

Eclairage scientifique

Projet 1, 2, 3...codez !

Cycles 1, 2, 3 et 4

Résumé

Voilà un siècle, les ordinateurs n'existaient pas. Ils sont aujourd'hui plusieurs milliards. Ces ordinateurs et autres machines numériques que sont notamment les réseaux, les téléphones, les télévisions, les baladeurs, les appareils photos et les robots ont changé la manière dont nous :

- concevons et fabriquons des objets,
- échangeons des informations entre personnes,
- gardons trace de notre passé,
- accédons à la connaissance,
- faisons de la science,
- créons et diffusons des œuvres d'art,
- organisons les entreprises,
- administrons les États,
- etc.

Si les ordinateurs ont tout transformé, c'est parce qu'ils sont polyvalents. C'est en effet le même objet « ordinateur » qui permet d'utiliser des logiciels de conception assistée par ordinateur, des logiciels de modélisation et de simulation, des logiciels de courrier électronique et de messagerie instantanée, des logiciels d'échange de fichiers, des logiciels de lecture de vidéos et musique, des tables de mixage numériques, des machines à commande numérique, des encyclopédies, des cours en ligne, des bases de données, des blogs, des forums, ou encore des archives numériques. Cette polyvalence s'illustre par le nombre d'outils que les ordinateurs ont remplacé : métiers à tisser, machines à calculer, machines à écrire, téléphones, télévisions, appareils photos, lecteurs multimédias...

Dans tous ces usages, les ordinateurs traitent de l'information, c'est-à-dire qu'ils appliquent, d'une manière systématique, des opérations à des objets symboliques. La recherche d'un mot dans un dictionnaire, le chiffrement et le déchiffrement d'un message secret, l'addition et la multiplication de deux nombres, la fabrication des emplois du temps des élèves d'un lycée ou des pilotes d'une compagnie aérienne, le calcul de l'aire d'une parcelle agricole ou encore le compte des points des levées d'un joueur au Tarot sont des exemples de traitements d'informations.

Un procédé systématique permettant de traiter des informations s'appelle un algorithme. Les ordinateurs sont en fait des machines à exécuter des algorithmes et l'informatique est le domaine d'activité scientifique, technique et industriel en rapport avec le traitement automatique de l'information par des machines.

Cet éclairage scientifique sur l'informatique propose, après avoir raconté l'histoire de cette discipline, d'en explorer les 4 concepts fondamentaux : **algorithme, langage, machine et information.**

Éclairage scientifique

Voilà un siècle, les ordinateurs n'existaient pas. Ils sont aujourd'hui plusieurs milliards. Ces ordinateurs et autres machines numériques que sont notamment les réseaux, les téléphones, les télévisions, les baladeurs, les appareils photo et les robots ont changé la manière dont nous :

- concevons et fabriquons des objets ;
- échangeons des informations entre personnes ;
- gardons trace de notre passé ;
- accédons à la connaissance ;
- faisons de la science ;
- créons et diffusons des œuvres d'art ;
- organisons les entreprises ;
- administrons les États ;
- etc.

Si les ordinateurs ont tout transformé, c'est parce qu'ils sont polyvalents. C'est en effet le même objet « ordinateur » qui permet d'utiliser des logiciels de conception assistée par ordinateur, des logiciels de modélisation et de simulation, des logiciels de courrier électronique et de messagerie instantanée, des logiciels d'échange de fichiers, des logiciels de lecture de vidéos et musique, des tables de mixage numériques, des machines à commande numérique, des encyclopédies, des cours en ligne, des bases de données, des blogs, des forums, ou encore des archives numériques. Cette polyvalence s'illustre par le nombre d'outils que les ordinateurs ont remplacé : métiers à tisser, machines à calculer, machines à écrire, téléphones, télévisions, appareils photo, lecteurs multimédias...

Dans tous ces usages, les ordinateurs traitent de l'information, c'est-à-dire qu'ils appliquent, d'une manière systématique, des opérations à des objets symboliques. La recherche d'un mot dans un dictionnaire, le chiffrement et le déchiffrement d'un message secret, l'addition et la multiplication de deux nombres, la fabrication des emplois du temps des élèves d'un lycée ou la construction du planning des pilotes d'une compagnie aérienne, le calcul de l'aire d'une parcelle agricole ou encore le compte des points des levées d'un joueur au tarot sont des exemples de traitements d'informations.

Un procédé systématique permettant de traiter des informations s'appelle un algorithme. Les ordinateurs sont en fait des machines à exécuter des algorithmes et l'informatique est le domaine d'activité scientifique, technique et industriel en rapport avec le traitement automatique de l'information par des machines.

Cet éclairage scientifique sur l'informatique propose, après avoir raconté l'histoire de cette discipline, d'en explorer les 4 concepts fondamentaux : algorithme, langage, machine et information.

Petite histoire de l'informatique

L'histoire de l'informatique est celle d'une confluence. Deux courants de pensée, centrés l'un autour de la notion d'algorithme et l'autre autour de celle de machine, ont conflué dans les années 1940, pour aboutir à la conception des premiers ordinateurs, machines à exécuter des algorithmes. Deux

autres courants de pensée, centrés autour des notions de langage et d'information, ont, par la suite, rejoint cette confluence.

L'histoire de la notion d'algorithme remonte aux débuts de l'écriture, puisque les premières traces écrites que nous avons retrouvées sont des calculs comptables : les scribes de Mésopotamie connaissaient donc des algorithmes pour effectuer des additions et des multiplications, mais aussi des opérations plus complexes comme des calculs d'intérêts composés, d'héritages... Il est même possible de faire remonter l'histoire de la notion d'algorithme plus tôt, en étendant la notion d'algorithme à des méthodes systématiques qui opèrent sur d'autres objets que de l'information, ce qui permet de concevoir les méthodes de fabrication du bronze ou de la céramique, les recettes de cuisine ou encore les procédés de tissage des étoffes comme des algorithmes. L'histoire de la notion de machine est ancienne, elle aussi, même si on la limite aux machines opérant sur de l'information, puisque les machines d'Héron d'Alexandrie, les carillons des cathédrales, les machines à calculer de Schickard, de Pascal, de Leibniz, le métier à tisser de Jacquard... contiennent déjà des ébauches de traitement de l'information.

L'histoire de l'informatique est foisonnante et s'étale sur des siècles. Voici une sélection de personnages associés à cette histoire qui marquent des jalons importants. Cette sélection permet d'expliquer des concepts de l'informatique en profitant de l'attrait de l'approche historique pour véhiculer des concepts. Elle doit cependant être accompagnée d'une modération : il est délicat d'attribuer clairement telle ou telle contribution à une personne en particulier. Les découvertes se font en effet de manière incrémentale, chaque « inventeur » ayant construit sa propre contribution grâce à celles de ses prédécesseurs et, souvent, de ses contemporains. Cette modération permet surtout de dire que ces personnes, aussi géniales ont-elles pu être, ne sont pas l'unique vecteur de ces contributions, mais qu'on peut tout à fait reconnaître l'importance majeure de leur travail comme une pierre angulaire.



Muhammad ibn Mûsâ al-Khwarizmi

Al-Khwarizmi est un savant perse, vraisemblablement né dans les années 780. Il évolue dans un empire qui, sous le règne du calife Al-Mamun, encourage les sciences et les arts. Avec d'autres penseurs de son époque, il travaille dans « la maison de la sagesse », une école fondée par le calife à Bagdad pour étudier la géométrie, l'algèbre et l'astronomie. Son travail consiste d'abord à traduire des manuscrits scientifiques ou philosophiques grecs et indiens. Avec ces travaux il transmet à ses contemporains les connaissances de ces civilisations, comme le système décimal de numération. Puis, Al-Khwarizmi écrit ses propres textes

scientifiques et en particulier l'ouvrage *Kitâb al-jabr wa al-muqâbala* dans lequel on peut entendre le son *al-jabr* qui est à l'origine du mot « algèbre ». Il cherche à fournir des outils permettant de régler des problèmes de la vie quotidienne. Le mieux est encore de laisser Muhammad al-Khwarizmi expliquer lui-même que ses travaux sont destinés à faciliter les calculs « dont ont besoin les gens dans leurs héritages, dans leurs donations, dans leurs partages, dans leurs jugements, dans leurs commerces et dans toutes les transactions qu'il y a entre eux à propos de l'arpentage des terres, du creusement des canaux, de la géométrie et d'autres choses relatives à ses aspects et à ses arts (...) ». La marque d'Al-Khwarizmi dans l'histoire est fortement liée à celle des mathématiques avec ses travaux sur l'algèbre et la géométrie. Mais son nom est également indissociable de l'informatique car chaque procédure qu'il décrit pour résoudre un calcul est un enchaînement d'instructions qui sont des manipulations sur des nombres. D'ailleurs, le nom d'Al-Khwarizmi a été traduit par *Algorismi* en latin, et donnera plus tard le mot « algorithme ».



Johannes Gutenberg

Johannes Gutenberg naît vers l'an 1400 à Mayence, en Allemagne. Il apprend l'orfèvrerie puis met son expérience dans le travail des métaux au service de son invention : l'imprimerie. L'histoire de l'imprimerie précède toutefois largement celle de Gutenberg. Cette histoire, qui reflète l'importance pour l'humanité de rendre ses œuvres durables et d'en assurer la diffusion, remonte au VII^e siècle. Des gravures sur planches de bois sont alors utilisées pour reproduire des textes et des images sur des pages de papier ou de tissu. Les procédés permettant de reproduire une page à l'identique évitent de faire ce travail manuellement et connaissent beaucoup de progrès. Le plus significatif de ces progrès est probablement l'invention par Bi Sheng du caractère mobile d'imprimerie, dont on retrouve les premières traces en Chine, vers 1040. Ces caractères sont alors en terre cuite puis ils évoluent vers le métal. En Europe, Gutenberg applique un principe habituel dans l'histoire des inventions : il perfectionne des techniques existantes et les combine. Ainsi, il crée un alliage métallique qui se révèle très adapté pour l'imprimerie : le « plomb typographique ». Il propose également la « presse à imprimer », plus efficace que les techniques existantes pour appliquer la page sur la planche. Enfin, il met au point une encre plus épaisse et adaptée pour l'imprimerie. Son invention connaît un succès grandissant et se propage rapidement en Europe. L'imprimerie, sans avoir directement influencé l'informatique, entretient avec elle des liens intéressants. Tout d'abord, parmi les innovations de l'imprimerie, penchons-nous sur le caractère mobile d'imprimerie. Il permet de remplir une page avec les caractères de son choix et de façon réversible, alors qu'avant il fallait graver la page tout entière de manière irréversible. La planche devient donc reconfigurable. On peut de surcroît fournir une planche à une autre presse qui pourra à son tour (re)produire la même page. D'une certaine manière, le simple fait de rendre une machine configurable peut être vu comme un ancêtre lointain de la notion de programmation qu'on trouvera bien plus tard dans le métier à tisser de Jacquard, dans les machines de Babbage, et finalement dans les ordinateurs. Un autre lien important avec l'informatique se retrouve dans la révolution culturelle apportée par l'imprimerie. On estime qu'au XVIII^e siècle, le nombre total de livres imprimés était d'environ un milliard, ce qui a totalement changé l'accès au savoir et donc l'équilibre social.

Joseph Marie Jacquard

Joseph Marie Jacquard est un inventeur et mécanicien français né en 1752 à Lyon. Pour avoir été tisseur dans l'atelier de son père, il connaît bien le métier à tisser, qui permet de fabriquer des étoffes. Cet outil est alors très répandu, source d'emplois, mais sa manipulation est complexe. Elle mobilise plusieurs personnes pour faire passer les navettes (qui déroulent les fils) au bon moment dans la machine. Jacquard veut mécaniser le métier à tisser pour le rendre plus facile d'emploi et plus performant. Il reprend alors le système de ruban perforé que Basile Bouchon a ajouté au métier à tisser et la mécanisation proposée par Jacques Vaucanson. En améliorant et en combinant ces deux technologies, il contribue à l'expansion du métier à tisser. Le système de cartes perforées permet de définir sur un support solide les instructions de passage des navettes. Les cartes perforées sont ensuite placées dans le métier à tisser, qui les fait défiler à chaque étape de la confection du tissu. Quand une carte perforée passe dans le métier, l'emplacement des trous dans la carte fait changer la position



des fils et guide le passage des navettes. Chaque façon de trouser la carte correspond à un passage différent, et donc à un motif précis sur le tissu. En multipliant les cartes perforées, il devient alors possible d'écrire les instructions pour obtenir le motif complet à l'échelle du lé entier. Comme pour l'imprimerie de Gutenberg, le métier à tisser devient configurable. Un ensemble de cartes peut être copié à l'identique et utilisé sur un autre métier et cela donne exactement le même tissu avec les mêmes motifs. Le même métier avec un jeu de cartes perforées différent donne un autre tissu avec d'autres motifs.

L'invention de Jacquard permet de produire plus de tissus, plus rapidement : elle connaît un grand succès avec plus de 20 000 métiers construits. Parce qu'elle utilise des cartes perforées pour donner des instructions à une machine configurable, cette invention est considérée comme un des ancêtres de l'ordinateur. Le travail de Jacquard illustre bien le schéma habituel des innovations. Jacquard cherche son inspiration dans les technologies les plus avancées de son époque. Il adapte et améliore plusieurs inventions pour les combiner en un seul mécanisme plus facile à installer et à utiliser. L'avancée majeure obtenue par Jacquard sur le métier à tisser permet sa manipulation par un seul ouvrier, au lieu de six précédemment. Ces ouvriers, appelés « canuts » se sont révoltés par la suite en raison du manque de travail, souvent attribué à cette invention.



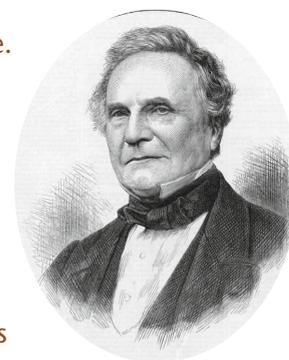
Mécanisme Jacquard (extrait d'un métier à tisser) exposé au musée des Arts et Métiers de Paris.

Charles Babbage

Charles Babbage, né en 1791, est un mathématicien et inventeur britannique.

À cette époque, bien avant l'invention du GPS, la navigation en mer se fait par l'observation de la position des étoiles dans le ciel, et par comparaison avec leur position attendue à cette date. Cette méthode permet de localiser le bateau sur le globe terrestre. Cela suppose de calculer, à l'avance, la carte du ciel jour après jour pour les semaines à venir. Malheureusement, les calculs sont longs et difficiles et les mathématiciens qui les réalisent risquent de se tromper, ce qui peut provoquer des accidents en mer. En 1821, Charles Babbage présente un projet de machine à la Société royale d'astronomie de Londres, pour accélérer ces calculs et les rendre plus fiables : la « machine à différences ». L'objectif de cette machine est de résoudre automatiquement, par un processus mécanique, les calculs complexes permettant de prévoir la position des étoiles.

