

Axiomatiques minimales et définitions : la thèse de Tarski sur le calcul biconditionnel

Pierre Joray

1. Introduction : axiomatique et définition explicite

Si l'on considère la méthode axiomatique sous un angle historique, on est frappé par les évolutions qui ont marqué l'usage, le rôle et la position des définitions dans les théories axiomatisées. Pourtant, si ces changements participent à bien des égards aux profondes mutations dont la méthode axiomatique a elle-même été l'objet – des *Eléments* d'Euclide, en passant par la théorisation de Pascal, les axiomatiques de Hilbert jusqu'aux systèmes contemporains – on ne peut que constater que la procédure de définition n'a pas été soumise à deux des importantes réformes qui ont mené à l'avènement de la théorie des systèmes formels au 20^e siècle. En examinant le rôle joué par les définitions dans les systèmes formels d'aujourd'hui, on remarque que l'outil définitoire n'y a été ni formalisé, ni même intégré au processus de symbolisation, caractéristique des approches formelles contemporaines (Blanché 1955 : 55-60). En effet, depuis les *Principia Mathematica* de Whitehead et Russell, et ensuite les axiomatiques de Hilbert, les définitions se sont trouvées rejetées hors des systèmes axiomatisés. Reléguées au titre d'un outil métathéorique informel, elles servent à fournir, de l'extérieur du système, des abréviations permettant une lecture raccourcie des formules officielles de la théorie et attirant l'attention sur des combinaisons de significations jugées importantes ou

privilegiées relativement aux visées particulières dans lesquelles la formalisation est effectuée.

B. Russell, qui fut l'un des promoteurs les plus influents de cette conception devenue aujourd'hui standard, jugeait lui-même qu'elle conduisait à une situation paradoxale des définitions. Dès 1903, il écrivait dans ses *Principles of Mathematics*:

C'est un curieux paradoxe, énigmatique pour la pensée symbolique, que les définitions ne soient rien d'autre que des énoncés d'abréviation symboliques, sans signification pour le raisonnement et introduites uniquement par commodité pratique. Pourtant, dans le développement d'un sujet, elles exigent toujours une contribution importante de la pensée et enveloppent souvent les plus grands résultats de l'analyse. (1903 : 63).

Paradoxale, la situation des définitions dans les axiomatiques modernes l'est aussi lorsqu'on la rapporte aux enjeux de la réforme de la méthode géométrique que préconisait Blaise Pascal dans son opuscule *De l'esprit géométrique*. Par ce texte, Pascal a souvent été considéré comme un précurseur de la conception contemporaine de la définition. Il est le premier à exposer clairement pourquoi seules des définitions nominales peuvent avoir leur place dans une axiomatique et donc ce genre de définitions que l'on qualifie aujourd'hui d'*explicites*. Dans la méthode géométrique, explique-t-il, on ne doit reconnaître comme définition «que les seules impositions de noms aux choses que l'on a clairement désignées en termes parfaitement connus» (Pascal 1658 : 577). Quant aux définitions de choses, qui avaient la faveur de la tradition, Pascal les rejetait comme n'étant pas à proprement parler des définitions. Il considérait que leurs énoncés devaient soit être posés à titre d'axiomes, soit être traités comme des propositions à démontrer à partir des axiomes (*ibid.* : 581).

Par ces préceptes méthodologiques, Pascal s'engageait vers une introduction clairement différenciées des termes primitifs d'une part – ceux qui, comme on le dira plus tard, avec Hilbert, avec Carnap aussi, sont *implicitement* définis par les axiomes –

et des termes dérivés d'autre part, dont seules des définitions *explicitites* permettent l'introduction dans le langage du système.

Plus encore que Russell, Pascal aurait-il sans doute été surpris de constater que dans les axiomatiques d'aujourd'hui aucun terme défini ne peut apparaître dans les théorèmes officiels. De façon théorique, tous les énoncés démontrés qui contiennent des termes dérivés ne sont à proprement parler que des pseudo-théorèmes ou, plus exactement, des énoncés abrégés, des formulations pratiques et souvent éclairantes figurant, pour le lecteur, en lieu et place des théorèmes officiels qui, eux, ne peuvent contenir que des termes primitifs. On ne reconnaît en effet comme *preuves* dans un système formel que des suites finies et réglées de *formules* de ce système. Or la conception même des systèmes formels exige que toute formule soit le résultat d'une construction finie, opérée sur la base de règles de formation qui ne permettent en aucune manière l'introduction d'un terme dérivé. La notion de preuve s'appuie de façon cruciale dans la théorie des systèmes formels sur l'exigence d'une détermination *préalable* d'un ensemble clos de formules, constitué à partir d'un alphabet, clos lui aussi, ne contenant qu'un nombre réduit de constantes.

Malgré cet état de fait proprement théorique, la pratique courante de la preuve semble s'être pleinement accommodée de l'idée que des termes définis puissent apparaître dans les théorèmes et dans leurs preuves. Ce hiatus entre théorie et pratique de la preuve semble parfaitement assumé par les théoriciens contemporains des formalismes. Pourtant, cette facilité ne va pas sans difficultés. Bien entendu, chacun comprend qu'il convient pour s'en prémunir d'exiger des termes définis qu'ils soient éliminables des formules. Les définitions explicites doivent en effet permettre à tout moment que les termes dérivés puissent être remplacés dans les formules par les termes primitifs qui ont servi à les définir. Dit autrement, l'exigence d'éliminabilité doit garantir que l'on puisse toujours retrouver

les formules officielles dont les énoncés à termes définis sont les abréviations. Pourtant l'éliminabilité des formules n'est pas toujours suffisante pour garantir, sur la base d'une démarche déductive usant de termes définis, l'existence d'une preuve formelle conforme aux exigences officielles et déclarées de la base axiomatique. Comme je l'ai montré ailleurs, dans les systèmes formels d'une certaine complexité, des problèmes – en particulier de substitution – peuvent donner lieu à des termes définis parfaitement éliminables des formules, mais néanmoins pas toujours éliminables des preuves (Joray 2005). En clair, les moyens symboliques, abrégatifs et éliminables des formules, introduits par des définitions explicites ne sont pas toujours sans incidence sur la puissance déductive d'un système formel. Souvent associé à une dimension abstraite (Joray 2008), l'introduction d'abréviations dans un langage n'est pas un acte neutre. Qu'il suffise pour s'en convaincre intuitivement de songer à la fertilité évidente de l'introduction dans le langage de l'arithmétique des signes de somme, de produit, de la notation des puissances, des racines, des logarithmes, etc.

