

Où va l'informatique ?

par

■ **Gérard Berry** ■

Professeur au Collège de France

En bref

Alors que les États-Unis tentent désormais de freiner le développement chinois, en particulier dans le domaine des hautes technologies numériques, le Vieux Continent reste à la traîne. En France, nous peinons depuis des décennies à prendre la mesure des enjeux numériques, considérant avec condescendance la science informatique. Il est donc illusoire de penser rattraper notre retard face aux GAFAs et autres Baidu. Néanmoins, grâce à la qualité des mathématiciens et des chercheurs que nous formons, qui trop souvent partent vers la Silicon Valley ou l'Asie, des champs entiers d'innovation nous restent ouverts. Les explorer nécessiterait cependant une volonté politique, nationale, mais aussi européenne, ferme et ambitieuse, une réforme en profondeur de l'enseignement de l'informatique pour les plus jeunes et une collaboration des scientifiques de toutes disciplines dans le cadre d'une véritable culture partagée du numérique.

Compte rendu rédigé par Pascal Lefebvre

L'Association des Amis de l'École de Paris du management organise des débats et en diffuse les comptes rendus, les idées restant de la seule responsabilité de leurs auteurs. Elle peut également diffuser les commentaires que suscitent ces documents.

Le séminaire Management de l'innovation est organisé avec le soutien de la Direction générale des entreprises (ministère de l'Économie et des Finances) et grâce aux parrains de l'École de Paris du management :

Algoé¹ • Carewan¹ • Conseil régional d'Île-de-France • Danone • EDF • Else & Bang • ENGIE • FABERNOVEL • Fondation Roger Godino • Groupe BPCE • Groupe Caisse des Dépôts • Groupe OCP • GRTgaz • HRA Pharma² • IdVectoR² • IPAG Business School • L'Oréal • La Fabrique de l'industrie • MINES ParisTech • Ministère de l'Économie et des Finances – DGE • RATP • Renault-Nissan Consulting • SNCF • Thales • UIMM • Ylios¹

1. pour le séminaire Vie des affaires
2. pour le séminaire Management de l'innovation

Invisibles!

J'ai débuté dans la recherche en intégrant l'École des mines en 1970. Je suis ensuite parti à Sophia Antipolis, dans un laboratoire commun des Mines et de l'Inria¹ où je suis resté jusqu'en 2001. J'ai beaucoup travaillé sur les logiciels embarqués, créé le langage et le système de programmation Esterel, participé à ses premières industrialisations, puis rejoint en 2001 Esterel Technologies, entreprise devenue depuis américaine et désormais leader mondial des logiciels de très haute sécurité, utilisés en particulier par l'aviation. Je suis revenu à l'Inria de 2009 à 2012.

En 2007, j'ai passé une première année au Collège de France, dans le cadre de la chaire Innovation technologique Liliane Bettencourt, avec un cours intitulé *Pourquoi et comment le monde devient numérique*. J'y suis revenu en 2009 pour intégrer la chaire Informatique et sciences numériques, nouvellement créée par l'Inria, avec le cours *Penser, modéliser et maîtriser le calcul informatique*. Enfin, en 2012, j'ai été nommé professeur titulaire de la première chaire d'informatique du Collège, avec pour thème de cours et de recherche *Algorithmes, machines et langages*.

Pour ma dernière année au Collège de France, j'ai proposé comme thème : *Où va l'informatique?* En effet, depuis mes débuts, j'ai constaté qu'il n'existait aucun moyen pour un humain normal de comprendre de quoi il retournait vraiment en matière d'informatique. Les journaux en parlent mal et les informaticiens en parlent peu au grand public. Il y a deux ans, j'ai donc écrit un livre² qui explique ce qu'est vraiment l'informatique et pourquoi, en Europe, nous avons consciencieusement raté son train depuis sa naissance. En 1983, Maurice Nivat et moi-même avons rendu à Laurent Fabius, alors ministre de la Recherche, un rapport intitulé « L'Enseignement de l'informatique », dans lequel nous soulignons l'urgence de la question. En 2013, nous avons à nouveau rendu un rapport, à l'Académie des sciences, intitulé « Enseignement de l'informatique : il est urgent de ne plus attendre », dont le contenu était sensiblement identique. En trente ans, il ne s'était donc rigoureusement rien passé! Nous le payons très cher à l'heure actuelle, car nous sommes tout simplement invisibles d'un point de vue industriel, même si nous prétendons, en 1983, être la deuxième puissance logicielle au monde en matière de recherche, ce qui était d'ailleurs le cas, mais que le reste de la planète semblait ignorer.

Un mode de pensée fondamentalement différent

L'informatique diffère, au plus haut point, de tout ce que nous avons pu réaliser auparavant. Les XIX^e et XX^e siècles étaient fondés sur le triangle classique matière-énergie-ondes, largement intuitif et sur lequel, jusqu'aujourd'hui, l'intégralité de l'enseignement a été fondée. Puis sont arrivés l'information et l'algorithme, qui conduisent à une façon radicalement nouvelle et presque inversée de penser et de faire.

