GÉRARD BERRY

L'Hyperpuissance de l'informatique

Algorithmes, données, machines, réseaux

POURQUOI L'INFORMATIQUE MET LE MONDE À L'ENVERS

Odile Jacob

Odile Jacob

L'Hyperpuissance de l'informatique

Algorithmes, données, machines, réseaux

Ce n'est que récemment que l'on a commencé à mesurer à quel point l'informatique est en passe de transformer notre société, et même de la bouleverser de fond en comble.

C'est à décrire et à analyser les fondements de l'hyperpuissance de l'informatique que Gérard Berry se consacre dans ce livre qui fera date.

Il montre en effet de façon non technique comment la science et la technologie informatiques mettent l'information au cœur de l'action, qu'elle soit produite par les hommes ou par les machines.

Algorithmes, données, machines et réseaux conduisent surtout à un nouveau schéma mental bien différent de celui des siècles précédents, qui confère un pouvoir étonnant à ceux qui le comprennent et l'organisent.

Pour donner concrètement à comprendre le mode de pensée inhérent à l'informatique, Gérard Berry passe en revue cinq domaines de transformations massives : les télécommunications, Internet, la photographie et la cartographie, l'informatisation de la médecine, et celle en cours de toutes les sciences.

Il analyse ensuite en détail deux dangers de l'informatique, les bugs et les trous de sécurité, qui peuvent parfois transformer des systèmes informatisés en dangers publics, et montre comment la science moderne permet de mieux contrôler ces dangers.

Enfin, l'auteur donne sa vision de l'évolution de l'informatique, bien loin des fantasmes trop souvent partagés.

Pour la première fois, un livre qui explique tout de l'informatique, de son monde, ses fondements, ses applications, et la révolution qu'elle représente.

Gérard Berry est professeur au Collège de France où il dirige la chaire Algorithmes, machines et langages. Il est médaille d'or du CNRS.

GÉRARD BERRY
L'Hyperpuissance de l'informatique

Ce livre se termine sur la page 512. Ce n'était pas voulu, mais, pour un informaticien, c'est une aubaine : nos nombres préférés sont les puissances de 2, et $512 = 2^9$ n'est pas la moindre puisqu'elle s'écrit 1 000 000 000 en binaire.

Pour mes 64 ans (2^6), probablement la dernière puissance de 2 pour moi, j'avais essayé d'impressionner mes petits-enfants en leur disant que j'avais un million d'années (en binaire), mais sans grand succès...

Sommaire

CHAPITRE 3

Les piliers de l'informatique

DEUXIÈME PARTIE

Études de cas

CHAPITRE 4

Les télécommunications : du fil à l'air

CHAPITRE 5

Internet, des hommes aux objets

TROISIÈME PARTIE

Rendre l'informatique plus sûre

CHAPITRE 9

Bugs et trous de sécurité : deux dangers de l'informatique

CHAPITRE 10

Comment rendre l'informatique plus sûre

CHAPITRE 11

Où va nous mener l'informatisation du monde ? Une vision personnelle

CHAPITRE 12

Annexes

Avant-propos

Le contenu de ce livre est largement issu des cours que j'ai donnés au Collège de France, d'abord dans le cadre de la chaire annuelle d'Innovation technologique Liliane Bettencourt en 2007-2008 (Berry, 2008), ensuite dans celui de la chaire annuelle Informatique et sciences numériques créée en partenariat avec l'Inria en 2009-2010 (Berry, 2009), et enfin sur la chaire de plein exercice « Algorithmes, machines et langages » que j'occupe depuis 2012-2013 (Berry, 2013). Il a aussi profité des séminaires associés aux cours qu'y ont donnés les collègues que j'y ai invités, des cours d'autres professeurs, et d'un grand nombre de conférences que j'ai données ou auxquelles j'ai assisté dans des milieux des plus variés, allant des grandes instances publiques, scientifiques ou industrielles de tous types jusqu'à des associations locales de curieux en passant par des classes de lycée.

Beaucoup des sujets que je développerai ici ont aussi eu pour moi comme source le Collège de France, à travers les cours des autres professeurs que j'y ai suivis. C'est en particulier le cas pour les instances successives de la chaire annuelle Informatique et sciences numériques, dont les titulaires successifs m'ont beaucoup appris. Je me suis aussi inspiré de ce qu'ont enseigné d'autres professeurs comme Françoise Combes en astronomie, José Sahel pour la vision et bien d'autres. J'ai aussi bien sûr bénéficié de bien d'autres liens avec de nombreux grands scientifiques qui exercent en France ou ailleurs.

C'est une chance extraordinaire que de travailler dans une telle institution, une chance maintenant partageable en vidéo si ce n'est en chair et en os : tous les cours et séminaires que je mentionnerai sont visibles et téléchargeables sur le site Web du Collège de France, avec leurs supports. J'y ferai de nombreuses références en fournissant des liens « http » directement cliquables sur la version électronique de ce livre.

J'ajoute que beaucoup des idées développées ici et des façons d'en parler sont nées en bonne partie lors d'un enseignement donné à l'école Montessori Les Pouces verts à Mouans-Sartoux, d'abord chez les 9-12 ans puis chez les 6-9 ans. Expliquer le monde aux enfants de cet âge est un exercice difficile. Curieux, ouverts et exigeants, ces enfants et leurs éducateurs ont largement modifié mes façons de comprendre, de faire et de communiquer, pour le mieux j'espère.

J'ai bien sûr aussi bénéficié de dialogues constants avec mes collègues informaticiens ou non, que je remercie vivement. Sans pouvoir les citer tous, je remercie particulièrement Serge Abiteboul, Martin Abadi, Nicholas Ayache, François Baccelli, Francis Besse, Jean-Daniel Boissonnat, François Bourdoncle, Françoise Combes, Gilles Dowek, Patrick Flandrin, Cédric Fournet, Georges Gonthier, Nicolas Halbwachs, Gérard Huet, Michael Kishinevsky, Xavier Leroy, Jean-Jacques Lévy, Jean-Christophe Madre, Philippe Manoury, Laurent Massoulié, Maurice Nivat, Marc Pouzet, Manuel Serrano, Laurent Thénié et Jean Vuillemin. Je voudrais aussi offrir un merci spécial à Frédéric Guichard, qui m'a initié à la photographie algorithmique, et Clément Narteau, qui a fait de même pour la physique des dunes. Enfin, je remercie tout particulièrement mes amis de la Société informatique de France, de l'Inria et de La Main à la pâte qui prennent à bras-le-corps les questions de la diffusion et de l'enseignement de l'informatique.

