LA LOGIQUE

Pierre Wagner

puf

QUE SAIS-JE ?

La logique

PIERRE WAGNER

Maître de conférences à l'université Paris-I
Panthéon-Sorbonne

Deuxième édition mise à jour

4e mille

Tous mes remerciements aux amis et collègues qui ont accepté de lire une première version de ce livre. Leurs commentaires m'ont permis d'apporter de nombreuses améliorations au texte initial.

BIBLIOGRAPHIE THÉMATIQUE
« QUE SAIS-JE ? »

Dominique Lecourt, *La Philosophie des sciences*, n° 3624
Jean-Paul Resweber, *La Philosophie du langage*, n° 1765
Pierre Oléron, *Le Raisonnement*, n° 1671
Jean-Michel Besnier, *Les Théories de la connaissance*, n° 3752
Patrice Hernert, *Les Algorithmes*, n° 2928

ISBN 978-2-13-058456-8

Dépôt légal – 1re édition : 2007
2e édition mise à jour : 2011, mars

6, avenue Reille, 75014 Paris

Chapitre I

LA LOGIQUE : DIRECTIONS, ORIENTATION

Qu'est-ce que la logique ? Quel est son objet ? Quelles sont sa fonction, ses frontières, son histoire ? Sur ces questions, il existe des opinions diverses dont plusieurs sont assez largement répandues.

Pour les uns, la logique est la science du raisonnement correct. Non une science empirique qui aurait pour objet de recenser, décrire et analyser les raisonnements réels, qu'ils soient écrits, exprimés verbalement ou pensés, mais une science des règles que tout raisonnement doit respecter afin d'être valide. À supposer qu'il existe de telles règles et que la logique en soit la science, elle est beaucoup plus qu'une science parmi les sciences. Car son but est alors d'étudier les présupposés communs à toutes les connaissances auxquelles on accède au moyen du raisonnement. Ainsi comprise, la logique n'a pas pour seule fonction de contrôler la validité des inférences ; elle a également pour tâche de structurer l'ensemble de notre savoir. Aussi lui accorde-t-on parfois une fonction constitutive pour la connaissance en général et pour notre système du monde.

Pour d'autres, la logique énonce les lois les plus générales de la pensée, en tant que celle-ci vise le vrai. Indépendantes de tout contenu, de tout objet particulier, ces lois valent alors non seulement pour tout ce qui est, mais également pour tout ce qui peut être pensé en général. Le principe du tiers exclu, par exemple, comme toutes les autres vérités logiques, s'applique à tout énoncé : selon ce principe, soit l'énoncé lui-même, soit sa négation est vrai. Soit il pleut, soit il ne pleut pas, il n'y a pas de troisième possibilité. Selon un autre principe logique, le principe de non-contradiction, un énoncé et sa négation ne sont pas l'un et l'autre vrais. En sorte que si quelqu'un affirmait à la fois que le monde a une fin et que le monde n'a pas de fin, nous serions conduits à penser soit qu'il ne donne pas le même sens à « monde » ou à « avoir une fin » dans les deux énoncés, soit qu'il donne à l'expression « ne… pas », en cette occurrence, un sens différent du sens usuel.

Sur l'histoire de la logique voici ce qu'on entend souvent dire. La logique remonte à l'Antiquité, aux travaux d'Aristote et des stoïciens, et résulte d'un travail de codification des procédés d'argumentation et de raisonnement en usage à cette époque. Au Moyen Âge, elle a fait partie intégrante des études scolastiques, à côté de la rhétorique et de la grammaire. Au XVIIe siècle, Descartes a critiqué son caractère formel et sa stérilité alors que Leibniz a cherché à en faire une science générale, applicable à tous les raisonnements, ainsi qu'un art d'inventer pour trouver la vérité dans les sciences. Au XIXe siècle, des auteurs comme Boole ont commencé à lui appliquer les techniques de l'algèbre, avant que la crise des fondements des mathématiques ne précipite son évolution et ne lui fasse connaître une véritable révolution. Les pères fondateurs de la logique moderne – Frege, Peano, Russell et d'autres – ont alors fait d'elle une logique symbolique, d'une nature comparable à celle des mathématiques, dont elle a fini par constituer l'une des branches. Essentiellement orientée, dans la première moitié du XX[e] siècle, vers le problème du fondement des mathématiques, elle a ensuite trouvé des applications inattendues en informatique et dans l'étude des langues naturelles, domaines dans lesquelles elle fait aujourd'hui florès.

Jadis pensée et enseignée par les philosophes, la logique est donc, selon une opinion courante, devenue une science dont les théorèmes et les démonstrations sont tout aussi peu discutables que ceux de l'arithmétique ou de la géométrie, et son caractère technique la rend difficilement accessible au profane. Comme toute science, elle a ses revues spécialisées, sa communauté de scientifiques et ses programmes de recherche. La logique ainsi comprise est parfois divisée en quatre grandes orientations principales : la théorie de la démonstration, la théorie des modèles, la théorie des ensembles et la théorie de la calculabilité. Comme la plupart des sciences, ses bases sont exposées dans des manuels d'introduction dont les titres précisent, le plus souvent, qu'il y est question de la logique *mathématique*, *symbolique* ou *formelle*.

Que penser d'un tel tableau, qui regroupe quelques-unes des représentations les plus communes touchant la nature, la fonction et l'histoire de la logique ? Bien que les opinions de ce genre soient souvent exposées comme des vérités sur la logique, chacune soulève davantage de questions qu'elle n'apporte de réponse. Leur vraisemblance dépend tantôt de l'état des connaissances à une époque donnée, tantôt de convictions philosophiques qui méritent d'être discutées.

En parlant de *la* logique, par exemple, on sous-entend qu'elle possède une unité qui traverse l'histoire depuis l'Antiquité. Il existe certes une tradition logique dont certains auteurs font incontestablement partie ; mais on serait bien en peine de trouver une définition que chacun d'eux eût jugée acceptable. Aristote, Leibniz, Kant et Frege ne se sont pas seulement fait de la logique des idées très diverses ; ils ne la définissaient tout simplement pas à partir du même projet intellectuel. Aristote, du reste, utilisait l'adjectif *logikos*, mais il ne disposait d'aucun substantif pour désigner quelque chose comme *la* logique, et la question de savoir quels traités, parmi ceux dont il est l'auteur, il convient d'inclure dans ce que nous appelons « la logique aristotélicienne » n'est rien moins qu'évidente. Leibniz, pour sa part, ambitionnait de réaliser une langue logique qui permette non seulement de formaliser les raisonnements mais également de trancher les disputes métaphysiques et théologiques. Quant à la distinction kantienne entre logique formelle et logique transcendantale, elle trouve sa justification dans le projet d'une philosophie générale de la connaissance. Les travaux logiques de Frege ont un sens encore différent puisqu'ils ont pour origine le projet d'un fondement de l'arithmétique qui ferait l'économie de tout recours à l'intuition. La logique qu'on trouve aujourd'hui dans les manuels d'introduction, quant à elle, est extrêmement éloignée, par son esprit comme par son contenu, de la logique telle que la comprenait chacun des auteurs qu'on vient de citer.

Si, pour la recherche d'une définition de la logique et de son unité, on laisse maintenant de côté la perspective historique pour se tourner vers la logique contemporaine, on s'aperçoit rapidement que les questions suivantes sont parmi les plus controversées : y a-t-il une ou plusieurs logiques ? Y a-t-il un critère de la « logicité », c'est-à-dire de ce qui relève de la logique proprement dite, par opposition à ce qui tombe en dehors de son domaine ? Comment définir le concept fondamental de *conséquence logique*, qui met en relation les prémisses et la conclusion d'un raisonnement ? Sur la question de la définition de la logique, on ne peut donc certainement pas se satisfaire de la réponse naïve qui consisterait à dire que la logique, c'est ce que font les logiciens « professionnels », membres d'une même communauté scientifique car, d'une part, la question de savoir qui fait partie de cette communauté n'admet pas de réponse simple, et, d'autre part, ceux qui en font indiscutablement partie ne sont d'accord entre eux ni sur la définition de la logique, ni sur son unité, ni sur la nature du projet intellectuel qui oriente leur travail, ni même sur l'intérêt qu'il pourrait y avoir à délimiter clairement la province de la logique.

La logique est souvent définie comme la science de l'inférence formellement valide. Mais cette expression est tellement générale qu'elle ne nous apprend pratiquement rien tant qu'on n'a pas expliqué ce qu'on entend par « inférence », par « formel » et par « valide » et tant qu'on n'a pas dit à quoi pourrait ressembler une telle science. Or, il s'agit là de questions qui dépassent la logique en tant que science, car toute esquisse de réponse présuppose qu'une solution a déjà été apportée à d'autres problèmes touchant la pensée, le langage, le jugement, la signification, la psychologie ou l'esprit. La logique telle que la concevaient Descartes, Leibniz ou Kant était non seulement indissociable de leur philosophie de la connaissance mais également ancrée dans leur système philosophique tout entier.

Il pourrait sembler, à première vue, qu'il en va différemment de la logique telle qu'on la trouve introduite dans les ouvrages contemporains de logique mathématique car ceux-ci se présentent souvent comme l'exposé d'une science indépendante de tout engagement philosophique. En réalité, cela vient seulement de ce que la plupart de ces ouvrages évitent tout simplement de discuter ce genre de question ; tout au plus mentionnent-ils, dans quelques paragraphes liminaires, un ensemble de présupposés qu'ils demandent au lecteur d'accepter. Fort heureusement, tous les logiciens ne sont pas partisans de ce genre d'économie intellectuelle, et les fondateurs de la logique moderne ont même consacré des ouvrages entiers à des recherches logico-philosophiques par lesquelles ils exposaient et justifiaient les travaux et programmes dans lesquels ils s'engageaient. Ce que montrent ces textes, cependant, comme ceux des auteurs contemporains qui traitent des mêmes sujets, c'est qu'il n'existe pas, aujourd'hui, de conception de la logique sur laquelle la communauté des logiciens soit entièrement d'accord, pas de définition universellement acceptée. Il existe au contraire des opinions divergentes touchant son extension, son unité, son but et son orientation.

Corrélativement, l'époque contemporaine connaît une prolifération de ce qu'on appelle *des* logiques, auxquelles sont attribués des noms plus ou moins évocateurs : logique intuitionniste, logique modale, logique déontique, logique temporelle, logique quantique, logique pertinente, logique IF, etc. Elles se présentent souvent comme des systèmes de signes qui permettent de former des expressions, de leur donner une interprétation, de définir un concept de conséquence logique et surtout de formaliser certains types de raisonnements afin d'en caractériser les règles. Mais l'usage du mot « logique » au pluriel suggère que ce mot reçoit ici une signification différente de celle qu'il a lorsqu'il est question de *la* logique, car l'article défini suppose évidemment qu'il y en a une et non plusieurs. Lorsqu'on donne à cet article son sens le plus fort, on ne pense pas à un ensemble de techniques de formalisation des raisonnements dans des domaines circonscrits ; ce qui est visé, ce sont les règles qui, au-delà des apparences du langage ordinaire, structurent la pensée ou le système de nos connaissances. Il reste à savoir s'il existe effectivement de telles règles universelles – et dans ce cas, quelles sont ces règles – ou si les recherches logiques ainsi comprises ne relèvent pas d'un mythe universaliste.

S'il existe de telles différences de points de vue et de sens, comment la logique peut-elle être couramment considérée comme une science, qui plus est comme l'une des branches de la science la plus assurée de toutes : les mathématiques ? Paradoxalement, l'absence d'accord sur la définition de la logique, sur son orientation générale et ses principes fondamentaux n'enlève rien au caractère scientifiquement contraignant des théorèmes qu'elle démontre, qui ne sont pas moins certains que des théorèmes mathématiques. Il existe effectivement quelque chose comme une doctrine logique commune dont la plupart des manuels exposent les concepts, les techniques et les résultats de base sous des formes variables. On y trouve généralement une définition de ce que les logiciens entendent aujourd'hui par variable, constante, connecteur, quantificateur, relation, fonction, formule, axiome, système formel, dérivabilité formelle, structure d'interprétation, vérité dans une structure, satisfaisabilité, théorie formelle, modèle d'une théorie, conséquence logique, cohérence, complétude, compacité, décidabilité, définissabilité, syntaxe, sémantique, métalangage, logique du premier ordre, logique d'ordre supérieur, pour ne citer que quelques-uns des concepts de base de la logique contemporaine ; on y trouve également la démonstration d'une série de théorèmes fondamentaux : théorèmes de complétude, de compacité, de Löwenheim-Skolem, etc. Mais que penser d'un tel appareillage conceptuel, si hautement technique ? Est-il de nature à répondre aux questions que nous nous posons touchant ce qu'est fondamentalement la logique ?

D'un côté, les résultats qui ont été obtenus en logique mathématique par la voie démonstrative laissent ouvertes un grand nombre de questions, notamment celles qui sont relatives à l'unité de la logique, à son fondement, à sa fonction dans le système du savoir ou à l'interprétation épistémologique des vérités logiques ; d'un autre côté, ces résultats sont si contraignants qu'ils ont effectivement mis en échec les conceptions de la logique qui avaient été défendues par certains des plus grands logiciens ; on pense au logicisme de Frege, à celui de Russell, ainsi qu'au programme de Hilbert dont la réalisation possible fut rendue extrêmement peu vraisemblable par les théorèmes d'incomplétude que Gödel publia en 1931. Aucune réflexion sur la logique ne peut donc ignorer les développements techniques de la logique contemporaine, bien qu'une connaissance de ces développements ne suffise certainement pas à répondre à toutes les questions susceptibles d'être soulevées relativement à ce qu'est la logique.

Ces quelques remarques devraient suffire à convaincre le lecteur qu'il est aujourd'hui difficile de prétendre connaître ne serait-ce qu'une petite partie de la logique si l'on ne maîtrise pas au moins certains de ses concepts de base et si l'on ignore tout de résultats comme les théorèmes de complétude, d'incomplétude, de compacité, de Löwenheim-Skolem, d'indécidabilité et d'indéfinissabilité du prédicat de vérité, pour ne citer que quelques exemples, parmi les plus fondamentaux. D'un autre côté, une étude non critique des bases de la logique mathématique dans un manuel qui en introduit les concepts, les méthodes et les théorèmes ne permet généralement pas d'apercevoir ce qui fait de la logique une discipline non seulement ouverte, vivante et féconde, mais également indéterminée dans sa définition, sa signification et son orientation générale.

Dans ce livre, nous n'aurons l'ambition ni de donner une définition dogmatique de la logique, ni d'ajouter, sur cette question, une nouvelle opinion à celles qui existent déjà. Il ne s'agit pas non plus d'une introduction à la logique mathématique, ni d'une histoire de la logique des origines à nos jours. Nous nous proposons plutôt de donner au lecteur une idée du genre de questions que se posent ou que se sont posés les logiciens, du genre de certitudes qu'ils ont acquises, des projets qui animent leurs recherches, des problèmes sur lesquels ils s'opposent et des raisons pour lesquelles il n'est pas facile de s'entendre sur ce qu'est la logique. On aura compris qu'un tel programme exige l'impossible : sans présupposer aucune connaissance de la part du lecteur, sans l'entraîner non plus dans une exposition formelle des techniques que les logiciens ont mises au point ou des résultats que leurs méthodes leur ont permis d'obtenir, lui donner une idée aussi précise que possible des bases de la logique contemporaine et de ses origines historiques, afin qu'il puisse distinguer quelques-unes des certitudes acquises et des questions ouvertes, et s'interroger lui-même sur ce qu'il est permis d'entendre par « logique ». L'exposition rigoureuse des concepts et théorèmes fondamentaux de la logique contemporaine exigerait un degré de précision et donc une prolixité qu'il est impossible de satisfaire dans un ouvrage de cette nature, qui n'a pas pour vocation de se substituer aux ouvrages d'introduction à la logique proprement dits[1].

On ne pourra évidemment pas discuter tous les usages qui sont faits du mot « logique » aujourd'hui, et certains de ces usages seront délibérément ignorés, par exemple dans des expressions comme « logique de l'économie de guerre », « logique de la découverte scientifique » ou même « logique inductive ». L'examen de ces différents cas, comme celui de bien d'autres, nous entraînerait dans des questions particulières trop éloignées de la logique en général. On ne pourra pas non plus espérer donner un panorama de l'ensemble des recherches qui sont menées dans la logique contemporaine. Nous présenterons tout au plus quelques questions choisies parmi celles qui peuvent être rendues accessibles au néophyte en un nombre limité de pages, en commençant par quelques-unes des transformations qui ont conduit à la logique contemporaine. Les concepts et théorèmes qui seront introduits appartiennent tous aux connaissances communes des logiciens contemporains. Sans qu'il soit question d'en retracer l'histoire, nous prenons le parti d'indiquer, autant qu'il est possible, le contexte dans lequel ils se sont imposés comme tels.

1. On pourra consulter F. Lepage, *Éléments de logique contemporaine*, Montréal, Presses de l'Université de Montréal, 1991 ; F. Rivenc, *Introduction à la logique*, Paris, Payot, 1989 ; ou R. Cori et D. Lascar, *Logique mathématique*, Paris, Masson, 1993 ; Dunod, 2003.

Malgré toutes les réserves qui ont été émises dans ce premier chapitre, nous n'hésiterons pas à faire précéder le mot « logique » de l'article défini, en gardant présent à l'esprit que ce présupposé d'unicité n'est rien moins que problématique.

Chapitre II

LES DÉBUTS DE LA LOGIQUE MODERNE

Frege est souvent considéré comme le père fondateur de la logique moderne. Son *Idéographie*, publiée en 1879, offre le premier exemple d'un système formel effectivement réalisé, et son œuvre introduit plusieurs innovations si importantes que l'histoire de la logique est souvent divisée en deux parties principales : la logique moderne et la logique « préfrégéenne ». Pourtant, la logique telle que la comprend Frege est, à maints égards, fondamentalement différente de la logique qui est communément enseignée et pratiquée aujourd'hui. Cela tient, au moins en partie, à ce que les recherches logiques de Frege étaient orientées par un projet philosophique qui s'est révélé irréalisable dans sa forme initiale, en sorte qu'il y a peu de logiciens qui travaillent encore, aujourd'hui, dans le prolongement de ce projet dit « logiciste ». Frege occupe donc une position tout à fait particulière puisqu'il est généralement reconnu comme l'un des principaux fondateurs de la logique moderne bien que sa conception de la logique ne soit pratiquement plus partagée par personne[1].

1. Sur la conception frégéenne de la logique, cf. W. Goldfarb, « Frege's Conception of Logic », *in* T. Ricketts et M. Potter, éds., *The Cambridge Companion to Frege*, Cambridge, Cambridge UP, 2010.

Le but de Frege n'était pas de réformer la logique mais d'établir les fondements de l'arithmétique, ce mot étant pris en un sens beaucoup plus large que la théorie des nombres entiers, sens qui inclut notamment toute l'analyse mathématique. Aux yeux de Frege, le manque de clarté touchant les principes de l'arithmétique et le caractère tout à fait insatisfaisant de la définition du nombre que pouvaient proposer les mathématiciens et les philosophes de son époque constituaient un véritable scandale pour l'esprit humain. Face à un tel état de fait, le problème n'était pas d'affermir les bases d'une science qui n'aurait pas été suffisamment assurée, mais de mettre en lumière ses fondements. Pour y parvenir, Frege se proposa d'établir par la voie démonstrative la thèse selon laquelle il n'y a pas de différence essentielle entre la logique et l'arithmétique, pas d'objet propre à l'arithmétique qui ne soit de nature logique ni aucune méthode de preuve spécifique dont l'arithmétique puisse se prévaloir ou à laquelle elle doive nécessairement avoir recours. L'un des sens du mot « logicisme » est celui d'une déduction de l'arithmétique à l'intérieur de la logique, ce qu'on exprime couramment en parlant d'une « réduction » de l'arithmétique à la logique.

Le projet de Frege se comprend mieux lorsqu'on le compare à d'autres conceptions des rapports entre logique et arithmétique. Le contraste est particulièrement frappant avec celle de Kant, pour qui la logique générale est la science des règles de l'entendement, science formelle qui, en tant que telle, ne nous donne la connaissance d'aucun domaine d'objets. Parce que, selon Kant, les mathématiques constituent, *a contrario*, un corps de connaissances et supposent que des concepts soient construits, c'est-à-dire représentés *a priori* dans l'intuition, les raisonnements mathématiques ne peuvent pas être déduits des règles logiques de l'entendement. De telles règles ne permettent, à elles seules, ni de construire les concepts ni de démontrer les théorèmes de l'arithmétique, qui est donc irréductible à la logique.

Pour préciser, il convient d'ajouter que Kant donne au mot « logique » des sens différents selon qu'il est question de la logique générale, transcendantale, spéciale, pure ou appliquée. Laissons de côté les cas d'une logique spéciale ou d'une logique appliquée et retenons les deux sens principaux. La logique *générale pure* contient les règles absolument nécessaires de la pensée ; elle fixe les normes de l'usage de l'entendement et de la raison d'un point de vue formel, indépendamment de toute référence à des objets particuliers. La logique *transcendantale*, quant à elle, donne les lois de la raison et de l'entendement en tant qu'elles se rapportent *a priori* à des objets. Elle ne fait donc pas abstraction de tout contenu de connaissance, bien qu'elle n'ait en elle-même le pouvoir de nous faire connaître aucun objet parce que son usage suppose que des objets nous soient donnés dans l'intuition. Ces deux définitions suffisent à indiquer pourquoi en aucun de ces sens, selon Kant, l'arithmétique ne peut être déduite à partir des principes de la logique. Elle ne peut l'être à partir des principes de la logique générale pure parce que celle-ci est entièrement formelle et n'exprime aucun contenu de connaissance, ni à partir des principes de la logique transcendantale parce que celle-ci n'a pas d'usage en dehors de l'intuition. Pour Kant, notre esprit ne peut recevoir de représentations que par cette faculté qu'il nomme « sensibilité », et nos pensées sont sans contenu si elles n'ont aucun rapport à l'intuition sensible. Toute possibilité de déduire les jugements arithmétiques – qui expriment un contenu de connaissance – à partir de jugements logiques est donc exclue.

Le projet de Frege est de montrer au contraire que pour les démonstrations arithmétiques, aucun recours à la construction de concepts dans l'intuition n'est nécessaire et que la pensée pure est capable, par elle-même, de définir les objets et de produire les connaissances de l'arithmétique. L'entreprise de formalisation qui est engagée et exposée dans l'*Idéographie* est un moyen d'atteindre ce but.

Que l'arithmétique se déduise de la logique et qu'il soit possible de faire l'économie de tout recours à l'intuition dans l'écriture d'une preuve ne signifie évidemment pas que les mathématiciens puissent se passer de cette intuition dans leur pratique scientifique. Cela signifie seulement qu'en rendant parfaitement explicites toutes les étapes d'un raisonnement arithmétique, on peut montrer que celles-ci ne requièrent que des moyens purement logiques, tant pour la définition ou la donnée initiale des objets – les nombres – que pour la conduite ultérieure des preuves. Un tel programme suppose que soient indiquées non seulement l'intégralité des propositions utilisées dans les démonstrations, sans qu'aucun présupposé ne demeure implicite, mais également les règles d'inférence selon lesquelles s'effectue, dans le cours de la démonstration, le passage d'une ou plusieurs propositions à une proposition ultérieure. Voilà précisément en quoi consiste la formalisation de l'arithmétique qui est entreprise par Frege, d'abord dans l'*Idéographie* en 1879 puis dans *Les Lois fondamentales de l'arithmétique* dont les deux volumes sont publiés, respectivement, en 1893 et en 1903.

Or, lorsqu'il tente de réaliser cette entreprise de formalisation, Frege constate que le langage usuel est tout à fait inadéquat. Cette difficulté est présentée dans la préface de l'*Idéographie* :

« Pour que rien d'intuitif ne puisse s'introduire ici subrepticement, tout devait reposer sur l'absence de lacune dans la chaîne d'inférence. Comme j'essayais de satisfaire à cette exigence le plus rigoureusement, je trouvai un obstacle dans l'inadéquation du langage : en dépit des lourdeurs de l'expression, plus les relations devenaient complexes, moins la précision que mon dessein exigeait pouvait être atteinte. De cette déficience surgit l'idée de la présente idéographie. »

L'idéographie est une langue auxiliaire spécialement conçue afin que puissent être exprimés en toute clarté les rapports logiques entre des contenus de pensée. Pour expliquer la fonction de cette « écriture conceptuelle » (tel est le sens littéral du mot allemand *Begriffsschrift*), Frege a recours à l'analogie suivante : l'idéographie est à la langue courante ce que le microscope est à l'œil. Si celui-ci, comme la langue usuelle, est mieux adapté aux multiples circonstances de la vie, le microscope, comme l'idéographie, répond beaucoup mieux aux exigences scientifiques de précision et de distinction. Car Frege conçoit son idéographie comme une langue de la pensée pure qui fait usage de signes et de formules spécifiques et qui, à la différence des langues naturelles, ne laisse place à aucune sorte d'ambiguïté.

Pour atteindre un tel but, réaliser un tel instrument pour l'expression de la pensée pure, il ne suffisait pas de rendre parfaitement explicites toutes les propositions présupposées et de veiller à ce que les chaînes déductives ne comportent aucune lacune. Il fallait également préciser quels étaient les signes dont la langue de la pensée autorisait l'usage et comment ces signes devaient être compris et utilisés ; quels étaient, en quelque sorte, les éléments de la pensée pure ; et il fallait proposer, corrélativement, une méthode d'analyse logique des propositions qui permette d'en révéler la structure, non en suivant les suggestions de la grammaire, qui sont toujours relatives à une langue particulière, mais dans le respect du sens de ces propositions et en montrant quelles sont leurs parties les plus élémentaires du point de vue logique.

