Publié dans Travaux de logique 16, 59-71, 2005,
source qui doit être utilisée pour toute référence à ce travail

Les cercles, les cercles vicieux et leur principe

Cédric Degrange

Introduction

A la suite de la publication des *Principia Mathematica*, le principe du cercle vicieux (pcv) est devenu le test classique pour déterminer la nature prédicative ou imprédicative d'une définition. Ce principe règle le rapport entre une collection et les éléments de celle-ci: «Whatever involves *all* of a collection must not be one of the collection» (Whitehead & Russell 1927: 37). Si une définition enfreint ce principe, elle est imprédicative et par là, considérée comme pouvant entraîner l'émergence d'une antinomie. Cependant, ce genre de définitions sont seulement *susceptibles* de produire une contradiction et n'entraînent pas systématiquement un tel résultat. Cette relativisation du caractère antinomique des définitions imprédicatives trouve un certain écho chez les auteurs même du pcv. Dans la première introduction des *Principia*, au chapitre consacré à la théorie des types logiques, Whithead et Russell notent:

> Arguments which are condemned by the vicious-circle principle will be called "vicious-circle fallacies". Such arguments, in certain circumstances, may lead to contradictions, but it often happens that the conclusions to which they lead are in fact true, though the arguments are fallacious. (1927: 37-38).

Bien qu'ils constatent que le pcv puisse être un principe trop fort, Whithead et Russell l'admettent tel quel – faute de mieux, dira-t-on. Cependant, sur ce passage, il convient de faire deux remarques qui permettent d'éclairer en quoi les arguments de forme circulaire (*the vicious-circle fallacies*) sont seulement *susceptibles* d'être problématiques. Premièrement, pour que ces arguments mènent à une contradiction, il faut que «certaines conditions» soient réunies. Autrement dit, de tels énoncés ne sont pas problématiques en eux-mêmes. Il ne suffit pas qu'une théorie contienne un argument circulaire pour qu'il y ait apparition automatique d'une contradiction. Celle-ci doit encore être construite. Ainsi la présence de la loi V rend la théorie des *Grundgesetze* de Frege inconsistante. Cependant, le caractère inconsistant de la théorie est apparue seulement lorsque Russell a montré qu'il était possible de construire une antinomie à l'intérieur de celle-ci. Il y a contradiction, si l'énoncé circulaire est utilisé dans certaines conditions. Par conséquent, une solution pour éviter les contradictions serait non d'écarter les énoncés circulaires, mais les conditions dangereuses dans lesquelles ils peuvent être utilisés. C'est dans une telle approche, de valeur pragmatique, que ce sont notamment engagées les logiques paraconsistantes[1].

Le second point est que ce passage des *Principia* relève l'existence d'un autre genre d'énoncés circulaires. A la différence des premiers, ceux-ci ne mènent à aucune contradiction, et cela quelles que soient les circonstances. En d'autres termes, bien que marqués du défaut de circularité, ces arguments n'entraînent pas d'inconsistance – en ce sens, leurs conclusions sont «vraies». Néanmoins, suivant les *Principia*, puisque ces énoncés transgressent le pcv, ils sont considérés comme imprédicatifs.

Le constat est donc le suivant: d'un côté, il y a des énoncés circulaires qui peuvent engendrer l'inconsistance de la théorie; de l'autre, il y a des énoncés qui, même s'ils sont circulaires, n'engendrent pas d'inconsistance. Suivant la théorie des *Principia*, il suffit qu'un énoncé transgresse le pcv pour être considéré comme imprédicatif et

[1] Cf. par exemple Volken (1997: 257-271).

donc comme devant être éliminé de la théorie. Le but de cette présentation est de montrer qu'un énoncé peut être à la fois circulaire et non problématique. Ce faisant, nous montrerons qu'il est possible d'intégrer de tels énoncés imprédicatifs au sein d'une construction logiciste comme celle qui est discutée dans le présent volume.

1. Un exemple propositionnel

Whitehead et Russell proposent un exemple d'énoncé qui, malgré son caractère circulaire, semble bien être non contradictoire.

> Take, for example, the law of excluded middle, in the form «all propositions are true or false». If from this law we argue that, because the law of excluded middle is a proposition, therefore the law of excluded middle is true or false, we incur a vicious-circle fallacy. (1927: 38).

