celui de Chivor en 1537, celui de Muzo dans les années 1550 et celui de Coscuez en 1646. Leur exploitation est intensive. La royauté espagnole perçoit une dîme d'un cinquième des pierres rapportées au Nouveau Monde. Le galleon *Nuestra Señora de Atocha*, qui coule en 1622 près des côtes de Floride, livre 120 émeraudes lorsqu'il est repêché. L'analyse de l'une d'entre elles par la sonde ionique indique non seulement son gisement d'origine (Muzo), mais la méthode est suffisamment précise pour que la mine elle-même soit identifiée : c'est celle

de Tecuandama, exploitée par les Espagnols à partir de 1594. L'afflux d'émeraudes rapportées du Nouveau Monde a de multiples conséquences économiques, mais aussi esthétiques : elles ornent tous les bijoux espagnols. Les croix en sont entièrement pavées. À la fin du XVII^e siècle, la production décline à cause des éboulements dans les mines et des révoltes des mineurs, traités en esclaves. En Europe, les émeraudes restent caractéristiques de la bijouterie espagnole et portugaise. Les pays voisins admirent ces beaux cristaux, qui ne sont plus givreux, comme ceux de l'Antiquité et du Moyen Âge. Les souverains en raffolent. Charles I^{er} d'Espagne aurait financé son élection à la tête du Saint-Empire avec des émeraudes colombiennes. Catherine de Médicis offre à son fils cadet un pendentif orné d'une grande table d'émeraude (conservée au cabinet des Médailles, à Paris). Parmi les autres pièces exceptionnelles de cette époque, un pot à onguents, sculpté par un artisan milanais dans un cristal d'émeraude, mesure dix centimètres de haut (voir la figure 4).

Une émeraude afghane en Inde

L'Inde regorge de gemmes. Les maharajas en sont friands et s'en parent à foison. D'où viennent les nombreuses émeraudes de leurs trésors ? Le pays possède les seules mines de diamants connues avant la découverte de celles du Brésil, en 1725, mais pas de mines d'émeraudes. Elles sont importées d'Égypte, et peut-être de Bactriane et du Bengale. La découverte récente de mines au Rajasthan et en Inde du Sud (à Coimbatore) relance le débat. Toutefois, aucune trace d'exploitation antique n'a été retrouvée. À partir du XVI^e siècle, les émeraudes viennent de Colombie. Après avoir voyagé sur les galleons espagnols jusqu'en Europe, elles reviennent sous pavillon portugais vers les comptoirs indiens de Goa et de Diu. Depuis que le navigateur portugais Vasco de Gama a découvert la route maritime des Indes par le cap de Bonne-Espérance, en 1498, les Portugais ont le monopole du transport de marchandises sur ce trajet. Moins fréquemment, les émeraudes empruntent aussi d'autres itinéraires, par la

Turquie ou par l'océan Pacifique (voir la figure 3).

Au XVI^e siècle, le voyageur français Jean-Baptiste Tavernier remarque qu'en Inde, les émeraudes sont si abondantes qu'elles se négocient à 20 pour cent au-dessous du prix français. Les Espagnols et les Portugais s'enrichissent en vendant leurs émeraudes sur le marché indien ou en échangeant ces pierres contre de l'or et des diamants. Les Moghols polissent les prismes d'émeraudes en perles irrégulières, afin de perdre le moins de matière possible. Ils les portent en pendeloque, en dégradé sur des colliers, ou seules, comme ornement de turban. Les gemmes de moins bonne qualité sont taillées en perles côtelées. Les grands cristaux d'émeraudes, plus rares, sont sectionnés latéralement, en « tables » à six côtes, ou taillés en rectangles aux angles émoussés gravés de rinceaux feuillagés et fleuris ou d'écritures coraniques (voir la figure 5). Au XIX^e et au XX^e siècle les parures des maharajas comportent d'innombrables émeraudes, diamants, perles et rubis, qui contribuent à leur prestige.

À partir des années 1920, avec l'art décoratif, ces émeraudes, des perles lisses ou gravées, en forme de feuille, ont une seconde vie : les joailliers de la place Vendôme, tels Cartier et Chaumet, les montent sur des bijoux identiques plus légers (en platine) ou les sertissent à la mode occidentale. Jusque dans les années 1960, Van Cleef propose toujours ces parures fastueuses d'inspiration indienne.

Quatre émeraudes taillées au XVIII^e siècle, de grandes dimensions et de très bonne qualité (translucides et d'un vert profond) du trésor de Nizam d'Hyderabad, le maharaja du plus puissant des États d'Inde du Sud, ont été analysées avec la sonde ionique. Ces belles émeraudes translucides, d'un vert profond, laissaient présager, par leurs inclusions triphasées caractéristiques (voir l'encadré des pages 62 et 63), une origine colombienne, que la sonde a confirmé pour trois d'entre elles : elles viennent du gisement de Muzo. Les rapports isotopiques obtenus indiquent même les mines dont elles sont issues : Peña Blanca, Coscuez et Tequendama.

La quatrième émeraude du trésor indien de Nizam d'Hyderabad vient d'Afghanistan, une provenance inattendue. Or, ce gisement, n'est connu



4. CE PONT À ONGUENTS, réalisé par un artisan milanais de la Renaissance, est taillé dans un cristal d'émeraude colombien haut de dix centimètres.



5. À L'ÉPOQUE MOGHOLE (du XVI^e au XVIII^e siècle), les émeraudes sont taillées en perles côtelées, comme sur ce collier. Les grands cristaux d'émeraudes sont sectionnés latéralement, en « tables » à six côtes et gravés de rinceaux feuillagés et fleuris ou d'écritures coraniques.

que depuis une vingtaine d'années, et on avait perdu toute trace d'exploitation antérieure. Les émeraudes afghanes peuvent être aussi belles que les émeraudes colombiennes, avec le même type d'inclusions. La provenance afghane de cette émeraude prouve que les pierres de ce gisement étaient déjà commercialisées au XVIII^e siècle (ce qui n'exclut pas une exploitation antérieure) et donne une nouvelle résonance à la thèse d'un gisement en Afghanistan connu dès l'Antiquité.

Nombre d'émeraudes indiennes se trouvent aujourd'hui en Iran, où

les descendants des Perses détiennent le plus gros trésor de gemmes du monde. Le shah Nader l'a constitué à partir des trésors moghols, pillés lors du sac de Delhi, en 1739. Le trésor est conservé à la banque Mélli de Téhéran, où les émeraudes sont stockées par tasques pleines. Lors du règne du dernier shah, deux gemmologues américains ont eu l'autorisation de les examiner et y ont observé les fameuses inclusions triphasées caractéristiques des pierres colombiennes et afghanes. Combien d'émeraudes afghanes se cachent-elles dans ce trésor ? Des analyses avec la sonde ionique permettraient de répondre à cette question et d'avoir ainsi une indication sur l'importance de l'exploitation du gisement afghan dans le passé.

Ces recherches à l'aide des résultats des fouilles archéologiques, des textes anciens et des références d'orfèvrerie ont permis de gommer nombre d'erreurs souvent répétées et d'établir la chronologie la plus précise de la route des émeraudes.

Costanza GIULIANI (Institut de recherche pour le développement) et Marc CHAUSSIDON (CNRS) sont géologues au Centre de recherches pétrographiques et géochimiques de Vandœuvre, près de Nancy. Michèle HEUZÉ est gemmologue diplômée d'Etat et historienne du bijou, une discipline qu'elle enseigne au Centre de formation des arts appliqués à Paris (ASAP).

John SINKANKAS, *Emeralds and other beryls*, Geoscience Press, 1989

D. GIARD, C. GIULIANI, A. CHEILLETZ, E. FRIESCHE et E. GONTHIER, L'émeraude. Connaissances actuelles et perspectives, AIG, CNRS, ORSTOM, eds, 1998.

G. GIULIANI, M. CHAUSSIDON, C. FRANCE-LANORD, C. ROLLION, D. MINGEN et P. COGET, *Analysis*, vol. 2, n° 3, pp. 203-206 : Application de l'analyse isotopique par spectrométrie de masse et sonde ionique de l'oxygène des émeraudes naturelles, 1999.

G. GIULIANI, M. CHAUSSIDON, P. J. SCHUBNEL, D.H. PIAT, C. ROLLION-BARD, C. FRANCE-LANORD, D. GIARD, D. DE NARVAEZ et B. RONDEAU, *Oxygen Isotopes and Emerald Trade Routes since Antiquity*, in *Science*, n° 287, pp. 631-633, 2000.

Michèle HEUZÉ, *Le jardin secret des émeraudes*, in *L'objet d'Art*, n° 345, pp. 52-65, 2000.

Une sélection naturelle de la culture

SUSAN BLACKMORE

Les comportements et les idées copiés par imitation de personne à personne – les mèmes – semblent avoir contraint les gènes humains à faire de nous ce que nous sommes aujourd’hui.

Savez-vous que vous passez la majeure partie de votre existence à copier et à transmettre des mèmes? Un mème est «une idée, un comportement, un style ou un usage qui se propage de personne à personne au sein d’une culture». Chaque fois que vous serrez une main, que vous chantez «Joyeux anniversaire» ou que vous votez lors d’une élection, vous transmettez des mèmes.

Personne ne conteste l’existence des mèmes, tels que nous venons de les décrire. Toutefois, la théorie présentée ici par la psychologue Susan Blackmore est controversée : elle soutient que l’aptitude de l’homme à l’imitation, et donc à la transmission des mèmes, le distingue des autres espèces. Les mèmes, affirme S. Blackmore, ont constitué (et constituent aujourd’hui encore) une force puissante qui a façonné notre évolution culturelle – et biologique. Pour situer le débat, nous présentons également trois points de vue qui divergent de celui de S. Blackmore : celui de l’écologiste behavioriste Lee Dugatkin, celui de l’anthropologue évolutionniste Robert Boyd et du biologiste des populations Peter Richerson, et celui du psychologue Henry Plotkin. À vous de vous faire une opinion.

L’être humain est un animal étrange. Bien que nous partagions de nombreuses caractéristiques avec les autres êtres vivants (ce que la théorie de l’évolution a expliqué brillamment), du code génétique qui gouverne la construction de notre corps jusqu’aux détails du fonctionnement de nos muscles et de nos neurones, nous nous singularisons sur d’innombrables points. Notre cerveau est exceptionnellement gros, nous sommes les seuls à posséder un langage véritablement grammatical, les seuls aussi à composer des symphonies, à conduire des voitures, à manger des spaghetti avec une fourchette et à nous émerveiller sur les origines de l’Univers.

Ces aptitudes semblent disproportionnées par rapport à nos besoins, car elles dépassent considérablement le minimum nécessaire à notre survie. Steven Pinker, de l’Institut de technologie de Massachusetts, l’a formulé ainsi : «Et ce qui concerne les causes et les effets biologiques, la musique est totalement inutile.» Il en est de même pour l’art, les échecs et les mathématiques pures.

La théorie darwinienne de l’évolution par sélection naturelle des individus les mieux adaptés au milieu n’explique pas toutes les compétences humaines. La théorie stipule que ces caractères sont commandés par les gènes ; au fil des générations, les gènes qui confèrent à leurs porteurs un avantage pour leur survie favorisent la production d’une abondante progéniture (héritière de ces gènes) et ils prolifèrent au détriment des autres gènes. Les

gènes sont ainsi en concurrence : ceux qui s’imposent sont ceux qui se transmettent le plus d’une génération à l’autre.

Malgré son immense intérêt, la théorie de Darwin n’explique pas pourquoi les humains accordent une si grande part de leurs ressources à des aptitudes sans intérêt pour la propagation des gènes. Quelle autre théorie expliquerait ces comportements?

Je pense que la réponse se trouve dans les mèmes : les histoires, chansons, habitudes, techniques, inventions et comportements que les individus acquièrent par imitation. La théorie de l’évolution explique la nature humaine si elle intègre la transmission des mèmes, en plus de celle des gènes.

La naissance des mèmes

Ce serait une erreur de considérer les mèmes comme de simples «idées» : ils correspondent plutôt à une forme d’information (les gènes sont également une forme d’information : ces instructions présentes dans l’ADN, comme dans la synthèse des protéines). Ainsi, le mème associé aux huit premières notes de la mélodie de *La Marseillaise* peut être enregistré non seulement dans les neurones d’une personne (qui reconnaît alors ces notes lorsqu’elle les entend), mais également sous forme magnétique, sur une cassette ou sous forme de signes sur une portée musicale.

Si le concept de mème date presque d’un quart de siècle, on ne lui a attribué un rôle déterminant dans l’évolution humaine que récemment.

Richard Dawkins, de l'Université d'Oxford, a proposé ce terme en 1976 : dans son livre intitulé *Le gène égoïste*, il décrit les principes fondamentaux de l'évolution darwinienne comme la somme de trois phénomènes, à savoir la réplication, la variation et la sélection – l'information est copiée encore et encore, avec des variations, puis certaines variantes sont sélectionnées par rapport à d'autres. Autrement dit, à mesure que ce cycle se répète, la population des copies survivantes acquiert progressivement de nouvelles propriétés qui la rendent globalement plus apte à triompher dans la compétition pour la production d'une descendance. Le cycle de réplication, variation et sélection engendre généralement une certaine organisation à partir du chaos.

R. Dawkins appelle «réplicateur» l'information qui se fait copier, et il

reconnaît que le réplicateur le plus familier est le gène. Toutefois, pour souligner que l'évolution peut reposer sur n'importe quel réplicateur, R. Dawkins a pris l'exemple du mème. La réplication des mèmes d'une personne sur une autre est imparfaite, tout comme la réplication des gènes, de parents à enfants : on peut embellir une histoire, changer un mot d'une chanson, adapter une vieille technique ou concevoir une nouvelle théorie à partir d'idées anciennes. Certaines de ces variantes seront à leur tour répliquées, tandis que d'autres disparaîtront. Les mèmes subissent ainsi les trois processus de l'évolution – réplication, variation, sélection –, à laquelle ils participent.

En imaginant les mèmes, R. Dawkins voulait seulement éviter que ses lecteurs ne réduisent la sélection

naturelle à celle des gènes. Cependant, le concept de mème est lourd de conséquences. Si les mèmes sont des réplicateurs, alors ils sont en concurrence et s'efforcent d'être répliqués pour leur propre compte, comme les gènes. Cette conclusion contredit l'hypothèse, soutenue par la majorité des psychologues évolutionnistes, selon laquelle la fonction ultime de la culture humaine est de servir les gènes en favorisant leur survie. Le fondateur de la sociobiologie, E. Wilson, a formulé l'idée en ces termes : «Les gènes tiennent la culture en laisse.» Parfois, la culture s'égare temporairement dans une direction défavorable à la propagation des gènes, mais,



1. COPIÉS DE CERVEAU EN CERVEAU, les mèmes proliféreraient dans les sociétés humaines. Les cultures seraient déterminées par leur propagation.

à long terme, la sélection naturelle génétique la remet dans le droit chemin, de même que le maître ramène son chien vers lui en tirant sur la laisse. De ce point de vue, les mèmes seraient les esclaves des gènes, car les gènes construisent les cerveaux qui copient les mèmes : les mèmes prospéreraient uniquement en aidant les gènes à proliférer. Cependant, si R. Dawkins a raison, c'est-à-dire si les mèmes sont des répliqueurs, ils servent alors égoïstement leurs propres objectifs, se répliquant dès qu'ils le peuvent. À mesure qu'ils se perpétuent, ils façonnent notre esprit, notre culture... et nos gènes.

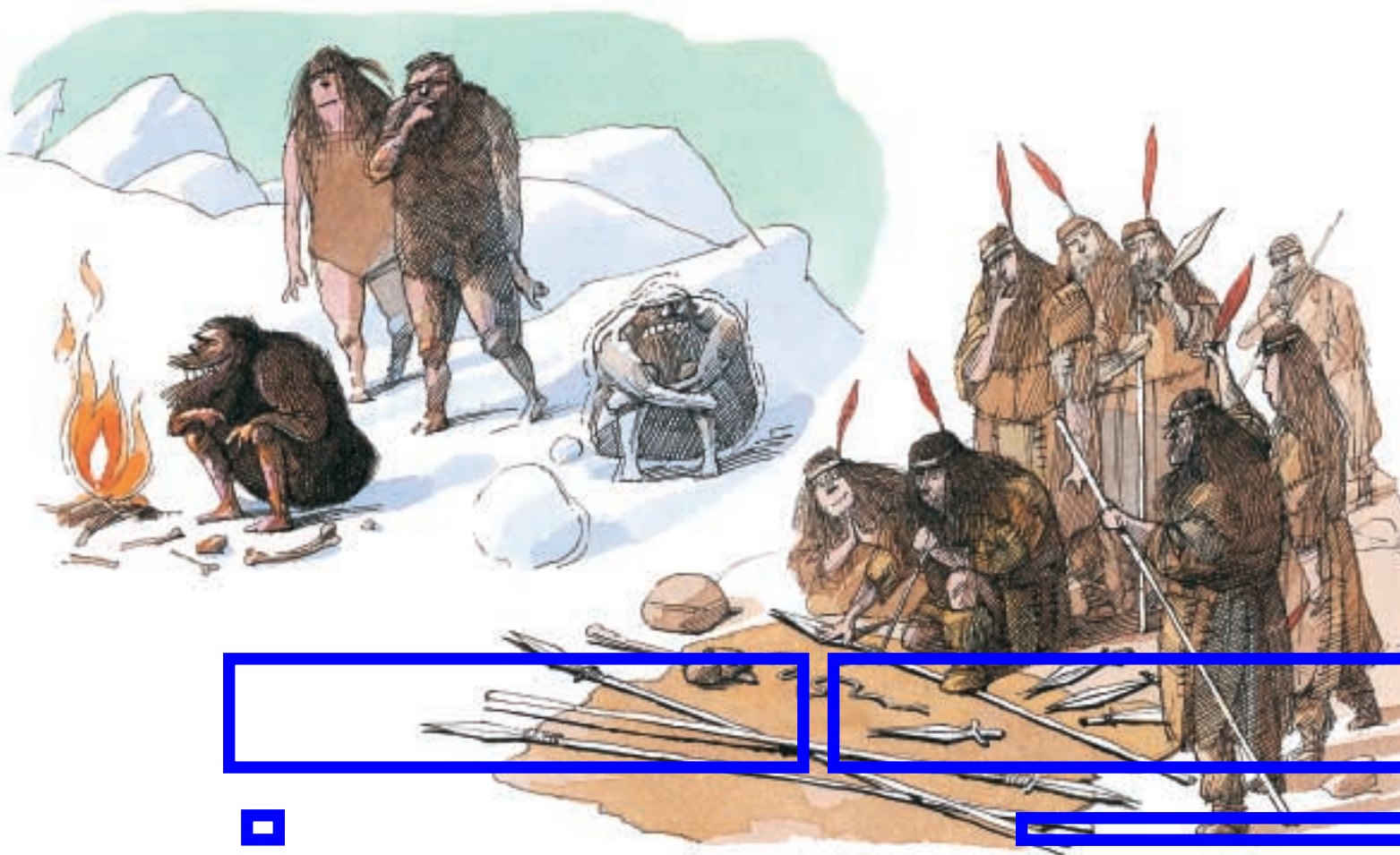
L'exemple des mèmes «viraux» illustre ce phénomène. Vous avez peut-être déjà reçu une lettre circulaire qui ne vous était pas personnellement adressée et qui vous enjoignait d'en faire des copies et de les renvoyer à 20 personnes de votre connaissance.

Ces lettres sont des éléments d'information écrite qui contiennent une instruction «copiez-moi», et une menace (si vous rompez la chaîne, cela vous portera malheur) ou une promesse (vous recevrez de l'argent). Évidemment la menace ou la promesse ne sont jamais suivies d'effets, et la recopie de la lettre est une perte de temps. Toutefois, nombre d'entre nous l'ont fait un jour. Ces mèmes possèdent une structure interne qui assure leur propre propagation.

Selon R. Dawkins, les grandes religions sont des mèmes viraux. Parmi les myriades de cultes apparus au cours de l'histoire humaine, quelques-uns seulement ont subsisté, grâce à des instructions du type «copiez-moi», accompagnées de menaces et de promesses. On menace les «mauvais» fidèles de mort ou de damnation éternelle, et on promet aux «bons» fidèles la béatitude éternelle. Les ins-

tructions du type «copiez-moi» sont variées : on demande aux fidèles de consacrer leur vie à la propagation de la bonne parole, de donner une partie de leurs revenus pour servir la religion, voire de consacrer leur fortune à la construction de magnifiques mosquées ou cathédrales destinées à promouvoir davantage les mèmes. Parfois les mèmes empêchent les gènes de se propager : par exemple, en imposant le célibat.

Tous les cultes (ou toutes les «chaînes de l'espoir») ne se propagent pas avec le même succès. Certaines menaces et certaines promesses sont plus efficaces (plus virulentes) que d'autres. Pour s'imposer à l'esprit humain, toutes s'opposent à l'expérience et au scepticisme qui, pour continuer la métaphore virale, se comportent comme des systèmes immunitaires. Certes, les religions ne sont pas entièrement virales : elles réconfortent et



2. LA CULTURE est probablement apparue quand nos ancêtres hominidés ont appris à s'imiter mutuellement (à gauche). Les individus les plus aptes à imiter de nouvelles techniques de survie – par exemple, allumer un feu – ont prospéré, propageant les gènes qui les avaient rendus plus capables d'une imitation généralisée. Plus

tard, lorsque l'imitation est devenue avantageuse en termes de sélection (au centre), les gènes ont «conçu» des stratégies pour garantir l'imitation des comportements les plus précieux. Des tactiques qui consistaient à imiter les meilleurs imitateurs ont abouti à de meilleures copies des nouvelles techniques de survie – et aussi à

donnent le sentiment d'appartenir à un groupe social. De surcroît, tous les mêmes ne sont pas viraux.

L'immense majorité des mêmes intervient tous les jours dans notre vie : les langues, les systèmes politiques, les institutions financières, l'enseignement, la science, la technique, etc. Toutes ces entités sont des mêmes (ou des groupes de mêmes) : elles se transmettent entre individus par imitation et rivalisent pour leur survie au sein de l'espace limité que constituent la mémoire et la culture humaines.

Le concept de même éclaire le fonctionnement de nos sociétés. Chaque être humain est une machine à fabriquer des mêmes : il en assure la propagation, il les réplique, et il représente une ressource qui attire les convoitises. Nous ne sommes ni les esclaves de nos gènes ni des êtres libres et rationnels qui créent la culture, l'art, la science et la technique pour notre propre bonheur ; nous par-

ticipons à une immense évolution dont les mêmes sont les répliqueurs évolutifs et où nous sommes leurs supports.

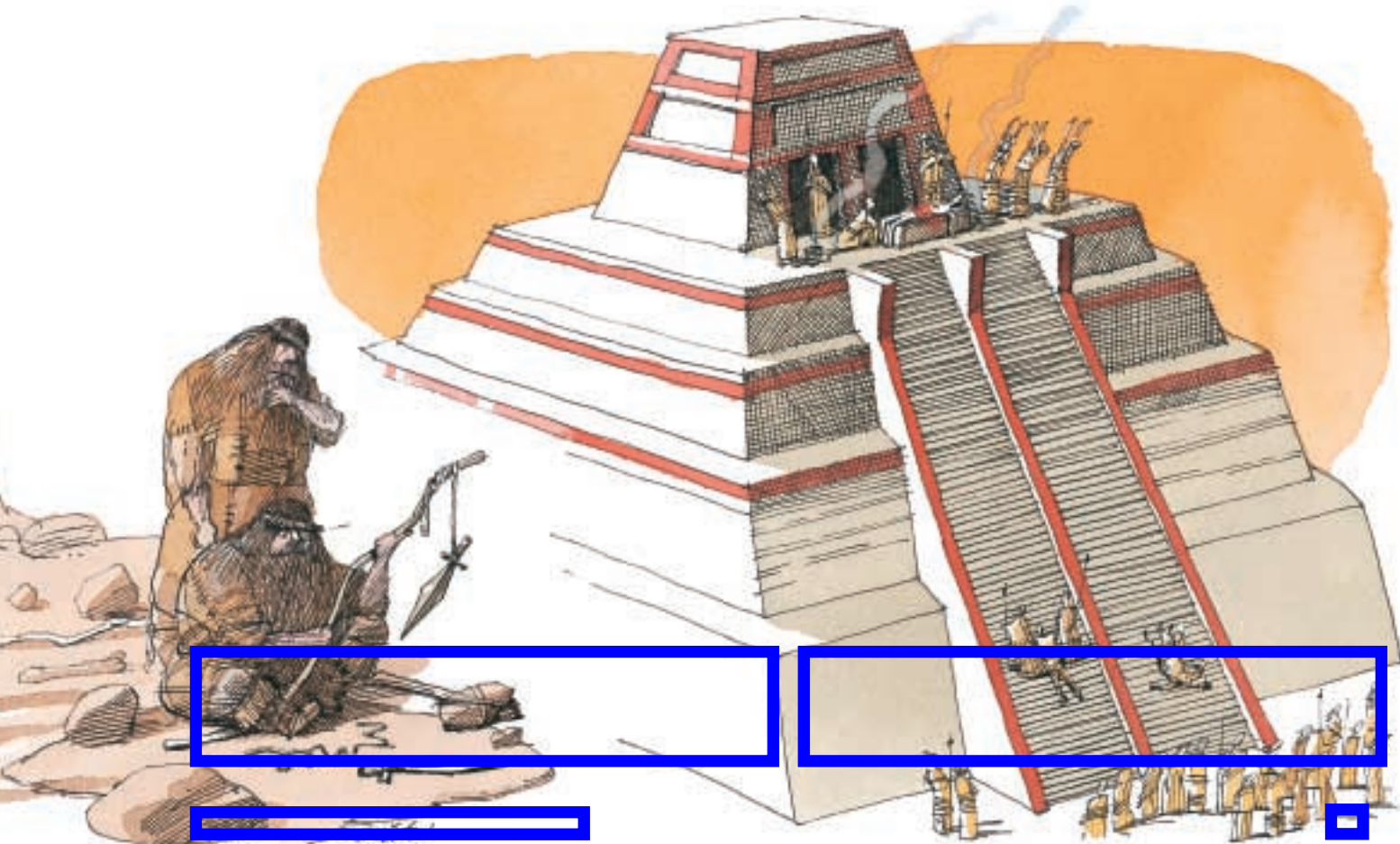
Cette vision est à la fois remarquable et inquiétante. Remarquable, car desormais, une théorie unique (celle de la coévolution des mêmes et des gènes) décrit la totalité de la culture et de la créativité humaines, ainsi que toute l'évolution biologique. Inquiétante, car elle semble réduire des pans considérables de notre humanité, de nos activités et de nos vies intellectuelles à un phénomène dépourvu de sens. La mémétique est-elle vraiment fondée ? Nous aide-t-elle à nous comprendre nous-mêmes et à prévoir l'avenir ? Est-elle le résultat d'une démarche véritablement scientifique ? Si tel n'est pas le cas, elle est sans intérêt.

Je pense que le même, en tant que répliqueur, est le concept qui manquait aux théories sur l'évolution humaine. La mémétique explique l'originalité de

notre espèce dans le monde animal et l'apparition des cultures complexes de l'humanité. Nous différons parce que, nous seuls, à une époque lointaine de notre passé, sommes devenus capables d'une imitation généralisée. Cette aptitude a engendré de nouveaux répliqueurs — les mêmes —, qui ont alors essaimé, nous utilisant comme des machines à reproduire, tout comme les gènes qui utilisent les mécanismes de réplication, à l'intérieur des cellules. Nous sommes uniques, parce que nous avons évolué sous l'action de deux répliqueurs : les gènes et les mêmes.

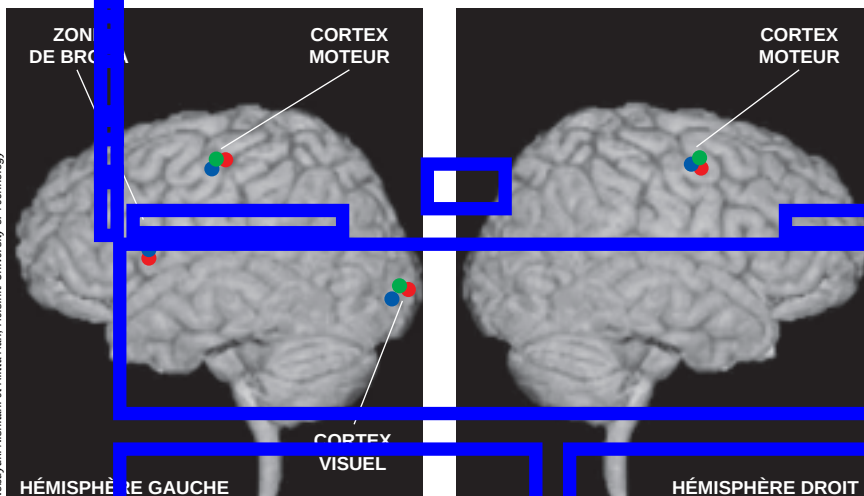
De gros cerveaux pour les mêmes

Cette évolution particulière nous a dotés d'un gros cerveau, d'un langage et de tout notre « excédent » d'aptitudes singulières. La mémétique explique la grosseur du cerveau



la propagation de comportements sans intérêt, tels que la décoration des vêtements. Les bons imitateurs ont acquis un statut social élevé, ont attiré des partenaires sexuels et ont eu une nombreuse progéniture. Ils ont poussé les gènes à développer de gros cerveaux capables d'une imitation sophistiquée. L'imitation est alors deve-

nue partie intégrante de la nature humaine. Les mêmes, sans cesse soumis à l'évolution, ont progressivement engendré des systèmes culturels (à droite) où l'on pratiquait des activités, telles que la construction de monuments et les sacrifices humains, dépourvues de valeur pour les gènes, mais propices à la transmission des mêmes.



3. PAR TOMOGRAPHIE, on déterminera si le cerveau humain a évolué de manière à imiter et à propager les mèmes. Sur ces cartes de l'activité neuronale associée à un mouvement particulier de la main, on voit que les mêmes aires du cerveau sont activées, que le sujet agisse selon sa propre volonté (en rouge), qu'il observe simplement quelqu'un d'autre (en bleu) ou qu'il l'imité (en vert). L'imitation correspond à l'activité la plus intense. Ces résultats indiquent que, chez l'homme, la zone de Broca commande un «système neuronal miroir» spécialement élaboré pour imiter les actions. Les singes ont un système analogue.

humain. Celui-ci a la taille maximale que lui donnent les gènes : il est trois fois plus gros, relativement au poids du corps, que le cerveau de nos plus proches parents, les grands singes. Sa construction et son entretien sont coûteux, et nombre de mères et de bébés meurent lors de l'accouchement, en raison de la grosseur de la tête de l'enfant. Pourquoi l'évolution a-t-elle laissé le cerveau atteindre des proportions aussi dangereuses ? Selon la théorie classique de l'évolution, la grosseur du cerveau est un avantage génétique : plus intelligente, l'espèce humaine trouve mieux sa nourriture et s'aggrave en groupes coopératifs plus grands, qui augmentent la survie de l'individu.