C'est peut-être sur ce dernier point que les réformes introduites par Frege sont les plus radicales si on les mesure à l'échelle de cette histoire de la logique qu'on fait habituellement remonter à l'Antiquité. Car l'ancienne analyse d'Aristote, selon laquelle toute proposition déclarative (par exemple « Socrate est musicien ») peut être analysée comme l'attribution d'un prédicat (musicien) à un sujet (Socrate) avait traversé les siècles. À cette analyse particulièrement peu éclairante pour les énoncés et les raisonnements mathématiques, Frege substitue une méthode dans laquelle il distingue deux catégories logiques fondamentales : les *objets* (par exemple Socrate, le fleuve qui traverse l'Égypte, la Seine) et les *fonctions*, dont l'expression linguistique (« ... est musicien », « ... est plus long que... », etc.) comporte une ou plusieurs places vides. C'est en comblant ces places vides qu'on forme des propositions déclaratives. « Socrate est musicien » peut alors être vu comme l'application de la fonction « ... est musicien » à l'argument « Socrate », et « le fleuve qui traverse l'Égypte est plus long que la Seine » comme l'application de la fonction « ... est plus long que... » aux deux arguments « le fleuve qui traverse l'Égypte » et « la Seine », pris dans cet ordre.

Contrairement à la logique traditionnelle issue d'Aristote, cette méthode d'analyse logique a recours aux relations (qu'elles soient à une place, comme « ... est musicien », à deux places, comme « ... est plus long que... », ou davantage comme dans « ... préfère... à... ») et s'avère, de ce fait, beaucoup mieux adaptée à la formalisation des raisonnements mathématiques. Mais surtout, combinée aux autres réformes introduites par l'*Idéographie*, elle produit une écriture conceptuelle dont la puissance d'expression est très supérieure à tout ce qui avait pu être proposé en logique jusqu'alors[1]. Parmi les apports majeurs que la logique contemporaine a retenus de l'œuvre de Frege, on compte notamment l'usage de symboles spécifiques pour représenter séparément des mots comme « tous », « aucun », « certains », « quelques » qui expriment une quantification dans « tous les hommes », « aucun nombre », « certains végétaux », etc., et l'analyse d'expressions comme « et », « ou », « ne... pas », « si... alors », qui sont nommés « connecteurs propositionnels » et qui permettent, à partir d'une ou plusieurs propositions, d'en former d'autres plus complexes.

Ces éléments rendent possible une analyse tout à fait nouvelle d'un énoncé comme « tous les hommes sont mortels ». Du point de vue de la logique moderne, cet

1. Il existe cependant des tentatives de réhabilitation de la logique traditionnelle. Cf. F. Sommers, *The Logic of Natural Language*, Oxford, Oxford UP, 1982 ; G. Englebretsen (éd.), *The New Syllogistic*, New York, Peter Lang, 1987 ; D. Oderberg (éd.), *The Old New Logic*, Cambridge, Mass., The MIT Press, 2005.

énoncé comporte deux expressions qui représentent des relations à une place (« ... est un homme » et « ... est mortel ») et qui sont liées l'une à l'autre par le connecteur propositionnel « si... alors ». Le tout est quantifié universellement. Pour exprimer cela en une formule concise, on utilise aujourd'hui les notations suivantes (qui ne sont pas celles qu'utilisait Frege) :

x est une variable, qui permet de désigner un individu quelconque.

H est un symbole de relation à une place et Hx se lit « x est un homme ».

M est un symbole de relation à une place et Mx se lit « x est mortel ».

$\rightarrow$ est un connecteur propositionnel qui peut se lire, ici, « si... alors ».

$\forall$ est le quantificateur universel et $\forall x$ se lit « pour tout individu x ».

« Tous les hommes sont mortels » est analysé ainsi : pour tout individu x, si x est un homme, alors x est mortel, ce qui s'exprime par la formule $\forall x(Hx \rightarrow Mx)$.

Nous sommes alors très loin de l'analyse traditionnelle selon laquelle le prédicat « mortel » est attribué au sujet « homme » pris universellement (tous les hommes). Selon cette analyse traditionnelle, il n'y a pas de différence fondamentale entre « tous les hommes sont mortels » et « Jean est plus âgé que René ». Dans les deux cas, un prédicat (« être mortel » ou « être plus âgé que René ») est attribué au sujet (« tous les hommes » ou « Jean »). Du point de vue de la logique moderne, au contraire, Jean et René sont deux individus, qu'on peut représenter, respectivement, par les signes a et b, et il existe entre eux une relation, « ... être plus âgé que... » qu'on peut noter R. La différence de structure logique entre les deux énoncés, que la logique traditionnelle n'aperçoit pas, apparaît clairement lorsqu'on compare les deux formules : $\forall x(Hx \rightarrow Mx)$ et Rab.

La possibilité d'utiliser des symboles de relations à une place, deux places, ou davantage, combinée à l'usage de la quantification universelle, donne à l'idéographie de Frege une capacité d'expression qui est d'autant plus forte que cette quantification peut porter non seulement sur des individus ou des objets (« $\forall x$ » : « pour tout individu x », « pour tout objet x ») mais également sur des fonctions (« $\forall f$ » : « pour toute fonction f ») ou sur des propositions (« $\forall p$ » : « pour toute proposition p »). En outre, la logique moderne offre la possibilité d'une quantification multiple (pour tous individus x et y) et alternée (pour tout individu x, il existe un individu y tel que, pour tout z...), très importante pour la formalisation des mathématiques et néanmoins ignorée par la logique traditionnelle.

Pourtant, en dépit de l'immense apport de la logique de Frege à la logique contemporaine, l'une ne saurait être identifiée à l'autre. Une différence essentielle apparaît par exemple dans l'interprétation des quantificateurs. Pour Frege, la quantification universelle (pour tout x) ou existentielle (il existe un x tel que) ne s'applique pas à des domaines particuliers d'individus (les nombres, les hommes, etc.) ou à des ensembles de fonctions ou de propositions qu'il serait possible de diversifier à l'envi ; elle s'applique à la *totalité* des individus – c'est-à-dire des objets – de l'univers, ou, respectivement, à la *totalité* des fonctions ou des propositions. Dans la logique telle que la conçoit Frege, l'univers du discours est unique : il s'agit de l'univers lui-même, la totalité de ce qui est, même si cet univers est hiérarchisé, divisé en niveaux, puisque la totalité des objets est clairement distinguée de la totalité des fonctions qui s'appliquent aux objets, elle-même différente de la totalité des fonctions de fonctions d'objets.

Une loi logique comme la loi du tiers exclu : « *p* ou non-*p* » (où *p* représente un énoncé quelconque) n'est donc pas conçue par Frege comme une expression formelle, vide de contenu, dans laquelle le symbole *p* pourrait être interprété dans différents domaines et qui aurait comme propriété d'être vraie pour chacune des interprétations possibles. Il ne s'agit pas non plus d'un schéma dans lequel *p* attendrait d'être remplacé par une véritable proposition. Les lois logiques sont bien plutôt des énoncés généraux et vrais qui ne sont pas moins descriptifs que des énoncés de la physique. La différence avec ces derniers tient uniquement à ce que les lois de la logique ont un caractère plus général : elles valent pour toute chose – tout individu, toute fonction, toute proposition – et ne contiennent aucun signe spécifique à un domaine particulier. Il n'y a pas non plus, pour Frege, plusieurs langages logiques, tel langage étant préféré à tel autre selon le domaine d'interprétation visé. Il n'y a qu'un seul langage, et dans ce langage s'expriment les lois de la logique.

On nomme « universalisme logique » une telle conception de la logique selon laquelle il n'y a pas plusieurs langages mais un seul et selon laquelle les énoncés n'ont pas à être interprétés dans des domaines d'objets différents mais portent simplement sur l'univers tout entier, sur tout ce qui est[1]. De ce point de vue, si les énoncés

1. Sur l'universalisme logique, cf. J. van Heijenoort, « Logic as Calculus and Logic as Language » (1967), *in* J. van Heijenoort, *Selected Essays*, Naples, Bibliopolis, 1985.

logiques ont un caractère de généralité, ce n'est pas non plus parce qu'ils seraient des formes de jugements obtenues par abstraction à partir de jugements particuliers, comme on pourrait obtenir « *p* ou non-*p* » à partir de « il pleut ou il ne pleut pas ». L'idéographie de Frege n'a pas pour but d'exprimer des rapports obtenus par abstraction à partir du concret, elle est au contraire conçue pour rendre possible l'expression des rapports logiques entre des contenus de pensée. « Je n'ai pas voulu donner en formules une logique abstraite, écrit Frege, mais donner l'expression d'un contenu au moyen de signes écrits, et d'une manière plus précise et plus claire au regard que cela n'est possible au moyen des mots. »[1]

Déduire l'arithmétique de la logique, ce n'est donc pas, pour Frege, la déduire de la logique traditionnelle, d'une logique dont les méthodes d'analyse et les concepts fondamentaux (sujet, prédicat, universel, particulier, etc.) étaient déjà connus ou avaient déjà été codifiés dans le passé, ni de ce que Kant entendait par « logique générale pure ». La logique dont Frege entendait déduire l'arithmétique demandait à être entièrement redéfinie. Comme le remarque Coffa, si l'on veut définir le logicisme comme le programme d'une déduction de l'arithmétique – ou, dans le cas de Russell, des mathématiques tout entières – à partir de la logique, il faut bien comprendre qu'à cette époque, « les mathématiques étaient une réalité et la logique un projet »[2].

1. G. Frege, « Sur le but de l'idéographie, 1882-1883 », trad. franç. *in* Frege, *Écrits logiques et philosophiques*, Paris, Le Seuil, 1971, p. 70-71.
2. A. Coffa, *The Semantic Tradition from Kant to Carnap*, Cambridge, Cambridge UP, 1991, p. 113.

On conçoit, dès lors, quel type d'objections pouvait être adressé au programme logiciste de Frege ou de Russell. Car si la « réduction » de l'arithmétique à la logique est obtenue au prix d'un enrichissement de la logique, on peut se demander quel est le gain espéré. Poincaré pose la question en ces termes : « On voit combien la nouvelle logique est plus riche que la logique classique ; les symboles se sont multipliés et permettent des combinaisons variées *qui ne sont plus en nombre limité*. A-t-on le droit de donner cette extension au sens du mot *logique* ? »[1] À cette question, qui fit l'objet d'un débat avec Russell, Poincaré donne une réponse négative.

Quoi qu'on puisse penser de la remarque de Poincaré, il importe, pour que le programme logiciste ait un sens, que les énoncés de la logique puissent être distingués des autres par quelque caractéristique particulière. On peut dire d'eux, par exemple, que la reconnaissance de leur vérité ne nécessite aucun recours à l'intuition sensible, ou encore qu'ils sont les énoncés les plus généraux dans la mesure où ils portent sur la totalité des choses, sans aucune restriction. Frege ne cherche cependant pas à donner une caractérisation de l'ensemble des vérités logiques comme le feront d'autres logiciens après lui. Son but est de montrer que les théorèmes de l'arithmétique peuvent être démontrés sur la seule base de lois qui devraient être reconnues comme relevant de la logique, par des preuves dépourvues de lacune.

1. H. Poincaré, « Les logiques nouvelles », in *Science et méthode*, 1908, Paris, Kimé, 1999, p. 140.

L'une des questions cruciales était cependant de savoir jusqu'à quel point la logique pouvait être enrichie, et en particulier si la logique pure était capable, par elle-même, de nous donner des objets. Frege pensait que c'était effectivement le cas puisque selon la loi V des *Lois fondamentales de l'arithmétique*, toute propriété permet de définir l'ensemble des objets qui satisfont cette propriété, exactement comme la propriété d'être bleu permet de définir l'ensemble des choses bleues et de poser ainsi l'existence de cet ensemble. Une telle supposition avait paru acceptable à Frege bien qu'il ait reconnu que les logiciens puissent avoir quelques scrupules à l'admettre au nombre de leurs principes.

En 1902, Russell adressa à Frege une lettre dans laquelle il montrait que la loi V conduisait de manière très directe à une contradiction, aujourd'hui appelée « paradoxe de Russell ». Considérons en effet la propriété, pour un ensemble, d'être élément de soi-même. Il existe des ensembles qui satisfont cette propriété, par exemple l'ensemble de tous les ensembles à plus de trois éléments, lequel a lui-même plus de trois éléments. Notons *P* la propriété, pour un ensemble, de ne pas être élément de soi-même. Considérons maintenant l'ensemble des ensembles qui satisfont *P*. Cet ensemble est-il élément de soi-même ? S'il l'est, alors il ne l'est pas (car il satisfait *P*) et s'il ne l'est pas, alors il l'est (car il ne satisfait pas *P*). Supposer l'existence d'un tel ensemble conduit donc à une contradiction, et le système de Frege, selon lequel toute propriété permet de définir l'ensemble des objets qui satisfont cette propriété, est contradictoire.

Il est difficile d'imaginer un sort plus tragique pour un système qui visait l'expression des vérités logiques, d'autant que son auteur avait consacré à la constitution de ce système une grande partie de sa vie. Frege crut pendant un certain temps que la difficulté pourrait être surmontée dans le respect des principes de son projet philosophique, mais il finit par renoncer au programme logiciste. Poincaré, grand critique des nouvelles logiques de l'époque, avait beau jeu d'ironiser : « La logistique n'est plus stérile, elle engendre l'antinomie ! »

Le programme logiciste ne fut pas abandonné pour autant. Après avoir pris connaissance des travaux de Peano, Russell conçut – avant même de connaître ceux de Frege – le projet d'une déduction des mathématiques dans la logique. Peano avait isolé un petit nombre de concepts qui semblaient avoir un très grand pouvoir d'expression et il avait imaginé pour ces concepts une notation particulièrement appropriée dont Russell s'inspira. Afin d'éviter la contradiction qui affectait le système de Frege, Russell introduisit la théorie des types, méthode de hiérarchisation des ensembles et des relations qui interdisait l'écriture d'une expression comme « x appartient à x ». Selon cette théorie, pour qu'on puisse écrire que x appartient à y, il faut que y soit d'un type immédiatement supérieur à celui de x dans la hiérarchie. Entre 1910 et 1913, Russell et Whitehead publièrent trois volumes d'une vaste entreprise de reconstruction des mathématiques sur une base logique. Ces trois volumes, intitulés *Principia Mathematica*, furent considérés, jusqu'au début des années 1930, comme le principal ouvrage de référence de la logique mathématique.

Chez Russell comme chez Frege, le programme logiciste fut conçu sans qu'une caractérisation préalable de l'ensemble des vérités logiques ait été proposée. La logique ressemblait plutôt à une idée directrice qui devait permettre d'orienter les recherches sur la question de savoir de quels éléments les mathématiques pouvaient être déduites, même si cette idée était suffisamment précise pour que Russell puisse éprouver certains scrupules à compter au nombre des vérités logiques des énoncés comme l'axiome du choix ou l'axiome de l'infini. L'axiome du choix dit que pour toute collection d'ensembles non vides et disjoints, même infinie, il existe un ensemble qui contient – qui « choisit », pour ainsi dire – exactement un élément de chacun des ensembles de la collection. Mais doit-on considérer comme « logique » un axiome qui pose l'existence d'un ensemble sans qu'il soit possible, en général, de donner aucune propriété qui caractérise ses éléments ni aucune loi selon laquelle ces derniers sont engendrés ? L'axiome de l'infini, quant à lui, dit qu'il existe un ensemble qui contient une infinité d'éléments. Si l'on suppose qu'une vérité, pour être logique, doit être parfaitement évidente, ou pouvoir être rendue parfaitement évidente par une démonstration dont toutes les étapes sont parfaitement évidentes, il est difficile de concevoir les deux axiomes cités comme des vérités logiques. Lorsque Russell s'aperçut qu'un grand nombre de preuves mathématiques exigeaient d'avoir recours à de tels axiomes, il comprit que le projet logiciste tel qu'il l'avait initialement conçu devait être révisé.

Le logicisme donne l'exemple d'un projet philosophique qui appelle une distinction entre des vérités qu'on s'accorde à reconnaître comme « logiques » et d'autres vérités qui sont, pour ainsi dire, « extralogiques ». L'histoire de ce projet comporte autant d'étapes que de réponses à la question : quels sens acceptables – c'est-à-dire suffisamment proche du sens usuel – peut-on donner aux mots « déduire », « logique » et « mathématique » en sorte qu'on puisse réussir à montrer que les mathématiques (ou, dans le cas de Frege, l'arithmétique) se déduisent de la logique ou, pourrait-on dire, *dans* la logique ?

Après les travaux de Frege et de Russell, cette histoire a montré qu'il fallait être prêt à faire de sérieuses concessions sur l'acceptabilité en question pour espérer satisfaire un programme qu'on puisse à juste titre qualifier de logiciste. On trouve néanmoins des travaux contemporains qui s'inscrivent explicitement dans la lignée de Frege ; tel est le cas, par exemple, du programme néologiciste de Bob Hale et Crispin Wright[1].

1. B. Hale et C. Wright, *The Reason's Proper Study*, Oxford, Oxford UP, 2001. Cf. aussi B. Linsky et E. N. Zalta, « What Is Neologicism? », *The Bulletin of Symbolic Logic*, 12, 2006, p. 60-99.

Chapitre III

LA LOGIQUE MATHÉMATIQUE

Les origines de la logique moderne ne se réduisent pas aux programmes logicistes de Frege et de Russell. Deux autres contributions décisives pour la formation de la logique contemporaine méritent d'être indiquées, fût-ce de manière sommaire. L'une et l'autre sont liées aux mathématiques, bien qu'à des titres différents. Il s'agit, d'une part, d'un ensemble de travaux logiques qu'on a coutume de réunir sous le nom d'« algèbre de la logique » et dont l'un des premiers représentants fut George Boole, et, d'autre part, d'une philosophie des mathématiques dite « formaliste », surtout connue par les travaux logiques de David Hilbert et de son école.

L'idée maîtresse de Boole fut d'utiliser les techniques mathématiques de l'algèbre pour créer un calcul logique destiné à résoudre des problèmes de déduction ou de décision : tel énoncé est-il déductible de tel ensemble de prémisses ? Quelles conséquences est-il possible de déduire d'un ou plusieurs énoncés donnés ? Boole recherchait les lois qui sont au principe du raisonnement en général, et il pensait pouvoir les formuler au moyen d'une généralisation de l'algèbre à des questions qui ne portaient plus uniquement sur le nombre ou sur la quantité[1].

1. G. Boole, *Les Lois de la pensée*, 1854, trad. franç., Paris, Vrin, 1992.

Ce projet n'a pas de rapport direct avec le problème de la formalisation des preuves dans un langage et les représentants de l'algèbre de la logique, comme Peirce ou Schröder, travaillaient généralement dans un esprit très différent de celui de Frege. Aussi soulevèrent-ils des problèmes d'un autre genre que ceux qui préoccupaient les logicistes. Partant d'une équation mathématique, ils s'interrogeaient par exemple sur les domaines dans lesquels cette équation pouvait être satisfaite, c'est-à-dire rendue vraie ; ou, étant donné une structure algébrique particulière – un ensemble muni d'opérations sur ses éléments –, ils cherchaient à déterminer quels étaient les axiomes qui permettaient de caractériser cette structure. C'est dans l'esprit de ce genre de recherche que Löwenheim démontra en 1915 un résultat qui, complété par Skolem en 1920, aura une importance capitale pour l'histoire ultérieure de la logique. Ce qu'on appelle aujourd'hui « théorème de Löwenheim-Skolem » dit que pour les langages dans lesquels la quantification porte uniquement sur les *individus* d'un domaine (par opposition à des *ensembles* de tels individus ou aux *fonctions* définies sur ce domaine), toute formule qui peut en général être satisfaite dans un domaine infini peut l'être dans un domaine dénombrable, c'est-à-dire dans un domaine qui ne contient pas plus d'individus qu'il n'y a de nombres entiers naturels.

Pour éclaircir le sens et expliquer l'importance de ce théorème, il convient de rappeler quelques idées de base de l'arithmétique des nombres infinis. Cantor avait montré que tous les ensembles infinis n'ont pas le même nombre d'éléments – le même nombre cardinal – et que les nombres cardinaux infinis peuvent être ordonnés. Par exemple, les nombres réels sont beaucoup plus nombreux que les nombres entiers, ce qui n'est pas le cas des nombres rationnels, ou fractions d'entiers. Que deux ensembles infinis aient le même nombre cardinal, cela signifie qu'il est possible d'apparier les éléments de ces deux ensembles par une bijection, c'est-à-dire par une relation qui à tout élément de l'un associe un unique élément de l'autre et inversement. Cantor montra qu'il existe une bijection entre les entiers naturels et les rationnels alors qu'il n'existe aucune bijection entre les entiers naturels et les réels. Les rationnels peuvent donc être appariés aux entiers, alors que les réels sont beaucoup trop nombreux pour cela. En outre, l'ordre croissant des nombres cardinaux infinis admet un plus petit élément – le nombre des nombres entiers – mais aucun plus grand élément. Un ensemble qui a autant d'éléments que l'ensemble des nombres entiers est dit « dénombrable ».

Ces distinctions étant établies, se pose la question de savoir s'il est possible de caractériser par une formule – ou par un ensemble de formules – le nombre cardinal d'un domaine d'individus. Existe-t-il, par exemple, une formule qui soit satisfaite uniquement dans les domaines qui ont autant d'individus que l'ensemble des nombres réels ? Le théorème de Löwenheim-Skolem apporte à cette question une réponse négative pour la logique du « premier ordre », ainsi nommée parce qu'elle a recours à des langages dans lesquels la quantification porte uniquement sur les *individus* du domaine, et non sur les *ensembles* formés de tels individus ou sur les *fonctions* définies dans ce domaine. Ce théorème affirme donc qu'il existe des limites à ce qu'il est possible d'exprimer dans un langage de cette nature : non seulement aucune formule, ou ensemble de formules, ne permet de caractériser un cardinal infini, mais encore, les objets qu'on cherche à caractériser à l'aide d'un système d'axiomes ne peuvent pas être distingués des nombres entiers. Le théorème de Löwenheim-Skolem dit en effet également que toute formule qui peut être satisfaite dans un domaine d'individus infini peut l'être aussi dans le domaine des entiers naturels. L'une des conséquences, par exemple, est qu'aucune formule – ni aucun ensemble de formules – d'un langage pour la logique du premier ordre ne permet de caractériser l'ensemble des nombres réels. (Il y a en fait deux théorèmes de Löwenheim-Skolem, l'un dit « montant », l'autre dit « descendant » ; c'est le second qui est mentionné ici, le premier le sera ci-dessous, p. 46.)

Historiquement, ce n'est certes pas en ces termes que Löwenheim et Skolem énoncèrent les résultats qu'ils avaient obtenus, respectivement, en 1915 et 1920. Mais tel est bien le sens de ce que la logique contemporaine énonce sous le nom de « théorème de Löwenheim-Skolem ». L'importance de ce résultat vient surtout de ce qu'il met en évidence certaines limites de la logique du premier ordre alors que, par ailleurs, d'autres théorèmes donnent de sérieux arguments en faveur de la thèse selon laquelle le cœur de la logique se trouve précisément dans cette théorie de la quantification du premier ordre et non dans les systèmes qui autorisent une quantification d'ordre supérieur. L'un des enjeux est de savoir si le sens de concepts mathématiques comme « fini », « infini », « nombre réel », « nombre entier naturel », etc., peut ou non être adéquatement exprimé dans un système formel et, le cas échéant, quel type de système est requis à cette fin. Plus généralement, la question est de savoir si le sens du discours mathématique peut, en général, être entièrement représenté dans un langage formalisé.

L'identification du cœur de la logique à la logique du premier ordre est une thèse fréquemment défendue. La compréhension des arguments qui justifient une telle opinion sera rendue plus aisée par quelques remarques préalables touchant l'autre courant de pensée qui contribua de manière décisive à la constitution de la logique mathématique contemporaine. Il s'agit du courant formaliste, dont le représentant le plus illustre est le mathématicien David Hilbert.

Comme Frege, Hilbert cherchait à résoudre la question du fondement des mathématiques. Mais l'orientation philosophique de son travail était très différente de celle de Frege ou de Russell. S'il admirait le travail d'axiomatisation accompli dans les *Principia Mathematica* par Russell et Whitehead, il adoptait sur cette tentative de reconstruction des mathématiques sur une base logique un point de vue qui n'avait jamais été celui des auteurs de l'ouvrage. Pour Hilbert, les multiples antinomies apparues dans les discours touchant l'infini (paradoxes de Russell, de Richard, de Burali-Forti, etc.) rendaient urgente une preuve de non-contradiction de l'analyse, et la formalisation des mathématiques pouvait servir de moyen en vue de cette fin. L'axiomatisation et la formalisation des mathématiques offraient en effet la possibilité de considérer les énoncés et les preuves de cette science comme des suites de signes et de les analyser en tant que telles, c'est-à-dire en faisant abstraction de leur signification et des interprétations possibles de ces suites. L'idée de Hilbert était d'adopter un point de vue métamathématique afin de pouvoir étudier la syntaxe des preuves formalisées et prouver qu'aucune contradiction du type « $0 = 1$ et $0 \neq 1$ » ne pouvait être déduite des axiomes de l'arithmétique, de l'analyse ou de la théorie des ensembles ; autrement dit : pour prouver que ces théories étaient non contradictoires. L'entreprise de formalisation acquérait par là un sens très différent de celui que lui avait donné Frege. Il ne s'agissait plus seulement de réécrire les preuves dans une langue logique en ayant soin de n'autoriser aucune lacune dans la chaîne démonstrative ; il s'agissait également de considérer ces preuves formelles comme des objets non interprétés – des suites de signes sans signification – et d'en proposer une analyse métathéorique afin d'obtenir une preuve de cohérence. La vérité des mathématiques n'était plus admise comme allant de soi[1].

