Publié dans Travaux de logique 19, 177-203, 2008,
source qui doit être utilisée pour toute référence à ce travail

DÉFINIR LA LOGIQUE, UNE QUESTION DE SIMPLICITÉ

Denis BONNAY

1. Que veut-on définir ?

Qu'est-ce qu'un raisonnement correct ? La logique vise traditionnellement à répondre à cette question, de sorte que, s'il s'agit de définir la logique, on doit commencer par comprendre les tenants et les aboutissants du problème. Considérons un argument particulier :

Tous les philosophes sont sages
Certains hommes sont philosophes
(1) --
Certains hommes sont sages

Intuitivement l'argument (1) est bien un argument valide. Les prémisses étant posées, la conclusion suit nécessairement. Mais en vertu de quoi (1) est-il correct ? Parce que la logique s'intéresse seulement aux arguments qui sont valides en vertu de leur forme, on dira que (1) est correct en tant qu'instance d'un schéma valide, donné par :

Tous les X sont Y
Certains Z sont X
(2) ----------------------------
Certains Z sont Y

On obtient (2) à partir de (1) en faisant abstraction du contenu particulier de l'argument, qui est véhiculé par les expressions comme « philosophes », « hommes » ou « sages », et en ne retenant que la *forme logique* des énoncés figurant comme prémisse ou comme conclusion. Cette forme logique est obtenue en remplaçant les expressions « non logiques » par des lettres schématiques et en ne gardant constantes que les expressions « logiques », précisément appelées « constantes logiques ». On a ainsi décomposé le travail du logicien en deux tâches distinctes. D'une part, il doit expliquer ce qu'est un schéma d'argument valide, d'autre part, il doit expliquer ce qu'est la forme logique d'un énoncé, et en particulier faire la différence entre les expressions non logiques et les expressions logiques. La définition par Tarski (1936) de la conséquence logique répond au premier problème ; c'est une définition sémantique, en termes de préservation de la vérité (pour toute interprétation possible des lettres schématiques, si les prémisses sont vraies, alors la conclusion est vraie), d'autres définitions sont possibles, notamment en termes preuves-théoriques (un argument sera valide s'il est possible de prouver la conclusion à partir des prémisses en utilisant un certain nombre de règles d'inférences élémentaires). La logique mathématique moderne commence comme discipline avec ces définitions, et les grands résultats méta-logiques classiques, comme le théorème de complétude, s'emploient à en tirer toutes les conséquences.

Reste le second problème. Qu'en est-il du choix du vocabulaire logique ? Tarski, lorsqu'il propose sa définition de la notion de conséquence, exprime son scepticisme quant à la possibilité d'y apporter une solution autre que conventionnelle. On peut simplement décider de considérer certaines expressions comme logiques – les connecteurs booléens de la logique propositionnelle et les quantificateurs existentiel et universel. Ce choix étant fait, les prédictions que l'on obtient concernant la validité des arguments sont en accord avec nos jugements préthéoriques. *Ex post*, le choix du vocabulaire logique s'en trouve validé.

On ne saurait sous-estimer l'importance de ce choix. Certains déterminants de la langue naturelle comme « tous » ou « certains », et leur contrepartie dans les langages artificiels de la logique, sont systématiquement considérés comme logiques. Mais en va-t-il de même pour tous les déterminants ? « La plupart », par exemple, est-il lui aussi un déterminant exprimant une notion logique ? A ne considérer que la langue naturelle, on serait tenté de répondre positivement, par parité. Pourtant « la plupart » n'est pas définissable dans la logique classique du premier ordre. Doit-on en conclure que la logique classique ne couvre pas la totalité du domaine de la logique, ou à l'inverse que les apparences sont trompeuses et que « la plupart » exprime une notion qu'on devrait considérer comme intrinsèquement arithmétique et pas comme logique au sens strict ? De même, on s'accorde pour considérer la quantification existentielle sur les variables du premier ordre comme logique, mais qu'en est-il de la quantification existentielle sur les variables du *second* ordre ? S'il s'agit là encore d'une expression logique, la logique du second ordre ne sera pas des mathématiques déguisées, mais une logique authentique, et les limites de la logique s'étendront bien loin, puisqu'on peut trouver par exemple un énoncé de la logique du second ordre qui est logiquement vrai si et seulement si l'hypothèse du continu est vraie.

Pour caractériser les limites de la logique, il est donc crucial de pouvoir définir la classe des expressions logiques, c'est-à-dire la classe des constantes logiques, afin de déterminer quelles sont les expressions qui contribuent de manière essentielle à la correction de nos raisonnements. Cette question se double d'une autre question : comment définit-on les constantes logiques elles-mêmes ? c'est-à-dire, comment définit-on telle ou telle expression particulière, afin de pouvoir l'utiliser dans la caractérisation indépendamment fournie de la notion de conséquence logique ou de schéma d'argument valide ? Les deux questions sont liées ; répondre à la première, c'est dire quels moyens on a le droit d'utiliser pour répondre à la seconde.

2. Simplicité des définitions et simplicité de la logique

Toute entreprise de définition est délicate, mais l'entreprise que l'on vient de décrire l'est tout particulièrement. Le problème vient de ce que l'on pourrait appeler le *réquisit de simplicité des définitions* : une bonne définition est une définition qui définit le moins élémentaire à l'aide du plus élémentaire. Il est possible et informatif de définir ce qu'est un nombre premier à l'aide de notions arithmétiques élémentaires, comme celles de nombre entier et de produit. Mais à l'inverse, il serait absurde de vouloir définir ce qu'est un nombre entier à l'aide de la notion de nombre premier, parce que la notion de nombre entier est plus élémentaire que la notion de nombre premier. Mais alors la simplicité de la logique va poser problème : les notions logiques sont 'on ne peut plus simples', et la logique est une discipline que l'on ne saurait fonder sur nulle autre, puisque toutes les autres disciplines utilisent les principes de la logique. C'est ce que l'on pourrait appeler le *principe de simplicité de la logique*. Le problème qui se pose lorsqu'on cherche à définir la logique – c'est-à-dire définir les limites de la logique et définir les expressions logiques – résulte alors du conflit entre le réquisit de simplicité des définitions et le principe de simplicité de la logique. Si l'on doit définir à l'aide du plus simple, et que rien n'est plus simple que la logique, comment pourrait-on définir la logique ?

