COMMUNIQUE DE PRESSE – 8 octobre 2018

Algorithmes : à la recherche de l'universalité perdue

Rachid GUERRAOUI, Professeur à l'Ecole polytechnique fédérale de Lausanne (EPFL),

occupera la chaire *Informatique et sciences numériques* (année 2018-2019)

Une chaire du Collège de France créée en partenariat avec Inria

- Leçon inaugurale le jeudi 25 octobre 2018 à 18h00 -

Rachid GUERRAOUI, a été nommé professeur invité au Collège de France pour l'année académique 2018-2019. Il occupera la chaire annuelle *Informatique et sciences numériques* où il donnera une série de cours sur l'algorithmique répartie. Sa leçon inaugurale se tiendra le jeudi 25 octobre à 18h00. Les cours qu'il y donnera porteront sur des questions telles que : ***l'atomicité dans un système réparti, les systèmes distribués dynamiques, le pouvoir de la mémoire partagée, l'impossibilité du consensus, la « blockchain »****, etc*. Ouverts à tous sans condition d'inscription préalable comme l'ensemble des enseignements du Collège de France, ils débuteront le vendredi 26 octobre (voir p 4) et seront rendus accessibles au plus grand nombre sur le portail www.college-de-france.fr.

Jusqu'à une date récente, les fondamentaux de l'informatique supposaient pour la plupart qu'un algorithme s'exécute sur un seul ordinateur en tant que séquence d'opérations élémentaires. Mais, avec l'accroissement constant de la puissance de calcul des ordinateurs et du débit de transmission de données par les réseaux, les applications informatiques modernes deviennent de plus en plus souvent *réparties sur plusieurs ordinateurs* calculant en même temps et communiquant constamment. Cela demande une évolution profonde de la théorie et de la pratique des algorithmes pour les rendre *parallèles* et géographiquement *distribués*.

Dans le cas de systèmes à échelle planétaire, les algorithmes distribués permettent de se passer de serveur central, ce qui simplifie beaucoup leur réalisation. Par exemple, dans les applications dites *pair à pair* (*peer-to-peer* en anglais) qui permettent de diffuser à grande échelle musiques et vidéos, chaque ordinateur est rendu simultanément client pour voir une vidéo et serveur pour la diffuser aux autres. La vidéo se trouve alors éclatée en fragments répartis et potentiellement répliqués dans les ordinateurs du réseau. Les algorithmes sont chargés d'en récupérer les bouts, de les mettre dans l'ordre, et de les rediffuser ailleurs. Dans les chaînes de blocs (*blockchain* en anglais) utilisés pour les monnaies virtuelles comme le *bitcoin*, les algorithmes s'exécutent sur un grand nombre d'ordinateurs géographiquement distants, qui varie constamment. Ils permettent de réaliser des transactions entre utilisateurs sans passer par un serveur central.

A l'autre bout de l'échelle, considérons le cas d'un algorithme qui s'exécute sur un seul circuit électronique, d'une taille de l'ordre du centimètre carré. Un tel circuit comporte désormais plusieurs processeurs, aussi nommés *cœurs* de calcul. Etant donné que la vitesse de chaque cœur ne peut quasiment plus s'accroître du fait de limitations physiques, le seul moyen d'accélérer l'algorithme est de *paralléliser* son calcul en en répartissant sur plusieurs cœurs des morceaux pouvant être calculés séparément.

Entre ces deux cas extrêmes, il existe aussi de nombreuses situations où les calculs sont répartis pour des raisons de tolérance aux défaillances. Par exemple, pour éviter les interruptions de service en cas de panne, on ne fait plus reposer toute une application informatique sur une seule machine ; on augmente sa disponibilité en dupliquant sur plusieurs machines les parties cruciales de ses calculs.

