PARADOXE POUR PARADOXE
(une critique du “ dialéthéisme ”)

Les logiques dites “ paraconsistantes ” refusent l'idée que d'une contradiction de la forme A et non-A, on puisse déduire n'importe quel énoncé B[1]. La justification de ce refus est aisée : admettre ce type d'inférence est contre-intuitif. Trouver, ou inventer, une explication de l'invalidité de ce type d'inférence est en revanche moins facile.

L'explication proposée par Graham Priest est particulièrement radicale. Nous devons nous habituer à l'idée que certains énoncés sont à la fois vrais et faux, et donc *paradoxaux* en ce nouveau sens : précisément les énoncés contradictoires, ceux qui étaient communément vécus comme paradoxaux au sens ancien, et donc comme insupportables. Typiquement, le Menteur, l'énoncé qui dit de lui-même qu'il est faux, est à la fois vrai et faux. Telle est la thèse du “ dialéthéisme ” : la vérité, telle Janus, peut offrir deux visages.

Puisque A peut être à la fois vrai et faux, sa négation non-A aussi. Dans un tel cas, la contradiction A et non-A est vraie (bien qu'elle soit fausse aussi, donc paradoxale). Prenez n'importe quel B faux, et sous une définition plausible de l'implication, vous avez que A et non-A n'implique pas B. Rejet de *Ex Falso Quodlibet Sequitur*.

Dans l'article de 1979 dont on lira ci-dessous une traduction, “ The Logic of Paradox ”, Priest construit un argument sophistiqué en faveur de cette thèse, argument basé sur l'analyse de la situation créée par le premier Théorème d'incomplétude de Gödel. Parmi tous les arguments apportés en faveur de la paraconsistance, c'est certainement le plus intéressant et le plus astucieux. Je voudrais néanmoins montrer qu'il est : 1) lacunaire ; 2) probablement circulaire ; et 3) un cas de *Non Sequitur*.[2]

***

On peut présenter l'argument de Priest en quatre temps :

1 - Imaginons un système formel S, qui formalise *toutes* nos procédures naïves de preuve, dont celles à l'oeuvre dans les mathématiques, procédures entendues comme des moyens d'aller du vrai-évident au vrai indirectement établi par les dites preuves. Ce système contient certainement le système P de Gödel (disons : l'arithmétique de Peano en second ordre, plongée dans une logique avec types) ainsi que certaines extensions considérées par Gödel, et tombe donc sous le coup du premier Théorème d'incomplétude. Il y a un énoncé g écrit dans le langage de P (et donc de S) qui, traduit conformément à l'interprétation attendue du système, “ dit de lui-même qu'il n'est pas prouvable ”. Cet énoncé, l'énoncé dit de Gödel, n'est pas prouvable formellement dans P, comme Gödel l'a montré (II.2), ni dans S.

2 - Cependant, on peut prouver “ par un raisonnement sémantique naïf ”, que g est vrai. En substance, le raisonnement qui établit ce point est le suivant. Soit $\mathbf{g}$ le nombre-code de Gödel de g ; on a montré que $g = \neg\text{Prov}(\mathbf{g})$ n'est pas prouvable dans P, où le prédicat “ $\text{Prov}(x)$ ” exprime la prouvabilité formelle de

[1] Autrement dit, on peut admettre des théories qui soient inconsistantes relativement à la négation, mais non triviales, i.e. absolument inconsistantes. Dans ce qui suit, je parlerai d'inconsistance (tout court) pour l'inconsistance relativement à la négation.

[2] Je m'inspire en partie dans ce qui suit des remarques critiques de Chihara 1984 dans “ Priest, The Liar, and Gödel ”.

x dans P. Maintenant, en vertu des équivalences-T de Tarski (qui disent en gros qu'un énoncé est vrai ssi ce qu'il dit est le cas), on a[3] :

$$\neg\text{Prov}(\mathbf{g}) \text{ est vrai } \underline{\text{ssi}}\ g \text{ n'est pas prouvable,}$$

et comme précisément g n'est pas prouvable, on a : g est vrai, par substitution des identiques (II.3).

Le raisonnement sémantique naïf intervient dans cette preuve de la vérité de l'énoncé indémontrable sous la forme des équivalences-T, qui régissent l'usage du prédicat de vérité.

Priest pense que cette situation est paradoxale : car d'une part g n'est pas prouvable dans S, c'est-à-dire n'est pas prouvable *du tout*, puisque S est censé représenter toutes nos procédures de preuve. Et pourtant, il y a une preuve de la vérité de g, preuve qui ne devrait pas pouvoir exister.

3 - En fait, c'est que le raisonnement sémantique naïf dont nous avons un exemple en 2 - n'a pas été formalisé dans S. Si maintenant nous ajoutons à S ce type de raisonnement, la situation paradoxale décrite plus haut disparaît, de même d'ailleurs que le résultat de Gödel : car g est finalement prouvable (II.6). Cela revient à faire de S une théorie sémantiquement close (i.e. qui contient son propre prédicat de vérité, et les équivalences qui le régissent).

Cela faisant, S est devenue bien sûr une théorie inconsistante, et cela montre que nos procédures naïves de preuve, prises dans leur totalité, sont elles aussi inconsistantes (II.6). C'est ainsi, et les antinomies sémantiques sont des “ *facts of life* ”.

4 - Il faut donc accepter l'idée que les antinomies soient des énoncés à la fois vrais et faux.

***

Deux remarques pour commencer.

Premièrement, un point doit d'emblée être précisé, sur lequel Priest n'est pas absolument clair. De la manière dont il présente la preuve de la vérité de g dans son point II.3, il pourrait sembler qu'intervient crucialement la proposition “ tout ce qui est prouvable est vrai ”, pour développer les conséquences de la supposition contrefactuelle : “ supposons que g soit prouvable ”. Et il est exact que Gödel raisonne exactement ainsi dans le sketch de la preuve (la présentation des “ Hauptgedanken des Beweises ”). Mais c'est pour ajouter dans les dernières lignes de ce même passage introductif :

“ La méthode de preuve qui vient d'être exposée peut clairement être appliquée à tout système formel (...) où, deuxièmement, toute formule prouvable est vraie contentuellement (“ inhaltlich richtig ”). Le propos de la preuve exacte qui suit est, entre autres choses, de remplacer cette seconde assomption par une qui soit purement formelle et plus faible. ” (Gödel 1931).

En un mot, aucune assomption concernant le concept de vérité n'est mobilisée dans la preuve purement “ métamathématique ” que g n'est pas prouvable.

Deuxièmement, Priest imagine une sorte d'idéalisation de P, le système formel S, qui ait la triple propriété de formaliser toutes nos procédures informelles de preuve, de mener du vrai reconnu comme tel, au vrai dérivé via des règles correctes, et de conserver les traits cruciaux du système P qui permettent de démontrer son incomplétude. C'est beaucoup demander, et il n'est pas sûr que toutes ces exigences puisent être simultanément satisfaites.

[3] ssi abrège : si et seulement si.

Le théorème de Gödel concerne des théories (des extensions de P) dont l'ensemble des axiomes est décidable. Si l'on prend à la lettre l'idée que nos méthodes de preuve partent de propositions vraies et évidentes, il est plus que douteux que " l'ensemble des évidences " soit décidable, ce qui reviendrait purement et simplement à dire que face à toute proposition, on devrait pouvoir décider si oui ou non elle est évidente (si oui, on la range au nombre des axiomes ; si non, non) ! Priest se débarrasse d'un trait de plume de cette question (II.2). Pourtant, cette considération rend très peu plausible l'idée que le système S soit sujet aux conditions exigées par le théorème de Gödel.

De plus, à moins que nos facultés cognitives soient grandement augmentées, et que soit enrichi le champ de l'évidence, il est invraisemblable que l'exigence que S contienne assez d'arithmétique pour exprimer sa propre syntaxe, soit compatible avec la demande que nos preuves informelles partent de l'évidemment vrai. Par exemple, on ne voit pas du tout comment " des faits de base sur l'addition " permettraient de prouver le principe d'induction complète. On sera donc, ou amené à restreindre l'arithmétique respectable à des procédures élémentaires, ou amené à aller chercher dans des axiomes dont " l'évidence " n'est que très indirecte (inductive, abductive, ou liée à la fécondité) les points de départs ultimes des preuves.

Ces considérations pour souligner ce qu'a de hasardeux la supposition de l'existence d'un tel système S, en tant que formalisation supposée de tout le prouvable.

***

Imaginons que le Malin Génie soit derrière certaines au moins de nos procédures de preuve informelles, et que pour cette raison le système S, qui les formalise *toutes*, soit inconsistant. Même l'exigence de partir de l'évidemment vrai peut nous mener au faux et au contradictoire, et cette possibilité ne peut être évacuée (on en a des exemples dans l'histoire). Dans ce cas, il n'est plus possible de prouver que g n'est pas prouvable, et l'inconsistance serait là avant toute intervention de la sémantique. Mais alors, la mise en scène proposée par Priest s'effondre : l'idée d'une preuve sans feu ni lieu de la vérité de g disparaît. Autrement dit, la nécessité pour laquelle plaide Priest d'intégrer l'inconsistance ne vaut que sous la supposition de la consistance de S. Ce qui m'amène à la question de la circularité de l'argument.

