Avis concernant cette traduction
Ce livre est une traduction automatique du texte original anglais, réalisée à l'aide d'outils de traduction automatique. Il n'a été ni vérifié ni corrigé par l'auteur ni par aucune autre personne. Des erreurs, des inexactitudes, des formulations maladroites ou des problèmes de mise en page peuvent donc s'y trouver. L'édition originale anglaise fait foi.

Les paradoxes et leurs résolutions

Une compilation thématique

Avi Sion, Ph.D.

© Copyright Avi Sion, 1990-2017.

Protégé par les conventions internationales et panaméricaines sur le droit d'auteur.
Tous droits réservés.
Aucune partie de ce livre ne peut être reproduite de quelque manière que ce soit.
ou stockés dans un système de récupération ou transmis,
sans l'autorisation expresse de l'auteur-éditeur,
sauf en cas de brèves citations accompagnées des mentions de source appropriées.

Fièrement autoédité par Avi Sion, 2017.
Genève, Suisse.

www.TheLogician.net
avi-sion@thelogician.net

Informations de catalogage de la bibliothèque :

Sion, Avi
Les paradoxes et leurs résolutions
Une compilation thématique

Pas d'index, pas de bibliographie

Abstrait

Les paradoxes et leurs résolutions Il s'agit d'un recueil thématique d'Avi Sion. Ce volume rassemble les essais qu'il a écrits sur ce sujet au cours des 27 dernières années. On y trouve des exposés et des explications de nombreux paradoxes anciens et modernes (mais pas tous), parmi lesquels : le paradoxe de Protagoras-Euathle (Athènes, Ve siècle av. J.-C.), le paradoxe du menteur et le paradoxe de Sorites (tous deux attribués à Eubulide de Milet, IVe siècle av. J.-C.), le paradoxe de Russell (Royaume-Uni, 1901) et ses dérivés, le paradoxe du barbier et le paradoxe du catalogue général (également de Russell), le paradoxe de Grelling (Allemagne, 1908), le paradoxe de confirmation de Hempel (États-Unis, années 1940) et le paradoxe de prédiction de Goodman (États-Unis, 1955). Ce volume présente et commente également certains discours antinomiques présents dans certains textes bouddhistes (notamment chez Nagarjuna, en Inde, au IIe siècle de notre ère ; et dans le Sutra du Diamant, date inconnue, mais probablement du début de notre ère).

A est A (et non-A est non-A) ;
Rien n'est à la fois A et non-A ;
Rien n'est ni A ni non-A.

(Les trois lois de la pensée d'Aristote.)

Contenu

Avant-propos 8
1. La vanité du tétralemme 13
   1. Les phénomènes sont positifs 13
   2. Il n'existe aucun phénomène négatif 14
   3. Une expérience mal interprétée 16
   4. Définition de la négation 18
2. Clarification de la contradiction 21
   1. Dialectique 21
   2. Contradiction 23
   3. La cohérence est naturelle 29
3. Clarification de la négation 32
   1. Négation dans l'adduction 32
   2. Phénomènes positifs et négatifs 35
   3. La négation est secondaire 37
   4. La négation est une intention 41
4. Clarification du paradoxe 44
   1. Incohérence interne 44
   2. Le sophisme du concept volé 46
   3. Systématisation 49
   4. Propriétés 53
5. Le paradoxe du menteur (première version) 59
   1. Double paradoxe 59
   2. Le paradoxe du menteur 61
   3. Suite sur le paradoxe du menteur 63
   4. L'utilité des paradoxes 65
6. Le paradoxe du menteur (redux) 70
   1. Première approche 70
   2. Deuxième approche 72
   3. Troisième approche 75
   4. Quatrième approche 77
   5. Cinquième approche 79
   6. Sixième approche 81

7. Le paradoxe de Russell (première version) 84
1. Auto-adhésion 84
2. Le paradoxe de Russell 86
3. Impermutabilité 88
4. Le paradoxe du barbier 92
5. Le paradoxe du catalogue principal 94
6. Le paradoxe de Grelling 98
8. Le paradoxe de Russell (suite) 103
1. Éléments de logique de classe 103
2. Un double paradoxe apparent 109
3. Un peu d'histoire 121
4. Un peu d'autocritique 128
9. Suite sur le paradoxe de Russell 131
1. Ma résolution du paradoxe de Russell 131
2. Pourquoi les résolutions de Russell sont inadéquates 134
3. Pourquoi la résolution de Rescher est inadéquate 138
4. Autres incarnations du paradoxe 140
5. À propos du paradoxe du barbier 142
10. Le paradoxe de la confirmation de Hempel ... 144
1. Analyse traditionnelle 144
2. Nouvelle analyse 146
3. Le problème et sa solution 152
11. Le paradoxe de la prédiction de Goodman 160
1. Le problème allégué 160
2. La solution logique 161
3. Le concept de « grue » 164
4. Jusqu'à quel niveau de l'échelle ? 167
12. Le paradoxe des Sorites 171
1. Qu'est-ce qu'un tas ? 171
2. L'utilisation de termes vagues 172
3. Raisonner avec des termes vagues 175
4. Inventer de faux paradoxes 178

13. Protagoras contre Euathlus 182
1. Un paradoxe antique 182
2. Première résolution 184
3. Deuxième résolution 186
4. Résolutions inadéquates 190
14. Discours antinomique bouddhiste 195
1. Le tétralemme 195
2. Ni réel ni irréel 200
15. Discours antinomique bouddhiste plus avancé ....
207
1. La thèse « Je n'ai pas de thèse » 207
2. Appeler un chat un chat 225
Références principales 230

Avant-propos

Les paradoxes et leurs résolutions Il s'agit d'un recueil thématique. Il rassemble en un seul volume les essais que j'ai écrits par le passé (sur une période d'environ 27 ans) sur ce sujet.

Il comprend des exposés et des résolutions de nombreux paradoxes anciens et modernes, loin d'être tous, parmi lesquels : le paradoxe de Protagoras-Euathle (Athènes, Ve siècle avant notre ère), le paradoxe du menteur et le paradoxe de Sorites (tous deux attribués à Eubulide de Milet, IVe siècle avant notre ère), le paradoxe de Russell (Royaume-Uni, 1901) et ses dérivés, le paradoxe du barbier et le paradoxe du catalogue principal (également de Russell), le paradoxe de Grelling (Allemagne, 1908), le paradoxe de confirmation de Hempel (États-Unis, années 1940) et le paradoxe de prédiction de Goodman (États-Unis, 1955).

Je présente et commente également ici certains discours antinomiques que l'on trouve dans certains textes bouddhistes (notamment chez Nagarjuna, en Inde, au IIe siècle de notre ère ; et dans le Sutra du Diamant, date inconnue, mais probablement du début de notre ère).

Malgré son titre, il convient de noter que le présent ouvrage n'a pas vocation à être une étude exhaustive ; il comporte de nombreux paradoxes qu'il n'aborde ni ne traite.[1]

---

[1] Sur Wikipédia (https://en.wikipedia.org/wiki/List_of_paradoxes), on trouve une liste intéressante de paradoxes, qui les regroupe sous différentes rubriques et propose des liens vers des articles consacrés à chaque paradoxe. L'Encyclopédie de philosophie de Stanford en ligne contient

Par « paradoxe », on entend ici, bien sûr, *le double paradoxe apparent* . Une proposition paradoxale a la forme hypothétique auto-contradictoire « si P, alors non-P », ou « si non-P, alors P », où P est une proposition quelconque. Une telle proposition, prise isolément, n'est pas antinomique, c'est-à-dire qu'elle ne contrevient pas aux lois de la pensée, simplement parce qu'il existe une solution formelle : « si P, alors non-P » implique la proposition catégorique « non-P » ; et « si non-P, alors P » implique la proposition catégorique « P ». Un *seul* paradoxe constitue donc un discours logiquement légitime. Un *double* paradoxe est très différent : il s'agit d'affirmer que *«* si P, alors non-P » et « si non-P, alors P » sont simultanément vrais pour une instance donnée de P. Une telle affirmation est diamétralement opposée au principe de non-contradiction, puisqu'elle conclut que les propositions catégoriques « non-P » et « P » sont simultanément vraies pour une instance donnée de P.

Il est clair que la résolution des paradoxes (doubles) est essentielle à la logique, à l'ontologie et à l'épistémologie, car accepter une telle antinomie remettrait gravement en question la connaissance humaine, et même les facultés cognitives qui la rendent possible. Les sceptiques se délectent des paradoxes, car ils cherchent malicieusement à invalider la connaissance et l'esprit humains. Les défenseurs de la raison humaine sont donc tenus de relever chaque défi de ce genre et de le neutraliser de manière convaincante. Plus positivement, les paradoxes sont de formidables occasions d'apprendre de nouvelles choses sur notre façon de penser ou sur ce que nous croyons, et de découvrir et corriger nos erreurs. Si nous n'étions pas confrontés au paradoxe, nous pourrions rester

---

également des articles sur les paradoxes. Un ouvrage de référence sur ce sujet est *Paradoxes: Their Roots, Range, and Resolution de Nicholas Rescher* (Chicago, Ill. : Open Court, 2001).

inconscients de nos erreurs ; le paradoxe nous ouvre la voie à leur correction.

La « résolution » d'un paradoxe consiste à démontrer que le double paradoxe apparent n'est en réalité, pour une raison précise, qu'apparent ; autrement dit, il est illusoire. Pour y parvenir, il est généralement nécessaire d'examiner très attentivement le discours sous-jacent ou environnant et d'y déceler l'erreur commise. On peut s'être appuyé sur une notion trop vague, avoir commis une équivoque, avoir formulé une supposition injustifiée, etc. La résolution d'un paradoxe n'est généralement pas un processus mécanique, mais exige une grande perspicacité et une profonde réflexion. Les principaux paradoxes sont loin d'être faciles à résoudre ; chacun requiert une attention particulière et un traitement spécifique. Souvent, même après avoir été découverts par un logicien ou un philosophe, ils demeurent longtemps irrésolus.

Concernant l'ordre des textes dans le présent recueil, je tiens à apporter la précision suivante. Dans mes précédents ouvrages thématiques (à savoir, *Les Lois de la pensée* , *Le Soi* , *L'Éthique* et *La Théologie*), j'ai généralement opté pour une présentation chronologique. Toutefois, dans le cas présent, l'ordre a été déterminé par des considérations à la fois didactiques et chronologiques.

- J'ai choisi comme premier chapitre un essai intitulé « La vanité du tétralemme », extrait de mon ouvrage **A Fortiori Logic** (2013). Les deux chapitres suivants, « Clarifier la contradiction » et « Clarifier la négation », sont tirés de mon livre **Ruminations** (2005). Ces trois premiers essais servent, d'emblée, à mettre en lumière et à réaffirmer les lois de la pensée.
- Ensuite, j'insère, en guise d'introduction au sujet des paradoxes, l'essai « Paradoxes », tiré de mon premier ouvrage, *Future Logic* (1990).

- Ce texte est suivi de quelques réflexions préliminaires sur les paradoxes en général et le paradoxe du menteur en particulier, dans l'essai « Le paradoxe du menteur (première version) », extrait de * *Logique future* et *ruminations** . J'insère ensuite une analyse plus approfondie de ce paradoxe dans « Le paradoxe du menteur (suite) », extrait de **Logique à tout prix** .
- Dans l'essai « Le paradoxe de Russell (première version) », également tiré de * *Logique future* et *réflexions** , je présente quelques réflexions préliminaires sur le paradoxe de Russell, ainsi que sur certains de ses dérivés, à savoir le paradoxe du barbier, le paradoxe du catalogue de livres et le paradoxe de Grelling. J'insère ensuite une analyse plus approfondie de ce paradoxe dans « Le paradoxe de Russell (suite) », extrait de * *Logique à tout prix** .
- J'ai ensuite inclus les essais « Le paradoxe de la confirmation chez Hempel » et « Le paradoxe de la prédiction chez Goodman », tous deux extraits de mes *Réflexions logiques et spirituelles* (2008-2009). Puis, les essais « Le paradoxe de Sorites est artificiel » et « La résolution du conflit entre Protagoras et Euathlus », tous deux tirés d'un ouvrage que je n'ai pas encore achevé ni publié (titre provisoire : « Sujets de logique, de philosophie et de spiritualité », partiellement publié en ligne en 2017).
- Enfin, le présent volume comprend deux essais, « Discours antinomique bouddhiste » et « Discours antinomique plus bouddhiste », tirés respectivement de mon *ouvrage Illogique bouddhiste* (2002) et de *Ruminations* et *réflexions logiques et spirituelles* .

Quant à l'importance relative des paradoxes présentés et traités ici, on peut dire ce qui suit.

À mon sens, le paradoxe du menteur et le paradoxe de Russell sont particulièrement importants, compte tenu de

leur impact potentiel sur la théorie logique ; et je recommande tout particulièrement mes réflexions les plus récentes à leur sujet (dans les essais intitulés « redux »). Bien que j'aie déjà abordé ces deux paradoxes dans des écrits antérieurs (ici intitulés « premiers »), j'ai pu, en 2013, alors que je finalisais mon ouvrage * *A Fortiori Logic* * [2], les approfondir considérablement. J'étais alors au sommet de ma forme, intellectuellement à mon apogée, et pouvais donc les appréhender avec une clarté inédite.

Dans la nouvelle étude du paradoxe du menteur, je démontre que ce paradoxe est une hydre à plusieurs têtes, qui ne peut être expliquée uniquement par l'autoréférence, comme on le fait généralement, mais qui soulève une variété de problèmes qu'il convient d'aborder. Dans la nouvelle étude du paradoxe de Russell, je montre que cette énigme de la logique des classes n'est pas due, comme on le suppose habituellement, à une difficulté avec l'idée de *non* -appartenance à soi, mais au contraire à l'impossibilité de l'appartenance *à soi* ; il s'agit d'une résolution nouvelle et définitive du paradoxe, dont j'expose en détail les nombreuses implications.

Les paradoxes modernes de Hempel et Goodman sont également importants pour la théorie logique, car ils soulèvent des doutes quant à l'induction ; toutefois, leur résolution est relativement aisée. Le paradoxe antique du sorite est lui aussi intéressant d'un point de vue logique, car il illustre l'importance de la précision intellectuelle et verbale dans le discours ; mais, en tant que paradoxe, il se résout facilement. Le paradoxe plus ancien de Protagoras et Euathle n'a pas une grande importance logique, si ce n'est peut-être pour nous enseigner que, lorsqu'on

[2] En fait, ces essais n'ont rien à voir avec un argument a fortiori ; mais il se trouve que je les ai écrits lorsque ce livre était presque terminé, et je voulais les publier dès que possible, alors je les ai placés dans une annexe.

s'efforce de résoudre des paradoxes, il ne faut pas s'enliser dans des questions logiquement non pertinentes (en l'occurrence, des questions juridiques).

En ce qui concerne le discours antinomique bouddhiste, il est important de l'aborder à cette époque de l'histoire, car le bouddhisme est à la mode dans certains milieux, et certaines de ses doctrines y sont reçues dans un esprit trop dogmatique.

Si j'écris à l'avenir d'autres essais sur les paradoxes, j'espère les inclure dans des éditions augmentées du présent volume.

1. La vanité du tétralemme

D'après la logique A Fortiori, appendice 7.3.

L'attaque la plus radicale contre la raison consiste à remettre en question les lois de la pensée, car elles constituent le fondement même de tout discours rationnel. On nie d'abord la loi d'identité en affirmant que les choses ne sont jamais tout à fait ce qu'elles paraissent être, ou que leur nature se rapproche davantage du gris que du noir et du blanc. Cette remarque est absurde, car elle prétend à une certitude et une clarté absolues. [3]On nie ensuite les lois de non-contradiction et du tiers exclu en prétendant que les choses peuvent être et ne pas être, ou n'être ni l'un ni l'autre. C'est le « tétralème », cette logique quadruple privilégiée dans les philosophies indienne et chinoise, dans le mysticisme religieux, et à laquelle certains « scientifiques » font de plus en plus référence. Pour saisir la vanité du tétralème, il est nécessaire de comprendre la nature de la négation et son rôle comme fondement de la logique humaine.

1. Les phénomènes sont positifs

La première chose à comprendre est que tout ce que nous expérimentons est *un phénomène positif.* Tout ce que nous percevons par nos sens, dont nous nous souvenons, que nous imaginons ou même que nous appréhendons par intuition, possède nécessairement un contenu pour être

[3] J'ai omis par inadvertance les mots « et du tiers exclu » dans les éditions originales de *A Fortiori Logic* , bien qu'il ressorte clairement du reste de la phrase que je les avais voulus.

perçu. Chaque organe sensoriel est une fenêtre ouverte sur un type distinct de phénomène positif. Nous voyons le ciel bleu, nous entendons le chant des oiseaux, nous sentons l'air frais, nous goûtons un fruit, nous sentons la texture et la chaleur de la terre, etc. De même, les images et les sons qui nous viennent à l'esprit, qu'ils proviennent de la mémoire ou de l'imagination, sont des phénomènes positifs ; et même les objets de l'intuition possèdent nécessairement un contenu que nous pouvons appréhender. Deuxièmement, il faut réaliser que de nombreux phénomènes positifs peuvent apparaître simultanément dans l'espace. Cela est vrai pour chaque type de phénomène. Ainsi, le ciel bleu peut n'occuper qu'une partie de notre champ de vision, bordé par des arbres verts et des bâtiments gris ; nous pouvons entendre simultanément le chant des oiseaux et le bruit des voitures ; et ainsi de suite. Troisièmement, de nombreux phénomènes positifs peuvent, à tout moment, partager l'espace que nous percevons. Ainsi, à des phénomènes visuels comme le ciel peuvent se superposer d'autres types de phénomènes : le chant des oiseaux dans les arbres, l'odeur de la circulation dans les rues, les sensations corporelles, etc. Il est même possible d'avoir des hallucinations, comme si des objets de la perception mentale étaient projetés dans l'espace physique. Par exemple, l'image de ses lunettes peut persister un certain temps après les avoir retirées. Enfin, chaque phénomène positif, quel qu'il soit, varie dans le temps, plus ou moins rapidement. Ainsi, le ciel bleu peut devenir rouge ou sombre, le chant des oiseaux ou le bruit de la circulation peuvent augmenter ou diminuer, voire s'interrompre momentanément, et ainsi de suite.

2. Il n'y a pas de phénomènes négatifs

Pour exprimer toutes ces possibilités perceptives – différences d'espace, de temps et autres –, nous avons besoin du concept de négation, ou plus précisément d'un acte de négation. Sans « négation », nous ne pouvons appréhender le monde de manière rationnelle ; elle constitue le fondement même de l'organisation logique de notre expérience. Ainsi, dans un champ visuel donné, où apparaissent (par exemple) un ciel bleu et des arbres, pour pouvoir dire « le ciel s'arrête ici, là où commencent les arbres », nous avons besoin de l'idée de « négation » : d'un côté d'une certaine limite, le ciel est visible et de l'autre, il ne l'est pas, tandis que d'un côté, les arbres ne sont pas visibles et de l'autre, ils le sont. De même, concernant le temps, pour pouvoir décrire un changement, par exemple du ciel bleu au ciel rose, nous avons besoin de l'idée de « négation » : auparavant, cette partie du ciel était bleue et non rose, et plus tard, elle était rose et non bleue. Là encore, la notion de « négation » nous est nécessaire pour exprimer des différences à d'autres égards – par exemple, pour dire que « le chant des oiseaux semble provenir des arbres plutôt que des bâtiments ». La négation est donc l'un des tout premiers outils de la logique, intervenant dès le tri des expériences.

De plus, la négation conserve un rôle central lorsqu'on aborde les abstractions. La connaissance conceptuelle, constituée de termes et de propositions fondés directement ou indirectement sur des phénomènes perceptifs, repose initialement sur notre capacité à percevoir des similarités entre les objets de la perception : « ceci semble ressembler un peu à cela ». Nous projetons ainsi mentalement l'idée que ceci et cela « ont quelque chose en commun », une propriété commune abstraite (c'est-à-dire non phénoménale, non perçue d'aucune manière), à laquelle nous pouvons choisir de donner un

nom. Cependant, pour aller plus loin dans ce processus conceptuel, nous devons être capables de nier, c'est-à-dire d'affirmer que « certaines choses autres que ceci et cela ne possèdent pas la propriété commune abstraite que ceci et cela semblent avoir », ou encore que « ceci et cela n'ont pas tout en commun ». Autrement dit, nous devons pouvoir affirmer non seulement qu'une chose ressemble à une autre d'une certaine manière, mais aussi que ces choses, ou d'autres, ne se ressemblent pas de cette manière ou d'une autre. Ainsi, la négation est essentielle pour donner un sens à l'information, même au niveau conceptuel de la conscience.

Qu'est-ce que la négation ? Pour répondre à cette question, il faut d'abord comprendre qu'il *n'existe pas de phénomènes négatifs dans le domaine de l'expérience* . Tout ce que nous percevons est un phénomène positif, car sinon, nous n'aurions rien à percevoir. Nous ne pouvons « percevoir » un état de choses négatif qu'après avoir défini mentalement un état de choses positif à rechercher, puis l'avoir recherché. Si, malgré une recherche assidue, nous ne le trouvons pas, nous concluons inductivement qu'il est « absent », c'est-à-dire « non présent ». Ainsi, les phénomènes positifs précèdent les phénomènes négatifs, et non l'inverse. L'existence précède logiquement la non-existence. Les phénomènes négatifs ne sont des « phénomènes » que métaphoriquement, par analogie avec les phénomènes positifs ; en réalité, les phénomènes négatifs n'existent pas. La « négation » n'est pas un concept au sens d'une abstraction issue de nombreuses expériences particulières partageant une certaine propriété. La négation est un outil de l'observateur pensant, comme décrit précédemment. C'est un acte, une intention.

3. Une expérience mal interprétée

Pour illustrer la confusion qui règne chez certaines personnes – même chez certains scientifiques – face à la négation, je vous propose l'exemple suivant, tiré de l'ouvrage de Richard Dawkins, * *Le Plus Grand Spectacle du Monde : La Preuve de l'Évolution* * [4]. Il y décrit une expérience menée par Daniel J. Simons, au cours de laquelle on demande à des participants de regarder une courte vidéo et de compter le nombre d'occurrences d'un certain événement. À la fin, on leur pose une toute autre question : ont-ils remarqué la présence d'un homme déguisé en gorille pendant la vidéo ? La plupart admettent ne pas l'avoir vu [5]. Selon Dawkins, cette expérience nous permet de déduire que « le témoignage oculaire, l'« observation directe », le « vécu » – tous ces éléments sont, ou du moins peuvent être, désespérément peu fiables ».

Mais cette conclusion est erronée à partir des données disponibles, car il confond expérience positive et expérience négative. Les personnes qui ont visionné la vidéo étaient trop occupées à chercher ce qu'on leur demandait d'observer pour remarquer le gorille. Plus tard, lors d'un second visionnage, elles l'ont effectivement aperçu. Il n'y a aucune raison de s'attendre à ce que nous percevions réellement tout ce qui est présenté à nos sens. Nos expériences sensorielles sont toujours, nécessairement, sélectives. La validité de la perception sensorielle en tant que telle n'est pas remise en cause par la portée limitée de certaines perceptions sensorielles. La preuve en est que c'est *grâce à des perceptions sensorielles supplémentaires* que nous découvrons ce qui nous avait échappé auparavant. La non-perception de

---

[4] Pages 13-14.

[5] La vidéo peut être visionnée à l'adresse suivante : www.theinvisiblegorilla.com/videos.html .

quelque chose ne constitue pas une perception erronée, mais simplement une perception incomplète. « Je n'ai pas vu X » n'implique pas déductivement « J'ai constaté l'absence de X », même si la répétition de la première affirmation tend à impliquer inductivement la seconde.

Un phénomène négatif n'est pas comparable à un phénomène positif, c'est-à-dire quelque chose que l'on peut percevoir ou pressentir directement. La négation est nécessairement le produit d'une cognition indirecte, c'est-à-dire d'un processus inductif (plus précisément, adductif). Nous émettons mentalement l'hypothèse qu'un tel phénomène positif est absent, puis nous testons et confirmons cette hypothèse en recherchant, sans succès, le phénomène positif [6]. Si nous venions à trouver le phénomène positif, l'hypothèse de négation serait immédiatement rejetée, car la fiabilité d'une négation est bien inférieure à celle d'une expérience positive. Nous ne formulerions même pas la négation si nous avions déjà perçu, par le passé ou dans l'immédiat, le phénomène positif. Et si nous la formulions, nous reviendrions naturellement sur notre affirmation si nous rencontrions ultérieurement le phénomène positif. Par conséquent, le contenu des phénomènes négatifs est nécessairement toujours hypothétique, c'est-à-dire provisoire dans une certaine mesure ; il n'est jamais ferme et certain comme celui des phénomènes positifs (expérimentés).

Les affirmations négatives, comme les affirmations positives, peuvent être justes ou fausses. Si l'on a cherché avec diligence un phénomène positif sans le trouver, on peut logiquement en affirmer la négation. Une telle

---

[6] Ne pas trouver, c'est ne pas trouver. Objectivement, il est important de noter que ne pas trouver est un phénomène négatif, et non positif. Mais subjectivement, quelque chose de positif peut se produire en nous : peut-être un sentiment de déception ou un soulagement persistant. Pour en savoir plus, consultez mon ouvrage *« Réflexions »* , chapitre 9.

affirmation est nécessairement inductive : elle n'est valable que tant que le phénomène positif est activement recherché et non trouvé. Dès que le phénomène positif est observé, la négation n'est plus justifiée. Si l'on n'a *pas* cherché le phénomène positif, ou si l'on ne l'a pas cherché avec toute la diligence requise, peut-être en raison d'une distraction (comme dans l'exemple cité plus haut), alors bien sûr, l'affirmation de négation est sujette à caution ; elle est certainement inductivement faible, et il est fort probable qu'une observation ultérieure puisse la contredire.

4. Définition de la négation

Comment définir la négation, alors ? On pourrait dire qu'elle se définit par les lois de non-contradiction et du tiers exclu. Autrement dit, pour tout terme « X » et sa négation « non-X », la relation entre eux est par définition la disjonction « soit X, soit non-X » – ce qui signifie ici que ces termes (X et non-X) ne peuvent être simultanément vrais et faux ; ils sont donc exclusifs et exhaustifs. Qu'entends-je par « définition » ? Est-ce un acte arbitraire ? Non, il s'agit de désigner une évidence, d'un choix délibéré. En l'occurrence, il s'agit de désigner l'instrument du discours rationnel dont nous avons besoin pour ordonner l'expérience et en produire des dérivés conceptuels cohérents. Cet instrument doit être ainsi construit ; une autre construction ne remplirait pas la fonction que nous lui attribuons. Autrement dit, la seule manière concevable d'ordonner logiquement notre connaissance est par la négation, définie par les lois de non-contradiction et du tiers exclu. Sans cet outil, l'analyse de l'expérience est impossible.

Supposons maintenant que quelqu'un conteste l'affirmation précédente. Il demande : « Comment savez-vous *que* le dilemme "soit X, soit non-X" est vrai ?

Vous avez arbitrairement *défini* les choses ainsi, mais cela ne signifie pas que c'est un fait ! Ne pourrait-on pas tout aussi bien affirmer que le tétralème “soit X, soit non-X, *soit les deux, soit aucun* ” est vrai ? » La réponse à cette objection est très simple. Supposons que j'accepte cette critique et que j'admette le tétralème. Décomposons alors cette disjonction quadruple, en plaçant d'un côté l'alternative unique « X » et de l'autre l'alternative triple « non-X, soit les deux, soit aucun ». J'ai alors de nouveau un dilemme : « soit X, soit non-X, soit les deux, soit aucun ». Définissons ensuite un nouveau concept de négation sur cette base, de sorte que nous obtenions une disjonction de deux alternatives au lieu de quatre. Appelons la seconde alternative complexe « non-X ou les deux ou aucun » de cette disjonction « NON-X » et appelons-la « la super-négation de X ».

Ainsi, l'objecteur et moi convenons désormais que la disjonction « soit X, soit NON-X » est exclusive et exhaustive. Nous convenons, vraisemblablement, que ce nouveau dilemme ne peut être opposé à son tour par un tétralème de la forme « soit X, soit NON-X, soit les deux, soit aucun des deux » – car une telle opposition pourrait assurément être contrée par une autre division et redéfinition. Nous ne pouvons raisonnablement répéter ce processus à l'infini ; cela reviendrait à bloquer toute pensée rationnelle pour toujours. Ayant ainsi bloqué toutes les voies de la pensée, l'objecteur ne pourrait prétendre avoir une meilleure pensée, ni même une pensée tout court. Il n'y a donc aucun profit à poursuivre l'objection. Le tétralème n'est donc qu'une provocation, car nous avons parfaitement pu parer l'attaque. Ayant convenu que la nouvelle disjonction « soit X, soit NON-X » est logiquement irréfutable, nous devons admettre que la disjonction initiale « soit X, soit non-X » était logiquement correcte dès le départ. Car je peux vous dire que ce que j'entendais par non-X, ou la « négation de X »,

était dès le début ce que signifie maintenant NON-X, ou la « super-négation de X » !

Je ne me suis jamais intéressé à une négation relative et faible, mais j'ai, dès le départ, recherché une négation absolue et forte. Car cette négation totale, et rien de moins radical, est l'outil dont nous avons tous besoin pour ordonner l'expérience et développer la connaissance conceptuelle de manière cohérente et efficace. Autrement dit, quelle que soit la version affaiblie de la négation que l'on tente d'inventer [7], nous pouvons toujours en proposer une version forte à laquelle s'appliquent sans aucun doute les lois de non-contradiction et du tiers exclu. Si une telle négation n'existait pas, il faudrait l'inventer. Nul ne peut la détruire en la niant ou en l'édulcorant. Ceux qui s'y essaient ne sont que des sophistes qui ignorent la source, la nature et la fonction de la négation dans le discours humain. Ils la réduisent à une simple manipulation symbolique et ne réalisent pas que son rôle dans le discours humain est bien plus fondamental et complexe. La négation est l'instrument indispensable à toute tentative de connaissance au-delà de la simple perception.

---

[7] Certains affirment que le principe de non-contradiction est logiquement nécessaire, mais que le principe du tiers exclu ne l'est pas. Il est clair que cette affirmation se réfute de la même manière. S'ils avancent les trois alternatives « soit X, soit non-X, soit "ni X ni non-X" », on peut à nouveau scinder la disjonction en deux, avec d'un côté « X » et de l'autre « non-X ou "ni X ni non-X" », puis procéder comme pour le tétralème. On peut faire de même si l'on accepte le principe du tiers exclu mais que l'on rejette le principe de non-contradiction. Toutes ces tentatives sont des sophismes.

2. Clarifier la contradiction

D'après les ruminations 1, 5 et 9.

1. Dialectique

Les trois « lois de la pensée » peuvent être brièvement expliquées comme suit :

(i) *Thèse* : il existe certaines apparences ; les apparences apparaissent.

(ii) *Antithèse* : il existe des incompatibilités entre certaines de ces apparences ; dans de tels cas, l'une ou les deux doivent être fausses.

(iii) *Synthèse* : certaines apparences restantes doivent être vraies ; découvrez lesquelles !

Dans cette perspective, on peut considérer *la dialectique* comme une forme fondamentale de la pensée, grâce à laquelle la connaissance progresse sans cesse. Elle n'est pas un simple détail, un processus de pensée occasionnel, mais une force motrice, un moteur de la pensée.

Ces lois ne sont pas de simples informations, mais des appels à l'action cognitive. Elles prescrivent des mesures cognitives proactives et correctives, afin d'assurer (autant que possible à tout moment) une vérification continue, la cohérence et l'exhaustivité.

(i) La loi d'identité nous enjoint de rechercher les faits et de les analyser avec la plus grande rigueur. Son but est d'inculquer aux individus un sain respect des faits, dans le cadre de l'observation et du jugement. Il s'agit essentiellement d'un appel à l'honnêteté et à la soumission

à la vérité. On croit souvent, ou on agit comme si l'on croyait, qu'ignorer ou nier les faits ou arguments déplaisants les fera disparaître ; or, la loi d'identité affirme : « Non, ils ne disparaîtront pas, vous devez les prendre en considération. »

Certains pensent qu'il nous est impossible d'ignorer que « A est A ». Loin de là ! Nous le faisons tous souvent : lorsque nous refusons de voir ou d'admettre une preuve ou une démonstration logique ; lorsque nous fuyons la réalité ou l'éludeons après l'avoir entrevue ; lorsque nous nous mentons à nous-mêmes ou aux autres ; et ainsi de suite. Si nous respections toujours la loi d'identité, il serait inutile de la formuler. La logique énonce l'évidence, car elle est souvent rejetée.

(ii) Lorsque la loi de non-contradiction nous dit « vous ne pouvez pas à la fois affirmer et nier une proposition », elle nous dit aussi que si jamais au cours d'un discours nous rencontrons une situation où une proposition semble à la fois vraie (pour une raison quelconque) et fausse (pour d'autres raisons), ***nous devons remonter en amont dans notre discours et découvrir où nous nous sommes trompés au cours de celui-ci*** [8]***, et nous devons effectuer une correction appropriée de manière à éliminer la difficulté*** .

Nous ne nous contentons pas de dire : « Ah, il y a une contradiction », et de rester là, déconcertés. Non, nous sommes poussés à chercher une solution au problème, c'est-à-dire à résoudre la contradiction. Nous en déduisons qu'il y a forcément une erreur dans notre raisonnement initial qui nous a conduits à cette impasse, une erreur d'observation ou de raisonnement qu'il convient de corriger. Tant que cette situation perdure et que nous ne parvenons pas à identifier la source de l'erreur, la crédibilité de toutes les connaissances connexes s'en

[8] « Vérifiez vos locaux », disait Ayn Rand.

trouve proportionnellement diminuée. La cohérence doit être rétablie au plus vite, sous peine de remettre en question toutes les connaissances ultérieures.

(iii) De même, le principe du tiers exclu ne se contente pas d'affirmer qu'« aucune proposition ne peut être déclarée ni vraie ni fausse ». Ce principe exige que, si nous nous trouvons dans une telle situation, et qu'il s'avère effectivement qu'une proposition et sa négation exacte semblent toutes deux fausses, nous ne pouvons ni en rester là ni espérer trouver un compromis ; nous devons, au plus vite, trouver une raison valable de choisir l'une ou l'autre position. Il n'existe aucun juste milieu logiquement acceptable, aucune échappatoire.

Ces implications pratiques inhérentes aux lois de la pensée peuvent également être caractérisées comme une pensée dialectique. Dans cette perspective, la « thèse » correspond à notre connaissance (ou opinion) telle qu'elle se présente à un instant donné ; l'« antithèse » est la découverte d'une faille logique dans cette thèse, qui nous amène à douter d'elle et à la réexaminer ; et enfin, la « synthèse » désigne les corrections que nous apportons à nos prémisses afin de résoudre la difficulté rencontrée et d'atteindre un nouvel état de connaissance moins problématique.

2. Contradiction

Beaucoup de gens se méprennent sur ce que nous, logiciens, entendons par « **contradiction** ». Le contradictoire d'un terme « A » est sa négation, « non A », qui désigne tout ce qui, dans l'univers, n'est pas A, c'est-à-dire précisément tout ce qui est absent du monde. La relation de contradiction entre A et non-A est réciproque, réversible et parfaitement symétrique.

La présence d'une chose (A) exclut son absence (c'est-à-dire non A) dans cette même chose, et réciproquement, si toutes les coordonnées d'espace et de temps sont identiques. Cependant, cela n'exclut pas la possibilité logique qu'une même chose soit en partie A et en partie non A. Ainsi, la loi de la pensée « soit A, soit non A » peut aussi s'exprimer plus quantitativement comme « soit entièrement A, soit entièrement non A, soit encore une partie A et une partie non A ».

Certains invoquent cette possibilité de trois alternatives comme argument *contre* les lois de la pensée ! Mais il s'agit là d'un malentendu, voire pire, d'un sophisme délibéré.

Si quelque chose, par exemple « B », implique mais n'est pas impliqué par non-A, alors B est aussi « incompatible » avec A que non-A, mais n'est pas contradictoire avec A : il est simplement *contraire* à A. Le non-A contradictoire de A diffère des contraires de A en ce que *l'absence* de non-A implique A, alors que dans le cas de simples contraires comme B (ou B1 ou B2… etc.), cette relation logique supplémentaire d'« exhaustivité » ne s'applique pas.

Lorsqu'on place des éléments contradictoires dans une disjonction, « soit A, soit non-A », cette disjonction signifie à la fois exclusion mutuelle (« ou », signifiant « pas ensemble ») et exhaustivité (« soit », signifiant « et il n'y a pas d'autre alternative »). Elle implique : si « A », alors non « non-A » ; et si non « A », alors « non-A ».

En revanche, on peut placer un nombre quelconque de contraires dans une disjonction : « A ou B ou B1 ou B2… etc. », de sorte que la présence d’un seul de ces contraires implique l’absence de tous les autres. Cependant, une telle disjonction n’est pas exhaustive, à moins de préciser que la liste des contraires qui la composent est complète. Si cette liste *est* effectivement complète, alors la négation de

tous les contraires, sauf un, implique l'affirmation du contraire restant. Ainsi, « non-A » peut être assimilé à la disjonction exhaustive de toutes les choses du monde « contraires à A ».

Ce qui *est différent* de A, par exemple C, n'est pas nécessairement contradictoire, ni même contraire à A. **La simple différence n'implique pas l'incompatibilité** . Des choses différentes (comme A et C) peuvent être compatibles, c'est-à-dire capables de coexister au sein d'une même chose, au même moment et au même endroit. La « différence » signifie simplement que nous sommes *capables de distinguer* les choses concernées : autrement dit, elles ne sont pas identiques lorsqu'elles se présentent à notre conscience. Des choses « similaires » peuvent être identiques en apparence, mais ne pas être une seule et même chose (par exemple, deux exemplaires du même type) ; ou bien elles peuvent être une seule et même chose (c'est-à-dire des parties d'un tout), sans pour autant être identiques.

Ainsi, par exemple, la relation logique entre les couleurs noir et blanc dépend de la précision avec laquelle on les perçoit. Elles sont différentes, puisqu'elles sont discernables. Puisqu'elles peuvent coexister sur différentes parties d'une même surface, elles sont globalement compatibles. Cependant, en tant que telles , elles sont contraires ; autrement dit : si je perçois une surface ou une partie de surface comme totalement blanche, et que vous percevez *le même* lieu et le même moment comme totalement noir, nos perceptions sont incompatibles [9]. Cette inconciliabilité n'est pas une

[9] Notre désaccord n'est pas d'ordre terminologique, notez-le bien. Par le passé, nous nous sommes accordés sur la correspondance entre les expériences « noires » et « blanches » ; or, ici, nos points de vue divergent soudainement.

contradiction, car une surface peut n'être ni noire ni blanche.

L'expression « **contradiction dans les termes** » désigne un terme composé d'éléments incompatibles, tels que « A et non A » ou « A et B (où B est contraire à A) ». Un tel terme confus peut être qualifié de paradoxal et d'incohérent, puisqu'il sous-entend la possibilité de la contradiction, niant ainsi les lois de la pensée et, par conséquent (par généralisation), autorisant tout, y compris la négation de la conjonction « A et non A ».

Par exemple, l'expression « réalité illusoire » est une contradiction dans les termes. En revanche, notez que des expressions comme « un être humain inhumain » ou « un Juif antisémite » ne sont pas, à proprement parler, des contradictions ; elles renvoient à des possibilités naturelles de conjonction, seule la terminologie employée leur donne une apparence superficielle de contradiction (c'est-à-dire qu'il existe des personnes qui se comportent de manière inhumaine, ou des Juifs qui haïssent leur propre peuple).

La proposition « A n'est pas A » (ou « ce qui est A n'est pas A »), étant contradictoire, implique « A est A », sa forme contradictoire. Il convient de le souligner, bien que cela paraisse évident, car cela met en relation deux concepts importants : « l'incohérence interne » et « la logique des paradoxes ».

L'énoncé « A n'est pas A » est logiquement impossible, car il affirme et nie simultanément la même chose. Par conséquent, l'énoncé contraire est vrai. Cet énoncé, c'est-à-dire « A est A », est logiquement nécessaire, car *même son contraire,* « A n'est pas A », l'implique.

Quiconque affirme « A n'est pas A » admet implicitement « A est A » – ***ipse dixit*** , il l'a dit lui-même ! En revanche, quiconque affirme « A est A » est cohérent avec lui-même.

L'auto-contradiction se compose de trois éléments :

(i) La proposition en question, appelons-la P.

(ii) L'aveu qu'il s'agit *d'une déclaration assertive* , c'est-à-dire une déclaration qui affirme ou nie quelque chose.

(iii) L'aveu que toutes les déclarations assertives *impliquent des prétentions* à la conscience, à la connaissance, à la vérité, etc.

Ainsi, étant donné P (par exemple « la réalité est inconnaissable »), admettons que P implique « ceci est une assertion » – or, toute assertion implique une certaine connaissance de la réalité – *donc* , P implique non-P. Il existe un processus allant de P à sa négation, que la logique nous enjoint de reconnaître. Refuser cette exigence reviendrait à commettre la même contradiction. Il ne s'agit pas d'une démonstration circulaire ou à l'infini, mais d'un appel à l'honnêteté, condition indispensable à tout dialogue.

Le fait que toutes les propositions assertives affirment quelque chose est un aspect du principe d'identité. Le principe de non-contradiction se manifeste dans l'argument suivant : toute assertion affirme quelque chose ; P est une assertion ; par conséquent, P affirme ; d'où, si P nie affirmer quelque chose, P implique non-P. Le principe du tiers exclu est également implicite ici, dans la conscience que nous n'avons d'autre choix que de rejeter fermement P.

Contradictions déguisées . Les contradictions apparaissent dans le discours sous de nombreuses formes. Elles ne sont pas toujours manifestes, mais peuvent être dissimulées dans le fait même de formuler une affirmation ou dans les critères de jugement utilisés.

Une affirmation peut être paradoxale car elle **implique intrinsèquement** sa propre contradiction, même si elle ne semble pas, de prime abord, incohérente. Cette

implication n'est pas toujours formelle et requiert une compréhension du sens des termes employés. Cette forme d'auto-contradiction indirecte est appelée « sophisme du concept volé » [10].

Par exemple, l'affirmation sceptique « Je ne sais rien » peut être rejetée comme contradictoire, car dès que quelqu'un la formule – quelqu'un *qui comprend et entend le sens* des termes « je », « savoir » et « rien » – cela constitue en soi une preuve absolue que la personne concernée « sait » *quelque* chose, ce qui démontre que l'affirmation initiale (d'ignorance totale) est inévitablement contradictoire et donc nécessairement fausse.

Ainsi, dans ce genre de cas, l'implication tacite est que l'un des termes employés (« ne rien savoir ») inclut implicitement l'acte en question (« savoir que je ne sais rien »), ce qui contredit l'affirmation explicite. (Il va sans dire que reformuler cette affirmation en « je ne *sais* rien » ne change en rien ses présupposés.)

Les exemples de telles contradictions intrinsèques sont innombrables. Dire « Je n'ai rien à dire » revient à dire quelque chose. Affirmer « Nous n'avons aucun souvenir » est contradictoire, car chaque terme de cette affirmation présuppose un mot, un concept et des expériences vécues dont se souviennent à la fois le locuteur et l'auditeur. Un exemple courant et amusant est : « Je ne parle pas un mot d'anglais ! »

Une autre forme importante d'incohérence inconsciente est le recours à **un double standard** . Cela consiste à appliquer des critères d'évaluation moins rigoureux à son propre discours qu'à celui de ses adversaires intellectuels. Nombre de discours philosophiques, et notamment politiques et religieux, ont recours à cette méthodologie inéquitable.

---

10 Par Ayn Rand et (je crois) Nathaniel Branden.

La contradiction inhérente à un double standard apparaît clairement dès lors qu'on prend du recul et qu'on considère l'ensemble des connaissances et de la méthodologie de celui qui l'applique. Dans cette perspective plus large, l'utilisateur d'un double standard se montre manifestement incohérent avec lui-même, même si son discours, pris individuellement, peut sembler cohérent en apparence.

Des philosophies entières peuvent reposer sur de tels raisonnements fallacieux. Par exemple, le phénoménalisme pose comme norme générale la limitation de la connaissance aux données sensorielles, sans autoriser d'extrapolations à des objets matériels externes supposés ; or, il ne remet pas en question ses propres déductions en appliquant cette même norme rigide.

Ce sophisme peut se manifester de deux manières : soit on applique des critères laxistes à son propre discours tout en semblant appliquer des normes universelles à celui de ses adversaires ; soit on semble appliquer des normes universelles à soi-même tout en élaborant des normes excessivement strictes pour les autres. On peut *s'exempter* des règles logiques habituelles ou imposer *à autrui des exigences logiques inhabituelles* .

Dans les deux cas, celui qui applique deux poids, deux mesures contrevient à l'exigence d'uniformité propre à la logique. La raison suppose que tous les êtres humains se situent épistémologiquement sur un même plan. L'équité relève du bon sens. Il convient de se convaincre soi-même et convaincre autrui par l'expérience et la logique, et non par la manipulation sophistique ou l'autorité.

Les critères d'évaluation doivent être justes et universels ; tout discours doit être traité de manière égale. Si des différences sont défendues, elles doivent être justifiées de façon convaincante. Le principe d'égalité implique certes

une généralisation ; mais la charge de la preuve incombe à toute particularisation proposée de ce principe.

Un exemple de double standard est le recours au relativisme culturel. On peut chercher à rationaliser des idées ou des processus de pensée contraires à la raison ordinaire en prétendant qu'ils appartiennent à un cadre culturel différent. Une telle tolérance semble, de prime abord, amicale et ouverte d'esprit, mais elle est proposée sans une pleine prise en compte de ses implications humaines et épistémologiques négatives.

3. La cohérence est naturelle

Il est important de réaffirmer ici le principe selon lequel la cohérence est naturelle , tandis que l'incohérence est exceptionnelle.

Certains logiciens modernes ont introduit la notion de « preuve de cohérence », mais cette notion est erronée. La cohérence est l'état naturel de la connaissance ; elle ne requiert aucune preuve (déductive) et nous sommes incapables d'en fournir une, car ce serait mettre la charrue avant les bœufs. La seule « preuve » possible de la cohérence est l'absence d'incohérence. La cohérence est un postulat inductif, très rarement remis en cause. On peut et on doit supposer que toute notre connaissance est cohérente, sauf raison de croire le contraire.

En résumé : l’harmonie règne généralement sans que l’on s’en aperçoive, tandis que des conflits éclatent parfois à notre grande surprise. On peut se demander si ce principe est compatible avec celui défendu ici, selon lequel les négatifs ne sont jamais des objets de connaissance en soi, mais n’existent que par négation des positifs correspondants. Notre principe selon lequel la cohérence est tenue pour acquise semble impliquer que nous avons

parfois des intuitions logiques sur *l'* incohérence, c'est-à-dire sur quelque chose de négatif !

Pour résoudre ce problème, il convient de souligner à nouveau la distinction entre l'expérience pure et les *interprétations* que nous, consciemment ou non, intégrons à nos expériences. Généralement, presque aussitôt que nous vivons une expérience, nous commençons à l'interpréter, en la reliant dynamiquement à l'ensemble de nos connaissances. Chaque expérience génère presque inévitablement en nous des suites d'associations, d'explications, etc.

Les contradictions que nous rencontrons parfois dans notre savoir ne concernent pas nos expériences pures (qui sont nécessairement harmonieuses, puisqu'elles coexistent – et l'on pourrait ajouter, de façon tout à fait harmonieuse). *Nos contradictions sont nécessairement des contradictions entre une interprétation et une expérience pure, ou entre deux interprétations* . À proprement parler, les contradictions ne révèlent pas de difficultés dans les données brutes de la connaissance, mais seulement dans les hypothèses que nous avons conçues à leur sujet.

Les contradictions sont donc imputables à la raison, et non à l'expérience. Cela ne signifie pas que la raison soit nécessairement défaillante, mais seulement qu'elle est faillible. Les contradictions ne doivent pas être perçues comme des preuves tragiques de notre ignorance et de notre stupidité, mais comme des indicateurs utiles d'une mauvaise interprétation, nécessitant une réinterprétation. Ces indicateurs constituent précisément l'un des principaux outils dont dispose la faculté de raison pour contrôler la qualité de nos croyances. La résolution d'une contradiction n'est autre qu'une nouvelle interprétation.

Comment savons-nous que deux théories, ou une théorie et des données brutes, sont « en contradiction » ? La

question reste ouverte. On l'élude un peu facilement en invoquant la « logique ». Cette logique est en partie « expérientielle », puisqu'elle repose sur l'examen des preuves et des doctrines disponibles. Mais elle n'est manifestement pas entièrement empirique et fait intervenir des facteurs abstraits. La « contradiction » est, après tout, une abstraction. Je crois que la réponse à cette question se trouve en grande partie dans l'analyse psychologique de la négation.

Il existe une intuition profonde de l' existence *d'intentions contradictoires* . Ainsi, la « prise de conscience logique » de cette incohérence ne porte pas essentiellement sur une perception négative (une non-cohérence), mais sur une perception positive (l'expérience intuitive d'un conflit d'intentions). Bien que le mot « incohérence » soit précédé d'une connotation négative, il évoque une réalité empiriquement positive : une tension ressentie entre deux thèses ou entre une thèse et des données.

C'est pourquoi affirmer que « la cohérence est présumée, jusqu'à ce que l'incohérence soit un jour constatée » est cohérent avec notre prétention selon laquelle « les négations ne sont pas purement empiriques ». (Remarquons d'ailleurs que nous n'avons pas « prouvé » ici la cohérence, mais l'avons simplement *retrouvée* en clarifiant les thèses en jeu.)

L'analyse précédente éclaire également la manière dont le principe de non-contradiction se manifeste concrètement. Il ne s'agit pas de trancher des expériences en tant que telles, mais de s'intéresser à des notions de connaissance plus abstraites. Pour le comprendre pleinement, il est indispensable de saisir les intentions sous-jacentes. Une analyse similaire peut être proposée pour expliquer le principe du tiers exclu.

Dans ce dernier cas, nous insisterions sur le fait que (par le principe d'identité) « les choses sont ce qu'elles sont,

quoi qu'il arrive ». On ne peut dire que les choses ne sont *ni* ceci *ni* la négation de cela, car de telles caractérisations sont négatives (et, par conséquent, doublement négatives) – et ne peuvent donc constituer ni être considérées comme une expérience positive. De telles situations renvoient à des incertitudes *chez le sujet connaissant* , qu'il est appelé à combler. Elles ne sauraient être proclamées savoir définitif (comme certains sophistes modernes ont tenté de le faire), mais doivent être considérées comme des postures temporaires dans la quête du savoir.

3. Clarifier la négation

Extrait de Ruminations 9.

1. Négation en adduction

Les concepts et les théories sont des constructions hypothétiques. Ils ne peuvent (généralement) être prouvés (définitivement, une fois pour toutes), mais seulement confirmés de manière répétée par l'expérience. C'est le côté positif de l'adduction : apporter des preuves à l'appui de constructions rationnelles. Cet aspect positif est bien sûr indispensable, car sans preuves concrètes, une abstraction n'est rien de plus qu'une pure invention, une spéculation hasardeuse. Plus nous apportons de preuves, plus notre concept ou notre théorie est fiable.

Mais, comme Francis Bacon l'a compris, l'explication de l'adduction proposée jusqu'ici ne lui rend pas justice. Tout aussi important que l'aspect positif de la fourniture de preuves est son aspect négatif : le rejet des hypothèses dont les prédictions contredisent l'expérience. Comme il l'a souligné, même si une hypothèse a reçu de nombreuses confirmations, *une seule* prédiction erronée suffit à la rejeter.

En d'autres termes, le raisonnement par adduction consiste à sélectionner l'hypothèse la plus probable parmi plusieurs explications (déjà conçues ou à concevoir) d'un fait. Chacune d'elles peut être étayée par de nombreux « exemples positifs » (c'est-à-dire des preuves empiriques) ; et tant qu'elles restent plausibles, nous pouvons privilégier celles qui en possèdent le plus.

Cependant, c'est par l' *élimination successive* des propositions qui s'avèrent réfutées (c'est-à-dire contredites par des preuves empiriques) que nous progressons de manière décisive dans notre connaissance conceptuelle et théorique.

Tout ce qui précède est désormais bien connu et n'a pas besoin d'être davantage explicité. Cette théorie du raisonnement inductif a démontré son immense succès à l'époque moderne, constituant le fondement de la méthode scientifique.

Mais à y regarder de plus près, la question n'est pas aussi simple et directe qu'elle le paraît au premier abord !

Prenons l'exemple de la question de la présence ou non d'eau sur Mars. Il semblerait que la proposition « Il y a de l'eau sur Mars » soit bien plus facile à démontrer par induction que la proposition contradictoire « Il n'y a pas d'eau sur Mars ». Ces deux propositions sont des hypothèses.

L'hypothèse positive serait en partie confirmée si l'on découvrait, grâce à certains instruments et à distance, des indices sérieux de la présence d'eau ; elle serait confirmée de manière plus solide si un échantillon martien, rapporté sur Terre, s'avérait contenir de l'eau après analyse. Dans les deux cas, la présence d'eau sur Mars resterait, dans une certaine mesure (même infime), incertaine, car nos hypothèses instrumentales pourraient être remises en question ultérieurement, ou l'échantillon rapporté pourrait s'avérer contaminé lors du voyage. Néanmoins, une quasi-certitude est envisageable.

L'hypothèse négative, en revanche, est beaucoup plus difficile à prouver par l'expérience. On peut aisément la retenir dans la mesure où l'hypothèse positive n'a pas encore été largement confirmée. Autrement dit, tant que nous n'avons pas trouvé de preuves en faveur de l'hypothèse positive (c'est-à-dire la présence d'eau sur

Mars), il vaut mieux privilégier l'hypothèse négative. Mais cette dernière n'est fiable que dans la mesure où nous avons tenté, sans succès, de confirmer la première. Si nous cherchions de l'eau avec acharnement et par tous les moyens possibles, sans en trouver, nous pouvons raisonnablement supposer qu'il n'y a pas d'eau.

Ainsi, dans notre exemple, la thèse négative est en réalité *plus difficile* à établir que la thèse positive. Elle *repose sur une généralisation* , un passage de la pensée « Où que nous ayons cherché de l'eau sur Mars, ***nous n'en avons trouvé aucune*** » à « ***Il n'y a pas*** d'eau sur Mars ». Cependant, il est important de noter qu'il reste concevable qu'une goutte d'eau soit découverte un jour ailleurs sur Mars, des siècles après que nous ayons conclu à son absence.