Ce problème est un cas particulier d'un problème classique, celui de l'indéfinissabilité des termes primitifs, que Pascal, dans son opuscule sur l'esprit géométrique, évoque en ces termes :

> [...] il est évident que les premiers termes qu'on voudrait définir en supposeraient de précédents pour servir à leur explication [...]. Aussi, en poussant les recherches de plus en plus, on arrive nécessairement à des mots primitifs, qu'on ne peut plus définir.

Et Pascal de se moquer de ceux qui vont contre cette évidence, et s'emploient à définir ce qui ne peut l'être, comme celui qui définissait la lumière comme un mouvement luminaire des corps lumineux.

Doit-on renoncer à définir les termes primitifs ? C'est la leçon que tire Pascal. A la suite de Hilbert[1], on en a tiré une autre, en distinguant définitions implicites et définitions explicites. On peut définir explicitement la notion de triangle à partir de notions primitives comme celle de droite. On ne peut pas définir explicitement la notion de droite (sauf à prendre d'autres termes primitifs, qui échapperaient à leur tour à la définissabilité explicite). Mais l'on peut considérer que les axiomes de la géométrie définissent *implicitement* les termes primitifs de la géométrie. On ne peut pas définir explicitement ce qu'est un point ou une droite, et Pascal a raison d'insister sur la vanité qu'il y aurait à tenter de le faire, mais on le fait implicitement en acceptant certains axiomes pour la géométrie : points et droites sont les (types d') objets qui satisfont ces axiomes.

La question du statut des définitions implicites – s'agit-il vraiment de définitions ? – excède de beaucoup l'objet du présent article. Le point important pour nous ici est à vrai dire indépendant de la manière dont cette question est tranchée. Il s'agit du point suivant : les définitions implicites présupposent les termes logiques. Les axiomes de la géométrie qui définissent implicitement les termes primitifs de la géométrie sont formulés à l'aide des expressions de la logique. Il ne semble donc pas que

1 Voir sur ce point la correspondance entre Hilbert et Frege à propos des *Grundlagen der Geometrie*, dans laquelle Hilbert envisage, sans employer le terme qui remonte à Gergonne, les axiomes comme des définitions implicites.

la possibilité de se rabattre sur des définitions implicites permette d'éviter la violation du réquisit de simplicité. En effet, pour pouvoir même formuler de telles définitions, on a besoin des notions logiques, de sorte que le réquisit de simplicité serait violé *même* pour des définitions seulement implicites.

Si l'on veut définir les constantes logiques (et la classe des constantes logiques), il semble donc nécessaire de renoncer soit au principe de la simplicité de la logique, soit au réquisit de simplicité des définitions. Y a-t-il des manières de le faire qui n'enlèvent pas à l'entreprise de définition son intérêt ? Imaginons une solution qui repose sur un renoncement au réquisit de simplicité des définitions : comment une définition de la logique à l'aide de notions plus complexes peut-elle saisir ce qui est propre à la logique, et quel intérêt épistémologique y a-t-il à définir le simple à l'aide du complexe ? Réciproquement, imaginons une solution qui repose sur un renoncement au réquisit de simplicité de la logique : qu'y a-t-il donc de plus simple que la logique, sur quoi une telle définition puisse se fonder, et quel intérêt y a-t-il à définir la logique dans une perspective théorique qui refuse le principe de simplicité ?

3. Approche modèle-théorique et approche preuve-théorique

Comme on l'a rappelé un peu plus haut, il existe deux approches rivales de la notion de conséquence logique : l'approche modèle-théorique et l'approche preuve-théorique. L'approche modèle-théorique est une approche sémantique, en termes de préservation de la vérité. Un ensemble d'énoncés Γ a pour conséquence logique au sens modèle-théorique un énoncé φ (notation $\Gamma \models \varphi$) si et seulement si toute structure qui rend vrais tous les énoncés de Γ rend également vrai φ. L'approche preuve-théorique est une approche syntaxique, en termes d'existence d'une preuve formelle. Un ensemble d'énoncés Γ a pour conséquence logique au sens preuve-théorique un énoncé φ (notation $\Gamma \vdash \varphi$) si et seulement si il existe une dérivation de φ sous hypothèses Γ, c'est-à-dire si et seulement il est possible de dériver φ à partir de Γ en utilisant les règles logiques élémentaires.

Les formulations que nous venons de donner sont incomplètes. Pour obtenir une définition modèle-théorique en bonne et due forme, il faut définir précisément ce que l'on entend par modèle – on peut utiliser des structures d'interprétation pour un langage classique du premier ordre et la notion classique de vérité dans un modèle, mais on pourrait également utiliser des modèles pour la logique intuitionniste. De manière analogue, pour obtenir une définition preuve-théorique en bonne et due forme, il faut définir ce que l'on entend par dérivation – on peut utiliser un calcul des séquents multi-conclusions pour la logique classique du premier ordre, mais on pourrait également utiliser la déduction naturelle intuitionniste par exemple.

Toutefois, même sans entrer dans ces détails, on comprend la forme que prend dans chaque cas une définition de la logique. Dans l'approche modèle-théorique, définir *telle* constante logique, c'est définir sa contribution compositionnelle à la valeur sémantique des formules dans un modèle. Par exemple, pour le cas propositionnel, on définit la conjonction en donnant une fonction de vérité. Définir la *classe* des constantes logiques, c'est donner la classe des interprétations possibles pour une expression logique. Pour le cas propositionnel à nouveau, il s'agit de la classe des fonctions de vérité. Dans l'approche preuve-théorique, définir *telle* constante logique, c'est donner les règles d'inférence qui lui sont associées. Pour reprendre l'exemple de la conjonction, et dans un format de déduction naturelle, il s'agit de la capacité à inférer A & B à partir de A et de B, de la capacité à inférer A à partir de A & B, et de la capacité à inférer B à partir de A & B. Définir la *classe* des constantes logiques, c'est caractériser les ensembles de règles d'inférence qui peuvent être considérées comme des ensembles des règles d'inférence proprement logiques, susceptibles de donner la signification d'une expression logique.

