Démonstrations mathématiques assistées par ordinateur

Assia Mahboubi

Chercheuse
Inria, Université Paris–Saclay

Les mathématiciens sont souvent représentés comme des espèces de sorciers, qui couvrent des feuillets ou des tableaux noirs de formules cabalistiques, illustrées par de mystérieuses figures. Et cette langue semble bien impénétrable à qui n'est pas mathématicien.

En fait, on apprend à l'école les rudiments de la langue des mathématiques : écrire des nombres, poser des équations simples qui permettent de résoudre des problèmes passionnants de trains qui se croisent et de baignoires qui se vident... Sans le savoir, les écoliers utilisent en fait déjà une langue mathématique très moderne, qui aurait sans doute semblé bien incompréhensible à un mathématicien de génie comme Euclide.

D'Euclide à Bourbaki : une longue maturation de la langue

Euclide est l'auteur des *Éléments*, texte fondateur des mathématiques, dont l'influence sur la pensée scientifique occidentale est immense. Vivant en Grèce probablement vers 300 avant notre ère, il représente les nombres en utilisant un système additif, c'est-à-dire sur le même principe que le système des chiffres romains, et sans zéro. En utilisant les chiffres romains, la multiplication 9×17 s'écrit IX fois XVII : rapidement, il devient difficile d'effectuer une telle opération sans l'aide d'un instrument mécanique de calcul !

Par ailleurs, la mise en équation d'un problème, et en particulier l'usage d'un symbole pour représenter et raisonner sur une quantité inconnue, est probablement tout à fait étrangère à Euclide. En effet, cette idée révolutionnaire est attribuée à Diophante d'Alexandrie, un mathématicien de langue grecque, qui a probablement vécu vers le II^{e} siècle de notre ère.

Portrait d'Euclide par Juste de Gand peint vers 1474.

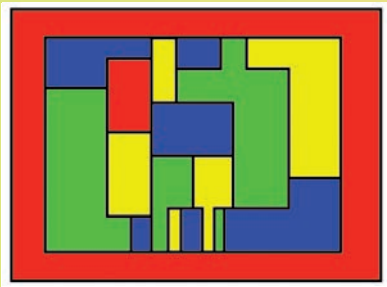
La langue des mathématiques, ses symboles, ses codes, évoluent au cours des âges et varient selon les lieux. Ils s'adaptent également à la branche spécialisée qu'elle doit servir, par l'usage de notations spécifiques adaptées. Ces notations sont là pour condenser le propos et le rendre intelligible, en le débarrassant des informations techniques sans contenu pertinent. Par exemple, l'expression $u + v$ aura un sens, que u et v soient des entiers, des nombres complexes, des vecteurs ou même des fonctions à valeurs réelles. Le lecteur reconnaît de lui-même quelle est l'opération sous-jacente, pourvu qu'il ait assez de culture mathématique et un contexte clair. Pour les chercheurs, qui inventent des concepts mathématiques nouveaux, trouver les « bonnes » notations, efficaces, lisibles et non ambiguës, peut être un sujet de réflexion en soi, qui va bien au-delà des considérations de goût. Le groupe de mathématiciens français connu sous le nom de Nicolas Bourbaki a ainsi marqué les mathématiques du XX^{e} siècle en proposant une série de traités couvrant les sujets enseignés dans les cursus universitaires. Ces traités rationalisent les présentations très disparates de l'époque et s'accompagnent d'une réflexion historique et méthodologique. La clarté de la terminologie et des notations qu'ils introduisent va de pair avec la recherche systématique des structures partagées par des objets mathématiques apparemment très différents.

Aujourd'hui, les étudiants du monde entier utilisent le même vocabulaire, les mêmes notations, les même symboles, à d'infimes variations près. Les textes d'enseignement et de recherche sont typographiés directement par leurs auteurs, en utilisant des logiciels de traitement de texte scientifique très puissants comme Latex ou des bibliothèques d'affichage de mathématiques pour les pages Web comme MathJax. Ces outils permettent de mettre en forme facilement des textes qui utilisent des symboles très variés et des formules arbitrairement sophistiquées. Il est de fait devenu très aisé d'écrire et d'échanger des textes mathématiques sous forme électronique.

