aussi défini par la formule $x^2=2$. Avec cet exemple, nous comprenons qu'une définition $P(x)$ de la théorie T est une formule utilisant la variable x et écrite dans le langage limité de la théorie. La formule $P(x)$ est considérée, par la métathéorie, comme la définition d'un objet précis si la théorie T peut démontrer avec ses axiomes que $P(x)$ n'est vérifié que par un seul élément. Formellement, cela signifie que la théorie démontre que $\forall x\ \forall y\ ((P(x) \wedge P(y)) \Rightarrow x=y)$ (pour tout x et pour tout y, si $P(x)$ et $P(y)$, alors $x=y$).

Dans le cas de l'arithmétique de Peano, si pour $P(x)$ on prend «$x=0$», c'est bien une définition de 0 et elle contient 3 symboles.

Avec cette définition de «définissable» il y a bien, en ce qui concerne la métathéorie, pour chaque entier k un plus petit entier non définissable dans T en moins de k caractères, mais cet entier ne gêne pas la théorie T elle-même car, en quelque sorte, elle ne le sait pas!

PAS DE NOTION ABSOLUE DE LA DÉFINISSABILITÉ

La notion métathéorique de définissabilité dépend du langage de la théorie qu'on considère, car elle dépend des formules $P(x)$ que la théorie peut écrire et de ce que la théorie peut en démontrer: il n'y a pas de notion absolue de définissabilité. La logique mathématique a fait apparaître au XX^e siècle des vérités d'une exceptionnelle profondeur dont on n'avait pas conscience auparavant. Parmi elles, il y en a plusieurs précisant qu'un concept est absolu ou au contraire relatif. Citons trois cas.

A. La notion de «calculable par algorithme» est absolue. C'est ce qui résulte des travaux d'Alonzo Church, Kurt Gödel et Alan Turing dans les années 1930. Toutes les méthodes raisonnables pour définir la notion de fonction calculable par algorithme, à la grande surprise des logiciens, aboutissent à la même notion mathématique, parfois dénommée «fonction récursive», ou «fonction calculable par machine de Turing».

B. La notion de «vérité mathématique démontrable», à l'inverse, est toujours relative à une théorie formelle qu'il faut préciser. C'est l'une des conséquences des théorèmes d'incomplétude de Gödel; ceux-ci montrent que toute théorie intéressante peut formuler des énoncés dont elle ne peut prouver ni qu'ils sont vrais ni qu'ils sont faux. Par exemple, l'énoncé «L'arithmétique de Peano ne produit aucune contradiction» n'est pas démontrable dans l'arithmétique de Peano (sauf si l'on découvre une telle contradiction), mais la théorie ZFC des ensembles, elle, peut le démontrer. Il résulte de cette vérité profonde sur les mathématiques que, pour progresser, il faut non seulement trouver de nouvelles démonstrations – c'est le travail de base du mathématicien –, mais aussi formuler

1 LES PARADOXES SUGGÈRENT DES DÉMONSTRATIONS

Nous allons examiner comment, paradoxalement, les paradoxes amènent des démonstrations.

Pour cela examinons une version analogue du paradoxe de Berry : « Le plus petit entier positif non définissable en français en moins de 100 caractères ».

Avec un alphabet de 200 symboles (minuscules, majuscules, chiffres, signes de ponctuation, lettres accentuées, symboles supplémentaires *, $, €, &, # @, etc.), il y a au plus $200 + 200^2 + ... + 200^{99} \approx 6 \times 10^{227}$ entiers définissables en moins de 100 caractères.

Il existe donc des entiers positifs non définissables en moins de 100 caractères et donc un plus petit. Le problème est que la phrase de définition n'a que 84 caractères et définit un entier positif non définissable en moins de 100 caractères !

Nous tirerons de ce paradoxe une démonstration produisant un résultat intéressant, dont voici le détail.

Nous nous fixons un langage de programmation, par exemple Python ou C++ et c'est aux programmes dans ce langage que nous nous référerons. Si, pour un entier n, il existe un programme de k caractères qui produit n comme résultat, nous dirons que « n est programmable en k caractères ».