Malheureusement pour Charles Babbage, sa machine demande des pièces mécaniques très fines, très précises, qui sont difficiles à fabriquer à son époque. Il n'arrive pas à construire sa machine, mais cela ne l'empêche pas de penser à une deuxième machine, encore plus performante. Babbage veut que cette deuxième machine ne se contente pas de calculer la position des étoiles. Il veut qu'elle soit capable d'exécuter n'importe quelle tâche qui pourrait lui être « décrite ». Pour cela, inspiré par le travail de Jacquard, il prévoit que la « description de la tâche » (une succession de calculs) soit réalisée avec des cartes perforées. Babbage ne réalise peut-être pas totalement le potentiel de son invention



et c'est Ada Lovelace qui va clairement comprendre tout ce que cette « machine analytique » rendrait possible. Malheureusement pour Babbage cette deuxième machine ne sera pas fabriquée non plus, faute de financements. Pourtant, avec elle, il avait posé tous les principes d'un ordinateur, dont la programmation par cartes perforées lui donnant les instructions à exécuter, comme ce sera le cas pour les premiers ordinateurs qui apparaîtront un siècle plus tard.

Ada Lovelace

Augusta Ada King (Ada Lovelace), est une mathématicienne britannique née le 10 décembre 1815 à Londres. Elle est la fille de lord Byron, poète britannique, qu'elle n'a pas connu. Sa mère, qui tient à ce qu'Ada reçoive une éducation scientifique soutenue, lui fait particulièrement étudier les mathématiques.

Cette éducation est assez atypique pour une jeune fille de cette époque. À 17 ans, Ada rencontre Charles Babbage et se passionne pour son travail sur la machine à différences, puis sur la machine analytique. Ada Lovelace et Charles Babbage forment un véritable duo scientifique. Ada écrit la plus complète des documentations sur la machine de Babbage et le soutient dans ses difficultés financières. Lui aide Ada à approfondir ses connaissances des mathématiques en



lui recommandant d'étudier avec les plus célèbres mathématiciens du royaume.

Alors que Babbage prévoit l'utilisation de sa machine pour des calculs numériques, Ada parle déjà d'une utilisation plus universelle. À l'époque, le calcul numérique est déjà traité de manière automatique par des machines comme la pascaline: il s'agit d'un ensemble de rouages permettant de faire uniquement des additions et des soustractions. Pour Babbage, sa machine était une sorte de « calculatrice programmable »: on pouvait d'abord lui expliquer le calcul qu'elle devait faire (avec des instructions, le « programme »)

puis lui donner les valeurs sur lesquelles appliquer ces instructions.

Visionnaire, Ada a su comprendre le potentiel de cette machine peut-être encore plus que Babbage lui-même. Elle décrit dans ses manuscrits la possibilité pour cette machine de manipuler des nombres, mais aussi des lettres ou n'importe quel type de symboles. Elle explique comment cette machine pourrait être programmée pour des tâches bien différentes du calcul numérique, comme la composition de morceaux de musique. Dans ses notes, Ada décrit l'ensemble des instructions qu'il faut donner à la machine de Babbage pour qu'elle réalise un calcul particulier: celui des termes successifs de la suite de Bernoulli. En réalisant cela, elle devient la première personne au monde à écrire un programme informatique. Autrement dit, le premier programmeur est une programmeuse!

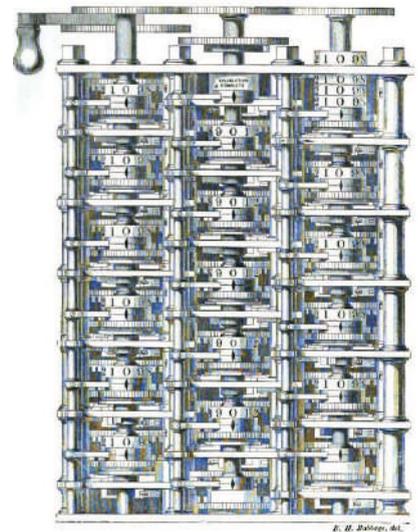


Schéma de la machine à différences de Babbage



Alan Turing

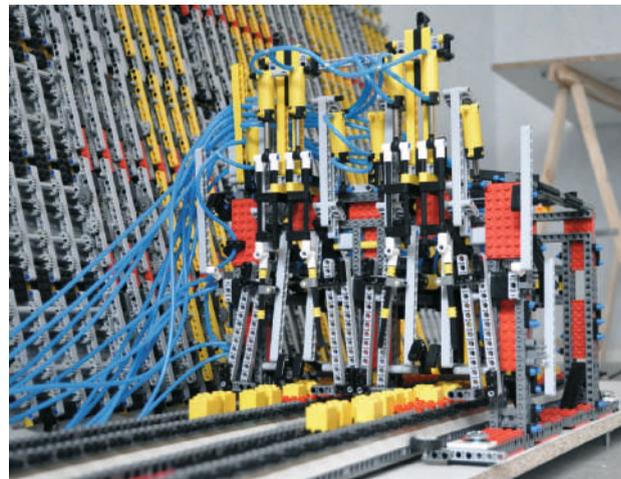
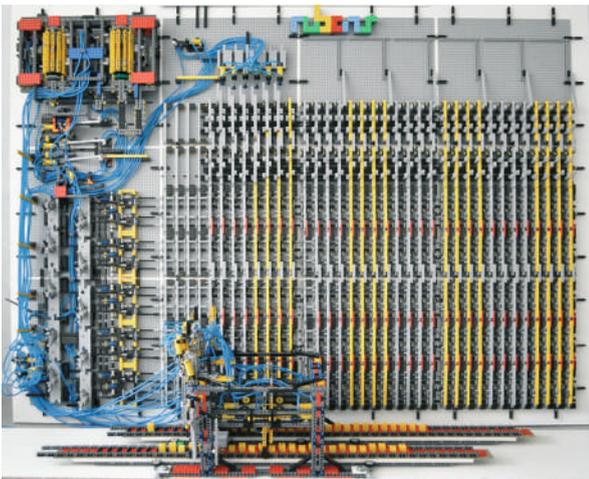
Alan Turing est un mathématicien et informaticien britannique, né en 1912. Il est souvent présenté comme le père fondateur de l'informatique – bien qu'il soit sans doute vain de chercher un tel père fondateur à ce qui est une création collective.

Ses contributions à l'informatique naissante sont multiples. Dans les années 1930 il donne l'une des premières définitions précises de la notion d'algorithme. Son objectif est alors d'étudier ce qui est « calculable ». Une multiplication, par exemple, est calculable car on peut la réaliser avec un certain nombre d'étapes. Mais quels sont les problèmes calculables ? C'est une question difficile, particulièrement étudiée dans les années 1930, sur

laquelle Turing se penche. En étudiant cette question, il propose une machine abstraite, c'est-à-dire une machine qui n'est pas destinée à être fabriquée concrètement, mais qui va lui servir à structurer son raisonnement. Imaginer son fonctionnement lui permet de mieux étudier ce qui est calculable ou pas. Cette machine se décrit très simplement. Elle utilise un ruban qui contient une suite de cases dans lesquelles sont inscrites des données (des 0 et des 1). La machine est capable de lire ce qu'il y a dans les cases, de se déplacer sur le ruban d'une case à gauche ou à droite, et d'écrire dans une case. Avec cette machine, Turing montre qu'il est possible de remplacer n'importe quelle autre machine, en écrivant sur le ruban le « programme » de la machine remplacée (par exemple la multiplication) et ensuite les « données » sur lesquelles appliquer ce programme. Cette façon de mettre ensemble le programme et les données dans une machine est fondamentale et se trouve aujourd'hui dans tous les ordinateurs.

Durant la Seconde Guerre mondiale, Turing joue un rôle majeur dans la construction de machines électromécaniques permettant de déchiffrer les messages secrets utilisés par l'armée allemande, sans en connaître la clé. Après la guerre, il travaille à la conception des premiers ordinateurs britanniques au National Physical Laboratory, puis à l'université de Manchester.

Il est aussi à l'origine de nombreuses autres idées sur le caractère phénoménologique de l'intelligence, sur la genèse des formes en biologie, sur la résolution approchée d'équations différentielles avec un ordinateur ou sur les preuves de programmes. Dans ce dernier cas, la problématique est la suivante: on peut faire des tests sur un programme pour vérifier sa stabilité et son exactitude. Si on trouve une erreur, on cherche alors à réparer le programme; mais si on ne trouve pas d'erreur par cette méthode, cela



Machine de Turing réalisée en Lego® (projet Rubens, ENS Lyon): vue globale (à gauche) et détail de la tête de lecture - écriture (à droite).

veut-il dire qu'il n'y a pas d'erreur ou qu'on n'a pas su la trouver? Existe-t-il une preuve mathématique que le programme est réellement sans erreur?

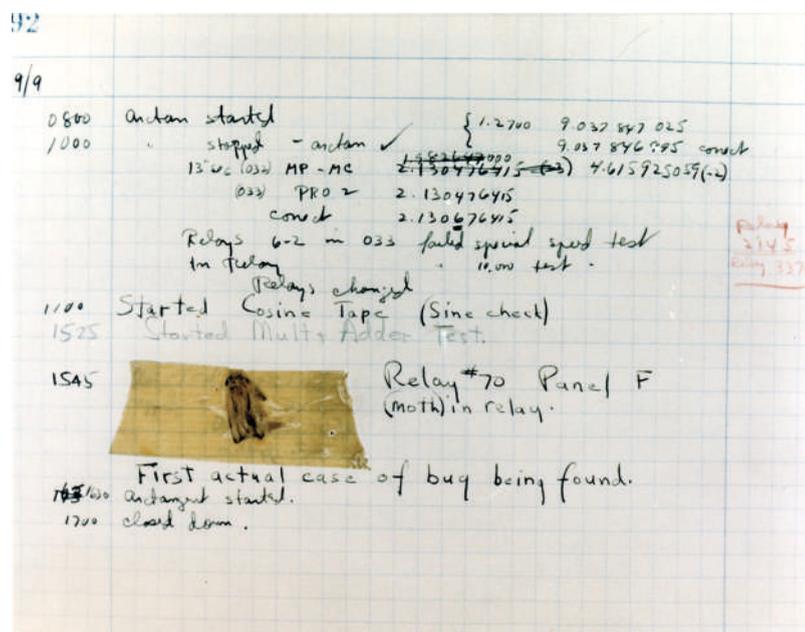
Condamné à la castration chimique en raison de son homosexualité, considérée comme un crime dans l'Angleterre de cette époque, Alan Turing s'est suicidé à 41 ans.



Grace Hopper

Grace Hopper est une mathématicienne américaine, née en 1906. Elle grandit en montrant une véritable curiosité pour les sciences et la technologie (par exemple, en démontant les réveils de sa maison les uns après les autres, jusqu'à arriver à en remonter un entièrement... à l'âge de 7 ans). Ses parents encouragent son goût pour les sciences tout en veillant à valoriser ses choix plutôt que de suivre la tradition en matière d'éducation des jeunes filles américaines. Ainsi, Grace sera une des rares femmes de son époque à étudier à l'université Yale et à obtenir un doctorat de mathématiques.

Pendant la seconde Guerre mondiale, elle entre dans la marine, où elle est nommée lieutenant, et rejoint l'équipe qui travaille sur le Mark 1, le premier ordinateur des États-Unis. Grace Hopper fait alors partie du premier groupe capable de programmer cette machine. Cet ordinateur prend ses instructions et ses données à l'aide de cartes perforées, comme cela était prévu sur la machine de Babbage, et comme le faisait le métier à tisser de Jacquard environ deux siècles auparavant. Dans les années 1950, Grace Hopper est motivée par l'idée que l'informatique devrait être accessible à un plus grand nombre de personnes, et pas seulement aux spécialistes des ordinateurs et aux mathématiciens. Elle défend l'idée que les cartes perforées sont un frein au développement de l'informatique et travaille sur un projet de « langage de programmation » qui serait proche de l'anglais et qui faciliterait l'écriture des programmes. La difficulté est alors de concevoir le programme capable de transformer des instructions de haut niveau (le langage proche de l'anglais et facilement compréhensible par un humain) en instructions de bas



Découverte d'un bug («insecte») par Grace Hopper, en 1947

niveau (le langage de la machine). Un tel programme est aujourd'hui appelé « compilateur » et c'est en partie à Grace Hopper que nous en devons l'invention. Son langage de programmation, appelé COBOL, sera le plus utilisé entre 1960 et 1980.

Grace Hopper n'est pas célèbre uniquement pour avoir rendu les ordinateurs plus accessibles aux non-spécialistes. Un jour, alors qu'elle programme ses ordinateurs avec des cartes perforées, l'un d'eux tombe en panne. Allant voir ce qui a pu se produire, elle trouve une mite coincée dans un des trous d'une carte perforée. Elle colle alors sur son rapport le morceau de carte avec un commentaire désignant l'insecte comme coupable de la panne. « Insecte » se disant *bug* en anglais, ce terme a été utilisé auparavant par les scientifiques, comme Thomas Edison, pour désigner une erreur due à l'intervention indésirée d'une de ces petites bêtes dans leurs expériences. En l'appliquant à l'exécution échouée d'un programme informatique, Grace en a fait le terme le plus utilisé pour qualifier cette situation.