NOTE. Dans ce livre, j'essaie de couvrir des sujets que je connais un peu mais dont je ne suis pas spécialiste. Il est

donc possible que j'aie laissé des erreurs, des contresens, ou des manques vraiment flagrants. Si vous en trouvez, signalez-les-moi s'il vous plaît à l'adresse suivante :

gerard.berry@college-de-france.fr.

CHAPITRE 1

Introduction

La construction d'un nouveau schéma mental

« Rien ne sert de penser, il faut réfléchir avant. »
Pierre DAC.

En janvier 2008, lors de ma première leçon inaugurale au Collège de France intitulée « Pourquoi et comment le monde devient numérique » (Berry, 2008), je disais que l'informatique serait désormais partout et que son expansion dans tous les secteurs de la société et de la science ne faisait que commencer. C'était précisément l'époque où la société commençait à percevoir confusément ses impacts, surtout à travers la constatation des expansions vertigineuses d'Internet et des téléphones portables. Beaucoup d'acteurs constataient déjà que les communications entre hommes et la structure du travail allaient être profondément modifiées, mais peu s'intéressaient encore à d'autres aspects plus cachés comme l'inéluctable informatisation massive des objets, dont je soulignais la très grande importance à terme, ou les problèmes qu'allaient poser la collecte des données personnelles et la sécurité informatique. On en parlait, mais en se demandant si c'était du lard ou du cochon, comme on dit à la campagne. Je mentionnais aussi une situation étonnante pour moi, qui reste encore largement actuelle : la plupart des acteurs politiques, industriels, médicaux, administratifs et juridiques que je rencontrais semblaient en permanence surpris par des évolutions tout à fait prévues, organisées et annoncées par la recherche et l'industrie informatiques. N'est-il pas surprenant

d'être surpris en permanence par du prévisible explicitement organisé ? Mais, heureusement, la curiosité commençait à apparaître.

Dix ans après, je continue à constater la force de l'informatisation du monde, qui va constamment en s'amplifiant. Mais je constate hélas la persistance de la surprise permanente jointe à certaines peurs, certaines raisonnables et d'autres moins. C'est dû je pense à une compréhension insuffisante par le public général et les décideurs en particulier de la nature même de l'informatique. Or cette compréhension est nécessaire pour bien apprécier les raisons et la dynamique de l'irruption de l'informatique dans la vie de tous les jours, et en particulier dans beaucoup d'activités et de métiers autrefois non touchés. Certes, les choses ont progressé, au moins en surface : les politiques s'y intéressent et essaient enfin d'agir, les médias en parlent bien plus qu'auparavant, et la science et la technique informatiques font partie depuis la rentrée 2016 du programme de l'enseignement général français. Mais, si l'on trouve beaucoup d'articles et d'ouvrages consacrés aux impacts et dangers du « numérique », adjectif maintenant promu en nom commun d'une façon que j'analyserai plus tard, la plupart se contentent d'analyser les effets sans vraiment parler de leurs causes, c'est-à-dire de l'informatique en tant qu'activité propre. Pour progresser dans la compréhension du nouveau monde qui se construit à grande allure, il me semble au contraire indispensable d'analyser bien mieux ces causes : quels sont les sujets d'étude de l'informatique, ses façons de penser et ses façons de faire ? D'où vient son étonnante puissance, et pourquoi gagne-t-elle le monde aussi vite et aussi profondément ? Comment interagit-elle avec les autres activités humaines et avec les systèmes physiques ou mécaniques traditionnels ?

Répondre à ces questions est précisément mon but dans ce livre. J'essaierai de le faire non pas de façon théorique, mais en analysant des exemples pris dans des domaines variés, et en n'entrant dans des détails un peu techniques qu'en de rares occasions. Je ne considérerai mon but comme atteint que si le lecteur parvient à se construire un nouveau

schéma mental mieux adapté à l'avenir, au lieu d'essayer de faire entrer l'informatique dans les schémas mentaux hérités des siècles précédents – XX^e^ siècle compris. Pour simplifier la discussion, et même si sa science et sa technologie se sont développées bien avant, je considérerai que l'impact généralisé de l'informatique n'a vraiment commencé qu'autour de l'an 2000 ; ce n'est pas faux quand on regarde son évolution et le nombre d'activités concernées.

Le schéma mental de l'informatique

Mon leitmotiv sera qu'on ne peut pas comprendre les bouleversements provoqués par l'informatique en restant dans les schémas mentaux traditionnels issus des sciences et techniques directement liées au monde physique. En effet, elle diffère profondément de toutes les activités scientifiques et techniques précédentes du fait même de ses objets d'étude et de ses méthodes d'action : l'informatique *calcule sur l'information* à l'aide d'*algorithmes*, de *programmes* et de *machines*, essentiellement des ordinateurs de toutes sortes. L'information est codée dans des *données* numériques, l'algorithme est le mécanisme conceptuel de calcul systématique, le programme constitue l'écriture précise de l'algorithme dans des langages appropriés, et la machine est l'objet matériel capable de faire les calculs nécessaires pour transformer les programmes en actions. Un point absolument essentiel est que la nature physique des objets matériels qui servent à stocker les données et faire les calculs est indifférente : la même information peut être stockée sur n'importe quel support, et n'importe quel ordinateur peut faire les mêmes calculs que n'importe quel autre. Aucun de ces objets conceptuels et pratiques ne ressemble donc à ceux des siècles précédents, tous pour la plupart reliés bien plus intimement au monde physique. Mais l'informatique n'est pas qu'une histoire

d'hommes et de pensée. Elle est très souvent en contact direct avec le monde physique, par exemple pour le mesurer et le contrôler, et tout autant avec le monde du vivant à travers les nouvelles avancées qu'elle permet en médecine et en biologie. La dématérialisation n'empêche pas le contact avec la matière ; au contraire, elle peut l'enrichir.