Au bout du compte, il nous faut reconnaître, dit Priest, que notre notion naïve de preuve, sémantique comprise, est inconsistante, ou exige une analyse qui conduit à une théorie inconsistante (II.7). Il faut accepter que le Menteur, par exemple, soit vrai ssi il est faux, soit à la fois vrai et faux (et pas simplement : nous paraisse tel dans une " évidence " trompeuse). Priest pense donc que nous devons conclure qu'il y a des contradictions vraies. Certes elles sont vraies, mais ce sont quand même des contradictions. Il faut donc que la contradiction vienne de quelque part. D'où peut-elle venir ? Des axiomes évidemment vrais, éventuellement sémantiques, ou des règles d'inférence qui leur sont appliquées. Si on regarde du côté de la logique que propose Priest après correction (III.9), il ne semble pas que des règles puissent à elles seules créer des contradictions : celles qui sont seulement " quasi-valides " ne sont incorrectes que dans un environnement déjà paradoxal (il est vrai que cette logique n'épuise pas la totalité des procédures de preuve : *quid* de l'induction arithmétique, ou de l'induction transfinie ?). Comment donc concilier la vérité évidente des axiomes, avec la vérité des contradictions qu'on peut en dériver ? Je ne vois pas d'autre solution que d'admettre tacitement que les axiomes, tout en étant vrais, peuvent également, du moins pour certains d'entre eux, être contradictoires. Bref, la thèse des contradictions vraies n'est compatible avec la description de nos procédures de preuve que sous le présupposé que le contradictoire peut être vrai, et l'argumentation repose alors sur ce qu'elle prétend démontrer.

Inutile de dire que si S est inconsistant, on est ramené à l'hypothèse du Malin Génie. Je crains donc que l'argument exige à la fois l'hypothèse de la consistance de S, et présuppose la thèse que la contradiction peut être vraie.

***

Acceptons la construction de Priest. Intégrons la sémantique, avec ses inconsistances, dans notre système total de preuves. Nous raisonnons, de fait, de manière à tomber sur des antinomies. Que faut-il conclure de cet état de choses ?

Qu'il faut accepter certaines contradictions ? Ou plutôt que dans notre système de preuves, il y a des " vérités évidentes " qui n'étaient des vérités qu'en apparence, et qu'il faut donc réformer notre entendement ? Ordinairement, on a choisi cette seconde voie. Priest nous invite à préférer la première. Pourquoi ce choix révolutionnaire, - abstraction faite du charme des révolutions ? Parce que c'est ainsi que, naturellement, informellement, avant tout contrôle critique, nous raisonnons. Est-ce une bonne raison ? Cela paraît en réalité terriblement conservateur. Et ce n'est évidemment pas un argument rationnel. En fait, Priest ne donne aucune raison pour conserver notre notion naïve de preuve, hormis l'appel aux " *facts of life* ". A ce compte, la guerre aussi est un fait de la vie (Héraclite).

La logique tri-valuée (à la fois vrai et faux joue le rôle d'une troisième valeur de vérité) qu'introduit Priest en III n'est bien sûr pas le seul moyen d'obtenir qu'une contradiction puisse être vraie. On peut aussi modifier les conditions d'évaluation de la négation, en introduisant des points (mondes possibles) tels qu'une négation non-A est vraie en un point ssi A est faux en un autre point lié au premier, tout en n'opérant qu'avec deux valeurs de vérité : A et non-A peuvent alors être simultanément vrais en un point, sans être à la fois vrais et faux.[4]

Pour des raisons conceptuelles, on peut aussi refuser qu'une contradiction implique n'importe quoi, sans avoir à admettre qu'une contradiction puisse être vraie. Il suffit pour, ce faire, de nier que l'implication logique consiste simplement en l'impossibilité de la combinaison : prémisses vraies et conclusion fausse. Une relation de contenu est nécessaire (implication pertinente).

En bref, il n'y a aucun moyen de passer de l'idée que le système de nos preuves est inconsistant à la thèse que le Menteur est à la fois vrai et faux : *non sequitur*.

***

Naturellement, en montrant que l'argumentation de Priest est fautive, je n'ai pas positivement montré que sa thèse est fausse, ni que le Principe de non-contradiction est vrai. Mais je ne peux ici faire mieux que renvoyer à Aristote :

" Il est absolument impossible de tout démontrer (...). Il est cependant possible d'établir par réfutation l'impossibilité que la même chose soit et ne soit pas, pourvu que l'adversaire dise seulement quelque chose. " (*Métaphysique*, Γ, 4).

[4] Voir par exemple la sémantique relationnelle de Routley-Meyer, dans Rivenc 2005.

Références.

Aristote,
Métaphysique, Tome I, trad. fr. par J. Tricot, Vrin, 1953.

Chihara, Charles S
1984 – 'Priest, the Liar, and Gödel', *Journal of Philosophical Logic* **13** p117-124.

Gödel, Kurt
1931 – "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I ", in *Kurt Gödel, Collected Works*, Vol. I, Oxford University Press, 1986.

Priest, Graham
1979 – "The Logic of Paradox", *Journal of Philosophical Logic,* **8**.
1987 – *In contradiction : a study of the transconsistent.*, Martinus Nijhoff.

Rivenc, François
2005 – *Introduction à la logique pertinente,* P.U.F.

LA LOGIQUE DU PARADOXE

"Oui, je le prédis dès maintenant : on en viendra à des recherches mathématiques sur des calculs contenant une contradiction et au fond on tirera vanité de s'être émancipé de la non-contradiction."

Wittgenstein[1930, p318][5].

0. Introduction

Le but du présent article est de suggérer une nouvelle façon de traiter les paradoxes logiques. Au lieu d'essayer de les dissoudre, ou d'expliquer ce qui allait mal, nous devrions les accepter et finir par apprendre à vivre avec. Les Sections I et II développent ce point. Pour des raisons évidentes, cela requerra l'abandon, ou au moins la modification de la logique 'classique'. Une façon de faire cela est suggérée dans la Section III. Les Sections IV et V discutent quelques implications de cette approche des paradoxes.

I. Paradoxes

I.1. Les paradoxes logiques (qui sont normalement subdivisés en paradoxes relatifs à la théorie des ensembles comme celui de Russell et en sémantiques comme celui du menteur) sont à présent en circulation depuis longtemps. Pourtant aucune solution n'a été trouvée. Il est vrai que le paradoxe du menteur, qui est connu depuis plus de 2000 ans, a souvent été ignoré comme étant une trivialité indigne de considération sérieuse. (Quoique des logiciens médiévaux comme Buridan ne l'ont nullement considéré ainsi.) Cependant, durant les trois derniers quarts d'un siècle, il y a eu probablement plus de travail intensif fait pour essayer de trouver une solution aux paradoxes logiques que pour n'importe quel autre sujet dans l'histoire de la logique. L'appel de ceux qui ont essayé de trouver une solution fait penser à une liste d'honneur de logiciens. Jusqu'ici, aucune solution largement acceptée n'a été trouvée.

I.2. Bien sûr, nous savons comment éviter les paradoxes formellement. Nous pouvons éviter les paradoxes sémantiques par ex., par une hiérarchie de métalangages à la Tarski, et ceux de la théorie des ensembles, par ex. par la distinction classe/ensemble de von Neumann. Mais ce ne sont pas des solutions. Un paradoxe est un argument avec des prémisses qui paraissent vraies et des étapes qui paraissent valides, qui néanmoins se termine par une conclusion qui est fausse. Une solution nous dirait quelle prémisse est fausse ou quelle étape est invalide ; mais en plus elle nous donnerait une *raison indépendante* de penser que la prémisse ou l'étape est erronée. Si nous n'avons pas d'autre raison de rejeter la prémisse ou l'étape que le fait que cela bloque les conclusions, alors la 'solution' est *ad hoc* et n'éclaire rien. Virtuellement toutes les 'solutions' connues aux paradoxes échouent à ce test et c'est pourquoi je dis que jusqu'ici aucune solution n'a été trouvée.

[5]Le texte se situe dans la deuxième appendice de l'édition française, qui constitue un extrait des notes sténographiques prises par F. Waismann sur les propos et conversations de Wittgenstein entre décembre 1929 et septembre 1931. Voir Wittgenstein [1964, p332] pour le texte que cite Priest. [NDT].

I.3. Cet échec massif de la part de la communauté logique suggère qu'essayer de résoudre les paradoxes n'est peut-être pas la bonne chose à faire. Supposons que nous arrêtions de nous prendre la tête en vain à essayer de trouver une solution, et que nous acceptions les paradoxes comme des faits bruts. C'est-à-dire, certains énoncés sont vrais (et seulement vrais), certains faux (et seulement faux), et certains vrais et faux à la fois ! ('Cet énoncé est faux' et 'L'ensemble de Russell s'appartient à lui-même' sont des exemples paradigmatiques de tels énoncés paradoxaux). Bien sûr, cela requiert d'abandonner la maxime d'Aristote 'Le plus ferme de tous les principes est qu'il est impossible pour la même chose d'appartenir et de ne pas appartenir à la même chose au même moment et sous le même rapport' (*Metaphysique* Γ, 3(1005^b 19-23)). Toujours est-il que tout progrès suppose qu'on se sépare de la tradition d'une manière ou d'une autre.