Toutes les propriétés de l'information sont en effet différentes de celles de la matière ou de l'énergie : elle se duplique parfaitement à l'identique, elle est indifférente à la nature de son support, etc. C'est cette pensée qui s'est désormais imposée, en particulier dans la Silicon Valley et en Asie, ce dont nous n'avons pas vraiment pris la mesure. Dès lors, nous continuons à nous raccrocher à des catégories mentales dépassées, comme lorsque nous demandons combien "pèse" un fichier, en nous rassurant par l'usage d'un vocabulaire familier.

1. Inria : l'Institut national de recherche en informatique et en automatique est un établissement public à caractère scientifique et technologique français spécialisé en mathématiques et informatique.

2. Gérard Berry, *L'Hyperpuissance de l'informatique – Algorithmes, données, machines, réseaux*, Odile Jacob, Paris, 2017.

Pendant très longtemps, en particulier dans le système d'enseignement, on a pensé que l'informatique était un outil et non une science. En 1985, alors que je siégeais au conseil d'enseignement de l'École polytechnique, c'était la question posée par les physiciens et les mathématiciens. Depuis, il y a été décidé qu'il s'agissait bien d'une science et qu'il était donc nécessaire de l'enseigner. L'expression "outil informatique" a formidablement pollué notre vision d'une informatique qui n'est pas un outil, mais un générateur d'outils, au même titre que la physique. Une autre idée désastreuse a été : « *Il suffit de savoir s'en servir!* », qui nous a très clairement mis en situation de dépendance totale vis-à-vis des États-Unis et de l'Asie!

Il est essentiel de comprendre que l'information et les algorithmes sont les mêmes partout. Quel que soit le domaine où ils s'appliquent – médecine, musique ou mécanique –, ils ont une forme unique et requièrent les mêmes méthodes, ce qui implique que ce soient les mêmes études qui y préparent.

Ensuite, pour traiter l'information, il existe une machine universelle. Cette notion fondamentale, révélée par Alan Turing, pose que toutes les machines informatiques sont équivalentes et qu'il n'y a pas de différence de nature entre le supercalculateur et le programmeur d'une machine à laver. L'un et l'autre font exactement la même chose, certes ni au même prix ni dans le même temps, mais cela s'arrête là. De plus, une difficulté mentale majeure n'a toujours pas été surmontée : le raisonnement et l'action en matière d'information sont complètement différents de ceux que l'on met en œuvre à propos de la matière et de l'énergie, sans toutefois être incompatibles.

L'informatique est donc extraordinairement puissante, mais son appropriation se fait au prix d'inversions mentales, façons de penser inverses de celles qui étaient les nôtres depuis des siècles. Ainsi, avant que le téléphone portable, objet profondément informatique, n'arrive en France, lorsque l'on appelait quelqu'un, la seule information dont on disposait était qu'il était chez lui, ou pas. Désormais, la question que l'on pose à son interlocuteur lorsqu'il décroche est « *T'es où?* », parce que celui-ci peut être n'importe où dans le monde. Une telle abolition de la géographie dans la communication n'a rien d'anodin.

Les industries informatiques ont délibérément construit ces inversions. Ainsi, au XX^e siècle, la photographie était une affaire assez simple : on amenait sa pellicule au labo et le lendemain, on repartait avec ses tirages en poche, puis, éventuellement, on les envoyait à quelqu'un par la poste. Désormais, la photo, aussitôt prise, est envoyée et instantanément reçue. Facebook traite ainsi deux milliards de photos par jour! On n'a donc pas seulement changé d'ordre de grandeur, on l'a fait à la puissance $n!$ À cette occasion, la photo aussi a changé de statut et, de souvenir précieusement gardé, elle est devenue un objet qui, comme la parole, a vocation à être éphémère.

Au XX^e siècle toujours, s'orienter sur une carte imposait de l'acheter et de connaître son échelle ainsi que le lieu où l'on se trouvait. Aujourd'hui, notre smartphone nous indique précisément où l'on est et le chemin à suivre vers la destination qu'on lui a indiquée, où que ce soit dans le monde. Quand il s'agit d'apprendre à des écoliers à s'orienter, cela n'a rien d'insignifiant, car la carte papier relève désormais pour eux davantage du cours d'histoire que de celui de géographie!

Les données et les algorithmes

À la différence des sciences dites naturelles, l'informatique est une science de construction. N'ayant pas à étudier la nature, puisqu'elle en fabrique une autre, cette nouvelle science progresse à grande vitesse. Elle repose sur quatre piliers : les données, les algorithmes, les programmes et les machines, auxquels s'adjoignent des interfaces et des interactions qui sont très souvent les points faibles des systèmes informatiques.

Quant aux usagers, ce sont désormais et majoritairement des objets. En effet, s'il se produit actuellement un peu plus de cent milliards de microprocesseurs par an, seulement 3% d'entre eux sont utilisés par des humains, sous forme d'ordinateurs ou de téléphones, les autres étant massivement dédiés à des objets. La 5G, censée améliorer la communication entre les hommes, a en réalité pour contrainte de permettre la connexion d'un million d'objets au kilomètre carré, avec un temps de latence d'une milliseconde et une très haute fiabilité.