Les processus classiques de validation trouvent leurs limites…

Néanmoins, un partie importante de la recherche en mathématique se transmet toujours par voie orale. Les chercheurs communiquent avec leurs semblables au moyen d'exposés, en utilisant le plus souvent le support d'un tableau pour compléter leur discours avec des formules, écrites au fil de l'exposé. Ils exposent ainsi leurs résultats et leurs preuves à la critique de l'audience, qui peut éventuellement détecter des failles. Résultats et démonstrations sont aussi le plus souvent consignés par écrit, dans des textes soumis à l'approbation de relecteurs anonymes, qui

décident de l'opportunité de leur publication. Ce processus d'évaluation, aussi bien oral qu'écrit, transforme les conjectures en théorèmes par la validation de la preuve proposée par l'auteur. Comme tout processus humain et social, cette évaluation est faillible. De fameuses conjectures ont connu une histoire à rebondissements et fait l'objet de « preuves » fausses, dont certaines furent publiées dans les journaux les plus sérieux ! Ce fut le cas du *théorème des quatre couleurs*, un résultat de théorie des graphes finalement démontré par Kenneth Appel et Wolfgang Haken en 1976, ou de la *conjecture jacobienne*, un problème de géométrie algébrique formulé en 1939 et toujours ouvert à ce jour.



Il est possible de colorier toute carte découpée en régions connexes en n'utilisant que quatre couleurs différentes de sorte que deux régions adjacentes reçoivent des couleurs distinctes.

© Hrvoje et joriki, 2015

Georges Gonthier (ci-dessus) et Benjamin Werner ont apporté en 2005 une démonstration formelle du théorème des quatre couleurs à l'aide de l'assistant de preuve Coq.

© É. Thomas, 2014

En théorie, il devrait être possible d'exposer les démonstrations candidates dans leurs moindres détails. En théorie toujours, il devrait être possible de vérifier ces démonstrations de façon parfaitement scrupuleuse. Mais ces suppositions ne sont pas réalistes. Comme déjà remarqué par Nicolas Bourbaki, une entreprise d'explicitation minutieuse des textes mathématiques n'apporterait rien à leur compréhension ; au contraire, elle ensevelirait très probablement les aspects essentiels sous un fatras de détails sans importance. Aucun lecteur humain ne pourrait rester suffisamment attentif pour suivre le fil d'un discours si verbeux et si long. Aucun lecteur humain… certes ; par contre, on peut envisager de s'aider d'un instrument mécanique (moderne) de calcul, à savoir un ordinateur.

Les *assistants de preuve* sont des logiciels qui permettent à leurs utilisateurs de décrire des énoncés mathématiques et leurs preuves, puis

de les vérifier. Énoncés et preuves sont décrits en utilisant un langage logique simple et parfaitement codifié, de sorte que la *correction* des preuves (au sens de leur caractère correct) peut être vérifiée de façon automatique.

À côté de ce vérificateur, et pour mériter son nom, un assistant de preuve fournit également un ensemble d'outils pour aider son utilisateur à composer ses définitions, ses énoncés et ses démonstrations candidates. Ici, c'est la vérification des preuves qui est automatisée : leur écriture, elle, reste *a priori* manuelle. Pour mener à bien cette tâche de réécriture, il est possible – et nécessaire – de modéliser, avec l'assistant de preuve, les usages de la langue mathématique. Il s'agit en effet de permettre à la fois à l'utilisateur d'utiliser ses notations usuelles et à la machine d'inférer le contenu implicite des formules.

Quelques applications des méthodes formelles

Le *bug* du Pentium IV, identifié en 1994 par le mathématicien américain Thomas Nicely, est resté célèbre pour le coup qu'il a porté à l'image de l'entreprise Intel, qui commercialisait ces processeurs. Il s'agissait d'une erreur (mathématique) dans l'algorithme de division des nombres flottants implanté dans ce processeur. Depuis, l'entreprise américaine emploie des chercheurs experts en preuves formelles pour travailler à la validation des algorithmes mis en œuvre sur les processeurs.

Le *bug* qui a provoqué l'explosion en vol de la fusée Ariane 5 reste probablement le plus cher de l'histoire de l'informatique. Il était dû *in fine* à une subtile erreur de choix dans la représentation en mémoire de la valeur de l'accélération de la fusée. Depuis, les méthodes formelles ont gagné leurs lettres de noblesse, typiquement dans le domaine de l'avionique (comme chez Airbus) ou dans celui des véhicules sans conducteurs (la ligne 14 du métro parisien utilise du code formellement vérifié).

