remarquable et étonnant théorème suivant : une théorie consistante donnée T ne peut démontrer qu'un nombre fini d'affirmations du type « $K(s) = n$ » ou du type « $K(s) > n$ ». Ainsi, au-delà d'un certain n_0 (qui dépend de la théorie T), tous les énoncés affirmant une complexité supérieure à n_0 sont des indécidables de la théorie T.

Éviter les indécidables

Toute théorie formalisée est donc gravement aveugle à la complexité, puisqu'elle ne sait pratiquement rien démontrer de la complexité des suites finies de 0 et de 1. Cela suggère, toujours pour éviter les indécidables, d'ajouter des axiomes de la forme « $K(s) = n$ » ou « $K(s) > n$ » pour des suites finies s et des entiers n choisis. Étudier la force axiomatique de la complexité et du hasard consiste donc à poser les questions « Est-ce que l'ajout de tels indécidables a des effets intéressants ? » et « Comment faire cela au mieux ? ». C'est exactement le sujet de l'article des chercheurs réunis autour de L. Bienvenu.

Insistons sur le fait que ces travaux, aussi abstraits qu'ils paraissent, renforcent notre compréhension philosophique et mathématique sur un sujet fondamental. On doit déjà beaucoup à la théorie algorithmique de l'information dans les progrès faits récemment concernant le hasard, car ce qu'elle montre va au-delà de la théorie des probabilités en établissant, ce qui était loin d'être évident, que les concepts de calculs, d'algorithmes et de programmes sont profondément liés à cette notion.

Les résultats obtenus par L. Bienvenu et ses collègues nous éclairent sur quatre points principaux.

1) Ajouter sans méthode des axiomes affirmant des propriétés presque certaines sur la complexité ne conduit pas à prouver des théorèmes nouveaux limitant l'incomplétude gödélienne, mais en revanche permet de raccourcir les démonstrations.

2) Ajouter comme axiomes des formules indiquant la complexité de séquences assez nombreuses du type « $K(s) = n$ » ou « $K(s) > n$ » supprime les indécidables de la forme « pour tout x : $P(x)$ » *(voir l'encadré 3 pour l'importance de telles formules)*. Cela constitue un renforcement considérable du pouvoir démonstratif de la théorie. L'affirmation « Tout nombre pair supérieur à 2 est somme de deux nombres premiers » (conjecture de Goldbach) est de la forme « Pour tout x : $P(x)$ » ; elle ne peut être indécidable dans une théorie complétée systématiquement par des axiomes concernant la complexité.

3) Tous les axiomes du type « $K(s) = n$ » n'ont pas la même force. Certains, assez rares, sont extrêmement puissants et rendent inutiles la plupart des autres.

4) Ajouter comme axiome qu'une séquence infinie s de 0 et de 1 est aléatoire au sens de Martin-Löf renforce la théorie, mais ne le fait jamais autant que la méthode précédente.

Parmi les 2^n suites de 0 et de 1 de longueur n (par exemple n = 100 000), on démontre qu'au plus une sur 1 024 ($= 2^{10}$) a une complexité de Kolmogorov inférieure à $n - 10$ (ici 99 990). Il n'y en a, au plus, qu'une sur un million (1 048 576 $= 2^{20}$ précisément) ayant une complexité inférieure à $n - 20$ (donc 99 980 pour notre exemple). L'énoncé général est : « Au plus une suite de longueur n sur 2^m a une complexité inférieure à $n - m$ » ; autrement dit, les suites peu complexes sont extrêmement rares *(voir l'encadré 2)*.

Si l'on choisit m = 100, que l'on tire uniformément une séquence s de 100 000 symboles 0 ou 1 (par exemple à l'aide d'une pièce de monnaie) et que l'on ajoute l'axiome « $K(s) > 99\,900$ » à une théorie T, alors sauf malchance extraordinaire (risque de $1/2^{100}$), on ajoute un indécidable vrai à la théorie T à laquelle on s'intéresse. Cela en accroît donc au moins un peu le pouvoir démonstratif : la théorie maintenant sait que $K(s) > 99\,900$.

On dispose ainsi d'une méthode de complétion qui, sauf au plus une fois sur 2^{100} (ce qu'on peut considérer comme totalement négligeable) améliore une théorie T en supprimant au moins un indécidable. La question est maintenant : avoir pris ce risque est-il utile, c'est-à-dire conduit-il à une théorie T' (comportant un axiome de plus) qui démontre des théorèmes nouveaux intéressants, que T était incapable de démontrer avant l'ajout ?