II. La clôture sémantique

II.1. Si mon argument pour accepter les paradoxes était simplement que personne ne les a encore résolus, alors ma position, bien que plausible, ne serait pas très tentante. Cependant, je crois qu'il y a des raisons théoriques pour lesquelles les paradoxes doivent être acceptés ; autrement dit, si nous en venons aux prises avec la notion de prouvabilité mathématique (dans le sens naïf).

II.2. Une preuve, au sens des mathématiciens (et non des logiciens) est ce processus par lequel nous établissons que certaines affirmations mathématiques sont vraies. En d'autres termes, supposons que nous ayons une affirmation (pour être précis, disons une affirmation arithmétique) dont nous voulions établir la vérité. Nous en recherchons une preuve ou une réfutation (i.e., une preuve de sa négation). Mais une preuve à partir de quoi ? Vraisemblablement à partir d'autres affirmations dont nous savons qu'elles sont vraies. Nous pouvons, bien sûr, demander comment nous avons su que ces affirmations étaient vraies, et c'est peut-être parce que nous en avons des preuves. Cependant, sous peine d'une régression infinie, on ne peut pas continuer à faire cela indéfiniment. Tôt ou tard, nous devons arriver à des affirmations dont nous savons qu'elles sont vraies sans preuve, où la question de la preuve, pour ainsi dire, ne se pose plus. Ce sont les axiomes dans le sens traditionnel : des vérités évidentes. (Pourquoi sont-elles évidentes, je n'ai pas besoin d'entrer dans ce débat. Il suffit qu'il doive y avoir de telles choses). Dans le cas présent, ce sont vraisemblablement des faits concernant les nombres, comme le fait que tout entier a un successeur différent de tout ce qui précède, ou des faits élémentaires sur l'addition. Ce sont les sortes de choses qui deviennent familières aux enfants quand ils apprennent à compter et à faire de l'arithmétique. (On peut, bien sûr, attendre des 'preuves' de ces axiomes dans un système fondationnel comme les *Principia Mathematica*. Cependant, ces 'preuves' ne sont pas des preuves dans le sens qui nous concerne −des moyens d'en venir à savoir que les choses prouvées sont vraies.)

Ainsi nous voyons que nous établissons les affirmations mathématiques, si elles ne sont pas des axiomes, en les prouvant (dans le sens naïf) à partir de ces axiomes. Tout cela semble évident au point d'être banal. Cependant, cela nous conduit tout droit à un problème majeur. Car il semble qu'il n'y ait aucun doute que cette procédure puisse être formalisée. Les axiomes pourraient, en principe, être écrits dans un langage formel et les preuves présentées comme des preuves formelles. Le système formel qui en résulterait encoderait nos méthodes naïves de preuves. De plus, il semble qu'il n'y ait pas de raison de douter que toutes les fonctions récursives soient représentables dans le système. Car certainement toutes les fonctions récursives sont définissables au sens naïf. Cependant, selon le théorème d'incomplétude de Gödel, dans n'importe quel système formel de ce genre, il y aura des énoncés qui ne sont ni prouvables, ni réfutables —du moins si cet ensemble d'axiomes est décidable. Si c'était tout ce qu'il y avait à dire sur le théorème de Gödel, le résultat serait peut-être surprenant, mais pas particulièrement inquiétant. L'incomplétude du système formel montrerait simplement qu'il y a des problèmes mathématiques à résoudre au-delà des capacités de nos procédures de preuves. Mais ce n'est pas tout. Car on peut *montrer* la vérité de ces énoncés non prouvables, i.e. on peut les prouver au sens naïf. Mais le système formel était construit de telle façon qu'il encodait nos méthodes de preuves informelles. Donc il ne peut y avoir de telle preuve. Le théorème de Gödel présente un problème épistémologique qui n'a jamais été directement affronté. Quelle peut en être la solution ?

Une isssue au problème est d'accepter que l'ensemble des axiomes ne soit pas récursif. Ainsi, si l'on accepte la thèse de Church, il n'y aurait pas de moyen effectif de décider si quelque chose est évident ! C'est évidement inadmissible. Une autre possibilité est que nos procédures naïves de preuves ne soient pas formalisables. Cela semble invraisemblable, bien qu'on puisse faire valoir le point ainsi. Gödel [1947] a suggéré que nous pouvons de temps en temps augmenter notre stock d'axiomes en en ajoutant un qui a un support inductif (au sens où il implique beaucoup de choses que l'on sait vraies et rien qui soit connu comme faux). Dans ce cadre nous pouvons supposer que nos méthodes naïves de preuves à un certain moment, disons t, peuvent être formalisées par un système S. Cependant, il peut arriver à un moment postérieur que, pour encoder nos procédures naïves de preuves, nous ayons besoin d'un autre système S' formé à partir de S en ajoutant à ses axiomes un énoncé A à partir duquel on peut prouver (dans S) un certain nombre de choses vraies et rien de faux. Quelle que soit la plausibilité de cet exposé, cela ne résoud pas le problème. Car considérons un des énoncés qui est indépendant de S, mais informellement prouvable. Alors il est déjà prouvable à t. Nous n'avons pas à rassembler de manière inductive des éléments en sa faveur, pour voir s'il devrait être ajouté aux axiomes de S. Il peut être établi déductivement par des moyens déjà disponibles. Par conséquent le problème apparaît maintenant relativement à nos méthodes de preuves naïves au temps t.

II.3. Nous n'avons donc pas trouvé de solution au problème. Cependant la suggestion de Gödel soulève la question, de savoir comment exactement se fait-il que nous soyons capables de prouver ces énoncés indépendants. C'est cela que nous allons maintenant examiner. Nous verrons que cela fournira une réponse au problème.

Supposons que P soit le système formel qui encode notre capacité naïve de preuve (peut-être à un certain moment) et soit g (le nombre-code de) l'énoncé de Gödel :

$$\neg\exists x \; \mathrm{Prov}(xg)$$

où $\mathrm{Prov}(xy)$ est le prédicat primitif récursif 'x est le d'une preuve (dans P) de l'énoncé y'. Maintenant, g n'est pas prouvable dans P, et cela peut être prouvé purement syntaxiquement. Cependant, la question n'est pas comment on montre que g est indépendant, mais comment on montre que g est vrai. Car on *peut* montrer que g est vrai. D'habitude, on dit (d'un geste) que g 'exprime sa propre nonprouvabilité' et est donc vrai. Quand cette indication plutôt vague est articulée précisément, on obtient l'argument suivant.

(1) $\exists x$ Prov$(xg) \Longrightarrow \ulcorner \exists x$ Prov$(xg) \urcorner$ est vraie.
(2) $\Longrightarrow g$ est prouvable
(3) $\Longrightarrow g$ est vraie
(4) $\Longrightarrow \neg\exists x$ Prov(xg)
Donc $\neg\exists x$ Prov(xg)

L'étape (2) dépend du fait que 'Prov' représente vraiment la relation de prouvabilité. L'étape (3) dépend du fait que tout ce qui est prouvable dans P est vrai. Les étapes (1) et (4) résultent du biconditionnel de Tarski $T \ulcorner\psi\urcorner \leftrightarrow \psi$ (où $\ulcorner\psi\urcorner$ est le code de ψ).

Or, le point important à noter à propos de ce raisonnement est qu'il exige un *détour* essentiel par le métalangage de P. L'utilisation de la notion de vérité, ses propriétés, et sa relation à la prouvabilité sont essentiels à la preuve. On ne peut pas prouver g dans P, mais on peut le prouver dans le métalangage de P.

II.4. Ainsi, nous voyons que notre notion naïve de preuve semble dépasser la notion axiomatique de preuve précisément parce qu'elle peut opérer avec des notions sémantiques. Bien sûr, on peut formaliser la sémantique axiomatiquement mais alors on peut naïvement raisonner à propos de la sémantique d'un *tel* système. Tant qu'une théorie ne peut pas formuler sa propre sémantique elle sera incomplète au sens de Gödel, i.e., il y aura des énoncés indépendants de la théorie dont nous pouvons établir la vérité par un raisonnement sémantique naïf.