Compte tenu de cette analyse, il apparaît clairement que le rasoir de Bacon, selon lequel « ce qui importe, c'est l'exemple négatif », est quelque peu simpliste. Il suppose qu'un exemple négatif est aussi accessible (voire plus accessible) qu'un exemple positif, ce qui n'est pas toujours le cas.

En pratique, un cas négatif peut être inductivement plus éloigné qu'un cas positif. Partant de ce constat, une question se pose : le cas négatif est- *il jamais* plus accessible empiriquement que (ou même aussi accessible que) le cas positif ? Autrement dit, *quand* la formulation de l'induction par Bacon entre-t-elle réellement en jeu ?

Si l'on considère des exemples historiques majeurs de rejet de théories, nos doutes peuvent persister. Par exemple, la mécanique newtonienne a prévalu pendant des siècles, jusqu'à ce qu'elle soit remise en question par la découverte de la constance de la vitesse de la lumière (qui a donné naissance à la théorie de la relativité) et, plus tard, par la découverte de divers phénomènes subatomiques (qui ont donné naissance à la mécanique

quantique). Dans cet exemple, les « cas négatifs » étaient en réalité des « cas positifs » : leur seul aspect « négatif » résidait dans leur réfutation de la vision newtonienne du monde !

Ces réflexions m'ont amené à penser que la « négation » évoquée par Bacon ne s'entend que *par rapport* à une abstraction choisie. Son rasoir ne saurait être interprété comme une apologie de la négation absolue. À y regarder de plus près, on constate que les données utilisées pour nier une idée sont, par essence, positives. Un examen plus approfondi de la nature de la négation s'impose donc.

2. Phénomènes positifs et négatifs

On a toujours considéré qu'il existait une différence entre un terme positif et un terme négatif. C'est d'ailleurs pourquoi les logiciens les ont nommés différemment. Mais ces mêmes logiciens ont eu du mal à exprimer clairement cette différence. Pourtant, il existe des différences phénoménologiques significatives entre les phénomènes positifs et négatifs.

a. Le monde matériel et mental concret est manifestement composé uniquement de phénomènes particuliers positifs, dont certains sont perceptibles (par les sens ou par l'esprit). Ces phénomènes existent au moins comme apparences, certaines paraissant réelles et d'autres illusoires. Il s'agit là d'une vérité phénoménologique, épistémologique et ontologique évidente.

Affirmer qu'un phénomène est « particulier » revient à reconnaître sa limitation spatio-temporelle. Il existe, certes, mais il est fini et éphémère, c'est-à-dire manifestement caractérisé par la diversité et le changement.

Nous ne faisons généralement l'expérience d'aucune réalité concrète qui s'étende uniformément à l'infini et à l'éternité (bien qu'une telle totalité de l'existence puisse exister, et que certains mystiques prétendent en avoir conscience lors de méditations profondes, la qualifiant d'« éternel présent »). Nous considérons généralement certaines choses comme étant ainsi omniprésentes. L'« existence » est perçue comme le substrat de tous les êtres existants ; l'« univers » désigne la somme totale de tous les êtres existants ; et nous concevons l'« espace-temps » comme définissant l'extension de tous les êtres existants. Mais seule l'« existence » peut être classée comme une expérience (une qualité présente dans tous les êtres existants) ; l'« univers » et l'« espace-temps » doivent être considérés comme des abstractions.

Cependant, les limites des choses particulières sont perceptibles sans qu'il soit nécessaire de nier ce qui se trouve au-delà, simplement en raison de la concentration variable de la conscience, c'est-à-dire de la direction du foyer de l'attention. Autrement dit, bien que « désigner » un phénomène positif (par exemple, pour le nommer) requière une certaine négation (nous voulons dire « ceci, mais pas cela »), on peut percevoir les limites de ce phénomène indépendamment de toute négation.

b. Les phénomènes négatifs (et de même les abstractions, qu'elles soient positives ou négatives), en revanche, dépendent pour leur existence d'un sujet/agent – une « personne » connaissante (ou synonyme : un soi, une âme ou un esprit) dotée de conscience et de volonté, à la recherche d'un phénomène positif mémorisé ou imaginé et incapable de le percevoir (ou, dans le cas des abstractions, comparant et contrastant des particuliers).

Ainsi, les phénomènes particuliers négatifs (et plus généralement, les abstractions) possèdent une existence particulière, plus « relative ». Ils ne sont pas aussi

indépendants du Sujet que les phénomènes particuliers positifs. Cela ne signifie pas qu'ils soient, au sens kantien, « a priori », « transcendantaux » ou purement « subjectifs », mais plutôt qu'ils constituent des potentialités ontologiques qui ne se réalisent que dans le contexte de la cognition (rationnelle).

Une autre forme d'expérience est nécessaire à cette réalisation : l'expérience de soi du sujet, sa connaissance intuitive de ses cognitions et de ses volitions. Cette expérience, étant immédiate, peut être positive ou négative sans difficulté logique. Le sujet raisonne inductivement comme suit :

Je recherche X ;

Je ne trouve pas X ;

Par conséquent, X « *n'est pas* » là.

La conclusion négative peut être « vraie » ou « fausse » , tout comme une perception ou une conclusion positive. Elle est vraie dans la mesure où les prémisses sont vraies – c'est-à-dire que la recherche alléguée de X a été diligente (intelligente, imaginative, bien organisée, attentive et approfondie), et que l'échec allégué de la recherche de X n'est pas malhonnête (un mensonge destiné à tromper soi-même ou autrui).

Il est donc juste d'affirmer que, contrairement à certains termes positifs, les termes négatifs ne reposent jamais *uniquement* sur la perception ; ils impliquent *nécessairement* un processus de pensée – la projection mentale préalable, ou du moins l'intention, du terme positif qu'ils nient.

Cette vérité épistémologique reflète bien une vérité ontologique : l'« absence » de phénomènes est dépourvue d'aspects phénoménaux. Un « non » n'est pas un « oui ».

Il est important de bien distinguer logiquement « **ne pas percevoir X** » et « **percevoir l'absence de X** ». Nous

n’avons pas d’expérience directe de cette dernière, mais ne pouvons l’affirmer qu’indirectement, par *inférence inductive* (ou extrapolation) à partir de la première. Dans le cas d’un phénomène positif, un tel raisonnement est superflu : on peut souvent percevoir X directement, et on le perçoit effectivement.

Supposons que nous représentions les propositions par un carré d'opposition (en les nommant par analogie avec les positions classiques) : « Je perçois X » (A), « Je ne perçois pas X » (I), « Je ne perçois pas X » (E), « Je ne perçois pas X » (O). La proposition A est connaissable par l'expérience, tandis que la proposition I ne l'est peut-être que par déduction. En revanche, la proposition E n'est pas connaissable par l'expérience, mais seulement par généralisation inductive à partir de la proposition O (qui, elle, repose sur l'expérience).

3. La négation est secondaire

La négation est un pilier de la logique déductive et inductive, et exige une analyse approfondie. Il nous faut comprendre que les termes négatifs sont fondamentalement différents des termes positifs si nous voulons commencer à sonder la nature de la logique. L'observation suivante me semble cruciale pour une telle analyse :

> ***Nous pouvons faire l'expérience de quelque chose de positif sans avoir d'abord fait l'expérience (ou pensé) à sa négation, mais nous ne pouvons pas faire l'expérience de quelque chose de négatif sans avoir d'abord pensé à (et donc avoir déjà fait l'expérience, même partielle) de l'aspect positif correspondant.***

a. La cognition, dans sa forme la plus simple, est la perception. Nos perceptions portent toujours *sur des particularités positives* . Le contenu de nos cognitions les

plus élémentaires est constitué de phénomènes visuels, auditifs, olfactifs, gustatifs et tactiles, ainsi que d'autres sensations corporelles qui semblent résulter de l'interaction de nos organes sensoriels avec la matière – ou d'équivalents mentaux de ces phénomènes qui semblent émerger de la mémoire d'expériences sensorielles, ou encore de recombinaisons imaginaires de tels souvenirs supposés.

Une expérience positive peut être vécue directement et passivement. Il suffit, en quelque sorte, de se détendre et d'observer avec réceptivité ce qui se présente à notre champ de vision ou d'écoute. C'est le principe de la méditation. Nul besoin de penser activement à autre chose (se souvenir, visualiser ou conceptualiser) pour vivre une telle expérience positive. Bien sûr, cette observation peut être compliquée par des pensées (préverbales ou verbales), mais il est possible, dans certains cas, de vivre une expérience pure. Il faut logiquement l'admettre si l'on considère que les concepts reposent sur des perceptions.

b. Dans le cas des *particularités négatives* , la situation est radicalement différente. Une particularité négative n'a *pas* de contenu phénoménal spécifique, mais se définit *entièrement* par l'« absence » des contenus phénoménaux qui constituent une particulière positive. Si j'observe mon environnement matériel ou mental, je verrai toujours des phénomènes présents. L'absence d'un phénomène n'est perceptible que si l'on pense d'abord à ce phénomène positif et que l'on se demande s'il est présent.

Il est exact d'affirmer que notre constat d'absence reflète une vérité ou un fait empirique – mais il est indéniable que nous ne remarquerions pas le négatif sans avoir préalablement envisagé le positif. La connaissance négative est donc nécessairement (par nécessité logique) plus indirecte et active. Elle demeure (dans le meilleur des cas) parfaitement ancrée dans l'expérience – mais une

telle expérience négative requiert un processus rationnel (verbal ou autre).

Pour ressentir une expérience négative, je dois d'abord imaginer (me souvenir ou inventer) une certaine expérience positive ; puis je dois regarder autour de moi et voir (ou entendre ou autre) si cette image correspond ou non à mon expérience actuelle ; et ce n'est qu'alors (si ce n'est effectivement pas le cas) que je peux conclure avoir « ressenti » une expérience négative.

Réfléchir à X peut être considéré comme se positionner d'un point de vue permettant, en quelque sorte, d'expérimenter le non-X. Si l'on ne porte pas d'abord son attention sur X, il est impossible d'expérimenter la négation de X. On peut certes faire l'expérience de toutes sortes de choses étranges et merveilleuses, mais pas spécifiquement du non-X.

De cette réflexion, nous pouvons dire que si les affirmations peuvent être vécues, les négations sont par nature des actes rationnels (faisant appel à l'imagination, à l'expérience et à l'intention). Une négation implique nécessairement une pensée : la pensée du positif correspondant (l'élément imaginatif), la vérification de sa présence ou de son absence (l'élément expérientiel) et la conclusion rationnelle de « négation » (l'élément intentionnel).

c. Le processus de négation peut impliquer des mots, mais ce n'est pas obligatoire.

Supposons que je fasse une expérience momentanée de perceptions visuelles, auditives, etc., et que je désigne cette expérience positive particulière par « X ». Le *contenu de la conscience* sur lequel je fonde le terme X est un ensemble spécifique d'expériences phénoménales positives, c'est-à-dire des perceptions physiques et/ou mentales. Chaque fois que je peux parler de ce X, je *me représente mentalement* un objet d'une certaine couleur et

d'une certaine forme qui se déplace d'une certaine manière, émet certains sons, etc.

La négation d'un terme aussi simple, « non-X », est tout à fait différente. Cette dernière n'est définissable par aucune perception spécifique ; elle *ne renvoie à aucune qualité perceptible* . Elle ne peut être identifiée aux phénomènes positifs qui se trouvent présents en l'absence de ceux qui constituent X. Ainsi, à proprement parler, non-X n'est définissable que par la « négation » de X.

Il est important de noter qu'il serait inexact d'affirmer (sauf a posteriori) que « non-X » désigne toutes les expériences autres que X (telles que Y, Z, A, B, etc.), car lorsque je recherche X ici et maintenant et que je ne le trouve pas, je me réfère uniquement à l'expérience présente dans mon champ d'action actuel, et non à toutes les expériences possibles de ce type. On ne qualifierait pas une situation dépourvue de X de « non-X » *sans penser à X* ; on la qualifierait plutôt positivement (par exemple, par « Y », « Z », etc.).

Ainsi, nous pouvons nommer (ou penser sans mots à) quelque chose de concret « X », *après* avoir expérimenté les phénomènes qui le constituent ; mais dans le cas de « non-X », nous devons nécessairement en évoquer le nom (ou une pensée sans mots) *avant* de l'expérimenter.

Le « non-X » est donc déjà un concept plutôt qu'une perception, même lorsque « X » renvoie à une simple perception (et d'autant plus lorsque « X » implique une abstraction, comme c'est généralement le cas). Le concept de « non-X » est d'abord construit hypothétiquement, puis confirmé par la tentative infructueuse de revivre X.

En bref, la négation – même au niveau le plus perceptif – implique un processus d'adduction. Elle ne se réduit jamais à une simple expérience. Un terme négatif ne vise jamais la simple perception d'une chose négative, mais consiste en une hypothèse assortie d'une confirmation

perceptive. La négation est toujours à la fois conceptuelle et perceptive.

On ne peut réfuter une théorie avant qu'elle ne soit formulée ; de même, on ne peut constater l'absence de X sans avoir d'abord pensé à X.

4. La négation est une intention

Or, le mot « non » ne renvoie à aucune expérience phénoménale spécifique. La négation n'a ni couleur, ni forme, ni son, ni odeur, ni goût, ni sensation particulière, qu'elle soit réelle ou illusoire ! Qu'est-ce donc ? Je propose la réponse suivante :

La négation, en tant que telle, renvoie à un « acte mental » – ou plus précisément, à un acte de volonté (ou encore plus précisément, de velleité) d'un sujet conscient. Plus précisément, la négation est *une intention* . Notez que notre volonté de nier est en elle-même *un acte positif* , même si notre intention, par là même, est de nier autre chose.

La négation exprime bien une expérience : l'« échec » de trouver ce que l'on a cherché. On poursuit volontairement un résultat cognitif (la perception d'un phénomène positif), mais celui-ci reste insatisfait (cette expérience est qualitativement une forme de souffrance, mais il s'agit toujours d'une intention positive, notons-le) ; de là, nous marquons mentalement (ou plus précisément, intentionnellement) la chose comme « absente », c'est-à-dire que nous construisons une idée de « négation » de la chose recherchée.

Ainsi, la négation n'est *pas un phénomène* (une perception physique ou mentale), *mais une intuition* (un acte de volonté au sein du Sujet connaissant). Il est important de noter qu'ici, « intuition » désigne la connaissance de soi du Sujet de la conscience et de l'Agent de la volition. Il

s'agit d'une expérience *non phénoménale* . Cette expérience de soi est immédiate : aucune distance spatio-temporelle ne nous sépare de cette expérience.

Lorsqu'un sujet nie la présence d'un phénomène matériel ou mental, après l'avoir cherché dans son expérience sans le trouver, cette « négation » consiste en un acte d'intention particulier. Cette intention est ce que l'on appelle « négation » ou « rejet d'une hypothèse ». Elle se produit chez le sujet, bien qu'elle porte sur l'objet.

Cette intention n'est toutefois pas un acte arbitraire. Si tel était le cas, elle serait purement subjective. Cet acte (dans le meilleur des cas) demeure suffisamment dépendant de la perception pour être considéré comme « objectif ». Le sujet doit toujours observer et constater la présence de X ; si cette expérience positive ne correspond pas à son observation empirique, il conclut à l'absence de X.

En effet, une négation initiale peut, après un examen plus approfondi, s'avérer erronée ; autrement dit, il nous arrive de penser que quelque chose « n'existe pas », puis, après des recherches complémentaires, de constater qu'au contraire, « existe ». Ainsi, cette théorie de la négation ne doit pas être interprétée comme affirmant que le fait de nier quelque chose le rende tel. La négation est régie par les principes de l'adduction : elle repose sur une apparence crédible tant qu'elle est confirmée, mais qui peut être ultérieurement infirmée.

On peut a posteriori parler d'une absence objective, mais on ne peut définir pleinement l'« absence » autrement que comme « non-présence », et le préfixe « non- » ici ne désigne pas un phénomène, mais une intention. L'« absence » est certes vécue, mais elle est *imperceptible* sans que le sujet ne pose préalablement la question : « X est-il présent ? »

L’absence n’est donc pas produite par le sujet, mais rendue perceptible par sa vaine quête de présence. Car,

pour le répéter, le non-X n’est pas vécu comme un contenu spécifique de la conscience, mais comme une incapacité persistante à percevoir les phénomènes positifs particuliers qui définissent X pour nous.

Bien que nous n'ayons conscience directement que des réalités apparentes, nous pouvons inférer par induction les réalités non apparentes à partir de l'expérience selon laquelle les apparences sont éphémères et peuvent changer. Sur cette base, nous considérons les catégories « existence » et « apparence » comme inégales, la première étant plus large que la seconde. De même, nous inférons par induction l'« absence objective » à partir du constat de « avoir cherché sans trouver », même si nous n'avons pas d'accès direct à la première, mais seulement un accès indirect par extrapolation à partir de la seconde. Une telle inférence est valide, avec un degré de probabilité proportionnel à la diligence raisonnable dont nous faisons preuve.

Pour ces raisons, je considère l'acte de négation comme une clé essentielle pour comprendre la nature et le statut de la logique. La négation est si fondamentale à la raison, un fait épistémique si crucial, qu'elle ne peut être réduite à autre chose.

On peut la décrire *approximativement* comme l'intention de « rayer » (sous l'influence d'une raison ou d'une autre) l'élément proposé de notre liste mentale des existants. Mais cela risque de paraître une définition circulaire, voire une répétition de la même chose à l'aide de synonymes. Il est évident qu'on *ne peut parler de négation sans y recourir* . Dès lors, il est préférable de considérer l'acte de négation comme un concept fondamental de la logique.

Remarque en passant : la théorie actuelle de la négation apporte à la biologie une distinction intéressante concernant les animaux rationnels.

Les êtres sensibles dépourvus de cette faculté de négation ne peuvent réagir qu'au présent, tandis que dès l'apparition de cette faculté chez un organisme (comme ce fut le cas chez l'espèce humaine), celui-ci peut mentalement transcender l'ici et maintenant. Un animal purement sensoriel se contente de réagir aux événements actuels, tandis qu'un être humain peut craindre les dangers et s'y préparer.

Dès l'apparition de la faculté de négation, l'esprit peut commencer *à abstraire, à concevoir des alternatives et à formuler des hypothèses* . La mémoire et l'imagination sont nécessaires pour projeter une idée positive, mais l'intention de nier est également indispensable pour rejeter les projections inadéquates. Sans cette capacité critique, nos fantasmes nous entraîneraient rapidement dans des situations destructrices.

4. Paradoxe éclaircissant

De Future Logic 31.

L'étude des paradoxes constitue un domaine fondamental de la logique, car elle nous offre un outil puissant pour établir certaines des certitudes les plus essentielles de cette science. Elle nous permet de revendiquer pour l'épistémologie et l'ontologie le statut de véritables sciences, et non de simples digressions spéculatives. Cette doctrine élégante peut être considérée comme faisant partie de l'étude des axiomes.

1. Incohérence interne

Considérons la forme hypothétique « Si P, alors Q », qui est un élément essentiel du langage logique. Elle a été définie comme « P et non Q est une conjonction impossible ».

Il est axiomatique que la conjonction d'une proposition P et de sa négation nonP est impossible ; par conséquent, une proposition P et sa négation nonP ne peuvent être simultanément vraies. Un corollaire évident, obtenu en considérant nonP comme la proposition à l'étude au lieu de P, est que la conjonction d'une proposition nonP et de sa négation non-nonP est impossible ; par conséquent, une proposition P et sa négation nonP ne peuvent être simultanément fausses.

Ainsi, le principe d'identité peut s'énoncer ainsi : « Pour toute proposition, “Si P, alors P” est vrai, et “Si non P, alors non P” est vrai. » Les principes de contradiction et du tiers exclu peuvent s'énoncer ainsi : « Pour toute

proposition, “Si P, alors non-non P” est vrai (P et non P sont incompatibles), et “Si non-non P, alors P” est vrai (non P et P sont exhaustifs). »

Considérons maintenant les propositions paradoxales « Si P, alors non P » ou « Si non P, alors P ». De telles propositions apparaissent à première vue comme des antinomies manifestement impossibles, nécessairement fausses.

Mais examinons de plus près leur signification. La première affirmation stipule « P et (non non)P est impossible », ce qui signifie simplement « P est impossible ». La seconde stipule « nonP et non P est impossible », ce qui signifie simplement « nonP est impossible ». Présentées sous cette forme, ces affirmations ne semblent plus antinomiques ! Elles nous indiquent simplement que la proposition P, ou nonP, selon le cas, contient un défaut intrinsèque, une contradiction interne, une propriété d'auto-négation.

Il ressort de ce qui précède qu'il peut exister des propositions logiquement autodestructrices et qui soutiennent logiquement leur propre négation. Définissons donc les points suivants : une proposition est ***contradictoire*** si elle se nie elle-même, c'est-à-dire si elle implique sa propre négation. Une proposition est donc ***évidente*** si sa négation est contradictoire, c'est-à-dire si elle est impliquée par sa propre négation.

Ainsi, la proposition « Si P, alors non P » nous informe que P est contradictoire (et donc logiquement impossible), et que non P est évident (et donc logiquement nécessaire). De même, la proposition « Si non P, alors P » nous informe que non P est contradictoire, et que P est évident.

L'existence de paradoxes n'indique en rien un défaut formel. Le *paradoxe* , c'est-à-dire la proposition hypothétique elle-même, n'est pas antinomique. Il peut être vrai ou faux, comme toute autre proposition.

Admettons-le vrai : c'est sa thèse antérieure qui est antinomique, et donc fausse, puisqu'elle se contredit ; la thèse conséquente est alors vraie.

Si la proposition paradoxale « Si P, alors non P » est vraie, alors sa contradiction « Si P, non-alors non P », signifiant « P n'est pas impossible », est fausse ; et si cette dernière est vraie, la première est fausse. De même, « Si non P, alors P » peut être contredit par « Si non P, non-alors P », signifiant « non P n'est pas impossible ».

Les deux paradoxes « Si P, alors non P » et « Si non P, alors P » sont contradictoires, puisqu'ils impliquent la nécessité d'incompatibles, respectivement non P et P. Ainsi, bien que de telles propositions prises individuellement ne soient pas antinomiques, le double paradoxe, situation où ces deux propositions paradoxales seraient vraies simultanément, est inacceptable pour la logique.

Contrairement aux propositions hypothétiques positives, les propositions hypothétiques négatives ne permettent pas d'exprimer de paradoxes. Les propositions « Si P, non-alors P » et « Si non-P, non-alors non-P » sont *dénuées* de sens, logiquement inconcevables et ne sont jamais vraies. Il est important de noter que de telles propositions sont formellement fausses. Puisqu'une forme comme « Si P, non-alors Q » est définie par rapport à une conjonction positive comme « {P et non-Q} est possible », on ne peut, sans antinomie, substituer P à Q (pour dire « {P et non-P} est possible ») ni non-P à P et Q (pour dire « {non-P et non-non-P} est possible »).

Il s'ensuit que la proposition « si P, alors non P » n'implique pas la forme minuscule « si P, non-alors P », et que la proposition « si non P, alors P » n'implique pas la forme minuscule « si non P, non-alors non P ». Autrement dit, dans le contexte du paradoxe, les propositions

hypothétiques se comportent anormalement, et non comme des formes basées sur la contingence.

Cela ne devrait pas nous surprendre, car ce qui est contradictoire est logiquement impossible et ce qui est évident est logiquement nécessaire. Les propositions paradoxales impliquant des thèses et des antithèses incontingentes, elles sont soumises aux lois propres à ce type de fondement.

Les implications et la cohérence de tout cela seront examinées prochainement.

2. Le sophisme du concept volé

Les propositions paradoxales se rencontrent bel et bien dans la pratique ; de plus, elles nous apportent des résultats très significatifs. En voici quelques exemples :

- Le déni, voire le doute, sur les lois de la logique dissimule un appel à ces mêmes axiomes, impliquant que c'est le déni plutôt que l'affirmation qu'il faut croire ;
- Le déni de la capacité de l'homme à connaître objectivement une quelconque réalité constitue en soi une prétention à la connaissance d'un fait de la réalité ;
- Le déni de validité de la perception, du pouvoir conceptuel ou du raisonnement de l'homme, toutes ces affirmations sceptiques présupposent l'utilisation et la confiance dans les facultés mêmes mises en doute ;
- Le refus, par principe, de toute généralisation, nécessité ou absolu, constitue en soi une prétention à une vérité générale, nécessaire et absolue.
- Le déni de l'existence des « universaux » ne résout pas en soi le problème des universaux, puisqu'il fait appel à certains d'entre eux, à savoir que les « universaux » « n'existent pas » et « n'existent pas ».

On trouvera davantage de détails sur ces paradoxes et d'autres encore disséminés dans le texte. Ainsi, la mise au jour des paradoxes constitue une technique logique importante et fréquemment employée. L'écrivaine Ayn Rand a beaucoup insisté sur cette méthode de réfutation des philosophies sceptiques, en démontrant qu'elles *font implicitement appel à des concepts qu'elles prétendent nier explicitement* ; elle a nommé ce phénomène « le sophisme du concept volé ».

Pour comprendre le fonctionnement des paradoxes, il est utile de les considérer dans le contexte des dilemmes. Une proposition P, qui semble aller de soi, pourrait s'énoncer ainsi : « Que l'on affirme ou que l'on nie P, elle est vraie ». Une vérité absolue est une chose qui se révèle vraie quelles que soient nos hypothèses initiales.

On peut formuler cela comme un argument constructif dont la branche gauche est la proposition axiomatique de l'identité de P avec lui-même, et la branche droite le paradoxe de l'auto-contradiction de nonP ; la prémisse mineure est l'axiome de contradiction absolue entre les antécédents P et nonP ; et la conclusion, le conséquent, la vérité absolue de P.

> Si P, alors P — et — si non P, alors P
>
> mais soit P soit nonP
>
> donc, P.

Une version destructive peut tout aussi bien être formulée, en utilisant la forme contraposée de l'identité, « Si nonP, alors nonP », comme corne gauche, avec le même résultat.

> Si nonP, alors nonP — et — si nonP, alors P
>
> mais soit non-nonP soit nonP
>
> donc, non-nonP, c'est-à-dire P.

La conclusion « P » signifie ici que P est logiquement nécessaire, et non simplement que P est vraie ; notez

bien ; cela découle de la nécessité formelle de la prémisse mineure, la disjonction de P et nonP, en supposant que la bonne prémisse soit bien établie.

Une autre façon de comprendre le paradoxe est de l'envisager en termes de contextes de connaissance. En lisant le paradoxe « si nonP, alors P » comme « tous les contextes avec nonP sont des contextes avec P », et l'identité « si P, alors P » comme « tous les contextes avec P sont des contextes avec P », on peut en déduire que « tous les contextes sont avec P », ce qui signifie que P est logiquement nécessaire.

On peut, de manière similaire, traiter le paradoxe « si P, alors non P » pour parvenir à la conclusion « non P », ou mieux encore : P est impossible. Le processus de résolution d'un paradoxe, par l'extraction de ses conclusions catégoriques implicites, peut être qualifié *de dialectique* .

Notons au passage que l'expression abrégée du dilemme simple, sous la forme d'une seule proposition, devient alors plus compréhensible. La proposition composée « Si P, alors {Q et non-Q} » signifie simplement « non-P » ; « Si non-P, alors {Q et non-Q} » signifie « P » ; « Si (ou si) P ou non-P, alors Q » signifie « Q » ; et « Si (ou si) P ou non-P, alors non-Q » signifie « non-Q ». De telles propositions pourraient également être qualifiées de paradoxales, même si la contradiction qu'elles engendrent concerne une autre thèse.

Cependant, rappelez-vous que les deux formes ci-dessus ne doivent pas être confondues avec les relations hypothétiques négatives mineures « Que P ou nonP, (non-alors non) Q » ou « Que P ou nonP, (non-alors non) nonQ », respectivement, qui ne sont pas paradoxales, à moins qu'il n'existe des conditions dans lesquelles elles atteignent le niveau d'hypothétiques positives.

3. Systématisation

Normalement, nous présumons que nos informations sont déjà exemptes de thèses évidentes ou contradictoires, tandis que dans des situations anormales, comme avec le paradoxe, les thèses nécessaires ou impossibles sont des éventualités formellement acceptables.

Un énoncé hypothétique de la forme primaire « Si P, alors Q » a été défini comme « P et non-Q sont impossibles simultanément ». Or, cette situation peut se présenter de plusieurs manières. Soit (i) les deux thèses, P et non-Q, sont individuellement contingentes, et seule leur conjonction est impossible – c'est le cas normal. Soit (ii) la conjonction est impossible car l'une ou l'autre des thèses est individuellement impossible, tandis que l'autre est individuellement possible, c'est-à-dire contingente ou nécessaire ; soit (iii) les deux sont individuellement impossibles – ces situations engendrent un paradoxe.

De même, une hypothèse de la forme primaire contradictoire « Si P, non-alors Q » a été définie comme « P et non-Q sont possibles simultanément ». Or, cette situation peut se présenter de plusieurs manières. Soit (i) les deux thèses, P et non-Q, ainsi que leur conjonction, sont contingentes – c'est le cas normal. Soit (ii) l'une ou l'autre est individuellement non seulement possible mais nécessaire, tandis que l'autre est individuellement contingente, de sorte que leur conjonction reste contingente ; soit les deux sont individuellement nécessaires, de sorte que leur conjonction est également non seulement possible mais nécessaire – ces situations engendrent un paradoxe.

Ces alternatives sont clarifiées par les tableaux suivants, pour ces formes primaires ainsi que pour leurs dérivés impliquant une ou les deux antithèses. Le terme « possible » signifie bien sûr « contingent ou nécessaire », il constitue le point commun entre les deux. Nous

utiliserons ici les symboles « **N** » pour nécessaire, « **C** » pour contingent (c'est-à-dire possible mais non nécessaire) et « **M** » pour impossible. Les combinaisons sont numérotées pour faciliter la consultation. Les symétries de ces tableaux garantissent leur exhaustivité.

Tableau 1. Modalités des thèses et des conjonctions.

No.	Theses				Conjunctions			
	P	nonP	Q	nonQ	P and Q	P and nonQ	nonP and Q	nonP and nonQ
Normal (P,Q both contingent)								
1.	C	C	C	C	C	C	C	C
2.	C	C	C	C	M	C	C	C
3.	C	C	C	C	C	M	C	C
4.	C	C	C	C	C	C	M	C
5.	C	C	C	C	C	C	C	M
6.	C	C	C	C	C	M	M	C
7.	C	C	C	C	M	C	C	M
Abnormal (one or both of P, Q not contingent)								
8.	M	N	C	C	M	M	C	C
9.	N	M	C	C	C	C	M	M
10.	C	C	M	N	M	C	M	C
11.	C	C	N	M	C	M	C	M
12.	N	M	N	M	N	M	M	M
13.	N	M	M	N	M	N	M	M
14.	M	N	N	M	M	M	N	M
15.	M	N	M	N	M	M	M	N

Le tableau suivant fait suite au précédent. « **Oui** » indique qu'une implication et sa contraposée sont implicites dans la forme concernée, tandis que « **non** » indique qu'elles en sont exclues. «□ » signifie « implique » et «□ » signifie « est impliqué par ».

Tableau 2. Hypothèses définies correspondantes.

No.	Name	Implications (→) and Contraposites (←)			
		P→nonQ	P→Q	nonP→nonQ	nonP→Q
		nonP←Q	nonP←nonQ	P←Q	P←nonQ
	Normal (P,Q both contingent)				
1.	Neutral	no	no	no	no
2.	Contrary	yes	no	no	no
3.	Subalternating	no	yes	no	no
4.	Subalternated	no	no	yes	no
5.	Subcontrary	no	no	no	yes
6.	Implicant	no	yes	yes	no
7.	Contradictory	yes	no	no	yes
	Abnormal (one or both of P, Q not contingent)				
8.	P impossible, Q contingent	yes	yes	no	no
9.	P necessary, Q contingent	no	no	yes	yes
10.	P contingent, Q impossible	yes	no	yes	no
11.	P contingent, Q necessary	no	yes	no	yes
12.	P, Q both necessary	no	yes	yes	yes
13.	P necessary, Q impossible	yes	no	yes	yes
14.	P impossible, Q necessary	yes	yes	no	yes
15.	P, Q both impossible	yes	yes	yes	no

La logique hypothétique normale suppose donc que les thèses des hypothèses sont toujours contingentes, et se limite ainsi aux cas n° **1** à **7** des tableaux précédents. Cependant, pour élaborer une logique hypothétique complète, il convient également de considérer les cas anormaux n° **8** à **15** , dans lesquels une ou les deux thèses ne sont pas contingentes (c'est-à-dire qu'elles sont évidentes ou contradictoires).

La définition de la forme positive primaire « Si P, alors Q », tout en restant inchangée comme « P plus nonQ n'est pas possible », comprend désormais plus précisément les situations suivantes : n° **3** , **6** , **8** , **11** , **12** , **14** ou **15** ,

c'est-à-dire tous les cas où « P et nonQ » est impossible (« M »), ou « P implique Q » est marqué « oui ».

La définition de la forme négative primaire « Si P, non-alors Q », tout en restant inchangée comme « P plus nonQ n'est pas impossible », est maintenant considérée comme comprenant plus précisément les situations suivantes : n° **1** , **2** , **4** , **5** , **7** , **9** , **10** ou **13** , c'est-à-dire tous les cas où « P et nonQ » est contingent (C), ou « P implique Q » est marqué « non ».

Les six autres formes hypothétiques, impliquant les antithèses de P et/ou Q, peuvent également se voir attribuer des définitions améliorées, en se référant aux tableaux ci-dessus.

Observez les symétries dans ces tableaux. Dans le cas n° **1** , toutes les conjonctions sont « C » et toutes les implications sont « non ». Dans les cas n ° **2** à **5** , une conjonction est « M » et une implication est « oui ». Dans les cas n° **6** à **11** , deux conjonctions sont « M » et deux implications sont « oui ». Dans les cas n ° **12** à **15** , trois conjonctions sont « M » et trois implications sont « oui ». Notez le statut correspondant des thèses individuelles dans chaque cas.

Le procédé de contraposition s'applique universellement à toutes les hypothèses, positives ou négatives, normales ou anormales, car il découle directement des définitions. C'est pourquoi, dans les tableaux ci-dessus, chaque implication est étroitement liée à sa contraposée. De même, la négation de toute implication engendre la négation de sa contraposée, de sorte que les tableaux ci-dessus concernent aussi indirectement les hypothèses négatives.

Lors de l'élaboration de notre théorie des propositions hypothétiques, il convient de formuler clairement la portée et les limites d'application de tout processus considéré, et de préciser les exceptions éventuelles à ses

règles. La validité ou l'invalidité des processus logiques dépend souvent du fait que l'on s'intéresse aux formes normales ou anormales, bien que, dans certains cas, ces deux catégories de propositions se comportent de manière identique. Si l'on néglige ces distinctions, on risque aisément de tomber dans des incohérences formelles.

4. Propriétés

Les propositions paradoxales obéissent aux lois de la logique qui s'appliquent à toutes les hypothèses, c'est-à-dire aux hypothèses de base indéterminée. Cependant, les paradoxales, étant des hypothèses fondées sur l'incontinence, possèdent des propriétés que les hypothèses normales n'ont pas, ou inversement. Dans de telles situations, où des différences de propriétés logiques apparaissent, la logique hypothétique générale suit le cas le plus faible.

Les similitudes et les différences en matière de comportement formel ont déjà été traitées en détail dans les chapitres correspondants, mais certaines sont rappelées ici afin de souligner le rôle joué par le paradoxe.

a. Opposition.

Dans la doctrine de l'opposition, nous affirmions que « Si P, alors Q » et « Si P, alors non Q » étaient nécessairement contradictoires, car si P était vrai, Q et non Q seraient tous deux vrais, ce qui est absurde. Cependant, en plaçant ces propositions dans un dilemme destructeur, comme ci-dessous, nous aurions obtenu un argument valable :

Si P, alors Q — et — si P, alors nonQ

mais soit nonQ, soit Q

donc nonP

De même, « Si P, alors Q » et « Si non P, alors Q » pourraient être intégrés dans un dilemme constructif simple et valide, donnant Q, au lieu de soutenir, comme nous l'avons fait, qu'ils doivent être contraires parce que leurs contrapositions aboutissent à l'absurdité de non Q impliquant non P et P.

Il s'ensuit que ces contradictions ne sont valides que sous certaines conditions, pour les hypothèses fondées sur la contingence. Il existe des circonstances exceptionnelles où elles ne s'appliquent pas, notamment en ce qui concerne les hypothèses anormales (y compris les paradoxales).

Cela ressort également clairement de l'observation des marques « oui » parallèles, dans les cas n° **8** , **14** , **15** (en tenant compte à la fois de « P implique non Q » et de « P implique Q », où P est impossible), et dans les cas n° **11** , **12** , **14** (en tenant compte à la fois de « P implique Q » et de « non P implique Q », où Q est nécessaire).

Des restrictions similaires s'appliquent automatiquement à la sous-contrarieté entre « Si P, non-alors nonQ » et « Si P, non-alors Q », ainsi qu'à la subalternation, par la majuscule « Si P, alors Q », de la minuscule « Si P, non-alors nonQ » (qui correspond à une obversion). Ces oppositions ne sont valables que pour les hypothèses normales ; lorsqu'il s'agit d'hypothèses anormales (et donc en logique générale), par souci de cohérence, nous devons considérer ces propositions comme neutres l'une par rapport à l'autre.

b. Éducation.

Il en va de même pour les déductions dérivées. Le processus primaire de contraposition est inconditionnel et applicable à toutes les hypothèses, mais les autres processus peuvent être critiqués de la même manière que précédemment, en formant des dilemmes simples valides,

utilisant la proposition source et la négation de la cible proposée, ou la ou les contraposées de l'une ou de l'autre, ou des deux, comme arguments.

Ces propositions peuvent également être combinées dans un syllogisme, aboutissant à une conclusion paradoxale. Ainsi :

Dans le cas d'obversion ou de conversion obvertie (dans le premier cas, nier la contraposée de la cible) :

Si Q, alors nonP (négation de la cible)

Si P, alors Q (source)

Donc si P, alors nonP (paradoxe = nonP)

Dans le cas d'une conversion par négation ou par inversion obvertie (dans ce dernier cas, négation de la contraposée de la cible) :

Si P, alors Q (source)

si nonQ, alors P (négation de la cible)

Donc, si nonQ, alors Q (paradoxe = Q)

Ainsi, les processus déductifs autres que la contraposition ne conviennent qu'aux hypothèses contingentes et ne peuvent être reproduits dans la logique anormale des paradoxes. Ceci est clairement illustré dans les tableaux ci-dessus.

Considérons la forme paradigmatique « Si P, alors Q ». Si nous nous limitons aux cas 1 à 7, elle n'apparaît que dans deux situations : la subalternation (**3**) et l'implication (**6**). Dans ces deux situations, « P implique non-Q » est systématiquement « non », donc la réciproque « Si P, non-alors non-Q » est vraie ; et la contraposée « Q implique non-P » est également « non », donc la réciproque par inversion « Si Q, non-alors non-P » est vraie ; « non-P implique Q » est systématiquement « non », donc l'inverse par inversion « Si non-P, non-alors Q » est vraie ; et la contraposée « non-Q implique P » est également « non », donc la réciproque par négation « Si

non-Q, non-alors P » est vraie. En ce qui concerne l'inversion et la conversion, elles ne sont pas applicables, car « nonP implique nonQ » et « Q implique P » sont « non » dans un cas, mais « oui » dans l'autre.

Cependant, si nous élargissons notre analyse aux cas n ° **8** à **15** , nous constatons que la relation « Si P, alors Q » se vérifie également si P est contradictoire et Q contingent (**8**), ou si P est contingent et Q est évident (**11**), ou si P et Q sont tous deux évidents (**12**), ou si P est contradictoire et Q est évident (**14**), ou encore si P et Q sont tous deux contradictoires (**15**). Les uniformités mentionnées précédemment, qui rendaient possibles les déductions énoncées, ne sont plus valables. La présence d'un mélange de réponses « non » et « oui » parmi les alternatives disponibles empêche de telles déductions.

c. Déduction.

En ce qui concerne le syllogisme, les *modes non subalternes* , validés par réduction à l'absurde, restent universellement valides, car une telle réduction indirecte est essentiellement une contraposition, et aucun autre processus déductif n'est supposé. En revanche, les *modes subalternes* des trois figures ne sont valides que pour les hypothétiques normaux. Puisque ces modes présupposent des subalternations pour leur validation, c'est-à-dire qu'ils dépendent de réductions directes par obversion ou inversion obvertie, ils ne sont pas valides pour les hypothétiques anormaux.

En ce qui concerne l'apodose, les modes à prémisse mineure modale nous offrent le point d'entrée vers la logique anormale. Quant au dilemme, il est l'instrument *par excellence* pour mettre au jour les paradoxes du raisonnement quotidien. Si l'on met sous forme syllogistique *un* dilemme simple, constructif (comme

ci-dessous) ou destructif (mutatis mutandis), on obtient une conclusion paradoxale :

Si P, alors R — et — si Q, alors R

mais P et/ou Q

donc, R

Cela implique les sorites :

Si nonR, alors nonP (corne gauche opposée)

si nonP, alors Q (mineur)

Si Q, alors R (klaxon droit)

par conséquent, si nonR, alors R (conclusion paradoxale = R)

Ainsi, les propositions paradoxales font partie intégrante de la logique hypothétique générale, et non un appendice étrange. Elles mettent en lumière la continuité essentielle entre le syllogisme et le dilemme simple, ce dernier étant réductible au premier.

Il s'ensuit incidemment que, puisque (comme on l'a vu précédemment) l'apodose peut être considérée comme un cas particulier et limite du dilemme simple, et le dilemme simple comme un cas particulier et limite du dilemme complexe, tous les processus inférentiels relatifs aux hypothèses sont étroitement liés.

Le paradoxe engendré par un dilemme simple dépend, bien entendu, de la vérité de ses prémisses. Il ne faut pas conclure hâtivement, du paradoxe inhérent à tout dilemme simple, que toutes les vérités sont fondamentalement évidentes et toutes les faussetés fondamentalement contradictoires. La connaissance n'est pas une entreprise purement rationnelle, mais repose en grande partie sur des observations empiriques.

Comme nous l'avons déjà souligné, un dilemme simple ne peut aboutir qu'à une nécessité ou une impossibilité catégorique, et ce, uniquement si toutes ses prémisses sont

elles-mêmes indubitablement incontingentes. S'il existe des conditions tacites ou un doute quant à l'inconditionnalité des hypothèses (les cornes) et/ou de la disjonction (la prémisse mineure), la conclusion s'en trouverait proportionnellement affaiblie quant à sa modalité logique.

Ainsi, en reprenant l'exemple précédent et en *admettant la validité* de la prémisse majeure : dans le cas particulier où notre prémisse mineure est une disjonction formellement donnée — si, par exemple, P et Q sont contradictoires (P ≠ Q, Q ≠ P) —, alors la conclusion R est effectivement nécessaire. Mais généralement, les alternatives P et Q énumérées ne sont exhaustives que contextuellement, de sorte que la conclusion R n'est que factuellement vraie.

Ainsi, bien que toute nécessité logique soit évidente et toute impossibilité logique contradictoire, formellement parlant, selon nos définitions, il serait judicieux d'affirmer que ces prédications ne sont pas réciproques en pratique et de distinguer le paradoxe apodictique du paradoxe factuel. Le premier est évident en soi ; le second découle de données plus empiriques et, par conséquent, bien que fiable dans son contexte, a moins de poids et de caractère définitif.

Enfin, notons que l'incohérence de deux dilemmes simples «également convaincants» peut maintenant être mieux comprise, car elle résulte des paradoxes contraires qu'ils impliquent.

5. Le paradoxe du menteur (première version)

De Future Logic 32 et Ruminations 1 et 5.

1. Double paradox

Nous avons vu que les propositions logiques de la forme « si P, alors non P » (qui équivaut à « non P ») ou « si non P, alors P » (qui équivaut à « P ») sont parfaitement valides. Elles signifient que l'antécédent est contradictoire et logiquement impossible, et que le conséquent est évident et logiquement nécessaire. En tant que propositions, elles ne sont aucunement antinomiques ; c'est l'un de leurs constituants qui est absurde.

Bien que chacune de ces propositions, prise isolément, soit formellement acceptable et susceptible d'être vraie, elles ne peuvent jamais être vraies simultanément : elles sont des contraires irréconciliables et leur conjonction est formellement impossible. Car si elles étaient toutes deux vraies, alors P et nonP seraient toutes deux implicitement vraies.

Il faut donc distinguer entre *le paradoxe simple* , qui a (plus précisément qu'on ne l'a suggéré précédemment) la forme « si P, alors non P ; mais si non P, alors non P ; d'où non P », ou la forme « si non P, alors P ; mais si P, alors non P ; d'où P » — et *le paradoxe double* , qui a la forme « si P, alors non P, *et* si non P, alors P ».

Le paradoxe simple, je le répète, relève du domaine de la logique, tandis que le double paradoxe le dépasse. Le

premier peut être vrai ; le second révèle toujours une erreur de raisonnement. Pourtant, pourrait-on objecter, le double paradoxe est fréquent en pratique ! Mais cela ne le rend pas pour autant juste, pas plus que la fréquence d'autres erreurs en pratique ne les rend vraies.

Le double paradoxe est rendu possible, comme nous le verrons, par un *usage détourné des concepts* . Il s'agit d'un sophisme par excellence, car il nous donne l'illusion superficielle d'un énoncé cohérent aboutissant à des résultats contraires à la raison. Mais un examen plus approfondi révèle la présence d'un sophisme, tel qu'une ambiguïté ou une équivoque, ce qui signifie qu'en réalité, la contradiction apparente n'a jamais eu lieu.

La logique exige que *l'une ou l'autre* des propositions hypothétiques constituant le double paradoxe, ou paradoxe sur paradoxe, *soit fausse* . Alors qu'un paradoxe simple est *résolu* en concluant catégoriquement à sa conséquence, sans nier le lien antécédent-conséquent, un double paradoxe est *dissous* en démontrant qu'un ou les deux paradoxes simples impliqués sont faux, voire inexistants. Il est important de noter la différence dans la résolution du problème : la résolution « explique » le paradoxe simple, tandis que la dissolution « élimine » le double paradoxe.

Le double paradoxe *révèle* que nous commettons une erreur ; la contradiction dont nous sommes confrontés constitue un indice suffisant pour prouver que nous avons formulé une hypothèse erronée. Notre capacité à percevoir correctement les liens logiques n'est pas remise en cause par un jugement initial trop hâtif, sans examen approfondi des termes employés. Une fois les concepts en jeu clarifiés, c'est la faculté rationnelle elle-même qui prononce le jugement contraire à son impression initiale de connexion.

Il faut bien comprendre que chaque double paradoxe (comme tout paradoxe simple, d'ailleurs) *nous apprend quelque chose* . De tels événements ne doivent pas être perçus comme des menaces pour la raison, remettant en cause la logique dans son ensemble, mais simplement comme des leçons. Ils sont des sources d'information ; ils nous révèlent certaines règles logiques de formation des concepts, que nous n'aurions pas remarquées autrement. Ils nous montrent les limites de la conformité linguistique.

2. Le paradoxe du menteur

Un exemple ancien de double paradoxe est le célèbre « paradoxe du menteur », découvert par Eubulide, un Grec du IVe siècle avant notre ère, appartenant à l'école mégarienne. Il se formule ainsi : « Un homme qui affirme mentir dit-il la vérité ? » Il semblerait donc que s'il ment, il dit la vérité, et que s'il ne ment pas, il dit la vérité.

L’erreur conceptuelle à l’origine de la difficulté réside ici dans le fait que la proposition est *définie par référence à elle-même* . Le paradoxe du menteur nous permet de constater que de tels concepts sont interdits.

Le mot « maintenant » (qui définit la proposition elle-même comme son propre sujet) est employé en référence à quelque chose qui n'existe pas encore, dont l'existence apparente n'est rendue possible que par lui. Ainsi, en réalité, le mot est dépourvu de référents spécifiques dans le cas présent. Le mot « maintenant » est certes généralement significatif, car dans d'autres situations, il possède des référents précis ; mais ici, il est employé avant même que nous ayons quoi que ce soit à désigner comme sujet de discours. Il ressemble à un mot, et il le sonne comme tel, mais il n'est rien de plus.

Une version plus moderne et plus claire de ce paradoxe est « cette proposition est fausse », car elle met en

évidence la fonction indicative du mot « maintenant » dans le mot « ceci ».

Le mot « ceci » accompagne nos désignations et présuppose l'existence d'un référent déjà présent. Il ne peut se créer un référent ex nihilo. C'est la leçon utile que nous enseigne le paradoxe du menteur. On peut parfaitement utiliser le mot « ceci » pour désigner un autre mot « ceci », mais pas lui-même. Ainsi, je peux vous dire sans difficulté : « ce “ceci”, qui figure dans la proposition “cette proposition est fausse” », car mon « ceci » a un référent, même s'il s'agit d'un symbole vide ; mais le « ceci » initial est dénué de sens.

De plus, cette version semble impliquer que « si la proposition est vraie, elle est fausse, et si elle est fausse, elle est vraie ». Cependant, un examen plus approfondi révèle que l'expression « la proposition » ou « elle » a une signification différente selon qu'il s'agit d'antécédents ou de conséquent.

Si, par hypothèse, nous comprenons ces implications comme suit : si cette proposition est fausse, alors cette proposition est vraie ; et si cette proposition est vraie, alors cette proposition est fausse — en prenant le « ceci » au sens d' *autoréférence* de chaque thèse —, alors nous constatons que les thèses n'ont en réalité pas un seul et même sujet, et ne sont que présumées contradictoires.

Elles ne le sont pas formellement, pas plus que, pour toutes propositions P1 et P2, « P1 est vrai » et « P2 est faux » ne sont contradictoires. Les implications ne sont pas nécessaires d'un point de vue logique, et les deux paradoxes sont ainsi résolus. Il n'y a pas de contradiction, ni dans « cette proposition est fausse », ni dans « cette proposition est vraie » ; elles sont simplement dénuées de sens, car les indicatifs qu'elles emploient sont sans référence.

Essayons aussi d'interpréter ces implications ainsi : si « cette proposition est fausse » est vrai, alors cette proposition est fausse ; et si cette proposition est fausse, alors cette proposition est vraie. Le premier « cette proposition » est une autoréférence, et les « cette proposition » suivants renvoient à la proposition initiale et non aux thèses suivantes elles-mêmes. Autrement dit : si « cette proposition est fausse » est vrai, alors « cette proposition est fausse » est fausse ; et si « cette proposition est fausse » est fausse, alors « cette proposition est fausse » est vraie.

Ici, les sujets des thèses sont identiques, mais les implications ne semblent plus nécessaires, comme le montre clairement la substitution du symbole P à « cette proposition est fausse ». La saveur du paradoxe disparaît : elle n'existait que tant que « cette proposition est fausse » semblait être impliquée par ou impliquer « cette proposition est vraie » ; dès que le sujet est unifié, les deux paradoxes s'effondrent.

On ne peut éluder le problème en formulant le paradoxe du menteur comme une généralité. La proposition « Je mens toujours » peut simplement être réfutée par « Tu mens parfois (comme dans le cas de « Je mens toujours »), mais parfois tu dis la vérité » ; cela ne donne lieu qu'à un double paradoxe à l'indicatif. De même, la proposition « Toutes les propositions sont fausses » peut être réfutée par « Certaines, d'autres non », sans difficulté.

Or, il convient de noter que ces deux propositions générales sont en réalité contradictoires ; elles engendrent chacune un paradoxe. Il s'ensuit qu'elles sont toutes deux fausses : on ne peut prétendre « toujours mentir », ni qu'« il n'existe aucune proposition vraie ». Il s'agit d'un raisonnement logique ordinaire, parfaitement légitime puisqu'il existe des alternatives logiques.

Concernant ces alternatives, la proposition « Je ne mens jamais » n'est pas incohérente en soi, sauf pour la personne qui affirme intentionnellement « Je mens toujours ». De même, la proposition « Toutes les propositions sont vraies » n'est pas incohérente en soi, mais elle est incompatible avec le fait logique que certaines propositions sont incohérentes ; elle est donc fausse. Dans ce cas, seule la proposition contingente « Certaines propositions sont vraies, d'autres fausses » peut être retenue.

3. Plus d'informations sur le paradoxe du menteur

Dès lors que l'on comprend que le sens des mots réside dans leur intention, individuellement et collectivement, la solution du paradoxe du menteur devient limpide. L'autoréférence est dénuée de sens, car une intention ne peut se porter elle-même sur elle-même, puisqu'elle n'existe pas encore ; une intention ne peut porter que sur quelque chose qui existe déjà, par exemple une autre intention dirigée vers une tierce chose.

De ce fait, la proposition « cette proposition est fausse » est dénuée de sens, tout comme la proposition « cette proposition est vraie ». On peut indifféremment les déclarer toutes deux à la fois vraies et fausses, ou ni vraies ni fausses ; cela ne change rien, car les mots « cette proposition » ne désignent absolument rien [11].

Bien que les mots employés dans ces phrases aient un sens pris individuellement et que leur structure grammaticale soit correcte, leur absence collective de contenu implique que leur valeur logique collective est nulle. L'autoréférence est syntaxiquement cohérente, mais sémantiquement incohérente. C'est comme un

---

11 Voir *Future Logic* , chapitre 32.2.

raisonnement circulaire, sans issue, qui ne mène nulle part.

L'affirmation exclusive « *Seule* cette proposition est vraie » implique à la fois : « *Cette* proposition est vraie » et « *Toutes les autres* propositions sont fausses », c'est-à-dire qu'elle est équivalente à l'affirmation exceptive « *Toutes* les propositions, *sauf celle*- ci, sont fausses ». Cette dernière est souvent avancée par certains philosophes, par exemple ceux qui affirment que « tout est illusion (sauf ce fait) ».