Selon la mémétique, l'explication est tout autre : la transition critique fut l'apparition de l'imitation, il y a 1,5 à 2 millions d'années, avant celle des outils de pierre et avant que le volume cérébral augmente. Une véritable imitation, d'un comportement ou d'un savoir-faire manifestés par un autre animal, est une opération difficile qui exige une grande intelligence, rare dans le règne animal. Certes, de nombreux oiseaux reproduisent des chants, et quelques baleines, quelques dauphins savent imiter des sons et des actions, mais la plupart des espèces sont incapables de tels exploits. Souvent, en présence d'une situation nouvelle, l'imitation animale est un simple recours à un comportement inné ; par exemple,

face à un nouveau prédateur, l'animal réagit de façon innée. Même les chimpanzés n'imitent que peu de comportements, tel l'épouillage. L'imitation généralisée de toutes sortes d'activités – apparemment si naturelle chez l'être humain – est bien plus difficile. Elle constitue une aptitude précieuse, car son détenteur bénéficie ainsi du savoir ou de l'ingéniosité des organismes imités. Par exemple, lors d'expériences effectuées en 1990 au Centre de recherche sur les primates de Géorgie, on a confronté des enfants et des orangs-outans aux mêmes problèmes : seuls les enfants ont utilisé l'imitation pour résoudre les problèmes.

Nos lointains ancêtres ont utilisé leurs capacités d'imitation pour allumer des feux, chasser, cuisiner. Puis, quand les premiers mèmes se propagèrent, l'aptitude à l'imitation est devenue un élément important pour la survie : les meilleurs imitateurs ont prospéré et les gènes responsables de la grosseur de leur cerveau, nécessaires à l'imitation, ont proliféré. Chacun a fini par acquérir de meilleures aptitudes à l'imitation, accentuant la nécessité d'agrandissement du cerveau, dans une sorte de course aux armements cérébraux.

Quand tous les individus ont été capables d'imitation, le second répliqueur – les mèmes – s'est répandu, modifiant sans retour l'évolution humaine. Les mèmes ont alors exercé leur emprise : outre les mèmes asso-

ciés à des savoir-faire utiles, telle la préparation du feu, les êtres humains ont copié des mèmes inutiles, tels ceux de l'ornementation du corps ou des danses de pluie (qui requièrent une dépense énergétique importante).

Les gènes et les mèmes se sont alors heurtés. Les mèmes, par exemple, conduisaient à des comportements inutiles, génétiquement néfastes. De nouveaux mèmes essaïmaient dans toute une population en l'espace d'une seule génération, sans que l'évolution génétique ait le temps de suivre. Pendant que les gènes favoriseraient les comportements génétiquement utiles, telle la maîtrise du feu pendant qu'ils limitaient les comportements néfastes, telle l'avarision envers les danses de la pluie, des engouements nouveaux apparaissaient et se propageaient. Notamment, l'évolution génétique a amélioré les capacités d'imitation, qui sont devenues plus sélectives.

Ainsi, les gènes semblent avoir conduit les individus à ne copier que les meilleurs imitateurs, ceux qui possédaient des versions non déformées des mèmes utiles. Aujourd'hui, les «meilleurs imitateurs» sont ceux qui lancent des modes. En plus de leur équipement personnel de survie (les mèmes utiles), ces meilleurs imitateurs ont acquis un statut social supérieur, qui améliore encore leurs chances de survie et favorise la propagation des gènes qui leur donnent leurs talents d'imitateurs, c'est-à-dire qui leur donnaient de gros cerveaux spécialisés dans l'imitation généralisée et efficace.

Les gènes ont alors accru la préférence innée pour l'imitation des meilleurs imitateurs, mais ces adaptations ont duré de nombreuses générations de sorte qu'ils étaient toujours à la traîne. J'ai nommé «pilotage mème-que» le mécanisme de commandement génétique assuré par les mèmes : les mèmes rivalisent entre eux, évoluent rapidement dans une direction, tandis que les gènes réagissent en améliorant l'imitation sélective, par l'augmentation régulière de la taille et de la puissance du cerveau. Les mèmes qui ont le plus de succès finissent ainsi par décider quels sont les gènes qui leur sont le plus favorables : ils tiennent la laisse.

Comme les individus ont intérêt à s'accoupler avec les meilleurs imitateurs, qui survivent mieux et se

reproduisent davantage, la sélection sexuelle, guidée par les mêmes, aurait contribué à augmenter la taille du cerveau. En coévoluant avec les mimes, les individus ont propagé les gènes nécessaires à l'exercice de la religion, à la pratique du chant, de la danse, de la

passée serait ainsi inscrite dans les structures de notre cerveau. Les mimes artistiques et religieuses. Le cerveau est un outil d'imitation sélective, construit par les mêmes, au bénéfice des mimes autant que des gènes.

L'origine du langage ?

coévolution des mimes et des gènes. Les origines et la fonction du langage sont très controversées. En 1866, la Société linguistique de Paris avait même décidé de ne plus examiner

Théorie des mimes et évolution culturelle

Les gènes sont des réplicateurs. Ils se transmettent fidèlement et commandent le fonctionnement des organismes vivants. Cette transmission fidèle est essentielle au fonctionnement de la sélection naturelle : les gènes qui permettent à leurs porteurs de se reproduire plus que les porteurs d'autres gènes finissent par s'imposer. D'autres mécanismes – telle la mutation – jouent des rôles déterminants dans l'évolution, mais l'adaptation résulte presque entièrement des gènes qui sont les plus rapides à se répliquer. Grâce à cette idée, les biologistes comprennent des phénomènes évolutifs variés, par exemple ceux qui ont gouverné la forme du pelvis humain et le rythme des changements sexuels chez les poissons hermaphrodites.

Selon S. Blackmore, les croyances et les idées – qu'elle désigne sous le nom de «mimes» – sont des réplicateurs qui, fidèlement transmis par le mécanisme d'imitation, commandent le comportement des individus qui les ont acquis. L'évolution des idées serait également façonnée par la sélection naturelle, et les changements culturels s'expliqueraient par les mimes qui se répliquent le plus rapidement.

Nous pensons qu'elle n'a, au moins, qu'à moitié raison. Les concepts biologiques sont utiles pour étudier l'évolution culturelle : la culture étant un ensemble d'idées dans une population de cerveaux humains, des mécanismes analogues à la sélection naturelle pourraient faire prospérer certaines idées et en faire disparaître d'autres. Cependant, contrairement à S. Blackmore, nous pensons que la seule sélection naturelle ne suffit pas à expliquer l'évolution culturelle. Les multiples phénomènes qui façonnent la culture humaine relèvent aussi de la psychologie, de l'anthropologie et de la linguistique.

Contrairement aux gènes, les idées ne se transmettent pas telles quelles, d'une personne à une autre. L'information présente dans le cerveau de quelqu'un engendre un comportement, puis une autre personne essaie de déduire l'information nécessaire pour reproduire ce comportement. Des erreurs surviennent dans la transmission des idées, parce que des différences génétiques, culturelles ou sociales entre deux personnes conduisent celle qui veut imiter l'autre à des hypothèses erronées sur ses motivations : les mimes sont transformés lorsqu'ils sont transmis. Ce mécanisme diffère complètement de la sélection naturelle, qui résulte de différences de vitesse de transmission des gènes. À cause de la transformation des mimes, les membres d'une génération acquièrent parfois un mème différent de celui que partage l'ensemble de la génération précédente.

À l'Institut Max Planck de psycholinguistique de Nimègue, David Wilkins a identifié un exemple simple de transformation d'un mème : il a découvert que les Américains de générations différentes n'interprètent pas le suffixe *gate* de la même façon.

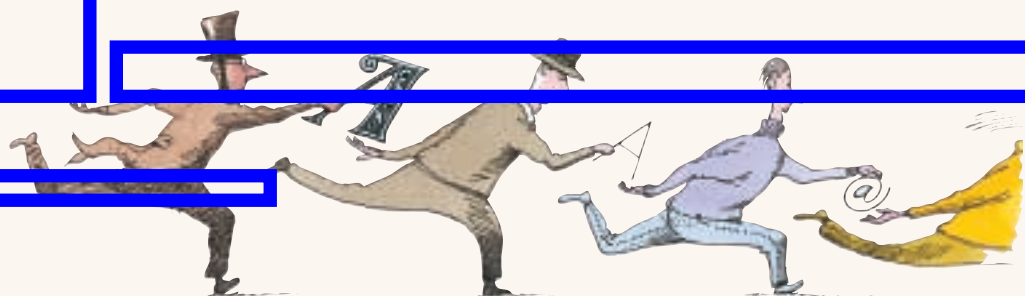
Les Américains âgés de plus de 40 ans l'associent à un scandale d'État et à une dissimulation. Pour eux, tous les mots construits avec le suffixe *gate* désignent des scandales analogues au Watergate. En revanche, les Américains plus jeunes ne font pas le lien avec ce scandale et associent le suffixe *gate* à n'importe quel type de scandale. Cette transformation ne correspond pas nécessairement à une rivalité entre mimes.

Les gènes se transforment aussi par mutations spontanées. Cependant, les mutations génétiques sont rares : elles ne surviennent en moyenne qu'une fois sur un million de répliquations. Leur action directe sur l'adaptation des organismes est souvent négligeable. Si des mutations survenaient plus souvent – par exemple, toutes les dix répliquations –, l'effet se manifesterait sur la sélection des gènes les plus courants. C'est ce qui se produit avec les idées, qui se transforment rapidement, à mesure qu'elles se propagent. Dans cette hypothèse, l'évolution culturelle est une combinaison de transformations et de sélection naturelle.

L'évolution des idées relève de processus non sélectifs variés : par exemple, il est fréquent qu'une personne emprunte une idée à quelqu'un d'autre, puis la modifie afin de l'améliorer ; ou que quelqu'un fasse la synthèse de diverses croyances après avoir été en contact avec d'autres personnes. Pour interpréter correctement l'évolution culturelle, nous devons tenir compte de ces multiples phénomènes qui orientent les différentes phases de cette évolution. Les sociologues qui explorent ces phénomènes ont déjà obtenu quelques résultats notables. Ainsi, à l'Université de Pennsylvanie, William Labov a décrit des processus psychologiques et sociaux qui engendrent des variations dans les dialectes, au fil des générations. À Stanford, Albert Bandura a étudié l'impact de l'imitation sur l'acquisition des idées.

Au X^e siècle, les biologistes ont mis au point des concepts et des outils mathématiques pour étudier l'interaction des divers phénomènes qui façonnent l'évolution des populations. Pour comprendre comment évolue la culture, les biologistes d'aujourd'hui et de demain devront sans doute tester la théorie des mimes et découvrir d'autres mécanismes qui la complètent.

Robert BOYD, anthropologue à l'Université de Los Angeles, et Peter RICHESON, biologiste des populations à l'Université de Davis



Les idées sont presque systématiquement transformées lorsqu'elles passent d'une personne à l'autre ou d'une génération à une autre.

aucune spéculation sur le sujet. Aujourd'hui, les théories qui sont le plus largement soutenues invoquent un avantage génétique. Ainsi, à l'Université de Liverpool, Robin Dunbar suppose que le langage s'est substitué au toilettage afin de maintenir la cohésion des groupes sociaux humains. Boston Terrence Deacon propose que le langage ait favorisé la communication symbolique, qui a amélioré les techniques de

chasse et renforcé les liens sociaux et la défense du groupe. La théorie du pilotage mémétique explique l'apparition du langage par les avantages sociaux qu'il procure aux êtres humains. Quels sont les mêmes qui ont le mieux survécu par le langage? La réponse serait la même pour n'importe quel réplicateur : ceux qui sont dotés d'une fécondité, d'une fidélité et d'une longévité élevées – autrement dit, ceux

qui sont capables d'effectuer des copies nombreuses, exactes et durables d'eux-mêmes. Les sons sont plus utiles que les images. Notamment, tout individu qui entend un «lé!» ou un «attention!» se trouve en alerte. La fidélité des mêmes parlés est supérieure lorsque ceux-ci sont faits d'unités sonores discontinues (les phonèmes) et divisés en mots : des erreurs de réplication sont alors évitées. À

Les animaux imitent aussi

Susan Blackmore tente de propager le même «l'imitation est importante» dans nos cerveaux ! J'applaudis à cette tentative, mais ne suis pas d'accord avec elle : je pense que les mêmes – les entités imitées – n'influencent pas uniquement le comportement des humains. Les animaux, des poissons aux primates, se copient mutuellement lorsqu'ils décident ce qu'ils vont manger ou qu'ils choisissent un partenaire sexuel. Les mêmes influent sur les habitudes de nombreux animaux, au même titre que sur le comportement humain. L'étude des merles montre que les mêmes ne sont pas l'apanage de l'espèce humaine ni même celle des primates, comme nous le verrons plus loin.

Qu'est-ce d'abord que l'«imitation»? Les psychologues ont largement discuté le sens de ce mot, mais puisque notre discussion concerne les mêmes, adoptons d'abord une définition que S. Blackmore donne dans son livre *La machine à mêmes*. L'imitation résulte de trois comportements complexes : l'individu décide d'imiter, il passe du statut d'observateur au statut d'imitateur, puis il effectue une action corporelle adéquate. Avec des critères aussi stricts, l'imitation n'existe pas chez les animaux, car ces derniers ne transforment pas leur point de vue en un autre. En outre, comment savoir ce qu'ils décident d'imiter?

S. Blackmore définit l'imitation de façon plus souple quand elle considère la propagation des histoires drôles : «L'individu qui retransmet une telle histoire ne rend pas fidèlement chaque geste et chaque parole de celui qui a initialement raconté l'histoire, mais quelque chose (le point central de l'histoire) a été conservé.» Avec cette seconde définition, des centaines de comportements animaux sont des imitations.

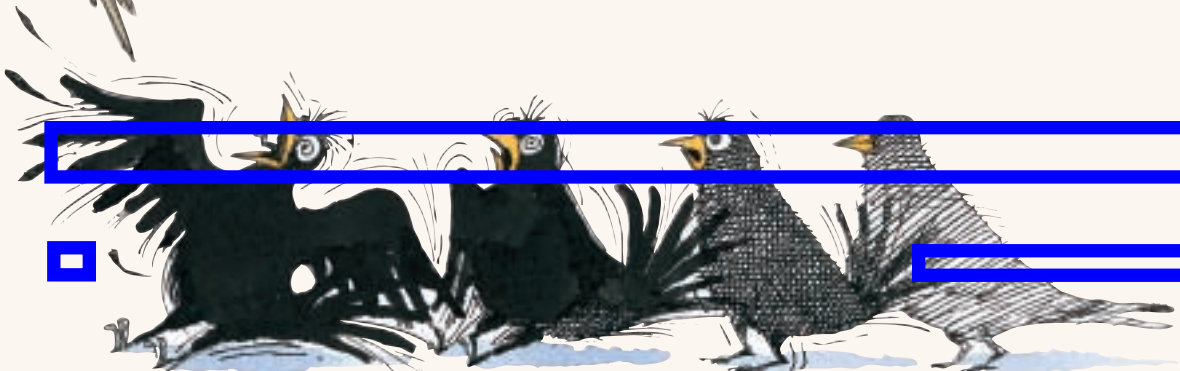
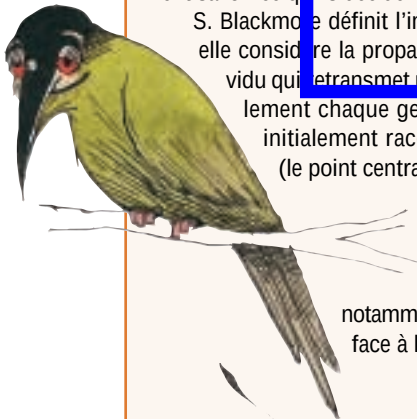
C'est par ce mécanisme, notamment, que les merles apprennent à réagir face à leurs prédateurs.

En 1971, Eberhard Curio et ses collègues de l'Université de Bochum ont réalisé une expérience où un merle voyait un second merle piailler et agitant la queue, en réaction à la présence d'un prédateur. Une ingénieuse manipulation faisait croire au premier merle que son compagnon réagissait à la présence d'un oiseau bruyant (un philémon), qui n'est pas un prédateur des merles. Ultérieurement, en présence d'un philémon, le premier merle s'est mis à son tour à piailler et à agiter la queue. E. Curio et ses collègues ont découvert que le message erroné «les philémons sont des prédateurs» pouvait être transmis en série à six autres merles au moins.

La simple copie d'un message ne signifie pas que ce message soit un même. Selon S. Blackmore, un message est un même s'il vérifie trois autres critères : ses copies doivent être fidèles, nombreuses et avoir une certaine longévité. Le message «les philémons sont des prédateurs» a été fidèlement transmis. L'évaluation expérimentale de sa longévité sera difficile, mais les ornithologues supposent qu'il subsisterait longtemps dans une population naturelle de merles.

Au cours de mes études, j'ai observé des dizaines d'autres exemples de comportement animal qui concordaient avec la définition d'un même. Les mêmes pourraient se révéler plus anciens et avoir joué un rôle plus déterminant dans l'évolution biologique que ne l'a affirmé S. Blackmore. En outre, la différence entre les mêmes animaux et les mêmes humains pourrait être plus quantitative que qualitative. S'ils reconnaissaient l'existence de mêmes animaux, les méméticiens comprendraient que les mêmes sont une force universelle dans l'évolution. Toutefois, si les mêmes ne nous distinguent pas des animaux, la théorie est affaiblie ; elle n'explique pas pourquoi la culture humaine est la seule qui soit si avancée.

Lee DUGATKIN, professeur de biologie à l'Université de Louisville



Les merles apprennent à reconnaître les prédateurs en observant ce qui effraie leurs compagnons. De ce fait, ils redoutent parfois des créatures qui ne les menacent pas.

mesure que diverses actions et vocalisations rivalisaient, au sein du groupe de mèmes préhistoriques, ces mots parlés prospérèrent et supplanterent les mèmes de communication formants. Puis, par association de mots dans des ordes différents, par addition de préfixes et de suffixes, des mèmes vocaux inédits et plus complexes sont apparus. En somme, les sons reproductibles et de qualité supérieure ont supplanté les sons les plus pauvres.

Ce phénomène s'est accompagné de modifications génétiques. Ici encore, les meilleurs imitateurs (les individus s'exprimant le plus clairement) ont accouplés avec les meilleures partenaires et ont eu plus de descendance. De fait, les sons les plus intelligibles et les plus reproductibles se sont propagés dans le groupe humain. Je pense que ces sons réussis – les fond

du langage parlé – ont progressivement conduit les gènes à créer un cerveau plus gros et aussi expert dans la replication des sons articulés : telle est l'aptitude humaine pour le langage, résultat

de la coévolution des mèmes et des gènes. Cet exemple de coévolution de réplicateurs (les mèmes) et de leurs machines à copier (les cerveaux) n'est

Nous ne sommes pas seulement des imitateurs

La mémétique que présente S. Blackmore pose deux grands problèmes. Premièrement, elle affirme que la culture n'est qu'une collection de mèmes. Dans le mot culture, elle inclut aussi bien les actions simples, telle l'utilisation d'un outil de pierre, et les institutions complexes, telles les banques. Deuxièmement, elle considère que tous les mèmes, donc tous les aspects de la culture, se propagent par imitation. Spécialiste de psychologie, je pense qu'aucune de ces deux hypothèses n'est admissible.

Au début du xx^e siècle, le psychologue américain Edward Thorndike a défini l'imitation comme l'apprentissage d'un geste fondée sur une observation de l'exécution de ce geste. Si l'on conserve cette définition, la théorie de S. Blackmore est évidente, car l'imitation de gestes ne transmet rien de culturellement important. Le mouvement des lacets ou le lancer de balles ne sont pas des gestes notables pour l'évolution culturelle de l'humanité.

En revanche, quand on adopte la définition de S. Blackmore (tout mode de communication entre personnes, de la transmission de l'argument central d'une histoire au souvenir des instructions lues dans un manuel, relève de l'imitation), le mot devient si vague qu'il se vide de sens. En outre, cette définition générale n'explique pas l'existence et l'évolution de la culture, qui est bien plus que la répétition machinale de gestes physiques. La culture humaine est liée au partage des savoirs, des croyances et des valeurs.

Toute culture repose sur un ensemble d'interprétations consensuelles de la marche du monde, que l'on désigne parfois sous le nom de «schémas». Les règles en vigueur dans les restaurants constituent un schéma classique : dans ces endroits, quelqu'un prépare votre repas, l'apporte à votre table et nettoie après votre départ, en échange d'une somme d'argent. Les enfants acquièrent les schémas qui caractérisent leur propre culture grâce aux indications des adultes, par des mécanismes psychologiques complexes qui nous font donner un sens aux abstractions. L'imitation, selon la définition des psychologues, n'intervient pas dans ces mécanismes. Les croyances et les valeurs communes, également nommées «constructions sociales», nous sont inculquées par un mécanisme mal compris, également complexe. Contrairement aux schémas, qui décrivent des entités tangibles, tels les restaurants, les constructions sociales existent uniquement parce que les individus conviennent qu'elles existent. L'argent est une construction sociale, tout comme la justice. Certaines de ces constructions sociales ont un support physique, comme les billets de banque ou les pièces de monnaie, mais leur sens dépasse largement leur apparence physique et repose sur un accord mental. Sans consensus sur les valeurs spécifiques des billets et des pièces de monnaie, l'argent ne présenterait aucun intérêt. Nombre de croyances et de valeurs régissent également les interactions sociales.

Dans la plupart des cultures occidentales, par exemple, la justice repose sur les concepts d'équité et de propriété. D'autres cultures définissent la justice par des notions telles que le service et la vengeance. Mais elle ne se réduit aux tribunaux, aux juges ou aux prisons.

Les mécanismes par lesquels les enfants comprennent et acceptent des abstractions ont été peu étudiés. Ils passent certainement par le langage, mais ils reposent également sur notre aptitude, que les psychologues nomment «théorie de l'esprit», à prendre conscience que les autres sont animés d'intentions et de désirs. La réactivité à la pression sociale – autre trait psychologique caractéristique de notre espèce – explique aussi l'adhésion aux croyances et aux valeurs communes. Ici encore, l'imitation n'intervient pas. Nous n'imitons pas et ne savons pas imiter la justice. Au contraire, nous découvrons progressivement sa signification, par les conversations, à l'école ou dans les livres.

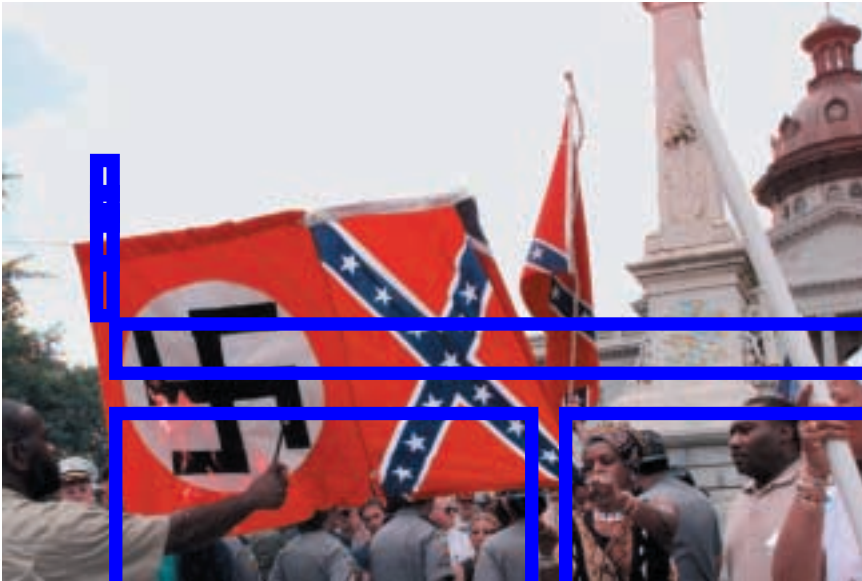


Les savoirs communs, telles les règles en vigueur dans un restaurant, ne sont pas susceptibles d'imitation.

Selon S. Blackmore, cette lente compréhension repose sur l'imitation. Cependant, les choses ne sont pas aussi simples. Des études récentes de neurobiologie indiquent que l'imitation résulte du déchiffrement de messages spécifiques dans des aires spécialisées du cerveau : lorsqu'un enfant comprend la signification des mots «restaurant» ou «justice», il passe par des étapes psychologiques différentes de celles qui lui permettent de nouer ses lacets.

Les schémas et les constructions sociales résultent de l'action de la mémoire et du phénomène d'abstraction. L'acceptation et la diffusion d'idées dans la société – en particulier de notions aussi complexes que la justice – est un processus lent, imprédictible, difficile à mesurer, et irréductible à la théorie restrictive des mèmes. La culture, en tant que collectif de cerveaux et d'esprits humains, est le plus complexe des phénomènes terrestres. Nous ne la comprendrons jamais en l'abordant par des méthodes naïves.

Henri PLOTKIN, psychologue au Collège universitaire de Londres



4. LES ÉMOTIONS déclenchées par les symboles idéologiques donnent une idée de l'importance des mèmes pour l'homme et de l'emprise qu'ils ont sur nos comportements.

d'années, un phénomène semblable s'est produit lorsque les premières molécules capables de se répliquer sont apparues dans la soupe primitive : elles ont conduit à l'ADN et à tout son arsenal de réplication cellulaire.

Après le langage, l'écriture s'est imposée. Elle est bien plus durable et fidèle. Puis, bien plus tard, l'imprimerie a renforcé l'efficacité de l'écriture ; du télégraphe au téléphone cellulaire, du courrier au courrier électronique, du phonographe au DVD et des ordinateurs à Internet, la machine à copier s'est perfectionnée, diffusant sans cesse plus loin et plus vite une multitude croissante de mèmes. La révolution informatique qui a commencé il y a quelques décennies est une conséquence de l'évolution mémétique.

La théorie mémétique se fonde sur des hypothèses que l'on peut tester expérimentalement. Elle stipule notamment que l'imitation – même si elle nous vient si facilement – exige une grande intelligence. Des études du cerveau par imagerie fonctionnelle (où l'on détermine les zones actives) devraient permettre une comparaison de l'activité cérébrale chez une personne qui effectue une action et l'activité d'une autre personne qui imite la première. Contrairement à l'intuition, la théorie mémétique stipule que l'imitation est l'activité la plus difficile à réaliser et que les régions du cerveau qui l'assurent sont récentes, en termes d'évolution. En outre, elle affirme que, dans tout groupe d'espèces animales apparen-

tées, l'espèce qui a la plus grande capacité d'imitation a le plus gros cerveau. Si la rareté des imitations chez les animaux limite le nombre de données disponibles, l'étude de certaines espèces d'oiseaux, de baleines et de dauphins devrait néanmoins permettre de tester cette théorie.

Les tests expérimentaux

Si le développement du langage humain résulte d'une coévolution des mèmes et des gènes, les linguistes devraient observer que le langage transmet mieux des mèmes dotés d'une fécondité, d'une fidélité et d'une longévité élevées qu'il ne transmet l'information sur des sujets spécifiques, telle la chasse, ou qu'il n'assure la cohésion des groupes sociaux. Des expériences de psychologie sociale devraient confirmer que les individus préfèrent imiter les personnes les plus éloquentes et les trouvent plus séduisantes que les personnes moins éloquentes.

On pourrait tester d'autres prévisions de la théorie mémétique avec des modèles mathématiques et des simulations numériques, comme de nombreux biologistes l'ont fait pour l'évolution génétique. On verrait alors comment l'apparition d'un second réplicateur, plus rapide, bouleverse la dynamique évolutive et comment il commande l'évolution du premier réplicateur. On pourrait alors utiliser ces modèles pour comprendre plus en détail la coévolution des mèmes et

des gènes. En outre, des simulations avec des robots imitateurs de bruits devraient montrer que le langage apparaît de façon spontanée au sein d'une population d'individus dotés d'aptitudes à l'imitation.

La mémétique est une théorie nouvelle. Certaines des critiques qui lui sont faites méconnaissent le concept de réplicateur : les mèmes, comme les gènes, ne sont que des éléments d'information qui réussissent à se faire copier ou ne se propagent pas. En ce sens seulement, ils sont «égoïstes» et possèdent un pouvoir de réplicateur. Les mèmes ne sont ni des entités magiques ni de vagues concepts abstraits, mais de l'information logée dans des souvenirs, dans des actes et dans des artefacts spécifiquement humains. Les contenus mentaux ne sont pas tous des mèmes : certains ne sont pas des copies. Sans mèmes, nous resterions incapables des perceptions, des émotions, des pensées et du savoir-faire qui nous sont propres.

On objecte fréquemment que les mèmes sont très différents des gènes. C'est vrai. Ils évoluent bien plus vite, et leur dynamique n'est pas encadrée par un système rigide, telle la machinerie biosynthétique des cellules.

De nombreuses critiques restent à examiner. La mémétique méritera de s'imposer si les explications qu'elle fournit sont meilleures que celles des théories concurrentes et si ses prévisions se révèlent justes. Contrairement aux religions, la science (qui est elle-même un groupe complexe de mèmes) a les moyens d'écarter les idées fausses : elle décide de l'intérêt de la mémétique.

Susan BLACKMORE est psychologue à l'Université de Bristol.

Richard DAWKINS, *The Selfish Gene*, Oxford University Press, 1976.

Robert BORD et Peter J. RICHERSON, *Culture and the Evolutionary Process*, University of Chicago Press, 1985.

Henry C. LOTKIN, *Darwin Machines and the Nature of Knowledge*, Penguin, 1992.

Daniel DENNETT, *Darwin's Dangerous Idea*, Penguin, 1998.

Susan BLACKMORE, *The Meme Machine*, Oxford University Press, 1999.

Darwinizing Culture : The Status of Memetics as a Science, sous la direction de Robert Aunger, Oxford University Press, à paraître.

Les premiers Américains

SASHA NEMECEK

On découvre en Amérique des outils fabriqués par des peuples indigènes vieux de plus de 14 000 ans. Ces découvertes revelent le mode de vie des premiers Amérindiens.

Dans les marais du Sud du Chili, il y a plus de 14 000 ans, une trentaine de personnes vivaient sur le site de Monte Verde, découvert par l'anthropologue Tom Dillehay dans les années 1970. Ces humains, parmi les premiers habitants de l'Amérique du Sud, étaient sans doute les descendants d'individus venus d'Asie. Ces derniers étaient-ils conscients d'être des pionniers? Ou bien avaient-ils simplement migré pour trouver à manger? Et quand sont-ils arrivés en Amérique?