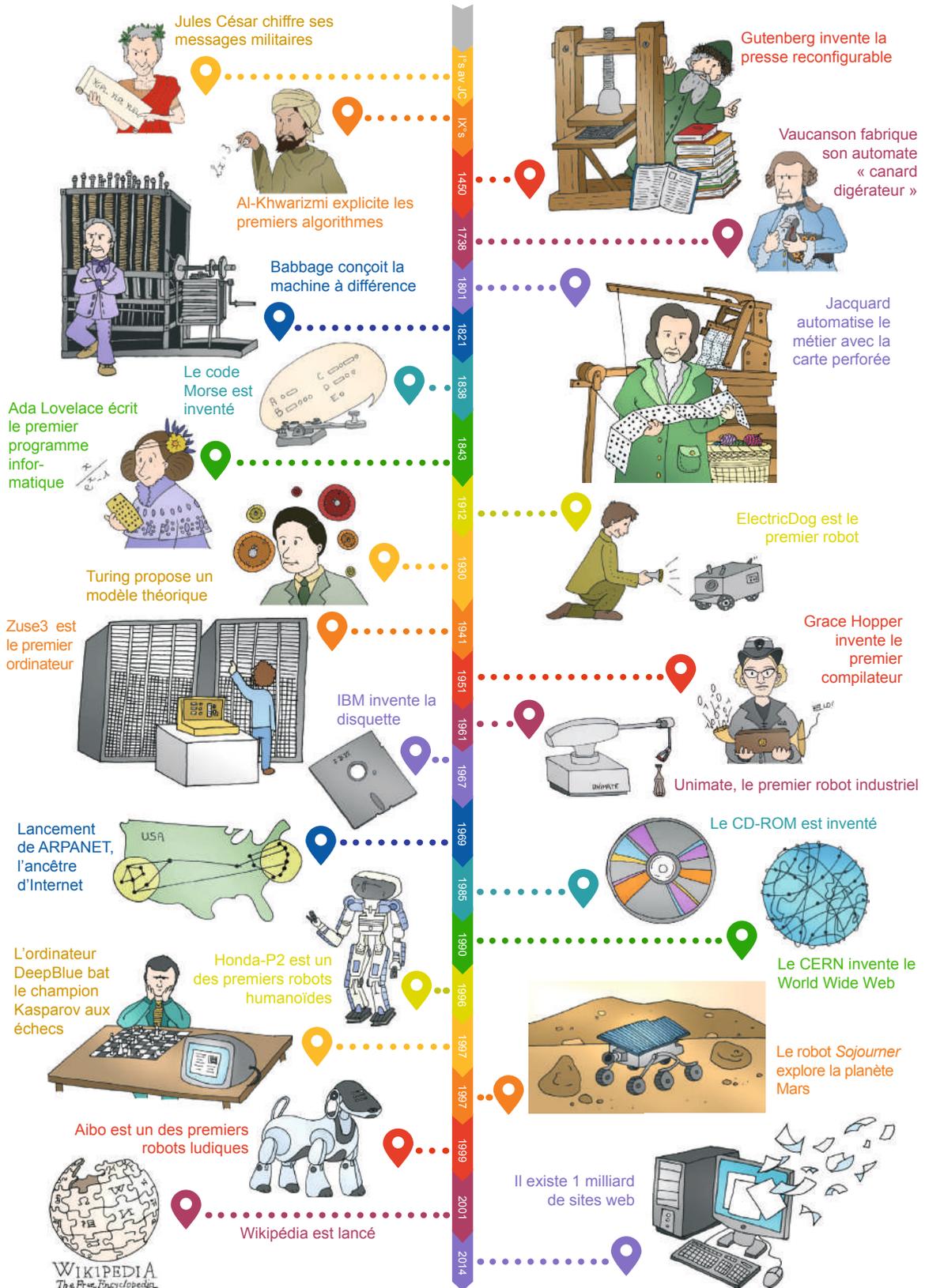
Une brève histoire du réseau Internet

L'idée d'une connexion de nombreux ordinateurs en réseau naît dans les années 1960 aux États-Unis : un chercheur, Leonard Kleinrock, formalise en 1961 un moyen de transmettre des données entre ordinateurs en découpant ces données en paquets, en transmettant les paquets par des voies qui peuvent être différentes et en les rassemblant dans l'ordre voulu au niveau du récepteur. C'est le principe de la méthode utilisée encore aujourd'hui pour transmettre des données – comme un message électronique – sur Internet (voir illustration page 32). Dès l'année suivante, Joseph Licklider promeut l'ordinateur comme un outil de communication et de partage des ressources, et propose la construction d'un réseau d'ordinateurs connectés entre eux. Licklider intègre l'ARPA (Agence des projets de recherche avancée), organisme créé en 1958 pour assurer la supériorité militaire et technologique des États-Unis, à la suite du lancement par l'URSS du premier *Sputnik*. Il devient le directeur du Bureau des techniques de traitement de l'information de l'ARPA, et c'est à ce poste qu'il lance la construction d'un réseau expérimental reposant sur le principe de la transmission des données par paquets. Ce réseau, nommé Arpanet, voit le jour en 1969. Il comporte seulement 4 ordinateurs, qui s'échangent des données à un rythme de 50 000 bits par seconde (6 ko/s).

Les premiers utilisateurs d'Arpanet développent des programmes permettant d'envoyer des messages d'un ordinateur à un autre, via le réseau. Le premier e-mail, envoyé en 1972, utilise déjà le caractère @ pour séparer le nom du destinataire et son adresse. Les principales fonctions des messageries électroniques – tri, réponse, transfert de messages – sont inventées dès cette époque. La première démonstration publique d'Arpanet a lieu en 1972, lors d'une conférence internationale. C'est un grand succès qui contribue à l'expansion du projet, à la naissance d'autres réseaux et fait naître l'idée d'une interaction entre réseaux. Une telle interaction nécessite des modifications dans le principe de fonctionnement d'Arpanet : la solution développée est le protocole TCP/IP, encore utilisé aujourd'hui sur le « réseau de réseaux » qu'est Internet.

Tim Berners-Lee, utilisateur du réseau du CERN (appelé Cernet), imagine en 1989 de faciliter la navigation au sein du réseau par l'usage d'un navigateur et de liens hypertextes. Il propose la création d'une toile permettant d'aller d'un contenu à un autre par des voies multiples. Ce projet, approuvé par le CERN, est nommé le « World Wide Web », ce qui signifie littéralement « toile d'araignée mondiale ». C'est l'un des réseaux d'Internet.

Frise chronologique



La séance « Bilan : qu'est-ce que l'informatique ? » pour le cycle 3 (page 335) propose aux élèves de reconstituer une frise chronologique similaire à partir d'études documentaires portant sur l'histoire de l'informatique

Une brève histoire du réseau Internet (suite)

Les projets foisonnent à partir des années 1990, entourés de spéculation boursière, jusqu'à éclatement de la « bulle » Internet en 2000 : le premier navigateur Web possédant une interface graphique voit le jour en 1993. Il sera rapidement supplanté par d'autres, eux-mêmes supplantés par d'autres ... L'annuaire Internet Yahoo ! est créé en 1994 puis il devient un portail international. La start-up Google est créée en 1998, puis ce sont eBay ou encore Amazon, entreprises géantes de l'Internet.

Algorithmes, langages et programmes

L'ambiguïté du verbe « coder »

Le verbe « coder » est très à la mode (on le trouve par exemple dans la formule « lire, compter, écrire et coder » au sujet de l'introduction de l'informatique à l'école). Ce mot pouvant être porteur de plusieurs ambiguïtés, discutons-en ici.

La première de ces ambiguïtés est certainement la plus facile à clarifier. En effet, il arrive parfois que le mot « coder » soit utilisé pour faire référence au fait de modifier le contenu d'un message pour le rendre illisible à quiconque ne connaîtrait pas le « code secret ». Dans la suite de ce manuel, cette action sera désignée comme du « chiffrement », soit l'action de « chiffrer » le message pour le rendre illisible à quiconque ne connaît pas la « clef de chiffrement » permettant non seulement de « chiffrer » le message, mais aussi de le « déchiffrer » pour le rendre à nouveau lisible. Plus généralement, le mot « coder » fait référence à l'action de donner des instructions à une machine, comme un ordinateur. Cette activité peut également être désignée par le terme de « programmation ». Les instructions seront alors interprétées par cette machine, qui va les exécuter pour produire un résultat.

Enfin, « coder » peut être utilisé pour représenter une information à l'aide de symboles (par exemple, écrire un texte en binaire, à l'aide de 0 et de 1). Selon les cas, on utilisera les termes « encoder » ou « décoder ». Dans ce guide pédagogique, c'est ce sens qui a été retenu pour le verbe « coder », bien que le grand public utilise plutôt « coder » pour « programmer ».

Le mot « coder » véhicule toutefois une seconde ambiguïté, qui ne concerne pas sa définition mais plutôt l'idée que l'on peut se faire de l'informatique à travers l'activité de programmation. Car l'informatique, ce n'est pas que de la programmation. Il s'agit certainement de l'activité la plus visible. Celle qui, dans la culture populaire, représente concrètement le pouvoir qu'ont les programmeurs, et qui illustre la séparation entre ceux qui subissent les machines en tant qu'utilisateurs (même tacites) et ceux qui les font fonctionner. C'est, par exemple, l'image souvent caricaturale du programmeur qui enchaîne des lignes de code très rapidement dans les films. C'est, encore, l'image d'un écran qui fait défiler

des pages entières de code en arrière-plan d'un informaticien dans un reportage ou un documentaire (probablement parce que cela donne un côté impressionnant, ou très spécialisé). Cette image pourrait participer à la diffusion de l'idée que l'informatique est difficile d'accès et réservée à une élite ou à des esprits qui lui sont prédestinés. Heureusement, ce n'est pas le cas. De toute évidence, programmer fait entièrement partie de l'informatique et, pourtant, cette activité n'en est pas le seul engrenage. Programmer permet effectivement de voir se concrétiser ses idées. Mais ces idées peuvent naître et mûrir en dehors de toute forme de programmation, en utilisant le raisonnement comme moteur principal. Ce raisonnement, associé à la nature mécanique des ordinateurs et de leur fonctionnement, constitue la « pensée informatique » et permet de mettre au point les algorithmes qui font aujourd'hui tourner une grande partie de notre monde.

Qu'est-ce qu'un algorithme ?

L'omniprésence des algorithmes n'est plus à démontrer. Quand on fait une requête sur un moteur de recherche, ce sont plusieurs algorithmes qui sont mis en jeu pour trouver les pages les plus pertinentes possible. Quand les marchés financiers voient des opérations se faire par milliers en une seconde, de la part d'un même trader, ce n'est pas un trader humain qui donne les ordres, mais un algorithme de trading à haute fréquence. Quand un gouvernement veut analyser de très grandes masses de données concernant les activités humaines, afin d'essayer d'y détecter des comportements suspects, alors il se tourne vers les algorithmes⁴. Les algorithmes jouent un rôle important dans nos vies, qu'on en fasse le choix ou pas.

Pour illustrer la notion d'algorithme, l'exemple de la recette de cuisine est assez bien adapté. Après tout, il s'agit d'une série d'instructions à appliquer pour obtenir un résultat. Et les analogies entre une recette de cuisine et un algorithme peuvent aller plus loin. Une recette peut être diffusée dans le monde ou bien elle peut être gardée secrète. Elle peut être recopiée à la main ou transmise oralement. Mais tant qu'elle n'est pas incarnée par des actions en cuisine, elle reste un concept. En informatique, ce concept est un algorithme. Il peut rester écrit à la main, sous une forme plus ou moins formelle mêlant des explications et des symboles mathématiques. Tant qu'il reste sous cette forme, sans être appliqué, il est un savoir qui ne produit aucun résultat. La puissance d'un algorithme se manifeste quand il est interprété et concrètement déroulé, que ce soit par un humain ou par une machine.

Pour la vitesse d'exécution, c'est autre chose... Un algorithme simple peut être interprété par un être humain, en suivant les étapes et en notant les résultats intermédiaires, s'il y en a. C'est le cas, par exemple, pour l'algorithme de la division que les élèves peuvent appliquer systématiquement une fois qu'ils l'ont appris. Ils peuvent alors dérouler n'importe quelle division à la main en appliquant rigoureusement les étapes de l'algorithme. Ils notent les valeurs intermédiaires et recommencent les opérations jusqu'à ce que les conditions d'arrêt du calcul soient atteintes. Pour un algorithme plus compliqué cela peut devenir délicat et pour la plupart des algorithmes qui se déroulent en ce moment autour de nous, cela devient inenvisageable. Les machines, elles, peuvent faire dérouler ces algorithmes à des vitesses chaque année plus grandes. Pour que cela soit possible il faut réunir au moins deux conditions :

1. Concevoir et écrire un algorithme qui permet d'atteindre le résultat visé.
2. Traduire cet algorithme dans un programme que la machine peut interpréter.

Ces deux activités sont certainement le cœur de la production d'objets numériques. La programmation

4. La question de la pertinence des algorithmes pour cet objectif n'est pas le sujet ici, mais ce sont bien des algorithmes qui sont envisagés pour traiter ces données colossales qui ne peuvent pas être analysées « à la main ».

peut connaître différents niveaux de qualité, en fonction de l'expérience du programmeur et de sa connaissance du langage utilisé. Un programme peut être écrit de plusieurs manières tout en suivant le même algorithme. Certains programmeurs connaissent tellement bien leurs langages préférés qu'ils peuvent obtenir des gains de performance importants, en écrivant des programmes plusieurs fois plus rapides que les programmes écrits par des programmeurs novices. Cela n'est pourtant pas toujours suffisant pour obtenir des programmes efficaces.

Car les algorithmes, eux aussi, peuvent être écrits de plusieurs manières. Simplement parce qu'ils peuvent être conçus de plusieurs manières. Un algorithme peut être particulièrement bien pensé, mais il peut aussi être basique et utiliser des choix malvenus. Ce que nous allons voir par la suite, c'est qu'un algorithme mal conçu ne peut pas être compensé par une programmation efficace, ou un ordinateur ultra performant. Un algorithme bien conçu, en revanche, même programmé par un novice sur un ordinateur vieillissant, peut offrir des performances radicalement supérieures. Cette conception de la méthode, ce raisonnement permettant de produire l'algorithme le plus efficace possible, est une véritable clef en informatique. La discipline qui lui correspond est l'algorithmique. Explorons ensemble quelques fondements de cette discipline, à la lumière de la petite histoire de l'informatique que nous venons de voir.

Langages

Les premiers ordinateurs étaient programmés avec des cartes perforées. C'est d'ailleurs ce que Babbage avait prévu pour sa machine analytique. Comme nous l'avons déjà vu, en voulant rendre la programmation plus accessible, Grace Hopper a compris les limites des cartes perforées comme interface entre la machine et ceux qui la programment. Pour dépasser ces limites, elle a contribué à la création des compilateurs. Il s'agissait de pouvoir exprimer un programme dans un langage proche du langage humain. Toute la difficulté étant de traduire ce programme dans le langage de la machine, pour qu'elle exécute les actions correspondantes. Cette traduction est connue sous le nom de « compilation ». Le programme qui se charge de la compilation est un compilateur. Il faut, pour compiler un programme, relever un grand nombre de défis, parmi lesquels :

- Reconnaître et séparer les mots et les éléments du langage pour en faire des entités avec lesquelles travailler par la suite. Par exemple la ligne de code `resultat=a+3` doit permettre de reconnaître la variable `resultat`, l'opérateur `=`, la variable `a`, l'opérateur `+` et le nombre `3`.
- Détecter les incohérences dans le langage et s'assurer que le programme est correct d'un point de vue syntaxique ou grammatical (donc vérifier que le programme respecte les règles du langage dans lequel il est écrit).
- Détecter les incohérences sémantiques. Par exemple, si deux variables portent le même nom alors il est impossible de traduire le programme pour la machine (il y a une ambiguïté qui rend impossible la compilation).
- Générer effectivement le code, dans le langage de la machine, à partir des éléments construits dans les étapes précédentes.

L'apport des compilateurs est majeur dans l'histoire de l'informatique. Il est évidemment beaucoup plus facile d'utiliser des mots, des nombres et des opérateurs pour programmer que d'inscrire les états de la machine, un par un, sur des cartes perforées. Cet apport se traduit aujourd'hui par une multitude de langages de programmation, parmi lesquels figure *Scratch*, le langage que nous avons choisi pour initier les élèves à la programmation. Les différences entre les langages peuvent être très importantes. Dans le projet pédagogique proposé pour le cycle 3, une séance (page 215), illustre la diversité des

ce sont donc des machines capables d'exécuter tous les algorithmes du monde : des **machines universelles**. Gageons que pour les pionniers de l'informatique, voir s'incarner les algorithmes dans des machines universelles devait être exaltant.

