



Table ronde sur la recherche en informatique

Serge Abiteboul¹

À l'occasion du dernier congrès de la Société informatique de France, on m'a demandé d'animer une table ronde dédiée aux principales avancées scientifiques en informatique au cours des dix dernières années, ainsi qu'aux défis des dix prochaines années. Le domaine progresse très vite et il est important d'interroger cette progression. J'ai réuni pour parler du sujet une véritable *dream team*. Je leur ai laissé une liberté complète. Dans un premier temps, ils regardaient le passé, et dans un second, le futur. Précision importante, je ne leur ai pas demandé de réaliser un podium. Il ne s'agit pas de la cérémonie des César, ou de quelque classement que ce soit. Ce ne sont que cinq points de vue, mais ils sont intéressants parce qu'ils viennent de cinq *leaders* de leurs domaines.

Les dix dernières années

Véronique Cortier, directrice de recherche CNRS, Loria

J'ai noté deux changements importants dans le domaine de la sécurité informatique, mon domaine actuel de recherche. Le premier est technique. Il y a 20 ans naissaient les premiers outils basés sur les méthodes formelles (logique du premier ordre, réécriture, etc.) pour analyser les protocoles cryptographiques, qui régissent les échanges sécurisés. Nous étions fiers de les appliquer sur de petits exemples-jouets, sortis de leur contexte. Ces outils ont acquis une grande maturité ces dix

1. Directeur de recherche Inria, Paris.

dernières années et sont désormais utilisés lors de la conception des nouvelles générations de protocoles comme TLS ou Signal. Un deuxième point est plus sociétal. La sécurité informatique peut tout autant protéger les citoyens et les institutions contre des attaques, protéger la vie privée, ou au contraire permettre une certaine forme de surveillance (pour la santé ou la sécurité des personnes par exemple). Le cas très médiatisé des applications de traçage en est un exemple. Définir un compromis acceptable entre surveillance et respect de la vie privée va au-delà du rôle de l'informaticien et implique de nombreuses considérations éthiques.

Serge Fdida, professeur, Sorbonne Université

L'histoire des technologies télécom est instructive à bien des égards, et questionne le bon niveau d'abstraction des objets à manipuler dans un réseau. En 1909, avec l'apparition de la téléphonie, le niveau d'abstraction est celui du circuit sur lequel vont s'écouler nos communications téléphoniques. En 1970, nous observons l'émergence de l'Internet suivant une architecture en rupture qui s'appuie sur le fameux paquet IP. Puis, on réalise qu'un certain nombre d'objets partagent des éléments communs et nous parlerons de flots, un ensemble de paquets IP liés par des adresses identiques, par exemple un flot TCP. Mais nous devons encore nous rapprocher de l'application qui fait sens en termes de service offert au client, nous parlerons alors de contenu, ce qui provoque le déploiement accéléré des CDN (réseaux de distribution de contenu). Enfin, pour généraliser les objets numériques manipulés, le concept de programme représente le meilleur candidat permettant de traiter une majorité des fonctions réseaux. Nous sommes donc passés en quelques années d'une architecture distribuée fondée sur internet, mais reposant sur des équipements propriétaires vers une architecture plus centralisée, reposant sur le *cloud*, et dont les fonctions pourront être déployées de façon continue et intégrées dans l'architecture grâce à la flexibilité du logiciel, en s'appuyant sur du matériel de commodité.

Nicholas Ayache, directeur de recherche Inria, Sophia Antipolis

En imagerie médicale et médecine numérique, on a observé ces 10 dernières années une percée spectaculaire des algorithmes d'apprentissage profond, notamment des réseaux de neurones à convolution, des réseaux génératifs antagonistes (GAN), des auto-encodeurs variationnels, etc. Ces algorithmes ont permis des avancées majeures en reconstruction d'images (réduction du bruit, meilleure résolution, accélération des calculs, etc.) et en analyse d'images que ce soit pour la détection, localisation et caractérisation d'anomalies (cancers, etc.) ou pour le recalage (temporel, entre sujets) et la fusion d'images (entre modalités) ainsi qu'en synthèse d'images (prédire une modalité à partir d'une autre). Enfin, ces algorithmes permettent une intégration efficace d'informations très hétérogènes sur le patient en combinant par exemple des images anatomiques et fonctionnelles avec des informations cliniques, biologiques, génétiques, voire comportementales et environnementales (exposome). Beaucoup de

ces avancées se traduisent progressivement par de nouveaux logiciels exploitables en clinique après leur validation par les organismes compétents.

