la lettre de l'Académie des Sciences
n° 19 / printemps 2006

l'Outil informatique

la lettre de l'Académie des sciences
n° 19 / printemps 2006

Sommaire

Éditorial

Après la loi sur la recherche
Édouard Brézin
page 2

Dossier

L'outil informatique

Les réseaux de communication
François Baccelli
page 3

La course à la puissance de calcul
Pierre Leca
page 7

La sécurité informatique
Jean-Jacques Quisquater
page 9

Progrès du matériel et progrès du logiciel
Entretien avec Gérard Berry par Paul Caro
page 12

Gilles Kahn et l'INRIA
Bernard Larrouturou
page 16

Question d'actualité

Le cerveau olfactif
Jean-Marie Lledo
page 17

La vie de l'Académie

Grands Prix scientifiques de l'Institut de France
page 19

La recherche spatiale française
page 20

Sciences et pays en développement
page 20

Editorial

Après la loi sur la recherche

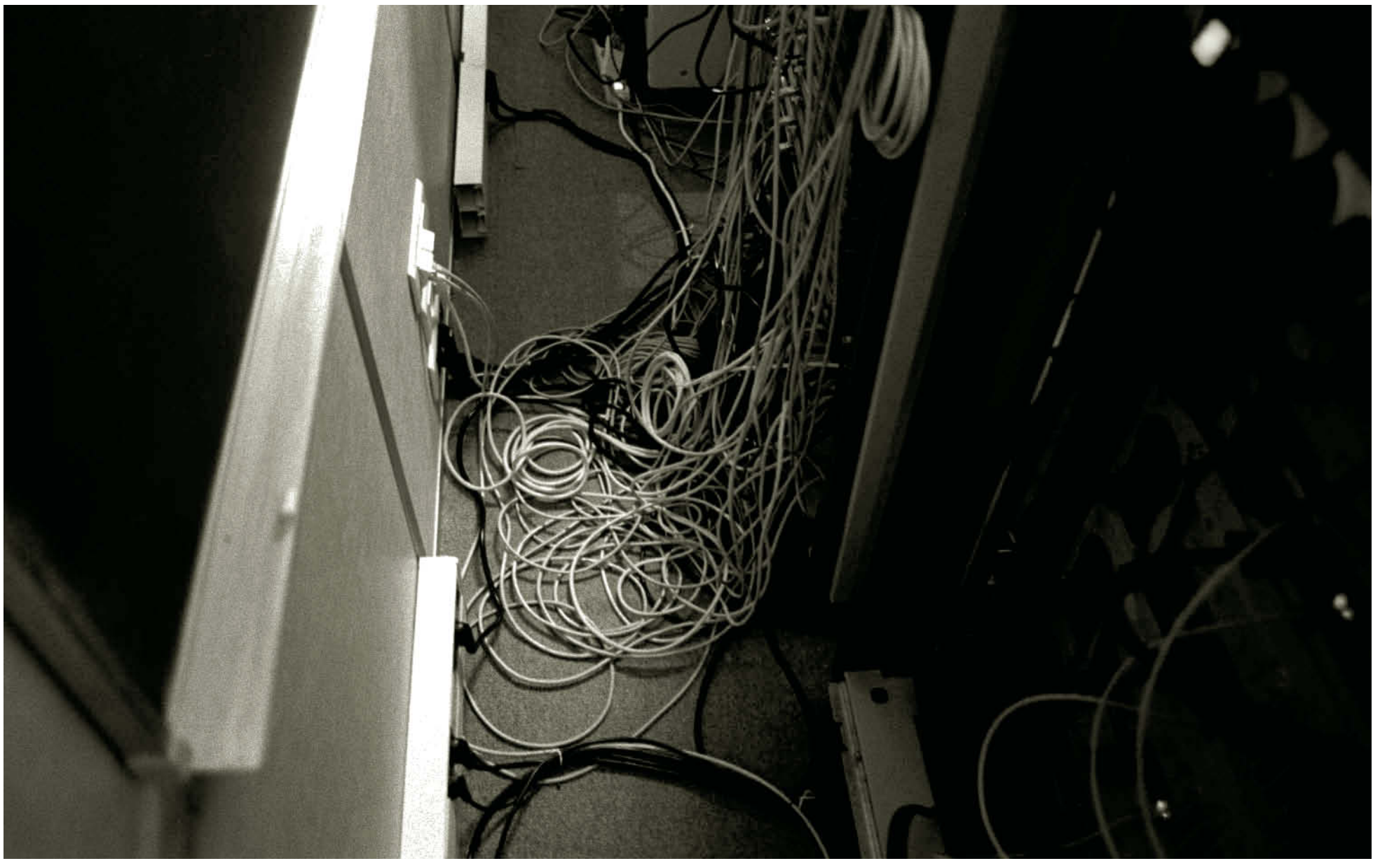
Une loi sur la recherche a été votée fin mars 2006, après plus de deux ans de tension entre chercheurs et pouvoirs publics. La disposition nouvelle la plus notable est la création officielle de l'Agence nationale de la recherche, qui a fonctionné en fait pendant l'année 2005 avant d'être pérennisée par cette loi. En ce qui concerne la recherche fondamentale, l'espoir est que cette agence donne des possibilités importantes à des jeunes équipes pour pouvoir lancer leurs propres projets, leur permettant ainsi de réaliser leurs idées et de devenir scientifiquement indépendantes. L'articulation avec les structures de recherche où sont implantées ces équipes va être délicate et il va nous falloir apprendre à utiliser ces nouveaux moyens sans détruire ce que peut avoir de positif la fédération d'équipes qui constituent un laboratoire. Nous avions plaidé pour un « overhead » destiné à cette fédération pour reconnaître le rôle de support matériel et intellectuel du laboratoire, et pour inciter ces derniers à attirer les meilleures équipes, celles dont les chances de succès auprès de l'ANR seront les plus grandes. Il semble qu'un abondement de 10 %, curieusement intitulé « préciput », soit bien prévu à cet effet dans les dispositifs futurs de l'ANR ; nous l'espérions au moins deux fois plus important. Rappelons qu'il est d'environ 50 % pour les grandes agences nord-américaines. Quoi qu'il en soit il faut qu'il apparaisse comme un supplément, et non un prélèvement sur les crédits attribués aux équipes retenues par l'ANR.
Autre élément annoncé, mais non inscrit dans la loi, la promesse d'une augmentation substantielle des allocations de recherche pour les doctorants, certes bienvenue puisque le rapport au SMIC de cette allocation n'avait cessé de baisser fortement depuis des années jusqu'à devenir inférieur à un. Ces mesures seront-elles suffisantes pour attirer à nouveau vers la recherche les meilleurs étudiants, souvent issus des écoles d'ingénieurs ou des facultés de médecine ? Quelques « bourses Descartes » devraient être créées à cet effet, mais je crains qu'elles ne suffisent pas à retourner la tendance, alors que c'est bien de notre capacité à attirer les meilleurs jeunes dans nos laboratoires publics et privés que dépend l'avenir de la science de notre pays et sans doute son avenir tout court. Nos actions afin d'accueillir des étudiants et des post-docs du monde entier sont évidemment essentielles, mais les mesures annoncées qui prétendent vouloir faire venir de bons esprits, tout en leur refusant *a priori* le droit à prolonger leur travail en France à l'issue de leur formation, ne sont certainement pas adaptées à la réussite dans la compétition internationale pour la matière grise qui régit aujourd'hui notre monde.
En revanche, le Gouvernement n'a pas voulu se pencher sur l'organisation des universités, sujet certes potentiellement explosif, mais les vraies réformes de notre système de recherche passent évidemment par des universités de grande qualité, dotées d'une organisation adaptée et de la capacité à conduire une vraie politique de recherche. L'Académie avait recommandé, dans son avis de 2005 sur la recherche, que l'on s'engage progressivement vers davantage de recrutements initiaux dans les universités, en gardant précieusement les postes dans les organismes tels que le CNRS ou l'INSERM, pour des périodes définies où la nécessité, correctement évaluée, de la recherche demanderait de pouvoir disposer de plus de temps. De tels mécanismes transfèreraient à terme la responsabilité essentielle des recrutements initiaux vers l'université et il est clair que les procédures actuelles devraient être notablement améliorées avant que cela puisse être envisagé avec sérénité. Aucun grand pays ne recrute ses professeurs comme nous le faisons. D'autre part le sujet de l'orientation initiale des étudiants, baptisée sélection pour la diaboliser, reste tabou, alors que chacun peut constater que les seules filières de notre enseignement supérieur qui conduisent à un emploi sont bien celles qui ont été autorisées à exiger des étudiants la démonstration de leur capacité à suivre le cursus choisi. Notre Académie dont les missions, inscrites dans nos statuts, concernent la recherche ainsi que l'enseignement des sciences, doit donc impérativement s'exprimer sur ces sujets ■

par **Édouard Brézin** [1]

1 Président de l'Académie des sciences, professeur à l'université Pierre et Marie Curie

L'outil informatique

dossier préparé par Olivier Pironneau[1] et Jacques Stern[2]



Université Pierre et Marie Curie, laboratoire d'Analyse numérique.

Les réseaux de communication

par **François Baccelli**[3]

De par les réalisations qui en ont découlé ces dix ou vingt dernières années, notamment dans le domaine des réseaux de communication, les sciences de l'information ont révolutionné notre environnement. Le but de cet article est de donner quelques éclairages sur les fondements scientifiques des transformations qui ont eu lieu dans ce domaine et d'esquisser certaines des évolutions possibles au vu des recherches en cours.

Nous aborderons d'abord le développement de l'Internet en décrivant les choix d'architecture et de protocoles qui ont rendu possible la mise en place très rapide d'un réseau global permettant de transférer l'information presque en tout lieu, quasi instantanément et à très faible coût.

Nous décrirons ensuite les bases du développement explosif de l'accès sans fil, avec maintenant plus d'un milliard d'utilisateurs du téléphone cellulaire. Une floraison de nouvelles technologies et architectures de réseaux d'accès sans fil tend à rendre la couverture du réseau universelle : réseaux cellulaires, Wifi, satellites, etc. Nous montrerons que la théorie de l'information y joue un rôle fondamental.

Dans le domaine du logiciel, nous décrirons brièvement les nouveaux modes d'organisation, d'archivage, de réplication et de transformation de l'information qui ont rendu possible tout un ensemble de nouveaux services. Nous nous concentrerons enfin sur la révolution de l'accès à l'information, entièrement transformé par l'hypertexte et les moteurs de recherche, dont nous expliquerons le fonctionnement.

1 Membre de l'Académie des sciences, professeur à l'université Pierre et Marie Curie.
2 Directeur du département d'informatique de l'École normale supérieure.
3 Membre de l'Académie des sciences, directeur de recherche à l'Institut national de recherche en informatique et en automatique.

L'Internet

Qu'est ce qu'un réseau de communication ?

De manière générale, un réseau peut être vu comme un graphe dont les nœuds sont de deux types, terminaux et commutateurs, et dont les arcs sont des liens. La fonction première d'un réseau est l'écoulement de communications entre les terminaux au travers des commutateurs. Dans sa version la plus simple, un commutateur est un calculateur ayant autant de liens d'entrée que de liens de sortie, ainsi qu'un état qu'on peut voir comme une bijection de l'ensemble des liens d'entrée vers celui des liens de sortie. La *commutation* a pour fonction d'associer à tout ensemble de demandes de communications une suite d'états des commutateurs, permettant d'écouler un sous-ensemble aussi grand que possible de ces demandes.

Dans le réseau Internet actuel, on distingue le *cœur* et l'*accès*. Le *cœur* est formé de commutateurs ou *routeurs* connectés par des liens à haut-débit, typiquement de l'ordre du Gigabit par seconde. L'accès concentre le trafic généré par les terminaux (stations de travail, portables, etc.) vers le cœur sur des lignes à plus faible débit, DSL ou radio. Cela se fait par des commutateurs de concentration appelés multiplexeurs.

Les premiers grands réseaux de communication

Avant l'Internet, l'archétype du grand réseau était le réseau téléphonique. Ce réseau est fondé sur la commutation de circuits : pour chaque communication, on y établit un circuit composé d'une suite de canaux à débit fixe et réservés à cette communication. Une fois le circuit établi, on peut garantir le débit requis pour le maintien de la communication. Ce débit dépend de la nature de l'application, par exemple voix ou vidéo. Ce type de commutation pose deux problèmes qui expliquent l'évolution vers d'autres modes :

- Avant toute communication, il faut établir le circuit. Ceci est algorithmiquement lourd lorsqu'il comporte beaucoup de liens, éventuellement localisés chez plusieurs opérateurs.
- Il faut aussi refuser l'accès lorsque certains des canaux requis ne sont pas disponibles : comme la demande est très variable dans le temps et l'espace, tous les canaux entre deux nœuds peuvent être occupés à un moment donné. Ce type de contrôle d'admission exige une connaissance assez précise de l'état du réseau à tout moment. Il requiert donc une gestion plutôt centralisée qui ne convient pas pour les très grands réseaux.
- le transfert des données est particulièrement inefficace. En effet, la plupart des applications de transfert de données donnent lieu à des trafics extrêmement sporadiques et donc mal adaptés à toute réservation de bande passante. Par exemple, la consultation de documents sur des pages de la toile génère une succession de transferts de fichiers séparés par des périodes d'inactivité.

Le fonctionnement de l'Internet

L'Internet fonctionne selon un mode de *commutation de paquets*. Un paquet est une suite de bits codant une partie de l'information à transporter et munie d'un en-tête. Dans ce mode, tout type d'information (voix, données, vidéo) est "paquétisé" et envoyé par la source dans le réseau sans aucune forme de réservation.

Ce mode de commutation s'implémente de manière efficace avec deux types d'algorithmes : *routage* et *contrôle de congestion*. Le routage remplace l'établissement du circuit. Il consiste en un aiguillage autonome des paquets de l'émetteur au destinataire en passant par un nombre de routeurs le plus petit possible. Le contrôle de congestion remplace le contrôle d'admission et permet d'éviter les engorgements. Les deux critères essentiels pour le choix de tout algorithme dans ce cadre sont son caractère décentralisé et son extensibilité, c'est-à-dire sa capacité à fonctionner pour des réseaux de très grande taille.