Nous nous proposons maintenant de considérer deux entreprises particulières de définition, l'une dans le style preuve-théorique, due à Sambin, l'autre dans le style modèle-théorique, due à Tarski, et de voir comment elles tentent de résoudre au moins partiellement le conflit entre le principe de simplicité de la logique et le réquisit de simplicité des définitions, les renoncements qu'elles opèrent et les conséquences qui en découlent quant au projet de la définition de la logique.

4. Approche preuve-théorique et simplicité de la logique

Considérons un format particulier pour les preuves formelles, le calcul des séquents. Un séquent est une suite de symboles de la forme $A_1,\ldots, A_n \vdash B_1,\ldots, B_n$ où les A_i et les B_i sont des formules, les unes (les A_i) servant de prémisses, les autres (les B_i) de conclusion. L'intréprétation intuitive d'un tel séquent est qu'à partir de A_1 *et* A_2 *et*... *et* A_n on peut obtenir B_1 *ou* B_2 *ou* ... *ou* B_n. On s'autorise de cette interprétation intuitive pour dire que la virgule à gauche du symbole $\vdash$ correspond à une conjonction, alors que la virgule à droite du même symbole correspond à une disjonction. Cela est corroboré par les règles suivantes :

$$\frac{A_1,A_2,\ldots, A_n \vdash B_1,\ldots, B_n}{A_1 \,\&\, A_2\ldots, A_n \vdash B_1,\ldots, B_n} \qquad \frac{A_1,\ldots, A_n \vdash B_1,B_2\ldots, B_n}{A_1,\ldots, A_n \vdash B_1 \vee B_2\ldots, B_n}$$

La première règle est la règle d'introduction à gauche de la conjonction &. La seconde est la règle d'introduction à droite de la disjonction $\vee$.

Sambin (voir Sambin, Battilotti et Faggian 2000) propose d'exploiter cette idée afin de définir systématiquement les constantes logiques. L'idée générale est que les constantes logiques reflètent les « signes de ponctuation »[2] du métalangage dans lequel est caractérisée la relation de conséquence preuve-théorique. A chaque constante est alors associée une « équation définitionnelle » qui la définit. Dans le cas de la conjonction, l'équation définitionnelle est :

$$\frac{A_1, A_2, \ldots, A_n \vdash B_1, \ldots, B_n}{A_1 \,\&\, A_2 \ldots, A_n \vdash B_1, \ldots, B_n}$$

Cette équation est la donnée d'une double règle. De bas en haut, on retrouve la règle habituelle d'introduction à gauche de &, qui nous dit comment obtenir & à partir de la virgule à gauche. De haut en bas, on a ce que Sambin appelle la règle de réflexion implicite, qui nous dit qu'on peut inversement retrouver la virgule à partir de &. La règle habituelle d'introduction à droite de la conjonction,

$$\frac{A_1, \ldots, A_n \vdash B_1 \qquad A'_1, \ldots, A'_n \vdash B_2}{A_1, \ldots, A_n, A'_1, \ldots, A'_n \vdash B_1 \,\&\, B_2}$$

est dérivable à partir de la règle de réflexion implicite. Etant fixé un métalangage dans lequel est décrite la relation primitive de conséquence logique, la classe des constantes logiques est la classe des expressions qui sont caractérisées par des équations définitionnelles du genre de celle que l'on vient de voir. Le paramètre constitué par le métalangage est important : plus celui-ci est riche, plus l'on pourra caractériser de constantes logiques (par exemple, le carré de la logique modale pourra être considéré comme logique si les séquents employés sont suffisamment

2 L'expression est de Dosen (1994), dont les idées sont proches de celles de Sambin.

structurés). Par contre, la constante « Tonk » imaginée par Prior, et caractérisée par les règles suivantes « A ⊢ A tonk B » et « A tonk B ⊢ B », ne sera pas considérée comme une constante logique admissible, car elle n'est pas la « solution » d'une équation définitionnelle[3].

Comment l'entreprise de définition preuve-théorique que l'on vient d'esquisser se situe-t-elle relativement au conflit entre le principe de simplicité de la logique et le réquisit de simplicité des définitions ? On peut d'abord remarquer que les équations définitionnelles ne sont que des définitions implicites : elles ne fournissent pas un *definiens* qui serait systématiquement substituable à la constante définie dans tous les contextes ; elles se contentent de contraindre la signification des constantes logiques à partir de la signification déjà attachée aux symboles en termes desquels les constantes sont définies, à savoir « , » et « ⊢ » dans le cas propositionnel. En effet, même à supposer que l'on considère seulement des formules intervenant dans des séquents, une constante logique ayant une occurrence à l'intérieur d'une formule dans un séquent ne peut pas être « remplacée » par une combinaison de « , » ou « ⊢ ». Le remplacement n'est possible que dans le contexte très particulier donné par l'équation définitionnelle définissant la constante – effectuer le remplacement du *definiendum* par le *definiens* correspond alors à utiliser la règle de réflexion implicite.

Cependant, on l'a dit, se contenter de définitions implicites ne suffit pas à résoudre le conflit entre simplicité des définitions

3 L'ajout de « Tonk » à un système de preuve dans lequel la relation '⊢' est transitif rend ce système inconsistant. C'est en ce sens que « Tonk » est une constante pathologique. Voir Bonnay & Simmenauer (2005) pour une analyse de la manière dont le problème des « fausses » constantes logiques du genre de Tonk est résolu grâce à l'idée d'équation définitionnelle. Le problème de Tonk est que les règles qui le définissent implicitement sont éminemment créatives ; l'on peut prouver que toutes les règles qui sont engendrées par des équations définitionnelles sous certaines conditions ne sont pas créatives. Joray (2008) montre que le problème de la créativité se pose également pour des constantes logiques introduites à l'aide de définitions explicites.