La réponse est négative pour l'essentiel. Cela résulte de la proposition suivante démontrée par le groupe des cinq chercheurs : si l'on ajoute un ou plusieurs axiomes du type envisagé en les choisissant au hasard (par exemple en tirant à pile ou face les suites s concernées) et en prenant un risque total inférieur à un seuil fixé ϵ (par exemple $1/2^{100}$), et si l'on ne considère que les théorèmes qui dans ce processus probabiliste de démonstration sont obtenus avec une probabilité supérieure à ϵ, alors les théorèmes qu'on découvre peuvent l'être sans ajout d'axiomes, autrement dit, pouvaient déjà l'être dans la théorie initiale.

Ajouter des axiomes peu risqués est inutile

Ajouter des informations presque certaines concernant la complexité n'est donc pas vraiment utile pour limiter l'incomplétude. En revanche, un second résultat sur ces processus de démonstration avec ajout aléatoire d'axiomes presque certains établit qu'on en tire un gain sur la taille des démonstrations : on n'a pas vraiment plus de théorèmes, mais les théorèmes de T sont maintenant obtenus par des démonstrations plus courtes.

La situation est analogue à celle déjà rencontrée en arithmétique à propos des tests de primalité probabilistes. Alors que démontrer sans accepter de prendre le moindre risque qu'un entier long (de 10 000 chiffres par exemple) est un nombre premier exige toujours un long calcul, on sait le faire rapidement si l'on accepte un risque d'erreur, même infinitésimal (par exemple inférieur à $1/2^{100}$). Ces tests probabilistes de primalité découverts dans les années 1980 sont d'une grande utilité en cryptographie.

Mis à part des gains sur la taille des démonstrations, ajouter des axiomes du type « $K(s) = n$ » ou « $K(s) > n$ » lorsque s est aléatoire n'est pas le bon moyen d'accroître le pouvoir de démonstration d'une théorie formalisée T. Qu'en est-il de l'ajout d'axiomes fournissant des informations précises sur la complexité, y compris concernant des suites finies s peu complexes ?

Cette fois, les résultats vont être positifs et très instructifs.

Le premier résultat dans cette direction indique ce que donne l'ajout, à une théorie T, de tous les axiomes concernant la complexité des suites finies : pour chaque suite finie s,

3. La classification des formules

En logique, on s'intéresse à la forme logique des affirmations mathématiques, ce qui conduit à une classification des formules. Chacun perçoit qu'en parcourant la liste suivante de formules, où chaque élément de la liste est comme un échelon, la difficulté de compréhension des affirmations mathématiques augmente.

Énoncé A : 2 + 11 = 13.

Énoncé B : Le carré de tout entier pair est aussi pair.

Énoncé C : Pour tout entier n donné, il existe un nombre entier $P > n$ qui est un nombre premier (autrement dit, il y a une infinité de nombres premiers).

Énoncé D : Pour tout entier k, il existe un entier n tel que pour tout entier $m > n$, on ait : $0 < 2^{1/m} - 1 < 1/k$ (autrement dit, la suite $2^{1/m}$ tend vers 1 quand m tend vers l'infini)

Énoncé E : Il existe un entier b et pour tout entier k, il existe un entier n tel que pour tout entier $m > n$: $b - 1/k < f(m) < b + 1/k$ (autrement dit, la suite $f(m)$ a une limite quand m tend vers l'infini et cette limite est un entier).

D'un énoncé au suivant, les formules contiennent de plus en plus de quantificateurs (« il existe », « pour tout ») et c'est ce qui rend leur compréhension plus difficile. Il y en a d'abord 0, puis 1, puis 2, puis 3, puis 4. Cette augmentation de difficulté est utilisée pour définir la hiérarchie arithmétique des formules.

Le premier niveau, noté Δ_0, ne contient que les affirmations arithmétiques immédiates à vérifier (même si parfois cela demande de la patience). D'autres exemples de formules Δ_0 sont « 1 048 576 est une puissance de 2 », « Il n'y a pas de nombres premiers entre 5 490 et 5 500 ». Pour chacune de ces affirmations, on peut entreprendre un calcul dont on sait évaluer la taille par avance et qui, achevé, indiquera si l'affirmation est vraie ou non.

Toutes les formules de Δ_0 vraies sont démontrables dans le système formel de l'arithmétique de Peano. En particulier, c'est le cas de toutes les affirmations vraies du type « n est un nombre premier ». De même, si une formule de Δ_0 est fausse (par exemple « 121 est un nombre premier »), l'arithmétique de Peano sait démontrer sa négation. Aucune des formules de Δ_0 n'est donc indécidable dans l'arithmétique de Peano (ou dans une théorie plus forte comme la théorie des ensembles).