En termes techniques, la question de la fertilité déductive des définitions concerne la dimension proprement logique des théories. Elle ne se pose avec précision que dans les axiomatiques dites *formelles*, à savoir celles qui en plus des termes et axiomes propres à une théorie particulière – par exemple, mathématique – incluent d'une manière expresse un appareil déductif, c'est-à-dire un ensemble réunissant les moyens logiques mobilisés par la théorie en question.

Pascal n'avait certes aucune idée de ce qu'on nomme aujourd'hui une axiomatique formelle. Pourtant, un des points importants – quelque peu oublié aujourd'hui – de sa réforme de la méthode géométrique consistait à reconnaître que pour aller des axiomes vers les théorèmes le cheminement devait procéder non pas selon un seul moteur, à savoir la déduction logique, mais selon deux moteurs : la déduction et la définition explicite.

Manifestement, l'évolution moderne de la méthode géométrique vers les axiomatiques formelles contemporaines n'a pas retenu cette idée importante de Pascal, puisque seuls les outils de déduction se sont trouvés réglés et intégrés aux axiomatiques. La définition, par contre, est restée un outil intuitif, disponible uniquement à l'extérieur du langage formel, dans le métalangage, un outil dont l'usage n'est en aucune manière réglementé par la base axiomatique.

Sept ans après avoir parlé de *paradoxe* au sujet des définitions, Russell publiait avec Whitehead le monumental édifice logiciste des *Principia Mathematica*, dont l'essentiel des résultats se trouve en effet exprimé dans une longue liste de définitions cumulées, accompagnées de preuves incluant à tous niveaux des termes dérivés. Malgré l'intérêt et l'immense influence de cet ouvrage, on en reconnaît aujourd'hui les nombreuses déficiences formelles. Il est surprenant que la conception des définitions explicites qui y est présentée ait conservé une telle influence sur la pratique des logiciens, car l'usage des définitions y manque singulièrement de rigueur. Laissant de côté les problèmes spécifiques liés aux définitions contextuelles (*definition in use*), on remarquera d'emblée que la distinction entre termes primitifs et dérivés n'y est tout d'abord pas soigneusement respectée. Les premiers axiomes et règles d'inférence (pour la logique des propositions) contiennent des conditionnels en lieu et place des seuls primitifs déclarés, à savoir négation et disjonction (Rickey 1975). On relèvera aussi le statut ambivalent du signe spécial de définition ($=_{df}$) dont on ne sait s'il est partie intégrante ou non du langage formel. Les auteurs affirment bien que les définitions ne font pas partie du sujet, mais ils ajoutent aussi qu'elles donnent lieu à des théorèmes d'équivalence qui les traduisent pour ainsi dire immédiatement dans le langage objet (Whitehead & Russell 1927 : Vol.1, 11sqq). Non seulement l'équivalence ou biconditionnel doit lui-même être défini, mais en dehors de ce

premier problème, on se demande également ce qu'un tel théorème exprime. Le terme défini qu'il contient doit-il être compris comme une simple abréviation éliminable? Par exemple, en termes de négation et disjonction, on peut définir la conjonction ainsi :

$$A \wedge B =_{df} \sim(\sim A \vee \sim B).$$

Cet énoncé n'est pas un théorème (il n'est même pas une formule du langage), mais il donne lieu au théorème biconditionnel suivant:

$$\vdash (A \wedge B) \equiv \sim(\sim A \vee \sim B).$$

Ici, de deux choses l'une : ou bien cet énoncé exprime un rapport entre conjonction, disjonction et négation (auquel cas la conjonction a bien été intégrée au langage objet), ou bien il s'agit d'une simple abréviation et l'énoncé n'est qu'une manière pratique d'écrire le théorème suivant qui ne dit rien de la conjonction et exprime une tautologie banale:

$$\vdash \sim(\sim A \vee \sim B) \equiv \sim(\sim A \vee \sim B).$$

Aucune de ces deux interprétations ne permet de donner pleine cohérence à la conception de la définition défendue par les auteurs des *Principia*.

Plus soucieux de rigueur formelle, les logiciens de l'Ecole de Varsovie furent des lecteurs attentifs des *Principia*. Ils développèrent deux conceptions rivales de la définition explicite. La première et la plus fidèle aux idées de Whitehead et Russell fut celle développée par Jan Łukasiewicz. Selon lui, imposer une définition consiste à ajouter au système une *règle de remplacement* permettant d'introduire ou d'éliminer à tout moment d'une déduction un symbole abrégatif (Łukasiewicz 1963 : chap. II.4). De leur côté, Stanisław Leśniewski et Alfred Tarski considèrent que la partie logique de la base axiomatique

doit inclure une procédure définitoire permettant l'inscription progressive de thèses définitoires. Leśniewski fut le seul logicien à avoir examiné avec précision les conséquences d'une telle option sur les notions de *langage* et de *système formels* (Leśniewski 1930)¹. Tout comme il le fit pour la substitution, il décrivit soigneusement les conditions formelles d'une bonne définition explicite. Il fut probablement le premier à remarquer que les définitions explicites sont susceptibles d'être fécondes ou *créatives*, selon le sens qui fut donné à ce terme dans l'Ecole de Varsovie.

Une définition, explique par exemple Jan Łukasiewicz en 1928, est *créative* si elle permet «d'inférer (...) des thèses qui ne contiennent que des termes primitifs et qui sont pourtant indépendantes des axiomes» (1928b). De façon plus précise:

Une définition D d'une constante d est *créative* dans un langage L si et seulement si il existe une formule F de L , sans occurrence de d et telle que F n'est prouvable dans L qu'avec l'aide de D (éventuellement, d'une autre définition).

