

LES ALGORITHMES PARTOUT... COMMENT ET POUR QUOI?

Les algorithmes imprègnent nos vies et l'ensemble de la société. Comprendre le fonctionnement et la puissance de ces outils est essentiel, explique leur spécialiste **Claire Mathieu**, pour éclairer les choix à faire en vue de les utiliser au service du bien commun.

Dans un contexte où les médecins, infirmières et autres soignants travaillent d'arrache-pied à sauver des vies, où des personnes de notre entourage se retrouvent hospitalisées en état critique, et où chacun court un risque de se retrouver soi-même atteint par le coronavirus, quel sens y a-t-il à faire de la recherche sur les algorithmes ? Tout ces calculs, ces modèles, ces constructions, ces théorèmes semblent soudain futiles, et on remet son travail en question.

INFORMATIQUE

Dans le cadre de notre partenariat avec l'Académie des sciences, les académiciens nouvellement élus fin 2019 présentent un éclairage sur leur discipline et ses enjeux scientifiques, éthiques, politiques et sociétaux, à travers leur expérience personnelle.

Et pourtant, l'une des sources historiques de l'étude des algorithmes – l'algorithmique – donne une piste encourageante : les débuts de la recherche opérationnelle, au cours de la Seconde Guerre mondiale, sous l'impulsion d'un groupe pluridisciplinaire de chercheurs menés par Blackett, prix Nobel de physique. Dans une note, en 1941, il explique que « pour des questions opérationnelles, les scientifiques peuvent encourager une réflexion basée sur les chiffres (de nos jours, on dirait sans doute "les données"), et aident ainsi à éviter de fonder des décisions de guerre sur des réactions émotionnelles ». Leurs travaux algorithmiques de localisation optimale des antennes ont doublé l'efficacité du système britannique de défense aérienne avec radars. Après la guerre, les chercheurs en recherche opérationnelle ont travaillé à la plani-

fication à grande échelle, pour des problèmes logistiques d'allocation de ressources, par des approches calculatoires et la conception de l'algorithme du simplexe, qui encore maintenant est l'un des grands succès de l'algorithmique. Il permet de résoudre des problèmes de type : j'ai une certaine quantité de tissu et d'élastiques, un certain nombre d'employés et de machines à coudre, un certain besoin en masques de tissu, surblouses et charlottes, comment répartir au mieux mes ressources pour produire au plus près des besoins ? Qu'il serait donc utile d'avoir une semblable équipe de choc !

PAS D'INTERNET SANS ALGORITHMES

Outre la conception d'algorithmes, l'analyse d'algorithmes aide à savoir à quel point les algorithmes proposés sont fiables. Longtemps le pire cas était roi : pour des algorithmes utilisés sur des données arbitraires, venant de contextes très divers, impossible de modéliser de façon fiable la structure de ces données, et les seules garanties sont celles qui sont valides pour n'importe quelle donnée. Avec les développements d'algorithmes plus spécialisés, on peut tenter de construire et valider des modèles probabilistes « réalistes » des données. Cette direction de recherche, traditionnelle dans les autres sciences, est relativement récente en algorithmique.

Et pourtant, dans le confinement actuel, les gens se tournent vers Internet pour étudier, travailler, s'informer, se distraire, garder le contact avec leurs proches. Rien de cela ne serait possible sans les algorithmes exploitant la structure d'Internet : l'accès



PROFIL

Informaticienne, membre de l'Académie des sciences, **Claire Mathieu** est directrice de recherche CNRS à l'Institut de recherche en informatique fondamentale (Irif). Ses recherches portent en particulier sur la conception d'algorithmes, pour trouver des solutions quasi optimales à des problèmes difficiles à résoudre exactement, et actuellement sur les algorithmes sociétaux.

à Internet est vu comme allant de soi, au point que, le jour où toute une génération se connecte simultanément aux sites d'enseignement à distance, les difficultés consécutives à la surcharge sont perçues comme « anormales ». Mais cette prétendue « normalité » repose sur une imbrication complexe de réseaux, de systèmes et d'algorithmes.