Les données numérisées sont en fait assez différentes des informations proprement dites et le passage des unes aux autres est très délicat. Si on dit qu'il fait "20 degrés", c'est une donnée brute. L'information qu'elle porte dépend du contexte : en Europe, où la mesure est le degré Celsius, il fait bon, aux États-Unis, avec le degré Fahrenheit, il gèle! Grâce aux progrès de la physique, les données deviennent désormais massives et sont majoritairement produites par les machines et des capteurs de toutes sortes. Les humains, là encore, n'en produisent qu'une infime partie – bien sûr, très importante pour eux.

Les algorithmes, quant à eux, sont des procédés de calcul dont le nom vient de Muhammad Ibn Mūsā al-Khwarizmi, savant persan du IX^e siècle, dont nous ignorons superbement que nous lui devons, entre autres, les chiffres dits arabes, l'algèbre et la formule de résolution de l'équation du deuxième degré! L'algorithme est en fait quelque chose de conceptuel, qu'il ne faut pas confondre avec le programme, texte beaucoup plus détaillé, écrit dans un langage différent et qui tourne sur un ordinateur avec un nombre considérable de détails en plus.

Les algorithmes évoluent considérablement et font de plus en plus des tirages au sort, les probabilités, avec leurs propriétés prodigieuses, faisant désormais la loi en la matière. C'est ainsi que Google, parmi le milliard de pages correspondant à votre requête, est capable de sélectionner, en un dixième de seconde, les dix réponses les plus susceptibles de vous intéresser.

Les langages de programmation se répartissent en deux groupes. Certains sont très rigoureux, comme Ocaml, et prennent de plus en plus de place dans l'industrie; la France est toujours le leader mondial pour ces langages. D'autres sont moins rigoureux, comme Python ou Javascript. Ils sont tout à fait utiles et très populaires, mais ne présentent guère de garanties de sécurité.

La révolution des programmes

En couplant informatique et physique, il est désormais possible de faire des choses complètement inaccessibles aux procédés physiques. Ainsi, un appareil photo argentique, du fait des propriétés de l'optique, courbe les lignes droites lors d'une prise de vue au grand angle, sauf à utiliser des objectifs hors de prix. Avec un simple smartphone, le résultat est immédiatement sans défaut. Connaissant les lois régissant ces courbures, qui constituent une information à partir de laquelle on peut appliquer des calculs, l'algorithme en corrige les effets. Tous les appareils et téléphones modernes savent le faire, tout comme ils savent arranger la lumière, la couleur, etc.

En photo argentique, quand vous appuyez sur le déclencheur, la photo est prise et non modifiable. En photo numérique, le processus commence seulement, car des dizaines d'algorithmes s'appliquent aussitôt à améliorer la prise de vue. Avec le petit capteur d'un smartphone, une telle qualité de résultat serait impossible du fait de la seule physique. En photographie classique, il est également impossible d'avoir simultanément le premier plan et l'arrière-plan nets. Avec le numérique, l'appareil peut prendre une rafale de clichés à des profondeurs de champ différentes pour que les algorithmes synthétisent le tout en une photo parfaitement nette, ce qui est très important en macrophoto, par exemple.

Les applications de ces mêmes technologies numériques ont révolutionné les pratiques en matière d'imagerie médicale. Désormais, en lieu et place d'une kyrielle de scans, spectrographies et autres IRM, analysés un à un durant des heures, les médecins ont immédiatement à leur disposition une image unique en 3D combinée. En dermatologie, il est attesté qu'une photo prise avec un téléphone et analysée par des algorithmes d'apprentissage est déjà plus performante qu'un médecin pour détecter un mélanome. De même, un chirurgien en cardiologie en train d'insérer une sonde dans une artère ne regarde plus son patient, mais des écrans sur lesquels il observe, de l'intérieur, la progression de la sonde puis de l'opération, grâce à une image reconstituée en 3D. De plus, en superposition à cette image virtuelle, il aura l'indication de toutes les mesures qu'il aura à faire et de leur localisation. Toutes ces pratiques relèvent désormais de la routine.

L'inconscience sociale de ces réalités est énorme. Lors d'une IRM, le patient pense ainsi que c'est la machine qui donne l'image, comme le ferait un appareil photo argentique. Or, les données physiques produites par un tel examen sont gigantesques et totalement parasitées par de multiples bruits. Là encore, ce sont des algorithmes extrêmement sophistiqués, dont le coût représente 50% de celui de la machine, qui réalisent le travail. Il en va

de même pour les séquenceurs d'ADN, les robots martiens, etc. Réussir à poser la sonde Philae sur la comète Tchouri n'est certes pas une réalisation triviale, pas plus que de réussir la récente photo d'un trou noir, qui a nécessité le traitement de dizaines de pétaoctets³ de données et des mois de calculs. Les astronomes sont donc, depuis longtemps, devenus de grands informaticiens.

La simulation et la modélisation numériques ont débuté avec la 3D et la mécanique. Désormais, on simule tout, que ce soient les turbulences subies en vol par Ariane 5, une explosion d'étoile ou une intervention cardiaque, toutes choses délicates à réaliser en laboratoire. En réalité, la simulation consiste à enlever toute la physique d'un phénomène pour la remplacer par de l'information et à substituer des algorithmes aux lois de la nature. Dès lors, grâce à notre machine universelle, nous pouvons remplacer le temps physique par du temps de calcul, ce qui donne des possibilités d'observation inédites. On peut ainsi simuler rapidement des phénomènes lents, comme la fusion de deux galaxies, simulation que l'on pourra ensuite corroborer par des observations en différents points de l'espace. À l'inverse, on peut simuler lentement des phénomènes rapides, ce qui est très utile en biologie pour étudier le repliement des protéines en quelques microsecondes, sujet central dans cette discipline et qui fait l'objet de grands programmes d'études à l'échelle mondiale. Les simulations en temps réel sont également indispensables pour l'entraînement des pilotes, des apprentissages divers ou les jeux vidéo.