II.5. Que se passe-t-il, cependant, si nous prenons une théorie sémantiquement close ? (Une théorie est sémantiquement close si elle peut formuler sa propre sémantique.) Si elle est sémantiquement close, il n'y a rien qui puisse nous empêcher d'utiliser le raisonnement de la Section II.3 dans le système lui-même. Par exemple, considérons la théorie PS suivante. Le langage de PS est celui de l'arithmétique de Peano avec en plus le prédicat à deux places 'Sat(x, y)'. (Intuitivement, cela sera lu 'la suite (finie) x satisfait la formule avec le nombre-code y'.) Nous pouvons coder les notions de formule, de $i^{ème}$ variable, de suite finie de nombres, de $i^{ème}$ membre d'une suite, de longueur d'une suite, de la façon ordinaire. Les axiomes de PS sont alors ceux de l'arithmétique de Peano plus toutes les instances du schéma de satisfaction de Tarski :

$$(\forall x)(x \text{ est une suite finie suffisamment longue} \rightarrow (\mathrm{Sat}(x \ulcorner\psi\urcorner) \leftrightarrow \psi'))$$

où ψ' est ψ avec toutes les variables libres 'v_i' dans ψ remplacées par x_i (le $i^{ème}$ membre de la suite x) et 'x est suffisamment longue' est '$(\forall i)$ si la $i^{ème}$ variable apparaît libre dans ψ alors la longueur de x est plus grande que i'. (Une autre manière de faire est de remplacer ce schéma par un nombre fini d'axiomes : un pour chaque prédicat atomique de PS $-$y compris 'Sat'$-$ et un pour chaque fonction de vérité et quantificateur. Pour les détails de ces sortes de constructions voir Hilbert et Bernays [1939, p334-335]. Toutes les instances du schéma de satisfaction seraient alors prouvables.) Maintenant PS est clairement sémantiquement close et nous pouvons exécuter la sorte de raisonnement que nous avons utilisée dans la Section II.3 pour prouver $\neg\exists x$ Prov(xg) *dans PS elle-même* (où 'Prov' est maintenant le prédicat de prouvabilité de PS).

II.6. Nous sommes maintenant en position de résoudre le problème posé dans la Section II.2. Nous avons vu que le problème se pose parce que nous pouvons établir la vérité d'énoncés qui ne sont pas prouvables dans quelque système particulier. Nous voyons maintenant que le problème est évité en utilisant une théorie formelle sémantiquement close. Car le raisonnement qui est utilisé pour montrer que cette sorte d'énoncé est vrai peut être représenté *à l'intérieur* d'une théorie sémantiquement close. Bien sûr, cela montre aussi que la preuve de Gödel de l'existence d'énoncés indépendants échoue. Car le raisonnement donne une preuve de l'énoncé 'indépendant' dans la théorie. Qu'il y ait ou non des énoncés indépendants est un sujet auquel je reviendrai dans la Section IV.13.)

Nous pouvons bien nous demander où exactement il y a une erreur dans la preuve de Gödel. L'endroit n'est pas difficile à localiser. Car, bien sûr, sa preuve ne fonctionne que si la théorie en considération est *consistante.* Il est bien connu que les théories closes sémantiquement sont inconsistantes. (Voir, p.ex., Fraenkel Bar-Hillel et Levy [1973, p312]). C'est de cette façon que le résultat de Gödel est évité.

Nous avons vu que la manière d'éviter le problème −en fait, la seule manière raisonnable[6]− est d'accepter que la formalisation correcte de nos méthodes de preuves naïves soit une théorie sémantiquement close et inconsistante. En fait, n'importe quelle analyse formelle de nos méthodes naïves de preuve doit utiliser une théorie sémantiquement close et inconsistante ou être inadéquate.

A ce point on peut objecter qu'il est impossible qu'une théorie inconsistante puisse codifier nos procédures naïves de preuves. Car dans une théorie inconsistante tout est prouvable. Cependant, cette objection présuppose que la logique sous-jacente soit 'classique'. Cela sera rejeté dans la Section III.1. Le fait qu'il y ait des preuves 'classiques' pathologiques (en utilisant, p. ex., la règle $(A\wedge\neg A)/B$) de n'importe quoi dans une théorie inconsistante montre justement que la logique des preuves naïves n'est pas classique.

II.7. Cependant, le fait qu'une théorie sémantiquement close est inconsistante est important. Car, en fait, les inconsistances engendrées par la clôture sémantique sont précisément les antinomies sémantiques. Depuis Tarski [1936] il est clair que la clôture sémantique engendre les paradoxes sémantiques. C'est donc là l'argument théorique auquel je faisais référence dans la Section II.1 pour en arriver à accepter certaines contradictions. Car n'importe quelle analyse acceptable de notre notion naïve de preuve requiert l'utilisation d'une théorie sémantiquement close et par conséquent inconsistante. Et donc n'importe quelle analyse adéquate de la notion naïve de preuve nous forcera à accepter les antinomies comme des faits de la vie.

II.8. Quoique bien des choses que j'ai dites dans cette Section soient techniques, l'esprit en est facile à saisir. Le logicien formel est essentiellement un mathématicien appliqué. C'est son métier de construire des systèmes mathématiques qui modélisent (dans le sens du physicien, pas du logicien) quelque phénomène naturel. Le phénomène qui intéresse particulièrement le logicien, est le raisonnement normal (naïf) mis en œuvre dans le langage naturel. (Après tout, les mathématiciens utilisent des langages ordinaires (non formels) −même s'ils sont enrichis par certains termes

[6]C'est cette affirmation qui sera précisément critiquée par Chihara [1984]. Voir la présentation [NDT].

techniques− et des raisonnements ordinaires). Et les systèmes mathématiques que le logicien utilise sont des langages formels, des sémantiques mathématiques, etc...

Or c'est une opinion standard dûe à Tarski que les langages naturels à la fois sont sémantiquement clos et contiennent des paradoxes. Comme il le dit :

> Si nous voulons maintenir l'universalité du langage de tous les jours en liaison avec les recherches sémantiques, nous devons, pour être consistant [*sic* !] admettre dans le langage, en plus de ses énoncés et autres expressions, aussi les noms de ces énoncés et expressions, et des énoncés contenant ces noms, aussi bien que des expressions sémantiques comme "énoncé vrai", "nomme", "dénote", etc... Mais c'est probablement exactement cette universalité du langage de tous les jours qui est la source fondamentale de toutes les antinomies sémantiques, comme les antinomies du menteur et des mots hétérologiques.
>
> Tarski [1944].

C'est un point de vue que je fais mien. (En fait, j'ai soutenu ailleurs, Priest [1974,Chapitre 5, Section 1], que toutes les antinomies logiques sont dûes à la clôture sémantique. Mais c'est une autre histoire.) Par conséquent, bien qu'une théorie formelle non sémantiquement close puisse être un modèle adéquat pour certains buts limités, une théorie formelle sémantiquement close −avec des paradoxes− est requise en général.

III. La logique du paradoxe (LP)

III.1. Dans les deux Sections précédentes j'ai soutenu que nous devrons apprendre à traiter des systèmes contenant des contradictions. Si nous avons à faire cela, nous devrons renoncer à la logique 'classique'. Car les contradictions ont une horrible façon d'infecter les théories classiquement formalisées. La population entière des énoncés mathématiques se trouve frappée par la contradiction, ce qui les rend inaptes au travail. Cependant, si nous pouvons isoler les paradoxes et les empêcher de contaminer tout le reste, tout ira bien. Dans cette Section, je vais construire une telle logique. Cependant, j'aimerais insister sur le fait que les arguments des deux Sections précédentes sont indépendants du destin de la logique construite. Même si elle s'avère insatisfaisante pour quelque raison, les considérations précédentes tiennent toujours bon.

III.2. En fait, des systèmes qui ne sont pas ruinés par les contradictions sont déjà apparus dans la littérature. Routley [1977] et da Costa [1974] ont tous deux soutenu que de telles logiques nous permettraient d'étudier des théories inconsistantes, mais non triviales. Cependant, aucun des systèmes proposés n'a la sémantique simple et intuitivement plausible de LP. Car outre qu'elles motivent la construction d'une telle logique, les Sections précédentes suggèrent en fait un moyen de le faire. Ceci assure que la logique a une sémantique très intuitive.

III.3. La logique classique erre en assumant qu'aucun énoncé ne peut être à la fois vrai et faux. Nous voulons corriger cette assomption. Si un énoncé est à la fois vrai et faux, appelons le 'paradoxal' (p). S'il est vrai et pas faux, nous l'appellerons 'seulement vrai ' (t)[7] et de manière analogue pour 'seulement faux ' (f). Mais cette assomption faite, nous continuerons à raisonner normalement.

III.4. Un énoncé est vrai ssi[8] sa négation est fausse. D'où, la négation d'un énoncé vrai et faux est fausse et vraie, i.e., paradoxale. La négation d'un énoncé seulement vrai est seulement fausse. (Si sa négation était vraie, il devrait être faux). De même, la négation d'un énoncé seulement faux est seulement vraie. Nous pouvons décrire ces résultats dans la table suivante.

$\neg$	
t	f
p	p
f	t

III.5. Un raisonnement semblable donne la table suivante pour la conjonction :

$\wedge$	t	p	f
t	t	p	f
p	p	p	f
f	f	f	f

Je donnerai seulement deux exemples. Si A est t et B est p, alors A et B sont vrais l'un et l'autre. Donc $A \wedge B$ est vrai. Cependant, comme B est faux, $A \wedge B$ est faux. Par conséquent, $A \wedge B$ est paradoxal.