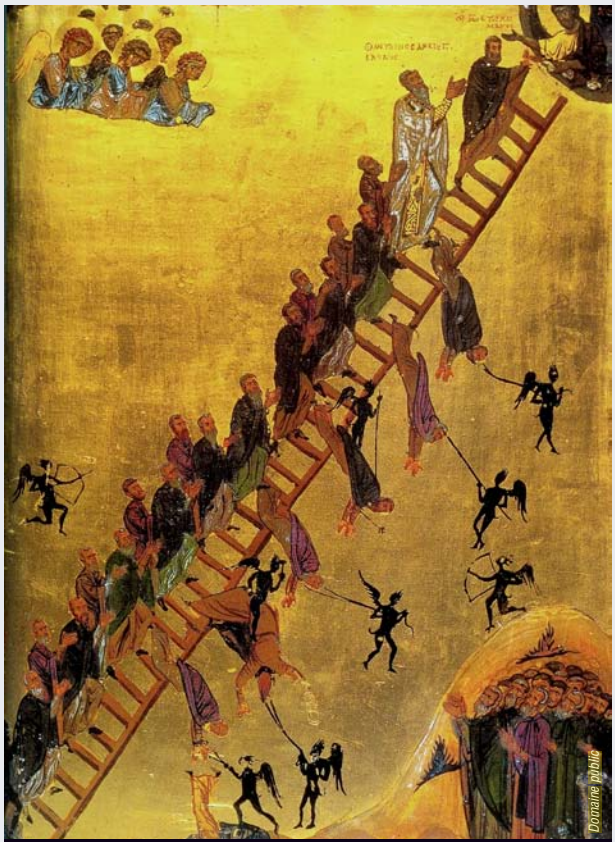
Au-dessus de Δ_0, il y a la classe très importante des formules Π_1. Ce sont les formules de la forme « Pour tout entier n, F(n) », où F(n) est une formule Δ_0. Au-dessus encore, il y a les formules Π_2 de la forme « Pour tout entier n, il existe un entier m tel que F(n, m) » où F(n, m) est une formule de Δ_0. Il y a ensuite les formules Π_3 de la forme « Pour tout entier n, il existe un entier m tel que, pour tout entier k, F(n, m, k) » où F(n, m, k) est une formule de Δ_0, etc.

La formule « Pour tout entier n, n est multiple d'un nombre premier » est Π_1, car si n est connu, la formule « n est multiple d'un nombre premier » est Δ_0. C'est une formule démontrable dans l'arithmétique de Peano.

Parmi les formules Π_1, certaines sont des indécidables. La formule exprimant que le système formel de l'arithmétique de Peano est consistant (c'est-à-dire non contradictoire) est Π_1, car elle signifie « Pour tout entier n, les démonstrations dans l'arithmétique de Peano de longueur au plus n ne conduisent jamais à démontrer que 0 = 1 ».

De nombreux énoncés sur lesquelles les mathématiciens s'interrogent sont équivalents à des formules Π_1. C'est le cas de la conjecture de Goldbach (tout nombre pair > 2 est somme de deux nombres premiers), mais aussi de la conjecture de Riemann (une fois correctement traduite). La formule arithmétique qui exprime le théorème des quatre couleurs (toute carte peut être coloriée avec quatre couleurs sans que deux pays voisins portent la même couleur) est aussi une formule Π_1.

Les énoncés du type « $K(s) \geq n$ » (s et n fixés) sont des formules Π_1. En effet, une telle affirmation indique qu'aucun programme de longueur inférieure à n ne donne s. À cause des programmes ne s'arrêtant pas, cet énoncé est du type « pour tout n, F(n) ».



« Échelle du paradis », selon la conception du moine Jean Climax (VIe siècle) : la vie monastique s'échelonne comme sur les 30 barreaux d'une échelle. Les formules mathématiques s'échelonnent, elles, sur une infinité de degrés.

on ajoute l'axiome « $K(s) = n$ » qui donne sa complexité. La nouvelle théorie T' a le pouvoir de démontrer tous les énoncés de la forme « Pour tout x : $F(x)$ » où F est une formule testable par programme (par exemple, « x est un multiple de 7 »). Les énoncés de ce type constituent ce qu'on nomme la classe de formules Π_1, et on montre que la plupart des questions non résolues en mathématiques sont équivalentes à des énoncés de la classe Π_1.

C'est le cas de la conjecture de Goldbach (« Tout nombre pair supérieur à 2 est somme de deux nombres premiers. ») ; c'est aussi le cas de l'hypothèse de Riemann, qui concerne la répartition des nombres premiers.

Notons bien que l'idée d'ajouter comme axiomes tous les énoncés de la forme « $K(s) = n$ » est étudiée ici comme on étudie ce qui se passe quand deux particules isolées interagissent. Il n'existe aucun moyen pratique d'isoler parfaitement deux particules, mais c'est intéressant quand même de faire le calcul. Ici, il n'existe aucun moyen pratique de connaître la complexité de toutes les suites finies, qui conduirait à ajouter véritablement tous les axiomes envisagés.