Mon propos est que de telles affirmations ne se limitent pas à l'erreur d'autoréférence (« cette proposition »). Elles impliquent en outre une référence à « tous les autres », ce qui est critiquable, car :

- Prétendre connaître « toutes les autres propositions » revient à revendiquer l' *omniscience* , à s'arroger le droit de tout savoir, tout ce qu'il y a à savoir, tout ce qu'il y aura jamais à savoir. Généralement, de telles affirmations sont faites sans justification crédible, bien qu'elles contredisent toutes les conclusions antérieures de l'expérience et de la raison.
- Certes, *d'* autres propositions sont en réalité considérées et admises comme vraies par ces philosophes. Ils sont généralement assez loquaces, voire verbeux ; ils ne se contentent pas d'affirmer systématiquement *cette* seule chose et refusent d'en dire d'autres.
- Et bien sûr, formellement, si « ceci » n'a aucun sens (comme démontré précédemment), alors « tous les autres », c'est-à-dire « tout autre *que ceci* », n'a également aucun sens !

Le paradoxe du menteur, soit dit en passant, est attribué aux Grecs anciens, soit à Eubulide de Milet (IVe siècle av. J.-C.), soit à Épiménide de Crète (VIe siècle av. J.-C.). J'ignore si sa résolution était évidente pour ces premiers

logiciens, mais un texte anonyme (européen ?) du XIVe siècle apr. J.-C. expliquerait que l'affirmation du menteur n'est ni vraie ni fausse, mais simplement dénuée de sens. Cette explication est donc historiquement bien antérieure à la logique moderne (Russell et al., même si ces logiciens plus tardifs ont certainement clarifié la question).[12]

4. L'utilité des paradoxes

Une *proposition paradoxale (simple)* se présente sous la forme « si P, alors non P » ou « si non P, alors P », où P est une proposition quelconque. Il est important de comprendre que ***de telles propositions sont parfaitement légitimes dans le discours : un paradoxe (simple) n'est pas une contradiction*** . En revanche, un *double paradoxe* , c'est-à-dire l'affirmation selon laquelle « si P, alors non P » *et* « si non P, alors P » sont tous deux vrais dans un cas donné de P, constitue bel et bien une contradiction.

Le principe de non-contradiction stipule que la conjonction « P et non-P » est logiquement impossible ; autrement dit, deux propositions contradictoires ne peuvent être vraies simultanément. De même, le principe du tiers exclu stipule que « non-P et non-non-P » est logiquement inacceptable. Ces principes se justifient par le fait que de telles situations d'antinomie nous placent face à un dilemme cognitif : nous sommes incapables de résoudre cette difficulté logique, de trouver une solution au problème inhérent.

En revanche, un paradoxe simple ne représente pas une telle menace pour la pensée rationnelle. Il nous offre une solution logique : la négation de l'antécédent (car contradictoire) et l'affirmation du conséquent (car évident). La proposition « si P, alors non P » implique logiquement « non P », et la proposition « si non P, alors

12 Voir *Future Logic* , chapitre 63, sections 3 et 6.

P » implique logiquement « P ». Ainsi, en l'absence de double paradoxe, *une proposition qui implique sa propre négation est nécessairement fausse, et une proposition impliquée par sa propre négation est nécessairement vraie* .

Il s'ensuit, soit dit en passant, que la conjonction de ces deux propositions hypothétiques, c'est-à-dire le double paradoxe, constitue une violation du principe de non-contradiction, puisqu'elle aboutit à la conclusion composée que « P et nonP sont toutes deux vraies ». Le double paradoxe viole également le principe du tiers exclu, puisqu'il implique tout autant que « P et nonP sont toutes deux fausses ».

Ces différentes inférences peuvent être prouvées et expliquées de diverses manières :

- Puisqu'une proposition hypothétique comme « si x, alors y » signifie « x et non y sont impossibles », il s'ensuit que « si P, alors non P » signifie « P et non non P sont impossibles » (c'est-à-dire que P est impossible), et « si non P, alors P » signifie « non P et non P sont impossibles » (c'est-à-dire que non P est impossible). Notez bien cette explication.

  Nous savons que la négation de P est équivalente à nonP, et que la négation de nonP est égale à P, grâce aux lois de non-contradiction et du tiers exclu. De plus, selon la loi d'identité, la répétition du nom d'un objet ne le duplique pas : il reste le même ; par conséquent, la conjonction « P et P » est équivalente à « P » et la conjonction « nonP et nonP » est équivalente à « nonP ».

  Il est important de noter que « si P, alors non P » signifie « (P et non non P) est *impossible* ». Ainsi, bien que cela implique que « non P est vrai », il ne s'ensuit *pas* que « si non P est vrai, P implique non P ». De même, *mutatis mutandis* , pour « si non P, alors

P ». Nous nous intéressons ici à l'implication stricte (nécessité logique), et non à ce qu'on appelle l'implication matérielle.

Cette position stricte se justifie par le fait qu'en pratique, la vérité et la fausseté dépendent du contexte : la plupart de ce que nous croyons vrai aujourd'hui pourrait s'avérer faux demain, et inversement. En revanche, la nécessité ou l'impossibilité logique renvoie à une relation bien plus forte qui, une fois établie, ne devrait en principe pas varier selon le contexte de connaissance : elle s'applique à *tous les* contextes imaginables.

- Puisqu'une proposition hypothétique comme « si x, alors y » peut être reformulée en « si x, alors (x et y) », il s'ensuit que « si P, alors non P » équivaut à « si P, alors (P et non P) », et que « si non P, alors P » équivaut à « si non P, alors (non P et P) ». Dans cette perspective, une proposition contradictoire implique une contradiction ; or, la contradiction étant logiquement inadmissible, une telle proposition est nécessairement fausse et sa contradictoire est nécessairement vraie. Ceci peut s'exprimer par apodose, où les lois de la pensée fournissent la prémisse mineure catégorique, nous permettant exceptionnellement de tirer une conclusion catégorique d'une prémisse hypothétique.

  Si P, alors (P et non P)
  mais : non(P et nonP)
  donc, pas P

  Si nonP, alors (nonP et P)
  mais : non (nonP et P)
  donc, pas notP

- On peut également traiter ces inférences sous forme de dilemme, en combinant le « si P, alors non P » avec « si non P, alors non P » (ce dernier découlant de la loi d'identité) ; ou encore, « si non P, alors P » avec « si P, alors P ». On obtient ainsi, de manière constructive :

  Si P alors nonP – et si nonP alors nonP
  mais : soit P, soit non P
  donc, pas P

  Si non P, alors P – et si P, alors P
  mais : soit nonP soit P
  donc, P

Les paradoxes ont parfois des conséquences indirectes. Par exemple, supposons que Q implique P et que P implique non-P (ce qui, comme nous l'avons vu, peut se réécrire sous la forme P implique à la fois P et non-P). En combinant ces propositions dans un syllogisme, on obtient la conclusion « si Q, alors P et non-P ». Cette dernière est également une proposition paradoxale, dont la conclusion est « non-Q », même si la contradiction dans le conséquent ne concerne pas directement l'antécédent. De même, la non-exclusion du milieu peut apparaître sous la forme « si Q, alors ni P ni non-P ». De telles propositions se rencontrent également dans la pratique.

Il est intéressant de constater que ces formes, « Q implique (P et nonP), donc Q est faux » et « Q implique (non P et non nonP), donc Q est faux », sont les arguments implicites dans notre application des lois de la pensée correspondantes.

Face à une antinomie dans la connaissance, nous cherchons dialectiquement à nous en débarrasser **en**

identifiant et en corrigeant une ou plusieurs erreurs d'observation ou de raisonnement *antérieures* . Ainsi, l'argument paradoxal n'est pas seulement une conséquence des lois de la pensée, mais plus largement la manière même dont nous les appliquons régulièrement dans la pratique.

Autrement dit, le processus dialectique que nous utilisons suite à la découverte d'une contradiction ou d'un tiers exclu (ou d'ailleurs d'une violation du principe d'identité) signifie que nous croyons que :

> **Toute apparente antinomie n'est, en réalité, qu'une illusion.**

Il s'agit d'une illusion *due au paradoxe* , c'est-à-dire que *certaines des prémisses* ayant conduit à cette conclusion apparemment contradictoire ou excluant le milieu sont erronées et nécessitent une correction. L'antinomie n'est jamais catégorique, mais hypothétique ; elle est le signe d'une supposition ou d'une hypothèse erronée antérieure, et en dépend. L'antinomie apparente contribue à la connaissance en révélant une faille dans sa totalité et en nous incitant à revoir notre raisonnement passé.

Contradiction et paradoxe sont étroitement liés, mais non synonymes. Le paradoxe (c'est-à-dire le paradoxe simple et non le double paradoxe) n'est pas équivalent à l'antinomie. On peut les considérer comme des difficultés cognitives de degrés différents. Dans cette perspective, tandis que l'antinomie catégorique constituerait une impasse, bloquant toute réflexion ultérieure, le paradoxe représente un degré de contradiction plus modéré (plus hypothétique), susceptible d'être résolu.

Il ressort de ce qui précède (et d'autres observations ci-dessous) le rôle crucial que joue le paradoxe en logique. La logique des propositions paradoxales ne se

limite pas à quelques cas particuliers et extrêmes comme le paradoxe du menteur. Elle constitue un outil essentiel à la quête de la connaissance, nous aidant à établir les fondements de la pensée et, d'une manière générale, à préserver notre raisonnement des erreurs logiques.

La compréhension des formes paradoxales n'est pas une découverte de la logique moderne [13], bien que relativement récente (remontant peut-être à la logique scolastique du XIVe siècle de notre ère) .

[13] Par exemple, Charles Pierce (États-Unis, 1839-1914) a remarqué que certaines propositions impliquent toutes les autres. J'ignore s'il avait conscience que cette propriété est propre aux propositions contradictoires ou logiquement impossibles, et que les propositions évidentes ou nécessaires ont la propriété inverse d'être impliquées par toutes les autres. Je suppose qu'il raisonnait en termes d'implication matérielle plutôt qu'en termes d'implication stricte.

6. Le paradoxe du menteur (version revisitée)

D'après A Fortiori Logic, appendice 7.4.

J'ai déjà abordé le paradoxe du menteur dans mon ouvrage * *Logique du futur* [14]*, mais je me rends compte qu'il convient de l'approfondir. Ce paradoxe est particulièrement complexe car il recourt simultanément à plusieurs artifices discursifs.

1. Première approche

L'affirmation « Cette proposition est fausse » semble concevable de prime abord, jusqu'à ce que l'on réalise que si on la tient pour vraie, il faut admettre qu'elle est fausse, et inversement, si on la tient pour fausse, il faut admettre qu'elle est vraie – ce qui paraît absurde. De toute évidence, un tel raisonnement est contradictoire, puisqu'une chose ne peut être à la fois vraie et fausse. Mais la question est : quelle en est la cause et comment la résoudre ? Nous ne « déduisons » pas la contradiction d'une « loi de la pensée » ; nous l'« observons » grâce à notre raison. Nous ne pouvons pas non plus « déduire » la solution de la contradiction d'une « loi de la pensée » ; il nous faut analyser le problème en détail et proposer une solution satisfaisante, c'est-à-dire une solution qui apaise notre angoisse intellectuelle. Comme nous le verrons, il s'agit d'une question loin d'être simple.

---

[14] Voir chapitre 32.2. (Voir aussi *Réflexions* 5.1.)

La proposition « Cette proposition est fausse » est un double paradoxe, car : *si elle est vraie, alors elle est fausse ; et si elle est fausse, alors elle est vraie* . On remarque la circularité du vrai au faux et du faux au vrai. Les implications que nous tirons de la proposition donnée semblent inévitables au premier abord. Mais il faut d'abord se demander *comment nous savons que ces implications* (les deux énoncés « si-alors ») sont vraies. Comment savons-nous que « c'est vrai » implique « c'est faux », et que « c'est faux » implique « c'est vrai » ? Apparemment, nous ne « déduisons » pas ces implications d'une proposition implicite. Nous utilisons plutôt une forme d'intuition rationnelle ad hoc, c'est-à-dire que nous « percevons » directement (nous appréhendons intellectuellement) les implications de la proposition donnée. Mais une telle intuition rationnelle, bien que fiable en principe, est clairement *inductive* , et non déductive, sur le plan épistémologique. Autrement dit, elle n'est digne de confiance que si, pour une raison ou une autre, elle s'avère erronée. Cela signifie qu'il peut y avoir une ou plusieurs erreurs dans notre raisonnement ; rien n'est immuable. Et en effet, il doit y avoir une ou plusieurs erreurs, puisque cela a conduit à un double paradoxe. Il nous faut donc les rechercher.

L’emploi du pronom « il » pose peut-être problème, car il s’agit d’un terme assez vague. Posons-nous donc la question : *à quoi se réfère plus précisément le pronom « il »* , ici ?

À première vue, le « il » dans « si c'est vrai, alors c'est faux ; et si c'est faux, alors c'est vrai » se réfère à *l'énoncé entier* : « Cette proposition est fausse ». Dans ce cas, il faut reformuler le double paradoxe comme suit : *si « cette proposition est fausse » est vraie, alors « cette proposition est fausse » est fausse ; et si « cette proposition est fausse » est fausse, alors « cette proposition est fausse » est vraie* . Ici, le sujet des deux propositions conditionnelles

est plus clairement identifié comme « cette proposition est fausse » et reste donc constant. Mais cette clarification révèle des changements anormaux de prédicat, de « vrai » à « faux » et de « faux » à « vrai », qui ne s'expliquent pas facilement. Normalement, on dirait : si « cette proposition est fausse » est vraie, alors « cette proposition est fausse » est *vraie* ; et si « cette proposition est fausse » est fausse, alors « cette proposition est fausse » est *fausse* . La raison pour laquelle nous inversons ici les prédicats est que nous considérons que la proposition originale, « cette proposition est fausse », impose une telle inversion.

Cependant, alors qu'une proposition de la forme « cette proposition est fausse est vraie » se traduit aisément par « cette proposition est fausse », une proposition de la forme « cette proposition est fausse est fausse » ne peut être simplifiée de la même manière. Comment exprimer la double négation sous-jacente ? Par « cette proposition est vraie » ? De toute évidence, le sens de cette dernière formulation diffère de celui de la première, puisque le sujet « cette proposition » se réfère à des propositions différentes dans chaque cas. Ainsi, la formulation complète du paradoxe du menteur, c'est-à-dire « si cette proposition est fausse est vraie, alors cette proposition est fausse est fausse ; et si cette proposition est fausse est fausse, alors cette proposition est fausse est vraie », ne permet pas de reproduire la formule initiale exprimée à l'aide du pronom « il ».

2. Deuxième approche

Essayons donc autre chose. Si le pronom « il » se réfère à *l'expression* « cette proposition », le double paradoxe devrait être reformulé ainsi : *si « cette proposition » est vraie, alors « cette proposition » est fausse ; et si « cette proposition » est fausse, alors « cette proposition » est vraie* . Or, on constate que dans chacune de ces deux

propositions conditionnelles, bien que le sujet (« cette proposition ») demeure constant, le prédicat (« vrai » ou « faux », selon le cas) diffère entre le conséquent et l'antécédent. Ces inversions du prédicat restent inexpliquées. En règle générale, la vérité d'une proposition P n'implique pas sa fausseté, et inversement.

Nous pourrions être tentés d'utiliser l'énoncé « Cette proposition est fausse » comme prémisse pour justifier l'inférence des antécédents aux conséquents. Nous pourrions formuler deux apodoses, comme suit :

> Si cette proposition est vraie, alors elle est fausse (hypothèse).
>
> et cette proposition est fausse (étant donné) ;
>
> par conséquent, cette proposition est vraie (conclusion présumée).

> Si cette proposition est fausse, alors elle est vraie (hypothèse).
>
> et cette proposition est fausse (étant donné) ;
>
> par conséquent, cette proposition est vraie (conclusion présumée).

De toute évidence, dans le premier cas, notre raisonnement est erroné : nous tentons de nier l'antécédent pour nier le conséquent, ou d'affirmer le conséquent pour affirmer l'antécédent. Dans le second cas, la conclusion supposée découle bien des prémisses ; mais nous pouvons toujours nous interroger sur l'origine de la prémisse principale (la proposition hypothétique), et nous n'avançons donc pas davantage. Par conséquent, cette approche est elle aussi inutile, c'est-à-dire qu'elle ne prouve rien.

On pourrait aussi essayer de formuler les deux syllogismes suivants :

Cette proposition est fausse (étant donné),

et cette proposition est vraie (supposition) ;

par conséquent, cette proposition est fausse (conclusion présumée).

> Cette proposition est fausse (étant donné).
>
> et cette proposition est fausse (supposition) ;
>
> par conséquent, cette proposition est vraie (conclusion présumée).

De toute évidence, ces arguments ne sont pas de forme syllogistique à proprement parler ; mais on peut les reformuler légèrement pour obtenir des syllogismes. Les deux premières prémisses aboutiraient alors à la conclusion « il existe une proposition qui est à la fois vraie et fausse » (3/RRI), ce qui est contradictoire (donc, l'une des prémisses est nécessairement fausse) ; les deux dernières prémisses, en revanche, ne formant qu'une seule et même proposition, ne permettraient d'aboutir à aucune autre conclusion syllogistique que « il existe une proposition qui est fausse et fausse » (3/RRI), ce qui est évident (et trivial). Mais ce ne sont pas là les conclusions que nous recherchons, qui doivent concerner « cette proposition » et non simplement « une proposition quelconque ».

Une meilleure approche consiste à considérer les deux derniers arguments comme suit. Dans le premier cas, les prémisses « cette proposition est fausse » (donnée) et « cette proposition est vraie » (supposition) semblent impliquer conjointement « cette proposition est à la fois vraie et fausse » ; et cette dernière conclusion paradoxale suggère effectivement que « cette proposition est fausse », la contradiction étant impossible. Dans le second cas, les prémisses « cette proposition est fausse » (donnée) et « cette proposition est fausse » (supposition) s'accordent sur le fait que « cette proposition est fausse », ce qui constitue leur conclusion logique. Puisque les deux arguments concluent à « cette proposition est fausse », la seconde conclusion générale s'impose.

Cependant, ce dernier résultat n'est pas aussi concluant qu'il n'y paraît, car un examen plus approfondi révèle que « cette proposition est fausse » et « cette proposition est vraie » ne se réfèrent pas au même sujet, puisque le prédicat change. Le premier « cette proposition » se réfère à la proposition « cette proposition est fausse » et le second « cette proposition » à la proposition « cette proposition est vraie ». Ainsi, ces deux propositions ont en réalité des sujets et des prédicats différents (à savoir, respectivement, faux et vrai). Les sujets semblent superficiellement identiques, car ils sont exprimés verbalement avec les mêmes mots ; mais leur signification sous-jacente diffère, puisqu'ils se réfèrent à des propositions significativement différentes (des propositions avec des prédicats manifestement différents, voire contradictoires). Cela signifie que lorsque le prédicat change, le sujet change effectivement lui aussi. Lorsque le prédicat est « vrai », le sujet signifie une chose ; et lorsque le prédicat est « faux », le sujet en signifie une autre. Bien que les mots « cette proposition » soient constants, leur signification sous-jacente varie. Autrement dit, l'expression « cette proposition » n'a pas une signification uniforme et ne peut donc pas servir de base aux inférences proposées ci-dessus.

3. Troisième approche

Essayons maintenant une autre approche. Si nous examinons plus attentivement notre raisonnement initial concernant le pronom « il », nous pouvons en saisir le véritable fonctionnement. Étant donné que « cette proposition est fausse » est vrai, nous pouvons dire plus brièvement : « cette proposition est fausse ». De même, étant donné que « cette proposition est fausse » est faux, nous pouvons déduire par négation que « cette proposition n'est *pas* fausse » est vrai, ce qui signifie que « cette

proposition est vraie » est vrai, ou plus brièvement : « cette proposition est vraie » [15]. De cette manière, il *semble* que nous passions, concernant le sujet « cette proposition est fausse », de « il est vrai » à « il est faux », et de « il est faux » à « il est vrai ». Mais en réalité, l'utilisation du pronom « il » ou de l'expression « cette proposition » comme sujet abrégé est une illusion, car le sujet sous-jacent change au cours de la seconde transition (celle qui aboutit à « cette proposition est vraie »). L'utilisation systématique d'abréviations donne l'impression qu'on parle d'une seule et même proposition, à la fois vraie et fausse. Mais comme il s'agit d'une équivoque sous-jacente, les paradoxes disparaissent.

Il est intéressant de noter que, lorsqu'on examine le raisonnement plus explicitement de cette manière, la proposition « cette proposition est vraie » devient elle aussi paradoxale ! On peut argumenter : si « cette proposition est vraie » est vraie, alors évidemment « cette proposition est vraie ». Et : si « cette proposition est vraie » est fausse, alors sa contradiction « cette proposition n'est *pas* vraie » doit être vraie, ce qui signifie que « cette proposition est fausse » est vraie, c'est-à-dire, plus succinctement : « cette proposition est fausse ». Ici, superficiellement, il ne semble pas y avoir de paradoxe, car il semble que l'on argumente, concernant le sujet « cette proposition est vraie », de « il est vrai » à « il est vrai », et de « il est faux » à « il est faux ». Mais si l'on considère la conclusion finale, à savoir « cette proposition est fausse », on constate qu'elle *correspond au* paradoxe

---

[15] Certains logiciens ont tenté de résoudre le paradoxe du menteur en niant la contradiction entre vrai et faux, c'est-à-dire que non-vrai = faux et non-faux = vrai. Une telle affirmation est absurde ; tenter de contourner le principe de non-contradiction du tiers exclu pour résoudre un paradoxe revient à se contredire.

du menteur ! [16]Et là encore, l'explication du double paradoxe réside dans le fait que le sujet apparent « il » ou « cette proposition » change de signification au cours du raisonnement.

Remarquez que, dans ces deux raisonnements, le premier est une simple tautologie, une auto-implication ordinaire, tandis que le second est une contradiction opérative, un paradoxe. Si la proposition donnée (que « cette proposition soit fausse » ou « cette proposition soit vraie ») est vraie, il suffit de la répéter telle quelle (sans ajouter la prédication « est vraie »). Mais si elle est fausse, on ne peut supprimer la prédication « est fausse » sans la modifier. Ainsi, on pourrait dire que les deux propositions, « cette proposition est fausse » ou « cette proposition est vraie », ne posent aucun problème lorsqu'on les considère comme vraies ; le problème n'apparaît que lorsqu'on les considère, par hypothèse, comme fausses. On pourrait donc dire que la solution au paradoxe du menteur (et à son analogue positif) consiste simplement à accepter les deux affirmations comme vraies, et à ne pas les imaginer fausses !

On pourrait en outre, si on le voulait vraiment, affirmer que « cette proposition est fausse » et « cette proposition est vraie » diffèrent en ce que la première semble explicitement se remettre en question, contrairement à la seconde. Sur cette base, on pourrait immédiatement rejeter la première et accepter plus ou moins la seconde, tout en admettant que cette dernière est tout aussi

---

[16] Le fait que « cette proposition est vraie » soit implicitement (ne serait-ce que potentiellement) aussi paradoxal que « cette proposition est fausse » est, à ma connaissance, une découverte récente. Il est important de noter que *ces deux* paradoxes se manifestent par des raisonnements tout à fait ordinaires : si l'on affirme « P est Q », alors P est Q ; et si l'on nie « P est Q », puis que l'on affirme « P n'est pas Q », alors P n'est pas Q (où P représente « cette proposition » et Q « faux » ou « vrai », selon le cas).

dépourvue d'information utile. Autrement dit, puisque la première paraît « plus paradoxale » que la seconde, il faudrait préférer cette dernière *en dernier recours* . Mais, il convient de le noter, cela ignore les difficultés tout aussi insurmontables qu'elle comporte. Il est préférable de rejeter résolument les deux formulations comme des constructions vicieuses.

4. Quatrième approche

Pour saisir l'illusion du paradoxe du menteur, il est essentiel de comprendre que les deux formes, « cette proposition est fausse » et « cette proposition est vraie », ne sont *pas* contradictoires ; et qu'en réalité, aucune d'elles *n'a* de contradictoire ! Il s'agit là *d'une anomalie logique* , d'une faille fatale dans le discours du paradoxe du menteur ; car, en principe, toute proposition bien formée et ayant un sens doit logiquement posséder une contradictoire. Si une forme propositionnelle est dépourvue de contradictoire, on ne peut la juger vraie ou fausse, car un tel jugement suppose l'existence d'un choix. Nous n'avons même pas besoin de limiter nos propositions aux prédicats « vrai » ou « faux » – tout prédicat X et sa négation non-X présenteraient la même propriété, étant donné le même sujet. Autrement dit, « cette proposition est X » et « cette proposition n'est pas-X » *ne sont pas* contradictoires, et sont donc *toutes deux* également dépourvues de contradictoire.

On pourrait certes remarquer que « cette proposition est X » peut être niée par « *cette* proposition (c'est-à-dire la précédente) n'est pas X », voire introduire *un symbole* pour la proposition originale dans la nouvelle proposition. Dans ce cas, bien que les sujets soient verbalement différents, leurs intentions seraient assurément les mêmes. Mais la forme « cette proposition n'est pas X » est plus proche de la forme « 'cette proposition est X' n'est pas X

», où la proposition originale entière joue le rôle de sujet et son prédicat celui de prédicat. Cependant, bien que ces deux formes soient en quelque sorte équivalentes en signification entre elles et à la proposition originale, leurs schémas de comportement logique ne sont pas identiques à celui de la proposition originale, comme nous l'avons déjà vu. Il n'en reste pas moins que « cette proposition n'est pas X » n'est pas le contraire de « cette proposition est X ».

De toute évidence, toute proposition impliquant le sujet particulier « cette proposition » présente une propriété très inhabituelle et peut être rejetée sur cette seule base. Si une telle proposition n'a pas de contradictoire, c'est parce que son sujet se réfère à la proposition *dans laquelle il se trouve* , et que cette proposition diffère manifestement selon que son prédicat est « faux » ou « vrai » (ou, plus généralement, selon la paire de prédicats « X » et « non-X »). Lorsque le prédicat change, *le sujet change également* ; il est donc insaisissable, variable et non constant comme il devrait l'être. L'expression « cette proposition » a une référence différente dans chaque cas, qui dépend du prédicat ; par conséquent, chaque sujet ne peut être associé qu'à un seul prédicat et jamais à l'autre (c'est-à-dire à sa négation).

On constate ainsi qu'au départ, face à l'affirmation « Cette proposition est fausse », nous pensions qu'en la prenant au pied de la lettre, elle devait être considérée comme fausse, et qu'il nous fallait donc lui préférer « Cette proposition n'est pas fausse », c'est-à-dire « Cette proposition est vraie », etc. Or, nous ne réalisions pas qu'en réalité, du fait de l'ambiguïté inhérente à l'expression « cette proposition » ou « elle », nous en modifiions constamment le sens. Ce changement de sens passe inaperçu, car le terme employé n'est pas, par nature, figé. Les pronoms « ceci » et « cela » peuvent s'appliquer à n'importe quoi et à son contraire sans que ce changement de sens soit explicitement

signalé. Ils ne sont pas attachés de manière permanente à un objet, mais constituent de simples désignations contextuelles. En linguistique, on les qualifie de « déictiques » ou « indexicaux ».

Ainsi, il apparaît que le paradoxe du menteur surgit, quelle que soit notre interprétation, d'une forme d'équivoque du sujet. Bien que l'antécédent et le conséquent de notre raisonnement « si-alors » semblent superficiellement désigner un seul et même sujet, un changement de sens imperceptible s'opère en réalité. Une fois perçu, ce changement dissipe l'apparence initiale de contradiction. Les implications impossibles ainsi suggérées sont remises en question, voire rendues invraisemblables. Les contradictions apparemment produites sont ainsi désamorcées, voire dissoutes, du fait de notre incapacité à les maintenir.

5. Cinquième approche

Une autre manière, complémentaire, d'aborder le paradoxe du menteur consiste à souligner *la difficulté logique de l'autoréférence* . Nombre de logiciens ont adopté cette approche, y compris moi-même lors de ma première exploration de ce sujet dans * *Future Logic** . L'argument avancé ici est que l'expression « cette proposition » se réfère à un objet (à savoir « Cette proposition est fausse » ou « Cette proposition est vraie ») qui inclut l'expression elle-même. Un doigt ne peut se pointer lui-même, et « ceci » est l'équivalent conceptuel d'un doigt. En effet, l'expression « ceci » est dénuée de contenu lorsqu'elle se réfère à elle-même ou à une phrase qui la contient. Elle est vide, sans substance. C'est comme si rien n'était dit lorsque nous nous adonnons à une telle autoréférence.

Ainsi, la phrase « Cette proposition est X » (où X représente faux, vrai, ou n'importe quoi) est en réalité

dénuée de sens ; or, une phrase dénuée de sens ne peut être ni vraie ni fausse. On peut raisonnablement la qualifier de ni vraie ni fausse, sans enfreindre le principe du tiers exclu, car aucune de ces évaluations logiques ne s'applique aux phrases dénuées de sens. La phrase « Cette proposition est fausse » semble avoir un sens car ses quatre constituants (« cette », « proposition », « est » et « fausse ») ont, pris séparément, un sens normal. Mais dans cette combinaison particulière, où l'un des éléments (« cette ») ne se réfère à rien d'existant, la phrase est dénuée de sens.

Les contradictions apparentes engendrées par l'autoréférence nous aident à en prendre conscience. C'est par cette prise de conscience intellectuelle de l'absurdité de l'autoréférence que nous expliquons et annulons ces contradictions apparentes. Dès lors, on peut affirmer que même si la phrase « Cette proposition est vraie » ne semble pas, à première vue, donner lieu à un paradoxe (contrairement à l'idée reçue : « si c'est vrai, c'est vrai ; si c'est faux, c'est faux »), elle n'en demeure pas moins, puisqu'elle implique autant d'autoréférence que « Cette proposition est fausse », qu'elle est tout aussi dénuée de sens et ne peut être qualifiée de vraie ou de fausse. En réalité, comme je l'ai démontré précédemment, « Cette proposition est vraie » engendre également un double paradoxe.

On pourrait objecter : qu'en est-il des propositions « cette proposition est autoréférentielle » et « cette proposition n'est pas autoréférentielle » ? Certes, on peut affirmer qu'elles ont un sens et que la première est vraie tandis que la seconde est fausse ! La réponse à cette objection est que les deux propositions « cette proposition se réfère à elle-même » et « cette proposition ne se réfère pas à elle-même » *ne sont pas* contradictoires, car (tout comme dans le paradoxe du menteur) leurs sujets diffèrent radicalement, chacun se référant à la proposition dans

laquelle il figure et non à l'autre. Ainsi, bien que la version positive puisse sembler plus cohérente que la négative, et donc préférable *en dernier recours* , elles sont en réalité toutes deux fondamentalement erronées, car (tout comme dans le paradoxe du menteur) aucune *n'a* de contraire, et sans la possibilité logique de nier un discours, il est impossible de juger de sa justesse ou de sa fausseté.[17]

6. Sixième approche

Peu après mes réflexions précédentes, je suis tombé sur un autre exemple intéressant d'autoréférence paradoxale :

[17] Une autre objection (qui m'a d'ailleurs été soumise par un lecteur) concerne les propositions telles que « cette affirmation comporte cinq mots » et « cette affirmation comporte six mots » : bien qu'elles contiennent le démonstratif « cette », la première semble vraie et la seconde fausse ! On pourrait alors faire remarquer que, même si les propositions « cette affirmation comporte cinq mots » et « cette affirmation ne comporte pas cinq mots » semblent avoir des significations opposées, elles ne peuvent être contradictoires, puisqu'elles paraissent toutes deux vraies. Comparons également : « cette affirmation comporte cinq mots » et « cette affirmation comporte bien cinq mots » – la première est vraie tandis que la seconde est fausse, bien que les deux *signifient* essentiellement la même chose. De toute évidence, le comportement de ces propositions est loin d'être normal, en raison de leur dépendance inhabituelle à la formulation employée. D'une certaine manière, nous comprenons le message de la proposition, nous comptons le nombre de mots qui la composent, puis nous vérifions si ce nombre correspond au nombre donné : si oui, la proposition est jugée « vraie », sinon, elle est jugée « fausse ». Mais il nous faut simultanément suivre l'évolution de la référence du démonstratif « ceci », ce qui complexifie la situation, comme expliqué précédemment. De plus, dans ce contexte précis, il convient d'être attentif à l'impact du choix des mots. Les Kneales citent « Ce que je dis maintenant est une phrase en anglais » comme exemple d'« autoréférence inoffensive » (p. 228).

« Désobéis-moi ! » [18]Cela implique un double dilemme : *si j'obéis, je désobéis, et si je désobéis, j'obéis* . Pour résoudre ce paradoxe, il faut d'abord formuler l'énoncé plus précisément : « Tu dois désobéir à cet ordre ! » On peut alors démêler l'écheveau en réalisant que l'ordre donné *a une forme impérative extérieure, mais qu'il est dépourvu de contenu* . Il ne définit aucune action spécifique et concrète à accomplir ou à ne pas accomplir. Si nous voulions obéir ou désobéir, nous ne saurions pas précisément ce que nous sommes censés faire ou ne pas faire ! C'est donc un ordre auquel on ne peut ni obéir ni désobéir. La réflexion sur ce cas m'a conduit à ce que je considère aujourd'hui comme l'argument décisif, qui résout de manière convaincante le paradoxe du menteur, même si les réflexions précédentes restent pertinentes.

Il m'est alors apparu que c'est précisément là le problème du paradoxe du menteur. Il affirme « cette proposition est fausse », mais *ne nous apprend rien sur le monde qui puisse être jugé vrai ou faux* . Une « proposition » est un énoncé qui formule une affirmation concernant le monde. Si l'énoncé ne formule aucune affirmation, s'il ne « propose » rien, on ne peut logiquement l'évaluer comme vraie ou fausse. Si il ne se réfère à rien – que ce soit du plan physique, mental ou spirituel, perceptif, intuitif ou conceptuel – il est dénué de sens. Un énoncé dépourvu de sens ne peut être qualifié de « proposition ». Les attributs de « vrai » ou de « faux » ne sont pas des prédicats ordinaires, comme « blanc » ou « noir », que l'on peut appliquer à n'importe quel sujet et dont on peut ensuite juger la véracité ou l'inexactitude. Les attributs de « vrai » et de « faux » requièrent une affirmation précise avant même de pouvoir être utilisés.

---

[18] J'ai trouvé cet exemple dans *La philosophie au jour le jour* de Robert Maggiori ; l'auteur ne dit pas s'il s'agit de sa propre invention ou de celle de quelqu'un d'autre (p. 438).

La validité de cette explication se vérifie à la lumière des « formes propositionnelles » utilisées en logique. Prenons par exemple « Tous les X sont Y ». Une telle forme propositionnelle ne peut être jugée vraie ou fausse, car elle est manifestement dépourvue de contenu. Ce n'est que lorsqu'on attribue à une telle abstraction un contenu précis, comme « Tous les hommes sont mortels », que l'on peut se demander si elle est vraie ou fausse. Une forme propositionnelle est trop vague pour être considérée comme une proposition. Elle ne nous apprend rien sur le monde, si ce n'est qu'elle implique l'existence (ou même la possibilité d'existence) de propositions concrètes de cette forme. De même qu'on ne peut désobéir à un impératif sans contenu, ni même lui obéir, on ne peut juger vraie ou fausse une expression purement formelle.

Il en va de même pour le paradoxe du menteur : à l'instar d'une proposition formelle, il est dépourvu de contenu concret et ne peut donc être jugé vrai ou faux. Ce manque de contenu tient en partie à son sujet autoréférentiel (« cette proposition »). Or, *même si son sujet n'était pas autoréférentiel* , son contenu resterait insuffisant. En effet, son prédicat « faux » (et de même son contraire, « vrai ») n'est pas un prédicat ordinaire ; il s'apparente davantage à un prédicat formel. Il ne peut être employé que si un autre prédicat, plus concret, a déjà été proposé pour le sujet considéré. Par exemple, la proposition « cette proposition est intéressante » pourrait être jugée vraie ou fausse (si elle n'était pas autoréférentielle) car elle possède déjà un prédicat (« intéressante »). Ainsi, le problème du paradoxe du menteur réside non seulement dans l'autoréférence qu'il implique, mais aussi dans l'absence d'un prédicat plus concret que le prédicat logique « faux » (ou « vrai »).

Tout ceci illustre que les « lois de la pensée » ne sont pas des axiomes au sens de prémisses fondamentales de la démarche intellectuelle à partir desquelles on déduit

mécaniquement d'autres prémisses. L'expression « lois de la pensée » désigne plutôt des intuitions récurrentes qui nous guident intellectuellement, mais ne peuvent à elles seules déterminer le résultat. L'individu en quête de connaissance, et en particulier le logicien, est *poussé* par l'évidence ou l'absurdité d'une situation à rechercher des solutions créatives aux problèmes. Il lui faut néanmoins envisager des solutions possibles et les mettre à l'épreuve.

7. Le paradoxe de Russell (première version)

De Future Logic 32, 42 et 45, et Ruminations 5.

1. Auto-adhésion

Concernant la question de l'auto-appartenance, il convient d'en dire plus. Intuitivement, du moins pour moi, l'idée qu'une chose puisse être à la fois contenante et contenue est difficile à accepter.

L'auto-appartenance signifie qu'un nominal est membre d'un nominal exactement identique. Ainsi, le fait que tous les X soient X, et donc membres de « X », ne constitue pas une auto-appartenance ; il s'agit simplement de la définition de l'appartenance à une classe de premier ordre par un non-classe.

Nous avons constaté qu'empiriquement, du moins avec des exemples courants, « X » (ou la classe de X) n'est jamais lui-même un X, ni donc un membre de « X ». Par exemple, « chiens » n'est pas un chien, ni donc un membre de « chiens ».

J'ai suggéré que cela pourrait être généralisé en un postulat inductif, en l'absence d'exemples contraires. Mon but ici est de montrer que tous les cas apparents d'appartenance à soi sont illusoires, dus uniquement à l'imprécision du langage.

Le fait que « X » soit une classe X, et donc un membre des « classes X », ne correspond pas à une auto-appartenance au sens littéral, mais simplement à la

définition de l'appartenance à une classe de second ordre par une classe de premier ordre. Par exemple, « chiens » est une classe de chiens, ou un membre des « classes de chiens », ou encore un membre de la classe des classes de chiens.

L'inférence formelle, à partir de « tous les X sont des X », selon laquelle toutes les classes de X sont des classes de X, et donc les membres des « classes de X » (ou de la classe des classes de X), ne nous donne pas un exemple de ce que l'on entend strictement par auto-appartenance ; il s'agit d'une tautologie. Par exemple, toutes les classes de chiens sont membres des « classes de chiens ».

Affirmer qu'une classe X peut être X, et donc appartenir à « X », est une affirmation plus générale que d'affirmer que « X » peut être X. Non seulement cela paraît tout aussi absurde et sans fondement empirique, mais cela ne constituerait de toute façon pas une auto-appartenance formelle. Par exemple, affirmer que les « retrievers » sont des chiens.

Quant à dire d'un X quelconque qu'il *est* « X », plutôt qu'un membre de « X » ; ou à dire qu'il *est* une autre classe X, et donc un membre des « classes X » — de telles affirmations ne semblent tout simplement pas être conformes aux intentions des définitions des classes et des classes de classes, et ne constituent en aucun cas une auto-appartenance.

La question se pose alors : « Classes X » est-elle elle-même membre de « Classes X » ? La réponse est non, même dans ce cas, il n'y a pas d'auto-appartenance. L'impression que « Classes X » pourrait être membre d'elle-même provient du fait qu'elle concerne X, bien que de manière moins directe que « X ». Par exemple, « Classes de chiens » désigne les « retrievers », les « terriers », et même les « chiens » ; et concerne donc, bien qu'indirectement, les chiens.

Plus formellement, l'appartenance à l'ensemble des « classes X » ne satisfait pas à la condition requise, à savoir que « toutes les classes X sont des X », de même que l'appartenance de « X » à l'ensemble des « classes X » repose sur l'affirmation que « tous les X sont des X ». Comme nous le verrons, cette impression ne saurait être considérée comme une généralité formelle, mais tout au plus comme une vérité contingente dans certains cas ; de ce fait, elle perd toute sa force et sa crédibilité.

Si l'on affirme que *pour tout* X, toutes les classes de X sont X, on sous-entend que pour tout X, « X » (qui est une classe de X) est X ; or, nous avons déjà présenté des cas empiriques contraires ; la relation ne peut donc être générale ni formelle. Par conséquent, nous pouvons seulement affirmer que, peut-être, *pour certains* X, toutes les classes de X sont X ; mais aucun exemple n'a été fourni concernant cette éventualité.

Puisque nous n'avons aucun fondement solide (exemples spécifiques) pour supposer que « X » ou « les classes X » soient jamais membres d'elles-mêmes, et que cette suggestion est semée d'embûches ; et que nous n'avons trouvé que des exemples crédibles où elles n'étaient pas membres d'elles-mêmes — nous sommes justifiés de présumer, par généralisation, *qu'aucune classe de quoi que ce soit, ou classe de classes de quoi que ce soit, n'est jamais membre d'elle-même.*

Je ne vois qu'une seule exception possible à ce postulat : les « choses » (ou les « classes de choses »). Mais je soupçonne que, dans ce cas, plutôt que de dire que la classe est membre d'elle-même, il faudrait considérer que la définition de l'appartenance est erronée. Autrement dit, bien que ce genre suprême *soit* une chose, il n'est « membre » de rien.

2. Le paradoxe de Russell

Le paradoxe de Russell est un exemple moderne de double paradoxe, découvert par le logicien britannique Bertrand Russell.

Il a demandé si la classe de « toutes les classes qui ne sont pas membres d'elles-mêmes » est ou non membre d'elle-même. Si « les classes qui ne sont pas membres d'elles-mêmes » n'est pas membre de « les classes qui ne sont pas membres d'elles-mêmes », alors elle est bien membre de « les classes qui ne sont pas membres d'elles-mêmes » ; et si « les classes qui ne sont pas membres d'elles-mêmes » est membre de « les classes qui ne sont pas membres d'elles-mêmes », alors elle est aussi membre de « les classes qui sont membres d'elles-mêmes ». Ainsi, nous sommes confrontés à une contradiction dans les deux cas.

En revanche, la classe de « toutes les classes qui sont membres d'elles-mêmes » ne présente pas une difficulté similaire. Si les « classes membres d'elles-mêmes » ne sont pas membres des « classes membres d'elles-mêmes », alors elles sont membres des « classes qui ne sont pas membres d'elles-mêmes » ; mais si les « classes membres d'elles-mêmes » sont membres des « classes membres d'elles-mêmes », il n'y a pas d'antinomie. Par conséquent, nous avons ici un paradoxe unique associé à une position cohérente, et une conclusion définitive peut être tirée : les « classes membres d'elles-mêmes » sont membres d'elles-mêmes.

Désormais, toute absurdité qui surgit dans le domaine de la connaissance doit être perçue comme une occasion de progrès, un stimulant à la recherche et à la découverte d'un détail jusque-là inconnu. Alors, quelle est la leçon cachée de cette énigme ?

Comme je le démontrerai, le paradoxe de Russell procède essentiellement d'une équivoque ; il s'apparente davantage

au sophisme du paradoxe du barbier qu'à celui du paradoxe du menteur. Car *la question n'est pas de savoir si l'auto-appartenance est possible ou non* . Russell pensait que certaines classes, comme les « classes », s'incluent elles-mêmes ; bien que je sois en désaccord avec cette idée, ce désaccord ne justifie pas la résolution du paradoxe de Russell. En effet, ce n'est pas le concept d'auto-appartenance qui engendre une incohérence réciproque, mais le concept de non-auto-appartenance. Or, chacun s'accorde à dire qu'au moins certaines classes (sinon toutes, comme je le crois) ne s'incluent pas elles-mêmes : par exemple, les « chiens » ne constituent pas un chien.

Ce qui a déconcerté tant de logiciens face au paradoxe de Russell, c'est l'hypothèse selon laquelle nous pourrions former des concepts à volonté, pourvu que nous en formulions une définition verbale. Or, ce point de vue est injustifié. Les mots doivent avoir une signification démontrable ; dans la plupart des cas, c'est le cas ; mais parfois, ils sont isolés ou assemblés sans égard pour leurs exigences structurelles intrinsèques. On ne peut pas, par exemple, utiliser le mot « supérieur » sans préciser « que quoi ? » ; de nombreux mots sont liés entre eux et ne peuvent être intervertis au hasard. Le fait que, dans le langage courant, nous utilisions souvent les mots de manière imprécise, pour éviter des constructions fastidieuses, ne donne pas aux logiciens la même latitude.

3. Impermutabilité

La solution au problème est si simple qu'elle en est drôle, même si je dois avouer avoir été assez perplexe pendant un certain temps. Il s'agit simplement du fait que ***les propositions de la forme « X (ou « X ») est (ou n'est pas) un membre de « Y » (ou « classes Y ») » ne peuvent pas***

être permutées *. Le processus de permutation est applicable à certaines formes, mais pas à toutes.*

a. Dans certains cas, lorsqu'il s'agit de relations relativement simples, la relation peut être rattachée au prédicat initial pour former un nouveau prédicat, sous la forme d'une proposition « S est P », où « est » a une signification strictement classificatoire. Ainsi, « X n'est pas Y » peut être permuté en « X est non-Y », ou « X est quelque chose qui n'est pas Y » ; « X possède (ou ne possède pas) la Y-ité » peut être permuté en « X est une chose possédant (ou ne possédant pas) la Y-ité » ; « X fait (ou ne fait pas) Y » peut être permuté en « X est une chose faisant (ou ne faisant pas) Y ». Dans de tels cas, cet artifice ne produit aucune erreur.

Mais dans d'autres cas, la permutation est impossible, car elle falsifie les propriétés logiques de la relation en jeu. Nous en avons vu des exemples clairs et indubitables dans l'étude des modalités.

Par exemple, la forme « X peut être Y » n'est pas permutable en « X est quelque chose qui peut être Y », car cela modifie le sujet de la relation « peut être » de « X » à « quelque chose », et transforme un potentiel « peut être » en un réel « est (capable d'être) ». De telles manipulations verbales entraînent des erreurs formelles. Ainsi, on pourrait penser que « X est Y, et tous les Y sont capables d'être Z » conclut « X est capable d'être Z », alors qu'en réalité, les prémisses sont tout à fait compatibles avec la contradiction « X ne peut pas être Z », puisque « X peut devenir Z » est une conclusion alternative valable, comme nous l'avons vu précédemment.

On peut également démontrer que « X peut devenir Y » n'est pas permutable en « X est quelque chose qui peut devenir Y », car alors le syllogisme « X est Y, tous les Y sont des choses qui peuvent devenir Z, donc X est quelque chose qui peut devenir Z » semblerait valide,

alors que sa conclusion correcte est « X peut être ou devenir Z », comme nous l'avons vu précédemment. Ainsi, la modalité est un facteur relationnel non permutable. Bien que l'on dise couramment « X est capable ou incapable de Y », ce verbe « est » n'a pas les mêmes propriétés logiques que le verbe « être » dans une proposition classique de type « S est P ».

b. Le paradoxe de Russell nous révèle l'information précieuse que la copule « est membre (ou n'est pas membre) de » n'est pas non plus ouverte à la permutation en « est quelque chose qui est membre (ou n'est pas membre) de ».

Le verbe « être » initial fait partie intégrante de la relation et n'a pas la même signification que le verbe « être » pris isolément. La relation « être ou ne pas être membre de » forme un tout indivisible ; on ne peut pas la dissocier arbitrairement. Le fait qu'elle soit composée d'une suite de mots plutôt que d'un seul est une particularité du langage ; la possibilité de séparer ses constituants verbaux n'implique pas que la relation objective elle-même puisse être divisée de la même manière.

La permutation est un procédé que nous utilisons, lorsque cela est possible, pour contourner les difficultés inhérentes à une relation particulière ; or, dans ce cas précis, nous ne pouvons nous affranchir des exigences spécifiques des relations d’appartenance par cet artifice. Le paradoxe de Russell nous contraint aux processus inférentiels précédemment décrits ; il nous indique qu’il n’en existe pas d’autres légitimes, il interdit les raccourcis conceptuels.

L'impermutabilité de « est (ou n'est pas) membre de » signifie qu'on ne peut former une classe de « classes auto-membres » ni une classe de « classes non auto-membres ». Ce ne sont pas des termes, mais des relations. Ainsi, le paradoxe de Russell est entièrement

résolu par la réfutation de la légitimité conceptuelle de ses termes. Il nous est impossible de former de tels concepts ; ils impliquent une permutation illicite. Les liens entre les termes sont donc purement verbaux et illusoires.

La définition de l'appartenance est : « si quelque chose *est* X, alors il est membre de X », ou encore : « si tous les X *sont* Y, alors X est membre des classes Y ». Le paradoxe de Russell nous montre que le verbe « être » dans la condition doit être un verbe simple et isolé ; il ne peut s'agir d'un verbe « être » isolé d'une suite de mots comme « est (ou n'est pas) membre de ». Si cette condition préalable n'est pas remplie, la règle conséquente ne peut être appliquée. Dans notre cas, la condition *n'est pas remplie* , et la règle ne s'applique donc pas.

c. Voici donc comment le paradoxe de Russell se manifeste formellement, étape par étape. Nous indiquerons les permutations par des crochets comme ceci : {}.

Soit « X » une classe quelconque, de n'importe quel ordre :

(i) Si « X » est un membre de « X », alors « X » est {un membre de lui-même}. Appelons la portion entourée Y ; alors « X » est Y, ce qui définit l'auto-appartenance.

(ii) Si « X » n'est pas un membre de « X », alors « X » n'est pas un membre de lui-même. Appelons la portion entourée nonY ; alors « X » est nonY, ce qui définit la non-appartenance à soi-même.

Ensuite, appliquons les définitions générales d'appartenance et de non-appartenance aux concepts de Y et de non-Y que nous venons de former :

(iii) tout ce qui n'est pas Y est non-Y et est donc un membre de « non-Y ».

(iv) tout ce qui est Y n'est pas membre de « nonY », puisque seules les choses qui ne sont pas Y sont membres de « nonY ».

Voici maintenant le double paradoxe :

(v) si « nonY » n'est pas membre de « nonY » :

— alors, en mettant « nonY » à la place de « X » dans (ii), « nonY » n'est pas un membre de lui-même, ce qui signifie qu'il est nonY ;

— alors, par (iii), « nonY » est un membre de « nonY », ce qui contredit la prémisse de départ.

(vi) si « nonY » est membre de « nonY » :

— alors, en mettant « nonY » à la place de « X » dans (i), « nonY » est {un membre de lui-même}, ce qui signifie qu'il est Y ;

— alors, d'après (iv), « nonY » n'est pas un membre de « nonY », ce qui contredit la prémisse de départ.

De tous les procédés utilisés pour développer ces arguments, un seul est d'une validité incertaine (non établie) : la permutation de « est membre de lui-même » en « est {membre de lui-même} », ou de « n'est pas membre de lui-même » en « n'est {pas membre de lui-même} ». Puisque tous les autres procédés sont valides, la source de l'antinomie réside nécessairement dans une telle permutation. CQFD

d. L'existence de relations impermutables suggère que nous ne pouvons considérer toutes les relations comme inhérentes *aux* choses liées, comme une composante intrinsèque de leur identité. Nous sommes amenés à considérer certaines relations, telles que la modalité ou l'appartenance, comme des liens extérieurs aux termes, qui ne font pas partie intégrante de leur être.

Ainsi, par exemple, l'affirmation « ce S peut être P » n'implique pas ontologiquement l'existence d'une « marque » inscrite dans l'identité même de ce S, qui

enregistrerait qu'il « peut être P ». C'est pourquoi la proposition verbale {peut être P} ne peut être considérée comme une unité ; rien ne lui correspond dans la réalité de ce S, la relation potentielle ne projette aucune trace concrète.

Ainsi, l'« existence potentielle » doit avoir une réalité, en dehors de l'« existence actuelle ». Lorsque nous disons que « ce S peut être P », nous considérons cette potentialité comme étant P, comme faisant partie intégrante de la « nature » de ce S. Mais le S dont nous parlons s'étend lui-même dans le temps, passé, « présent » et futur ; il possède également une existence « potentielle » et est plus vaste que le S actuel.

On peut en dire autant de « ne peut pas », « doit » ou « ne peut pas ». Ainsi, les modalités naturelles (et temporelles) renvoient à différents degrés, ou niveaux, d'existence.

De même, l'impermutabilité des relations d'appartenance signifie qu'elles sont extérieures à leurs termes, ne laissant aucune trace sur eux, même lorsqu'elles sont effectives.

Cela semble une position raisonnable, car si toute relation d'une chose à toute autre impliquait une caractéristique correspondante au sein de cette chose, alors chaque chose dans le monde devrait contenir une infinité de messages, un message pour chacune de ses relations. Il est beaucoup plus simple de considérer les relations (du moins celles qui sont irréductibles) comme ayant une existence distincte de leurs termes, comme d'autres contenus de l'univers.

4. Le paradoxe du barbier

Le paradoxe du barbier [19]peut s'énoncer ainsi : « Si un barbier rase tous les habitants de sa ville qui ne se rasent pas eux-mêmes, se rase-t-il lui-même ou non ? S'il se rase lui-même, il ne se rase pas ; s'il ne se rase pas lui-même, il se rase. »

Ce double paradoxe résulte de la confusion entre les expressions « ne se rase pas lui-même » et « se fait raser par quelqu'un d'autre que lui-même ».

On peut diviser les habitants d'une ville en trois grands groupes : (a) ceux qui ne se rasent pas eux-mêmes, mais se font raser par d'autres ; (b) ceux qui ne se rasent pas eux-mêmes et ne se font pas raser par d'autres ; (c) ceux qui se rasent eux-mêmes et ne se font pas raser par d'autres. Le postulat est que notre barbier rase tous les habitants du groupe (a). Il est implicitement suggéré, sans être formellement affirmé, que personne n'appartient au groupe (b), de sorte que personne ne porte la barbe ni n'a besoin d'être rasé. Quoi qu'il en soit, le postulat ne nous apprend rien sur le groupe (c).

Ensuite, subdivisons chacun des groupes précédents en deux sous-groupes : (i) les personnes qui rasent les autres et (ii) les personnes qui ne rasent pas les autres. Il est clair que chacune des six combinaisons obtenues est logiquement acceptable, puisque le fait de savoir qui me rase n'a aucune incidence sur qui je peux raser. De toute évidence, seul le groupe (i) concerne les coiffeurs, et notre

---

[19] Ce paradoxe a été proposé par B. Russell (dans son ouvrage de 1918-1919, * *La Philosophie de l'atomisme logique* *) comme illustration du paradoxe de Russell ; mais il ne le revendiquait pas comme étant de lui, affirmant qu'il lui avait été « suggéré » par quelqu'un d'autre. Russell estimait cependant que « sous cette forme, la contradiction n'est pas très difficile à résoudre », car on peut simplement nier le sujet (c'est-à-dire dire qu'un tel barbier n'existe pas).

prémisse peut être interprétée comme signifiant que notre coiffeur est le seul coiffeur de la ville.