Dans un système formel compris en ce sens non frégéen, les formules sont des suites de signes dont on ne considère pas la signification, et les preuves formelles sont des suites de formules. Les règles qui déterminent la progression dans le cours d'une preuve formelle, donc le passage d'une ou plusieurs formules à une formule ultérieure, sont purement syntaxiques et peuvent être appliquées mécaniquement, sans considération de la signification des formules. Pour souligner la particularité de ce point de vue, on peut alors parler de la *dérivation formelle* d'une formule ψ à partir des formules $\phi_1, \phi_2, \ldots, \phi_n$ plutôt que de la *déduction* de ψ à partir

1. Cf. les articles de D. Hilbert traduits *in* J. Largeault (éd.), *Intuitionisme et théorie de la démonstration*, Paris, Vrin, 1992, ainsi que D. Hilbert, « Sur l'infini », 1926, tr. fr. *in* J. Largeault (éd.), *Logique mathématique. Textes*, Paris, A. Colin, 1972.

de ϕ_1, ϕ_2, …, ϕ_n. L'idée de Hilbert était de développer ce qu'il nommait « théorie de la démonstration » afin de prouver, au niveau du métalangage, la cohérence des systèmes formels destinés à formaliser l'arithmétique et l'analyse.

On peut évidemment s'interroger sur l'intérêt d'une telle méthode si les raisonnements destinés à prouver la cohérence d'une théorie mathématique formalisée s'effectuent dans un métalangage qui n'est pas lui-même formalisé. Pour répondre à cette objection, Hilbert distingua deux sortes de raisonnements et de principes mathématiques. Les mathématiques dites « réelles » ne font usage que de méthodes jugées suffisamment élémentaires pour pouvoir être admises sans être justifiées. Les mathématiques dites « idéales », en revanche, ont recours à des méthodes abstraites qui appellent une justification. Moyennant quoi, Hilbert exige qu'une preuve de cohérence ne fasse usage, dans le métalangage, que des mathématiques « réelles », et donc de méthodes qu'il juge suffisamment sûres et qu'il qualifie de « finitaires ». Ce qui justifie une telle restriction des moyens démonstratifs utilisés dans la métamathématique, c'est que le processus de formalisation avait précisément pour fonction de réduire les preuves mathématiques à des suites de suites de signes qu'on puisse considérer comme dépourvues de signification, en sorte que le raisonnement métamathématique ne porte que sur des objets finis, concrets, donnés immédiatement à l'intuition.

Le programme de Hilbert sur le fondement des mathématiques comportait un second aspect. Il s'agissait d'établir que tout théorème des mathématiques « réelles » pouvait être démontré en ne faisant appel qu'à des méthodes élémentaires, et à montrer que si, de fait, les mathématiciens utilisaient des méthodes hautement abstraites pour démontrer des énoncés concrets de la théorie des nombres, un tel détour s'expliquait par le fait que ces méthodes permettaient d'abréger considérablement les preuves. Le programme de Hilbert visait donc à montrer que le recours à des méthodes abstraites pour la démonstration de théorèmes dont l'énoncé ne contient que des concepts concrets n'est, en droit, jamais indispensable, même si l'élimination de ces détours est souvent, de fait, irréalisable parce qu'elle aurait comme effet d'engendrer des preuves si longues qu'elles seraient, en pratique, impossibles à écrire.

La théorie de la démonstration pouvait également être utilisée pour l'étude d'autres propriétés des systèmes formels que la propriété de cohérence. Pour le problème du fondement des mathématiques, il importait de savoir si un système formel destiné à formaliser, par exemple, l'arithmétique ne laissait pas ouvertes certaines questions que le langage de cette théorie permettait de formuler. Supposons qu'on dispose d'un système formel dans lequel les énoncés usuels de l'arithmétique sont représentés par des formules. Est-ce que pour toute formule ϕ qui représente formellement un tel énoncé dans ce système, soit ϕ soit la négation de ϕ (non-ϕ) est formellement dérivable des axiomes du système, c'est-à-dire des formules qui représentent formellement les axiomes de l'arithmétique ? Si la réponse est affirmative, on dit que le système est syntaxiquement complet ; si la réponse est négative, le système est syntaxiquement incomplet.

Les logiciens de l'école formaliste estimaient que le problème du fondement de l'arithmétique serait résolu si, d'une part, on disposait d'un système formel permettant de représenter formellement tous les énoncés de l'arithmétique et si, d'autre part, on pouvait prouver par les moyens de la théorie de la démonstration que ce système est cohérent et syntaxiquement complet. Au cours des années 1920, des logiciens comme Ackermann, von Neumann ou Herbrand avancèrent dans la réalisation d'un tel programme et jusqu'au début des années 1930, Hilbert et les logiciens formalistes purent croire au succès prochain de leur entreprise. Cet optimisme fut sérieusement mis à mal par la démonstration des célèbres théorèmes d'incomplétude de Gödel publiés en 1931. Le premier de ces théorèmes dit en effet que tout système formel cohérent qui permet de formaliser une certaine partie de l'arithmétique élémentaire est syntaxiquement incomplet. Autrement dit, dans un tel système, il existe au moins une formule qui représente formellement un énoncé de l'arithmétique telle que ni elle ni sa négation n'est formellement dérivable des axiomes. Les énoncés en question sont donc indécidables dans ce système. Le second théorème d'incomplétude dit que, sous les mêmes conditions, la cohérence du système ne peut pas être démontrée sans faire appel à d'autres principes démonstratifs que ceux qui peuvent être formalisés dans le système lui-même. Historiquement, la version des théorèmes d'incomplétude qui fut publiée en 1931 est plus faible que l'énoncé qu'on vient d'en donner. Mais des versions généralisées furent démontrées et publiées dans les années qui suivirent.

La publication de ces deux théorèmes ne mit cependant pas fin aux recherches en théorie de la démonstration : si les restrictions imposées par le programme de Hilbert sur les moyens démonstratifs auxquels le logicien pouvait légitimement avoir recours dans la métamathématique étaient trop fortes pour que la cohérence de l'arithmétique puisse être démontrée par ces moyens, il restait la possibilité d'étendre le point de vue finitiste qui avait été celui de Hilbert, avec l'espoir de rendre ainsi possible une démonstration de cohérence.

Comme le remarqua Gödel lui-même, le programme de Hilbert n'était donc pas réfuté, car une version libéralisée de ce programme était encore concevable, en donnant une interprétation du finitisme qui soit moins restrictive que celle de Hilbert. C'est dans cette voie que s'engagèrent, par exemple, les recherches de Gentzen au cours des années 1930. Dans un article de 1958, Gödel proposa lui aussi « une extension du point de vue finitiste non encore utilisée »[1] qui permettait d'apporter une nouvelle preuve de cohérence de l'arithmétique.

Tous les logiciens ne partageaient donc pas la conception hilbertienne du fondement des mathématiques et tous ne pensaient pas devoir respecter les restrictions que Hilbert avait imposées aux démonstrations métamathématiques. Gödel, notamment, ne voyait aucune raison de principe pour laquelle il faudrait se priver, pour des raisonnements *sur* les systèmes formels, de la puissance démonstrative des mathématiques abstraites non finitaires. C'est précisément cette distance à l'égard du finitisme de Hilbert qui lui permit non seulement de démontrer, mais également simplement d'énoncer, le théorème de complétude (au sens sémantique) qui fit l'objet de sa thèse d'habilitation en 1929 et dont la démonstration fut publiée en 1930. Outre la preuve du théorème, cette publication apportait une

1. Cf. Gödel, « Über eine noch nicht benützte Erweiterung des finiten Standpunkts », *Dialectica*, 12, 1958.

importante clarification touchant certains concepts dont les logiciens de l'époque n'avaient pas toujours bien saisi l'importance ou la signification. Tel est le cas de la différence entre logique du premier ordre et logique d'ordre supérieur. Avant 1930, la plupart des logiciens ne percevaient pas l'intérêt qu'il pouvait y avoir à limiter la quantification aux objets ou individus du domaine de discours et à ne pas l'étendre aux ensembles d'objets ou aux fonctions et relations définies sur ces objets. Les travaux de Gödel sur la complétude sémantique clarifient ce point, car cette propriété, démontrée pour la logique du premier ordre, ne vaut pas, en général, pour les systèmes logiques d'ordre supérieur.

Gödel apporte également une clarification des concepts de validité et de satisfaisabilité d'une formule, concepts dont dépend l'énoncé du théorème de complétude. La définition que nous en donnons maintenant est, en substance, celle qui figure généralement dans les ouvrages d'introduction à la logique. Soit $\mathcal{L}$ un langage pour la logique du premier ordre, langage non interprété qui comporte des symboles de relation : par exemple R, un symbole de relation à une place, et S, un symbole de relation à deux places. Une structure d'interprétation de ce langage est définie par un domaine d'individus (l'ensemble des êtres humains, l'ensemble des nombres entiers naturels, etc.) et par une interprétation de R et de S dans le domaine en question. Par exemple, si le domaine d'individus est l'ensemble des êtres humains, R peut être interprété par « ... est âgé de plus de 30 ans » et S par « ... est plus âgé que... » ; si le domaine d'individus est l'ensemble des entiers naturels, R peut être interprété par « ... est un nombre premier » et S par « ... est inférieur ou égal à... ». Pour chaque langage du premier ordre, il existe une infinité de structures d'interprétations possibles. Or si un énoncé déclaratif du langage ordinaire est, en général, vrai ou faux, les formules qui, dans $\mathcal{L}$, correspondent à ces énoncés ne sont, en tant que telles, ni vraies ni fausses tant que le langage n'a pas été interprété. Leur vérité ou fausseté n'a de sens que relativement à une structure d'interprétation. Par définition, une telle formule est *satisfaisable* s'il existe au moins une structure d'interprétation du langage dans laquelle elle est vraie ; elle est *valide* si elle est vraie dans toutes les structures d'interprétation du langage.

Avant d'énoncer le théorème de complétude sémantique, remarquons que les concepts de formule valide et de formule satisfaisable sont très loin d'être finitaires. Leur définition fait intervenir une quantification sur les structures d'interprétation du langage considéré (« ... dans *toutes* les structures d'interprétation du langage », « ... *il existe au moins une* structure d'interprétation du langage... ») et implique donc la considération d'une infinité d'ensembles qui peuvent être eux-mêmes infinis. Sur ce point, ils se distinguent nettement des concepts de cohérence syntaxique, de complétude syntaxique ou de dérivabilité formelle dont les définitions ne font intervenir que des suites finies de signes. Dans les années 1930, certains logiciens rejetèrent le point de vue sémantique au motif qu'il avait recours à des concepts qu'ils jugeaient trop obscurs.

Étant donné un langage pour la logique du premier ordre et une méthode de dérivation formelle soigneusement choisie, le théorème de complétude sémantique dit qu'*une formule est valide si, et seulement si, elle est formellement démontrable* (c'est-à-dire prouvable par la méthode de dérivation formelle choisie, sans présupposer aucune autre formule). Ou, selon une autre formulation du théorème, *un ensemble de formules est satisfaisable si, et seulement si, il est syntaxiquement cohérent.*

Ce théorème marque ainsi une claire distinction entre des concepts syntaxiques (dérivation formelle, cohérence syntaxique) et des concepts sémantiques (structures d'interprétation d'un langage, validité et satisfaisabilité d'une formule), en même temps qu'il établit un lien entre ces deux catégories de concepts : pour la logique du premier ordre, nous disposons de méthodes de dérivation formelle qui soient telles que toutes les formules formellement démontrables soient valides (c'est-à-dire vraies dans toutes les structures d'interprétation du langage) mais aussi et surtout, telles que toutes les formules valides soient formellement démontrables par ces méthodes (c'est cette seconde partie de l'énoncé qu'on appelle, à proprement parler, « théorème de complétude »).

Pour la logique du premier ordre, ce théorème circonscrit un ensemble de formules défini à la fois d'un point de vue sémantique (l'ensemble des formules valides) et d'un point de vue syntaxique (l'ensemble des formules formellement démontrables par une certaine méthode de dérivation formelle). Aussi le théorème de complétude a-t-il pu servir d'argument en faveur d'une double thèse : premièrement, l'ensemble des formules valides se présente comme un candidat naturel pour une caractérisation des vérités logiques ; deuxièmement, le premier ordre se présente comme un candidat naturel pour une délimitation du domaine de la logique parce que la propriété de complétude sémantique ne vaut pas, en général, pour la logique d'ordre supérieur.

Le théorème de complétude apparaît généralement comme l'un des résultats centraux des ouvrages d'introduction à la logique mathématique[1], dans lesquels sa démonstration est donnée après qu'un certain nombre de concepts de base et de théorèmes ont été exposés.

Typiquement, on commence par définir les langages pour la logique du premier ordre à partir d'un alphabet composé de deux ensembles de signes. Premièrement, les signes logiques, qui comprennent une infinité de variables d'individus, des connecteurs propositionnels (négation, conjonction, disjonction, connecteur de l'implication matérielle, etc.), les quantificateurs universel (pour tous) et existentiel (il existe au moins un), les parenthèses et la virgule, à quoi l'on ajoute souvent le signe d'égalité ; deuxièmement, les signes non logiques qui peuvent comprendre des constantes d'individus, des symboles de fonctions et des symboles de relations. L'ensemble des termes (c'est-à-dire des expressions qui ont vocation à désigner un individu d'un domaine) est défini inductivement par des règles formelles, de même que l'ensemble des formules. On distingue les formules closes, dans lesquelles toutes les variables sont liées par un quantificateur et qui correspondent aux énoncés déclaratifs de la langue usuelle (par exemple $\forall x\, \exists y Rxy$: pour tout individu x, il existe un individu y qui est dans la relation R avec x) et les formules ouvertes, dans lesquelles certaines variables ne sont pas quantifiées (par exemple : $\exists y Rxy$). Un langage formel non interprété, $\mathcal{L}$, est entièrement défini par l'ensemble des formules.

1. Parmi les classiques du genre : S. C. Kleene, *Logique mathématique*, 1967, trad. franç., Paris, A. Colin, 1971 ; H. B. Enderton, *A Mathematical Introduction to Logic*, New York, Academic Press, 1972, 2e éd., 2001.

Dans une exposition de la logique mathématique, la partie sémantique commence par introduire le concept de structure d'interprétation. Étant donné un langage $\mathcal{L}$, une structure d'interprétation de $\mathcal{L}$ se définit par un ensemble non vide, $\mathcal{D}$, qui forme le domaine d'individus ou l'univers du discours, et par une fonction d'interprétation qui associe à chaque constante d'individu de $\mathcal{L}$ un élément de $\mathcal{D}$, à chaque symbole de fonction de $\mathcal{L}$ une fonction définie sur $\mathcal{D}$ et à valeurs dans $\mathcal{D}$, et à chaque symbole de relation de $\mathcal{L}$ une relation définie sur $\mathcal{D}$. Moyennant quelques compléments techniques (notamment le concept d'assignation de valeur aux variables), on peut alors définir, pour une formule quelconque, « être satisfait par l'assignation α dans la structure d'interprétation $\mathcal{J}$ » et, dans le cas particulier des formules closes, « être vrai dans $\mathcal{J}$ ». Si une formule close ϕ est vraie dans $\mathcal{J}$, on dit que $\mathcal{J}$ est un *modèle* de ϕ (on dit aussi que $\mathcal{J}$ est modèle d'un ensemble Γ de formules closes si $\mathcal{J}$ est modèle de toutes les formules de Γ). On dit alors qu'une formule est *valide* si elle est satisfaite par toute assignation de valeur aux variables, dans toute structure d'interprétation du langage. Dans le cas particulier des formules closes, « valide » signifie simplement « vrai dans toutes les structures d'interprétation du langage ».

Avant de pouvoir énoncer le théorème de complétude, il faut encore définir une relation de dérivation formelle entre les formules, généralement à l'aide d'un ensemble d'axiomes (appelés « axiomes logiques ») et de règles d'inférence (comme la règle de détachement – ou *modus ponens* – selon laquelle si ϕ est dérivable et si $(\phi \rightarrow \psi)$ est dérivable, alors ψ est dérivable). Cette relation relève de la partie syntaxique de la logique et permet de définir l'expression « la formule ϕ est formellement dérivable de l'ensemble de formules Γ ». Ainsi se trouve constitué le système formel qu'on nomme souvent « calcul des prédicats », cadre méthodologique de la logique du premier ordre dans lequel peuvent être formalisées différentes théories scientifiques.

Imaginons par exemple qu'on veuille formaliser l'arithmétique. Il faut premièrement veiller à ce que le langage comporte des symboles non logiques adéquats comme une constante d'individu destinée à être interprétée par 0, un symbole de fonction unaire destiné à être interprété par la fonction successeur sur les entiers naturels, deux symboles de fonction binaires destinés à être interprétés, respectivement, par l'addition et la multiplication sur les entiers naturels, un symbole de relation à deux places destiné à être interprété par la relation « strictement inférieur à ». Notons $\mathcal{L}_A$ ce langage. Il faut, deuxièmement, représenter les axiomes de l'arithmétique usuelle – les axiomes de Peano – par des formules de $\mathcal{L}_A$. L'ensemble de ces formules constitue une théorie formelle de l'arithmétique qu'on notera T_A. L'un des axiomes de l'arithmétique, appelé « principe d'induction », est en réalité ce qu'on appelle un « schéma d'axiomes ». Il dit en effet que *pour toute propriété P*, l'énoncé suivant est un axiome : si 0 satisfait P et si, pour tout entier naturel n, si n satisfait P alors le successeur de n satisfait P, alors, tout entier naturel satisfait P. Comme la logique du premier ordre permet de quantifier sur les individus (ici : les nombres) mais non sur leurs propriétés, le principe d'induction doit être exprimé soit par un schéma d'axiomes, dans le métalangage, soit par un ensemble infini de formules du langage (une formule pour chaque propriété exprimable dans le langage), auquel cas T_A est un ensemble infini. Seule une logique d'ordre supérieur, qui autorise la quantification sur les propriétés, permet d'exprimer le principe d'induction en une formule unique : $\forall P((P(0) \wedge \forall n(P(n) \rightarrow P(succ(n)))) \rightarrow \forall nP(n))$.

Si l'on note $\mathbb{N}$ l'ensemble des entiers naturels et *succ* la fonction successeur sur $\mathbb{N}$, la structure $\mathcal{N} = (\mathbb{N}, 0, succ, +, \times, <)$ est une structure d'interprétation de $\mathcal{L}_A$ qui est un modèle de T_A. $\mathcal{N}$ est ce qu'on appelle le « modèle standard de T_A ».

Le premier théorème d'incomplétude syntaxique nous apprend, en outre, que si T_A est cohérent, il existe des formules de $\mathcal{L}_A$ qui ne sont ni démontrables (la formule n'est pas formellement dérivable de T_A) ni réfutables (la négation de la formule n'est pas formellement dérivable de T_A). Quant au second théorème d'incomplétude, il nous apprend qu'il est impossible d'établir la cohérence de T_A en utilisant uniquement T_A et les ressources démonstratives du système formel précédemment défini. Enfin, selon le théorème de Löwenheim-Skolem « montant », il existe des structures d'interprétation qui satisfont toutes les formules de T_A et dont le domaine d'individus est indénombrable, ce qui signifie que T_A (comme toute théorie définie par un ensemble d'axiomes formulés dans un langage pour la logique du premier ordre) échoue à caractériser le concept de nombre entier naturel.

Apparaissent alors clairement les différences radicales qui existent entre la conception de la logique qui vient d'être esquissée et celle que défendait Frege. Ici, on construit un ou des systèmes formels non interprétés – des calculs – spécialement conçus pour la formalisation de l'arithmétique ou d'une autre théorie ; chez

Frege, la logique est un langage universel interprété, écriture conceptuelle de la pensée pure en général. Ici, un privilège est accordé aux langages du premier ordre ; Frege ne voit aucune raison de séparer la partie du langage dans laquelle la quantification porte uniquement sur les individus. Ici, les axiomes de l'arithmétique sont soigneusement distingués des axiomes logiques qui servent à constituer le système formel lui-même ; chez Frege, les énoncés de l'arithmétique sont déduits des vérités logiques. Ici, la structure des entiers naturels est présupposée, elle sert à interpréter le langage de l'arithmétique ; chez Frege, cette structure doit être construite et déduite non seulement *à l'aide de* la logique, mais *dans* la logique. Ici, le langage peut être interprété dans une infinité de domaines d'individus différents ; chez Frege, il n'y a qu'un univers du discours, l'univers lui-même, et la signification des signes est invariable. Ici, l'universalité des formules logiques vient de ce qu'elles sont vraies dans toutes les interprétations du langage ; chez Frege, les vérités logiques sont des vérités sur le monde qui ont comme particularité d'être les vérités les plus générales qui soient.

Chapitre IV

VÉRITÉ, CONSÉQUENCE LOGIQUE, THÉORIE DES MODÈLES

Après les travaux de Gödel sur la complétude sémantique et l'incomplétude syntaxique, la distinction entre syntaxe et sémantique s'est imposée comme l'une des plus fondamentales en logique, alors qu'elle n'était pas clairement perçue comme telle auparavant. Dans son orientation syntaxique, la logique conduit à la théorie de la démonstration, à l'étude des propriétés des preuves et des systèmes de dérivation formelle ; dans son orientation sémantique, à la théorie des modèles, à l'étude des structures d'interprétation des langages, au concept d'énoncé vrai *dans une structure*, et à l'étude des relations entre des ensembles d'énoncés et les structures d'interprétation qui rendent ces énoncés vrais.

La formalisation de l'arithmétique telle que l'avait conçue Frege ne reposait nullement sur l'opposition entre syntaxe et sémantique. Cette opposition appelait en effet une distinction entre langage et métalangage que le point de vue universaliste excluait. Pour Frege, cette opposition n'était donc pas tant inutile que dépourvue de sens. Les axiomes du système frégéen devaient être vrais, chaque nouvelle étape d'une démonstration devait préserver la vérité des propositions précédentes et le système dans son ensemble devait permettre de prouver formellement tous les théorèmes de l'arithmétique. Contrairement à Frege, Hilbert adoptait ce qu'on allait nommer ultérieurement le « point de vue syntaxique », écartant le point de vue sémantique qui contredisait les exigences finitistes de la métamathématique. Il espérait pouvoir construire un système formel non interprété, adéquat pour l'arithmétique, et dont on puisse prouver qu'il était cohérent et complet. Les théorèmes d'incomplétude syntaxique mirent cependant en évidence que les concepts syntaxiques de preuve (« ... est démontrable ») et de dérivation formelle (« ... est formellement dérivable à partir de... ») ne pouvaient pas servir de substituts aux concepts sémantiques de vérité (« ... est vrai ») et, respectivement, de conséquence logique (« ... est conséquence logique de... »). Ces théorèmes montrent en effet que, dans tout système formel cohérent (dans lequel une partie de l'arithmétique est formalisable), il existe des formules indémontrables qui représentent pourtant des énoncés vrais. En ce sens précis, le concept sémantique de vérité n'a pas la même extension que le concept syntaxique de démontrabilité. Par ailleurs, une propriété des entiers naturels, exprimable par une formule $\phi[x]$ où x est une variable libre, peut être formellement démontrable pour n'importe quel entier n sans que $\forall x \phi[x]$ le soit, bien que d'un point de vue sémantique, si la variable x parcourt l'ensemble des entiers naturels, cette formule soit, logiquement, une conséquence de l'ensemble infini des $\phi[n]$. Au cours des années 1930, les concepts de preuve, de conséquence logique et de vérité firent l'objet d'un réexamen qui apporta de nouvelles lumières sur le sens de l'opposition entre syntaxe et sémantique.

Du côté de la théorie de la démonstration, de nouvelles méthodes de formalisation des preuves – la déduction naturelle et le calcul des séquents – furent introduites par Gentzen qui, en dépit du second théorème d'incomplétude, poursuivit ses recherches sur la cohérence de l'arithmétique. Par ailleurs, on considéra la possibilité d'introduire des règles d'un genre différent, peu conformes à l'orthodoxie alors en vigueur. Par exemple, la règle ω permet d'inférer $\forall n\phi[n]$ (« tout entier possède la propriété exprimée par la formule ϕ ») à partir d'une infinité de prémisses : $\phi[0]$, $\phi[1]$, $\phi[2]$, ..., $\phi[n]$... (la propriété exprimée par ϕ est vérifiée pour chaque entier n). Mais l'introduction de la règle ω modifie la notion même de système formel, initialement conçue dans le but de codifier les étapes d'un raisonnement. Nous ne sommes en effet jamais en position d'appliquer effectivement une règle qui comporte une infinité de prémisses ni de vérifier une « preuve » qui ferait usage d'une telle règle (bien qu'il soit parfaitement possible d'étudier les propriétés des systèmes formels auxquels cette règle appartient). Les systèmes formels, qui avaient pour vocation le traitement d'objets finis, intègrent la possibilité de preuves infinies lorsqu'on y introduit la règle ω.

Du côté de la sémantique, le logicien dont les contributions furent les plus décisives au cours des années 1930 est assurément Tarski. Ses travaux les plus connus touchent à la définition de la vérité ainsi qu'à l'analyse du concept de conséquence. Cette analyse, très éloignée des méthodes de formalisation des preuves et du concept de dérivation formelle, est à l'origine de la relation de « conséquence sémantique » ou de « conséquence logique » telle qu'elle est conçue aujourd'hui.

Une définition de cette relation requiert quelques remarques préalables sur la vérité et sur les conditions de possibilité d'une définition du prédicat « ... est vrai » pour les énoncés d'un langage.

Le concept de vérité divise les logiciens, qui le considèrent tantôt comme un concept central en logique, tantôt comme négligeable ou susceptible d'être éliminé du discours scientifique et des recherches logico-philosophiques. Les logiciens universalistes estimaient quant à eux qu'il était indéfinissable, toute relation entre langage et réalité étant par principe ineffable. D'autres logiciens s'en méfiaient en raison des contradictions auxquelles il conduisait aisément[1].