Pour Rachid Guerraoui, *« l'algorithmique répartie est la discipline scientifique qui étudie les réseaux d'ordinateurs. Les théories classiques de calculabilité et de complexité algorithmiques, inventées pour les calculs séquentiels et centralisés, ne peuvent pas s'appliquer au cas de l'algorithmique répartie. La raison en est très simple. Alors que dans les théories classiques, on suppose que la machine de Turing utilisée fonctionne correctement, émettre une telle hypothèse pour un ensemble de machines est déraisonnable. On souhaite qu'un algorithme qui s'exécute sur un réseau de machines se termine correctement malgré la défaillance d'un sous-ensemble de ces machines. Cette distribution est l'une des motivations les plus fréquentes de l'algorithmique répartie ».*

Retrouver l'universalité de Turing en algorithmique répartie

Par Rachid GUERRAOUI

Les algorithmes existent depuis que l'humain essaye de calculer. Dans les temps anciens, le calculateur qui exécutait l'algorithme était l'humain lui-même. Au moyen âge, des machines exécutant des algorithmes à notre place firent leur apparition. Les premières étaient très spécialisées et chacune ne pouvait exécuter qu'une petite classe d'algorithmes. Mais en 1936, Alan Turing proposa l'architecture d'une machine qui pouvait calculer tout ce qui est calculable. Sa machine *universelle,* exécutant tous les algorithmes possibles et imaginables, donna naissance à un outil fantastique que l'on a appelé depuis l'*ordinateur*, et à une science consacrée à sa compréhension, l'*informatique*. Sans cette universalité, il nous serait impossible aujourd'hui d'étudier et d'enseigner cette science, en faisant abstraction des détails technologiques sous-jacents qui évoluent à très grande vitesse depuis près d'un siècle.

A partir des années 60, l'invention des réseaux a permis d'aller encore plus loin avec l'*informatique répartie*. Il a été possible de connecter des ordinateurs dans de grands réseaux comme *Internet*, et des processeurs dans des petits réseaux à l'intérieur de chacun des ordinateurs, devenu désormais un calculateur *parallèle*. L'objectif global était de créer l'abstraction d'une *s*uper machine répartie, indestructible et ultra-rapide. Mais la recherche de ces super-pouvoirs a eu comme conséquence majeure la perte de l'universalité. Certes, quelques algorithmes peuvent être exécutés de manière robuste et efficace sur la machine répartie, mais pas tous, et on ne peut plus faire abstraction des technologies sous-jacentes à ces réseaux.

L'algorithmique répartie est la discipline scientifique qui identifie les conditions nécessaires et suffisantes sur les réseaux, grands ou petits, permettant de retrouver l'universalité de Turing. Lorsque ces conditions ne sont pas satisfaites, il s'agit de définir les formes d'universalités restreintes qu'il est possible de réaliser. Sans cela, il est impossible d'appréhender ce que font les algorithmes exécutés sur une *Blockchain*, un *Cloud*, un *Data Center* ou sur l'*Internet des Objets,* ni de définir ce que permettent de calculer exactement des architectures *Multi-Processeurs.* Autrement dit, il est impossible de comprendre l'informatique moderne, fondamentalement répartie.

En informatique répartie, les algorithmes sont constitués, outre les instructions élémentaires des algorithmes classiques centralisés, d'instructions permettant de faire communiquer plusieurs machines, comme des envois de messages ou des accès à des zones mémoires partagées. Ces instructions « réseau » ont un impact fondamental sur la nature des algorithmes. Leur complexité s'en trouve profondément affectée et de nouvelles métriques sont nécessaires pour mesurer leur efficacité, basées par exemple sur le nombre de messages envoyés en fonction du nombre d'ordinateurs connectés. Ces métriques permettent d'étudier le compromis entre les super-pouvoirs de la machine répartie obtenue : *robustesse* et *efficacité*. Les algorithmes d'apprentissage sous-jacents à l'intelligence artificielle moderne illustrent ce compromis. Du fait de la grande quantité de données disponibles, l'apprentissage est réparti sur plusieurs machines. Une moyenne des analyses obtenues est alors effectuée. Une seule machine donnant des valeurs extravagantes rend le tout néanmoins inutilisable.

Au-delà des considérations technologiques qui motivent l'étude de l'algorithmique répartie pour apprivoiser les inventions *humaines* que sont les ordinateurs et les réseaux, cette étude est tout aussi fondamentale à la compréhension de phénomènes *naturels.* La modélisation de la synchronisation du clignotement des lucioles, du mouvement coordonné d'un banc de poissons, des formes géométriques dessinées par une nuée d'oiseaux, ou du comportement collaboratif d'un réseau de neurones, passe par des algorithmes répartis.