Nous pouvons maintenant aborder la question posée. Notre coiffeur ne peut appartenir au groupe (a)(i), puisqu'il n'est pas rasé par d'autres. Il pourrait appartenir au groupe (b)(i) s'il était autorisé à porter la barbe ou s'il était chauve ; mais admettons, pour les besoins de la démonstration, que ce ne soit pas le cas. Il peut néanmoins appartenir au groupe (c)(i), ce qui signifie qu'il se rase lui-même (plutôt que d'être rasé par d'autres), bien qu'il rase également d'autres personnes.

Il n'y a donc pas de double paradoxe. Ce paradoxe n'apparaît que parce que nous avons supposé à tort que « il rase tous ceux qui *ne* se rasent pas » exclut « il rase certains (comme lui-même) qui *se* rasent ». Or, *« X rase Y » ne contredit pas formellement « X ne rase pas non-Y »* ; rien ne permet de supposer que la copule « se raser » soit observable, de sorte que « X rase Y » implique « X ne rase pas non-Y ».

Si le principe était reformulé comme « il se rase tous ces ... » *Si l'on exclut* ceux qui ne se rasent pas (afin d'exclure « il se rase lui-même »), on pourrait toujours s'en sortir en disant « il ne se rase jamais ». Si l'on élargissait et restreignait davantage cette prémisse en insistant sur le fait qu'« il se rase d'une manière ou d'une autre ou qu'on le rase », elle deviendrait tout simplement contradictoire (à la manière d'un paradoxe absolu).

On pourrait encore enrichir ce qui précède, par exemple en prenant en compte les personnes qui se rasent dans d'autres villes, ou en faisant la distinction entre toujours, parfois/parfois pas, et jamais. Mais je pense que l'idée est claire. La leçon à tirer du « paradoxe » du barbier est que, sans catégorisations précises, des équivoques peuvent apparaître (comme celle entre « se rase » et « est rasé »), ce qui donne l'illusion d'un double paradoxe.

5. Le paradoxe du catalogue principal

On peut concevoir une classe comme une enveloppe imaginaire qui, de manière flexible, englobe tous ses membres supposés, aussi dispersés soient-ils dans l'espace et le temps, à l'exclusion de tout le reste. La question se pose alors : cette enveloppe figurative de la classe « classes » peut-elle s'envelopper elle-même ?

En examinant le **paradoxe de Russell** [20], force est de constater que *tous les « objets-mots » ne sont pas des « choses »* : les mesures des choses ne sauraient être considérées comme des choses en soi. Puisque la classification est l'expression de notre mesure des choses, elle ne peut être elle-même considérée comme une chose. Procéder ainsi conduit à un paradoxe qu'il convient d'éviter.

En d'autres termes, le problème réside dans le fait que la forme itérative (« classe de classes ») n'est identique à la forme simple (« classe »), que de manière superficielle et verbale ; la première ne peut donc être logiquement subsumée sous la seconde. La relation sujet-prédicat, modifiée par l'itération du même terme, est suffisamment significative pour exclure toute subsomption. Le paradoxe qui surgit si l'on ne maîtrise pas cette impulsion est précisément ce qui nous incite à nous en abstenir.

Le mot « choses », notez-le, a de nombreuses significations. Parfois, nous entendons par là tous les objets de pensée possibles. Parfois, nous entendons par là exclure les mots [21]. Parfois, nous entendons exclure des classes ; ou plus précisément, comme nous l'avons souligné, des classes de classes ; idem, en ce qui concerne les concepts ou les concepts de concepts. Parfois, le mot «

---

[20] Voir *Future Logic* , chapitres 43-45, sur la logique des classes.

[21] Bien entendu, cette distinction peut paraître paradoxale, puisque le mot « mot » se réfère aux mots.

choses » inclut uniquement les objets matériels, quelle que soit leur catégorie. Parfois, nous entendons par là des « entités [22]» (corps matériels, mentaux ou spirituels, ou substances délimitées, dont les cas particuliers sont généralement sujets de propositions) par opposition à leurs « propriétés » (les prédicats de lieu, de temps, de qualité, d'action, de quantité, de relation, etc.). Parfois, dans le langage courant, nous parlons de « choses » par opposition à « personnes » – c'est-à-dire que « choses » désigne ici des entités inanimées ou non volitionnelles. Et il existe encore d'autres sens du mot.

Ainsi, chaque fois que les logiciens parlent de « choses », ils devraient s'efforcer de préciser d'abord ce qui doit être inclus sous cette appellation.

Par ailleurs, pire encore que le concept d'« auto-appartenance », il y a la notion de « *classes qui semblent contradictoires avec ce qu'elles englobent* » – cette dernière paraît inconcevable de prime abord, du moins à première vue ! Ainsi, par exemple : « absence de relation » est une forme de relation ; « non-classes » est en quelque sorte une classe. Il y a forcément une erreur de raisonnement derrière ces termes, qu'il convient de clarifier. S'agit-il d'une hyperbole ou d'un terme impropre ?

La réponse à cette question est que nous avons ici encore affaire à *des classes de classes* , et que celles-ci ne sont pas nécessairement cohérentes avec leurs classes membres. Ainsi, la classe des non-relations implique toujours une relation. La classe des non-classes est néanmoins une classe. La classe des classes vides ou nulles possède bien des membres. La classe des classes dénuées de sens ou contradictoires n'est elle-même ni dénuée de sens ni contradictoire. Et ainsi de suite.

---

[22] Le mot « entité », bien sûr, est parfois employé dans un sens plus général, en référence à tout ce qui existe.

Bertrand Russell illustre son paradoxe en faisant référence à :

(a) un catalogue de tous les livres qui se mentionnent eux-mêmes, et

(b) Un catalogue de tous les livres qui ne se mentionnent *pas* eux-mêmes.

Le cas (a) ne pose aucun problème : le catalogue peut s'y inscrire sans contredire sa propre définition ; en revanche, s'il ne s'y inscrit pas, il contredit cette définition [23]. Le cas (b), par contre, est problématique : s'il ne s'y inscrit pas, conformément à sa propre définition, il devient de ce fait éligible à son propre inclusion[24] ; mais s'il s'y inscrit, il contredit sa propre définition. C'est ce dernier point qui constitue le double paradoxe dont il est question.

Mon premier grief serait le suivant : le titre du catalogue (et peut-être même une brève description de son contenu, un résumé) pourrait figurer dans le livre lui-même, mais un tel livre ne saurait contenir la reproduction intégrale du catalogue (sans parler de tous les autres livres qu'il recense ou reproduit), pour la simple raison que la tâche serait infinie (un livre dans un livre dans un livre… etc., ou la même chose au pluriel).

Le livre n'est donc pas *lui-même* un membre de lui-même ; à proprement parler, seuls les mots *qui* le concernent peuvent y être mentionnés. Les termes « inclusion » ou « appartenance », tels qu'employés ici, ont donc une signification très limitée. Ainsi, la plausibilité de l'exemple de Russell est très superficielle, voire fallacieuse ; son raisonnement est sophistique, il suggère l'impossible.

---

23 Autrement dit, le catalogue n'est pas éligible à l'inclusion en lui-même – mais cela n'affecte pas son exhaustivité.

24 De sorte que, si elle n'est pas immédiatement incluse en elle-même, elle ne peut plus être considérée comme complète.

De plus, tout livre « s'inclut lui-même » en ce sens qu'il est constitué de son propre contenu, et rien de plus. Or, si l'on conçoit un livre comme incluant un certain nombre d'autres livres, définis par une certaine affirmation (par exemple, tous les livres en anglais), il ne peut s'inclure lui-même au sens où ce contenu ne serait *qu'une partie de* lui-même. Cela signifierait non seulement une régression à l'infini (un livre contenant d'autres livres et lui-même, ce dernier contenant à son tour d'autres livres et lui-même, et ainsi de suite), et une taille infinie, mais constituerait une contradiction au sein même de la définition. Un livre ne peut être à la fois tout son contenu et seulement une partie de son contenu.

Dans cette perspective, en définissant le livre comme « le catalogue de tous les livres qui ne s'incluent pas eux-mêmes », le paradoxe de Russell s'apparente au paradoxe du menteur, puisque le livre projeté est une entité sans dimension finie ; il ne peut jamais être circonscrit.

Une seconde objection serait la suivante. Même si l'on considère la construction de Russell comme une simple liste de livres, définie comme « le catalogue de tous les livres qui ne *se mentionnent pas* eux-mêmes », cette définition est absurde, car *elle est logiquement irréalisable* . On ne peut tout simplement pas écrire un livre listant tous les livres qui ne se mentionnent pas eux-mêmes (*Lord Jim de Conrad, Notre-Dame* de Hugo , etc.), compte tenu du dilemme énoncé : que l'on y inclue ou non le livre lui-même, on se trouve face à une contradiction. Par conséquent, ce concept est nécessairement une classe nulle et dénué de sens.

La logique n'est pas déconcertée par ce paradoxe ; elle vient précisément d'apprendre que le concept proposé est erroné et inutilisable ; il faut donc l'abandonner ou, à tout le moins, le modifier. Ce paradoxe n'a rien de

dramatique ; il illustre l'une des fonctions de la logique. Nous pourrions tenter de proposer un concept modifié, comme suit. Peut-être devrions-nous plutôt nous référer à une bibliothèque.

(a) Considérons un catalogue de tous les livres d'une certaine bibliothèque, qui doit y être placé. Si le livre se mentionne lui-même, cela ne pose aucun problème. S'il ne se mentionne pas, il est simplement incomplet et doit être complété ; ou son titre est incorrect et doit être modifié (« tous les livres sauf celui-ci ») ; ou il doit être retiré de la bibliothèque.

b) Concernant le catalogue de tous les livres *qui ne figurent pas* dans notre bibliothèque : un tel livre ne peut à la fois se mentionner lui-même et être conservé dans la bibliothèque. Si nous voulons le garder, nous devons supprimer toute mention de lui-même. Si nous voulons qu'il se mentionne lui-même, nous devons l'exclure de la bibliothèque. Ce sont là des solutions pratiques qui ne posent aucun problème.

Dans cette perspective, et dans la recherche d'une expression pratique, le paradoxe de Russell se rapproche davantage du paradoxe de Barber.

6. Le paradoxe de Grelling

Pour développer son paradoxe [25], Kurt Grelling [26]qualifie un mot d'« homologique », s'il possède la qualité à laquelle il se réfère (par exemple, le mot « court » est court, ou le mot « polysyllabique » est polysyllabique), ou

---

25 Ce paradoxe s'inspire du paradoxe de Russell.

26 Allemagne, 1886-1942. Ce paradoxe est également appelé le paradoxe de Grelling-Nelson, car il a été présenté dans un article de 1908 écrit conjointement avec Leonard Nelson (Allemagne, 1882-1927).

d'« hétérologique », s'il en est dépourvu (par exemple, « long » est dépourvu de longueur, ou encore « monosyllabique » n'est pas monosyllabique). Il se demande ensuite si ces deux mots, eux-mêmes, doivent être classés de cette manière ou de l'autre, en argumentant :

- Si « hétérologique » est homologique, alors il est hétérologique (prédicats contradictoires).
- Si « hétérologique » est hétérologique, alors il est homologique (prédicats contradictoires).

Mais c'est méconnaître le sens de ces mots que de tenter de les appliquer à eux-mêmes. Dans leur cas, les références sont *trop abstraites* pour avoir des corollaires visibles ou audibles. *Aucun* des deux termes ne leur est applicable.

Il convient de noter tout d'abord que les contradictions apparentes dans la prédication, dans les deux sens, ne s'appliquent qu'au mot « hétérologique ». En effet, en suivant un raisonnement similaire concernant le mot « homologique », bien qu'il puisse paraître plus cohérent de dire qu'« homologique » est homologique plutôt qu'hétérologique, la séquence des prédicats semblerait cohérente dans les deux cas, c'est-à-dire :

- Si «homologique» est homologique, alors il est homologique (prédicats cohérents).
- Si «homologique» est hétérologique, alors il est hétérologique (prédicats cohérents).

On pourrait en déduire que le terme « homologique » est en quelque sorte mieux construit, tandis que le terme « hétérologique » présente un défaut structurel. Mais là n'est pas la véritable question.

La véritable difficulté réside dans la distinction entre les mots « homologique » et « hétérologique » et leurs significations respectives, à savoir homologique et

hétérologique. Traditionnellement, on emploie les guillemets lorsqu'on emploie un mot en tant que tel ; et lorsqu'on emploie son sens propre, on l'utilise simplement. Dans les propositions ci-dessus, d'où semble surgir un paradoxe, les sujets sont les mots eux-mêmes (entre guillemets) et les prédicats sont les significations de ces mots.

Dans cette perspective, rien ne justifie l'affirmation selon laquelle « hétérologique » impliquerait « hétérologique » et « homologique », ou inversement. Ce raisonnement est superficiel, car il confond le mot en tant que tel (utilisé comme sujet) avec sa signification (utilisée comme prédicat). Autrement dit, les guillemets placés autour du sujet ne sont pas employés de manière intentionnelle ; on sous-entend que le sujet est la signification du mot.

Comment avons-nous déduit les conséquents des antécédents ? Pouvions-nous voir d'un coup d'œil que la première thèse implique la seconde ? Examinons de plus près les propositions hypothétiques en question :

Si, dans l'antécédent, nous mettons l'accent sur *la propriété désignée* par le mot « hétérologique », à savoir une qualité présumée appelée hétérologicalité, nous formulerions les paradoxes comme suit :

- Si le mot « hétérologique » possède la propriété à laquelle il se réfère (c'est-à-dire qu'il est hétérologique), alors il semble qu'il ne possède pas la propriété à laquelle il se réfère (c'est-à-dire qu'il est homologique).
- Si le mot « hétérologique » est dépourvu de la propriété à laquelle il se réfère (c'est-à-dire qu'il est homologique), alors il possède apparemment la propriété à laquelle il se réfère (c'est-à-dire qu'il est hétérologique).

Si, en revanche, dans l'antécédent, nous mettons l'accent sur le fait que le mot « hétérologique » *possède ou non* la

propriété à laquelle il se réfère, nous formulerions alors les paradoxes comme suit :

- Si le mot « hétérologique » possède la propriété à laquelle il se réfère (c'est-à-dire qu'il est homologique), alors il semble qu'il ne possède pas la propriété à laquelle il se réfère (c'est-à-dire qu'il est hétérologique).
- Si le mot « hétérologique » est dépourvu de la propriété à laquelle il se réfère (c'est-à-dire qu'il est hétérologique), alors il possède apparemment la propriété à laquelle il se réfère (c'est-à-dire qu'il est homologique).

Dans chacun de ces cas, le conséquent est construit en comparant le sujet « hétérologique » au prédicat antécédent « hétérologique » ou « homologique ». S'il s'agit du même mot, on « infère » « homologique » comme prédicat conséquent ; s'ils diffèrent verbalement, on « infère » « hétérologique ». En réalité, en effectuant ces comparaisons entre sujet et prédicat antécédent, nous n'avons pas décelé de qualité intrinsèque au mot « hétérologique », mais nous avons implicitement fait référence à sa signification sous-jacente et l'avons confrontée au prédicat hypothétique.

Autrement dit, l'affirmation selon laquelle « hétérologique » est homologique (ou inversement) n'est pas aussi contradictoire qu'il n'y paraît ; elle pourrait même être cohérente. En réalité, elle est indéterminée et donc dénuée de sens.

Plus précisément, pour résoudre ce paradoxe, il faut se rappeler comment nos termes ont été initialement induits. On peut affirmer que « court » est court *simplement en voyant ou en entendant le mot* « court » (en supposant que chaque syllabe, quelle que soit sa prononciation, soit considérée comme courte). Mais dans le cas d'un terme comme « hétérologique », on ne peut déterminer si le mot

possède ou non la propriété à laquelle il se réfère, car cette propriété n'est pas une qualité concrète (visible ou audible) du mot, mais une abstraction que nous appliquons aux composantes visibles ou audibles des mots. Si la qualité recherchée n'est ni visible ni audible, elle est inconnaissable et il nous est impossible de déterminer quel prédicat s'applique.

Autrement dit, nos définitions initiales de ces termes, qui mentionnent « un mot possédant/dépourvu d'une certaine qualité à laquelle il se réfère », manquent de clarté et de précision, car elles n'indiquent pas, comme il se doit, que les qualités visées sont phénoménales, c'est-à-dire des aspects perceptibles du mot. Si le mot désigne quelque chose qui n'est pas inclus dans ses aspects physiques, les termes « homologique » et « hétérologique » ne s'appliquent tout simplement pas. Les employer relèverait d'un artifice verbal. Par conséquent, aucun de ces prédicats n'est applicable à ces mots en tant que tels.

On pourrait objecter que les mots possèdent des attributs non phénoménaux. Par exemple, nous considérons souvent un mot comme utile ou inutile. Dans ce cas, on pourrait se demander : le mot « utile » est-il utile ou non ? Oui, je répondrais. Par conséquent, « utile » est homologique. De même, « inutile » est utile, donc « inutile » est hétérologique. Dans cette perspective, on peut douter de l'exactitude de ce que nous venons de proposer, à savoir que les termes homologique et hétérologique présupposent des prédicats concrets (plutôt qu'abstraits).

À cette objection, on pourrait rétorquer que l'utilité d'un mot est fondamentalement concrète : un mot est utile s'il apporte une différence pratique perceptible au développement des connaissances. Dans ce cas, notre définition pourrait être légèrement modifiée, en précisant que les termes « homologique » ou « hétérologique » ne

sont applicables que si l'on peut les rattacher *, directement ou indirectement,* à une propriété concrète.

En résumé, *ces termes doivent se référer à autre chose qu'eux-mêmes avant de pouvoir être utilisés* . L'erreur en cause est similaire à celle du paradoxe du menteur, où le terme « ceci » est utilisé par rapport à lui-même, alors qu'il n'acquiert de sens que lorsqu'il se réfère à autre chose. Ces termes sont *relationnels* et ne peuvent donc pas se référer à d'autres relations de manière circulaire ou infinie : ils doivent nécessairement être rattachés à un terme non relationnel.

Remarquez, au passage, que si l'on remplaçait le mot « court » par « courtissime », tout en conservant le même sens, le mot changerait de statut et deviendrait hétérologique, puisque « courtissime » n'est pas « courtissime ». Par ailleurs, quel que soit le mot que l'on substitue à « hétérologique », le paradoxe de Grelling demeure manifeste. Ce test montre que, dans ce dernier cas, ce n'est pas le mot en lui-même qui importe, mais plutôt son sens sous-jacent. Concernant le mot « utile », on pourrait également dire qu'il est utile de par son contenu, ou tout au plus de par le fait qu'il soit un mot (une unité de langage), et non en raison de sa forme ou de sa sonorité spécifiques.

8. Le paradoxe de Russell (version améliorée)

D'après A Fortiori Logic, appendice 7.5.

La logique nous permet de transformer des perceptions concrètes éparses en concepts abstraits et structurés. La connaissance humaine, ou opinion, repose sur l'expérience, l'imagination et l'intuition rationnelle. Cette dernière est une forme d'« expérience » au sens large, une expérience non phénoménale, que l'on pourrait appeler « intuition » logique. C'est pourquoi les anciens, en Occident comme en Orient, nommaient la raison le « sixième sens » ou « sens commun », c'est-à-dire l'organe sensoriel qui relie les cinq autres sens, ceux qui nous mettent en contact empirique avec l'expérience phénoménale : couleurs, formes, sons, odeurs, goûts, sensations tactiles, etc., qu'elles soient perçues physiquement ou imaginées mentalement. Sans le sixième sens, les cinq sens produisent un chaos incohérent (ils sont dénués de sens, on ne peut leur donner de sens) ; et inversement, le sixième sens est inutile sans les cinq autres, car il n'a rien sur quoi fonder une intuition rationnelle. L'imagination remanie les données et raisonnements expérimentaux antérieurs, rendant possible la formation de nouvelles idées et théories qui seront ensuite testées au regard d'expériences et de raisonnements ultérieurs.

1. Éléments de la logique des classes

La logique s'est initialement développée comme une science principalement en référence au discours naturel, donnant naissance à ce que nous appelons aujourd'hui la logique des prédicats. Dans le discours humain naturel, nous (vous, moi et tous les autres) pensons et discutons couramment des choses que nous avons perçues, ou finalement conçues, au moyen de propositions catégoriques impliquant un sujet (disons S) et un prédicat (disons P) liés par la copule « est ». Ces propositions ont la forme « S est P », qui peut être au singulier ou au pluriel, et dans ce dernier cas général (ou universel) ou particulier, et positif ou négatif, et peut en outre impliquer divers modes et catégories de modalité [27].

Une proposition de la forme « S est P » est en réalité une double prédication : elle nous indique qu'une chose qui est S est aussi P. Ainsi, S et P sont tous deux des prédicats, bien que l'un (le sujet S) soit privilégié afin de « prédiquer » l'autre (le prédicat P) [28]. S désigne généralement un ou plusieurs phénomènes concrets (qu'ils soient physiques, mentaux ou spirituels), c'est-à-dire une entité individuelle ou un ensemble d'entités, et P une propriété de cette entité ou de ses entités. Par exemple, « Jean est un homme » et « Tous les hommes sont des êtres humains » sont respectivement une prédication singulière (concernant un seul homme, Jean) et une prédication plurielle (concernant tous les hommes).

---

[27] Il n'est pas nécessaire d'entrer ici dans le détail de ces distinctions, car elles sont bien connues. Il existe également de nombreuses distinctions subtiles entre les différents types de termes pouvant apparaître dans les propositions comme sujets ou prédicats ; mais restons simples.

[28] Le terme « prédication » désigne la copule et le prédicat ensemble, comme s'il s'agissait d'une action du locuteur (ou de l'énoncé fait) sur le sujet.

De plus, toujours dans le discours naturel, le sujet de nos pensées peut être constitué de prédicats *en tant que tels* , c'est-à-dire de prédicats en leur qualité de prédicats ; par exemple : « “hommes” peut être le sujet ou le prédicat d’une proposition ». Ce dernier cas se rencontre spécifiquement dans le discours philosophique (ou logique ou linguistique) ; par exemple, dans le présent essai.

Or, à travers les âges, et plus particulièrement à l'époque moderne, les logiciens ont constaté que le discours naturel était insuffisant pour leurs besoins et ont progressivement développé un langage plus artificiel : celui des « classes » [29]. Ce type de discours est très proche du discours naturel, mais plus abstrait et descriptif, ce qui facilite le discours philosophique (ou logique ou linguistique) et le rend plus précis. Dans ce langage, au lieu de dire « ce S *est* P », on dit « ce S *est un membre (ou une instance) de* P » (notez bien l'allongement de la copule de « est » à « est un membre (ou une instance) de »). Si « ce S » symbolise un individu concret, alors « P » est appelé une « *classe* » ; mais si « ce S » symbolise une classe abstraite, alors « P » est appelé une « *classe de classes* ».

Une classe est donc une abstraction, une construction mentale par laquelle nous regroupons, de manière figurative, des éléments concrets (qu'ils soient physiques, mentaux ou spirituels). Bien que nous puissions classer mentalement des choses, et que nous le fassions temporairement, sans nommer la classe à laquelle elles

---

[29] L'exposé suivant sur la logique des classes s'appuie sur ma présentation dans **Future Logic** , chapitres 43 à 45. Le mot « classe » vient du latin * *classis* * , qui désigne un « groupe appelé au service militaire » (Merriam-Webster). J'ignore si les Anciens employaient ce mot dans son sens logique, ou un terme similaire, dans leurs discours, mais il est certain qu'ils raisonnaient selon une logique de classes. La distinction d'Aristote entre espèces et genres, ainsi que l'arbre de Porphyre, illustrent ce type de pensée.

appartiennent, nous nommons généralement les classes (c'est-à-dire que nous leur attribuons un mot ou une expression distinctive) car cela facilite la mémorisation et la communication. Nommer n'est pas l'essence de la classification, mais c'est un outil précieux pour les classifications à grande échelle. Le nom d'une classe d'éléments ne les représente pas simplement, mais les désigne plutôt, attire notre attention sur eux ; autrement dit, c'est un instrument d'intention.

Une classe, au sens premier, est une classe d'objets en général ; une classe, au sens secondaire, est plus spécifiquement une classe *de classes* . L'appartenance est donc de deux ordres : l'appartenance de non-classes à une classe, ou l'appartenance de classes à une classe de classes. On peut également parler de classes du premier ordre et de classes du second ordre pour distinguer ces deux types. Il n'existe pas d'autres ordres de classes. Lorsque nous pensons ou discutons d'objets plus concrets, nous utilisons la logique des classes du premier ordre ; lorsque nous pensons ou discutons de classes du premier ordre, nous utilisons la logique des classes du second ordre, et cette dernière s'applique également aux classes du second ordre, puisqu'elles sont elles aussi des classes. Il ne faut pas confondre les deux ordres de classes avec la hiérarchie des classes au sein de chaque ordre.

La relation entre classes de classes et classes est analogue à la relation entre classes et objets concrets ; c’est une relation de subsomption. Lorsqu’une classe inférieure (de premier ordre) appartient à une classe supérieure (de second ordre), il ne s’ensuit *pas* que les membres de la classe inférieure appartiennent également à la classe supérieure ; en effet, s’ils appartiennent à l’une, ils n’appartiennent certainement pas à l’autre. Ainsi, par exemple, vous et moi, bien que nous appartenions à la classe « hommes » parce que nous sommes des hommes, nous n’appartenons pas à la classe « classes d’hommes »

parce que nous ne sommes pas des « hommes ». De même, la classe « hommes » n'est pas un homme, mais appartient à la classe « classes d'hommes ». Les membres de la classe « classes d'hommes » (ou plus brièvement, « classes d'hommes »), qui est une classe de classes, sont, outre la classe générale « hommes », les classes plus restreintes « jardiniers », « ingénieurs », « sages », « névrosés », etc.[30]

La hiérarchisation, en revanche, désigne des classes, *au sein d'un ordre donné* , qui partagent des instances non pas par un simple chevauchement partiel, mais de telle sorte que tous les membres d'une classe appartiennent à l'autre (et dans certains cas, mais pas tous, réciproquement). Par exemple, puisque tous les hommes sont des animaux, mais que tous les animaux ne sont pas des hommes, la classe « hommes » est *une sous-classe (ou espèce) de* la classe « animaux », et la classe « animaux »… Il s'agit *d'une surclasse (ou d'un genre) de* la classe « hommes ». Si deux classes ont exactement les mêmes instances, ni plus ni moins, on dit qu'elles sont *coextensives* (une classe qui sert à la fois d'espèce et de genre dans un certain contexte est dite *sui generis*). Si deux classes partagent seulement quelques instances, on dit qu'elles sont *intersectantes (ou chevauchantes)* , mais elles ne sont pas hiérarchisées (par exemple, « jardiniers » et « ingénieurs »). Si deux classes du même ordre n'ont

---

30 Notez que le mot « hommes », employé sans guillemets, se rapporte à un *prédicat* . Pour désigner la *classe correspondante* , on dit « la classe des hommes » ou « la classe des hommes ». À l'écrit, on peut utiliser les guillemets avec ou sans guillemets, ou simplement écrire « hommes » entre guillemets. Lorsqu'il s'agit de classes de classes, on dit « la classe des classes d'hommes » ou « la classe des classes d'hommes », et on peut utiliser les guillemets avec ou sans guillemets, ou simplement écrire « classes d'hommes » ou « classes d'hommes » entre guillemets.

aucune instance en commun, on dit qu'elles sont *mutuellement exclusives* .

Il est important de saisir et de garder à l'esprit la distinction entre hiérarchie et ordre. Puisque nous sommes des hommes, chacun de nous appartient à la classe « hommes » ; il s'agit d'une subsomption de ses instances concrètes par une classe de premier ordre. Puisque tous les hommes sont des animaux, la classe « hommes » est une sous-classe de la classe « animaux » ; il s'agit d'une hiérarchie entre deux classes de premier ordre. Puisque « hommes » est une classe d'animaux, il appartient à la classe « classes d'animaux » (ou « classes animales ») ; il s'agit d'une subsomption de ses instances de classe de premier ordre (c'est-à-dire de simples classes) par une classe de second ordre (c'est-à-dire une classe de classes). Puisque toutes les « classes d'hommes » sont des « classes d'animaux », la classe « classes d'hommes » est une sous-classe de la classe « classes animales » ; il s'agit d'une hiérarchie entre deux classes de second ordre, c'est-à-dire entre deux classes de classes. La relation entre les classes de premier ordre et les classes de second ordre n'est jamais une relation de hiérarchie, mais toujours une relation de subsomption. Autrement dit, les premiers sont toujours des membres (instances) des seconds, jamais des sous-classes. Les hiérarchies n'existent qu'entre des classes de même ordre.

Ainsi, en logique des classes, nous avons deux plans d'existence à considérer. Au niveau le plus fondamental se trouve le plan relativement objectif des phénomènes empiriques (qu'ils soient de substance physique, mentale ou spirituelle) ; au-dessus, résidant dans notre esprit, se trouve le plan relativement subjectif des idées (conçues comme insubstantielles, mais possédant des aspects phénoménaux – à savoir des images mentales ou physiques, des mots prononcés ou écrits, et les intentions de ces signes), comprenant des idées sur les phénomènes

empiriques et des idées sur ces idées. Les classes sont développées pour faciliter l'étude des phénomènes empiriques et les classes de classes sont développées à leur tour pour faciliter l'étude des classes – car les classes (y compris les classes de classes) sont bien sûr elles-mêmes des phénomènes empiriques d'une certaine sorte. La classification est une invention humaine utile pour l'ordonnancement cognitif des choses observées par nos sens, notre imagination ou nos intuitions introspectives. Bien que les classes soient des produits de l'esprit, cela ne signifie pas qu'elles soient arbitraires – elles sont formées, organisées et contrôlées par notre faculté rationnelle, c'est-à-dire à l'aide de la logique.

Pour être qualifiée de classe, une classe doit clairement posséder au moins un membre (auquel cas ce membre unique est « unique en son genre »). Généralement, une classe possède deux membres ou plus, voire un nombre incalculable. Une classe est finie si elle comprend un nombre spécifié d'instances ; si ce nombre est difficile à dénombrer, la classe est dite ouverte (c'est-à-dire infinie ou du moins indéfinie). Ce qui unit les instances d'une classe, c'est la possession d'une propriété distinctive commune ; la classe est définie par cette propriété (qui peut bien sûr être une combinaison complexe de plusieurs propriétés). Une classe sans instances est appelée classe nulle (ou vide) ; cela signifie que sa propriété définissante est arbitraire, de sorte qu'à proprement parler, il ne s'agit pas d'une classe.

Il est donc important de noter que le terme « classe » est quelque peu ambigu, car il peut désigner une classe de premier ordre (une classe de non-classes, c'est-à-dire d'éléments autres que des classes) ou une classe de second ordre (une classe de classes, c'est-à-dire une construction mentale regroupant deux ou plusieurs de ces constructions mentales). Une classe (de premier ordre) n'est pas, et ne peut pas être, une classe de classes (c'est-à-dire une classe

de second ordre). Il existe, bien sûr, une classe appelée « non-classes » ; ses instances sont principalement toutes les choses concrètes, qui ne sont pas elles-mêmes des classes ; par exemple, vous et moi sommes des non-classes. Les « non-classes » constituent simplement une classe, et non une classe de classes, puisqu'elles n'incluent aucune classe. Ainsi, on peut dire que les « non-classes » sont une classe de premier ordre, mais elles ne constituent pas une classe de second ordre.[31]

Le domaine des classes de classes est très limité comme objet d'étude, comparé à celui des classes simples. Quelles distinctions peut-on établir entre les classes ? Peu nombreuses. On peut distinguer les classes des classes de classes, les classes finies des classes ouvertes, les classes positives des classes négatives [32], et peut-être quelques autres choses, mais guère plus.

---

31 Il convient de noter que, si les termes positifs se traduisent aisément dans le langage de la logique des classes, les termes négatifs présentent une réelle difficulté. Par exemple, alors que le terme « hommes » ne désigne que les non-classes, son antithèse stricte, le terme « non-hommes » au sens le plus large, inclut à la fois les non-classes (c'est-à-dire les choses concrètes autres que les hommes) et les classes (c'est-à-dire les choses plus abstraites). De même, alors que le terme « classes finies » ne désigne que les classes, son antithèse stricte, le terme « classes non finies » au sens le plus large, inclut à la fois les classes ouvertes (abstraites) et les non-classes (concrètes). Ainsi, par souci de cohérence, nous devons admettre que certains termes couvrent à la fois les non-classes et les classes (y compris les classes de classes). Concrètement, cela signifie que nous devons utiliser des *disjonctives* qui révèlent les alternatives implicites. Ceci, bien entendu, complexifie considérablement la logique des classes.

32 Les classes positives sont définies par une propriété positive, et les classes négatives par une propriété négative. Par exemple, « hommes » est défini par référence aux animaux rationnels (positif), tandis que « célibataires » est défini par référence aux hommes non mariés (négatif).

2. Un double paradoxe apparent

Bertrand Russell (Grande-Bretagne, 1872-1970) a proposé une distinction entre « une classe qui est membre d'elle-même » et « une classe qui n'est pas membre d'elle-même ». Bien que toute classe soit nécessairement coextensive à elle-même (et, en ce sens, incluse dans elle-même), il ne s'ensuit pas que toute classe soit membre d'elle-même (certaines le sont, d'ailleurs, et d'autres non). On peut justifier cette distinction par des exemples convaincants. Ainsi, la classe « classes positives » est membre d'elle-même, puisqu'elle est définie par une propriété positive ; tandis que la classe « classes négatives » n'est pas membre d'elle-même, puisqu'elle est également définie positivement (bien que de manière générale par rapport à la négation). De même, la classe « classes finies » n'est pas membre d'elle-même, puisqu'elle possède une infinité d'éléments ; tandis que la classe « classes ouvertes » est membre d'elle-même, puisqu'elle possède également une infinité d'éléments.

Qu'en est-il de la classe « classes » ? Est-elle membre d'elle-même ou non ? Puisque « classes » est une classe, elle est nécessairement membre de « classes », c'est-à-dire d'elle-même. Cette affirmation est faite sans tenir compte de la distinction entre les classes du premier et du second ordre. Si l'on pose la question plus précisément, la réponse se nuance. La classe « classes du premier ordre », étant une classe de classes et non une simple classe, ne peut être membre d'elle-même, mais seulement membre des « classes du second ordre » ; les membres des « classes du premier ordre » sont tous de simples classes. En revanche, puisque la classe « classes du second ordre » est une classe de classes, elle est membre d'elle-même, c'est-à-dire membre des « classes du second ordre ». Ainsi, la classe « classes du second ordre » inclut à la fois elle-même et la classe « classes du premier ordre », de

sorte que lorsque l'on dit que la classe plus large « classes » est membre d'elle-même, on entend plus précisément qu'elle est membre de la classe plus restreinte « classes de classes ». Quant à la classe des « non-classes », puisqu'il s'agit d'une classe et non d'une non-classe, elle n'est pas membre d'elle-même. Notons toutefois que le paradoxe de Russell ne fait pas de distinction entre les classes du premier et du second ordre, mais s'intéresse aux « classes » en général.

Russell se demandait **si « la classe de toutes les classes qui ne sont pas membres d'elles-mêmes » était ou non membre d'elle-même** . Il semblait logiquement impossible de répondre à cette question, car dans les deux cas, une contradiction en découlait. En effet, si la classe « classes non membres d'elles-mêmes » *n'est pas* membre de la classe « classes non membres d'elles-mêmes », alors elle est bien membre de « classes non membres d'elles-mêmes » (c'est-à-dire d'elle-même) ; et si la classe « classes non membres d'elles-mêmes » *est* membre de « classes non membres d'elles-mêmes », alors elle est aussi membre de « classes qui sont membres d'elles-mêmes » (c'est-à-dire de sa contradiction). Cela ressemblait à un double paradoxe stupéfiant.

La solution du problème . La quête du savoir est une entreprise humaine, et par conséquent, elle procède par tâtonnements. Le savoir est bien plus inductif que déductif ; la déduction n'est qu'un outil parmi d'autres de l'induction. Il existe des absolus dans le savoir humain, mais ils sont rares. Lorsqu'on formule une théorie, il s'agit toujours essentiellement d'une hypothèse, qui pourrait ultérieurement nécessiter une révision ou une réfutation. Tant qu'elle paraît utile et solide, et plus que toute autre théorie concurrente, on l'adopte ; mais si jamais elle se révèle contredite par des faits ou source d'antinomie, on est contraint de la reformuler ou de l'abandonner. C'est le principe de l'induction. Face à une contradiction, il faut «

vérifier ses prémisses » et les modifier si nécessaire. Dans le cas présent, puisque le paradoxe de Russell révèle une faille dans notre conception de la logique des classes, il nous faut revenir en arrière et identifier précisément notre erreur. Revenons donc attentivement sur notre démarche. Nous avons défini une classe et l'appartenance à une classe en transformant la prédication en classification, ce qui revient à dire :

> Si quelque chose est X, alors c'est un membre de la classe 'X', et non un membre de la classe 'nonX'.
>
> Si quelque chose n'est pas X, alors ce n'est pas un membre de la classe 'X', mais un membre de la classe 'nonX'.

D'où nous vient cette définition ? Il ne s'agit pas d'un absolu qui nous aurait été imposé par la pensée. Nous l'avons inventée ; c'était une convention grâce à laquelle nous avons élaboré l'idée de classes et d'appartenance à celles-ci. La connaissance peut très bien progresser sans recourir à cette idée, et elle l'a fait pendant des millénaires, et continue de le faire dans l'esprit de nombreuses personnes. C'est une idée qui a une histoire, ajoutée à l'arsenal des techniques de raisonnement par des logiciens d'une époque relativement récente. Ces logiciens ont constaté que, comme d'autres, ils raisonnaient par classification, et ils ont réalisé qu'il s'agissait d'un artifice utile, distinct de la prédication, mais qui, d'une certaine manière, s'en inspire. Ils ont donc formellement proposé la définition ci-dessus et ont entrepris d'étudier la question plus en détail afin d'en maximiser l'utilité. La « logique des classes », ou « logique de classes », était née.

Cependant, à un moment donné, un logicien, Bertrand Russell, réalisa l'existence d'une incohérence inhérente à notre conception de la classification, ce qui mit sérieusement en doute l'édifice même de la logique des classes. Ce fut la découverte du paradoxe qui porte son

nom. Ce fut une découverte majeure, car rien n'est plus important pour le développement des connaissances, et en particulier pour celui de la logique formelle, que le maintien de la cohérence. Chaque découverte d'incohérence nous incite à affiner et à perfectionner notre savoir. Russell mérite d'être reconnu pour cette découverte, même s'il lui fut difficile de résoudre le paradoxe de manière pleinement convaincante. Essayons ici d'aller plus loin, en explorant plus profondément que lui les processus de pensée impliqués dans la classification. Qu'est-ce que la classification, plus précisément ?

Si l'on examine de plus près notre définition de la classe « X » et de l'appartenance à cette classe du seul fait d'être X, il convient de se demander : concrètement, à quoi sert cette définition ? Se contente-t-on de substituer l'expression « est membre de » à la copule « est », et la classe « X » au prédicat X ? Si tel est le cas, cela n'a aucun sens, car il est évident que changer *le nom* d'une relation ou d'un terme ne le modifie en rien. Les mots sont accessoires à la connaissance ; ce qui importe, c'est leur intention sous-jacente, leur signification. Si les mots changent sans que le sens ne change, rien d'essentiel n'a changé. Non, il ne s'agit pas ici de simplement changer les mots employés ; nous proposons une image mentale.

Notre conception de la classification repose sur des *entités mentales* appelées classes, au sein desquelles des choses autres que des classes (ou des classes inférieures, dans le cas de classes de classes) sont *métaphoriquement rassemblées et contenues* . Lorsque nous disons que des choses appartiennent à la classe « X », nous entendons par là que la classe « X » est une sorte de boîte *dans* laquelle ces choses sont, par le biais de l'imagination, rangées (à un instant donné, de façon temporaire ou permanente). Autrement dit, notre « définition » de la classification est

en réalité une convention formelle servant à instituer cette image. Ce qu'elle signifie réellement est le suivant :

> Si quelque chose est X, alors il *appartient à* la classe 'X', et n'appartient pas à (c'est-à-dire *n'est pas dans*) la classe 'nonX'.
>
> Si quelque chose n'est pas X, alors il est hors de (c'est-à-dire pas dans) la classe 'X', mais dans la classe 'nonX'.

Il est clair qu'« être » quelque chose et « être dans » (à l'intérieur de) quelque chose d'autre *ne sont pas synonymes* . Notre définition conventionnelle (c'est-à-dire par accord mutuel) stipule que si X est prédiqué de quelque chose, alors nous pouvons concevoir cette chose comme étant contenue dans l'entité mentale appelée classe « X ». Mais cette définition n'est pas absolue ; elle n'affirme pas qu'être soumis à la prédication de X *implique naturellement et nécessairement* être membre de la classe « X ». Car l'idée même de classification, et donc cette définition de ce qui constitue une classe et l'appartenance à celle-ci, est une invention humaine. Cette invention peut être, et est en effet, très utile – mais elle reste soumise aux lois de la nature. Si nous constatons que notre conception, c'est-à-dire notre définition, conduit inévitablement à une contradiction, nous devons la modifier afin d'éviter toute contradiction. C'est ainsi que nous raisonnons et agissons dans toutes les situations similaires.

Comme nous le montrerons plus loin, puisque la contradiction découle de l'implication définissante mentionnée précédemment, nous devons modifier cette implication. Autrement dit, nous devons lui imposer des limites. Bien entendu, nous ne pouvons nous contenter d'affirmer vaguement qu'elle a des limites ; nous devons définir précisément ces limites afin de restaurer la valeur pratique de notre concept de classification. Nous pouvons

y parvenir en réalisant que notre définition de la classification, en référence à l'« appartenance » d'une chose à une autre, implique que la *logique des classes est conçue comme liée à la logique géométrique* . Cela devient évident lorsqu'on considère le fait que nous « représentons » souvent les classes par des figures géométriques (notamment des cercles) et leurs membres par des points à l'intérieur de ces figures. Cette pratique n'est pas fortuite, mais constitue l'essence même de notre idée de classification. Classer, c'est imaginer que nous plaçons certains éléments, identifiés par la possession d'une propriété commune et distinctive, dans un contenant étiqueté [33].

Examinons maintenant le concept d'auto-appartenance à la lumière de ces réflexions. Qu'est-ce que l'idée d'auto-appartenance ? C'est la présupposition qu'une classe puisse être membre d'elle-même. Mais cette notion est-elle véritablement concevable ? Si nous laissons de côté un instant la question de la logique des classes et la reformulons en termes de logique géométrique, nous constatons qu'elle est absurde. Un contenant peut-il se contenir lui-même ? Bien sûr que non. Il n'existe aucun exemple connu d'un contenant se contenant lui-même dans le monde physique ; et en effet, nous ne pouvons même pas nous le représenter visuellement. Ainsi, l'idée d'auto-contenu n'a aucun fondement empirique, pas même dans la sphère mentale. Ce n'est qu'une conjonction fantaisiste de deux mots, sans base expérimentale. Pour cette raison, l'idée nous paraît illogique et nous pouvons affirmer sans risque comme un « axiome » universel et éternel que l'auto-contenu est impossible. Un terme absurde comme « la collection de toutes les collections » est nécessairement un terme vide de sens. Nous ne

[33] Il s'agit d'une « représentation » picturale, une image analogique à ne pas prendre au pied de la lettre.

sommes pas obligés de l'accepter, et en réalité, nous n'y sommes même pas autorisés logiquement ; nous ne pouvons parler de manière cohérente que de « l'ensemble de toutes *les autres* collections » [34].

Un contenant est toujours coextensif à lui-même, c'est-à-dire qu'il occupe exactement l'espace qu'il occupe. Mais cette « coextension » n'est pas un contenant, et encore moins un auto-conteneur, car elle ne concerne pas réellement (si ce n'est verbalement) deux choses, mais une seule ; il n'y a pas de préfixe « co- », il s'agit simplement d'une extension, une seule fois. On parle de contenant lorsqu'un objet plus petit s'insère dans un objet plus grand (ou, à la limite, lorsqu'un *autre* objet de même taille s'insère parfaitement dans un certain objet). Le concept de contenant se réfère à deux objets, et non à un seul. Il doit y avoir deux objets distincts ; il ne suffit pas de désigner le même objet de deux manières différentes. Imaginer un « auto-conteneur » revient à imaginer qu'un objet entier puisse en quelque sorte s'insérer dans lui-même en tant qu'objet plus petit (ou qu'il puisse en quelque sorte se diviser en deux, l'un des deux étant contenu dans l'autre). C'est *inconcevable* . Un tout ne peut être une partie (totale ou partielle) de lui-même ; rien ne peut être à la fois entier et partie. Une seule chose ne peut être deux choses (qu'elles soient de même taille ou de taille différente) à la fois ; rien ne peut exister simultanément en deux choses.

On ne peut décider par convention qu'une chose est à la fois entière et partielle, ni qu'une chose est double. On ne peut réunir ce qui est naturellement impossible. On ne peut réunir que ce qui est naturellement possible, même si cela est superflu. Ainsi, le concept d'auto-contenu est

---

[34] Pour donner une image concrète : un sac de billes (seul ou, pire encore, avec les billes à l'intérieur) ne peut pas être mis dans lui-même, même si le sac dans son ensemble, avec tout son contenu, peut être roulé comme une bille et donc être appelé une bille.

dénué de sens ; c'est un concept inévitablement vide, car il présuppose la possibilité de l'impossible. L'auto-contenu n'existe pas ; un contenant ne peut jamais se contenir lui-même. Si cela est vrai, alors il est tout aussi vrai qu'aucune classe ne s'inclut elle-même, car (comme nous l'avons vu) la classification est essentiellement une idée géométrique. Étant donné qu'un contenant ne peut se contenir lui-même, il s'ensuit que la réponse à la question de savoir si une classe peut être un membre d'elle-même est indubitablement et définitivement : non. Car affirmer d'une classe qu'elle est un membre d'elle-même revient à impliquer qu'un contenant peut être son propre contenu. De même qu'aucun contenant n'est son propre contenu, *aucune classe n'est son propre contenu !*

Voici une idée révolutionnaire pour la logique des classes. Elle s'applique à toute classe, et non seulement à la classe des « classes qui ne sont pas membres d'elles-mêmes » à l'origine du paradoxe de Russell. De plus, il est important de noter que nous nions ici la possibilité qu'une classe soit membre d'elle-même, mais non la possibilité qu'elle *ne* le soit pas. Dire qu'aucun contenant ne se contient lui-même implique que tout contenant ne se contient pas lui-même. De même, dire qu'aucune classe n'est membre d'elle-même implique que toute classe n'est pas membre d'elle-même. Autrement dit, tout en reconnaissant l'objet du paradoxe de Russell, à savoir la classe des « classes qui ne sont pas membres d'elles-mêmes », nous rejetons l'idée qu'une telle classe puisse un *jour* , même hypothétiquement, être membre d'elle-même (et donc paradoxale) ; car, affirmons-nous, aucune classe n'est jamais membre d'elle-même.

Comment cela est-il possible, vous demanderez-vous peut-être ? N'avons-nous pas déjà démontré par l'exemple que certaines classes sont membres d'elles-mêmes ? N'avons-nous pas convenu, par exemple, que la classe « classes », étant une classe, est nécessairement membre de

la classe « classes », c'est-à-dire d'elle-même ? Comment nier une évidence pareille ? Certes, objecterez-vous, si la classe « classes qui ne sont pas membres d'elles-mêmes » n'est pas membre d'elle-même, alors elle l'est incontestablement ; et si elle l'est, alors elle ne l'est incontestablement pas ? Pour répondre à ces questions légitimes, revenons à notre définition de la classification et à ce que nous avons dit à son sujet. Comme je l'ai souligné précédemment, notre définition des classes et de l'appartenance à celles-ci prend la forme d'une implication conventionnelle. Elle stipule :

> Si et seulement si quelque chose est X, alors c'est un membre de la classe 'X'.

Puisque cette implication conventionnelle nous conduit inéluctablement au paradoxe, nous devons la réviser, c'est-à-dire en restreindre la portée, en précisant les conditions exactes de son application et de ses limites. Son principal défaut, comme nous l'avons vu, réside dans l'idée qu'une classe peut être un élément d'elle-même. Par exemple, puisque la classe « classes » est une classe, elle est un élément de « classes » ; dans cet exemple, la variable X a la valeur « class » et la variable « X » a la valeur « classes ». Or, comme nous l'avons démontré, affirmer qu'une classe peut être un élément d'elle-même implique logiquement une impossibilité géométrique : qu'un contenant puisse être son propre contenu. Ainsi, pour éviter l'apparition du paradoxe de Russell, nous devons prévenir les conséquences indésirables de notre définition. Étant donné que notre concept de classification est problématique en l'état, quelles sont les conditions à spécifier pour le délimiter et résoudre le problème ?

La réponse à cette question est que lorsque le sujet et le prédicat de la proposition subordonnée sont identiques, la proposition subordonnée n'est plus sous-entendue. Autrement dit, si la proposition subordonnée est de la

forme « si la classe 'X' est X », la proposition subordonnée « alors la classe 'X' est un membre de 'X' (et donc d'elle-même) » *ne s'ensuit pas* . Ce « ne s'ensuit pas » est une convention, tout comme l'était le « s'ensuit » en général. Nous avons simplement établi une distinction, en précisant que la proposition subordonnée suit *généralement l'antécédent, sauf dans le cas particulier* où le sujet et le prédicat de l'antécédent sont identiques (en ce sens que le prédicat X s'applique à la classe 'X', elle-même fondée sur le prédicat X). Il s'agit là d'une mesure logiquement tout à fait acceptable. Si l'on constate qu'une proposition générale induite comporte des exceptions, il est tout à fait légitime, voire obligatoire, de la rendre moins générale, en ne la restreignant que juste assez pour tenir compte de ces exceptions.

Puisque la définition initiale de la classification était une convention générale, il est tout à fait permis, lorsqu'on constate que cette convention conduit à une contradiction, de convenir d'une convention légèrement plus restrictive. Ainsi, bien que, dans la grande majorité des cas, il demeure vrai que si quelque chose est X, alors il appartient à la classe « X », et plus précisément, si une classe (disons « Y ») est X, alors elle (c'est-à-dire « Y ») appartient à la classe des classes « X », il n'en demeure pas moins que, ***exceptionnellement, dans le cas particulier où la classe qui est X est la classe « X » (c'est-à-dire lorsque « Y » = « X »), on ne peut affirmer qu'elle appartient à « X »*** , car cela reviendrait à dire qu'elle appartient à elle-même, ce qui est impossible puisque cela impliquerait qu'un contenant puisse être son propre contenu. Notons bien que nous ne nions pas, par exemple, que la classe « classes » *soit* une classe ; Nous contestons seulement *l'implication* généralement admise, à savoir que la classe « classes » *est un membre de* la classe « classes ». Nous pouvons continuer à utiliser l'adjectif « est » (car il s'agit d'une simple prédication),

mais nous ne sommes pas autorisés ici à transformer ce « est » en « est un membre de » (car cela constituerait une classification illicite).

Ainsi, le paradoxe de Russell est évité. Autrement dit, concernant la classe « classes non membres d'elles-mêmes » : premièrement, il est tout à fait légitime de supposer que cette classe n'est pas membre d'elle-même, puisque nous savons avec certitude (par la logique géométrique) qu'aucune classe n'est membre d'elle-même ; mais il n'est *pas* légitime d'affirmer que ce fait (qu'elle n'est pas membre d'elle-même) implique qu'elle en est membre, car une telle implication est conventionnellement exclue. Deuxièmement, il n'est *pas* légitime de supposer, même par hypothèse, que la classe « classes non membres d'elles-mêmes » est membre d'elle-même, puisque nous savons déjà (par la logique géométrique) qu'aucune classe n'est membre d'elle-même, et nous ne pouvons donc pas établir par une telle supposition qu'elle n'en est pas membre, même s'il est vrai qu'elle n'en est pas membre.

Comme on peut le constater, notre correction de la définition de la classification, qui la rend moins générale qu'à l'origine en précisant le cas particulier où l'implication sous-jacente ne s'applique pas, permet d'éliminer le paradoxe de Russell. On peut affirmer que la classe « classes qui ne sont pas membres d'elles-mêmes » n'est pas membre d'elle-même, mais on ne peut pas affirmer qu'elle l'est ; par conséquent, les deux volets du double paradoxe sont bloqués. Dans le premier volet, nous avons bloqué l'inférence de « ne pas être membre d'elle-même » à « être membre d'elle-même » ; dans le second volet, nous avons interdit la supposition d'être membre d'elle-même, même si l'inférence de ne pas l'être serait sans conséquence. Ainsi, la réponse à la question posée par Russell – à savoir : « La classe de toutes les classes qui ne sont pas membres d'elles-mêmes est-elle ou non membre d'elle-même ? » – est que cette classe n'est

pas membre d'elle-même, et que le fait qu'elle ne le soit pas ne la rend pas, contrairement aux apparences, membre d'elle-même, car aucune classe n'est membre d'elle-même.

Ainsi, certes, bien que la classe « classes » soit une classe, il ne s'ensuit pas qu'elle soit membre d'elle-même ; bien que la classe « classes de classes » soit une classe de classes, il ne s'ensuit pas qu'elle soit membre d'elle-même ; bien que la classe « classes positives » soit une classe positive, il ne s'ensuit pas qu'elle soit membre d'elle-même ; bien que la classe « classes ouvertes » soit une classe ouverte, il ne s'ensuit pas qu'elle soit membre d'elle-même ; bien que la classe « classes qui ne sont pas membres d'elles-mêmes » soit une classe qui n'est pas membre d'elle-même, il ne s'ensuit pas qu'elle soit membre d'elle-même. Quant à la classe « classes membres d'elles-mêmes », elle ne possède aucun membre. Il convient de souligner que la restriction de classification que nous avons introduite ici est d'une portée très limitée ; elle n'affecte pratiquement pas la logique des classes, car elle ne concerne que quelques cas très particuliers.