et simplicité de la logique. Si l'on veut sauvegarder le réquisit de simplicité des définitions, il faut encore se donner quelque chose de plus primitif que les constantes logiques que l'on définit. Ce quelque chose de plus primitif, c'est, dans le cadre de Sambin, la relation de conséquence logique et les signes de ponctuation dont les propriétés sont analysées dans le métalangage. L'ordre des raisons – et des définitions – est alors le suivant : on se donne une relation de conséquence que l'on décrit dans le métalangage en termes de séquents. Typiquement cette relation est supposée satisfaire certaines propriétés, qui correspondent à ce que l'on appelle les règles structurelles, comme l'échange :

$$\frac{A_1, A_2, \ldots, A_n \vdash B_1, \ldots, B_n}{A_2, A_1, \ldots, A_n \vdash B_1, \ldots, B_n}$$

ou encore l'affaiblissement :

$$\frac{A_1, \ldots, A_n \vdash B_1, \ldots, B_n}{C, A_1, .., A_n \vdash B_1, \ldots, B_n}$$

Ensuite, et ensuite seulement, on peut, sur la base de cette relation de conséquence, définir les différentes constantes logiques comme « reflétant » dans le langage objet telles ou telles expressions utilisées dans le métalangage dans lequel sont décrits les séquents. Les constantes logiques que l'on définit se retrouvent alors dépendantes des choix qui ont été faits à propos de la relation de conséquence logique. Par exemple, en l'absence de règles d'affaiblissement et de contraction, on obtiendra les constantes de la logique linéaire. En présence de ces règles au contraire, on obtiendra les constantes de la logique classique. Cette dépendance est dans une certaine mesure bienvenue. Elle permet de ressaisir ce qu'il y a de commun aux différentes « versions » d'une même constante logique déclinées à travers différents systèmes logiques – une même équation définitionnelle caractérise la conjonction multiplicative de la logique linéaire et la conjonction habituelle de la logique classique, la différence entre les deux venant du contexte dans lequel s'opère la définition.

On voit néanmoins également ce que l'on perd. Au lieu d'utiliser les constantes logiques afin de définir la notion de conséquence logique, on traite la notion de conséquence elle-même comme primitive afin de se donner le « plus élémentaire » permettant de définir les constantes logiques sans violer le réquisit de simplicité des définitions. Le prix à payer est alors de laisser inanalysée la notion de conséquence logique ; notamment, les différentes propriétés structurelles possibles restent comme des paramètres libres, qui engendrent les différences entre les systèmes logiques, mais pour lesquels on ne dispose pas de justification indépendante. Dans une certaine mesure, s'il s'agit de définir la logique, le problème n'a été que repoussé : il devient celui d'une caractérisation de la relation de conséquence logique. Pour le dire autrement, le contournement du principe de simplicité de la logique n'est qu'apparent. La stratégie adoptée par Sambin consiste en réalité à isoler *au sein de la logique* quelque chose de plus élémentaire, à savoir les séquents et la relation donnée par « $\vdash$ », pour définir (implicitement) ensuite des choses moins élémentaires, à savoir les constantes logiques qui reflètent les expressions utilisées pour décrire les séquents dans le métalangage. Plutôt qu'une définition de la logique, ce qui a lieu est donc une réorganisation au sein de la logique. Le problème d'une définition des expressions « , » et « $\vdash$ », et d'une analyse de la relation de conséquence logique, reste, lui, entier.

Indépendamment des lumières que jette l'entreprise de Sambin sur les propriétés inférentielles des constantes logiques, il se pourrait que, du point de vue de l'entreprise de définition, la leçon à tirer concerne les limites de toute stratégie visant à renoncer au principe de simplicité de la logique. Si l'on définit les constantes logiques à l'aide de quelque chose de plus simple, qui relève encore du domaine de la logique, la démarche est tout à fait sensée. En réalité, elle ne viole pas le principe de simplicité de la logique. Toutefois, elle ne fait que repousser le problème d'un cran. Si l'on choisissait au contraire de définir les constantes logiques à l'aide de quelque chose de plus simple, qui ne relève plus du domaine de la logique, alors on pourrait espérer faire mieux que repousser le problème. Mais cette fois le principe de simplicité de la logique, qui semble consubstantiel à l'idée que nous nous faisons de la logique, serait bel et bien violé.

5. Approche modèle-théorique et invariance

Peut-être alors convient-il de renoncer au réquisit de simplicité des définitions plutôt qu'au principe de simplicité de la logique. La définition des constantes logiques dans l'approche modèle-théorique de la conséquence nous semble correspondre à ce choix, et c'est vers elle que nous nous tournons maintenant.

Nous avons déjà donné l'exemple élémentaire de la définition de la conjonction comme fonction de vérité dans le cas propositionnel. Cette manière de faire s'étend-elle au-delà de la logique propositionnelle ? La réponse est positive, et l'on peut voir que, dans le cas de la quantification existentielle par exemple, la clause récursive habituelle de la satisfaction dans un modèle contient une définition du quantificateur en jeu. Une formule $\exists x\varphi(x)$ est satisfaite dans un modèle M relativement à une assignation σ (notation $M \models \exists x\varphi(x)\ \sigma$) si et seulement si, intuitivement, il y a des objets du domaine qui ont la propriété $\varphi(x)$, c'est-à-dire si et seulement si il existe un objet a tel que $M \models \varphi(x)\ \varphi[x{:=}a]$ où $\varphi[x{:=}a]$ est l'assignation qui est identique à

σ, sauf peut-être sur la variable x à laquelle elle assigne la valeur a. Ce faisant, on n'a fait qu'affirmer une certaine propriété de l'ensemble des objets qui satisfont $\varphi(x)$ (dans M, relativement à σ), à savoir que cet ensemble est non vide – puisqu'on peut choisir un a tel que... Autrement dit, le quantificateur existentiel $\exists$ peut être vu comme un prédicat de second ordre, qui est interprété sur M par une fonction $Q_\exists$ de l'ensemble des parties de M dans M, qui rend la valeur vrai si et seulement si on l'applique à un ensemble non vide.