Le routage IP (*Internet Protocol*)

Le routage IP est fondé sur l'allocation d'une adresse IP à chaque machine et sur une organisation hiérarchique de ces adresses en *systèmes autonomes*. Dans IP V6, en cours de déploiement, on disposera de 2^{128} adresses et il y a actuellement plusieurs dizaines de milliers de systèmes autonomes.

Les routes optimales entre deux points du graphe sont déterminées au moyen des équations de la *programmation dynamique*, dont voici un exemple d'implémentation distribuée. Chaque routeur maintient une table de routage calculée par échange de messages avec ses voisins ; pour toute destination le routeur demande à chacun de ses voisins sa distance à cette destination et il choisit comme "étape suivante" pour cette destination celui des voisins qui déclare la distance la plus courte ; en faisant fonctionner cet algorithme de manière récursive, les tables des divers routeurs convergent vers un point fixe qui donne le routage optimal. Si l'en-tête de chaque paquet contient l'adresse IP de sa destination, tout routeur est bien en mesure d'orienter les paquets de manière autonome par simple consultation des tables de routage. Cette description est très simple et un peu idéalisée. Il existe notamment une hiérarchie dans les systèmes autonomes, des mécanismes de routage spécifiques entre systèmes autonomes, etc.

Dans les architectures actuelles de type *store and forward*, un routeur reçoit les paquets sur les divers liens d'entrée, les stocke dans une mémoire tampon pour absorber les pointes de trafic, lit leurs en-têtes, consulte ses tables de routage, et commute les paquets vers les bons liens de sortie. Les algorithmes de calcul de suite d'états-bijections adaptés aux contenus des mémoires tampons sont particulièrement intéressants. Ils doivent tourner très vite : des routeurs commutant un Térabit par seconde sont d'actualité. L'architecture et l'algorithmique des routeurs font l'objet de recherches très poussées dans lesquelles s'impliquent notamment les meilleures universités de la côte ouest des États-Unis. Grâce à Cisco, l'industrie US est en position dominante dans ce domaine. Les multiplexeurs d'accès pour DSL ont des architectures et des fonctions plus complexes. Alcatel est très bien placé en ce qui concerne ce type de matériel.

Le contrôle de congestion

Comme il n'y a pas de contrôle d'admission, le réseau peut être sujet à des phases de congestion conduisant à son effondrement. Il faut donc un contrôle de congestion. Celui utilisé dans l'Internet est le protocole TCP (*Transmission Control Protocol*) sur IP, aussi noté TCP/IP, initialement proposé par R. Kahn et V. Cerf. Il s'agit d'un mécanisme de contrôle adaptatif et décentralisé du débit de l'information émise par toute source. Dans la version utilisée actuellement, la congestion est évitée par un algorithme qui augmente linéairement le débit d'émission au cours du temps tant qu'il ne se produit pas d'engorgement, l'idée de base étant d'essayer d'utiliser la bande passante disponible. En cas d'engorgement, l'algorithme réduit de moitié le débit d'émission. Chaque paquet fait l'objet d'un accusé de réception par la destination, dans un but premier de fiabilité du transport. La source sait qu'il y a engorgement lorsque certains de ses paquets sont perdus. Elle évalue donc expérimentalement l'état de congestion du réseau et s'y adapte à tout instant. Le protocole est installé sur chaque terminal et contrôle plus de 90 % du trafic à chaque instant de façon entièrement décentralisée.

Le développement de l'Internet

On peut maintenant répondre à une question souvent posée : comment un réseau comme l'Internet a-t-il pu croître si rapidement depuis sa création dans les années 70 et supplanter les réseaux d'opérateurs classiques ? La réponse réside dans l'extensibilité et la décentralisation : en connectant deux réseaux IP contrôlés par TCP et en laissant interagir leurs tables de routage, on obtient tout simplement un nouveau réseau IP contrôlé par TCP. Ce réseau fonctionne en corollaire de l'extensibilité des protocoles. Ainsi, en s'inspirant de situations où les structures ultimes sont fondées sur des interactions locales et très simples, les concepteurs de l'Internet sont parvenus à créer des réseaux ayant ce pouvoir de composition qui seul permet une croissance rapide.

Quelques axes de recherche actuels

Mathématiques de l'Internet

Une des conséquences surprenantes de ces choix d'architecture et de protocoles est l'impossibilité pour quiconque de connaître et *a fortiori* de contrôler le réseau dans son ensemble. Même pour un grand opérateur, le découpage en systèmes autonomes et l'extrême distribution des mécanismes mis en œuvre rendent quasi-impossible, l'observation complète de l'évolution du réseau. Au niveau macroscopique, la topologie du graphe d'interconnexion des routeurs ou des systèmes autonomes varie constamment. Il en va de même au niveau microscopique : le nombre et la nature des applications présentes sont extrêmement variables.

La modélisation mathématique des grands réseaux est un des seuls moyens actuellement disponibles pour l'inférence statistique de leurs propriétés au vu d'observations partielles. Il peut paraître étrange d'avoir à modéliser un système conçu par l'homme, mais il faut garder à l'esprit que la taille de ce système devient comparable à celle de certains systèmes de la physique statistique : la mole n'est pas un ordre de grandeur inapproprié quand on cherche à évaluer le nombre de bits transférés chaque année par le réseau.

La modélisation mathématique de TCP présente de nombreuses difficultés, en raison du caractère décentralisé, sto-

chastique, non linéaire et complexe de la situation. L'encombrement et les pertes évoluent aléatoirement ; le réseau est très étendu, ce qui implique des interactions entre des routeurs distants. L'élaboration de modèles mathématiques intégrant tous ces éléments permettra de prédire et contrôler les variations aléatoires de la qualité de service offerte par le réseau à ses utilisateurs. En effet, si on pouvait garantir un débit fixe à toute communication dans le cadre du contrôle d'admission, le débit Internet offert à un utilisateur donné fluctue fortement. Ceci est dû au partage dynamique et imprévisible de la bande passante opéré par TCP entre les utilisateurs utilisant le réseau à un instant donné.

Par ailleurs, la preuve d'extensibilité est une étape importante pour faire adopter puis finalement imposer tout nouveau protocole.

À première vue, les outils mathématiques permettant cette modélisation sont proches de ceux utilisés dans le cadre de l'analyse des systèmes de particules en physique statistique : il s'agit de l'interaction d'un très grand nombre de systèmes stochastiques simples, comme par exemple des applications de transfert de données interagissant *via* TCP. Cependant, alors que les structures étudiées en physique ou en chimie sont plutôt contrôlées de l'extérieur, celles de l'Internet sont en un sens contrôlables de l'intérieur : si on ne peut pas changer la manière qu'ont les particules élémentaires d'interagir entre elles, on peut changer les protocoles du routage IP ou TCP et donc le mode d'interaction des flots de bits associés aux diverses applications. Le contrôle adaptatif de ces grands systèmes d'interaction est donc devenu un axe de recherche majeur, avec comme objectif l'optimisation systématique des protocoles.

Une autre science en cours de développement est celle de l'observation et de la métrologie du réseau ou plus généralement de l'expérimentation sur le réseau. Un exemple intéressant est celui des méthodes de sondes actives : on envoie des trains de paquets en un ou plusieurs points d'entrée du réseau et on collecte ces paquets-sondes en un ou plusieurs points de sortie. Par exemple, la *tomographie Internet* vise à développer des méthodes permettant d'inférer les propriétés de l'état du cœur de réseau à partir des informations de topologie et de délais observées par des sondes émises et collectées par des terminaux situés à sa périphérie. De nombreuses questions de nature statistique se posent alors : quelles variables d'état du réseau sont observables ? Quelles méthodes d'inférence doit-on mettre en œuvre pour définir de bons estimateurs de ces quantités ?

Les réseaux applicatifs

Pour tout transfert de données point à point (entre deux terminaux), l'Internet offre un mode de transport fiable, avec un partage dynamique assez équitable des ressources en bande passante. Un des axes de recherche actuellement les plus actifs dans le domaine de l'algorithmique concerne la conception de réseaux applicatifs extensibles fondés sur ces mécanismes de transport point à point. Un réseau applicatif (*overlay network*) est une structure d'interaction entre un grand nombre de partenaires visant à l'accomplissement d'une tâche commune à laquelle chacun apporte son concours, soit comme relais, soit comme contributeur. Citons des exemples simples de réseaux applicatifs : la diffusion d'un fichier d'une source vers un grand nombre de destinataires, comme la diffusion de la dernière version de Linux ; un échange auto organisé de fichiers de très grandes tailles entre un grand nombre d'utilisateurs par exemple dans le contexte des réseaux dits "pair à pair" ; un calcul distribué sur réseau qui vise à exploiter la gigantesque capacité de calcul disponible à tout instant dans l'Internet, comme SETI@home. Dans ces contextes, le but est de nouveau de trouver des structures pouvant croître de manière autonome et très rapide. Décentralisation et extensibilité sont donc aussi les principes guidant la recherche en algorithmique dans ce domaine en forte expansion : comment organiser et accomplir une tâche commune sans "chef d'orchestre" ? Comment prendre en compte la dynamique de la structure en fonction de l'arrivée ou du départ des partenaires ? Comment s'assurer que ces structures restent respectueuses des transferts point à point en ce qui concerne le partage équitable de la bande passante ?

Les réseaux d'accès sans fil

La révolution de l'Internet en cache une autre qui est celle de la diversification des modes d'accès à haut-débit. Nous nous concentrerons ici sur les accès radio cellulaire et Wifi qui ont connu un développement particulièrement rapide ces dernières années. Parmi les autres modes d'accès importants, il convient de citer l'accès DSL, l'accès satellite et l'accès par fibre optique.

Une technique dominante : le CDMA

L'accès radio majoritaire est celui offert par les réseaux cellulaires qui peuvent dès aujourd'hui assurer le transfert de données multimédia en plus des fonctions de téléphonie. La norme CDMA ou *Code Division Multiple Access*, est la base des réseaux UMTS (Système Universel de Communications Mobiles) en cours de déploiement. Elle est fondée sur une technique d'*étalement de spectre*.

Dans les communications radio, les interférences entre les diverses communications en cours entre des mobiles et la station de base engendrent les principales limitations. Une première idée naturelle pour limiter ces interférences est le FDMA ou *Frequency Division Multiple Access*, qui consiste en une division du spectre radio disponible en bandes sans recouvrement et une utilisation de bandes différentes pour des canaux géographiquement proches. L'idée de base du CDMA est, qu'au contraire, il vaut mieux que chaque émetteur utilise tout le spectre disponible et considère les autres transmissions en cours comme du bruit. Comme le trafic de données est très sporadique, une transmission peut avoir peu de compétiteurs à un instant donné. Dans le CDMA, un bruit faible conduit immédiatement à un débit élevé, alors que la pré-allocation de bandes de fréquence aux communications rend moins facile un tel partage dynamique du spectre radio.

Mais comment peut-on considérer les autres transmissions comme du bruit ? Dans le cas du CDMA, chaque émetteur emploie une *fonction de signature* qui lui est propre et par laquelle il module chaque bit d'information envoyé. Un bit est répété plusieurs fois et modulé par la signature en question. Le récepteur la superposition de tous les signaux émis, atténués et réfléchis par les obstacles, et estime le signal qui lui est envoyé par un émetteur donné en multipliant à son tour la superposition des signaux reçus par la signature de l'émetteur. On peut imaginer une salle où plusieurs paires d'interlocuteurs parlent à voix haute, et où chaque paire utilise une langue différente. On utilise la théorie de l'information pour évaluer la capacité d'un tel canal en fonction du "rapport signal sur bruit", (rapport de la puissance reçue de l'émetteur sur celle de l'interférence effective, c'est-à-dire de la somme après filtrage de tous les autres signaux reçus). Il faut faire appel à la théorie des matrices aléatoires pour évaluer cette capacité de manière précise.

Le contrôle de ce type de réseau pose de nombreuses difficultés tant mathématiques qu'algorithmiques. Il y a d'abord la question de la faisabilité : un ensemble donné de canaux en interférence peut ne pas être réalisable si on exige un débit minimal pour chaque canal. Dans ce cas, il n'existe aucun réglage des puissances qui permette d'atteindre simultanément les débits demandés. Lorsqu'une configuration est faisable, il y a ensuite les questions de contrôle proprement dit. La ressource radio étant très limitée, il est souhaitable de mettre en œuvre divers mécanismes qui visent à faire en sorte que cette ressource soit bien utilisée et notamment bien partagée. Pour la voix, on utilise un contrôle de puissance, qu'on peut décrire comme suit. Considérons un ensemble de mobiles réussissant à transmettre de la voix vers leurs stations de base respectives ; le besoin est d'approximativement 12 Kilobits par seconde ; supposons qu'un mobile, qui dispose d'un canal de capacité suffisante, décide de baisser sa puissance d'émission jusqu'au point lui donnant un rapport signal sur bruit correspondant à ce besoin ; l'interférence créée par ce mobile sur les autres canaux va baisser ; chacun des autres mobiles pourra alors baisser à son tour sa puissance d'émission ; on entre alors dans un cercle vertueux qui fait converger les puissances vers un point d'équilibre optimal, au sens de Pareto, où chaque mobile a exactement le débit qu'il lui faut. Un réseau contrôlé de cette manière peut admettre plus de mobiles qu'un réseau non contrôlé. Il convient de nouveau de mettre ce contrôle en œuvre au moyen d'algorithmes distribués, qui s'exécutent avec aussi peu d'échanges d'information que possible.

Les réseaux d'accès sans fil du futur

Dans le domaine des réseaux cellulaires, les travaux de recherche actuels portent sur des modes d'accès que le CDMA, notamment l'OFDM qui est une sorte de FDMA dynamique. D'autres types d'accès connaissent aussi une très forte expansion.