Ce qui précède constitue, à mon avis, la *résolution correcte et définitive* du paradoxe de Russell. Nous avons reconnu l'existence d'un problème : le paradoxe de Russell. Nous en avons diagnostiqué la cause : l'hypothèse de la possibilité d' *une auto-appartenance* . Nous avons démontré que cette hypothèse est inconcevable, puisqu'elle implique qu'un contenant puisse se contenir lui-même ; il ne s'agissait pas d'un bricolage arbitraire, notez-le bien, mais d'un raisonnement logique. Nous avons proposé une solution au problème, une solution qui le cible précisément et l'élimine chirurgicalement. Notre remède consiste à *dissocier* la prédication de la classification dans tous les cas où l'auto-appartenance est supposée, et uniquement dans ces cas. Cette solution relève du simple bon sens et non de la spéculation ; elle est simple et élégante ; elle est

convaincante et sans complication ; elle ne modifie pas fondamentalement le concept d'appartenance à une classe, mais en limite légèrement l'application ; elle introduit une restriction, certes, mais clairement circonscrite et minime ; elle n'entraîne aucun dommage collatéral sur les domaines non problématiques de la logique des classes, ni sur la logique en général, et ne requiert donc aucune adaptation supplémentaire de la doctrine logique. Notez en outre que notre solution ne recourt à aucun « système » obscur de logique symbolique moderne, mais est entièrement développée à l'aide d'un langage ordinaire et de concepts et processus largement connus et acceptés.

3. Un peu d'histoire

Examinons maintenant brièvement l'histoire du paradoxe de Russell et voyons comment lui et d'autres logiciens symboliques modernes l'ont traité [35].

Georg Cantor avait déjà mis en évidence, dès 1895, une antinomie dans sa propre théorie des ensembles. En 1902, alors que Gottlob Frege (Allemagne, 1848-1925) s'apprêtait à publier le second volume de ses Principes fondamentaux de l' *arithmétique (Grundgesetze*), Russell l'informa de ce paradoxe. Frege fut totalement convaincu et ne voyait pas comment résoudre les contradictions inhérentes à « la classe des classes qui ne s'appartiennent pas à elles-mêmes ». Il considérait cela comme un problème très grave, déclarant : « La question est de savoir… si l'arithmétique peut véritablement avoir un fondement logique. » Frege tenta d'abord de résoudre le problème en suggérant l'existence possible de « concepts sans classes correspondantes », ou encore en ajustant l'un de ses « axiomes » de manière à ce que :

---

[35] Je fais ici principalement référence au récit de William et Martha Kneale dans *The Development of Logic* , ch. XI.1-2.

> « On dit que deux concepts ont la même extension si, et seulement si, tout objet qui relevait du premier *mais qui n'était pas lui-même l'extension du premier* relevait également du second et vice versa [36]. »

De toute évidence, la suggestion initiale de Frege, selon laquelle il pourrait exister des « concepts sans classes correspondantes », peut être considérée comme une anticipation de ma dissociation de la prédication et de la classification dans certains cas. Cependant, Frege n'a pas précisé dans quels cas une telle dissociation devait se produire. Cela transparaît dans sa suggestion suivante qui, tout en soulignant de manière troublante la difficulté de la notion d'auto-appartenance, ne la rejette pas d'emblée, mais tente plutôt de l'atténuer. Il parle de deux concepts au lieu d'un, et s'efforce d'exclure, par convention, l'extension de chacun dans son ensemble de l'autre, tout en continuant bien sûr d'inclure les objets relevant de cette extension ; ceci montre qu'il n'a pas réalisé que l'auto-inclusion par extension est tout simplement inconcevable.

Il convient de souligner que le paradoxe de Russell concerne une certaine classe (à savoir l'ensemble des classes qui ne sont pas membres d'elles-mêmes) qui est ou n'est pas membre d' *elle-même* , et non d' *une autre* classe. Frege tente de résoudre ce paradoxe en se référant non pas à une seule classe, mais à une paire de classes égales, bien que (à ma connaissance) il n'ait pas démontré que des classes coextensives engendrent un paradoxe comparable à celui de Russell. Il s'ensuit que sa solution proposée n'est pas pertinente. De plus, Frege semble avoir pensé que si tous les éléments appartenant à une classe (disons

---

[36] Kneale et Kneale, p. 654. Italiques ajoutés par eux.

« Y ») appartiennent à une autre classe (disons « X »), alors on peut supposer que *la classe* « Y » appartient à la classe « X », et réciproquement en cas de coextension. Ceci est suggéré par sa tentative d'empêcher une telle supposition, afin d'éviter (selon lui) le paradoxe de Russell qui en résulterait. En réalité, l'affirmation selon laquelle tous les Y sont des X n'implique pas que la classe « Y » appartienne à la classe « X » ; elle implique seulement que la classe « Y » est une sous-classe de la classe « X », ou une classe égale si la relation est réciproque. Il apparaît donc que Frege a confondu les relations d'appartenance à une classe et de hiérarchisation des classes, utilisant un terme vague comme « chute » pour les caractériser toutes deux.

On peut légitimement se demander si une classe égale, une sous-classe, voire une surclasse, peut appartenir à sa classe hiérarchiquement apparentée. De prime abord, cela semble possible. Par exemple, toutes les classes positives sont des classes et appartiennent donc à la classe « classes », et la classe « classes positives » est une sous-classe de la classe « classes ». Cependant, bien que toutes les classes ne soient pas positives (certaines sont négatives), la classe « classes » est néanmoins une classe positive (étant définie par un énoncé positif) et appartient donc à la classe « classes positives ». Mais bien que cet exemple suggère qu'une surclasse puisse appartenir à sa sous-classe (et par conséquent, qu'une classe égale ou une sous-classe puisse appartenir à sa classe hiérarchiquement apparentée), on peut néanmoins exprimer un doute par *analogie* , comme Frege l'a peut-être envisagé. On pourrait argumenter par induction, par généralisation, que si une classe ne peut être membre d'elle-même, alors elle ne peut peut-être être membre d'aucune classe coextensive (comme le suggère Frege), ni même d'une sous-classe ou d'une surclasse. En effet, la question est de savoir si les instances désignées par la première classe peuvent être

considérées comme apparaissant deux fois dans la seconde (à la fois comme membres à part entière et comme constituants d'un membre). Frege a donc peut-être soulevé une question pertinente, susceptible d'entraîner de nouvelles restrictions en logique des classes. Cependant, il n'est pas nécessaire de s'y attarder dans le contexte présent, car (comme expliqué précédemment) cette question n'est pas directement liée à la résolution du paradoxe de Russell.

Russell a décrit son paradoxe dans son ouvrage *Principes des mathématiques* , publié peu après. Bien qu'initialement enclin à privilégier la seconde approche de Frege, il a par la suite préféré la première, proposant l'existence de « fonctions propositionnelles ne déterminant pas de véritables classes ». On notera ici encore l'incapacité à identifier précisément la source et la solution du problème. Par la suite, Russell a pensé que « le problème ne pourrait jamais être complètement résolu tant que toutes les classes ne seraient pas éliminées de la théorie logique ». À mon avis, ce serait jeter le bébé avec l'eau du bain – une réaction excessive. Mais il a ensuite découvert (ou plutôt, il pensait l'avoir découvert, ou s'en est convaincu) que le même paradoxe pouvait être généré sans « parler de classes », c'est-à-dire en se référant à de simples prédicats – autrement dit, en termes de logique des prédicats plutôt qu'en termes de logique des classes. Comme l'ont formulé Kneale et Kneale[37] :

> « Au lieu de la classe censée contenir toutes les classes qui ne sont pas membres d'elles-mêmes, considérons la propriété d'être une propriété qui ne s'illustre pas elle-même. Si cette propriété s'illustre elle-même, alors elle ne peut pas s'illustrer elle-même ; et si elle ne s'illustre pas elle-même,

[37]William et Martha Kneale, p. 655.

> alors elle doit s'illustrer elle-même. De toute évidence, la nature du problème est la même ici que dans le paradoxe originel, et pourtant il n'est pas question de classes. »

Même si les classes ne sont pas explicitement mentionnées ici, il est clair qu'elles sont sous-entendues. Comment une propriété pourrait-elle « s'illustrer » elle-même ? Vraisemblablement, la propriété X serait « une propriété qui s'illustre elle-même » si elle fait partie des choses qui la possèdent. Autrement dit, X illustre X si X appartient à la classe des choses qui sont X. On ne peut parler de propriétés sans recourir à la prédication ; et une fois la prédication effectuée, on peut (étant donné la définition initiale de la classification) classifier, assurément. Cette tentative n'est donc qu'une manipulation verbale ; la même idée est sous-entendue, mais présentée sous un autre jour. C'est malhonnête. De plus, la manière dont le paradoxe est prétendument évoqué ici n'en crée pas un.

On ne peut affirmer, même hypothétiquement, « si cette propriété [c'est-à-dire la propriété d'être une propriété qui ne s'illustre pas elle-même] s'illustre elle-même », car cela serait déjà contradictoire. Pour reconstruire un paradoxe de Russell en termes de « propriété », il faudrait parler de « *la propriété de toutes les propriétés* qui ne s'illustrent pas elles-mêmes » ; nous aurions alors un nouveau terme à analyser, comme en logique des classes. Mais de toute évidence, ce nouveau terme est artificiel et dénué de sens. Là encore, il faut entendre « *la classe de toutes les propriétés* qui ne s'illustrent pas elles-mêmes » – et dans ce cas, nous revenons à la logique des classes. Ainsi, il est important de noter que si Russell avait raison de chercher à déterminer si son paradoxe était un problème spécifique à la logique des classes, ou s'il se

posait également en logique des prédicats, et s'il prétendait avoir établi qu'il se manifestait dans les deux domaines, en réalité (comme nous venons de le démontrer), il n'y est pas parvenu. En réalité, le paradoxe *était* spécifique à la logique des classes. Et il aurait mieux fait d'admettre le fait que d'essayer de l'ignorer.

En réponse à certaines critiques de ses pairs, Russell finit par admettre que tous les paradoxes étaient dus à des cercles vicieux et énonce comme principe, pour les éviter, que « ce qui inclut l'ensemble d'une collection ne doit pas en faire partie ». On peut donc dire que Russell concédait le principe que j'ai également utilisé, à savoir qu'une collection ne peut s'inclure elle-même parmi ses éléments, même si, en réalité, sa formulation suggère qu'il le concevait comme une convention destinée à bloquer des cercles vicieux incompréhensibles plutôt que comme un absolu logique (remarquez qu'il emploie « ne doit pas » plutôt que « ne peut pas »). Il considérait les paradoxes de la théorie des ensembles comme étant « essentiellement de même nature que le vieux paradoxe d'Épiménide (ou du Menteur) ». Cela laisse supposer qu'à ce stade, il percevait son propre paradoxe comme résultant, d'une manière ou d'une autre, d'une autoréférence. À première vue, la proposition « la classe de toutes les classes qui ne sont pas membres d'elles-mêmes est (ou n'est pas) membre d'elle-même » semble contenir une forme d'autoréférence, puisque la proposition « membre d'elle-même » est répétée (positivement ou négativement, au singulier ou au pluriel) dans le sujet et le prédicat [38]. Cependant, on ne peut pas affirmer que l'autoréférence soit précisément le problème.

---

[38] Il convient de noter que si l'autoréférence était au cœur du problème, alors la proposition « la classe de toutes les classes qui sont membres d'elles-mêmes est (ou n'est pas) membre d'elle-même » serait tout aussi problématique, même si elle ne semble pas aboutir à un paradoxe similaire.

Quelques années plus tard, dans un article publié en 1908, Russell proposa une explication plus élaborée du paradoxe de Russell, fondée sur sa « théorie des types ». Il affirmait alors qu'« aucune fonction ne peut avoir parmi ses valeurs quoi que ce soit qui la présuppose, car si tel était le cas, nous ne pourrions considérer comme définis les objets ambigus désignés par la fonction tant que celle-ci ne serait pas définie, et inversement… la fonction ne peut être définie tant que ses valeurs ne le sont pas » [39]. Autrement dit, la question « la classe de toutes les classes qui ne sont pas membres d'elles-mêmes, est-elle ou non membre d'elle-même ? » est fondamentalement viciée, car le sujet demeure à jamais hors d'atteinte. Nous ne pouvons le saisir tant que nous n'avons pas déterminé s'il est ou non membre d'elle-même, et nous ne pouvons faire cette dernière détermination tant que nous n'avons pas résolu la première ; ainsi, le problème est insoluble, c'est-à-dire que la question est insoluble. En réalité, le sujet est un terme cognitivement impossible à formuler, en raison du double dilemme que sa définition implique pour tout penseur.

Il convient de noter ici que l'objectif de l'explication de Russell était en réalité d'invalider la classe négative « classes qui ne sont pas membres d'elles-mêmes », car c'est cette classe qui est à l'origine du double paradoxe qu'il cherchait à résoudre. La classe positive « classes qui sont membres d'elles-mêmes » ne conduit manifestement pas à un double paradoxe : si l'on suppose qu'une classe n'est *pas* membre d'elle-même, il en résulte une contradiction, mais on peut néanmoins affirmer sans contradiction qu'elle l' *est* . En fait, si l'explication de Russell était correcte, la classe positive devrait être aussi illicite que la négative. Car si l'on affirme l'impossibilité qu'une classe se réfère à quelque chose qui n'est *pas*

---

39 Cité par les Kneales, p. 658.

encore établi , comme Russell l'a fait à propos de la classe négative, alors il faut admettre que cette caractéristique se retrouve également dans la classe positive, et il faut la rejeter aussi. Russell ne semble pas l'avoir compris, c'est-à-dire que son remède ne distinguait pas techniquement les deux classes et pouvait donc s'appliquer à l'une comme à l'autre. Pour cette raison, sa tentative de résoudre le paradoxe de Russell par la circularité ou l'infini doit être considérée comme un échec. Dans ma propre théorie, en revanche, c'est la classe positive (celle de l'auto-appartenance) qui est invalide (et vide), puisqu'elle est géométriquement impensable, tandis que la classe négative (celle de la non-auto-appartenance) reste tout à fait légitime (et instanciée), comme on pourrait s'y attendre selon le principe que toutes les affirmations (y compris celle de l'auto-appartenance) devraient être réfutables.

En résumé, Russell conclut qu'une fonction ne peut être une valeur en soi et propose de différencier fonction et valeur en deux « types » distincts, non interchangeables. Or, cette théorie est, à mon sens, trop générale et complexifie considérablement les choses. Comme nous l'avons vu, nous ne pouvons nier que, par exemple, « classes » *soit* une classe ; tout au plus pouvons-nous nier que cela implique que « classes » *soit un membre d'* elle-même. Il s'agit d'une négation de l'auto-appartenance, et non de l'auto-prédication ou de l'auto-référence. Quant aux notions de « types », puis d'« ordres au sein des types », il ne faut pas les confondre avec les idées plus classiques de hiérarchies et d'ordres de classes, exposées précédemment. En réalité, la ressemblance entre les concepts de Russell et ces derniers confère à sa théorie une apparence de crédibilité ; mais cette apparence est illusoire : il s'agit de concepts très différents. La notion de « types » chez Russell est hautement spéculative et loin d'être intuitive ; bien qu'elle puisse sembler résoudre le

paradoxe de Russell, ses implications vont bien au-delà et, de surcroît, invalident des idées traditionnelles qui, en apparence, ne posent pas de problème [40]. En bref, il s'agit d'une mesure approximative et de fortune, et peu convaincante.

Chaque paradoxe auquel nous sommes confrontés est, bien sûr, un signal d'alarme indiquant que nous nous égarons. Ainsi, le paradoxe de Russell peut être considéré comme un signal d'alarme pour Frege, Russell et d'autres logiciens modernes, leur révélant une erreur dans leur approche. Ils se sont efforcés d'en trouver la source, mais sans parvenir à la localiser précisément. Toute la complexité de leur approche symbolique et axiomatique de la logique ne leur a été d'aucune aide, mais a au contraire obscurci la solution. Ceci démontre qu'avant toute tentative de symbolisation et d'axiomatisation, il est essentiel pour les logiciens de bien comprendre le sujet traité, en termes simples et par le bon sens. À ma connaissance, la solution proposée dans cet essai est originale, c'est-à-dire inédite. Si cela s'avère exact, la théorie de la logique des classes développée par les logiciens symboliques modernes, qui constitue encore aujourd'hui le cœur de l'enseignement universitaire, doit être profondément réexaminée et révisée.

4. Un peu d'autocritique

Concernant la résolution du paradoxe de Russell que j'ai proposée il y a plus de vingt ans dans mon ouvrage **Logique du futur** , il convient de préciser ce qui suit. Si je maintiens globalement ma théorie de la logique des classes présentée dans cet ouvrage (chapitres 43 et 44), je dois désormais prendre quelque peu de recul par rapport à

---

[40] Voir pour commencer la critique de Kneales de la « théorie des types » au chapitre XI.2.

ma tentative de résolution du paradoxe de Russell (chapitre 45).

Dans mes travaux précédents, j'ai, à mon crédit, exprimé un grand scepticisme quant à la notion d'auto-appartenance ; mais je ne suis pas parvenu à l'exclure totalement. J'ai déclaré : « Intuitivement, du moins pour moi, l'idée qu'une chose puisse être à la fois contenant et contenu est difficile à accepter », et j'ai même postulé, par une généralisation tirée de plusieurs cas examinés, qu'« aucune classe d'une chose, ni aucune classe de classes d'une chose, n'est jamais membre d'elle-même », à l'exception possible des « choses » ou des « classes de choses » (bien que l'on puisse dire de ces classes qu'elles ne sont membres d'aucune classe, et encore moins d'elles-mêmes [41]) ; mais je n'ai pas pour autant rejeté l'auto-appartenance par principe, ni utilisé ce rejet pour expliquer et résoudre le paradoxe de Russell, comme je le fais dans le présent essai.

Cela se manifeste, par exemple, dans mon acceptation de l'idée que « les classes qui sont membres d'elles-mêmes » sont membres d'elles-mêmes. J'ai fait ce choix en partant du principe que « la possibilité ou l'impossibilité de l'auto-appartenance n'est pas le problème ». En apparence, c'est bien sûr vrai : le paradoxe de Russell concerne la « classe de toutes les classes qui ne sont pas membres d'elles-mêmes », et non « la classe de toutes les classes qui sont membres d'elles-mêmes ». Mais en réalité, comme je l'ai démontré aujourd'hui, c'est faux ; l'acceptation de l'auto-appartenance est la véritable cause du paradoxe de Russell, et la non-auto-appartenance n'est pas problématique en soi.

---

[41] Il convient de noter que c'est dans ce contexte que j'évoque l'idée que la définition de l'appartenance pourrait parfois être erronée. Mais je n'ai pas approfondi cette idée à l'époque, car je n'avais pas analysé ce qu'un tel échec impliquerait formellement.

N'ayant pas dûment écarté l'auto-appartenance, je me suis rabattu sur la seule solution qui me paraissait prometteuse à l'époque : le rejet de la « permutation » de « est (ou n'est pas) membre de lui-même » à « est (ou n'est pas) {membre de lui-même} » (notez l'ajout des accolades). Autrement dit, j'ai proposé l'interdiction logique de transformer la *relation* d'auto-appartenance ou de non-auto-appartenance en un *terme prédicable* . Je comprends maintenant que c'était une erreur : cette action est intervenue *trop tard* dans le processus de réflexion qui a conduit au paradoxe de Russell. C'était une tentative superficielle, traitant un symptôme plutôt que la cause. J'ai agi ainsi, bien sûr, car je pensais que, « de tous les processus utilisés pour développer ces arguments, [le] seul dont la validité était incertaine (non établie) ». Mais en réalité, ce n'était pas la seule cause possible de l'effet ; il y avait un processus *antérieur* , d'une importance plus profonde : la transition de « est » à « est membre de ». Je n'avais pas remarqué à l'époque ce processus antérieur, et encore moins réalisé sa vulnérabilité ; et pour cette raison, je ne l'ai pas attaqué.

De toute évidence, j'étais sur la bonne voie en cherchant un moment du processus de pensée où bloquer le développement du paradoxe de Russell. Mais mon erreur a été de choisir ce moment trop tard. En réalité, le moment opportun se situe plus tôt, comme je le défends dans cet essai. Le paradoxe de Russell ne résulte pas d'une permutation illicite, mais de la transformation illicite d'un prédicat en une classe dans les cas où une affirmation d'auto-appartenance pourrait en découler. Et bien que le remède proposé soit encore aujourd'hui, d'une certaine manière, « conventionnel », le défaut qu'il vise à corriger est bien réel : l'auto-appartenance est en fait impossible et ne peut donc jamais être considérée comme vraie. Ma solution précédente ne faisait qu'empêcher le paradoxe de Russell ; elle n'empêchait pas l'auto-appartenance, qui en

est la véritable cause. Ainsi, la solution que je propose dans cet essai est plus profonde et plus juste.

9. Plus d'informations sur le paradoxe de Russell

Extrait de Topics (publié uniquement en ligne pour le moment).

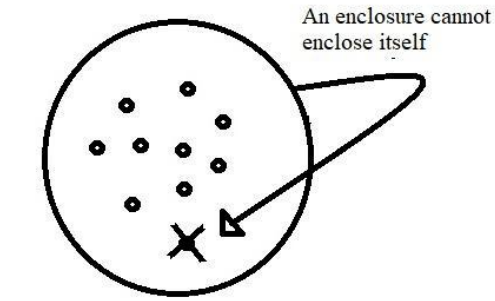
1. Ma résolution du paradoxe de Russell

Ma résolution du paradoxe de Russell est, en termes simples, *qu'aucune* classe n'est, ni ne pourra jamais être, membre d'elle-même. L'auto-appartenance est *inconcevable* . Il n'est donc pas surprenant que l'adoption irréfléchie de cette idée par Russell l'ait conduit au paradoxe. Le concept d'appartenance à une classe diffère sensiblement de celui de prédication. La prédication est un mode de pensée naturel ; l'appartenance à une classe est une construction artificielle, inventée par les logiciens. Le concept d'appartenance à une classe est utile, mais il a ses limites. L'idée d'appartenance à une classe est très ancienne, puisqu'elle est déjà sous-entendue dans les discours d'Aristote ; mais elle a été utilisée plus consciemment par les logiciens à l'époque moderne.

L'idée originale est que si un individu (appelons-le X) « possède » une certaine propriété (appelons-la Y-ité), alors il est « subsumé » sous l'espèce ou le genre des choses possédant Y ; c'est-à-dire que ***si X est Y, alors X est un Y*** (Remarquez l'emploi de « un » pour indiquer qu'une espèce ou un genre, et non une propriété, est visé par Y). Cette réflexion a ensuite mené à l'idée d'une « classe avec des membres », définie comme suit : ***si X est un Y, alors X est un membre de la classe « Y ».*** (Remarquez l'emploi

des guillemets pour désigner une classe, par opposition à un prédicat). Sont implicitement sous-entendus ici les compléments : « si X n'est pas Y, alors X n'est pas un Y » et « si X n'est pas un Y, alors X n'appartient pas à la classe « Y » ».

Il convient toutefois de préciser que ces deux idées, bien que liées, sont tout à fait distinctes. On peut dire de tout ce qui est Y qu'il est *un* Y ; en revanche, on ne peut dire d'une classe « Y » quelconque, qui se trouve être un Y, qu'elle est membre de la classe « Y » (c'est-à-dire membre d'elle-même). En effet, l'idée d'appartenance à une classe est, au sens figuré, celle d'un enclos (la classe) dans lequel sont inclus des éléments (les membres) qui se ressemblent d'une certaine manière (c'est-à-dire qui partagent une propriété, par exemple Y) ; or, il est *impossible d'imaginer visuellement* que cet enclos soit inclus en lui-même (c'est-à-dire que la classe puisse être membre d'elle-même). On peut l'exprimer verbalement, certes ; mais on ne peut le représenter graphiquement.



A class cannot be a member of itself

Ainsi, les notions de subsomption sous une espèce ou un genre (de « Y » sous Y) et d'appartenance à une classe (de « Y » dans « Y ») ne sont que partiellement équivalentes

et ne doivent pas être confondues. Par exemple, bien que la classe « classes » soit effectivement une classe, elle ne peut logiquement pas être membre de la classe « classes ». Les caractérisations « est une classe » et « est membre de la classe 'classes' » ne sont pas interchangeables. La classe « classes » peut cependant être membre d'autres classes, notamment de la classe « classes non membres d'elles-mêmes ». En effet, aucune classe ne peut être membre des « classes membres d'elles-mêmes », car une telle classe n'existe pas (comme expliqué précédemment) ; par conséquent, *toutes* les classes sont membres des « classes non membres d'elles-mêmes ».

Il est important de noter que, même si certaines classes (comme la classe « classes ») *s'incluent* elles-mêmes, on ne peut raisonnablement dire qu'elles sont *membres d'* elles-mêmes. Cela signifie que, quelle que soit la valeur de Y (***que la classe « Y » soit ou non une classe Y), « Y » n'est* pas *membre de la classe « Y »,*** c'est-à-dire qu'il n'est pas membre d'il y a une classe. Ainsi, la proposition par laquelle nous avons initialement défini l'appartenance à une classe, à savoir si X est une classe Y, alors X est membre de la classe « Y », comporte *une exception formelle* , à savoir le cas où la classe « Y » est visée par le terme X (qui est par ailleurs général).

(Bien que nous n'ayons pas traité ici des sous-classes et des surclasses, notez que l'exception mentionnée n'empêche pas la classe « Y » d'être membre de la *surclasse* « Y-classes », qui, par définition, inclut « Y » ainsi que toutes ses sous-classes. Ceci est vrai car « Y » et « Y-classes » se situent à des niveaux différents, ce qui exclut une auto-appartenance. Plus de détails ci-dessous.)

Une autre particularité de la logique des classes est que, de même que les « classes » constituent une classe, les « non-classes » peuvent être considérées comme une classe,

sans contradiction. En logique des prédicats, une telle contradiction est impossible ; autrement dit, une non-classe (c'est-à-dire une chose qui n'est pas une classe) ne peut être considérée comme une classe : seule une classe peut être considérée comme telle. Ceci confirme la divergence entre ces deux modes de pensée. Par ailleurs, il est important de noter que l'on ne peut pas affirmer systématiquement que les « classes » constituent une non-classe ; et de même, on ne peut pas affirmer que les « non-classes » constituent une non-classe.

Concernant le paradoxe de Russell, on prétend que la classe de « toutes les classes qui ne sont pas membres d'elles-mêmes » constitue un double paradoxe, car : si l'on dit qu'une classe n'est pas membre d'elle-même, alors elle *est membre d'elle-même* , et si l'on dit qu'elle *est membre d'elle-même* , alors elle n'est pas membre d'elle-même. Or, il est clair, au vu de ce qui précède, que le problème réside dans la conséquence de la première proposition conditionnelle et dans l'antécédent de la seconde, c'est-à-dire dans l' affirmation *erronée* selon laquelle elles contiennent l'idée que « toutes les classes qui ne sont pas membres d'elles-mêmes » peuvent être membres d'elles-mêmes. En réalité, *aucune* classe n'est membre d'elle-même ; par conséquent, les deux propositions conditionnelles sont fausses (non-sequiturs).

En d'autres termes, la résolution du paradoxe de Russell est que la classe « toutes les classes qui ne sont pas membres d'elles-mêmes », *tout comme toute autre classe* , n'est *pas* membre d'elle-même, et cette affirmation ne *donne* lieu à aucune contradiction.

2. Pourquoi les résolutions de Russell sont inadéquates

Or, Bertrand Russell a formulé une explication assez similaire de son paradoxe, en disant :

> « Aucune collection (dans son ensemble ou sa totalité) ne peut contenir des membres qui sont définis en fonction d'elle-même ; plus précisément, aucune collection existante ne peut jamais être une partie constitutive d'elle-même [42]. »

En quoi cette formulation diffère-t-elle de mon refus de la possibilité logique de l'auto-appartenance ? La formulation de Russell se réfère manifestement indirectement à la collection elle-même ; autrement dit, il veut dire, plus directement : une collection ne peut se contenir elle-même comme membre, car une telle appartenance ne pourrait être déclenchée qu'en *définissant la collection en fonction d'elle-même* , ce qui est impossible. Où se trouve la « définition », ici ? C'est la formulation même du terme, à savoir « toutes les classes qui sont membres d'elles-mêmes » dans notre cas.

La définition proposée repose sur le terme qu'elle définit ; autrement dit, elle comporte une forme d'autoréférence, de circularité. On peut parler d'autodéfinition ; or, elle est manifestement inutile si l'on ignore déjà, par d'autres moyens, ce qui est sous-entendu. Mais la définition de la classe en fonction d'elle-même est-elle ici le véritable enjeu ? Je dirais que non : ce terme n'est que le résultat d'un raisonnement ; il n'apparaît pas ex nihilo. C'est ce raisonnement qu'il convient d'examiner et d'évaluer.

La classe de « toutes les classes qui sont membres d'elles-mêmes » est -*elle produite par* auto-définition ? Non – le prédicat sous-jacent « toutes les classes qui *sont*

---

42 Citation tirée de *Principia Mathematica* de Nicholas Rescher dans *Paradoxes: Their Roots, Range, and Resolution* , p. 172. Ce « principe du cercle vicieux » a été formulé auparavant par Henri Poincaré.

des classes » n'est pas problématique : en effet, toutes les classes sont des classes. Cette prédication engendre le terme classificatoire « toutes les classes qui *sont membres des* "classes" », qui n'est pas problématique en soi non plus. Ce qui est problématique, c'est que ce dernier terme, utilisant par exemple le fait que « classes » est une classe, et donc apparemment membre d'elle-même, donne inductivement naissance au terme « classes qui sont membres d' *elles-mêmes* ». Ainsi, la genèse de ce dernier terme problématique ne résulte pas d'une simple auto-définition, comme Russell semble le penser, mais de la traduction illégitime (ou supposée traduction) d'une prédication parfaitement légitime en une proposition d'appartenance à une classe.

La traduction en question est connue et prouvée illégitime, car elle engendre le paradoxe de Russell. Ma solution se démontre aisément à l'aide de l'expression négative « toutes les classes qui *ne sont pas* membres d'elles-mêmes ». Si la question portait sur l'autodéfinition, comme le suggère Russell, cette expression serait elle *aussi* problématique, puisqu'elle se réfère également à elle-même (ne serait-ce que négativement) ; or, elle ne l'est pas ! Il est [43]donc clair que le problème n'est pas l'autodéfinition, mais tout autre chose : l'inférence naïve que je viens de décrire.

---

[43] Autrement dit, si la seconde formulation est jugée problématique, alors la première ne peut l'être. Quoi qu'il en soit, l'une des deux doit être non problématique pour que le discours soit possible. Je mentionne cette alternative car, *à première vue,* il peut sembler que la formulation négative « toutes les classes non membres d'elles-mêmes » soit problématique, puisqu'elle semble aboutir à un paradoxe (alors que la formulation positive, elle, ne semble pas en engendrer). En réalité, cette formulation négative est non problématique, et c'est la formulation positive « toutes les classes membres d'elles-mêmes » qui est à l'origine du paradoxe. Il est important de bien le noter.

Plus juste est la conclusion de Russell, à savoir qu'« aucune collection existante ne peut être une partie constitutive d'elle-même » – mais cela ne constitue pas une explication en soi : c'est une conséquence évidente du paradoxe de Russell, c'est-à-dire l'affirmation qu'il existe manifestement un problème avec l'idée qu'une « collection existante » puisse être « une partie constitutive d'elle-même », ce qui implique qu'une telle collection ne peut exister. Ce n'est pas une explication, mais une observation qui nécessite une explication, et l'explication proposée par Russell s'articule autour de l'autodéfinition, comme nous l'avons vu. Or, l'autodéfinition n'est pas, en réalité, d'ordre génétique.

Je dirais que, bien que Russell ait entrevu un problème inhérent à l'idée d'auto-appartenance (ou, alternativement, peut-être, de non-auto-appartenance) d'une classe, il n'en a pas saisi précisément l'origine. Il a, à tort, identifié le problème comme une auto-définition, c'est-à-dire une auto-référence dans la définition. Or, bien que cette caractéristique soit présente et significative, elle n'en constitue pas le cœur.

En réalité, le problème ne peut être résolu que par l'analyse conceptuelle plus fine que j'ai formulée. Il découle de l'hypothèse erronée selon laquelle la subsomption sous un concept et l'appartenance à une classe seraient des notions parfaitement parallèles (pratiquement interchangeables) – or, elles ne le sont pas. C'est cette hypothèse initiale erronée qui est à l'origine du paradoxe de Russell. Une fois ce point compris, il devient aisé de saisir pourquoi l'« auto-appartenance » est inconcevable et nous devons affirmer catégoriquement qu'aucune classe n'est membre d'elle-même ; dès lors, le paradoxe de Russell disparaît.

Plus tard, Russell, apparemment insatisfait de la solution qu'il avait initialement proposée à son double paradoxe, a

élaboré sa théorie des types pour y remédier. Dans sa forme la plus simple, cette théorie s'intéresse aux différences entre les éléments individuels inclus dans les classes, les classes elles-mêmes, les classes de classes qui les englobent, et ainsi de suite. Il s'agit là de « types » successifs, formant une hiérarchie d'éléments allant des objets concrets à leurs dérivés de plus en plus abstraits.[44]

En résumé, cette théorie postule qu'une proposition peut être universellement vraie pour un type inférieur, sans pour autant s'appliquer à un type supérieur. Par exemple, on peut dire que toutes les classes appartiennent à la classe « classes », sans pour autant que « classes » en fasse également partie, simplement parce que (selon cette théorie) elle est d'un type supérieur à toutes les classes qu'elle inclut. Ainsi, « toutes les classes » ne signifie pas littéralement *toutes* les classes ; cette expression peut (en raison de la différence de type) exclure la classe « classes » elle-même. Bien sûr, cette affirmation est tout à fait arbitraire et ne fait que masquer le paradoxe de Russell.[45]

Cette idée comporte une part de vérité, mais elle engendre aussi une certaine confusion. Il existe bien une hiérarchie entre les éléments individuels concrets, les regroupements abstraits d'éléments (les classes) et les groupes de groupes d'éléments encore plus abstraits (les classes de classes). Cependant, le paradoxe de Russell ne se pose pas lorsqu'on tente d'inclure une classe dans une classe de classes, c'est-à-dire une sous-classe dans une surclasse. Et bien sûr, le paradoxe de Russell ne se pose pas non plus

---

44 Je n'entrerai pas ici dans une analyse détaillée de cette théorie. Je renvoie les lecteurs à mes analyses précédentes, plus approfondies, dans mes ouvrages **Future Logic** (chapitres 43 à 45) et *A Fortiori Logic* (Annexe 7.5).

45 Je renvoie les lecteurs à l'excellente exposition et critique du paradoxe de Russell et de ses tentatives de résolution, dans *The Development of Logic de William et Martha Kneales* , chapitre 11.

lorsqu'on tente d'inclure un élément individuel dans une classe. Le paradoxe de Russell n'apparaît que lorsqu'on tente d'inclure une classe dans elle-même ; ou (par extension) une classe de classes dans elle-même.

Le paradoxe de Russell n'apparaît *que* lorsqu'on tente d'inclure une classe, ou une classe de classes, *dans elle-même* . La hiérarchie n'est donc pas en cause ici, et la théorie des types ne peut apporter de solution. Il est seulement fortuit qu'une classe ne puisse regrouper que des objets concrets, ou qu'une classe de classes ne puisse regrouper que des classes moins abstraites. La raison du paradoxe ne réside pas dans les différences de types, mais dans le fait qu'une classe, ou une classe de classes, ne peut, même par imagination, s'inclure elle-même.

Notons au passage que, dans notre exemple précédent, la classe « classes » (c'est-à-dire la classe de toutes les classes) est en réalité une classe de classes, puisque les unités qu'elle regroupe sont toutes des classes. Néanmoins, elle demeure une classe à part entière et, de ce fait, semble s'inclure elle-même, c'est-à-dire se subsumer. La solution n'est pas qu'elle ne puisse pas se subsumer elle-même, mais seulement qu'elle ne puisse pas être membre d'elle-même. Il ne s'agit pas d'une question de hiérarchie (de type), mais d'une question de transformation de la prédication en classification.

3. Pourquoi la résolution de Rescher est inadéquate

La solution proposée par Nicholas Rescher au paradoxe de Russell diffère quelque peu, mais pas fondamentalement, de la première tentative de Russell. Il applique au cas présent un principe général qu'il a formulé, appelé le Principe d'Introduction Réussie. Il souligne que « ce n'est que lorsqu'un élément est correctement identifié qu'il peut faire l'objet d'une

discussion pertinente », ajoutant : « Avec un pseudo-objet mal identifié… la porte est grande ouverte à la contradiction et au paradoxe » [46]. Selon lui, le paradoxe de Russell résulte donc d'un défaut d'identification propre à une classe définie par elle-même.

Ainsi, Rescher, tout comme Russell, se concentre sur le problème de l'autodéfinition. Son approche diffère, comme il le souligne, en ce que, tandis que Russell conçoit l'autodéfinition comme une question ontologique, Rescher la perçoit comme une question de communication. Pour Russell, une classe ne peut exister si elle est définie par rapport à elle-même ; tandis que pour Rescher, on ne peut même pas commencer à utiliser ou à discuter d'une telle classe, car c'est comme si rien n'avait encore été dit. Ces deux intuitions sont, à mon avis, justes. Mais elles ne résolvent pas le paradoxe de Russell, comme nous l'avons déjà expliqué. Pour rappel, il ne s'agit pas d'un problème d'autodéfinition, mais du passage d'un discours prédicatif à un discours classificatoire. Une

---

[46] Op. cit., p. 172-173. Bien que ce principe SIP soit intuitivement pertinent et adapté au cas présent, il faut préciser qu'il ne s'agit pas d'une vérité générale. En réalité, une grande partie du discours humain se construit à l'aide de termes plus ou moins vagues. Rares sont les termes, voire aucun, à être définis de manière complète et définitive. Cela tient au fait que notre connaissance est essentiellement inductive, et non déductive (comme l'imaginent de nombreux logiciens). Nous utilisons un mot (lorsque nous en avons besoin) et une définition opérationnelle (lorsque nous en avons une) comme outils pragmatiques. Notre compréhension de leur référence est toujours provisoire et temporaire. À mesure que nos connaissances s'enrichissent et que nos analyses s'affinent, nos termes gagnent (en principe) en précision et deviennent plus spécifiques. Le principe proposé ici par Rescher ne doit donc pas être interprété de façon trop générale. Des discussions pertinentes ont lieu constamment avec des mots imparfaitement définis ; c'est la norme plutôt que l'exception, et nous parvenons généralement très bien à éviter les contradictions. Le sens n'est pas essentiellement une question de verbal, mais une question d'intuition et de compréhension.

classe comme « classes » *est* bien une classe, mais « classes » ne peut pas être *membre de* la classe « classes ». Ce sont deux systèmes de pensée différents, qui ne se comportent pas de manière identique en toutes circonstances.

On ne peut pas former des concepts à volonté, simplement en juxtaposant des mots dans une définition. Une définition est une hypothèse soumise à vérification logique. Dans le cas de la définition de l'appartenance à une classe, à savoir la proposition « si X est un Y, alors X est un membre de la classe Y », un double paradoxe apparaît dans le cas particulier où « si Y est un Y, alors Y est un membre de Y ». Par conséquent, la règle que nous avons tenté de formuler (cette définition de l'appartenance à une classe) comporte nécessairement une exception : lorsque X représente la classe Y.

Le paradoxe de Russell est précieux car il nous enseigne qu'une exception doit être admise dans la formule de définition. Nous supposons initialement la définition générale valide ; mais lorsqu'on découvre qu'elle conduit à une contradiction et à un paradoxe, on est contraint de la rétracter partiellement. Logiquement, nous sommes forcés de rejeter totalement la notion d'auto-appartenance. Cette rétractation, très limitée et précise, rétablit la cohérence de la logique des classes. Notre définition hypothétique a été ajustée pour éviter toute incohérence. C'est la résolution correcte du paradoxe de Russell.

4. Autres incarnations du paradoxe

Il convient de mentionner que Russell avait conscience que la notion de classe était en jeu, mais il s'efforça ensuite de démontrer que ce même paradoxe pouvait être

formulé au sein de la prédication. Comme l'ont formulé Kneale et Kneale[47] :

> « Au lieu de la classe censée contenir toutes les classes qui ne sont pas membres d'elles-mêmes, considérons la propriété d'être une propriété qui ne s'illustre pas elle-même. Si cette propriété s'illustre elle-même, alors elle ne peut pas s'illustrer elle-même ; et si elle ne s'illustre pas elle-même, alors elle doit s'illustrer elle-même. De toute évidence, la nature du problème est la même ici que dans le paradoxe originel, et pourtant il n'est pas question de classes. »

Il en découle que certaines propriétés s'« illustrent » d'elles-mêmes, tandis que d'autres non. Autrement dit, une propriété Y : dans certains cas, Y est Y (c'est-à-dire qu'elle possède la propriété Y) et dans d'autres, Y n'est pas Y (c'est-à-dire qu'elle ne possède pas la propriété Y). Si Y est Y, alors Y s'illustre d'elle-même ; et si Y n'est pas Y, alors Y ne s'illustre pas d'elle-même. Mais est-il jamais vrai que Y est Y ? Peut-on désigner quelque chose qui ait précisément « lui-même » comme prédicat ? Je dirais que non. De toute évidence, « Y est Y » ici ne renvoie pas à une tautologie prédicative, comme « le blanc est blanc ». Ce que nous entendons par « Y est Y » ici, c'est que « la blancheur est blanche », c'est-à-dire que la Y- *ité* (*la propriété* Y) est Y (possède la propriété Y). Il s'agit d'une affirmation très différente, qui ne découle pas de la précédente.

On peut lui appliquer la même critique qu'à la version classificatoire. On peut affirmer que si X est Y, alors X « exemplifie » la Y-ité (c'est-à-dire qu'il fait partie des

---

47 *Le développement de la logique* , p. 655.

choses qui possèdent la Y-ité). Cependant, dans le cas particulier où X est remplacé par la Y-ité (c'est-à-dire la propriété Y), il faut faire une exception et nier que « la Y-ité est Y » implique « la Y-ité exemplifie la Y-ité (c'est-à-dire elle-même) ». Pour éviter un éventuel paradoxe de Russell, il faut au moins court-circuiter cette implication, voire rejeter totalement le conséquent (en affirmant qu'il n'est jamais vrai que la Y-ité exemplifie la Y-ité), voire même l'antécédent (en affirmant qu'il n'est jamais vrai que la Y-ité est Y).

Russell affirmait que si la « propriété de non-auto-exemplification » s'exemplifie d'elle-même, alors elle est dépourvue de « la propriété de non-auto-exemplification » et ne peut donc s'exemplifier ; et si elle ne s'exemplifie pas d'elle-même, alors elle possède « la propriété de non-auto-exemplification » et doit donc s'exemplifier d'elle-même. À mon sens, cela signifie que nous devons rejeter l'idée même d'auto-exemplification et affirmer que *rien ne* s'exemplifie d'elle-même (même si nous trouvions par hasard une certaine Y-ité qui soit Y). Dans ce cas, nous pouvons parfaitement affirmer que la « propriété de non-auto-exemplification » ne s'exemplifie pas d'elle-même, sans qu'aucune contradiction n'apparaisse.

On peut exprimer la même chose autrement : certains concepts sont des « instances » d’eux-mêmes, tandis que d’autres ne le sont pas. Autrement dit, on peut définir l’instanciation ainsi : si X est un Y, alors X est une instance de Y (c’est-à-dire que X est un individu appartenant au groupe appelé Y) ; mais il faut faire une exception dans le cas particulier où X est remplacé par Y, et nier que si Y est un Y, alors Y est une instance de Y. Il est clair qu’ici, Y désigne spécifiquement un concept, et non une chose quelconque. Par exemple : « Le concept « concept » est un concept ; par conséquent, le concept « concept » est une instance de lui-même. »

Cela revient manifestement à dire : « La classe de toutes les classes est elle-même une classe ; par conséquent, la classe de toutes les classes est un membre d'elle-même. » Un paradoxe similaire en découle inévitablement, mais il peut être neutralisé de façon analogue. Autrement dit, après réflexion, il faut admettre que *rien* n'est une instance de soi-même, de sorte que, de fait, le « concept de non-auto- instanciation » n'est pas une instance de lui-même. Un raisonnement similaire peut être appliqué à l'« inclusion » et à toute autre relation analogue.

Il est clair qu'aucun de ces problèmes ne concerne la prédication en tant que telle, mais plutôt des notions plus géométriques (classification, possession/appartenance, instanciation, inclusion) par lesquelles nous tentons de la représenter. Ces relations sont conçues comme des métaphores de la prédication ; or, il s'avère que ces analogies ne sont pas parfaites. La leçon du paradoxe de Russell est que toutes ces relations dérivées comportent des exceptions (comme détaillé précédemment). Le paradoxe de Russell n'affecte pas la prédication en tant que telle ; la prédication, en tant que telle, y est insensible. Bien que Russell ait eu raison de se demander si son paradoxe influençait la prédication en tant que telle, il a eu tort de conclure que c'était le cas ; ce n'est pas le cas.

5. À propos du paradoxe du barbier

Il existe un village avec un seul barbier. Certains hommes s'y rasent eux-mêmes ; les autres ne le font pas, mais se font raser par le barbier. Qu'en est-il du barbier ? Se rase-t-il lui-même ou non ? S'il se rase lui-même, il fait partie de ceux qui se rasent eux-mêmes (aucun problème) ; s'il ne se rase pas lui-même, il fait partie de ceux qui se font raser par le barbier, c'est-à-dire par lui-même (apparente contradiction). De toute évidence, dans son cas, les deux catégories d'hommes se

chevauchent : il « se rase lui-même » et il est « rasé par le barbier ». Cela ne pose guère de problème : cela nous indique simplement que la prétendue dichotomie entre les deux groupes d'hommes n'en est pas une ; son cas constitue une exception logique à cette dichotomie. Autrement dit, dans son cas, contrairement aux autres, se faire raser par le barbier n'implique pas de ne pas se raser soi-même.

En quoi cela se compare-t-il au paradoxe de Russell, dont le paradoxe du barbier est souvent cité comme illustration ? Je résous le paradoxe de Russell en affirmant qu'aucune classe n'appartient à une seule classe. Il s'ensuit que l'on peut affirmer que la classe des « classes qui ne sont pas membres d'elles-mêmes » n'est nécessairement pas membre d'elle-même, sans que cela ne constitue une contradiction. Je résous le paradoxe du barbier différemment, en disant que, contrairement à tous les autres hommes de la ville, le barbier peut parfaitement se raser lui-même et être rasé par lui-même. Dans le paradoxe de Russell, l'une des classes conflictuelles est totalement éliminée. Dans le paradoxe du barbier, le conflit entre les classes est atténué dans un seul cas (celui du barbier lui-même). Puisque les deux paradoxes ne sont pas résolus de la même manière, ils doivent être considérés comme logiquement distincts.

10. Le paradoxe de la confirmation de Hempel

Extrait de Réflexions logiques et spirituelles 1.8.

1. Analyse traditionnelle

Carl Gustav Hempel [48]a mis en évidence un prétendu « paradoxe de la confirmation », suggérant l'impossibilité d'une logique inductive formelle parfaitement cohérente. Ce paradoxe, communément appelé « paradoxe du corbeau », peut être décrit comme suit :

a) L'observation que certains corbeaux sont noirs (certains A sont B) confirme l'hypothèse que tous les corbeaux sont noirs (tous les A sont B).

Cette dernière proposition peut être opposée à « Toutes les choses non noires ne sont pas des corbeaux » (Tout ce qui n'est pas B n'est pas A).

b) Considérons ensuite l'observation suivante : « Certaines pommes sont vertes (certains C sont D) ». Ceci se traduit par : « Certaines choses vertes sont des pommes (certains D sont C) ».

Il découle de cette proposition que certaines choses non noires ne sont pas des corbeaux (certains non-B ne sont pas des A), puisque les choses vertes ne sont pas noires et les pommes ne sont pas des corbeaux.

De même que le fait que certains corbeaux soient noirs (certains A sont B) confirme l'hypothèse que tous les corbeaux sont noirs (tous les A sont B), de même le fait

48 Allemagne-États-Unis, 1905-1997.

que certaines choses non noires ne sont pas des corbeaux (certains non-B ne sont pas A) confirme l'hypothèse que toutes les choses non noires ne sont pas des corbeaux (tous les non-B ne sont pas A).

Cette proposition induite peut à son tour être opposée à « Tous les corbeaux sont noirs » (Tous les A sont B), et c'est là que réside la difficulté, car il semble que la simple observation de quelques pommes vertes suffise à confirmer l'hypothèse que tous les corbeaux sont noirs ! Notons bien que, pour parvenir à ce résultat, nous n'avons même pas eu besoin d'observer de corbeaux noirs.

c) Il découle de ce qui précède que nous pouvons tout aussi bien, en utilisant le même processus logique, étant donné que certaines pommes sont vertes, confirmer l'hypothèse que tous les corbeaux sont roses, ou de toute autre couleur (sauf le vert) d'ailleurs.

d) Ceci constitue en soi un mystère : comment les pommes peuvent-elles nous renseigner sur les corbeaux ? Intuitivement, il faut considérer cela comme un *non-sens* . De plus, dans le cas des corbeaux noirs, leur existence est étayée par des observations empiriques, comme indiqué précédemment ; par conséquent, la conclusion selon laquelle tous les corbeaux sont noirs à partir de l'observation de pommes vertes semble encore plausible. En revanche, nous n'avons jamais observé de corbeaux roses ; de ce fait, la conclusion selon laquelle tous les corbeaux sont roses à partir de l'observation de pommes vertes paraît tout à fait injustifiée.

De plus, sachant par observation que certains corbeaux sont noirs, comment pouvons-nous « conclure » que tous les corbeaux sont roses ? Même si nous n'affirmons pas que tous les corbeaux sont noirs, mais seulement qu'ils sont roses, nous soutiendrions alors des propositions contradictoires : la proposition particulière selon laquelle certains corbeaux sont noirs et la proposition générale selon laquelle tous les corbeaux sont roses.

De plus, même si nous n'avons jamais observé la couleur d'aucun corbeau, nous pouvons, selon le processus inductif ci-dessus, conclure simultanément à de nombreuses affirmations contradictoires telles que « Tous les corbeaux sont noirs », « Tous les corbeaux sont roses », « Tous les corbeaux sont oranges », etc. Ceci aussi est un résultat qui contrevient au principe de non-contradiction.
De plus, on peut faire le même raisonnement non seulement pour les pommes vertes, mais aussi pour les pommes de toutes les couleurs (sauf les noires ou les roses, etc.), et même pour des choses autres que des pommes (autres que les corbeaux). Dès lors, presque tout est permis en matière de connaissance, et *nous pouvons à volonté affirmer ou nier à peu près n'importe quoi sur à peu près tout* !
Selon les présentations traditionnelles [49], il s'agit donc du paradoxe de Hempel. Il ressort de cette analyse que les processus inductifs de confirmation d'hypothèses (tels que les généralisations tirées directement de l'expérience ou indirectement de ses dérivés logiques) sont fondamentalement erronés. L'analyse sous-jacente est très *formelle* ; elle peut être effectuée à l'aide de symboles comme A, B, C, D, et possède donc une portée universelle.
Il s'ensuit que l'induction aboutit inévitablement à diverses absurdités : des non-sequiturs apparents, de nombreuses contradictions, et implique finalement l'arbitraire de toute connaissance humaine. De toute évidence, Hempel a mis au jour ici un défi de taille pour la logique inductive et la logique en général.

---

[49] Voir par exemple : l'article de Wikipédia à l'adresse http://en.wikipedia.org/wiki/Raven_paradox.

2. Analyse novatrice

Comme je vais le démontrer en détail, *l'analyse précédente comporte des inexactitudes importantes* . Je montrerai que, bien que Hempel ait effectivement mis au jour un problème formel intéressant pour les logiciens, ce problème ne conduit pas à ce que nous appellerions un paradoxe. Autrement dit, le paradoxe de Hempel (comme on peut continuer à le nommer par convention) recèle de précieux enseignements, mais il ne place pas la logique face à une difficulté insurmontable.

a) La première opération décrite ci-dessus est le processus inductif de *généralisation , couramment utilisé* . Une proposition particulière (Certains A sont B) est transformée en une proposition générale (Tous les A sont B). Le particulier soutient le général de la même manière qu'une preuve positive confirme une hypothèse. Leur relation logique est adductive. « Certains » signifie ici « au moins certains, éventuellement tous et éventuellement seulement certains » ; en généralisant, nous privilégions l'hypothèse « tous » à l'hypothèse « seulement certains ».

Il serait toutefois erroné de considérer que l'affirmation « Certains A sont B » suffit *à elle seule* à justifier par induction l'affirmation « Tous les A sont B ». Une telle généralisation est une inférence inductive admissible *à condition que nous ayons cherché, sans en trouver jusqu'à présent, aucun* A qui ne soit *pas* B. Car si nous *avions* constaté (par observation directe ou par un raisonnement quelconque) que certains A ne sont pas B, nous n'aurions certainement pas généralisé. De plus, si nous rencontrions *ultérieurement* un A qui n'est pas B, nous devrions *particulariser l'affirmation «* Tous les A sont B » pour revenir à « Certains A sont B ».

Cette condition *sine qua non* de la généralisation, à savoir la nécessité de rester vigilant face aux cas contradictoires

et d'ajuster son jugement en conséquence, n'est ni soulignée ni même mentionnée dans la présentation précédente ; il convient de le noter. Or, il s'agit d'une règle reconnue et acceptée de la pensée scientifique au moins depuis l'époque de Francis Bacon, qui insistait sur l'importance du « contre-exemple » dans le raisonnement inductif. Négliger cette condition conduit inévitablement à des contradictions.

Concernant l'opposition entre « Tous les A sont B » et « Tous les non-B ne sont pas A », il s'agit bien sûr d'un raisonnement déductif. Toutefois, il faut garder à l'esprit que la conclusion « Tous les non-B ne sont pas A » découle uniquement de l'inférence inductive préalable « Tous les A sont B » à partir de « Certains A sont B ». Aucune observation n'est requise pour la déduction, mais nous restons tenus de vérifier constamment le raisonnement inductif précédent, c'est-à-dire de rester attentifs aux éventuels cas où A ne serait pas B.

b) Admettons maintenant que « Certains C sont D », comme indiqué précédemment. L'affirmation « Certains C sont D » se convertit aisément en « Certains D sont C ». Cependant, « Certains D sont C » n'implique pas formellement que « Certains non-B ne sont pas A » ; une inférence syllogistique est ici tacitement impliquée et mérite d'être explicitée. De toute évidence, nous tenons tacitement pour acquis les prémisses que *le vert n'est pas noir* et *que les pommes ne sont pas des corbeaux* , d'où découlent les deux syllogismes successifs suivants :

1ère figure, EIO :

Rien de vert n'est noir (Aucun D n'est B)

Certaines pommes sont vertes (Certaines C sont D)

Par conséquent, certaines pommes ne sont pas noires (certaines C ne sont pas B).