De prime abord, bien qu'il y ait cercle, il semble que cette formulation de la loi du tiers exclu doive plutôt être vraie. Le problème est la présence du cercle.

Précisons que la difficulté ici n'est pas liée à la loi du tiers exclu mais à une formulation possible de celle-ci. C'est la formulation ou l'expression de cette loi qui, du fait qu'elle enfreint le pcv, est problématique.

Une manière de résoudre la difficulté serait de faire appel à la distinction entre langage et métalangage. Dans ce cadre, une certaine proposition peut très bien porter sur l'ensemble des propositions d'un langage, sans appartenir à ce langage. Ainsi, lorsqu'il est affirmé que «Toutes les propositions sont vraies ou fausses», les propositions dont il est question appartiennent à un premier langage L. Par contre, l'expression, elle, appartient à un autre langage, $L + 1$, qui porte sur L. En d'autres termes, $L + 1$ est un métalangage. La formulation de la loi du tiers exclu se comprend alors comme signifiant: «Toutes les propositions du langage L sont vraies ou fausses» et cette formulation appartient au métalangage $L + 1$. Formellement, cela donne «$\square \vee \neg \square$», où $\square$ est une métavariable à laquelle peut être substituée n'importe quelle proposition appartenant à L. Quant à la vérité de cette formule

– ou, le cas échéant, sa fausseté – elle est affirmée au niveau du métalangage. C'est là la position généralement adoptée après les travaux de Tarski (1936).

Cependant, la conception de la logique adoptée ici n'autorise pas une telle solution. En effet, dans un paradigme comme celui des *Principia*, la logique est conçue comme un langage universel, de sorte que, fondamentalement, il n'y a qu'un seul langage, celui de la logique. Il n'est donc pas possible de distinguer un langage qui porterait sur la logique car cela signifierait que ce langage est plus fondamental que celui de la logique. Tout doit se faire à l'intérieur de cette dernière. La difficulté que représente la formulation de la loi du tiers exclu doit ainsi être résolue à l'intérieur du langage de la logique.

Dans ce paradigme, pour exprimer l'idée que toutes les propositions sont concernées, il nous faut faire usage d'une quantification qui porte sur les propositions. Ainsi, nous obtenons l'expression suivante:

(1) $(\forall p)\,(p \vee \neg p)$

Si on considère que l'expression (1) est une proposition – ce qui semble bien être le cas, bien que son ordre ne soit pas déterminé –, elle doit être comptée parmi cette totalité. Dès lors, pour évaluer la proposition (1), il est nécessaire de connaître l'évaluation de la proposition suivante:

(2) $[(\forall p)\,(p \vee \neg p)] \vee \neg[(\forall p)\,(p \vee \neg p)]$

Cette nouvelle proposition est obtenue par la substitution de la proposition (1) à la variable p. Ici, nous nous fondons sur une conception intuitive de la substitution.

Si nous cherchons à évaluer cette dernière proposition, une difficulté apparaît. Nous nous trouvons en effet dans une situation telle que l'évaluation de la proposition (2) requiert l'évaluation de la proposition (1) qui elle-même requiert l'évaluation de (2). Il y a donc cercle.

2. La solution des *Principia Mathematica*

La solution des *Principia* consiste à diminuer la portée de l'expression «toutes les propositions»: «"All propositions" must be in some way limited before it becomes a legitimate totality [...]» (1927: 38) Cela signifie limiter la totalité sur laquelle porte le quantificateur et donc restreindre le domaine de quantification selon certaines conditions. Suivant la théorie des types formulée dans les *Principia*, ces conditions sont notamment définies relativement au type et et à l'ordre de la proposition dans laquelle apparaît la quantification.

Cela ne va évidemment pas sans conséquence. Ainsi, il suit que *la formulation* de la loi du tiers exclu ne peut être interprétée comme une généralisation illimitée. La formulation doit être comprise non plus comme s'appliquant à la totalité des propositions, mais seulement à une totalité de propositions de type et d'ordre déterminés. Cela signifie qu'elle ne peut être traduite par l'expression (1). Dans les *Principia*, cette dernière ne peut être considérée comme une expression du langage logique et donc être considérée comme une proposition[2].