Structure du livre

Dans la première partie, aux chapitres 2 et 3, j'expliquerai la nature de ce nouveau schéma mental et sa mise en œuvre pratique. Dans la deuxième partie, des chapitres 4 à 8, j'étudierai de façon approfondie cinq exemples pris dans des domaines variés de la vie courante, autrefois très indépendants les uns des autres : télécommunications, Internet, photographie et cartographie numériques, médecine, sciences naturelles et mathématiques. Cette énumération n'a aucunement l'intention d'être exhaustive, mais elle touche aux limites de mon champ personnel de compétences et suffira je pense à illustrer mon propos. J'insisterai sur le fait que l'informatique est *la même partout*. Pour elle, il n'y a pas de différence fondamentale entre un système de gestion bancaire, un appareil photo numérique, une expérience de physique, un dispositif médical ou un réseau social. Même si leurs détails sont bien sûr différents, tous procèdent du même type de pensée et d'action et demandent désormais une formation générale commune.

Comme je n'ai pas du tout l'intention de présenter une vision bêtement idyllique de l'informatique, je consacrerai la troisième partie à deux de ses dangers intrinsèques, qui sont majeurs en pratique mais restent souvent méconnus ou sous-estimés du public. Le chapitre 9 sera consacré aux *bugs* qui cassent nos systèmes et aux *trous de sécurité* qui permettent aux intrus d'y pénétrer. Le chapitre 10 présentera les méthodes modernes pour les contrôler ou les éviter,

qui forment un pan majeur de la grande science informatique à laquelle j'ai consacré ma carrière. La dernière partie, constituée du seul chapitre 11, présentera quelques projections personnelles (et modestes) sur les développements et impacts futurs. Enfin, des annexes préciseront dans le chapitre 12 quelques points scientifiques et techniques pour le lecteur curieux.

Sauf en ce qui concerne l'enseignement, je serai peu disert sur des questions que traitent la plupart des articles ou livres actuels destinés au grand public, c'est-à-dire liées aux impacts sociaux de l'informatique à travers Internet, aux réseaux sociaux, à la propriété des données, à l'éthique des algorithmes, etc. Ces ouvrages discutent certes de questions sociétales et éthiques intéressantes, mais le plus souvent sans vraiment entrer dans les causes. Une exception notable est le beau livre *Le Temps des algorithmes* écrit par mes collègues et amis informaticiens Serge Abiteboul et Gilles Dowek (Abiteboul & Dowek, 2017), dont je recommande chaudement la lecture. Je n'aurais pas grand-chose à ajouter à son contenu. Et je recommande aussi la lecture du blog « Binaire[1] » du journal *Le Monde*, coordonné par Serge Abiteboul avec la Société informatique de France (SIF).

Informatique ou numérique ?

Avant d'entrer dans la chair des chapitres, je souhaite continuer une discussion un peu générale dans cette introduction.

On a vu se produire récemment deux glissements de vocabulaires dans les médias, les discours politiques et l'enseignement : « informatique » est devenu « numérique » et « programmation » est devenu « codage », les deux anciens mots ayant pratiquement disparu. C'est surtout sensible depuis qu'on parle de plus en plus du sujet, en reconnaissant

1. http://binaire.blog.lemonde.fr/.

maintenant qu'il faut « y aller » : il faut enseigner « le numérique » (maintenant devenu un substantif), et les enfants doivent apprendre « le code ». D'où vient ce glissement ? Mon idée personnelle, peut-être un peu brutale, est la suivante : changer les mots a permis à ceux qui le souhaitent de se construire des formes de compétence, d'adhésion ou de rejet sans avoir à entrer dans le cœur du sujet, donc en gardant l'intention de ne se renseigner vraiment ni sur l'informatique, ni sur la programmation. Un bon exemple est donné par les très nombreux hommes politiques et commentateurs de la politique qui associent systématiquement numérique avec Internet, les réseaux sociaux et maintenant l'intelligence artificielle. Ce sont certainement des sujets très visibles et très importants, surtout depuis que les réseaux sociaux jouent un rôle essentiel dans les discussions politiques et la transmission des rumeurs et fausses nouvelles de tous types. Mais l'informatique est beaucoup plus que cela, comme nous allons le voir tout au long du livre. Puisque mon but est précisément de montrer à l'inverse qu'on ne peut pas comprendre le monde numérique dans sa totalité sans comprendre suffisamment ce qu'est son cœur informatique, je ferai très attention aux mots pour que l'arbre ne cache pas la forêt.

Le mot « numérique » a pour moi un sens précis que je souhaite conserver. Je l'utilisais comme adjectif dans le titre « Pourquoi et comment le monde devient numérique » de ma leçon inaugurale de 2008 au Collège de France, où j'expliquais comment la numérisation du monde et l'utilisation des algorithmes, programmes et ordinateurs conduisaient à une modification profonde du monde. Ce mot me permettait aussi d'inclure d'autres communautés scientifiques et techniques. Les mathématiciens appliqués qui montrent comment résoudre des équations complexes à l'aide de grands calculs numériques sont appelés depuis toujours des « numériciens ». La science qui invente les algorithmes de contrôle des avions, satellites, voitures ou robots est appelée l'automatique ; si ces algorithmes étaient autrefois réalisés par des

machines analogiques, ils le sont désormais par numérisation des données et calcul à l'aide de circuits et logiciels informatiques qui permettent beaucoup plus de souplesse et évitent complètement l'accumulation du bruit qui limitait les machines analogiques. Le traitement du signal, fondamental dans des domaines aussi divers que la musique et l'analyse des images, des tremblements de terre ou des signaux émis par le cerveau, est un domaine en soi. Tous ces domaines et bien d'autres sont devenus numériques car ils s'appuient sur l'informatique pour leurs réalisations. De même, l'expression « économie numérique », maintenant d'usage constant, est justifiée par les faits, alors que l'expression « économie informatique » ne serait pas appropriée.