Quoique l'on veuille simuler, pour l'astronomie comme pour les jeux vidéo, ce sont les mêmes ordinateurs qui le permettent et ils travaillent tous de la même façon. Cette machine universelle est donc réellement une invention prodigieuse du XX^e siècle.

Les machines universelles et leurs circuits

Les machines universelles sont très variées, allant des grandes fermes de données aux smartphones et aux lave-linges, mais toutes utilisent la même technologie. Les ordinateurs classiques sont les seules machines, par ailleurs extrêmement performantes, dont nous disposons à ce jour dans l'attente d'éventuelles machines quantiques. En outre, des évolutions technologiques extrêmement intéressantes se profilent, qui vont nous permettre de penser enfin à autre chose qu'au silicium.

Le développement des circuits est un autre sujet d'émerveillement. Nous sommes là dans le domaine d'une physique classique, bien que le transistor soit déjà, en lui-même, un objet éminemment quantique. La loi de Moore, qui énonce que « *la puissance des ordinateurs double tous les 18 mois* », n'est qu'une approximation des médias, complètement inconnue des spécialistes. Ce n'est en effet pas une loi physique, mais une décision industrielle concertée qui a fait en sorte que le nombre de composants électroniques par unité de surface double tous les deux ans. Ainsi, en 1970, le premier microprocesseur comportait 2 400 transistors; aujourd'hui, on approche des 30 milliards et un circuit de smartphone en comporte désormais 9 milliards. Tout cela va encore doubler dans un futur proche.

On ne manquera sans doute pas de toucher un jour les limites de la physique, mais rien n'est évident, car si l'on sait à présent réaliser des transistors à un atome, rien ne dit que l'on ne saura pas faire plus petit. Aujourd'hui, la technologie de 5 nanomètres est atteinte et le 3 nanomètres est annoncé, le nombre de transistors étant inversement proportionnel au carré de cette taille. Néanmoins, les vraies limites seront industrielles et économiques, celles de la physique risquant de ne jamais être atteintes faute de moyens. Cette miniaturisation a de multiples avantages : plus les composants sont petits, moins ils consomment et moins ils ont besoin d'être refroidis; leur développement se fait désormais de façon modulaire, ce qui le rend plus aisé; le processus de fabrication reste à peu près le même quel que soit le nombre de transistors, de même que le prix par unité produite, etc.

Ces avancées ont permis un progrès conceptuel et technologique extraordinaire en substituant aux anciennes cartes électroniques un composant unique dont la taille, de quelques millimètres carrés, reste constante. Pour l'instant, les mémoires restent physiquement séparées et l'un des grands enjeux actuels est de les insérer

3. 1 pétaoctet = 1 million de milliards d'octets.

à l'intérieur des circuits. Il est intéressant de noter que le terme employé pour un tel circuit est IP, *Intellectual Property*, car ce qui est vendu par ses concepteurs, c'est uniquement son plan. Une entreprise anglaise, ARM Holdings, l'un des plus extraordinaires designers d'IP, ne vend ainsi que des "microfilms" logiciels de CAO (conception assistée par ordinateur) électronique et se garde bien de vouloir fabriquer des circuits.

Concevoir un circuit est devenu un enfer, car, l'objet final étant fabriqué en une seule fois, tout bug doit absolument être évité lors de la conception. Tout se passe comme si l'on devait fabriquer un Airbus en une seule journée sans qu'aucune modification ne puisse ensuite lui être apportée !

Le coût des usines est également un facteur limitant important. L'usine Samsung dédiée au 5 nanomètres a coûté aux alentours de 12 milliards de dollars et, pour le 3 nanomètres, qui ne sera produit qu'à deux ou trois endroits dans le monde, chaque usine coûtera de 20 à 30 milliards de dollars. Les États-Unis eux-mêmes peinent à suivre alors que l'Europe reste bloquée à 28 nanomètres. Global Foundries (anciennement AMD), un des plus grands fondeurs historiques, n'a pas réussi à réaliser du 7 nanomètres, non pour des raisons techniques, mais faute de moyens suffisants.

Où en est alors l'Europe ? Le leader des machines de fabrication est Hollandais. L'un des leaders mondiaux du design est, comme on l'a vu, l'Anglais ARM, mais, quelle que soit la qualité du design, personne en Europe n'est en capacité de fabriquer les circuits ultrafins actuels. L'Europe développe un projet de mégacalculateur, stratégiquement important, mais peu signifiant d'un point de vue commercial. Cette situation européenne, pour insatisfaisante qu'elle puisse être, n'est cependant pas en voie d'amélioration.