L'existence de définitions explicites créatives devait naturellement susciter la méfiance des logiciens qui, à la suite de Whitehead et Russell, concevaient la définition explicite comme purement abrégative. En faisant de la définition un outil formel déclaré de la base axiomatique – au même titre que les règles d'inférence – Leśniewski se donnait au contraire la possibilité d'en faire un outil déductif positif, concourant à la puissance démonstrative de ses systèmes axiomatiques.

Pour Leśniewski, c'est même ce qui constitue l'intérêt principal des définitions explicites. En 1928, pour répondre aux critiques faites par son collègue Łukasiewicz à l'encontre de la créativité des définitions, il affirmait: «les définitions conduisent effectivement à des thèses indépendantes des axiomes» et il ajoutait «ceci n'est pas un vice, bien au contraire. Si on inscrit

¹ Cf. aussi l'article de D. Miéville dans le présent volume.

des définitions, celles-ci devraient être les plus créatives possible» (cité dans Łukasiewicz 1928a). Désireux d'obtenir une base déductive incluant une procédure interne de définition et reposant sur une base axiomatique minimale, Leśniewski développa le système propositionnel quantifié le plus étendu qui soit, à partir d'une axiomatique ne contenant comme connecteur primitif que le biconditionnel.

Dans les pages qui viennent mon but n'est pas de présenter l'édifice achevé de ce calcul auquel Leśniewski donna le nom de *Protothétique*², mais d'examiner un travail qui reste relativement méconnu et qui fut mené à bien par Alfred Tarski lorsqu'il était le doctorant de Leśniewski. Les résultats présentés par Tarski dans cette thèse de doctorat publiée en 1923 constituent tout d'abord une pierre de touche de la Protothétique; sans leur découverte le système de Leśniewski eut difficilement pu avoir la forme remarquablement aboutie qu'on lui connaît aujourd'hui. Enfin, il donne à voir un excellent exemple de construction où l'inscription préalable de définitions explicites ouvrent les possibilités déductives liées à la présence d'une quantification.

2. La thèse de doctorat de A. Tarski

Alfred Tarski soutint sa thèse de doctorat en 1923 à l'Université de Varsovie. Ce travail fut dirigé par Stanisław Leśniewski, qui était à ce moment là en pleine élaboration du plus fondamental de ses systèmes de logique : la Protothétique, un calcul des propositions quantifié et d'ordre supérieur. Les résultats de la thèse de Tarski nous sont connus par un article publié la même année sous le titre «Sur le terme primitif de la logistique». Le but du travail y est présenté dès les premiers paragraphes :

² Pour une telle présentation, cf. Miéville 1984 et 2001.

Le problème pour lequel j'offre une solution est le suivant: *est-il possible de construire un système de la logistique comportant comme seul signe primitif, le signe d'équivalence* (en plus, bien sûr, des quantificateurs)? (1923 : 3)

Tarski montre qu'il est possible de construire un système fonctionnellement complet du calcul quantifié des propositions sur une base comprenant le biconditionnel comme unique connecteur primitif. Comme on le verra, la démarche de Tarski ne consiste pas à construire effectivement un tel système, mais à montrer qu'il existe entre les connecteurs principaux de la logique des propositions des rapports qui permettent d'exprimer la négation et la conjonction par une expression biconditionnelle quantifiée. Le texte poursuit alors ainsi:

Ce problème me semble intéressant pour les raisons suivantes. Nous savons qu'il est possible de construire le système de la logistique à l'aide d'un seul terme primitif, utilisant pour cela, soit le signe d'implication, si l'on veut suivre l'exemple de Russell, soit l'idée de Sheffer qui introduit spécialement à cet effet le signe d'incompatibilité. (*ibidem* : 4)³

Mais Tarski ajoute une remarque qui montre bien que son travail se situe au centre des réflexions de Leśniewski sur les définitions:

Pour réellement parvenir à ce but [sous-entendu: ni Russell, ni Sheffer n'y parviennent *réellement*], il faut se garder de faire entrer dans les énoncés des définitions tout terme constant particulier, distinct à la fois du terme primitif adopté, des termes préalablement définis, et du terme à définir. (*ibidem* : 4)

Pour Tarski, tout comme pour Leśniewski, l'énoncé d'une définition ne doit comprendre que trois types de symboles de constantes:

³ Dans le cas de l'implication (conditionnel), une quantification sur les variables de propositions est requise. Cf. l'article de D. Miéville dans ce volume.

1. Le terme primitif adopté (ici le biconditionnel, pour Russell le conditionnel et pour Sheffer la barre d'incompatibilité).
2. Les termes éventuellement déjà définis.
3. Un symbole nouveau pour le terme à définir.

Les raisons de ces conditions nous sont données dans une note où Tarski écrit :

Dans cet article, je considère les définitions comme appartenant au système de la logistique. Si nous avons utilisé un terme spécial pour formuler les définitions nous ne pourrions pas prétendre que nous n'acceptons qu'un seul terme primitif. On notera que dans [les *Principia Mathematica*] de A. N. Whitehead et B. Russell, toutes les définitions sont de la forme « $a = b \text{ Df}$ » et donc contiennent réellement un symbole spécial qui n'apparaît ni dans les axiomes, ni dans les théorèmes ; il semble cependant que ces auteurs ne considèrent pas les définitions comme des propositions appartenant au système. (*ibidem* : 4, n. 6)

Comme on le sait, toute bonne définition explicite doit être constituée d'un *definiendum* α , d'un *definiens* β et l'énoncé de la définition doit stipuler que α et β sont en une relation d'équivalence (une relation réflexive, symétrique et transitive). Le terme à définir ne doit avoir qu'une seule occurrence, dans le *definiendum*. Le *definiens* doit être construit uniquement avec ce que le système contient déjà, à savoir le(s) terme(s) primitif(s) et éventuellement des termes déjà définis. Enfin, on doit retrouver dans le *definiendum* et le *definiens* les mêmes éventuelles variables, chacune avec une seule occurrence dans le *definiendum*⁴.