On vient de définir $\exists$, et cette définition, que l'on pourrait donner sous la forme d'une définition explicite, s'effectue au sein de la métathéorie que l'on utilise pour définir la sémantique du système logique que l'on étudie, typiquement une version de la théorie des ensembles. On n'a pas encore pour autant isolé la classe des constantes logiques, puisqu'on n'a pas caractérisé l'ensemble des fonctions qui pouvaient être utilisées comme interprétation de symboles logiques. Dans le cas propositionnel, la question se pose à peine : les constantes habituelles sont interprétées par des fonctions de vérité, et toute fonction de vérité est l'interprétation potentielle d'un connecteur propositionnel. Dans le cas des quantificateurs, la chose est moins évidente. Par exemple, on pourrait considérer un quantificateur existentiel « il existe un ami de Jean tel que » qui donnerait la valeur vrai à tous les sous-ensembles du domaine d'interprétation qui sont non vides « en termes d'amis de Jean », c'est-à-dire à tous les sous-ensembles du domaine dont un élément au moins est un ami de Jean. Intuitivement, le quantificateur existentiel spécialisé en amis de Jean n'est pas un quantificateur logique, précisément parce qu'il est « spécialisé » et qu'il est lié à une propriété empirique, à savoir être un ami de Jean. Parmi tous les prédicats de second ordre que l'on peut imaginer, seuls ceux qui sont suffisamment généraux et suffisamment formels devraient être considérés comme logiques[4].

Tarski (1986), revenant sur le scepticisme qu'il avait exprimé dans son article sur la notion de conséquence logique, a proposé un critère[5] permettant de distinguer les opérations qui sont susceptibles de servir d'interprétation à des symboles logiques et celles qui ne le sont pas. Selon ce critère, on dira qu'une opération sur un domaine est logique si et seulement si elle est invariante par permutation. Dans le cas d'une opération Q servant à interpréter sur un domaine M un prédicat unaire de second ordre – Q est alors une fonction de l'ensemble des parties de M dans l'ensemble {V,F} des valeurs de vérité – on obtient la définition suivante : Q est *logique* si et seulement si pour toute permutation p sur M, et pour tout sous-ensemble A de M, Q(A)=Q(p(A)). Une permutation p est ici simplement une bijection de M dans M, et p(A) est l'ensemble des éléments x du domaine tels que $p^{-1}(x)$ est dans A. Typiquement, le quantificateur existentiel passe le test. Si $Q_\exists$(A)=V, c'est que A est non vide. Prenons alors une permutation arbitraire p. L'image de A par p est encore un sous-ensemble non vide, donc $Q_\exists$(p(A))=V. Inversement si A est vide, p(A) le sera tout autant. Il n'en va pas de même pour le pseudo quantificateur existentiel « il existe un ami de Jean tel que... ». En effet, on peut par exemple considérer un ensemble {Henri} où Henri est un des amis de Jean, et une permutation p qui envoie Henri sur Laurent, Laurent n'étant pas un ami de Jean. Dans ce cas, on a Q({Henri})=V, Q({Laurent})=F et p({Henri})={Laurent}. Donc Q n'est pas invariant par permutation.

4 Voir Bonnay (2008) pour une analyse conceptuelle détaillée des rapports entre généralité et formalité.

5 Ce critère a été utilisé indépendamment dans la littérature sur les quantificateurs généralisés. Voir Mostowski (1957). Sher (1991) présente une défense systématiquement de l'invariance par permutation comme critère de logicité.

Les quantificateurs de la logique du premier ordre sont interprétés par des opérations invariantes par permutation, et les prédicats de second ordre ayant un contenu empirique sont écartés par le critère de Tarski. Mais, à l'instar de « il existe une infinité de x tels que », de « il existe un nombre indénombrable de x tels que », ou encore de « il existe exactement $aleph_1$ x tels que » (où $aleph_1$ est le premier cardinal indénombrable), il y a beaucoup de quantificateurs qui ne sont pas des quantificateurs de la logique du premier ordre, qui ne sont pas non plus définissables au premier ordre, et qui sont cependant invariants par permutation et donc logiques selon le critère de Tarski. Si l'on accepte le critère de Tarski, cela suggère simplement que la logique du premier ordre n'épuise pas le domaine de la logique : on peut l'enrichir de nouveaux quantificateurs, comme les quantificateurs de cardinalités infinies précédents, et obtenir un langage que l'on peut encore qualifier à bon droit de logique. L'entreprise de définition de la logique nous amènerait à redéfinir les limites de la logique – si l'on considère que la logique du premier ordre marque les limites « classiques » de la logique, mais cela pourrait être précisément les vertus de l'entreprise que de nous montrer que ces limites doivent être reconsidérées.

6. La revanche de la simplicité

Qu'en est-il de la stratégie de définition que l'on vient d'esquisser relativement au problème particulier que pose la définition de la logique, à savoir le conflit entre le principe de simplicité de la logique et le réquisit de simplicité des définitions ? Les définitions que l'on donne pour les constantes logiques sont données dans la théorie des ensembles – les expressions logiques sont interprétées par des opérations sur les ensembles qui servent de domaine aux structures d'interprétation. Le réquisit de simplicité des définitions est violé d'emblée, pour autant que la théorie des ensembles est moins simple que la logique. En effet, la théorie des ensembles est une théorie mathématique particulière que l'on peut formuler grâce à la logique et il semble que ses axiomes et théorèmes jouissent d'une évidence moindre que les lois logiques – l'hypothèse du continu est un principe discuté, sur la validité duquel les théoriciens des ensembles peuvent être en désaccord, alors que la plupart des mathématiciens[6] s'accorderaient sur la validité des lois de la logique classique. Néanmoins, l'entreprise de définition ne perd pas tout son intérêt, dans la mesure où on accepte de ne pas la considérer uniquement dans une perspective fondationnelle. En adossant la logique à la théorie des ensembles, *via* une définition ensembliste des constantes logiques, on prend acte de ce que logique et mathématiques forment un tout. La logique fournit l'arrière-plan dans lequel les théories mathématiques sont développées. Réciproquement, les mathématiques fournissent l'arrière-plan dans lequel les langages logiques sont définis et étudiés[7], et l'on peut ce faisant obtenir des résultats métathéoriques hautement non triviaux portant sur les systèmes étudiés.