Le développement de l'accès Wifi (norme IEEE 802.11) est sans doute en partie lié au fait que les réseaux à base de Wifi ne sont pas nécessairement des réseaux d'opérateurs, mais plutôt des réseaux de communautés nécessitant peu d'infrastructure : campus, quartier, secours, opérations militaires etc. Ils doivent être adaptés à la mobilité et fondés sur une coopération auto-organisée des divers éléments du réseau : points d'accès, téléphones portables, stations de travail mobiles, agendas

électroniques etc. Nous ne parlerons pas ici des questions de transmission radio, ni du codage, mais plutôt des questions d'auto-organisation qui se posent dans ce contexte. Par auto-organisation, on entend une algorithmique où toutes les décisions concernant l'accès au canal radio partagé, le routage et le contrôle de congestion sont aussi distribués que possible, c'est-à-dire centrées autant que faire se peut sur l'utilisateur final.
Il y a plusieurs types de problèmes : un des plus importants est celui du routage en plusieurs sauts ; il s'agit d'utiliser, de façon opportuniste, la présence des autres éléments du réseau comme relais pour transmettre de l'information vers une destination donnée. Un autre est celui des algorithmes de choix d'un point d'accès parmi plusieurs en fonction de leurs charges respectives. De manière plus générale, les recherches sur l'auto-organisation visent à définir des architectures d'accès globales utilisant au mieux les divers types d'accès radio disponibles à tout instant et en un lieu donné : cellulaire, Wifi, Wimax etc.

Les architectures envisagées dans les réseaux mobiles "ad hoc" sont radicalement nouvelles. Elles sont fondées sur un effacement de la différence entre terminal et routeur : chaque nœud du réseau est l'un et l'autre à la fois ; on peut aussi les voir comme fondées sur la disparition des routeurs et l'appropriation de leurs fonctions par les terminaux. Certains chercheurs voient ces architectures comme une étape inéluctable dans l'évolution des grands réseaux ; une autre étape possible, et même probable à court terme, est celle correspondant à une hybridation entre l'architecture actuelle du cœur de l'Internet et celle envisagée ci-dessus.

Accès sans fil et théorie de l'information

Les modes d'accès radio évoqués ci-dessus peuvent être considérés comme des réalisations de grande ampleur fondées sur la *théorie de l'information* et du codage, développé il y a plus d'un demi-siècle par C. Shannon aux Bell Labs.
Cette théorie est déterminante dans des domaines très divers : l'enregistrement magnétique de haute densité utilisé dans les disques durs ou encore les méthodes de compression pour l'audio et la vidéo telles que MPEG, JPEG et MP3. C'est aussi le codage qui permet d'exploiter la capacité réelle d'une ligne téléphonique. Alors que le fil de cuivre utilisé pour ces lignes était initialement prévu pour un signal vocal qui correspond à une dizaine de Kilobits par seconde, un modem DSL permet d'écouler sur cette ligne, un tel signal, et en plus, quelques dizaines de Mégabits par seconde de trafic de vidéo et de données sur la voie descendante ainsi que quelques Mégabits sur la voie montante. Les recherches sur la théorie de l'information sont très actives dans le cadre des communications radio. La compétition internationale est particulièrement vive sur *la théorie de l'information multi-utilisateurs* qui étudie les problèmes liés au codage conjoint du signal sur plusieurs antennes et au décodage conjoint par plusieurs récepteurs. DARPA, l'agence de recherche du ministère de la défense des USA, vient par ailleurs de lancer un grand programme de recherche visant à déterminer la capacité des réseaux mobiles *ad hoc* évoqués ci-dessus.
Tous ces exemples font partie d'un immense effort international, tant académique qu'industriel. Le but est de calculer explicitement les limites dont l'existence est prouvée par la théorie de l'information et de concevoir des dispositifs d'une efficience la plus proche possible de ces limites.

Représentation, stockage et recherche des données

Représentation et stockage des données

Chacun connaît le principe de l'hypertexte (HTML) et de l'organisation de la toile en un gigantesque graphe extrêmement dynamique : les nœuds de ce graphe sont les pages, caractérisées par leur URL (*uni/oTm resource locator*) les arcs sont les références HTML. Ces pages HTML sont initialement stockées sur des serveurs puis répliquées dans des mémoires caches. La réplication permet un accès rapide aux pages les plus populaires : l'existence de plusieurs copies plus ou moins uniformément distribuées sur la planète permet notamment de paralléliser la lecture. Cette réplication augmente aussi la pérennité des documents.

Notons au passage que cette question de la pérennité des pages est particulièrement importante dans la mesure où une partie de l'information, même scientifique, n'est plus transmise et conservée que sous cette forme. Les textes grecs ou latins nous sont parvenus ou non en fonction de la pérennité du support sur lequel ils avaient été couchés. Qu'adviendra-t-il des données et des textes actuellement archivés sur la toile ?

Les moteurs de recherche : l'exemple de Google

Considérons un mot du langage courant ou technique. Imaginons que l'on donne à un robot la tâche de déterminer, dans l'ensemble des pages de la toile qui contiennent ce mot, celles qu'un lecteur devrait consulter en premier. En d'autres termes, comment classer automatiquement (sans faire appel à un expert du sujet) les pages en question, par ordre de pertinence décroissante ? Les premiers moteurs de recherche ont proposé des réponses assez différentes à cette question, fondées sur diverses définitions de la pertinence. Dans la pratique, Google a supplanté ses concurrents. Voyons comment ce logiciel procède.
Il envoie un robot qui parcourt les pages de la toile selon l'algorithme probabiliste suivant : si le robot est sur une page contenant le mot en question, il est envoyé ensuite au hasard vers l'une des pages qui contiennent aussi ce mot et qui sont citées en références hypertexte sur la page de départ ; appliquant ce principe de page en page, le robot électronique effectue un parcours aléatoire sur le sous-ensemble de la toile concernant ce mot. Supposons pour simplifier que le graphe des pages en question est tel qu'on peut trouver une suite de références qui nous conduise de toute page contenant ce mot à toute autre ; le théorème ergodique des chaînes de Markov montre alors que si le parcours aléatoire est assez long, la fréquence empirique de visite de chaque page converge vers une limite déterministe qui ne dépend pas de la position initiale du robot. Google classe les pages de la manière suivante : une page qui a une fréquence empirique plus grande qu'une autre est jugée plus pertinente.

Le parcours de la toile ne pouvant être fait en temps réel, les robots de Google l'explorent en temps continu pour chaque mot de chaque langue et calculent sans cesse les classements associés à chacun. Pour les recherches comportant plusieurs mots clés, Google n'a pas à lancer un robot pour chaque assemblage de mots ; la combinatoire rendrait ceci impossible ; une fois connu le classement des pages contenant un premier mot et celui de celles en contenant un deuxième, des techniques de bases de données permettent de trouver les pages contenant l'un et l'autre mot ainsi que d'établir le classement de ces dernières. Une autre composante importante du procédé est donc le maintien et la mise à jour de bases de données archivant, pour chaque mot, le classement des pages les contenant.

Quelques réflexions sur l'accès à la connaissance

On justifie l'approche de Google en postulant que la pertinence d'une page est proportionnelle à la somme pondérée des pertinences des pages qui pointent vers elle ; cette dernière définition, auto-référentielle à première vue, est cependant justifiable dans le contexte de la théorie spectrale des matrices : elle revient à dire que la pertinence correspond au vecteur propre, dit de Perron-Frobenius, de la matrice de transition de la chaîne de Markov décrivant la dynamique du robot et codant la forme du graphe des références ; l'évaluation des fréquences empiriques par le robot est juste un moyen simple d'estimer ce vecteur propre. Cette propriété spectrale est une propriété globale du graphe des références, fondée sur une définition acceptable de la pertinence. C'est sans doute pourquoi Google est aussi efficace et quasi universellement utilisé.

Quand on considère le nombre des transactions exécutées par Google chaque jour, on se dit à juste titre que jamais le théorème de Perron-Frobenius (sur le vecteur propre associé au rayon spectral d'une matrice positive), ou encore le théorème ergoqique sur les chaînes de Markov, n'auront autant servi la connaissance humaine !

Les changements radicaux dans l'organisation générale des réseaux de communication qui sont à la base des grandes transformations technologiques de ces dernières années s'appuient sur des travaux scientifiques novateurs, pluridisciplinaires, venant de domaines aussi divers que l'algorithmique, les probabilités, le contrôle, la théorie de l'information, la théorie des graphes, etc. Notons pour conclure que l'effort international de recherche dans ces domaines est considérable, tout particulièrement en Amérique du Nord et en Asie, tant dans les universités que dans l'industrie. La plupart des grands industriels des technologies de l'information ont créé des centres de recherche propres ou immergés dans le tissu académique. Les grandes universités d'Amérique du Nord ont notamment réussi à constituer des groupes pluridisciplinaires d'une créativité et d'un niveau scientifique exceptionnels sur ces questions ■

Université Pierre et Marie Curie, laboratoire d'Analyse numérique.

La course à la puissance de calcul

par **Pierre Leca** [1]

Le calcul intensif et ses évolutions

Bien que le calcul scientifique représente, par le nombre des utilisateurs, une toute petite partie de l'informatique, il joue un rôle phare, comparable à la compétition automobile, dans le développement de l'informatique.

L'arrivée des calculateurs CRAY dans les années 70 a marqué l'émergence d'une nouvelle discipline : le calcul intensif. Depuis lors celui-ci a pris une place essentielle dans la plupart des secteurs de la recherche et de l'industrie et est considéré comme un secteur capital voire stratégique par certains pays qui conduisent des politiques à long terme pour en développer la maîtrise.

En tout premier lieu c'est le cas des États-Unis où, reconnu à nouveau comme un des 3 domaines prioritaires de la recherche en 2006 (cf. « discours sur l'état de l'Union du Président Bush »), il fait l'objet d'un consensus bi-partisan puisque qu'il s'agit d'un domaine prioritaire consigné par la loi depuis 1991.

Le Japon, dont le résultat le plus éclatant est la mise en service dès 2002 de l'*Earth Simulator*, s'est fixé un nouveau défi en annonçant au milieu de l'année dernière l'entrée dans la course au Petaflops avec un projet de calculateur de 10 Pflops pour 2011 à destination des bio et nano technologies.

Un nouveau venu a rejoint ce club, la Chine, qui a mis en place un programme visant à construire un calculateur d'une puissance de 1 Pflop dans le cadre du 11ème plan quinquennal.

Afin de mieux comprendre ce secteur, la liste des 500 plus puissants calculateurs au monde (cf. www.top500.org), remise à jour tous les 6 mois depuis 1993, constitue un bon indicateur de certaines de ses évolutions et tendances. L'Europe, à égalité avec l'Asie, ne possède qu'environ 20 % de la puissance de calcul du Top 500. L'évolution des catégories d'utilisateurs est par contre tout à fait significative avec une augmentation constante du secteur industriel dont la part est passée de 20 % en 1993 à près de 50 % en 2005. Mais le plus frappant est la croissance exponentielle, non démentie depuis l'origine, de la puissance de calcul que représente l'ensemble de ces 500 calculateurs. En effet la puissance installée s'est accrue d'un facteur 10 tous les 4 ans, Cette croissance est supérieure d'un facteur 16 à ce que l'on pouvait attendre des conséquences de la Loi de Moore et a été obtenue par l'architecture parallèle, c'est-

1 Chef du Service Numérique, Environnement et Constantes, du département CSA, CEA Bruyères le Chatel

à-dire l'interconnexion et la coordination de plusieurs processeurs au sein du calculateur. Mais en se concentrant uniquement sur la puissance de calcul le Top 500 ne rend pas compte de certaines évolutions majeures du calcul intensif, notamment la tendance grandissante à intégrer puissance de calcul et traitement du flot de données. Deux nouveaux facteurs sont en effet à l'œuvre depuis quelques années :

- La réduction du coût des médias de stockage de l'information : songez qu'un disque de 1 Toctet coûte aujourd'hui moins de 1 000 euros, et permet de mettre en place les moyens de stocker plusieurs Petaoctets (pour fixer un ordre de grandeur, la numérisation de la très Grande Bibliothèque représenterait environ 30 à 40 Toctets seulement),
- L'évolution des applications du calcul intensif : celui-ci est devenu un grand producteur, les simulations sont tridimensionnelles et instationnaires, et consommateur, nous y reviendrons, de volumes de données considérables (à titre d'exemples le calculateur TERA1 du CEA/DAM, produit en moyenne 4 Toctets de données par jour et le détecteur Babar de l'accélérateur linéaire de Standford conduit au traitement de 1 Poctet de données).

Néanmoins, lorsque les traitements sur les données sont indépendants et doivent être mis à disposition d'une vaste communauté de chercheurs, la Grille (Grid computing en anglais) est un formidable outil de travail et de coopération, complémentaire aux grands centres de calcul. Dans ce domaine le CERN est un des pionniers qui a organisé la distribution et le traitement des données issues de ses instruments. L'ensemble de la communauté mondiale du domaine a ainsi été structurée de manière hiérarchique avec un centre Tier0 au CERN effectuant un premier traitement, un ensemble de grands centres Tier1 qui assurent les services pour l'archivage des données et une première analyse, et des centres Tier2 et Tier3 qui fournissent les ressources pour les communautés locales.

Loi de Moore et architecture des calculateurs

Un autre paramètre caché de l'accroissement de la performance est l'accroissement proportionnel de la puissance électrique consommée et de la chaleur dissipée. Un processeur de dernière génération peut consommer jusqu'à 130 W. Facteur négligeable jusqu'en 2004, c'est pourtant ce qui va conduire à une remise en question fondamentale de l'architecture des processeurs au cœur de tous les systèmes, du PC au supercalculateur. En effet si d'après les technologues la Loi de Moore n'est pas remise en cause pour les 10 ans à venir : à savoir le doublement des transistors sur une puce tous les 18 mois, ce sont les conséquences qui doivent être revues. Jusqu'à présent la dynamique d'amélioration de la performance était obtenue par l'augmentation de la fréquence d'horloge, conjointement à l'optimisation de l'exécution des instructions et la mise en œuvre de mémoires caches Depuis 2004, la dynamique de croissance est conduite par la mise en œuvre simultanée de « cœurs multiples » au sein du processeur (ie un multiprocesseur sur la puce) et de la capacité à gérer simultanément plusieurs flots d'exécution (mécanisme d'«hyperthreading»), tout en utilisant les mémoires caches.