3ème figure, AII :

Toutes les pommes ne sont pas des corbeaux (Tous les C ne sont pas des A)

Certaines pommes ne sont pas noires (certaines C ne sont pas B)
Par conséquent, certaines choses non noires ne sont pas des corbeaux (certains non-B ne sont pas des A).
Par généralisation, on obtient : tout ce qui n'est pas noir n'est pas un corbeau (tout ce qui n'est pas B n'est pas A) ; et par contraposition : tous les corbeaux sont noirs (tout ce qui est A est B). Notons que les prémisses qui ont conduit à cette conclusion générale n'incluent pas « Certains corbeaux sont noirs » ; autrement dit, cette conclusion ne repose sur aucune observation empirique de corbeaux noirs.
Notons également que nous aurions pu obtenir le même résultat avec les prémisses « Aucun corbeau n'est vert (Aucun A n'est D) » et « Aucune pomme n'est noire (Aucun C n'est B) ». Notons aussi que, bien que les syllogismes en jeu soient des raisonnements déductifs, toutes ces prémisses implicites nécessitent des observations et des généralisations préalables (c'est-à-dire des inductions) pour être valides.
De plus, il est important de noter que ces syllogismes seraient impossibles à construire si les couleurs des corbeaux et des pommes considérées étaient identiques (tous deux verts ou tous deux noirs), ou si les corbeaux et les pommes ne constituaient pas des classes mutuellement exclusives. Nous supposons également ici qu'un corbeau ne peut avoir qu'une seule couleur (par exemple, être partiellement noir et partiellement vert, ou tantôt d'une couleur, tantôt de l'autre) ; il en va de même pour une pomme.
L'étape suivante consistait à généraliser « Certains non-B ne sont pas des A » à « Tous les non-B ne sont pas des A ». Mais là encore, la généralisation n'est permise que si nous n'avons aucune preuve ni déduction d'aucune autre source indiquant que certains non-B ne sont pas des A. Autrement dit, dans notre exemple, nous devons garder à

l'esprit qu'il est possible que certains êtres non noirs ne soient pas des non-corbeaux, c'est-à-dire des corbeaux, ce qui signifie que nous pourrions encore trouver des corbeaux non noirs (albinos).
Ici aussi, nous devons nous assurer, conformément au principe fondamental d'adduction de Bacon, qu'aucune observation contradictoire ne vienne entraver notre *élan d'expansion* . Ceci est d'autant plus nécessaire que la prémisse de généralisation « Certains non-B ne sont pas-A » a été obtenue indirectement par déduction à partir de résultats d'induction antérieurs, alors que précédemment, notre prémisse « Certains A sont B » était (supposément) observée directement.

Il convient de noter que ces deux généralisations sont soumises aux conditions régulatrices suivantes : « Certains A ne sont pas B » ou « Certains non-B ne sont pas non-A », respectivement, ne doivent pas être vraies. Or, ces conditions sont identiques puisque ces deux propositions sont logiquement équivalentes par contraposition. Cela signifie que, dans les deux cas, que l'on raisonne directement à partir de corbeaux noirs ou indirectement à partir de pommes vertes, la condition implicite de généralisation reste la même : jusqu'à présent, aucun corbeau non noir n'a été observé dans notre expérience ou notre raisonnement.

Ainsi, quelle que soit la généralisation choisie, la condition d'absence d'instances connues de A qui ne soient pas B demeure inchangée, et la vérité de l'affirmation « Tous les A sont B » dépend de cette condition. Notons également que cette même condition est valable avant et après de telles généralisations. Autrement dit, même après un tel processus inductif, si nous découvrons de nouvelles preuves du contraire, nous pouvons et devons logiquement *revenir sur notre conclusion* .

Comme indiqué précédemment en c) et d) : en utilisant le même raisonnement, étant donné que certaines pommes sont vertes, on peut tout aussi bien confirmer l'hypothèse que tous les corbeaux sont roses, et bien d'autres hypothèses farfelues qui se contredisent [50]. De toute évidence, nous commettons une erreur quelque part et devons agir pour éviter de telles conséquences absurdes ou les corriger si elles se produisent. Je vais maintenant expliquer la solution à ce problème.

La généralisation n'est jamais irréversible. Ainsi, si une généralisation conduit à des contradictions, nous sommes libres, et même tenus, de la particulariser. La question demeure : *dans quelle direction précise et jusqu'où devons-nous remonter ?* Cela signifie qu'il n'y a pas de « paradoxe » en logique inductive comme en logique déductive ; presque tout (à l'exception de la logique elle-même, et notamment des lois de la pensée sur lesquelles elle repose) est et restera toujours « négociable ».

En déduction, une contradiction est un événement bien plus grave, car le processus qui y conduit est présumément nécessaire. *En induction, en revanche, nous savons d'emblée que le lien entre la ou les prémisses et la conclusion est conditionnel ; des contradictions sont donc prévisibles, et c'est précisément le rôle de la logique inductive de déterminer comment y répondre.*

Le traitement des contradictions relève de la logique inductive et s'appelle ***l'harmonisation*** ou la résolution des conflits. Ce n'est ni rare ni exceptionnel : c'est un processus constant dans le développement des

---

[50] Pour démontrer que des propositions ayant des prédicats différents sont contradictoires, on utilise le syllogisme. Par exemple, « Tous les corbeaux sont noirs » et « Tous les corbeaux sont roses » sont incompatibles, car sachant que rien de noir n'est rose, on obtient, par syllogisme (1re figure, EAE) : « Aucun corbeau n'est rose », ce qui est contraire à « Tous les corbeaux sont roses ».

connaissances. *Parfois, les conflits sont résolus avant même qu'ils n'apparaissent, parfois après. Si nous les anticipons, nous les prévenons ; sinon, nous procédons aux corrections nécessaires.*

La particularisation d'une proposition générale est une rétractation. Plus largement, la rétractation signifie le rejet ou la modification d'une théorie à la lumière de nouvelles preuves. Ainsi, par exemple : jusqu'à présent, je n'ai vu que des corbeaux noirs et j'ai supposé qu'ils étaient tous noirs ; demain, je pourrais apercevoir un corbeau blanc et revoir mon opinion sur les couleurs possibles des corbeaux.

Hempel semble ignorer cet aspect crucial du raisonnement inductif, sans quoi il n'aurait pas considéré comme paradoxales les contradictions surgissant au cours de l'induction. Néanmoins, la situation qu'il décrit est intéressante dans ce contexte, pour des raisons dont il n'a pas (je crois) perçu la portée.

Après la première généralisation, partant de « Certains corbeaux sont noirs (Certains A sont B) », si l'on découvre tardivement que « Certains corbeaux ne sont pas noirs (Certains A ne sont pas B) », *on revient simplement à notre observation initiale* : « Certains corbeaux sont noirs (Certains A sont B) ». En revanche, après la seconde généralisation, partant de « Certaines choses non noires ne sont pas des corbeaux (Certains non-B ne sont pas A) », si l'on découvre tardivement la même preuve contradictoire, on ne peut pas simplement nier que « Tous les corbeaux sont noirs (Tous les A sont B) ».

Pourquoi ? Parce que *cela nous laisserait encore seuls. Avec une partie de notre généralisation* , à savoir l'affirmation que certains corbeaux sont noirs (certains A sont B). Autrement dit, on s'attendrait à ce que « tous les A sont B » plus « certains A ne sont pas B » aboutisse à la conclusion harmonieuse « certains A sont B et certains A ne sont pas B ». Le cas particulier négatif n'élimine pas le

cas particulier positif sous-jacent à la généralité positive ; puisque nous croyions auparavant (en raison de cette généralisation), il nous reste un élément à expliquer.
Dans le cas de l'affirmation « Tous les corbeaux sont noirs », une telle rétractation n'est pas remarquable, puisque nous savons par expérience que certains corbeaux sont noirs ; mais dans le cas de l'affirmation « Tous les corbeaux sont roses », nous avons un problème sérieux, car il n'existe pas la moindre preuve que certains corbeaux soient roses ! Autrement dit, la rétractation proposée ne peut suffire dans la situation présentée par Hempel, c'est-à-dire lorsque l'affirmation « Tous les A sont B » est déduite de l'affirmation « Certains non-B ne sont pas A ».

3. Le problème et sa solution

À mon avis, c'est là le nœud du problème révélé par l'étude de Hempel. Il ne s'agit pas à proprement parler d'un paradoxe, puisque *la validité de la généralisation repose formellement sur le fait que ce processus ne donne lieu à aucune contradiction* .[51]
Le fait que l'observation de quelques pommes vertes nous permette de généraliser et de conclure que tous les corbeaux sont noirs et que tous les corbeaux sont roses,

[51] Un paradoxe est une thèse qui se contredit formellement ou qui conduit déductivement à des propositions contradictoires. Face à un tel paradoxe, on peut conclure que la thèse en question est fausse ; la logique en tant que telle n'est pas remise en cause, car la contradiction est simplement conditionnelle. Un double paradoxe, en revanche, constitue une menace sérieuse pour la logique ; ici, la thèse *et* sa contradiction sont toutes deux paradoxales, la contradiction est donc *inconditionnelle* . Dans ce cas, la logique ne peut déclarer ni l'une ni l'autre vraie ou fausse, mais doit y déceler soit un non-sequitur (comme dans le paradoxe du barbier), soit un terme dénué de sens (comme dans le paradoxe du menteur). Autrement dit, la logique doit contester une ou plusieurs implications, et/ou un ou plusieurs termes ou thèses. Le scénario de Hempel ne donne pas lieu à un paradoxe inconditionnel/double.

ainsi que bien d'autres conclusions contradictoires, est amusant, mais non inquiétant. Car, face à une telle contradiction, nous pouvons, par une rétractation, retrouver le moyen d'harmoniser nos connaissances. Le problème est temporaire.

En revanche, Hempel a mis en évidence que l'on ne peut pas toujours réfuter une thèse par une simple particularisation. La particularisation semble acceptable dans certains cas (par exemple, avec les corbeaux noirs), mais dans d'autres, elle aboutit à des résultats inacceptables (par exemple, avec les corbeaux roses), car la suite logique de cette réfutation est dépourvue de fondement empirique.

Supposons, en utilisant la méthode de Hempel et en partant des pommes vertes, que nous induisions les deux généralités suivantes : « Tous les corbeaux sont roses » et « Tous les corbeaux sont oranges ». Ces deux conclusions sont contradictoires. Supposons que nous décidions de résoudre ce conflit en les niant toutes les deux ; il nous reste alors deux propositions : « Certains corbeaux sont roses » et « Certains corbeaux sont oranges ».

Ces deux propositions ne sont pas contradictoires – et admettons qu'aucune ne repose sur des bases empiriques –, et pourtant, elles se sont toutes deux insidieusement « implantées » dans notre savoir ! Introduites par des généralisations à partir de pommes vertes, elles n'ont pas été délogées lorsque nous avons abandonné les généralisations correspondantes. Nous y sommes contraints, même si les processus complexes qui y ont conduit ont été remis en question.

Il est inconcevable que de tels *détails* (vrais ou faux) puissent émerger de l'observation sans lien apparent de pommes vertes (ou de tout autre élément). C'est là, à mon avis, le problème fondamental mis en lumière par Hempel. Il ne s'agit pas d'un conflit entre généralités ou entre propositions générales et particulières, ni donc d'un

paradoxe. Le problème concerne les « dommages collatéraux » causés à la connaissance par une correction incomplète des erreurs.

Je propose la solution suivante : lorsqu'on généralise de « Certains A sont B » à « Tous les A sont B », puis qu'on découvre que certains A ne sont pas B, on particularise « Tous les A sont B » pour revenir à « Certains A sont B ». C'est la procédure habituelle, que nous pratiquons tous couramment.[52]

En revanche, lorsqu'on obtient « Tous les A sont B » par généralisation de « Certains non-B ne sont pas A » à « Tous les non-B ne sont pas A » (suivie de la contraposition de cette dernière), alors, lorsqu'on découvre que « Certains A ne sont pas B », on ne peut pas simplement particulariser « Tous les A sont B » pour revenir à « Certains A sont B », mais il faut également rétracter l'intermédiaire prémisse de la proposition Tous les A sont B, à savoir Tous les non-B ne sont pas A, et retour à Certains non-B ne sont pas A.

Compte tenu de cette dernière rétractation, nous n'avons en réalité plus de fondement pour affirmer que certains A sont B (ceci ne peut être déduit de l'affirmation que certains non-B ne sont pas A). Ce serait une erreur d'induction que d'oublier la *source même* de notre croyance que tous les A sont B. Il est essentiel de garder à l'esprit la distinction entre les fondements inductifs « certains A sont B » et « certains non-B ne sont pas A », afin que, dans le cas de la découverte de preuves contradictoires, à savoir que certains A ne sont pas B, nous puissions revenir *exactement à notre position initiale* dans chaque cas.

Nous pouvons donc formuler la nouvelle loi suivante de la logique inductive, que l'on peut appeler ***la loi de la rétractation commensurable*** : un produit de la

---

52 Symboliquement, A + O = IO.

généralisation comme « Tous les A sont B » ne peut être traité sans égard à sa source particulière ; lorsqu'il est un jour réfuté par de nouvelles preuves, nous devons revenir au même particulier initial et non à *un autre* particulier qui était impliqué par la généralité lorsqu'elle semblait vraie, mais qui ne l'est plus maintenant puisqu'elle n'est plus vraie.

Autrement dit, si nous rencontrons une contradiction de ce type, nous devons comprendre qu'elle ne discrédite pas seulement la généralité précédemment induite, mais plus profondément encore l'acte inductif qui l'a engendrée. Par conséquent, nous ne devons pas nous rétracter par simple particularisation, mais vérifier soigneusement si le particulier restant repose sur un fondement indépendant et, dans le cas contraire, remonter suffisamment loin dans le temps jusqu'à l' *état initial* pour nous assurer qu'aucun particulier non confirmé ne subsiste.

Cela semble être une instruction parfaitement raisonnable : annuler et effacer toute trace d'un acte inductif jugé illicite, c'est-à-dire qui nous a conduits à une impasse logique.

Tout cela signifie qu'en utilisant les procédures logiques ordinaires, nous ne nous retrouverions jamais dans une situation contradictoire (par exemple, affirmer paradoxalement que tous les corbeaux sont noirs et que tous les corbeaux sont roses). Le fait que des généralisations puissent aboutir à des résultats incompatibles est courant ; nous gérons quotidiennement de tels conflits sans difficulté. Lorsque de tels conflits surviennent, nous sommes logiquement tenus de les harmoniser. Si nous ne trouvons pas de solution spécifique au conflit, celui-ci est résolu de manière générique, c'est-à-dire que toutes les généralisations concernées sont remises en question.

Dans une situation où deux propositions ou plus sont mises en doute par un conflit mutuel, nous accorderons

naturellement plus de crédit à celle qui repose sur une base empirique directe (comme « Tous les corbeaux sont noirs ») qu'à celle qui résulte simplement d'une projection indirecte (comme « Tous les corbeaux sont roses »). Il n'est pas nécessaire de traiter toutes les propositions contradictoires avec le même degré de doute ; nous pouvons être sélectifs quant à leur origine inductive.

Concernant les arguments en faveur d'une thèse contradictoire : il est évident que, sans données sur les corbeaux noirs ou roses, nous ne saurions quelle généralisation adopter, et les deux approches seraient problématiques. En revanche, si nous avons observé des corbeaux noirs et jamais de corbeaux roses, nous opterions naturellement pour la généralisation « Tous les corbeaux sont noirs » (tout A est B). Par contre, si nous avons observé à la fois des corbeaux noirs et des corbeaux roses, nous ne formulerions aucune généralisation et nous nous contenterions de combiner les deux observations.

Concernant les processus inductifs utilisés, la généralisation directe serait naturellement privilégiée par rapport à la méthode indirecte envisagée par Hempel. Si le conflit en question peut être résolu par des moyens ordinaires, par exemple par des considérations empiriques, il n'est pas nécessaire de revenir sur le processus. Mais lorsqu'aucun autre moyen de décision n'est possible, le processus ferait naturellement l'objet d'une révision.

On pourrait objecter à la loi de rétractation proportionnelle qu'en pratique, il est rare de pouvoir retracer précisément les sources de nos généralisations. Une telle ignorance pourrait se produire et engendrer des perturbations du type de celles décrites par Hempel dans notre savoir.

Cependant, il convient de souligner qu'en *pratique, nous ne nous trouvons quasiment jamais dans la situation décrite* par Hempel. Combien de fois généralise-t-on à

partir d'une proposition comme « Certains non-B ne sont pas-A » ? Statistiquement, la réponse est « probablement jamais » : le paradoxe de Hempel n'est *qu'une possibilité formelle, certes lointaine* , que les logiciens doivent prendre en compte, mais son impact pratique est pratiquement nul.

De plus, il est peu probable que nous parvenions à une proposition de la forme « Certains non-B ne sont pas-A » autrement que par le raisonnement décrit précédemment, c'est-à-dire par l'intermédiaire d'autres termes comme C et D. Nous ne pouvons pas *observer directement* que certains non-B ne sont pas-A. L'observation concerne principalement les phénomènes positifs ; elle peut concerner les phénomènes négatifs, mais seulement indirectement. Cela suggère que si nous étions confrontés à une situation relevant du paradoxe de Hempel, nous serions probablement conscients de son origine.

Il convient également de noter que la solution proposée au problème soulevé par le paradoxe de Hempel peut être qualifiée d'heuristique ; il s'agit d'un travail de réparation par tâtonnement. J'ai cependant déjà proposé dans mon ouvrage *la Logique du futur*:[53] un détaillé, systématique, Traitement formel de l'induction, au moyen de la factorisation et de la révision des formules. Je crois que ce traitement est exempt du problème de Hempel, puisque toutes les possibilités formelles sont incluses dans les formules factorielles développées.

En ce qui concerne les solutions au paradoxe de Hempel proposées par d'autres logiciens, par exemple celles de Goodman et de Quine décrites dans l'article Wikipédia mentionné précédemment :

Nelson Goodman a suggéré d'imposer des restrictions à notre raisonnement, par exemple en ne considérant jamais une instance comme un argument en faveur de « Tous les

[53] Publié pour la première fois en 1990, quelques années avant la mort de Hempel. Voir la partie VI.

P sont Q » si elle pouvait également soutenir « Aucun P n'est Q ». Goodman, puis un autre philosophe, W.V. Quine, ont utilisé le terme *de prédicat projectible* pour décrire les expressions, telles que *« corbeau »* et *« noir »* , qui *permettent* une généralisation inductive ; les prédicats *non projectibles* sont, par opposition, ceux tels que *« non noir »* et *« non corbeau »* qui, apparemment, ne le permettent pas. Quine suggère que la question de savoir quels prédicats, le cas échéant, sont projectibles est une question empirique ; et il note que dans un domaine infini d'objets, le complément d'un prédicat projectible devrait toujours être non projectible. Il en résulterait que, bien que « Tous les corbeaux sont noirs » et « Toutes les choses non noires ne sont pas des corbeaux » doivent être également soutenus, ils tirent tous deux leur soutien des corbeaux noirs et non des non-corbeaux non noirs.

Je trouve ces propositions raisonnables et compatibles avec les miennes. Cependant, je pense que les miennes cernent avec plus de précision le problème et sa solution.

La suggestion de Goodman de restreindre le raisonnement inductif à partir d'une proposition si ce processus aboutit à des conclusions contradictoires est logiquement fondée. Toutefois, cette instruction ne peut être appliquée préventivement, mais seulement *après* avoir constaté que le processus aboutit à des conclusions contradictoires. Il ne s'agit donc pas d'une mesure préventive, comme il semble le penser, mais d'une correction a posteriori. On peut par conséquent la considérer comme équivalente à la loi de rétractation commensurable que j'ai proposée précédemment. La seule différence réside dans le fait qu'il ne semble pas avoir fait de distinction entre le conflit des généralités et les cas particuliers sous-jacents.

Concernant les « prédicats non projectibles », je conviens que les termes négatifs (compléments) posent un problème général en induction. Si la logique déductive ne fait aucune distinction entre termes positifs et négatifs, la

phénoménologie, elle, distingue la présence des phénomènes positifs de leur absence. Alors que l'on peut observer un phénomène positif (comme un corbeau noir) sans tenir compte de sa négation, on ne peut mentionner un terme négatif (comme non noir ou non corbeau) sans avoir d'abord cherché le phénomène positif correspondant, sans le trouver.[54]

Ainsi, un terme véritablement négatif ne peut jamais être véritablement empirique. Son contenu n'est jamais « j'ai vu quelque chose », mais toujours « j'ai cherché avec diligence quelque chose et je ne l'ai pas trouvé ». Une négation est « empirique » dans un sens plus restreint et plus dérivé qu'une affirmation. Elle implique déjà une forme de généralisation, passant de « introuvable » à « n'était pas là pour être trouvé ».

Il découle de cette observation que toute généralisation à partir de termes négatifs, tels que « Certains non-B ne sont pas-A », ne peut être entreprise qu'avec une prudence et un scepticisme extrêmes. Le scénario de Hempel justifie d'ailleurs davantage cette prudence. On pourrait même être tenté, comme solution radicale, d'interdire systématiquement toute généralisation pour un sujet véritablement négatif. Si un discours quelconque a des conséquences illogiques probables, les logiciens ont intérêt à formuler une loi logique préventive de ce type.

Une autre tentation consiste à nier tout contenu significatif aux propositions de la forme « Certains non-B ne sont pas-A ». Une telle proposition est formellement impliquée par « Tous les A sont B », et compatible avec « Certains A sont B », « Aucun A n'est B » et « Certains A ne sont pas B » – mais nous apprend-elle réellement quelque chose ? En effet, puisqu'on ne peut rien inférer sur A ou B (en tant que sujets) à partir de « Certains non-B ne sont pas-A », quelle information contient une

[54] Voir mon essai sur ce sujet dans *Ruminations* (partie I, chapitre 9).

telle proposition ? Ne pourrait-on pas concevoir d'affirmer une telle proposition en utilisant deux termes presque quelconques choisis au hasard ? Ce type de doute pourrait servir à justifier davantage l'interdiction de généraliser à partir de telles propositions.[55]

Toutefois, puisqu'une solution moins radicale est possible, à savoir la loi de rétraction commensurable proposée ci-dessus, il n'est peut-être pas nécessaire d'aller aussi loin. Plutôt que d'interdire d'emblée certains procédés douteux en toutes circonstances, je préfère les autoriser au cas où ils fonctionneraient occasionnellement, et prévoir le mécanisme correctif approprié lorsqu'ils échoueraient.

En résumé, je crois que nous avons démontré ici de manière convaincante que le prétendu paradoxe de Hempel ne pose aucun problème insurmontable à la logique ; il est présenté comme quelque peu plus complexe qu'il ne l'est réellement. Il n'en demeure pas moins une contribution intéressante à la réflexion des logiciens.

---

[55] Ces questions s'éclaircissent si l'on considère la négation éventuelle de « Certains non-B ne sont pas-A », c'est-à-dire la forme « Aucun non-B n'est non-A », ce qui implique « Tous les non-B sont A ». Si cette dernière proposition est vraie, nous aurions un terme négatif (non-B) inclus dans un terme positif (A). On pourrait en déduire que presque tout le monde (à l'exception des choses qui sont B) relève de A. Pour cela, A devrait être un concept très vaste. Un tel concept serait très exceptionnel et presque dénué de sens. Dès lors, on peut dire que « Certains non-B ne sont pas-A » est presque toujours vrai, et en même temps peu informatif.

11. Le paradoxe de la prédiction de Goodman

Extrait de Réflexions logiques et spirituelles 1.9.

1. Le problème présumé

Nelson Goodman [56]proposa une « énigme de l'induction » (selon sa propre appellation [57]) ou un « paradoxe de la prédiction » (selon d'autres appellations) , qui semblait mettre en évidence une difficulté formelle de généralisation. On peut la formuler ainsi :

Goodman a introduit la couleur « grue », qui s'applique à toutes les choses examinées avant un certain instant *t,* au cas où elles seraient vertes, mais aussi à d'autres choses, au cas où elles seraient bleues et n'auraient pas été examinées avant cet instant *t* . Si l'on examine des émeraudes avant l'instant *t* et que l'on constate que l'émeraude *a* est verte, l'émeraude *b* est verte, et ainsi de suite, chaque observation confirmera l'hypothèse selon laquelle toutes les émeraudes sont vertes. Cependant, les émeraudes *a, b, c, ...* etc. confirment également l'hypothèse selon laquelle toutes les émeraudes sont « grue ». Dans ce cas, les émeraudes *a, b, c, ...* examinées après l'instant *t* devraient être « grue », **et donc bleues** ! (Soulignement ajouté)[58]

---

56 États-Unis, 1906-1998.

57 Ou, plus prétentieusement, « le nouveau problème de l'induction ».

58 Je cite ici : http://en.wikipedia.org/wiki/Nelson_Goodman . Ailleurs, on apprend que « s'applique à toutes les choses examinées

L'importance de cet artifice, selon ses partisans, réside dans le fait que, bien que « vert » et « grue » aient la même forme linguistique et devraient donc être soumis aux mêmes processus logiques (en l'occurrence, le processus inductif de généralisation), il s'agit de concepts fondamentalement différents, puisque le premier implique une similarité entre ses instances passées et futures, tandis que le second suggère un changement de couleur au fil du temps, de sorte que le résultat est paradoxalement très différent si l'on généralise en se référant à l'un ou à l'autre.

2. La solution logique

Cependant, comme je vais le démontrer formellement, il ne s'agit là que d'un tour de passe-passe. En effet, bien que la généralisation soit valable pour le vert et le gris, il ne s'ensuit pas que l'on puisse déduire que toutes les émeraudes sont bleues de la proposition générale induite selon laquelle toutes les émeraudes sont gris. Autrement dit, la conclusion « et donc bleues » présentée ci-dessus est une déduction erronée.

Pour mettre en évidence cette simple erreur, il faut reformuler le scénario donné avec plus de précision (les symboles X, A, B, C sont de mon invention) :

- Supposons que nous examinions toutes les émeraudes disponibles (X), jusqu'à un certain moment (t), et que nous les trouvions toutes vertes (A), nous concluons généralement par généralisation que toutes les émeraudes sont vertes (tous les X sont A), bien que nous sachions [par expérience passée avec l'induction en général] que la prochaine émeraude que nous trouverons,

---

avant *t* juste au cas où elles seraient vertes, mais à d'autres choses juste au cas où elles seraient bleues » est la formulation même de Goodman dans sa présentation originale dans *Fact, Fiction, and Forecast* (http://en.wikipedia.org/wiki/Grue_%28color%29).

après le temps t, pourrait bien s'avérer être bleue (B) [ou en effet, être d'une autre couleur [59]] !

- Introduisons maintenant, à la suite de Goodman, un nouveau concept « grue » (C) à définir comme le groupe de classe regroupant toutes les choses qui ont été examinées avant un certain temps t et qui se sont révélées être vertes (A) et toutes les choses *qui n'ont pas* été examinées avant le temps t et qui se trouvent *être* bleues (B) [ou en effet, être d'une autre couleur].[60]

- En appliquant cette définition, tous les X (émeraudes) examinés avant l'instant t et jugés A (verts) sont également C (gris) ; autrement dit, par syllogisme, on peut déduire que certains X sont C. Quant aux cas restants de X, ceux qui ne seront examinés qu'après l'instant t (le cas échéant), *chacun sera soit A (vert), soit B (bleu) (ou d'une autre couleur)* ; en ce sens, ces derniers X sont également C. Par conséquent, la conclusion que tous les X sont C semble raisonnable.

- Mais il ne découle certainement pas logiquement de ce qui précède que toutes les émeraudes seront d'une

---

[59] Il s'agit d'une interpolation personnelle visant à préciser la thèse de Goodman. En effet, rien ne permet de supposer a priori que seules des émeraudes bleues pourraient être trouvées. Nous ne faisons qu'émettre une hypothèse quant à la possibilité de trouver des émeraudes bleues, sans nous appuyer sur des observations concrètes ; par conséquent, toute autre couleur est tout aussi probable (ou improbable). Néanmoins, ma réfutation de Goodman reste valable même sans ce commentaire.

[60] Il convient de noter que ces dernières choses sont simplement décrites comme « non examinées jusqu'à l'instant t [si jamais elles le sont] » ; il ne faut pas confondre cela (comme certains commentateurs l'ont fait) avec « examinées après l'instant t », car quel que soit le nombre de choses que nous examinons finalement, nous n'atteindrons évidemment jamais (ni ne saurons jamais avoir atteint) une énumération complète de toutes ces choses dans l'univers. Notons également que le concept de « grue » est ici défini comme un prédicat général pour tout sujet éventuel (« choses »), plutôt que spécifiquement pour les émeraudes.

autre couleur que le vert, c'est-à-dire que tous les X seront B (c'est-à-dire bleus, ou toute autre couleur non verte) ! Car, correctement comprise, la catégorie C n'est *pas* formulée comme une disjonction de A ou B qui actualiserait nécessairement à la fois certains cas de XA et certains cas de XB.

- En y regardant de plus près, on constate que C inclut d'une part des éléments déjà *identifiés* comme étant de type A (des émeraudes vertes déjà observées) et d'autre part un ensemble d'éléments dont la nature reste *inconnue* , c'est-à-dire soit A, soit B (bleu) [ou même d'une autre couleur]. Cette dernière disjonction représente des résultats *possibles* , et non des résultats *inévitables* . En déduire que XB est un résultat *réel* serait donc un *raisonnement fallacieux* .
- Le fait que nous ignorions si un X futur sera A ne nous permet pas de déduire de cette disjonction de possibilités qu'un X futur sera nécessairement B. Nous ignorons également si un X futur sera B. Il se peut fort bien que l'affirmation « Tous les X sont A » (Toutes les émeraudes sont vertes) reste valable indéfiniment après l'instant t comme avant l'instant t (conformément à la généralisation initiale).
- Les prémisses « Tous les X sont C » et « Tous les C sont A ou B » aboutissent bien à la conclusion syllogistique « Tous les X sont A ou B ». Cependant, la disjonction « A ou B » ne peut être interprétée différemment dans la prémisse majeure et dans la conclusion. La disjonction de la prémisse n'étant pas extensionnelle, celle de la conclusion ne peut être

considérée comme telle [61]. Ce serait commettre le sophisme des quatre termes.

Il ressort donc clairement de notre exposé que *l'introduction du concept de « grue » n'a absolument rien changé aux possibilités inductives offertes par les données fournies* . La conclusion inductive correcte demeure intacte malgré les jeux de logique de Goodman. Ce dernier n'est parvenu qu'à dissimuler habilement son raisonnement *déductif fallacieux* (une mauvaise interprétation du type de disjonction en cause) ; il ne s'agit là que de sophismes.

Dans le brouillard épais de la rhétorique de Goodman, on pourrait croire que les émeraudes bleues sont aussi faciles à prédire que les vertes. Or, ce n'est absolument pas la conclusion logique selon le raisonnement inductif. Pourquoi ? Parce que, dans le cas de l'hypothèse selon laquelle les futures émeraudes observées seront vertes, nous disposons de données concrètes à l'appui : toutes les émeraudes observées, passées et présentes, se sont révélées vertes.

En revanche, concernant l'hypothèse de Goodman selon laquelle des émeraudes bleues apparaîtront, nous ne disposons à ce jour d'aucune preuve empirique. Nous pouvons seulement affirmer qu'il n'est pas impossible que des émeraudes bleues soient un jour découvertes, sans pour autant que cela soit certain. « Pratique » ne justifie pas une prédiction. Cela signifie simplement « imaginable dans le contexte actuel des connaissances ».

En d'autres termes, nous ne disposons que d'un principe épistémologique général nous enjoignant de rester ouverts

---

[61] Autrement dit, la base de la disjonction donnée est « Certains C *pourraient être* B », tandis que la base correspondante de la disjonction prétendument inférée est « Certains X *sont* B ». Or, imaginer un événement ne prouve pas qu'il se produise nécessairement. La conclusion ne découle pas de la prémisse.

à toutes les éventualités, en nous appuyant sur l'expérience passée relative à toutes sortes d'objets, et en reconnaissant que la nouveauté apparaît parfois. Mais cette ouverture d'esprit scientifique n'équivaut pas à une prédiction positive de changements spécifiques. Il s'agit simplement d'un appel, au nom du réalisme, à éviter les préjugés et la rigidité.

3. Le concept de «grue»

Une question que nous devrions nous poser est de savoir si le concept de «grue» de Goodman est un concept bien formé ?

Le terme courant de « vert-bleu » se définirait simplement par « vert et/ou bleu ». Il peut s'avérer utile d'introduire une telle définition, ne serait-ce que pour souligner la proximité des couleurs vertes et bleues, ou encore que certaines choses sont à la fois vertes et bleues, ou parfois vertes et parfois bleues, ou encore que certaines teintes intermédiaires sont difficiles à classer comme étant clairement vertes ou clairement bleues. La frontière entre ces couleurs est, après tout, assez arbitraire.

Étant donné que certaines émeraudes sont vertes, on pourrait en déduire que certaines sont grises. Il serait tout aussi valable d'en déduire que toutes les émeraudes sont vertes ou que toutes les émeraudes sont grises. Cela n'impliquerait aucune contradiction intrinsèque, car affirmer que toutes les émeraudes sont grises n'implique (ni n'exclut) pas qu'il puisse exister des émeraudes bleues. L'affirmation « toutes les émeraudes sont grises » est formellement compatible avec l'éventualité que toutes les émeraudes soient vertes. Il n'y a donc pas de « paradoxe de prédiction ».

Le concept de « grie » de Goodman ne diffère en rien de ce concept ordinaire quant à son implication logique. Sa

différence ne réside pas dans l'utilisation de la disjonction (vert ou bleu), car une telle disjonction est tout à fait courante ; par exemple, le concept de « coloré » signifie (en gros) « rouge, orange, jaune, vert, bleu, indigo ou violet ». La différence significative du concept de Goodman tient à son implication de conditions spatio-temporelles. Ceci sert un objectif rhétorique : brouiller les pistes.

Définir le concept de « vert » comme l'ensemble des choses examinées avant l'instant t et jugées vertes, et de toutes les choses *non* examinées avant t qui se trouvent *être* bleues, implique une certaine contradiction. Si je n'ai pas encore examiné les choses après l'instant t, comment puis-je affirmer avec certitude que certaines d'entre elles sont bleues ? Je ne pourrais faire une telle affirmation *qu'a posteriori* , après avoir examiné certaines de ces choses après l'instant t et les avoir trouvées bleues.

Autrement, il faudrait *qu'un tiers* , ayant examiné certains objets avant l'instant t et les ayant trouvés bleus, l'affirme, et qu'il observe ma situation avant que je n'aie fait de même. Or, l'ensemble des observateurs actuels ne peut logiquement adopter une telle hypothèse concernant des objets qui se trouvent être bleus alors qu'ils n'ont pas encore été observés comme tels. Nous ne pouvons parler de manière cohérente que des objets qui *pourraient encore* être trouvés bleus. C'est pourquoi le concept de « grue » de Goodman est mal formé.

Le terme « grue » se définit principalement comme l'union des choses vertes et des choses bleues ; mais il ne découle pas de cette définition que si certaines choses (comme les émeraudes) sont vertes, alors d'autres choses *de même nature* (c'est-à-dire d'autres émeraudes) doivent être bleues. Dire qu'une catégorie de choses (les émeraudes) est « grue » ne signifie pas que ses exemples doivent couvrir l'ensemble des possibilités incluses dans

le terme « grue ». Le concept de « grue » reste légitime tant que l'on retrouve les prédicats qu'il rassemble (vert, bleu) dispersés dans diverses catégories de choses (émeraudes, mer, etc.).

Ainsi, tout prédicat ordinaire comporte une part d'incertitude quant à son application à des sujets spécifiques. De plus, cette applicabilité peut varier dans le temps : selon notre contexte de connaissance et selon les changements survenant dans les objets observés. Par conséquent, il n'est pas nécessaire d'intégrer de tels facteurs épistémiques et temporels dans la définition des concepts que nous proposons. Ces facteurs sont inhérents à la conceptualisation.

Si Goodman a introduit de telles complications dans sa définition de « grue », c'est parce qu'il voulait réfuter (ou donner l'impression de réfuter) le processus de généralisation que nous utilisons couramment pour développer nos connaissances sur la base d'observations limitées.

Selon le raisonnement inductif, si l'on observe que certains X sont A, et que l'on n'a jusqu'à présent trouvé aucun X qui ne soit pas A, on peut généraliser et affirmer que tous les X sont A. Cela reste vrai tant que l'on n'a aucune preuve de l'existence d'un X qui ne soit pas A. La généralisation implique la prédiction, c'est-à-dire l'affirmation d'une chose concernant des cas de X que nous n'avons pas encore observés et que nous n'observerons peut-être jamais.

Goodman souhaitait démontrer que nous sommes tout aussi justifiés de prédire un résultat négatif (c'est-à-dire différent de A, par exemple B) qu'un résultat positif (c'est-à-dire A) [62]. Il n'a pas saisi la *justification logique*

---

62 Pour ce faire, il lui fallait construire un concept englobant à la fois A et non-A, afin de démontrer formellement que la généralisation pouvait s'effectuer dans les deux sens. Cependant, un

de nos généralisations [63]. Nous ne prédisons pas arbitrairement que les cas de X que nous observerons à l'avenir seront A plutôt que différents de A. Nous nous en tenons simplement à *la même polarité* (A), car c'est la seule pour laquelle nous disposons, à ce jour, de preuves empiriques. En comparaison, prédire *la polarité opposée* (différente de A, dans ce contexte) serait purement arbitraire – une affirmation hasardeuse. Concernant X en particulier, la première hypothèse bénéficie d'un certain soutien empirique, tandis que la seconde n'en bénéficie d'aucun.

Goodman n'a tout simplement pas saisi cette différence de justification entre les deux approches, pourtant évidente pour quiconque prend le temps d'y réfléchir. Il a ainsi omis d'appliquer le principe inductif selon lequel une hypothèse confirmée est toujours préférable à une hypothèse non confirmée. De plus, comme nous l'avons vu, dans son empressement à invalider le raisonnement inductif, il a commis l'une des erreurs les plus élémentaires du raisonnement déductif !

4. Jusqu'où sur l'échelle ?

Derrière l'énigme de Goodman se cache une autre question importante pour la logique inductive : jusqu'à quel point une généralisation peut-elle être légitimement poussée sur une échelle de classification ? Si l'on a généralisé un certain prédicat pour un sujet donné, pourquoi ne pas généraliser davantage à un prédicat plus large ?[64]

---

concept incluant des contradictoires étant non informatif, il y a inclus les contraires, à savoir « A ou B » (où B n'est pas A). Cela masque légèrement le problème, sans pour autant le résoudre.

63 Voir mon ouvrage **Logique du futur** , chapitre 50.

64 J'ai abordé ce sujet (indirectement, en ce qui concerne la logique éthique) dans mon *ouvrage Logique judaïque* , chapitre 13.3.

Considérons un sujet X et deux prédicats quelconques S et G, liés comme *espèce et genre* , c'est-à-dire tels que tout S est G mais que tout G n'est pas S (certains non-S sont aussi G). Remarquons ici que S et G sont des concepts ordinaires, comme vert et coloré.

- Si tous les cas de X que nous avons observés jusqu'à présent sont de type S, et que nous n'avons rencontré aucun cas de X qui ne soit pas S, nous pouvons inférer par induction que tous les X sont S. Cette généralisation reste valable tant qu'aucun cas de X qui ne soit *pas* S n'est trouvé ; mais si des cas de X non S apparaissent finalement, la logique inductive nous oblige à réviser notre jugement précédent et à le particulariser en : certains X sont S et certains X ne sont pas S. Car l'induction procède conditionnellement [65].
- Le même raisonnement s'applique à G. [66]Par ailleurs, admettant que tous les S sont G, on peut déduire, par syllogisme (1re figure, AAA), que tous les X sont G à partir de l'affirmation que tous les X sont [S.] *Ainsi* , on peut postuler que si la généralisation est justifiée pour un sujet X donné, jusqu'au prédicat spécifique S, elle l'est également pour un niveau de classification supérieur, jusqu'au prédicat plus général G. Ce raisonnement est logiquement valide s'il est correctement compris et appliqué.

---

[65] L'adduction, ou généralisation, se justifie par *deux* principes essentiels : (1) la confirmation d'une hypothèse par un exemple positif et (2) la non-réfutation de cette même hypothèse par un exemple négatif. Ces *deux* principes doivent être respectés sans distinction pour que le raisonnement soit logique. Bien entendu, de nombreuses autres conditions entrent en jeu – voir mon essai « Principes de l'adduction » dans *Phénoménologie* (chapitre VII.1).

[66] Autrement dit, étant donné que certains X sont G (ou en déduisant cela de certains X sont S), nous pouvons généraliser à tous les X sont G, à condition qu'il n'existe aucun cas négatif connu (X-nonG) pour le contredire.

● Toutefois, ce serait une grave erreur de jugement [67]que de *déduire* d'une telle généralisation valide qu'il pourrait exister des X qui sont G mais pas S (même si nous savons qu'il existe *des choses (À l'exception des X* qui sont G mais pas S). À ce stade, le contenu de « Tous les X sont G » est identique (dans son extension et implicitement dans son intension) à celui de « Tous les X sont S » dont il est dérivé [68]. La différence entre les deux énoncés réside uniquement dans une éventuelle *particularisation corrective* …

● Supposons que demain nous découvrions *un X qui, bien que toujours G, n'est pas S* (par exemple, une émeraude d'une autre couleur que le vert). Dans ce cas, nous devrions préciser la première affirmation (la plus spécifique) en « Certains X sont S et certains X ne sont pas S » ; mais la seconde affirmation (la plus générique) « Tous les X sont G » resterait inchangée.[69]

● Mais, du fait de cette particularisation, l'affirmation « Tous les X sont G » prend *un sens plus vague* , puisque G ne désigne plus seulement l'espèce S de G, mais également une autre espèce (non-S) de celui-ci.

---

[67] C'est là, comme nous l'avons vu, l'une des erreurs commises par Goodman dans la formulation de son « énigme ». Cette erreur est de nature déductive plutôt qu'inductive.

[68] Cela est évident si l'on considère que l'on peut tout aussi bien obtenir « Tous les X sont G » : (a) par généralisation à partir de « Certains X sont G », que l'on déduit de « Certains X sont S », ou (b) par déduction à partir de « Tous les X sont S », que l'on généralise à partir de « Certains X sont S ». En réalité, on pourrait arguer que ces deux approches diffèrent légèrement, puisque (a) exige de s'assurer qu'il n'existe aucun cas de X qui ne soit pas G, tandis que (b) exige de s'assurer qu'il n'existe aucun cas de X qui ne soit pas S. Cette différence apparaît cependant lors de l'étape suivante de particularisation éventuelle.

[69] Notons que si nous découvrons un X qui n'est pas G, il n'est nécessairement pas non plus S, étant donné que tous les S sont G. Dans ce cas, les deux propositions générales devraient bien sûr être particularisées.

Ainsi, bien que la règle inductive nous pousse à généraliser aussi loin que possible, il faut garder à l'esprit que plus on généralise, plus la portée de notre généralisation s'en trouve diluée.[70]

Ainsi, bien qu'en principe la généralisation vers le haut soit libre, en pratique nous procédons relativement lentement afin de préserver l'utilité noétique de nos idées et énoncés. À titre d'exemple formel : la proposition « Tous les X sont S » peut servir de prémisse mineure dans un syllogisme où S est le moyen terme, tandis que la proposition « Tous les X sont G » – même si elle est identique en extension et en intension à la précédente – serait inutile dans ce même contexte (c'est-à-dire avec S comme moyen terme).

De plus, considérer « Tous les X sont G » comme une généralisation plus avantageuse que « Tous les X sont S », en ce sens qu'elle nous fournit des informations sur davantage de choses pour le même prix en termes de données données, révèle une confusion [71]entre la généralisation d'un sujet donné d'un prédicat plus étroit à un prédicat plus large, et la généralisation d'un prédicat donné d'un sujet plus étroit à un sujet plus large.

Ce dernier cas représente la forme de généralisation véritablement profitable. Supposons que tous les X soient P et que Y soit une surclasse de X (c'est-à-dire que tous les X sont Y, mais que tous les Y ne soient pas X) ; alors, cela consisterait à induire que tous les Y sont P — du moins jusqu'à ce qu'un Y non P soit découvert. Les règles

---

[70] Dans ce contexte, on pourrait comparer le concept de « grue » de Goodman au concept d'« oomph » de Feynman. Ce dernier, défini (avec une pointe d'ironie) comme « une sorte de tendance au mouvement », pourrait sembler utile pour « expliquer » divers phénomènes, mais il est si vague qu'il ne peut rien prédire et est donc sans valeur (p. 19).

[71] Ce dont Goodman s'est rendu coupable en formulant son « énigme », soit dit en passant.

de cette généralisation sont traitées en détail dans mon ouvrage **Logique du futur**, sous le titre « Induction factorielle » (Partie VI).

12. Le paradoxe des Sorites

Extrait de Topics (publié jusqu'à présent uniquement en ligne).

1. Qu'est-ce qu'un tas ?

Le paradoxe des sorites n'est pas un paradoxe au sens strict du terme, mais une question. Cette question est parfois posée de manière sophistique, afin de lui donner un aspect paradoxal. Cependant, elle peut être formulée plus simplement, auquel cas elle apparaît simple, bien que non dénuée d'importance. Le terme *sorites* est latin et dérive du grec *sōros* , qui signifie tas.

Une façon d'exprimer le paradoxe de Sorite est la suivante : qu'est-ce qu'un « tas » (de cailloux, par exemple) ? Ou encore, combien de cailloux (par exemple) constituent un « tas » de cailloux ? La réponse évidente est qu'il doit y avoir au moins un caillou. S'il n'y a pas de cailloux, il n'y a pas de tas, mais rien. Mais un seul caillou suffit-il ? La réponse évidente est : non, il en faut *au moins Deux* cailloux forment un tas, car « tas » est un terme collectif qui sous-entend que les cailloux sont empilés les uns sur les autres (et on ne peut empiler un objet qui n'est pas un caillou sur un caillou, ni un caillou sur un objet qui n'est pas un caillou). Un caillou seul ne constitue pas un tas ; « tas » est intrinsèquement pluriel.

Formulé ainsi, le problème n'est pas très complexe. Mais formulé de la manière suivante, il se complique. Si l'on empile une centaine de cailloux, on obtient évidemment un tas. Que se passe-t-il si l'on retire un caillou ?

Avons-nous toujours un tas ? Oui, 99 cailloux forment un tas. Que se passe-t-il si l'on retire encore un caillou ? Avons-nous toujours un tas ? Oui, 98 cailloux forment un tas. Et ainsi de suite, jusqu'à de petits nombres, moment où le sophiste se demande si deux cailloux constituent un tas, puis un seul, puis aucun. À ce stade, la question paraît paradoxale, voire artificielle : faut-il conclure qu'aucun caillou ne forme un tas, ou qu'un seul caillou en forme un ? Le sophiste le pense ; le logicien, lui, sait que non.

2. L'utilisation de termes vagues

De toute évidence, le problème ici, s'il y en a un, réside dans la formulation précise de termes initialement vagues. Si nous ne prenons pas d' *emblée* le recul nécessaire pour réfléchir à l'intention exacte d'un terme vague de ce genre (un terme suggérant une quantité, en l'occurrence) ou de tout autre, nous risquons de rencontrer des difficultés *par la suite* . Il est donc indispensable de réfléchir avant d'employer de tels termes et d'anticiper les difficultés qu'ils pourraient engendrer. Dans le cas présent, comme nous l'avons vu, le terme « tas » est intrinsèquement pluriel (et donc inapplicable à moins de deux cailloux). Un sophiste préfère compliquer les choses, afin de semer le doute sur la raison et la connaissance humaines, comme à son habitude ; mais la question est en réalité assez simple.

Certes, dans certains cas, la définition d'un terme vague comporte une part *de convention* . On peut l'illustrer par une autre version du paradoxe dit de Sorites : à partir de combien de cheveux un homme n'est-il pas chauve ? De toute évidence, un homme totalement chauve l'est assurément. Mais un homme avec un seul cheveu, ou très peu, ne serait-il pas considéré comme chauve dans le langage courant ? Peut-être, mais seulement par convention. Si l'on définit *strictement la calvitie* , elle

implique l'absence totale de cheveux ; mais si l'on adopte une définition *plus large* , elle peut inclure des cas impliquant un nombre *arbitraire* , quoique de préférence faible, de cheveux, déterminé par convention – par exemple, cinq ou vingt cheveux [72]. De toute évidence, pour éviter toute confusion, rien ne nous empêche de désigner ce sens élargi de « chauve » par l'expression plus précise « chauve ou presque chauve ».

Ce problème d'imprécision n'a rien d'exceptionnel ; il ne se limite pas aux termes à l'origine du paradoxe de Sorites. Par exemple, un terme relatif comme « petit » (ou son relatif « grand ») est *intrinsèquement vague* et ne peut être employé avec précision que dans des situations spécifiques, au moyen d'une grandeur conventionnelle. De même, lorsqu'on travaille avec des continuums, il peut être nécessaire d'établir des lignes de séparation arbitraires. Par exemple, il n'existe pas de ligne de séparation objective entre deux couleurs adjacentes dans le spectre, et il peut être nécessaire, dans certaines circonstances, d'en imaginer une (par exemple, à des fins juridiques ou autres [73]). Les distinctions conventionnelles font partie de la pensée humaine ; mais, il est important de

---

[72] Ou toute pilosité minimale qui nous semble subjectivement si proche de la calvitie qu'elle en devient effectivement une calvitie.

[73] Par exemple, en droit juif (*halakha*), une grande importance est accordée aux définitions quantitatives, notamment aux valeurs maximales et minimales des durées, des heures, des distances, des longueurs, des volumes, des poids, des températures, des valeurs monétaires, etc. Initialement, ces mesures étaient souvent exprimées par les rabbins en termes vagues (par exemple, « le volume d'un œuf »), mais par la suite, des formulations plus précises ont été nécessaires (que différentes autorités pouvaient estimer différemment). Cependant, certaines mesures restent subjectives (par exemple, l'estimation du moment où l'on est rassasié après un repas). Pour plus de détails, voir : http://halachipedia.com/index.php?title=Reference_of_Measurements_in_Halacha .

le souligner, elles ne constituent pas la seule pensée humaine. Il existe toujours des éléments objectifs sous-jacents aux distinctions conventionnelles. Par exemple, la ligne de séparation entre le bleu et le vert se situe quelque part entre ce que nous percevons comme clairement bleu et ce que nous percevons comme clairement vert ; elle ne sera jamais très éloignée d'un côté ou de l'autre, et encore moins entre le vert et le jaune, ou entre le bleu et l'indigo, ou encore plus loin.

Pour revenir à notre prétendu paradoxe, quelques précisions s'imposent. Comme indiqué précédemment, le concept de tas n'a rien de paradoxal, pourvu qu'il soit correctement défini. En termes simples, un tas peut être défini comme un ensemble d'objets matériels empilés les uns sur les autres, quelle que soit la manière, ce qui implique la présence d'au moins deux objets. Une définition plus complexe supposerait qu'un tas soit pyramidal, c'est-à-dire qu'il nécessite au moins quatre objets (trois à la base et un au sommet) ; mais cela ne pose aucun problème, car on peut plus précisément parler de tas pyramidal. De même, la calvitie *au sens strict* désigne l'absence totale de cheveux ; elle s'applique néanmoins, de façon approximative, à la présence de quelques cheveux.

Notre langage courant regorge de termes vagues de ce genre [74], dont certains nécessitent une définition précise

[74] En effet, cela est inévitable pour deux raisons. Premièrement, il est impossible de définir précisément de nombreux objets concrets. Où, par exemple, s'arrête précisément une orange ? Le parfum ou la chaleur qui émanent du fruit en font-ils partie ? À quels moments et dans quels espaces ces émanations peuvent-elles logiquement être considérées comme distinctes ? Deuxièmement, la connaissance humaine étant inductive, nous ne pouvons pas toujours commencer un concept par une définition précise, mais avons tendance à le laisser ouvert, à se préciser au fur et à mesure que l'expérience se déploie. Dans cette perspective, la majorité des termes abstraits que nous utilisons sont ouverts, y compris des termes qui

pour des raisons pratiques. Par exemple, le terme « foule » peut désigner un rassemblement de trois personnes ou plus, conformément à l'adage populaire « à deux, c'est bien, à trois, c'est la foule ». On peut aussi, dans le cadre d'un logiciel utilisé par la police pour surveiller de grands groupes de personnes, opter pour un seuil minimal plus élevé (par exemple, 50 ou 500), fixé arbitrairement comme motif d'alerte. Le terme « émeute » sera alors privilégié lorsque cette dernière foule se déchaîne.

En tout état de cause, ces expressions quantitatives imprécises ne posent aucun problème logique. Elles ne constituent pas un défaut du langage courant, nous obligeant à construire un « langage idéal » où chaque terme aurait une signification unique et précise [75]. Elles requièrent encore moins un traitement au moyen de logiques symboliques abstruses (cherchées à nombre de logiciens modernes). Au contraire, elles témoignent de la versatilité et de la flexibilité de la pensée et du langage ordinaires ; elles attestent également que notre communication linguistique comporte souvent des

---

peuvent servir à définir plus précisément d'autres termes ; ainsi, là encore, le flou est inévitable. Mais ces difficultés ontiques et épistémiques n'impliquent pas de paradoxe ; elles appellent simplement à une réflexion philosophique.

[75] Le recours à un « langage idéal » par certains logiciens modernes pour résoudre un problème de logique est vain. Incapables de comprendre comment nous, êtres humains, abordons logiquement certaines difficultés cognitives, ils tentent de nous imposer une pensée superficielle, artificielle et impraticable. Le rôle du véritable logicien n'est pas d'imposer des logiques imaginaires, mais de comprendre nos moyens logiques naturels afin de les perfectionner et de les renforcer. Le raisonnement humain devrait être au cœur des préoccupations des logiciens. Pour traiter les paradoxes de Sorites, il convient d'utiliser un langage plus précis : par exemple, au lieu de qualifier de « chauves » les personnes ayant très peu de cheveux, il faut dire « presque chauves » ; ou encore, plus précisément si nécessaire pour des raisons pratiques, « ayant (disons) un à dix cheveux ».

sous-entendus non verbaux, que nous comprenons généralement très bien. La plupart du temps, notre usage quotidien de termes vagues *ne nécessite pas* de définition plus précise. Ils servent à suggérer des choses de manière approximative et ne prétendent pas être des affirmations exactes et véridiques. Si le besoin de précision et de vérité se fait sentir, il est alors comblé de manière à préserver la cohérence (par convention explicite si nécessaire).