Parmi les paradoxes liés au concept de vérité, le plus célèbre est probablement celui du menteur. Si je dis « je suis en train de mentir », est-ce que je dis la vérité ? Si je dis la vérité, alors je suis en train de mentir, et je ne dis donc pas la vérité. Si je ne dis pas la vérité, alors je ne suis pas en train de mentir, et je dis donc la vérité. Lorsqu'on tente de répondre à la question précédente, qui semble parfaitement légitime, on se trouve ainsi pris dans une contradiction. Voici un second paradoxe, qui résulte des deux phrases suivantes :

(*a*) La phrase (*b*) est vraie.
(*b*) La phrase (*a*) est fausse.

On demande si la phrase (*a*) est vraie ou fausse. Si elle est vraie, alors la phrase (*b*) est vraie, et la phrase

1. Cf. S. Candlish et N. Damnjanovic, « A Brief History of Truth », *in* D. Jacquette (ed.), *Philosophy of Logic*, Amsterdam, North-Holland, 2006. Voir aussi les textes de H. Field et de S. Kripke traduits *in* D. Bonnay et M. Cozic (éds.), *Philosophie de la logique*, Paris, Vrin, 2009.

(*a*) est donc fausse. Si elle est fausse, alors la phrase (*b*) est fausse, et la phrase (*a*) n'est donc pas fausse, mais vraie. Ici encore, on se trouve pris dans une contradiction dès qu'on tente de répondre à la question précédente, qui semble pourtant parfaitement légitime.

Tarski montre que si ce genre de paradoxe peut apparaître dans les langues naturelles, c'est parce qu'elles autorisent un usage non restreint du prédicat « ... est vrai ». Il montre aussi que l'usage du prédicat de vérité n'a pas les mêmes conséquences paradoxales lorsque sa définition est limitée à des langages qui satisfont certaines conditions (notamment que les lois de la logique classique soient respectées : principe du tiers exclu, loi de la bivalence, etc.) et que son application obéit à certaines restrictions. Tarski donne une méthode de définition de ce prédicat pour les langages en question, mais il montre également – et ce résultat est souvent considéré comme l'un des plus importants de la logique contemporaine – que toute définition du prédicat « ... est vrai » pour un langage $\mathcal{L}$ requiert un langage dont le pouvoir d'expression est supérieur à celui du langage $\mathcal{L}$ lui-même[1]. Si « vrai dans $\mathcal{L}$ » est définissable, cette définition ne peut donc pas être donnée dans $\mathcal{L}$, mais uniquement dans un métalangage pour lequel la définition d'un prédicat de vérité requiert un troisième langage, et ainsi de suite. Dans les paradoxes précédents, l'exigence d'une distinction entre les énoncés d'un langage (dans lequel on peut exprimer une proposition *p*) et ceux d'un métalangage (dans lequel

1. A. Tarski, « Le concept de vérité dans les langages formalisés », 1933, trad. franç., *in* Tarski, *Logique, sémantique, métamathématique*, vol. I, Paris, Armand Colin, 1972.

on peut dire « "p" est vrai ») n'est effectivement pas respectée.

Ce que donne Tarski n'est pas une définition de la vérité en général, ni même de la vérité pour une certaine classe de langages. Étant donné $\mathcal{L}$ un langage déterminé qui doit satisfaire certains critères, il indique une méthode permettant de définir le concept d'énoncé vrai, relativement à $\mathcal{L}$. La méthode qui est généralement utilisée aujourd'hui diffère à plusieurs égards de celle que Tarski avait introduite en 1933. L'une et l'autre ont cependant comme point commun de reposer sur un concept auxiliaire : la satisfaction.

Étant donné $\mathcal{L}$ un langage non interprété, ϕ une formule de $\mathcal{L}$, $\mathcal{J}$ une structure d'interprétation de $\mathcal{L}$ dont le domaine d'individus est $\mathcal{D}$, et α une assignation de valeurs aux variables (fonction qui associe un élément de $\mathcal{D}$ à chaque variable d'individu), on définit inductivement « ϕ est satisfait par α dans $\mathcal{J}$ » puis, dans le cas particulier où ϕ est une formule close, « ϕ est vraie dans $\mathcal{J}$ ».

L'une des conditions pour que la définition d'un prédicat de vérité pour $\mathcal{L}$ soit recevable est qu'elle reflète de manière adéquate notre compréhension intuitive de ce prédicat. Or selon Tarski, cette exigence d'adéquation est satisfaite uniquement si pour tout énoncé p de $\mathcal{L}$, la condition « "p" est vrai si et seulement si c » peut être déduite de la définition, où c est une traduction de p dans le métalangage. Par exemple, si « *der Schnee ist weiß* » est un énoncé du langage et que le métalangage est la langue française, « "*der Schnee ist weiß*" est vrai si et seulement si la neige est blanche » doit pouvoir être déduit de la définition de « ... est vrai » pour le langage en question. Il est clair qu'on ne doit pas espérer tirer d'une définition tarskienne de la vérité un quelconque critère de vérité pour les énoncés. Telle n'est pas la vocation d'une définition de la vérité ainsi comprise. Cette définition a cependant le mérite de donner un sens précis à l'expression « la formule (close) ϕ est vraie dans la structure d'interprétation $\mathcal{J}$ », expression dont dépend le concept sémantique de conséquence qui est, lui aussi, dû à Tarski[1].

À quelles conditions un énoncé ϕ est-il conséquence logique d'un ensemble d'énoncés Γ ? Ce problème essentiel demeure, aujourd'hui encore, une question logique controversée[2]. Une réponse classique, qui remonte, en substance, aux *Premiers Analytiques* d'Aristote, donne la définition suivante : ϕ est conséquence logique de Γ si, les énoncés de Γ étant posés, ϕ ne peut pas ne pas être posé. En d'autres termes, si les énoncés de Γ sont vrais, ϕ est *nécessairement* vrai. Cette condition stipule, premièrement, que la relation de conséquence préserve la vérité (si les énoncés de Γ sont vrais, ϕ est vrai), mais elle ajoute, deuxièmement, que cette préservation a un caractère de nécessité. Les logiciens sont généralement d'accord sur le premier point, mais nombre d'entre eux se sont montrés beaucoup plus réticents à accepter le second, estimant que la signification du terme modal « nécessairement » était bien trop obscure pour qu'on puisse espérer en tirer une clarification de la relation de conséquence logique. On

1. Tarski, « Sur le concept de conséquence logique », 1936, trad. franç., *in* Tarski, *Logique, sémantique, métamathématique. 1923-1944*, vol. II, Paris, Armand Colin, 1974.
2. Cf. J. Etchemendy, *The Concept of Logical Consequence*, Stanford, Cal., CSLI, 1999.

verra que la définition tarskienne de la conséquence logique prend soin d'éviter tout recours aux concepts modaux.

Les logiciens s'accordent également sur le caractère formel de cette relation. Que l'énoncé A soit conséquence logique des énoncés B et C ne dépend pas de ce dont il est question dans ces énoncés, mais uniquement de leur forme. Considérons les deux raisonnements suivants : « Tous les nombres entiers ont un successeur. Or, 23 est un nombre entier. Donc 23 a un successeur » et « Toutes les planètes ont un satellite. Or, Vénus est une planète. Donc Vénus a un satellite. » Ils ont une forme commune qui peut être représentée ainsi : « Tous les P sont Q. Or a est P. Donc a est Q. » En outre, ils sont l'un et l'autre valides : on reconnaît aisément que le troisième énoncé est conséquence logique des deux précédents (que certains de ces énoncés soient faux n'y change rien). Considérons maintenant, plus généralement, un ensemble d'énoncés qu'on représente, dans un langage formel, par un ensemble Γ de formules closes, et un énoncé qu'on représente, dans le même langage, par une formule close ϕ. Typiquement, Γ représente l'ensemble des prémisses d'un raisonnement, et ϕ la conclusion. La définition tarskienne de la conséquence logique stipule que ϕ *est conséquence logique de* Γ *si, et seulement si, toute structure d'interprétation du langage dans laquelle les formules de* Γ *sont vraies donne aussi la valeur vraie à* ϕ. *En d'autres termes,* ϕ *est conséquence logique de* Γ *si, et seulement si, tout modèle de* Γ *est modèle de* ϕ.

Pour un langage donné, la vérité logique se définit comme le cas limite où l'ensemble Γ ne contient aucun élément : les vérités logiques sont les formules qui sont conséquences logiques de l'ensemble vide (ou, de manière équivalente, de tout ensemble de formules).

Par la définition de la conséquence logique, comme par celle de la dérivabilité formelle, on tente de caractériser en termes précis et scientifiquement acceptables la relation d'*inférence déductive logiquement valide*. Typiquement, cette relation lie les prémisses d'un raisonnement à sa conclusion, lorsque le raisonnement est correct. Les relations de conséquence logique et de dérivabilité formelle sont néanmoins conçues selon des points de vue fondamentalement différents. La première est sémantique, elle fait fond sur les relations entre les formules d'un langage et les structures d'interprétation de ce langage, qui sont infinies en nombre ; sa définition, abstraite et infinitaire, a recours à une quantification sur les structures d'interprétation (« tout modèle... ») et elle ne dit absolument rien des étapes successives par lesquelles une preuve pourrait conduire des prémisses à la conclusion. *A contrario*, la dérivabilité formelle est clairement syntaxique ; elle porte sur des objets finis : des suites de suites de signes ; elle est relative à un système formel qui détermine une méthode particulière pour la codification des preuves. En dépit de ces différences, le théorème de complétude (dans une version plus générale que celle qu'on a donnée au chapitre III) établit néanmoins une forme d'équivalence entre les deux concepts, au sens suivant : étant donné un langage pour la logique du premier ordre, il est possible de définir un système de dérivation formelle tel que, pour tout ensemble de formules Γ et toute formule ϕ, ϕ est conséquence logique de Γ si, et seulement si, ϕ est formellement dérivable à partir de Γ.

Pour prouver que l'inférence déductive qui conduit de Γ à ϕ est logiquement valide, il existe donc deux voies : selon une première méthode – sémantique – on raisonne, dans le métalangage, sur les structures d'interprétation du langage afin de montrer que tout modèle de Γ est modèle de ϕ ; selon une seconde méthode – syntaxique – on cherche une dérivation formelle de ϕ à partir de Γ dans un système formel particulier. Le théorème de complétude montre que, dans la logique du premier ordre, la relation sémantique de conséquence logique, infinitaire et abstraite, est équivalente, ou se réduit, à une relation syntaxique de dérivation formelle. Une telle réduction a cependant un coût : une dérivation formelle de ϕ à partir de Γ, par exemple d'un théorème de la théorie des nombres à partir des axiomes de cette théorie, est généralement beaucoup plus longue que la preuve classique qui est habituellement donnée en mathématiques.

L'un des corollaires les plus importants du théorème de complétude est le théorème de compacité : *un ensemble* Γ *de formules a un modèle si, et seulement si, tout sous-ensemble fini de formules de* Γ *a un modèle*. Pour donner un exemple d'application de ce théorème, nous montrons maintenant que la théorie de l'arithmétique, T_A, écrite dans le langage $\mathcal{L}_A$ (cf. *supra*, p. 45) admet des modèles non standard.

Considérons en effet le langage $\mathcal{L}_A$ pour l'arithmétique auquel on ajoute une constante d'individu, c, et notons $\mathcal{L}'_A$ le langage ainsi obtenu. À la théorie T_A, on ajoute une infinité d'axiomes : $\underline{0} < c$, $\underline{1} < c$, $\underline{2} < c$… et $\underline{n} < c$ pour chaque entier n et on note T'_A la théorie ainsi obtenue (ici, « $\underline{n}$ » est la représentation de l'entier n dans le langage $\mathcal{L}'_A$). Notons $\mathcal{N}$ la structure d'interprétation de $\mathcal{L}_A$ qui est le modèle standard de l'arithmétique.

Soit $\mathcal{N}_c$ une structure d'interprétation de qui est identique à $\mathcal{N}$, à ceci près que dans $\mathcal{N}_c$ figure l'interprétation de la constante c, par un entier. Toute partie finie de T'_A a un modèle : il suffit en effet de prendre $\mathcal{N}_c$ et d'interpréter la constante c par $m + 1$, où m est le plus grand entier tel que la formule $\underline{m} < c$ appartient à la partie finie de T'_A qui est considérée. Donc, d'après le théorème de compacité, T'_A a un modèle. Notons $\mathcal{M}$ ce modèle et $\mathcal{D}$ son domaine d'individus. $\mathcal{D}$ contient un élément qui est plus grand que tous les entiers naturels (car $\mathcal{D}$ est modèle de $\underline{n} < c$ pour tout entier n). Appelons k l'élément de $\mathcal{D}$ qui interprète c. Puisque $\mathcal{M}$ est modèle de T'_A, $\mathcal{M}$ est aussi modèle de toute partie de T'_A, et donc en particulier de T_A. Il existe donc un modèle de la théorie de l'arithmétique dont le domaine d'individus contient un élément k qui est plus grand que tous les entiers naturels. Le « nombre » k est un entier non standard, et $\mathcal{M}$ est un modèle non standard de l'arithmétique. Comme l'un des axiomes de T_A est que tout nombre a un successeur, k a un successeur et $\mathcal{D}$, le domaine d'individus de $\mathcal{L}$, contient non seulement 0, 1, 2, 3... mais aussi k, $k + 1$, $k + 2$... qui sont des entiers non standard, tous supérieurs à tous les entiers naturels. On peut montrer, en outre, que l'ensemble des entiers standard d'un modèle non standard n'est pas définissable dans la logique du premier ordre.

La théorie des modèles examine typiquement des problèmes de ce genre : étant donné une structure mathématique comme celle de la théorie des nombres, existe-t-il un langage du premier ordre et un ensemble d'axiomes formels de ce langage qui permettent de caractériser la structure en question ? Le théorème de compacité permet de voir que la réponse à cette question précise est négative : aucune axiomatisation de l'arithmétique ne permet d'exclure les modèles non standard, ce qui montre que la logique du premier ordre échoue à caractériser la propriété mathématique « être un nombre entier naturel », de même qu'elle échoue à caractériser d'autres propriétés comme « être fini » ou « être dénombrable ».

La logique du second ordre offre une puissance d'expression supérieure, car dans un langage du second ordre, la quantification ne s'applique pas uniquement à des variables d'individus : elle porte également sur des propriétés de ces individus, représentées par des sous-ensembles de l'univers du discours (si $\mathcal{D}$ est l'ensemble des entiers naturels, un sous-ensemble de $\mathcal{D}$ – par exemple l'ensemble des nombres premiers – représente une propriété des entiers naturels, comme « être premier »). Comme on l'a remarqué (*supra*, p. 45), il est alors possible d'exprimer que *pour toute propriété P*, si *P* est satisfaite par zéro et par le successeur de tout entier qui la satisfait, tous les nombres entiers naturels possèdent la propriété en question. Par cet axiome, la logique du second ordre permet de caractériser la structure des entiers naturels. Plus généralement, elle permet d'axiomatiser bien d'autres classes de structures pour lesquelles la puissance d'expression de la logique du premier ordre est insuffisante.

Une partie des débats contemporains touchant la nature et le domaine de la logique s'est concentrée sur la question de savoir si ses frontières coïncidaient avec celles des langages du premier ordre ou si elles enveloppaient un plus vaste champ, allant jusqu'à inclure les langages du second ordre. Ce type de questions ne suscitait pas autant d'intérêt avant que les théorèmes de complétude sémantique, d'incomplétude syntaxique, de compacité, de Löwenheim-Skolem et d'indéfinissabilité du prédicat de vérité n'aient révélé que le choix d'un langage et d'une méthode d'interprétation avait des conséquences décisives sur les propriétés qu'on pouvait attendre d'un système logique. Après que les théorèmes d'incomplétude eurent montré qu'une axiomatisation et une formalisation de l'arithmétique dans un seul et même langage laissaient nécessairement ouvertes un certain nombre de questions, plusieurs conceptions de la logique, comme celles qu'avaient formées Frege, Russell ou Hilbert, durent être, sinon abandonnées, du moins sérieusement révisées. Plus généralement, toute conception de la logique qui visait l'idéal d'une formalisation unifiée des raisonnements dans un seul langage se voyait contrainte de faire des concessions sur les effets attendus d'un tel programme. La question était alors de savoir à quelles propriétés le logicien était prêt à renoncer : cohérence, complétude, compacité, axiomatisation de concepts mathématiques comme « fini », « dénombrable » ou « entier naturel », possibilité de définir un prédicat de vérité, caractère finitaire des preuves, décidabilité (sur cette propriété, cf. chap. VI), définition tarskienne de la conséquence logique, interprétation d'un langage en termes de structure ?

Typiquement, un logicien qui défend la thèse selon laquelle le domaine propre de la logique est celui du premier ordre invoquera les propriétés de complétude et de compacité qui ne sont pas satisfaites, en général, pour une logique d'ordre supérieur. Quine, qui défend cette thèse, ajoute que les langages du second ordre permettent certes de définir des structures mathématiques comme celle des entiers naturels mais que, précisément pour cette raison, elles reposent sur des principes qui ne sont logiques qu'en apparence. Un défenseur du programme logiciste qui se proposerait de déduire une partie des mathématiques de la logique du second ordre devrait commencer par exposer les vertus supposées d'un tel programme, ce qui implique qu'il explique en quoi la logique du second ordre est, à proprement parler, logique.

La logique du premier ordre et la sémantique tarskienne se sont imposées comme les éléments de base de la logique contemporaine et elles forment, de fait, la partie centrale de la plupart des manuels d'introduction. Les recherches logiques qui reposent sur d'autres principes ont souvent, du reste, été jugées négativement (« logiques révisionnistes »). Ce genre de jugement peut s'expliquer par le fait que la théorie des ensembles, qui permet de codifier toutes les mathématiques contemporaines, peut être formalisée dans un langage du premier ordre dont le seul signe non logique est le symbole d'appartenance « $\in$ », même si cette possibilité ne fait évidemment pas disparaître pour autant les limitations de la logique du premier ordre précédemment signalées[1]. Il y a maintenant plus d'un demi-siècle, cependant, que la question du fondement des mathématiques n'est plus l'affaire principale de la logique et que celle-ci a trouvé non seulement d'autres domaines d'application, mais également d'autres sources d'inspiration, notamment dans l'étude des langues naturelles et des machines informatiques. De fait, l'analyse des raisonnements non mathématiques et des capacités des machines est à l'origine d'une grande partie des innovations introduites en logique au cours des dernières décennies.

1. Sur la théorie des ensembles, cf. H. B. Enderton, *Elements of Set Theory*, New York, Academic Press, 1977 ; M. Machover, *Set Theory, Logic and their Limitations*, Cambridge, Cambridge UP, 1996.

Chapitre V

LOGIQUE ET LANGAGE

Selon des textes anciens, Eubulide de Milet serait l'auteur du dialogue suivant : « Tu as ce que tu n'as pas perdu ? – Oui. – Mais as-tu perdu des cornes ? – Non ! – Donc tu as des cornes ? » Les logiciens grecs de l'Antiquité exerçaient leur sagacité sur des raisonnements de ce genre qu'ils créaient à dessein comme des jeux logiques. L'exercice consistait alors à trouver les règles que l'auteur du sophisme n'avait pas respectées. La logique n'a jamais cessé de considérer les multiples manières dont le langage occasionne des erreurs de jugement et de raisonnement. Mais la résolution d'un paradoxe ou l'examen de la validité d'une inférence ne sont pas toujours aussi aisés que dans le cas du Cornu et l'on comprend, pour cette raison, la nécessité d'une analyse logique du langage ainsi que le projet d'une langue logiquement parfaite, libre de tout ce qui, dans la langue usuelle, obscurcit la structure logique des énoncés et la chaîne déductive des raisonnements.

Les langues que nous parlons et écrivons n'ont pas pour seule fonction l'expression d'inférences valides. Elles servent à persuader, ordonner, prier, traduire des émotions, émettre des hypothèses, transmettre des informations, poser des questions, créer des illusions… Aussi n'est-il pas étonnant que même dans le cas d'énoncés déclaratifs relativement simples, le travail des logiciens ne consiste pas uniquement à rendre parfaitement explicites des expressions ambiguës ou incomplètes. Les analyses qu'ils produisent ont également pour vertu de montrer que la forme logique réelle d'un énoncé est parfois très éloignée de sa structure grammaticale. Lorsque Frege préconisa une rupture par rapport à la logique traditionnelle et le remplacement des concepts *sujet* et *prédicat* par *argument* et *fonction*, il justifia cette réforme en arguant que la logique pâtissait de sa dépendance à l'égard de modèles empruntés au langage : « La logique s'est jusqu'ici toujours rattachée trop étroitement à la langue et à la grammaire. » Il pensait que son idéographie, *a contrario*, pourrait contribuer à « rompre la domination du mot sur l'esprit humain »[1].

Alors que l'analyse traditionnelle suggère de voir aussi bien dans « Félix est le voisin de Marion » que dans « tous les hommes sont mortels » l'attribution d'un prédicat à un sujet (respectivement : « Félix » et « tous les hommes »), Frege voit dans le premier énoncé une égalité du type $a = b$ où « a » et « b » sont des expressions qui désignent des objets, tandis qu'il représente la structure logique du second par la formule $\forall x(Hx \rightarrow Mx)$ (« pour tout individu x, si x est un homme alors x est mortel »). Dans cette formule, aucun constituant ne correspond au groupe nominal « tous les hommes » et la condition « si x est un homme » ne correspond à rien, au moins à première vue, dans l'énoncé initial. De même, dans la formule $\exists x(Mx \wedge Ax)$ qui se lit « il existe

1. Frege, *Idéographie*, 1879, trad. franç., Paris, Vrin, 2000, p. 8, pour les deux citations.

au moins un individu x tel que Mx et Ax » et qui sert à transcrire un énoncé comme « certains mammifères sont aquatiques » dans la logique du premier ordre, on chercherait en vain un constituant qui traduirait « certains mammifères ».

Dans l'analyse frégéenne, une description définie qui sert à désigner un individu (par exemple « le voisin de Marion ») est considérée comme un terme référentiel, au même titre qu'un nom propre comme « Socrate ». Pourtant, si l'un des attendus de toute langue logique est de désambiguïser les expressions de la langue usuelle afin de pouvoir rendre compte des inférences déductives, cette analyse n'est pas irréprochable, car elle ne permet pas de faire la différence entre les deux lectures possibles de « Loïc veut voir le voisin de Marion », selon que Loïc veut voir le voisin de Marion parce qu'il s'agit de Félix (lecture *de re* de la phrase), ou que Loïc veut voir le voisin de Marion sans savoir de qui il s'agit (lecture *de dicto*).

Russell, qui contribua au moins autant que Frege à montrer combien la structure logique réelle d'un énoncé pouvait être éloignée de sa structure logique apparente, exposa une stratégie d'analyse qui permet non seulement de lever cette difficulté mais également de résoudre certaines apories liées aux descriptions définies qui n'ont pas de dénotation. Comment, demande Russell, analyser une phrase comme « l'actuel roi de France est chauve » dans laquelle une description définie (« l'actuel roi de France ») semble désigner un individu qui, pourtant, n'existe pas ? Comment éviter de représenter cette description définie par une expression qui laisse supposer qu'elle possède une dénotation ? La phrase initiale est-elle fausse ou dépourvue de signification ? La

solution que Russell expose en 1905 dans « On denoting »[1] est un paradigme de l'analyse logique. Selon les principes exposés dans cet article, « l'actuel roi de France est chauve » doit être analysé ainsi :
il existe un individu x tel que

1. x est roi de France ;
2. il n'existe pas d'individu y différent de x tel que y soit roi de France ;
3. x est chauve.

Premièrement, comme il n'existe aucun individu qui ait les trois propriétés énumérées, la proposition en question est fausse. Deuxièmement, l'analyse proposée ayant pour vertu de faire disparaître la description définie « l'actuel roi de France », le problème de sa dénotation ne se pose tout simplement plus. Troisièmement, l'analyse de Russell permet de lever l'ambiguïté de « l'actuel roi de France n'est pas chauve » qui est, respectivement, vraie ou fausse selon qu'elle est entendue comme « il est faux qu'il existe actuellement un individu qui soit roi de France, unique roi de France et chauve » ou comme « il y a actuellement un individu qui est roi de France, unique roi de France, et qui n'est pas chauve ». Enfin, l'application des principes d'analyse de Russell permet de distinguer les lectures *de re* et *de dicto* de « Loïc veut voir le voisin de Marion. »

L'analyse russellienne des descriptions définies apporte donc une solution à plusieurs difficultés logiques et ontologiques et c'est elle qui est généralement retenue dans le cadre de la logique classique du premier

1. Russell, « De la dénotation », 1905, trad. franç., *in* Russell, *Écrits de logique philosophique*, Paris, PUF, 1989.

ordre. Elle n'est cependant pas la seule analyse possible. Surtout, elle conduit à creuser un tel écart, au point de vue de la structure, entre une phrase déclarative simple comme « le voisin de Marion est musicien » et la formule qui est censée en donner l'analyse logique qu'on peut légitimement se demander si la formule en question en révèle la forme logique réelle ou si elle n'en constitue pas plutôt une paraphrase, c'est-à-dire une transposition qui vise à en supprimer les ambiguïtés éventuelles, à en exprimer aussi clairement que possible les présupposés implicites, et à rendre compte de la validité des inférences déductives dans lesquelles elle est susceptible de figurer. Plus généralement, que doit-on attendre de l'analyse logique d'un énoncé : qu'elle mette en évidence la forme logique sous-jacente qui serait dissimulée par la forme grammaticale ou, plus simplement, qu'elle opère une transcription des énoncés de la langue usuelle dans une langue artificielle choisie en raison de son pouvoir de clarification ? Se pose également la question connexe : parmi toutes les ressources du langage naturel, quelles sont celles qui doivent servir à former les éléments syntaxiques constitutifs de la langue logique à laquelle on a recours ?