LES ROIS DE LA COMPRESSION DE DONNÉES

Sans algorithmes de routage de paquets sur le réseau, Internet s'écroulerait. Pour le comprendre, imaginons une mairie chargée de restructurer le réseau de bus qui dessert l'agglomération. Il faut tenir compte des lieux où les gens habitent, des lieux particulièrement fréquentés comme les gares ou centres commerciaux, et du nombre prévisionnel de personnes se déplaçant par bus pour se rendre dans ces lieux ; il faut aussi tenir compte de la durée du trajet en heures creuses ou en heures de pointe, des délais maximaux

L'algorithmique permet de résoudre des problèmes de type : j'ai une certaine quantité d'élastiques et de tissu, un certain nombre d'employés et de machines à coudre, un certain besoin en masques, surblouses et charlottes. Comment répartir au mieux mes ressources ?

supportables ; il peut aussi être pertinent de donner des priorités lorsqu'il y a risques d'embouteillages, par exemple pour l'accès aux hôpitaux, aux universités, ou pour les personnes sans voiture. Un réseau bien conçu fera gagner du temps à tous les utilisateurs et assure un service minimal même en période de saturation. La même logique prévaut au routage des paquets qui contiennent les informations téléchargées sur nos portables.

Sans algorithmes de compression de données, on ne pourrait pas télécharger musique et vidéos. Nous utilisons naturellement des méthodes de compression dans certains cas simples, par exemple, au lieu de lire 010101010101010101, nous dirons « répéter 01 dix fois ». Les algorithmes de compression repèrent répétitions et redondances, et les exploitent pour construire une représentation plus compacte de l'information, quitte à la simplifier quelque peu.

LE CARBURANT DES MOTEURS DE RECHERCHE

Sans algorithmes de chiffrement de données, on ne pourrait ni consulter son compte en banque ni faire des achats en ligne. La plupart de ces algorithmes reposent sur des opérations mathématiques faciles à faire mais difficiles à défaire : par exemple, il est facile de multiplier deux nombres, même grands, mais beaucoup plus difficile de décomposer un grand nombre en produit de facteurs premiers. Pour chiffrer un message, on se sert d'un grand nombre, mais pour le déchiffrer, il faut connaître la factorisation de ce nombre. Ainsi, lorsque les données sur notre carte de crédit circulent sur Internet, elles sont chiffrées pour qu'une personne malhonnête ne puisse pas les lire si elle ne sait pas factoriser.

Sans méthodes sophistiquées de sélection des réponses les plus pertinentes, ni les moteurs de recherche ni les algorithmes de recommandation n'existeraient. Les moteurs de recherche sur le Web suivent le principe du mouton de Panurge : si quelqu'un a déjà suivi un lien particulier, c'est sans doute une bonne idée, alors faisons pareil. Au lieu d'analyser l'intérêt d'une page Web directement, il suffit de se reposer sur l'avis des utilisateurs passés, et le fait pour eux d'avoir mis un lien sur une page est un puissant indicateur de son intérêt. Ainsi se construit une théorie du comportement des usagers qui, en ricochet, a des implications sur le choix des pages à présenter aux futurs utilisateurs. Cette approche algorithmique, qui complète une analyse directe du contenu de la page par une analyse indirecte de ce que les utilisateurs en font, a révolutionné les moteurs de recherche, mais risque d'être manipulée. On)))

Qu'a-t-on le droit de faire? Qu'est-ce qui est juste? Y répondre demande d'entamer une réflexion éthique très nouvelle pour le domaine. Un algorithme qui contribue à structurer la société doit emporter l'adhésion de ses citoyens et, pour cela, la transparence est essentielle.

» voit apparaître des « fake news » qui se propagent à grande vitesse par un phénomène de propagation d'information semblable à une contagion. La détection et le contrôle de tels phénomènes sont des questions largement ouvertes.