3. Raisonner avec des termes vagues

Certes, les termes vagues peuvent être périlleux si l'on tente de raisonner avec eux. Mais ils sont souvent employés sans pour autant être intégrés à un processus argumentatif. De plus, raisonner avec des termes vagues n'est pas toujours invalide : il existe des contextes où le caractère vague n'empêche pas de parvenir à une conclusion fiable.

En syllogisme catégorique (ou hypothétique), la règle concernant les termes (ou thèses) vagues est la suivante : le terme ou la thèse du milieu ne peut être vague, car cela ne garantirait pas que son sens soit identique dans les deux prémisses. Si le terme du milieu est vague, on ne peut être certain de la concordance et la conclusion est invalide. En revanche, les termes majeur et mineur peuvent être vagues sans que cela n'affecte la validité de l'argument, et ce, dans les quatre cas. Il existe cependant une exception à cette règle : lorsque le terme du milieu est vague, mais pas au point de compromettre la concordance, un argument valide peut néanmoins être construit.

Ce dernier point est évident lorsqu'on considère les deux arguments suivants de la troisième figure : le terme médian assure un chevauchement même si aucune des prémisses n'est universelle. Les expressions « la plupart » (plus de la moitié) et « peu » (moins de la moitié) sont

vagues, car elles ne précisent pas de nombres exacts. Mais notons la particularité (par opposition à la majorité ou à la minorité) de la conclusion, c'est-à-dire son imprécision accrue.

La plupart des M sont des P

et la plupart des M sont S ;

Par conséquent, certains S sont P

Peu de M sont P (ce qui implique que la plupart des M ne sont pas P)

et la plupart des M sont S ;

Par conséquent, certains S ne sont pas des P

En apodose, les règles concernant les termes ou thèses vagues sont les suivantes :

Modus ponens:

Si A est B, alors C est D, et A est B (affirmant l'antécédent) ; alors, C est D (le conséquent est affirmé).

L'antécédent et la prémisse mineure ne doivent pas être vagues ; sinon, la conclusion est invalide. En revanche, le conséquent peut être vague, et la conclusion reste valide (bien que vague elle aussi).

Modus tollens:

Si A est B, alors C est D, et C n'est pas D (ce qui nie le conséquent) ; alors, A n'est pas B (l'antécédent est nié).

Le conséquent et la prémisse mineure ne doivent pas être vagues ; sinon, la conclusion est invalide. En revanche, l’antécédent peut être vague, et la conclusion reste valide (bien que vague elle aussi).

Là encore, dans les deux cas, une exception est concevable, si l'on sait que les prémisses majeures et mineures se chevauchent, même si l'on ne sait pas précisément dans quelle mesure ce chevauchement est important.

Des règles similaires peuvent être formulées pour d'autres types d'arguments, tels que le dilemme ou le a fortiori.

Ceux qui prétendent que les termes vagues sont intrinsèquement paradoxaux font preuve d'une mauvaise foi malhonnête, cherchant à impressionner autrui ou à se mettre en avant par leur capacité à déceler et résoudre des paradoxes (artificiels), ou pire encore, à démontrer que la connaissance humaine est inévitablement paradoxale et donc vaine. De même qu'il est malhonnête de qualifier de tas un caillou isolé ou l'absence de caillou, il est malhonnête de qualifier de chauve une personne ayant un ou plusieurs cheveux [76]. Si l'on s'adonne d'emblée à de telles contradictions, on aboutit forcément à des paradoxes. Ceux qui agissent ainsi ne sont pas de véritables logiciens, mais des sophistes. Ils manipulent et inventent des histoires ; la recherche du véritable savoir ne les intéresse pas.

---

[76] Comme je l'explique dans **A Fortiori Logic** , chapitres 1.4 et 2.5, il est parfois utile de formuler des termes de manière si inclusive qu'ils englobent toutes les valeurs positives, nulles et négatives. Cette pratique est courante dans le discours scientifique car elle facilite certains calculs et graphiques. Il est cependant essentiel de bien comprendre que de tels termes inclusifs sont intrinsèquement indéniables – autrement dit, ils englobent déjà à la fois une affirmation et sa négation. Dans le contexte présent, nous pourrions choisir d'élargir des termes comme « tas » ou « chauve » pour y inclure leurs contraires, pour une raison ou une autre ; mais ce faisant, nous devons rester pleinement conscients de nos actes. Si nous ne traitons pas ces termes avec la prudence requise, nous ne devons pas nous étonner de nous retrouver face à des contradictions.

4. Inventer de faux paradoxes

Les formulations originales des deux énigmes décrites ci-dessus, relatives à un tas et à la calvitie, sont attribuées à Eubulide de Milet (*actif* au IVe siècle av. J.-C.), le logicien mégarien à qui l'on doit également le paradoxe du menteur [77]. Élève d'Euclide, maître de Diodore de Cronos, il était contemporain et rival d'Aristote. Ces énigmes n'étaient peut-être pas initialement présentées comme des paradoxes, mais plutôt comme des illustrations d'une question (à savoir : où tracer la limite ?). Cela reflétait sans doute une prise de conscience naissante de l'existence de termes vagues qui peuvent nécessiter une définition arbitraire dans certaines circonstances. Comme indiqué précédemment, ce problème se résout assez facilement. Cependant, des penseurs postérieurs ont eu tendance à le complexifier inutilement et l'ont présenté sous la forme d'une chaîne d'arguments (ou sorites, où la conclusion de l'argument précédent sert de prémisse au suivant).

Ainsi, le paradoxe de la calvitie se formule, de manière positive : un homme avec un seul cheveu est assurément aussi chauve qu'un homme sans cheveux ; et si un homme avec un seul cheveu peut être qualifié de chauve, alors un homme avec deux cheveux l'est aussi ; et si un homme avec deux cheveux est chauve, alors un homme avec trois cheveux l'est également ; etc. ; par conséquent, un homme avec mille cheveux peut encore être considéré

---

[77] Quelques siècles auparavant, Épiménide de Cnossos déclarait : « Les Crétois sont toujours des menteurs », bien qu'étant lui-même Crétois, apparemment sans se rendre compte de la contradiction inhérente à une telle affirmation. Eubulide a peut-être perçu le paradoxe et cherché à le nuancer et à le renforcer, car il ne s'agissait pas d'un double paradoxe, mais d'un paradoxe facilement résolu en disant que tous les Crétois ne sont peut-être pas des menteurs ou que tous les Crétois ne mentent pas toujours (Épiménide étant une exception notable).

comme chauve (paradoxe). On peut aussi formuler l'argument de manière négative : si un homme avec mille cheveux n'est pas considéré comme chauve, alors celui qui en a 999 ne l'est pas non plus ; et si 999 ne suffit pas, alors 998 non plus ; et ainsi de suite… d'où, un homme avec un seul cheveu n'est pas chauve ; par conséquent, un homme sans cheveux n'est pas chauve (paradoxe).

De toute évidence, ces arguments sont *forcés* ; ils reposent sur des prémisses très douteuses et trompeuses. Dans la version positive, la prémisse erronée est que « un homme sans cheveux peut être qualifié de chauve » *implique* « un homme avec un cheveu peut être qualifié de chauve ». Dans la version négative, la prémisse erronée est que « un homme avec un cheveu ne peut pas être qualifié de chauve » *implique* « un homme sans cheveux ne peut pas être qualifié de chauve ». Il s'agit du même sophisme d'implication (sous forme contraposée). Le paradoxe naît du refus d'admettre que « chauve », à proprement parler, signifie « sans cheveux », refus qui ne repose pas sur une logique honnête, mais sur une volonté délibérée de contourner la logique. De même, concernant un tas, le piège consiste à sous-entendre qu'un simple caillou, voire aucun caillou, peut être considéré comme un tas.

Les inventeurs de ces arguments ne prennent pas le temps de réfléchir rationnellement au problème sous-jacent (où se situe la limite ?) avant de se lancer dans un raisonnement fallacieux, mais tentent plutôt de nous piéger. La chaîne d'arguments proposée n'est qu'un écran de fumée dissimulant la fausseté fondamentale qui est véhiculée. C'est l'équivalent logique d'une vente pyramidale, chaque vente soutenant la suivante sans fondement solide. Ceux qui tombent dans le piège sont tout simplement de mauvais logiciens, voire d'une malhonnêteté crasse. Ils élaborent ensuite avec prétention des théories complexes et alambiquées autour de ce

phénomène, sans se soucier de l'erreur ou du mensonge initial dans leur discours.

À ceux qui prétendent qu'un cheveu ou un caillou ne change rien, je suggère de réaliser cette simple expérience physique : prenez une balance précise, équilibrée de part et d'autre, et ajoutez un cheveu ou un caillou sur l'un des plateaux. Vous verrez alors la balance pencher ! À ceux qui restent sceptiques, car ils croient dogmatiquement que la logique n'est qu'une question d'imagination et de convention, je suggère (avec une pointe d'ironie) de placer sous le plateau le plus lourd un détonateur relié à un dispositif explosif fixé sur leur nez, et de regarder la réalité leur exploser au visage ! Cet argument, je pense, pourrait bien les convaincre, s'ils survivent.

Concernant Eubulide, mentionnons au passage ses autres paradoxes. Le plus significatif est bien sûr le paradoxe du menteur, qui, comme je l'ai démontré en détail ailleurs, [78]est exceptionnellement puissant en raison de la variété des difficultés qu'il soulève (mais reste néanmoins tout à fait résoluble). Trois autres paradoxes [79]portent sur les équivoques liées au terme « savoir », notamment l'incapacité à reconnaître immédiatement une personne que l'on reconnaît habituellement sans hésitation (comme un proche ou un vieil ami), lorsque cette dernière est masquée, est absente depuis trop longtemps ou n'est pas observée avec suffisamment d'attention. Un autre encore, le paradoxe des cornes, affirme que ce que l'on n'a pas perdu est nécessairement en notre possession ; par conséquent, si l'on n'a pas perdu de cornes, on doit en avoir. Cette apodose repose sur une prémisse majeure erronée, car quelque chose que l'on n'a « jamais eu » peut tout aussi bien (que quelque chose que l'on « possède

---

78 Voir mon *A Fortiori Logic* , Annexe 7.4.

79 À savoir, les paradoxes de l'Homme masqué, d'Elektra et de l'Homme négligé.

encore ») être qualifié de « non perdu » – le conséquent ne découle donc pas de l'antécédent ; par conséquent, si l'on n'a pas perdu de cornes, on ne peut supposer que l'on en possède (puisqu'on peut très bien n'en avoir jamais eu).

Cette courte liste révèle que les questions d'Eubulide suscitent toutes une réflexion sur la logique : réflexions sur les termes vagues, les définitions conventionnelles, les équivoques, les négations, l'autoréférence et diverses autres difficultés inhérentes au discours humain. Il serait erroné, à mon avis, de supposer que ces questions visaient à provoquer ou à obscurcir le propos (bien que, comme mentionné précédemment, certains sceptiques ultérieurs aient exploité ces énigmes à des fins malveillantes). J'estime plutôt qu'elles ont servi de tremplin à des réflexions et des discussions approfondies sur la logique, car il semble improbable qu'elles aient été formulées sans chercher à résoudre les problèmes qu'elles soulevaient. Dans certains cas, des explications ou des solutions valables ont sans doute été proposées (même si elles ne nous sont pas parvenues), tandis que dans d'autres, les difficultés ont pu paraître insurmontables. En d'autres termes, je doute qu'Eubulide ait été un simple sophiste ; je le considérerais plutôt comme un logicien sérieux.

13. Protagoras contre Euathlus

Extrait de Topics (publié uniquement en ligne pour le moment).

1. Un paradoxe ancien

L'un des nombreux paradoxes supposés qui nous sont parvenus des Grecs est la dispute entre Protagoras (d'Abdère, vers 480-410 av. J.-C.) et son élève Euathlus (dont on ne sait rien de plus). Selon Aulus Gellius (Romain, vers 125-180 apr. J.-C.) [80], Protagoras, célèbre sophiste et professeur réputé, accepta de former Euathlus à la rhétorique, discipline essentielle à l'époque pour plaider devant les tribunaux. L'accord stipulait qu'Euathlus ne devrait payer à Protagoras les honoraires convenus (ou la partie restante, selon certaines sources) qu'une fois sa formation achevée et après avoir plaidé et gagné sa première affaire [81].

On raconte qu'après avoir terminé ses études, Euathlus, pour une raison inconnue, n'a pas choisi d'utiliser ses nouvelles compétences devant un tribunal. Il n'a donc jamais gagné ni perdu de procès et n'était contractuellement tenu à aucune rémunération envers Protagoras. Néanmoins, ce dernier, pour des motifs que nous examinerons plus loin, l'a poursuivi en justice devant

---

80 Quelque 550 à 600 ans plus tard, dans ses *Nuits du grenier* (vers 150 C500 av. J.-C.).

81 Le contrat en question était vraisemblablement verbal plutôt qu'écrit à cette époque ; mais nous pouvons supposer que les deux parties étaient d'accord sur sa clause stipulée.

l'Aréopage d'Athènes. Euathlus a choisi d'assurer lui-même sa défense. Les arguments suivants auraient été présentés au tribunal :

a. **Protagoras** a fait valoir qu'il devait certainement être payé, car (i) si le tribunal statuait en sa faveur, il pourrait sur cette base exiger le paiement ; et (ii) si le tribunal statuait en faveur du défendeur, alors ce dernier aurait gagné sa première affaire et serait donc contractuellement obligé de payer les honoraires de toute façon.
b. **Euathlus** a répondu qu'il ne devait certainement *pas* payer les frais, car (i) si le tribunal se prononçait en faveur du demandeur, Euathlus aurait perdu son premier procès et ne serait donc pas contractuellement obligé de payer les frais ; et (ii) si le tribunal se prononçait en faveur du défendeur, il serait de toute façon exempté de paiement.

Ainsi, tandis que le demandeur soutenait, avec une conviction apparente, qu'il était certain d'avoir droit au paiement quelle que soit la décision du tribunal, le défendeur, de son côté, pouvait affirmer, avec la même conviction apparente, qu'il était certain d'en être exempté dans les deux cas. C'est pourquoi cette affaire est considérée comme paradoxale. On raconte que le tribunal, déconcerté par ces arguments et contre-arguments, a choisi d'ajourner l'affaire *sine die* (ou, selon certains, pour 100 ans) afin d'éviter de rendre un jugement.

L'importance de ce différend juridique pour la logique et la philosophie réside dans l'impression qu'il donne que deux personnes peuvent argumenter de manière dilemmatoire et paradoxale, en sens opposés, et avoir toutes deux raison. Les ennemis de la raison humaine se réjouissent de ce genre de conclusion, car elle remet en question la fiabilité et l'efficacité de la raison. Mais comme nous allons le montrer, cette impression est très

superficielle. Il existe, à mon sens, au moins deux solutions possibles à ce que l'on appelle le paradoxe de la Cour. En fait, bien que j'aie pensé à la seconde avant la première, cette dernière précède logiquement la première. J'ai appris plus tard, grâce à l'étude de la littérature sur le sujet réalisée par Peter Suber, [82]que la première solution avait été proposée bien avant moi par Aulus Gellius ; la seconde semble quant à elle inédite.

2. Première résolution

La solution la plus simple à ce problème est de supposer que le rusé Protagoras, voyant qu'Euathlus tardait à se mettre au travail, décida d'accélérer les choses. Protagoras piégea son élève en utilisant l'argument susmentionné, sachant pertinemment qu'il perdrait le premier procès, mais gagnerait le second. Il savait qu'il perdrait le premier procès car l'accord entre les parties n'obligeait Euathlus à payer les honoraires qu'après avoir gagné sa première affaire ; il ne l'obligeait pas à exercer le droit dans l'immédiat, ni même jamais. Euathlus tomba imprudemment dans le piège et plaida lui-même l'affaire devant le tribunal. Si le tribunal n'avait pas ajourné l'audience *sine die* , il lui aurait logiquement donné gain de cause, lui permettant ainsi de remporter sa première affaire. Par la suite, en supposant qu'un second procès fût légalement autorisé – et les arguments des deux parties reposent sur cette hypothèse –, Protagoras aurait été le grand vainqueur. Bien entendu, aucun second procès ne

---

[82] Dans **Le paradoxe de l'auto-amendement* * , la section consacrée au paradoxe de Protagoras et Euathlus est disponible en ligne ; sa lecture, y compris les notes, est intéressante et relativement courte. Voir : https://dash.harvard.edu/bitstream/handle/1/10288413/Peter%20Suber%2c%20Paradox%20of%20Self-Amendment%2c%20Section%2020.html?sequence=1#A

serait nécessaire si Euathlus reconnaissait qu'ayant gagné le premier procès, il était certain de perdre le second, et réglait l'affaire sur-le-champ.

Autrement dit, le premier argument de Protagoras (i) n'était qu'un leurre ; il s'appuyait en réalité sur son second argument (ii). Euathlus, grisé par sa vanité, élabora deux contre-arguments sophistiqués, croyant surpasser son maître. Mais Protagoras était plus rusé. Le seul moyen pour Euathlus d'éviter la défaite était d'engager un avocat [83]. Si ce dernier remportait le premier procès, comme on pouvait logiquement s'y attendre puisque la seule condition de défaite n'était pas remplie, Euathlus ne serait pas considéré comme ayant personnellement gagné son premier procès, et Protagoras ne pourrait obtenir gain de cause lors du second. Ainsi, le maître était bel et bien supérieur à l'élève. Il n'y avait pas de véritable paradoxe, puisqu'il existait une solution concrète à ce qui semblait être un paradoxe.

Certes, on pourrait ajouter qu'Euathlus, dans les faits, sortit vainqueur, car, par son contre-argumentation, il parvint à tellement embrouiller le tribunal que celui-ci renonça à juger l'affaire, et il ne fut pas contraint de payer. Peut-être l'espérait-il et la chance lui sourit. Mais, d'un point de vue théorique et logique, avec le recul, il se révéla peu perspicace. Non seulement il commit l'erreur de ne pas engager d'avocat pour le défendre, mais il supposa aussi, à tort, que son argumentation était efficace pour contrer celle de Protagoras. Il s'illusionna en pensant qu'une victoire au premier procès l'exempterait de la seconde. Il aurait dû examiner plus attentivement l'argumentation de son maître.

---

[83] En supposant que cela fût possible à Athènes à cette époque. C'est une hypothèse raisonnable (même si certains commentateurs la contestent), car Euathlus aurait lui-même été formé par Protagoras pour devenir avocat.

Examinons maintenant les arguments en termes plus formels, afin de les clarifier. J'introduirai les symboles suivants : soit P = Protagoras, Q = Euathlus ; A = la condition convenue selon laquelle Q doit gagner son premier procès pour être tenu responsable, et C = le résultat convenu selon lequel Q doit payer des honoraires à P. Nous savons d'emblée que si A est vrai, alors C est vrai, et si A n'est pas vrai, alors C n'est pas vrai : ce sont les termes de l'accord. Les arguments sont les suivants (avec mon commentaire critique en *italique*) :

a. Selon P : (i) si le tribunal décide que P gagne, alors C est vrai (*mais objectivement le tribunal ne peut pas décider en faveur de P, puisque A n'est pas encore vrai, donc cette question ne peut pas être soulevée*) ; tandis que si (ii) le tribunal décide que Q gagne (*comme il le doit logiquement*), alors non-C est vrai (*à la fin d'un premier procès*) ; mais si Q gagne ainsi, alors A devient vrai et C doit (*dans un procès ultérieur*) en découler, auquel cas P gagne finalement.
b. Selon Q : (i) si le tribunal décide que P gagne, alors non-A est vrai (*à la fin d'un premier procès*) ; mais si non-A est vrai, alors non-C doit (*lors d'un procès ultérieur*) s'ensuivre, donc Q gagne finalement (*mais objectivement, le tribunal ne peut pas décider en faveur de P, puisque A n'est pas encore vrai, donc cette question ne tient pas*) ; tandis que si (ii) le tribunal décide que Q gagne (*comme il le doit logiquement*), alors non-C est vrai (*mais ici Q omet de mentionner les conséquences ultérieures que P a justement soulignées*).

En conclusion, P est logiquement le vainqueur ; les arguments de Q sont en réalité insuffisants pour empêcher ce résultat. P prétend vouloir gagner immédiatement, mais en réalité, son objectif est une victoire à plus long terme (au second round). Q s'imagine qu'il pourrait perdre le premier round mais gagner le second, ou qu'il pourrait

gagner le premier round sans avoir à disputer de second round ; mais ce ne sont que des chimères. Il est difficile de comprendre pourquoi le tribunal a trouvé cette affaire trop confuse – le ou les juges ne devaient pas être des esprits brillants.

3. Deuxième résolution

Une solution plus complexe au problème est la suivante. Il est possible que Protagoras ait poursuivi Euathlus en invoquant *une clause tacite* de l'accord. L'accord ne contenait *qu'une seule clause explicite* , à savoir qu'Euathlus ne devrait payer les honoraires à Protagoras que s'il gagnait son premier procès. Dans ce cas, Protagoras n'aurait aucun fondement pour demander un procès, puisque cette condition n'avait manifestement pas été remplie. Mais puisqu'il a intenté une action en justice, il a dû penser et argumenter que l'accord comportait une *clause tacite. (ou peut-être implicitement) Sachant* qu'Euathlus pratiquerait le droit dans un délai raisonnable [84], après quoi ses nouvelles compétences seraient mises à l'épreuve devant un tribunal et il devrait payer en cas de victoire, Protagoras ne pouvait imaginer que le juge autoriserait un procès, et encore moins qu'il statuerait en sa faveur, sans raison valable [85].

Il est clair que le juge devait trancher, dans ce procès (et cela ne pouvait être que le cas, compte tenu du contexte actuel), la question de savoir si l'affirmation de

---

[84] Protagoras soutiendrait également, bien entendu, que le délai raisonnable était expiré. Si le tribunal reconnaissait l'existence d'une clause tacite mais contestait son exécution, il ajournerait simplement l'affaire pour une durée déterminée (selon son estimation).

[85] Il va sans dire que la présente analyse est entièrement faite d'un point de vue logique, même si un tribunal pourrait raisonner différemment ou même ne pas raisonner du tout (par exemple, corruption, favoritisme, pression).

Protagoras, selon laquelle l'accord comportait une clause tacite, était fondée. Il aurait pu la justifier en considérant que si Euathlus n'avait pas été autorisé à exercer le droit ou à l'exercer aussi tard qu'il le souhaitait, l'accord aurait certainement précisé cette réserve. Il aurait tout aussi bien pu la rejeter en considérant que Protagoras avait tenu pour acquis un élément sur lequel il aurait dû obtenir un accord explicite. Ainsi, l'affaire repose sur une question tacite, et non exclusivement sur la clause explicite de l'accord ; autrement dit, les faits sont plus complexes que ce qui a été révélé.

De plus, il ressort clairement des arguments présentés par les deux parties lors de ce procès que chacune d'elles envisageait la possibilité d'un second procès, susceptible d'infirmer le jugement du premier. Ceci est logiquement sous-entendu dans le second argument de Protagoras et dans le premier argument d'Euathlus. Dans ces deux cas de figure, un second procès serait nécessaire pour rendre un jugement définitif, à moins, bien sûr, que la partie perdante n'accepte librement l'issue inévitable et ne règle ses comptes à l'avance. Le juge du premier procès ne pouvait se prononcer en faveur d'une partie, puis contre elle dans la même phrase. Un premier jugement de victoire ou de défaite devait être établi avant qu'un second puisse être rendu, conformément aux termes du contrat. Il convient donc de prendre en compte le facteur temps dans l'analyse de ces arguments.

Il convient de noter que, tandis que le premier procès a une issue incertaine, puisqu'il dépend de la décision (le jugement) du juge concernant une clause tacite alléguée de l'accord, le second procès a une issue prévisible, puisqu'il dépend uniquement de la clause explicite de l'accord.

Comme nous allons le démontrer, si l'on tient compte des considérations précédentes, l'apparence paradoxale

disparaît aisément. J'utiliserai ici les symboles suivants : P = Protagoras ; Q = Euathlus ; A = la condition explicite selon laquelle Q doit gagner son premier procès ; B = la condition implicite alléguée selon laquelle Q devait exercer dans un délai raisonnable ; et C = les honoraires que Q doit verser à P. Les arguments sont les suivants :

a. Selon P : (i) si B alors C ; et (ii) si non-B alors non-C, *mais si non-C alors (plus tard) si A alors C* ; donc, C de toute façon.
b. Selon Q : (i) si B, alors C, *mais si C alors (plus tard) si non-A alors non-C* ; et (ii) si non-B, alors non-C ; donc, non-C de toute façon.

On constate ainsi que les arguments des deux parties sont très similaires, mais chacune a astucieusement omis une partie des conséquences (indiquées en *italique*). P a tronqué les conséquences de (i) que Q souligne, et Q a tronqué les conséquences de (ii) que P souligne. De ce fait, leurs arguments semblent s'équilibrer. Cependant, si l'on prend en compte *tous* les événements ultérieurs (dans le cadre d'un éventuel second procès), on obtient l'argument commun (c) plus objectif, qui est clairement non paradoxal :

(i) si B alors C, mais si C alors (plus tard) si non-A alors non-C ; et
(ii) si non-B alors non-C, mais si non-C alors (plus tard) si A alors C .

Ainsi, dans (i), la conclusion finale est non-C, tandis que dans (ii), elle est C – ce qui signifie qu'il n'y a pas de paradoxe. Ceci montre également que, bien qu'il soit impossible de prédire objectivement si P ou Q remportera le premier procès, on peut affirmer que (i) si P remporte le premier procès, il perdra le second, et si Q remporte le premier procès, il perdra le second. De toute évidence, P ne peut pas prétendre avoir droit à un second procès (s'il perd le premier) alors que Q ne le pourrait pas ; de même,

Q ne peut pas prétendre avoir droit à un second procès (s'il perd le premier) alors que P ne le pourrait pas. Il est donc nécessaire de prendre en compte tous les événements ultérieurs pour concilier logiquement tous les arguments. Notons que si, pour une raison quelconque, il n'y a pas de second procès, il n'y a pas non plus de paradoxe, puisque la conclusion sera soit (i) C, soit (ii) non-C.

Quoi qu'il arrive, il n'y a pas de paradoxe car aucune des parties ne peut en réalité, contrairement aux apparences, prétendre à une victoire inévitable ; la victoire ne se présente pas dans les deux sens pour les deux parties, mais d'une seule façon pour chacune d'elles.

Il est clair qu'ici, les deux parties ont eu recours à la ruse sophistique courante consistant à dissimuler au tribunal un aspect gênant de leur argumentation impartiale. Euathlus était un bon disciple de Protagoras, puisque son contre-argument est le reflet exact de celui de ce dernier. Autrement dit, leurs deux arguments comportaient une part de malhonnêteté ; tous deux étaient des imposteurs intellectuels, exposant sciemment des demi-vérités. Par conséquent, ce faux paradoxe ne représente aucun défi majeur pour la logique, contrairement aux affirmations des relativistes. En particulier, l'affirmation générale de Protagoras selon laquelle « il y a deux côtés à chaque question » (*duo logoi*) se révèle fallacieuse dans le contexte présent.

Il est important de toujours garder à l'esprit que certaines personnes, qu'elles soient philosophes, logiciennes ou autres, sont parfois sujettes à des penchants maléfiques ; certaines, en effet, bien plus que d'autres. Elles peuvent mentir consciemment ou déformer inconsciemment les faits et les arguments pour des raisons très diverses. Le plus souvent, la soif de pouvoir, de gloire et de fortune y contribue. Un universitaire peut rechercher l'admiration

de ses pairs ou de ses étudiants ; un mari peut vouloir impressionner sa femme ; un chômeur peut espérer trouver un emploi ; et ainsi de suite. Il est erroné de considérer tous les propos philosophiques comme désintéressés. Les philosophes et les logiciens ne sont pas tous de purs scientifiques ni des saints.

J'ajouterais que le secret de la réussite dans la résolution des problèmes philosophiques et logiques, et notamment des paradoxes, réside dans le désir sincère de les résoudre. Nombre de philosophes et de logiciens abordent les problèmes avec une attitude négative, cherchant non pas réellement à les résoudre, mais plutôt à justifier leur aversion pour la raison humaine. Le chercheur honnête, quant à lui, est guidé par ses meilleurs instincts ; il désire sincèrement confirmer l'efficacité des facultés cognitives humaines – et c'est précisément pourquoi il réussit.

4. Résolutions inadéquates

Comme mentionné précédemment, d'après l'analyse de la littérature proposée par Suber, la première résolution serait attribuable à Aulus Gellius, tandis que la seconde semble originale. L'analyse de Suber montre que le paradoxe judiciaire a été abordé dans de nombreux ouvrages au fil du temps, mais plus souvent, semble-t-il, d'un point de vue juridique que logique. Les questions juridiques soulevées sont multiples, mais la plupart n'ont pas à être, ni ne devraient, être considérées dans une perspective purement logique. Pourquoi ? Parce que le but du logicien n'est pas ici de trancher le litige, c'est-à-dire de déterminer qui doit gagner ou perdre, mais simplement d'expliquer et de lever l'apparence de paradoxe.

Cette remarque peut être illustrée par la résolution (non mentionnée par Suber) proposée il y a plusieurs siècles par Lorenzo Valla (Italie, vers 1406-1457). Cette tentative de résolution est insatisfaisante car elle repose sur une

distinction hâtive entre le paiement dû suite à un jugement et celui dû en vertu d'un accord. Je cite un article de la SEP qui la décrit[86] :

> Si Euathlus perd le procès, il devra payer le reste des honoraires, conformément au verdict des juges ; mais s'il gagne, il devra également payer, cette fois en vertu de son accord avec Protagoras. Euathlus, cependant, retourne habilement la situation : dans les deux cas, il n'aura rien à payer, ni en raison de la décision du tribunal (s'il gagne), ni en vertu de l'accord avec Protagoras (s'il perd). En résumé, Valla affirme qu'Euathlus ne peut pas avoir le beurre et l'argent du beurre et doit choisir l'une ou l'autre des solutions : il doit se conformer soit à son accord avec Protagoras, soit au verdict des juges. Si ces derniers statuent contre Protagoras, il pourra tenter de récupérer son argent lors d'une seconde action en justice.

Analysons la résolution proposée par Valla en utilisant les symboles suivants : soit P = Protagoras ; Q = Euathlus ; et C = les frais que Q doit payer à P.

a. Selon P : Si Q perd (pour quelque raison que ce soit), alors C (par verdict) ; mais si Q gagne, alors C (par accord).
b. Selon Q : Si Q gagne (pour quelque raison que ce soit), alors non-C (par verdict) ; mais si Q perd, alors non-C (par accord).

---

86 Voir *Stanford Enc. of Phil.* , en ligne à l'adresse https://plato.stanford.edu/entries/lorenzo-valla/. Cette question est apparemment traitée dans *Repastinatio dialectice et philosophie.* L'auteur de l'article serait Lodi Nauta (2013).

La conclusion de Valla, telle que présentée ici [87], est obscure. Il semble faire peser sur Euathlus la responsabilité de « choisir entre l'accord et le verdict », et de se conformer soit à l'accord, soit au verdict. Cela ne nous apprend rien, car l'argument d'Euathlus montre qu'il est disposé à se conformer aux deux, même s'il les présente toutes deux comme étant à son avantage. Valla ajoute que si Protagoras perd un premier procès, il peut gagner le second. Mais là encore, cela ne résout pas le paradoxe, mais n'en reproduit qu'une partie.

Il apparaît donc que Valla n'a pas suffisamment approfondi la question : il n'envisage pas les motifs pour lesquels Q pourrait gagner ou perdre « par décision de justice » ; et il ne réalise pas que le second argument avancé par chaque partie, selon lequel Q pourrait gagner ou perdre « par accord », renvoie en réalité à une seconde action en justice, puisque l'accord n'est manifestement pas sur le point d'être mis en œuvre volontairement (la mention d'une seconde action en justice par Valla est placée après les quatre propositions conditionnelles, qui elles-mêmes ne mettent pas l'accent sur la chronologie des événements). En définitive, Valla ne parvient donc pas à une résolution simple de ce paradoxe.

Le problème de l'approche de Valla est qu'elle considère de fait la décision initiale du tribunal comme arbitraire, c'est-à-dire sans lien avec l'accord entre les parties. Elle

---

87 Je n'ai pas lu l'œuvre de Valla. Je dois supposer que l'auteur de l' article de la *SEP* sur Valla a correctement présenté son raisonnement. J'en doute cependant, car les détails fournis sont plutôt vagues et peu concluants ; ils ne précisent pas quelle solution au paradoxe Valla avait en tête. La logique n'est pas innée. (Notez que j'ai écrit à l'auteur pour lui demander de vérifier le résumé de la pensée de Valla présenté dans l'article de la SEP, mais je n'ai reçu aucune réponse. Son absence de réponse confirme mon impression que cette source n'est pas très fiable : les personnes intelligentes accueillent volontiers les questions pertinentes.)

n'examine pas sur quelle base le tribunal pourrait juger qu'Euathlus puisse perdre le premier tour. Si cela peut être acceptable juridiquement, ce n'est pas acceptable logiquement – et c'est précisément la logique qui nous intéresse ici. Logiquement, l'idée qu'Euathlus puisse perdre le premier tour est inconcevable, si l'on s'en tient à la clause explicite de l'accord. Il ne pourrait perdre le premier tour que si le tribunal admettait l'existence d'une clause tacite dans l'accord, par exemple l'obligeant à exercer le droit prochainement.

En tout état de cause, le verdict du premier tour ne saurait être arbitraire – c'est-à-dire irrationnel, injustifié – mais doit être directement lié à l'accord. Il en va de même pour le second tour. Par conséquent, aucun décalage entre le verdict et l'accord n'est logiquement admissible.

Il est intéressant de constater que ma première réaction face à ce paradoxe, il y a quelques années, [88]était très similaire à celle de Valla. Mais dès que j'ai entrepris d'examiner sérieusement les questions soulevées dans le cadre de cet essai, j'en ai réalisé la faiblesse et l'inutilité. Le paradoxe semble neutralisé si l'on introduit une distinction entre « paiement suite à un jugement » et « paiement conformément aux termes du contrat » et si l'on attribue des symboles différents à ces deux conséquences, par exemple C1 et C2. Ces deux termes peuvent être quantitativement identiques ou non ; mais ils se réfèrent de toute façon à des événements distincts. Cette mesure semble annuler le paradoxe, car les conséquences des première et deuxième affirmations conditionnelles seraient différentes pour chaque partie. Ainsi :

a. Selon P : Si P gagne, alors C1 ; mais si P perd, alors C2.

---

88 En 2009, dans un courriel adressé à quelqu'un.

b. Selon Q : Si Q gagne, alors non-C1 ; mais si Q perd, alors non-C2.

Il s'ensuit que si P gagne et Q perd, alors C1 et non-C2 sont vrais ; et si P perd et Q gagne, alors non-C1 et C2 sont vrais. Puisque C1 et C2 sont des termes différents, même s'ils se réfèrent par hasard au même montant monétaire, les conséquences de l'argument de chaque partie ne sont pas formellement contradictoires (puisque le motif déterminant est différent), il n'y a donc pas de paradoxe. Mais, je le répète, cette approche n'explique pas pourquoi le tribunal rendrait un verdict non conforme à l'accord, elle est donc artificielle. Pour une résolution véritablement logique, nous devons nous concentrer sur l'accord et ne pas accepter de verdicts arbitraires. Les verdicts arbitraires ne font que brouiller les pistes.

D'autres tentatives de résolution que j'ai pu observer sont également imparfaites, soit parce qu'elles confondent elles aussi questions juridiques et logiques, soit parce qu'elles n'examinent pas toutes les conséquences de toutes les hypothèses. Quoi qu'il en soit, chacune des deux solutions que je propose suffit à démêler le paradoxe ; mais les deux, combinées, permettent de couvrir toutes les éventualités. Notons que, dans la première solution, nous ne nous contentons pas de résoudre le paradoxe, mais nous tranchons incidemment le cas (en faveur de Protagoras) ; tandis que dans la seconde solution, nous nous contentons, en tant que logiciens, de résoudre le paradoxe, laissant au juge le soin de trancher le litige.

En conclusion : le paradoxe du tribunal est dû à plusieurs facteurs qu'il convient d'éclaircir et de prendre en compte pour le résoudre. (a) Les termes « gagner/perdre » ne sauraient désigner des jugements arbitraires ; à défaut, le paradoxe risquerait de se perpétuer. (b) Il faut comprendre que les arguments avancés par les deux parties impliquent que le processus de résolution comporte deux phases : un

premier procès suivi d'un second (ou d'une concession et d'un règlement amiables sans second procès) ; l'apparence de paradoxe tient en partie à la négligence de ce facteur temporel. (c) La notion d'« accord » est quelque peu imprécise, et toutes les interprétations possibles doivent être envisagées : si l'accord se réfère exclusivement à la clause explicite, le premier procès porte uniquement sur celle-ci et se résout aisément ; en revanche, si l'accord inclut une entente tacite, le premier procès vise à statuer sur cette clause tacite, tandis que le second porte sur la clause explicite.

14. Discours antinomique bouddhiste

Extrait de « Illogique bouddhiste 1 et 2 ».

1. Le tétralemme

La pensée philosophique et scientifique occidentale repose sur la logique aristotélicienne, dont les principes fondateurs sont les trois « Lois de la pensée ». On peut les résumer ainsi : « A est A » (Identité), « Rien n'est à la fois A et non-A » (Non-contradiction) et « Rien n'est ni A ni non-A » (Exclusion du milieu). Il est important de noter que ces lois ne sont pas de simples hypothèses, mais des prémisses incontestables de toute pensée humaine rationnelle [89].

Des religions comme le judaïsme, le christianisme et l'islam, tout en adhérant à ces lois dans une grande partie de leur discours et en les respectant du bout des lèvres, ont souvent ignoré ces mêmes lois dans leurs efforts pour interpréter leurs textes sacrés et rendre leurs doctrines acceptables aux yeux de leurs convertis. Cela est particulièrement vrai des courants mystiques au sein de ces religions, mais on pourrait citer de nombreux exemples dans les écrits traditionnels. On peut en dire autant de certains aspects de la philosophie bouddhiste.

---

89 Voir mon ouvrage *Future Logic* , chapitres 2 et 20, ainsi que des essais ultérieurs sur le sujet (publiés sur mon site web www.thelogician.net).

Le *tétralemme* [90]est un dérivé des lois de la pensée, relatif à deux termes ou propositions quelconques, notés A et B, et à leurs contraires non-A et non-B. Quatre combinaisons de ces quatre termes sont concevables : « A et B » (tous deux), « non-A et non-B » (aucun), « A et non-B » et « non-A et B » (l'un ou l'autre seulement). Selon la logique aristotélicienne, ces quatre énoncés sont incompatibles entre eux (*un seul peut être vrai* , car si deux ou plus étaient affirmés, alors « A et non-A » ou « B et non-B » ou les deux seraient vrais, ce qui est contradictoire) et exhaustifs (*au moins l'un d'eux doit être vrai* , car s'ils étaient tous niés, alors « non A et non non-A » ou « non B et non non-B » ou les deux seraient vrais, ce qui contrevient au principe du tiers exclu).

Nagarjuna remplace B par A (il considère donc le cas où B = A) et affirme que les quatre combinaisons logiques possibles s'appliquent dans ce cas particulier, rendant ainsi les propositions « A et A (= B) », « non-A et non-A (= non-B) », « A et non-A (= non-B) » et « non-A et A (= B) » logiquement acceptables. Il soutient ensuite qu'il existe quatre possibilités existentielles : **l'affirmation de A (A + A = A), la négation de A (non-A + non-A = non-A), l'affirmation et la négation simultanées de A (A et non-A), et l'absence d'affirmation ou de négation de A (ni A ni non non-A)** . Il utilise donc apparemment les principes et la terminologie de la logique commune pour aboutir à un résultat diamétralement opposé. Ceci donne à lui et à ses lecteurs l'impression qu'il est tout à fait raisonnable d'affirmer et de nier, ou de ne ni affirmer ni nier.

---

[90] Voir Cheng, pp. 36-38, à ce sujet. Il y fait référence à l'énoncé introductif du MT, ainsi qu'aux versets XVII:12a et XXIII:1a. Étymologie : du grec *tetra* (quatre) et *lemme* (alternatives). Terme forgé par opposition au dilemme « A ou non-A ».

En logique aristotélicienne, les deux dernières alternatives sont d'emblée exclues : « A et non-A » par le principe de non-contradiction et « ni A ni non-A » par le principe du tiers exclu. Il ne reste donc que « A » ou « non-A ». La position anti-aristotélicienne peut être perçue, d'un point de vue positif, comme une position anti-nominaliste, nous rappelant que les choses ne sont jamais tout à fait ce qu'elles paraissent et qu'il est impossible de les classer ou de les étiqueter précisément. Mais, en définitive, elle vise à abolir la logique ; car sans lois de la pensée, comment distinguer le vrai du faux ?

La loi d'identité « A est A » est la conviction que toute chose possède une identité (quelle qu'elle soit précisément) plutôt qu'une autre, ou qu'aucune identité du tout. Elle affirme la possibilité ultime de la connaissance et rejette le relativisme pur et simple ainsi que l'obscurantisme. Nagarjuna vise à nier l'identité.

Il convient de noter ici qu'Aristote est très précis dans sa formulation du principe de contradiction, affirmant dans sa *Métaphysique :* « Un même attribut ne peut appartenir et ne pas appartenir *simultanément au même sujet sous le même aspect* » (italiques ajoutés). Ainsi, une formulation alternative des lois de la pensée serait le « trilemme » (appelons-le ainsi) : « *soit entièrement A, soit entièrement non-A, soit à la fois partiellement A et partiellement non-A* », ce qui exclut la quatrième alternative : « à la fois entièrement A et entièrement non-A ». La critique bouddhiste des lois de la pensée tire une partie de sa crédibilité du fait que l'on se réfère inconsciemment à ce « trilemme », en pensant superficiellement que des choses opposées peuvent effectivement se produire au même endroit à des moments différents, ou simultanément en différents lieux ou sous divers aspects, sans pour autant engendrer de difficulté logique insoluble. Mais il est clair que la position bouddhiste est bien plus radicale, puisqu'elle accepte une antinomie totale.

De même, concernant le principe du tiers exclu, qui affirme l'impossibilité de la situation « ni A ni non-A », l'incertitude quant à la réalité (A ou non-A) peut induire en erreur *et faire* croire que ce principe est sujet à débat. Or, il faut comprendre que ce principe logique est inductif, et non déductif : il affirme qu'au *terme* du processus d'acquisition des connaissances, on aboutira soit à « A », soit à « non-A », sans possibilité de troisième voie. Il n'exclut pas la possibilité d'une situation d'incertitude *intermédiaire* . La position de Nagarjuna exploite cette confusion.

Nagarjuna interprète la limitation inhérente au dilemme « A ou non-A » comme un « dualisme » arbitraire propre aux penseurs ordinaires [91]. Cela démontre simplement

---

[91] Il est trompeur de parler de « dualité » ou de « dichotomie », comme le font souvent les bouddhistes, car cela suggère qu'une chose unitaire a été arbitrairement divisée en deux – et, incidemment, qu'elle aurait tout aussi bien pu être divisée en quatre. Or, au niveau perceptif, il n'y a ni choix ni « découpage ». Un phénomène qui apparaît est *une seule et même* chose, appelons-le « **un** » (un nom propre ou un « ceci » indicatif), et non une disjonction. La question de la « dichotomie » ne se pose qu'au niveau conceptuel. *La négation* est un acte rationnel : on ne peut parler de « non **-un** », de ce qui n'apparaît pas, qu'en se rappelant d'abord quelque chose de « **un** » qui est apparu précédemment (dans la sensation ou l'imagination). Lors de *la conceptualisation initiale* , deux phénomènes sont comparés et opposés, l'un à l'autre et à d'autres choses, sous un ou plusieurs aspects. La question est alors de savoir si ces phénomènes sont suffisamment similaires entre eux et suffisamment différents des autres pour être considérés comme « identiques » et désignés par un terme général (par exemple « A »), ou s'ils doivent être considérés comme « différents », ou encore s'il existe une incertitude. À *un stade ultérieur de la reconnaissance* , il nous faut décider si un troisième phénomène s'intègre à la classe formée pour les deux précédents (c'est-à-dire s'il relève de la catégorie « A »), s'il n'y intègre pas (c'est-à-dire s'il relève de la catégorie « non-A ») ou s'il demeure incertain. Dans ce dernier cas, nous nous demandons s'il s'agit d'un phénomène « A » ou « non-A », et nous précisons qu'il ne peut appartenir ni à l'un ni à l'autre.

qu'il comprend mal la formalisation (ou qu'il feint de la comprendre, dans le but de tromper les lecteurs crédules). Lorsque les logiciens utilisent une variable comme « B » et admettent que « non-A et B » et « A et non-B » sont tous deux possibles en principe, ils n'entendent pas cela comme une généralité applicable à *toutes* les valeurs de B (telles que « A »), mais seulement comme une affirmation générique applicable à *toute* valeur compatible de B. Dans le cas particulier où B = A, ces deux combinaisons doivent être éliminées car elles sont illégales (c'est-à-dire qu'elles enfreignent deux lois de la pensée).

La propriété des symboles mentionnée ci-dessus, à savoir leur applicabilité conditionnelle et limitée aux contraintes de cohérence, est manifeste dans toute la logique formelle, et Nagarjuna l'ignore totalement ici. Son objectif était bien sûr de verbaliser et de rationaliser la doctrine du Bouddha selon laquelle la vérité ultime transcende *le nom* et *la* forme (c'est-à-dire la discrimination et le discours), et n'est accessible qu'à une conscience transcendantale (la doctrine de la Double Vérité). Plus précisément, comme le souligne Cheng, Nagarjuna voulait démontrer que la logique est intrinsèquement incohérente et que, par conséquent, la raison n'est qu'une folie confuse à rejeter. Autrement dit, il ne cherchait pas (ici et ailleurs) à défendre un tétralemma avec B égal à A – ni même à affirmer que les choses sont à la fois A et non-A, ou ni A ni non-A – mais souhaitait nous amener à dépasser les distinctions conceptuelles et les jugements de la logique.

Mais comme nous l'avons vu, il n'atteint pas son but. Sa critique repose en effet sur une interprétation erronée de la science logique. Il prétend démontrer que la logique est confuse et contradictoire, mais en réalité, ce qu'il présente comme la thèse de la science logique n'est pas ce qu'elle affirme, mais précisément ce qu'elle interdit explicitement. De plus, même si la théorie logique menait effectivement à des contradictions, comme il le prétend,

ce fait ne nous conduirait pas à son rejet sans un appel implicite à notre préférence pour la logique dans la pratique. Si la logique était fausse, les contradictions seraient acceptables. Ainsi, paradoxalement, Nagarjuna fait appel à notre habitude logique lorsqu'il nous recommande précisément d'ignorer la logique. En somme, bien qu'il donne l'illusion qu'il est raisonnable d'abandonner la raison, il est facile de constater que sa conclusion est inévitable et ses moyens erronés.

2. Ni réel ni irréel

Mais Nagarjuna conçoit également la réalité ultime (« vacuité » [92]) comme une « voie du milieu ». [93]– de sorte que le monde de l'expérience ne saurait être considéré ni comme réel, ni comme irréel (« il n'y a rien, ni mental ni non mental, qui soit réel » et « cela ne peut être conçu comme irréel », rapporte Cheng). Dans ce contexte, Nagarjuna s'appuie manifestement sur l'une des disjonctions logiquement impossibles mentionnées plus haut, à savoir « ni A ni non-A » (précisons-le au passage). Je souhaite maintenant démontrer pourquoi l'affirmation de Nagarjuna paraît superficiellement raisonnable et vraie.

Comme je l'ai souvent précisé et expliqué [94], la connaissance doit être envisagée de manière phénoménologique (c'est la seule thèse épistémologique cohérente). Il faut commencer par reconnaître et observer

---

92 Au-delà de la conscience de « Shunyata » se trouve une conscience plus vive appelée « Mahamudra », selon Chögyam Trungpa, dans *Le Jeu des illusions* . Mais ces subtilités ne nous concernent pas ici.

93 Voir Cheng, pp. 38-39, sur ce sujet. Il y fait référence à MT XIII,9a et XVIII,7.

94 Voir mon *ouvrage Future Logic* , chapitres 60 à 62, et mes essais ultérieurs sur le sujet.

les apparences en tant que telles, sans porter de jugement initial sur leur réalité ou leur illusion. À première vue, toutes les apparences semblent suffisamment *réelles* . Mais après un certain temps, il faut admettre que certaines apparences entrent en conflit avec d'autres et juger ces apparences (c'est-à-dire une ou plusieurs de celles en conflit) comme *illusoires* . Puisqu'il n'y a rien dans notre « monde » que des apparences, toutes les apparences restantes qui ne sont pas jugées comme illusoires (c'est-à-dire tant qu'elles ne sont pas logiquement invalidées par des conflits avec d'autres apparences) conservent leur statut initial de réalités.

Autrement dit, la distinction entre apparences réelles et illusions émerge au sein même du monde des apparences, se limitant à classer certaines d'une manière et d'autres d'une autre. Nous n'avons de concept de réalité ou d'illusion qu'en référence à l'apparence. Utiliser la catégorie de réalité pour désigner quelque chose *qui dépasse* l'apparence constitue un détournement de concept, une extrapolation qui ignore la genèse même du concept dans le contexte de l'apparence. Appliquer le concept d'illusion à *toutes* les apparences, sous prétexte que certaines apparences sont des illusions, est une généralisation injustifiée qui ignore comment ce concept émerge en référence à un événement spécifique (à savoir, l'incohérence entre certaines apparences et la diminution conséquente de leur crédibilité intrinsèque). De plus, affirmer qu'aucune apparence n'est réelle ou que toutes sont des illusions est contradictoire, car une telle affirmation relève logiquement de la catégorie de l'apparence.

L'illusion existe sans pour autant être réalité ; elle existe en tant qu'apparence. Le réel est également apparent, du moins en partie. Par conséquent, l'apparence en soi ne saurait être comprise comme la réalité (puisque certaines apparences sont illusoires), ni assimilée à l'illusion

(puisque toutes les apparences n'ont pas été, ou ne peuvent être, illusoires). L'apparence est ainsi le *terrain d'entente* des réalités et des illusions, leur caractéristique commune, la synthèse dialectique de ces thèses et antithèses. Elle est un genre, dont les apparences et les illusions sont des espèces mutuellement exclusives. (La différence entre apparence et existence est une autre question, que j'ai traitée ailleurs ; en résumé, l'existence est un genre d'apparence et de non-apparence, ces dernières étant relatives à la conscience, tandis que la première est supposée indépendante.)

Aucune de ces observations ne permet de conclure que les apparences ne sont « ni réelles ni irréelles » (en supposant que « irréel » signifie « non réel »). On peut seulement affirmer que certaines apparences sont réelles et d'autres irréelles. Formellement, la relation logique correcte entre les trois concepts est la suivante. Déductivement , l'apparence est impliquée par la réalité et l'illusion, mais ne les implique pas ; car la réalité et l'illusion sont contradictoires, de sorte qu'elles ne peuvent être à la fois vraies et à la fois fausses. De plus, *inductivement* , l'apparence implique la réalité, jusqu'à ce qu'elle soit jugée illusoire (en raison de la découverte d'une incohérence).

Plus précisément, toutes les apparences sont initialement considérées comme réelles. Toute apparence contradictoire est (déductivement) illusoire, et sa contradiction est par conséquent évidente et (déductivement) réelle. Toutes les autres apparences restent considérées comme réelles tant qu'elles ne sont pas contestées. Celles qui sont contestées doivent être évaluées de manière dynamique. Lorsqu'une apparence est contredite par une autre, le conflit entre elles les met toutes deux en doute, et elles deviennent donc initialement *problématiques* . Par la suite, leur crédibilité relative doit être provisoirement évaluée dans le contexte

global des connaissances empiriques et rationnelles disponibles – et réévaluée de manière répétée, à mesure que ce contexte évolue. Sur cette base, l'une de ces apparences peut être jugée plus crédible que l'autre, de sorte que la première est qualifiée de *probable* (proche du réel) et la seconde de relativement *improbable* (proche de l'illusion). À la limite, elles peuvent être caractérisées comme étant respectivement (inductivement) réelles ou illusoires. Ainsi, la réalité et l'illusion sont les extrêmes (respectivement 100 % et 0 %) dans une large gamme de probabilités avec de nombreux degrés intermédiaires (y compris la problématique au point médian).

Plus précisément, *les perceptions pures* (c'est-à-dire les apparences concrètes, les phénomènes) ne sont jamais illusoires. Le jugement de valeur relatif à l'« illusion » ne concerne que les concepts (c'est-à-dire les apparences abstraites, les « universaux »). Dire d'une perception qu'elle était illusoire signifie simplement que nous l'avons mal interprétée. Autrement dit, ce que nous considérions initialement comme une perception pure contenait en réalité *un mélange de concept* , qui s'est avéré erroné. Par exemple, je vois certaines formes et couleurs au loin et je pense « voilà une fille à vélo », mais en m'approchant, je réalise que je n'ai vu qu'un tas d'ordures au bord de la route. La perception pure, ce sont les formes et les couleurs que je vois ; la fausse interprétation est « fille à vélo », l'interprétation plus juste est « tas d'ordures ». La perception initiale n'a pas changé, mais ma proximité accrue y a ajouté des détails perceptifs. Ma première impression était correcte, seul mon jugement initial était erroné. Je revois ce dernier concept, non pas par des moyens supérieurs à la connaissance, mais simplement par une *perception et une conception plus poussées* .