Je continuerai donc à m'appuyer sur les définitions que mes collègues et moi-même avions utilisées lors de la rédaction du rapport de l'Académie des sciences *L'Enseignement de l'informatique en France. Il est urgent de ne plus attendre* (Académie des sciences, 2013), que j'ai eu l'honneur de coordonner :

> 1. Le mot « informatique » désignera spécifiquement la science et la technique du traitement de l'information, et, par extension, l'industrie directement dédiée à ces sujets.
> 2. L'adjectif « numérique » peut être accolé à toute activité fondée sur la numérisation et le traitement de l'information : photographie numérique, son numérique, édition numérique, sciences numériques, art numérique, etc.
> On parle ainsi de « monde numérique » pour exprimer le passage d'un nombre toujours croissant d'activités à la numérisation de l'information et d'« économie numérique » pour toutes les activités économiques liées au monde numérique, le raccourci « le numérique » rassemblant toutes les activités auxquelles on peut accoler l'adjectif numérique. Puisque toute information numérisée ne peut être traitée que grâce à l'informatique, celle-ci est le moteur conceptuel et technique du monde numérique. Par rapport à l'anglais, notre acception du mot « informatique » recouvre *computer science*, *information technology* et ce que l'on entend souvent par *informatics*, alors que l'adjectif « numérique » correspond à *digital*, par exemple dans la correspondance entre « monde numérique » et *digital age*.

Pour ce qui est de « programmation » et « codage », je garderai aussi le mot initial, bien que les informaticiens utilisent volontiers le mot « code » : *montre-moi ton code* est une expression très répandue. On dit cependant « langage de programmation » et pas « langage de codage », et « code » est ambigu car il a d'autres sens assez différents, comme « le code secret Enigma » pour le chiffrement allemand pendant la Seconde Guerre mondiale. Mais les informaticiens aussi changent les mots. Maintenant, on ne programme plus, on « développe », et les programmeurs sont devenus des développeurs. À suivre...

Enfin, je me tiendrai loin de la récente et étonnante formule à la mode « le logiciel des hommes politiques ». Quiconque sait vraiment ce qu'est un logiciel espère que les hommes politiques ont bien plus que du logiciel dans la tête, et que leur taux de bugs est moins élevé que celui de beaucoup de logiciels...

La pensée algorithmique

La communauté des informaticiens est unanime : le cœur du nouveau schéma mental réside dans une nouvelle façon de penser, la *pensée algorithmique*, qui est quasi indépendante du domaine d'application, et qui conduit à de nouvelles formes d'action dont les implications sont nombreuses et profondes. Cette pensée met en avant les notions d'information et d'algorithme, sans pour autant être éthérée puisque ses applications sont la plupart du temps dans la vie matérielle. Le leitmotiv sur lequel mon discours sera construit est simple :

> Avec l'information et par le calcul sur elle, on peut faire assez simplement des choses extraordinaires qu'il est très difficile, voire impossible, de faire avec les moyens physiques traditionnels.

Ce n'est pas la première fois que cette maxime s'applique : avec le télégraphe et le téléphone, les télécommunications

CHAPITRE 2

Les grands principes

Avant de parler d'information et d'informatique, il me paraît utile de rappeler brièvement quelle était la pensée scientifique et technique dominante au XX^e^ siècle. Cela nous permettra de comprendre pourquoi l'informatique n'en est vraiment pas une conséquence directe, ce qui est probablement une des raisons pour lesquelles elle n'a pas été facilement acceptée par les communautés scientifiques de l'époque.

Le triangle du XX^e^ siècle

Nos sens nous donnent immédiatement accès à deux aspects de la nature : *l'espace* et *le temps*, que nous pouvons qualifier et quantifier assez directement de façon plus ou moins précise suivant les besoins. Nous évoluons dans l'espace, et tel arbre est perçu comme plus ou moins loin. Le soleil se lève tous les matins, et les plantes, les animaux et les hommes vieillissent de façon bien visible. L'espace et le temps font l'objet de spéculations religieuses, philosophiques puis scientifiques depuis la nuit des temps. Au XX^e^ siècle, les physiciens nous ont expliqué qu'espace et temps sont bien plus subtils qu'il n'y paraît, et que plus le temps passe moins ils comprennent quelle est sa nature. Mais, sauf dans le chapitre 8 où nous parlerons de simulation des phénomènes physiques, nous en resterons à nos impressions macroscopiques

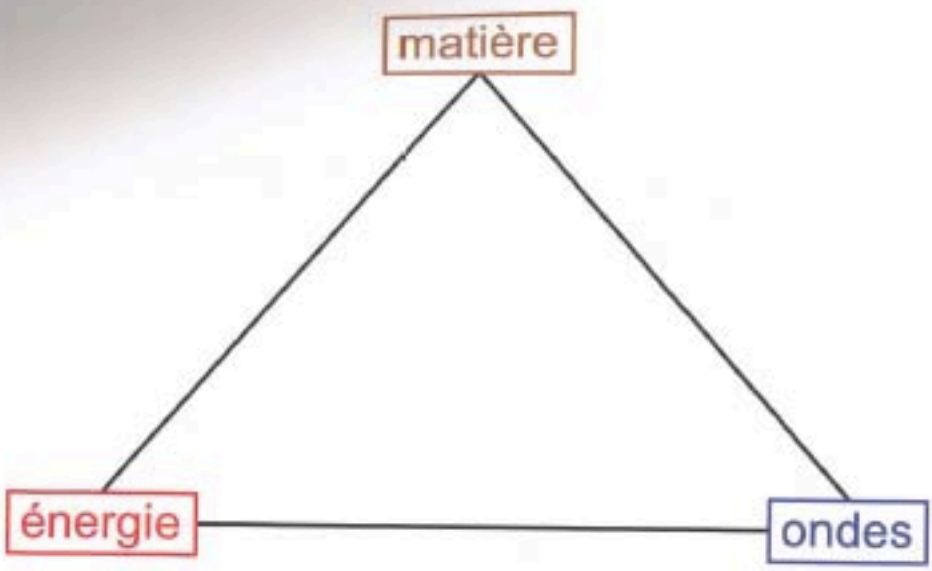

Figure 1. Le triangle du XX^e^ siècle.

et ne déborderons pas de la physique newtonienne – celle qu'on apprend à l'école, avec un temps linéaire et un espace à trois dimensions.