Ainsi, on peut avoir l'impression qu'en ce moment, la société est à l'arrêt, mais combien plus extrême serait notre situation présente sans Internet et les algorithmes sous-jacents! Les algorithmes imprègnent la société sans toujours qu'on s'en rende compte.

Outre ces algorithmes omniprésents mais invisibles, il y a actuellement une montée en puissance d'algorithmes qui eux sont très visibles et qui aident à prendre des décisions traditionnellement dévolues aux humains.

CAPABLE DE POSER UN DIAGNOSTIC VALIDE

Les méthodes d'analyse d'image permettent d'étudier des radios pour détecter des anomalies et signaler au radiologue les images suspectes. Beaucoup de ces méthodes relèvent de la science des données : elles « apprennent » par l'expérience à partir des radios d'anciens patients, les uns sains, les autres malades. L'apprentissage profond, à partir d'énormes quantités de radios, ajuste les paramètres de la fonction calculée de façon que, pour les anciens patients, elle distingue au mieux les personnes saines des malades. Cela donne une mystérieuse combinaison de coefficients, une sorte de recette magique qui pose un diagnostic valide pour les patients de la base de données. On espère alors qu'elle posera également un diagnostic correct pour les nouveaux patients. Cette méthode nécessite d'avoir des données fiables, mais aussi de faire rapidement de gros calculs d'optimisation, brique purement algorithmique dans la construction.

Les affectations d'étudiants dans les établissements de l'enseignement supérieur par la plateforme Parcoursup utilisent des algorithmes de « mariage stable » pour chercher une affectation qui réponde au mieux

aux souhaits des étudiants et des formations qui les recrutent. L'algorithme de mariage stable est bien plus satisfaisant que de laisser chacun faire son marché.

Mais, dès que les décisions algorithmiques touchent directement à la vie des personnes, des lièvres sont levés : qu'a-t-on le droit de faire ? Qu'est-ce qui est juste ? N'y aura-t-il pas des personnes « indûment » avantagées par l'algorithme ? Y répondre demande d'entamer une réflexion éthique très nouvelle pour le domaine. Un algorithme qui contribue à structurer la société doit emporter l'adhésion de ses citoyens et, pour cela, la transparence est un élément essentiel. Pour un algorithme, que signifie la transparence ? Quelques principes préliminaires émergent, mais c'est dans leur application pour chaque problème, au cas par cas, que pourra s'approfondir la réflexion.

L'an prochain, le budget de la recherche sera certainement en bonne partie consacré à la santé, et avec raison. Une partie des chercheurs en informatique et en algorithmique réorienteront leur recherche pour apporter un soutien à la recherche médicale. Mais il restera toujours ceux dont les travaux permettent qu'on trouve « normal » de faire fonctionner une grande partie de la société par le biais de canaux numériques et en passant par l'intermédiaire d'un monde informatique et numérique si bien intégré à notre quotidien qu'il en devient presque invisible. ★

Beaucoup de méthodes d'analyse d'images médicales relèvent de la science des données. Le principe : un « apprentissage » par l'expérience à partir des radios d'anciens patients, les uns sains, les autres malades.

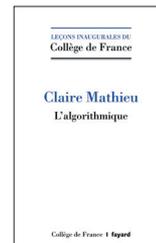


LUCAS BARROUET / ANP

EN SAVOIR PLUS

LE SITE DE L'ACADÉMIE DES SCIENCES :

www.academie-sciences.fr



« **L'Algorithmique** », de Claire Mathieu, Fayard, coll. « Les leçons inaugurales du Collège de France », 2018.

Le texte de cette leçon inaugurale, prononcée le 16 novembre 2017, très accessible, est disponible en ligne : <https://books.openedition.org/cdf/5618?lang=fr>

« **Le Temps des algorithmes** », de Serge Abiteboul et Gilles Dowek, le Pommier, 2017.