Néanmoins, elle nous permet de faire de la recherche de haut niveau dans des domaines où la concurrence est moindre, en particulier sur les alternatives au silicium pour l'informatique classique et sur l'informatique quantique, qui pourrait permettre dans certains cas (rares, mais pratiquement importants) des gains de vitesse exponentiels. L'Académie américaine des sciences explique, dans un récent rapport, qu'il n'est pas certain que nous puissions réaliser des machines quantiques pour des applications autres que de recherche. Par ailleurs, cette technologie encore potentielle remet en cause certains des systèmes de sécurité actuels, en particulier pour les connexions Internet. Tous les documents encryptés pourraient être à réencrypter, ce qui prendrait un temps infini, et ce pour un coût pharaonique. Nous ne sommes donc pas du tout sûrs qu'une telle machine existe un jour, mais c'est néanmoins dès aujourd'hui que nous devrions agir afin d'anticiper son apparition potentiellement dévastatrice.

Des millions de lignes de code

Il est courant de dire que le smartphone est devenu un véritable couteau suisse, ce qui est parfaitement erroné. En effet, un couteau suisse est moyen partout, alors qu'un smartphone est excellent partout : c'est un excellent appareil photo, un excellent magnétophone, une excellente caméra vidéo, un excellent moyen pour naviguer sur Internet et, accessoirement, un excellent téléphone ! Jamais nous n'avons fabriqué un objet présentant simultanément de telles qualités. Il n'est donc pas étonnant qu'il ait rencontré un tel succès. Cependant, pour réaliser les logiciels de ce smartphone, il a fallu quelques centaines de millions de lignes de programme écrites par des milliers de personnes, ce qui a pris énormément de temps.

Le problème de l'infrastructure logicielle n'est donc pas simple. Un système d'exploitation, par exemple, comporte plusieurs millions de lignes de code et Linux, réputé être le plus petit d'entre eux, en compte déjà plus de 30 millions. Un navigateur est, lui aussi, un objet extraordinairement complexe, en particulier en matière de sécurité. Hélas, leur qualité n'est toutefois pas équivalente à celle requise lors des développements matériels, pour lesquels le droit à l'erreur n'existe pas, alors que les usagers sont d'une tolérance incroyable en matière de logiciels.

Aujourd'hui, 4 milliards d'humains utilisent Internet, et les réseaux sociaux, qui n'existaient pas au début de ce siècle, touchent la moitié de l'humanité. Il y a désormais plus d'abonnements à la téléphonie mobile que d'habitants sur la planète. Dans le même temps, l'arrivée de la 5G devrait encore renforcer le secteur des télécoms.

Naguère, la France avait développé le Minitel, qui n'était qu'un réseau géré par un opérateur de télécommunication unique. Aujourd'hui, par construction, ce sont les utilisateurs eux-mêmes qui gèrent le réseau et c'est pour cela qu'il fonctionne efficacement. Il n'y a pas de gestion centrale d'Internet, juste des routeurs qui font transiter des paquets d'information circulant sur le réseau à l'aide d'algorithmes probabilistes sophistiqués.

Le *cloud computing*, récemment apparu, est un phénomène intéressant, car oscillant dans un dialogue permanent entre centralisation et décentralisation. Ainsi, dans un pays dont la langue vous est inconnue, il vous suffit de vous adresser à votre smartphone, qui traduira instantanément vos propos. Ce n'est évidemment pas lui qui effectue ce travail, mais une ferme informatique du *Cloud*, quelque part dans le monde. Grâce aux nouveaux algorithmes d'apprentissage par *deep learning*, cette traduction fonctionne extraordinairement bien, alors que c'était jusque-là jugé parfaitement impossible par les linguistes.

Le passage de l'ordinateur au smartphone a induit un changement d'usage radical, en abandonnant l'affichage multifenêtre au profit d'une fenêtre unique, ce qui modifie totalement l'interaction et a fait disparaître les programmes intégrés classiques au profit d'innombrables applications. Mais certains sites ne sont pas parvenus à s'adapter et ne sont toujours pas visibles sur les smartphones. Le monde s'est donc clivé, entre ceux qui ont compris que l'ordinateur était un objet en voie de disparition et les autres.

Désormais, l'internet des objets a fait son apparition. Sur Internet, il y a aujourd'hui plus d'objets qui communiquent entre eux que d'humains. Pour ne parler que des voitures, si l'aspect extérieur n'a guère changé au fil du temps, l'intérieur a été révolutionné avec une bonne trentaine d'ordinateurs embarqués et une centaine de millions de lignes de code. Mais hélas, aucune mesure de certification de sûreté et de sécurité ne leur est imposée, à l'inverse de l'aéronautique où il existe des certifications légales très fortes.

Le recentrage sur les données, terme que je préfère à celui d'intelligence artificielle, est porteur d'un grand changement, mais sa particularité est que les scientifiques n'ont qu'une idée très vague de pourquoi et comment marchent les algorithmes d'apprentissage profond. La reconnaissance d'image est un domaine dans lequel on obtient des résultats très performants, de même que celui de la traduction, mais l'on reste très loin encore d'une véritable intelligence.