Sonia Ben Mokhtar, directrice de recherche CNRS, LIRIS

Durant les dix dernières années, l'ubiquité de l'informatique est devenue une réalité. Nous sommes en effet submergés de dispositifs informatiques communicants (téléphones, tablettes, montres connectées et de plus en plus d'ordinateurs sont embarqués dans nos objets du quotidien). Ceci est dû à des avancées matérielles majeures : développement de processeurs de plus en plus petits et puissants et de moins en moins chers, connectés à travers des réseaux de plus en plus rapides. En plus de permettre d'aller vers de l'extrêmement petit (informatique embarquée, Internet des objets), ces évolutions matérielles ont permis de faire émerger des infrastructures de stockage et de calcul de plus en plus grandes (centres de calculs, informatique dans les nuages) menant ainsi à des évolutions logicielles majeures (logiciels qui s'appuient sur ces infrastructures tels que les réseaux sociaux, les services géo-localisés, les systèmes de recommandation). Ces logiciels sont également devenus ubiquitaires et s'appuient massivement sur la collecte et l'analyse de données personnelles.

Claire Mathieu, directrice de recherche CNRS, IRIF

Dans le domaine de la conception d'algorithmes polynomiaux qui résolvent des problèmes d'optimisation avec un facteur d'approximation, le problème du voyageur de commerce est un problème phare. La meilleure heuristique connue auparavant, celle de Christofides, avait un facteur d'approximation de 1.5. L'an dernier a été publié la première légère amélioration depuis 50 ans, dans un article de Nathan Klein, Anna Karlin et Shayan Oveis Gharan. Au lieu de commencer par construire un arbre couvrant minimum, ils construisent un arbre aléatoire tiré d'une certaine distribution définie à partir de la relaxation linéaire de Held-Karp pour le voyageur de commerce. Cela permet d'éviter les situations de pire cas pour l'heuristique. Ce résultat montre



une sophistication croissante du domaine qui lui permet, en combinant probabilités, programmation mathématique et outils combinatoires, de progresser enfin dans la compréhension de problèmes qui forment le socle de notre domaine.

Les dix prochaines années

Véronique Cortier

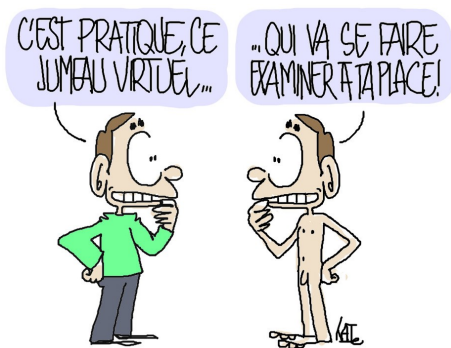
Je pense qu'un enjeu non scientifique nous attend pour les dix prochaines années, à savoir le système de sélection des publications. Le nombre de soumissions explose et avec lui, le nombre de *reviews*, de discussions, la taille des comités de programme. Les conférences jouent un rôle important, à savoir présélectionner les articles à lire dans chaque domaine. Elles sont également un outil central d'évaluation, du recrutement du jeune chercheur à sa promotion tout au long de sa carrière. Cet aspect est plus pervers puisqu'il pousse à une course à la publication et à présenter ses résultats sous le meilleur angle possible. Ces dernières années ont déjà vu des changements importants avec le passage de plus en plus fréquent à l'anonymisation des soumissions (*blind* ou *double blind*), la multiplication des cycles de soumissions (possibilité de soumettre trois fois dans l'année pour une même conférence). La crise Covid ainsi que, de façon plus profonde, la crise écologique obligent à revoir le mode de fonctionnement des conférences : est-il soutenable de se rendre à cinq ou six conférences chaque année, dans le monde entier ? Les 10 prochaines années nous amèneront certainement à revoir notre façon de publier.