Depuis Frege, et jusqu'au milieu du XX^e^ siècle, la réponse à cette dernière question a été largement influencée par le problème du fondement des mathématiques. De fait, dans cette période, une grande part du travail des logiciens s'est concentrée sur la question de savoir quelles étaient les ressources minimales dont un langage logique devait disposer afin de pouvoir formaliser tous les théorèmes et toutes les preuves des mathématiciens.

Parmi les ressources du langage naturel qu'on intègre généralement aux langues logiques figurent des connecteurs propositionnels (le plus souvent : « ne... pas », « et », « ou », « si... alors »), des quantificateurs (« tous » et « au moins un »), la prédication (« x est P »), les relations (« x est dans la relation R avec y »), les variables d'individus (x, y, etc.) et les termes généraux. Il est possible de donner une interprétation claire et précise de ces éléments syntaxiques, pourvu qu'on ne cherche pas à retenir tous les aspects de leur usage dans le langage courant.

Parmi les propriétés attendues des systèmes destinés à la formalisation des mathématiques, l'*extensionalité* est l'une de celles qui sont recherchées pour la clarté qu'elles confèrent à la structure des langages dans lesquels ces systèmes sont exprimés. Pour comprendre ce qu'elle signifie, considérons tout d'abord le cas de la logique propositionnelle.

Un langage formel pour le calcul propositionnel comporte des lettres de propositions (p, q, r, etc.) et des connecteurs propositionnels : la négation ($\neg$), la conjonction ($\wedge$), la disjonction inclusive ($\vee$), le connecteur de l'implication matérielle ($\rightarrow$). Le plus souvent, le langage comporte également la parenthèse ouvrante et la parenthèse fermante qui servent à éliminer les cas d'ambiguïté. Si p représente la proposition « Emma lit Corneille » et q représente « Emma aime le théâtre classique », la formule $(p \wedge q)$ (qui se lit « p conjonction q » ou « p et q ») peut servir à formaliser « Emma lit Corneille et elle aime le théâtre classique. » Du point de vue sémantique, une *interprétation* du langage pour le calcul propositionnel est une fonction qui donne à chaque lettre de proposition la valeur vrai ou la valeur faux. Les connecteurs sont interprétés par des fonctions de vérité. Par exemple $(p \wedge q)$ est vrai si p est vrai et q

est vrai, et faux dans tous les autres cas. Les connecteurs sont donc vérifonctionnels : la valeur de vérité de $(p \wedge q)$ ne dépend pas du sens des propositions que représentent p et q, mais uniquement de leur valeur de vérité. Supposons que « Emma lit Corneille » et « Emma aime le théâtre classique » soient vraies ; étant donné que « 2 + 2 = 4 » est également vrai, « Emma lit Corneille et elle aime le théâtre classique » a la même valeur de vérité que « Emma lit Corneille et 2 + 2 = 4 ». Si à la place de « et » nous considérons maintenant « parce que », nous voyons aisément – en reprenant le même exemple – que la valeur de vérité de « p parce que q » ne dépend pas uniquement de la valeur de vérité de p et de q mais également de leur sens. La conjonction « parce que » ne peut pas être représentée par un connecteur vérifonctionnel.

L'extensionalité d'un langage suppose donc que le remplacement d'un énoncé A par un énoncé A' dans un énoncé B produise un énoncé B' qui a même valeur de vérité que B si A' a même valeur de vérité que A. Plus généralement, si l'on ne se limite plus à la logique propositionnelle, elle suppose que soit respecté le principe de substituabilité des identiques. Selon ce principe, si $a = b$, tout énoncé dans lequel figure « a » a même valeur de vérité que l'énoncé obtenu en remplaçant une occurrence de « a » par « b », « a » et « b » étant ici des termes qui ont pour vocation de désigner un individu ou un objet.

Bien des logiciens estiment que la logique classique du premier ordre – qui *est* extensionnelle – possède des propriétés si remarquables et une structure si claire qu'elle n'est pas seulement le cadre le plus adéquat pour la formalisation des mathématiques mais qu'elle marque également les limites de la logique elle-même. Pour Quine, par exemple, la logique du premier ordre avec identité définit le cadre d'une notation canonique dans laquelle tous les énoncés scientifiquement acceptables doivent pouvoir être paraphrasés. Cette notation canonique austère a recours uniquement aux variables d'individus, aux symboles de relations, à un quantificateur (universel ou existentiel), à des connecteurs propositionnels (par exemple la négation et la conjonction) et au signe d'égalité, à l'exclusion de toutes les autres ressources que peuvent offrir les langues naturelles.

L'extrême parcimonie des moyens utilisés dans la notation canonique s'explique par l'objectif philosophique qui est celui de Quine : « La recherche d'une notation canonique universelle qui ait la structure la plus simple et la plus claire possible ne doit pas être distinguée d'une recherche des catégories dernières ou d'un effort de reproduction des traits les plus généraux de la réalité. »[1] Pour être recevables, les expressions du langage usuel doivent pouvoir être paraphrasées dans cette notation logique qui a donc valeur de norme : un énoncé, un argument ou un raisonnement qui ne pourraient absolument pas être paraphrasés dans la notation canonique seraient réputés obscurs et non scientifiquement acceptables.

Entendue au sens quinien, la paraphrase dans la notation canonique ne prétend pas produire un synonyme des énoncés initiaux. Elle a une fonction de clarification et de désambiguïsation. Considérons l'exemple d'un nom propre comme « Pégase ». Il ne peut pas être représenté

1. W. V. Quine, *Le Mot et la Chose*, 1960, trad. franç., Paris, Flammarion, 1977, p. 232.

par une constante d'individu (la notation canonique n'a pas recours à cette catégorie de signes) et il soulève tous les problèmes liés aux désignateurs qui renvoient à des objets inexistants. Pour transcrire « Pégase vole » dans la notation canonique, Quine commence par paraphraser cette phrase ainsi : « il existe un x tel que x pégasise et x vole », le verbe « pégasiser » étant un moyen *ad hoc* dont la fonction est d'exprimer la propriété « être Pégase » et qui peut être représenté par un symbole de relation à une place « P ». D'où la formule : $\exists x(Px \wedge Vx)$ qui sert à paraphraser « Pégase vole ».

Cette conception exigeante de la logique doit faire face aux critiques, au motif, par exemple, que la faiblesse des moyens d'expression dont elle dispose la rend difficilement apte à rendre compte d'un certain nombre de raisonnements que nous effectuons couramment et que chacun juge tout à fait recevables. Comment justifier, par exemple, l'inférence de « Jean sait que Stendhal est romancier » à « il y a quelqu'un dont Jean sait qu'il est romancier » ? Le contexte « Jean sait que » n'est pas extensionnel. Jean peut en effet savoir que Stendhal est romancier sans savoir qu'Henri Beyle est romancier, bien que Stendhal et Henri Beyle soient une seule et même personne. Mais dans l'exemple précédent, la difficulté principale vient du fait que l'inférence introduit, à l'extérieur du contexte non extensionnel « Jean sait que », une quantification dont la portée va jusqu'à l'intérieur de ce contexte : « *il y a quelqu'un* dont Jean sait qu'il est romancier »[1].

1. Sur la critique quinienne de la quantification dans les contextes modaux, cf. Quine, « Référence et modalité », *in* Quine, *Du point de vue logique*, 1953, trad. franç., Paris, Vrin, 2003.

Que la notation canonique ne permette pas de paraphraser toutes les tournures de la langue usuelle ne constitue pas, aux yeux de Quine, une objection sérieuse contre elle. Car le but qu'il assigne à la paraphrase dans une langue logique n'est pas de parvenir à une description de la structure des langues naturelles mais à une clarification logique et ontologique de notre discours ou, le cas échéant, à la mise en évidence des obscurités irréductibles de certains de nos énoncés. Elle doit permettre, notamment, de rendre parfaitement transparentes les implications ontologiques de nos théories : savoir ce qui doit exister pour qu'elles soient vraies. Le célèbre critère quinien selon lequel « être, c'est être la valeur d'une variable liée »[1] explique l'importance d'une paraphrase de nos énoncés dans une langue logique aussi claire et simple que possible.

L'austérité des ressources linguistiques auxquelles la notation canonique quinienne a recours se justifie donc par un projet philosophique qui touche non seulement la logique mais également la philosophie de la connaissance, l'ontologie et la philosophie du langage, projet global que les logiciens ne partagent évidemment pas tous. Aussi peuvent-ils avoir de multiples raisons de ne pas limiter leurs recherches à la logique extensionnelle ou à la logique classique du premier ordre. Si celle-ci jouit assurément de propriétés remarquables, elle impose en effet également de sévères contraintes, notamment parce que son pouvoir d'expression n'est pas illimité. Le recours à une logique du second ordre peut se justifier par la volonté de formaliser des concepts couramment

1. Cf. Quine, De ce qui est, 1948, in *Du point de vue logique*, 1953, trad. franç., Paris, Vrin, 2003.

utilisés en mathématiques et qui ne peuvent pas être caractérisés au premier ordre. Tel est le cas de « être fini » ou « être dénombrable »[1]. Quant au recours à des langages qui ne sont pas purement extensionnels, il ne se justifie pas par la formalisation des mathématiques ou des sciences de la nature mais par d'autres projets, notamment celui d'une description des structures (syntaxiques ou sémantiques) des langues naturelles. Dans la suite de ce chapitre, nous nous proposons d'examiner quelques-unes des raisons que les logiciens – ou les linguistes qui font appel, pour leurs travaux, aux techniques de la logique – ont pu avoir de ne pas s'en tenir au cadre extensionnel ou à la logique classique du premier ordre. Nous considérerons, à titre d'exemples, trois cas particuliers : les énoncés conditionnels, l'expression de la modalité et les quantificateurs généralisés.

Dès le début du XX[e] siècle, C. I. Lewis proposa de renoncer à l'extensionalité du langage logique des *Principia Mathematica* et d'introduire, au moyen de son « implication stricte », une analyse des expressions de la forme « si *p* alors *q* » qui soit plus respectueuse des usages de la langue usuelle que l'analyse qu'en avaient donnée Russell et Whitehead ou Frege. L'expression « si... alors... » soulève en fait des difficultés dont les logiciens discutent depuis l'Antiquité[2]. Dans le calcul propositionnel issu de l'idéographie frégéenne, « si *p* alors *q* » est interprété par

1. Cf. *supra*, p. 58 et 59. S. Shapiro, *Foundations without Foundationalism. A Case for Second Order Logic*, Oxford, Clarendon Press, 1991, développe une argumentation en faveur de la logique du second ordre.
2. Cf. J.-B. Gourinat, *La Dialectique des stoïciens*, Paris, Vrin, 2000, p. 217-233.

« on n'a pas p sans avoir q », c'est-à-dire « non (p et non-q) », ce qui conduit à une transcription par le connecteur vérifonctionnel de l'implication matérielle ainsi défini : $(p \rightarrow q)$ est faux si, et seulement si, p est vrai et q est faux. Il est clair, cependant, qu'une telle interprétation ne permet pas de donner une formalisation satisfaisante des énoncés qui expriment un conditionnel irréel : « S'il n'avait pas plu ce jour-là, nous aurions pu mener à bien notre projet. » Elle a, en outre, des effets paradoxaux parce qu'elle contraint à donner la valeur vrai à tous les énoncés de la forme « si p alors q » dans lesquels p est faux, y compris, par exemple « si la Terre est immobile, alors Galilée a raison », que nous aurions plutôt tendance à considérer comme faux ! Elle donne aussi, ce qui n'est pas moins fâcheux, la valeur vrai à des énoncés dans lesquels il n'existe aucun rapport de sens entre l'antécédent et le conséquent, par exemple « si 3 est un nombre pair, alors la Terre tourne ». Si ce genre de difficulté ne se présente guère tant qu'il n'est question que de formaliser les démonstrations des mathématiciens, il n'en va pas de même lorsque les connecteurs du calcul propositionnel sont mis au service d'une analyse logique d'énoncés ou de raisonnements quelconques.

En 1912, C. I. Lewis suggéra de remplacer le connecteur de l'implication matérielle par le concept d'implication stricte, obtenu par l'adjonction d'un opérateur exprimant la nécessité. « p implique strictement q » signifie « il est nécessaire que (p implique [au sens matériel] q) ». Afin de parvenir à une transcription plus satisfaisante de « si p alors q », Lewis introduisit donc des systèmes de logique qui permettaient l'expression de modalités comme « il est possible que » ou « il est nécessaire que »[1] et qui ne jouissaient pas de la propriété d'extensionalité. Ultérieurement, d'autres logiciens montrèrent cependant que la solution de Lewis ne permettait pas d'éliminer tous les paradoxes de l'implication. En particulier, la formule $(p \rightarrow (q \rightarrow q))$ est logiquement vraie, que la flèche soit interprétée au sens du connecteur de l'implication matérielle ou de l'implication stricte. Or, la vérité de cette formule est paradoxale dès que p et q représentent des propositions dont les significations respectives n'ont aucun rapport entre elles (par exemple p : la Terre tourne ; q : $2 + 2 = 4$). La logique pertinente introduite par Anderson et Belnap dans les années 1960 se propose précisément de pallier ce genre d'inconvénient en introduisant une nouvelle manière de réformer la logique extensionnelle[2].

Pour les systèmes de logique modale introduits par Lewis, aucun théorème de complétude ne put être démontré avant que les travaux de Hintikka, Kanger et Kripke ne définissent une sémantique adéquate, en termes de mondes possibles. Considérons le cas, relativement simple, de la logique propositionnelle. On introduit deux opérateurs modaux $\Box$ et $\Diamond$ dans un langage classique. $\Box p$ peut être interprété par « il est nécessaire que p » et $\Diamond p$ par « il est possible que p ». En quoi consiste une interprétation pour ce nouveau langage ? À quelles conditions $\Box p$ et $\Diamond p$ sont-ils vrais dans une telle interprétation ? Comment définir la validité

1. C. H. Langford et C. I. Lewis, *Symbolic Logic*, Londres, The Century, 1932.
2. Cf. A. R. Anderson, N. D. Belnap et J. M. Dunn, *Entailment: The Logic of Relevance and Necessity*, Princeton NJ, Princeton UP, vol. I, 1975, vol. II, 1992 ; F. Rivenc, *Introduction à la logique pertinente*, Paris, PUF, 2005.

d'une formule qui contient ces opérateurs ? Des affirmations comme « il est possible que la royauté soit rétablie en France » ou « Socrate aurait pu ne pas être philosophe » peuvent être comprises comme signifiant qu'un état de chose est possible dans lequel la royauté est rétablie en France, ou dans lequel Socrate n'est pas philosophe. De telles variantes du monde réel peuvent alors être conçues comme figurant dans un monde possible, plus ou moins différent du monde réel qui est le nôtre. Partant de cette idée, les logiciens introduisirent une sémantique des mondes possibles pour l'interprétation des énoncés modaux.

Pour un langage de la logique modale propositionnelle, une interprétation est donnée par un ensemble $\mathcal{W}$ (un ensemble d'états ou de mondes possibles) et une fonction v (une assignation de valeurs) qui, pour chacun des éléments w de $\mathcal{W}$, associe une valeur de vérité à chaque lettre de proposition du langage. L'intérêt de $\mathcal{W}$ tient à ce que la valeur de vérité d'un énoncé peut être différente dans w et dans w'. La formule $\Diamond p$ interprétée par « il est possible que p » est alors vraie dans le monde w pour l'assignation v s'il existe un élément de $\mathcal{W}$ dans lequel v donne la valeur vrai à p. Mais dans la sémantique des mondes possibles, la structure d'interprétation comporte, outre $\mathcal{W}$ et v, une relation d'accessibilité entre les éléments w de $\mathcal{W}$: sans pouvoir justifier ici davantage l'introduction de cette relation, disons qu'elle détermine, pour chaque w, quels sont les éléments de $\mathcal{W}$ qui sont accessibles à partir de w. Dans une structure de Kripke, la formule $\Box p$ est vraie pour l'assignation v dans le monde ou dans l'état w si v donne à p la valeur vrai dans *tous* les w' accessibles à partir de w. La formule $\Diamond p$ est vraie pour l'assignation v dans w si v donne à p la valeur vrai dans *au moins l'un* des w' accessibles à partir de w. Si l'on passe maintenant de la logique modale propositionnelle à la logique modale du premier ordre, la sémantique est plus complexe, notamment parce qu'un domaine d'individus est associé à chaque élément de $\mathcal{W}$, parce qu'il faut préciser comment les symboles de relations sont interprétés, et parce qu'il existe de multiples options possibles pour la définition de cette sémantique[1]. Ajoutons que la sémantique des mondes possibles autorise l'interprétation des opérateurs $\Box$ et $\Diamond$ par d'autres modalités que le nécessaire et le possible (logiques épistémiques, temporelles, déontiques, etc.).

Les applications de la sémantique de Kripke – ou de ses variantes – vont bien au-delà de la logique modale proprement dite. Cette sémantique a été utilisée, par exemple, dans le cadre de la « nouvelle théorie de la référence », opposée à celle de Frege et de Russell pour qui un nom propre n'était pas essentiellement différent d'une description définie. Alors que la référence d'une description définie varie avec les mondes possibles, Kripke introduit le concept de désignateur rigide comme celui d'une expression dont la référence ne varie pas d'un monde possible à l'autre[2]. Cette théorie a elle-même des répercussions qui dépassent l'analyse linguistique des noms propres et qui touchent, par exemple, la distinction entre vérité nécessaire et vérité *a priori*. La sémantique de Kripke a également permis

1. Cf. P. Gochet, P. Gribomont, A. Thaysse, *Logique*, vol. III, Paris, Hermès, 2000 ; P. Blackburn, M. de Rijke, Y. Venema, *Modal Logic*, Cambridge, Cambridge UP, 2001.
2. Cf. S. Kripke, *La Logique des noms propres*, 1970, trad. franç., Paris, Minuit, 1982.

de progresser dans l'analyse des conditionnelles contrefactuelles (« si Napoléon avait gagné à Waterloo… ») et dans la compréhension des expressions linguistiques de la croyance, du savoir, du devoir, de la temporalité… Plus généralement, elle s'est révélée extrêmement utile pour les recherches sur les logiques intensionnelles et a ainsi contribué au tournant qu'a connu la logique dans la seconde moitié du XX^{e} siècle, lorsque les travaux des logiciens cessèrent d'être orientés de manière quasi exclusive par le problème du fondement des mathématiques.

Ce tournant s'accentua d'autant plus que la linguistique faisait elle-même usage de méthodes formelles empruntées à la logique[1]. Partant de l'idée que les enfants possèdent une connaissance innée de certains principes universels de grammaire et afin de trouver une explication de la compétence linguistique des locuteurs – comment expliquer leur capacité de produire un nombre potentiellement infini d'énoncés ? – Chomsky se proposa d'étudier la syntaxe des langues naturelles en faisant appel aux techniques de la logique formelle[2]. À la fin des années 1960, Montague s'engagea, pour sa part, dans le programme d'une analyse sémantique de fragments de l'anglais qui avait l'ambition d'être aussi précise et rigoureuse que les analyses sémantiques pratiquées par les logiciens sur des langues formelles. La grammaire universelle dont il conçut le projet devait être

1. Sur les rapports entre logique et linguistique, cf. A. Lenci et G. Sandu, « Logic and Linguistics in the Twentieth Century », *in* L. Haaparanta (ed.), *The Development of Modern Logic*, Oxford, Oxford UP, 2009. F. Rivenc et G. Sandu, *Entre logique et langage*, Paris, Vrin, 2009.

2. N. Chomsky, *Structures syntaxiques*, 1957, trad. franç., Paris, Le Seuil, 1969.

applicable indifféremment aux langues formelles et aux langues naturelles[1]. Cette orientation de la sémantique formelle a pour but de mener à bien une tâche que Tarski avait jugée irréalisable : déterminer des conditions de vérité pour tous les énoncés d'une langue naturelle, ou au moins pour un large fragment d'une telle langue. Parmi les moyens logiques utilisés pour la réalisation de ce programme figurent la théorie des types, le lambda-calcul, la sémantique des mondes possibles, ainsi que la théorie des quantificateurs généralisés.

Rétrospectivement, on peut dire que la logique aristotélicienne prend en considération quatre quantificateurs : « tous », « quelques », « aucun » et « pas tous ». Frege, quant à lui, avait retenu les quantificateurs universel (« tous ») et existentiel (« il existe au moins un »), qu'on retrouve, interprétés différemment (cf. *supra*, p. 22), dans la logique du premier ordre. Il s'agit alors d'opérateurs qui servent à lier des variables : si la variable x a une occurrence libre dans la formule ϕ, toutes les occurrences de x sont liées dans $\forall x\phi$. Dans les années 1950, Mostowski[2] généralisa le concept de quantificateur qui était devenu standard dans la logique mathématique. Il fut suivi par Lindström[3] en 1966 puis, du point de vue de la linguistique, par Barwise et Cooper[4] en 1981.

1. R. Montague, *Formal Philosophy*, New Haven, Londres, Yale UP, 1974.
2. A. Mostowski, « On a Generalization of Quantifiers », *Fundamenta Mathematicae*, 44, 1957.
3. P. Lindström, « First-Order Predicate Logic with Generalized Quantifiers », *Theoria*, 32, 1966.
4. J. Barwise et R. Cooper, « Generalized Quantifiers and Natural Language », *Linguistics and Philosophy*, 4, 1981.

Si les quantificateurs de la logique classique du premier ordre ($\forall x$ et $\exists x$) permettent de formuler d'autres quantifications comme « un seul x », « au moins trois x », « au plus quatre x », « aucun x » pourvu qu'on dispose du signe d'égalité dans le langage, ils sont impuissants à exprimer « un nombre fini de x », « une infinité de x », « la plupart des x », « un nombre indénombrable de x », « un nombre pair de x », etc. Du point de vue syntaxique, il est toujours possible d'introduire une expression comme « Q_0x » avec l'intention d'exprimer « une infinité de x ». Le problème est de savoir quelle sera la sémantique du langage ainsi étendu. Un langage dans lequel « Q_0x » est interprété par « une infinité de x » ne jouit plus des mêmes propriétés sémantiques qu'un langage pour la logique du premier ordre.

Dans la logique d'Aristote, la quantification établit une relation entre le sujet et le prédicat : dans « tous les hommes sont mortels », « tous » est une manière de lier « hommes » à « mortels » qui diffère du lien établi entre ces deux termes par « aucun », « quelques » ou « pas tous ». L'une des difficultés soulevées par cette approche est de comprendre ce que signifient les expressions quantifiées. Si l'on décide d'interpréter « tous les hommes » par l'ensemble de tous les hommes, on voit mal en effet quelle sera l'interprétation de « aucun homme », qu'on souhaiterait pouvoir distinguer de « aucun cheval ». La logique du premier ordre adopte une stratégie d'analyse tout à fait différente, tant au point de vue syntaxique (les quantificateurs sont nécessairement attachés à une variable d'individus : $\forall x$, $\exists y$) qu'au point de vue sémantique, qui fait intervenir le concept de structure d'interprétation. La théorie des quantificateurs généralisés, quant à elle, oriente l'analyse dans une autre direction, qu'il est possible d'indiquer sommairement à partir de quelques exemples. Considérons des groupes nominaux comme « tous les hommes », « certains nombres » ou « la plupart des chiens ». Pour une structure d'interprétation déterminée dont le domaine d'individus est $\mathcal{D}$, ces expressions sont interprétées comme des ensembles de sous-ensembles de $\mathcal{D}$. « Certains nombres », par exemple, est l'ensemble des sous-ensembles de $\mathcal{D}$ qui contiennent au moins un nombre. « Un nombre pair de chaises » est l'ensemble des sous-ensembles de $\mathcal{D}$ qui contiennent un nombre pair de chaises. En suivant ce principe, il devient possible de définir des quantificateurs comme « la majorité des retraités », « la plupart des États », « un nombre fini de couleurs », « une infinité de points », etc.

L'intérêt d'une telle théorie de la quantification ne tient pas uniquement au plus grand pouvoir d'expression qu'elle est capable de donner à des langages logiques. Du point de vue linguistique, son intérêt majeur est d'offrir un traitement uniformisé de certaines structures syntaxiques comme les groupes nominaux. Un nom propre comme « Socrate » tombe en effet, lui aussi, sous le concept de quantificateur généralisé. Dans une structure d'interprétation dont le domaine d'individus est $\mathcal{D}$, « Socrate » est interprété par l'ensemble de tous les sous-ensembles de $\mathcal{D}$ qui contiennent Socrate ou, si l'on préfère, par l'ensemble de toutes les propriétés de Socrate. Grâce à cette analyse, dont l'origine se trouve chez Montague, « Socrate » peut alors être combiné d'une manière très naturelle à d'autres quantificateurs, ce qui autorise à voir également « Socrate et Platon » ou « tous les Athéniens sauf Socrate » comme des quantificateurs. La théorie des quantificateurs généralisés donne ainsi de ce concept une définition et une sémantique qui s'accordent avec l'une des principales catégories syntaxiques des langues naturelles, celle de groupe nominal. Encore ces quelques exemples ne donnent-ils qu'une faible idée des généralisations envisagées dans cette théorie qui offre également une analyse des déterminants du nom, à l'intérieur du groupe nominal : « tous », « une infinité de », « la plupart des », « une majorité de », etc.[1].

L'esprit dans lequel s'effectue le genre de travaux qu'on vient de décrire est donc très éloigné de toute défiance à l'égard de la grammaire ou des structures et catégories syntaxiques des langues naturelles. Dans ce contexte, la logique n'a pas pour fonction de révéler ce que les logiciens ont parfois appelé « la structure réelle des énoncés », censée être dissimulée par les idiosyncrasies lexicales et grammaticales, mais de contribuer à comprendre quelles sont les structures syntaxiques et sémantiques des langues que nous parlons et à expliquer les étonnantes capacités linguistiques des locuteurs que nous sommes. La question de savoir jusqu'à quel point la logique – qui répond, par ailleurs, à d'autres motivations – pourra guider ou éclairer la recherche en linguistique, et en particulier en sémantique formelle, demeure cependant largement ouverte.