Dans la perspective des *Principia*, l'expression «toutes les propositions» ne signifie dès lors plus «la totalité des propositions» mais seulement «la totalité d'une partie des propositions». La lecture de la formulation de la loi du tiers exclu s'en trouve altérée ainsi qu'une compréhension intuitive de celle-ci.

La raison de limiter la portée de l'expression «toutes les propositions» est d'obtenir «une totalité légitime». On peut se demander si c'est là le seul moyen d'atteindre ce but. Il n'est peut-être pas nécessaire de procéder à une limitation du domaine de quantification. On peut aussi se demander en quoi consiste une totalité légitime, plus précisément ce qui fait la légitimité d'une telle totalité.

Suivant la théorie des types des *Principia*, le langage est hiérarchisé de telle manière que les énoncés d'un certain ordre et d'un certain type ne portent que sur des énoncés d'ordre et de type inférieurs. Dans le cadre propositionnel, c'est surtout l'ordre de la

[2] Dans les *Principia*, de telles expressions sont considérées comme dénuées de signification.

proposition qui importe. La hiérarchisation des propositions dérive de celle des fonctions dans laquelle l'ordre et le type doivent être pris en considération. Cependant, les propositions sont obtenues par la généralisation de toutes les variables libres présentes dans les fonctions, de sorte que la distinction de type n'est plus pertinente[3]. Ainsi, une proposition d'ordre n doit contenir au moins une variable liée d'ordre $n - 1$. Ce qu'il importe de remarquer, c'est que chaque niveau de langage ainsi obtenu, et donc de domaine de quantification, est homogène du point de vue des propositions qu'il peut contenir – celles-ci ne peuvent dépasser l'ordre de la proposition quantifiée. Il suit qu'une expression ne peut appartenir à un domaine de valeur d'ordre et de type inférieurs aux siens. Suivant les *conditions posées*, le domaine de valeur s'avère ainsi homogène.

Le caractère légitime d'une totalité est assuré par l'homogénéité du domaine de quantification. Cependant, nous ne sommes pas tenus d'imposer des conditions d'homogénéité sur l'ordre et le type des expressions. De même, nous ne sommes pas non plus tenus de restreindre le domaine de quantification. Ainsi, nous pouvons adopter d'autres conditions d'homogénéité. Le choix de ces conditions dépend des buts à atteindre – si ce qui importe est l'élimination de toutes les formes circulaires, la théorie des types est, somme toute, bien adaptée. Ici, nous voulons notamment pouvoir admettre certains énoncés circulaires. En outre, une solution qui ne remette pas en question la portée générale de l'expression «toutes les propositions» est souhaitable.

Suivant notre alternative, nous ne chercherons pas à restreindre le domaine de quantification. De même, plutôt que de considérer la forme syntaxique des termes susceptibles d'être quantifiés, nous

[3] A partir d'une *matrice* d'ordre 2 – $f(\phi!z, x)$ – toutes les fonctions d'ordre 2 engendrées par généralisation sont soit de type 1 – $(\phi)f(\phi!z, x)$, une fonction d'individu – soit de type 2 – $(x)f(\phi!z, x)$, une fonction de fonction d'individu. Les propositions d'ordre 2 sont obtenues à partir des fonctions du même ordre par généralisation de toutes les variables libres. La hiérarchie des types perdure dans le cas des propositions. Puisque les propositions dérivent des fonctions, elles respectent de manière atavique cette hiérarchie mise en place dans le cadre des fonctions. Cependant, cela n'empêche pas que la distinction de type perde de sa pertinence dans le cadre propositionnel, ex: $(x)(\phi)f(\phi!z, x)$ et $(\phi)(x)f(\phi!z, x)$.
Pour une présentation de la théorie voir: Vernant (1993), de Rouilhan (1996), Linsky (1999). Pour un exposé plus technique: Hatcher (1982).

privilégierons la signification de ces termes. De cette manière, comme nous le verrons, il est possible d'obtenir un domaine de quantification homogène tout en préservant le caractère général de la formulation de la loi du tiers exclu.

