

Première démonstration « open source » d'un théorème mathématique

Est-ce uniquement par leur licence qu'un noyau Linux et une encyclopédie Wikipédia sont identifiés comme étant libres ?



Juridiquement parlant oui, mais s'en tenir là serait passer à côté du modèle collaboratif particulier que ce sont donnés ces deux fleurons de la culture libre pour développer leur projet.

Conséquence de la licence, c'est aussi voire surtout la puissance de ce modèle qui caractérise le Libre. Et ce modèle commence à se diffuser un peu partout dans la société...

Les mathématiques sont « libres » depuis la nuit des temps (enfin depuis que les pythagoriciens ont cessé de se cacher, pour être plus précis)^[1]. Chacun est libre des les étudier, les copier et les améliorer en étant fortement encouragé à rendre évidemment publiques ces améliorations, dans la plus pure tradition universitaire.

C'est absurde, mais si il fallait a posteriori leur accoler une licence issue de la culture libre, ce pourrait être la plus simple des Creative Commons, la CC-By, puisque cette notion de paternité est très importante dans un monde scientifique avant tout motivé par la reconnaissance des pairs et la place laissée dans l'Histoire de leur champ disciplinaire.

On retrouve d'ailleurs souvent les crédits et les hommages dans les noms que l'on donne aux résultats. On parlera ainsi de la preuve selon Euclide du théorème de

Thalès.

Les mathématiques seraient donc « libres », les professeurs également (enfin surtout en France dans le secondaire avec Sésamath), mais quid des mathématiciens eux-même et de leurs pratiques ? Et si ils s'organisaient à la manière d'un projet de développement d'un logiciel libre pour chercher et éventuellement trouver ensemble des résultats importants ?

Entendons-nous bien. Les mathématiciens ne vivent bien entendu pas dans une tour d'ivoire. Ils sont habitués à travailler et échanger ensemble au sein d'un laboratoire, lors d'un séminaire... et évidemment sur Internet. Mais l'aventure intellectuelle et collective du « Projet Polymath » que nous avons choisi de vous narrer ci-dessous est un peu différente, en ce sens que ses auteurs eux-mêmes la définissent comme une expérience « open source ».

Ne vous arrêtez surtout pas à la complexité mathématique de ce qu'ils cherchent à accomplir (à savoir une « meilleure » preuve d'un théorème déjà démontré !), c'est totalement secondaire ici. Ce qui nous intéresse en revanche c'est comment ils y sont arrivés (désolé d'avoir vendu la mèche) et les enseignements qu'ils en tirent pour l'avenir.

Parce qu'il se pourrait que cet mode opératoire, original et efficient, fasse rapidement des émules, et ce bien au delà des mathématiques.

Remarque : C'est le terme « open source » qui a été choisi tout du long par les auteurs et non « free software ». Contrairement au domaine logiciel, il ne s'agit pas d'un choix ou d'un parti pris mais tout simplement d'une question de sens : « open source mathematics » étant largement plus signifiant que « free software mathematics » !

Mathématiques massivement collaboratives

Massively collaborative mathematics

Timothy Gowers^[2] et Michael Nielsen^[3] - 24 octobre 2009 - Nature.com (Opinion)
(Traduction Framalang : Olivier et Goofy)

Le « Projet Polymath » est la preuve que l'union de nombreux cerveaux peut résoudre des problèmes mathématiques complexes. Timothy Gowers

et Michael Nielsen nous livrent leurs réflexions sur ce qu'ils ont appris des sciences open source.

Le 27 janvier 2009, l'un d'entre nous, Gowers, a lancé une expérience inhabituelle par l'intermédiaire de son blog (*NdT : avec ce billet Is massively collaborative mathematics possible?*). Le Projet Polymath s'était fixé un but scientifique conventionnel : s'attaquer à un problème mathématique irrésolu. Mais son but plus ambitieux était d'innover dans la recherche en mathématiques. Reprenant l'idée des communautés open source comme Linux et Wikipédia, le projet s'appuyait sur des blogs et des wikis pour canaliser une collaboration complètement ouverte. Libre à tout un chacun d'en suivre la progression et s'il le désire d'apporter sa contribution. Les blogs et wikis étaient en quelque sorte une mémoire à court terme collective, un brainstorming ouvert à grande échelle dédié à l'amélioration des idées.

Le résultat de la collaboration dépassa de loin les espérances de Gowers, offrant une belle illustration de ce que nous pensons être une dynamique formidable pour les découvertes scientifiques : la collaboration de nombreux cerveaux connectés par Internet.