On bouche un gros trou grâce à la complexité

Cependant, le résultat de cette expérience de pensée nous apprend quelque chose de profond sur l'indécidabilité : les trous que comporte toute théorie T correspondent (quand on se limite aux énoncés Π_1) à un manque d'information sur la complexité. L'indécidabilité de Gödel provient donc de ce qu'une théorie donnée T n'a jamais une assez bonne connaissance de la complexité et qu'il est impossible de mettre toute cette connaissance dans des axiomes simples. En deux mots : l'indécidabilité gödélienne provient de l'ignorance de la complexité.

L. Bienvenu et ses collègues se sont aussi interrogés sur ce que donnerait une information partielle sur la complexité, fournie par un nombre limité d'axiomes. Ils montrent que pour qu'une théorie soit capable de démontrer toutes les formules Π_1 de taille au plus n, il suffit d'ajouter un axiome unique de la forme $K(s) > n$ pour une suite finie s bien choisie de longueur n. Les suites s utilisables pour cet ajout concentrent en elles l'information nécessaire pour connaître la complexité de toutes les suites de longueur au plus n.

Quand on veut rendre décidables toutes les formules Π_1, cette façon de procéder est beaucoup plus économique en axiomes que la précédente (qui ajoutait tous les « $K(s) = n$ »). Les cinq chercheurs en déduisent le beau théorème suivant : pour tout entier n, il existe un indécidable dont la taille est approximativement n, qui est une formule Π_1 et qui, ajoutée aux axiomes de l'arithmétique de Peano, supprime tous les indécidables Π_1 de taille inférieure à n. En clair, pour boucher tous les trous correspondant à des formules Π_1 précisée et qui est donc une séquence concentrant toute l'information jusqu'au niveau n qu'il faut connaître pour supprimer les indécidables Π_1 de complexité inférieure à n.

Un tel théorème, contrairement aux résultats envisageant d'ajouter une infinité d'axiomes pour supprimer tous les indécidables Π_1, permet d'envisager réellement une complétion progressive de l'arithmétique de Peano dans le but de supprimer petit à petit les indécidables Π_1.

Kurt Gödel a défendu l'idée qu'il fallait tenter de compléter les axiomes de nos théories pour en limiter l'incomplétude. Ce théorème indique une voie possible et montre qu'il existe des méthodes systématiques pour ce faire (au moins pour les formules Π_1).

4. Le rôle des formules Π_1

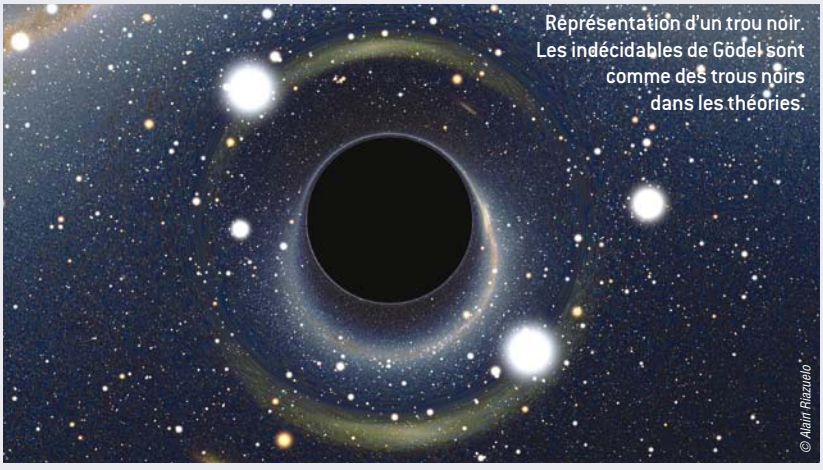
Un théorème central figure dans le travail de Laurent Bienvenu, Andrei Romashchenko, Alexander Shen, Antoine Taveneaux et Stijn Vermeeren. Ce théorème indique non seulement que pour une théorie T, connaître la complexité de toutes les séquences finies permet de supprimer tous les indécidables Π_1 et réciproquement ; il indique aussi ce qui se passe quand T ne connaît que partiellement ce type d'information.

Le théorème utilise la notion de complexité d'une formule (fondée sur la même idée que la définition de la complexité de Kolmogorov) et il énonce que les trois méthodes suivantes pour compléter une théorie T sont équivalentes :

a) Ajouter aux axiomes toutes les formules indiquant la complexité de séquences s qui ont une complexité $K(s)$ inférieure à n.

b) Ajouter aux axiomes toutes les formules Π_1 vraies dont la complexité (comme formule) est inférieure à n ;

c) Ajouter un seul axiome $K(r_n) > n$ pour une séquence r_n dont la construction est



Représentation d'un trou noir. Les indécidables de Gödel sont comme des trous noirs dans les théories.

© Alain Riazuelo