Mais alors le principe de simplicité de la logique est-il respecté ? Une fois qu'on a accepté de travailler dans un cadre ensembliste, il appartient précisément au critère de logicité, destiné à isoler parmi les opérations sur un domaine donné celles qui sont logiques, d'assurer que la spécificité des notions logi-

6 Bien sûr un mathématicien intuitionniste sera en désaccord avec un mathématicien classique par exemple quant à la validité du tiers-exclu. Il n'en restera pas moins que le développement des théories mathématiques présuppose un accord sur la logique dans laquelle ces théories sont développées, puisque ce qui compte comme un théorème d'une théorie mathématique donnée dépend des lois logiques acceptées. Au moins dans cette mesure, l'accord sur les vérités logiques apparaîtra comme plus fondamental que l'accord sur les vérités de telle ou telle théorie mathématique.

7 En particulier, Sher (1991) place cette idée d'échange théorique entre logique et mathématiques au cœur de sa conception de la logique associée à l'adoption du critère de Tarski.

ques est bien prise en compte. L'analyse conceptuelle de Tarski qui le mène au critère d'invariance par permutation vise bien dans une certaine mesure à rendre compte du caractère élémentaire de la logique, où l'élémentaire est assimilé au général. En effet, Tarski justifie l'adoption du critère d'invariance par permutation en référence au classement par Klein des théories géométriques à l'aide de leurs invariants. La géométrie euclidienne étudie les notions invariantes relativement aux transformations de l'espace qui préservent les ratios de distance. La topologie étudie les notions invariantes relativement aux transformations de l'espace qui préservent la continuité. Intuitivement, la topologie est plus élémentaire que la géométrie euclidienne ; la géométrie euclidienne s'intéresse à un type bien particulier d'espace, les espaces euclidiens, alors que la topologie ne s'intéresse qu'aux propriétés les plus générales de l'espace. Le point est alors que cette différence dans les degrés de généralité peut être mesurée à travers la relation d'ordre correspondant aux classes de transformation associées. Plus une théorie est invariante relativement à une large classe de transformations, c'est-à-dire plus elle fait abstraction des propriétés particulières des domaines étudiées, plus elle est générale. Toute transformation qui préserve les ratios de distance est continue, mais la réciproque n'est pas vraie. Pour cette raison, la topologie est plus générale que la géométrie euclidienne. Tarski, parce qu'il identifie la logique à la théorie la plus générale est alors conduit à l'identifier à l'étude des notions invariantes par la plus grande classe de transformations, qui n'est autre que la classe de toutes les permutations

Si l'on accepte l'équation entre généralité et simplicité, et la formalisation des degrés de généralité en termes d'ordre engendré par les classes de transformation associées, on devrait considérer que le critère de Tarski permet de sauvegarder le principe de simplicité de la logique alors même que le réquisit de simplicité des définitions a été abandonné. Ce critère sauvegarderait le principe de simplicité de la logique en isolant, parmi toutes les opérations définissables de manière ensembliste, celles qui ont la généralité qui sied aux opérations logiques. Mais on a vu que des propriétés comme « être le premier cardinal indénombrable » passaient le test de Tarski. Intuitivement, de telles notions sont pourtant typiquement des notions ensemblistes qui « manquent de simplicité » : la question de savoir ce qui compte comme premier cardinal indénombrable est disputée, puisque l'hypothèse du continu l'est. Il y aurait donc une sorte de revanche de la simplicité. En abandonnant le réquisit de simplicité des définitions, on a accepté de travailler dans un langage ensembliste pour définir la logique. Ce faisant, on a cherché à capturer la spécificité de la logique, sa généralité, à travers le critère de Tarski en termes d'invariance par permutation. Mais, sans doute parce que la théorie d'arrière-plan que l'on utilise, la théorie des ensembles, est une théorie tout sauf simple, on a néanmoins laissé s'infiltrer, pour ainsi dire, des notions éminemment complexes au sein des notions logiques.

Si le critère de Tarski capture bien l'idée de généralité, ce qui précède suggère que l'équation entre généralité et simplicité n'est pas correcte. Certaines opérations invariantes par permutations sont tout sauf simples. Donc si les opérations invariantes par permutation sont bien les opérations les plus générales, il y a des opérations parmi les plus générales qui ne sont pas pour autant simples. Conceptuellement, ceci n'est d'ailleurs pas tellement surprenant : la généralité d'une opération la caractérise relativement à son domaine d'application, une opération générale est indifférente aux propriétés particulières des objets. Mais pourquoi devrait-on s'attendre à ce que toutes les opérations de ce genre soient élémentaires ? Afin de ne pas faire perdre à l'entreprise de définition une grande partie de son intérêt, et afin d'obtenir une caractérisation de la classe des expressions logiques qui soit en accord avec le principe de simplicité de logique, il conviendrait donc de compléter le critère de Tarski par une contrainte *sui generis* de simplicité.

7. La simplicité en théorie des ensembles

La situation est alors la suivante. Dans une métathéorie ensembliste, on définit les opérations sur les domaines d'interprétation susceptibles de constituer des interprétations admissibles pour des symboles logiques. Parmi toutes les opérations possibles, seules les plus générales nous intéressent, ce sont celles qui satisfont le critère de Tarski. Mais parmi les plus générales, on ne voudrait encore retenir que les plus simples. Mais comment capturer la simplicité d'une opération ? On pourrait même objecter que la simplicité d'une opération est une propriété épistémique subjective, qui est relative à celui qui considère l'opération en question. Et que cette propriété épistémique dépend notamment du « niveau d'entrainement » du sujet en question. « Etre dénombrable » est certainement une notion complexe pour l'homme de la rue, elle cesse de l'être pour l'étudiant en mathématiques, et elle est « triviale » pour le théoricien des ensembles. Il serait alors douteux qu'il soit possible de capturer une bonne notion de simplicité pertinente pour la délimitation de la classe des expressions logiques. En effet, il n'existerait pas de notion objective de simplicité à capturer, et les notions subjectives variables de simplicité qui existent seraient à la fois difficiles à formaliser et peu pertinentes quant à la délimitation de la logique, s'il est vrai que l'on doit donner une réponse objective à la question de savoir ce qui compte comme logique.

