

Comprendre les algorithmes pour les apprivoiser

Claire MATHIEU, directrice de recherche au CNRS,

est nommée titulaire de la chaire *Informatique et sciences numériques* 2017/2018

Une chaire du Collège de France créée en partenariat avec Inria

- Leçon inaugurale le jeudi 16 novembre 2017 à 18h00 -

Faut-il avoir peur des algorithmes ? Comment utilisent-ils les milliers de données qu'ils collectent sur nous et surtout dans quels buts ? Sont-ils discriminants ? Jusqu'où empiètent-ils sur notre libre arbitre ? Les questions sont vastes face à ces méthodes de résolution de problèmes et de traitement des données qui s'immiscent dans notre quotidien. Leur méconnaissance génère aujourd'hui inquiétudes et appréhensions.

« *Le problème est qu'en matière d'algorithmes la pratique a pris énormément d'avance sur la théorie et la compréhension en profondeur des mécanismes mathématiques qui les régissent. Ce retard induit effectivement des risques de manipulation sans contrôle* », estime Claire Mathieu, « *mais s'ils sont bien conçus et compris, les algorithmes sont aussi de puissants outils d'optimisation qui peuvent contribuer au bien commun, comme par exemple dans l'analyse des données médicales* ».

Comprendre, décrire, modéliser ou encore concevoir et développer de nouvelles techniques : les champs de recherche de l'algorithmique sont vastes et subissent de profonds développements. Ils constituent un enjeu fondamental de l'informatique du 21^{ème} siècle, notamment face à des données de plus en plus massives et dispersées, ainsi qu'un enjeu éthique au regard de l'utilisation de ces données. « *Une question importante du domaine pour la décennie à venir est de comprendre le 'pourquoi' et le 'dans quel but' des algorithmes qui ont des retombées sociétales directes* », souligne Claire Mathieu.

Directrice de recherche au CNRS, Claire Mathieu est l'une des plus éminentes représentantes de ce domaine. Ses travaux portent notamment sur les algorithmes d'approximation permettant de calculer une solution approchée à un problème trop difficile à résoudre exactement, ou encore sur les algorithmes en ligne qui doivent gérer des flux constants de données. Elle s'est récemment intéressée à la modélisation mathématique de réseaux sociaux.

Algorithmes et discrimination / Etude du plafond de verre dans les réseaux sociaux

Peut-on prouver mathématiquement qu'il existe effectivement un plafond de verre dans les réseaux sociaux - à savoir « une barrière invisible mais inébranlable qui empêche les femmes ou une minorité de s'élever jusqu'aux échelons supérieurs de la hiérarchie » - et quelles conditions doivent être réunies pour qu'un tel plafond émerge ? Claire Mathieu et ses équipes ont modélisé ce phénomène, très largement étudié en sociologie et l'ont mathématiquement démontré.

Claire Mathieu, professeur invitée sur la chaire annuelle *Informatique et sciences numériques* pour l'année académique 2017/2018, partagera au Collège de France son expertise de l'algorithmique : problèmes phares, enjeux, questions fondamentales, diversité des techniques, nouvelles problématiques, nouvelles méthodes de calcul. Sa leçon inaugurale se tiendra le jeudi 16 novembre à 18h00. Ses cours, ouverts à tous sans inscription, débuteront le mardi 28 novembre (*voir p 4*). L'ensemble de son enseignement sera disponible sur le site www.college-de-france.fr



Comprendre les algorithmes pour les apprivoiser

Par Claire Mathieu

Les algorithmes, quand ils sont bien conçus car bien compris, contribuent au bien social et sont un outil de transformation de la société. Ils y prennent une place grandissante, que ce soit pour des applications informatiques telles que gérer efficacement des données, pour les transmettre rapidement et de façon fiable, pour faire des calculs, pour optimiser l'allocation de tâches, ou de plus en plus pour des usages en société comme le calcul d'itinéraire, les recommandations de lectures ou de films, l'affectation des étudiants avec APB, etc.

Très souvent, ces algorithmes sont opaques, mystérieux, voire effrayants. Comment marchent-ils ? Quelles sont les étapes du calcul ou de la prise de décision ? Pourquoi marchent-ils ? Quelles sont les propriétés mathématiques qui expliquent leurs performances ? Le "comment" et le "pourquoi" sont les deux facettes de l'étude des algorithmes : la conception (comment ça marche) et l'analyse (pourquoi ça marche).

En général, l'analyse d'algorithmes sert d'abord à montrer qu'un algorithme résout correctement un problème, puis qu'il le fait en un temps raisonnable, de façon fiable et en restant économe en ressources.