Si A est f et B est p, alors A et B sont faux l'un et l'autre. Donc $A \wedge B$ est faux. Si $A \wedge B$ était aussi vrai, alors A et B seraient vrais, mais A est seulement faux. Par conséquent, $A \wedge B$ est f.

III.6. En raisonnant d'une manière semblable, on peut justifier la table de la disjonction.

$\vee$	t	p	f
t	t	t	t
p	t	p	p
f	t	p	f

Autrement, '$A \vee B$' peut être défini par '$\neg(\neg A \wedge \neg B)$'.

On peut définir '$A \rightarrow B$' par '$\neg A \vee B$' et '$A \leftrightarrow B$' par '$A \rightarrow B \wedge B \rightarrow A$'. Ce qui donne les tables suivantes :

$\rightarrow$	t	p	f
t	t	p	f
p	t	p	p
f	t	t	t

$\leftrightarrow$	t	p	f
t	t	p	f
p	p	p	p
f	f	p	t

[7] t comme "true" [NDT].

[8] "ssi" signifie "si et seulement si". [NDT].

Dans quelle mesure '$\rightarrow$' peut être considéré comme une forme d'implication, nous y reviendrons à la Section IV.4.

III.7.[9] En fait, ce sont les matrices de Kleene [1952, p332 et suiv.]. Cependant, son interprétation est très différente de la nôtre. Les matrices pour '$\vee$','$\wedge$', '$\neg$', sont aussi celles de Lukasiewicz [1920]. Cependant, $'t'$ est la seule valeur désignée dans son système. Nous prendrons t et p comme valeurs désignées, puisqu'elles sont toutes les deux les valeurs des énoncés vrais. Formellement, soit L un langage propositionnel dont l'ensemble des variables propositionnelles est P. Soit $v : P \rightarrow \{t, p, f\}$ (i.e., v est une évaluation des variables propositionnelles). Soit v^+ l'extention naturelle de v à tous les énoncés de L utilisant les tables de vérité ci-dessus. Si Σ est un ensemble d'énoncés de L, on définit :

$\Sigma \models A$ ssi il n'existe pas de v tel que $v^+(A) = f$
mais pour tout $B \in \Sigma, v^+(B) = t$ ou p.

$\models A$ ssi $\varnothing \models A$ (i.e., pour tout $v, v^+(A) = t$ ou p).

III.8. THÉORÈME : A est une tautologie bivaluée ssi $\models A$.

Preuve. La preuve de droite à gauche est immédiate puisque toute évaluation bivaluée est une évaluation trivaluée.

Pour la converse, si v est une évaluation trivaluée, soit v_1 l'évaluation bivaluée formée en changeant tous les p par des t. En vérifiant les tables de vérité, on peut voir que si $v^+(A) = f, v_1^+(A) = f$. D'où le résultat.

Les matrices de LP avec une interprétation semblable à celle que j'ai donnée se trouvent chez Asenjo [1966]. Quoiqu'il ne le formule pas explicitement, il semble aussi prendre t et p (ses 0 et 2) comme valeurs désignées. Cependant, il ne prend pas sa sémantique très au sérieux. Car immédiatement après avoir donné la sémantique, il affirme que $\neg(A \wedge \neg A)$ (ce qu'il appelle la loi de contradiction) ne doit pas être un théorème d'un système de logique approprié à la formalisation des théories inconsistantes. Cependant, par ce théorème on peut voir que selon ces sémantiques, $\neg(A \wedge \neg A)$ est logiquement valide.

III.9. Nous avons vu que la propriété d'être une tautologie est préservée dans LP. Cependant, la relation de déductibilité est changée. On peut vérifier :

$A \models A \vee B$	$A, B \models A \wedge B$	$A \rightarrow B \models \neg B \rightarrow \neg A$
$A \rightarrow (B \rightarrow C) \models B \rightarrow (A \rightarrow C)$	$A \models B \rightarrow A$	$\neg A, \neg B \models \neg(A \vee B)$
$\neg A \rightarrow \neg B \models B \rightarrow A$	$\neg(A \vee B) \models \neg A$	$A \models \neg\neg A$
$\neg\neg A \models A$	$\neg A \models \neg(A \wedge B)$	$\neg(A \rightarrow B) \models A$
$A \wedge B \models A$	$A, \neg B \models \neg(A \rightarrow B)$	$A \rightarrow B \models A \wedge C \rightarrow B \wedge C$
$\neg A \models A \rightarrow B$	$A \rightarrow (A \rightarrow B) \models A \rightarrow B$	$A \rightarrow \neg A \models \neg A$

Cependant, il est facile de trouver des évaluations qui montrent que les relations suivantes n'ont pas lieu.

[9] Priest [1984, note 23, p174] reviendra sur ce paragraphe : "Dans *LP* §III.7 j'ai défini la validité en termes de préservation de la vérité. Je pense maintenant que c'est faux, juste parce qu'une telle validité ne supporte pas le détachement [...]." [NDT].

$$\begin{array}{ll} A \wedge \neg A \models B & A \to B, B \to C \models A \to C \\ A, \neg A \vee B \models B & A, A \to B \models B \\ A \to B \wedge \neg B \models \neg A & A \to B, \neg B \models \neg A \end{array}$$

III.10. THÉORÈME : Si A et B n'ont pas de variables propositionnelles en commun et si B peut prendre la valeur f alors $A \not\models B$.

Preuve. Soit v une évaluation telle que $v^+(B) = f$.

Soit v_1 identique à v sauf que pour toutes les variables propositionnelles q qui figurent dans A, $v_1(q) = p$. On vérifie facilement que $v_1^+(A) = p$ et $v_1^+(B) = f$. D'où le résultat.

III.11. THÉORÈME : Si $A_1...A_n \models B$, alors $A_1...A_{n-1} \models A_n \to B$.

Preuve. Triviale.

III.12. Il est facile d'étendre LP en une logique quantifiée LPQ. Soit L un langage du premier ordre. Pour simplifier, nous supposons qu'il n'a ni constantes, ni symboles de fonctions.

Soit **A**=$<DI>$.

D est un domaine d'objets. Si P_n est un prédicat de L à n-places, et $\bar{x}$ un n-uplet membre de D, I applique $<P_n\bar{x}>$ dans $\{t, p, f\}$.

Soit S une évaluation des variables de L (i.e., une application des variables dans D). Nous définissons la 'valeur de vérité' d'un énoncé sous S comme suit :

Si A est de la forme $P_n(v_i...v_j)$, A est t, p, ou f, selon que $I < P_n < Sv_i...Sv_j >>$ est t, p, ou f.

Si A est de la forme $\neg B$ ou $B \wedge C$, les conditions de vérité de A sont données par les matrices des Sections III 6-7.

Si A est de la forme $(\forall v)B$, alors

A est t ssi pour tout $d \in D$, B est t sous $S(v/d)$
(où $S(v/d)$ est $S- < v\ S(v) > \cup < v\ d >$).

A est f ssi pour quelque $d \in D$, B est f sous $S(v/d)$.

A est p sinon (i.e., pour tout $d \in D$, B est t ou p sous $S(v/d)$, et pour quelque $d \in D$, B est p sous $S(v/d)$).

Cette dernière définition peut être justifiée par une extension évidente des arguments utilisés pour justifier les tables de vérité des Sections III 4-6. Je vais donner un exemple. Si B est t ou p pour tout $S(v/d)$ alors B est vraie pour tout $S(v/d)$. Par conséquent, $(\forall x)B$ est vraie. Si B est p pour quelque $S(v/d)$, alors elle est f pour quelque $S(v/d)$. Ainsi, $(\forall x)B$ est fausse. Par conséquent, A est paradoxale.

Les 'conditions de vérité' de $(\exists x)B$ peuvent être données comme suit. (Autrement, '$(\exists x)B$' peut être défini par '$\neg(\forall x)\neg B$').

$(\exists x)B$ est t ssi pour quelque $d \in D$, B est t sous $S(v/d)$

$(\exists x)B$ est f ssi pour tout $d \in D$, B est f sous $S(v/d)$

$(\exists x)B$ est p sinon (i.e., pour tout $d \in D$, B est f ou p sous $S(v/d)$ mais pour quelque $d \in D$, B est p sous $S(v/d)$.

Comme d'habitude on peut prouver que si A est un énoncé clos, alors la valeur de vérité de A est indépendante des S choisis.

Si Σ est un ensemble d'énoncés de L, on définit :

$\Sigma \models A$ ssi il n'y a pas de $\mathbf{A}$ et pas de S tels que A est f sous S et pour tout $B \in \Sigma$, B est t ou p sous S.

$\models A$ ssi $\varnothing \models A$ (i.e., pour tout $\mathbf{A}$ et S, A est t ou p sous S).

III.13. THÉORÈME : A est une vérité logique bivaluée ssi $\models A$.

Preuve. La preuve de droite à gauche est immédiate puisque tout modèle bivalué est un modèle trivalué.

Inversement, si $\mathbf{A}$ et S sont respectivement, un modèle trivalué et une évaluation, soit $\mathbf{A}_\perp$ le modèle bivalué obtenu en changeant tous les p de I en t. On vérifie facilement que si A est f sous S, dans $\mathbf{A}$, A est f sous S dans $\mathbf{A}_\perp$. Le résultat s'ensuit.