Le Projet Polymath visait à trouver une preuve élémentaire d'un cas particulier du théorème de Hales-Jewett sur la densité (*NdT : DHJ pour Density Hales-Jewett*), théorème central de l'analyse combinatoire, une branche des mathématiques qui étudie les structures discrètes (voir Morpion à plusieurs dimensions). Ce théorème avait déjà été démontré, mais les mathématiciens ne se contentent pas toujours d'un seul chemin. De nouvelles preuves peuvent être déterminantes pour une meilleure compréhension du théorème. Trouver une nouvelle démonstration du théorème DHJ était important pour deux raisons. Tout d'abord il fait partie d'un ensemble de résultats fondamentaux possédant plusieurs démonstrations, alors que le théorème DHJ n'en avait jusqu'alors connu qu'une seule, longue, complexe, et très technique. Seul un élan de nouvelles idées pouvait amener à une preuve plus simple, s'appuyant sur des concepts de base plutôt que sur des techniques compliquées. Ensuite, du théorème DHJ découle un autre théorème important, le théorème de Szemerédi. La découverte de nouvelles preuves de ce théorème a conduit à de grandes avancées au cours de la dernière décennie, on pouvait donc raisonnablement s'attendre à ce qu'un même phénomène accompagne la découverte d'une nouvelle démonstration du théorème DHJ.

À l'origine du projet on retrouve Gowers. Il publia une description du problème, indiqua quelques ressources et établit une liste préliminaire de règles régissant la collaboration. Grâce à ces règles, les échanges sont restés polis et respectueux, elles encourageaient les participants à proposer une et une seule idée par commentaire, même sans la développer complètement. Ces règles stimulaient ainsi les contributions et aidaient à conserver le caractère informel de la discussion.

Mettre le projet sur les bons rails

La discussion collaborative a vraiment commencé le 1er février, doucement : il a fallu attendre plus de 7 heures avant que Jozsef Solymosi, un mathématicien de l'Université de la Colombie Britannique à Vancouver, fasse le premier commentaire. Quinze minutes après, un nouveau commentaire était posté par Jason Dyer, enseignant dans un lycée de l'Arizona. Trois minutes après, c'est Terence Tao (lauréat de la médaille Fields, la plus haute distinction en mathématiques), de l'Université de California à Los Angeles, qui écrivit son commentaire.

Au cours des 37 jours qui suivirent, ce sont pas moins de 27 personnes qui ont apporté leur contribution au travers d'environ 800 messages, pour un total de 170 000 mots. Personne n'a été spécialement invité à participer : la discussion était ouverte à tout le monde, des doctorants aux experts mathématiciens. Nielsen a mis en place un wiki pour mettre en avant les contributions importantes apparues dans les commentaires du blog. Au moins 16 autres blogs ont parlé du projet, il s'est hissé à la première page de l'agrégateur Slashdot technology et a donné naissance à un projet assez proche sur le blog de Tao.

Tout s'est déroulé sans accroc : pas de « trolls » (ces adeptes des posts non constructifs, voire malveillants), ni de commentaires bien intentionnés mais complètement inutiles (on peut malgré tout signaler que le wiki a été spammé). Le rôle de modérateur pris par Gowers se résumait essentiellement à corriger les fautes.

Le projet progressa bien plus rapidement qu'attendu. Le 10 mars, Gowers annonça qu'il était assez sûr que les participants au projet Polymath avaient découvert une preuve élémentaire d'un cas particulier du théorème DHJ et, qu'étonnamment (compte tenu des expériences avec des problèmes similaires),

cette preuve pouvait être généralisée assez facilement pour prouver le théorème entier. La rédaction d'un article décrivant cette preuve est entreprise, ainsi que celle d'un second papier décrivant des résultats liés. De plus, alors que le projet était encore actif, Tim Austin, doctorant à l'Université de Californie, Los Angeles, publia une autre preuve (non élémentaire celle-ci) du théorème DHJ en s'appuyant largement sur les idées développées dans le projet Polymath.

Les archives du projet Polymath constituent une ressource exceptionnelle pour les étudiants en mathématiques, les historiens et autres philosophes des sciences. Pour la première fois, on peut suivre le cheminement intellectuel complet à l'origine d'un résultat mathématique sérieux. On y voit les idées naître, grandir, changer, s'améliorer ou être abandonnées, et on y découvre que la progression de la compréhension ne se fait pas nécessairement par un unique pas de géant, mais plutôt par le regroupement et le raffinement de plusieurs petites idées.

C'est un témoignage direct de la persévérance dont il faut faire preuve pour résoudre un problème complexe, pour progresser malgré les incertitudes et on réalise que même les meilleurs mathématiciens peuvent faire des erreurs simples et s'entêter à creuser une idée vouée à l'échec. Il y a des hauts et des bas et on ressent une vraie tension à mesure que les participants s'approchent d'une solution. Les archives d'un projet mathématique peuvent se lire comme un thriller, qui l'eût cru ?

Des implications plus larges

Le projet Polymath ne ressemble pas aux collaborations à grande échelle traditionnelles, comme on peut en trouver dans l'industrie ou dans les sciences. En général, ces organisations sont divisées hiérarchiquement. Le projet Polymath était complètement ouvert, chacun était libre d'apporter sa pierre à l'édifice... n'importe quelle pierre. La variété des points de vue a parfois apporté des résultats inattendus.