3. Vers une alternative catégorielle

Nous quittons désormais le cadre des *Principia* mais conservons une conception de la logique comme langage universel. De plus, nous considérons les expressions (1) et (2) comme des propositions. En fait, cette alternative[4] requiert de les considérer comme telles – ce qui ne me semble pas une requête par trop singulière.

Nous l'avons déjà vu, l'évaluation de la proposition (2) impose l'évaluation de la proposition (1). Si nous considérons cette dernière du point de vue de ses significations possibles, elle est soit vraie, soit fausse. En affirmant cela, nous supposons que nous nous trouvons dans un cadre bivalent et que les propositions ne désignent pas des états de choses. Dès lors, quelle que soit la valeur attribuée à la proposition (1), l'évaluation de la proposition (2) donne le vrai. Formellement, cela se présente comme suit:

$$(3) \quad \begin{pmatrix} V \\ F \end{pmatrix} \vee \neg \begin{pmatrix} V \\ F \end{pmatrix}$$

Ici, la forme syntaxique de la proposition (1) n'est pas pertinente. Seule importe qu'elle soit de la catégorie des propositions et que les significations liées à cette catégorie soit le vrai ou le faux.

Dans le cas de la proposition (1), la difficulté supplémentaire pour l'évaluer provient de la présence d'une quantification. Nous pouvons cependant étendre notre stratégie en adoptant une nouvelle lecture de

[4] Cette alternative s'appuie en grande partie sur les travaux liés à l'interprétation de l'Ontologie de Leśniewski, produits par Simons (1985), Miéville (1999) et Joray (1999, 2001, 2005e).

la quantification[5]. Dans cette perspective, la proposition se lit comme suit:

> Quelle que soit la signification qui peut être associée à une expression de la catégorie des propositions, si cette signification est attribuée à la variable *p*, alors l'expression «$p \vee \neg p$» est vraie / fausse.

Ici, le domaine de quantification est celui des significations attachées à la catégorie de la variable. Autrement dit, selon cette lecture, lorsque la quantification lie une variable, son domaine est l'ensemble des significations susceptibles d'être associées à un quelconque élément de la catégorie à laquelle appartient cette variable.

On procède de la manière suivante. Tout d'abord, il s'agit de déterminer la catégorie C à laquelle appartient la variable liée par le quantificateur (dans notre exemple, p est de la catégorie S des propositions). Le domaine de quantification lié à cette catégorie, $\mathrm{Dom}(C)$, consiste en les valeurs associées à cette catégorie. Si, par exemple, nous choisissons d'associer à la catégorie des propositions comme significations les valeurs du vrai ou du faux, le domaine de quantification s'exprimera:

$$\mathrm{Dom}(S) = \{V, F\}$$

Il est ensuite possible d'évaluer la proposition suivant les différentes combinaisons des valeurs possibles de S et suivant les clauses d'évaluations associées aux constantes de l'expression. Plus simplement, une fois spécifié le domaine de quantification, l'évaluation de (1) se présente sous la forme d'une table de vérité:

$$(4)\ \mathrm{Dom}(S) = \{V, F\}$$

$$\begin{pmatrix} V \\ F \end{pmatrix} \vee \neg \begin{pmatrix} V \\ F \end{pmatrix}$$

Rapidement, il apparaît que l'évaluation donne toujours le vrai. Formellement, la différence entre les expressions (3) et (4) – respectivement les interprétations de (1) et (2) – réside dans la

[5] Nous reprenons ici ce que Joray (2001: 220-230; 2005e) nomme l'interprétation *catégorielle* de la quantification.

présence pour (4) de la description du domaine de quantification. Notons que cette seule description marque la présence d'une quantification.

Bien que la formulation de la loi du tiers exclu transgresse le pcv – et demeure ainsi circulaire – par rapport à l'interprétation des *Principia*, elle possède une évaluation déterminée. Ainsi, nous admettons qu'il y a des énoncés circulaires qui ont une signification déterminée de sorte qu'ils peuvent être intégrés dans notre langage.

4. Extensionnalité

La solution présentée repose sur une interprétation particulière de la quantification. Celle-ci possède pour domaine l'ensemble des significations associées à la catégorie de la variable de sorte que nous obtenons une totalité homogène. Cette homogénéité du domaine de quantification est, comme nous l'avons remarqué plus avant, l'un des points requis pour éviter les *vicious-circle fallacies*.