1. Cf. E. L. Keenan et D. Westerståhl, « Generalized Quantifiers in Linguistics and Logic », *in* J. van Benthem et A. ter Meulen, *Handbook of Logic and Language*, Amsterdam, Elsevier, 1997.

Chapitre VI

CALCUL, DÉCISION, COMPLEXITÉ

L'expression « calcul des prédicats » est fréquemment utilisée pour désigner la logique du premier ordre. Dans ce contexte, le mot « calcul » ne désigne pas une opération arithmétique ou algébrique. Il indique un certain usage des systèmes formels devenu si courant aujourd'hui qu'il est souvent confondu avec la pratique ordinaire de la logique. Pour un logicien, travailler dans le cadre du calcul des prédicats signifie : se donner un langage formel non interprété du premier ordre ; considérer la possibilité d'interpréter ce langage par diverses structures d'interprétation ; choisir une méthode de dérivation formelle qui permette de démontrer le théorème de complétude sémantique ; considérer la possibilité de formaliser des théories dans le système formel ainsi défini. Le calcul des prédicats constitue donc un cadre formel pour la formalisation des théories.

Si utile que puisse être cette méthode, elle reste néanmoins assez modeste au regard d'un autre projet qui assigne à la logique une tâche bien plus ambitieuse et donne au mot « calcul » un sens beaucoup plus fort : les systèmes formels pourraient-ils fournir non seulement une méthode de formalisation et de vérification des preuves mais également une méthode de décision ? Plus précisément : étant donné un ensemble de formules Γ

et une formule ϕ quelconques, existe-t-il une méthode permettant de déterminer mécaniquement – par application de règles de calcul – si ϕ est conséquence logique de Γ ? Par exemple, si $\mathcal{L}_A$ est un langage pour la théorie des nombres et si T_A est un ensemble de formules qui représentent les axiomes de Peano dans ce langage, existe-t-il une méthode mécanique permettant de déterminer si une formule quelconque est conséquence logique de ces axiomes ? Existe-t-il, par ailleurs, une méthode de décision pour d'autres théories ?

L'une des fonctions que Leibniz assignait à la langue logique à laquelle il travaillait était précisément de servir de méthode de décision. Dans une lettre au duc de Hanovre, il décrit son projet d'une écriture universelle qui remplirait également les fonctions d'un calcul général :

« [...] cette même écriture serait une espèce d'algèbre générale et donnerait moyen de raisonner en calculant, de sorte qu'au lieu de disputer, on pourrait dire : comptons. Et il se trouverait que les erreurs du raisonnement ne seraient que des erreurs de calcul qu'on découvrirait par des épreuves comme dans l'arithmétique. »[1]

Le projet leibnizien de langue universelle devait permettre de soumettre les raisonnements écrits dans cette langue à un *calculus ratiocinator*, calcul algébrique qui aurait rempli les fonctions d'une méthode de décision capable de trancher toutes sortes de débats.

Dans ce qui peut être considéré comme l'un des premiers manuels de la logique moderne, les *Principes*

1. C. J. Gerhardt (éd.), *Die philosophischen Schriften von G. W. Leibniz*, Berlin, Weidmann, 1890 ; Hildesheim, Olms, 1978, t. VII, p. 26.

de la logique théorique (Grundzüge der theoretischen Logik) parus en 1928, Ackermann et Hilbert donnèrent – dans des termes qui ne sont pas exactement ceux que les logiciens utilisent aujourd'hui – un sens bien particulier à l'expression « problème de la décision » : existe-t-il une méthode mécanique permettant de déterminer si une formule du calcul des prédicats est universellement valide ? Pour distinguer ce sens précis du problème, les logiciens utilisent le terme allemand d'*Entscheidungsproblem*. En 1928, Ackermann et Hilbert pouvaient encore imaginer la possibilité d'une solution positive de ce problème. Ils pouvaient également imaginer la possibilité d'un système formel adéquat pour l'arithmétique qui soit syntaxiquement complet. Un tel système aurait permis de trancher par des moyens purement mécaniques tous les problèmes de la théorie des nombres. Pour déterminer si ϕ est un théorème, il aurait en effet suffit d'énumérer de manière systématique toutes les dérivations formelles qui peuvent être écrites dans le système jusqu'à ce qu'apparaisse une dérivation formelle de ϕ ou de non-ϕ à partir des axiomes de la théorie des nombres.

En 1931, la démonstration du premier théorème d'incomplétude établit qu'on ne pouvait espérer résoudre positivement ni le problème de la décision pour la théorie des nombres ni l'*Entscheidungsproblem* par un tel procédé d'énumération des dérivations formelles (en réalité, le premier problème se réduit au second : toute solution de l'*Entscheidungsproblem* est convertible en une solution du problème de la décision pour la théorie des nombres). Bien que cette démonstration ait rendu très peu vraisemblable une solution positive du problème, elle n'excluait pas entièrement la découverte d'une autre méthode de décision. Pour prouver qu'il n'existe aucune méthode mécanique permettant de résoudre l'*Entscheidungsproblem*, il était nécessaire de disposer d'une définition précise du concept de méthode mécanique.

Les logiciens de l'époque avaient aussi une autre raison de chercher une définition de ce concept. La démonstration des théorèmes d'incomplétude qu'avait donnée Gödel en 1931 s'appliquait en effet uniquement aux systèmes apparentés aux *Principia Mathematica.* Afin de prouver que ces théorèmes valent universellement, il fallait disposer d'un concept de système formel qui soit le plus général possible. Or, l'une des conditions nécessaires pour qu'un système de signes soit un système formel est qu'on puisse déterminer *par des moyens mécaniques* si une suite de signes quelconque appartient à l'ensemble des formules du système, si une formule quelconque appartient à l'ensemble des axiomes et si une suite de formules quelconque est une dérivation formelle.

Dans la démonstration des théorèmes d'incomplétude, Gödel avait utilisé un procédé d'arithmétisation permettant de coder les formules et les démonstrations d'un système S par des entiers naturels. Ce procédé avait pour but de représenter certaines propriétés des formules de S (par exemple « être démontrable dans S ») par des propriétés arithmétiques (par exemple $\exists x Dem(x,y)$, où $Dem(x,y)$ est une formule qui signifie « x est le code d'une démonstration de la formule codée par y »). À condition qu'une certaine partie de l'arithmétique puisse être formalisée dans S, un énoncé du métalangage comme « la formule ϕ n'est pas démontrable dans S » pouvait donc être exprimé *dans* S lui-même par la formule $\neg \exists x Dem(x,a)$, où a était le code de ϕ. Gödel montra, en outre, qu'il existe un entier b qui est le code de $\neg \exists x Dem(x,b)$. Cette formule ϕ_G de l'arithmétique élémentaire signifiait donc « ϕ_G n'est pas démontrable dans S » ; ϕ_G exprimait ainsi au sujet d'elle-même qu'elle n'était pas démontrable dans S. La démonstration du premier théorème d'incomplétude établit alors que si S est cohérent, ϕ_G n'est ni démontrable ni réfutable dans S et qu'aucune extension de S obtenue par ajout de nouveaux axiomes n'est complète si elle est cohérente.

L'usage d'un codage par des entiers naturels se généralisa par la suite et la question d'une définition du concept de procédure effective ou de procédure mécanique put être ramenée à la recherche d'une caractérisation des fonctions numériques calculables, c'est-à-dire des fonctions d'entiers dont les valeurs pouvaient être calculées de manière effective. L'*Entscheidungsproblem* se ramenait donc au problème suivant : étant donné un codage gödélien des formules d'un langage pour le calcul des prédicats, existe-t-il une fonction calculable qui à chaque entier n associe 0 si n est le nombre de Gödel d'une formule valide et 1 dans le cas contraire ? Toute solution négative supposait une définition précise du concept de fonction numérique calculable.

Une fonction numérique f est calculable s'il existe un algorithme qui permet de calculer $f(n)$ pour chaque entier n. Mais qu'est-ce qu'un algorithme ? Les réponses habituelles restent très informelles : un algorithme pour la fonction f est une méthode générale qui décrit de manière parfaitement déterminée chacune des opérations élémentaires successives qu'il convient d'effectuer afin de calculer $f(n)$ pour un entier naturel n quelconque. Au cours des années 1930, plusieurs logiciens imaginèrent des formalismes par lesquels ils proposaient de caractériser précisément le concept de fonction calculable. Les plus connus sont les machines de Turing, les termes du lambda-calcul, dus à Church, et, grâce aux travaux de Herbrand et Gödel, les fonctions récursives. Les recherches de ces logiciens, comme celles de Kleene, Rosser et d'autres, aboutirent à une série de résultats tout à fait remarquables.

Premièrement, les différentes méthodes proposées pour caractériser l'ensemble des fonctions calculables se révèlent équivalentes : l'ensemble des fonctions calculables par une machine de Turing est exactement l'ensemble des fonctions définissables par un terme du lambda-calcul, qui est exactement l'ensemble des fonctions récursives. Cette convergence est l'un des arguments en faveur de la thèse selon laquelle l'ensemble ainsi défini donne une caractérisation adéquate de ce qu'on entend informellement par « fonction calculable ». La thèse de Church, publiée en 1936, identifie les fonctions calculables aux fonctions récursives. La thèse de Turing, publiée la même année, identifie les fonctions calculables aux fonctions calculables par une machine de Turing. Les deux thèses étant équivalentes, on parle aussi de la thèse de Church-Turing.

Deuxièmement, il n'existe aucune fonction récursive – et donc, si l'on accepte la thèse de Church, aucune fonction calculable – de l'ensemble des entiers naturels dans $\{0,1\}$ qui à n associe 0 si n est le nombre de Gödel d'une formule valide et 1 dans le cas contraire. Autrement dit : la solution de l'*Entscheidungsproblem* est négative. Ce théorème d'indécidabilité fut démontré par Church en 1936 et par Turing, de manière indépendante, la même année[1]. Le problème de la validité d'une formule est néanmoins *décidable* dans le cas particulier du calcul des prédicats monadiques, c'est-à-dire lorsque le langage ne comporte aucun symbole de relation à deux places ou plus. En outre, certaines théories mathématiques *sont* décidables : c'est le cas de la théorie des ordres denses, de la théorie des corps réels fermés ou de l'arithmétique de Presburger (dans laquelle on ne dispose pas de la multiplication)[2].

Troisièmement, l'ensemble des fonctions calculables (dans tout ce paragraphe, en vertu de la thèse de Church-Turing, on pourra remplacer à sa guise l'expression « fonction calculable » par « fonction récursive », « fonction calculable par une machine de Turing » ou « fonction représentable par un terme du lambda-calcul ») est un ensemble de fonctions *partielles*. Si les fonctions calculables sont vues comme des procédés de calcul applicables à des nombres, cela signifie que certains de ces procédés, appliqués à certains nombres, se poursuivent indéfiniment sans qu'aucun résultat ne soit jamais produit. Une fonction calculable est *totale* si, à l'inverse, elle est définie pour chacun de ses arguments. Or, l'ensemble des fonctions totales n'est pas décidable, ce qui signifie qu'il n'existe aucune procédure mécanique permettant de déterminer, pour une fonction calculable quelconque, si elle est totale. De tels résultats

1. Les textes originaux de Church et Turing (qui exposent également des arguments en faveur des thèses, respectivement, de Church et de Turing) sont traduits dans une anthologie de la théorie de la calculabilité dirigée par J. Dubucs et J. Mosconi, à paraître à Paris, chez Cassini.

2. Cf. R. David, K. Nour et Ch. Raffalli, *Introduction à la logique*, Paris, Dunod, 2001, 2e éd., 2004.

conduisirent à la définition de nouveaux concepts en théorie de la calculabilité, comme celui d'ensemble *récursivement énumérable*. Un sous-ensemble E de l'ensemble des entiers naturels est récursivement énumérable, ou semi-décidable, s'il existe une fonction calculable partielle dont le domaine de définition est exactement E ; ou, de manière équivalente : s'il existe un algorithme qui énumère les éléments de E. Par ailleurs, E est non seulement récursivement énumérable mais également *récursif*, ou décidable, s'il existe une fonction calculable totale f telle que $f(n) = 1$ si n appartient à E et $f(n) = 0$ dans le cas contraire. À titre d'exemple, l'ensemble des formules valides du calcul des prédicats, bien que non récursif, est néanmoins récursivement énumérable, de même que l'ensemble des théorèmes de la théorie des nombres formalisée dans un système formel du premier ordre.

Le théorème d'indécidabilité donnait une raison précise de penser que le projet leibnizien de *calculus ratiocinator* était utopique ; il permettait aussi de comprendre pourquoi l'*Entscheidungsproblem* ne pouvait recevoir de solution positive et pourquoi la question de la vérité des énoncés mathématiques ne pouvait pas être tranchée par une procédure générale de calcul. Mais les résultats des recherches relatives au concept de calcul sont loin d'avoir été entièrement négatifs et ce n'est certainement pas un hasard si des logiciens comme Turing et von Neumann sont à l'origine non seulement de la théorie de la calculabilité, mais aussi des premiers calculateurs universels programmables qu'on commença à construire dans les années 1940. Les machines de Turing et les termes du lambda-calcul sont des outils conceptuels mathématiquement définissables qui ont servi à résoudre des problèmes de logique mais qui, dans le même temps, peuvent être interprétés comme des programmes informatiques[1].

Une machine de Turing peut être représentée comme un mécanisme doté 1) d'un ruban infini divisé en cases dans lesquelles peuvent être inscrits des symboles d'un alphabet ; 2) d'une tête de lecture-écriture capable de se déplacer de case en case le long du ruban ; 3) d'un programme qui détermine le comportement de la machine : en fonction du symbole qui se trouve dans la case lue et d'un paramètre appelé « état de la machine », la tête de lecture-écriture peut effacer ce symbole, en écrire un autre à la place, se déplacer d'une case à droite ou à gauche, ou effectuer plusieurs de ces opérations. Chaque machine de Turing est donc définie par un programme qui, appliqué à un argument n (une suite de signes inscrits sur le ruban), détermine une succession d'opérations à effectuer afin de trouver la valeur en n d'une fonction calculable f (cette valeur est donnée par les signes inscrits sur le ruban lorsque la machine s'arrête). Dans le cas d'une fonction partielle, si $f(n)$ est indéterminé, la suite de ces opérations est sans fin et la machine ne s'arrête pas, elle ne produit aucun résultat. En 1936, Turing montre d'une part qu'aucune machine de Turing ne permet de déterminer si une machine de Turing quelconque s'arrêtera et d'autre part que l'*Entscheidungsproblem* n'admet de solution que si ce problème de l'arrêt en admet une.

1. G. Boolos, R. Jeffrey, *Computability and Logic*, Cambridge, Cambridge UP, 1974, 5e éd., 2007 ; P. Wolper, *Introduction à la calculabilité*, Paris, InterÉditions, 1991 ; 3e éd., Paris, Dunod, 2006 ; P. Wagner, *La Machine en logique*, Paris, PUF, 1998.

Les termes du lambda-calcul, quant à eux, sont des expressions d'un langage formel qui peuvent représenter aussi bien une fonction f qu'un argument n. Or, l'application d'un terme qui représente f à un terme qui représente n correspond très précisément à l'application d'un programme pour f à la donnée n, tandis que l'opération qu'on nomme « β-réduction » correspond précisément au calcul de $f(n)$ par ce programme.

Dans les systèmes de lambda-calcul typés, on accole aux termes une expression qui indique leur type : entier, fonction d'entiers, fonction de fonction d'entiers, etc. Par exemple, si N représente le type « entier », l'expression « $t : N$ » indique que le type du terme t est celui des entiers alors que « $t' : N \rightarrow N$ » indique que le type de t' est celui des fonctions d'entiers dans les entiers. Les règles du lambda-calcul typé autorisent alors l'application de t' à t (ce qui produit un terme de type N), mais non celle de t à t' ou à t. On réduit ainsi le pouvoir d'expression du système en question, mais on s'assure, dans le même temps, qu'un terme qui est appliqué à un autre est bien du type attendu.

La caractérisation de l'ensemble des fonctions calculables par les machines de Turing ou par d'autres formalismes apporte une clarification conceptuelle importante et jette un pont entre la logique et l'informatique, car la théorie de la calculabilité peut être conçue comme une théorie des capacités de ces calculateurs universels programmables que sont les ordinateurs. Mais les fonctions calculables présentent des différences considérables. Leur ensemble compte en effet aussi bien des fonctions dont les valeurs, pour chaque argument, se calculent très facilement (les fonctions constantes par exemple) que des fonctions dont les valeurs ne peuvent être obtenues qu'au moyen de ressources de calcul si énormes qu'il est pratiquement, et même physiquement, tout à fait impossible de les calculer. Afin de prendre en considération ce genre de différence, la théorie de la complexité introduit des concepts plus fins que celui de fonction calculable : la complexité d'un programme pour la fonction f est mesurée par le temps et l'espace nécessaires à l'obtention de la valeur de f pour l'argument n. On ne s'intéresse pas seulement à ce qui est calculable en principe, indépendamment de toute contrainte physique, mais également à la caractérisation des algorithmes qui peuvent effectivement être mis en œuvre dans les conditions réelles[1].

L'origine des théories de la calculabilité et de la complexité remonte aux questions de décidabilité soulevées par les logiciens. Ces théories ne donnent pourtant qu'une première idée des rapports étroits qui se sont tissés, surtout au cours des dernières décennies, entre logique et informatique. Ils trouvent aujourd'hui leur point culminant dans l'isomorphisme de Curry-Howard qui établit une correspondance entre les preuves et les programmes, les formules et les types de données, la normalisation des preuves et le déroulement d'un programme qui effectue un calcul. Bien que l'importance de cette correspondance n'ait commencé à être comprise qu'à la fin des années 1960, ses origines remontent aux questions soulevées par le programme de Hilbert, aux travaux de Gentzen sur la cohérence de l'arithmétique

1. Cf. R. Lassaigne et M. de Rougement, *Logique et complexité*, Paris, Hermès, 1996.

et à la conception intuitionniste des mathématiques défendue par Brouwer.

En 1934, dans le cadre de ses recherches sur la théorie de la démonstration, Gentzen introduit deux méthodes de dérivation formelle : le calcul des séquents et la déduction naturelle. L'intérêt de ces formalismes réside dans la clarification qu'ils apportent sur les preuves mathématiques. Ils constituent de puissants instruments pour l'étude des propriétés, de la structure et de la complexité des preuves. L'un des principaux résultats obtenus par Gentzen est le théorème d'élimination des coupures pour le calcul des séquents. Sa signification intuitive est que toute preuve d'une formule valide (de la logique du premier ordre) dans le calcul des séquents peut être obtenue de manière directe, c'est-à-dire sans avoir recours à la règle du calcul qu'on nomme « règle de coupure ».

Typiquement, selon une version simplifiée de la règle de coupure, si l'on a une déduction de C à partir de B, et par ailleurs une déduction de B à partir de A, on peut inférer directement qu'il existe une déduction de C à partir de A. Éliminer les coupures utilisées dans une preuve a pour effet, en général, d'augmenter la longueur de la preuve. Gentzen utilisait le théorème d'élimination des coupures pour prouver que les règles du calcul des séquents, et donc les règles utilisées en mathématiques, ne sont pas contradictoires. Du point de vue du programme de Hilbert, ce théorème répond à un idéal de pureté des méthodes : l'un de ses corollaires est en effet que pour prouver la formule A, il n'est nullement besoin d'avoir recours à des formules autres que les sous-formules de A. En particulier, si A est un énoncé élémentaire, il existe une démonstration de *A* qui ne fait aucun détour par des énoncés non élémentaires[1].

Du point de vue mathématique, l'intérêt du théorème principal de Gentzen est très limité : d'une part, les démonstrations sans coupure sont généralement beaucoup trop longues pour être utiles en pratique ; d'autre part, le théorème vaut pour la logique du premier ordre mais n'est pas vérifié de manière générale dans le cas d'un système qui comporte des axiomes non logiques. Le théorème est en revanche du plus haut intérêt lorsqu'il est établi pour des systèmes logiques d'ordre supérieur qui permettent de représenter les entiers et de formaliser l'arithmétique ainsi qu'une grande partie des mathématiques sans adjonction d'axiomes non logiques. Il apparaît alors que l'élimination des coupures – ou « normalisation des preuves » – peut être interprétée comme le déroulement d'un programme qui effectue un calcul. Pour éclaircir cet aspect de l'isomorphisme de Curry-Howard, il convient d'apporter quelques précisions relatives au caractère constructif de la logique intuitionniste.

Brouwer, mathématicien de génie, n'aurait certainement pas apprécié qu'on choisisse de parler de sa conception des mathématiques, l'intuitionnisme, à l'occasion d'une réflexion sur la logique et l'informatique. Dans les premières décennies du XX^e siècle, il s'opposa à la fois au logicisme et au formalisme et ne pensait nullement que les mathématiques, qui étaient, pour lui, une construction mentale du mathématicien,

1. Sur la déduction naturelle, le calcul des séquents et l'élimination des coupures, cf. R. David, K. Nour, Ch. Raffalli, *Introduction à la logique*, Paris, Dunod, 2001, 2^e éd., 2004.

devaient être fondées par la logique. La formalisation de la logique intuitionniste, réalisée par Heyting[1], est donc une entreprise qui va à l'encontre de la philosophie de Brouwer. Elle n'en est pas moins éclairante pour la question qui nous occupe car elle met en lumière une interprétation constructiviste des mathématiques et des constantes logiques. Par exemple, si p et q sont deux énoncés, $(p \rightarrow q)$ n'est pas interprété en termes de conditions de *vérité* comme dans la logique classique (dans laquelle $(p \rightarrow q)$ est vrai si p est faux ou q est vrai) mais en termes de conditions de *prouvabilité* : $(p \rightarrow q)$ est prouvable si l'on dispose d'une méthode pour convertir toute preuve de p en une preuve de q. De même, $\forall x \exists y Rxy$ est prouvable s'il existe une méthode permettant de produire un y tel que Rxy pour n'importe quel x. Alors que la loi du tiers exclu est universellement valide en logique classique, la formule $(\phi \vee \neg \phi)$ n'est prouvable, en logique intuitionniste, que si l'on dispose d'une preuve de ϕ ou d'une preuve de non-ϕ. De manière similaire, $\exists x P(x)$ n'est prouvable en logique intuitionniste que si l'on peut exhiber un objet a et montrer qu'il satisfait la propriété P.

Or, le caractère constructif de la logique intuitionniste, dont le sens originel se trouve dans la philosophie des mathématiques de Brouwer, peut être interprété en termes de calcul. Ainsi, une preuve intuitionniste de $\forall x \exists y Rxy$ donne une méthode effective, et donc un algorithme, qui, appliqué à un objet a,

1. A. Heyting, *Intuitionism: an Introduction*, Amsterdam, North-Holland, 1956.

produit un objet b tel que Rab. En 1969, Howard[1] remarquait l'existence d'une correspondance précise entre les preuves intuitionnistes et les termes du lambda-calcul. Plus précisément, il est possible de concevoir des systèmes de lambda-calcul typé qui soient tels qu'un terme t de type A corresponde précisément à la preuve de l'énoncé A, et tels que la β-réduction de t corresponde précisément à l'élimination des coupures dans la preuve de A. Pour prendre un exemple précis, l'application du terme t de type $A \rightarrow B$ au terme t' de type A produit le terme $(t)t'$ de type B, ce qui correspond à l'application de la règle logique du *modus ponens* selon laquelle si l'on a A et $(A \rightarrow B)$, on peut en déduire B. Loin d'être arbitraire, la correspondance entre preuves et programmes et entre formules et types peut être généralisée. L'isomorphisme de Curry-Howard met ainsi en évidence d'étroites relations entre logique et informatique[2].

L'ensemble de ces relations a un intérêt évident au point de vue informatique. D'une part, les méthodes logiques sont appliquées, de manière générale, à l'étude, au contrôle et à la création des programmes et des langages de programmation ; d'autre part, l'étude des systèmes de lambda-calcul typé et la découverte de l'isomorphisme de Curry-Howard a eu d'importantes conséquences sur un problème majeur en informatique : garantir la fiabilité et la terminaison des programmes.

1. W. A. Howard, « La notion de construction où les formules sont des types » [1969], trad. franç. à paraître dans l'anthologie de la calculabilité signalée *supra*, p. 88, n° 1.
2. Cf. J.-Y. Girard, Y. Lafont, P. Taylor, *Proofs and Types*, Cambridge, Cambridge UP, 1989.

Bien que nous sachions, par le théorème d'indécidabilité de Church, qu'il n'existe aucun programme capable de déterminer si un programme quelconque s'arrêtera, il existe des systèmes de lambda-calcul typé qui permettent de programmer un très grand nombre de fonctions et de garantir la terminaison des programmes qui les calculent.

Qu'en est-il, inversement, pour la logique ? Que nous apprend, de ce point de vue, l'isomorphisme de Curry-Howard ? La découverte d'une correspondance entre preuves et programmes a-t-elle un intérêt purement technologique (exercer un meilleur contrôle sur les programmes, augmenter les performances des machines) ou modifie-t-elle notre compréhension de ce que sont la logique et les mathématiques ? On sera d'autant plus enclin à donner une réponse positive à la seconde branche de l'alternative qu'on place la théorie de la démonstration au cœur de la logique ou qu'on est prêt à définir celle-ci comme la science des preuves – ce qui ne va, évidemment, nullement de soi. La réponse dépend également de l'extension qu'il sera possible de donner à l'isomorphisme de Curry-Howard dans l'avenir. Le fait que cet isomorphisme ne se limite pas à la logique intuitionniste et qu'il ait été possible, depuis les années 1990, de l'étendre non seulement à la logique classique, mais également aux axiomes de la théorie des ensembles conduit certains logiciens, comme Jean-Louis Krivine, à donner une nouvelle interprétation de l'activité du mathématicien : celle-ci pourrait être vue comme un travail d'écriture de programmes dans un langage de bas niveau.