La communauté des anthropologues et des archéologues discute, parfois âprement, de l'âge des sites anciens et de la date de la colonisation de l'Amérique. Les différentes datations des sites prêtent souvent à confusion. Toutefois, les découvertes de Monte Verde et les fouilles en cours dans l'Est des États-Unis et en Amérique du Sud semblent reculer la date d'entrée admise auparavant: on croyait que la colonisation datait d'il y a 14 000 ans, mais on sait aujourd'hui que des humains ont foulé le sol américain il y a plus de 20 000 ans!

Comment vivaient ces colons? Le site de Monte Verde a livré des pointes de lance de moins de dix centimètres de longueur et deux centimètres de large, qui attestent un travail soigneux. Notamment, les séries de minuscules encoches qui forment les bords aigui-

sés sont parfaitement symétriques. Grâce à de tels objets, les anthropologues reconstituent la vie quotidienne des premiers Américains et bousculent l'image du chasseur de mammouths qui prévalait naguère.

En 1581, le missionnaire jésuite Jose de Acosta proposa que les peuples américains soient venus de Sibérie quelques milliers d'années auparavant (voir la figure 4). Son hypothèse s'imposa, et les études archéologiques la corroborèrent. Au début du XX^e siècle, les plus anciens vestiges humains trouvés étaient datés de quelque 13 000 ans.

De surcroît, les outils découverts en différents endroits, du Canada jusqu'en Amérique centrale, se ressemblaient: on y voyait l'indication d'une culture commune et d'une colonisation unique du continent. Les archéologues ont nommé «Clovis» les colons et leur culture «Clovis», d'après le site de Clovis, au Nouveau-Mexique, où le premier objet ancien a été mis au jour.

Dans certains sites des États-Unis, notamment dans les régions désertiques du Sud-Ouest, ces pointes de type Clovis sont aussi communes que les cactus. Pourquoi les populations Clovis avaient-elles eu besoin de tant d'armes? À nouveau, la réponse semblait évidente: ils auraient été de redoutables chasseurs, qui poursuivaient du gros gibier, tel le mammouth, il y a 14 000 ou 15 000 ans, à travers l'actuel détroit de Bering qui à l'époque était

1. CHASSEUR DE MAMMOUTHS OU PÊCHEUR? Les premiers habitants du Nouveau Monde n'étaient pas uniquement des chasseurs de gros gibier, vêtus de peaux de bêtes (à gauche), mais aussi des chasseurs de petit gibier, des pêcheurs et des cueilleurs de plantes (à droite). La petite pointe de lance montrée (à droite) a été découverte au Chili; elle est l'un des plus anciens objets façonnés par l'homme découverts en Amérique.

un isthme nommé Bérिंगie, et à travers une bande de terre qui s'étendait du pôle Nord au Sud du Canada, entre les plaques de glace. Ces chasseurs auraient colonisé l'Amérique du Nord et du Sud en 1 000 ans environ, parce qu'ils poursuivaient le gibier sans relâche. L'importance accordée à la chasse s'accordait avec les données climatologiques : pendant cette période glaciaire, la viande de mammoth ou de bison apportait une grande partie des graisses et des protéines nécessaires ; les fourrures fournissaient des vêtements chauds.

Il était une fois en Amérique

Les premiers Américains ont-ils aussi été des chasseurs de la période glaciaire ? Selon l'anthropologue David Meltzer, cette hypothèse est simpliste : de petits groupes nomades de 15 à 30 personnes auraient été menacés d'extinction par les unions consanguines ou par les catastrophes naturelles. En outre, la chasse au mammoth était sans doute trop périlleuse pour constituer l'unique source de nourriture. Aussi les humains ont-ils également exploité d'autres sources, tel le petit gibier, les fruits, le poisson et les tortues de mer. Par exemple, des archéologues ont découvert des os de cerfs, de lapins et de serpents sur les sites de culture Clovis.

Les outils en bois, les filets et les paniers utilisés pour la chasse du petit gibier, pour la pêche et pour la cueillette sont moins durables que les objets en pierre avec lesquels on chassait le gros gibier. Toutefois, James Adovasio, qui a étudié pendant 30 ans l'abri sous roche de Meadowcroft, en Pennsylvanie, a découvert les paniers utilisés pour vraisemblablement transporter des plantes ou des mollusques pêchés dans la rivière Ohio ; ces objets datent d'environ 13 000 ans. Il a également mis au jour des collets, avec lesquels les Amérindiens capturaient peut-être du petit gibier, et des poinçons en os qui servaient sans doute au travail du tissu.

Les interprétations et les dates avancées par J. Adovasio ont été contestées. Cependant, celui-ci répond notamment en signalant que les aiguilles en os trouvées n'ont pu être utilisées pour coudre des peaux : elles casseraient ; elles ont probablement permis le tissage d'étoffes légères, d'origine végétale. Pendant la dernière glaciation, comme aujourd'hui, le climat était saisonnier :

les vêtements de peau devaient être trop chauds en été.

Qu'étaient ces vêtements estivaux ? Selon J. Adovasio, l'hypothèse de colons chasseurs de mammoths néglige le rôle des femmes, des enfants et des grands-parents. L'étude des objets en pierre ne rend compte que de cinq pour cent des activités humaines. En outre, dans les sociétés de chasseurs-cueilleurs récentes, les femmes, les enfants et les personnes âgées des deux sexes fournissent la plus grande partie de la nourriture et s'acquittent de tâches essentielles, telle la fabrication de vêtements, de filets et de paniers. Les premiers Américains n'étaient sans doute pas différents.

Aujourd'hui, les archéologues découvrent enfin des objets qui ne sont

ni en pierre ni en os. Par exemple, à Monte Verde, T. Dillehay a mis au jour des fragments de cordes qui, selon lui, fixaient des tentes constituées de peaux d'animaux ; les restes de ces tentes ont également été retrouvés.

Ces objets corroborent la nouvelle image des premiers Américains : des groupes humains connaissant bien leur environnement, qui tuaient les mammoths à la lance et qui attrapaient également du poisson, cueillaient des baies, tissaient des fibres végétales pour des vêtements et des paniers, construisaient des bateaux pour leurs déplacements.

De nouvelles questions se posent alors. Par exemple, si le site de Monte Verde date de 14 700 ans, quel fut le premier homme de la culture Clovis ?



Qui étaient les premières populations des Amériques? Quand et comment sont-elles arrivées? Sont les populations de Clovis les premières en Amérique, soit les populations clairsemées se trouvaient en Amérique du Nord avant la culture Clovis, mais n'ont laissé aucune trace, soit une importante culture pré-Clovis reste à découvrir.

La conquête du Sud

On croyait naguère que les populations de la culture Clovis se déplaçaient uniquement à pied. Toutefois, les sites de la culture Clovis découverts sur les deux rives du Missouri indiquent qu'ils naviguaient. Les bateaux étaient indispensables aux personnes âgées, aux enfants, aux femmes enceintes... En outre, des populations humaines ont construit des embarcations il y a au

moins 40 000 ans : leur présence est attestée en Australie.

Les premiers bateaux, construits avec des peaux d'animaux et du bois, ont laissé peu de vestiges. Cependant, à partir des techniques de construction des peuples indigènes modernes en Amérique du Nord, Margaret Jodry a proposé des marqueurs archéologiques de construction de bateaux : par exemple, une configuration précise de trous de poteaux – en ovale ou en forme de bateau –, entourée de pierres témoignerait d'un ancien atelier nautique. Se fondant sur de telles hypothèses, les archéologues modifient leurs méthodes de fouilles et présentent une nouvelle analyse des sites déjà fouillés, afin d'y découvrir des indices de la présence de bateaux.

Ce travail suscite une nouvelle question : les premiers colons seraient-

ils arrivés par bateau? Pendant des décennies, les archéologues ont rejeté cette idée, mais, ces dernières années, l'hypothèse est revenue avec d'autant plus de force qu'elle résout des questions sur la durée de migration, de l'Alaska jusqu'à la Terre de Feu.

Comment expliquer que le site de Monte Verde soit âgé de 14 700 ans, alors que, selon la théorie dominante, des humains ne seraient venus d'Asie que de l'ordre de 15 700 ans (auparavant, des plaques de glace qui recouvraient totalement l'Alaska et le Canada bloquaient le passage). Les premiers colons auraient-ils parcouru environ 20 000 kilomètres à travers deux continents en 10 000 ans seulement?

Ils n'ont pu être si rapides que s'ils ont longé le littoral pacifique en bateau. Selon Knut Fladmark, de l'Université de Burnaby, qui a proposé cette hypothèse, dans les années 1970, les colons sont arrivés en Amérique parce qu'ils cherchaient un climat moins rude ; ils auraient traversé les deux continents en une centaine d'années. Les colons s'arrêtaient sans doute durant les mois d'hiver et s'installaient, le temps d'une génération, lorsque les ressources locales étaient suffisantes.

Bien que séduisante, la théorie de K. Fladmark ne sera pas facile à établir, car l'élévation du niveau de la mer due à la fonte des glaciers a inondé des milliers de kilomètres carrés, le long des côtes pacifiques des deux continents.

L'Amérique profonde

C'est dans l'eau que l'on trouvera des confirmations de cette théorie. L'exploration a commencé. En 1997, Daryl Fedgen et ses collègues archéologues des Parcs canadiens ont remonté un petit outil en pierre, qui gisait à 50 mètres de profondeur, au bord de la côte de Colombie-Britannique. Cet outil montre que des couples occupaient autrefois des zones littorales qui sont aujourd'hui submergées. Quand se sont-ils installés? Des dates comprises entre 15 000 et 50 000 ans sont avancées.

Une conquête des deux continents en 100 ans, par bateau, laisse peu de temps aux colons pour s'adapter au nouvel environnement. Or, selon T. Dillehay, les humains qui vivaient à Monte Verde, savaient exactement où ils se situaient : ils étaient dans la région depuis assez longtemps pour établir un camp fixe, à une heure de marche des marais voisins, luxuriants et riches en plantes



Kenneth Garrett, National Geographic

2. MONTE VERDE, au Chili, abritait une trentaine de personnes il y a environ 14 700 ans. C'est le plus ancien site connu d'habitation humaine en Amérique. L'archéologue Tom Dillehay (debout) et ses collègues ont découvert plusieurs objets d'origine humaine dans cet ancien campement.



Kenneth Garrett, National Geographic



Avec l'aimable autorisation de Tom Dillehay

3. PARMI LES OBJETS mis au jour sur le site de Monte Verde, on trouve un « bâton-pioche » en bois (à gauche) et des restes de troncs qui fixaient les tentes où vivaient la trentaine d'habitants de ce site (à droite).

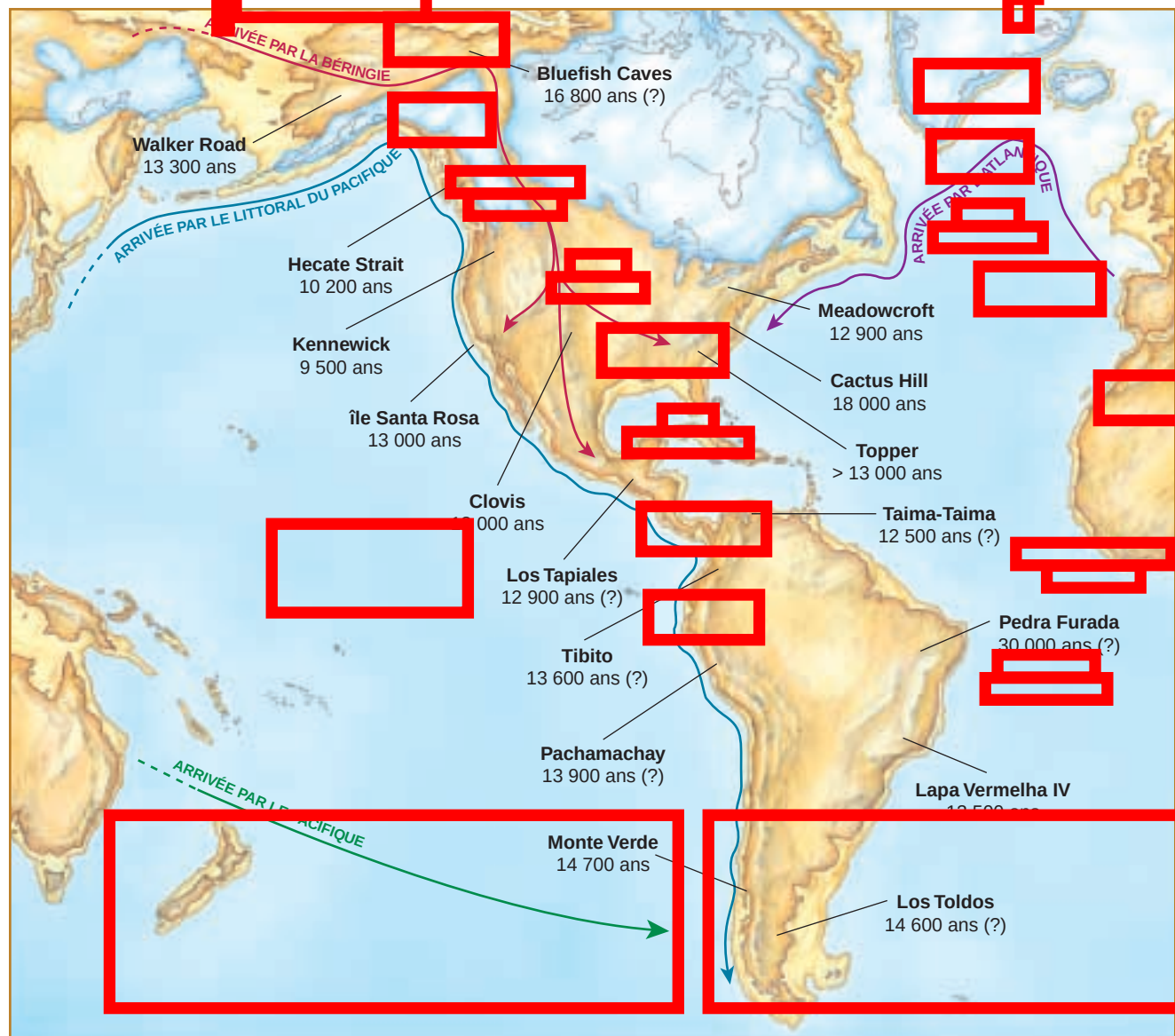
comestibles. L'océan et les collines andines étaient à une journée de marche. Le groupe s'était soigneusement fixé à proximité de trois environnements différents, qu'il exploitait. Par exemple, les habitants de Monte Verde suçaient des gâteaux faits d'algues. Ils tuaient ou récupéraient des animaux piégés dans les marécages voisins, et ils utilisaient les côtes de ces animaux en guise de pioches pour déterrer des tubercules et des rhizomes, dans ces mêmes marécages.

Une telle connaissance de l'environnement n'a pu s'acquiescer que par

plusieurs générations. Combien? La question est difficile, car on connaît peu de situations historiques comparables : l'arrivée de êtres humains «modernes» dans un continent désert n'a eu lieu que deux fois : en Australie et aux Amériques. Toutefois, T. Dillehay remarque que, dans la vallée de l'Indus ou en Chine, des civilisations complexes sont apparues en plusieurs dizaines de milliers d'années ; il conclut que les Américains n'ont pas posé les fondements d'une civilisation en quelques milliers d'années seulement. De ce fait, il pro-

pose que la colonisation de l'Amérique ait eu lieu il y a 20 000 ans.

Cette date est étayée par deux autres types d'indices. D'une part, la linguiste Johanna Nichols, de l'Université de Berkeley, soutient que la diversité des langues américaines indigènes n'a pu apparaître qu'après au moins 20 000 ans – peut-être même 30 000 ans – d'occupation du Nouveau Monde par des humains. D'autre part, des généticiens ont comparé plusieurs marqueurs d'ADN d'Américains indigènes modernes et de Sibériens modernes.



4. LA CONQUÊTE DE L'AMÉRIQUE EST aujourd'hui racontée par quatre théories. Selon la théorie dominante (flèches rouges), des émigrants du Nord-Est de l'Asie auraient traversé l'isthme qui unissait la Sibérie et l'Amérique du Nord durant la dernière période glaciaire, puis ils auraient parcouru le Canada par un corridor non gelé et colonisé les deux continents. D'autres archéologues pensent aujourd'hui (flèches bleues) que des explorateurs du Sud-Est de l'Asie ont plutôt suivi le littoral, à bord de petits bateaux, et atteint la Patagonie en une centaine d'années. La théorie, peu

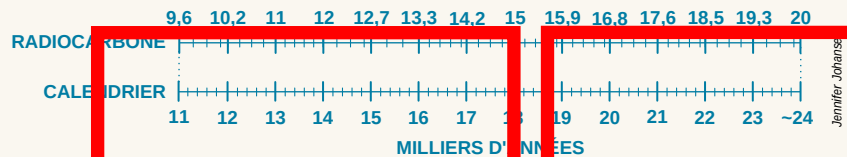
probable, de la traversée du Pacifique (flèches vertes), stipule que des habitants de l'Australie et des îles du Pacifique Sud ont poursuivi leur route vers l'Est jusqu'en Amérique du Sud. Enfin, selon la théorie de la traversée de l'Atlantique (flèches violettes), des habitants de la péninsule ibérique ont suivi en bateau le bord des glaciers qui recouvraient alors la mer du Nord. Cette dernière théorie est fondée sur une similitude entre les pointes de lances de la culture Clovis et la technique solutréenne en Europe, il y a 16 000 à 24 000 ans.

La datation des sites américains

La datation de l'arrivée des premiers êtres humains en Amérique se heurte à un problème de correspondance : une année sur l'échelle de la désintégration du carbone n'est pas toujours égale à une année calendaire. Les êtres vivants absorbent du carbone sous deux « isotopes » : le carbone 12 et le carbone 14, qui ont les mêmes nombres de protons et d'électrons (six pour chacun) mais des nombres différents de neutrons (respectivement six et huit). Dans un organisme vivant, les proportions de carbone 14 et de carbone 12 correspondent à celles de l'atmosphère. Toutefois, après la mort d'un organisme, ce rapport change, car le carbone 14, radioactif, se désintègre en carbone 12, stable. Ainsi, le rapport du carbone 14 au carbone 12 diminue progressivement. Connaissant la vitesse de désintégration du carbone 14, on estime la quantité de carbone 14 d'un objet, puis on détermine son âge.

Toutefois, le rapport variable du carbone 14 au carbone 12 dans l'atmosphère fausse le calibrage des dates. Ce déséquilibre s'explique par les variations du taux de croissance des arbres, les éruptions volcaniques et la datation à l'uranium-thorium (l'uranium, radioactif, se désintègre en thorium).

Malheureusement, la période comprise entre 10 000 et 20 000 ans est la plus difficile à dater. Pendant de nombreuses années, les archéologues ont donné leurs résultats en années de radiocarbone non corrigées. Cependant, grâce à des découvertes récentes, on ajuste plus facilement les dates de cette époque : toutes les dates de cet article ont été calibrées en années calendaires.



LES ANNÉES DE RADIOCARBONE (en haut) sont transcrites en années calendaires (en bas). Par exemple, un objet qui a 11 000 années de radiocarbone a en fait 13 000 ans pour le calendrier.

(voir *L'origine des Amérindiens*, page 82) : ils estiment que les ancêtres des premiers ont quitté la Sibérie pour le Nouveau Monde il y a au moins 30 000 ans.

Cette hypothèse est corroborée par un des plus vieux squelettes américains connus : celui d'une femme adulte qui vivait il y a 13 500 ans. Découvert dans le Sud-Est du Brésil, par Walter Neves, de l'Université de São Paulo, ce squelette ressemble plus à celui d'Africains et d'aborigènes australiens qu'à ceux d'Asiatiques ou d'Américains indigènes modernes. Cette observation, ainsi que des analyses d'une cinquantaine de crânes vieux de 8 900 à 11 600 ans, montre que des migrants non monoloïdes figuraient parmi les premiers colons américains.

Ainsi, le Nouveau Monde est pluriethnique depuis des millénaires : des chasseurs de la période glaciaire ont traversé la Beringie pour s'établir sur le continent américain, mais ils n'étaient pas les premiers à le faire ainsi, et ils n'étaient pas les seuls. Les archéologues découvrent d'autres populations telle celle de pêcheurs longeant en bateau la côte pacifique ou celle de chasseurs-cueilleurs vivant dans les forêts tempérées de l'Amérique du Nord et du Sud.

Parallèlement, les archéologues fouillent l'Alaska, à la recherche de vestiges des premiers habitants ; des géologues étudient la fonte des glaciers et le moment où elle a permis le déplacement des colons vers le centre du Canada et les États-Unis. Dans l'Est des États-Unis, on découvre qu'un groupe, peut-être de la culture Clovis, était déjà établi il y a 18 000 ans.

Al Godoyear, de l'Université de Caroline du Sud, est récemment retourné sur un site de la culture Clovis pour y chercher des traces humaines encore plus anciennes. Les objets qu'il a découverts, de minuscules lames de pierre et des outils de type racloir associés à l'utilisation du bois et de l'os, diffèrent totalement de ceux de la culture Clovis. Les datations récentes montrent que ces objets sont sans doute antérieurs à environ 100 ans à la culture Clovis (voir l'encadré ci-dessus).

D'autres archéologues ont décrit des campements qui pourraient être vieux de 30 000 ans, mais les datations restent à confirmer.

Sasha NEMECEK est journaliste à *Scientific American*.

L'origine des Amérindiens

SERGIO PENA • FABRÍCIO SANTOS

Les études génétiques du chromosome Y montrent que les premiers Amérindiens sont venus de Sibérie centrale.

Au XV^e siècle, lorsque les Européens débarquent aux «Indes occidentales», ils rencontrent des peuples que l'Ancien Testament ne mentionne pas. Sont-ils les descendants d'une tribu perdue d'Israël? Le jésuite José de Acosta propose qu'ils soient plutôt d'anciens chasseurs venus d'Asie. Quatre siècles plus tard, en 1840, Ernst Haeckel (1834-1919) reprend cette hypothèse en la précisant : les Amérindiens descendraient de populations sibériennes (voir la figure 1).

Aujourd'hui, la question reste discutée : la plupart des paléanthropologues admettent que des ancêtres des indigènes américains sont arrivés par la Beringie, un pont de terres qui existait, il y a 40 000 à 13 000 ans, entre l'Amérique du Nord et la Sibérie. D'autres hypothèses subsistent (voir *Les premiers Américains*, page 76), mais il est avéré qu'il y a 10 000 ans des êtres humains occupaient les trois Amériques, de l'Alaska à la Patagonie.

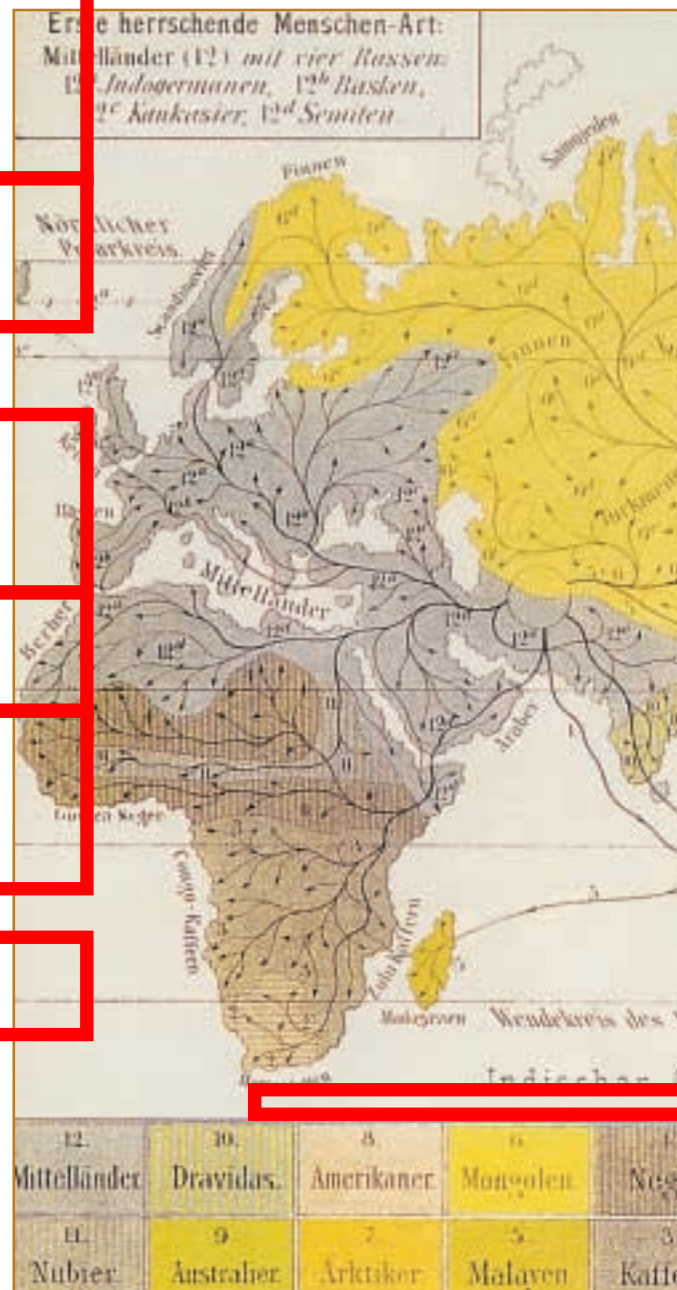
Les Américains de souche se répartissent en trois groupes linguistiques distincts : les Amérindiens (la grande majorité, dispersés à travers les deux continents) ; les Na-Déné (à l'Ouest de l'Amérique du Nord ; les Apaches et les Navajos appartiennent à ce groupe) ; et les Aléoutiques esquimaux (au Nord de l'Amérique du Nord).

Tous égaux, tous différents

Ces peuples ont été beaucoup étudiés, mais bien des questions restent ouvertes. Sont-ils arrivés de manière régulière ou par vagues successives? Combien y a-t-il eu de ces migrations? Quand ont-elles eu lieu? D'où sont-elles parties? Quelles ont été les migrations à l'origine des populations respectives actuelles? Certains groupes, à l'origine des migrations, se sont-ils éteints?

Depuis longtemps, la paléanthropologie et l'archéologie linguistique explorent ces questions. Aujourd'hui, nous utilisons la génétique moléculaire pour tester leurs hypothèses et faire de nouvelles découvertes.

La génétique moléculaire dispose de deux méthodes pour comprendre l'histoire de l'évolution humaine et, notamment, son installation dans les deux continents américains. La première consiste à prélever de l'ADN humain sur des momies ou sur des ossements, afin de reconstituer la structure génétique des populations du passé. La seconde, que nous avons choisie, est l'étude



1. LA MAPPEMONDE D'ERNST HAECKEL, publiée en 1884, situe l'origine de l'humanité dans une région hypothétique, nommée Lémurie. Aujourd'hui, on pense que l'*Homo sapiens* provient d'Afrique, puis qu'il a colonisé les Amériques via l'Asie.

génétique des populations actuelles, puis l'établissement d'arbres qui expliquent la répartition actuelle des gènes.

Pour une telle étude, nous avons utilisé le polymorphisme de l'ADN : certaines régions du génome humain diffèrent d'un individu à l'autre. La transmission de ces régions dépend de leur composition et de leur localisation sur le génome humain. Les polymorphismes des autosomes, ceux des séquences présentes sur les chromosomes non sexuels, sont d'excellents marqueurs, très polymorphes, car ces chromosomes sont en double exemplaire dans chaque cellule, et les brassages de gènes sont fréquents d'une génération à l'autre chez l'être humain.

En revanche, le chromosome sexuel Y et l'ADN mitochondrial (contenu dans les mitochondries, petits organites intracellulaires qui fournissent l'énergie aux cellules) proviennent d'un seul géniteur ; le chromosome Y se transmet de père à fils, tandis que l'ADN mitochondrial passe de mère à fils et à filles. Ces ADN n'échangent pas de gènes avec d'autres segments génomiques ; autrement dit, ils ne se recombinaient pas : les gènes du chromosome Y et de l'ADN mitochondrial se transmettent en blocs, nommés haplotypes.

Seules les mutations modifient les haplotypes patrilinéaires (le chromosome Y) ou matrilinéaires (l'ADN mitochondrial). Ces haplotypes constituent

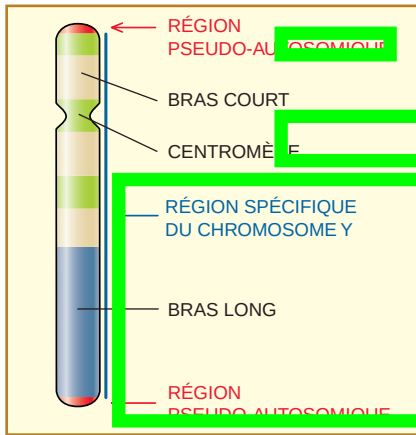
donc des marqueurs importants pour les reconstructions évolutives.

Un chromosome idéal

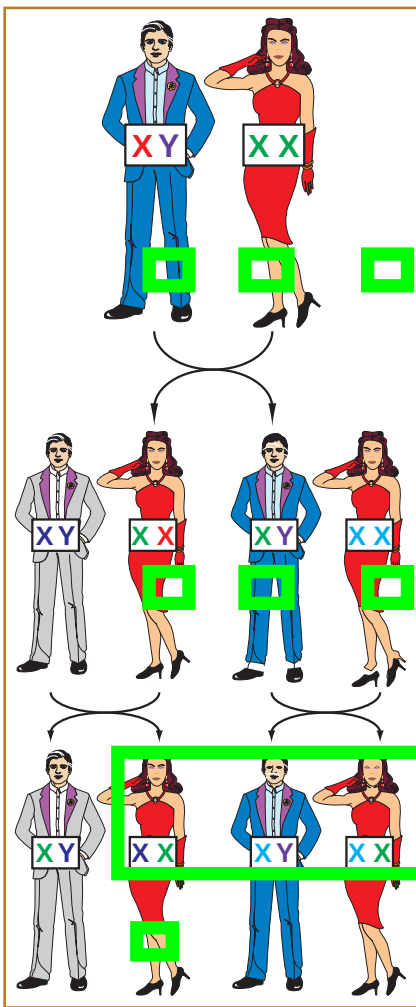
Le chromosome Y humain se compose de trois parties (voir la figure 2) : deux petites régions (nommées régions pseudo-autosomiques, car elles se recombinaient parfois) occupent les extrémités des bras courts et longs et ressemblent à celles du chromosome X ; la troisième région, qui constitue plus de 90 pour cent du chromosome, est spécifique du chromosome Y et ne se recombine pas. Les haplotypes de cette dernière région se transmettent, inchangés, de



Illustrations de Sergio D.J. Pena et Fabricio R. Santos



2. LE CHROMOSOME Y HUMAIN est divisé en trois parties. Deux régions pseudo-autosomiques (en rouge), ressemblent à celles du chromosome X, l'autre chromosome sexuel. Chez les hommes, ces régions s'apparient au cours de la méiose (la formation des gamètes) et échangent du matériel génétique. Le fragment spécifique du chromosome Y ne se recombine pas et se transmet en bloc (nommé **haplotype**) de père à fils.