La pensée scientifique et technique s'est petit à petit structurée autour du triangle matière-énergie-ondes, illustré par la figure 1, en commençant par le couple matière-énergie. La matière est un aspect de la nature que nous percevons très directement : un objet solide a une forme bien définie, est lourd ou léger, régulier ou non ; un liquide mouille et prend la forme des récipients, alors qu'un gaz ne se voit pas mais fait du vent et peut sentir bon, mauvais, ou pas du tout. L'énergie reste directement perceptible, bien que plus complexe : on sent bien que monter une côte est plus fatigant que la descendre, mais dire de combien n'est pas simple. Pour les Grecs, cette perception duale conduisait à penser que notre univers était fait de quatre éléments : la terre, l'eau et l'air pour la matière, et le feu pour l'énergie. Cette vision sommaire s'est évidemment perfectionnée petit à petit, et l'explosion des sciences du XVII^e^ au XIX^e^ siècle est précisément due à une meilleure compréhension des relations entre matière et énergie. C'est ce qui a conduit aux moteurs,

cœurs de la grande révolution de l'industrie et des transports aux XVIII[e] et XIX[e] siècles, ainsi qu'aux débuts de la chimie.

Vers le milieu du XIX[e] siècle, la compréhension de plus en plus fine de l'électricité, de l'électromagnétisme, de l'acoustique et des vagues sur la mer a conduit à la mise en avant des *ondes*. Cela a provoqué les grandes révolutions de la communication à distance avec le télégraphe électrique (vers 1840), le téléphone (vers 1870), l'éclairage électrique (1879) puis la radio (vers 1890) et la télévision (vers 1925). Ces nouveautés ont conduit à des changements très profonds dans l'organisation de la vie de tous les jours et celle de la société. Qu'on se rappelle simplement que les messages entre l'Angleterre et les Indes ou l'Amérique mettaient désormais quelques secondes au lieu de quelques mois (Standage, 2014), que la nuit n'a plus été l'opposé absolu du jour, et que les ombres sont devenues fixes au lieu de bouger constamment à cause de l'instabilité des flammes[1].

Au XX[e] siècle, la pensée scientifique et technique est restée largement ancrée sur ce triangle matière-énergie-ondes. Les sciences physiques, qui ont fondé les progrès technologiques majeurs de ce siècle, ont profité d'une modélisation mathématique efficace, alors que les sciences de la vie sont restées plutôt descriptives.

INFORMATION ET ALGORITHMES : DU TRIANGLE AU TÉTRAÈDRE

À partir du premier tiers du XX[e] siècle, deux notions bien différentes ont été lentement théorisées puis rapidement mises en pratique à la fin du siècle : celles d'*information* et de *calcul algorithmique*, qui fondent les sciences et techniques informatiques auxquelles cet ouvrage est consacré. Cela a conduit à étendre le triangle en un tétraèdre illustré par la figure 2, avec les conséquences considérables que

1. Merci à Jean-Marc Lévy-Leblond pour cette belle remarque. L'ambiance devait être bien différente quand les ombres bougeaient !

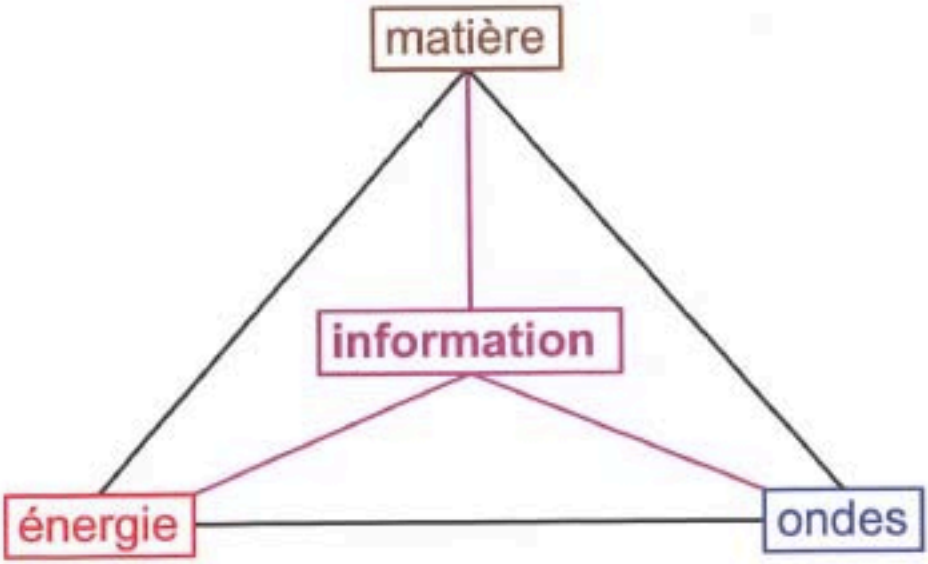

Figure 2. Le tétraèdre du XXI^e^ siècle.

nous observons maintenant. Dans la figure, j'ai volontairement placé l'information au centre, car c'est son exploitation systématique à l'aide de l'ordinateur qui a modifié en profondeur les façons de faire traditionnelles et permis de réaliser des choses tout à fait nouvelles et insoupçonnables au XX^e^ siècle.

L'ancien triangle n'a pas disparu pour autant : l'ordinateur lui-même est bien sûr fait avec de la matière, de l'énergie et des ondes. Mais, comme l'a montré Alan Turing dès son premier article fondateur de 1936 (Turing, 1936), c'est une *machine universelle* : ses mêmes composants physiques peuvent être utilisés identiquement *pour toutes les tâches liées à l'information et au calcul*, quels que soient leurs objectifs. De ce point de vue, même s'ils n'ont pas les mêmes performances, tous les ordinateurs se valent, quelle que soit la façon dont ils sont construits.