Certains logiciens vont beaucoup plus loin et proposent une réinterprétation globale de la logique elle-même, de sa nature et de sa fonction. Tel est le cas de Jean-Yves Girard, théoricien de la démonstration à qui l'on doit notamment la logique linéaire[1] et le concept de réseaux de preuves. La logique linéaire introduit de nouveaux connecteurs propositionnels qui sont interprétés en termes de *ressources*. Ainsi, l'implication $A \multimap B$ signifie qu'on a B en *utilisant* A, ce qui a pour conséquence qu'au cours d'une démonstration, on ne peut pas utiliser deux fois la même formule $A \multimap B$ pour obtenir B à partir de A. Le connecteur de la conjonction, quant à lui, est remplacé par deux connecteurs : $A \otimes B$ (qui se lit « A fois B ») signifie qu'on dispose des ressources pour A et pour B, alors que $A \& B$ (qui se lit « A avec B ») signifie qu'on ne dispose des ressources que pour l'un des deux. Pour prendre un dernier exemple de connecteur, l'expression « !A » signifie qu'on dispose de A autant de fois qu'on veut. La logique linéaire, ainsi que ses variantes (logique linéaire bornée, logique linéaire allégée) a de nombreuses applications informatiques. Mais le programme de recherche de son inventeur est bien plus ambitieux : proposant une réinterprétation de tout le développement de la logique depuis le programme de Hilbert, il vise une reconstruction globale de la logique en tant que telle[2]. Une telle perspective va alors bien au-delà des questions évoquées dans le présent chapitre, centré sur le calcul, la décision et la complexité.

1. J.-Y. Girard, « Linear logic », *Theoretical Computer Science*, 50, 1987, p. 1-102.
2. Cf. J.-Y. Girard, « La théorie de la démonstration : du programme de Hilbert à la logique linéaire », *in* Godefroy, Girard *et al.*, *Leçons de mathématiques d'aujourd'hui*, Paris, Cassini, vol. II, 2003, p. 37-100 ; « From Foundations to Ludics », *Bulletin of Symbolic Logic*, 9. 2, 2003.

Chapitre VII

LOGIQUE ET CONNAISSANCE

Un lecteur uniquement formé à la logique mathématique et qui aurait la curiosité d'ouvrir l'un des ouvrages, souvent volumineux, qui furent publiés au XIXe siècle et jusqu'au début du XXe siècle sous le titre de *Logique* aurait toutes les chances d'être étonné par ce qu'il pourrait y trouver. Cet étonnement ne tiendrait pas seulement à la forme non mathématisée de l'exposition ou à la division tripartite des matières fréquemment adoptée avant la révolution frégéenne : analyse du concept, du jugement et du raisonnement. L'étonnement résulterait également du but assigné à la logique dans la plupart de ces livres ainsi que de la manière dont elle y est ancrée dans des problèmes de méthodologie des sciences ou de théorie de la connaissance.

La logique moderne, postfrégéenne, n'a certes pas rompu avec les questions de méthodologie scientifique, comme en témoignent la publication de nombreux ouvrages sur le sujet[1], l'existence de groupes de

1. Parmi les classiques, citons M. Cohen et E. Nagel, *An Introduction to Logic and Scientific Method*, Londres, Routledge & Kegan Paul, 1934, et Tarski, *Introduction à la logique*, Paris, Gauthier-Villars, 1960, trad. franç. d'après la version anglaise révisée et augmentée parue sous le titre *Introduction to Logic and to the Methodology of Deductive Sciences* en 1941 (la version originale en polonais était parue en 1936).

recherche qui travaillent sur ces questions (par exemple le « Group in Logic and the Methodology of Science » de l'université de Berkeley) ainsi que la réunion quadriennale, depuis 1960, du Congrès international de logique, méthodologie et philosophie des sciences. D'un autre côté, les développements actuels de la logique épistémique proposent un traitement formel spécifique des concepts de croyance et de connaissance. Mais alors que les orientations méthodologique ou épistémique ne représentent aujourd'hui que quelques ramifications particulières parmi les innombrables branches de la logique contemporaine, les questions de théorie de la connaissance ou de méthodologie scientifique apparaissent au contraire, dans nombre d'ouvrages publiés entre 1830 et 1920, comme les principaux cadres dans lesquels s'inscrivait la logique formelle de l'époque et ce qui constituait sa principale raison d'être. Le point important n'est pas seulement que le rapport de la logique à la connaissance a beaucoup changé depuis l'avènement de la logique moderne ; il est aussi que ce rapport est aujourd'hui encore en pleine mutation et que les différentes manières dont il est ou a pu être envisagé correspondent souvent à des conceptions fondamentalement différentes de ce qu'est la logique.

Bien que les deux dimensions, méthodologique et épistémique, touchent l'une et l'autre aux relations entre logique et connaissance et qu'elles n'aient pas toujours été conçues indépendamment l'une de l'autre, elles doivent être nettement distinguées. Dans ce qui suit, il sera question de l'une et de l'autre.

Les questions de méthodologie scientifique sont abondamment discutées dans le *Système de logique* de

Mill[1], qui traite non seulement du nom, de la proposition, de la définition ou du rapport entre les expressions du langage et ce qu'elles signifient, mais également de l'induction, des lois de la nature, de la loi de causalité et des méthodes de la recherche expérimentale, tant dans les sciences humaines ou sociales que dans les sciences de la nature. Pour Mill, « la logique est la science des opérations intellectuelles qui servent à l'*estimation* de la preuve »[2], ce qui place le logicien en position de juge pour toutes les recherches particulières qui visent à établir la vérité dans quelque domaine de la connaissance que ce soit. De ce point de vue, la logique est elle aussi une science, mais son étendue est précisément limitée par l'utilité pratique des connaissances qu'elle produit. La finalité des études logiques, pour Mill, est d'ordre pédagogique et méthodologique : trouver les règles qui nous permettent de porter un jugement sur les arguments par lesquels nos connaissances se trouvent être justifiées, afin d'améliorer notre capacité de raisonner.

Wilhelm Wundt, connu comme l'un des fondateurs de la psychologie expérimentale, est également l'auteur d'un traité de logique en deux volumes plusieurs fois réédité et dont le titre exact est *Logique : une recherche des principes de la connaissance et des méthodes de la recherche scientifique*[3]. Pour Wundt, la logique a pour

1. J. S. Mill, *Système de logique déductive et inductive : exposé des principes de la preuve et des méthodes de recherche scientifique*, 1843 ; trad. franç. de la 6e éd., 1866 ; rééd. Bruxelles, Mardaga, 1995.

2. J. S. Mill, *Système de logique*, p. 11-12.

3. W. Wundt, *Logik. Eine Untersuchung der Principien der Erkenntnis und der Methoden wissenschaftlicher Forschung*, Stuttgart, Enke, vol. I : *Erkenntnislehre*, 1880 ; vol. II : *Methodenlehre*, 1883 ; rééd. en trois volumes : 1893, 1906, 1921, 1924.

tâche de fonder les lois de la pensée qui se présentent, de fait, comme des normes de la recherche scientifique. La logique est donc normative dans la mesure où elle énonce des règles qui régissent la recherche dans les sciences, mais la connaissance de ces règles suppose une investigation de type psychologique. Avant Wundt, le philosophe et logicien Christoph von Sigwart avait publié une *Logique* en deux volumes[1] dans laquelle il définissait cette discipline comme une technologie de la pensée ayant pour but de donner les critères de la pensée vraie. Sigwart assignait à la logique une tâche essentiellement pratique et une orientation clairement psychologique. Plus tôt encore, Friedrich Beneke avait également défendu une position psychologiste dans son *Système de la logique comme technologie de la pensée*[2].

C'est dans la perspective d'une critique de telles conceptions psychologistes de la logique, très courantes à l'époque, que Paul Natorp fit paraître en 1887 l'article intitulé « Fondation objective et fondation subjective de la connaissance »[3]. Pour Natorp, qui fut, avec Hermann Cohen et Ernst Cassirer, l'un des principaux représentants du néokantisme de l'École de Marbourg, la logique est en effet tout autre chose qu'une doctrine de la méthode ou une discipline de la raison, et si elle doit être étendue au-delà de la syllogistique et de la logique formelle, ce n'est certainement pas dans la

1. Ch. von Sigwart, *Logik*, Tübingen, Laupp, vol. I, *Die Lehre vom Urteil, von Begriff und vom Schluss*, 1873 ; vol. II, *Die Methodenlehre*, 1878.

2. F. Beneke, *System der Logik als Kunstlehre des Denkens*, Berlin, Dümmler, 1842 ; rééd. Hildesheim, Olms, 2005.

3. Cet article est traduit dans H. Cohen, P. Natorp *et al.*, *Néokantismes et théorie de la connaissance*, Paris, Vrin, 2000.

direction d'une logique inductive ou d'une méthodologie de la science. Car la logique, selon Natorp, est de nature transcendantale, ce qui signifie qu'elle est la science des relations les plus générales qui existent entre pensée et objet et, à ce titre, constitutive de la connaissance scientifique.

Kant avait distingué la logique générale pure, qui expose les règles formelles de toute pensée, indépendamment de tout contenu de connaissance, et la logique transcendantale qui détermine « l'origine, la valeur et l'étendue » des connaissances rationnelles par lesquelles « nous pensons des objets tout à fait *a priori* »[1]. Les philosophes néokantiens de l'École de Marbourg reprennent l'idée de ce que Kant avait appelé « logique transcendantale », mais ils estiment, contrairement à Kant, que ce n'est pas dans une philosophie des facultés que doivent être recherchés les fondements *a priori* de la connaissance objective. Pour ces philosophes, c'est sur la science elle-même considérée comme un fait, et sur le devenir des sciences, qu'il convient de s'appuyer pour déterminer les conditions de possibilité de la connaissance des objets en général. La « logique pure » qu'ils appellent de leurs vœux ne se réduit donc ni à la logique formelle, ni à la logique transcendantale au sens où l'entendait Kant, ni à l'énoncé d'un ensemble de normes utiles au point de vue de la méthodologie des sciences. Sa tâche est la détermination des lois de l'objectivité scientifique et des principes qui justifient et rendent possible toute connaissance en général. Cette logique pure vise les fondements de toute connaissance

1. E. Kant, *Critique de la raison pure*, 1781, A 57 ; 2e éd., 1787, B 81. Cf. *supra*, p. 15-16.

rationnelle, ce qui explique qu'elle ne puisse reposer ni sur la psychologie ni sur une quelconque enquête empirique. Ayant affaire au contenu objectif des pensées et non à la subjectivité des états mentaux, elle ne part ni des états psychologiques du sujet connaissant, ni de ses facultés de connaissance, mais d'une analyse critique de la science prise comme un fait, de ses lois, de ses méthodes et de son évolution en tant qu'elle est déterminée par des règles[1].

En 1900, Husserl critique lui aussi les conceptions psychologistes de la logique et la plus grande partie de ses *Prolégomènes à la logique pure*[2] est consacrée à cette tâche. Chez lui, comme chez les philosophes néokantiens, la question de la logique est profondément ancrée dans le problème général du fondement de la connaissance. Aussi défend-il, lui aussi, l'idée d'une logique pure. Ce qu'il entend par là est cependant très différent car la fonction de cette « théorie des théories » est de dire quels sont les concepts et les lois qui forment les conditions de possibilité de toute théorie en général. Ce qu'il nomme également « science des sciences »[3] précède, pour cette raison, toute autre science et vise *« l'essence idéale de la science en tant que telle »*[4]. Afin d'expliciter la tâche de la logique pure, Husserl distingue trois niveaux d'investigation. Le premier concerne les lois par lesquelles le sens se distingue du non-sens et qui appartiennent à la grammaire logique pure ; ces lois

1. Cf. P. Natorp, *Die logischen Grundlagen der exakten Wissenschaften*, Leipzig, Teubner, 1910, p. 12-13.
2. Husserl, *Recherches logiques*, vol. I : *Prolégomènes à la logique pure*, 1900, trad. franç., Paris, PUF, 1958.
3. Husserl, *Prolégomènes à la logique pure*, chap. XI, § 66, p. 269.
4. *Ibid.*, p. 267.

a priori déterminent « les formes possibles de signification »[1]. À un deuxième niveau, la logique pure traite des formes possibles du jugement vrai. Ici, les lois logiques qui sont élucidées sont celles qui servent à prévenir la contradiction et elles ne font plus, comme au niveau précédent, abstraction de la validité objective des significations. À un troisième niveau, enfin, la logique pure est « théorie des formes de théories possibles »[2].

Selon quelle méthode cette logique pure procède-t-elle ? Husserl ne part ni de la factualité des sciences, ni des facultés du sujet connaissant mais des actes mentaux donateurs de sens par lesquels une science se constitue. Selon Husserl, en effet, une science est un ensemble déductivement clos de significations données par des actes cognitifs mais qui n'ont pas, en tant que telles, d'existence séparée. Pour autant, cette méthode n'implique aucun retour à la psychologie, car si ces actes mentaux ont un aspect subjectif et si Husserl parle effectivement de « psychologie descriptive », ce qui est visé, dans la description de ces actes, est un corrélat idéal qui a valeur objective. L'apparente incompatibilité entre la subjectivité de l'acte de connaissance donateur de sens et le caractère objectif du contenu de la connaissance soulève une difficulté que la méthode phénoménologique doit permettre de résoudre. La logique pure exige donc, selon Husserl, une élucidation de l'origine de la signification, et par conséquent un examen des

1. Husserl, *Recherches logiques*, vol. II, 2e partie, Recherche logique IV.
2. Husserl, *Prolégomènes à la logique pure*, chap. XI, § 69, p. 272. Cf. aussi J. Benoist, « Husserl et la fascination du formel », *in* P. Wagner (dir.), *Les Philosophes et la science*, Paris, Gallimard, 2002.

vécus intentionnels et de leur contenu, examen qu'il distingue soigneusement de toute étude psychologique de la genèse des représentations.

Frege se livre, lui aussi, à une sévère critique du psychologisme en logique. Mais contrairement à Husserl, il ne conçoit pas la logique comme une théorie des théories et ne l'envisage pas dans son rapport à une critique des sciences ou à une théorie de la connaissance. Son objectif initial est de montrer que, contrairement à ce que pensait Kant, il n'est pas nécessaire d'avoir recours à l'intuition pure pour fonder la science des nombres. À la différence de Husserl, Frege ne croit pas non plus devoir chercher l'origine des contenus de signification idéaux dans des actes de connaissance. Pourtant, la question de la signification est bien au cœur de la logique frégéenne ; mais Frege cherche à éclaircir directement le rapport du langage à ce à quoi il se rapporte, sans revenir à l'analyse d'actes donateurs de sens. La fonction de l'idéographie, à cet égard, est de rendre aussi transparents que possible les rapports logiques entre contenus de signification. Si cela est possible, c'est parce que Frege, comme Bolzano avant lui, pose l'existence d'un domaine autonome, celui des pensées, qui ne sont, dans ce contexte, ni des représentations mentales ni des choses physiques. Il s'agit d'entités qui appartiennent à un « troisième royaume » et que Frege identifie au sens des phrases déclaratives. De telles phrases « expriment » les pensées en question. Ce qu'on nomme « théorème de Pythagore », par exemple, est une entité objective qui ne saurait être réduite aux diverses représentations mentales de ce théorème que chacun peut former en son esprit. Aussi n'est-il pas *produit* par l'activité de pensée, mais *saisi* par cette activité[1]. Or, parce que le prédicat *vrai*, selon Frege, s'applique aux pensées objectives ainsi comprises et non aux représentations ni aux phrases, c'est à elle qu'a affaire la logique, comme science de l'être-vrai.

Si l'objectif de Frege est de fonder l'arithmétique dans la logique, il ne cherche pas, en revanche, à répondre à la question de la fondation de la logique elle-même : « À la question de savoir pourquoi et de quel droit nous reconnaissons une loi de la logique comme vraie, la logique ne peut répondre qu'en la ramenant à une autre loi de la logique. Lorsque cela n'est pas possible, la logique ne peut apporter aucune réponse. »[2] En ce sens, bien que Frege considère encore les lois de la logique comme de véritables connaissances – elles sont les vérités les plus générales qui soient – ce n'est plus du tout dans le contexte d'une recherche sur la théorie de la connaissance qu'il pose le problème de la nature de la logique. Cela explique, du reste, que certains de ses critiques aient pu lui reprocher de ne donner aucun éclaircissement sur la manière dont nous serions capables, selon lui, de saisir par notre activité intellectuelle ces entités objectives qu'il nomme « pensées ».

Au cours des premières décennies du XX[e] siècle, de nouvelles conceptions de la logique contribuèrent à séparer davantage encore la question de la nature de

1. Cf. Frege, « Logique », 1897, trad. franç., *in* Frege, *Écrits posthumes*, Nîmes, Jacqueline Chambon, 1994, p. 149-177, et Frege, Recherches logiques. 1. « La pensée », 1918, trad. franç., *in* Frege, *Écrits logiques et philosophiques*, Paris, Le Seuil, 1971.

2. Frege, *Die Grundgesetze der Arithmetik*, vol. I, 1893, Hildesheim, Olms, 1962, p. XVII.

la logique et le problème général d'une théorie de la connaissance. Sur ce point, l'influence du *Tractatus logico-philosophicus* de Wittgenstein fut décisive. Selon cet auteur, loin d'être les vérités les plus générales, les propositions de la logique ne disent rien, ne « traitent » de rien et sont de pures tautologies[1]. Ainsi, une instance du principe logique du tiers exclu comme « il pleut ou il ne pleut pas » ne nous apprend rien sur l'état des choses dans le monde et elle est, pour cette raison, « vide de sens ». Wittgenstein ne veut pas dire par là que les propositions de la logique sont absurdes, ni même qu'elles n'ont aucun usage ; car si elles ne *disent* rien, et si, par conséquent, elles « ne sont pas des images de la réalité », si elles « ne figurent aucune situation possible » (4.462), elles montrent, en revanche, quel est l'« échafaudage du monde » (6.124). Selon Wittgenstein, la forme logique du monde peut être *indiquée* ou *montrée* au moyen du langage, mais elle ne saurait être *exprimée* par des propositions. Une telle expression supposerait que nous puissions nous placer « en dehors de la logique, c'est-à-dire en dehors du monde » (4.12). La position de Wittgenstein a donc, comme celle de Frege, un caractère universaliste (cf. *supra*, p. 23) : « Il faut que les règles de la syntaxe logique se comprennent d'elles-mêmes » (3.334) parce que toute tentative pour les exprimer dans le langage lui-même ne produirait qu'absurdité. Mais Wittgenstein insiste, beaucoup plus que Frege, sur le caractère négatif de cette position, sur ce qui ne peut être exprimé.

1. L. Wittgenstein, *Tractatus logico-philosophicus*, 1922, trad. franç., Paris, Gallimard, 1993. Cf. par exemple les aphorismes 5.43, 6.1, 6.11 et 6.124.

Carnap lut très attentivement Wittgenstein et il est probable que l'une des idées maîtresses de *La Syntaxe logique du langage*, qu'il fit paraître en 1934, fut inspirée du § 3.341 du *Tractatus* : « Si nous connaissons la syntaxe logique d'un symbolisme quelconque, alors nous sont déjà données toutes les propositions de la logique. » Mais contrairement à Wittgenstein, Carnap estimait qu'en ayant recours à un métalangage, il était tout à fait possible d'exprimer cette syntaxe sans verser dans le non-sens. Carnap pensait en effet que la métalogique développée par les logiciens polonais (Ajdukiewicz, Lesniewski, Lukasiewicz, Tarski), la métamathématique de Hilbert et la technique d'arithmétisation de la syntaxe utilisée par Gödel avaient abondamment justifié, contre Wittgenstein, la possibilité d'exprimer de manière sensée les règles de syntaxe d'un langage. La méthode syntaxique introduite par Carnap en 1934 avait pour but de permettre l'analyse des différents cadres linguistiques possibles pour une reconstruction rationnelle de la science[1].

De ce point de vue, la distance à l'égard des philosophes néokantiens (dont l'influence était clairement perceptible dans les premiers écrits de Carnap) est extrême : non seulement les propositions de la logique n'expriment plus aucune connaissance, mais elles n'ont plus pour fonction de fonder la science. Au niveau du langage objet, les propositions de la logique sont des propositions analytiques sans objet et dépourvues de contenu. Elles servent de propositions auxiliaires pour la déduction des propositions synthétiques, les seules

1. Cf. P. Wagner (ed.), *Carnap's Logical Syntax of Language*, Basingstoke, Palgrave Macmillan, 2009.

qui expriment une véritable connaissance[1]. Au niveau du métalangage, la logique est une méthode d'analyse pour le langage de la science et pour les cadres linguistiques dans lesquels les théories scientifiques peuvent être formalisées. On comprend que la position de Carnap ait pu paraître inacceptable à quiconque estimait que la logique ne devait pas être coupée des actes de connaissance ou du devenir historique de la science. Cavaillès, par exemple, parle de *La syntaxe logique du langage* comme de « notre vieille ennemie »[2], estimant que « l'imagination syntaxique semble se perdre dans le vide d'une abstraction radicale »[3].

Carnap lui-même renonça aux exigences d'une méthode strictement syntaxique après avoir pris connaissance des travaux sémantiques de Tarski qui rendaient possible la définition d'un prédicat de vérité pour les langages extensionnels et la définition d'un concept de conséquence logique. Dans les années qui suivirent, ces travaux devaient largement contribuer au déclin de l'universalisme logique au profit d'une conception de la logique dans laquelle l'analyse sémantique reposait sur le concept de structure d'interprétation d'un langage. La base de cette conception était formée de la logique classique du premier ordre, de la théorie des modèles et de la sémantique tarskienne. Son cadre

1. Cf. R. Carnap, « Science formelle et science du réel », 1935, trad. franç., *in* Ch. Bonnet et P. Wagner (éds.), *L'Âge d'or de l'empirisme logique*, Paris, Gallimard, 2006.

2. Lettre à Lautman du 4 novembre 1942 (cf. J. Cavaillès, *Sur la logique et la théorie de la science*, Paris, Vrin, 1997, *postface* de Jan Sebestik, p. 118).

3. J. Cavaillès, *Sur la logique et la théorie de la science*, Paris, Vrin, 1997, p. 49.

métathéorique était celui de la théorie des ensembles, amplement suffisante pour la formalisation des mathématiques au moyen d'une codification de ses objets par des ensembles et des ensembles d'ensembles. De fait, certains résultats obtenus dans le cadre de cette conception devenue standard (cf. *supra*, chap. III et IV) ont fini par être considérés comme de véritables dogmes, notamment l'impossibilité – censée avoir été démontrée par Tarski en toute généralité – de définir sans contradiction un prédicat de vérité pour un langage $\mathcal{L}$ dans le langage $\mathcal{L}$ lui-même. Aussi était-il devenu habituel, lorsqu'un logicien remettait en question les présupposés d'un tel résultat, et donc le cadre standard lui-même, de qualifier son travail de manière négative (logique « non classique ») ou désobligeante (logique « déviante »). On parle également parfois, sans que l'expression ait un sens très précis, de logique « philosophique ». Ce qu'on désigne par là, le plus souvent, est un ensemble de logiques dans lesquelles interviennent des concepts utiles à l'analyse des langues naturelles et que discutent généralement les philosophes (par opposition à ceux des logiciens qui se préoccupent essentiellement des mathématiques et de leurs fondements) : la connaissance, la croyance, le devoir, la nécessité, etc. Il faut cependant bien reconnaître que l'histoire de la logique des dernières décennies doit énormément aux multiples manières dont ce cadre standard et les théorèmes qui en découlent ont été remis en question. Il existe par exemple des langages logiques dans lesquels un prédicat de vérité pour ce langage *est* définissable[1].

1. Cf. J. Hintikka, *Les Principes des mathématiques revisités*, 1996, trad. franç., Paris, Vrin, 2007.

Une des caractéristiques de la logique mathématique, dans sa version classique, est que le concept de connaissance n'y intervient pas et qu'elle est entièrement détachée du problème général d'une théorie de la connaissance. Cette remarque, qui vaut pour la logique mathématique proprement dite, est cependant loin d'être généralisable à l'ensemble de la logique contemporaine. Après les travaux précurseurs de von Wright[1], Hintikka est, en 1962, l'un des premiers logiciens à prendre au sérieux la possibilité d'un traitement logique de concepts épistémiques comme la croyance et la connaissance, et donc d'expressions comme « *A* sait que *p* », « *A* sait si *p* », « *A* sait qui est *i* », « *A* croit que *p* » ou encore « *p* est compatible avec tout ce que croit *A* »[2]. Il introduit également une approche formelle de questions apparentées : savoir implique-t-il savoir que l'on sait ? Ne pas savoir implique-t-il savoir qu'on ne sait pas ?, etc. Pour cela, Hintikka introduit dans un langage formel les opérateurs modaux *B* et *K*, destinés à modéliser, respectivement, la croyance (*belief*) et la connaissance (*knowledge*), et de les interpréter à l'aide d'une sémantique des mondes – ou des scénarios – possibles (cf. *supra*, p. 75). Comme on va le voir, bien qu'il soit ici question de connaissance, le changement de perspective par rapport à la théorie de la connaissance de la philosophie traditionnelle est radical.

Si *A* est un agent épistémique et *p* une proposition, l'expression $K_A p$ se lit « *A* sait que *p* ». L'interprétation

1. G. H. von Wright, *An Essay in Modal Logic*, Amsterdam, North-Holland, 1951.
2. J. Hintikka, *Knowledge and Belief. An Introduction to the Logic of the Two Notions*, Ithaca et Londres, Cornell UP, 1962.

qu'en donne Hintikka est que p est vrai dans tous les mondes possibles compatibles avec ce que A sait. L'ensemble $\mathcal{W}$ des mondes possibles est donc divisé en deux : ceux qui sont compatibles avec les connaissances de A, et ceux qui ne le sont pas. Cette dichotomie est représentée par la relation d'accessibilité R définie sur $\mathcal{W}$. Si l'on suppose que w et w' sont deux éléments de $\mathcal{W}$ et que A se trouve dans le monde w, Rww' se lit « w' est accessible depuis w », ce qui signifie que le monde possible w' est compatible avec toutes les connaissances de A. Pour l'agent A, avoir une connaissance revient donc à exclure un certain nombre de mondes possibles, ce qui correspond au fait qu'ils soient inaccessibles pour lui depuis w.