3. LE CHROMOSOME Y se transmet de père à fils, de sorte que les généticiens l'utilisent pour retrouver des filiations patrilinéaires. Les individus en bleu appartiennent à une même filiation patrilinéaire.

père à fils, jusqu'à ce qu'apparaisse une mutation. Ainsi, on obtient des filiations patrilinéaires en remontant des dizaines de générations dans le passé. En revanche, les haplotypes du chromosome X, inchangés, ne représentent la contribution génétique que d'une très petite partie des ancêtres d'un individu (compromis par la branche qui va au père du père du père... du père).

Les deux types de polymorphisme les plus utilisés dans les études de génétique évolutive sont les polymorphismes d'un seul nucléotide (SNP pour *single nucleotide polymorphism*) et les polymorphismes de microsatellites. Les SNP sont des sites de l'ADN où alternent deux bases (les molécules, nommées bases, que sont l'adénine A, la thymine T, la cytosine C et la guanine G) sont les composants élémentaires de l'ADN) avec des fréquences supérieures à un pour cent dans les chromosomes d'une population. Ce polymorphisme est plus commun.

Une unité génétique

Les SNP mutent rarement : ils constituent donc des marqueurs moléculaires d'évolution lente ; chaque mutation correspond à un seul épisode évolutif.

Les polymorphismes d'un seul nucléotide sur le chromosome Y, faciles à analyser, ont été décrits pour la première fois en 1994. Puis, récemment, Peter Underhill et ses collègues de l'Université de Stanford ont montré comment étudier systématiquement le chromosome Y au moyen des SNP.

D'autre part, les microsatellites sont des répétitions de motifs qui comprennent une à six bases. La fréquence de mutation des microsatellites est quatre à cinq fois supérieure à celle des SNP : ces microsatellites sont donc des marqueurs d'évolution rapide. Ainsi, les SNP fournissent une information moléculaire sur les épisodes évolutifs les plus anciens, tandis que les microsatellites renseignent sur des événements récents.

En 1997, Jörg Epplen et ses collègues de l'Université de Munich ont décrit le premier microsatellite polymorphe du chromosome Y humain. Nommé *DYS19*, il est composé de répétitions du motif GATA. Nous avons démontré la présence de cinq allèles (des variantes d'un même microsatellite, qui diffèrent par le nombre de répétitions) dans la population brésilienne, avec de nombreuses petites variations (une seule base) d'un individu à un autre.

À la même époque, Roewer et ses collègues ont étudié le microsatellite *DYS19* sur un échantillon de 11 Indiens yanomamos, parmi lesquels 10 avaient le même allèle nommé A (13 répétitions du motif GATA). Or, chez les Européens, l'allèle B (14 répétitions GATA) domine, alors que, chez les Africains et les Asiatiques, les allèles C et D sont les plus fréquents. Ainsi, l'analyse du microsatellite *DYS19* révèle une homogénéité de la population amérindienne.

Parallèlement, nous avons étudié un polymorphisme « alphasid » du centromère du chromosome Y (la région resserrée du chromosome). Ce nouveau type de polymorphisme a été découvert à Oxford par l'un d'entre nous (F. Santos) et par Chris Tyler-Smith, à l'aide d'une technique d'amplification des gènes fondée sur l'emploi d'ADN polymérase (l'enzyme qui duplique l'ADN). Aujourd'hui, le nombre de polymorphismes alphasides décrits dépasse la cinquantaine.

À partir du microsatellite *DYS19* et des polymorphismes alphasides, nous avons mis en évidence 46 haplotypes distincts dans les populations de divers continents.

En collaboration avec des collègues d'Argentine, du Brésil et du Chili, nous avons établi les haplotypes de 73 Amérindiens originaires de douze tribus d'Amérique du Sud : les Mapuches, les Wichits, les Chorotis, les Chulupis, les Tobas, les Huilliches, les Atacames, les Suruis, les Caritianas, les Quechuas, les Aucas et les Mayas. L'haplotype nommé IIA (une combinaison du type alphasid et de l'allèle A du microsatellite *DYS19*), était présent dans 74 pour cent des échantillons. En éliminant les Mapuches, qui se sont beaucoup mêlés aux Européens, on obtient une fréquence de l'haplotype IIA qui atteint 91 pour cent. Nous avons alors conclu que l'haplotype IIA était un bon marqueur génétique des Amérindiens du Sud. Puis nous avons confirmé cette conclusion en étudiant 37 Amérindiens originaires de cinq tribus d'Amazonie et du centre du Brésil : les Waiwais, les Garijóns, les Zorós, les Suruis et les Xavantes : là encore, l'haplotype IIA était prépondérant (87 pour cent).

Une hypothèse confirmée

Évidemment, nous avons alors cherché qu'elle était la fréquence de cet haplotype chez les Amérindiens d'Amérique du Nord. Avec Kenneth

Weiss, de l'Université d'État de Pennsylvanie, et John Moore, de l'Université de Floride, nous avons étudié les échantillons d'ADN d'Amérindiens Muskokes qui vivent aujourd'hui dans l'Oklahoma. L'haplotype le plus fréquent est encore l'haplotype IIA, mais il n'est présent que chez 38 pour cent des individus étudiés. Cette faible fréquence s'explique par l'histoire de la tribu Muskoke : elle est entrée pour la première fois en contact avec les Européens en 1541 et s'est ensuite croisée largement avec des populations européennes et avec des esclaves noirs.

Au total, la prédominance de l'haplotype IIA confirme que les Amérindiens des trois Amériques proviennent d'une même onde migratoire, d'origine asiatique.

À Stanford, Peter Underhill et Luca Cavalli-Sforza ont corroboré nos résultats en étudiant également l'haplotype *DYS19* et aussi un SNP nommé *DYS199* : sur ce site de l'ADN, les individus ont soit une base C, soit une base T.

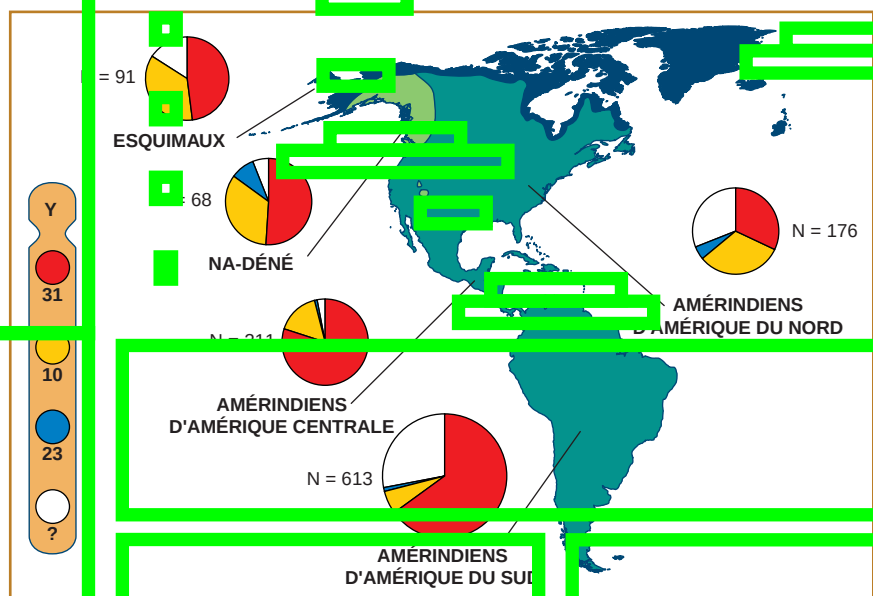
On trouve l'allèle C de *DYS199* chez tous les Européens, les Asiatiques, les Africains et les primates supérieurs étudiés. En revanche, l'allèle T apparaît chez 91 pour cent des 54 individus d'Amérique du Nord étudiés, soit 15 Caribéens, 17 Surois, 8 Mayas, 2 Colombiens, 6 Esquimaux et 6 Navajos. De plus, une trentaine des 46 individus qui avaient l'allèle T de l'haplotype *DYS199* étaient porteurs de l'allèle A de l'haplotype *DYS19*.

Puis nous avons confirmé que tous les Amérindiens IIA que nous avons étudiés portaient également l'allèle T de l'haplotype *DYS199*. Toutefois, cet allèle était absent des échantillons provenant de Mongolie ou de Sibérie : la mutation de C en T a probablement eu lieu au cours de la migration des peuples asiatiques qui ont peuplé l'Amérique, ou immédiatement après celle-ci.

P. Underhill a mis en évidence les allèles A et T de l'haplotype *DYS19* et T de l'haplotype *DYS199* chez 66 pour cent des Esquimaux et chez 33 pour cent des Navajos. Or, les trois groupes linguistiques proviendraient de trois vagues migratoires différentes.

D'autres études ont confirmé la présence de l'allèle T chez les Na-Déné et chez les Esquimaux Nord-américains : cette présence est-elle due au croisement de ces peuples avec des Amérindiens ? L'hypothèse précédente est-elle fautive ?

Par ailleurs, la faible fréquence de l'allèle T dans les populations sibé-



4. LES PRINCIPAUX GROUPES D'INDIGÈNES AMÉRICAINS ont été identifiés par des analyses génétiques. Pour chaque groupe, les principaux types de chromosomes Y sont indiqués (haplotypes 31, 10 et 23). L'haplotype 31 est l'haplotype fondateur principal, l'haplotype 10 est l'haplotype fondateur secondaire, ancêtre de l'haplotype 31, duquel il diffère par une base C sur le site *DYS199*. L'haplotype 23 diffère des haplotypes 10 et 31 par une mutation sur le site *RPS4Y*. Quelques individus représentent des chromosomes qui n'ont pas encore été caractérisés (en blanc) ou qui proviennent d'autres groupes ethniques (Européens, Africains, etc.).

rienne (Esquimaux de Sibérie, Chukchis et Evenks) est surprenante : y a-t-il eu des migrations inverses, de l'Amérique vers l'Asie ?

Environ 90 pour cent des Amérindiens d'Amérique du Sud et 50 à 70 pour cent des Amérindiens d'Amérique du Nord ont un même haplotype principal de chromosome Y. Peut-on alors identifier la population asiatique à l'origine de la majorité des Amérindiens ?

Pour répondre à cette question, nous avons étudié 306 hommes de diverses populations du monde entier, en mettant l'accent sur celles originaires de Mongolie et de Sibérie. Nous avons enquêté sur cinq groupes linguistiques différents parmi les Sibériens : les Bouriat, les Yakoutes, les Evenkis, les Altaïens et les Kètes. Nous avons employé sept types de polymorphismes et identifié 30 polymorphismes sur le chromosome Y humain.

Les populations asiatiques les plus proches génétiquement des Amérindiens sont les groupes originaires de Sibérie : les Kètes (de la vallée de l'Ienisseï) et les Altaïens (des monts Altaï). Les trois autres groupes sibériens sont regroupés avec les Mongols.

Un haplotype du chromosome Y observé chez les Altaïens et les Kètes conserve une proche parenté avec le principal chromosome Y des Amérindiens. Les haplotypes similaires à celui-ci sont rares dans la plus grande partie

de l'Asie (Chine, Japon, Indonésie), mais ils sont abondants dans toute l'Europe. Ainsi un type de chromosome Y ancestral, vieux de plus de 30 000 ans en Eurasie, serait originaire de la région des chromosomes Y européens, des chromosomes Y des Kètes et des Altaïens de Sibérie et des chromosomes Y des indigènes américains. Autrement dit, des peuples auraient migré vers l'Amérique en passant par le Nord de l'Asie, à travers la Sibérie et la Béringie.

Sergio PENA et Fabrício SANTOS sont professeurs à l'Université fédérale de Minas Gerais.

S.D. PENA, F.R. SANTOS, N.O. BIANCHI, C.M. TRAVES, F.R. CARNESE, F. ROTH-HAMMER, T. GERELSAIKHAN et al., *A Major Founder Y-Chromosome Haplotype in Americans*, in *Nature Genetics*, n° 11, pp. 15-16, 1995.

P.A. UNDERHILL, L. JIN, R. ZEMANS, P.J. O'NEILL et L.L. CAVALLI-SFORZA, *A Pre-Columbian Y-Chromosome-Specific Transition and Its Implications for Human Evolutionary History*, in *Proceedings of the National Academy of Sciences, U.S.A.*, n° 93, pp. 196-200, 1996.

F.R. SANTOS, A. PANDYA, C. TYLER-SMITH, S.D. PENA, M. SCHANFIELD, W.R. LEONARD, L. OSIPOVA et al., *The Central Siberian Origin for Native American Y-Chromosome*, in *American Journal of Human Genetics*, n° 64, pp. 619-628, 1999.

Lasers à impulsions ultracourtes

JOHN-MARK HOPKINS • WILSON SIBBETT

Les lasers dont les impulsions ne durent que quelques millièmes de nanoseconde se sont imposés dans plusieurs sciences et techniques : les chirurgiens les utilisent pour de délicates opérations des yeux, les ingénieurs des télécommunications pour transmettre rapidement des données, les chimistes pour observer le comportement des molécules ou réagissent.

En combien de temps avez-vous lu cette phrase? Reconnaître la première lettre n'a pris que quelques millisecondes (10^{-3} seconde). Des molécules de neuromédiateur ont transporté les signaux d'un neurone à un autre, dans le cerveau, en 0,05 milliseconde, soit 50 microsecondes (10^{-6} seconde). Teniez-vous votre revue à la bonne distance? La lumière a mis une ou deux nanosecondes (10^{-9} seconde) de la page à vos yeux, et environ 20 picosecondes (10^{-12} seconde) pour traverser le cristallin de l'œil. Ces événements très brefs sont toutefois des éternités par rapport aux impulsions laser les plus courtes que l'homme ait aujourd'hui produites : ces impulsions laser durent seulement quelques femtosecondes (ou millièmes de milliardième de seconde).

On conçoit difficilement la brièveté d'une femtoseconde (10^{-15} seconde). Une femtoseconde est à une seconde ce qu'une seconde est à 32 millions d'années, et il y a dix fois plus de femtosecondes dans une seconde qu'il n'y a d'heures écoulées depuis le Big Bang. De nombreux mécanismes fondamentaux de l'Univers, tels le mouvement des électrons entre les atomes ou la formation de liaisons entre deux molécules, se produisent en des temps inférieurs à quelques centaines de femtosecondes. Les physiciens utilisent des impulsions femtosecondes pour enregistrer

et étudier en détail ces événements. À la fin du XIX^e siècle, Jules Étienne Marey utilisait des flashes de lumière qui duraient quelques microsecondes pour produire de superbes photographies de chevaux au galop. À l'Institut de technologie de Californie, Ahmed Zewail a observé le déroulement de réactions chimiques de quelques centaines de femtosecondes (ce travail a été récompensé par le prix Nobel de chimie en 1999). Plus récemment, les physiciens ont «figé» le mouvement des électrons dans des matériaux semi-conducteurs, de sorte que, comprenant mieux le fonctionnement des dispositifs optoélectroniques, ils ont mis au point des systèmes rapides de traitement du signal.

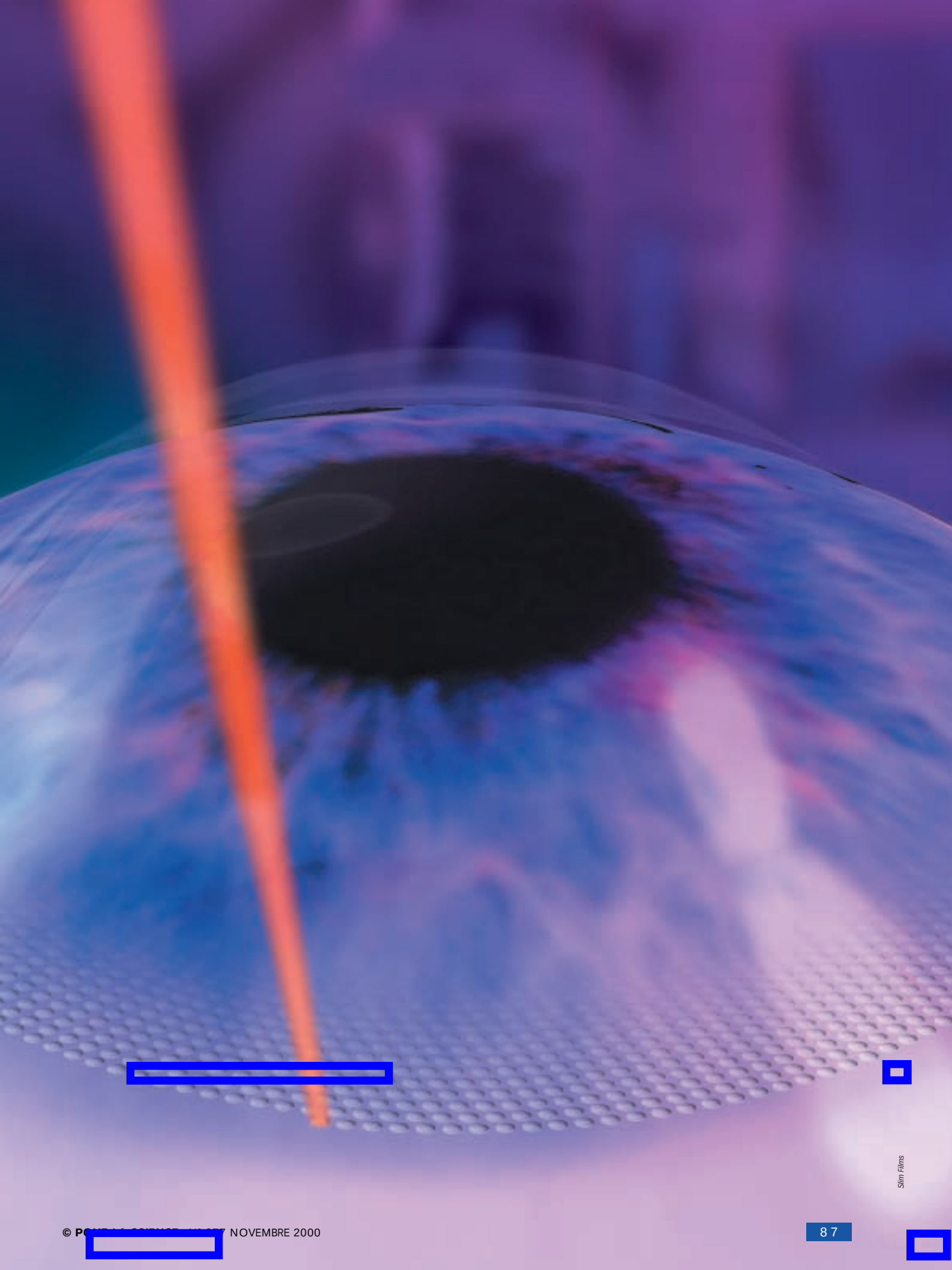
Au milieu des années 1960, les premiers lasers étaient volumineux, fragiles et gros consommateurs d'énergie. Aujourd'hui les lasers émettant des impulsions ultracourtes (moins de dix femtosecondes) sont compacts, polyvalents et fiables. L'intensité et la diversité de ces lasers ne cessent d'augmenter : les faisceaux obtenus aujourd'hui couvrent tout le spectre électromagnétique, des rayons X aux rayons T (rayonnement térahertz, situé au-delà de l'infrarouge), et les puissances maximales atteignent plusieurs pétawatts (plusieurs milliards de mégawatts). Ces systèmes ont de nombreuses applications en physique, chimie, biologie, médecine, optique...

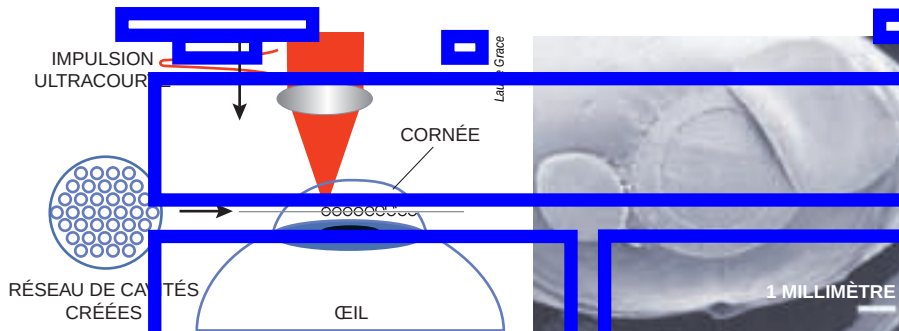
Le plus pointu des scalpels

Les impulsions ultracourtes des lasers sont très intenses. Bien que la puissance moyenne du laser soit modérée et que l'énergie totale d'une impulsion soit faible, la puissance instantanée est considérable, parce que les impulsions sont brèves (la puissance est égale à l'énergie divisée par le temps pendant lequel cette énergie est transmise). Dans les lasers à impulsions ultracourtes les plus courants, l'intervalle entre chaque impulsion est 100 000 fois plus long que la durée de l'impulsion elle-même, de sorte que la puissance maximale est 100 000 fois supérieure à la puissance moyenne. Par exemple, une impulsion de 100 femtosecondes avec une énergie de trois microjoules (insuffisante pour élever la température d'une goutte d'eau de un milliardième de degré) délivre une puissance maximale de 30 mégawatts.

Focalisées sur une minuscule surface, de telles puissances subliment de nombreux matériaux (elles les font passer de l'état solide à l'état gazeux sans qu'ils aient le temps de fondre), transformant les impulsions ultracourtes en d'excellents outils de micro-usinage, de perçage, de découpe et de soudure.

1. LES IMPULSIONS LASER ULTRACOURTES peuvent, en chirurgie des yeux, découper la cornée en provoquant une série de perforations à profondeur constante.





2. UN TRAITEMENT EXPÉRIMENTAL DE L'ŒIL PAR LASER dégage une épaisseur de la cornée par la création de minuscules trous au point focal du laser (gauche). D'autres découpes du même type détachent un disque de cornée intermédiaire (droite). Quand on rabat la première épaisseur, la cornée est aplatie et la myopie corrigée. En chirurgie classique, la cornée est coupée avec une lame : la procédure est plus risquée et la découpe moins fine.

On obtient une résolution plus fine que la largeur du faisceau laser en réglant le laser de sorte que seule l'intensité de la partie centrale, la plus brillante, soit supérieure au seuil d'ablation du matériau. L'impulsion communique alors l'énergie au point focal si rapidement que la chaleur ne diffuse pas dans le matériau environnant : les contours découpés sont lisses et précis (voir la figure 3). Par cette méthode, les physiciens ont usiné avec précision du diamant, du carbure de titane et de l'émail dentaire. Les lasers découpent sans danger des explosifs, sublimant le matériau qui se trouve au point de découpe sans faire détonner le matériau adjacent : ils semblent utiles pour des opérations de désarmement.

Les applications chirurgicales sont légion. En chirurgie microvasculaire, un laser à impulsions ultracourtes perce de petits trous dans la paroi du cœur pour fournir du sang oxygéné aux muscles cardiaques dont les vaisseaux sanguins sont irrémédiablement bouchés. En médecine, la faible puissance moyenne des impulsions ultracourtes est utile, parce que les modifications sont limitées et les tissus avoisinants épargnés. Des médecins du Laboratoire Lawrence Livermore, en Californie, ont utilisé des impulsions ultracourtes associées d'un système de commande astucieux pour enlever des excroissances osseuses dans la colonne vertébrale d'un patient, sans endommager les tissus nerveux adjacents. De même, des ophtalmologues de l'Université d'Heidelberg et de l'Université du Michigan ont mis au point un procédé de chirurgie oculaire qui focalise un faisceau d'impulsions ultracourtes au niveau de la cornée, créant un fin réseau de bulles interconnectées à une profondeur précise ; ces bulles sont formées par cavitation, c'est-à-dire par évaporation de l'eau présente

au point focal (voir la figure 2). Ils parviennent alors à décoller la surface externe de la cornée, découvrant le disque de tissu profond à enlever ; puis ils replacent la couche supérieure, restée intacte. Le changement de courbure de la cornée obtenu par l'enlèvement du disque de tissu corrige la myopie ou l'hypermetropie. Cette technique préserve la surface externe lisse de la cornée. Les impulsions femtosecondes produisent une incision plus régulière et plus précise que les impulsions des lasers classiques.

L'imagerie

Une source de lumière forme des images caractéristiques parce qu'elle excite des marqueurs spécifiques : des colorants spéciaux ou des molécules naturelles qui absorbent la lumière, puis la réémettent. En imagerie, l'utilisation de lumière de grande longueur d'onde est avantageuse, car les photons (ou grains de lumière) sont moins énergétiques, de sorte qu'une molécule n'est excitée que si plusieurs photons la heurtent simultanément. Cette excitation « multiphotonique » n'apparaît donc qu'à de fortes densités.

Les lasers à impulsions ultracourtes procurant les intensités requises, tout en délivrant une puissance moyenne suffisamment faible pour éviter d'endommager les tissus. De surcroît, les grandes longueurs d'onde pénètrent plus dans les tissus et abîment moins les cellules que les longueurs d'onde plus courtes. Comme pour le micro-usinage de haute précision, on obtient une meilleure résolution en ajustant l'intensité de telle sorte que la fluorescence n'ait principalement lieu qu'au point focal du faisceau. Cette technique produit des images tridimensionnelles de tissus vivants avec une clarté et une définition inégalées. Une

méthode similaire procure des images de circuits intégrés en silicium dans leur emballage plastique : on détecte alors les courants électriques provoqués par l'excitation multiphotonique.

L'imagerie à rayons T, d'autre part, utilise le rayonnement térahertz, dont la longueur d'onde est comprise entre 15 micromètres et un millimètre. Les rayons T pénètrent profondément dans de nombreux matériaux qui sont opaques aux longueurs d'onde plus courtes, comme celles de la lumière visible. Ces matériaux opaques deviennent ainsi « transparents ». Les images ont une meilleure résolution que celles qui sont obtenues avec les micro-ondes, dont les longueurs d'onde sont pourtant supérieures (plus de un centimètre). Les micro-ondes induisent une agitation des molécules et, notamment, des molécules d'eau, provoquant un accroissement de la température. On connaît l'exemple du four à micro-ondes.

Dans les années 1980, David Auston et ses collègues de l'Université de Columbia ont créé des rayons T en envoyant les impulsions ultracourtes sur des structures semi-conductrices : les impulsions provoquent alors des courants d'électrons qui engendrent les rayons T. Les rayons T produits par des impulsions laser sont nettement distincts de ceux qui proviennent de la radioactivité naturelle des corps et du bruit de fond, de sorte que les systèmes d'imagerie à rayons T différencient des couches de matériaux très semblables, tels que différents types de tissus mous. Par exemple, l'imagerie à rayons T distingue le tissu brûlé et le tissu sain.

L'imagerie par impulsions térahertz a été présentée pour la première fois en 1997 par Martin Nuss et ses collègues des laboratoires AT&T Bell (devenu depuis Lucent Technologies). Cependant, des physiciens l'ont utilisée pour inspecter des composants électroniques pour détecter des caries dentaires ou d'autres dégénérescences des tissus. Dans les aéroports, les rayons T identifieraient les armes fabriquées en céramique, aujourd'hui difficilement repérables avec les équipements standards à rayons X.

Dans l'industrie micro-électronique, les lasers à impulsions ultracourtes facilitent la fabrication de circuits intégrés complexes : ils contrôlent l'épaisseur du métal et des couches de semi-conducteurs, sans interrompre la production, permettant d'obtenir une fabrication de meilleure qualité

et un rendement supérieur. Dans l'industrie des télécommunications, des impulsions femtosecondes transmettent simultanément plus de 200 canaux de données dans une seule fibre optique.

De telles applications étaient inenvisageables il y a seulement dix ans, parce que les lasers à impulsions femtosecondes étaient trop encombrants, inefficaces et peu fiables. Aujourd'hui, ils sont fiables, compacts, de sorte qu'ils peuvent être intégrés dans des équipements d'imagerie ou de diagnostic portables.

La production d'impulsions ultracourtes

Les principes de base du fonctionnement d'un laser sont l'amplification et la rétroaction. À la fin des années 1960, Jimi Hendrix subjuguait les foules en plaçant sa guitare *Stratocaster* devant un amplificateur pour obtenir une longue rétroaction plaintive. Au début de la même décennie, au Laboratoire de recherche Hughes, Theodore Maiman obtenait un phénomène similaire avec de la lumière : il avait montré qu'un bâtonnet de rubis excité et placé

dans une cavité pour produire une rétroaction optique produisait un faisceau intense de lumière laser.

La cavité, également nommée cavité résonante, est l'espace entre les deux miroirs ; le bâtonnet joue le rôle de milieu amplificateur (voir la figure 4). Dans la cavité, la lumière rebondit sur les miroirs et gagne en intensité à chaque traversée de l'amplificateur optique. Un des miroirs, partiellement transparent, laisse échapper une partie de la lumière sous la forme d'un faisceau laser. Seules les longueurs d'onde, ou modes, correspondant à des



3. LE MICRO-USINAGE LASER met à profit les propriétés des impulsions les plus courtes. Une impulsion de 200 femtosecondes produit par ablation un trou parfaitement lisse dans l'acier (à gauche). Sans techniques particulières, des impulsions d'une puissance

deux fois moindre et d'une durée seize fois supérieure, font fondre la zone environnante (au centre). Le micro-usinage laser peut produire des endoprothèses coronaires de tantale lisses (à droite) servant à traiter des maladies cardiovasculaires.

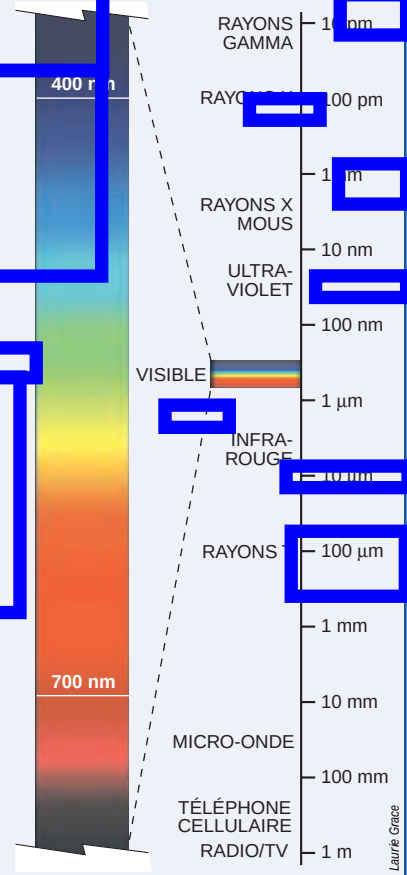
Principes de base

Le rayonnement électromagnétique présente de nombreuses facettes, dont la plus familière est la lumière visible. Toutes les couleurs de l'arc-en-ciel sont caractérisées par un paramètre : la longueur d'onde, allant de 380 nanomètres environ, en ce qui concerne le violet le plus profond, jusqu'à 750 nanomètres environ pour le plus rouge des rouges. La plupart des couleurs que nous voyons autour de nous, en particulier les bruns, les gris et les couleurs pastel ne figurent pas dans l'arc-en-ciel ; elles sont le résultat de la combinaison de nombreuses longueurs d'onde. De part et d'autre du domaine visible, le spectre électromagnétique s'étend jusqu'aux rayons gamma à une extrémité (les fréquences les plus hautes et les longueurs d'onde les plus courtes), et jusqu'aux ondes radio à l'autre. Les ondes de faible énergie émises par le réseau électrique à 60 hertz des machines particulières se situent environ sept graduations au-dessous du trait le plus en bas de l'échelle de droite.