Ces conditions générales ne sont pas sujettes à polémique, mais Tarski souligne dans sa note que si l'on veut que les définitions appartiennent au système, il convient d'exprimer l'équivalence entre *definiendum* et *definiens* à l'aide des outils

⁴ Cf. Miéville 1984, 2001 et son article dans le présent volume.

linguistiques à disposition dans la base du système, à savoir le(s) terme(s) primitif(s). C'est précisément ce que les auteurs des *Principia* manquent de faire lorsqu'ils recourent au symbole spécial « =_{df} ». Pour Tarski, qui reprend ici clairement une idée de Leśniewski:

Le signe d'équivalence [i.e. le biconditionnel] employé comme terme primitif présente à ce point de vue l'avantage de permettre d'observer strictement cette règle et de donner en même temps aux définitions une forme aussi naturelle que commode, celle d'une équivalence [i.e. une formule biconditionnelle]. (*ibidem* : 4)

En effet, si l'on reprend l'exemple évoqué plus haut de définition de la conjonction en termes de disjonction et de négation, plutôt que de poser hors du système l'expression

$$A \wedge B =_{df} \sim (\sim A \vee \sim B)$$

et d'ajouter qu'elle donne lieu au théorème biconditionnel suivant, dont on ne peut clairement déterminer le statut:

$$\vdash (A \wedge B) \equiv \sim (\sim A \vee \sim B),$$

il suffit de considérer ce second énoncé comme l'expression définitoire elle-même. Son statut devient alors celui d'une thèse d'équivalence, inscrite dans le système en vertu d'une règle déclarée dans la base axiomatique. Cette règle, qui fut soigneusement décrite par Leśniewski (1930), doit alors imposer les conditions formelles exactes sous lesquelles une thèse définitoire peut être ajoutée. Dans un système quantifié comme la Protothétique, l'expression définitoire d'une constante prendra la forme générale suivante:

$$(\forall v_1 \cdots v_n)(Dum(C, v_1, \cdots, v_n) \equiv Diens(v_1, \cdots, v_n))$$

ou, avec la notation de Leśniewski pour la quantification :

$$\lfloor v_1 \cdots v_n \rfloor \lceil Dum(C, v_1, \cdots, v_n) \equiv Diens(v_1, \cdots, v_n) \rceil$$

En endossant cette conception de la définition explicite, Tarski entend montrer la faisabilité du projet de Leśniewski: construire une logique des propositions quantifiée, munie d'une règle définitoire. Le choix du biconditionnel comme unique connecteur primitif s'impose alors comme celui qui permet l'expression la plus simple et la plus directe des thèses d'équivalence définitoires⁵.

Dans le cadre non quantifié, on sait que le biconditionnel est trop faible. Contrairement à la barre de Sheffer, il ne permet pas l'expression de l'ensemble classique des connecteurs unaires et binaires. Dans le cadre quantifié, Leśniewski sait comment obtenir la négation, il lui suffit pour cela de poser la thèse définitoire suivante:

$$\lfloor p \rfloor \lceil \sim p \equiv (p \equiv \lfloor r \rfloor \lceil r \rceil) \rceil$$

Il lui manque cependant encore l'accès à certains connecteurs binaires, en particulier la conjonction, la disjonction, le conditionnel.

C'est ce dernier obstacle que Tarski va franchir en construisant une définition de la conjonction. Sa solution est constituée par le théorème 11 de son travail, qui présente effectivement la forme requise d'une thèse définitoire de la Protothétique:

Th. 11 :

$$\lfloor pq \rfloor \lceil (p \wedge q) \equiv \lfloor f \rfloor \lceil p \equiv (\lfloor r \rfloor \lceil p \equiv f(r) \rceil \equiv \lfloor r \rfloor \lceil q \equiv f(r) \rceil) \rceil \rceil$$

⁵ On peut également utiliser à cette fin d'autres primitifs, mais les thèses définitoires sont alors plus complexes et demandent toujours plusieurs inscriptions du *definiens* et du *definiendum*. Cf, à ce sujet, Joray 2004, Lejewski 1958.

où l'inscription

$$\lfloor f \rfloor \lceil p \equiv (\lfloor r \rfloor \lceil p \equiv f(r) \rceil \equiv \lfloor r \rfloor \lceil q \equiv f(r) \rceil) \rceil$$

est le *definiens*, qui sera abrégé dans la suite par Π_{pq} .

Tout le début de l'article de Tarski est consacré à montrer que cette expression constitue une définition possible de la conjonction en termes de biconditionnels et de quantification portant sur des variables propositionnelles ainsi que sur des variables de connecteurs unaires. Dans une seconde partie, Tarski va également montrer que, si le système dans lequel on définit la conjonction présente un axiome d'extensionnalité⁶, alors la définition de la conjonction peut être nettement simplifiée. C'est ce que montre son théorème 17, où l'antécédent *Sb* du conditionnel exprime l'extensionnalité (ou la vérifonctionnalité) des connecteurs unaires et le conséquent constitue l'expression biconditionnelle pour une définition simplifiée de la conjonction:

$$\text{Th. 17: } Sb \supset \lfloor pq \rfloor \lceil (p \wedge q) \equiv \lfloor f \rfloor \lceil p \equiv (f(p) \equiv f(q)) \rceil \rceil$$

Pour montrer ces résultats, Tarski va se placer dans une logique propositionnelle riche: un langage quantifié dans lequel il suppose comme étant déjà à disposition, en plus du biconditionnel, le conditionnel, la conjonction et la négation. En usant des règles d'inférence usuelles de ces connecteurs et du quantificateur, il va alors donner des preuves des expressions concernées.

Comme le Th. 11 exprime une relation d'équivalence entre la conjonction d'une part et une expression biconditionnelle quantifiée de l'autre, Tarski aura dès lors établi qu'il peut servir de

⁶ A savoir un axiome qui impose que tous les connecteurs expriment des fonctions de vérité. Dans sa thèse, Tarski n'utilise pas le terme «extensionnalité», mais parle de «loi de substitution».

définition de la conjonction dans une logique à base plus modeste, caractérisant uniquement le biconditionnel et la quantification. Dans le cas du Th. 17, la base axiomatique adéquate devra encore inclure Sb au nombre de ses axiomes (ou de ses théorèmes). Mais laissons pour l'instant la question de l'extensionnalité et examinons la solution générale du Th. 11.