Des outils enfin disponibles pour certains contextes industriels

Les premiers assistants de preuve ont vu le jour dans les années 1960 et n'ont pas cessé d'évoluer depuis. Ils ont d'abord été utilisés pour étudier les propriétés de programmes informatiques, ou pour la vérification de résultats d'informatique théorique. En effet, il est très difficile d'être certain qu'un programme informatique réalise bien ce que l'on attend de lui ; les sources d'erreurs potentielles sont multiples : erreur mathématique dans la conception de l'algorithme, *bug* dans la transcription de cet algorithme dans un langage de programmation, subtilités dans les caractéristiques techniques du matériel électronique dont est constitué l'ordinateur…

Le moyen le plus sûr de garantir le comportement attendu d'un programme est d'écrire un énoncé mathématique qui exprime sa correction – et surtout de le démontrer ! Hélas, de telles démonstrations sont très difficiles à obtenir et à vérifier, car elles sont terriblement longues et fastidieuses à lire pour un être humain. On est bien loin ici du corpus mathématique des traités de Nicolas Bourbaki : de telles preuves enchaînent typiquement des analyses par cas prolifiques et des combinaisons gigantesques de trivialités arithmétiques.

Néanmoins, avec l'aide d'outils informatiques, il est possible de remplacer avantageusement les tests par des preuves vérifiées par ordinateur. Les parties trop fastidieuses des preuves peuvent être générées automatiquement, ne laissant à l'utilisateur que des parties de la preuve significatives d'un point de vue mathématique.

La vérification des preuves peut certes être totalement automatisée, mais le coût en temps et en expertise d'une telle validation formelle reste élevé. Heureusement, il existe maintenant des techniques et des outils suffisamment matures pour avoir été adoptés dans des contextes industriels, pour des applications où la fiabilité des programmes est critique.

Confidentialité des données : les méthodes formelles à la rescousse

L'actualité récente a jeté à plusieurs occasions une lumière crue sur l'importance prise par les protocoles cryptographiques dans la vie quotidienne et la préservation des données sensibles. En 2014, on découvre qu'une faille dans la bibliothèque de cryptographie OpenSSL rend vulnérables d'innombrables programmes, logiciels et sites Web qui en dépendent. L'impact de ce *bug*, surnommé Heartbleed, est considérable. Il a exposé pendant de longues semaines à qui en avait connaissance de nombreuses informations confidentielles, sans que leur consultation indue ne laisse de traces. Ainsi, le site de l'Agence du revenu du Canada (équivalent canadien du Trésor public) a été ainsi fracturé, permettant la collecte frauduleuse des données personnelles relatives à plusieurs centaines de contribuables.

La recherche sur la sécurisation de ces protocoles grâce aux méthodes formelles est actuellement très active, en contact rapproché avec des acteurs gouvernementaux comme le NIST (Institut national des normes et de la technologie, États-Unis).

Aujourd'hui, les assistants de preuve sont suffisamment évolués pour que l'on puisse écrire des bibliothèques de mathématiques digitales représentant n'importe quel résultat (dont la preuve est évidemment supposée connue). Ils ont de fait été utilisés pour vérifier des résultats mathématiques récents et hautement sophistiqués comme la *conjecture*

En 1998, Thomas Callister Hales (ci-dessus), de l'université de Pittsburgh aux États-Unis, prouve la conjecture de Kepler, vieille de quatre siècles, sur la densité maximale d'un empilement de sphères. Il a ensuite consacré douze années à la vérifier, aidé d'une équipe internationale… et d'un assistant de preuve.

de Kepler ou le redoutable *théorème de Feit–Thompson* en théorie des groupes (par Georges Gonthier et son équipe). Incidemment, la réalisation de ces bibliothèques permet en retour d'étudier la correction de programmes reposant sur des mathématiques avancées, comme les primitives cryptographiques ou les algorithmes de l'analyse numérique.

Améliorer les assistants de preuves pour les rendre utilisables par des chercheurs ou étudiants non spécialistes pose néanmoins des problèmes, qui sont autant de sujets de recherche actifs. Cependant, ils seront peut-être demain les outils informatiques privilégiés pour l'enseignement et la recherche en mathématiques…

A. M.

Pour en savoir (un peu) plus :

Notices Of The American Mathematical Society 55 (11). Numéro spécial *«A Special Issue On Formal Proof»,* 2008, disponible en ligne (en anglais) : ***http://www.ams.org/notices/200811/***

Les leçons d'un algorithme délinquant. Jean-Michel Muller, *Interstices*, 2004, disponible en ligne : ***https://interstices.info/jcms/c_5737/les-lecons-d-un-algorithme-delinquant***

Comment faire confiance à un compilateur ? Xavier Leroy, *Interstices*, 2010, disponible en ligne : ***https://interstices.info/jcms/n_52365/comment-faire-confiance-a-un-compilateur?mediego_ruuid=9e796110-19fb-11e7-bdd1-51b6d0d02713***