Les 3 révolutions en marche

Depuis novembre 1997 il n'y a plus eu de machine monoprocesseur dans la liste du Top 500 et avec la diffusion généralisée des processeurs multi-cœurs, on peut prédire qu'à peine 10 ans plus tard c'est l'ensemble des utilisateurs de l'informatique qui devra prendre en considération ce nouveau paradigme. C'est la première révolution du parallélisme à tous les étages. Par ailleurs la non-linéarité dans la relation entre fréquence et consommation peut encore accroître cette dynamique et conduire, à l'extrême comme pour l'architecture BlueGene d'IBM, à choisir d'interconnecter jusqu'à 130 000, processeurs d'architecture simplifiée et moins puissants que la technologie ne le permettrait, dans ce cas précis des processeurs à 700 MW, pour atteindre plus de 360 Tflops pour une consommation de seulement 1,5 MW, soit environ 7 W par processeur. Ceci étant néanmoins obtenu au prix de faibles capacités de mémoire et d'entrées/sorties.

Un autre exemple plus représentatif de supercalculateurs généralistes est la gamme des calculateurs TERA installés au CEA/DAM dans le cadre du Programme Simulation : le calculateur TERA1 installé en 2001 possède 2 500 processeurs pour une puissance de 5 Tflops. Le calculateur TERA10 construit par BULL, installé 4 ans plus tard, utilise 8 700 cœurs. Le facteur 10 en puissance a été obtenu autant par l'augmentation du nombre de processeurs que par l'accroissement de performance de celui-ci. Pour TERA100, prévu à l'horizon 2010, on peut s'attendre à atteindre 500 Tflops avec 60 000 cœurs de calcul, soit un facteur 10 en puissance obtenu essentiellement par l'augmentation d'un facteur 7 du nombre de processeurs !
Tout indique donc qu'une machine d'une puissance de 1 Pflop pourra être construite avant 2010 et que ce calculateur utilisera entre 100 000 et 1 000 000 cœurs selon le compromis qui sera effectué entre fréquence du processeur et consommation électrique. De plus ces dizaines de milliers de cœurs seront disposés dans le calculateur de manière hiérarchique, tout d'abord au sein du processeur, puis à l'intérieur d'un multiprocesseur, enfin au sein de grappes de multiprocesseurs. Et il s'agit bien entendu là d'un potentiel de puissance qui ne sera exploitable qu'au prix d'une adaptation des algorithmes et des méthodes numériques, tant il est vrai, rappelons-nous l'époque de la « vectorisation », que l'architecture se reflète dans les algorithmes et les méthodes de programmation utilisés.
Ces nouvelles contraintes pesant tant sur les algorithmes que sur le développement du logiciel vont provoquer la seconde révolution, celle du logiciel et des algorithmes. Pour certains, avec la diffusion du parallélisme dans toute la pyramide de l'informatique, il s'agit du plus grand changement depuis la révolution de la programmation orientée objets. Pour les spécialistes du calcul scientifique qui seront au cœur de la tourmente, la prise en compte de contraintes multiples telles que parallélisme massif, prise en compte de la localité, gestion des flots d'information va conduire obligatoirement à la composition de plusieurs approches tant du point de vue des méthodes numériques que des modèles de programmation ou des architectures logicielles.

Cependant le calcul intensif n'est pas réduit à l'utilisation de tout le calculateur pour une simulation unique, effectuée le plus rapidement possible. Au contraire, c'est le résultat et sa sensibilité à différents paramètres qui vont être de plus en plus recherchés. On peut donc penser qu'un calculateur de 1 Pflops pourra être rempli par 1 000 simulations requérant chacune 1 Tflop, dont les résultats seront « fusionnés » pour procéder à une analyse de sensibilité ou dans le cadre d'un processus d'optimisation. À nouveau on voit ici toute la nécessité d'une architecture équilibrée intégrant puissance de calcul et flot de données.

Mais revenons à nos grands calculateurs. Si on suppose disposer d'un processeur à 8 cœurs cadencés à 2 Ghz, pour obtenir vers 2010 d'une machine de 1,2 Pflops, et en supposant une consommation de 30 W par cœur, un bref calcul montre que sa consommation serait de 13 MW, Au coût actuel de l'électricité ceci représente 6,5 Meuros par an, soit 32 Meuros pour 5 ans d'utilisation, auxquels il faut ajouter le budget de maintenance. Alors même que d'autres coûts annexes, tels que la nécessaire adaptation de la salle machine, n'ont pas été pris en compte, cette estimation démontre que dans ce cas l'investissement représente au mieux 50 % du coût total de possession.
Ainsi, l'intégration de la puissance de calcul avec la gestion du flot de données, la nécessité de capitaliser la connaissance dans le logiciel, la nécessaire pérennité et portabilité de celui-ci (« les calculateurs passent, le logiciel reste ») ainsi que la part importante prise par les coûts annexes dans la mise en place de grands moyens de calcul conduira à les considérer et à les administrer de manière analogue aux très grands équipements. Ceci, conjointement aux deux autres révolutions, a des conséquences sur les communautés d'utilisateurs qui devront s'organiser en équipes pluridisciplinaires : c'est la troisième révolution, celle des communautés scientifiques.

Participer à la course ou mourir

Les avancées de la simulation promises par les ordinateurs de grande puissance ne pourront être obtenues qu'au prix d'une maîtrise de la complexité : complexité de la modélisation, complexité des méthodes mathématiques et algorithmiques, complexité de l'informatique, des techniques de validation et de production du logiciel. Ce qui fait la force des USA, les leaders actuels du domaine, c'est une telle approche intégrée et c'est le véritable défi auquel les communautés du calcul en Europe sont aujourd'hui confrontées.
Ce travail qui devra se fédérer autour des grands équipements de calcul exige la collaboration d'ingénieurs et de chercheurs, spécialistes des applications, du génie logiciel et des architectures informatiques.
Au-delà, d'ici une vingtaine d'années une autre révolution pourrait apparaître, le calcul quantique, lui-même intrinsèquement parallèle, qui pourrait conduire à la conception d'algorithmes radicalement nouveaux ■

Mélanger les lettres du message à envoyer.

La sécurité informatique

par **Jean-Jacques Quisquater** [1]

De la sécurité physique à celle de l'information et des ressources

On parlera intuitivement de sécurité dès qu'il s'agira de protéger de l'information, des ressources, des équipements ou des utilisateurs contre des menaces ou des attaques : il faudra en évaluer les risques. Il y a aussi le problème du faux document et de la manipulation de l'information. Les conquêtes ne purent se faire sans un excellent réseau de communications et de messagers (Marathon) en se protégeant des interceptions. L'information et sa protection vont se trouver du côté du pouvoir, des militaires et des diplomates.

Pour communiquer, les deux parties vont se mettre d'accord, à l'avance, sur une transformation des messages : l'expéditeur va appliquer au message la transformation directe et le destinataire applique le procédé inverse au message reçu pour avoir le message initial. C'est le procédé du *chiffrement-déchiffrement* du t*exte clair* vers le *texte chiffré* et vice-versa.

Une première idée - aujourd'hui, on parle d'algorithme cryptographique - est de mélanger les lettres du message à envoyer (scytale spartiate, décrite par Plutarque), c'est le *procédé de transposition*. Une deuxième, la méthode de César, est de substituer à une lettre donnée de l'alphabet une autre lettre de façon à former pour tout l'alphabet une table qui permute toutes les lettres, c'est le *procédé de substitution*. Jusqu'au début du XX[e] siècle, pratiquement toutes les méthodes de chiffrement proposée ou en usage sont des généralisations de ces deux méthodes (Viète, Vigénère,...). Les intercepteurs vont essayer de percer le sens : on parlera alors de cassage, ou mieux de *cryptanalyse*. La dissimulation d'objets ou de textes à l'intérieur d'autres textes fut aussi utilisée : c'est la *stéganographie* que l'on retrouve aujourd'hui (watermarking par exemple) pour protéger les supports multimédias (films, enregistrements sonores, ...) et permettre de tracer les copies illicites.

De l'analogique au digital

Sous l'impulsion de Kerkhoffs (1883) on aura des principes, les axiomes de la cryptographie : en termes modernes, tous les détails de l'algorithme sont publics, seul un paramètre, appelé la *clé*, facile à changer et à échanger, est gardé secret.

Le passage de la télégraphie avec fil à celui sans fil change la donne de la sécurité : en effet, il n'est plus possible de détecter ni d'empêcher les écoutes. En 14-18, ceci devient une nécessité. Vernam (l'ingénieur) et Mauborgne (le militaire) proposent un système en 1917. Il deviendra assez vite le système des militaires, vu ses qualités (incassable si bien implémenté) et malgré ses défauts (lourdeur d'implantation et de gestion). Shannon fut l'artisan d'une révolution : pour la sécurisation d'un système analogique, entre Roosevelt et Churchill, il montre qu'il n'est pas possible de le concevoir sûr, basé sur le téléphone et le système de Vernam. Il propose donc de passer de l'analogique au digital, donc en binaire pour la voix et toute information, puis d'appliquer le système de

1 Professeur à l'université catholique de Louvain, Belgique, directeur du groupe de recherche en cryptographie

Université Pierre et Marie Curie, laboratoire d'Analyse numérique.

Vernam pour transmettre sur un canal, au destinataire de déchiffrer et de faire la conversion du binaire vers la voix. Puis viennent, internet, GSM, web, exploration spatiale, chargement gratuit, ... La théorie de l'information (Shannon, 1949), y compris sa protection introduira les notions de confusion et diffusion pour les systèmes de chiffrement, la version moderne de la scytale et du système de César combinés pour le mieux.

La cryptographie moderne au service de la sécurité

Au cours des années 70, IBM proposera un algorithme de chiffrement, le DES, à clé courte, qui dominera la sécurité durant plus de 20 ans. Mais les papiers fondateurs de Diffie-Hellman (1976) et Rivest-Shamir-Adleman (RSA, 1978) vont donner les bases de la cryptographie actuelles pour résoudre les problèmes de sécurité posé par les objets comportant des microprocesseurs et communicant avec l'extérieur (aujourd'hui cela va du téléphone portable, carte de crédit, baladeur, etc).

Les buts classiques de la sécurité sont la *confidentialité*, l'*intégrité*, la *continuité du service* et l'*audit* dans les différents contextes : communications, stockage, calculs. Les nouvelles applications ajouteront par exemple l'anonymat, le calcul partagé, etc. L'intégrité sera résolue par la *signature*, apport immédiat de ces nouvelles idées.

Le chiffrement ne résout, qu'une partie de la confidentialité : le message à protéger circulera bien en clair à un moment donné, il faudra faire confiance aux communications et aux équipements : des paramètres sont à initialiser et à distribuer.

Le système à clé publique (Diffie-Hellman-Merkle) et son instance réussie, le RSA, vont compter parmi les résultats les plus extraordinaires de la conjonction des mathématiques (théorie des nombres) et de la théorie de la complexité. L'idée : supposons que vous disposez d'un algorithme de chiffrement avec une clé suffisamment longue : vous chiffrez un message en un temps t, très petit. Par contre, pour déchiffrer, le temps sera de l'ordre de *2100 x t*, infaisable. Ceci résout le problème de distribution des clés : chaque correspondant aura une paire de clés, publique et privée, au lieu d'une clé par paire de correspondants (ce qui augmente avec le carré du nombre d'utilisateurs). De nouvelles notions sont introduites : les fonctions à sens unique, faciles à calculer, difficiles à inverser, qui aideront à résoudre les problèmes d'intégrité.

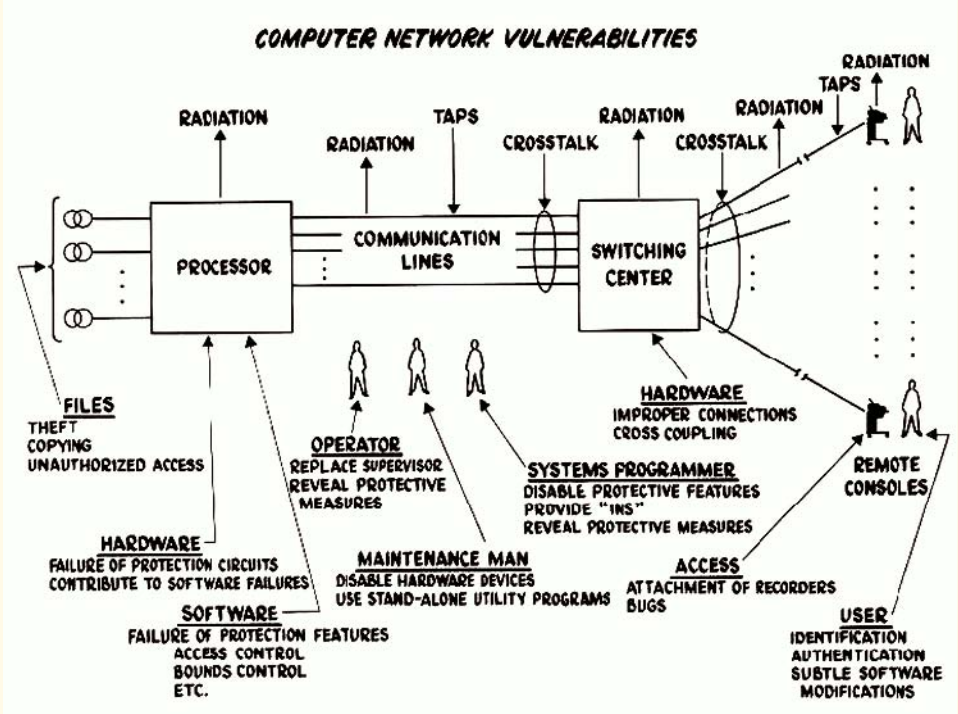
Réseaux et ordinateurs en besoin de sécurité

Durant les années 1960-1964, Baran publie sur les communications distribuées. Prémonitoire de l'internet, le mémoire « *Security, Secrecy, and Tamper-Free Considerations* » peut être vu comme le papier fondateur de la sécurité informatique. Ainsi, il invente le paquet (le concept utilisé dans les protocoles IP) pour des raisons de fiabilité et de confidentialité, propose le concept de réseau, discute des problèmes de distribution de clés et de leur changement, parle des problèmes de rayonnement

des câbles et des équipements, introduit le concept de résistance à l'intrusion physique (tamperproof), et, finalement, toujours utilisée aujourd'hui, décrit la procédure de login. La liste est bien plus longue quant aux vulnérabilités recensées pour un réseau d'ordinateurs. En fait, la figure peut faire office de programme de recherches en sécurité de 1965, pour les 50 années, au moins, à venir... Et tout semble encore à faire.