Comme ces pionniers, l'apprenti programmeur découvre les capacités de son ordinateur de façon progressive. Il commence avec des algorithmes simples bien connus comme des opérations mathématiques (l'addition, la multiplication, etc.). Ensuite, les algorithmes se complexifient, reproduisant des suites mathématiques comme celle de Fibonacci ou de Bernoulli. Justement, c'est en écrivant un programme pour calculer la suite de Bernoulli sur la machine analytique de Babbage qu'Ada Lovelace est devenue la première personne au monde à écrire un programme informatique. Il existe bien d'autres programmes encore. La machine universelle nous permet d'appliquer des algorithmes connus, mais également de rechercher des algorithmes pour résoudre des problèmes qu'on ne sait pas encore résoudre. Cependant, le fait que l'on trouve un algorithme signifie-t-il forcément que cet algorithme nous donnera toujours une solution ? Pour y voir plus clair, explorons ce que l'on appelle la « pensée informatique ».

La **pensée informatique** est un ensemble de connaissances et d'attitudes mises en œuvre pour comprendre le monde qui nous entoure et nous aider à diriger nos actions. Elle couvre une vaste étendue de concepts, et nous allons tenter d'en éclairer ici quelques-uns⁵.

Tout un chacun applique, tous les jours, des algorithmes... parfois sans le réaliser vraiment. Nous avons établi des stratégies, des habitudes, des raisonnements que nous appliquons à nos activités du quotidien. Dans un magasin, un samedi de courses, on peut choisir de parcourir les allées du début à la fin du magasin en vérifiant, pour chaque allée, si elle contient un produit qui est sur sa liste de courses. Mais pourquoi ne pas se faciliter la vie en rédigeant cette liste au préalable, en fonction de sa connaissance du magasin, pour que la liste présente les produits dans le bon ordre ? Pour aller travailler, on peut prévoir son trajet en fonction du jour de la semaine parce qu'il est connu que le mardi et le jeudi il faut absolument éviter la rue Kolmogorov qui est saturée. Enfin, quand on cherche un mot dans un dictionnaire, on l'ouvre à l'endroit le plus probable en fonction de la position de la première lettre de ce mot dans l'alphabet (vers le début pour la lettre « c » ou vers le milieu pour la lettre « p », etc.)... puis, par tâtonnements, on recule ou on avance de quelques pages, en fonction de la lettre qui débute les mots de la page courante. Et lorsque la première lettre du mot est trouvée, on recommence avec la deuxième lettre, et ainsi de suite.

Les ordinateurs peuvent nous aider à traiter ces tâches quotidiennes ! Il s'agit alors de poser le problème de manière formelle, puis d'écrire un algorithme basé sur les quatre ingrédients vus plus haut (instructions, boucles, tests et variables) et permettant de le résoudre. Pour automatiser l'exécution de ces algorithmes, il faut les traduire en programmes et laisser une machine dérouler leurs instructions. Une partie de la pensée informatique consiste à rationaliser de tels procédés pour qu'une machine soit capable de les mettre en œuvre.

Voyons maintenant quelles sortes de défis nous pouvons résoudre grâce à la pensée informatique et aux algorithmes. Il y a les défis que l'on peut résoudre avec des algorithmes qui s'inspirent du raisonnement humain. Ce sont des algorithmes comme, par exemple, celui qui établit la liste des courses dans l'ordre des allées du magasin. Beaucoup d'algorithmes sont conçus de cette manière. La personne qui met au point l'algorithme essaye de formaliser son propre raisonnement. C'est un peu comme expliquer la méthode à quelqu'un d'autre, en voulant s'assurer qu'il n'y aura aucune erreur d'interprétation.

5. L'article « La pensée informatique » https://interstices.info/jcms/c_43267/la-pensee-informatique donne une vue d'ensemble de ce que recouvre cette expression.

Ensuite, il y a les défis que l'on va résoudre avec un algorithme inspiré du raisonnement ou des pratiques humaines, mais en y ajoutant une étude qui permet de rendre l'approche plus rigoureuse et plus générale. Reprenons l'exemple de la recherche d'un mot dans le dictionnaire. Il existe un exercice très similaire qui consiste à chercher son nom, ou son prénom, dans une liste (par exemple, pour savoir si on est reçu à un concours). Essayons de formaliser la manière dont on réalise cette tâche.

Imaginons une liste, non triée, de quelques centaines de prénoms. Il faudra à tout un chacun quelques minutes pour y retrouver son nom (ou être sûr qu'il n'y est pas). Dans un second temps, si l'on fournit la même liste triée par ordre alphabétique, l'exercice est beaucoup plus rapide!

Cette différence d'efficacité vient du fait que, quand la liste est triée, il existe une méthode qui est bien plus efficace que si la liste n'est pas triée. Quand elle n'est pas triée, on peut chercher un peu partout au hasard, ou bien on parcourt systématiquement la liste du début à la fin en vérifiant les prénoms un par un. En revanche, quand la liste est triée, on repère l'endroit où se trouvent les prénoms qui commencent par la première lettre de son prénom. Ensuite la deuxième lettre, etc.

Cette approche fonctionne, mais peut encore être perfectionnée. Car si le nombre d'objets est très grand (comme dans un dictionnaire, voire un répertoire téléphonique national avec des millions d'entrées) alors on peut faire de grands bonds en avant ou en arrière pour gagner du temps. Il existe un algorithme de recherche dans une liste triée très efficace (le plus efficace connu pour répondre à ce défi): la « dichotomie », qui tire son nom du grec pour l'expression « couper en deux ». Pour trouver un item dans une liste triée, on coupe la liste en son milieu; si l'item était au milieu, on l'a trouvé, sinon on sait dans quelle moitié de la liste triée il se trouve; on recommence donc la dichotomie sur la demi-liste restante. Et ainsi de suite. À chaque étape, la liste dans laquelle on cherche l'information est deux fois plus petite.

Une liste de 20 prénoms est fouillée en 4 coups dans le pire des cas: le nombre d'opérations dans le pire des cas est donné par le nombre de fois qu'on peut diviser 20 par 2 avant d'arriver à 1 (on arrondit la première division si le nombre de cases est impair). Sans dichotomie, il aurait fallu (dans le pire des cas) 20 itérations pour lire successivement chaque prénom, du premier au dernier, pour se rendre compte que le sien n'y était pas. La puissance de la dichotomie se voit vite dans le cas de listes bien plus longues: une liste de 300 000 items est fouillée par dichotomie en 19 opérations seulement. Imaginons maintenant que chaque opération de lecture dans le tableau demande une milliseconde à l'ordinateur. Le tableau suivant estime le temps d'application des deux algorithmes dans le pire des cas:

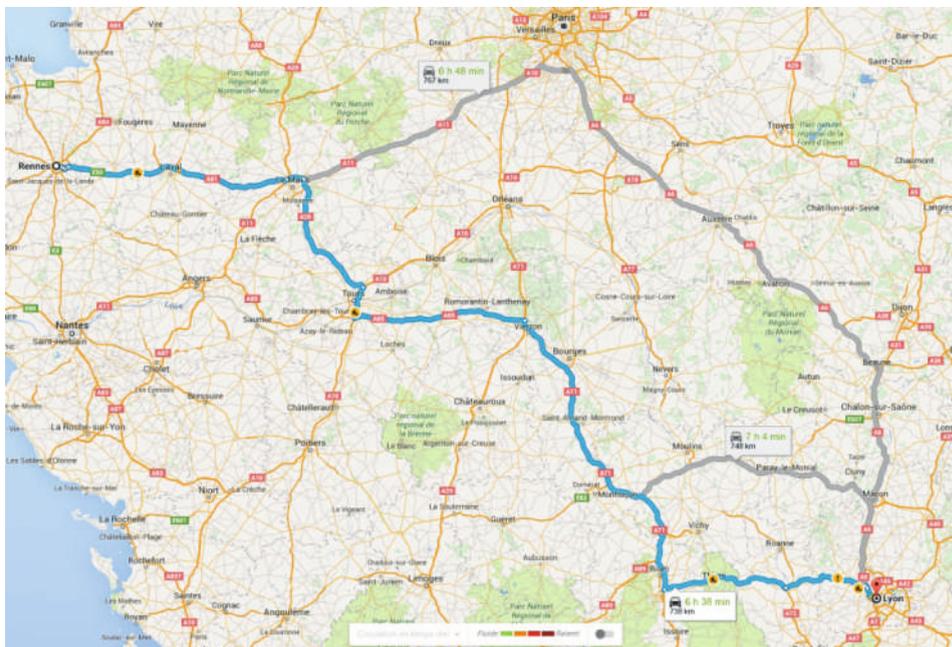
Nombre de prénoms dans la liste	300	3 000	30 000	300 000	3 000 000
Algorithme 1 : parcours de toute la liste, du premier au dernier	0,3 s	3 s	30 s	300 s	3 000 s
Algorithme 2 : dichotomie	0,008 s	0,012 s	0,015 s	0,018 s	0,022 s

Comparaison du temps mis, dans le pire des cas, pour rechercher un prénom dans une liste triée, à l'aide de 2 algorithmes différents. L'algorithme le plus simple (on parcourt la liste du début à la fin) met 10 000 fois plus de temps pour s'exécuter si la taille de la liste est multipliée par 10 000 (on passe de 3 dixièmes de seconde à 1 heure), tandis que la recherche dichotomique n'est rallongée que d'un centième de seconde.

Les approches pour mettre au point des algorithmes sont nombreuses. Nous ne pouvons pas les passer toutes en revue ici. En revanche, il est important de se pencher sur des défis d'un type particulier, pour lesquels on sait écrire un algorithme correct, mais qui ne connaîtront pourtant jamais de solution. Ou plutôt devrions-nous dire qu'on n'obtiendra pas la solution en un temps raisonnable. Il s'agit là

d'un aspect fondamental de l'algorithmique. Considérons par exemple le « problème du voyageur de commerce » que nous décrivons lors d'une séance débranchée pour le cycle 3 (page 273). Il s'agit de trouver, pour un ensemble de villes dans lesquelles doit se rendre un voyageur de commerce, l'ordre dans lequel visiter ces villes de façon à passer le moins de temps possible sur la route. S'il n'y a que 2 villes à visiter, la réponse est simple, avec un seul trajet possible pour les relier. Avec 3 villes, il y a 6 trajets possibles à comparer (dans ce décompte, on distingue des trajets « identiques » mais parcourus dans 2 sens différents, par exemple Lyon-Bordeaux-Marseille et Marseille-Bordeaux-Lyon). Avec 4 villes, il y a 24 trajets possibles. Avec 20 villes, il y a déjà plus de deux milliards de milliards de possibilités. Un ordinateur mettrait plusieurs milliards d'années à tester toutes les possibilités pour y trouver à coup sûr la meilleure ! Ce problème se retrouve dans de nombreuses applications comme la tournée du facteur (ou des éboueurs, des déneigeuses...) : dans quel ordre parcourir les rues dans lesquelles il faut passer ? Il s'agit du même problème que dans les usines où des bras articulés (des robots) doivent visser des boulons sur une plaque à différents endroits : dans quel ordre visser les boulons pour minimiser les déplacements et donc économiser de l'énergie et d'allonger la durée de vie du robot ? Quand le nombre « d'objets » (villes à visiter, boulons à visser...) augmente alors le nombre de possibilités explose. On parle d'« explosion combinatoire » ou de « croissance exponentielle ». Au-delà d'un nombre d'objets relativement petit, la solution à tous ces problèmes est... tout simplement inconnue !

Bien entendu, pour réduire le temps de calcul, on peut envisager de prendre un ordinateur plus performant ou un autre langage plus efficace. Le nouveau programme, écrit par un programmeur expérimenté dans un langage plus performant, sera peut-être dix fois ou même cent fois plus rapide. Mais, quand le temps de calcul est estimé à 10 milliards d'années, on passera peut-être à « seulement » cent millions d'années. Imaginons même, pour la beauté du raisonnement, qu'on trouve une architecture informatique, un langage de programmation, un système qui permette de calculer la solution exacte au trajet entre 20 villes en une minute. Que se passera-t-il si le problème initial change



Ce chemin proposé pour relier Nantes et Lyon n'est pas le meilleur chemin possible : le trouver nécessiterait des milliards d'années de calcul. C'est une heuristique, un compromis entre la qualité du chemin et le temps de calcul alloué à sa recherche. Le meilleur chemin n'est probablement pas beaucoup plus court que celui-ci. © Google Maps

un peu ? Ajoutons une 21^e ville : il y a 21 fois plus de trajets à tester qu'avec 20 villes, ce qui nécessite 21 fois plus de temps, soit 21 minutes pour trouver la solution. Ajoutons une 22^e ville : le super-ordinateur tourne pendant 7 heures. Une 23^e ville ? Le calcul dépasse 7 jours. Cette machine imaginaire, et déjà totalement hors de portée, se trouve ainsi mise à mal avec seulement quelques items supplémentaires. Toutefois, la situation n'est pas désespérée car il existe plusieurs méthodes pour répondre à ces défis, mais elles ne calculent pas la solution exacte. On propose une solution qui paraît convenable, on l'améliore plusieurs fois, et quand on pense que c'est trop difficile de faire mieux alors on en reste là en estimant que ce n'est déjà pas si mal. En informatique une telle approche, donnant un résultat satisfaisant mais non optimal, est appelée une « heuristique ». Cela revient un peu à appliquer l'adage qui veut que « le mieux est l'ennemi du bien ». Savoir reconnaître ces problèmes pour lesquels on ne peut pas calculer de solution en un temps raisonnable est un aspect majeur de l'algorithmique. Si un programme est très long à s'exécuter, ce n'est peut-être pas à cause d'un bug ou d'une mauvaise connaissance du langage. C'est peut-être simplement parce que l'algorithme est correct, tout comme le programme, mais... qu'il met beaucoup de temps à s'exécuter.

Qu'est-ce que l'intelligence artificielle ?

Le terme « intelligence artificielle » (IA) est utilisé pour caractériser des algorithmes et des machines dotées de formes d' « intelligence ». Il désigne aussi le domaine de recherche qui étudie ces algorithmes, et qui a été créé dans les années 1950. Les machines étudiées en IA sont très diverses, tant sur le plan de leurs mécanismes, de leurs fonctionnalités et du but de leurs concepteurs.