3. La démonstration du théorème 11

Avant d'examiner la démarche de Tarski, il convient de faire deux remarques concernant le rôle des définitions. Tout d'abord, ce que montre clairement la preuve du Th. 11, c'est que l'introduction de la quantification renforce le calcul propositionnel puisque cet ajout modifie les possibilités définitoires des ensembles de connecteurs primitifs. Le Russell des *Principles of Mathematics* avait déjà remarqué cette caractéristique, car pour développer le calcul propositionnel à partir du seul conditionnel, il introduisait une quantification sur les propositions. Enfin, ce qui est moins évident et qui, à mon sens, n'a pas été particulièrement souligné dans la littérature, c'est le rôle que jouent les définitions préalables de connecteurs unaires dans la démonstration des propriétés caractéristiques de la conjonction introduite par l'expression biconditionnelle de Tarski. Comme on le verra, la possibilité d'instancier des connecteurs unaires est nécessaire pour obtenir à partir de la définition tarskienne les deux versions de l'élimination de la conjonction.

Mais voyons maintenant comment Tarski procède. Je rappelle que la démonstration du Th. 11 est donnée dans une logique quantifiée comprenant biconditionnel, conditionnel, conjonction et négation. Par ailleurs, Tarski ne donne pas une preuve formelle de son théorème, mais s'appuie sur les règles admises de ces connecteurs. De fait, son texte ne présente que ce qu'il qualifie lui-même de «commentaires montrant les

étapes du raisonnement». La démarche est néanmoins très claire et peut se résumer en six étapes.

1. Tout d'abord, trois nouveaux connecteurs unaires sont introduits par le biais des expressions définitoires suivantes:

Déf. 1 $\lfloor p \rfloor \lceil \forall p \equiv (p \equiv p) \rceil$ le *Verum*

Déf. 2 $\lfloor p \rfloor \lceil Ap \equiv p \rceil$ l'*Assertion*

Déf. 3 $\lfloor p \rfloor \lceil Fp \equiv (p \equiv \sim p) \rceil$ le *Falsum*

Tarski dispose alors des quatre connecteurs unaires qui épuisent les significations vérifonctionnelles. On notera que la négation n'a pas besoin d'être définie puisqu'elle est supposée déjà à disposition. On sait cependant que celle-ci pourrait être exprimée en termes du biconditionnel et de la quantification. On peut en effet montrer aisément que l'équivalence suivante, déjà citée plus haut, est un théorème du calcul propositionnel quantifié où Tarski présente ses preuves:

$$\lfloor p \rfloor \lceil \sim p \equiv (p \equiv \lfloor r \rfloor \lceil r \rceil) \rceil$$

2. Dans la deuxième étape, Tarski démontre que le *definiens* qu'il propose pour la conjonction implique le second conjoint, à savoir q . C'est le théorème 5 de Tarski:

Th. 5 $\lfloor pq \rfloor \lceil \Pi_{pq} \supset q \rceil$

3. Dans la troisième étape (théorème 8), il montre que son *definiens* implique également le premier conjoint, c'est-à-dire p :

Th. 8 $\lfloor pq \rfloor \lceil \Pi_{pq} \supset p \rceil$

4. La quatrième étape consiste à réunir les deux résultats précédents par une conjonction:

$$\text{Th. 9} \quad \lfloor pq \rfloor \left[\Pi_{pq} \supset (p \wedge q) \right]$$

5. Dans la cinquième étape, Tarski montre que son *definiens* est impliqué par la conjonction de p et q :

$$\text{Th. 10} \quad \lfloor pq \rfloor \left[(p \wedge q) \supset \Pi_{pq} \right]$$

6. Enfin la dernière étape consiste à réunir les théorèmes conditionnels 9 et 10 en un théorème biconditionnel unique:

$$\text{Th. 11} \quad \lfloor pq \rfloor \left[(p \wedge q) \equiv \Pi_{pq} \right]$$

De fait, si on examine le détail des moments de la démonstration, on constate que les étapes 4, 5 et 6 ne présentent aucune difficulté et que le nœud de la démarche se trouve aux étapes 2 et 3. Celles-ci correspondent à l'élimination de la conjonction ; Tarski y montre que chacun des conjoints est impliqué par son *definiens*. Or, c'est précisément lors de ces étapes que se trouvent exploitées les définitions des connecteurs unaires posées en 1. En effet, pour déduire d'une part p , d'autre part q , à partir du *definiens* Π_{pq} , il convient d'éliminer le quantificateur principal de cette expression:

$$\Pi_{pq} : \lfloor f \rfloor \left[p \equiv (\lfloor r \rfloor \left[p \equiv f(r) \right]) \equiv \lfloor r \rfloor \left[q \equiv f(r) \right] \right].$$

Le quantificateur en question lie une variable f dont la catégorie est celle des connecteurs unaires. Son élimination nous amène donc à substituer à f un symbole ou une expression de même catégorie. Vu les éléments à disposition dans le langage, il n'y a ici que deux possibilités: soit substituer à f une variable qui sera alors en occurrence libre, soit substituer une constante de connecteur unaire. La première des possibilités

étant manifestement sans intérêt pour la démonstration recherchée, il reste quatre instanciations possibles pour f , qui correspondent aux quatre connecteurs unaires disponibles. Par l'élimination du quantificateur, on obtient alors les quatre formules suivantes:

1. $p \equiv (\lfloor r \rfloor \lceil p \equiv Vr \rceil \equiv \lfloor r \rfloor \lceil q \equiv Vr \rceil)$ $\Pi_{pq}, \lfloor \rfloor e, f / V$
2. $p \equiv (\lfloor r \rfloor \lceil p \equiv Fr \rceil \equiv \lfloor r \rfloor \lceil q \equiv Fr \rceil)$ $\Pi_{pq}, \lfloor \rfloor e, f / F$
3. $p \equiv (\lfloor r \rfloor \lceil p \equiv \sim r \rceil \equiv \lfloor r \rfloor \lceil q \equiv \sim r \rceil)$ $\Pi_{pq}, \lfloor \rfloor e, f / \sim$
4. $p \equiv (\lfloor r \rfloor \lceil p \equiv Ar \rceil \equiv \lfloor r \rfloor \lceil q \equiv Ar \rceil)$ $\Pi_{pq}, \lfloor \rfloor e, f / A$