Le développement des réseaux, avec l'accès large par un public varié, donne lieu à nouveaux problèmes internet semble plier mais résiste. L'histoire des virus informatiques est, cependant, un peu à part. En effet, l'idée ne provient pas au départ d'une attaque plus ou moins sophistiquée mais bien de la conduite d'un séminaire (1984) où Adleman (le A de RSA) et Cohen élaborent le concept à partir de la théorie des automates reproducteurs de von Neumann. Un concept théorique, vite implanté, qui continue ses exploits aujourd'hui. Les abus divers (courriers non sollicités, spam, phishing) vont conduire à une réflexion sur la protection de la vie privée. Des comparaisons prises au domaine médical ou à l'écologie abondent mais ne donnent pas toujours les meilleurs résultats. On parle d'attaques automatiques à grande échelle, utilisant la reproduction d'objets nuisibles (*malware*), les dénis de service ou de ressources et d'intrusion. La recherche, et les produits dérivés, aura à cœur de détecter, prévenir, trouver des contremesures, et la tolérance (continuité malgré des attaques). De là, les concepts de sécurité adaptative et de gestion dynamique des réseaux et des systèmes.

Les cartes à puce (Moreno, 1974, et d'autres), la cryptographie dans votre poche, se développent d'abord comme applications sécurisées, carte de paiement, carte d'identité, passeport, comme preuve d'accès (SIM, TV à péage, ...) : ici, le produit sera utilisé avant que la recherche ne s'empare.



Avancées récentes de la sécurité

La sécurité se retrouve partout : on parlera de cybercrime, commerce électronique, réseaux de senseurs, de la sécurité Wifi, pervasive, ubivoque et des vignettes RFID. L'apparition de la cryptographie asymétrique, dispensant de contacts préalables rendra les applications (commerciales) possibles. On imagine donc des systèmes de vote, des enchères électroniques, des paiements...

La définition de langages de programmation de sécurité, ou dont la sécurité est une préoccupation essentielle, : le langage Java est présenté ainsi, le langage Oz et son système de programmation Mozart aussi. La sécurité par des spécifications rigoureuses dès le début viendra au secours des problèmes les plus compliqués.

La cryptographie symétrique, à clé secrète, (chiffrement en bloc ou en flux) continue à exister avec les nombreuses questions en suspens : de nouvelles attaques, fondatrices de la conception de ces algorithmes, se publient : les cryptanalyses différentielles (Biham-Shamir) et linéaire (Matsui), après avoir cassé tout sur leur passage, excepté le DES, permettent l'élaboration de nouvelles familles d'algorithmes, aboutissant à l'AES, résultats de toutes ces recherches, combinées avec celles sur les fonctions booléennes.

Preuves de sécurité

On procédera par contradiction sur des problèmes connus plutôt que de procéder par essai-erreur ou par examen des attaques connues, ce qui est sans fin. On essaiera aussi d'avoir des techniques pour convaincre (et se convaincre) de la sécurité. Ces preuves ont des limites dans la mesure où le modèle rend compte de la réalité.

Les preuves de connaissance, avec non transférabilité, conduiront à des preuves interactives par opposition aux preuves mathématiques classiques, preuves avec une grande probabilité plutôt qu'absolues,...

Les chercheurs (Goldreich, GMW) nous apprennent que :

- étant donné un protocole, une tâche et ses propriétés de sécurité, la question de la sécurité est indécidable en général, mais
- étant donné une tâche (réalisable) et les propriétés de sécurité associées, il est toujours possible de construire un protocole prouvé sûr permettant de réaliser la tâche en garantissant les propriétés de sécurité.

Récemment, on a pu ainsi réaliser l'analyse formelle de la norme IEEE 802.11i pour les réseaux sans fil, et du protocole TLS.

La mécanique quantique et la sécurité

Sur le plan théorique, l'apport de la mécanique quantique en algorithmique (Shor, 1994) et en cryptographie (Bennett et Brassard, 1982) a été une avancée considérable. Utilisant les effets liés au principe d'Heisenberg, à la polarisation des photons et au principe de superposition des ondes, la cryptographie quantique propose des systèmes incassables pour la confidentialité et la transmission de clés. L'identification semble bien plus difficile à résoudre dans ce contexte. Des démonstrations convaincantes ont eu lieu récemment et des applications démarrent. Cela reste un domaine encore en devenir, malgré la théorie élaborée, par les lourdeurs matérielles. Le problème du confinement de l'information, à la source et à la réception, reste : la sécurité quantique est encore à découvrir.

La mécanique quantique pourrait jouer un (grand) tour à la cryptographie : d'une part, comme on vient de le voir, des propositions pour résoudre des problèmes fondamentaux de la sécurité, et d'autre part, une contribution étonnante et, si réalisée, aux conséquences encore mal perçues aujourd'hui. En effet, Shor nous explique comment on peut réaliser des ordinateurs quantiques et comment les programmer pour résoudre efficacement des tâches de cryptanalyse. On pourrait alors imaginer qu'il soit possible de factoriser rapidement des grands nombres entiers. Le résultat est dur : on peut oublier le RSA, le logarithme discret autre que ceux basés sur les bonnes courbes elliptiques et on doit augmenter la longueur des clés secrètes au double. Il faudrait revoir bien des protocoles. Cette liste peut grandir. Il est devenu très difficile pour un chercheur de conseiller une méthode de protection des données à long terme (données médicales, par exemple, avec un terme d'une centaine d'années).

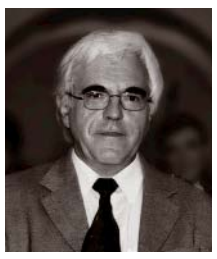
Depuis 1996 (Kocher), on sait trouver des clés secrètes cachées dans des circuits par des attaques où on mesure le temps, le courant, le rayonnement électromagnétique, etc. Ceci a donné un nouvel éclairage à des concepts anciens, de confinement, Tempest, canaux cachés. Le traitement statistique et du signal ont permis de grands progrès et tout matériel est en danger. D'autres attaques, plus ou moins intrusives, par fautes, par débordement, par pollution de cache, permettent aussi d'accéder aux secrets. Récemment on a montré que les fonctions de hachages ont un impact immédiat sur l'usage quotidien de la cryptographie, car de nouvelles attaques ont été trouvées par ce biais. Un vrai domaine de recherches actuelles où les fondements se sont révélés plus faibles que prévus.

La recherche en sécurité ne sera jamais finie. À la rigueur des chercheurs succède à chaque fois l'imagination des cryptanalystes ou pirates, et aux implantations peu soigneuses de multiples trous. L'apport de la logique, de la cryptographie la plus récente, et d'une science en devenir de la vérification des protocoles de sécurité et de leur implantation, par leur construction systématique, correcte et sûre, permet d'espérer une plus grande résistance à tous les problèmes. Il faut continuer à investir dans une recherche de qualité,... et à l'appliquer ■

Le marché le plus important, c'est le téléphone

Progrès du matériel et progrès du logiciel

Entretien avec Gérard Berry[1]



par **Paul Caro**[2]

Quelle est l'évolution du matériel le « hardware » en informatique ?

Nous vivons à l'heure actuelle une révolution profonde, celle de la numérisation de tous les types d'information : textes, sons, images, films, sous, forces, etc. Tout est transformé de façon homogène en langage binaire 0 / 1 et traité par des circuits électroniques eux-mêmes commandés par des logiciels. Ceci a été rendu possible par l'extraordinaire progrès des matériels électroniques, circuits, mémoires, et stockage de masse. Ces trois types de matériel évoluent de façon conjointe, mais je parlerai surtout du premier. L'évolution des circuits est encore dirigée par la loi de Moore, qui dit en gros que la puissance de calcul double tous les deux ans. C'est dû aux progrès de la micro-électronique et aussi aux progrès en matière d'architecture logique des circuits. A l'heure actuelle, il y a deux genres assez fondamentalement différents de circuits. Les plus connus sont les microprocesseurs généralistes des PC ou des supercalculateurs pour lesquels il y a très peu de fabricants. Ils sont passés en moins de 30 ans du millier au milliard de transistors . Ils ont atteint des vitesses extrêmement grandes, et les accélérer n'est plus si capital sauf pour des applications importantes mais minoritaires (super-calculateurs). Le deuxième genre est embarqué un peu partout dans les objets du quotidien, avec des fonctions tout à fait différentes de celles des ordinateurs : communiquer, contrôler, mesurer, faire de la musique. Ces circuits appelés System on Chips (SoC) sont plus petits, moins rapides, et davantage dédiés à des applications particulières. Ils peuvent contenir des processeurs, divers types d'accélérateurs, des capteurs et actionneurs, des émetteurs-récepteurs, des étiquettes informatiques, etc. Il s'en fabrique beaucoup, plusieurs milliards par an à l'heure actuelle, et leur nombre est en augmentation absolument astronomique. Le marché le plus important est le téléphone, mais arrivent derrière tout l'audiovisuel, maintenant intégralement numérique, les transports, les automobiles, etc. Dans l'avenir, beaucoup plus d'objets recevront des étiquettes électroniques d'identification unique et des puces électroniques reliées en réseau plus ou moins sophistiqué. Par exemple, des dispositifs assureront des contrôles médicaux en continu et la surveillance des personnes âgées. Après l'Internet, l'embarqué est l'autre grand domaine de l'informatique du futur. Il sera peut-être encore plus explosif. L'industrie électronique s'en trouvera profondément modifiée.

Cette loi se ralentira, probablement pas à cause des difficultés physiques bien réelles de la miniaturisation des transistors, mais bien avant à cause d'autres

1 Membre de l'Académie des sciences, Membre de l'Académie des technologies,directeur scientifique d' Esterel Technologies.
2 Correspondant de l'Académie des sciences, directeur de recherche honoraire au CNRS.

difficultés : les problèmes de chaleur dissipée (on ne veut pas ventiler pour des questions de bruit, et il ne faut pas user la pile), les énormes coûts d'investissement et aussi le fait que le besoin d'augmentation de puissance décroît (mais pas partout : une télévision haute définition devra faire des téra-opérations – millions de milliards – par seconde, pour quelques Watts). La conception d'un circuit devient extrêmement dure, parce qu'elle demande de définir l'assemblage de dizaines ou centaines de millions d'éléments fabriqués en une fois, sans pouvoir en changer aucun, ni au cours de la fabrication, ni après. Le design et la vérification deviennent des problèmes gigantesques, utilisant des logiciels de CAO extrêmement compliqués. On est bel et bien dans l'industrie lourde : les coûts marginaux de production sont très faibles, la diffusion est considérable, mais les investissements en conception et fabrication sont énormes. Les circuits spécifiques ne sont donc plus fabriqués à moins de plusieurs millions d'exemplaires. Heureusement, il existe une nouvelle race de circuits qui se développe beaucoup pour les applications de moindre volume, les circuits programmables. Les plus connus sont les FPGA (Field Programmable Gate Array), dont on peut télécharger la logique par logiciel.

Où en est le progrès en termes de logiciels dans ce contexte ?

Le matériel est complexe sur le plan de la physique et de la logique interne mais reste encore relativement simple et régulier sur le plan fonctionnel: il s'agit de faire très vite des choses individuellement stupides, comme prendre un nombre, lui ajouter 1, et mettre le résultat dans une mémoire. Le logiciel est considérablement plus varié. On lui demande des fonctions beaucoup plus riches avec des contraintes extrêmement variables. Un simulateur de vol sur PC ou un vrai logiciel de vol Airbus n'obéissent pas aux mêmes contraintes de sécurité alors que les matériels sur lesquels ils s'exécutent ne sont pas si différents. Il y a aussi plusieurs classes de logiciels. Les logiciels classiques (gestion, etc.) évoluent vers une plus grande sophistication et vers la répartition géographique des fonctions. Les logiciels de communication ou de gestion des réseaux cherchent à transférer les informations à grande échelle. Les très grands moteurs de recherche qui gèrent des milliards d'informations n'existaient pas il y a dix ans ; ils emploient des algorithmes originaux, que ce soit pour la collecte des données ou pour la réponse ultra-rapide aux questions. Les logiciels embarqués dans les objets deviennent de taille et d'importance considérable : il y a plusieurs millions de lignes de code dans un téléphone ou dans un avion. Enfin, pour les logiciels embarqués dans les avions, les voitures et les appareils médicaux, la sécurité est absolument fondamentale. Il n'est pas question de diffuser des versions plus ou moins cuites en demandant aux utilisateurs de télécharger les nouvelles versions quand on trouve un problème, pratique courante dans d'autres domaines. Dans les avions, le logiciel est d'ailleurs certifié au même titre que n'importe quel autre composant, c'est à dire très très sérieusement et par des autorités indépendantes. Dans les voitures ce n'est malheureusement pas encore le cas et le niveau de qualité n'est pas encore le même, comme peuvent le constater les utilisateurs à leurs dépends. Pour beaucoup d'applications, la conception de l'interface homme-machine est aussi un endroit à problèmes.

Comment éviter les erreurs, les « bugs » ?