Des exemples de tels algorithmes sont ceux qui permettent à des ordinateurs de battre les meilleurs joueurs du monde aux échecs ou d'aider des mathématiciens à trouver automatiquement la preuve de certains théorèmes. Ces algorithmes reposent sur la manipulation astucieuse de symboles, et le but de leurs concepteurs est de construire des machines efficaces capables de nous aider dans des tâches quotidiennes comme la recherche d'information sur Internet ou la résolution de problèmes de logistique. Bien qu'ils soient très performants sur des tâches particulières, ces algorithmes fonctionnent de manière très différente du cerveau humain, et ne sont pas capables de s'adapter de manière autonome à des tâches nouvelles (c'est-à-dire sans l'aide d'un ingénieur qui les reprogrammerait).

D'autres travaux visent au contraire à construire des algorithmes qui modélisent les processus cognitifs et neuronaux humains, afin de contribuer à mieux les comprendre. Certains laboratoires modélisent par exemple les processus d'apprentissage et de développement des savoir-faire sensoriels et moteurs de l'enfant, ou l'acquisition du langage, souvent en les expérimentant sur des robots. Cependant, les mécanismes qui sont mis en œuvre ont encore des capacités cognitives et d'adaptation qui sont très éloignées de celles de jeunes enfants, et des défis fondamentaux de compréhension de la cognition restent à résoudre.

Enfin, on peut mentionner depuis quelques années l'importance croissante du big data (masses énormes de données) dans le développement de nouvelles formes d'intelligence artificielle. Des chercheurs ont, par exemple, fait visualiser des milliers d'images de chats à un programme qui s'est alors construit, seul, le « concept » de chat. La traduction automatique s'inscrit également dans ce champ : l'algorithme de traduction de Google, par exemple, apprend au fur et à mesure qu'il réalise des traductions.

Bien entendu, cela ne signifie pas que l'on peut se passer de programmeurs expérimentés qui connaissent bien le fonctionnement de leurs langages préférés, ni que les progrès technologiques sont sans importance. Beaucoup seront satisfaits de voir les performances de leur logiciel préféré améliorées par une programmation judicieuse ou un processeur plus rapide. Mais les progrès de l'algorithmique sont indispensables à une meilleure connaissance de la nature des problèmes que l'on aborde et de la manière dont on peut les résoudre... ou pas!

Les objets informatiques : ordinateurs, robots, réseaux...

Comme nous l'avons évoqué précédemment, l'informatique est née lorsqu'on a pu concevoir des machines capables d'exécuter des algorithmes. Ces machines « universelles » sont appelées aujourd'hui des ordinateurs.

Qu'est-ce qu'un ordinateur ?

Un ordinateur est une machine capable de manipuler automatiquement des informations et de les mémoriser, et que l'on peut programmer. Les manipulations de l'information sont notamment des suites d'opérations arithmétiques et logiques. Bien que le mot « ordinateur » soit souvent associé à l'image de nos ordinateurs personnels, avec écran et clavier, les ordinateurs au sens large sont aujourd'hui présents dans de très nombreux objets quotidiens sous la forme de processeurs « embarqués » dans ces objets⁶ : téléphones, tablettes, montres électroniques, voitures, avions, feux de circulation, caméras, appareils photo, télévisions, appareils ménagers, jouets interactifs, appareils auditifs, thermostats... et tous les robots dont nous parlerons plus loin.

Des ordinateurs mécaniques aux microprocesseurs

Aujourd'hui, la plupart des ordinateurs sont des machines électroniques; cependant, les premiers ordinateurs ont d'abord été mécaniques. Précurseur, Charles Babbage a conceptualisé le premier ordinateur au début du XIX^e siècle, sous la forme d'une machine composée d'un jeu complexe de rouages mécaniques (voir ci-avant, page 11). On fournissait à cette machine des instructions de calcul sous la forme de cartes perforées, selon un principe similaire à la programmation des machines à tisser. La machine de Babbage était en particulier capable de réaliser des boucles et des branchements conditionnels, c'est-à-dire de décider de la suite d'instructions à réaliser en fonction du résultat d'un calcul.

6. Il y a environ 10 fois plus d'ordinateurs « embarqués » que d'ordinateurs « classiques » (avec écran, clavier...) dans le monde.

Les premiers ordinateurs électromécaniques sont apparus pendant la Seconde Guerre mondiale. Ils étaient utilisés pour les calculs balistiques des missiles. Ces ordinateurs utilisaient l'électricité pour activer et désactiver des relais mécaniques afin de réaliser des calculs. Les ordinateurs entièrement électroniques sont apparus peu après, basés sur l'utilisation de lampes à vide: citons le Colossus en Grande-Bretagne qui servit à décoder les messages secrets échangés par l'armée allemande (voir ci-avant, page 13, à propos d'Alan Turing). Les lampes à vide furent ensuite remplacées par les transistors, dont la petite taille, la faible consommation d'énergie et la robustesse permirent leur accumulation sur des « circuits intégrés ». La miniaturisation constante des circuits intégrés a donné naissance aux microprocesseurs dès la fin des années 1950.



Lampe à vide de 1947



Transistor PNP de 1953



Microprocesseur Intel© 4040 de 1974. Contient 3 000 transistors



Microprocesseur Intel© Core i7 de 2014. Contient 2,6 milliards de transistors

Aujourd'hui, des laboratoires de recherche imaginent des alternatives aux circuits électroniques pour les ordinateurs de demain. En particulier, ils étudient comment utiliser des propriétés de la matière afin d'accélérer certains types de calculs (comme les états quantiques dans le cas des « ordinateurs quantiques »), ou comment utiliser certaines molécules complexes (comme l'ADN dans le cas du « DNA computing »).

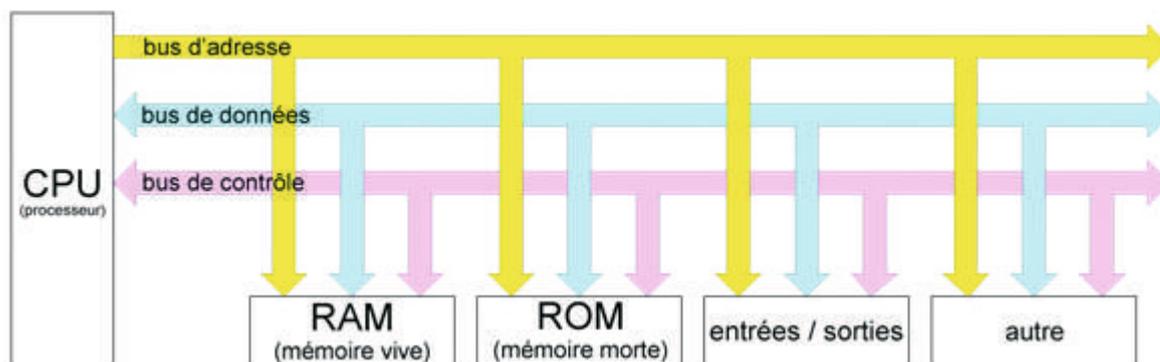
Les composants des ordinateurs

La plupart des ordinateurs actuels reposent sur quatre entités: l'unité de contrôle, l'unité arithmétique et logique (ALU), la mémoire, et les dispositifs d'entrée-sortie. Ces entités, souvent constituées elles-mêmes de plusieurs composants, sont interconnectées via des conducteurs électriques (pistes en cuivre, câbles) que l'on appelle de façon générique des « bus ». Ces bus transmettent des informations suivant un codage précis, le « protocole de communication ».

L'unité de contrôle est le chef d'orchestre qui dirige l'ensemble des autres entités: elle lit et décode les instructions et les données de la mémoire, les transforme en signaux qui vont activer d'autres entités (par exemple envoyer à l'ALU un calcul à effectuer, récupérer le résultat et le ranger à un emplacement particulier dans la mémoire).

On appelle souvent « CPU » (*central processing unit* en anglais) l'ensemble composé de l'unité de contrôle, de l'ALU et de cellules mémoire spéciales, les « registres », utilisées pour mémoriser l'emplacement en mémoire de la prochaine instruction.

La mémoire d'un ordinateur peut être vue comme un ensemble de cases dans lesquelles on peut stocker et écrire des nombres (encodés selon le codage binaire: un circuit élémentaire allumé code la valeur « 1 », et un circuit élémentaire éteint code la valeur « 0 »). Chaque case a une adresse qui permet aux autres composants de la localiser.



Les ordinateurs sont souvent associés à des dispositifs d'entrée-sortie, les « périphériques », qui leur permettent d'échanger des informations avec l'extérieur. Au-delà des claviers, souris, disques durs, écrans, haut-parleurs ou imprimantes, les périphériques de la plupart des ordinateurs embarqués dans notre quotidien sont des capteurs (par exemple de lumière, de mouvement, de chaleur, de signal GPS...), des actionneurs (par exemple des moteurs, des valves, des radiateurs,...), et surtout d'autres ordinateurs, situés à l'intérieur du même objet, ou à l'extérieur. Par exemple, un ordinateur personnel ou un téléphone contiennent plusieurs ordinateurs dédiés au traitement du son ou de l'image et qui communiquent les uns avec les autres, ainsi qu'avec les autres ordinateurs, externes, connectés au réseau Internet.

Les robots

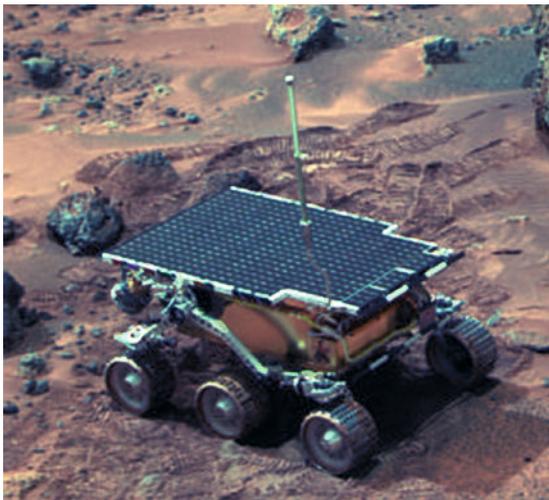
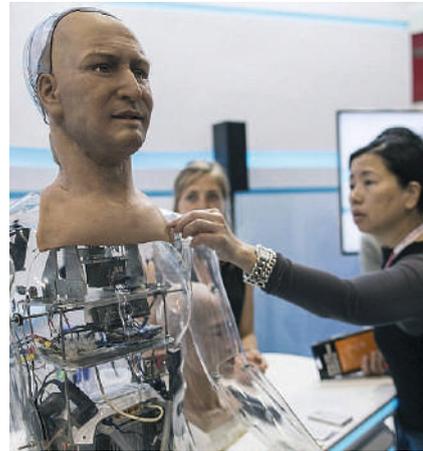
Parmi les objets qui embarquent des ordinateurs, les « robots » ont une importance scientifique, sociétale et économique grandissante. Les robots sont partout : dans les usines et dans les champs, au fond des mers et dans l'espace, dans les jardins et les salons. En outre, ils ont pénétré notre culture et certains d'entre eux participent au renouvellement de la vision que nous avons de nous-mêmes.

D'un point de vue technique, un robot est une machine dotée de capteurs (de contact, de distance, de couleur, de force...) qui lui permettent de percevoir son environnement, de moteurs l'autorisant à bouger et à agir sur cet environnement, et d'un système qui contrôle ce qu'effectue le robot en fonction de ce qu'il perçoit. Une caractéristique fondamentale des robots, qui les distingue des automates, est cette rétroaction entre perception et action. Les automates (tels que ceux de Jacques de Vaucanson ou de Pierre et Henri-Louis Jaquet-Droz au XVIII^e siècle) ne sont pas des robots car leurs mouvements ne dépendent pas de ce qui se passe autour d'eux : ils n'ont pas de capteurs et leurs enchaînements sont entièrement prédéterminés par le programme.

En pratique, cette définition d'un robot recouvre une très vaste diversité de machines : les bras articulés programmables des usines automobiles, les voitures (en mode d'assistance à la conduite) et les avions qui sont aujourd'hui largement automatisés, les aspirateurs qui font le ménage tout seuls, certains objets électroniques des magasins de jouets ou les robots biomimétiques en forme d'animaux (singes, poissons...) que l'on rencontre parfois dans les laboratoires de recherche.

Cette diversité n'est pas seulement une diversité de formes, c'est aussi une diversité d'usages, qui sous-tendent une diversité de logiques de fonctionnement. Ainsi, les robots se spécialisent selon deux axes majeurs : l'autonomie et les capacités d'adaptation et/ou d'apprentissage.

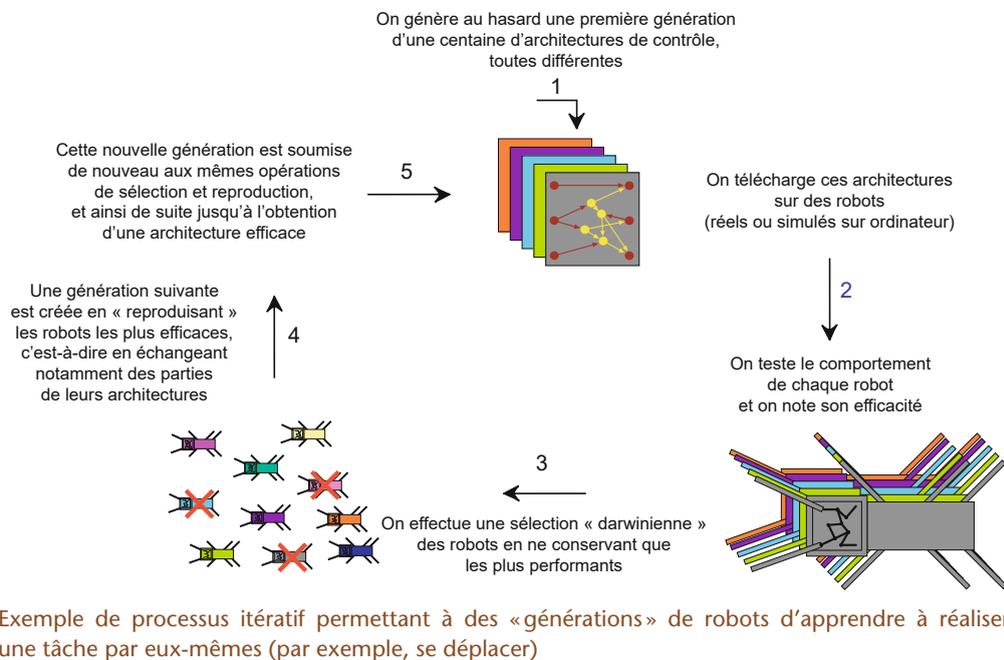
L'autonomie : il existe des robots qui agissent sans qu'un humain les guide, et d'autres dont le comportement est soit influencé par un humain, soit presque totalement contrôlé par un humain. Par



Quelques exemples de robots. De gauche à droite et de haut en bas : bras mécaniques utilisés dans l'industrie automobile; robot humanoïde dédié à la reconnaissance et à la reproduction des émotions; robot explorateur sur Mars et robot militaire destiné à transporter des charges lourdes sur des terrains accidentés

exemple, dans une usine, les robots qui travaillent à la chaîne et répètent toujours le même geste le font souvent de manière autonome. En revanche, les robots utilisés dans les centrales nucléaires (pour opérer dans les zones à forte radiation) sont typiquement téléguidés par un humain qui leur dit où aller et quoi faire après chaque action.