Un sous-domaine fondamental largement développé depuis une vingtaine d'années concerne les algorithmes d'approximation. Un tel algorithme, plutôt que de rechercher une solution exacte et précise, se contente d'une solution approchée. On les utilise pour résoudre les problèmes trop difficiles pour être résolus complètement en un temps raisonnable. On s'autorise alors à accepter des solutions dont la valeur serait, disons, à 5% de la valeur optimale. Par exemple, prenons le problème de la livraison de marchandises. Comment optimiser les trajets d'un camion qui doit livrer des marchandises, connaissant le lieu de l'entrepôt, les lieux et volumes des commandes des clients, ainsi que la capacité du camion ? C'est un problème difficile lorsque le nombre de clients est élevé et on ne sait le résoudre que dans des cas particuliers qui ne correspondent pas toujours aux applications réelles rencontrées : aujourd'hui, même dans les cas les plus simples, calculer une solution approchée à 1% prend un temps déraisonnable. L'un de nos axes de recherche est donc de continuer à développer et d'améliorer ces algorithmes d'approximation.

Modélisation mathématique et évolutions technologiques

Pour les chercheurs qui comme moi travaillent en informatique théorique, l'élément non-négociable est la recherche du théorème. Les problèmes sont variés, les méthodes de résolution aussi, mais à la fin, ce qui est produit, ce qui apparaît dans le paragraphe intitulé "nos résultats" dans nos introductions, est toujours un théorème. La rigueur mathématique prime. *C'est un principe de base.*

Cependant, avec les évolutions technologiques, de nouvelles problématiques surgissent et se traduisent de façon naturelle par des évolutions des questions **mathématiques** étudiées. Ainsi, les modèles de calcul évoluent, avec par exemple de plus en plus de modèles de calcul ou de communication qui limitent l'accès aux données ou dont les données évoluent de façon dynamique, ou encore avec une incertitude sur les données.

Il nous faut donc non seulement travailler sur des problèmes reconnus du domaine, mais également proposer des problèmes algorithmiques nouveaux. Il s'agit par exemple de définir une question algorithmique de façon formelle, qui soit suffisamment simple dans sa structure pour que l'on puisse l'étudier avec un espoir de démontrer des résultats rigoureux. C'est ainsi qu'on peut rigoureusement démontrer l'effet de « petit monde » dans les réseaux sociaux (soit le nombre limité de degrés de séparation entre tous ces membres) : bien que leur taille aille grandissant, il demeure toujours possible à un membre du réseau social de prendre contact avec presque n'importe quel autre membre grâce à une courte chaîne de connaissances. Ici, la démonstration mathématique se fait sur un modèle simplifié, inspiré du modèle dit « de feu de forêt ». C'est en passant par ce type de modèles simplifiés que l'on peut espérer arriver à des intuitions pour une meilleure compréhension du monde.

“Pourquoi” et “Dans quel but” : des questions d'éthique

Par ailleurs, comme les algorithmes agissent de plus en plus non seulement sur de simples octets mais également sur les personnes, les questions d'éthique comme l'équité ou la transparence deviennent prégnantes.

Une constante du domaine est qu'un algorithme est développé dans un but précis et analysé en relation avec la façon dont il atteint son but. Mais ces années dernières ont vu apparaître des algorithmes qui résolvent un problème imprécis, et qui ne “marchent” que dans un sens très flou – le choix d'informations à présenter au lecteur en ligne ou les recommandations de produits à acheter, par exemple. Qu'est-ce qui fait qu'une sélection donnée est le meilleur choix ? Un résultat donné est-il, ou non, satisfaisant ? Quel est le but recherché ? Ces questions sont à la périphérie de l'algorithmique classique et en même temps centrales pour l'évolution du domaine dans la décennie à venir.

Un des nouveaux enjeux de l'algorithmique est donc comprendre le “pourquoi” et le “dans quel but” des algorithmes de la famille des méthodes d' « apprentissage profond », qui se retrouvent appliqués un peu partout avant qu'on ait vraiment compris leurs aspects fondamentaux. Ici, la théorie a pris du retard sur la pratique, et ce retard de compréhension induit des risques de manipulation sans contrôle.

Il ne faut pas avoir peur des algorithmes, mais il faut apprendre à les connaître pour les apprivoiser : c'est la mission de l'algorithmique.

Claire Mathieu



Présentation du cycle d'enseignement du Pr Claire MATHIEU

Leçon inaugurale le jeudi 16 novembre 2017, à 18h00

Cours les mardis à 10h (à partir du 28 novembre)

Panorama de techniques modernes de conception et analyse d'algorithmes

La recherche en conception et analyse d'algorithmes a beaucoup évolué ces dernières années. De nouveaux modèles de calcul sont apparus, car les données, désormais trop massives pour tenir en mémoire en un seul lieu, sont d'accès plus difficile que dans les modèles classiques ; ou elles sont accessibles partiellement, modulo certaines incertitudes (algorithmes stochastiques). Pour les problèmes les plus difficiles, on apprend à se contenter de solutions approchées, ou de solutions qui ne marchent en temps raisonnable qu'en posant des hypothèses supplémentaires. Des méthodes de conception plus sophistiquées se sont également développées : méthodes de type Monte-Carlo, méthodes de type primal-dual de la programmation linéaire, ou hiérarchie de relaxations semi-définies.