En choisissant de concevoir le domaine de quantification comme l'ensemble des significations associées à la catégorie de la variable, implicitement, nous avons choisi une conception extensionnelle. Par là, il faut entendre que nous avons pris le parti de concevoir la notion de catégorie comme s'appliquant à un ensemble d'expressions individuelles. Ce choix a pour conséquence que des expressions telles que $p \supset q$, $p \wedge q$, $p \vee q$ ou $(p \wedge q) \supset m$ sont de catégories S, de même que les symboles de propositions atomiques p, q, m. Quant aux opérateurs, $\supset$, $\wedge$, $\vee$, ils sont de catégories S/SS (formateurs de proposition – de catégorie S – à partir de deux arguments de catégorie S).

Si nous admettons ce cadre extensionnel, il est ensuite possible d'associer à chaque catégorie une manière de signifier. Cela se fait selon le principe suivant:

> Deux expressions peuvent avoir la même signification *ssi* elles sont de la même catégorie.

Par exemple, à la catégorie S, nous pouvons choisir d'associer – comme nous l'avons fait précédemment – l'ensemble de significations

$\{V, F\}$. Nous dirons que la manière de signifier d'une expression de la catégorie S est de prendre pour signification une des valeurs V ou F. Ainsi, une expression de la catégorie S peut être vraie ou fausse. Suivant le principe adopté, cet ensemble est spécifique à cette catégorie. L'ensemble spécifique de significations associé à la catégorie *S/SS* est plus complexe. Il contient seize valeurs:

$$\{(V, V, V, V)^1, (V, V, V, F)^2, (V, V, F, V)^3, \ldots, (F, F, F, V)^{15}, (F, F, F, F)^{16}\}$$

Ces valeurs correspondent aux seize signatures possibles pour un opérateur binaire. Un foncteur de la catégorie *S/SS* prendra ainsi pour signification l'une de ces seize valeurs.

Cette correspondance univoque entre une catégorie et l'ensemble de significations qui lui est associé assure une homogénéité du domaine de quantification. Il suffit de connaître la catégorie de la variable pour connaître l'ensemble de significations qui est le domaine de quantification. Cette homogénéité est assurée même dans le cas d'un domaine de quantification infini. Considérons le cas d'une logique multivalente dans laquelle est associée à la catégorie de base S un ensemble composé d'une infinité de valeurs.

$$\mathrm{Dom}(S) = \{v_1, v_2, v_3, \ldots v_{n-1}, v_n, v_{n+1}, \ldots\}$$

Une expression de catégorie S pourra prendre pour signification une des valeurs v_i appartenant à cet ensemble infini. Dans le cas d'une expression de catégorie *S/SS*, on représentera l'ensemble d'attributions possibles de la manière suivante:

$$\mathrm{Dom}(S/SS) = \{\varphi \mid \varphi : \mathrm{Dom}(S) \times \mathrm{Dom}(S) \rightarrow \mathrm{Dom}(S)\}$$

Il s'agit de l'ensemble des applications possibles φ, telle que φ a pour domaine $\mathrm{Dom}(S) \times \mathrm{Dom}(S)$ et pour codomaine $\mathrm{Dom}(S)$. A la catégorie *S/SS* est ainsi associée de manière univoque un ensemble d'attributions de valeurs. Le domaine de quantification d'une expression d'une quelconque catégorie peut bien être infini, il n'en demeure pas moins qu'il est propre, ce qui nous assure de l'homogénéité du domaine.

Un point à préciser est celui de la substitution. Suivant notre approche de la quantification, celle-ci porte sur les significations associées à la catégorie de la variable présente dans le quantificateur. Pour que la substitution soit possible, la quantification doit porter sur les mêmes significations associés, ce qui entraîne que p, de catégorie S, ne peut être substitué que par une expression de catégorie S. Ce faisant, le domaine de quantification est assuré de conserver son homogénéité.

Les conditions d'extensionnalité et de substitution doivent pouvoir nous assurer de toujours avoir des totalités homogènes, c'est-à-dire légitimes. Ainsi, alors que la théorie des *Principia* fait peser des conditions d'homogénéité par le biais des notions d'ordre et de type, dans notre perspective celles-ci sont réglées par les notions de significations et de catégorie[6].