Le « bug » est le grand ennemi de l'informaticien. Je cite toujours une phrase historique de Maurice Wilkes, l'un des créateurs des ordinateurs, en 1949 : « Dès que nous avons commencé à programmer, nous avons découvert, à notre grande surprise, qu'il n'était pas aussi facile d'obtenir des programmes corrects que nous le pensions. J'ai alors réalisé qu'une grande partie de ma vie serait consacrée à découvrir des fautes dans mes propres programmes ». Voyons pourquoi. Nous avons vu qu'un circuit fait parfaitement et très vite des choses très simples. Pour faire quelque chose de complexe, comme téléphoner ou prendre une photo, le logiciel doit définir le nombre gigantesque d'actions simples qu'il faut faire exécuter par le circuit, qui se comptent en milliards. S'il y a la moindre erreur de spécification et de codage, le circuit exécute l'ordre erroné avec parfaite obéissance et totale absence de regret, rendant éventuellement un résultat complètement erratique. Au contraire du matériel où on sait bien corriger les erreurs aléatoires liées à la physique, le logiciel a un comportement chaotique : changer un seul 0 en 1 peut passer de la chose juste au n'importe quoi. Malheureusement, la taille et la complexité rendent l'absence d'erreurs très improbable, car on parle de millions de lignes de code et de dizaines de logiciels communiquant entre eux. La notion de bug est assez unique à l'informatique, ce qui fait qu'elle est souvent ignorée ou redécouverte à la dure par les ingénieurs et scientifiques des autres disciplines, avec de nombreux échecs à la clef.

Le matériel est complexe sur le plan de la physique et de la logique interne mais reste encore relativement simple et régulier sur le plan fonctionnel: il s'agit de faire très vite des choses individuellement stupides, comme prendre un nombre, lui ajouter 1, et mettre le résultat dans une mémoire.

Université Pierre et Marie Curie, laboratoire d'Analyse numérique.

ESTEREL, c'est un langage de programmation défini mathématiquement depuis le début de son histoire, la définition mathématique a précédé le langage. Ce genre de technologie est utilisée pour faire voler les Airbus et le fait que le langage soit rigoureux et simple est un avantage énorme sur les langages classiques qui sont faits par des praticiens très bons bricoleurs

Pour éviter les bugs, la première chose à faire est de reconnaître et d'accepter leur caractère inéluctable par défaut. Il faut ensuite employer des techniques de spécification, développement, et vérification très rigoureuses, qui sont maintenant bien définies mais pas encore utilisées partout. Par exemple, pour la certification des logiciels d'avionique, tout ce qu'on écrit doit être expliqué et justifié, et avoir une correspondance explicite entre les exigences formulées au plus haut niveau et chaque ligne du code. Cette méthode donne assez satisfaction, mais elle permet encore de certifier des choses fausses, car on atteint les limites de la possibilité humaine de comprendre ce que l'on a vraiment écrit! Pour aller au-delà, il faut une approche scientifique, dont les fondements sont développées depuis 40 ans et qui passe maintenant à la pratique sous le nom de méthodes formelles.

Quels sont les principes mathématiques, les méthodes formelles qui peuvent être utilisées pour les éviter ?

Il y a deux questions complémentaires : faire moins d'erreurs et mieux détecter celles qui restent. La première est

profondément liée à la notion de langage de programmation. Il faut éloigner les modes de spécification des fonctions de la structure de la machine pour les rapprocher du mode pensée de celui qui conçoit et écrit les programmes. Il faut aussi donner un sens parfaitement précis aux programmes, ce qu'on appelle une sémantique mathématique ; on ne peut pas espérer prouver quoi que ce soit sur un programme si le langage dans lequel il est écrit est imprécis. De bons langages réduisent le taux d'erreur de façon majeure. La deuxième question est fondée sur un idée simple : puisque les ordinateurs peuvent calculer sur tout, pourquoi ne pas les faire calculer sur leurs propres programmes ? Notre regretté confrère Gilles Kahn a largement contribué à cette approche, qui permet d'explorer beaucoup plus de comportements potentiels que ne peuvent le faire les hommes et de vérifier leur correction. En couplant les deux techniques, on peut dans certains cas aller jusqu'à la vérification formelle complète du comportement souhaité d'un programme ou d'un circuit.

Scientifiquement parlant, il y a deux approches complémentaires, l'approche logique et l'approche systèmes finis. L'approche logique repose sur tout ce qui a été fait depuis les années 30 sur l'étude du calcul (Turing, Gödel, Scott, etc.), et sur la relation très profonde qui existe entre les notions de calcul et de déduction logique. Des langages de programmation comme CAML (généraliste) ou Esterel / Lustre (pour l'embarqué) ont des sémantiques logiques bien définies ; les derniers cités sont utilisés industriellement pour la programmation des Airbus et de bien d'autres objets volants. On étudie maintenant beaucoup les logiques intuitionistes, créées dans les années 1930 mais largement rejetées par les mathématiciens classiques comme trop complexes pour leurs besoins. Elles s'avèrent bien adaptées à l'informatique car elles d'attachent à la construction des objets et pas seulement à leur existence : de la preuve d'existence d'un objet on peut effectivement extraire le programme qui le fabrique. Elles s'implémentent également bien en machine. Un des systèmes les plus évolués à l'heure actuelle est Coq, créé dans les années 1980 par G. Huet, de l'Académie, et T. Coquand. Lui ou ses semblables commencent à être utilisés dans l'industrie pour prouver des applications variées comme les protocoles de cartes à puces, des compilateurs, ou la correction des calculs flottants dans les circuits. Coq sert également de support à une nouvelle formalisation des mathématiques constructives. G. Gonthier a ainsi pu démontrer en Coq le fameux théorème des 4 couleurs, dont la preuve humaine était gigantesque et encore sujette à caution.

L'approche systèmes finis repose sur des principes différents. Lorsque les objets sur lesquels on calcule sont de taille finie, même très grande, on peut utiliser l'ordinateur pour analyser l'ensemble de leurs configurations potentielles. On peut par exemple démontrer qu'un ascenseur ne peut pas voyager la porte ouverte. Ceci ne se fait pas par énumération brutale, mais en calculant symboliquement sur des formules représentant les configurations. Une des clefs est le calcul Booléen, i.e. le calcul sur 0 / 1 avec et, ou, non. Ce calcul apparemment tout bête reste un mystère mathématique. Théoriquement, il n'y a aucune chance de calculer efficacement sur les très grosses formules dont on a besoin (dizaines ou centaines de milliers de variables). Pratiquement, on travaille souvent sur des formules un peu spéciales pour lesquelles des algorithmes heuristiques marchent bien, sans qu'on sache réellement pourquoi. Ces techniques sont utilisées en grand pour la vérification des circuits, de protocoles, ou de logiciels embarqués. Il leur reste un petit côté alchimique qui demande encore beaucoup de compréhension scientifique.

Il existe une approche intermédiaire, l'interprétation abstraite, due à Patrick Cousot et à son équipe. Au lieu de faire des calculs exacts sur des domaines potentiellement infinis, on fait des calculs approchés sur des enveloppes finies en respectant la sémantique mathématique du langage de programmation utilisé. Par exemple, pour démontrer qu'un dénominateur ne sera jamais zéro, on va simplement démontrer qu'il est toujours compris entre 2 et 43, sans jamais faire le calcul exact. L'interprétation abstraite a été mise en valeur par le célèbre bug de l'explosion d'Ariane 501, dû à un débordement arithmétique. Elle permet de le trouver facilement sans même exécuter le programme en simulation. Plus récemment, elle a été utilisée pour vérifier l'absence de défaut arithmétique dans les commandes de pilotage de l'Airbus A380.

Je ne parlerai que peu de l'autre grande branche de l'informatique, l'algorithmique, qui consiste à organiser les calculs pour les faire de la façon la plus économique possible. Il s'agit d'un sujet très large et très difficile, allant des algorithmes de stockage et de recherche d'information aux algorithmes de cryptage mathématique en passant par l'analyse des protocoles de réseaux de communication et de leurs mécanismes de sécurité. Comme la différence entre un algorithme juste et un algorithme faux est très ténue, il sera indispensable d'avoir de vraies bibliothèques d'algorithmes certifiés et bien analysés, ce qui est actuellement en gestation.

Malheureusement, il faut souvent attendre des bugs majeurs pour promouvoir la recherche sur ces sujets. Un crash massif du système téléphonique interurbain d'ATT a provoqué la création d'un grand labo de recherche aux Bell Labs. Le fameux bug de division du Pentium d'Intel a fait de même à Intel ; l'arithmétique des Pentium est maintenant prouvée formellement, y compris pour le flottant. Le crash d'Ariane 501 a mis l'interprétation abstraite au goût du jour. Il est temps de prendre la chose avant et pas après les ennuis, comme en matière de bâtiment où on fait les calculs de résistance d'un pont avant et pas après la construction. J'ai cité plus haut de nombreux travaux français. La France est effectivement très bien placée dans ce domaine, que ce soit dans la recherche, la création de sociétés technologiques, ou l'utilisation industrielle. Mais il reste encore à transformer l'essai.

Que pensez-vous de l'enseignement de l'informatique ?

Il y a en France un retard psychologique considérable et beaucoup de confusion, surtout dans l'enseignement primaire et secondaire. Beaucoup de gens, même chez les scientifiques, pensent que l'informatique consiste à savoir se servir d'un traitement de texte ou d'Internet et qu'il n'y a donc pas besoin de l'enseigner. D'autres pensent urgent d'enseigner à l'école l'usage de la souris, alors que ce sont souvent les enfants qui l'enseignent aux adultes. D'autres encore disent qu'il faut enseigner très tôt la programmation, chose pas particulièrement intéressante si on n'a pas un but bien compris et apprécié. Résultat, on ne fait pas grand-chose, et le vrai sujet n'est quasiment jamais abordé. Ce qu'il faut enseigner, c'est pourquoi toutes les informations sont devenues numériques, pourquoi et comment elles se manipulent de façon homogène, que ce soit des textes, des sons, des images, des forces, ou des sous, et comment faire des choses qui marchent. Il faut fonder l'enseignement sur une vision, pas sur un ensemble de détails.

Enfin, il reste une forte résistance de beaucoup de scientifiques ou d'enseignants à penser que l'informatique est un sujet scientifique. Ils pensent plutôt que c'est une technique utilisée pour le calcul numérique, et ne perçoivent ni la richesse des applications ni l'ensemble des questions scientifiques nouvelles qu'elle pose et résout. Une question m'a souvent été posée par de hauts responsables de tous types : l'informatique est-elle une science ? La réponse est pour moi évidemment oui. L'informatique a développé une façon autonome d'identifier et de résoudre les problèmes de calcul à grande échelle, en s'appuyant bien sûr sur un passé mathématique et physique considérable, et en perfusant à son tour dans quasiment tous les autres domaines de la science (voir l'entretient de Gilles Kahn sur le site). Mais, au fond, la question a-t-elle véritablement un intérêt ? En 20 ans, l'informatique a déjà complètement remodelé le monde dans des domaines aussi variés que la communication, la publication, l'outillage scientifique et médical, les transports, et ce n'est qu'un début. Faut-il enseigner aux enfants le monde d'hier ou celui de demain ? ■

Gilles Kahn[1] et l'INRIA

La disparition prématurée de Gilles Kahn laisse un vide profond dans l'informatique de notre pays, pour laquelle il fut tout à la fois un scientifique créatif et important et un penseur doté d'une vision sur l'avenir et sur l'organisation de cette discipline à laquelle il s'est consacré avec une énergie et une efficacité inégalées. C'est avec une admiration teintée d'une immense tristesse que nous fûmes témoins de son courage face à la maladie cruelle qui rongeait progressivement ses forces. Il n'avait pourtant en rien ralenti son activité inlassable et, au sein de l'Académie, il avait conservé jusqu'à ses dernières heures toutes les tâches auxquelles il avait accepté de se consacrer, membre du Comité restreint, président de la Fondation Franco Chinoise, etc.
Il nous a semblé que la republication de cet entretien que Gilles Kahn avait accordé à la « Lettre » il y a cinq ans, dans lequel il exposait sa vision de l'informatique, permettrait de retrouver pour quelques instants la pensée de ce très grand scientifique.
Ce texte sera disponible sur le site de l'Académie (http :www.academie-sciences.fr).

par **Bernard Larrouturou**[2]

Président-directeur général de l'INRIA, membre de l'Académie des sciences et de l'Académie des technologies, Gilles Kahn est décédé le 9 février 2006, dans sa soixantième année. Son décès interrompt brutalement un parcours d'exception, entièrement dédié à la science et au service de la recherche française.
Après ses études à l'École Polytechnique et à l'Université de Stanford, Gilles Kahn passe quelques années au CEA et rejoint à trente ans l'IRIA, futur INRIA, où il mènera toute sa carrière. Dès les années 60, il s'engage dans la recherche en informatique. Il a une influence très profonde dans le domaine de la sémantique des langages de programmation, qui vise à donner un sens précis à un programme informatique et à ses spécifications. Ses recherches personnelles comprennent plusieurs idées fondamentales pour fournir des méthodes et des outils permettant de développer des logiciels conformes à leurs spécifications. Les « réseaux de Kahn », sa première contribution marquante, fournissent un cadre conceptuel pour décrire le calcul distribué asynchrone. Gilles est aussi le père de la « sémantique naturelle » qui fournit un cadre élégant pour manipuler les programmes et calculer leurs propriétés, et ouvre la voie au développement des environnements de programmation. Depuis plus d'un quart de siècle, Gilles Kahn est reconnu partout dans le monde comme un des grands pionniers de la recherche en informatique.
En parallèle avec ses recherches, Gilles exerce aussi de nombreuses activités de direction de recherche. Dès la fin des années 70, Jacques-Louis Lions, premier président de l'INRIA, confie des responsabilités à ce jeune chercheur dont le jugement s'exerce déjà avec acuité sur un spectre bien plus large que ses propres recherches. Pierre Bernhard et Gilles Kahn sont, en 1983, les fondateurs du laboratoire de l'INRIA à Sophia-Antipolis ; ensemble, ils le conduiront à un succès exemplaire, qui en fera en quelques années un des tout meilleurs centres français en informatique. À partir de 1994, Gilles Kahn est directeur scientifique de l'INRIA, auprès d'Alain Bensoussan puis de moi-même, et il est nommé en 2004 président directeur général de cet organisme.
Dans ce parcours, Gilles se révèle un dirigeant de premier ordre. Sa remarquable ouverture d'esprit, son goût pour les échanges d'idées avec tous les chercheurs, qu'ils soient académiciens ou doctorants, son enthousiasme et sa grande chaleur lui donnent un rayonnement exceptionnel. Sa mémoire prodigieuse, ses contacts internationaux incroyablement nombreux, sa soif insatiable d'apprendre, ses lectures encyclopédiques font de lui une source inépuisable de savoir et de conseils. Assumant tous les aspects du rôle de dirigeant d'organisme, au point de se former au management, il est toujours attentif à l'équilibre entre la réflexion stratégique, où se forme l'avenir d'un établissement, et la place laissée à l'initiative des chercheurs, dont il reste très proche.
Scientifique reconnu pour ses contributions à son propre domaine de recherche, Gilles est aussi un homme de science à la vision très large. Très vite, sa vision embrasse tous les domaines présents à l'INRIA, jusqu'au calcul scientifique, à l'automatique, au traitement d'images, et aux aspects scientifiques et technologiques des télécommunications. Plus largement, il est passionné par l'apport des sciences de l'information et de la modélisation mathématique aux autres sciences, notamment dans les domaines de la biologie et de la médecine dans lesquels il investit beaucoup. Premier informaticien élu à l'Académie des sciences en 1997, il est très présent dans le dialogue avec des scientifiques d'autres domaines... et rien ne l'agace plus que les situations où, en ce début du XXI ème siècle, il faut encore expliquer à des collègues de disciplines plus anciennes que l'informatique est une vraie science, avec ses approches spécifiques et ses idées profondes, et dont l'interaction avec les autres sciences est une perspective de recherches extrêmement prometteuses !
Comme chercheur et comme dirigeant, Gilles Kahn est aussi animé par une vision très positive des applications de la recherche. Il participe à plusieurs aventures de créations d'entreprises issues de l'INRIA. En 1996, il joue un rôle de premier plan dans l'enquête sur la défaillance logicielle de la première mission de la fusée Ariane 5. Plus récemment, il rédige plusieurs rapports pour l'État français sur l'impact du développement des technologies de l'information et de la communication. De nombreux industriels, français ou étrangers, bénéficient de ses conseils.