La numérisation progressive de la société

Il est remarquable que l'évolution de la numérisation des services publics français soit très rapide, bien que non uniforme. Il est également intéressant de considérer l'apparition de nouveaux acteurs, telles les plateformes, pour qui l'information importe avant tout. Sur ce point, la règle du *winner takes all* est un problème redoutable. Booking.com est ainsi un acteur qui a considéré que, pour être le roi de l'hôtellerie, il ne fallait surtout pas avoir d'hôtel, car un hôtel, c'est de la matière, donc des ennuis. Pour ce site, ce qui importe, c'est de savoir qui veut aller où, et ce partout dans le monde, informations indispensables dont les hôteliers eux-mêmes ne peuvent pourtant pas disposer. Objectivement, pour la musique ou la vidéo, il n'est nul besoin d'avoir des supports physiques. Cependant, dans notre société, le problème est que l'on parle beaucoup des effets indésirables de cette numérisation sans vraiment se pencher sur leurs causes.

Certains sites, comme Marmiton.com, sont des réussites considérables, de même que ceux promouvant des projets à petite échelle. Wikipedia est absolument remarquable, en particulier dans le domaine des sciences, et a supplanté toute autre encyclopédie. Linux réunit 20 000 contributeurs autour de son logiciel collaboratif. NTP, qui synchronise tous les ordinateurs du monde en 2 millisecondes dès leur mise sous tension, est également un projet collaboratif, extraordinaire quoique méconnu, associant des milliers de gens depuis vingt ans à travers le monde. Internet Archive, sauvegarde volontaire du Web, archive 330 milliards de pages et déjà plus de livres que la bibliothèque du Congrès américain. Software Heritage, de Roberto di Cosmo, chercheur détaché à l'Inria, est la grande mémoire des logiciels et a réuni à ce jour 4,5 milliards de fichiers sources et 83 millions de projets logiciels, artefacts dont il s'agit de préserver tout l'historique.

Toutes ces réalisations et bien d'autres sont nées de la propagation de la pensée informatique dans la société.

Des succès, mais aussi des échecs

Dans les domaines de l'aéronautique, de la photographie, de la génomique, entre autres, la numérisation est un réel succès. Dans l'automobile, des avancées significatives sont d'ores et déjà opérationnelles, bien qu'il reste encore beaucoup à faire. En revanche, il existe aussi de nombreux échecs, dont la plupart sont dus à des bugs. Or le bug est quelque chose que les gens ont du mal à comprendre.

J'ai naguère été mandaté par l'IRSN⁴ pour conduire une expertise sur les systèmes d'arrêt d'urgence des centrales nucléaires, processus qui ne peut être que logiciel. Au sein du conseil scientifique, on m'a demandé la probabilité qu'il reste un bug dans le programme. Ma réponse a été : « Zéro, s'il n'y en a pas ; un, s'il en reste. » Personne n'a semblé comprendre et l'on m'a rétorqué que toute panne ayant une probabilité de survenue, elle ne pouvait être ni de zéro ni de un. Je leur ai alors demandé quelle était la probabilité que le théorème de Pythagore soit vrai. « Question idiote ! », m'a-t-on rétorqué. « Pas du tout, c'est la même que la vôtre ! Un programme informatique est un théorème de maths, ce n'est pas une machine, et un théorème ne tombe jamais en panne ! » Un logiciel fait ce qu'on lui a dit de faire, rien d'autre. Un bug est donc une panne humaine, celle du programmeur. Tant qu'on ne certifiera pas sérieusement ces productions, on sera confronté aux bugs. C'est là un très sérieux problème, auquel est particulièrement confronté le secteur automobile.

Nous ne sommes pourtant pas inéluctablement confrontés à ce problème. Les logiciels des lignes 1 et 14 du métro automatique parisien n'ont jamais eu besoin d'être testés, car ils ont fait l'objet d'une vérification mathématique. Un professeur au Collège de France, Xavier Leroy, a aussi pu créer un compilateur C, l'un des objets les plus complexes de l'informatique, dont il a été prouvé mathématiquement qu'il ne contenait aucune erreur, alors que tous les concurrents en ont encore un bon nombre.

La sécurité informatique

Dans son classement des enjeux de sécurité les plus importants, le *World Economic Forum* a placé en 3^e et 4^e positions les attaques informatiques contre les infrastructures et celles contre les systèmes, juste après les événements climatiques extrêmes et les catastrophes naturelles.

Ainsi, sur un unique site de hackers, on a pu trouver 773 millions de paires adresses-mots de passe. Une loi extrêmement simple explique pourquoi ceux-ci sont si faciles à craquer : si l'on s'en souvient, c'est qu'ils ne valent rien ! La seconde loi est qu'il ne faut avoir qu'un seul mot de passe, celui du coffre-fort informatique, géré par un gestionnaire spécifique, où tous les autres mots de passe, que l'on ne verra jamais, sont en sécurité.

En ce qui concerne les attaques sur les données, il faut savoir que le virus Wannacry a attaqué 83 hôpitaux en Angleterre, certains d'entre eux devant évacuer leurs patients, car plus rien ne fonctionnait. Tous leurs ordinateurs tournaient sous Microsoft XP, que Microsoft recommande pourtant depuis longtemps de ne plus utiliser. De son côté, Saint-Gobain a perdu 200 millions d'euros dans cette même attaque. Lors d'une autre attaque, l'Ukraine a vu une partie de son système électrique paralysée.