Nous allons démontrer que « la fonction $n \to K(n)$ qui à tout entier n associe la taille $K(n)$ du plus court programme produisant cet entier comme résultat n'est pas une fonction programmable ».

Supposons le contraire. Soit P le programme calculant K.

Pour tout entier k donné, on sait grâce à P trouver le plus petit entier n_k non programmable en moins de k caractères. Pour ce faire, nous calculons $K(0)$, puis $K(1)$, etc. jusqu'à trouver le premier entier n tel que $K(n) > k$.

Puisque K est programmable, cette méthode donne un programme P′ qui, quand nous lui donnons k, calcule n_k. Le programme P′, qui dépend du paramètre k, donne pour chaque k un programme P'_k calculant n_k, qui est le même que P′, sauf que k est spécifié, par exemple dès le début du programme.

Dans le programme P'_k, l'entier k n'est présent qu'une fois (on écrit par exemple k = 32 431 au début). On peut supposer qu'on écrit la valeur de k en notation décimale. La longueur du programme P'_k est de la forme $[\log_{10}(k)] + C$, car, pour écrire un entier en décimal, il suffit d'utiliser $[\log_{10}(k)] + 1$ caractères où $[m]$ désigne la partie entière de m. La constante C est la longueur de ce qui, dans le programme P'_k, n'est pas l'écriture de k.

Nous savons que $([\log_{10}(k)] + C)/k$ tend vers 0 quand k tend vers l'infini. Donc pour un k_0 assez grand, $([\log_{10}(k_0)] + C)/k_0$ est inférieur à 1, d'où $[\log_{10}(k_0)] + C < k_0$.

C'est absurde car, pour un tel k_0, le programme P_{k_0} a pour longueur $[\log_{10}(k_0)] + C$, qui est strictement inférieur à k_0, et calcule n_{k_0}, qui, par définition, est un entier non calculable par un programme de longueur inférieure à k_0.

Conclusion : la fonction $n \to K(n)$ n'est pas une fonction calculable par programme.

Cette utilisation du paradoxe de Berry a été découverte par Gregory Chaitin, qui a aussi proposé une démonstration alternative à celle de Gödel utilisant le paradoxe de Berry pour démontrer le théorème d'incomplétude (voir G. J. Chaitin, The Berry paradox, *Complexity*, vol. 1, pp. 26-30, 1995).

Gregory Chaitin

de nouveaux axiomes, qui augmentent le pouvoir déductif des théories utilisées et qui conduisent à construire une hiérarchie complexe de théories de plus en plus puissantes.

C. La notion de «définissable» est toujours relative à une théorie. C'est la leçon du paradoxe de Berry, mais c'est aussi un résultat du logicien polonais Alfred Tarski. Confirmant les résultats sur le «démontrable», l'étude de la notion de «définissable» oblige à considérer une hiérarchie de théories de plus en plus puissantes, chacune servant de métathéorie à celles placées sous elles, pour lesquelles elle peut exprimer la notion de définissable.

UTILITÉ DES PARADOXES

Notons que les paradoxes des philosophes ou créés pour le plaisir du jeu, même si on ne peut pas les utiliser pour introduire des contradictions dans les théories mathématiques (du moins jusqu'à présent!), sont malgré tout très importants. Il est souvent arrivé que, en s'en inspirant, les logiciens découvrent de nouveaux procédés de démonstration. Le paradoxe du menteur «Je mens» a donné la formule autoréférentielle «Je ne suis pas démontrable dans la théorie T», qui est l'idée de base de la démonstration des théorèmes d'incomplétude de Gödel. Il se trouve que le paradoxe de Berry, outre qu'il permet de retrouver le théorème d'incomplétude, est aussi l'idée de la démonstration d'un résultat important en théorie de la complexité. Ainsi, d'une variante du paradoxe de Berry (*voir l'encadré 1*), on déduit une démonstration.

Cette récupération des paradoxes de la langue commune pour en faire des démonstrations est utile et amusante, mais certaines subtilités mises en évidence par les paradoxes introduisent des difficultés plus délicates à circonscrire que celles créées par le paradoxe du menteur ou le paradoxe de Berry. On se trouve en effet parfois confronté à des situations où on a du mal à croire que l'on va s'en sortir et où la solution proposée par les logiciens ne laisse pas totalement intacte la confiance qu'on avait dans certains concepts mathématiques.