III.14. Comme en III.9 la relation de déductibilité est changée dans LPQ. Car bien que nous ayons

$$(\forall x)A, (\forall x)B \models (\forall x)(A \wedge B)$$
$$(\forall x)A \models (\forall x)(A \vee B)$$
$$(\forall x)A \models A(x/y)$$
$$(\forall x)(A \to B) \models (\forall x)(\neg B \to \neg A)$$
$$(\forall x)(A \to B) \models (\exists x)A \to (\exists x)B$$
$$(\forall x)A \models \neg(\exists x)\neg A$$
$$(\forall x)(A \to B) \models (\forall x)A \to (\forall x)B$$
$$(\forall x)A \models (\forall x)(B \to A)$$
$$(\forall x)A \models (\forall x)(\neg A \to B)$$

il est facile de trouver des contre-exemples qui montrent que nous n'avons pas :

$$(\forall x)A, (\forall x)(A \to B) \models (\forall x)B$$
$$(\forall x)(A \to B), (\forall x)\neg B \models (\forall x)\neg A$$
$$(\forall x)(A \to B), (\forall x)(B \to C) \models (\forall x)(A \to C)$$

III.15. THÉORÈME (substituabilité des équivalents) : Si A, A^1 ont des variables parmi $y_1 ... y_n$ et B^1 est comme B sauf qu'il contient A^1 là où B contient A, alors

$$(\forall y_1 ... y_n)(A \leftrightarrow A^1) \models B \leftrightarrow B^1.$$

Preuve. La preuve est par induction sur la formation de B.

Si A est B alors le résultat s'ensuit puisque $(\forall y)C \models C$.

Si une fonction de vérité est utilisée dans la construction, alors les règles suivantes sont utilisées pour établir le résultat

$$A \leftrightarrow B \models \neg A \leftrightarrow \neg B$$
$$A \leftrightarrow B \models A \wedge C \leftrightarrow B \wedge C.$$

Si un quantificateur est utilisé dans la construction, alors le résultat s'ensuit à partir des deux règles

Si $A \models B$ et x n'apparaît pas libre dans A
alors $A \models (\forall x)B$.

et

$(\forall x)(A \leftrightarrow B) \models (\forall x)A \leftrightarrow (\forall x)B$.

III.16. Il n'est pas difficile de donner un système de règles ou d'axiomes pour LP et LPQ. Cependant je laisserai ces problèmes techniques pour un autre article.

LP est aussi dans une relation intéressante à la sémantique pour la conséquence du premier degré[10] présentées par Dunn[1976] et les Routley[1972]. Là encore, j'examinerai ce problème dans un autre article.

IV. Conséquences

IV.1. La chose la plus évidente à propos de la logique du paradoxe est qu'elle nous force à abandonner, comme invalides, certains principes de déduction que l'on ne suspecterait pas normalement. Par exemple les règles

$$\frac{A \quad \neg A \vee B}{B} \qquad \frac{A \rightarrow \neg B \quad B}{\neg A} \qquad \frac{\forall x A \quad \forall x(A \rightarrow B)}{\forall x B}$$

préservent la validité mais pas la vérité. Nous pouvons un peu amortir le choc pour nos intuitions en faisant remarquer que bien que ces inférences ne soient généralement pas valides, elles sont valides à condition que toutes les valeurs de vérités concernées soient classiques (i.e., vraies seulement ou fausses seulement). Appelons une règle qui préserve la vérité seulement sous de telles conditions 'quasi-valide'. Alors les inférences quasi-valides sont parfaitement O.K., dans la mesure où nous évitons des formules louches comme 'Cette phrase est fausse' et 'L'ensemble de Russell est membre de lui-même'. Il est clair que toutes les règles d'inférences classiquement valides sont soit valides soit quasi-valides.

IV.2. Cependant, les règles quasi-valides sont généralement invalides. On pourrait donc raisonnablement dire qu'elles ne doivent pas être utilisées. Mais ce serait alors l'échec de notre projet. Car la finalité de l'opération était de construire une logique qui puisse être utilisée (en connexion avec une théorie sémantiquement close) pour capturer le raisonnement mathématique naïf. Cependant, en évitant les règles quasi-valides cela aurait à l'évidence un effet handicapant sur le raisonnement mathématique. Comment peut-on raisonner sans *modus ponens* ? Il y a deux moyens se tirer de ce problème. Je vais les considérer chacun leur tour.

IV.3. Restons d'abord dans le fragment de langage qui ne contient que '$\neg$', '$\wedge$', '$\vee$'. La principale perte est ici le syllogisme disjonctif. Cependant, nous n'avons pas ici le choix. Car la preuve bien connue de C.I. Lewis montre que cela nous conduit directement à $(A \wedge \neg A)/B$ comme suit.

$A \wedge \neg A$	(1) Hypothèse
A	(2) de (1)
$A \vee B$	(3) de (2)
$\neg A$	(4) de (1)
B	(5) de (3), (4) par le syllogisme disjonctif

Donc il faut y renoncer. Cependant, si nos pertes s'arrêtaient là, nous pourrions penser nous en tirer.

IV.4. Les pertes suivantes sont liées à '$\rightarrow$'.

Comme nous l'avons vu, les règles de *modus ponens*, *modus tollens* et *reductio ad absurdum* doivent être abandonnées. Cependant, en exprimant '$\rightarrow$' en termes

[10] First degree entailment [NDT].

de '¬' et '∨' on voit que ce sont en fait des variantes du syllogisme disjonctif. Cela suggère que c'est peut-être notre identification de '$A \rightarrow B$' avec '$\neg A \vee B$' qui est la cause du problème.

Or, il y a de nombreux arguments bien connus contre la lecture de l'implication matérielle comme une implication quelle qu'elle soit. Voir Anderson et Belnap [1975]. Je ne vais pas les répéter de nouveau. Je dirai juste que notre sémantique peut bien être considérée comme adéquate pour '¬', '∨', '∧' mais a besoin d'une manière ou d'une autre d'être étendue pour se charger de '→'. Au vu de tout le travail accompli dans ce domaine, cela semble tout à fait plausible.

IV.5. Cependant le problème de savoir comment étendre LP n'est pas facile. En particulier, il est bien connu que tout système contenant le connecteur '→' de H, S4, R, E ou T (voir Anderson et Belnap [1975, Chapitre 1]) est réduit à la trivialité si on y ajoute le schéma de compréhension naïf de la théorie des ensembles. (Voir, par ex., Routley [1977, Section 6]). Une caractéristique semblable peut être obtenue pour des théories sémantiquement closes. Pour être précis, dans toute théorie sémantiquement close dans un langage qui contient un connecteur '→' satisfaisant les règles

$$\frac{A \qquad A \rightarrow B}{B} \qquad\qquad \frac{A \rightarrow (A \rightarrow B)}{A \rightarrow B} \qquad \text{(Absorption)}^{11}$$

tout peut être prouvé. La preuve est comme suit. Considérons l'énoncé 'Si cet énoncé est vrai alors A l'est'[12]. Symboliquement soit (1) :

(1) $\qquad T(1) \rightarrow A$

où 'Tx' est 'x est vrai'.

Alors par le schéma de vérité de Tarski, nous avons :

$$T(1) \leftrightarrow (T(1) \rightarrow A).$$

En utilisant l'Absorption de gauche à droite on obtient

$$T(1) \rightarrow A$$

et *modus ponens* de droite à gauche

$$T(1).$$

D'où encore par *modus ponens* on a

$$A.$$

Ainsi soit *modus ponens* ou absorption doit être abandonné. Il est plausible qu'on puisse se passer d'Absorption, mais tous les systèmes que nous avons mentionnés la contiennent.

IV.6. Les pertes suivantes touchent à la quantification. D'après notre définition de '→',

$$\frac{\forall x(A \rightarrow B) \qquad (\forall x)A}{(\forall x)B}$$

[11] Priest appelle "Absorption" la règle qu'on nomme souvent "Contraction". [NDT].

[12] Priest examine ici une forme du paradoxe de Curry. [NDT].

est invalide. Ainsi, si nous identifions 'Tous les A sont B avec '$(\forall x)(A \to B)$' , nous devons abandonner ceci.

$$\frac{\text{Tous les } A \text{ sont } B \qquad \text{Tout est } A}{\text{Tout est } B}$$

Si nous trouvons un autre moyen pour obtenir '$\to$' alors cela, bien sûr, pourrait changer. Cependant, il est intéressant d'observer que de quelque manière que nous définissions 'Tous les A sont B' (écrivons le de manière neutre $[A,B]_x$) une des règles suivante doit être abandonnée

$$\text{A1}\ \frac{[A,B]_x}{[\neg B, \neg A]_x} \qquad \text{A2}\ \frac{(\forall x)A}{[B,A]_x} \qquad \text{A3}\ \frac{(\forall x)A \qquad [A,B]_x}{(\forall x)B}$$

Car si nous les gardons toutes on peut obtenir $(A\wedge\neg A)/B$ de la manière suivante. Soit x une variable ni dans A ni dans B.

