3. Les axiomes de fondation et d'antifondation

L'axiome de fondation de ZFC indique que partant d'un ensemble (une boîte), si l'on prend l'un de ses éléments (une boîte-ensemble dans la boîte) et qu'on prend un élément dedans, et qu'on prend encore un élément dedans, etc., alors nécessairement on est obligé de s'arrêter : il n'existe pas de chaînes infinies descendantes x_0, x_1, x_2, x_3, x_4,... avec $x_1 \in x_0$, $x_2 \in x_1$, $x_3 \in x_2$, $x_4 \in x_3$,...

Cet axiome correspond à l'idée que lorsqu'on passe des êtres vivants aux cellules, des cellules aux molécules, des molécules aux atomes, des atomes aux quarks etc., on finit nécessairement par arriver à un ultime niveau qui n'est plus décomposable. L'axiome interdit aussi qu'un ensemble puisse se retrouver à l'intérieur de lui-même : $x \in x$ et interdit donc l'ensemble de tous les ensembles qui aurait cette propriété.

Le monde de la théorie classique des ensembles ZFC de Zermelo-Fraenkel correspond donc à une vue hiérarchisée du monde qui est facile à appréhender. C'est ce côté hiérarchisé et rassurant que la théorie des hyperensembles propose d'abandonner.

L'axiome d'antifondation qui définit la théorie des hyperensembles exprime que tout système d'équations, fini ou infini, de la forme : $x = \{1, \{7, 8\}, y, z\}$; $y = \{2, x, z\}$; $z = \{5\}$ (à gauche des variables représentant des ensembles, à droite des expressions utilisant les mêmes variables-ensembles) possède une solution unique : un hyperensemble.

L'équation $x = \{x\}$ définit l'hyperensemble le plus simple. Il est noté Ω et contient un élément : lui-même. C'est une sorte de bouteille de Klein. L'hyperensemble infiniment profond $\{1, \{2, \{3, \{...\}\}\}\}$ est le x_1 de la solution du système d'équations : $x_1 = \{1, x_2\}$, $x_2 = \{2, x_3\}$, $x_3 = \{3, x_4\}$, etc.

Cette possibilité de construire des ensembles infiniment profonds est un outil puissant de modélisation des structures autoréférentes souvent rencontrées en informatique où des objets dépendent mutuellement et cycliquement les uns des autres (jeux de pointeurs, flux de données, machines reliées en réseaux, etc.). Jon Barwise et John Etchemendy, dans leur livre intitulé *The Liar* (« Le menteur »), traitent les phrases autoréférentes et, par exemple, la phrase « je mens » est associée à l'hyperensemble $m = \{\text{faux}, \{m\}\}$ (une valeur de vérité, suivie d'un contexte simple composée dans ce cas de la phrase elle-même).

caractérisé par ses éléments, les connaître c'est le connaître. Les axiomes qui suivent donnent des moyens de construire de nouveaux ensembles à partir d'ensembles déjà connus. Ces axiomes sont les règles « légales » de définition et de manipulation de ces êtres mathématiques que sont les ensembles.

L'axiome de la réunion indique que si E est un ensemble, le regroupement des éléments d'ensembles appartenant à E, c'est-à-dire des x appartenant à des y appartenant à E, constitue encore un ensemble. De $E = \{\{a, b\}, \{\pi, 1/e\}, \{1, 2, 3\}\}$, on déduit l'existence de l'ensemble $\{a, b, \pi, 1/e, 1, 2, 3\}$. Ainsi, l'ensemble des nombres de Fibonacci $(1, 1, 2, 3, 5, 8, 13, 21,...)$ et l'ensemble des nombres triangulaires $(1, 3, 6, 10, 15,...)$ forment un nouvel ensemble. Vous pouvez « parler » de l'ensemble des éléments présents dans l'un au moins des deux ensembles.

L'axiome de l'ensemble des parties affirme que le regroupement des sous-ensembles d'un ensemble E constitue lui aussi un ensemble. De $E = \{0, 1, 2\}$, on déduit l'existence de $\mathcal{P}(\mathcal{E}) = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0, 1\}, \{0, 2\}, \{1, 2\}, \{0, 1, 2\}\}$. Notons « pour le plaisir » que si l'ensemble E a k éléments, il y a C_k^1 éléments individuels, C_k^2 paires d'éléments, etc., et la formule du binôme donne $(1 + 1)^k = 1 + C_k^1 + C_k^2 + ... + 1 = 2^k$. L'ensemble des parties de E a ainsi 2^k éléments.