A y réfléchir davantage, cette objection n'est peut-être pas aussi dirimante qu'il peut sembler au premier abord. En effet, on ne cherche pas à caractériser une notion « absolue » de simplicité. On a accepté de travailler dans un cadre métathéorique ensembliste, et *dans ce cadre*, on se demande ce qui compte comme des opérations élémentaires. La simplicité que l'on cherche à capturer est la simplicité relativement à la théorie des ensembles qui nous sert de théorie d'arrière-plan. La simplicité ne devrait pas alors être indexée sur les capacités cognitives variables des sujets, mais sur les propriétés de la théorie d'arrière-plan.

Qu'est-ce alors qu'une opération simple relativement à la théorie des ensembles dans laquelle cette opération est définie ? Pour esquisser une réponse à cette question, nous allons nous inspirer des travaux réalisés à partir des années 1970 sur les définitions ensemblistes des logiques abstraites et sur les logiques absolues[8]. Nous allons en fait voir deux lignes de réponses, convergentes. La première est sémantique : étant donné un modèle de la théorie des ensembles, une opération définie à l'intérieur de ce modèle pourrait légitimement être considérée comme simple si elle ne dépend pas trop des caractéristiques spécifiques de ce modèle. Imaginons que l'on passe d'un modèle particulier à un modèle plus 'gros' qui le contient, une opération simple devrait donner les mêmes résultats lorsqu'on passe du 'petit' au 'gros' modèle, et inversement. Considérons une opération qui détermine si un ensemble est dénombrable, c'est-à-dire s'il existe une bijection entre cet ensemble et l'ensemble des entiers. Plaçons-nous dans un modèle dénombrable de ZFC et considérons dans ce modèle un ensemble indénombrable particulier. Il n'existe pas *dans ce modèle* de bijection entre cet ensemble et l'ensemble des entiers du modèle. Mais rien n'empêche – et on peut le faire par forcing – de se placer dans un modèle plus grand dans lequel une telle bijection existe. Le même ensemble apparaîtra alors comme dénombrable dans ce nouveau modèle. En ce sens, « être indénombrable » n'est pas une propriété simple ; c'est une propriété qui dépend de

8 Voir la présentation générale, voir Väänänen, (1985).

l'existence de telle ou telle bijection dans le modèle, et des bijections qui n'existent pas dans un 'petit' modèle dénombrable peuvent exister dans un plus 'gros'. La notion technique d'absoluité[9], introduite par Gödel (1940), vise précisément à capturer cette idée. Relativement à une théorie T, une formule est dite persistante si elle est préservée lorsqu'on passe d'un modèle de T à une extension finale de ce modèle qui est encore un modèle de T. Une formule absolue relativement à T est une formule qui est persistante et dont la négation est encore persistante. Une opération absolue relativement à une théorie T est une opération définissable à l'aide d'une formule absolue. Si l'absoluité reflète bien le desideratum de simplicité, nous devrions donc nous intéresser aux opérations qui sont absolues relativement à la théorie ensembliste d'arrière-plan, disons aux opérations absolues relativement à ZFC.

Selon une autre ligne de réponse, la simplicité d'une opération définie de manière ensembliste pourrait être évaluée en fonction de la simplicité syntaxique de la formule qui la définit[10]. Typiquement, la complexité d'une formule est liée aux alternances de quantificateurs de cette formule. Soient $\varphi(x)$, $\varphi'(x,y,z)$ des formules sans quantificateurs non bornés, on considèrera volontiers que la formule $\exists x\varphi(x)$ est plus simple que la formule $\forall x\exists y\forall z\ \varphi'(x,y,z)$. On dit qu'une formule est Σ_1 si elle est équivalente à une formule comme $\exists x\varphi\ (x)$, c'est-à-dire à une formule constituée par un préfixe de quantificateurs existentiels et une formule sans quantificateurs non bornés. On dit qu'une formule est Π_1 si elle est équivalente à une formule comme $\forall x\varphi(x)$, c'est-à-dire à une formule constituée par un

9 Le fait que certaines opérations invariantes par permutation ne soient pas absolues est également mentionné par Feferman (1999), qui y voit un argument contre le critère de Tarski.

10 La pertinence d'une caractérisation syntaxique de la simplicité des opérations logiques, indépendamment du critère d'absoluité, nous a été suggérée par J. van Benthem.

préfixe de quantificateurs universels et une formule sans quantificateurs non bornés. Une formule est Δ_1 si elle est à la fois Π_1 et Σ_1. D'un point de vue syntaxique, une formule Δ_1 peut raisonnablement être considérée comme une formule exprimant une propriété « simple » puisqu'on peut la voir soit comme une formule purement existentielle, soit comme une formule purement universelle.

Cette approche syntaxique de la simplicité a ses limites. Si, intuitivement, la complexité quantificationnelle du *definiens* apparaît certes comme une bonne mesure du degré de complexité du *definiendum*, nous n'avons sans doute pas donné d'arguments complètement convaincants quant à la nécessité d'identifier simplicité et définissabilité *par une formule Δ_1*. Après tout, être définissable par une formule Δ_0 (sans quantificateurs non bornés), c'est être encore plus simple qu'être définissable par une formule Δ_1, et il en va *a fortiori* de même pour les opérations définissables sans quantificateurs du tout (même bornés). Néanmoins, un résultat technique, prouvé par Feferman et Kreisel (voir Feferman 1968), établit une corrélation remarquable entre simplicité syntaxique et absoluité : les formules absolues relativement à une théorie T sont précisément les formules Δ_1 relativement à cette théorie. Malgré le vague qui entoure le passage de la comparaison de la complexité des formules par la comparaison de leur complexité quantificationnelle à l'identification des formules Δ_1 comme les formules suffisamment simples, la convergence de la notion sémantique de simplicité – en termes d'absoluité – et de la notion syntaxique – être définissable par une formule Δ_1 – suggère que le concept de simplicité que l'on capture ainsi est relativement robuste. En effet, on peut dire, de manière quelque peu relâchée, que ce qui est simple lorsque l'on regarde les modèles est le même que ce qui est simple lorsque l'on regarde les formules.