Désormais, en exploitant les définitions 1-3, et les règles d'inférence usuelles, on obtient, à partir de Π_{pq} , les quatre déductions suivantes (dont je ne donne ici que des esquisses):

I. Cas du *Verum*

1. $\lfloor f \rfloor \lceil p \equiv (\lfloor r \rfloor \lceil p \equiv f(r) \rceil \equiv \lfloor r \rfloor \lceil q \equiv f(r) \rceil) \rceil \Pi_{pq}$
2. $p \equiv (\lfloor r \rfloor \lceil p \equiv Vr \rceil \equiv \lfloor r \rfloor \lceil q \equiv Vr \rceil)$ 1, $\lfloor \rfloor e, f / V$
3. $Vr \equiv (r \equiv r)$ Déf. 1
4. $p \equiv (\lfloor r \rfloor \lceil p \equiv (r \equiv r) \rceil \equiv \lfloor r \rfloor \lceil q \equiv (r \equiv r) \rceil)$ 2, 3
5. $p \equiv (p \equiv q)$ 4
6. $(p \equiv p) \equiv q$ 5, *Ass* \equiv
7. q 6

II. Cas du *Falsum*

1. $\lfloor f \rfloor \lceil p \equiv (\lfloor r \rfloor \lceil p \equiv f(r) \rceil \equiv \lfloor r \rfloor \lceil q \equiv f(r) \rceil) \rceil \Pi_{pq}$
2. $p \equiv (\lfloor r \rfloor \lceil p \equiv Fr \rceil \equiv \lfloor r \rfloor \lceil q \equiv Fr \rceil)$ 1, $\lfloor \rfloor e, f / F$
3. $Fr \equiv (r \equiv \sim r)$ Déf. 3
4. $p \equiv (\lfloor r \rfloor \lceil p \equiv (r \equiv \sim r) \rceil \equiv \lfloor r \rfloor \lceil q \equiv (r \equiv \sim r) \rceil)$ 2, 3
5. $p \equiv (\sim p \equiv \sim q)$ 4
6. $p \equiv (p \equiv q)$ 5
7. $(p \equiv p) \equiv q$ 6, *Ass* \equiv
8. q 7

III. Cas de la *Négation*

1. $\lfloor f \rfloor \lceil p \equiv (\lfloor r \rfloor \lceil p \equiv f(r) \rceil \equiv \lfloor r \rfloor \lceil q \equiv f(r) \rceil) \rceil \quad \Pi_{pq}$
2. $p \equiv (\lfloor r \rfloor \lceil p \equiv \sim r \rceil \equiv \lfloor r \rfloor \lceil q \equiv \sim r \rceil)$ 1, $\lfloor _ \rfloor e, f / \sim$
3. $\left| \begin{array}{l} \lfloor r \rfloor \lceil p \equiv \sim r \rceil \\ p \equiv \sim p \\ \sim(p \equiv \sim p) \end{array} \right. \quad \text{Hyp.}$
4. $\left| \begin{array}{l} \lfloor r \rfloor \lceil p \equiv \sim r \rceil \\ p \equiv \sim p \end{array} \right. \quad 3, \lfloor _ \rfloor e, r / p$
5. $\left| \begin{array}{l} \lfloor r \rfloor \lceil p \equiv \sim r \rceil \\ p \equiv \sim p \\ \sim(p \equiv \sim p) \end{array} \right. \quad \text{Th.}$
6. $\sim \lfloor r \rfloor \lceil p \equiv \sim r \rceil \quad 3, 4, 5, \sim i$
7. $\left| \begin{array}{l} \lfloor r \rfloor \lceil q \equiv \sim r \rceil \\ q \equiv \sim q \\ \sim(q \equiv \sim q) \end{array} \right. \quad \text{Hyp.}$
8. $\left| \begin{array}{l} \lfloor r \rfloor \lceil q \equiv \sim r \rceil \\ q \equiv \sim q \end{array} \right. \quad 7, \lfloor _ \rfloor e, r / q$
9. $\left| \begin{array}{l} \lfloor r \rfloor \lceil q \equiv \sim r \rceil \\ q \equiv \sim q \\ \sim(q \equiv \sim q) \end{array} \right. \quad \text{Th.}$
10. $\sim \lfloor r \rfloor \lceil q \equiv \sim r \rceil \quad 7, 8, 9, \sim i$
11. $\sim \lfloor r \rfloor \lceil p \equiv \sim r \rceil \equiv \sim \lfloor r \rfloor \lceil q \equiv \sim r \rceil \quad 6, 10, \equiv i$
12. $\lfloor r \rfloor \lceil p \equiv \sim r \rceil \equiv \lfloor r \rfloor \lceil q \equiv \sim r \rceil \quad 11$
13. $p \quad 2, 12, \equiv e$

IV. Cas de l'Assertion

1. $\lfloor f \rfloor \left[p \equiv (\lfloor r \rfloor \left[p \equiv f(r) \right] \equiv \lfloor r \rfloor \left[q \equiv f(r) \right]) \right] \Pi_{pq}$
2. $p \equiv (\lfloor r \rfloor \left[p \equiv Ar \right] \equiv \lfloor r \rfloor \left[q \equiv Ar \right])$ 1, $\lfloor _ \rfloor e, f / A$
3. $\left| \begin{array}{l} \lfloor r \rfloor \left[p \equiv Ar \right] \end{array} \right.$ Hyp.
4. $\left| \begin{array}{l} p \equiv A \sim p \end{array} \right.$ 3, $\lfloor _ \rfloor e, r / \sim p$
5. $\left| \begin{array}{l} p \equiv \sim p \end{array} \right.$ Déf. 2
6. $\left| \begin{array}{l} \sim (p \equiv \sim p) \end{array} \right.$ Th.
7. $\sim \lfloor r \rfloor \left[p \equiv Ar \right]$ 3, 5, 6, $\sim i$
8. $\left| \begin{array}{l} \lfloor r \rfloor \left[q \equiv Ar \right] \end{array} \right.$ Hyp.
9. $\left| \begin{array}{l} q \equiv A \sim q \end{array} \right.$ 8, $\lfloor _ \rfloor e, r / \sim q$
10. $\left| \begin{array}{l} q \equiv \sim q \end{array} \right.$ Déf. 2
11. $\left| \begin{array}{l} \sim (q \equiv \sim q) \end{array} \right.$ Th.
12. $\sim \lfloor r \rfloor \left[q \equiv Ar \right]$ 8, 10, 11, $\sim i$
13. $\sim \lfloor r \rfloor \left[p \equiv Ar \right] \equiv \sim \lfloor r \rfloor \left[q \equiv Ar \right]$ 7, 12, $\equiv i$
14. $\lfloor r \rfloor \left[p \equiv Ar \right] \equiv \lfloor r \rfloor \left[q \equiv Ar \right]$ 13
15. p 2, 14, $\equiv e$