Serge Fdida

L'évolution du logiciel est transformante, tant sur la méthodologie, que les compétences, la conception ou les opérations d'une infrastructure numérique qui intègre désormais des objets variés du *cloud* à l'*Internet of Things* (IoT). Le réseau devient une plate-forme virtualisée programmable qui exploite des concepts et des outils utilisés dans ce cadre. L'impact est énorme dans la mesure où l'objectif est de déconstruire (on parle de désagrégation) un routeur, un réseau d'accès radio (RAN) ou l'ensemble d'un réseau pour ensuite le reconstruire en utilisant si possible des briques logicielles open-source déployées à façon. Ce principe s'inspire de celui des PCs s'appuyant sur une base matérielle (*chipset* et *hardware*) et un OS sur lequel s'exécutent des programmes indépendants, ou encore ceux de l'Appstore/Android Market tournant sur un OS ad hoc et un environnement matériel. L'impact industriel qui s'ensuit sera certainement important. Cette évolution illustre l'impact croissant du logiciel dans la conception des systèmes en raison de leur évolution rapide et de leur flexibilité. Par ailleurs, l'ouverture de ces systèmes devient un terreau pour l'innovation. Cette transformation illustrée ici dans le monde des télécoms est observable dans de nombreux autres domaines.

Nicholas Ayache

Les 10 prochaines années devraient voir le développement généralisé de la médecine numérique sur plusieurs fronts. D'abord, pour soulager les médecins spécialistes, dans le cadre des dépistages à grande échelle, une première lecture automatique des images et données du patient pour réduire la masse des dossiers à analyser plus finement ; par ailleurs, des outils d'aide au diagnostic exploitant des appareils portés par les patients eux-mêmes ou opérés par des non-spécialistes (techniciens, généralistes) pour guider l'orientation vers des spécialistes. Ensuite, le développement généralisé du jumeau numérique du patient pour guider le diagnostic, le pronostic et la thérapie, notamment grâce à des outils de simulation numérique. Ceci nécessitera de compléter les modèles statistiques d'apprentissage profond (entraînés sur des bases de données mondiales et multicentriques, centralisées ou distribuées) par des modèles biophysiques de l'anatomie et de la physiologie du patient pour mieux interpréter, prédire et comprendre les raisons d'un dysfonctionnement et les effets attendus d'une thérapie.



Sonia Ben Mokhtar

L'extrême centralisation des services en ligne d'aujourd'hui ainsi que les nombreuses fuites de données qu'ils subissent de manière régulière forcent le questionnement sur la préservation de la vie privée de leurs usagers. Aussi bien le législateur (commission européenne avec des textes de loi tels que la RGPD) que le monde technologique (chercheurs et entreprises) travaillent sur cette problématique. Une piste prometteuse pour les années à venir consiste en la décentralisation (d'une partie) du Web. Plusieurs initiatives ont débuté depuis l'apparition des cryptomonnaies et plus généralement de la Blockchain (par exemple la finance décentralisée), mais d'autres travaux qui ne seront pas forcément liés à la Blockchain vont probablement voir le jour. Je pense en particulier à des travaux autour de la décentralisation des algorithmes d'apprentissage qui sont au cœur de nombreux services en ligne. Décentraliser les algorithmes d'apprentissage soulève néanmoins de nombreux défis et nécessitera certainement la collaboration entre chercheurs de différentes communautés (apprentissage, optimisation, réseaux, algorithmique et systèmes distribués, sécurité et vie privée).

Claire Mathieu

De plus en plus, les rouages de la société tournent grâce au numérique et aux algorithmes. Parcoursup a généralisé APB à presque toutes les formations de l'enseignement supérieur en France. Waze, Google maps et les applications de calcul de plus court chemin, avec évolution dynamique au fil des embouteillages et des accidents, ont révolutionné la façon dont les usagers appréhendent les itinéraires. L'algorithmique peut-elle aider au découpage de circonscriptions en France ? À une rationalisation des micro-règles du calcul de l'impôt sur le revenu ? Comment incorporer les aspects algorithmiques dans les discussions entre les divers partenaires du débat public ? Ces questions guideront une partie de la recherche en algorithmique pour des problèmes « sociétaux ».