Les différences de longueurs d'onde ne sont qu'un des aspects du conditionnement de la lumière. En général, la lumière est incohérente, comparable à la foule d'un stade criant tous sens. Les lasers produisent de la lumière cohérente : la foule chante à l'unisson. La lumière peut se présenter sous forme d'ondes continues, comme les vagues de l'océan qui arrivent sans cesse sur la plage, ou bien sous forme d'impulsions, comme un tsunami. La compression de la lumière en impulsions cohérentes les plus courtes possible (le record est d'environ quatre femtosecondes) crée un outil à la fois puissant et délicat.

Infiniment petit, infiniment grand

Préfixe	Facteur	Préfixe	Facteur
milli (m)	10^{-3}	kilo (k)	10^3
micro (μ)	10^{-6}	méga (M)	10^6
nano (n)	10^{-9}	giga (G)	10^9
pico (p)	10^{-12}	tera (T)	10^{12}
femto (f)	10^{-15}	peta (P)	10^{15}
atto (a)	10^{-18}	exa (E)	10^{18}
zepto (z)	10^{-21}	zetta (Z)	10^{21}
yoto (y)	10^{-24}	yotta (Y)	10^{24}



multiple de la distance entre les miroirs restent dans la cavité. Ces modes sont les équivalents optiques des harmoniques musicaux produits par les cordes de guitare.

Les lasers continus émettent un faisceau constant de lumière dont la couleur correspond à la fréquence de l'un de ces modes. En revanche, une impulsion laser courte contient une large gamme de fréquences, comme un accord constitué de plusieurs notes. Cette caractéristique est décrite par le principe d'incertitude d'Heisenberg : plus une impulsion est courte, plus elle a une large gamme de fréquences. Une impulsion lumineuse de 10 femtosecondes, par exemple, a un spectre qui s'étale sur environ 100 nanomètres. Cet étalement est considérable : décalé dans le visible, il correspondrait au spectre de toutes les couleurs, du violet au cyan, soit un tiers de la lumière visible (le faisceau rouge d'un laser à diode classique a une gamme de fréquences inférieure au nanomètre).

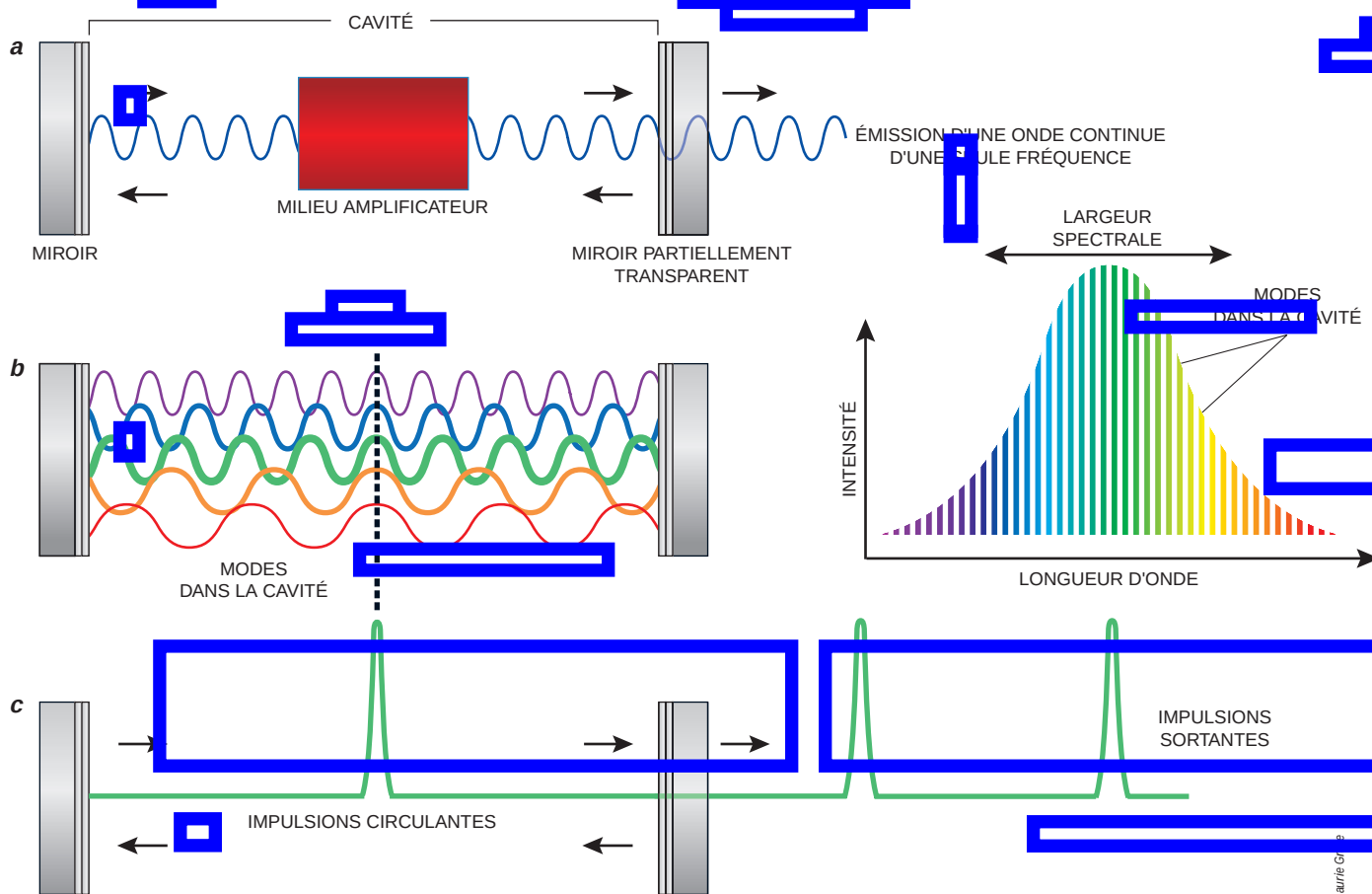
Dans les impulsions ultracourtes, ces milliers de modes sont en phase. Imaginez une rangée de cloches, chacune avec un son différent, qui se balancent en fait d'un clocher. Si les cloches battent au hasard, une longue et terrible cacophonie retentit. Au contraire, si les cloches se balancent de façon synchronisée à intervalles réguliers, elles produisent des accords très puissants, régulièrement espacés. De la même manière, dans le cadre du fonctionnement dit « en verrouillage de mode », toute la lumière de la cavité optique d'un laser est confinée dans une impulsion ultracourte qui circule entre les miroirs (voir la figure 4). Le laser émet une séquence périodique d'impulsions lumineuses à travers le miroir semi-transparent.

L'amplification d'une large gamme de fréquences nécessite évidemment des amplificateurs optiques à large bande. La première génération de lasers à impulsions ultracourtes réglables utilisait pour l'amplification, des jets de colorants organiques dissous dans des

solvants visqueux. Ces lasers à colorants étaient volumineux, difficiles à utiliser et à entretenir. Consommant beaucoup d'énergie électrique, ils ne produisaient qu'un faisceau de quelques milliwatts. De surcroît, les systèmes produisant les impulsions les plus courtes (20 à 30 femtosecondes) n'étaient pas assez fiables pour que l'on envisage des applications. Les dosages de colorant et de solvant étaient délicats, d'autant qu'ils devaient circuler sous pression dans des tuyaux très fins et que leur rapide dégradation imposait leur renouvellement régulier. Les pionniers de ces lasers se rappellent avec émotion leurs débuts... qui ont parfois taché les murs et le plafond de leur laboratoire.

Les cristaux amplificateurs

La découverte de cristaux à large bande, capables d'émettre et d'amplifier des longueurs d'onde comprises entre le visible et l'infrarouge, représenta une avancée décisive. Le plus célèbre de ces matériaux est le saphir

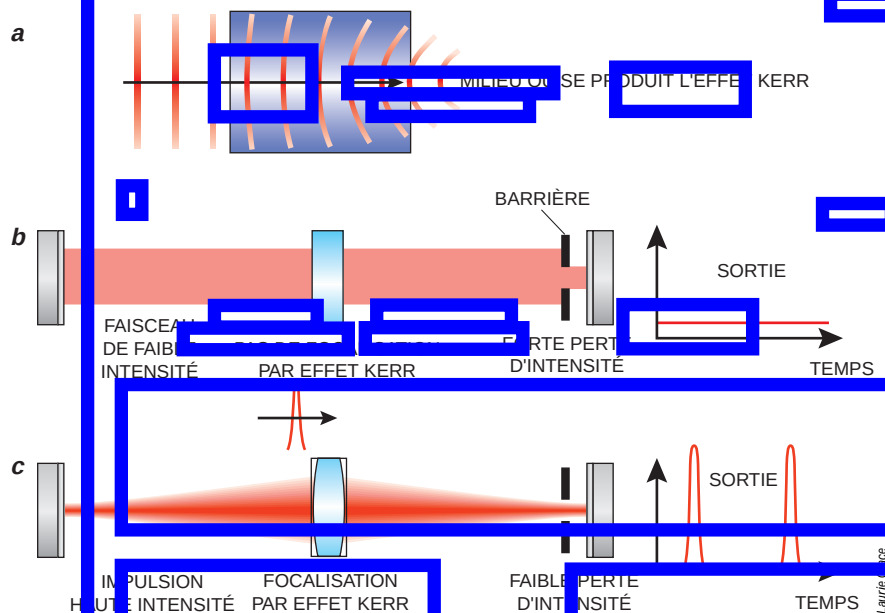


4. LA BOUCLE DE RÉTROACTION ET L'AMPLIFICATION sont les deux principes de base du laser. La lumière se réfléchit, faisant des allers et retours entre les deux miroirs (a) ; elle est amplifiée à chaque passage dans le milieu amplificateur, qui est continuellement excité. Un faisceau continu émerge d'un miroir partiellement trans-

parent situé à une extrémité. Quand de nombreuses longueurs d'onde (ou modes) de lumière se propagent dans la cavité en parfait synchronisme (b), elles s'additionnent pour permettre la création d'une impulsion intense (c). Un peigne d'impulsions émerge alors du miroir partiellement transparent.

dopé au titane, mis au point par Peter Moulton, à l'Institut de technologie du Massachusetts. Ce cristal amplifie des longueurs d'onde comprises entre 700 et 1 100 nanomètres, et il émet des impulsions qui durent moins de quatre femtosecondes.

Avec ces nouveaux matériaux, la difficulté était la production d'impulsions périodiques où des millions de modes sont synchronisés. Un des premiers systèmes de production d'impulsions laser ultracourtes à l'aide de saphirs dopés au titane comprenait deux cavités optiques couplées dont on ajustait la longueur. Alors qu'ils mettaient au point cette technique, en 1989, les physiciens de l'Université de Saint Andrews, en Écosse, ont eu la surprise de constater que des impulsions étaient produites avec une seule cavité. Avec d'autres équipes de recherche, ils ont montré que cet «auto-blocage du mode» était lié à un effet optique, à l'intérieur du cristal de saphir dopé au titane. Ce phénomène, nommé effet Kerr, se produit dans un matériau transparent comme le verre ou le saphir dopé. Le champ électrique du faisceau focalisé sur le matériau produit un changement de son indice de réfraction, phénomène appelé biréfringence. De par ses changements de propriétés, la lumière de haute intensité se propage plus lentement que celle de faible intensité. L'intensité lumineuse du faisceau, maximale en son centre, décroît au fur et à mesure que l'on s'en éloigne, ainsi la lumière à la périphérie du faisceau se déplace plus vite que celle située en son centre. Ce phénomène, équivalent à celui du passage d'un faisceau dans une lentille convexe, s'appelle autofocalisation, ou focalisation résonante de Kerr (voir la figure 5). La focalisation résonante de Kerr favorise la formation de l'impulsion au détriment de l'impulsion lumineuse continue. Les modes s'ajoutent et l'impulsion devient de plus en plus importante. La focalisation résonante de Kerr a produit des impulsions ultracourtes réglables sur une large gamme de fréquences, à des niveaux de puissance encore jamais atteints (voir la figure 6). Cette technique est aujourd'hui le moyen le plus simple et le plus élégant de produire des impulsions de quelques femtosecondes. Les lasers à impulsions ultracourtes actuels ne sont que des perfectionnements du système initial.



5. L'EFFET KERR OPTIQUE entraîne une propagation plus lente de la lumière de forte intensité que de la lumière de faible intensité. Un milieu où se produit l'effet de Kerr agit comme une lentille convergente sur un faisceau de forte intensité (c). Une barrière munie d'un petit trou (b) élimine la partie laser continue, qui ne peut pas gagner suffisamment en intensité pour déclencher l'effet Kerr. Le milieu focalise une impulsion de forte intensité (c), de telle sorte qu'elle traverse le trou et que son amplification est alors maximale.

Nouveaux miroirs et nouvelles pompes

Une des innovations les plus importantes est l'utilisation de miroirs absorbants à base de semi-conducteurs : l'absorption supprime la réflectivité aux faibles intensités lumineuses ; en revanche, quand l'intensité est suffisante, l'absorbant sature et la réflectivité du miroir augmente. En réfléchissant alors les fortes intensités plus efficacement, les miroirs favorisent la production d'impulsions dans la cavité plutôt que la sortie d'ondes continues. Parmi les nombreux modes formés dans la cavité, seuls ceux qui vibrent en phase, c'est-à-dire ceux dont l'intensité est maximale au même instant, ont ensemble une intensité suffisante pour saturer l'absorbant et le traverser (leurs amplitudes s'ajoutent). Les autres modes peuvent être partiellement ou complètement en opposition de phase et se détruire, conduisant à une intensité insuffisante pour traverser l'absorbant. L'avantage d'un tel mécanisme est que de faibles intensités qui vont s'additionner peuvent être à l'origine des impulsions. Ces lasers ont été mis au point, en 1991, par les chercheurs des Laboratoires AT&T Bell.

La découverte de ces mécanismes a simplifié l'utilisation des lasers à impulsions ultracourtes et a assuré un succès immédiat aux lasers femtose-

condes. Toutefois, les sources d'excitation, ou pompes, de ces lasers restaient insatisfaisantes. Tout comme un amplificateur de guitare a besoin de puissance électrique, un laser a besoin d'énergie électrique pour amplifier la lumière de son faisceau. Les cristaux de saphir dopés au titane étaient excités par un laser de pompage à onde continue : le faisceau de pompage excite des atomes du cristal, qui se dés excitent ensuite en émettant des photons qui s'ajoutent au faisceau pulsé. Les premiers lasers à saphir dopé étaient pompés par des lasers à argon, volumineux, qui consommaient une puissance électrique de plusieurs dizaines de kilowatts pour produire 10 à 20 watts d'une lumière bleu-vert. Les coûts étaient prohibitifs.

Aujourd'hui, le pompage du saphir dopé au titane et des matériaux apparentés s'effectue à l'aide de lasers à semi-conducteurs, eux-mêmes pompés par des lasers à diode compacts de forte puissance. Ce système de pompage «en cascade» donne un rendement bien supérieur à celui des lasers à argon : on obtient jusqu'à dix watts de lumière verte dans un faisceau de meilleure qualité, à l'aide d'un système dix fois plus petit. En outre, le coût de fonctionnement est bien inférieur, et le système peut être branché sur une simple prise de courant.

Le temps, l'espace et l'énergie

Que de chemin parcouru, de l'énoncé par Einstein en 1917 du principe de l'émission stimulée à la mise en évidence expérimentale de l'effet laser, en 1960, et enfin à la génération des lasers à impulsions femtosecondes. En Europe, les premières impulsions optiques d'une centaine de femtosecondes ont été obtenues à la fin des années 1970 par une équipe de physiciens du Laboratoire d'optique appliquée de l'Ecole nationale supérieure des techniques avancées et de l'École polytechnique. Aujourd'hui, les efforts se concentrent sur la caractérisation d'impulsions de plus en plus courtes, entre 10 et 30 femtosecondes, et sur l'optimisation des phénomènes d'amplification. Les techniques de plus en plus perfectionnées fournissent une densité d'énergie par impulsion considérable, ciblée sur un échantillon solide, gazeux ou liquide. La matière est ainsi sondée ou manipulée dans tous ses états.

Les effets engendrés par les impulsions femtosecondes dépendent de la façon dont l'énergie des photons est déposée dans le milieu. Une impulsion femtoseconde de dix gigawatts émise dans le rouge contient jusqu'à trois millions de milliards (3×10^{15}) de grains de lumière. Dès lors, pourquoi une impulsion femtoseconde focalisée sur un millimètre carré de peau et dont la puissance instantanée serait de quelques gigawatts (un surgénératuer nucléaire délivre une puissance moyenne 1 000 fois inférieure) ne provoque-t-elle rien de plus qu'un petit picotement ? Pourquoi un matériau inerte soumis à une impulsion de quelques terawatts (10^{12} watts) se transforme-t-il en plasma (gaz dense constitué d'électrons et d'ions très localisé dans l'espace) avant d'émettre des rayons X ?

À la vitesse de la lumière, un photon traverse une molécule de 0,3 nanomètre de diamètre en un centième de femtoseconde. Si la molécule vibre à la fréquence de 10^{14} cycles par seconde, le photon n'explore qu'un millième de la vibration. En revanche, les innombrables photons d'une impulsion d'une centaine de femtosecondes peuvent entrer en résonance avec cette molécule et en modifier le comportement. Les physiciens de l'École polytechnique et de l'Université Paris 6 ont ainsi étudié, dans l'eau, l'excitation des liaisons hydrogène par des impulsions infrarouge femtosecondes. Ils ont suivi, en temps réel, le mouvement de ces liaisons, qui confèrent à l'eau ses propriétés physiques étonnantes.

La durée de l'impulsion est un paramètre essentiel, au même titre que l'énergie ou la densité de photons. Des rayons X mous aux rayons T, l'énergie s'étend du milliélectronvolt au kiloelectronvolt, gamme d'énergie qui autorise de nombreuses expériences : une fraction de milliélectronvolt modifie la rotation d'une molécule, quelques électronvolts arrachent un électron périphérique d'un atome, ou cassent une liaison interatomique dans une molécule, quelques kiloelectronvolts perturbent l'arrangement des électrons des couches profondes d'un atome. En dépouillant progressivement les atomes de leur cortège électronique, les physiciens découvrent de nouveaux comportements de la matière. Ainsi, grâce à une impulsion laser de 120 femtosecondes focalisée sur du

silicium, et dont l'intensité était de l'ordre de 10^{16} watts par centimètre carré, des chercheurs du laboratoire pour l'utilisation des lasers intenses de l'École polytechnique ont réalisé une spectroscopie détaillée d'un plasma de silicium. Ils ont observé de nouveaux comportements électroniques du plasma dont la dynamique est comparable à celle de l'impulsion laser. Ils ont également produit

par l'irradiation d'impulsions femtosecondes sur un gaz, des boffes ultrabrèves d'électrons relativistes et très énergétiques. Avec d'autre méthode, par exemple la source européenne synchrotron de Grenoble, on obtient des paquets d'électrons, également très énergétiques, mais d'une durée supérieure (de 150 à 190 picosecondes).

Un siècle après la découverte des rayons X et de la radioactivité, on maîtrise les énergies considérables libérées par le noyau atomique pour des applications médicales ou industrielles. Une des applications les plus intéressantes de l'interaction des lasers femtosecondes intenses avec la matière est la génération d'impulsions femtosecondes dans le spectre X. Des chercheurs du Laboratoire d'optique appliquée ont ainsi obtenu des images résolues en temps montrant des mouvements d'atomes ou de molécules. Les impulsions femtosecondes de plus en plus courtes et de plus en plus énergétiques complètent la panoplie des outils disponibles. Ces impulsions laser sont des micro-outils dont les caractéristiques sont conditionnées par l'énergie des photons et par leur densité spatio-temporelle. Chaque type d'impulsion femtoseconde contenant un paquet d'ondes est adapté à l'étude ou à la manipulation d'objets tels que des atomes, des molécules, des puces de semi-conducteurs ou le cœur des cellules vivantes.

Des impulsions femtosecondes infrarouges sont utilisées pour le développement d'un microscope à balayage laser par des physiciens du Centre de physique moléculaire de Bordeaux. Dans le microscope à balayage classique, un faisceau d'électrons « balaie » la surface de l'échantillon ; ici, le faisceau d'électrons est remplacé par un faisceau laser ultracourt. Le faisceau est focalisé, à différentes profondeurs, selon des plans de coupe différents de l'objet observé. Grâce à des logiciels de reconstruction, on obtient des images tridimensionnelles de l'objet. L'équipe de Bordeaux a ainsi reconstitué, à partir de plans de coupe séparés de 0,2 micromètre, des images tridimensionnelles, de bonne qualité d'organismes unicellulaires ou de cellules nerveuses embryonnaires. Cette méthode a l'intérêt notable de préserver l'échantillon : c'est une méthode intéressante pour procéder à des analyses non destructives.

La femtochimie et la femtobiologie suivent, en temps réel, les transformations des atomes et des molécules lors de réactions. Les sources femtosecondes offrent un accès à une visualisation temporelle des électrons, des atomes et des molécules. Dans une relation espace-temps-énergie, les lasers femtosecondes « visualisent » la transformation de la matière en temps réel et à l'échelle de l'infiniment petit.

Yann GAUDUEL, INSERM, Laboratoire d'optique appliquée ENSTA-École polytechnique

Pompage direct

Aujourd'hui, les physiciens cherchent encore des perfectionnements. Notamment, ils voudraient éviter l'emploi du laser à semi-conducteurs intermédiaire et pomper directement le laser à impulsions ultracourtes avec un laser à diode. Les systèmes seraient alors plus petits et plus efficaces. Toutefois, des difficultés subsistent : les lasers à diode, initialement utilisés pour le pompage d'autres cristaux, produisent des longueurs d'onde et des puissances inadaptées au pompage direct du saphir dopé au titane. Pourrait-on remplacer le saphir dopé par d'autres cristaux ? On étudie aujourd'hui les « colquirites », qui doivent leur nom aux mines d'étain de Colquibambesi, en Bolivie, où elles ont été découvertes. Les colquirites dopées au chrome oscillent sur une gamme de longueurs d'onde comparable à celle du saphir dopé et elles peuvent être pompées directement par les classiques lasers à diode rouges. Des colquirites d'excellente qualité optique, synthétisées à Livermore, ont permis la mise au point de cavités à impulsions ultracourtes plus efficaces.

Dans notre laboratoire, nous avons notablement amélioré l'efficacité des lasers à impulsions ultracourtes pompés directement par diode en réduisant la puissance de pompage nécessaire pour le fonctionnement du laser en verrouillage de mode et en pompant les cristaux à l'aide de diodes de faible puissance, semblables à celles des lecteurs de DVD. En optimisant les optiques de la cavité laser, nous avons minimisé les pertes. Nous avons ainsi fait fonctionner des systèmes de moins de 20 centimètres de longueur à l'aide de piles LR6 standards (voir la figure 6). Nos lasers produisent des impulsions de 100 femtosecondes avec des puissances moyennes de 10 milliwatts. Leur consommation électrique (un watt) est 100 fois inférieure à celle des lasers à colorants et à verrouillage de mode de puissance équivalente.

De petits lasers à fibres optiques émettent également des impulsions de quelques femtosecondes non ajustables ; leur milieu amplificateur est une courte fibre optique dopée. Les lasers de ce type sont de bonnes sources compactes d'impulsions femtosecondes et sont adaptés à certaines applications industrielles émergentes. Au début de l'an 2000, des ingénieurs de la Société IMRA America ont annoncé

des puissances moyennes de sortie supérieures à dix watts, pour des lasers à fibres à impulsions femtosecondes.

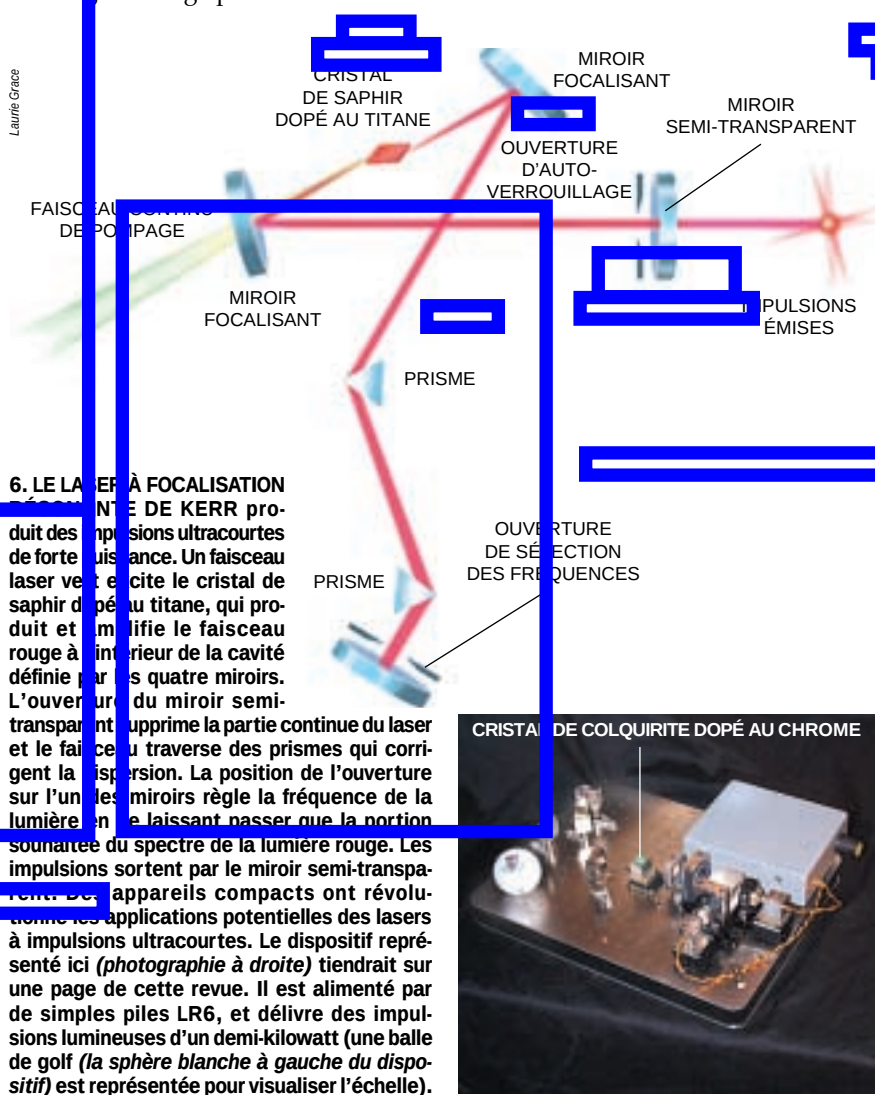
Le record *plus ultra* en matière de compacité sera l'arrivée de lasers à diodes qui produisent directement des impulsions de quelques femtosecondes. Cependant, la génération actuelle de lasers à impulsions ultracourtes à semi-conducteurs a déjà permis à cette catégorie de lasers de sortir des laboratoires spécialisés pour devenir des outils aux applications concrètes.

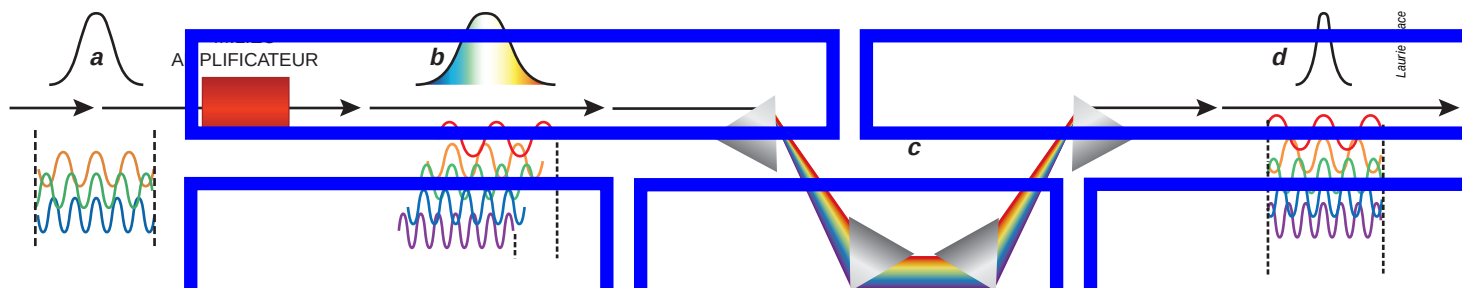
La puissance suprême

Aux hautes énergies, l'intensité maximale des impulsions ultracourtes endommage l'optique classique des lasers en modifiant les propriétés du milieu qu'elle traverse, en le cassant par émission de décharges électriques dans la cavité optique (de ce fait, on espère utiliser des impulsions ultracourtes de forte puissance pour provoquer la décharge de nuages d'orage par ionisation d'un

chemin conducteur dans l'air). Aussi les intensités des lasers focalisés sont-elles longtemps restées limitées à environ 10^{15} watts par centimètre carré. En 1985, à l'Université de Rochester, Gérard Mourou et Donna Strickland ont mis au point une technique d'étirement de l'impulsion par dérive de fréquence (ou CPA en anglais pour *Chirped-Pulse Amplification*), qui évite les intensités excessives par allongement des impulsions. Ainsi allongées, les impulsions sont amplifiées sans risque pour le milieu amplificateur et pour les autres éléments d'optique ; puis les impulsions étirées et amplifiées sont recompressées grâce à des systèmes nommés grilles de diffraction dans le vide.

Les premiers modèles comportaient plusieurs étages de verrouillage, d'amplification puis de compression ; aujourd'hui, les systèmes étaient encore plus petits, nous disposons de lasers plus petits, de quelques térawatts (10^{12} watts) avec des faisceaux de sortie dont l'intensité





7. LES IMPULSIONS INTENSES ÉVOLUENT à chaque passage dans la cavité d'un laser à impulsions ultracourtes sous l'influence de plusieurs effets. Quand une impulsion (a) traverse un milieu amplificateur, des effets non linéaires ajoutent des modes de fréquence supérieure aussi bien qu'inférieure, et la dispersion étale les différentes longueurs d'onde de la lumière (b). Un système de prismes (c) peut réaligner tous les modes, produisant une impulsion plus courte.

est considérable (10^{18} watts par centimètre carré). À l'aide de ces lasers, on obtient des plasmas qui peuvent à leur tour produire des impulsions ultracourtes de rayons X mous et cohérents, utiles pour la microscopie et la lithographie. À l'intérieur de ces plasmas, les champs électriques sont très élevés et les électrons ont un comportement relativiste : des explorations de physique fondamentale sont envisageables.