Les piliers de l'informatique

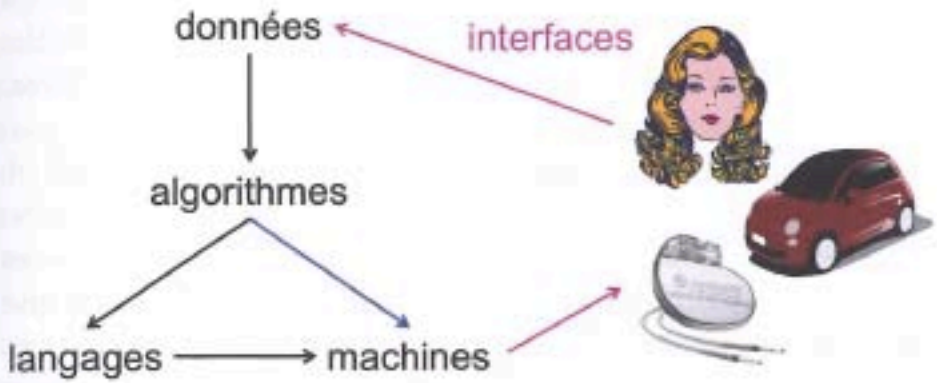

Figure 10. Les piliers de l'informatique.

Il est temps d'entrer plus directement dans la structure interne de l'informatique moderne. Comme illustré par la figure 10, elle repose sur quatre piliers : *données*, *algorithmes*, *langages* et *machines*, auxquels s'ajoutent les *interfaces* qui permettent les communications homme-ordinateurs, ordinateurs-ordinateurs et ordinateurs-objets informatisés, ici une voiture et un pacemaker. La flèche bleue allant des algorithmes aux machines est là pour illustrer le fait que les machines elles-mêmes sont entièrement conçues à l'aide d'algorithmes.

Les données

Les *données* sont l'objet de travail de l'informatique. Ce sont des quantités brutes numérisées (montants financiers, mesures physiques, textes, etc.), ou des groupements structurés

CHAPITRE 4

Les télécommunications : du fil à l'air

La communication entre les personnes et les institutions a toujours été un enjeu majeur du développement des sociétés. C'est incontestablement un des domaines où l'informatique a apporté le plus, au point que l'industrie des télécommunications, autrefois forte et indépendante, n'est plus vraiment dissociable de l'industrie informatique en général. Au début des années 1970, il était difficile et long d'avoir un téléphone fixe. Maintenant, il devient banal d'avoir un smartphone connecté à Internet, et les objets communiquent entre eux autant que les hommes. Nous allons étudier ici quelques-unes des causes centrales de cette évolution aussi rapide que profonde. Mais, auparavant, je pense utile de s'intéresser à une révolution bien plus ancienne mais tout aussi puissante, celle du télégraphe, surnommé par Tom Standage l'« Internet victorien » (Standage, 2014).

Le télégraphe, une révolution méconnue

Le télégraphe manuel, introduit par Chappe à la fin du XVIII[e] siècle, avait donné à la France une bonne avance : les informations, militaires en particulier, pouvaient être

CHAPITRE 5

Internet, des hommes aux objets

Internet, dont le nom signifie littéralement le « réseau des réseaux », est la partie la plus visible de l'informatique. Pour la plupart des gens, il se réduit au Web, la toile d'araignée des sites avec lesquels ils peuvent interagir, bien qu'il contienne d'autres fonctions importantes. Internet existe depuis longtemps pour les chercheurs, mais il n'a réellement atteint le public qu'à la fin des années 1990, avec l'explosion du Web puis l'entrée fracassante des moteurs de recherche qui allaient tellement modifier les usages qu'on n'arrive même plus à se souvenir de comment on faisait les choses avant eux.

Les impacts d'Internet sont constamment discutés dans les cercles les plus variés. Il homogénéise le monde et offre des possibilités entièrement nouvelles dont on découvre tous les jours la puissance. Il pose aussi des problèmes fondamentaux dont beaucoup ne sont pas vraiment nouveaux mais se trouvent amplifiés de plusieurs ordres de grandeur par l'universalité et la vitesse du grand réseau : l'accès à l'information et les questions relatives à sa pertinence et à sa possession, la diffusion ultrarapide et à grande échelle d'informations justes ou fausses et de courrier « électronique » important ou ennuyeux, la facilité des contacts mais aussi les atteintes possibles à la vie privée, et, de façon générale, la sécurité des données et objets qui y sont connectés. Les discussions associées sont difficiles, les avis souvent tranchés et pas

Figure 18. Résultat de PageRank (source Felipe Micaroni Lalli, CC BY-SA 2.5, https://commons.wikimedia.org/w/index.php?curid=2776582).

La bonne idée, illustrée par la figure 18, a été de dire qu'une page doit être vue comme populaire *si et seulement si elle est pointée par de nombreuses pages elles-mêmes populaires*.

Tout cela s'exprime par une formule mathématique simple mais profonde, qu'il est utile de présenter ici au moins pour les amateurs de formules. Selon les notations de Brin et Page, appelons N le nombre total de pages, A la page courante, $T1$, $T2$, ..., Tn les pages pointant sur A, et $C(Ti)$ le nombre de liens sortant de la page Ti. La probabilité (popularité) $PR(A)$ est définie par l'équation suivante en fonction de celles des $PR(Ti)$:

$$PR(A) = (1 - d)/N + d \times (PR(T1)/C(T1) + \ldots + PR(Tn)/C(Tn))$$

Vu dans l'autre sens, chaque page Ti distribue sa probabilité multipliée par d équitablement à toutes ses pages

CHAPITRE 6

Images, sons et cartes numériques

Comme je l'ai déjà expliqué au chapitre 2, les domaines de la production, de la diffusion, et de l'analyse de l'image et du son font partie de ceux qui ont été le plus vite et le plus profondément révolutionnés par l'informatique. Ce sont aussi des domaines essentiels de notre vie, car la vue et l'ouïe sont nos sens principaux. La révolution informatique de l'image et du son ne touche pas qu'aux moyens de stockage et de diffusion de photos, musiques et films dont nous avons déjà parlé : les domaines d'action s'y sont aussi rapidement étendus.