Se pose alors la question de la paternité : difficile de décréter une règle de paternité stricte sans heurt ou sans décourager des contributeurs potentiels. Quel crédit accorder aux contributeurs apportant uniquement une idée perspicace, ou à un contributeur très actif mais peu perspicace ? Le projet a adopté une solution provisoire : il signe ses articles du pseudonyme « DHJ Polymath » suivi d'un lien vers les archives (*NdT : cf cette première publication ainsi signée A new proof of*

the density Hales-Jewett theorem). Grâce à la collaboration ouverte privilégiée par le projet Polymath, on sait exactement qui a fait quoi. Au besoin, une lettre de recommandation peut isoler les contributions d'un membre du projet en particulier, comme c'est déjà le cas en physique des particules où il est courant de voir des articles avec des centaines d'auteurs.

Se pose aussi le problème de la conservation. Les archives principales du projet Polymath sont dispersées sur deux blogs (1 et 2) et un wiki, le rendant très dépendant de la disponibilité de ces sites. En 2007, la bibliothèque du Congrès américain a initié un programme de conservation des blogs tenus par les professionnels du droit. De la même manière, mais à plus grande échelle, un programme similaire est nécessaire pour conserver les blogs et wikis de recherche.

D'autres projets, ayant vu le jour depuis, permettront d'en apprendre plus sur le fonctionnement des mathématiques collaboratives. Il est en particulier crucial de savoir si ce genre de projet peut être élargi à plus de contributeurs. Bien qu'il ait rassemblé plus de participants qu'une collaboration classique en mathématiques, le projet Polymath n'est pas devenu la collaboration massive que ses créateurs espéraient. Ceux qui y ont participé s'accordent sur le fait que pour grandir, il faudrait que le projet adapte sa manière de procéder. La narration linéaire du blog posait, entre autre, problème aux nouveaux arrivants. Difficile pour eux, parmi la masse d'informations déjà publiées, d'identifier où leur aide serait la plus précieuse. Il y avait également le risque qu'ils aient loupé un « épisode » et que leur apport soit redondant.

Les projets de logiciels libres utilisent des logiciels de suivi de version pour organiser le développement autour des « problèmes », des rapports de bogues ou des demandes de fonctionnalités en général. Ainsi, les nouveaux arrivants ont une vision claire de l'état du projet, ils peuvent concentrer leurs efforts sur un « problème » particulier. De même, la discussion est séparée en modules. Les futurs projets Polymath pourraient s'en inspirer.

Bientôt, les sciences ouvertes

L'idée derrière le projet Polymath est potentiellement applicable à tous les défis scientifiques, mêmes les plus importants comme ceux pour lesquels le Clay Mathematics Institute de Cambridge, Massachusetts offre un prix d'un million de

dollars. Même si certaines personnes souhaiteraient toujours garder tout le mérite pour elles-mêmes et être pour ainsi dire refroidies par l'aspect collaboratif, d'autres au contraire pourraient y voir une excellente opportunité d'être associés à la résolution de l'un de ces problèmes.

Dans d'autres disciplines, l'approche open source ne gagne que très lentement du terrain. Un domaine s'y est ouvert avec succès : la biologie synthétique. Les informations ADN d'organismes vivants sont enregistrées numériquement et envoyées à un dépôt en ligne, comme celui du MIT Registry of Standard Biological Parts. D'autres groupes peuvent utiliser ces informations dans leur laboratoire et, s'ils le souhaitent, reverser leurs améliorations au dépôt. qui compte actuellement plus de 3 200 contributions, apportées par plus de 100 entités différentes. Les découvertes qui en découlent ont fait l'objet de nombreux articles comme par exemple Targeted Development of Registries of Biological Parts. La biologie open source et les mathématiques open source montrent comment la science peut progresser grâce aux diverses contributions apportées par des gens aux compétences variées.

On pourrait, de la même manière, employer ces techniques open source dans d'autres domaines, telle la physique et l'informatique théoriques, où les données brutes contiennent de nombreuses informations et peuvent être partagées librement en ligne. Appliquer les techniques open source aux travaux pratiques est plus délicat, les conditions expérimentales étant difficilement reproductibles. Quoiqu'il en soit, le partage libre des données expérimentales permet néanmoins leur analyse libre. Adopter ces techniques open source à grande échelle ne sera possible que grâce à un changement profond des mentalités en science et au développement de nouveaux outils en ligne. Nous croyons à un fort développement de la collaboration de masse dans de nombreux domaines des sciences, et que cette collaboration massive repoussera les limites de nos capacités à résoudre des problèmes.

Notes

[1] Crédit photo : Robynejay (Creative Commons By-Sa)

[2] Timothy Gowers appartient au Département of Pure Mathematics and Mathematical Statistics de l'Université de Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK et il est Royal Society 2010 Anniversary Research

Professor.

[3] Michael Nielsen est écrivain et physicien, il habite à Toronto et travaille sur un livre sur le futur de la science.