À travers l'exemple de quelques problèmes phares, le cours de Claire Mathieu présentera les dernières avancées, la diversité des techniques et les nouvelles directions de recherche en matière de conception et d'analyse d'algorithmes.

Les cours seront suivis à 11h d'un séminaire animé par un spécialiste du domaine évoqué

Programme complet : www.college-de-france.fr/site/claire-mathieu/_course.htm

Colloques le Jeudi 07 juin 2018 de 9h30 à 18h

Programme : www.college-de-france.fr/site/claire-mathieu/symposium-2017-2018.htm

L'ensemble de l'enseignement de Claire Mathieu sera disponible sur www.college-de-France.fr



Biographie



Informaticienne et mathématicienne, Claire Mathieu est, depuis 2012, directrice de recherches au CNRS.

Ancienne élève de l'ENS, née en 1965, elle est titulaire d'une thèse en sciences informatiques de l'université Paris-Sud (sous la direction de Claude Puech) et a été de 1990 à 1997 chercheur CNRS à l'ENS.

Elle a enseigné à l'Université Paris-Sud (1997/2004), à l'École polytechnique (2002/2004), à l'Université de Brown - USA (2004/2012). Elle a été professeur associée au département des sciences informatiques de l'ENS (2014 /2016).

Claire Mathieu est notamment reconnue pour ses travaux sur les algorithmes d'approximation et les algorithmes en ligne. Elle s'est récemment intéressée à la modélisation de réseaux sociaux, à la reconstruction de graphes cachés (circulation des données ...), et aux graphes qui peuvent être dessinés dans le plan. Elle s'intéresse également à l'analyse théorique de problèmes liés à des domaines connexes de l'informatique, voire à d'autres sciences : Modèle de calcul biologique, Physique statistique, Économie ou encore Sociologie.

La chaire *Informatique et sciences numériques*

La chaire annuelle « *Informatique et sciences numériques* » a été créée en 2009 dans le cadre d'un partenariat entre le Collège de France et Inria. Elle a été inaugurée par Gérard Berry, nommé depuis professeur titulaire d'une chaire pérenne d'informatique, *Algorithmes, machines et langages*, et accueille chaque année un nouveau titulaire spécialiste reconnu d'un domaine (langages de programmation, sécurité informatique, Big Data, etc).

Les précédents titulaires de cette chaire ont été :

- Jean-Daniel Boissonnat (2016/2017) : *Géométrie algorithmique*
- Yann LeCun (2015/2016) : *L'apprentissage profond : une révolution en intelligence artificielle*
- Marie-Paule Cani (2014/2015) : *Façonner l'imaginaire, de la création numérique 3D aux mondes virtuels animés*
- Nicholas Ayache (2013/2014) : *Des images médicales au patient numérique*
- Bernard Chazelle (2012/2013) : *L'algorithmique et les sciences*
- Serge Abiteboul (2011/2012) : *Sciences des données : de la logique du premier ordre à la Toile*
- Martin Abadi (2010/2011) : *La sécurité informatique.*
- Gérard Berry (2009/2010) : *Penser, modéliser et maîtriser le calcul informatique.*

À propos du Collège de France

Le Collège de France est un grand établissement public d'enseignement supérieur et de recherche. Institution unique en France et sans équivalent à l'étranger, il répond à une double vocation : être à la fois le lieu de la recherche la plus audacieuse et celui de son enseignement. Un enseignement ouvert à tous et gratuit. La grande majorité des cours et séminaires qui y sont dispensés sont librement accessibles sur internet.

Voué à la recherche fondamentale, le Collège de France possède une caractéristique singulière : ses professeurs partagent avec le public leur travaux de recherche ; une recherche libre et un savoir vivant, dans tous les domaines des lettres, des sciences ou des arts. Les chaires, et par conséquent les disciplines enseignées, y sont sans cesse renouvelées en fonction de l'évolution des connaissances. Le Collège de France accueille également dans ses laboratoires et auprès de ses professeurs, de nombreuses équipes de recherche.

À propos d'Inria

Inria, institut national de recherche dédié au numérique, promeut « *l'excellence scientifique au service du transfert technologique et de la société* ». Inria emploie 2400 collaborateurs issus des meilleures universités mondiales, qui relèvent les défis des sciences informatiques et mathématiques. Son modèle ouvert et agile lui permet d'explorer des voies originales avec ses partenaires industriels et académiques. Inria répond ainsi efficacement aux enjeux pluridisciplinaires et applicatifs de la transition numérique. Inria a contribué à façonner le monde numérique que nous connaissons aujourd'hui et est à l'origine de nombreuses innovations créatrices de valeur et d'emplois et s'attache à relever les défis des sciences du numérique du XXI^e siècle.

Contact presse Inria : Laurence Goussu / laurence.goussu@inria.fr - tel : 01 39 63 57 29