De gros appareillages ont délivré des faisceaux qui ont déposé une puissance lumineuse de 2×10^{21} watts par centimètre carré sur des cibles, avec une puissance de quelques pétawatts (10^{15} watts), 1 000 fois supérieure à la puissance dissipée par un éclair. Par des effets de pression de radiation, notamment, ces lasers accélèrent la matière 10^{22} fois plus que la gravitation terrestre, ce qui correspond à une accélération bien supérieure à celle que des corps subiraient à proximité d'un trou noir de la masse du Soleil ! Ils procurent des températures et des pressions comparables à celles de l'intérieur d'une étoile, laissant envisager des études de dynamique stellaire. En travaillant avec des intensités focalisées de 10^{18} watts par centimètre carré, les physiciens espèrent créer des paires de particules et d'antiparticules, et ils prévoient d'étudier le domaine de l'électrodynamique quantique, qui décrit le comportement quantique de particules chargées et du champ électromagnétique quantifié.

D'autre part, les physiciens tirent parti de la brièveté des impulsions. Certains ont utilisé la large bande spectrale des impulsions de moins de dix femtosecondes pour la mesure précise des fréquences optiques. Les modes d'un laser en verrouillage de mode sont espacés uniformément, comme les dents d'un peigne (voir la figure 4). Ce

«peigne» peut servir de référence pour la mesure de fréquences difficiles à évaluer directement. L'utilisation de cette technique permet de mesurer des énergies de transition atomique optique avec une précision mille fois meilleure que tout ce qui se faisait jusqu'à maintenant. Certaines constantes de la physique peuvent ainsi être déterminées : une de ces expériences a déterminé la «constante de structure fine» qui caractérise l'intensité de la force électromagnétique avec une précision de 120 pour un milliard.

Le record de brièveté des impulsions est aujourd'hui inférieur à cinq femtosecondes : il est détenu par des lasers à saphir dopé au titane, à l'École polytechnique de Zurich et à l'Institut de technologie du Massachusetts. Jusqu'à présent, les impulsions ultracourtes étaient décrites par une théorie où l'enveloppe des impulsions variait lentement comparée aux oscillations de la lumière, mais l'enveloppe des impulsions de cinq femtosecondes ne contient que deux périodes, on atteint donc les limites de la théorie. Bientôt, la durée de l'impulsion sera inférieure à celle d'une période.

Les lasers à impulsions ultracourtes, avec une gamme de puissance qui s'étale sur plus de 17 ordres de magnitude, sont présents dans de nombreuses disciplines scientifiques et techniques, de l'imagerie en profondeur des tissus vivants à l'élucidation de questions fondamentales sur l'Univers. Il y a quelques années, les impulsions femtosecondes ont permis l'observation de mécanismes moléculaires et de mouvements d'atomes ; de même, les impulsions de quelques attosecondes (une attoseconde vaut un millième de femtoseconde) montreront les mouvements des électrons, des protons et des neutrons.

Les nouveaux lasers gagneront progressivement tous les domaines. Après la médecine avec l'imagerie biologique et la chirurgie, ils devraient conduire à des découvertes pharmaceutiques importantes, révélant par exemple les mécanismes de certains médicaments. En micro-électronique, ils assureront la commande cohérente du courant électrique dans les composants à semi-conducteurs et dans les réseaux de communications optiques numériques.

Quand le laser fut inventé, certains l'ont qualifié de «solution à la recherche d'un problème». Aujourd'hui, les lasers sont omniprésents dans les lecteurs de codes à barres des supermarchés, dans les lecteurs de disques compacts, etc. Depuis leur apparition au milieu des années 1960, les lasers à impulsions ultracourtes s'imposent également, du laboratoire à la vie courante. Ils sont aujourd'hui la meilleure solution à nombre de problèmes concrets et participeront demain à l'amélioration de la qualité de vie.

John-Mark HOPKINS est physicien à l'Université de Saint Andrews, en Écosse. Wilson SIBBETT est professeur dans la même université et membre du Conseil d'ingénierie et de sciences physiques du Royaume-Uni.

Gérard A. MOUROU, Christopher P.J. BARTY et Michael D. PERRY, *Ultra-high-Intensity Lasers : Physics of the Extreme on a Tablet in Physics Today*, vol. 51, n° 1, pp. 22-28, janvier 1998.

Henry KAPTEYN et Margaret MURNANE, *Ultrashort Light Pulses : Life in the Fast Lane*, in *Physics World*, vol. 12, n° 1, pp. 31-35, janvier 1999.

Don ARNONE, Craig CIESLA et Michael PEPPER, *Terahertz Imaging Comes into View*, in *Physics World*, vol. 13, n° 4, pp. 35-40, avril 2000.



Mathématiques et philosophie

JEAN-PAUL DELAHAYE

Que les leçons philosophiques peut-on tirer des mathématiques? Les résultats difficiles à démontrer mathématiquement sont-ils utiles?

Les rapports entre science et philosophie sont ambigus. La philosophie adopte souvent une attitude normalisatrice envers la science, fixant les critères de ce qui est «scientifique»: le philosophe se propose d'aider la science en lui donnant des conseils.

La philosophie a pu rendre d'authentiques services à la science, et de grands savants comme Albert Einstein, Kurt Gödel, Max Planck ou Jacques Monod ont accordé de l'importance à la pensée philosophique qui les a aidés à formuler et à défendre les audacieuses théories qu'ils proposaient.

Hélas, trop souvent, les conceptions avancées par la philosophie ne tiennent compte de la science ont freiné ou retardé les nouveautés plutôt qu'elles ne les ont favorisées. Lors des débuts de la théorie de l'évolution, de la théorie de la relativité, de la mécanique quantique et de la logique mathématique, la philoso-

phie a constitué un obstacle plus qu'elle n'a été une aide à la découverte.

Cette affirmation provocante nécessiterait de longues analyses pour être pleinement justifiée. Évoquons simplement Henri Poincaré, qui, au nom d'une certaine intuition du temps s'opposait à la théorie de la relativité et se mit dans l'impossibilité de la comprendre. Pensons aussi aux conceptions *a priori* de ce que sont les mathématiques: Henri Poincaré et d'autres refusèrent les progrès de la logique mathématique – en particulier, de la théorie des ensembles, aujourd'hui universellement adoptée et dont aucun mathématicien ne peut se passer.

Pensons encore aux difficultés rencontrées par la théorie de l'évolution au XIX^e siècle, au point de vue, par exemple, de la transmission des caractères acquis (dans l'œuvre de Lyssenko, c'est même une science prétendument fondée sur un matérialisme dialectique qu'on condamna la

génétique moderne). Aujourd'hui encore, aux États-Unis, des idées philosophiques fausses sur la vie interdisent à la pensée scientifique de faire valoir ses vues et entravent l'innovation.

Rapportons aussi que la conception de l'espace que défendait Kant est incompatible avec celle aujourd'hui universellement admise en physique, où l'on considère que la nature de la géométrie (euclidienne, hyperbolique ou sphérique) doit se déterminer par observation et l'expérimentation, et non par la simple analyse *a priori* des données immédiates de la raison.

La vérité est que la science enrichit la philosophie et la renouvelle, plus que l'inverse.

Ce constat que certains philosophes acceptent peu (voir, par exemple, *La philosophie d'aujourd'hui* de Gaston Bachelard), fait apparaître la philosophie comme une entrave à la science, en même temps qu'elle rend inacceptables les prétentions de certains philosophes à dire aux chercheurs comment ils doivent travailler.

DE LA PHILOSOPHIE AUX MATHÉMATIQUES

En mathématiques, en logique et en théorie de la calculabilité, des problèmes considérés comme philosophiques deviennent, à la suite d'un progrès scientifique, des problèmes techniques – en l'occurrence des problèmes mathématiques – dont les solutions (parfois inattendues) sont riches en contenus philosophiques; l'impression qui se dégage est que la science fait de la philosophie et de la logique.

Parmi les avancées incontestables des mathématiques sur des sujets classés autrefois «purement philosophiques», il y a, bien sûr, celles concernant l'infini. Envisageant celui-ci comme un domaine réservé des philosophes ou même des théologiens, les mathématiciens ont réussi en plusieurs étapes (géométrie, théorie des nombres, calcul infinitésimal, théorie

1. Science et philosophie

La science fait de la philosophie en remettant en cause les conceptions du sens commun. La philosophie ne peut permettre de telles remises en cause, car son support concret

est limité, ce qui la rend incapable de poursuivre son interrogation audace que la science est, et qui la rend, contrairement à ce qu'on croit, capable de franchir les obstacles qu'elle rencontre. Ces audaces de la science constituent des avancées dont les conséquences philosophiques sont parfois importantes.

Voici quatre exemples de champs philosophiques qui ont été bouleversés par des progrès scientifiques :

- la conception de la vie et de l'homme, par la théorie Darwinienne;
- la conception du temps et de l'espace, par la théorie de la relativité;
- la conception du déterminisme et des objets physiques, par la mécanique quantique;

- la conception du raisonnement, de ses propriétés et de ses limitations, la conception du cerveau et de l'intelligence, par les avancées de la logique mathématique.

Bachelard avait perçu cette importance philosophique de la science: «Les progrès de la pensée scientifique contemporaine ont déterminé des transformations dans les principes mêmes de la connaissance.»

cantorien des ensembles, théorie des grands cardinaux) à nourrir la notion d'infini et à constituer une série de domaines mathématiques nouveaux où le concept d'infini dévoile une richesse que la philosophie n'avait su déceler.

Au début du xx^e siècle, la logique mathématique a précisé la notion de preuve mathématique et a permis à la pensée philosophique sur les mathématiques de s'appuyer sur une multitude de notions précises (systèmes formels, calcul propositionnel et calcul de prédicats, modèles, etc.) et de résultats précis (théorèmes de complétude, théorèmes de limitation, etc.). La philosophie des mathématiques en a été transformée, et si le débat y est devenu plus technique, la réflexion des philosophes qui acceptent ces avancées mathématiques n'en a été que plus pertinente et plus incisive (voir, par exemple, *The Philosophy of Mathematics*, W. H. Hart ed., Oxford University Press, 1996).

L'éclaircissement de la notion de calcul mécanique vers 1936 (par K. Gödel, A. Church, A. Turing) a posé le socle d'une discipline nouvelle (appelée théorie de la récursivité ou théorie de la calculabilité) qui a récemment conduit à de miraculeuses applications dans le domaine de la philosophie des sciences. La théorie de la complexité de Kolmogorov, un développement récent de la théorie de la calculabilité, est, de ce point de vue, d'une remarquable efficacité. On lui doit nombre de propositions dont la portée philosophique est incontestable et qui, dans certains cas, ont résolu des problèmes posés par les philosophes eux-mêmes.

Avant de donner quelques détails sur l'une de ces théories, la théorie de la complexité, «est-ce que les résultats mathématiques difficiles sont pratiquement utiles?», question incontestablement philosophique et intéressante, nous allons survoler rapidement les idées de base de la théorie de la complexité de Kolmogorov.

LE COMPLEXE, LE SIMPLE, LE HASARD

Formulée par R. Solomonoff, A. Kolmogorov et G. Chaitin vers 1965, la théorie de la complexité algorithmique, dénommée aujourd'hui complexité de Kolmogorov, a pour ambition de définir ce qu'est un objet simple, et propose une mesure absolue de la complexité.

L'idée est de donner un sens à l'évidence : la suite 0101010101010101 est plus simple que la suite 10000100110101110110. Pourquoi donc? Parce que la première suite peut se définir brièvement, c'est «dix fois 01», alors que la seconde n'est pas exprimable avec la même concision : pour la décrire, on ne

2. L'utilité des résultats mathématiques difficiles

Dans la grande majorité des cas, les résultats mathématiques possédant des preuves compliquées sont d'une faible utilité pratique. Les résultats que l'on n'a pas encore prouvés aujourd'hui (qu'on dénomme des conjectures) n'ont sans doute pas de preuves simples. Malgré la fascination qu'exercent ces énigmes sur les mathématiciens et les amateurs, on ne peut attendre que peu de retombées concrètes de leurs solutions. Parmi les conjectures les plus étudiées dont l'inutilité pratique est presque certaine, indiquons :

- la conjecture de Goldbach : tout nombre pair est somme de deux nombres premiers (un million de dollars à gagner) : $24 = 11 + 13$;
- la conjecture des nombres premiers jumeaux : il existe une infinité de paires de nombres premiers espacés de deux unités (comme 11 et 13 ou 17 et 19) ;
- la conjecture des nombres parfaits impairs : il n'existe aucun nombre parfait impair (un nombre parfait est un nombre comme $6 = 1 + 2 + 3$ ou $28 = 1 + 2 + 4 + 7 + 14$, qui est égal à la somme de ses diviseurs autres que lui-même). C'est sans doute la plus vieille de toutes les conjectures mathématiques ;
- la conjecture de l'équipartition des chiffres de π : la fréquence de chacun des chiffres dans le développement décimal de π est 1/10.

Plus étrange, il semble que la démonstration de la grande conjecture de l'informatique « $P = NP$?» (qui concerne la classification des problèmes en fonction de la difficulté de leur résolution et qui a été dotée d'un prix d'un million de dollars cette année) n'aurait pas de conséquences concrètes. Si on démontre que P est différent de NP (ce que les spécialistes pensent en général), aucun algorithme nouveau n'en résultera. Si l'on montre que $P = NP$, il apparaît probable que cela se fera d'une telle façon que rien de concrètement utile n'en sortira.

pourra que faire mieux que d'en énumérer les éléments : «1, puis 0, puis 0, puis 0, etc.», ce qui est beaucoup plus long.

La théorie de la calculabilité, pour définir plus précisément la simplicité en termes de programmes, utilise les machines de Turing universelles, qui sont des sortes d'ordinateurs abstraits dont chaque ordinateur réel est une réalisation particulière, et un détail précis : l'ordinateur réel ne possède qu'une mémoire de travail finie, alors que la machine de Turing est munie d'un ruban de mémoire indéfiniment extensible.

L'idée est simple : choisir une machine de Turing universelle, U , et de dire qu'un objet O représenté sous forme d'une suite de 0 et de 1 (ce qui est toujours possible comme l'informatique le démontre chaque jour), est simple s'il existe un programme court P pour la machine U qui engendre l'objet O . Un objet sera complexe si, au contraire, aucun programme court pour U ne donne O .

Cette définition ne dépend pratiquement pas de la machine U choisie, ce qui est rassurant et garantit l'invariance de la notion. La complexité de Kolmogorov de O , $K(O)$, est la taille (mesurée en nombre de chiffres binaires) du plus petit programme qui permet à la machine U d'engendrer O .

Cette idée très naturelle : «est simple ce qui s'exprime brièvement» a en particulier permis de définir la notion de suite aléatoire, problème que la théorie des probabilités avait contourné en utilisant une notion collective de probabilité qui interdisait qu'on puisse dire, d'une suite don-

née, quelle est aléatoire. La définition individuelle d'objet aléatoire avait été longuement recherchée par les philosophes, et par exemple Karl Popper, dans son ouvrage fondamental de 1934, *La logique de la découverte scientifique* (que Jacques Monod admirait beaucoup) y consacre plusieurs pages.

D'après la théorie de la complexité de Kolmogorov, une suite infinie de chiffres binaires est aléatoire si, par définition, à chaque fois qu'on en prend n chiffres, on obtient un objet dont la complexité de Kolmogorov est approximativement n . Dit autrement encore, il est impossible de comprimer les données contenues dans $(a_0, a_1, a_2, \dots)$: être aléatoire, c'est être incompressible.

Des liens ont alors été établis entre l'incompressibilité et l'imprévisibilité : une suite aléatoire est, par nature, imprévisible, et donc incompressible. Ces définitions ont donné un sens mathématique à tout un ensemble d'idées philosophiques vagues, et ce succès remarquable est apparu comme un petit miracle après les nombreuses tentatives infructueuses passées.

La théorie de la complexité de Kolmogorov a attiré une remarquable médiatité dans les années 1990, et les lecteurs intéressés trouveront un tour d'horizon complet de cette merveilleuse théorie dans le livre profond et synthétique de Ming Li et Paul Vitányi cité en bibliographie, et une présentation plus concise se trouve dans mon ouvrage *Information, complexité et hasard*, également cité en bibliographie. Parmi les progrès philosophiques que la

théorie de la complexité de Kolmogorov a apporté, on remarque ceux-ci.

1) Une nouvelle compréhension du sens des théorèmes d'incomplétude de Gödel : les propositions indécidables (ni elles, ni leur négation ne sont démontrables) dans une théorie donnée doivent leur existence au fait que le contenu en information de la théorie (qu'on mesure par la complexité de Kolmogorov des axiomes de la théorie) limite son pouvoir de preuve. L'incomplétude est un problème de contenu en information des théories, ce qui n'apparaissait pas clairement dans les travaux antérieurs des logiciens.

2) Une conception objective de l'entropie physique : l'entropie physique, lorsqu'on la définit traditionnellement, comporte une composante subjective (elle semble dépendre de l'observateur, ce qui est gênant) que évite la définition proposée par le physicien W. Zurek, définition fondée sur la prise en compte de la complexité de Kolmogorov des informations mémorisées par l'observateur.

3) Le nombre *Oméga* de Chaitin est le seul nombre réel qu'on peut définir précisément et dont on peut montrer qu'il est équiréparti dans toutes les bases (en base 10, par exemple, son écriture comporte autant de 0 que de 1, que de 2, etc.), ce qui éclaire à nouveau les phénomènes d'incomplétude et d'indécidabilité algorithmique et rassure sur la cohérence de la théorie de la mesure (laquelle indique qu'il existe des nombres équirépartis, mais ne propose aucune méthode pour en définir).

4) Une nouvelle compréhension du problème de l'induction scientifique (que le philosophe Karl Popper considérait comme impossible et qui, d'ailleurs, l'est toujours, de K. Popper), de la règle de Bayes et du principe du rasoir d'Ockam (qui indique

que, dans une explication, il faut éviter autant qu'on le peut de multiplier les entités et les concepts) et des mises en relation précises de toutes ces notions.

5) La proposition par Charles Bennett d'une distinction fondamentale entre complexité désordonnée et complexité organisée (utilisant la complexité de Kolmogorov), distinction qui semble essentielle pour expliquer la complexification des structures de l'Univers au cours du temps qui passe (évolution stellaire, évolution du vivant, etc.).

LA QUESTION DE L'UTILITÉ DES RÉSULTATS DIFFICILES

L'énumération pourrait se poursuivre, mais nous allons nous arrêter sur une des applications toute récente de la théorie de la complexité à l'étude d'un problème d'épistémologie des sciences : est-ce que les résultats mathématiquement difficiles sont ceux qu'on a le plus de chances d'être utiles dans les applications ou, au contraire, un résultat mathématiquement difficile est-il toujours condamné à avoir peu d'incidences concrètes ?

Bien sûr, tous les mathématiciens s'accrochent et tentent de défendre l'idée que démontrer des résultats difficiles est important et très utile. Bien sûr aussi, personne ne conteste que les théories abstraites et les résultats complexes ou très longs à démontrer sont souvent d'une grande valeur et constituent des accomplissements humains remarquables. De même, personne ne met en doute que la compréhension profonde du monde nécessite des concepts et des théories qui peuvent être d'une grande abstraction et d'une grande difficulté.

À un moment (ce qui est sans doute exagéré) que seul Einstein comprenait la

théorie de la relativité, dont l'utilité est incontestable.

Et alors, il s'agit ici c'est seulement la question de savoir si les résultats difficiles parce qu'ils sont difficiles à démontrer sont utiles, sans prendre en compte la difficulté de la compréhension, qui est un autre problème (chaque sait d'ailleurs que certains résultats de démonstration simple sont difficiles à comprendre).

Il faut reconnaître que les résultats incontestablement difficiles n'ont que peu d'applications concrètes. Les résultats les plus utiles des mathématiques semblent être tous assez simples à prouver. Citons quelques exemples confirmant cette remarque, que le mathématicien J. Hartmanis a récemment développée.

FACILE ET TRÈS UTILE, OU DIFFICILE MAIS INUTILE ?

Rares sont ceux qui prétendent qu'on trouvera une application concrète au théorème de Fermat (on n'existe pas de solution entière de l'équation $x^n + y^n = z^n$, pour n supérieur à 2), mais on chercha la démonstration pendant des siècles avant de la trouver, il y a six ans, à l'aide d'un raisonnement d'une extrême difficulté, à la fois long et faisant usage de concepts d'une grande abstraction. Aujourd'hui le théorème de Fermat est l'exemple type d'un résultat mathématiquement difficile, mais sans utilité concrète.

Parmi les résultats difficiles des mathématiques, il y a le fameux théorème de raréfaction des nombres premiers, qui indique que la proportion de nombres premiers parmi les entiers proches de n est environ $1/n$. Longtemps recherchée, la preuve de ce théorème fut découverte en 1896 par Jacques Hadamard et Charles de La Vallée-Poussin. Cette preuve, longue et complexe, a été très soigneusement étudiée. Des variantes évitant le recours aux concepts abstraits, dont il semblait étrange que la preuve fasse usage, ont été mises au point, mais ces variantes sont moins utiles, dans leurs concepts, sont toujours longues et complexes. On peut donc dire, sans risque de se tromper, que ce résultat fait bien partie de la classe des résultats complexes à démontrer. Pourtant son utilité pratique semble réduite, et je crois qu'aucun commerçant, assureur, chimiste ou physicien ne trouvera jamais la moindre utilité à ce théorème, même s'il aide à la conception de certains algorithmes, ni n'est jamais utile pour garantir le bon fonctionnement.

Les résultats d'arithmétique qui sont les plus utiles, par exemple en cryptographie, sont tous d'une relative simplicité et, par exemple, les résultats qui justifient l'algorithme RSA (voir *La cryptographie RSA vingt ans après*, Pour la

3. Les théorèmes mathématiques utiles sont faciles à démontrer

On a constaté que les résultats mathématiques les plus utiles sont souvent relativement simples à démontrer, ce qui n'empêche pas qu'ils peuvent passer inaperçus devant des difficultés. Le cas le plus étonnant est celui de la règle de calcul de Karatsuba, qui permet de réaliser des multiplications entre grands nombres plus rapidement que la méthode classique. Cette règle s'écrit :

$$(a + b \times 10^k)(c + d \times 10^k) = ac - [(a - b)(c - d) - ac - bd] \times 10^k + bd \times 10^{2k}$$

On peut la démontrer en classe de troisième. Cette identité, lorsqu'on multiplie deux entiers de longueur $2k$, ramène le calcul à 3 multiplications entre entiers de taille k , ce qui fait gagner 25 pour cent du temps de calcul par rapport à l'identité classique à la base de la méthode usuelle :

$$(a + b \times 10^k)(c + d \times 10^k) = ac + (ad + bc) \times 10^k + bd \times 10^{2k}$$

Les additions sont plus nombreuses dans la formule de Karatsuba, mais elles sont d'un coût négligeable à côté des multiplications.

Directement sous cette forme ou sous des formes un peu plus élaborées (mais toujours faciles à démontrer), l'identité de Karatsuba est l'un des résultats les plus concrètement utiles des mathématiques, car de nombreux logiciels de calcul en font usage. Or elle est à la portée d'un élève de troisième...

Science, février 2000) sont élémentaires, comparés à ceux nécessaires au théorème de réciprocité des nombres premiers. Revenons en passant un paradoxe : les résultats pour le RSA, bien plus simples et utiles que le théorème de réciprocité des nombres premiers, ont pourtant été formulés et démontrés bien après : ce qui était difficile, ce n'était pas démontrer le résultat clef du RSA, mais de le découvrir et d'en comprendre l'intérêt.

La règle de calcul de Karatsuba est exactement dans la même situation, mais d'une façon plus nette encore. Elle permet des multiplications rapides de grands nombres, et c'est l'un des résultats les plus utiles de toutes les mathématiques (à chaque seconde qui passe, elle est utilisée par des milliers d'ordinateurs). Pourtant cette règle est véritablement facile à prouver.

Inversément, de nombreux autres grands résultats mathématiques (classification des groupes simples, démonstration de la transcendance de π , théorèmes de Gödel, etc.), et dont les démonstrations sont longues, semblent d'une utilité pratique nulle.

De même, les grandes conjectures indémontrées (dont on peut penser qu'elles sont indémontrées parce qu'elles n'ont pas de démonstrations simples) ne paraissent pas avoir d'applications potentielles. Si l'on réussissait à établir qu'il y a une infinité de nombres premiers jumeaux (conjecture des nombres premiers jumeaux), ou que tout nombre pair est somme de deux nombres premiers (conjecture de Goldbach), ou que l'algorithme de Syracuse conduit toujours à 1 (conjecture de Syracuse), rien de concret n'en résulterait, aucune *start-up* n'en sortirait et, abstraction faite des récompenses offertes par des mécènes sensibles aux mystères des mathématiques et des positions universitaires auxquelles pourront prétendre les génies qui résoudront ces problèmes, ce sont de purs défis sans conséquences économiques ou matérielles attendues.

Les réflexions menées sur la célèbre conjecture de P vs NP (qui pourtant porte sur une question d'efficacité d'algorithmes) ont conduit le mathématicien J. Hartmanis à croire que sa solution (si on la trouve) ne changera concrètement pas grand-chose à la situation de l'informatique. Et encore, et même en informatique, on croit découvrir l'indémontrabilité que ce qui est difficile à prouver ne peut pas être concrètement utile.

Cette constatation ne doit pas nous décourager de faire des mathématiques, mais au contraire doit nous rassurer : si ce qui est utile est toujours assez facile à démontrer, c'est une raison de plus pour

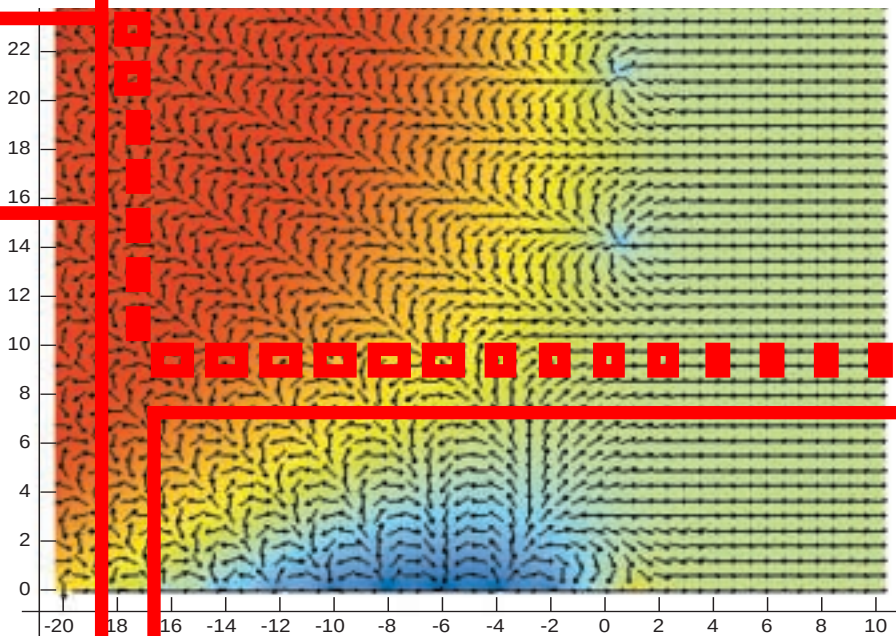
faire des mathématiques qui deviennent, pour ce qui est applicable, une science à la portée du plus grand nombre.

LA LOI DE L'INUTILITÉ DES PREUVES COMPLEXES

Les résultats mathématiques utiles semblent donc toujours simples, ou – ce qui revient au même – ce qui est mathématiquement complexe à démontrer appa-

raît toujours inutile sur le plan concret. Peut-on établir de telles affirmations générales et en apparence philosophiques ? Oui, et c'est ce que viennent de faire Vladik Kreinovich et Luc Longpré, deux chercheurs de la théorie de la complexité. Ils ont prouvé que, si un résultat est potentiellement utile, alors il n'est pas possible qu'il possède une preuve complexe.

Le résultat technique est impossible à expliquer dans tous ses détails, mais



4 La conjecture de Riemann est-elle utile ?

Le théorème suivant est dû à G. Miller et L. Bach.

Son hypothèse de Riemann généralisée (une conjecture liée à la répartition des nombres premiers, qui stipule que les zéros de cette fonction, représentée ci-dessus, sont sur la droite d'abscisse $1/2$) est vraie et si pour tout entier a entre 1 et $2 \cdot \ln^2(n)$ le nombre n est *a-fortement-probablement premier*, alors n est premier. (Le nombre n est *a-fortement-probablement premier* si, après avoir écrit $n - 1 = 2^q \cdot r$ avec q impair, on constate que $a^q \equiv 1 \pmod{n}$ ou que $a^{q \cdot 2^i} \equiv -1 \pmod{n}$ pour un entier $i = 0, 1, \dots, r - 1$.)

Considérons alors l'algorithme suivant :

Algorithme de Miller et Bach

- Pour tout $a = 2, 3$, jusqu'à $2 \cdot \ln^2(n)$, tester si n est *a-fortement-probablement premier* ;
- Dès que la réponse est non, répondre : n est composé ;
- Si la réponse a toujours été oui, répondre : n est premier.

Le résultat de Miller et Bach signifie que, si l'hypothèse de Riemann généralisée est vraie, cet algorithme ne se trompe jamais. Cet algorithme fonctionne en temps polynomial et est à dire rapidement. Nous avons donc un algorithme rapide pour tester la primalité des nombres entiers (ce qui est incontestablement utile), mais la garantie de son bon fonctionnement n'est pas assurée, car l'hypothèse de Riemann généralisée reste non démontrée. Il semble donc que l'hypothèse de Riemann généralisée est un résultat à la fois difficile et utile concrètement : sa démonstration garantirait que l'algorithme de Miller et Bach fonctionne toujours.

Un théorème récent de Vladik Kreinovich et Luc Longpré indique que, malgré cette situation, l'hypothèse de Riemann généralisée n'est pas vraiment utile, car, en pratique, sans l'hypothèse de Riemann, on montre que, sauf pour des cas très rares, l'algorithme de Miller et Bach fonctionne correctement.

Leur théorème, dont le sens est plus général, signifie que, si un résultat ne possède que des démonstrations complexes, alors il ne peut être concrètement utile.

5. La loi de l'inutilité des preuves complexes

Vladik Kreinovich et Luc Longpré ont prouvé, en utilisant les concepts et méthodes de la théorie de la complexité, que :

- Si un résultat est potentiellement utile, alors il n'est pas possible qu'il possède une preuve complexe.

Leur travail a comporté plusieurs étapes.