Relativement au *definiens* Π_{pq} de la conjonction, on voit que les quatre connecteurs unaires vérifonctionnels se rangent en deux groupes. D'une part, le *Verum* et le *Falsum* (qui expriment des fonctions constantes) permettent lorsqu'ils sont substitués à la variable f de déduire le second conjoint q . D'autre part, la *Négation* et l'*Assertion* (dont le résultat est fonction de la valeur de leur argument) permettent la déduction du premier conjoint p . Ainsi pour disposer des deux versions de l'élimination de la conjonction à partir de la définition de Tarski, il convient que le langage où celle-ci est posée contienne au moins deux connecteurs unaires, l'un choisi dans le premier groupe, l'autre dans le second⁷.

Dans sa thèse, Tarski ne s'appuie que sur deux des déductions présentées ici, celle qui exploite le *Verum* et celle utilisant l'*Assertion*. Il faut cependant noter que, si les quatre déductions s'appuient largement sur les règles concernant le biconditionnel et le quantificateur, les deux déductions menant au premier conjoint p (cas de la *Négation* et de l'*Assertion*) s'appuient de façon essentielle sur les propriétés de la négation. Ainsi, si l'on s'interroge sur la construction d'une base axiomatique quantifiée, ne contenant que le biconditionnel et permettant d'exploiter la définition exprimée par le Th. 11, on peut d'ores et déjà conclure que celle-ci devra bien entendu garantir les propriétés usuelles du biconditionnel et du quantificateur, mais qu'elle devra également donner accès à la définition de la négation (le *Verum* et l'*Assertion* étant très évidemment accessibles dans un tel système et le *Falsum* aisément définissable une fois la négation à disposition).

⁷ On trouve une généralisation de ce résultat dans Hiž 1952. Malheureusement, il s'agit seulement d'un résumé de conférence qui ne contient qu'une esquisse de démonstration.

4. Théorème 17 et extensionnalité

Comme je l'ai indiqué plus haut, la démarche de Tarski ne fait que montrer la *possibilité* de construire une logique propositionnelle quantifiée complète sur la base du seul connecteur biconditionnel. Si l'examen des preuves menant au théorème 11 montre bien cette possibilité et permet aussi de dégager les caractéristiques générales auxquelles doit répondre une telle base axiomatique, on doit bien constater que ni Tarski, ni Leśniewski ne présentent une construction explicite permettant d'exploiter directement le résultat du Th. 11. La raison est assez évidente pour celui qui connaît la Protothétique. Leśniewski désirait en effet disposer d'une logique très large et en particulier d'ordre supérieur. Cela signifie que toute catégorie syntaxico-sémantique devait y être accessible et que la définition d'une constante d'une quelconque catégorie devait ouvrir la possibilité d'une quantification sur des variables de cette catégorie⁸. Pour réaliser un projet d'une telle ambition, il était nécessaire de restreindre le système de manière à ce que chaque catégorie ne donne lieu qu'à un nombre fini de significations possibles. L'idée était alors la suivante: si dans chaque catégorie il devait rester possible de définir une infinité de constantes (syntaxiquement) distinctes, des théorèmes d'extensionnalité devaient être disponibles, afin de garantir que toutes les constantes se rangent en un nombre fini de classes d'équivalence sémantiques (dans les cas les plus connus, 2 valeurs propositionnelles, 4 valeurs de connecteurs unaires, 16 valeurs pour les binaires, etc).

On comprend dès lors qu'après avoir montré son théorème 11, Tarski entreprenne d'examiner une forme simplifiée de définition de la conjonction, efficiente dans un système qui,

⁸ Cf. l'article de Miéville dans le présent volume.

comme le dit Tarski, «contient la loi de substitution parmi ses axiomes ou ses théorèmes» (1923 : 7), à savoir un axiome ou un théorème garantissant la vérifonctionnalité ou l'extensionnalité au moins pour la catégorie des connecteurs unaires. Dans les termes que Tarski reprend à Whitehead et Russell, une telle loi prend la forme de l'expression:

$$Sb : \lfloor p q f \rfloor \lceil ((p \equiv q) \wedge f(p)) \supset f(q) \rceil$$

Si l'on veut se restreindre à l'usage du biconditionnel, l'extensionnalité des connecteurs unaires s'exprime alors:

$$Ext : \lfloor p q \rfloor \lceil (p \equiv q) \equiv \lfloor f \rfloor \lceil f(p) \equiv f(q) \rceil \rceil$$

Cette expression étant prouvable dans la Protothétique de Leśniewski, il est alors possible d'y accéder à la conjonction par la définition simplifiée indiquée par le Th. 17 de Tarski, à savoir:

$$\lfloor p q \rfloor \lceil (p \wedge q) \equiv \lfloor f \rfloor \lceil p \equiv (f(p) \equiv f(q)) \rceil \rceil$$

Contrairement à son collègue Łukasiewicz, Leśniewski était un défenseur d'une conception fermement classique de la logique. Il considérait comme absolument valides les principes de bivalence, du tiers-exclu et de contradiction. Dans cette perspective, il tenait également les théorèmes d'extensionnalité de la Protothétique pour des expressions de vérités logiques. Il n'avait ainsi aucune raison de se passer de la simplification importante que lui procurait une base axiomatique incluant la formule *Ext* (ou les moyens de sa preuve).