Les failles sur les données tiennent aussi à la faiblesse des cryptages, parfois obsolètes sur des machines anciennes, ce qui a conduit des chercheurs (heureusement pas des hackers) à construire des attaques applicables à 18 % des serveurs sur Internet. Il existe également des failles logiques dans la conception des systèmes : la nouvelle norme WIFI WPA3 vient ainsi d'être totalement cassée par des chercheurs. Quant à fabriquer une attaque, vous pouvez trouver sur Internet, pour une somme modeste, des kits logiciels et toute l'expertise d'informaticiens plus ou moins mafieux.

Aux États-Unis, il a été montré que le contrôle total des Jeep Cherokee pouvait être pris par le biais de leurs autoradios. Une autre étude a mis en évidence que la plupart des pacemakers contenaient des milliers de trous de sécurité généralement triviaux. Un autre danger est celui des canaux cachés au sein même de tous

4. IRSN : Institut de radioprotection et de sûreté nucléaire.

les microprocesseurs et qui permettent, à qui sait en user (ce qui est difficile), d'accéder à la totalité des mémoires de l'utilisateur.

L'hyperpuissance de l'informatique n'est donc pas un vain mot. Bien mise en œuvre, ses réussites sont extraordinaires, mal conçue, ses conséquences peuvent être graves, mais la différence entre les deux ne se voit pas. À l'avenir, il semble donc inconcevable de continuer à tolérer cet état de fait et urgent d'imposer des contraintes de sécurité plus strictes.

Débat



L'émotion et l'algorithme

Un intervenant : *L'intelligence artificielle est annoncée par certains comme devant se substituer à celle de l'homme dans bien des domaines, mais j'en doute. En effet, pour prendre un exemple trivial, le titulaire du permis de conduire le plus moyen sait anticiper un risque, bien mieux que ne le font les voitures autonomes, car il utilise sa capacité d'imagination et d'émotion dont sont dépourvus les systèmes embarqués.*

Gérard Berry : Je ne suis pas d'accord avec vous pour deux raisons. La première est que l'humain est plutôt inapte à la conduite. Si le freinage a été la première fonction informatisée sur une voiture, c'est que, pour être efficace, il doit agir différemment sur chacune des quatre roues en fonction de l'état du sol. Il faudrait donc quatre pédales de frein pour un individu qui ne possède que deux jambes! Si les moteurs consomment moins, c'est également parce que les décisions de l'homme ont été remplacées par celles d'algorithmes qui en optimisent bien mieux les réglages.

La seconde raison est que conduire est partout considéré comme une tâche sans intérêt, sauf dans nos pays latins. L'anticipation n'est donc pas la même partout ni pour tous et, de ce fait, la difficulté vient de la cohabitation entre véhicules classiques et véhicules autonomes. Dès lors qu'il n'y aura plus que des véhicules autonomes communiquant entre eux, ce problème ne se posera plus. Par ailleurs, il ne faut pas oublier que les accidents nucléaires de Three Mile Island et de Tchernobyl n'ont pas été causés par des défaillances techniques, mais bien par des erreurs humaines!

Il est également faux de penser que ces systèmes sont rigides et n'évoluent pas, car l'apprentissage existe pour eux aussi. C'est certes compliqué, mais dire que quelque chose est impossible est très imprudent, ce genre d'assertion étant souvent démentie par les avancées en la matière. Les linguistes pensaient qu'il serait impossible à un système de traduire correctement une langue du fait de sa spécificité humaine? Désormais, ils ne le disent plus! Le jeu de go était réputé invincible? Le programme Alpha Go Zéro a démontré le contraire en connaissant seulement les règles et en apprenant ensuite de ses erreurs.

Mais le go est un jeu où l'on peut perdre et rejouer, ce qui n'est pas le cas dans la vie. L'être humain est bon pour inventer, le reste du temps, il est souvent très médiocre. Ceci étant, les ordinateurs manquent cruellement d'humour et peut-être est-ce là la chose la plus résistante de toutes à la numérisation!

Int. : *Quels sont les liens entre informaticiens et musiciens?*

G. B. : Je préside le conseil scientifique de l'Ircam (Institut de recherche et coordination acoustique/musique) et je travaille ainsi directement avec les musiciens. Ce sont désormais des amateurs avisés de procédés informatiques pour composer, voire pour improviser. Bach, en son temps, était lui aussi un grand utilisateur d'algorithmes et aurait été comblé de disposer d'ordinateurs! Beaucoup de travaux sont également en cours

sur les rapports entre les émotions et la musique, les transformations de la voix, etc. Néanmoins, les musiciens se heurtent actuellement à des problèmes insondables. L'ordinateur pouvant désormais générer tous les sons et toutes les suites de sons, la notion fondamentale de partition, issue directement des tablatures d'instruments, s'efface. Chaque compositeur invente donc sa propre notation, mais, comme elle disparaît avec lui, cela rend parfois son œuvre indéchiffrable.