1)	$A\wedge\neg A$	Hypothèse
2)	$(\forall x)(x = x)$	Vérité évidente (!)
3)	A	1)
4)	$A\wedge(\forall x)(x = x)$	2) & 3)
5)	$(\forall x)(A \wedge x = x)$	4)
6)	$[\neg B \vee x \neq x, A \wedge x = x]_x$	5) & A2
7)	$\neg A$	1)
8)	$(\forall x)\neg(A \wedge x = x)$	7)
9)	$[\neg(A \wedge x = x), \neg(\neg B \vee x \neq x)]_x$	6) & A1
10)	$(\forall x)\neg(\neg B \vee x \neq x)$	8) & 9) & A3
11)	$(\forall x)(B \wedge x = x)$	10)
12)	$B \wedge (\forall x)(x = x)$	11)
13)	B.	12)

Toutes les inférences à l'exception de A1, A2 et A3 sont LPQ valides.
Par conséquent une de celles-ci doit être supprimée. Peut-être la plus plausible est A2, en vertu de son lien avec le paradoxe de l'implication $A \to (B \to A)$. Mais ce n'est qu'une conjecture. Tout de même, si nous pouvions contenir nos pertes au syllogisme disjonctif, absorption et A2, nous pourrions toujours espérer (avec un peu d'optimisme) y arriver.

IV.7. L'approche précitée au problème (réduire nos pertes) est possible mais requiert évidemment beaucoup plus de travail avant que nous puissions raisonnablement être assurés de sa viabilité.

Il y a une approche bien plus radicale. Cependant, elle est si radicale que l'on est tenté de la rejeter d'emblée. Tout de même considérons la. Supposons que *modus ponens*, *reductio ad absurdum*, etc... soient vraiment seulement quasi-valides. On ne peut pas les abandonner sans paralyser le raisonnement classique. Pourquoi ne pas continuer à les utiliser quand même ?

IV.8. La proposition est que nous nous permettions les inférences quasi-valides même si elle ne sont pas généralement valides. Nous savons que les inférences quasi-valides préservent la vérité pourvu qu'il n'y ait pas d'énoncés paradoxaux concernés (voir Section IV.1). Par conséquent, si nous étions certains que nous n'avions pas affaire à des énoncés paradoxaux, nous pourrions utiliser les règles quasi-valides avec la conscience tranquille.

IV.9. Bien sûr, les énoncés paradoxaux ne portent pas leur marque sur leur front et il semble n'y avoir aucune raison de supposer que la classe des énoncés paradoxaux est décidable (i.e., nous n'avons aucun moyen effectif pour dire, en général, quand un énoncé est paradoxal). Cependant, les énoncés paradoxaux semblent être une proportion assez petite des énoncés avec lesquels nous raisonnons. (J'affirmerais qu'ils ne se présentent que dans des circonstances très spécifiques : quand il y a quelque genre d'autoréférence sémantiquement close, voir Section II.8.) C'est pourquoi, il semble raisonnable de formuler la maxime méthodologique suivante.

MM A moins que nous ayons des raisons particulières pour penser que des énoncés paradoxaux se présentent dans notre argument, nous pouvons nous autoriser à utiliser à la fois les inférences valides et quasi-valides.

Il semblerait plausible d'affirmer que dans nos raisonnements quotidiens nous présupposons (tout à fait correctement) que nous n'avons pas affaire à des affirmations paradoxales. (Cela expliquerait pourquoi les non-logiciens sont normalement si perplexes quand on leur présente des paradoxes logiques. Car il est alors clair qu'une présupposition du raisonnement ordinaire est en train d'être violée). Par conséquent MM a l'effet de légitimiser le statut quo. Cependant, là où l'homme de la rue refuse de continuer à raisonner avec des énoncés paradoxaux, nous savons maintenant qu'il est parfaitement correct de continuer tant que nous nous restreignons aux inférence valides (non quasi-valides). Puisque celles-ci sont très faibles, cela a précisément l'effet de contenir les singularités dangereuses que sont les paradoxes.

IV.10. Peut-être que la réaction immédiate est que si nous nous autorisions des inférences invalides nous ne serions plus sûrs de nos conclusions. C'est correct, mais ce n'est pas si impressionnant que ça en a l'air à première vue. Même si l'on croit dur comme fer que les règles d'inférences sont valides, on ne peut pas être plus certain des conclusions que l'on ne l'est des prémisses. En fait la nouvelle situation n'est guère différente. Si nos prémisses ne sont pas paradoxales, les règles d'inférences quasi-valides préservent la vérité. Par conséquent nous pouvons être aussi sûrs de la vérité de nos conclusions que nous sommes sûrs de la valeur seulement vrai de nos prémisses. Ce sera normalement seulement un petit peu moins que (si ce n'est le même) le niveau auquel nous sommes sûrs de la vérité de nos prémisses.

IV.11. Bien que ce ne soit pas un argument très fort contre notre présent projet, cela en indique une importante conséquence épistémologique. En fait nous nous trouvons devant un nouvel argument pour le faillibilisme en général, et en mathématiques en particulier. Supposons que nous ayons un argument pour une certaine assertion mathématique. Supposons de plus que l'argument emploie quelques inférences quasi-valides (comme la plupart des arguments). Alors pour autant que nous n'avons pas de raisons spécifiques de croire que nous avons quelque énoncé paradoxal dans la preuve, nous pouvons invoquer MM et déclarer que nous avons prouvé l'assertion. Mais qu'est-ce qui constituerait une raison spécifique de croire que nous avons un énoncé paradoxal dans la preuve ? Evidemment, si nous pouvons montrer que si un certain énoncé est vrai, il est faux et vice et versa (comme dans le cas du menteur) alors nous avons un énoncé paradoxal dans nos mains. Si l'énoncé contient des termes sémantiques et comporte de l'auto-référence, de l'auto-applicabilité, etc..., nous n'avons peut-être pas trouvé une démonstration de situation paradoxale, néanmoins nous avons d'excellentes raisons d'être suspicieux. Nous devrions avancer avec précaution. Cependant, il est toujours possible que des termes sémantiques soient entrés clandestinement, implicitement dans un énoncé à notre insu ; ou que le caractère paradoxal provient d'une autre raison dont nous n'avons pas conscience. Nous pouvons alors invoquer la maxime mais découvrir plus tard de quoi nous convaincre qu'il y a des énoncés paradoxaux dans la preuve et qu'ils y figurent de façon à invalider une règle d'inférence quasi-valide. Nous aurons alors à rejeter la preuve que nous acceptions auparavant. Bien que cela puisse paraître peu réaliste, il n'y a en fait rien d'essentiellement nouveau dans cette situation pour ceux qui sont un peu familiers de l'histoire des mathématiques.

IV.12. Ainsi, cette approche n'est peut-être pas aussi invraisemblable qu'elle le semblait à première vue. En fait, parmi les alternatives considérées, il semble qu'elle est plutôt la plus plausible. Cela nous permet d'avoir le beurre et l'argent du beurre, pour ainsi dire. Tandis que la première alternative souffre de la forte suspicion qu'il n'y aura peut-être pas assez de beurre.

IV.13. Finalement dans cette Section j'aimerais reconsidérer la question de l'incomplétude. Dans la Section II.6 nous avons vu que sous réserve que nous utilisions une théorie sémantiquement close nous pouvons éviter qu'il y ait des énoncés indépendants de la théorie, dont nous pouvons établir la vérité. Cependant, cela laisse ouverte la question de savoir si, pour toute arithmétique axiomatique il y a des énoncés indépendants.

Nous avons vu dans la Section II.6 que le théorème de Gödel ne peut pas être appliqué aux théories sémantiquement closes pour montrer qu'il y a des énoncés indépendants, puisque le théorème de Gödel ne s'applique qu'aux théories consistantes. Il y a cependant un autre théorème dû à Tarski qui, s'il est correct, montrerait qu'il doit y avoir de tels énoncés. Le théorème de Tarski affirme que l'ensemble des énoncés vrais de l'arithmétique n'est pas arithmétique. Si cela est vrai, alors l'ensemble des assertions vraies de l'arithmétique n'est certainement pas axiomatique. Cependant la preuve du théorème de Tarski échoue s'il y a des énoncés paradoxaux. La preuve standard du théorème de Tarski est la suivante :

Supposons que l'ensemble des énoncés vrais soit arithmétique, i.e., qu'il y ait un énoncé arithmétique à une variable libre Tx telle que toute instance de $Tk \leftrightarrow A$ soit vraie, où k est le nombre-code de A. Alors par l'argument diagonal usuel nous pouvons trouver une formule $\neg Tj$ dont le nombre-code est j. En substituant celle-là à 'A' on obtient

(1) $Tj \leftrightarrow \neg Tj$

Cela ne peut pas être vrai. Par conséquent la supposition est incorrecte.

Cette preuve ignore la possiblité d'énoncés paradoxaux. Si Tj est t ou f alors (1) est effectivement fausse. Mais si Tj est paradoxale (ce qu'elle est évidemment, puisqu'elle est la version arithmétique de 'Cette phrase est fausse') alors (1) est vrai ! Par conséquent la preuve est invalide.