L'adaptation et l'apprentissage : le comportement de certains robots est figé au départ et une fois pour toutes par le programmeur, alors que d'autres robots sont capables d'acquérir de nouveaux comportements et de nouvelles connaissances par leurs expériences : leur comportement évolue en fonction de l'histoire de leurs interactions avec l'environnement. Ainsi, certains robots sont capables d'apprendre à reconnaître des objets dans des images ou encore d'apprendre à marcher en expérimentant et en évaluant eux-mêmes différentes stratégies. Ces mécanismes d'adaptation sont permis par des « algorithmes d'apprentissage », qui reposent sur la détection automatique de régularités dans des flux de données captées par le robot, et sur des méthodes d'« optimisation » permettant de raffiner de manière progressive et itérative les paramètres des solutions à un problème.



Grâce à ces algorithmes d'apprentissage, certains robots sont capables d'inventer des solutions et des comportements non prédits par leur concepteur, et même de sélectionner pour eux-mêmes des objectifs qui ne sont pas préprogrammés. Il est par exemple possible de programmer un robot en lui donnant des instructions lui demandant de rechercher des situations nouvelles afin d'augmenter ses connaissances sur le monde qui l'entoure. Ainsi, certains algorithmes permettent de doter ces machines de formes d'apprentissage et de créativité. Néanmoins, les capacités et les performances de ces algorithmes sont aujourd'hui, et encore pour très longtemps probablement, très faibles en comparaison des capacités d'adaptation et de raisonnement de nombreux animaux, et en particulier de l'humain.

Comme nous l'avons vu, ces différentes logiques de fonctionnement répondent à différents besoins : il existe en effet une grande variété de raisons, et donc de fonctions, pour lesquelles les robots sont construits et utilisés. On peut considérer trois familles de fonctions : *travailler et explorer*, *assister l'humain*, *modéliser les mécanismes cognitifs et comportementaux du vivant*.

Travailler et explorer. La plupart des robots en service dans le monde sont industriels (on en compte aujourd'hui environ 9 millions). Très tôt, les entreprises se sont intéressées à ces machines pour deux raisons :

- D'abord, les robots peuvent être utilisés pour remplacer les travailleurs humains dans les tâches répétitives, pénibles et nécessitant de faibles compétences, comme les montages, la peinture ou les soudures de pièces ;
- En outre, ces machines sont capables de réaliser des travaux à la chaîne bien plus rapidement et efficacement que les humains.

Le premier robot industriel, Unimate, est apparu en 1961 : installé dans une usine automobile de General Motors, ce bras articulé manipulait de lourdes pièces de fonderie. Dans les années 1970, l'usage des robots dans l'industrie a pris son envol. Aujourd'hui, les robots ont pénétré toutes les branches de l'industrie et ne sont plus restreints au seul secteur de l'automobile. Par exemple, dans l'agriculture et l'agro-alimentaire, les robots vont dans les champs cueillir fruits et légumes, certains coupent, pressent et mettent en bouteille ; d'autres trient et mettent en cartons ; d'autres encore groupent en palettes. Dans certains aéroports, des flottes de robots s'occupent de transporter les bagages et de les charger dans les soutes.

Les robots ne sont pas seulement utiles dans l'industrie pour les tâches simples et répétitives, ils sont aussi utilisés pour travailler dans des environnements dangereux pour l'homme. L'industrie nucléaire est un exemple typique. Qu'ils soient autonomes ou partiellement téléguidés, les robots des centrales nucléaires peuvent déambuler dans les enceintes confinées et radioactives, ils peuvent manipuler les substances dangereuses et s'occuper de la maintenance des autres machines. Un autre exemple est l'industrie pétrolière : les robots sous-marins sont par exemple utilisés pour contrôler l'état de la coque des navires afin de prévenir les accidents et d'identifier les « navires poubelles ».

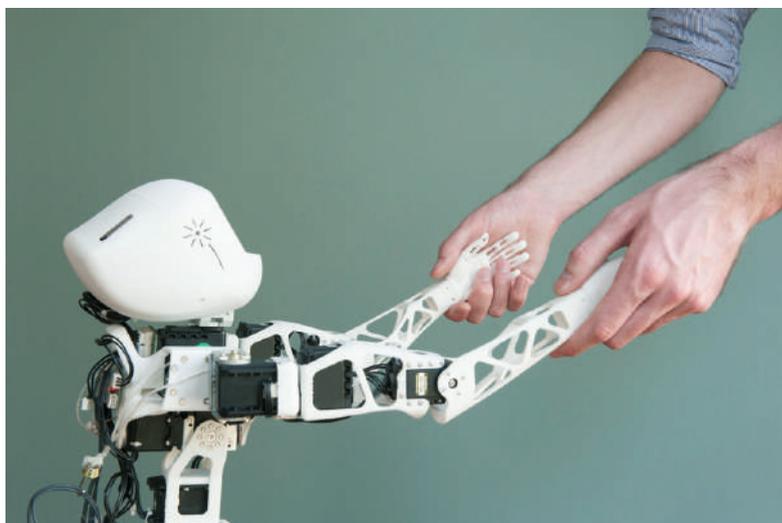
Les robots sont enfin cruciaux pour l'exploration des endroits où l'homme ne peut pas aller, au premier rang desquels l'espace et les corps du Système solaire. C'est en 1966 que le premier robot mobile arrive sur la Lune, embarqué dans la sonde *Surveyor*. Suivront le soviétique *Lunokhod*, puis toute la série des américains *Mariner*. En 1997, un robot atterrit sur la planète Mars : *Sojourner*, qui est propulsé par l'énergie qu'il capte grâce à ses panneaux solaires ; il envoie à la Terre des milliers de clichés et provoque un engouement du grand public. *Sojourner* navigue en partie de manière autonome car, étant donné la distance avec la Terre, il est très difficile de le téléguider en temps réel. En 2004, une nouvelle mission robotisée concentre l'attention du monde entier : *Spirit* et *Opportunity*, équipés de spectromètres et d'un bras qui leur permet de creuser la surface, apportent la preuve que de l'eau a coulé sur Mars.

Assister l'humain dans son quotidien. Si le xx^e siècle a vu l'avènement des robots travailleurs et explorateurs, à l'aube du xxi^e siècle, une autre grande famille de robots prend son essor : les robots d'assistance à la personne. Dans nos maisons, les robots d'assistance aux travaux ménagers se démocratisent. On retrouve aussi des assistants robotisés dans les magasins ou sur nos lieux de travail. L'activité médicale est particulièrement transformée par la robotique : si les robots d'assistance chirurgicale sont quotidiennement utilisés depuis une quinzaine d'années, on voit émerger aujourd'hui de nouveaux usages. Ainsi, des robots accompagnent les personnes qui ont des difficultés physiques ou cognitives, par exemple pour les aider à se lever et à s'asseoir, pour les stimuler cognitivement quand elles ont des problèmes de mémoire, ou pour jouer le rôle de facilitateur des contacts avec la famille ou l'entourage médical. Demain, des chercheurs introduiront d'autres robots, beaucoup plus petits, dans les blocs opératoires : ce sont des capsules endoscopiques miniaturisées capables d'explorer conduits intestinaux, artères ou veines pour aider le chirurgien dans son diagnostic. Apparaissent aussi depuis quelques années des prothèses robotisées de mains ou de bras entières destinées aux amputés.



Prothèse robotisée ©BeBionic

Modéliser le vivant et la cognition. Les robots sont devenus essentiels comme outils pour penser et modéliser les systèmes complexes dans le domaine du vivant. En particulier, les robots sont aujourd'hui utilisés dans des laboratoires de recherche pour comprendre comment les êtres vivants s'adaptent à leur environnement naturel, en termes de comportement et de déploiement de capacités cognitives. En effet, ce comportement est le résultat de la dynamique des interactions entre un cerveau, un corps physique et un environnement, dynamique qui se redéfinit en permanence car le cerveau est modifié à chaque nouvelle interaction. Les robots embarquent à la fois un « cerveau » (des programmes qui permettent de traiter l'information acquise par le robot, selon des règles précises), des « organes sensoriels » (les senseurs) et des « systèmes moteurs » (les actuateurs). De cette manière, les robots peuvent, comme les êtres vivants, modifier l'environnement physique dans lequel ils agissent et être modifiés par lui. Leur



Le robot humanoïde Poppy est *open source*. Il est développé par l'Inria à la fois pour des recherches sur la cognition et pour des projets éducatifs. Plus de détails sur : <https://www.poppy-project.org>

« cerveau » aussi en est modifié, car le robot acquiert de nouvelles connaissances qu'il peut réutiliser dans ses interactions ultérieures.

Les chercheurs peuvent étudier la complexité de l'interaction cerveau-corps-environnement, grâce à des expérimentations possibles sur des robots et impossibles sur des êtres vivants – comme par exemple « éteindre » une partie de leur cerveau artificiel pour voir comment le comportement est modifié ou altérer des parties du corps. Certains laboratoires étudient ainsi le contrôle moteur, la perception visuelle, le repérage dans l'espace, ou encore les mécanismes de l'apprentissage et de l'évolution de la parole et du langage chez l'humain. Dans le cadre de ces projets, les interactions avec les neurosciences, la biologie, la psychologie ou encore l'éthologie jouent un rôle central.

Les réseaux

De nos jours, la plupart des ordinateurs sont interconnectés entre eux, et rares sont ceux coupés du réseau. On les connecte pour leur permettre d'échanger des informations, qu'il s'agisse de transmettre un courriel ou de télécharger des données stockées sur les ordinateurs de Wikipédia pour les lire.

On interconnecte aussi les ordinateurs pour leur permettre de collaborer à une tâche trop volumineuse pour être réalisée sur une seule machine. De nos jours, des « fermes » d'ordinateurs regroupent des milliers voire des millions de machines – qui, prises individuellement, ont des performances assez classiques – dans le même hangar climatisé, pour faire toujours plus de calculs à la seconde.

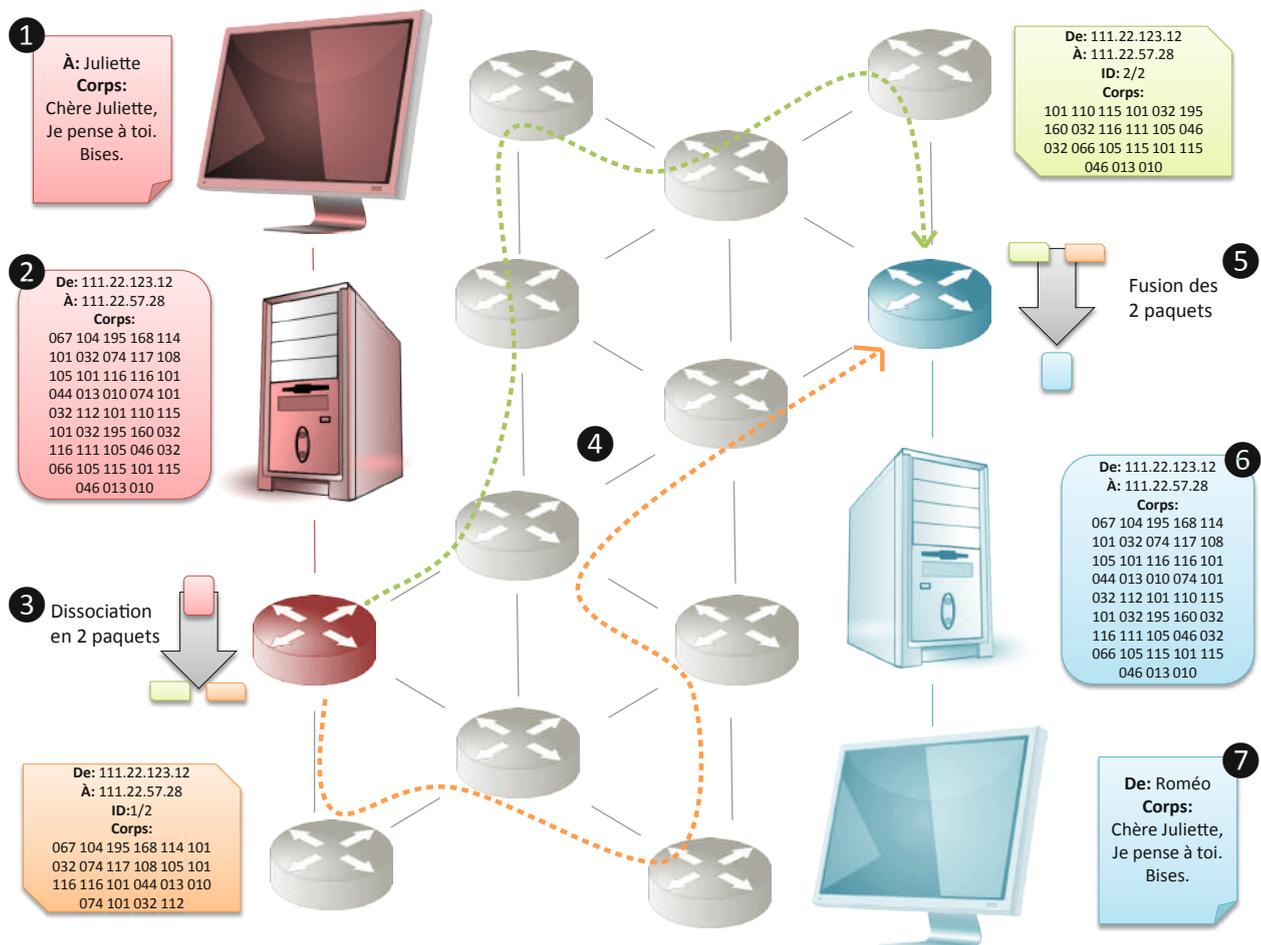
Les premiers réseaux à grande distance sont apparus à la fin des années 1960. On ne comptait qu'une quinzaine de machines interconnectées en 1971, un millier en 1984, un million en 1992, et elles sont plusieurs milliards aujourd'hui. Le premier e-mail date de 1972, le premier site Web de 1991, le premier tweet de 2006.

Dans les anciens réseaux téléphoniques, il fallait connecter physiquement deux téléphones de bout en bout pour qu'ils puissent communiquer. Des opérateurs dans des centraux téléphoniques réalisaient ces branchements, qui demandaient d'avoir autant de lignes que de discussions en cours.

La plupart des réseaux informatiques actuels suivent une autre approche plus efficace, nommée « commutation de paquets ». Les informations sont coupées en paquets, qui sont envoyés les uns derrière les autres. À chaque nœud du réseau (qui est maillé), des « routeurs » reçoivent ces paquets et les font suivre au suivant, jusqu'au destinataire.