Pour l'image, citons la photographie numérique, que nous étudierons en détail dans ce chapitre, l'imagerie satellitaire de la planète, utile de la météorologie à la cartographie en passant par l'agriculture, l'imagerie médicale multimodale (radiographie, scan, échographie, différentes modalités d'IRM) que nous étudierons au chapitre 7, l'imagerie de l'univers discutée au chapitre 8, la surveillance, la reconnaissance automatique de visages ou d'objets, la conduite automatique des voitures, etc. De son côté, la synthèse d'images réalistes, techniquement assez différente, sert au cinéma à la simulation des phénomènes naturels et à la réalité virtuelle ; voir par exemple le cours de Marie-Paule Cani sur le sujet au Collège de France (Cani, 2015). Pour comprendre la taille du champ de l'imagerie numérique, il suffit de jeter un œil

CHAPITRE 7

L'informatisation de la médecine

Médecine et informatique ont des interactions variables suivant les domaines. Les aspects informatiques les mieux connus du grand public sont encore sous-représentés dans l'interaction des patients avec le système médical : il n'est pas souvent possible de prendre des rendez-vous par Internet, et, selon mon expérience, il est virtuellement impossible de faire suivre le dossier de son ancien généraliste au nouveau quand on déménage. La carte Vitale a mis un temps considérable pour s'imposer dans le paysage, et j'ai encore trouvé récemment des médecins qui ne la prennent pas. C'est bien sûr différent au sein des hôpitaux qui gèrent informatiquement les données des patients, y compris les données lourdes de leurs examens, mais je ne suis pas sûr qu'il soit uniformément facile de les faire passer d'un hôpital à un autre. Cependant, Internet est très utilisé en dehors des circuits médicaux officiels. De nombreux sites discutent des effets secondaires des médicaments ou de thérapies alternatives, mais lire le *Vidal* qui n'est pas fait pour l'individu standard a tendance à lui faire ressentir tous les effets secondaires d'un coup. Et il n'est vraiment pas facile de savoir où trouver des informations accessibles et fiables.

À l'opposé, certains domaines de la médecine (et de la dentisterie, dont je ne parlerai pas) sont de grands consommateurs et créateurs d'informatique. Les minuscules

CHAPITRE 8

Vers une informatisation massive des sciences

Le but de ce chapitre est de montrer à l'aide de quelques exemples comment la pensée informatique s'introduit en grand dans les sciences, avec des impacts déjà considérables qui vont aller en s'amplifiant rapidement. Nous discuterons des sciences naturelles et des mathématiques. Mais, avant d'entrer dans le vif du sujet, il faut faire un détour par une brève analyse des rôles respectifs des instruments et des mathématiques dans les sciences naturelles.

LE RÔLE FONDAMENTAL DES INSTRUMENTS DANS LES SCIENCES NATURELLES

Puisqu'une grande partie des phénomènes intéressant les chercheurs des sciences naturelles ne sont pas visibles à l'œil nu, l'instrumentation y est essentielle : les lunettes astronomiques puis télescopes et radiotélescopes ont permis de scruter le ciel, les microscopes optiques puis électroniques ont permis de scruter l'infiniment petit, les chronomètres puis horloges atomiques ont permis de mesurer finement le temps, etc. Jusque dans les années 1970, presque tous les instruments ont reposé sur les principes de la physique classique, et donc sur le triangle matière-énergie-ondes. Le recours à

CHAPITRE 9

Bugs et trous de sécurité : deux dangers de l'informatique

Passons maintenant à deux côtés sombres mais incontournables de l'informatique, ceux liés aux bugs et aux trous de sécurité des systèmes informatisés, avant de montrer au chapitre suivant que rien n'est désespéré car nous disposons de méthodes de plus en plus efficaces pour en réduire le nombre et l'impact, voire les éliminer réellement dans certains cas importants. Comme déjà expliqué au chapitre 3, les propriétés qui nous intéressent sont la sûreté et la sécurité. La sûreté est la garantie que les systèmes font ce qu'ils doivent faire et ne font pas ce qu'ils ne doivent pas faire, alors que la sécurité est la garantie qu'ils ne peuvent pas être pénétrés de l'extérieur pour modifier leur fonctionnement ou voler, détruire ou chiffrer des données. Les deux aspects sont fortement reliés, car les trous de sécurité sont le plus souvent liés à des microbugs passés inaperçus car n'étant pas reliés au fonctionnement normal et donc pas détectés par les tests fonctionnels classiques, souvent restreints aux cas considérés comme normaux. On réduit souvent la sécurité aux problèmes concernant les données. Mais ceux concernant le fonctionnement des systèmes sont tout aussi dangereux, voire plus : qu'on pense aux catastrophes que peut engendrer une prise de pouvoir sur les réseaux de distribution électrique ou sur les feux rouges d'une ville informatisée. Les récentes attaques mettant hors service des

CHAPITRE 10

Comment rendre l'informatique plus sûre

Après avoir présenté cette liste d'échecs ou de problèmes dus à des bugs identifiés comme tels après coup, abordons une des questions centrales de l'informatique : comment concevoir et mettre en place des systèmes informatisés qui marchent comme nous le souhaitons, sans être affectés, voire détruits, par les bugs que nous y avons nous-mêmes introduits ? Ceci suppose de répondre à trois questions : pouvons-nous vraiment caractériser ce que nous appelons un bug, pouvons-nous détecter et éradiquer les bugs dans tous les cas, ou devons-nous accepter de vivre avec ? Aucune de ces questions n'est simple et toutes méritent d'être abordées ici.

À l'heure actuelle, on leur applique trois types de réponses : d'abord, mieux travailler pour réduire le nombre de bugs dès la conception ; ensuite, vérifier en permanence son travail pour attraper les bugs avant la mise en service, soit par des batteries de tests, soit à l'aide de méthodes formelles présentées dans ce chapitre ; troisièmement, mettre des garde-fous dans les systèmes pour détecter les bugs résiduels suffisamment tôt avant que leurs conséquences ne deviennent catastrophiques, ce qui permet aussi de transmettre de bons rapports de bugs aux concepteurs. Tout projet bien conçu applique ces trois principes, mais, dans certains domaines

CHAPITRE 11

Où va nous mener l'informatisation du monde ? Une vision personnelle

J'espère que les analyses présentées dans les chapitres précédents ont permis au lecteur de se construire un meilleur schéma mental sur ce qu'est réellement l'informatisation du monde, ses apports, ses difficultés intrinsèques, et d'où vient sa puissance. J'ai surtout voulu illustrer le fait fondamental que la pensée et l'action informatiques sont les mêmes partout même si leurs effets sont très variés : elles font appel à la même science et aux mêmes outils quel que soit le champ d'application.