Présentation de l'enseignement qui sera dispensé par le Pr Rachid GUERRAOUI

Leçon inaugurale le jeudi 25 octobre 2018 à 18h00

Cours les vendredis à 10h tous les quinze jours (ouverture le 26 octobre)

Algorithmique répartie

L'objectif de cette série de cours est de présenter les fondements de l'algorithmique répartie. Les cours présenteront de manière incrémentale les résultats les plus importants obtenus dans ce domaine depuis près d'un demi-siècle et souligneront les nombreux problèmes encore ouverts. Aucune connaissance préalable en informatique répartie n'est requise. Les cours couvriront en particulier les notions d'exécution d'un algorithme réparti et revisiteront les notions de calculabilité, d'universalité et de complexité dans le contexte réparti. Ils aborderont les théorèmes d'universalité et d'impossibilité du consensus, tout comme les hypothèses permettant de contourner cette impossibilité ainsi que les algorithmes associés. Ils feront le lien avec les architectures utilisatrices d'algorithmes répartis comme la *Blockchain*, le *Cloud,* le *Data Center* et les multi-processeurs.

Les cours seront suivis à 11h00 d'un séminaire animé par un spécialiste du domaine évoqué

Lien vers le programme complet : https://www.college-de-france.fr/site/rachid-guerraoui/_course.htm

Colloque le vendredi 12 avril 2018 (9h-18h00)

Lien vers le programme : https://www.college-de-france.fr/site/rachid-guerraoui/symposium-2018-2019.htm

L'ensemble de l'enseignement de Rachid GUERRAOUI sera rendu accessible sur www.college-de-France.fr

Biographie

Rachid GUERRAOUI est professeur à l'Ecole Polytechnique Fédérale de Lausanne (EPFL) où il dirige le Laboratoire de calcul distribué (DCL). Titulaire d'un doctorat en informatique de l'Université d'Orsay (aujourd'hui Université Paris-Saclay), il a été affilié au MIT à Boston et aux laboratoires *Hewlett Packard* en Californie (Palo Alto). Il a été professeur invité à Paris-VI, Paris-VII et à l'Université de Rennes.

Rachid Guerraoui a été élevé en 2012 au rang de « fellow » au sein de l'*Association for Computing Machinery* (ACM), avant de recevoir en 2013 le prestigieux *Google Focus Award* pour son projet sur le *Web Alter ego*. Ce prix est attribué chaque année pour soutenir les travaux les plus ambitieux en informatique dans les domaines jugés prometteurs. Il a obtenu une bourse « senior » auprès de l'ERC, le Conseil européen de la recherche, dans le cadre du projet AOC, explorant le pouvoir inhérent et les limites d'un nouveau paradigme, l'informatique orientée vers l'adversaire (*Adversary-Oriented Computing*). Il a rédigé de nombreux articles et ouvrages sur l'informatique répartie et est éditeur associé de la revue *Journal of the ACM*. La vingtaine d'étudiants qui ont obtenu leur doctorat sous sa direction sont aujourd'hui professeurs ou chercheurs dans de grands laboratoires industriels. Rachid Guerraoui a aussi participé aux lancements des plateformes d'enseignement en ligne Wandida et Zetabytes, bibliothèques de contenus pour les étudiants et les spécialistes rassemblant plusieurs centaines de 500 vidéos sur la résolution de problèmes en informatique, en mathématiques et en physique (près de 2 millions de vues et un taux de rétention de 50%).

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 :

- 2017-2018 : Claire Mathieu - *Algorithmes*
- 2016-2017 : Jean-Daniel Boissonnat - *Géométrie algorithmique : données, modèles, programmes*
- 2015-2016 : Yann LeCun - *L'apprentissage profond : une révolution en intelligence artificielle*
- 2014-2015 : Marie-Paule Cani - *Façonner l'imaginaire, de la création numérique 3D aux mondes virtuels animés*
- 2013-2014 : Nicholas Ayache - *Des images médicales au patient numérique*
- 2012-2013 : Bernard Chazelle - *L'algorithmique et les sciences*
- 2011-2012 : Serge Abiteboul - *Sciences des données : de la logique du premier ordre à la Toile*
- 2010-2011 : Martin Abadi - *La sécurité informatique.*
- 2009-2010 : Gérard Berry - *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, l'institut national de recherche dédié aux sciences du numérique, promeut l'excellence scientifique et le transfert pour avoir le plus grand impact.
Il emploie 2400 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3000 scientifiques pour relever les défis des sciences informatiques et mathématiques, souvent à l'interface d'autres disciplines.
Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 160 start-up.
L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

Contact presse Inria : Mme Laurence Goussu – mél : laurence.goussu@inria.fr - tél. : 01 39 63 57 29