Plusieurs technologies ont été proposées pour échanger ces paquets entre ordinateurs, au risque de faire ressembler les réseaux informatiques à la tour de Babel. Heureusement, Internet permet à tous les ordinateurs qu'il connecte de communiquer malgré leurs différences. Pour parvenir à ce tour de force, les scientifiques et ingénieurs ont appliqué une approche classique en informatique : le problème est découpé en sous-problèmes, qui sont résolus séparément, puis les réponses partielles sont combinées. Ainsi, le problème consistant à échanger des informations entre des ordinateurs est découpé en couches.

C'est ce découpage qui rend possible un réseau à l'échelle mondiale : les technologies sont interchangeables au sein de chaque couche, et l'ensemble collabore grâce à des interfaces standardisées entre les couches. Internet Protocol (IP) est le nom d'un protocole permettant d'interconnecter n'importe quels réseaux locaux entre eux. C'est pour cela qu'on dit qu'Internet est le « réseau des réseaux ».



Roméo utilise son logiciel de messagerie pour envoyer un email à Juliette (1). Ce message est d'abord encodé en ASCII (2) puis en binaire. Un routeur le découpe en plusieurs paquets (3: dans la réalité, il y a bien plus de paquets!). Chaque paquet transite sur le réseau via plusieurs autres routeurs (4). Les paquets sont finalement fusionnés (5), puis le message est transmis à l'ordinateur de Juliette (6), avant d'être décodé et affiché à l'écran (7)

Internet ou Web ?

Dans le langage courant, et très souvent dans les médias, il n'existe pas de différence entre ces deux termes. Erreur !

Internet est le « réseau des réseaux », et le Web n'est que l'un des réseaux qu'Internet relie (voir page 15 pour l'historique).

Le World Wide Web, inventé en 1990 au CERN, à Genève, est un réseau de pages liées entre elles exclusivement par liens hypertextes (on reconnaît facilement leur nature car leur URL –*Uniform Resource Locator* –, leur adresse, commence souvent par « www. ». Mais il existe d'autres réseaux sur Internet, comme par exemple :

- les services de courrier électronique (les e-mails), qui utilisent les protocoles SMTP (*Simple Mail Transfer Protocol*), POP3 (*Post Office Protocol version 3*), IMAP (*Internet Message Access Protocol*)... ;
- les services de messagerie instantanée (les chats), sur protocole IRC (*Internet Relay Chat*), XMPP (*Extensible Messaging and Presence Protocol*)... ;
- les serveurs de forum Usenet (newsgroups), sur protocole NNTP (*Network News Transfer Protocol*)...

Représentation de l'information⁷

L'idée de représenter un objet, concret ou abstrait, par un autre est très ancienne, puisque toutes les langues reposent sur une telle représentation des objets par des sons.

Comme les objets sont trop nombreux pour être chacun représenté par un son différent, ils sont représentés, non par un son unique, mais par une suite de sons, choisis dans un ensemble de petite taille. Par exemple, le français classique utilise trente-six sons, ou phonèmes, pour former tous les mots. Dans une langue, un message est donc une suite finie d'unités choisies dans un ensemble fini. Cette idée n'est cependant pas complètement aboutie dans les langues orales où le volume, le débit, l'intonation du message contribuent à sa signification, tout autant que la suite des unités dont il est constitué.

Combien d'information contient un texte ?

L'écriture, et surtout son utilisation pour exprimer les langues, élimine ces éléments accessoires, mais elle constitue, au moins dans un premier temps, une régression, puisque les premières écritures des langues associaient un symbole différent à chaque objet. Ce n'est qu'avec l'alphabet que se retrouve l'idée d'associer à chaque objet non une unité, mais une suite d'unités prises dans un ensemble de petite taille : l'alphabet. L'alphabet du français, par exemple, contient 110 caractères environ : 42 minuscules

7. Le projet « 1, 2, 3... codez ! » propose plusieurs activités aux cycles 2 et 3 pour travailler la représentation de l'information (texte et image), notamment les séances situées aux pages 121, 129 (cycle 2) et 221, 227, 233, 302, 310 et 315 (cycle 3).

(les 26 lettres fondamentales, 13 voyelles accentuées, le c cédille et les 2 ligatures æ et œ), autant de majuscules, 10 chiffres, et une vingtaine de signes de ponctuation.

Le nombre de suites de n caractères que l'on peut former en les choisissant dans un alphabet de k éléments est k^n : en effet, il y a k possibilités pour le premier caractère, k possibilités pour le deuxième, ... et k possibilité pour le n -ième, donc le nombre total de possibilités est $k * k * \dots * k$ (n fois) c'est-à-dire k^n . Plus ce nombre est grand, plus chacune de ces suites contient d'information. Par exemple, un texte de 4000 caractères choisis parmi les 110 caractères de l'alphabet français contient plus d'information qu'un texte de deux caractères choisis parmi trois.

L'unité d'information est appelée le « bit ».

Pour des raisons historiques on utilise souvent une autre unité d'information, l'octet qui vaut huit bits, ainsi que ses multiples: le kilooctet (ko) qui vaut mille octets, soit huit mille bits, le mégaoctet (Mo) qui vaut un million d'octets, le gigaoctet (Go) qui vaut un milliard d'octets, le téraoctet (To) qui vaut mille milliards d'octets, et le pétaoctet (Po) qui vaut un million de milliards d'octets.

La quantité d'information contenue dans un caractère de l'alphabet est de l'ordre de 1 octet. Celle contenue dans une page de 2000 caractères est d'environ 2 kilooctets, celle contenue dans un livre de six cents pages est de l'ordre de 1 mégaoctet, celle contenue dans une petite bibliothèque de mille volumes est de l'ordre de 1 gigaoctet. Celle contenue dans la collection de livres imprimés de la Bibliothèque nationale de France, qui compte 14 millions de volumes (en laissant de côté les images, les films...) est de l'ordre du téraoctet (précisément 14 téraoctets). Un téraoctet est aussi la capacité d'un disque vendu quelques dizaines d'euros. Un pétaoctet correspond à cent fois la taille de la BNF. Chaque année, le CERN (Centre européen pour la recherche nucléaire) produit de l'ordre de 15 pétaoctets de données.

Le binaire : un alphabet à 2 lettres

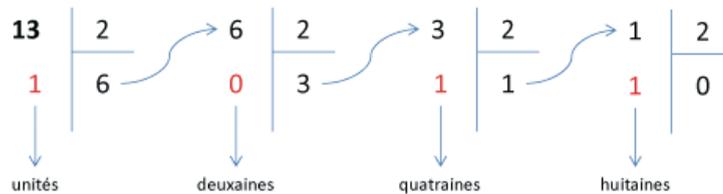
Comment représenter un nombre ?

Une branche entière de l'informatique s'intéresse à la représentation des données sous la forme d'une suite de symboles pris dans un alphabet fini, souvent dans un alphabet qui ne contient que deux caractères : **0** et **1**. Ainsi, pour représenter les nombres entiers dans un alphabet qui ne contient que ces deux caractères, on peut associer une suite de quatre caractères « 0 ou 1 » à chaque chiffre, par exemple :

- 0 : 0000
- 1 : 0001
- 2 : 0010
- 3 : 0011
- 4 : 0100
- 5 : 0101
- 6 : 0110
- 7 : 0111
- 8 : 1000
- 9 : 1001

Puis, pour représenter un nombre, la première idée, intuitive, qu'on peut avoir consiste à juxtaposer la représentation de chacun de ces chiffres. Le nombre 13 par exemple se représente alors en juxtaposant la représentation 0001 du chiffre 1 avec la représentation 0011 du chiffre 3 : 00010011. Cette notation, peu compacte, n'est pas utilisée en pratique pour les nombres.

Une autre représentation, plus concise, consiste à représenter le nombre en base deux. Le nombre 13 par exemple, qui se décompose en une unité, zéro paquet de deux unités, un paquet de deux paquets de deux unités et un paquet de deux paquets de deux paquets de deux unités ($1 + 4 + 8$) s'écrit de droite à gauche: 1-0-1-1, soit 1101 de gauche à droite.



Pour écrire le nombre 13 en base 2, on doit le décomposer comme une somme de puissances de 2. Pour cela, il faut effectuer une série de divisions par 2, jusqu'à obtenir un quotient égal à 0. Les restes des divisions successives (qui sont des 0 et des 1) correspondent aux chiffres des unités, deuxaines, etc. Pour écrire 13 en binaire, il faut remettre les chiffres dans l'ordre (avec les unités à droite).

Une succession de 64 caractères «0 ou 1» permet alors de représenter les nombres de 0 à $2^{64} - 1 = 18\,446\,744\,073\,709\,551\,615$. La question de la représentation des nombres ne s'arrête cependant pas là, car les nombres négatifs et les nombres décimaux ont leur propre mode de représentation.

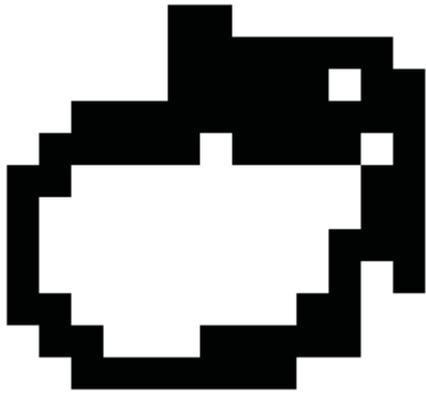
Comment représenter un texte ?

La représentation des caractères textuels – et des textes qui sont des successions de caractères textuels – est aussi une question très riche. Une première idée pour représenter les caractères textuels était d'associer un groupe de 7 caractères «0 ou 1» à chaque caractère d'un alphabet qui en contenait 95 (26 minuscules, 26 majuscules, 10 chiffres et 33 autres symboles, incluant les signes de ponctuation). Le caractère textuel «a» était représenté par 1100001, «b» par le 1100010, etc. Et un mot était simplement représenté par la juxtaposition de la représentation de ses caractères – ses lettres. Le mot «Le» par exemple était représenté par la juxtaposition de la représentation 1001100 du caractère «L» et la représentation 1100101 du caractère «e»: 10011001100101.

Mais cette représentation, appelée American Standard Code for Information Interchange (ASCII), était adaptée à l'anglais et non aux langues qui, comme le français, utilisent des signes diacritiques: accents, cédilles,... La représentation ASCII a donc été étendue plusieurs fois pour arriver à la représentation Unicode, qui permet de représenter les quelque 110 000 caractères utilisés par les différentes écritures du monde, y compris les écritures non alphabétiques, comme le chinois.

Comment représenter une image ?

La représentation des images est plus difficile et elle repose sur un découpage de l'image en pixels, puis par un codage de la teinte de chaque pixel. Dans une image en noir en blanc, chaque pixel est représenté par un seul «0 ou 1»: 1 s'il est noir, 0 s'il est blanc.



0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0
0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	1
0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	1	1	1	1	1	0	1	1	1	1	0	1	0
0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0
0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0
0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1
0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0	1	1	1	1	1	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

À gauche, une image en noir et blanc de 16 x 16 pixels. À droite, le codage binaire de cette image. Pour faciliter la lecture, on a disposé les bits sur une grille 16 x 16.

Dans une image en niveaux de gris chaque pixel est classiquement représenté par un nombre compris entre 0 et 255, lui-même représenté en base deux par un groupe de huit «0 ou 1», 0 = 00000000 s'il est noir, 1 = 00000001 s'il est gris très foncé... 254 = 11111110 s'il est gris très clair, 255 = 11111111 s'il est blanc. Bien entendu, le choix d'utiliser 255 niveaux de gris est arbitraire et d'autres formats existent. Dans une image en couleurs, chaque pixel est classiquement représenté par trois nombres, chacun compris entre 0 et 255 qui indiquent la quantité de rouge, de vert et de bleu dans ce pixel. Par exemple un pixel rose bonbon est obtenu en additionnant 249 unités de rouge, 66 unités de vert et 158 unités de bleu; le nombre 249 se représente en base deux par 11111001, le nombre 66 par 01000010 et le nombre 158 par 10011110, si bien qu'un pixel rose bonbon se représente par la succession de 24 «0 ou 1»: 111110010100001010011110. Une image se représente naturellement en juxtaposant les représentations de chacun de ses pixels. Si bien qu'une image formée de deux pixels rose bonbon se représente par: 111110010100001010011110111110010100001010011110.

Organiser l'information

Ces principes permettent de représenter de petits objets: un nombre, un texte, une image, un son, une vidéo... On peut, par exemple, les utiliser pour stocker sur le disque d'un ordinateur un seul texte, une seule image, un seul son ou une seule vidéo. Mais, bien souvent, on veut stocker sur un tel disque plusieurs objets de cette nature. Cela mène à d'autres principes qui permettent non de représenter, mais d'organiser l'information.

Un disque est un espace sur lequel il est possible de stocker un grand nombre de «0 ou 1», typiquement huit mille milliards pour un disque d'un téraoctet. Pour pouvoir y stocker plusieurs textes, images,... on commence par découper cet espace en plusieurs zones appelées «fichiers» qui permettent chacune de stocker un texte, une image, un son, une vidéo... Mais dès que ces fichiers sont nombreux il est nécessaire de les regrouper en dossiers, eux-mêmes regroupés en dossiers..., c'est-à-dire de les structurer en un arbre. Toutefois, retrouver un fichier dans un tel arbre devient vite difficile, si bien que l'on a inventé d'autres principes pour organiser l'information à grande échelle. Le premier est l'utilisation de liens hypertextes.

Un tel lien permet d'indiquer, dans un premier fichier, le nom et la position d'un second fichier situé sur le même disque, ou sur un autre.

Ainsi, on peut indiquer dans le premier fichier que le second se trouve à l'adresse « disque3 : archives/2016/lettre.txt », c'est-à-dire qu'il porte le nom « lettre.txt » qu'il est situé dans le dossier « 2016 », lui-même situé dans le dossier « archives », sur un disque qui porte le nom « disque3 ». De cette manière, en utilisant un navigateur pour lire le premier fichier, il suffira de cliquer sur ce lien pour accéder au second. Cette manière d'organiser l'information à l'aide de liens hypertextes, à travers un réseau de disques et d'ordinateurs, est l'idée du Web (voir plus haut, page 33).

Cette organisation de l'information en fichiers, puis en dossiers, et son enrichissement par des liens hypertextes est elle-même insuffisante quand on veut stocker beaucoup d'information, et il existe deux autres manières d'organiser de grandes quantités d'information. La première, qui est adaptée à l'information structurée, est la constitution d'une base de données. Par exemple, un carnet d'adresses est une petite base de données, formée de quintuplets (5 données) constitués chacun d'un nom, d'un prénom, d'une adresse, d'un numéro de téléphone et d'une adresse électronique. Il est possible d'interroger cette base de données en formulant une requête, dans laquelle on demande la liste des quintuplets qui contiennent par exemple un nom, ou alors un prénom, voire un numéro de téléphone. Pour les données moins structurées, on peut laisser en vrac un ensemble de fichiers sur un ou plusieurs disques et laisser un moteur de recherche les indexer, puis accéder à ces fichiers en formulant une requête au moteur de recherche, qui indiquera alors où se trouvent tous les fichiers qui contiennent, par exemple, un certain mot. Cette manière d'organiser l'information demande moins de travail que leur structuration dans une base de données, mais donne des résultats de moins grande qualité. Chercher le numéro de téléphone de Mme Marchand en formulant une requête dans un moteur de recherche risque d'indiquer que le mot « Marchand » est présent dans de nombreux textes, dont beaucoup ne parleront que du métier de marchand.