- Formulation d'une définition mathématique de la notion de résultat potentiellement utile. Selon eux, un résultat est utile s'il permet de prouver le bon fonctionnement d'un algorithme et qu'on ne peut pas substituer à cet algorithme un algorithme probabiliste (c'est-à-dire fonctionnant avec un risque d'erreur aussi petit qu'on le souhaite) dont le bon fonctionnement soit simple à établir.

- Formulation d'une définition mathématique de la notion de preuve complexe.

Leur définition prend en compte à la fois le nombre d'étapes élémentaires du raisonnement de la preuve et les possibilités de simplification dues à d'éventuelles parties répétitives. Leur définition est telle qu'une preuve simple est une preuve longue, et qui le reste même après exploitation des parties répétitives.

- Utilisation de la théorie de la complexité pour prouver leur affirmation.

Le résultat de Kreinovich et Longpré est un exemple récent provenant de la théorie de la complexité de résultats mathématiques ayant un contenu philosophique intéressant et inattendu. Il est probable qu'il déplaça aux mathématiciens qui aiment juger leurs travaux, même les plus difficiles, en évoquant de possibles applications futures.

tentons tout de même de le présenter, il en vaut la peine.

Les deux chercheurs, pour arriver à leur résultat, ont d'abord donné une définition mathématique de la notion de résultat potentiellement utile.

Pour eux, un tel résultat est un énoncé S qui permet d'établir une formule du type «pour tout x , $P(x)$ », où $P(x)$ est facilement vérifiable (par exemple, une formule du type «si la somme des chiffres de x est un multiple de 9, alors x est un multiple de 9»).

Leur définition est justifiée, car ce sont les résultats de ce type qui assurent que les méthodes algorithmiques fonctionnent correctement : algorithmes de codage, algorithmes de recherche dans une base de données, algorithmes de tri de données, algorithmes de calcul numérique, algorithmes de simulation, etc.

Ensuite les deux chercheurs définissent une preuve complexe d'une manière fine et subtile (voir la figure 5).

Leur résultat, formulé avec plus de précision, est que si S n'est pas une preuve complexe et permet de montrer que «pour tout x , $P(x)$ », alors pour la quasi-totalité des données d'entrée raisonnables (les seules qu'on rencontre en pratique) $P(x)$ est vrai indépendamment de S . En quelques mots : si S est complexe, et que S est une preuve (pour tout x , $P(x)$) alors, sauf de manière exceptionnelle, on n'a pas besoin de S pour savoir que $P(x)$ est vrai.

En très court, si S est complexe, alors ce n'est pas un résultat potentiellement utile.

Ce résultat s'applique en particulier à la conjecture de Riemann (un résultat qui concerne la répartition des nombres pre-

miers et qui est considéré comme l'énoncé mathématique irrésolu le plus recherché des mathématiques aujourd'hui). Il a été prouvé par G. Miller et E. Bach que, si une forme généralisée de l'hypothèse de Riemann était vraie, alors il existait un algorithme efficace pour tester si un entier est un nombre polynomial, c'est-à-dire efficacement. Cette situation a conduit de nombreux mathématiciens à croire que les recherches sur l'hypothèse de Riemann étaient utiles, même dans une optique d'applications pratiques.

EST-IL VRAIMENT UTILE DE MONTRER RIEMANN ?

La situation créée par l'algorithme de Miller est qu'on dispose sans doute d'un algorithme efficace de test de primalité dont le bon fonctionnement dépend de la vérité de l'hypothèse de Riemann généralisée. Ce qu'énonce le résultat de Vladik Kreinovich et Luc Longpré est : si l'hypothèse de Riemann généralisée ne possède pas de preuve simple, alors l'algorithme de Miller pour des entiers raisonnables fonctionne correctement pour la quasi-totalité des cas qu'on peut lui soumettre. On n'a donc pas besoin en pratique de la démonstration de l'hypothèse de Riemann généralisée si l'on désire utiliser cet algorithme.

Cette inutilité pratique de l'hypothèse de Riemann dans le cas des tests de primalité est d'ailleurs corroborée par l'usage informatique réel : dans les applications cryptographiques où l'utilisation des nombres premiers est intensive, on se sert le plus souvent des méthodes probabi-

listes de test de primalité, c'est-à-dire de méthodes dont le bon fonctionnement est justifié par un raisonnement relativement simple, mais qui ne donnent un résultat que d'une manière probabiliste, autrement dit avec un risque, jugé négligeable, d'erreur.

TOUT N'EST PAS TERMINÉ

Notons, pour l'amusement, que le résultat de Vladik Kreinovich et Luc Longpré possède une démonstration assez courte (environ une page) : il n'est donc pas impossible d'après ce qu'il nous apprend, qu'il soit un résultat concrètement utile !

Nous ne saurions qu'en toute rigueur malgré les quelques recherches faites à son sujet, il n'est pas logiquement exclu que la conjecture de Riemann possède une démonstration simple. Si c'est le cas, l'hypothèse de Riemann pourrait être pratiquement utile. Nous devons continuer à en rechercher la preuve, même si l'intérêt mathématique par lui-même ne semble pas une raison suffisante de nous y intéresser.

Précisons que, même si, après une traduction mathématique sous une certaine forme particulière, on a réussi à établir que le difficile est inutile, il ne faut pas donner une importance excessive au résultat de Vladik Kreinovich et Luc Longpré qui, dans leurs définitions, le sens des mots peut-être un peu rapidement.

Dans le cœur de tous les mathématiciens est inscrite la conviction que plus c'est difficile, plus c'est beau et important, et donc utile au moins par l'approfondissement et une meilleure vision que cela procure à notre perception du monde des entités abstraites. «Le but unique de la science, c'est l'honneur de l'esprit humain», écrivait le mathématicien Jacobi dans une lettre adressée à son collègue Legendre en 1830. Il est peu vraisemblable que le formalisme mathématique s'empare un jour de cette vérité philosophique-là !

G. BACHELARD, *La philosophie du non*, Presses Universitaires de France, Paris, 1940.

W. H. ZUREK, *Algorithmic Randomness and Physical Entropy*, in *Physical Review A*, n° 8, vol. 40, pp. 4731-4751, 1989.

The Philosophy of Mathematics, sous la direction de W. H. Hart, Oxford University Press, 1999.

MING-CHENG CHEN et PAUL VITANYI, *An Introduction to Kolmogorov Complexity and Its Applications*, Springer-Verlag, New York, 1997.

J.-P. DELORME, *Complexité, hasard et hasard*, Hermès, Paris, 1999.

V. KREINOVICH et L. LONGPRÉ, *How Important Is Theory for Practical Problems?*, in *Bulletin of the European Association for Theoretical Computer Science (EATCS)*, n° 71, pp. 160-164, 2000.

Le droit au mélange

BERNARD CABANE - SYLVIE LIÉNON

L'eau se déplace à travers les membranes par osmose. Ce phénomène permet de détruire les bactéries par déshydratation ou, en l'inversant, de dessaler l'eau de mer.

Vivendi

Aux temps héroïques de l'exploration pétrolière des régions tropicales, on protégeait de la gangrène les plaies infectées en les recouvrant de sucre, traitement encore utilisé aujourd'hui à l'hôpital pour les plaies profondes. Le sucre préserve aussi les confitures des proliférations bactériennes et le sel conserve viandes et poissons. Comment sel et sucre empêchent-ils la croissance bactérienne ? Par le phénomène d'osmose, important dans les mécanismes biologiques et industriels.

La plupart des bactéries sont des organismes unicellulaires. Remplies d'une solution aqueuse d'ions, d'acides aminés et de protéines, elles ne sont protégées de l'extérieur que par une membrane constituée d'une mince couche de lipides et de protéines, perméable à l'eau. Lorsque les bactéries sont plongées dans un milieu aqueux, l'eau entre et sort de la cellule à travers la membrane.

Plongeons une bactérie dans du sucre : comme il y a beaucoup plus de molécules de sucre à l'extérieur de la bactérie que de molécules à l'intérieur, la quasi totalité de l'eau traverse la membrane afin de diluer le sucre. La bactérie se déshydrate et meurt. En effet, selon le second principe de la thermodynamique, la situation la plus favorable est celle où les molécules d'eau se mélan-

gent équitablement avec toutes les autres molécules (les molécules internes à la bactérie et le sucre externe).

Inversément, le contact avec l'eau pure est dangereux pour les cellules de notre corps, parce qu'elles contiennent des liquides salés. Dans notre corps, elles sont en contact avec les fluides physiologiques, salés eux aussi. Si nous introduisons de l'eau pure dans un de ces fluides, nous créons un déséquilibre. Par exemple, si, pour nettoyer nos yeux ou nos lentilles de contact, nous utilisons de l'eau pure au lieu d'une solution saline, en vertu de son « droit au mélange », l'eau passe dans les cellules de la cornée, ce qui crée un œdème.

L'ÉQUILIBRE OSMOTIQUE

Pour éviter que nos cellules ne gonflent ou ne dégonflent par des transferts d'eau, nous contrôlons les mélanges de l'eau avec les sels, à l'intérieur et à l'extérieur des cellules, doivent être équivalents. Comme le métabolisme des cellules produit de l'eau, une régulation est nécessaire afin que notre teneur en eau n'augmente pas : cette régulation est assurée par les reins.

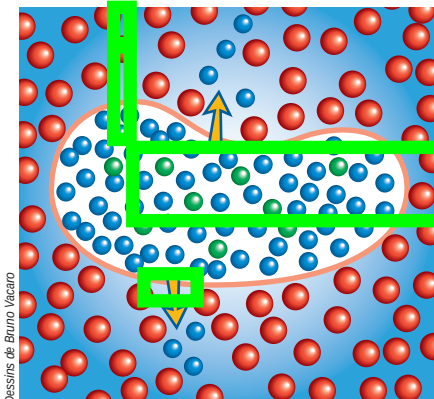
Si l'eau passe facilement à travers les membranes des cellules, le pas-

sage des sels et des molécules plus grosses est restreint. Ainsi, l'important pour l'équilibre osmotique n'est pas que les fluides interne et externe contiennent les mêmes composés en solution, mais que la proportion d'eau, en nombre de molécules, soit égale des deux côtés de la membrane. Par exemple, l'équilibre osmotique doit être assuré pour le sérum administré lors des perfusions, afin que les globules rouges du sang n'éclatent pas. On utilise des solutions salées, ou des solutions de glucose, qui nourrissent l'organisme tout en assurant son équilibre en eau.

Loin du monde médical, la culture des tomates dans le désert illustre le maintien de l'équilibre osmotique ; comme l'eau douce y est rare, on arrose les plantations avec des saumures. La forte concentration en sels à l'extérieur des cellules végétales pourrait entraîner la sortie de l'eau cellulaire : afin d'éviter la déshydratation, les cellules végétales sécrètent une grande quantité de sucres ou d'acides aminés qui équilibrent les concentrations en eau à l'intérieur et à l'extérieur des cellules. Les tomates ainsi cultivées dans le désert sont plus sucrées que celles qui sont alimentées avec de l'eau douce.

Les cellules végétales diffèrent toutefois des bactéries ou de nos propres cellules car leur concentration interne en espèces dissoutes reste supérieure à celle du milieu extérieur : l'eau externe devrait alors entrer dans les cellules et les faire gonfler, mais la pression à l'intérieur de la cellule, supérieure à la pression extérieure, l'en empêche. L'application d'une pression s'oppose ainsi à l'osmose.

Pour-on reproduire ce phénomène à grande échelle ? Oui, à l'aide de membranes qui laissent passer l'eau, mais ni le sel, ni le sucre. Imaginons un récipient rempli d'eau pure divisé par une membrane en deux compartiments. Ajoutons du sucre dans un compartiment. L'eau se déplace alors vers le compartiment sucré toujours en vertu du « droit au mélange ».



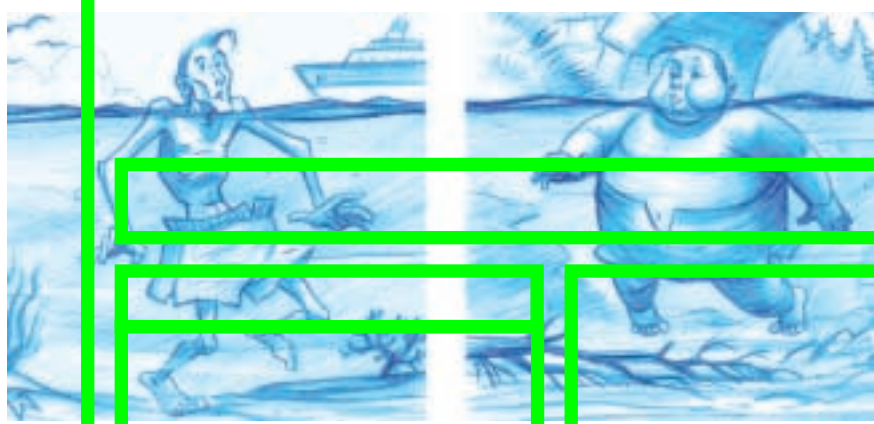
Dessins de Bruno Vacaro

Quand une bactérie est plongée dans une solution sucrée, la proportion d'eau à l'extérieur est inférieure à la proportion à l'intérieur (à gauche) : les molécules d'eau (boules bleues) traversent la paroi cellulaire jusqu'à rétablir l'équilibre. La bactérie se déshydrate et meurt (à droite).

Le niveau dans le compartiment sucré monte. À une certaine hauteur, la colonne d'eau soulevée crée une surpression qui s'oppose au flux des molécules d'eau diffusant dans le compartiment sucré. En appliquant au compartiment sucré cette même surpression, on s'oppose au passage de l'eau. Cette surpression mesure la force qu'il faut exercer pour équilibrer celle du mélange de l'eau et du sucre : c'est la pression osmotique de la solution de sucre, qui dépend de sa concentration.

Quelle est la valeur de cette pression? En ajoutant une pincée de sucre dans l'eau, peut-on soulever des colonnes d'eau sur de grandes hauteurs? Pour l'estimer il faut mesurer la variation d'une quantité thermodynamique dénommée entropie. L'entropie est une mesure du désordre et les systèmes évoluent vers une augmentation de cette grandeur. Initialement, elle était faible, car le sucre en forte concentration était confiné dans un compartiment : une égalisation des concentrations augmente le désordre, donc l'entropie. On calcule combien d'entropie l'ensemble des molécules gagne grâce au mélange.

On retrouve pour les solutions, ou pour le mélange de deux liquides, les mêmes formules que pour un gaz. La pression osmotique s'exprime ainsi, à faible concentration, par une sorte de loi des gaz parfaits. Pour un gaz, s'il y a une mole de molécules à 25 °C dans un volume de 24,5 litres, la pression est égale à une atmosphère. Pour une solution de sucre, s'il y a une mole de molécules de sucre (soit environ 333 grammes) dans un volume d'eau égal à 24,5 litres, la pression osmotique sera aussi de une atmosphère. Comme c'est le nombre de molécules d'espèces dissoutes qui compte, il suffit de mettre une demi-mole de sel (environ 30 grammes) dans 24,5 litres d'eau pour obtenir la même pression osmotique de une atmosphère, car le cristal de sel



Les liquides physiologiques contiennent plus de sels que l'eau douce, mais moins que l'eau de mer. Si notre peau était plus perméable à l'eau, un bain de mer nous déshydraterait. À l'inverse, si nous baignant dans une rivière, nous gonflerions rapidement.

se dissout en deux espèces ioniques Na^+ et Cl^- .

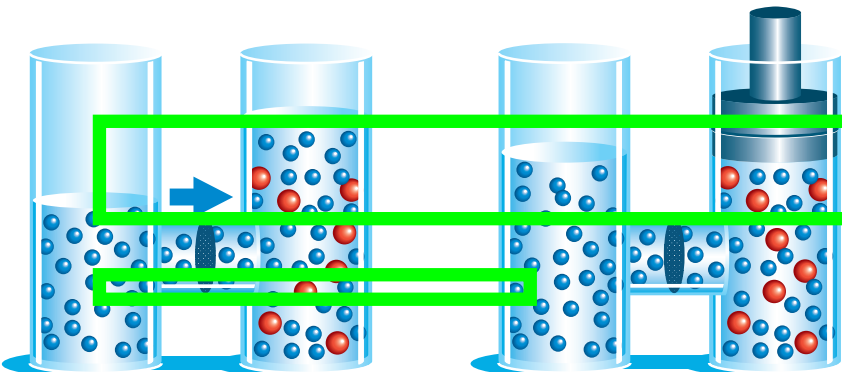
Dans notre corps la pression osmotique des fluides est élevée : en comptant tous les ions des sels, on arrive à une pression osmotique égale à plus de sept atmosphères, qui pourrait faire monter une colonne d'eau à 70 mètres de hauteur! Cette force énorme nous enjoint d'équilibrer la pression des fluides qui viennent en contact avec nos fluides internes. Nous pouvons toutefois nager dans une rivière ou dans la mer, mais nous sommes alors protégés par notre peau, qui est une barrière à travers laquelle l'eau passe trop lentement pour endommager nos cellules.

L'OSMOSE INVERSE

Au lieu de laisser monter la colonne d'eau, on peut pousser dessus. Si la pression appliquée est suffisante, l'eau quitte le mélange pour repasser à travers la membrane vers le compartiment non sucré, ou non salé. Ce mécanisme de passage forcé contre le droit au mélange est l'osmose inverse. Pourquoi ne pas dessaler l'eau de mer

ainsi, en appuyant dessus? C'est l'une des méthodes utilisées, mais la pression à exercer est élevée, 25 atmosphères environ, car l'eau des mers contient encore plus d'ions que les fluides physiologiques (le cartouche représente une usine de dessalement). Le maintien de cette pression pour obtenir des quantités notables d'eau douce nécessite une énergie considérable. De surcroît, les membranes doivent résister au passage d'eau sous des pressions de plusieurs dizaines d'atmosphères, tout en laissant passer un flux d'eau important. Aujourd'hui, on dispose de membranes en polymère organique qui laissent passer des flux d'eau bien supérieurs à ceux des membranes en boyaux utilisées par les chimistes du XIX^e siècle.

L'osmose inverse est l'une des méthodes les plus efficaces pour le dessalement de l'eau : l'énergie mécanique à fournir est environ dix fois inférieure à l'énergie thermique nécessaire pour vaporiser la même quantité d'eau. Le coût du dessalement par osmose inverse reste toutefois élevé, comparable à celui de la distillation qui utilise du pétrole, abondant et peu onéreux dans certains pays du Golfe. Mais les techniques évoluent : la production d'un mètre cube d'eau à partir d'eau de mer par osmose inverse qui nécessitait autrefois une heure il y a cinq ans, elle n'en requiert que trois aujourd'hui.



Dans deux récipients séparés par une membrane perméable à l'eau mais pas au sucre (en rouge), une différence de concentration de l'eau entraîne le passage de l'eau pour rétablir l'équilibre des concentrations : c'est l'osmose. L'eau traverse la membrane pour se mélanger au sucre, excepté si l'on applique une pression à la solution sucrée (à droite).

B. CABANE, nouvelle au Laboratoire physique et mécanique des milieux hétérogènes de l'ESPCI (CNRS), S. HÉNON, au Laboratoire de biorhéologie et hydrodynamique physico-chimique à Jussieu.

R.P. RAND, *Raising water to New Heights*, in *Science*, vol. 256, p. 618, 1992.

C. BONNET-GONNET, L. BELLONI, B. CABANE, *Osmotic pressures of latex dispersions*, in *Langmuir*, vol. 10, p. 4012, 1994.



La sphère à cornes

ADRIEN DOUADY • LOIC MANGIN

Une sculpture enchevêtrée nous entraîne dans les méandres de la topologie.

Dans le patio Sud du bâtiment du Centre de mathématiques et d'informatique de l'Université de Chateaubriant à Nantes se trouve une sculpture inhabituelle. Composée de matériaux polymères composites, elle est l'œuvre de Jean-Louis Lhermitte qui l'a nommée *La sphère à cornes*. Pourquoi une sphère, alors qu'elle ne ressemble en rien à la forme connue d'une orange ou de la Terre. Doit-on voir dans cette dénomination une bizarrerie d'artiste ? Non, elle est l'œuvre du mathématicien américain

James Alexander (1888 - 1971) qui a conçu cet objet en 1915.

D'une masse oblongue s'étendent deux prolongements, de la même façon que les bras prolongent le corps. Les deux extrémités de ces prolongements, en forme de portion d'anneaux, se ramifient également en deux plus petits arcs d'anneaux qui s'enlacent comme les doigts formant deux maillons ouverts d'une chaîne. À nouveau, les extrémités des "doigts" se ramifient, les extensions s'enlacent... (voir la figure 1), l'opération est répétée à l'infini pour l'objet théorique d'Alexander, d'un volume fini, qui a donc un caractère fractal : un agrandissement d'une région aux alentours d'un des points limites (les «extrémités ultimes» des ramifications) a la même apparence que l'ensemble tout entier. Dans la sculpture de J.-L. Lhermitte, les itérations à l'infini ne sont pas représentées, seules les cinq premières le sont : elles évoquent l'ensemble, comme toute représentation d'une fractale.

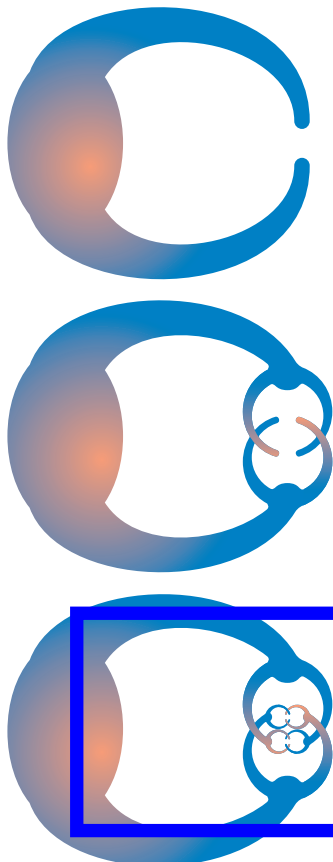
La topologie établit le rapport entre cet objet à l'allure végétale et une sphère ordinaire. La topologie est la science des objets en caoutchouc, donc déformables. Pour la topologie, une chambre à air et une tasse à café sont équivalents ! Pourquoi ? Parce qu'il est possible de passer de l'une à l'autre par une transformation continue en poussant les parois en caoutchouc : la tasse, la chambre à air devient l'anse de la tasse ; les deux objets sont des tores.

Dans des termes plus mathématiques, deux objets sont topologiquement équivalents (ou appartiennent à la même classe) lorsqu'il existe une bijection entre les points de la surface des deux objets : à chaque point de l'un correspond un point de l'autre. Cette définition est dite «intrinsèque», elle devient «plongée» quand une telle bijection existe aussi pour les régions délimitées — l'intérieur et l'extérieur — par les ensembles de points constituant les objets. Selon l'espace considéré, la première définition n'entraîne pas obli-

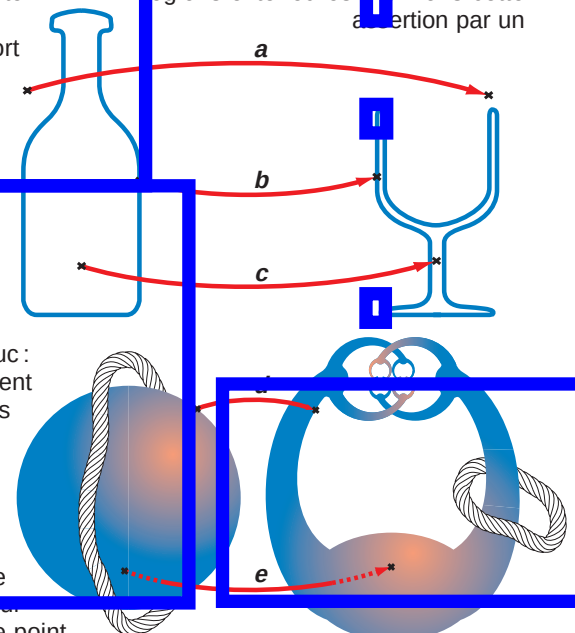
gatoirement la seconde : c'est ce que montre *La sphère à cornes*.

Dans un espace à deux dimensions, un plan les deux définitions vont de pair. Cette alliance est au cœur du théorème de Jordan, du nom du mathématicien ardennais Camille Jordan (1838 - 1922), qui stipule qu'une courbe fermée simple, c'est-à-dire sans croisement, sépare ce plan en deux régions. On peut établir une bijection entre les deux régions délimitées par la courbe et celles délimitées par un cercle. En apparence évidente, la démonstration de ce théorème est ardue. Dans un espace à trois dimensions, le théorème n'est plus vrai : la sphère à cornes en est une illustration.

La surface d'une sphère ordinaire est en bijection avec la surface de la sphère à cornes. La bijection existe aussi entre le volume renfermé par ces deux surfaces. En revanche, ce n'est plus vrai pour les régions extérieures. Vérifions cette affirmation par un



1. LA SPHÈRE À CORNES est obtenue par déformation d'une sphère ordinaire dont deux extensions en arcs d'anneaux se ramifient, les ramifications se scindent à leur tour en deux, et ce jusqu'à l'infini : on obtient ainsi un objet fractal. En vignette à côté du titre, la représentation de la sphère à cornes symbolise les dédales du cerveau prolifique de James Conway, mathématicien à l'Université de Cambridge.



2. DEUX OBJETS, TENUS DEUX COURBES fermées (en haut), sont topologiquement équivalents lorsqu'il existe une bijection entre leurs périmètres et les deux régions du plan qu'ils délimitent (a, b et c). Cependant, l'extérieur d'une sphère ordinaire n'est pas en bijection avec celui de la sphère à cornes (en bas) : une ficelle fermée peut être séparée de la première, mais pas de la seconde.

raisonnement par l'absurde. Si une telle bijection existe, à un point de l'extérieur d'une sphère ordinaire correspond un point de l'extérieur de la sphère à cornes. Donc, à un ensemble de point correspond un autre ensemble de points. Prenons le cas où l'un de ces ensembles de points est une «ficelle» qui entoure les objets. Cette ficelle peut être séparée de la sphère ordinaire par une suite de positions. La bijection imposerait l'existence d'une autre suite de positions qui permettrait de séparer une ficelle de la sphère à cornes. Or, une ficelle qui entoure l'un des bras de la sphère à cornes ne peut pas quitter l'ensemble. Il n'y a donc pas de bijection entre les régions extérieures des deux objets (voir la figure 2).

En topologie, les mathématiciens classent des objets en fonction de différentes équivalences : selon la catégorie des objets étudiés et la relation d'équivalence choisie, on obtient différentes classifications. Pour les courbes fermées d'un plan, le théorème de Jordan montre qu'il n'y a qu'une seule classe de courbes fermées simples. Pour des surfaces considérées de façon intrinsèque (on ne s'intéresse pas aux régions qu'elles délimitent), les classes se distinguent selon le nombre p de trous des tores auxquels les surfaces sont équivalentes. Ainsi le tore (la chambre à air) à un trou diffère de la sphère à zéro trou et du bretzel à trois trous. Pour étudier les surfaces plongées dans un espace à trois dimensions, c'est-à-dire les régions délimitées par ces surfaces, l'exemple de la sphère à cornes montre que la classification devient très complexe en l'absence d'hypothèse de régularité : la sphère ordinaire est régulière, ou lisse, car un agrandissement d'une région quelconque est plan, tandis qu'un tel agrandissement d'une région de la sphère à cornes aux alentours des points limites montre de nombreux pics : la sphère à cornes est irrégulière, voire «sauvage».

Pour des objets lisses, on définit trois relations d'équivalence : l'équivalence topologique (des bijections existent pour les objets eux-mêmes), l'équivalence différentiable (la bijection qui transforme un objet en l'autre est différentiable), et l'équivalence homotopique (elle requiert la notion de déformation continue). Pour les surfaces, des objets équivalents selon ces trois relations appartiennent à la même classe. En est-il de même pour des objets en trois dimensions ? Cette question est toujours non résolue. Il s'agit de l'hypothèse de Poincaré : Si vous la résolvez, vous gagnez un million de dollars.

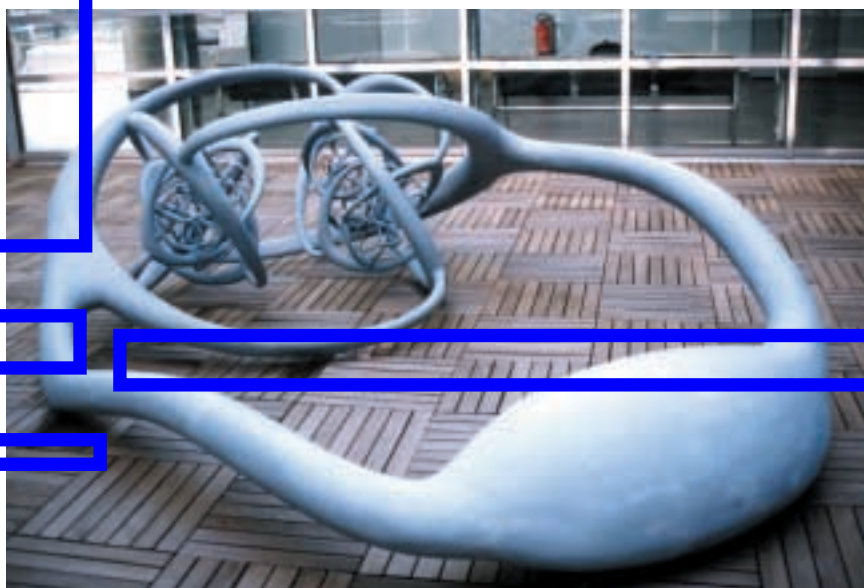
Adrien DOUADY est mathématicien à l'Université d'Orsay.



Jean-Louis Lhermitte



Jean-Louis Lhermitte



Jean-Louis Lhermitte

3. La sphère à cornes (ici, vue sous plusieurs angles), trône dans un patio du Centre de mathématiques et d'informatique de Chateau-Gombert, à Marseille.

Chimie

Un chimiste au passé simple

Jean Jacques. Éditions Odile Jacob, 2000 (240 pages, 145 francs).

Jean Jacques, directeur de recherche émérite au CNRS, a fait toute sa carrière scientifique au Collège de France. Spécialiste de stéréochimie organique et, s'il fallait caractériser son œuvre par un seul mot, ce serait certainement «originalité». À travers ses analyses très personnelles, son refus du conformisme, il a exercé une influence profonde sur ses élèves directs, dont j'ai eu l'honneur de faire partie, et sur plusieurs générations de chimistes.

Voici presque 30 ans, J. Jacques publiait les *Confessions d'un chimiste ordinaire*, un ouvrage conçu pendant et après un accident cardiaque, période propice pour un retour sur soi et un regard en arrière.

Il récidive, avec *Un chimiste au passé simple*. Dans cette nouvelle version de ses mémoires, on retrouve sa verve, son humour caustique, la distance qu'il sait prendre par rapport aux événements, notamment ceux qui le concernent, mais le style s'est affirmé, et le document est bien différent des *Confessions* : c'est une autobiographie plus littéraire. L'âge avançant, ses méditations sur la vieillesse et la mort sont plus présentes, ce qui rend ce texte plus humain.