Il me faut maintenant aborder une question qu'on me pose toujours, comme à tous mes collègues : où va nous mener l'informatisation galopante du monde ? Cette question est évidemment légitime, mais je serais bien présomptueux de prétendre y répondre de façon vraiment fiable. Je vais cependant m'y essayer, mais après quelques bonnes pages de précautions pour que le lecteur comprenne clairement que je vais formuler surtout des hypothèses et pas des certitudes, ainsi que quelques pistes concrètes d'action. Contrairement aux chapitres précédents, je prendrai donc soin de ne jamais mélanger le « je » et le « nous ». Je parlerai aussi beaucoup de la situation de mon pays, la France, celui que je connais le mieux et dans lequel je m'investis le plus à l'heure actuelle. Mais la situation ne me paraît pas très différente dans les autres pays européens (aux États-Unis et en Asie, c'est évidemment très différent).

Je ne sais pas ce qu'il faut faire pour réduire l'emprise écologique de l'informatique, sinon mettre les gens au courant de son existence. Ce n'est pas vraiment fait pour l'instant, contrairement à d'autres dangers dont je parlerai plus loin.

QUID DE L'ORDINATEUR QUANTIQUE ?

Un sujet très à la mode est l'ordinateur quantique, qui pourrait, selon ce que j'entends dans les médias, « être *infiniment* plus rapide que les ordinateurs actuels ». S'ensuit une explication confuse de ce qu'est la superposition quantique, sujet que même les meilleurs physiciens n'ont vraiment réussi à expliquer sans équations au commun des mortels, pour de bonnes raisons d'ailleurs : ça ne fait vraiment pas partie de la vie commune.

Quelle est la situation réelle ? Je n'en suis pas un spécialiste, donc je ne répondrai à la question qu'avec précaution. D'après mes renseignements, il y a deux problèmes assez distincts, l'un physique et l'autre algorithmique. Sur le plan physique, l'ordinateur quantique n'existe pas encore, même s'il y a beaucoup de recherches sur le sujet. Beaucoup de chercheurs essaient différentes technologies pour fabriquer des *qubits*, ces fameux « bits quantiques » qui permettent la superposition quantique de valeurs. Mais la superposition qubit à qubit n'est pas très intéressante, car c'est la coordination de l'ensemble des superpositions dans le système global considéré qui est pertinente.

Considérons la question du cassage du chiffrement RSA décrit dans l'annexe 12.1. Il s'agit de factoriser un nombre $n = pq$ où p et q sont deux grands nombres premiers, en ne connaissant ni p ni q. Avec un ordinateur classique, on ne connaît aucun algorithme capable de le faire en temps raisonnable pour de très grands nombres (2 048 bits par exemple). En 1994, le mathématicien Peter Shor, professeur au MIT, a réalisé une avancée fondamentale en montrant qu'un ordinateur quantique de taille suffisante pourrait effectivement réaliser cette factorisation très rapidement. Ce résultat,

qui s'applique aussi à presque tous les chiffrements actuellement utilisés, est scientifiquement remarquable et potentiellement grave en pratique car casser tous les chiffrements d'Internet provoquerait un chaos de première grandeur. On pourrait certes remplacer en théorie les chiffrements actuels par d'autres résistants au quantique, mais ce ne serait pas une mince affaire.

Le premier problème est qu'on ne dispose pas encore d'ordinateur quantique, sinon tout petits. Les bits quantiques sont très délicats et tendent à perdre très vite leur capacité globale de superposition, surtout dès qu'il y en a beaucoup. C'est ce qu'on appelle la décohérence. On ne sait pas quand il sera possible d'avoir des machines universelles de taille raisonnable. Certains annoncent cinq ans, d'autres quarante, mais je ne crois pas du tout au sens de ce genre de chiffres, qu'il faut bien fournir pour être financé. En pratique, pour la factorisation, le record actuel implémenté est la factorisation de 143 sur 4 qubits. Elle peut être étendue facilement à 56 153 qui n'est en fait pas plus compliqué car ce nombre appartient à une classe facilement factorisable avec un ordinateur standard ; 175 étant plus compliqué pour des raisons mathématiques inexplicables ici, il n'est pas encore factorisé en pratique.

De plus, contrairement à ce qu'on entend souvent, un ordinateur quantique général ne sera pas *infiniment* plus rapide qu'un ordinateur classique, mais seulement exponentiellement plus rapide dans les bons cas (ce qui est déjà excellent[1]). Un autre problème majeur est que l'algorithmique quantique est difficile, et que, d'après mes informations, on ne connaît que peu d'algorithmes réellement utilisables. Enfin, un ordinateur quantique est un objet physique très complexe, qui ne va pas tenir dans votre poche et qui ne pourra peut-être servir que de fantastique coprocesseur à

1. La fameuse machine quantique D-Wave, effectivement assez grosse, a été souvent présentée à tort comme un ordinateur quantique général. Elle n'exécute en fait qu'un seul algorithme, l'optimisation par recuit simulé. C'est déjà tout à fait intéressant, mais pas magique.

CHAPITRE 12

Annexes

12.1. La science des algorithmes

L'algorithmique, science des algorithmes, est centrale en informatique. Je parlerai d'abord des algorithmes classiques, ceux qui prennent des données, calculent, puis rendent des résultats. Ils sont très nombreux, et l'algorithmique a pour objectif d'en découvrir de nouveaux ou d'améliorer les anciens, en veillant toujours à garantir leur *terminaison* sur les données concernées, leur *correction*, qui exprime que l'algorithme fait bien ce pour quoi il a été conçu, et leur *complexité*, qui est leur coût en temps de calcul, mémoire, ou énergie. Analyser ces propriétés fait toujours appel aux mathématiques, et y conduit souvent à de nouveaux développements.

DE L'INDÉCIDABILITÉ À LA COMPLEXITÉ ALGORITHMIQUE

Dès 1936, Turing avait formalisé la notion d'algorithme à l'aide de sa fameuse machine. Il avait aussi montré que tout n'était pas faisable par des algorithmes : son théorème d'*indécidabilité du problème de l'arrêt* établissait l'impossibilité de construire un algorithme T se terminant toujours et décidant si un algorithme précis A travaillant sur des données D va se terminer ou non. Ce résultat fondamental établit la limite la plus dure de l'informatique : si on peut construire