Former du capital intellectuel aujourd'hui

Int. : *Au trio classique en économie capital-travail-matières premières, ne faut-il pas désormais ajouter l'information ?*

G. B. : Le véritable changement est que le capital se déplace vers les hommes. Pour les GAFAs, les machines sont relativement peu onéreuses, mais, dans un domaine comme l'intelligence artificielle, ce sont les individus qui détiennent la connaissance et comme cette ressource est rare, elle coûte très cher. Pour faire un avion, le nombre d'ingénieurs nécessaires est relativement faible. En revanche, le nombre d'ingénieurs compétents requis pour faire des systèmes informatiques tels ceux que réalise Apple est extrêmement élevé. C'est pour cela que le fossé va, de plus en plus, s'élargir entre les entreprises qui peuvent se payer leurs services et les autres. Cela se ressent jusqu'au Collège de France, qui arrive de moins en moins à attirer des mathématiciens tant ils sont chassés par les grandes universités et attirés par des rémunérations inenvisageables en France. Or, dans le même temps, un élève français arrive encore au bac sans avoir étudié l'informatique une seule seconde dans son cursus, ce qui n'est sans doute pas la meilleure façon de former du capital intellectuel aujourd'hui.

Int. : *Si l'informatique est une science, dans quelle mesure les informaticiens peuvent-ils alors coopérer avec les autres sciences, en particulier sociales ?*

G. B. : Il faut que tous les scientifiques fassent comme les astronomes et certains physiciens, qui abandonnent petit à petit leurs préjugés sur l'informatique, et que les informaticiens fassent de même de leur côté à l'égard des sciences autres que la leur.

En ce qui concerne les sciences sociales, il se prend aujourd'hui beaucoup de décisions concernant la nature d'une donnée privée et son anonymisation. Il faudra bien trouver une réponse conjointe à ces questions qui ne relèvent pas de la seule compétence en informatique. Il en va ainsi, par exemple, dans le domaine médical où des épidémiologistes se voient refuser le droit de rechercher des corrélations entre un type de cancer, sur lequel des masses de données sont disponibles, et le lieu de vie des malades, pour lequel les données sont inaccessibles au titre de la protection de la vie privée. Il y a donc des équilibres à négocier rapidement, sauf à laisser les informaticiens en décider arbitrairement. On a alors besoin de connaissances mutuelles afin que chacun assimile la logique de l'autre.

Voici un autre exemple : à l'heure actuelle, les ministères concernés par l'éducation générale et supérieure souhaitent faire quelque chose d'inédit en France, consistant à corréliser la façon dont les lycées pensent l'orientation de leurs élèves avec ce que le supérieur attend d'eux. On dispose de toutes les données nécessaires dans les deux cas, mais, outre le manque de moyens pour les traiter, ce projet a rencontré beaucoup de résistances au sein de l'Éducation nationale, qui le percevait comme une remise en cause de ce qu'elle faisait jusque-là. Désormais, les moyens algorithmiques et techniques existent. Faut-il alors le faire ? Sans une compréhension mutuelle minimale, le travail ne peut même pas débiter.

Heureusement, l'attitude sociale face à l'informatique, qui était d'avoir une opinion bien arrêtée sans rien vouloir comprendre, est en train de lentement changer. C'est important, car pour comprendre, il faut d'abord vouloir comprendre. Or, dans la Silicon Valley ou en Corée, on entretient soigneusement le fait que les gens restent des utilisateurs passifs.

Int. : *Quelle serait aujourd'hui la transposition dans le monde de l'informatique de grandes ambitions nationales, comme la dissuasion nucléaire en son temps ?*

G. B. : Malheureusement, nous sommes déjà hors-jeu, car les forces en présence sont désormais gigantesques. L'arme nucléaire n'était pas chère en comparaison des enjeux actuels. En matière de sécurité informatique,

on a beau avoir d'excellents chercheurs, les mafias sont particulièrement puissantes et leur business extrêmement profitable, pour ne rien dire des États. Et, comparaison n'étant pas raison, l'arme nucléaire était faite pour ne pas avoir à s'en servir alors que le numérique est d'un usage universel et constant.

Il serait donc temps qu'en France, on songe à financer la recherche et l'enseignement en informatique et que l'on crée les postes qui font cruellement défaut aujourd'hui. La première étape est un nécessaire changement mental, fort loin d'être effectif aujourd'hui, puisque le cours d'informatique enfin mis au programme de la classe de seconde a été baptisé du titre abscons et fort peu incitatif *Sciences numériques et technologie*, le mot *informatique* paraissant être une grossièreté bannie du vocabulaire de l'Éducation nationale. L'urgence est aussi là !

■ Présentation de l'orateur ■

Gérard Berry : Ingénieur de l'École polytechnique (diplômé en 1970), ingénieur des mines (diplômé en 1973), il est l'auteur d'une thèse de doctorat d'État en mathématiques, soutenue en 1979. Il a été chercheur puis directeur de recherche à l'École des mines de Paris de 1970 à 2001 et directeur scientifique d'Esterel Technologies de 2001 à 2009. Il est directeur de recherche à l'Inria (Institut national de recherche en informatique et en automatique) depuis 2009. Il est également membre du conseil d'administration de l'Agence nationale de la recherche et a été membre de l'Academia Europaea en 1993, membre de l'Académie des sciences en 2002, membre de l'Académie des technologies en 2005. Il a reçu plusieurs prix : le prix Montpetit de l'Académie des sciences 1990, le prix Science et Défense 1999, le grand prix informatique de la fondation EADS 2005 et la médaille d'or du CNRS en 2014. Il est professeur au Collège de France depuis 2007.

Diffusion juillet 2019