J. Jacques est fasciné par les histoires de l'Histoire, et cela dans les domaines où, quand il exerce au Collège de France, il est le plus à l'aise : le métier favori, celui d'historien des sciences et, notamment, de la chimie, il se passionne pour la vie de ses acteurs, recherche l'anecdote qui révèle l'homme à travers le travail intellectuel. Quand il parle de sa vie, de son histoire, c'est toujours en les situant par rapport à des repères historiques, à des courants de pensée...

Il nous conte son parcours, son enfance dans la fonderie paternelle à Épinay-sur-Seine, ses vacances dans les Ardennes chez le grand-père qui l'amenait pêcher à la ligne et... chanter la

messe le dimanche, ses activités de lycéen «hors normes».

Fasciné par son métier de chimiste, il était attiré par une carrière littéraire ; sa famille le convainc cependant qu'il pousserait un métier plus concret, et c'est ainsi qu'il se retrouva élève ingénieur à l'Institut de chimie de Paris, l'actuelle École nationale supérieure de chimie de Paris. Ce qui est moins classique, c'est qu'il mena de front des études de philosophie et obtint un diplôme d'ingénieur chimiste et même temps qu'une licence de philosophie.

Après tout, que, grâce à des amis d'enfance ou de jeunesse, il fréquenta les milieux littéraires et artistiques en vue. On voit défilier dans son récit Mario et Gisèle Grassino, César – qui réalisa ses premières compositions dans l'atelier du «Père Jacques» –, Clavé, Giacometti... On comprend pourquoi ce «chimiste pas ordinaire» a un charme unique, par son regard si différent de celui de la plupart des scientifiques «purs et durs».

Où est la chimie, dans tout cela ? Elle est omniprésente, à toutes les étapes de la vie de l'auteur, à travers une présentation bien à lui. Ses activités de «chimiste militaire» en 1939-1940 dans le laboratoire de Marcel Delépine, au Collège de France, sont l'occasion de pages caustiques et savoureuses à propos de la guerre chimique. J'ai relu avec grand plaisir ses histoires d'ancien combattant, si souvent entendues, l'épopée de son évvasion manquée d'un camp de prisonniers, sa libération, son retour au Collège de France, l'épisode de chimie pratique que fut la fabrication de cocktails Molotov, pendant la libération de Paris, avec Frédéric Joliot-Curie.

Sa mission militaire en Allemagne, en 1945, pour espionner le laboratoire d'Adolf Butenandt (prix Nobel de chimie en 1939, le découvreur des hormones sexuelles stéroïdes, puis des phéromones d'insectes), nous vaut un chapitre bien documenté sur la chimie des substances naturelles. Ce faisant, le lecteur ne pourra rester insensible à la narration des histoires d'amour de papillons. Toujours à propos d'hormones, la rencontre avec Gregory Pincus, le père de la pilule, est l'occasion d'une intéressante page d'histoire scientifique. Le thème de stratégie industrielle.

Un autre exemple de chimie pratique, sinon amusante, nous est narré à travers le fameux «casern» imaginé par J. Jacques introduisant avec du vinyle, au Palais de justice de Pointe-à-Pitre, acide chlorhydrique et permanganate de potassium afin de prouver expérimentalement, devant un tribunal médusé, que leur mélange n'est pas explosif. Il s'agissait de défendre Yves Jasor, un ancien

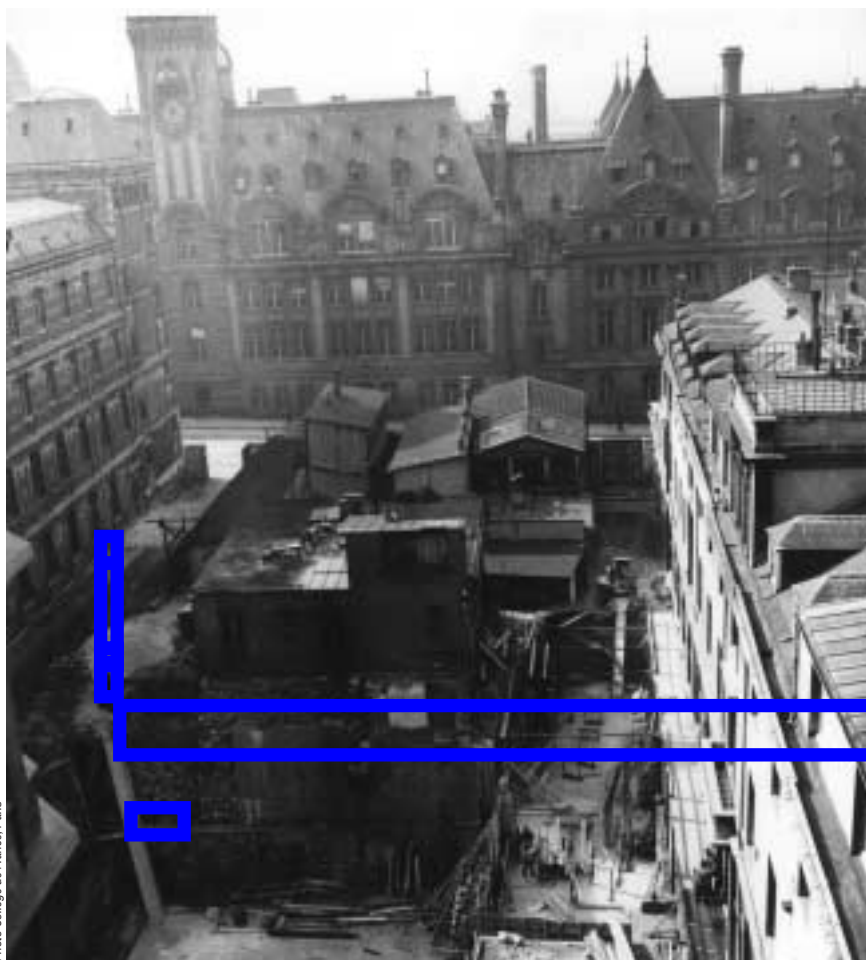


Photo Collège de France, Paris

Le Collège de France : dans le trou, au premier plan, on essayait des grenades expérimentales durant la Seconde Guerre mondiale.

collègue revenu enseigner la chimie en Guadeloupe, accusé de terrorisme pour avoir transporté des produits dans sa voiture... Cet épisode est aussi prétexte à discuter les difficiles problèmes d'«expertise» et à jeter un regard plein de sympathie sur les problèmes des Antilles.

J'ai particulièrement apprécié le dernier chapitre «Un tour du Luxembourg à pas comptés». Cette méditation est une petite merveille. Je laisse au lecteur le plaisir de la découvrir.

Ce livre comporte quelques longueurs, par exemple dans le chapitre consacré à l'affaire Jasor, mais il se dévore malgré tout comme un roman. Le lecteur scientifique appréciera d'y trouver, plaisamment analysés, les us et coutumes de son milieu. Tous y apprendront un peu d'histoire de la chimie et tous passeront quelques heures délicieuses à écouter un sage parler, avec humour, de sa vie et de sa carrière. L'avertissement du livre en donne un avant-goût.

J. Jacques dit d'abord l'histoire des orphelins. «Les orphelins n'imaginent pas l'acharnement à survivre dont sont capables certains orphelins... Ils ne savent pas le plaisir de raconter leurs congés payés au Tréport en 1936 à des gens qui s'en foutent.» Il ajoute : «Ce livre est mon Tréport à moi, en 1999. Ça ne s'est jamais allé au Tréport. J'aurais pu, mais j'étais ailleurs.»

Andrée MARQUET

Entomologie

Coléoptères phytophages d'Europe

Gaétan du Chatenet. Éditions N.A.P., 2000 (368 pages, 330 francs).

Ce livre est consacré à seulement quelques familles sur la centaine que comptent les Coléoptères. Ce sont essentiellement les Cléridés, les Élatéridés, les Buprestidés et les Cérambycides qui sont examinés ici.

Leurs représentants ont un régime alimentaire varié. Certains sont des prédateurs, mais d'autres mangent des racines, du bois ou, même, des matériaux en décomposition (on les trouve dans les arbres morts ou mourants, par exemple). Autrement dit, la majorité des espèces examinées dans le livre de G. du Chatenet se rencontre dans les forêts, où ils participent, pour la plupart, à la dégradation des bois. Dans cette mesure, peu sont foncièrement nuisibles.

Sur ces onze familles de coléoptères représentées, sept sont rares : les Eymnoptéridés, les Cérophytidés, les Eucnemidés, les Lissomidés, les Throscidés (tous des décomposeurs de bois), les Déracéridés (des prédateurs de pucerons), les Cébrionidés (qui vivent essentiellement dans le sol). Tous les Cléridés sont prédateurs d'insectes variés, de pontes d'acridiens, de nids d'hyménoptères, et l'un des genres, *Necrobia* (du grec «mort et vie»), doit son nom à l'une de ces aventures qui font l'histoire de l'entomologie. Après la Révolution française, le prêtre et entomologiste Pierre-André Latreille (1762-1833) fut considéré comme réfractaire, bien que ne deservant pas de paroisse. Il devait être déporté en Guyane en 1794, mais, dans sa cellule, il découvrit un insecte qu'il croyait nouveau. Il communiqua sa trouvaille au médecin qui soignait un codétenu, lequel transmit l'animal à l'un de ses amis, Bory de Saint-Vincent, qui le nomma Latreille par ses écrits. Bory le reconnut et élargit celui qu'on sur-nomme aujourd'hui le «prince de l'entomologie». Ce dernier fut doublement sauvé, car le bateau qui aurait dû l'em-mener comme prisonnier du régime, ayant tous les déportés (l'histoire n'est pas contée dans le livre).

Les Élatéridés ont des larves qui vivent surtout dans le bois décomposé dont elles se nourrissent ou dans lequel elles trouvent des proies. Quelques-unes sont nuisibles, vivant dans le sol, où elles attaquent aux racines des plantes cultivées, comme les *Agriotes*.

Les Buprestidés, aux couleurs brillantes, occupent essentiellement les

lieux exposés, à l'état adulte, et sont xylophages à l'état larvaire.

Les Cérambycides, aux nombreuses espèces, ont certaines de grande taille, ont des larves presque toutes xylophages. Pour les Buprestidés et les Cérambycides, la quasi-totalité des espèces d'Europe figurent dans le livre, pour les Élatéridés, les genres y sont tous considérés. Chaque représentant est décrit avec des notes sur ses mœurs, son habitat et sa répartition géographique. Les planches en couleurs, au nombre de 43, sont dues à l'auteur dont la maîtrise artistique est réputée. Plus de 700 animaux sont ainsi représentés.

Cet ouvrage nous fait regretter qu'il ne soit pas une vraie suite du *Guide des coléoptères*, paru en 1986, du même auteur, et resté inachevé. Cet ouvrage suivait la classification et avait l'avantage, tout en excluant les espèces de petite taille, de passer en revue les différentes familles. C'est ainsi que le sous-ordre des Archostemata et la famille des Tétrapaleridés étaient quand même mentionnés, avec une espèce d'à peine plus d'un millimètre. Bien qu'il soit prévu un autre volume, avec notamment les Chrysomélides et les Curculionides, des groupes entiers risquent de manquer, tels les Cucujoïdés ou les Dermestoïdés, qui comptent des espèces courantes de bonne taille.

Ce regret exprimé, ne boudons pas notre plaisir de disposer d'un ouvrage de vulgarisation, réservé à quelques familles de coléoptères prisées des collectionneurs et fréquemment rencontrées.

Jacques d'AGUILAR



Ce bupreste se rencontre au soleil, sur les troncs de pins fraîchement abattus. Les buprestes, dont on connaît près de 200 espèces en Europe, occupent surtout les parties les plus chaudes de la France.

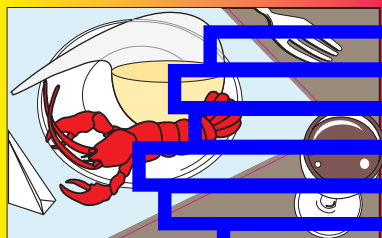
FRANCE INFO

POUR LA SCIENCE

vous invitent
à écouter la chronique
de Marie-Odile Monchicourt :

Info Sciences

Tous les jours
sur **France Info**
à 15 heures 41,
17 heures 10,
20 heures 12,
22 heures 12,
et 23 heures 42



Les prochains
résultats présentés
dans la rubrique
Science et Gastronomie
seront annoncés sur
France Info
le 29 novembre 2000.



Biologie

Biologie et physiologie animales (bases moléculaires, cellulaires, anatomiques et fonctionnelles ; orientations comparée et évolutive)

Rüdiger Wehner et Walter Gehring.
Traduit de l'allemand par Christiane
Meyer. Éditions De Boeck, 1999
(364 pages, 385 francs).

Physiologie animale (mécanismes et adaptations)

Sous la direction de Roger Eckert
(traduit de la 4e édition américaine par
François Math). Éditions De Boeck,
1999 (840 pages, 485 francs).

De nombreux livres ont un titre
voisin de celui des deux ouvrages
ci-dessus. Aussi, ce titre galvaudé,
quel que peu obscur, ne renseigne
pas suffisamment sur l'orientation que les
auteurs souhaitent donner à leur ouvrage.
Pour y remédier, un sous-titre devient
nécessaire. C'est ce qu'a compris la tra-
ductrice de l'ouvrage de R. Wehner et
W. Gehring, qui a modestement apporté
des précisions où apparaissent le mot
«bases», tandis que disparaissait le titre
Zoologie de l'ouvrage original, vraisem-
blablement trop restrictif pour elle.

Aucun des deux ouvrages, parus qua-
siment en même temps, ne fait de l'ombre
à l'autre : ils sont complémentaires.
R. Eckert et ses collègues ont opté pour
une division de leur ouvrage en trois
parties : la première est consacrée aux
principes de base de la biophysique et de
la biochimie, aux techniques expé-
rimentales utilisées en physiologie ; les
deux autres, qui constituent le cœur de
l'ouvrage, traitent respectivement des pro-
cessus physiologiques et de l'inté-
pendance des systèmes physiologiques
et leur intégration dans les différents envi-
ronnements. En revanche, R. Wehner
et W. Gehring ont préféré, dans les cinq
parties qui constituent leur ouvrage, une
cohabitation harmonieuse des aspects
techniques et scientifiques (même de
base) et des aspects historiques, ce qui
donne à celui-ci un ton original.

En plus des thèmes traités dans l'ou-
vrage de R. Eckert et de ses collègues,
R. Wehner et W. Gehring y intègrent aussi
deux domaines de la biologie qui connais-
sent aujourd'hui un développement spec-
taculaire : celui formé par le programme
génétique et sa réalisation dans l'orga-
nisme, et celui formé par l'éthologie, l'éco-
logie et la biologie de l'évolution. Enfin,

ils y ajoutent «la diversité des organismes»,
traduite par la systématique, où se mani-
festent, ou particulièrement des lois éta-
blies à partir des diverses représentations
marquées par les philosophes de la clas-
sification. En effet, pour ces deux auteurs,
la déclinaison du monde animal sous
toutes ses formes est essentielle, car,
comme ils le disent bien dans leur pré-
face : «La recherche biologique des der-
nières années démontre clairement que
l'écologie qui assure sa progression nais-
sent de plus en plus à l'interface des dif-
férentes disciplines.»

En résumé, R. Eckert et ses collègues
ont consacré un volume de 840 pages
environ à la physiologie comparée, la
développant un peu plus dans le que
R. Wehner et W. Gehring, dont l'ou-
vrage est pourtant d'un volume équiva-
lent. Ces derniers ont voulu donner une
représentation unitaire de la zoologie, en
montrant le monde animal dans toute sa
dimension, dans toute son amplitude,
sans «réductionnisme» aucun.

On regrettera cependant, pour l'ou-
vrage de R. Wehner et W. Gehring, la
présence de certains chapitres, tel celui qui
porte sur le contrôle hormonal, mais on
apprécie le fait que ceux-ci restent
équilibrés : l'historique des connaissances
est conservé, les généralités restent étayées
par de nombreux exemples, la biologie
est encore perçue comme une science
expérimentale (!) et enfin, des tableaux
récapitulatifs apportent encore leur lot d'in-
formations. Une remarque : le choix des
thèmes traités de manière plus succincte
serait mieux compris si une présentation
des auteurs – que j'ai retrouvée dans
l'œuvre originale – accompagnait l'ouvrage !

La concision des deux ouvrages donne
la mesure du remarquable développement
des connaissances. Sans en faire une ana-
lyse approfondie, il en ressort clairement
que l'approche de la biologie par leurs
auteurs est faite avec élégance : ni les uns,
ni les autres n'ont cédé à la tentation de
cloisonner les différents domaines des
Sciences de la Vie qui se voit de plus en
plus renforcée par une spécialisation quasi
inévitables. Ils ont pris avec brio cette ten-
dance à contre-pied, estompant la «dicho-
tomie» entre la biologie cellulaire et la
biologie animale en refusant de se placer
exclusivement dans l'un ou l'autre des deux
«camps». Ces deux livres viennent donc
à point nommé, alors que le développe-
ment de la «physiologie intégrative» devient,
du point de vue de l'effort de recherche
consenti, la priorité de l'«après génome».

Comparons à présent l'organisation des
deux ouvrages. Elle diffère quelque peu.

Dans celui de R. Eckert et de ses
collègues, chaque chapitre se termine
par un résumé, des questions de révision
et une liste des lectures conseillées, ce

qui est d'ailleurs l'une des caractéristiques des ouvrages anglo-saxons.

Dans les deux ouvrages, les termes fondamentaux et les expressions de spécialistes apparaissent en caractères gras et se retrouvent dans un glossaire situé à la fin de l'ouvrage. Dans celui de R. Wehner et W. Gehring, la définition des termes est complétée par leur étymologie, des renvois vers d'autres entrées permettent un recoupement de certains aspects des termes concernés. Un autre point intéressant à noter dans ce dernier : les termes anglais sont accompagnés de leur terme français correspondant, pour lequel une entrée est créée. Cette démarche est d'autant plus louable qu'une utilisation intempestive de la terminologie anglaise dans nos écrits français nuit profondément aux étudiants, qui finissent par perdre leurs repères dans les deux langues et ne se retrouvent plus dans la terminologie scientifique française. On retrouve cette démarche dans le corps de l'ouvrage : conservation des termes anglais, mais placés à la suite de la nomenclature française.

Les « notes du traducteur » sont relativement nombreuses dans l'ouvrage de R. Wehner et W. Gehring ; elles sont clairsemées dans celui de R. Eckert. Ces notes apportent un éclairage supplémentaire au texte original, ainsi qu'un complément d'informations jugé utile à une bonne compréhension d'un phénomène.

On doit saluer au passage la minutie, la rigueur avec laquelle l'ouvrage de R. Wehner et W. Gehring a été traduit : cette traduction se distingue parmi les différentes traductions publiées en langue française par le soin extrême apporté au style de la rédaction qui, sans être un style académique lourd, est d'une grande qualité, ainsi que par la fidélité au sens.

Ce compliment ne peut être adressé dans le cas de l'autre ouvrage, où l'esprit des auteurs n'est pas respecté. Il y a aussi des fautes typographiques, des fautes grammaticales, un manque de rigueur dans la présentation des légendes des figures, un manque de rigueur sur le plan de la langue française proprement dite, sur le plan de la terminologie scientifique, l'utilisation d'une terminologie scientifique obsolète, des gallicismes, des traductions hâtives... Il est vrai que certaines de ces erreurs n'altèrent en rien la bonne compréhension du texte ou des figures, mais elles nuisent par ailleurs.

Ce qui caractérise aussi les ouvrages anglo-saxons et que l'on repère immédiatement dans l'ouvrage de R. Eckert, c'est le travail graphique : le support iconographique est remarquable ; de nombreuses données sont illustrées par des figures au trait démonstratives, tout particulièrement lorsqu'elles sont présentées en perspective, ce qui est souvent le cas

et ce qui aide grandement à la compréhension du processus physiologique présenté. Pour l'ouvrage de R. Wehner et W. Gehring, un regret : la reproduction des figures originales n'est pas toujours très réussie et les images sont parfois bien trop foncées, sans doute parce que le livre a été imprimé sur du papier recyclé.

Enfin un appendice précieux et une bibliographie exhaustive accompagnent le glossaire dans les deux ouvrages. L'index est détaillé dans les deux cas, avec en plus dans l'ouvrage de R. Wehner et W. Gehring, des renvois vers des explications et des commentaires situés dans le corps de l'ouvrage et vers le glossaire.

On aura compris que ces deux ouvrages méritent une large diffusion auprès du public universitaire et étudiant (pour les étudiants, les enseignants, pour la préparation aux concours de recrutement CAPES et agrégation des sciences de la Vie et de la Terre, classes préparatoires à vocation scientifique, écoles d'été, etc.) et auprès des professeurs de l'enseignement secondaire.

Yvon LE MAHO

Sociologie

L'invention des populations : biologie, idéologie et politique

Sous la direction d'Hervé Le Bras.
Éditions Odile Jacob, 2000
(264 pages, 155 francs).

L'essai sur le principe de population de l'économiste britannique Thomas Robert Malthus (1766-1834) a changé la vision du monde de Charles Darwin (1809-1882) : ce dernier a été impressionné qu'il a « inventé » la sélection naturelle. Malthus s'est intéressé à la population humaine et il avait constaté qu'elle obéit à une croissance géométrique (de la forme $a, 2a, 4a, 8a, \dots$), tandis que les ressources suivent une croissance arithmétique (selon une loi de la forme $a, a+b, a+2b, a+3b, \dots$), bien plus lente. En raison de ce décalage, les savants ont jugé utile de quantifier les populations, de les analyser par des méthodes statistiques : c'est la démographie. Ils ont aussi cherché à gérer les biens : c'est l'économie.

Cependant, comme l'indique Hervé Le Bras, qui a dirigé les auteurs du livre *L'invention des populations*, la définition d'une population « représente un coup de

POUR LA SCIENCE

8, rue Férou 75278 PARIS CEDEX 06
Tél : 01-55-42-84-00
<http://www.pourlascience.com>

POUR LA SCIENCE Directeur de la rédaction : Philippe Boulanger, Françoise Cinotti (Rédactrice en chef), Philippe Pajot, Loïc Mangin, Claire Le Poulennec, François Savatier, Emmanuel Leclerc (Rédacteurs). Secrétariat de rédaction/Maquette : Annie Tacquenot, Pascale Thiollier, Céline Lapert, Laurence Cristoforo, Sylvie Sobelman. Site Internet : Cyril Lamotte. Marketing et Publicité : Stéphane Montouchet, assisté de Séverine Merviel et Christophe Vitiello. Direction financière : Anne Gusdorf. Direction du personnel : Jean-Benoît Boutry. Fabrication : Jérôme Jalabert, assisté de Géraldine Le Coz. Presse et communication : Sylvie Gillet, assistée de Nathalie Choplain. Directeur de la publication et Gérant : Olivier Brossollet. Conseiller scientifique : Hervé This.

Ont également collaboré à ce numéro : Pierre-Marie Baudonnière, Eric Braunberger, Sylvie Chevret, Claire Gaston, Yann Gauduel, Évelyne Heyer, Évelyne Host Platret, Pascal Jouxte, Marie-Thérèse Landouzy, Valérie Martin-Rolland, Claude Olivier, Christophe Pichon, Daniel Tacquenot.

SCIENTIFIC AMERICAN Editor: John Rennie. Board of editors: Michelle Press, Mark Alpert, Carol Ezzell, Wayt Gibbs, Justin Leutwyler, Madhusree Mukerjee, George Musser, Sasha Nemecek, Ricki Rusting, Sarah Simpson, Gary Stix, Philip Yam, Glenn Zorpette. Chairman Emeritus: John Hanley. Chairman: Rolf Gruebach. President and Chief Executive Officer: Gretchen Teichgraber. Vice-President: Chuck McCullough and Frances Newburg.

PUBLICITÉ France

Chef de Publicité : Jean-François Guillotin (jf.guillotin@pourlascience.fr), assisté d'Irène Limon
8 rue Férou 75278 Paris Cedex 06
Tél. : 01 55 42 84 28 ou 01 55 42 84 97
Télécopieur : 01 43 25 18 29

Étranger: 415 Madison Avenue, New York, N.Y. 10017 - Tél. (212) 754.02.62

SERVICE ABONNEMENTS

Anne-Claire Ternois : 01 55 42 84 04

SERVICE DE VENTE RÉSEAU NMPP

Christophe Vitiello : 01 55 42 84 81.

DIFFUSION DE LA BIBLIOTHÈQUE

POUR LA SCIENCE

Canada: Éditions : 945, avenue Beaumont, Montréal, Québec, H3N 1W3 Canada.

Suisse: Services : Chemin des châlets, 1979 Chavannes - 2 - Bâle

Belgique: La Caravelle : 303, rue du Pré-aux-oies - 1130 Bruxelles.

Autres pays: Éditions Belin : 8, rue Férou - 75278 Paris Cedex 06.

Toutes demandes d'autorisation de reproduire, pour le public français ou francophone, les textes, les photos, les dessins ou les documents contenus dans la revue «Pour la Science», dans la revue «Scientific American», dans les livres édités par «Pour la Science» doivent être adressées par écrit à «Pour la Science S.A.R.L.», 8, rue Férou, 75278 Paris Cedex 06.

© Pour la Science S.A.R.L.

Tous droits de reproduction, de traduction, d'adaptation et de représentation réservés pour tous les pays. La marque et le nom commercial «Scientific American» sont la propriété de John Wiley & Sons, Inc. Licence accordée par John Wiley & Sons, Inc.

En application de la loi du 11 mars 1957, il est interdit de reproduire intégralement ou partiellement la présente revue sans autorisation de l'éditeur ou du Centre français de l'exploitation du droit de copie (20, rue des Grands-Augustins - 75006 Paris).

Nos lecteurs trouveront des bulletins d'abonnement en pages 18A, 18B, 98A et 98B.

force...». Certes, on peut considérer, avec le dictionnaire, que c'est un «ensemble de personnes qui habitent un espace», mais les individus peuvent peupler la Terre, un continent, un pays, une région ou un village. Sont-ils mélangés ou triés selon des critères particuliers, ethniques, religieux, professionnels, scolaires, sanitaires ? En outre, les populations peuvent augmenter ou décroître. Montesquieu (1689-1755) avait insisté sur la dépopulation et écrit : «Cent mille causes peuvent agir, capables de les détruire et, à plus forte raison, d'augmenter ou de diminuer leur nombre.» Enfin, les individus se déplacent, les «populations bougent».

Tous ces phénomènes incitent à considérer les recensements comme des images temporelles et évolutives. Malgré la difficulté d'établir des modèles, le démographe américain Alfred Lotka a développé une théorie des populations stables, outil de base de la démographie mathématique (*Théorie analytique des associations biologiques*, 1939). Cette théorie ne s'applique qu'à des populations fermées dont la mortalité est stable sur 25 à 30 ans, la fécondité et la nuptialité constantes sur deux générations, et qui ne migrent pas ni ne reçoivent de migrants. Les conditions d'application de la théorie sont si strictes que son utilisation devient rapidement abusive. Pourtant l'Américain turcologue Justin McCarthy l'a employée pour décrire le génocide arménien de 1915. À cette époque, la mortalité n'était pas constante, la fécondité était mal connue et, en sous-estimant la population arménienne avant le massacre et en surestimant les survivants, J. McCarthy détermina un nombre de victimes bien inférieur au chiffre officiel de 1,5 million de morts. Sa manipulation relève du négationnisme démographique.

Lorsque les critères instaurés lors des recensements sont orchestrés par le démographe nazi Friedrich Burgdörfer (1890-1960), c'est l'escalade de l'horreur : obsédé par la baisse de la natalité, par le vieillissement de la population germanique et par la peur de l'immigration, il organise sa hiérarchie selon des choix clairs : le juif, l'«homme héréditairement vain», la «femme apte à enfanter». Afin de «préserver le patrimoine biologique», il veut inciter ceux «qui doivent faire vivre éternellement la race» à se reproduire. C'est l'incitation à «produire à vivre», le devoir de se perpétuer. Dans

le grand plan de l'extermination, il veut éliminer les juifs qui, selon lui, ont détourné

la sexualité et les valeurs familiales ; leurs influences spirituelles sont néfastes et destructrices, ajoute-t-il. Il envisage d'abord de regrouper les quelque 6,5 millions de juifs du monde sur l'île de Madagascar (1940), puis propose une

élimination spirituelle et physique, qui ira jusqu'à traquer le judaïsme «caché» (conversions ou mariages mixtes). Eugénisme et racisme ont trouvé là leurs illustrations parmi les plus ignobles.

Chercher à identifier une catégorie particulière d'individus, observés ensemble, peut avoir une motivation plus noble, mais dont la justification n'est pas évidente. Dans le domaine sanitaire, la notion de groupe à risque est associée à une forte prévalence : c'est par exemple le rapport entre le nombre de personnes infectées par le virus VIH dans une population sur un effectif total. Or, dans l'épidémie de SIDA, dès les années 1980, on a identifié des groupes à risque, devenus des groupes de transmission dans les années 1990. Dès le début de l'épidémie, la politique a été de placer les jeunes (pour simplifier, les moins de 25 ans) au cœur des préoccupations sanitaires et des campagnes de sensibilisation. Les jeunes, qui ne constituent pas un groupe de transmission selon les statistiques, sont considérés comme un groupe à risque. Or la notion de prévalence ne s'applique pas à la population des jeunes. Ils ne sont que potentiellement à risque, et l'on a transformé les risques en dangers réels. En effet, la mortalité de cette classe d'âge se situe nettement du côté des accidents de voiture et des suicides, et non du SIDA. Ainsi on a voulu rendre homogène une population qui est censée dans sa globalité prendre des risques, faire des expériences, se droguer et pratiquer le multipartenariat sexuel : c'est abusif et faux. Il n'est pas question de déclarer inutile la prévention qui a été faite, mais un message uniformisé à ce point risque de ne jamais atteindre la fraction marginale à laquelle il est vraiment destiné.

Ce livre propose beaucoup d'autres chapitres historiques, écrits avec objectivité, par exemple sur l'Algérie coloniale ou sur les travailleurs immigrés à la fin de la Grande Guerre. Il faut se réjouir de la qualité des analyses, de la façon qu'a le livre de «requadrer le contexte», selon la formule de Maurizio Gribaudi, mais est-on sûr que les recensements cachés d'aujourd'hui échappent aux dérives dénoncées dans l'ouvrage ? Les comités d'éthique ont de beaux jours devant eux, et le principe de précaution devrait s'appliquer au concept de population, afin de «qualifier» le contexte et non de le revisiter *a posteriori*.

La liste de tous les livres reçus en service de presse par la rédaction de Pour la Science est donnée sur notre site Internet, à l'adresse : <http://www.pourlascience.com>