Ainsi les théorèmes standards montrant qu'aucune arithmétique axiomatisée peut être complète échoue. Il n'y a aucune raison de croire que l'ensemble des énoncés vrais de l'arithmétique n'est pas axiomatique. Je conjecture qu'il l'est.

V. Postface autoréférentielle finale.

V.1. Il est toujours difficile d'admettre que quelque chose que vous avez écrit est faux. Mais c'est la position dans laquelle je dois admettre que je suis. Car ce que j'ai dit n'est pas sans portée sur ce que j'ai dit. En particulier, si ce que j'ai dit est vrai, alors certaines des choses que j'ai dites sont fausses (également). En particulier, j'ai fais certaines affirmations à propos des conditions de vérité d'énoncés et c'est précisément le genre d'auto-référence sémantique qui conduit au paradoxe.

V.2. Pour voir cela, considérons les conditions de vérité de $\ulcorner A$ est vrai$\urcorner$. Le biconditionnel de Tarski nous donne que

$$A \text{ est vrai ssi } A.$$

Par conséquent si A est vrai, $\ulcorner A$ est vrai$\urcorner$ est vraie. Si A est faux, $\ulcorner A$ est vrai$\urcorner$ est faux. Il suit que A est seulement vrai (vrai et pas faux) ssi $\ulcorner A$ est vrai$\urcorner$ est seulement vrai. A est seulement faux ssi $\ulcorner A$ est vrai$\urcorner$ est seulement faux. Résumons l'information comme ci-dessous.

A	A est vrai
t	t
p	p
f	f

Des considérations symétriques nous donnent la table suivante pour $\ulcorner A$ est faux$\urcorner$.

A	A est faux
t	f
p	p
f	t

Maintenant considérez l'affirmation métalinguistique

(1) Quelques énoncés sont vrais et faux.

(i.e. '$\exists s$(s est vrai et s est faux)' où le quantificateur porte sur tous les énoncés vrais ou faux $-$ce qui inclu bien sûr les paradoxaux.)

Alors en utilisant les tables ci-dessus et les conditions de vérité pour les quantificateurs données dans la Section III.12, on peut voir que (1) est vraie, en fait paradoxale. Ainsi sa négation

Aucun énoncé n'est vrai et faux

est vraie aussi. Aussi bien mon affirmation qu'il y a des énoncés paradoxaux que l'affirmation d'Aristote (rapportée dans la Section I.3) selon laquelle il n'y en a pas, sont vraies !

Peut-être encore plus surprenant, les deux affirmations

Tous les énoncés sont soit vrais soit faux
Quelques énoncés ne sont ni vrais ni faux

sont vraies ! On s'attendrait à ce que la première soit vraie mais pas la seconde puisqu'aucune clause n'a été faite pour les énoncés sans valeur de vérité. Cela devrait servir d'avertissement que l'on ne peut pas tirer des faits métalinguistiques à propos de LP à partir de ses matrices, d'une façon cavalière.

V.3. Bien sûr, le point est qu'une fois que nous avons renoncé à exiger que la théorie objet soit consistante, il n'y a pas de raison d'exiger que la métathéorie soit consistante. Effectivement cela s'impose à nous si nous voulons donner un exposé cohérent de ce qui est paradoxal. Toute inconsistance $A \wedge \neg A$ de la théorie-objet est simplement transformée par le schéma-T en une contradiction métathéorique $\ulcorner A$ est vraie et A n'est pas vraie$\urcorner$.

Peut-être l'une des inconsitances du métalangage les plus intéressantes est fournie par l'énoncé suivant :

(2) Cet énoncé n'est pas seulement vrai.

Si (2) est seulement vrai, alors il est certainement vrai et par conséquent pas seulement vrai. Ainsi (2) n'est pas seulement vrai. Mais si (2) n'est pas seulement vrai il est faux. Par conséquent il est seulement vrai. Donc (2) est seulement vrai. Nous voyons que (2) est seulement vrai et pas seulement vrai.

Ce raisonnement est tout à fait correct et souligne la différence entre cette approche et l'approche superficiellement semblable prise par van Fraassen [1968] et d'autres consistant à appeler les énoncés paradoxaux ni vrais ni faux. On pourrait penser que notre p peut être assimilé à *ni vrai ni faux*. Mais ce serait une erreur. En appelant les énoncés paradoxaux ni vrais ni faux on peut échapper au paradoxe du menteur. Cependant cela n'évite pas le paradoxe du menteur renforcé. (Car tel est (2).) Puisque le but de cette sorte d'approche est d'*éviter* les paradoxes, quelque chose de nouveau (et habituellement d'*ad hoc*) doit être fait pour l'éviter. En revanche le but de cet article n'a pas été d'éviter les paradoxes mais de voir comment ils peuvent être acceptés sans avoir d'accident. Une fois que nous avons accepté (comme j'ai montré que nous le devons) certains paradoxes comme des faits de la vie, alors les propriétés paradoxales de (2) apparaissent juste comme un autre fait.

V.4. Cette dernière Section illustre le fait que le sujet des assertions paradoxales est un sujet plein de surprises. Cependant *que* cela soit le cas n'est pas particulièrement surprenant. Après tout, comme nous en avons discuté dans la Section IV.9, nous présumons tous normalement que nous ne raisonnons pas au sujet d'une situation paradoxale : quand nous sommes face à une contradiction nous la prenons comme un signe que quelque chose s'est mal passé et nous refusons d'aller plus loin. (Et laissez moi ajouter encore une fois, avant d'être accusé d'accepter la première contradiction venue, que la plupart des contradictions sont un signe que quelque chose s'est mal passé : que nous avons une prémisse non vraie.) C'est précisément quand nous allons plus loin que notre monde familier disparaît et que nous nous trouvons dans un étrange nouveau décor. Le nouveau terrain a clairement besoin d'être exploré. Où cela nous mènera n'est pas encore clair. Toutefois une conséquence pour l'histoire des mathématiques apparaît déjà. La découverte par Russell d'un ensemble qui à la fois est et n'est pas élément de lui-même, est la plus grande découverte mathématique depuis $\sqrt{2}$.

Université d'Australie-Occidentale.

BIBLIOGRAPHIE

Anderson, A. R. et Belnap, N. D.,
1975 - *Entailment*, Princeton U.P.

Asenjo, F. G.,
1966 - 'A Calculus of Antinomies', *Notre Dame Journal of Formal Logic* **XV**, 497-509, **VII**, 103-105.

Curry, H. B.,
1942 - 'The Inconsistency of Certain Formal Logics', *Journal of Symbolic Logic* 7 (1942), 115-117.

da Costa, N.,
1974 - 'On the Theory of Inconsistent Formal Systems', Notre *Dame Journal of Formal Logic* **XV**,497-509.

Dunn, J. M.,
1976 - 'Intuitive Semantics for First-Degree Entailments, and "Coupled Trees" ', *Philosophical Studies* **29**, 149-168.

van Fraassen, B.,
1968 - 'Presupposition Implication and Self Reference', *Journal of Philosophy* **65**, 136-52.

Fraenkel, A., Bar Hillel, Y. and Levy, A.,
1973 - *Foundations of Set Theory*, North Holland, Amsterdam.

Gödel, K.,
1947 - 'What is Cantor's Continuum Problem', *American Mathematical Monthly* **54**, 515-25. Reprinted in P. Benacerraf and H. Putnam (eds.), Philosophy of Mathematics, Blackwell, 1964.
Trad. fr dans *Intuitionnisme et théorie de la démonstration*, Vrin, 1992, p514-531.

Hilbert, D. and Bernays, P.,
1939 - *Grundlagen der Mathematik* Vol. II, Springer Verlag,

Kleene, S. C.,
1952 - *Metamathematics*, North Holland, Amsterdam.

Lukasiewicz, J.,
1920 - 'On 3-valued Logic' in S. McCall *Polish Logic*, Oxford U.P.,

Priest, G.,
1974 - 'Type Theory in which Variables Range Over Predicates', Ph.D. thesis, University of London, 1974.

Routley, R.,
1977 - 'Ultralogic as Universal', *Relevance Logic Newsletter* **2** (1977).

Routley, R. and Routley, V.,
1972 - 'The Semantics of First Degree Entailment', *Nous* **VI** (1972), 335-359.

Tarski, A.,

1936 - 'Der Wahrheitsbegriff in den formalisierten Sprachen', *Studia Philosophica* **1**, 266-405. Réimprimé en anglais dans *Logic Semantics and Metamathematics*, Oxford U.P., 1956.

1944 - 'The Semantic Conception of Truth and the Foundations of Semantics', *Philosophy and Phenomological Research* **4**, 341-376. Réimprimé dans L. Linsky (éd.), *Semantics and the Philosophy of Language*, University of Illinois, 1952

Wittgenstein, L.,

1964 - *Philosophical Remarks*, Blackwell.

1930 - *Remarques philosophiques*, éd. posthume dûe à Rush Rhees. Trad. Jacques Fauve, Gallimard, 1975, rééd. 1984.