L'axiome de séparation ou **axiome de compréhension restreint** indique que si $P(x)$ est une propriété (telle $x > 3$) et que E est un ensemble, alors les x qui sont dans E et qui vérifient cette propriété restrictive $P(x)$ constituent aussi un ensemble. Il est essentiel de noter que cet axiome permet de définir un ensemble à partir d'une propriété, ce qui est bien sûr souhaitable, mais qu'il impose la restriction de ne prendre que des x présents dans un ensemble déjà connu E.

• Cette définition permet d'envisager l'ensemble des nombres premiers. Pour cela, on part de l'ensemble des entiers (construit préalablement) et de la propriété $P(x)$: « x n'a pas de diviseurs autres que 1 et lui-même. »

• La propriété $x = x$, seule, ne définit pas un ensemble. Si c'était le cas, la formule définirait l'ensemble de tous les ensembles, ce qui contredirait le théorème de Cantor de ZFC.

• « Posséder deux éléments exactement » ne définit pas un ensemble dans ZFC.

C'est la clef de la résolution des paradoxes : elle empêche la définition d'ensembles trop gros. On ne peut pas parler de l'ensemble de tous les ensembles ni de l'ensemble des x qui possèdent la propriété de ne pas être éléments d'eux-mêmes. Les paradoxes de Russell et Cantor sont donc évités, de même d'ailleurs que le paradoxe de l'ensemble des ordinaux et celui de l'ensemble des cardinaux.

L'axiome de remplacement indique que si, étant donnée une fonction R telle que pour tout élément x d'un ensemble E, la fonction R détermine un élément y, alors les éléments y constituent un nouvel ensemble, F. En clair, l'image d'un ensemble E par une fonction R est un ensemble. Contrairement à l'axiome de séparation, l'axiome de remplacement peut engendrer (d'une manière contrôlée) des ensembles F qui ne sont pas inclus dans E.

L'axiome de l'infini indique qu'il existe au moins un ensemble infini E. Il existe plusieurs façons de définir un ensemble infini. L'une d'elles, due à Dedekind, consiste à dire que E est infini si E possède un sous-ensemble F différent de E, et que l'on peut mettre en bijection avec E. Selon cette définition, l'ensemble des entiers est infini, car l'ensemble des entiers peut être mis en bijec-

tion avec l'ensemble des entiers pairs en doublant chaque entier (fonction $n \to 2n$).

L'axiome de fondation (AF), proposé par von Neumann en 1925, est explicité sur la figure 3 : il n'existe pas de chaînes infinies de boîtes-ensembles les unes dans les autres et pas de boîtes-ensembles qui se contiennent elles-mêmes !

L'axiome du choix (le « C » de ZFC) indique que si E est un ensemble d'ensembles non vides, alors il existe un ensemble F qui a un élément commun exactement avec chaque ensemble de E. De {{0, 1, 2}, {π, 1/e}, {A, B}}, on déduit l'existence d'un ensemble F comme {2, 1/e, B}. L'axiome du choix crée certaines difficultés *(voir l'article d'octobre 2009 de cette rubrique)* et c'est pour cela qu'on le fait apparaître dans le nom ZFC de la théorie.

Il est remarquable que ces simples axiomes, tous assez naturels, permettent l'expression de pratiquement toutes les idées et démonstrations mathématiques et que depuis plus d'un siècle qu'on explore les conséquences de ZFC, jamais aucune contradiction n'y a été découverte. Cependant, la méthode adoptée par ZFC opère une distinction entre ensembles et regroupements (définis seulement par une propriété). Cette distinction interdit de considérer que certains regroupements simples et naturels sont d'authentiques ensembles… ce qui est étrange. En un mot, la solution exprimée dans ZFC apparaît comme une rustine : elle répare la théorie initiale de Cantor, mais elle lui enlève sa beauté et sa généralité. Elle en limite la puissance en interdisant en particulier de mentionner un ensemble universel contenant la totalité des ensembles ou de définir le nombre *n* comme le regroupement de tous les ensembles possédant *n* éléments, selon l'élégante idée de Cantor et de Frege.