À proprement parler, la perception n'est donc jamais en cause ; ce sont nos conceptions que nous évaluons. Il est vrai qu'en pratique, il est souvent très difficile d'isoler une

perception de son interprétation, c'est-à-dire des concepts qui y sont rattachés. Notre perception des choses est, en effet, dans une large mesure « eidétique ». Ce fait ne doit cependant pas nous amener à rejeter toute perception (comme le font rapidement nombre de philosophes occidentaux, ainsi que des bouddhistes), ni même toute conception. Les « impuretés » conceptuelles des perceptions ne sont pas nécessairement erronées. Nous savons qu'elles étaient erronées lorsque nous décelons un motif précis de contestation : une contradiction logique ou empirique. Tant que nous ne trouvons pas un tel défaut précis, elles peuvent être considérées comme justes. Cela signifie simplement que nous devons appliquer les règles de l'adduction [95]à nos interprétations immédiates des perceptions individuelles, tout comme nous le faisons pour les théories complexes relatives à des ensembles de perceptions. Ces règles sont universelles : aucun jugement n'échappe à l'exigence d'un examen attentif et d'une réévaluation.

Or, à la lumière du récit de Cheng et de certaines citations de Nagarjuna qu'il contient, on pourrait interpréter ce dernier comme ayant voulu exprimer précisément ce que j'ai dit. Par exemple, Cheng écrit[96] : « Ce que Nagarjuna voulait nier, c'est que les phénomènes empiriques… soient absolument réels… Cependant, [cela] ne signifie pas que rien n'existe. *Cela n'annule rien dans le monde* » (c'est moi qui souligne). J'interprète cette non-annulation comme une reconnaissance de l'apparence comme fondement minimal de la connaissance. Nagarjuna a peut-être éprouvé des difficultés à élaborer une terminologie appropriée (distinguant existence, apparence

---

[95] L'adduction considère toute connaissance conceptuelle comme hypothétique, à tester sans cesse – en concurrence avec toutes les hypothèses alternatives concevables – en se référant à toute la logique et à l'expérience disponibles.

[96] P. 42.

et réalité, comme je le fais ci-dessus), sans doute influencées par son penchant pour les affirmations paradoxales qui semblaient exprimer et confirmer la doctrine mystique bouddhiste.

Mais si tel est son propos, alors il n'est pas parvenu à une « voie médiane » (un rejet du principe du tiers exclu), mais seulement à une « voie commune » (un terrain d'entente admis). À mes yeux, il ne s'agit pas d'un mince accomplissement – la découverte philosophique de la phénoménologie ! Mais pour lui, ce serait trivial, voire contre-productif, car ce qu'il cherche, c'est à nier la conscience ordinaire et ses raisonnements limitants, pour accéder ainsi à une conscience différente, supérieure, capable d'atteindre la vérité transcendantale ou la réalité ultime.

Il est intéressant de noter que le rejet effectif de la réalité par l'école Madhyamika de toute apparence ne fut pas accepté par une école ultérieure de la philosophie Mahayana, le Yogacara (VIIe- VIIIe siècle) . Cheng décrit ainsi la position de ce dernier[97] : « Tout objet, mental ou non, peut être logiquement ou dialectiquement démontré comme illusoire. Mais pour être illusoire, il faut qu'il existe une certaine pensée sujette à l'illusion. *Le fait même de l'illusion prouve l'existence et la réalité d'une certaine conscience ou d'un certain esprit* . Affirmer que tout ce qui est mental et non mental est irréel relève du suicide intellectuel. Il faut au moins admettre la réalité de quelque chose pour que parler d'illusion ait un sens » (italiques ajoutés). Tel est le sens de l'argument phénoménologique que je présente ci-dessus, bien que ma conclusion finale diffère clairement de celle du Yogacara, selon laquelle tout est conscience ou esprit (une forme d'idéalisme), mais laisse ouverte la possibilité de juger et de classer les

97 P. 25.

apparences comme matière ou esprit en fonction de diverses considérations.

Le rejet du dualisme par l'école Madhyamika va jusqu'à affirmer que la vacuité ne se trouve pas dans le nirvana, antithèse du samsara (selon la conception bouddhiste antérieure), mais dans ce qui n'est ni le samsara ni le nirvana. En réalité, des affirmations similaires se retrouvent dans le Canon Pali, c'est-à-dire dans les écoles Theravada bien plus anciennes, ce qui démontre qu'il ne s'agit pas d'une construction spécifiquement Mahayana. La différence réside dans l'accentuation : ces affirmations, relativement rares dans la période antérieure, sont devenues la norme et sont fréquemment reprises dans la période postérieure. On en trouve un exemple dans le *Dhammapada* , un sutra datant du IIIe [siècle] avant notre ère [98], soit quatre ou cinq siècles avant Nagarjuna. Ici, le samsara est comparé à un cours d'eau ou à sa rive actuelle, et le nirvana à la rive opposée ; et il nous est demandé de transcender ces deux extrêmes.

> *Lorsque tu auras traversé le fleuve du Samsara, tu atteindras le Nirvana… Celui qui a atteint l'autre rive accède alors à la vision suprême et tous ses liens sont brisés. Celui pour qui il n'y a ni cette rive ni l'autre, ni les deux…*

---

98 Il s'agit vraisemblablement de la date de composition, bien que le traducteur, Juan Mascaro, dans son introduction, indique « compilé » à cette époque, ce qui semble suggérer une composition antérieure. Ce commentaire ne précise pas à quelle date le sutra aurait été mis par écrit pour la première fois. Si cette mise par écrit est beaucoup plus tardive, par exemple durant la période de cristallisation de la pensée mahayana, entre 100 avant notre ère et 100 de notre ère, cette dernière a pu influencer les moines qui l'ont transcrit. Voir le chapitre 26 (383-385) pour la citation.

Une telle formule est légitime si on l'interprète comme un avertissement : *la poursuite* du nirvana (l'éveil et la libération) constitue un obstacle à son accomplissement, une forme subtile de samsara (l'ignorance et l'attachement). Il n'y a aucune contradiction à affirmer que *la conception du* nirvana comme but de l'action nous maintient dans le samsara – il s'agit d'une simple affirmation causale. Cette formule est également logiquement acceptable si on la considère comme un rappel qu'aucun mot ni concept – pas même « samsara » ou « nirvana » – ne peut saisir ou transmettre toute la portée des significations voulues (autrement dit, « pas » devrait être plus précisément remplacé par « pas tout à fait »). Il n'y a pas non plus de contradiction à dire que celui qui a atteint le nirvana n'a pas besoin de quitter le monde de ceux qui sont prisonniers du samsara, mais peut continuer d'y exister et d'y agir, d'une manière distinctement détachée de tout attachement.

Mais il serait contradictoire de parler de la « vacuité » comme n'étant « ni le samsara ni le nirvana », puisque le nirvana, en tant que concept, est originellement défini comme non-samsara ; la vérité ne saurait être une troisième voie. Tout au plus, on pourrait dire que la vacuité est un niveau supérieur de nirvana (au sens élargi), à ne pas confondre avec le niveau inférieur désigné par le terme originel de nirvana, ni bien sûr avec le samsara. Dans ce cas, le nirvana (au sens générique du terme, signifiant littéralement non-samsara) englobe à la fois une catégorie supérieure et une catégorie inférieure ; et l'affirmation « ni le samsara ni le nirvana inférieur » est alors compatible avec l'affirmation « nirvana supérieur ». Il existe une différence fondamentale entre un langage poétique, expressif et dramatique, et son interprétation littérale.

15. Discours antinomique bouddhiste plus avancé

Extrait de Ruminations 5 et Réflexions logiques et spirituelles 3.10.

1. La thèse « Je n'ai pas de thèse »

a. Le [99]philosophe bouddhiste Nagarjuna (Inde, *vers* 150-250 apr. J.-C.) s'attaquait à toutes les thèses qu'il considérait comme rationnelles par tous les moyens qu'il considérait comme logiques, et déclarait son propre discours à l'abri de tout examen et de toute critique, en disant (selon une traduction) :

> « *Si j'avais une thèse, je serais en faute ; puisque je suis le seul à n'avoir aucune thèse, je suis le seul à être sans faute* » (VV 29) [100].

---

[99] Il va sans dire que les commentaires qui suivent ne constituent pas une attaque contre le bouddhisme, mais contre la rhétorique de Nagarjuna. Le bouddhisme tire de mauvais enseignements de tels jeux de pouvoir. Je pense à Nagarjuna chaque fois que je lis le verset 306 du *Dhammapada* : « Celui qui dit ce qui n'est pas… et celui qui prétend ne pas avoir fait ce qu'il sait pertinemment avoir fait… a péché contre la vérité. » Pour moi, il n'est qu'un philosophe parmi d'autres ; son intérêt pour le bouddhisme est secondaire (tout comme son statut de saint aux yeux de beaucoup).

[100] Nagarjuna, dans *Vigraha Vyavartani* (*Éviter les arguments*), verset 29. La traduction utilisée ici est celle de « Namdrol » sur le forum E-Sangha Buddhism (http://www.lioncity.net/buddhism/index.php?s=d8946a5bcb1f56f3e9

Le premier aspect de la déclaration de Nagarjuna est une **affirmation audacieuse : celle de n'avoir aucune prétention.** Ceci est bien sûr contradictoire. Toute proposition qui se prétend significative et vraie (qu'elle porte sur une expérience concrète ou sur une abstraction, qu'elle soit positive ou négative) est une assertion, une prétention. Prétendre ne pas avoir de prétention tout en en formulant une clairement constitue une violation du principe d'identité : c'est nier qu'un fait soit un fait.

Il est impossible de réfuter ou de critiquer logiquement les thèses ou les méthodologies d'autrui sans soumettre son propre discours à l'évaluation. Toute réfutation ou critique est un discours, et tout discours est susceptible d'examen logique. Prétendre à la possibilité logique d'une dispensation est malhonnête (et si une telle prétention est déjà grave, elle l'est d'autant plus lorsqu'elle est formulée explicitement).

Le discours de Nagarjuna était en réalité (comme je le démontre dans **L'Illogisme Bouddhiste**) un véritable festival d'arguments fallacieux, une simple parodie de la logique se faisant passer pour telle. Mais il savait que les personnes non initiées à la logique se laisseraient prendre au piège, et il scella leur destin intellectuel par cette affirmation trompeuse. Afin d'étouffer toute discussion ultérieure, il les induisit en erreur en leur faisant croire qu'il avait simplement mis en évidence l'absurdité logique de la logique et de toutes les doctrines qui en découlent, sans pour autant avoir lui-même proposé aucune méthodologie ni doctrine propre.

---

e21a108125823f&showtopic=5604&st=100&#entry82577). Il convient toutefois de noter que le mot « seul » dans cette traduction pourrait ne pas figurer dans l'original, d'après d'autres traductions que j'ai consultées, même si cela semble correspondre à l'intention de Nagarjuna.

Non seulement sa prétendue réfutation de la raison était truffée d'erreurs de raisonnement, mais sa conclusion, « affirmation sans fondement », était une parodie de logique et de sincérité. Il se contente de dire « Je n'affirme rien » et nie obstinément que cette déclaration constitue une affirmation. J'qualifie cela de manipulation psychologique éhontée, motivée par un désir de supériorité. Il exploite cyniquement la crédulité de certains pour les dominer intellectuellement.

Le second aspect de la déclaration de Nagarjuna ci-dessus peut être interprété comme une version atténuée du paradoxe du menteur, puisqu'il affirme : « **Tous, sauf moi, sont dans l'erreur »** . Bien qu'une telle affirmation ne soit pas en soi incohérente (Dieu pourrait concevablement la prononcer en toute vérité), elle prête logiquement à caution, car elle *s'auto-exempte* .

En substance, cela revient à dire : « Je suis le seul être humain à posséder la connaissance ; je sais que tous les autres en sont incapables. » Seul un imbécile se laisse abuser par une telle prétention, dénuée de fondement, à s'arroger un privilège. La raison considère que tout être humain est, en théorie, susceptible d'acquérir la connaissance moyennant un effort suffisant, même si tous n'atteignent pas leur plein potentiel de la même manière. La raison exige que le discours soit raisonné et équitable, c'est-à-dire fondé sur *des normes générales communes* permettant de déterminer le vrai et le faux.

Si Nagarjuna fondait sa critique des moyens humains ordinaires d'accès à la connaissance sur l'affirmation d'avoir atteint un *« niveau supérieur » de conscience* (c'est-à-dire l'éveil bouddhiste ou la prophétie biblique), nous ne pourrions pas le contredire de manière convaincante (étant incapables de prouver ou de réfuter de telles affirmations empiriques). Mais il ne se base pas sur une telle affirmation ; il tente de réfuter la raison par un

discours logique *ordinaire* . Dès lors, il devient une cible légitime pour la logique.

L'affirmation d'infaillibilité est alors perçue comme une arrogance manifeste, un manque de respect pour les autres penseurs. En déclarant « Je suis le seul à être exempt de toute critique », l'auteur s'octroie de manière agressive un privilège : il serait le seul détenteur du savoir ; tous les autres seraient des idiots ou des malhonnêtes. Il s'agit d'un discours totalitaire et dictatorial.

Comparez ce méprisant « vous ne savez rien » à l'affirmation inclusive « je (ou nous) ne savons rien ». Cette dernière – même si elle sous-entend « je sais que je ne sais rien » et est donc incohérente – est au moins modeste ; à tel point qu'un tel aveu est largement considéré comme une marque de sagesse (et il est louable, sous une forme modifiée, c'est-à-dire « je ne sais *presque* rien, très peu »).

L'auto-exemption est une forme insidieuse d'incohérence, car elle repose sur *un double discours* . Celui qui s'en prévaut présente des arguments en apparence rationnels contre l'expérience et la logique humaines, sans se demander comment lui-même (un être humain ordinaire) est parvenu (par les moyens cognitifs mêmes qu'il rejette) à ce savoir prétendument vrai. L'auteur critique autrui, mais ne s'autocritique pas avec la même rigueur.

Il s'agit d'un raisonnement fallacieux, fréquent chez les prétendus philosophes sceptiques. Il se manifeste sous de nombreuses formes subtiles. Il est sage de toujours rester vigilant face à de telles pratiques, en appliquant le test réflexif présenté ici.

b. En examinant plus en détail la déclaration de Nagarjuna ci-dessus, on peut ajouter ce qui suit.

Pour commencer, que signifie ici « ***avoir une thèse*** » ? Cela renvoie à toute croyance, explicite ou même tacite, à toute opinion, claire ou même vague, défendue (et

considérée comme constituant une connaissance), à toute proposition que l'on soutient ou que l'on approuve implicitement par la logique. Le sujet abordé par Nagarjuna est ici tout résultat de la cognition rationnelle humaine, toute croyance, opinion ou doctrine à laquelle on peut parvenir, à tort ou à raison, par la conscience ordinaire, c'est-à-dire par l'expérience, la négation, l'abstraction, la formulation d'hypothèses, le raisonnement inductif ou déductif.

Que signifie ici « *être en faute* » ? Cela fait référence au fait de commettre une erreur dans le processus d'observation ou de raisonnement, de sorte qu'une thèse à laquelle on a adhéré se révèle en réalité une illusion plutôt qu'une réalité, fausse plutôt que vraie, erronée plutôt que correcte.

Comment déterminer le statut approprié à une thèse ? Nous le déterminons (à mon avis) *par l'application holistique de toute la science de la logique à la totalité des données de l'expérience* . Nos concepts cognitifs de bien et de mal sont eux-mêmes entièrement construits par la logique et l'expérience, sans recours à une justification extérieure extraordinaire (telle qu'une révélation prophétique ou une réalisation mystique, ou simplement l'autorité d'un grand personnage, d'un texte ou d'une institution religieuse).

Nagarjuna en est manifestement parfaitement conscient, mais il s'attache à nier la fiabilité indépendante de l'expérience et de la raison ordinaires. Sa stratégie, dans l'ensemble de son discours, comme je l'ai démontré dans mon ouvrage « *L'Illogisme bouddhique* » , consiste à donner l'impression (aussi paradoxale soit-elle) que la logique peut être invalidée par la logique elle-même. Et cette double phrase, « Si j'avais une thèse, je serais en faute ; puisque je suis le seul à n'en avoir aucune, je suis

le seul à être sans faute » , s'intègre parfaitement à son programme philosophique destructeur.

À première vue, cette phrase pourrait être interprétée comme un argument unique :

Si X (une proposition est proposée), alors Y (une erreur est commise)

mais pas X (pas de proposition)

Donc, non Y (pas d'erreur)

Bien que cette apodose soit logiquement invalide, puisqu'elle nie l'antécédent pour nier le conséquent, Nagarjuna ne s'interdit pas de la laisser passer sans commentaire, sachant qu'elle suffira à convaincre certains, même s'il est parfaitement conscient que les esprits logiques la décèleront et la contesteront. Mais pour ce dernier public, il réserve une forme de manipulation plus subtile.

Il faut bien comprendre que le but de ce fameux verset 29 du discours de Nagarjuna est *de donner l'illusion d'une cohérence logique* . Il souhaite ainsi faire croire que son discours anti-rationnel est justifiable, qu'il bénéficie de l'approbation de la logique. Certes, il s'attaque en réalité à la logique ; mais simultanément, il doit feindre de l'utiliser, car il sait que cette manœuvre est nécessaire pour convaincre. Pour la plupart des gens, un vernis de logique (c'est-à-dire une simple rhétorique) suffit à apaiser leur esprit critique. Nous allons maintenant voir comment il s'y prend.

La première partie de l'affirmation de Nagarjuna, à savoir : « Si j'avais une thèse, je serais en faute », ne vise pas (contrairement à ce que certains ont supposé) à justifier l'ensemble de son discours. Elle n'y figure pas comme un argument d'une force intrinsèque, qui étayerait

ou prouverait directement sa philosophie. Il ne s'agit certainement pas d'un principe logique évident, ni d'un axiome faisant l'unanimité, à partir duquel son discours pourrait être déduit ou même généralisé. Non, elle est elle-même une déduction et une application de la thèse principale de Nagarjuna, à savoir l'affirmation que « Toute connaissance humaine fondée sur l'expérience et la raison ordinaires est nécessairement erronée ».

Cette dernière affirmation sous-jacente constitue sa prémisse majeure dans une éducation productive (ici tacite), c'est-à-dire une éducation qui déduit une proposition hypothétique particulière à partir d'une proposition catégorique plus générale [101]. Cet argument est formellement valide et se présente comme suit :

Tous les X (opinions) sont nécessairement Y (erronés) ;
donc,
Si ceci est X (une proposition est proposée), alors ceci est Y (une erreur est commise).

Ainsi, la première partie de l'affirmation de Nagarjuna apparaît comme une déduction plutôt que comme une assertion arbitraire. Elle est habilement présentée comme une application d'informations déjà admises, et non comme une assertion isolée. Admettant la prémisse, la conclusion en découle logiquement (c'est le vernis de la logique) – mais la prémisse est-elle pour autant admise ? Non. Notons également qu'une fois la conclusion apparemment tirée, elle peut, par généralisation, servir à renforcer la prémisse ; bien qu'il s'agisse d'un raisonnement circulaire, l'effet psychologique est saisissant.

---

101 Voir *Future Logic* , chapitre 29.3.

De plus, Nagarjuna parvient, par cet argument productif implicite, à feindre la cohérence : il nous dit en substance : « Voyez, je ne me contente pas d'attaquer le savoir d'autrui, mais je suis prêt à appliquer la même rigueur critique au mien ! » Cette déclaration vertueuse n'est bien sûr qu'une illusion, car il ne remet pas en cause le principe général, mais le réaffirme simplement. Il n'a jamais établi que « Toutes les propositions sont fausses ». Sa posture est pseudo-logique.

Comme le précise la suite de sa déclaration, il ne considère pas que son discours relève de la règle critique qu'il a formulée. L'affirmation « Si j'avais une thèse, je serais en faute » est une hypothèse contrefactuelle ; son propre discours n'est jamais soumis au débat. Il semble ouvert d'esprit, mais il s'agit d'un jugement préconçu. Son intention est d'« éviter toute controverse » et de se placer d'emblée hors de la mêlée. Il semble d'abord admettre, puis nier avec véhémence que son propre discours soit le fruit d'une conscience ordinaire. Cette habile dérobade à la responsabilité cognitive a trompé bien des personnes.

Passons maintenant à la seconde partie de l'affirmation de Nagarjuna, à savoir : « Puisque je suis le seul à n'avoir aucune thèse, je suis le seul à être sans faute. » Comme nous l'avons déjà souligné, on peut y voir la prémisse mineure et la conclusion d'une apodose invalide, la première partie de l'affirmation constituant la prémisse majeure. On pourrait toutefois supposer, avec plus de bienveillance, que Nagarjuna entendait une apodose valide, dont la prémisse majeure implicite serait la proposition évidente : « Si l'on n'a pas de thèse, on ne peut se tromper. »

On peut affirmer sans exagérer que cette prémisse est restée tacite simplement parce qu'elle est évidente et acceptée de tous. Il est vrai que si l'on ne prononce aucune parole, ne pense à aucune chose, ni même n'a aucune

intention, si l'on ne nourrit aucune opinion, ne prétend à aucune connaissance, si l'on demeure silencieux intérieurement et extérieurement, on ne commettra jamais d'erreur. *Car le statut de vérité ou de fausseté ne s'applique qu'aux assertions significatives.*

Une pierre ne se trompe jamais, car elle est dépourvue de pensées. De même, une personne insouciante peut, par ignorance, cécité ou stupidité, commettre de nombreuses erreurs dans sa vie, mais ne commet aucune erreur au sens logique du terme, c'est-à-dire aucune erreur en proposant une idée inappropriée. Tout cela est si évident (et si vide de sens) qu'il est inutile de le débattre. L'apodose suivante est donc implicite dans la déclaration de Nagarjuna :

Si ce n'est pas X (aucune proposition n'est proposée),
alors Y (aucune erreur n'est commise)
mais pas X (pas de proposition)
Donc, non Y (pas d'erreur)

Cet argument repose sur une prémisse majeure véritable, et sa forme est valide. Cela confère à son discours une apparence de logique, lui permettant de convaincre davantage de personnes. Cependant, sa prémisse mineure demeure sujette à caution, voire réfutable ! (Par conséquent, sa conclusion est elle aussi sujette à caution.) Il tient pour acquis qu'il « n'a pas de thèse », mais cette affirmation est loin d'être acquise. La prémisse majeure tacite masque la prémisse mineure.

De plus, il convient de noter que, bien qu'« avoir raison » implique « ne pas être en faute », l'inverse n'est pas nécessaire. Nagarjuna suggère que sa prétendue innocence implique la justesse de sa position, mais cela n'en découle pas ! Ce n'est que si sa critique de toutes les thèses opposées était correcte (ce qui est loin d'être vrai),

et si sa thèse n'était pas susceptible d'une critique similaire et constituait donc la seule possibilité logique restante, qu'une telle conclusion pourrait être tirée.

Nagarjuna a bel et bien une thèse. Sa thèse principale, qui constitue l'objectif de tout son discours philosophique, est, comme mentionné précédemment, l'affirmation que « toute connaissance humaine fondée sur l'expérience et la raison ordinaires est nécessairement erronée ». Ceci, pour commencer, constitue une thèse – et une thèse sceptique de grande envergure, riche en implications négatives. C'est un principe de logique que *nier une thèse revient à affirmer une thèse contraire.* Son affirmation selon laquelle sa doctrine ne serait pas une thèse, dans sa prémisse mineure, n'est qu'une simple assertion arbitraire.

De plus, il « émet une thèse » à chaque fois qu'il formule une assertion, quelle qu'elle soit, y compris celle qui est ici examinée : « Si j'avais une thèse, je serais en faute ; or, puisque je suis le seul à n'en avoir aucune, je suis le seul à être irréprochable. » Notons que Nagarjuna considère qu'énoncer une affirmation négative revient en quelque sorte à « ne pas avoir de thèse » – or, la polarité d'une affirmation ne diminue en rien la nécessité de la justifier ; au contraire, on pourrait même affirmer que les affirmations négatives sont plus difficiles à établir que les affirmations positives !

Il convient d'inclure dans ses « thèses » non seulement ses affirmations explicites, mais aussi toutes les présuppositions et suggestions implicites de son discours (comme la prémisse majeure implicite et l'apodose qui en résulte, que nous venons de souligner). Peu importe que ces éléments explicites, non énoncés et non admis, constituent de l'information, une méthode logique, un contenu ou un processus.

Tous ces éléments du discours, qu'ils soient verbaux ou implicites, correspondent en toute équité à notre

compréhension et à notre définition communes de ce que signifie « avoir une thèse ». Aucune de ces propositions catégoriques ou hypothétiques (à l'exception peut-être de « le silence n'est pas une erreur ») n'est une évidence. Elles ne sont pas apparues *ex nihilo* dans l'esprit de Nagarjuna, toutes faites et auto-justifiées.

Ce sont tous des produits complexes de la cognition humaine ordinaire, fondés sur l'expérience et élaborés par la raison (même si, dans le cas de Nagarjuna, l'esprit en question est dérangé). Ils forment indéniablement ensemble une philosophie spécifique, une théorie de la logique, une épistémologie et une ontologie. Le simple fait que nous puissions (comme cela a été fait ici) les examiner et les débattre prouve qu'il s'agit de thèses.

Le principe d'identité (A est A) doit être respecté : les faits sont les faits et il est vain de prétendre le contraire. Nagarjuna peut nier indéfiniment l'existence d'une thèse, mais nous l'affirmons avec conviction. Ses arguments n'ont en rien réussi à réfuter ce jugement juste et incontestable. Par conséquent, sa doctrine est contradictoire. Non seulement il prétend avoir une thèse, mais puisque sa thèse est que « avoir une thèse, c'est être dans l'erreur », il doit, de par sa propre formulation, être reconnu comme étant dans l'erreur.

En conclusion, l'affirmation de Nagarjuna, « Si j'avais une thèse, je serais en faute ; or, puisque je suis le seul à n'en avoir aucune, je suis le seul à être irréprochable », tisse une toile complexe de tromperie. Elle induit en erreur par des ambiguïtés subtiles et des imitations superficielles de la logique. Une fois sa malhonnêteté révélée, elle doit être catégoriquement rejetée.

Le simple fait historique que Nagarjuna soit célèbre et admiré par beaucoup ne justifie pas de s'accrocher *obstinément à sa doctrine* , en tentant a posteriori de la rendre logique. La célébrité n'est pas la preuve d'une

vérité cachée ; c'est de la vanité. La plupart de ceux qui s'y adonnent ne cherchent qu'à se glorifier par procuration. De toute façon, l'attachement à l'autorité est un argument *ad hominem* . Les « groupies » religieux et universitaires qui lui ont conféré et perpétuent son autorité sont, aussi nombreux soient-ils, dépourvus de logique. C'est un cas où l'aveugle guide l'aveugle.

c. Nagarjuna défend son idée de « non-thèse » dans le verset suivant (VV 30), la décrivant comme « ***une non-appréhension des non-choses*** » (selon une traduction [102]). Or, cette expression est pour le moins surprenante. Pour les esprits impressionnables, elle paraît d'une grande profondeur, chargée de sens. Elle semble suggérer que cet homme possède une forme de connaissance supérieure et privilégiée, qui transcende l'expérience et le raisonnement ordinaires.

Mais à vrai dire, au sens littéral, nous sommes tous parfaitement capables de « non-appréhender le non-être » et nous le pratiquons quotidiennement, pour la simple raison que le non-être ne peut être appréhendé ! Logiquement, c'est tout ce que signifie cette expression, notez bien. Que manigance donc le vieux renard ?

---

102 Par Frederick J. Streng. Le texte intégral de sa traduction semble être celui publié sur Internet à l'adresse suivante : http://www.orientalia.org/article491.html . Il est à noter que Plamen Gradinarov considère l'expression « non-appréhension des non-choses » comme une traduction incorrecte. Toutefois, tout en admettant son objection, je ne partage pas l'avis selon lequel la traduction plus libre de Streng serait totalement inadmissible. À mon sens, même si elle n'est pas littéralement exacte, elle saisit l'esprit et l'intention paradoxaux de Nagarjuna. Voir notre discussion à ce sujet à l'adresse suivante : http://nyaya.darsana.org/topic3.html . Quoi qu'il en soit, même s'il est préférable de ne pas se fier à l'expression « non-appréhension des non-choses », les critiques formulées à l'encontre de Nagarjuna dans cette section peuvent être justifiées par d'autres arguments.

Nagarjuna cherche à éloigner le plus logiquement possible sa position de « non-thèse » de notre « thèse » plébéienne, c'est-à-dire de la conscience ordinaire, qui consiste en « l'appréhension des choses ». Logiquement, il n'a que trois alternatives :

- la « non-appréhension des choses » (inconscience) ;
- la « perception des non-choses » (une conscience d'un autre monde) ;
- ou la « non-appréhension des non-choses ».

Doté d'un goût prononcé pour la surenchère et les effets dramatiques extrêmes, Nagarjuna choisit naturellement le troisième de ces termes comme véhicule. Bien que le sens premier de cette expression soit puéril, elle possède une ampleur et un attrait poétiques. Elle semble impliquer à la fois « connaissance sans conscience » et « conscience de l'inconnaissable ».

Ainsi, sa « non-appréhension » est un mélange d'appréhension et de non-appréhension, ou encore autre chose. De même, ses « non-choses » sont à la fois des choses et des non-choses, ou peut-être quelque chose de tout à fait différent.

En d'autres termes, les termes négatifs de l'expression « non-appréhension des non-choses » ne sont ni conçus par Nagarjuna ni perçus par ses disciples et étudiants comme de simples négations des termes positifs correspondants, mais comme **des termes paradoxaux** qui peuvent (conformément au schéma du tétralemme) être à la fois *positifs ou négatifs, les deux ou ni l'un ni l'autre* .

C'est (et ce n'est pas) « l'appréhension/la non-appréhension des choses/des non-choses » tout en un.

Nagarjuna se distingue dans l'histoire de la philosophie mondiale comme l'opposant le plus virulent aux lois de la pensée. Non seulement il recourt librement à des

propositions contradictoires ou contenant un juste milieu, mais il utilise même des termes chargés de contradiction et d'inclusion d'un juste milieu.

Certains pourraient dire : « Quel est le problème ? » Ils argumenteront : « Le monde réel est extrêmement subtil, et nous ne pouvons espérer l'exprimer par la pensée que de manière très approximative ; Nagarjuna essaie simplement de prendre cette incertitude en considération dans son discours ; les lois de la pensée ne sont que des exigences arbitraires, nous obligeant à enfermer nos pensées dans des carcans préjudiciables. »

Mais le laxisme logique n'est pas la bonne attitude face à un monde réel extrêmement complexe et difficile à exprimer. C'est précisément en raison de la grande difficulté de la tâche cognitive à accomplir qu'il convient d'être très clair et rigoureux. Éviter de contrôler et de contrebalancer nos jugements n'en accroît pas l'efficacité, mais les rend moins fiables.

Dans le cas présent, si Nagarjuna possède effectivement une forme privilégiée de conscience transcendante, il peut tout simplement l'affirmer. Les lois de la pensée ne lui interdisent en aucun cas de formuler une telle affirmation. Il n'a pas besoin de tourner autour du pot ni de prétendre détenir quelque chose d'indicible et d'intouchable. Il peut et doit être franc et défendre sa position avec la même équité que tout le monde.

S'il considère que les termes « appréhension » et « choses » présentent un défaut logique intrinsèque, il peut défendre son point de vue ouvertement ; il n'a pas besoin de recourir à l'allusion, à la suggestion ni à des raisonnements fallacieux. La plupart des penseurs sont ouverts d'esprit et disposés à corriger leurs erreurs : si ces termes sont erronés, nous ne leur sommes pas attachés ; nous sommes flexibles, prêts à les modifier ou à les

remplacer si nécessaire, à la lumière de nouvelles preuves et d'un raisonnement nouveau.

Mais Nagarjuna est comme un accusé qui, contraint de comparaître, refuse de décliner son identité, de reconnaître l'autorité de la loi et des juges, de plaider coupable ou non coupable, ou de présenter sa défense. Pire encore, dans un mépris total du tribunal, il nie même que son refus soit un refus – il le qualifie de « non-thèse ». Cela suffit-il à interrompre la procédure ou à obtenir son innocence ? Certainement pas.

Nagarjuna se méprend sur la nature de la négation. Il pense que si une personne affirme « X » et une autre « non X », la charge de la preuve incombe davantage à la première qu'à la seconde. Il considère qu'énoncer une affirmation positive est plus exigeant logiquement qu'énoncer une affirmation négative. *Dans sa confusion, il imagine que dire « non » équivaut à ne rien dire, c'est-à-dire à ne rien dire du tout* . La plupart des logiciens seraient en désaccord avec lui et soutiendraient que toute thèse avancée (même par simple insinuation) nécessite une preuve, quelle que soit sa polarité.

J'irais même jusqu'à dire qu'au contraire, une affirmation négative est plus exigeante qu'une affirmation positive. On peut prouver une affirmation positive assez facilement, en apportant des preuves suffisantes en sa faveur. Mais comment prouver une affirmation négative ? C'est beaucoup plus difficile, car les négations ne sont pas vécues directement ; elles ne le sont que par l'absence de positivité. Une négation ne peut finalement être prouvée qu'indirectement, par l'impossibilité de prouver toute affirmation positive contraire.

Ainsi, en réalité, non seulement la prétendue limitation de Nagarjuna aux négatifs ne l'exempte pas de toute démonstration, mais au contraire, elle alourdit la charge de la preuve qui lui incombe. Il a raison de considérer les

négatifs comme fondamentalement différents des positifs, mais il ne se rend pas compte que cette différence lui est défavorable. Il prétend n'avoir aucun fondement épistémologique ou ontologique, et pourtant il s'arroge le droit de rejeter d'emblée toutes les théories de la connaissance et de la réalité. Une telle prétention grandiose et fantaisiste exige assurément une justification bien plus solide que toute autre !

Il convient de souligner, au passage, que la « non-appréhension des non-choses » de Nagarjuna ne doit pas être interprétée (comme certains le font) comme une défense de l'expérience ou de la perspicacité méditative non verbale. Ce n'est pas là l'essence de sa philosophie anti-rationnelle, même si son affiliation bouddhiste revendiquée pourrait laisser penser le contraire.

Si Nagarjuna était un homme profondément absorbé par la méditation, il n'écrirait pas de philosophie. Si son intention était de promouvoir la méditation, il se contenterait d'enseigner des méthodes et n'attiserait pas de polémiques. Non, cet homme nourrit des ambitions philosophiques. Il prétend vouloir exprimer par écrit certains des « raisonnements » que le Bouddha aurait suivis avant d'atteindre l'éveil. Nagarjuna présuppose d'emblée que ces « raisonnements » sont nécessairement antilogiques, un rejet de la raison.

Mais il faut bien comprendre que cette supposition n'est qu'un préjugé issu d'un esprit déformé. C'était un révolutionnaire de la philosophie, convaincu que la raison devait être renversée, transcendée. Or, une perspective évolutionniste est plus crédible : elle considère la méditation comme un moyen de progresser, au-delà des limites de la pensée discursive, sans pour autant renier cette dernière dans son cadre propre.

Plaider pour le respect de la logique ne revient pas à alimenter des bavardages interminables, mais plutôt à

exiger que toute pensée soit soumise à une évaluation cognitive rigoureuse. La logique est possible sans aucun mot, par le biais d'intentions silencieuses. Même en méditation profonde, une forme de vérification de la réalité par la logique s'opère, et celle-ci ne requiert aucun mot. C'est sans doute uniquement par ce moyen qu'un futur Bouddha peut se préserver des illusions communes et des conceptions délirantes, et ainsi atteindre la pleine réalisation.

Contrairement à ce que pensait Nagarjuna, rationalité et spiritualité ne sont pas nécessairement en conflit. Raison et méditation peuvent même, dans une certaine mesure, se renforcer mutuellement. Ce n'est pas la pensée en tant que telle, et encore moins la logique, mais plutôt l'excès de pensée, notamment les bavardages inutiles, qui entrave la concentration méditative et la contemplation. Une certaine dose de réflexion appropriée est souvent nécessaire pour préparer l'esprit à la méditation.

d. En fait, comme je vais le démontrer, la phrase « Si j'avais une thèse, je serais en faute » [103]est une **impossibilité formelle** . Je l'ai précédemment interprétée et symbolisée comme « Si X (une proposition est proposée), alors Y (une erreur est commise) », en attribuant à l'antécédent et au conséquent deux symboles distincts, X et Y. Mais examinons maintenant ces constituants de plus près.

Que signifie ici « faire une erreur » ? Il ne s'agit pas d'un prédicat ordinaire. Le conséquent Y ne se réfère pas simplement à une erreur en général, mais spécifiquement à une erreur *dans l'antécédent* X. Y nous indique que X

---

[103] Deux autres traductions de cette phrase confirment et développent cette interprétation : « Si je faisais une proposition quelconque, je commettrais une erreur logique » (Streng). « Si j'avais avancé une thèse, l'erreur logique aurait été de mon fait » (Gradinarov).

est faux . Par conséquent, Y implique formellement la négation de X, c'est-à-dire non X ! Dès lors, la phrase de Nagarjuna se lit désormais : « **Si X, alors non X** », c'est-à-dire « Si X est vrai, alors X est faux » – une proposition hypothétique paradoxale, dont la conclusion serait l'affirmation catégorique « X est faux » (comme suggéré précédemment).

Cependant, la question ne s'arrête pas là. Si l'on considère la signification de X – à savoir « une proposition est proposée » – on peut raisonnablement supposer qu'il s'agit de *n'importe quelle proposition* . Dans ce cas, la proposition en question pourrait même être la négation de X ; on peut donc substituer nonX à X dans toute l'hypothèse. On obtient ainsi « Si nonX, alors non nonX », c'est-à-dire « **Si non X, alors X** », ou en d'autres termes « Si X est faux, alors X est vrai ». Il s'agit là aussi, bien sûr, d'une proposition paradoxale, dont la conclusion formelle est « X est vrai ».

Ainsi, par une interprétation universelle de « posséder une thèse » incluant « ne pas en posséder », nous nous trouvons non pas face à un simple paradoxe, mais à un *double paradoxe* ! Autrement dit, notre conclusion n’est pas seulement que X est faux, mais que X est à la fois vrai et faux. Cette dernière conclusion est bien sûr contraire au principe de non-contradiction, comme dans le cas du paradoxe du menteur.

Cela signifie que l'affirmation de Nagarjuna est formellement impossible : c'est une contradiction dans les termes ; elle est non seulement fausse, mais aussi dénuée de sens. Elle ne constitue en rien un discours légitime, et encore moins une position ou une théorie philosophique tenable. Les mots ou symboles employés sont logiquement inconcevables, si bien qu'il ne dit rien. Il semble dire quelque chose d'intelligible, mais ce n'est qu'une illusion.

On pourrait objecter que Y ne signifie pas nécessairement que X *est* faux, mais simplement que X *pourrait être* faux. Autrement dit, « commettre une erreur » pourrait signifier que X est incertain plutôt que définitivement réfuté. Dans ce cas, nous aurions les deux hypothèses suivantes : « Si X est vrai, peut-être pas X » et « Si non X est vrai, peut-être X » ; ou, en résumé : « Que X soit proposé ou non, le résultat est incertain ». De fait, cette posture plus modale et ambiguë pourrait bien être considérée comme l'intention exacte de Nagarjuna (que certains ont interprétée comme une « illocution » non engageante).

À première vue, l'emploi de termes vagues ou de symboles peut rendre cette objection crédible et faire disparaître les conclusions contradictoires. Cependant, à y regarder de plus près, un conflit sous-jacent persiste : affirmer X ou le nier est contraire à une position qui ne prend pas position sur X. Une affirmation affirmative (affirmer ou nier X) est *incompatible avec* une affirmation problématique (dire que X peut être vrai ou faux). On ne peut prétendre simultanément *connaître* X et ne pas *le connaître* (la question de la vérité ou de la fausseté restant ouverte). C'est une contradiction aussi flagrante que de déclarer la même chose (X) à la fois vraie et fausse.

Une personne non familiarisée avec la logique des propositions hypothétiques pourrait objecter que X, ou non-X, n'est proposé qu'hypothétiquement dans l'antécédent, et pourrait donc poser problème dans le conséquent. Or, cette objection est logiquement infondée, du fait du processus d'addition (décrit dans le chapitre sur la logique formelle) ; autrement dit, du fait que « Si X, alors Y » implique « Si X, alors (X et Y) ». En l'espèce, cela signifie : « Si X est affirmé, alors X est à la fois affirmé et incertain ». Il suffit que la contradiction survienne conditionnellement, comme ici, pour que la condition soit réfutée ; par conséquent, notre conclusion

est parfaitement formelle : « X ne peut être affirmé ». CQFD.

On pourrait ici, enfin, objecter que la certitude de l'antécédent et l'incertitude du conséquent ne sont pas nécessairement simultanées, et ne produisent donc pas de conflit logique. Cette objection serait valable, admettant qu'un processus de pensée sépare le début et la fin de la proposition hypothétique. Cependant, dans le cas examiné, Nagarjuna affirme clairement que l'acte même de « proposer quelque chose » constitue une « erreur » ; autrement dit, l'erreur n'est autre que la proposition elle-même. Dès lors, aucune séparation temporelle ne peut être raisonnablement défendue, et la thèse de Nagarjuna demeure illogique.

Il convient de noter que la discussion précédente n'a porté que sur la première partie du verset 29, c'est-à-dire la prémisse majeure : « Si j'avais une thèse, je serais en faute. » Nous avons constaté que cette proposition hypothétique est logiquement erronée, que Nagarjuna admette ou non avoir une thèse. Revenons donc sur cette prémisse mineure et sur sa conclusion selon laquelle il n'est pas en faute.

Nous avons ici introduit un nouvel élément dans l'analyse, en constatant que « Si X, alors Y » (compris comme « Si X, alors non X ») implique « Si non X, alors Y » (puisque ce dernier est impliqué par « Si non X, alors X », lui-même impliqué par le premier en remplaçant X par non X). Ainsi, nous avons maintenant une nouvelle prémisse majeure pour Nagarjuna, à savoir « Si non X, alors Y », signifiant : « Si je n'ai pas de thèse, je serai en faute ».

Partant de cette prémisse majeure implicite et de la prémisse mineure de Nagarjuna, à savoir « Je n'ai pas de thèse », la conclusion est « Je suis en faute ». Cette conclusion est, notez-le, l'inverse de la sienne (« Je ne suis

pas en faute »). Ainsi, même si Nagarjuna se vante d'avoir une pensée irréprochable, elle est manifestement erronée !

En clair, abstraction faite de toute sa rhétorique, il affirme simplement : « aucune thèse n'est vraie », ce qui n'est qu'une autre version du paradoxe du menteur. Sa tentative d'atténuer son propos, par la précision « sauf ma thèse », n'est en réalité qu'une affirmation supplémentaire : un cas particulier qui relève sans équivoque de la règle générale. De plus, avant qu'une exception puisse être appliquée, la règle elle-même doit pouvoir être formulée de manière cohérente, ce qui n'est manifestement pas le cas ici (comme nous venons de le démontrer).

Enfin, il convient de noter que cette réfutation n'implique en rien que le silence soit impossible ou dénué de valeur. Si, comme le prétendent certains commentateurs, Nagarjuna souhaitait promouvoir la cessation du discours, il s'y est assurément pris de la mauvaise manière. Il n'avait pas besoin d'élaborer une philosophie controversée et antilogique. Il lui aurait suffi de poser, comme un fait psychologique, que le silence (intérieur et extérieur) est propice à la méditation profonde.

2. Appeler un chat un chat

Le bouddhisme, sans aucun doute depuis ses origines, mêle logique et illogisme dans son discours. Parmi ses quatre principales écoles philosophiques – Abhidharma, Prajnaparamita, Madhyamika et Yogacara –, Madhyamika est celle qui tend le plus à rejeter les trois lois de la pensée (identité, non-contradiction et exclusion du juste milieu) [104]. Toutefois, cette tendance était déjà présente dans

---

104 Voir mon ouvrage **L'Illogisme Bouddhiste** sur ce sujet, ainsi que mes commentaires sur le discours de Nagarjuna dans mes **Réflexions **, partie I, chapitre 5. Je tiens à souligner que mon propos, dans les critiques précédentes comme dans la présente, n'est pas de rejeter le bouddhisme en tant que tel, mais de montrer qu'il

l'école Prajnaparamita, comme en témoignent des exemples du *Sūtra du Diamant* [105].

Dans ce sutra, nous trouvons des exemples d'arguments logiques valides. Par exemple, on trouve un argument a fortiori bien formé dans la section 12[106] : « Partout où ce sutra, ou même quatre de ses versets, sont prêchés, ce lieu sera respecté de tous les êtres… Combien plus digne de respect est celui qui peut mémoriser et réciter ce sutra… ! » Mais nous y trouvons aussi des antinomies flagrantes, comme « le Dharma… n'est ni saisissable ni insaisissable » (affirmé alors même que « insaisissable » signifie insaisissable, et « non insaisissable » signifie saisissable).

Mais le *Sūtra du Diamant* recourt à plusieurs reprises à un raisonnement que, en tant que logicien, je qualifierais de nouvelle aberration de l'illogisme bouddhiste. Ce raisonnement affirme : « **Ce que l'on appelle X n'est pas X en réalité ; par conséquent, on l'appelle X** » (ou parfois : « **Ce que l'on appelle X n'est pas X ; ce n'est qu'un nom, c'est pourquoi on l'appelle X** »).

---

peut s'harmoniser avec la raison. Je considère comme tout à fait inutile et contre-productive l'attitude de nombreux philosophes bouddhistes qui semblent considérer la réalisation (c'est-à-dire l'éveil, la libération, la sagesse) comme impossible sans rejeter la logique. Mon principe directeur est que ces deux dimensions sont parfaitement compatibles, et que la raison est même un moyen essentiel (avec la morale et la méditation) pour atteindre cette fin souhaitable.

105 À en juger par sa langue sanskrite, la place centrale qu'elle occupe dans l'idéal du bodhisattva et d'autres éléments importants, ce sutra est un texte mahayana. On estime qu'il a été composé et écrit en Inde vers 350 de notre ère, bien qu'au moins une autorité suggère une date aussi ancienne que 150 de notre ère. À titre de comparaison, Nagarjuna, le fondateur de la philosophie madhyamika, était actif entre 150 et 200 de notre ère ; par conséquent, ce texte de la Prajnaparamita a été écrit à peu près durant la même période, voire bien plus tard.

106 Mu Soeng, p. 111.

On trouve plus de vingt exemples de cet argument dans le sutra en question. En voici un : « Ce que le Tathagata a appelé la Prajnaparamita, la sagesse suprême et transcendante, n'est en réalité pas la Prajnaparamita, et c'est pourquoi on l'appelle Prajnaparamita. » En voici un autre : « … ceux que l'on appelle êtres ne sont en réalité pas des êtres. Ce n'est qu'un nom. C'est pourquoi le Tathagata en a parlé comme d'êtres. »[107]

Ce que je conteste ici, concernant ce type de discours, c'est uniquement la conjonction « *donc* » ou « *c'est pourquoi* » [108]. Je ne nie pas qu'on puisse désigner quelque chose par un nom inapproprié, ni même que les mots ne puissent jamais exprimer qu'une approximation de ce que l'on veut vraiment dire. Mais affirmer que l'on nomme quelque chose X *parce* qu'il n'est pas X est assurément absurde et intenable.

Il ne s'agit pas simplement de « ne pas appeler un chat un chat », mais d'appeler un chat un chat tout en croyant qu'il n'en est rien ! C'est, du moins en apparence, contraire à toute logique. Si l'étiquette n'est pas appropriée, pourquoi l'utiliser ? De plus, pourquoi se vanter de cette inversion absurde en disant « donc » ?

Affirmer que quelque chose « n'est pas en fait ou véritablement X » sous-entend que le mot X a un sens auquel la chose en question ne correspond pas ; dès lors,

---

107 Dans Mu Soeng : pp. 145 et 151 respectivement. J'ai repéré un argument similaire dans un autre texte Mahayana : « Et c'est parce que pour eux [les bodhisattvas] l'entraînement consiste à ne pas s'entraîner qu'on dit qu'ils s'entraînent » (ma traduction d'une traduction française) – que l'on trouve au chapitre 2, v. 33 du « Sūtra des paroles du Bouddha sur la Sagesse suprême » (voir Eracle, p. 61).

108 En supposant, bien sûr, que la traduction de cette édition soit correcte (et elle semble tout à fait respectable ; voir p. ix de la préface). Mon propos est qu'aucun logicien n'a jamais formellement validé un tel argument ; et, de fait, il est formellement invalide, puisque la conclusion contredit une prémisse.

pourquoi appeler cette chose « X », contre toute logique ? Pourquoi ne pas simplement l'appeler « non X » (ou lui trouver un autre nom plus précis) et éviter ainsi le paradoxe ?

Un discours comme « ce n'est qu'un nom » est de toute façon contre-productif, puisqu'il utilise précisément des noms porteurs de sens. Cette phrase sous-entend qu'aucun mot n'a de référence valable, tout en s'appuyant sur l'efficacité des mots employés pour communiquer ses diverses intentions. C'est une affirmation qui tente de s'exonérer des critiques qu'elle adresse à *toutes* les affirmations de ce type.

Dans les exemples précédents, l'argument repose sur notre compréhension de termes tels que « Prajnaparamita » (c'est-à-dire la perfection de la sagesse) ou « êtres », tout en cherchant à invalider une telle compréhension. On ne peut donc affirmer qu'il communique quoi que ce soit d'intelligible.

Il est indéniable que nous ne pouvons exprimer pleinement la réalité ultime (ou Dieu) par les mots. Mais il n'en demeure pas moins vrai que nous pouvons exprimer verbalement le fait de l'ineffabilité (comme cela vient d'être fait dans la phrase précédente). Il n'est pas nécessaire de dévaloriser les mots en tant que tels pour admettre leurs limites.

De plus, il est fort douteux que de telles affirmations paradoxales (comme « nommez ceci X *car* ce n’est pas X ») soient psychologiquement efficaces pour atteindre l’éveil ; elles ne font que semer la confusion et paralyser l’esprit rationnel. Au lieu de faire taire la curiosité intellectuelle, elles ne font en réalité que l’exciter par des questions inconscientes restées sans réponse. Ces affirmations absurdes sont le fruit d’une mode regrettable

qui s'est développée au sein du bouddhisme à une certaine époque [109].

Ce genre de perversité intellectuelle a fini par paraître profond, comme c'est encore le cas pour certains penseurs postmodernes occidentaux aujourd'hui, précisément parce qu'une antinomie logique n'implique rien – et que ce vide de sens est (à tort) assimilé au Vide sous-jacent à tous les phénomènes. Le gouffre béant que représente l'antinomie pour la connaissance donne l'illusion d'être riche de sens, alors qu'en réalité, il n'est que la preuve de l'ignorance. Il est important d'y prêter attention.

Il convient d'ajouter qu'il existe bel et bien une sorte de paradoxe structurel dans l'acte méditatif – mais il ne s'agit pas de l'habitude du *Sūtra du Diamant* de « ne pas appeler un chat un chat ». Le paradoxe réside dans le fait que si nous poursuivons l'éveil par la méditation, nous ne pouvons espérer l'atteindre, car alors notre ego (qui s'accroche à cette valeur transcendantale comme à un objet matériel) se maintient ; or, la méditation est la meilleure voie vers l'éveil. Aussi, il nous faut-il simplement nous y mettre – nous asseoir et laisser notre éveil inné (notre « nature de Bouddha ») se manifester peu à peu.

Il convient également de rappeler que le bouddhisme est originellement motivé par un réalisme profond. Il s'agit essentiellement d'une quête de la Réalité. Dans cette perspective, la notion bouddhiste de « tel-tel » peut être considérée comme un engagement envers la Loi d'Identité. L'homme éveillé est celui qui perçoit les

---

[109] Bien qu'elles ne soient pas totalement absentes de la littérature Abhidharma antérieure et de la littérature Yogacara postérieure, elles ne sont pas rares dans certaines littératures Prajnaparamita (y compris le *Sutra du Diamant*) et sont plutôt courantes dans la littérature Madhyamika.

choses, en particulier et en général, *telles qu'elles sont réellement* .

Ceci est illustré, par exemple, dans l'échange zen suivant. Un moine demanda à Li-shan : « Quelle est la raison [de la venue de Bodhidharma de l'Occident, c'est-à-dire de l'Inde en Chine] ? », ce à quoi le maître zen répondit : « Simplement parce que les choses sont telles qu'elles sont », et dans le commentaire de D.T. Suzuki, cela fait référence à la « Telité » [110].

110 *La doctrine zen du non-esprit* , p. 93.

Références principales

Cheng, Hsueh-li. *Logique vide : le bouddhisme Madhyamika à partir de sources chinoises* . Delhi : Motilal Banarsidass, 1991.

Dawkins, Richard. *Le plus grand spectacle du monde : les preuves de l'évolution.* New York : Free Press, 2009.

Le Dhammapada. Trad . Juan Mascaro. Harmondsworth, Mx. : Penguin, 1973.

Éracle, Jean. *Enseignements du Bouddha* . Paris : Librio, 2005.

Kneale, William et Martha. *Le développement de la logique* . Oxford, Londres : Clarendon, 1962.

Maggiori, Robert. *La philosophie au jour le jour.* Paris : Flammarion, 1994.

Mu Soeng. *Le Sutra du Diamant : transformer notre façon de percevoir le monde* . Boston : Wisdom, 2000.

Rescher, Nicholas. *Les paradoxes : leurs racines, leur portée et leur résolution.*

Suzuki, DT. *La doctrine zen du non-esprit.* Éd. Christmas Humphreys. York Beach, ME : Weiser, 1972.

Trungpa, Chögyam. *Le jeu des illusions.* Shambhala : Boston, 1994.

Œuvres d'Avi Sion

Avi Sion est l'auteur de plusieurs ouvrages sur la logique et la philosophie, notamment : *Future Logic* (1990) ; *Judaic Logic* (1995) ; *Buddhist Illogic* (2002) ; *Phenomenology* (2003) ; *The Logic of Causation* (1999, 2003, 2010) ; *Volition and Allied Causal Concepts* (2004) ; *Ruminations* (2005) ; *Meditations* (2006) ; *Logical and Spiritual Reflections* (2008-9) ; *A Fortiori Logic* (2013).

Ces ouvrages peuvent être consultés gratuitement sur les sites suivants : www.TheLogician.net et www.AviSionBooks.com.

Les paradoxes et leurs résolutions peuvent être consultés en ligne à l'adresse suivante : www.avisionparadoxes.bravesites.com/

La liste complète des sites web d'Avi Sion est consultable à l'adresse suivante : www.avisionlinks.bravesites.com.

Pour toute question ou commentaire, vous pouvez contacter Avi Sion à l'adresse suivante : avi-sion@thelogician.net

Références principales