Transformer et manipuler l'information⁸

Représenter et organiser ainsi les informations permettent de les transmettre d'un lieu à un autre (c'est-à-dire de les faire voyager dans l'espace), de les archiver pour les restituer plus tard (c'est-à-dire de les faire voyager dans le temps), mais surtout de les transformer. De nombreux algorithmes permettent de transformer des données particulières ; par exemple l'algorithme de la multiplication ne s'applique qu'à des nombres. Mais trois types d'algorithmes s'appliquent à des données quelconques : les algorithmes de compression, de correction et de chiffrement.

Compression

Nous avons vu qu'un pixel rose bonbon se représente par 111110010100001010011110 et qu'une image monochrome d'un million de pixels de cette même couleur s'obtient en juxtaposant 1 million de fois ce bloc de 24 « 0 ou 1 ». Au lieu de produire un texte très répétitif de vingt-quatre millions de signes, on peut utiliser une représentation plus concise qui consiste à indiquer que le bloc 111110010100001010011110 doit être répété un million de fois. C'est ce que l'on appelle une compression.

8. Le projet « 1, 2, 3... codez ! » propose deux activités pour aborder la correction et le chiffrement au cycle 3 (pages 322 et 330).

Correction

Les algorithmes de correction d'erreurs permettent de repérer, et parfois de corriger, une erreur dans une information, par exemple une erreur introduite lors de la transmission de cette information. Une méthode simple, mais coûteuse en espace mémoire, est de répéter trois fois chaque information unitaire. Par exemple 101 est transformé en 111000111. Si une erreur se produit, par exemple dans le troisième triplet: 111000110, il est facile non seulement de repérer cette erreur, mais aussi de la corriger: comme il y a une majorité de 1 dans 110, la lettre initiale était sans doute un 1. Cet algorithme de correction est mis en échec uniquement quand deux erreurs se produisent dans le même triplet.

Chiffrement

Enfin les algorithmes de chiffrement permettent de rendre une information inintelligible, sauf aux personnes qui possèdent la clef de chiffrement. Une méthode simple, mais peu efficace, était déjà utilisée par Jules César pour communiquer avec ses armées. Elle consiste à décaler chaque lettre d'un nombre fixe de places dans l'alphabet. Ainsi en décalant les lettres de trois places dans la phrase «VENI VIDI VICI» on obtient «YGPL YLPL YLEL». Pour déchiffrer ce message il faut connaître la clé, c'est-à-dire le nombre de places dont il est nécessaire de décaler chaque lettre pour retrouver le message initial. Comme il n'y a que vingt-trois lettres en latin, il n'y a que vingt-trois clefs possibles, si bien que ce message n'est pas très difficile à casser (ou à «cryptanalyser»), c'est-à-dire à déchiffrer sans connaître au préalable la clé.

L'analyse fréquentielle permet de trouver rapidement la clé: dans une langue donnée, toutes les lettres ont une fréquence d'apparition connue (en français, le E – dans toutes ses déclinaisons – est de loin la lettre la plus fréquente; les lettres A, I, N, O, R, S, T, L sont ensuite assez fréquentes... et à l'autre bout du spectre les lettres Z, W et K sont très peu utilisées). Si, la plupart du temps, la lettre la plus fréquente donne directement la clef d'un chiffrement de César, il peut être nécessaire d'utiliser un histogramme plus complet, qui fournit les fréquences de toutes les lettres. Il suffit, pour connaître la clé, de décaler l'histogramme du texte chiffré jusqu'à le faire correspondre au mieux avec l'histogramme de la langue considérée. Dans les 3 exemples ci-dessous, la clef de chiffrement est toujours la même (+3, la même que celle utilisée par Jules César): même si l'on constate de petits écarts dans les histogrammes, on retrouve un même motif (même pour le texte issu du roman *La Disparition* de G. Perceval qui a été écrit sans le moindre E).

Message chiffré (clef+3)	Analyse fréquentielle	Message déchiffré
MH MHWWH DYHF JUDFH PRQ IHXWUH MH IDLV OHQWHPHQW O DEDQGRQ GX JUDQG PDQWHDX TXL PH FDOIHXWUH HW MH WLUH PRQ HVSDGRQ		Je jette avec grâce mon feutre, Je fais lentement l'abandon Du grand manteau qui me calfeutre, Et je tire mon espadon (Cyrano de Bergerac, E. Rostand)

<p>O HWUH KXPDLQ FURLUD WRXMRXUV TXH SOXV OH URERW SDUDLW KXPDLQ SOXV LO HWW DYDQFH FRPSOAH HW LQWHOOLJHQW</p>		<p>L'être humain croira toujours que plus le robot paraît humain, plus il est avancé, complexe et intelligent. (Les robots de l'aube, I. Asimov)</p>
<p>LO DEDQGRQQD VRQ URPDQ VXU VRQ OLW LO DOOD D VRQ ODYDER LO PRXLOOD XQ JDQW TX LO SDVVD VXU VRQ IURQW VXU VRQ FRX</p>		<p>Il abandonna son roman sur son lit. Il alla à son lavabo; il mouilla un gant qu'il passa sur son front, sur son cou. (La Disparition, G. Perec)</p>

Le chiffrement de César fait partie de la famille des chiffrements « par substitution monoalphabétique » : Il existe d'autres méthodes bien plus complexes : par exemple, pendant la Seconde Guerre mondiale, le chiffrement allemand Enigma (cassé par A. Turing et son équipe, voir page 13) se basait lui aussi sur des substitutions, mais la clef de chiffrement changeait régulièrement au fil du texte ! Historiquement, des méthodes de chiffrement qui n'étaient pas des substitutions monoalphabétiques ont aussi été utilisées, notamment en temps de guerre (par exemple, le chiffrement ADFGVX durant la Seconde Guerre mondiale).

Des méthodes de chiffrement bien meilleures existent et sont utilisées tous les jours pour protéger nos numéros de cartes de crédit, et autres secrets que nous souhaitons ne partager qu'avec un petit nombre de personnes...

Informatique et enjeux sociétaux

Enjeux actuels, scientifiques, économiques et sociétaux

La construction d'ordinateurs depuis les années 1940 a d'abord transformé le monde de la recherche : en permettant de simuler des phénomènes physiques complexes – telles les évolutions du climat – et en permettant de traiter de grandes masses de données, comme la séquence du génome humain, les ordinateurs ont transformé la méthode scientifique. Les ordinateurs et les réseaux ont parallèlement

transformé la circulation des informations dans les entreprises. Enfin, l'arrivée des ordinateurs et des réseaux dans tous les foyers a transformé la manière dont nous communiquons dans notre vie privée, avec nos « amis » et nos familles.

On qualifie parfois de « monde numérique » le monde dans lequel nous vivons, qui se caractérise par l'omniprésence des objets informatiques. Ces objets sont partout, de manière visible : ordinateurs, tablettes, téléphones... et invisible. Dans l'informatique « embarquée » (dans les avions, les trains, les voitures, les régulateurs cardiaques, les appareils électroménagers...) des objets informatiques remplacent en effet de plus en plus de systèmes mécaniques ou électriques.

L'accès aux données, grâce à Internet, se fait désormais de façon extrêmement simple et rapide. Une donnée, dès lors qu'elle est numérisée, occupe une place négligeable, pour un coût minime, et peut être dupliquée instantanément. De ce fait, les systèmes informatiques sont dits « hypermnésiques » : ils sont capables de collecter des quantités énormes d'information et de les conserver sur de longues durées. Un exercice simple permet de prendre la mesure de cette hypermnésie : la Bibliothèque nationale de France contient 14 millions de volumes. Supposons que chaque volume comporte cinq cents pages et chaque page deux mille signes (oublions, pour cet exercice de pensée, les images) et qu'un signe s'exprime sur un octet. La « taille » de cette bibliothèque est donc de $14\,000\,000 \times 500 \times 2000 = 14\,000\,000\,000\,000$ octets soit quatorze mille milliards d'octets, 14 téraoctets. Or un disque de 10 téraoctets coûte quelques centaines d'euros. On voit donc que, grâce aux techniques numériques, la capacité de stockage d'une petite entreprise peut rapidement dépasser la taille de la Bibliothèque nationale de France.

D'un point de vue économique, ce monde favorise bien évidemment l'industrie informatique, qui produit tous ces matériels et logiciels à haute valeur ajoutée, ainsi que la recherche en science informatique, dont les avancées théoriques vont de pair avec les avancées technologiques. Le monde numérique est destructeur d'emploi : acheminer le courrier d'un bout à l'autre de la Terre demandait des armées de facteurs, qui ne sont plus nécessaires aujourd'hui que le courrier est électronique. Mais il est aussi source d'emploi, car il faut sans cesse développer de nouvelles applications et de nouveaux logiciels pour suivre l'évolution rapide des besoins. Enfin, troisième pilier de l'importance économique de l'informatique, auquel on ne pense pas forcément de prime abord : la formation professionnelle. Tous ceux qui manipulent, de près ou de loin, les ordinateurs et les logiciels (scientifiques, techniciens, commerciaux, juristes, enseignants, architectes, artistes...) ont besoin de formation, initiale ou continue. La formation professionnelle et le conseil en informatique représentent une part de plus en plus importante de l'activité liée au numérique.

Malgré ces opportunités, force est de constater que la France et l'Europe accusent un retard vis-à-vis de l'Amérique du Nord et de l'Asie. À l'exception de quelques niches (services informatiques, recherche, robotique...), nos pays ont trop longuement relégué l'informatique au rang d'outil, enseignant comment utiliser tel logiciel (logiciel qui devient souvent obsolète en quelques années) plutôt que d'enseigner la pensée informatique, étape indispensable pour favoriser l'émergence d'innovation. On assiste depuis quelques années à une prise de conscience de ce retard et à un souhait de le combler, notamment par l'éducation.

Informatique et éthique

Le monde numérique offre une liberté sans pareille : s'exprimer, accéder à l'information, se former, communiquer, échanger... Mais là où il y a tant de libertés, les risques et les dangers sont grands. Il convient donc d'accompagner les jeunes en leur montrant les limites de ce nouveau monde.

Accéder à l'information de manière immédiate et gratuite aurait semblé une utopie aux bâtisseurs de la bibliothèque d'Alexandrie. Pourtant, comme nous l'avons vu plus haut, un disque dur de faible coût peut contenir presque autant de données que la BNF. Mais si les informations sont accessibles en quantité, sont-elles pour autant de qualité? Quelle fiabilité donner aux informations que l'on trouve sur le Web? Comment garder un certain esprit critique, vérifier les sources (elles aussi accessibles par Internet), croiser les témoignages...? En classe, on peut par exemple faire réaliser une page Web par les élèves, afin qu'ils prennent conscience du fait qu'ils peuvent eux-mêmes écrire n'importe quoi sur le Web, et donc que, s'ils le peuvent, vraisemblablement d'autres le peuvent aussi. Les informations que l'on trouve sur le Web y ont été déposées par des gens comme eux.

Parmi les informations numérisées qui sont stockées ici et là, nous trouvons également nos données personnelles. Nos données de sécurité sociale sont stockées sur un serveur en France, mais nous n'y pouvons rien. Nos données bancaires sont stockées, probablement sur un serveur étranger, et nous n'y pouvons rien non plus. En revanche, nos photos de vacances sont postées sur Facebook, car nous les y avons mises nous-mêmes. Dans ces trois cas, la prudence s'impose : quel est le degré de sécurisation de nos données? Qui peut avoir accès à cette collection astronomique d'informations personnelles? Souvenons-nous que les ordinateurs sont hypermnésiques et que toute activité peut être enregistrée et traitée très rapidement. En 2008, l'entreprise Google a lancé le projet « Google Flu » : en accumulant, région par région, toutes les recherches des mots-clefs « grippe », « fièvre »... il est possible de décrire en temps réel la progression de l'épidémie, et même de la prévoir un peu à l'avance. De telles informations sont évidemment précieuses pour les autorités de santé, mais aussi pour les entreprises pharmaceutiques. De telles informations doivent-elles être rendues publiques, ou peuvent-elles être gardées privées et vendues aux compagnies les plus offrantes? Ce qui est vrai pour la grippe l'est pour de nombreux autres sujets, et comme Google enregistre non seulement les requêtes mais aussi l'adresse IP de l'ordinateur émetteur il n'est plus trop difficile de transformer cette donnée d'apparence anonyme en véritable espionnage. Et que dire de toutes ces applications qui usent et abusent de la géolocalisation, permettant de nous localiser au mètre près?

Tous ces outils, que nous utilisons pour nous simplifier la vie, recèlent des dangers insoupçonnés. Chaque pays doit donc légiférer en connaissance de cause. S'assurer que nos libertés ne sont pas compromises, mais garantir la sécurité de nos données voire un droit à l'oubli. Cela passe par la reconnaissance légale des délits numériques, la recherche en cryptologie et virologie informatique, la formation à la sécurité... Ce qui est déjà difficile à l'échelle nationale, faute de « bon sens informatique » ou de connaissances approfondies des risques, devient encore plus difficile sur la scène internationale. Les réseaux sont en effet supranationaux, les lois diffèrent d'un pays à l'autre et la vente de certains produits comme les drogues est autorisée dans certains pays et non dans d'autres. De même, la publication de certains textes – diffamatoires, blasphématoires... – est autorisée dans certains pays et non dans d'autres. Quelles lois régissent la publication d'un texte : celles du pays où est situé l'ordinateur qui héberge le texte ou celle du pays où est situé l'ordinateur qui permet d'accéder à ce texte?

L'éducation des jeunes mais aussi des adultes (citoyens comme législateurs) est nécessaire pour favoriser une utilisation éclairée et respectueuse de ces nouveaux outils.

Auteurs

David WILGENBUS, Mathieu HIRTZIG, Claire CALMET

Cette ressource a été produite avec le soutien des Editions Le Pommier

Date de publication

2016

Licence

Ce document a été publié par la Fondation *La main à la pâte* sous la licence Creative Commons suivante : Attribution + Pas d'Utilisation Commerciale + Partage dans les mêmes conditions.



Le titulaire des droits autorise l'exploitation de l'œuvre originale à des fins non commerciales, ainsi que la création d'œuvres dérivées, à condition qu'elles soient distribuées sous une licence identique à celle qui régit l'œuvre originale.

Fondation *La main à la pâte*

43 rue de Rennes
75 006 Paris
01 85 08 71 79
contact@fondation-lamap.org

Site : www.fondation-lamap.org