5. Cercles et cercles vicieux

Il nous faut encore éclairer un point qui concerne la distinction entre cercle et cercle vicieux. Nous avons vu que suivant notre solution, il est possible d'admettre au sein du langage formel des énoncés qui transgressent le pcv. Ainsi, bien qu'ils soient circulaires, ils ne s'avèrent pas être vicieusement circulaires – au sens où, les énoncés qui transgressent le pcv sont vicieusement circulaires pour la théorie des *Principia*.

La situation est similaire à celle de l'exemple de F. P. Ramsey: «l'homme le plus grand du groupe» (1925: 41). Cette description est circulaire et enfreint le pcv. Cependant, le cercle formé réside au niveau de l'expression et ne concerne pas les individus du groupe. Il s'agit d'une description possible d'un individu déjà donné. Le cercle serait vicieux, si cette description signifiait la création d'un nouvel individu, ce qui raisonnablement n'est pas le cas ici.

Dans le cas de la formulation de la loi du tiers exclu, l'aspect circulaire concerne l'expression de celle-ci. Par contre, la signification de l'expression n'est pas affectée. Il n'y a pas création d'une nouvelle

[6] L'opposition doit être nuancée du fait que l'usage de la notion de catégorie donne lieu à une hiérarchie des énoncés telle que celle issue de la théorie des types *simple*. La notion de catégorie correspond en ce sens à celle de type. Sur les rapports entre catégories et types, cf. ici même l'article de N. Gessler.

signification mais seulement la description circulaire d'une signification déjà déterminée.

Ce faisant deux niveaux sont au moins à distinguer. Le premier concerne l'expression de cette loi comme, par exemple, l'énoncé (1). Il s'agit de ce que nous appellerons le niveau syntaxique. Le deuxième niveau est celui sémantique et concerne la signification de l'énoncé (1). Plus précisément, il concerne les différentes combinaisons de valeurs qui sont la signature (sémantique) de l'expression et qui, dans notre exemple, est le vrai pour les deux combinaisons possibles. Le niveau sémantique ne doit pas être confondu avec celui auquel appartient, par exemple, l'énoncé (4) qui est *l'expression de la signification*[7] de l'énoncé (1).

Le caractère circulaire de la formulation du tiers exclu se situe ainsi au niveau syntaxique. Cependant, le caractère vicieux est désamorcé du fait qu'il ne concerne pas le niveau sémantique où notre interprétation de la quantification situe le domaine de valeur des variables. A ce niveau-ci, il n'y a pas de collusion entre l'ensemble des significations d'une catégorie et une quelconque de ces significations. En fait, au niveau sémantique le rapport entre les membres d'une totalité et cette totalité respecte le pcv. Autrement dit, si le pcv ne s'avère pas pertinent au niveau syntaxique, il le reste au niveau sémantique, là où se situe notre quantification et donc la où se règle le rapport entre une totalité et un quelconque de ses membres.

Conclusion

La solution présentée repose principalement, comme nous venons de le voir, sur une interprétation particulière de la quantification. Une telle interprétation est présente dans les systèmes de Leśniewski, même si elle ne leur est pas spécifique. Cette lecture catégorielle de la quantification peut se faire dans un système classique pour autant qu'il soit conçu dans un cadre extensionnel. Le choix de l'extensionnalité apparaît être la clef de cette solution aux énoncés circulaires.

[7] Nous parlerons d'une syntaxe de la sémantique.

L'un des aspects de cette solution est qu'elle éclaire l'intérêt de la distinction syntaxe / sémantique dans un paradigme logiciste. En mettant en évidence, ces différents niveaux de langue au sein d'un langage, nous avons vu qu'il est possible de s'en servir pour gérer le cas d'énoncés circulaires en les intégrant au langage. Ce dernier peut ainsi être enrichi et sa portée affinée. Notons qu'il n'est pas question de nier le caractère circulaire de ces énoncés. Au contraire de la solution des *Principia* qui évacue toute circularité, notre interprétation supporte sans difficulté les cercles en jouant de ces différents niveaux de langues.