On voit que la logique épistémique ainsi comprise ne dépend pas de la question centrale de toute théorie épistémologique : savoir ce qu'est la connaissance. Son but est de donner une représentation formelle de la *logique* de la connaissance, et donc des raisonnements dans lesquels cette modalité figure de manière essentielle. Il n'en reste pas moins que pour son axiomatisation, plusieurs systèmes non équivalents se trouvent en concurrence. L'axiome de distribution, selon lequel si A sait que $(p \rightarrow q)$, alors, s'il sait que p, il sait que q, est généralement accepté, ainsi que l'axiome de vérité selon lequel p est vrai si A sait que p. Si l'on ajoute à ces deux énoncés l'axiome d'introspection positive, selon lequel A sait qu'il sait que p s'il sait que p, on obtient le système nommé S4. Le système S5, quant à lui, est obtenu à partir de S4 en ajoutant l'axiome d'introspection négative, plus controversé que les précédents, selon lequel si A ne sait pas que p, il sait qu'il ne sait pas que p. Les axiomes qu'on vient de citer (il en existe d'autres, grâce auxquels d'autres axiomatisations de la connaissance ont pu être formulées) s'écrivent :

$K_A(p \to q) \to (K_A\, p \to K_A\, q)$
axiome *K*, ou axiome de distribution.

$K_A\, p \to p$
axiome *T*, ou axiome de la vérité.

$K_A\, p \to K_A\, K_A\, p$
axiome 4, ou axiome d'introspection positive.

$\neg K_A\, p \to K_A \neg K_A\, p$
axiome 5, ou axiome d'introspection négative.

L'un des points remarquables est que leur validité correspond à des propriétés de la relation d'accessibilité *R*. Par exemple, l'axiome *T* est valide dans toutes les structures d'interprétation du langage dans lesquelles *R* est réflexive, tandis que l'axiome 4 est valide si *R* est transitive. Ainsi, les propriétés de *R* sont directement liées aux caractéristiques du concept de connaissance modélisé par les axiomes qu'on aura retenus.

L'une des principales difficultés que soulève la logique épistémique ainsi conçue est le problème de l'*omniscience logique* : la caractérisation formelle de la connaissance qu'on a donnée a en effet pour conséquence que l'agent épistémique est censé connaître non seulement toutes les vérités logiques (celles-ci sont vraies dans tous les mondes possibles), mais également toutes les conséquences logiques de ses propres connaissances, ce qui est évidemment tout à fait irréaliste pour des agents épistémiques humains, à moins de considérer que ce qui est modélisé n'est pas la connaissance au sens usuel du terme mais la connaissance implicite.

Diverses solutions ont été proposées pour faire face à cette difficulté, mais il s'agit d'une question qui reste largement ouverte.

Au début des années 1980, la logique épistémique connut un renouveau lorsque les informaticiens l'utilisèrent pour modéliser le fonctionnement de systèmes à bases de connaissances ou le comportement de programmes capables de gérer les flux d'informations entre les multiples composants d'un système. D'autres applications et extensions sont également considérées. Par exemple, la modélisation des interactions entre plusieurs agents épistémiques : lorsque A et B savent l'un et l'autre que p, il est possible que A ne sache pas que B sait que p, et que B, en revanche, sache que A le sait ; A peut alors apprendre que B le sait. Ainsi, le traitement formel de la connaissance vise la modélisation d'une multiplicité de concepts ou relations épistémiques comme l'apprentissage, la révision des croyances ou la connaissance commune, c'est-à-dire la situation dans laquelle tous les agents épistémiques savent que p, savent que tous savent que p, savent que tous savent que tous savent que p, etc. D'autres extensions intègrent spécifiquement l'activité des agents en tant qu'elle détermine leurs connaissances, ou permettent la combinaison des opérateurs épistémiques à des opérateurs destinés à modéliser d'autres modalités comme la nécessité, la possibilité, l'obligation, les relations temporelles, etc.[1].

L'intérêt des développements qui viennent d'être esquissés ne tient pas seulement à la meilleure com-

1. R. Fagin, J. Halpern *et al.*, *Reasoning about Knowledge*, Cambridge (Mass.), Londres, MIT Press, 1995, donne de nombreux exemples d'applications et d'extensions de la logique épistémique.

préhension qu'ils apportent de la logique de la connaissance, mais également aux interactions entre la logique et des théories qui, à l'origine, répondent à des motivations différentes (théorie des systèmes informatiques à bases de connaissances, intelligence artificielle, théorie de la révision des croyances, théorie des jeux, etc.). La logique ne bénéficie pas moins de ces interactions que les théories en question ne bénéficient, inversement, des développements de la logique contemporaine. On voit que les rapports entre logique et connaissance s'en trouvent profondément modifiés. Dans ce contexte, il n'est plus question des facultés d'un sujet connaissant, ni d'une logique intégrée à une théorie de la connaissance en quête de fondement. Les connaissances modélisées appartiennent indifféremment à des agents épistémiques humains ou à des systèmes informatiques et les problèmes soulevés ne sont pas ceux du scepticisme ou de la certitude[1]. Ce qui est en question, outre l'analyse logique d'expressions telles que « *A* sait que *p* », est la représentation des connaissances, la communication, l'apprentissage, la recherche d'informations dans des bases de données, la gestion des flux d'informations, l'interaction entre agents épistémiques, l'inférence en situation d'incertitude, etc. On voit également que les problèmes de logique de la connaissance ainsi conçus peuvent difficilement être séparés des questions de méthodologie.

L'examen d'autres développements actuels permettrait de confirmer que la remise en question de ce qui

1. Sur les relations entre théories de la connaissance et logique épistémique, cf. les articles de J. Van Benthem, V. F. Hendricks et J. Symons, et R. Stalnaker, in *Philosophical Studies*, 128, 2006.

a longtemps été considéré comme la conception standard de la logique (logique classique du premier ordre, sémantique tarskienne, métalangage ensembliste) renouvelle en profondeur les rapports entre logique, connaissance et langage. La théorie sémantique des jeux (connue sous l'appellation « GTS », ou *game-theoretical semantics*) par exemple, est une sémantique non tarskienne dans laquelle la vérité d'une formule est définie à partir d'un jeu logique entre deux joueurs : l'un, qu'on nomme souvent, par convention, *E* (ou Héloïse, ou $\exists$), cherche à montrer que la formule considérée est vraie, et le second, A (ou Abélard, ou $\forall$), que cette formule est fausse. Le déroulement du jeu est déterminé par la structure syntaxique de la formule en question. Étant donné $\mathcal{J}$ une structure d'interprétation du langage et ϕ une formule, l'expression « ϕ est vrai dans $\mathcal{J}$ » est définie par l'existence d'une stratégie gagnante pour Héloïse.

L'intérêt de cette sémantique (dont on peut montrer, en utilisant l'axiome du choix, qu'elle est équivalente, pour la logique du premier ordre, à la sémantique tarskienne) vient des multiples extensions et applications auxquelles elle ouvre la voie[1]. L'une des plus remarquables, qui est à la base de la logique IF *(independance friendly logic)*, touche aux relations de dépendance entre quantificateurs. Dans une formule comme $\forall x \forall y \exists z \exists u \phi[x,y,z,u]$ (où *x, y, z* et *u* sont les variables libres de la formule ϕ), *z* et *u* dépendent l'une et

1. Cf. J. Hintikka et G. Sandu, Game-theoretical semantics, *in* J. Van Benthem et A. Ter Meulen (ed.), *Handbook of Logic and Language*, Amsterdam, Elsevier, 1997, et J. Hintikka, *Les Principes des mathématiques revisités*, 1996, trad. franç., Paris, Vrin, 2007.

l'autre de x et de y. La logique IF ouvre une possibilité qui n'existe pas dans la logique classique du premier ordre, à savoir que z soit dépendant de x sans l'être de y, et que u soit dépendant de y sans l'être de x. Dans le jeu sémantique que jouent Abélard et Héloïse, Abélard choisit deux valeurs dans le domaine d'individus de la structure d'interprétation considérée : une pour x et une pour y, en ayant pour objectif de montrer que la formule $\exists z \exists u \phi[x,y,z,u]$ ne peut pas être satisfaite lorsqu'on assigne à x et à y les valeurs qu'il choisit ; Héloïse choisit alors à son tour deux valeurs – une pour z et une pour u – en ayant pour objectif de montrer que $\phi[x,y,z,u]$ peut être satisfaite pour la valeur de x et la valeur de y choisies par Abélard. Que z dépende de x sans dépendre de y signifie que lorsque Héloïse choisit une valeur pour z, elle connaît le choix d'Abélard pour x sans connaître son choix pour y. Sachant que la vérité d'une formule se définit par l'existence d'une stratégie gagnante pour Héloïse, on voit que cette définition fait intervenir un concept que les théoriciens des jeux connaissent bien, celui de jeu à information incomplète. Les applications sont multiples : analyse logique du langage naturel, définition d'un prédicat de vérité pour le langage $\mathcal{L}$ dans $\mathcal{L}$ lui-même, logique épistémique, etc.[1].

1. Outre les références citées dans la note précédente, cf. R. E. Auxier et L. E. Hahn (éds.), *The Philosophy of Jaakko Hintikka*, Chicago, La Salle (Ill.), Open Court, 2006.

Chapitre VIII

MUTATIONS DE LA LOGIQUE

La logique classique du premier ordre et la sémantique tarskienne forment aujourd'hui la base de l'enseignement de la logique et le cadre méthodologique à partir duquel celle-ci est exposée. Pourtant, d'un côté, les fondateurs de la logique moderne ne la concevaient pas en ces termes et, d'un autre côté, la logique actuelle, consciente des limites de ce cadre, est de plus en plus engagée dans un travail de reconceptualisation.

Deux questions viennent donc naturellement à l'esprit : d'une part, comment est-on parvenu à ce noyau de la logique contemporaine qu'est la logique du premier ordre et comment expliquer sa stabilité ? Les chapitres II, III et IV ont tenté de donner quelques éléments de réponse à cette question. D'autre part, comment comprendre la volonté de dépasser ce cadre et de concevoir la logique selon d'autres principes ? Sur cette seconde question, les chapitres V, VI et VII contiennent quelques indications qui demandent cependant à être prolongées.

La logique classique du premier ordre peut être vue comme un cadre formel qui résulte d'un compromis entre des exigences qu'aucun système logique ne peut toutes satisfaire : cohérence, complétude sémantique, compacité, complétude syntaxique, puissance d'expression élevée, décidabilité de l'ensemble des formules valides, extensionalité, faible degré de complexité... Ainsi, elle permet de formaliser l'arithmétique, mais non de caractériser la structure des entiers naturels. Or, si l'on passe du premier au second ordre, on parvient, certes, à une meilleure caractérisation des objets dont s'occupe le théoricien des nombres, mais on perd du même coup la propriété de complétude sémantique, en sorte qu'une formule peut être conséquence logique des axiomes de l'arithmétique sans qu'il existe pour autant aucune preuve de cette formule par une méthode de dérivation formelle. Pour prendre un second exemple, la logique classique du premier ordre permet de formaliser toutes les mathématiques, mais ni les modalités ni les quantificateurs dont il est fait usage dans la langue naturelle n'y reçoivent un traitement logique satisfaisant. Or, s'il est certes possible d'intégrer dans le langage des opérateurs de modalité, le prix à payer est celui de la propriété d'extensionalité ; quant à la théorie des quantificateurs généralisés, elle exige elle aussi une logique plus puissante que celle du premier ordre. D'une manière générale, le choix d'un système formel est le résultat d'un compromis entre deux exigences : une plus grande puissance d'expression et une plus faible complexité, logique ou algorithmique.

Ces quelques remarques risquent cependant de donner de la logique une image déformée : celle d'une sorte de laboratoire métalinguistique dans lequel seraient testés et éprouvés les différents systèmes formels sortis de l'imagination des logiciens. Une telle conception, technicienne et relativiste, réduit la logique à une science des systèmes formels et de leurs usages possibles. Telle est effectivement l'image que peuvent donner certains ouvrages d'introduction dans lesquels s'enchaînent les expositions de différents formalismes : pour le calcul propositionnel, la logique du premier ordre, la logique d'ordre supérieur, la logique modale, etc. Or, la logique ne se réduit nullement à cela.

Le logicien ne s'intéresse pas tant à la collection des systèmes formels qu'à un ensemble de concepts transversaux dont il voudrait pouvoir donner, idéalement, une caractérisation indépendante de toute contrainte syntaxique particulière : la validité, la conséquence logique, la négation, la quantification, la complexité... Comment définir, par exemple, le concept d'algorithme en toute généralité, et donc indépendamment des différents modèles particuliers du calcul (machine de Turing, termes du lambda-calcul, etc.) ?

Si la logique ne se réduit ni à un système formel particulier, ni à une collection de systèmes, ni à une méthode générale destinée à analyser les formalismes, de quoi traite-t-elle ? Quel est son objet ? Selon une réponse traditionnelle qui, comme telle, reste extrêmement vague, la logique est la science formelle du raisonnement correct. Dans la période comprise entre les premiers travaux de Frege et le milieu du XX^e siècle, cette définition générale a été peu à peu précisée, mais aussi tellement particularisée que son objet a fini par se réduire, pour l'essentiel, au cas particulier des mathématiques, comme si le paradigme du raisonnement se trouvait dans les preuves que produisent les mathématiciens et que la tâche principale de la logique pouvait se ramener à la recherche d'une solution pour le problème du fondement des mathématiques. Dans le même temps, la logique atteignit un tel degré de technicité que bien des logiciens eurent tendance à oublier qu'elle pouvait avoir – et avait effectivement eu – d'autres horizons. Eu égard à cette période, le point le plus important tient probablement moins, toutefois, à l'objet même de cette spécialisation qu'au fait que la logique ait été conduite, par là même, à réformer ses méthodes d'analyse et à développer des outils conceptuels aussi précis et fins que ceux dont disposent les mathématiciens.

Depuis les années 1950, les recherches logiques ont pris de nouvelles directions comme l'étude des langues naturelles et de leurs structures syntaxiques et sémantiques, l'informatique, l'intelligence artificielle, la révision des croyances, la théorie des jeux, la logique de la connaissance, les sciences cognitives... en sorte que la période précédente apparaît, rétrospectivement, comme un épisode qui a permis de mettre au point des méthodes d'analyse techniquement très élaborées qui peuvent certes, aujourd'hui encore, permettre d'avancer sur le problème des fondements des mathématiques mais dont l'application ne se limite plus du tout à cette question. Si les ouvrages d'introduction générale ne reflètent encore qu'assez rarement ce changement d'orientation, il n'en est pas moins réel et tout à fait notable.

L'une des caractéristiques de la logique actuelle est l'intérêt qu'elle porte à des questions dont le rejet ou le dépassement avait précisément permis à la logique mathématique de se constituer : les singularités de la grammaire, les langues naturelles, la psychologie du raisonnement, la méthodologie scientifique, l'analyse de la connaissance... Elle connaît donc indiscutablement un élargissement de son champ d'investigation sans qu'on puisse pour autant décider clairement si elle doit être considérée comme une méthode générale dont le domaine d'application s'étend maintenant bien au-delà des mathématiques ou comme une discipline dont l'objet n'est plus seulement la preuve formelle ou le fondement de la connaissance en général mais plus généralement l'ensemble des phénomènes qui touchent à l'information, son partage, son évaluation, son traitement ou son échange[1]. « La logique s'est principalement occupée de propriétés d'objets éternels comme des propositions, ou les relations d'inférence qui existent entre elles. Mais de tels objets sont le résultat d'activités telles que l'apprentissage, la mise à jour des informations dont on dispose, la révision d'une croyance, le fait de poser une question, le test ou la mise à l'épreuve d'une affirmation donnée, etc. Cette dynamique est devenue un objet d'étude logique en soi. »[2] De cette évolution témoignent la logique dialogique, la sémantique des jeux, ou la logique dynamique, dans laquelle l'évaluation des propositions est conçue comme un processus[3]. Dans cette perspective, la logique peut encore être définie comme la « science du raisonnement correct », mais le raisonnement dont il est question ne se réduit plus à une suite de signes car il peut aussi bien être vu comme un processus dynamique auquel prennent part des agents rationnels placés dans un environnement spécifique, susceptibles de réviser leurs croyances au vu d'informations nouvellement

1. Cf. J. Van Benthem, « Wider still and Wider... Resetting the Bounds of Logic », *in* A. Varzi (ed.), *The European Review of Philosophy*, vol. IV, Stanford, CSLI, 1997.
2. J. Van Benthem, « Where Is Logic Going, and Should it? », *Topoi*, 2006, 25, p. 120.
3. Cf. P. Lorenzen et K. Lorenz, *Dialogische Logik*, Darmstadt, Wissenschaftliche Buchgesellschaft, 1978 ; *Philosophia Scientiae*, « Logique et théorie des jeux », vol. VIII, cahier 2, 2004. Sur la logique dynamique, cf. J. Van Benthem, *Exploring Logical Dynamics*, Stanford, CSLI, 1996.

acquises et qui disposent de ressources d'évaluation et de calcul qui ne sont pas illimitées.

Bien entendu, les logiciens sont loin de partager tous la vision d'une logique contemporaine qui réintégrerait en son sein ce que les fondateurs de la logique moderne, en quête d'une logique pure, s'étaient efforcés de mettre à l'écart, et on ne peut certainement pas affirmer avec certitude que l'orientation qu'on vient d'esquisser dans le paragraphe qui précède représente la logique de l'avenir. Si un grand nombre de logiciens s'accordent à reconnaître que la logique actuelle connaît un tournant, tous ne sont pas d'accord sur la manière dont il convient de le comprendre et ce qui vient d'être indiqué donne tout au plus l'une des visions possibles de l'orientation de la logique contemporaine, telle que la voient certains de ses représentants.

BIBLIOGRAPHIE

Aristote, *Catégories*, Paris, Le Seuil, 2002.

Aristote, *Seconds Analytiques*, Paris, Flammarion, 2005.

Arnault A., Nicole P., *La Logique, ou l'art de penser*, 1662, Paris, Gallimard, 1992.

Blackburn P., Rijke M. de, Venema Y., *Modal Logic*, Cambridge, Cambridge UP, 2001.

Blanché R., Dubucs J., *La Logique et son histoire*, Paris, A. Colin (1970), 1996.

Bonnay D., Cozic M. (éds.), *Philosophie de la logique. Conséquence, preuve et vérité*, Paris, Vrin, textes clés, 2009.

Boolos G., Jeffrey R., *Computability and Logic*, Cambridge, Cambridge UP, 1974, 5e éd., 2007.

Cori R., Lascar D., *Logique mathématique*, 2 vol., Paris, Masson, 1993 ; Paris, Dunod, 2003.

David R., Nour K., Raffalli Ch., *Introduction à la logique. Théorie de la démonstration*, Paris, Dunod, 2001, 2e éd., 2004.

Enderton H. B., *A Mathematical Introduction to Logic*, New York, Academic Press, 1972, 2e éd., 2001.

Enderton H. B., *Elements of Set Theory*, New York, Academic Press, 1977.

Fagin R., Halpern J. *et al.*, *Reasoning about Knowledge*, Cambridge (MA), MIT Press, 1995.

Frege G., *Écrits logiques et philosophiques*, Paris, Le Seuil, 1971.

Frege G., *Idéographie*, 1879, trad. franç., Paris, Vrin, 2000.

Frege G., *Écrits posthumes*, 1969, trad. franç., Nîmes, Jacqueline Chambon, 1994.

Gabbay D., Guenthner F. (eds.), *Handbook of Philosophical Logic*, Dordrecht, D. Reidel, 4 vol., 1983-1989; 2e éd., Dordrecht, Kluwer, vol. I à XI, 2001-2005, Dordrecht, Springer, vol. XII à XVIII, 2005-.

Gabbay D., Woods J. (eds.), *Handbook of the History of Logic*, Amsterdam, Elsevier, 2004-, 11 vol. prévus.

Gamut L. T. F., *Logic, Language and Meaning*, 2 vol., Chicago, University of Chicago Press, 1991.

Girard J.-Y., *Proof Theory and Logical Complexity*, Naples, Bibliopolis, 1987.

Goble L. (ed.), *The Blackwell Guide to Philosophical Logic*, Oxford, Blackwell, 2001.

Gochet P., Gribomont P., Thayse A., *Logique*, 3 vol., Paris, Hermès, 1990-2000.

Gödel K., *Collected Works*, 5 vol., New York, Oxford UP, 1986-2003.

Haaparanta L. (ed.), *The Development of Modern Logic*, Oxford, Oxford UP, 2009.

Hintikka J., *Les Principes des mathématiques revisités*, 1996, trad. franç., Paris, Vrin, 2007.
Jacquette D. (ed.), *A Companion to Philosophical Logic*, Oxford, Blackwell, 2002.
Jacquette D. (ed.), *Philosophy of Logic. An Anthology*, Malden, Mass., Blackwell, 2002.
Jacquette D. (ed.), *Philosophy of Logic*, Amsterdam, North-Holland, 2006.
Kleene S. *Introduction to Metamathematics*, Amsterdam, North-Holland, 1952.
Kleene S. *Logique mathématique*, 1967, trad. franç., Paris, A. Colin, 1971.
Kneale W., Kneale M., *The Development of Logic*, Oxford, Clarendon Press, 1962.
Lepage F., *Éléments de logique contemporaine*, Montréal, Presses de l'Université de Montréal, 1991.
Largeault J. (éd.), *Logique mathématique. Textes*, Paris, A. Colin, 1972.
Largeault J. (éd.), *Intuitionisme et théorie de la démonstration*, Paris, Vrin, 1992.
MacFarlane J., *What Does it Mean to Say that Logic Is Formal?*, Ph. D, université de Pittsburgh, 2000 (disponible en ligne).
Machover M., *Set Theory, Logic and their Limitations*, Cambridge, Cambridge UP, 1996.
Quine W.V., *Du point de vue logique*, 1953, trad. franç., Paris, Vrin, 2003.
Quine W.V., *La Philosophie de la logique*, 1970, trad. franç., Paris, Aubier, 1975 ; 2008.
Quine W.V., *Méthodes de logique*, 1952, trad. franç., Paris, A. Colin, 1973.
Read S., *Thinking about Logic*, Oxford, Oxford UP, 1995.
Ricketts T., Potter M. (eds), *The Cambridge Companion to Frege*, Cambridge, Cambridge UP, 2010.
Rivenc F., *Introduction à la logique*, Paris, Payot, 1989.
Rivenc F., *Introduction à la logique pertinente*, Paris, PUF, 2005.
Rivenc F., Rouilhan Ph. de (éds.), *Logique et fondements des mathématiques. Anthologie (1850-1914)*, Paris, Payot, 1992.
Russell B., *Introduction à la philosophie mathématique*, 1919, trad. franç., Paris, Payot, 1991.
Russell B., *Écrits de logique philosophique*, Paris, PUF, 1989.
Smullyan R., *Le Livre qui rend fou*, 1982, tr. fr. Paris, Dunod, 1984.
Smullyan R., *Logical Labyrinths*, Wellesley, AK Peters, 2009.
Tarski A., *Logique, sémantique, métamathématique*, 2 vol., trad. franç., Paris, A. Colin, 1972, 1974.
Van Benthem J., Meulen A. Ter (eds.), *Handbook of Logic and Language*, Amsterdam, Elsevier, 1997.
Van Dalen D., *Logic and Structure*, Berlin, Springer, 1980, 4e éd., 2004.
Van Heijenoort J. (ed.), *From Frege to Gödel. A Sourcebook in Mathematical Logic, 1879-1931*, Cambridge (Mass.), Harvard UP, 1967.

TABLE DES MATIÈRES

Mis en pages et imprimé en France
par JOUVE
1, rue du Docteur Sauvé, 53100 Mayenne
mars 2011 - N° 639841A

JOUVE est titulaire du label imprim'vert®

PHILOSOPHIE

Par conséquent…

LA LOGIQUE

Il est devenu si courant aujourd'hui de parler *des* logiques qu'on ne sait bien souvent plus ce qu'est *la* logique. Cet ouvrage donne à comprendre cette discipline en décrivant le genre de questions que se posent ou que se sont posées les logiciens, le genre de certitudes qu'ils ont acquises et la variété des projets qui animent leurs recherches.

S'il expose avec clarté les bases de la logique contemporaine et ses origines historiques, l'auteur montre aussi que, depuis les années 1950, les recherches logiques ont pris de nouvelles directions comme l'étude des structures syntaxiques et sémantiques des langues naturelles, l'informatique théorique, l'intelligence artificielle, la théorie des jeux, l'analyse dynamique des croyances et de la connaissance, ou encore les sciences cognitives.

Pierre Wagner

Pierre Wagner est maître de conférences à l'Université Panthéon-Sorbonne (Paris 1), où il enseigne la logique et la philosophie des sciences, et membre de l'IHPST (Institut d'histoire et philosophie des sciences et des techniques). Il a notamment dirigé Les philosophes et la science *(Gallimard, « Folio essais », 2002).*

À lire également en *Que sais-je ?…*

- **La philosophie du langage** *de Sylvain Auroux*
- **La philosophie des sciences** *de Dominique Lecourt*

2e édition 2011

9 €

ISBN : 978-2-13-058456-8

9 782130 584568

Que sais-je ?

www.quesais-je.com

COLLECTION ENCYCLOPÉDIQUE

fondée par Paul Angoulvent

N°