Les révisions possibles

Une extension de ZFC donne un statut théorique aux regroupements problématiques (ensemble de tous les ensembles, ensembles des ordinaux, etc). Elle consiste à introduire, à côté de la notion d'ensemble, la notion de classes en définissant certaines opérations entre classes, dont en particulier la définition d'une classe à partir d'une propriété. La classe de tous les x qui vérifient $x = x$ correspond à l'ensemble de tous les ensembles. Cette théorie de von Neumann, Bernays et Gödel, notée NBG, est une première méthode pour disposer du droit de mentionner les gros ensembles problématiques interdits dans ZFC. Cependant, cette façon d'étendre ZFC n'est pas jugée satisfaisante, car elle se fonde sur une distinction artificielle et injustifiée entre deux catégories de regroupements, les ensembles et les classes, un peu comme dans une société qui considérerait deux types de citoyens : ceux qui jouissent pleinement de tous les droits (les ensembles), et ceux qui n'en ont qu'une partie sans qu'on sache trop bien pourquoi.

Une méthode plus élégante pour augmenter le pouvoir de ZFC consiste à revoir l'axiome de fondation qui est le dernier historiquement ajouté à ZFC et que d'ailleurs Cantor n'avait pas envisagé du tout.

En interdisant directement l'ensemble de tous les ensembles, l'axiome de fondation est rassurant. Cependant, on s'était déjà débarrassé des paradoxes avant de l'introduire. L'utilisation de l'axiome de fondation ressemble un peu à la méthode consistant à pendre un noyé pour être certain qu'il est bien mort ! Notons aussi que même si AF à lui seul évite le paradoxe de l'ensemble de tous les ensembles, il est inopérant contre d'autres paradoxes et ne constitue donc pas une alternative complète à la doctrine de la limitation de la taille que met en œuvre ZFC.

Le Britannique Peter Aczel, et indépendamment les Italiens Marco Forti et Furio Honsell, ont proposé dans les années 1980 de renoncer à AF et de le remplacer. Acceptons l'idée qu'il y a des ensembles qui appartiennent à eux-mêmes, comme on accepte l'idée qu'il y a des nombres dont le carré est –1, et imposons par un axiome particulier qu'il y en existe de toutes les sortes imaginables.

L'axiome d'antifondation, noté AFA *(voir la figure 3)*, indique non seulement qu'il y a toutes sortes d'ensembles appartenant à eux-mêmes, mais que précisés par une loi adéquate, ils sont uniques. Cette double affirmation d'existence et d'unicité confère à l'axiome AFA toute sa force et en fait l'intérêt. L'axiome AFA crée une révolution comparable à l'introduction des nombres complexes. En plus de tous les « ensembles habituels » qui bien sûr restent des ensembles dans cette théorie, il y a maintenant une multitude d'autres ensembles « antifondés » qu'on dénomme *hyperensembles*.

Les hyperensembles

En particulier, dans la théorie des hyperensembles ZFC – AF + AFA (tous les axiomes de ZFC, moins l'axiome de fondation, plus l'axiome AFA), il y a un hyperensemble unique noté Ω qui vérifie $x = \{x\}$. L'hyperensemble Ω est un être étrange possédant un seul élément, lui-même. D'après AFA, il est le seul à posséder cette propriété. Il représente une sorte de monde autiste, refermé en boucle sur lui-même !

Il y a aussi un hyperensemble unique qui vérifie $x = \{x, \emptyset\}$. C'est un objet étrange qui possède deux éléments : l'ensemble vide et lui-même. Il y aussi l'hyperensemble infiniment profond $\{0, \{1, \{2, \{3, \{...\}\}\}\}\}$ *(voir la figure 3)*.

Ajouter un nouvel axiome comme AFA fait prendre des risques : si ZFC ne conduit jamais à une contradiction, en est-il de même de ZFC – AF + AFA ? La réponse est oui : on ne prend aucun risque avec AFA. En effet, P. Aczel a prouvé un résultat de « consistance relative » analogue à un autre obtenu par von Neumann pour AF : si la théorie usuelle des ensembles, ZFC, ne conduit jamais à une contradiction, alors il en est de même de ZFC – AF + AFA. Les hyperensembles n'introduisent aucun danger particulier de contradiction.

L'intérêt d'une théorie se prouve en l'appliquant. C'est aujourd'hui chose faite pour la théorie des hyperensembles. Les domaines dans lesquels elle a été utilisée sont principalement liés à l'informatique et à l'analyse des paradoxes. Plusieurs livres *(voir la bibliographie)* en donnent le détail. Notons que P. Aczel avait été amené à introduire les hyperensembles alors qu'il étudiait le problème informatique de la représentation formelle des flux de données et la synchronisation entre systèmes communicants fonctionnant