Avec le décès de Gilles Kahn, nous avons perdu un homme de science de premier plan, un visionnaire et un grand dirigeant de la recherche française. Beaucoup d'entre nous ont aussi perdu un ami, un « grand frère » dont la chaleur restera toujours dans nos mémoires et dans nos cœurs ■

1 Membre de l'Académie des sciences, ancien président directeur général de l'Institut nationnal de recherche en informatique et en automatique.
2 Président directeur général de l'Institut nationnal de recherche en informatique et en automatique.

Mise au point sur l'olfaction

par **Pierre-Marie Lledo** [1]

À sa naissance, l'art des parfums reste indissociable des religions. De l'Égypte ancienne au christianisme, il reste omniprésent dans le rapport de l'homme avec ses Dieux. L'histoire montre qu'ensuite de complexes cheminements sociaux et culturels conduiront à désacraliser l'emploi des senteurs puis à les ignorer. Après plus de 10 000 ans de civilisation, voici que le XXI ème siècle s'ouvre de nouveau à l'odorat pour explorer les relations intimes que l'Homme entretient avec la matière de son environnement. Longtemps négligée eu égard à sa connotation sensuelle, l'olfaction se trouve aujourd'hui revisitée. Scientifiques, artistes et industriels s'affèrent ardemment à déchiffrer cet univers aussi complexe qu'évanescent.

Une renaissance

De nos cinq sens, l'olfaction est restée sans doute le sens le plus méconnu, le plus flou. Ce dernier a longtemps été considéré comme une modalité sensorielle secondaire, pas très utile à l'homme, en un mot peu digne d'intérêt pour les recherches scientifiques. Depuis l'avènement de la biologie moléculaire portée par Richard Axel et Linda Buck dans le champ de l'olfaction, cette situation s'est renversée. Aujourd'hui, les recherches sur l'olfaction sont en plein essor. Ce sens mystérieux livre peu à peu ses secrets et l'opinion publique se fascine à découvrir ce sens mal connu. On réapprend que les odeurs influeraient sur nombre de nos comportements, le plus souvent à notre insu. On découvre que sans cesse refoulé, l'odorat inspire nos humeurs et dirige nos actions artistiques ou prosaïques. Les psychiatres nous apprennent que ne plus percevoir d'odeur nous plongerait inéluctablement dans la dépression. Au contraire, certaines senteurs nous relaxeraient. Ces exemples traduisent l'importance des odeurs imprégnant la vie d'*Homo Sapiens*. Pourquoi donc l'olfaction se trouva pendant fort longtemps, sinon réprouvée, du moins constamment dévalorisée par une communauté scientifique prompte à suivre la grande tradition de la philosophie occidentale[2] ?

Sans doute est-ce parce que l'odorat est un sens hérité de notre évolution — proche des pulsions nées de la part animale de l'Homme et notamment de sa sexualité — que de nombreuses sociétés l'ont refoulé. Elles ont voulu juguler ce sens souvent capable de réveiller les tabous relégués dans l'ombre de nos consciences. Elles y sont parvenues, à tel point que Freud voyait dans cette répression la preuve ultime de civilisation. Or, en se méfiant de l'olfaction, ces mêmes sociétés se privent de connaissances susceptibles d'améliorer les conditions mêmes de notre vie[3]. On découvre que les molécules volatiles odorantes ne sont pas les seuls à porter un message important. Les phéromones[4] capables de délivrer secrètement des messages au sexe opposé révèlent aussi leur lot de secrets. À partir de molécules émises par notre corps, on s'aperçoit que nous décryptons des informations relatives à l'appartenance sexuelle d'un individu, de son âge (pré ou post-pubère), de sa filiation génétique, de son rang hiérarchique ou de son état de santé. Autant l'admettre, le pouvoir des odeurs, par le contenu émotionnel qu'elles véhiculent, par les images qu'elles font surgir de notre mémoire, reste central dans l'adaptation de l'être humain à son environnement.

Une anthropologie de l'olfaction

Chez l'Homme, trois caractéristiques définissent l'odorat : un univers complexe des stimuli (plus de 10.000 [5] molécules sont odorantes), le déclenchement instantané d'émotions (agréables ou désagréables) et enfin, un lien privilégié avec la mémoire. Par leur forte tonalité émotionnelle et leur faible contenu cognitif, les odeurs restent pour les humains le lien le plus directe avec la matérialité des substances qui les entourent. Ce lien fort est sans aucun doute la traduction d'une très ancienne histoire phylogénétique. Les odeurs généralement qualifiées d' "agréables" évoquent des histoires de nourritures, de nature ou de partenaire sexuel, les odeurs dites "désagréables" sont en rapport direct avec les dangers : incendie, maladie ou prédateur. Cette observation conduit à formuler l'hypothèse selon laquelle l'appareil olfactif fonctionnerait comme un système d'alarme inné vis-à-vis de sources potentiellement dangereuses et d'un système capable d'associer des expériences fort diverses. Bien avant l'apparition de l'Homme, l'olfaction fournissait aux animaux terrestres et aquatiques les indices les plus sûrs pour tirer le meilleur parti de leur environnement grâce à des comportements bien adaptés à leur finalité. Aujourd'hui, l'histoire récente du marketing olfactif ne rend-elle pas compte de ce lointain héritage ?

La découverte des récepteurs aux molécules odorantes

L'odeur se définit comme la sensation produite par des émanations de molécules volatiles qui activent un capteur sensoriel. L'olfaction est donc une sensibilité chimique qui nécessite le contact physique de certaines molécules avec des récepteurs localisés dans l'organe sensoriel. Chez les mammifères, cet organe est la muqueuse olfactive située dans les fosses nasales. Chez l'Homme, il occupe une surface de 2 à 3 cm^2. Si l'accès à cette région pour les molécules odorantes contenues dans l'air inspiré est optimal lors du flairage, il peut se faire aussi par voie rétro-nasale. Ainsi, l'arôme souvent confondu avec la saveur d'un aliment est l'ensemble des substances odorantes perçues selon cette voie rétro-nasale et non directement par les narines.

Les particularités de l'odorat que nous venons de citer soulignent l'importance des travaux conduits par Richard Axel et Linda Buck, récompensés le 4 octobre 2004 par l'Académie suédoise. Cette plus haute des distinctions scientifiques couronne un ensemble de travaux qui ont conduit une partie de la communauté scientifique à repenser la physiologie de l'odorat. Ces deux chercheurs ont réussi à décrypter, à l'échelon génétique et moléculaire, les mécanismes responsables de la reconnaissance des odeurs. Ils ont de la sorte élucidé quelques-uns des secrets entourant la fonction olfactive qui demeurait la plus méconnue de nos fonctions sensorielles. Limitée jusqu'alors aux mondes des poètes, gastronomes, et autres parfumeurs, l'olfaction s'est progressivement transformée en véritable objet de recherche. Nous sommes aujourd'hui assurés que les récepteurs aux molécules odorantes sont des protéines. Nous avons accumulé nombre de connaissances précises sur leur composition et leur fonctionnement. Il s'agit de récepteurs formés à partir de protéines transmembranaires dont le nombre avoisine le millier chez les rongeurs et le demi-millier chez l'homme. Les récepteurs olfactifs forment une gigantesque famille de molécules constituées d'une chaîne d'acides aminés qui traverse la membrane plasmatique des cellules à sept reprises[6]. Toutes les protéines de cette famille partagent des similitudes de

1 Directeur de recherche au CNRS, Unité « Perception et mémoire », Institut Pasteur, CNRS URA 2182

2 L'odorat apparaissait trop vulgairement sensuel pour figurer au rang des sens supérieurs dans la hiérarchie qui s'est imposée. C'est Aristote qui ouvrait la voie vers une relégation de l'odorat dans son traité de l'*âme*. L'olfaction y est dit manquer de finesses ou de discernement. Sens grossier parce que utilitaire, reprendra Saint Thomas d'Aquin (1225-1274). Lorsque Descartes (1596-1650) en vient à analyser le sens de l'odorat dans son *Traité de l'Homme*, il apporte bien des précisions qui ne manquent pas de saveur : "Le sens de l'odorat dépend de plusieurs petits filets qui s'avancent de la base du cerveau vers le dessous de ces deux petites parties toutes creuses, et qui ne diffèrent en rien des nerfs qui servent à l'attouchement et au goût". L'anti-aristotélisme de Descartes ne va pas jusqu'à contester la hiérarchie scolastique : l'odorat reste le dernier des sens. John Locke (1632-1704) à son tour suit l'ordre aristotélicien et, si hostile qu'il se veut au cartésianisme, reprend l'essentiel de son explication physiologique. Les odeurs se trouvent rangées parmi les "qualités secondes" des corps, que le philosophe anglais considère comme imputables à l'action de quelques particules insensibles sur les organes de nos sens. La question des sensations olfactives restera donc secondaire aux yeux des philosophes. La dépréciation de l'odorat, si manifeste encore chez Kant et Hegel, sera enfin rejetée par un élève de Hegel (Feuerbach) puis Nietzsche qui entamèrent une véritable guerre contre les philosophes qui méprisaient le corps.

3 Voir "*Pour une nouvelle physiologie du goût*", de Jean-Didier Vincent et Jean-Marie Amat, Odile Jacob (2000).

4 Du grec *pherein* porter et *horman* exciter. On appelle phéromone un signal chimique libéré par un sujet qui procure aux autres individus de l'espèce des informations sur son genre, son statut social ou reproducteur. Ce signal agit à distance de son lieu d'origine et à très faible concentration, pour déclencher des comportements sexuels ou sociaux stéréotypés.

5 À vrai dire, ce chiffre n'a pas grand sens car l'industrie chimique synthétise chaque jour de nouvelles molécules qui se révèlent odorantes ; l'espace des stimulus olfactifs n'est donc pas limité.

séquenceux niveaux de certains domaines transmembranaires. Il est particulièrement satisfaisant pour les biologistes de voir que les premières étapes de la perception olfactive, malgré ses propriétés singulières, suivent des principes de fonctionnement analogues à ceux qui régissent d'autres propriétés de la communication cellulaire[7]. Quelle belle confirmation de l'unité du Vivant !
Sur le plan fondamental, les recherches actuelles montrent combien le système olfactif reste unique. Ce dernier a adopté une stratégie bien différente de celle employée par les autres organes sensoriels. Pour analyser les informations gustatives, visuelles, auditives ou tactiles, les systèmes sensoriels ont développé un nombre limité de récepteurs puis les ont répartis sur la surface de l'organe sensoriel. Ainsi une partie du codage de l'information afférente réside sur la localisation précise des récepteurs activés : on parle de codage spatial. En révélant l'existence de la plus large famille de protéines dédiées à l'olfaction, les travaux de Richard Axel et de Linda Buck indiquent que le système olfactif code différemment l'information. Il est vrai que non seulement le stimulus olfactif n'a pas de dimension spatiale, mais ses paramètres sont trop nombreux pour être correctement transposés dans les deux dimensions d'une surface sensorielle. La grande variété des récepteurs moléculaires capables de reconnaître les molécules odorantes montre qu'une association précise et particulière de récepteurs activés participe au traitement de l'information olfactive. Ainsi, contrairement aux autres modalités sensorielles, le code des odeurs est de nature combinatoire.

Genèse de la perception olfactive

Avec son grand nombre de neurones interconnectés, le premier relais central, nommé bulbe olfactif, présente l'organisation requise d'un réseau qui servirait de support aux traitements et stockage de l'information sensorielle. Les recherches menées par notre groupe s'efforcent d'évaluer la nature de cet encodage et de déterminer les paramètres fondamentaux utiles pour cet encodage. Nos travaux neurophysiologiques montrent que les neurones principaux de ce relais forment des assemblées synchroniques. Fait remarquable, ces assemblées de neurones sont transitoires[8] et se distribuent dans l'ensemble du réseau.
Il semble qu'un lien existe entre la formation d'assemblée synchronique et la perception olfactive. Lorsque nous sentons un parfum, l'activité synchronisée des neurones principaux produit des rythmes rapides qui se manifestent par des bouffées d'ondes d'électroencéphalogramme de grande amplitude. Les travaux expérimentaux et théoriques sur le fonctionnement du bulbe olfactif et, au-delà, sur la perception olfactive, reposent largement sur la conception selon laquelle la répartition spatiale de l'amplitude de ces ondes est la représentation neuronale la plus pertinente de l'odeur. Remarquons que l'information sensorielle portée par le message nerveux olfacto-bulbaire est, comme celle véhiculée par le message des neurones sensoriels, une information distribuée. Ainsi, l'odeur est représentée par l'activité électrique d'une assemblée cellulaire et non par des neurones individuels. Selon cette combinatoire les possibilités d'encodage sont évidemment très nombreuses et remarquablement adaptées à la diversité également considérable du monde des odeurs.