Pourtant, le Th. 11 de la thèse de doctorat de Tarski montre clairement qu'il est possible d'accéder à la conjonction dans un calcul biconditionnel quantifié ou d'une Protothétique non soumise au principe d'extensionnalité. A ma connaissance, cette possibilité n'a jamais été exploitée.

5. Epilogue

Au terme de ce parcours, il serait sans doute prématuré de conclure par quelque résultat ferme et j'espère seulement avoir montré que l'intégration des définitions explicites au sein des appareils déductifs ouvre de nombreuses perspectives à l'axiomatisation. Ce que nous montre la thèse trop méconnue de Tarski, c'est d'abord que les questions liées à l'axiomatisation du plus élémentaire des systèmes de logiques sont dépendantes de la conception des définitions que l'on adopte. La démonstration du Th. 11 met en évidence le rôle fructueux des définitions préalables de connecteurs unaires dans l'accès qu'une base axiomatique restreinte au biconditionnel peut offrir à la conjonction et donc à l'entière des connecteurs usuels. Bien entendu, montrer la créativité, au sens stricte, des différentes définitions examinées dans ces pages nécessiterait de travailler sur des bases axiomatiques explicites: celle de la Protothétique dans le cas où c'est la définition simplifiée de la conjonction qui est exploitée, une base plus restreinte, sans axiome d'extensionnalité, dans le cas plus général concerné par le Th. 11 de Tarski.

L'existence de définitions explicites créatives dans les systèmes quantifiés de logique des propositions est un fait démontré (Joray 2005), dont l'idée remonte aux logiciens de Varsovie. Les démarches examinées plus haut montrent que les définitions des connecteurs unaires sont sans doute créatives dans le contexte des calculs biconditionnels quantifiés. Il reste encore à se demander si la définition de la conjonction elle-même serait créative dans ces systèmes axiomatiques.

Dans le cas de la Protothétique (à savoir lorsqu'on dispose d'un axiome d'extensionnalité), Tarski donne dans sa thèse des éléments qui inclinent à une réponse négative à cette question. En effet, à la fin de la seconde section de son article, il indique

comment montrer que le produit logique des deux thèses définitoires de la conjonction et de la disjonction est équivalent à l'expression de l'extensionnalité des connecteurs unaires (dont les axiomes garantissent déjà la preuve). En revanche, dans le cas d'une base sans axiome d'extensionnalité, la question reste pleinement ouverte. Pour ce dernier cas, il serait à mon sens intéressant d'examiner si l'inscription de la définition de la conjonction exprimée par le Th. 11 de Tarski n'offre pas précisément un accès purement logique à l'extensionnalité des connecteurs unaires.

Les résultats de Tarski sur les systèmes biconditionnels sont connus dans les études leśniewskiennes comme ayant constitué une pierre de touche de la construction de la Protothétique. Mais que le Th. 11 n'ait pas été directement exploité dans toute sa généralité laisse encore de nombreuses questions ouvertes, relatives au rôle des définitions et de la quantification dans les calculs propositionnels. Parmi ces questions, celle par exemple de l'existence d'une preuve logique du principe d'extensionnalité est d'un intérêt philosophique qui dépasse largement les problèmes techniques liés à la recherche d'une axiomatique minimale pour la logique des propositions.

Bibliographie

- BLANCHÉ R. 1955. *L'axiomatique*, Paris : PUF.
- HIŻ H. 1952. On primitive terms of logic (abstract of a paper), *The Journal of Symbolic Logic* **17**, 156-7.
- JORAY P. 2004. A note on definitions in propositional calculi, *Travaux de logique* **17**, 170-182.
- JORAY P. 2005. Should definitions be internal? in Bilkova M. & Behounek L. (eds) *The Logica Yearbook 2004*, Praha: Filosofia, 189-199.
- JORAY P. 2006. La définition dans les systèmes logiques de Łukasiewicz, Leśniewski et Tarski, dans Pouivet R. & Rebuschi M. (éds) *La philosophie en Pologne 1918-1939*, Paris :Vrin. 203-222.
- JORAY P. 2008. Définitions explicites et abstraction, *Travaux de logique* **19**, 135-157.
- LEŚNIEWSKI S. 1930. Ueber Definitionen in der sogenannten Theorie der Deduktion, *Comptes rendus des séances de la Société des Sciences et des Lettres de Varsovie* **iii.24** (1931), 142-70. [Trad. angl. dans McCall 1967 et Leśniewski 1992].
- LEŚNIEWSKI S. 1992. *Collected Works I, II*, S.J. Surma, J.T. Szrednicki, D.I. Barnett (eds), Varsovie : PWN, Polish Scientific Publisher ; Dordrecht : Kluwer.
- ŁUKASIEWICZ J. 1928a. O definicyach w teori dedukcyi (Sur les définitions dans les systems déductifs), *Ruch Filozoficzny* **11**, 177-8. [Trad. fr. par Błaszczuk M. en annexe de Joray 2006].
- ŁUKASIEWICZ J. 1928b. Rola definicyj w systemach dedukcyjnych (Le role des definitions dans les systems déductifs), *Ruch Filozoficzny* **11**, 164. [Trad. fr. par Błaszczuk M. en annexe de Joray 2006].
- ŁUKASIEWICZ J. 1963. *Elements of mathematical logic*, Warszawa: PWN; Oxford: Pergamon Press.
- MCCALL S. (ed) 1967. *Polish Logic 1920-1939*, Oxford: Clarendon Press.

- MIÉVILLE D. 1984. *Un développement des systèmes logiques de S. Leśniewski. Protothétique-Ontologie-Méréologie*. Berne : P. Lang.
- MIÉVILLE D. 2001. *Introduction à l'œuvre de S. Leśniewski. Fascicule I : La protothétique*. Université de Neuchâtel : *Travaux de logique*.
- RICKEY V. F. 1975. On creative definition in the *Principia Mathematica*, *Logique et Analyse* **18**, 175-182.
- TARSKI A. 1923. Sur le terme primitif de la logistique, dans Tarski 1972, *Logique, sémantique, métamathématique 1923-1944*, Paris : A. Colin. Vol. 1, 3-25.
- WHITEHEAD A. N. & RUSSELL B. 1927. *Principia Mathematica*, Cambridge University Press. [First ed. 1910-13]