Arrivés là, nous devrions considérer que les opérations logiques sont les opérations qui sont à la fois générales et simples. Cette caractérisation est une caractérisation précise, dans la mesure où le test de Tarski nous permet d'isoler les opérations générales et où la notion gödelienne d'absoluité nous permet d'isoler les opérations simples. Cette caractérisation n'est pour autant, en l'état, qu'une suggestion. Contentons-nous ici d'esquisser certains développements attendus. D'un point de vue technique, on voudrait identifier les systèmes qui vont compter comme logiques si l'on accepte de délimiter sur ces bases la classe des constantes logiques. Les résultats de Barwise (1972) sur les logiques « absolues » donnent des éléments de réponse, en pointant le rôle particulier joué par les logiques infinitaires comportant des conjonctions et disjonctions infinies mais des quantifications seulement sur des ensembles finis de variables,. D'autres questions sont ouvertes. La notion d'absoluité est relative à la théorie des ensembles sous-jacentes. ZFC est une théorie « riche ». On pourrait considérer qu'il serait davantage pertinent, si c'est la simplicité qui nous intéresse, de considérer une théorie des ensembles elle-même plus simple comme KP. De plus, l'articulation technique et conceptuelle entre invariance, généralité et simplicité demanderait à être analysée plus finement. Bonnay (2008) soutient, sur la base d'une conception plus large de la notion d'invariance, que l'invariance par permutation ne permet pas de capturer adéquatement la généralité, et propose d'utiliser la contrainte d'absoluité comme une contrainte non pas sur les opérations définies mais sur les notions d'invariance utilisées. La correspondance avec les logiques absolues au sens de Barwise n'est que partielle.

8. Conclusion

Nous avons cherché à préciser la nature du projet d'une définition de la logique, compris comme définition des constantes logiques, et à mettre en relief la difficulté conceptuelle propre d'un tel projet, à travers la tension entre le réquisit de simplicité des définitions et le principe de simplicité des définitions. Nous ne pensons pas que cette tension constitue un obstacle insurmontable, qui établirait la vanité de l'entreprise. Mais nous pensons que l'on peut évaluer toute tentative de réalisation à la lumière de la manière dont elle résout cette tension.

Nous avons discuté deux approches particulières, une approche preuve-théorique et une approche modèle-théorique, dont nous avons suggéré qu'elles correspondaient à deux manières différentes de négocier la résolution de la tension entre simplicité des définitions et simplicité de la logique. Nous avons cherché à éclairer des travaux antérieurs (Bonnay, 2008) à la lumière de cette problématique, travaux qui se plaçaient spécifiquement dans le cadre modèle-théorique. Définir de manière ensembliste les notions logiques, c'est renoncer d'emblée au réquisit de simplicité des définitions. Dans une perspective non fondationnelle visant à comprendre les échanges entre logique et mathématiques, un tel renoncement fait sens. Tout l'enjeu devient alors de voir si, et comment, l'on peut retrouver l'idée de simplicité de la logique alors même qu'on a accepté de travailler dans une théorie d'arrière-plan ensembliste « complexe ». Utiliser pour ce faire la notion d'absoluité, ou ses équivalents syntaxiques, constitue notre suggestion, qui se situe ainsi dans le prolongement des travaux mathématiques sur les logiques absolues.

Références bibliographiquess

BARWISE J. (1972). Absolute logics and $L_{\infty,\omega}$. *Annals of Mathematical Logic vol. 4*, 309-340.

BONNAY D. (2008). Logicality and invariance. *Bulletin of Symbolic Logic, vol. 14*.1, 29-68.

BONNAY D. & SIMMENAUER B. (2005). Tonk strikes back. *Australasian Journal of Logic vol. 3*, 33-44.

DOSEN K. (1994). Logical constants as punctuation marks. *Notre Dame Journal of Formal Logic vol. 30*, 362-81.

FEFERMAN S. (1968). Persistent and invariant formulas for outer extensions. *Compositio Mathematica vol. 20*, 29-52.

FEFERMAN S. (1999). Logic, logics, and logicism. *Notre Dame Journal of Formal Logic vol. 40*, 31-54.

GÖDEL K. (1940). The consistency of the axiom of choice and of the generalized continuum hypothesis with the axioms of set theory. *Annals of Mathematical Studies vol. 3*, Princeton, Princeton University Press. [Réimpr. in : S. Fefeferman *et alii* (eds), *Kurt Gödel, Collected Works*. Oxford : OUP, 1990, vol. II, 33-101].

HILBERT D. (1899). Lettre à Frege du 29 décembre 1899. In : F. Rivenc & Ph. Rouilhan de, *Logique et fondement des mathématiques*. Paris : Payot, 1992 [Tr. fr. J. Dubucs].

JORAY P. (2008). Définitions explicites et abstraction, ici-même, 135-157.

MOSTOWSKI A. (1957). On a generalization of quantifiers. *Fundamenta Mathematicae vol. 44*, 12-36.

PASCAL B. (1963) *De l'esprit géométrique et de l'art de persuader*. In : B. Pascal, *Œuvres complètes*. Paris : Seuil, 348-358.

SAMBIN G., BATTILOTTI G. & FAGGIAN C. (2000). Basic logic : reflection, symmetry, visibility. *Journal of Symbolic Logic vol. 65*, 979-1013.

SHER G. (1991). *The Bounds of Logic*. Cambridge : MIT Press.

TARSKI A. (1936). Über den Begriff der logischen Folgerung (On the concept of logical consequence). *Actes du Congrès International de Philosophie Scientifique vol. 7*, 1-11 [Trad. ang. in : Tarski 1983, 409-20].

TARSKI A. (1986). What are logical notions. *History and Philosophy of Logic vol. 7*, 143-154.

VÄÄNÄNEN J. (1985). Set-Theoretic definability of logics. In : J. Barwise & S. Feferman (eds), *Model-Theoretic Logics*. New York : Springer, 599-644.