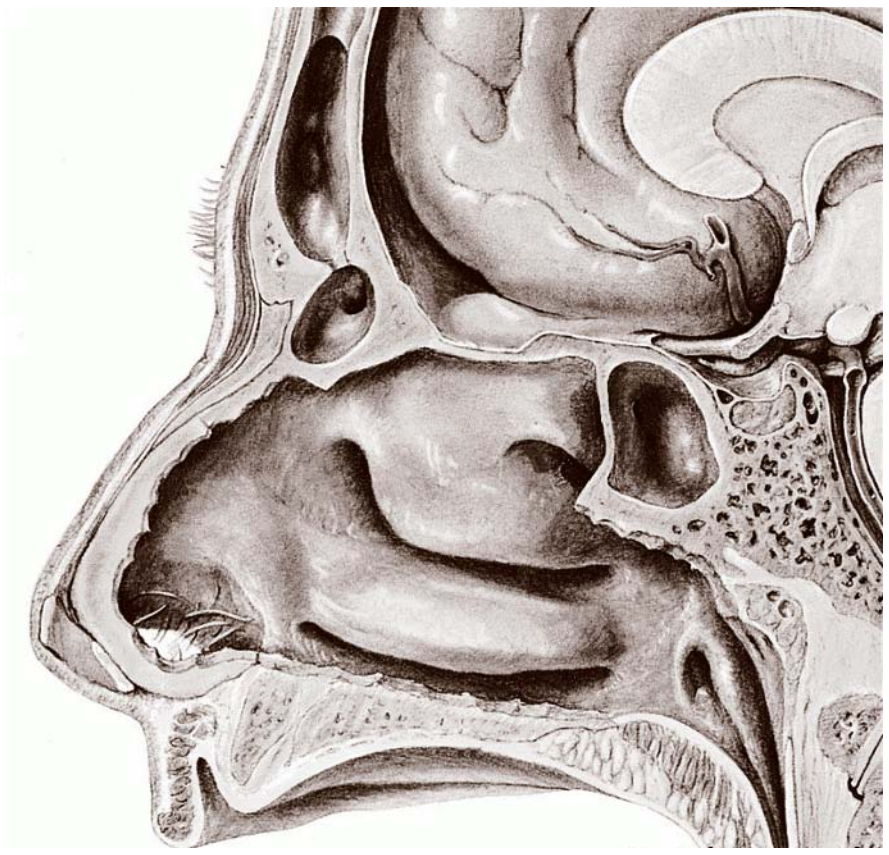


Planche anatomique de l'organe sensoriel de l'odorat, chez l'Homme.

Des rythmes à l'émergence des sensations olfactives

Rappelons que l'intégralité du monde physique qui nous entoure ne nous est accessible que par l'intermédiaire de nos sens. Des philosophes se sont penchés sur les relations qui relient le monde extérieur à nos perceptions, autrement dit sur les rapports de la pensée à son objet. Différentes écoles de psychologie et de philosophie se sont affrontées sur la question de savoir si la perception pouvait être comprise comme pure association de sensations élémentaires ou comme totalité d'emblée organisée. Ce débat philosophique s'ouvre maintenant aux Sciences de l'olfaction, il intéresse les spécialistes du cerveau et les conjectures font maintenant place à l'expérimentation. Il devient évident que les psychologues et les neurobiologistes qui souhaitent étudier le monde du réel doivent se placer, avant tout, dans le monde subjectif, et ceci au sens propre du terme c'est-à-dire "qui relève du sujet". Cette distinction n'est cependant pas toujours respectée ; nous confondons allègrement les objets et leurs représentations. En matière d'olfaction, ces notions théoriques qui relèvent de la phénoménologie sont importantes. Par exemple, nous pensons que le goût d'un bon vin réside dans le verre. Ainsi l'œnologie s'est-elle centrée en permanence sur la caractérisation des composés possédant des propriétés organoleptiques, c'est-à-dire celles qui sont susceptibles de stimuler les organes des sens. Un rapide inventaire des composés volatiles du vin révèle plus d'un millier de composés. Mais rappelons que cet inventaire ne témoigne en aucune manière du goût du vin. Pour que ce dernier existe, il faut un organe sensoriel connecté au cerveau, donc un dégustateur !

De nouveaux neurones dans le système olfactif : un nouveau mystère levé

Contrairement au dogme établi, des neurones néoformés apparaissent en permanence dans deux régions du système olfactif adulte. L'une de ces régions concerne l'organe sensoriel, l'autre le premier relais central. Si l'on commence à comprendre les détails cellulaires et moléculaires responsables de cette production neuronale tout au long de la vie, les conséquences fonctionnelles de ce renouvellement restent encore obscures. Nous avons montré que l'aptitude des souris à discriminer entre deux odeurs nécessite la production de ces nouveaux neurones. Les travaux les plus récents mettent en évidence le rôle important des processus permanents de prolifération et de migration des neurones olfactifs tant pour la perception que la mémoire olfactive. Il semble donc qu'un nombre critique de nouvelles cellules soit nécessaire aux facultés de discrimination et de mémorisation des odeurs.

L'organe olfactif, loin de jouer un rôle accessoire parmi les appareils sensoriels, apparaît essentiel pour la survie d'un sujet. La discrimination et la sélection des aliments, la détection des substances toxiques et des nourritures avariées, l'attraction et la reconnaissance des partenaires sexuels, l'établissement de liens parentaux et sociaux, manifestes chez l'animal et encore très présents chez l'humain, reposent sur une organisation neuronale complexe et bien particulière. Celle-ci suppose que les cellules réceptrices de l'organe sensoriel de l'odorat soient capables de contacter les aires spécialisées de notre cerveau. Fait remarquable, ces structures cérébrales sont également concernées par les processus émotionnels et mnésiques.
Une meilleure connaissance des caractéristiques de l'odorat permet non seulement de faciliter la compréhension des bases neurobiologiques de l'olfaction, ce sens mal-aimé plus important chez l'homme que ne le veut sa réputation, mais aussi d'ouvrir une porte sur les phénomènes inconscients qui régissent nos instincts et nos souvenirs ■

6 On parle de "protéines à sept segments transmembranaires".
7 De nombreuses substances actives, comme les hormones ou des neurotransmetteurs, agissent sur l'organisme grâce à des récepteurs couplés aux protéines G. Dans ces systèmes, la fixation du ligand conduit à un changement de conformation du récepteur et entraîne l'activation de la protéine G. Ce phénomène permet la mise en route d'une cascade de réactions intracellulaires qui peuvent notamment conduire à l'obtention d'une dépolarisation de la membrane cellulaire.
8 De l'ordre de centaines de millisecondes.

Grands prix scientifiques

La Séance solennelle de remise des Grands Prix de l'Institut de France a eu lieu le mercredi 14 juin 2006 dans la grande salle des séances. Les lauréats ont présenté leurs travaux et programmes scientifiques sur le thème « Imagerie, architecture et physiologie des cellules vivantes ».

Grand prix scientifique de la Fondation Louis D. de l'Institut de France

« Biologie aux temps ultra-courts : dynamique femtoseconde et microscopie non-linéaire »
par **Jean-Louis Martin**, directeur de recherche à l'INSERM, professeur à l'École polytechnique, laboratoire d'optique et biosciences, École polytechnique, CNRS UMR 7645, INSERM U696, Palaiseau.

« Biologie aux temps ultra-courts : contrôle cohérent »
par **Manuel Joffre**, directeur de recherche au CNRS, professeur chargé de cours à l'École polytechnique, laboratoire d'optique et biosciences, École polytechnique, CNRS UMR 7645, INSERM U696, Palaiseau

Grands prix scientifiques Simone et Cino del Duca de l'Institut de France

« Épigénétique et dynamique de la chromatine »
par **Geneviève Almouzni**, directeur de l'UMR 218 CNRS/Institut Curie « Dynamique nucléaire et plasticité du génome », Paris

« Guidage des vaisseaux sanguins au cours du développement embryonnaire »
par **Anne Eichmann**, directeur de recherche au CNRS, Direction de l'Équipe de recherche « Avenir », INSERM U36 « Développement vasculaire », Collège de France, Paris

« Les canaux ioniques dans tous leurs états »
par **Eric Honoré**, directeur de recherche au CNRS, UMR 6097, Institut de Pharmacologie moléculaire et cellulaire, Sophia Antipolis

« Dynamique des interactions hippo-campo-corticales : de la formation à la stabilisation des souvenirs »
par **Bruno Bontempi**, chargé de recherche au CNRS, UMR 5106, laboratoire de neurosciences cognitives, Université de Bordeaux ■

La recherche spatiale française[1]

La recherche spatiale revêt un caractère stratégique, ses applications sont multiples et son rôle dans la formation d'une communauté de haut niveau est indiscutable. La nécessité d'une politique spatiale ambitieuse à l'échelle de l'Union élargie, gardant un accès autonome à l'Espace, a donc été affirmée par le *Livre blanc pour une politique spatiale européenne*, élaboré en 2004 par la Commission européenne et l'Agence spatiale européenne.
Cependant, il apparaît à l'Académie des sciences qu'il faut insister sur la nécessité de faire les efforts nécessaires en termes de moyens humains, financiers et d'évolution des structures, afin que la recherche spatiale européenne puisse conserver sa compétitivité.

1 Rapport sur la Science et la Technologie n° 22, EDP Sciences, 17, avenue du Hoggar, Parc d'activités de Courtabœuf, BP 112, 91 944 Les Ulis Cedex A, www.edpsciences.org

Le présent rapport constitue une explicitation de la position de l'Académie, qui a entrepris une réflexion sur les recherches scientifiques utilisant les moyens spatiaux, réflexion divisée en deux étapes :

- la première partie de cet ouvrage propose une série de recommandations sur l'organisation générale de la recherche spatiale et son financement ;
- la seconde partie analyse la recherche spatiale française en dégageant les forces et faiblesses, et en s'appuyant sur des rapports préliminaires par discipline rédigés par les spécialistes impliqués dans ces recherches. Pour chaque discipline, un rapport synthétique et des recommandations spécifiques ont été élaborés.

Des recommandations générales sont également présentées.
Les activités liées à l'Espace dépassent aujourd'hui largement la recherche spatiale, même si celle-ci a été l'élément moteur de leur développement initial : c'est pourquoi ce rapport apporte de larges développements sur la recherche utilisatrice de l'espace et ne traite des infrastructures spatiales, de leur technologie et des programmes opérationnels que dans la mesure où ils ont des liens et/ou des implications forts avec la recherche spatiale ■

Sciences et pays en développement
Afrique subsaharienne francophone[1]

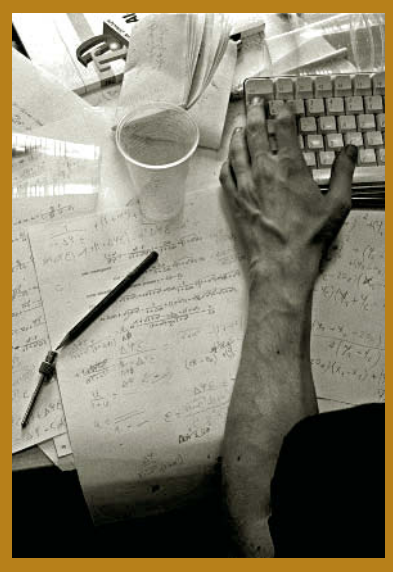
L'Académie des sciences vient de conduire une importante étude concernant l'effort national en matière de formation et de recherche avec et pour les pays francophones de l'Afrique subsaharienne.

Cet ouvrage, sans omettre de rappeler le contexte socio-économique et humanitaire global de cette région du continent africain, et les objectifs du millénaire, s'efforce de faire le point sur la situation précaire, des universités et centres de recherche africains. Cette situation est en effet à l'origine d'un exode massif, souvent irréversible, des étudiants et d'une dislocation plus ou moins étendue du tissu de la recherche locale, en dépit d'une prise de conscience générale aux plans panafricain et international.

Toutefois, l'accent est mis sur les projets de coopération actuellement développés au sein du réseau français des universités, établissements publics de recherche, écoles d'ingénieurs, etc., pour répondre aux défis posés.

Des secteurs d'activité aussi variés que : l'éducation de base, la formation aux sciences mathématiques et expérimentales, le rôle des technologies de l'information et de la communication, la recherche en santé publique et l'épidémiologie, la recherche agronomique et l'élevage, l'accès à l'eau et aux différentes sources d'énergies, la préservation et l'exploitation des ressources géologiques et minières ainsi que la place des sciences sociales et humaines, font l'objet d'analyses approfondies. On s'intéresse, ici, tant aux coopérations bilatérales (France-Afrique) qu'à la participation de la France aux grands programmes internationaux.

Le rapport comprend enfin une série de recommandations qui plaident en faveur d'une meilleure coordination des efforts déployés par les institutions scientifiques françaises, ainsi que pour un regard nouveau sur la formation universitaire et pour une politique plus affirmée en faveur du développement ■



la lettre de l'Académie des sciences
n° 19 / printemps 2006

Publication de l'Académie des sciences

23, quai de Conti 75006 PARIS
Tel : 01 44 41 43 68
Fax : 01 44 41 43 84
http : www.academie-sciences.fr

Directeur de publication :
Jean-François Bach

Directoire :
Jean-François Bach
Jean Dercourt

Rédacteur en chef :
Jean-Didier Vincent

Secrétariat général de rédaction :
Marie-Christine Brissot

Conception graphique
Nicolas Guilbert

Photographies :

p. 1, 3, 7, 9, 10, 12, 13, 14, 19, 20, photos N. Guilbert

p. 2, 3, 7, 11, 12, 16, 17, 18, photos (DR)

Photogravure & impression :
Edipro/Printreference™
01 41 40 49 00

n° de C.P. : 0108 B 06 337