Récemment, Joel David Hamkins, professeur à l'université d'Oxford et élève du logicien universellement admiré Hugh Woodin, a insisté sur des exemples troublants de la relativité des concepts et de la nécessité absolue de ne pas confondre les concepts externes (métathéoriques) à une théorie et les concepts analogues qu'elle tente de retrouver en interne. L'impossible coïncidence de la métathéorie et de la théorie est réellement déconcertante. Les remarques et exemples de Joel Hamkins sont liés au théorème de Löwenheim-Skolem et au paradoxe qui en résulte, et en aggravent la portée.

Le théorème de Löwenheim-Skolem a été énoncé en 1915 par le mathématicien allemand Leopold Löwenheim et démontré de manière complète en 1920 par le mathématicien norvégien Thoralf Skolem. Énonçons-le avant d'expliciter les termes qui le composent:

Si une théorie formulée dans le cadre de la logique du calcul des prédicats du premier ordre possède un modèle infini, alors elle possède un modèle dénombrable, c'est-à-dire dont les éléments constitutifs sont aussi nombreux que les entiers, mais pas plus.

DES MODÈLES D'UNE THÉORIE

Expliquons cela. Un modèle pour une théorie est une structure qui vérifie les axiomes de la théorie. Si, par exemple, on considère la théorie T ayant pour unique axiome A: $\forall x\, \forall y\, (x \spadesuit y = y \spadesuit x)$, alors la donnée des entiers $\mathbb{N}$ et de l'addition entre entiers en guise d'opération $\spadesuit$ constitue un modèle de T, car on sait que l'addition entre entiers est commutative. La donnée des nombres réels $\mathbb{R}$ et de la multiplication en guise de $\spadesuit$ est un autre modèle de la théorie T, et il y a encore bien d'autres possibilités. Les formules vraies d'une théorie sont celles vérifiées par tous les modèles de la théorie. Il s'agit d'une notion mathématique parfaitement précise dans la métathéorie, puisque la notion de modèle l'est.

Dans notre exemple, la formule f: $\forall x\, (x \spadesuit x) \spadesuit x = x \spadesuit (x \spadesuit x)$ est une formule vraie de la théorie T: vous ne trouverez jamais un ensemble et une opération sur cet ensemble qui vérifie l'axiome A mais pas la formule f. En revanche, la formule g: $\forall x\, \forall y\, \forall z\, (x \spadesuit y) \spadesuit z = x \spadesuit (y \spadesuit z)$ n'est pas une formule vraie de la théorie T, car même si elle est vraie dans certains modèles de T, elle ne l'est pas dans tous. Par exemple, la formule g n'est pas vraie en prenant pour modèle la structure définie par les entiers et l'opération $x \spadesuit y = x^2 + y^2$, car $(0 \spadesuit 1) \spadesuit 2 = 1 + 4 = 5$ tandis que $0 \spadesuit (1 \spadesuit 2) = 0 \spadesuit 5 = 25$.

Le «calcul des prédicats du premier ordre» est la logique utilisée pour donner l'exemple ci-dessus de la théorie T avec un seul axiome A. Dans une théorie formulée en calcul des prédicats du premier ordre, on peut utiliser des symboles logiques, des variables, des quantificateurs portant sur les variables, des symboles de prédicats ou d'opérations (comme le $\spadesuit$) ainsi qu'indiqué dans l'encadré 2. Le théorème de complétude de Gödel, démontré par ce dernier en 1929 dans sa thèse de doctorat, indique que la logique du calcul des prédicats du premier ordre est telle que les formules démontrables à partir des axiomes habituels du calcul des prédicats du premier ordre (notion syntaxique) sont les mêmes que les formules vraies de la théorie (c'est-à-dire vraies dans tous les modèles des axiomes).

C'est un résultat central en logique et c'est lui qui justifie qu'on formalise les théories mathématiques en utilisant le calcul des prédicats du premier ordre. Il a en particulier la