

LES LIMITES DU SYLLOGISME PRATIQUE EXAMEN DE DEUX CAS DIFFICILES

Introduction

Pourquoi s'intéresser, en philosophie pratique, à des « cas difficiles » (*hard cases*) ? À elle seule, cette question demanderait un exposé entier. Je me contenterai ici de partir d'un exemple, qui n'est d'ailleurs pas sans rapport avec l'une des études de cas dont je vais traiter.

Dans ma critique de la *Théorie de la justice* de John Rawls, j'ai naguère eu recours à des exemples comme le « choix de Sophie » (William Styron) ou le « choix de Jim » (Bernard Williams) – faut-il livrer à la mort, ou même exécuter, un innocent pour éviter que deux innocents ou plus ne périssent ? –, lesquels sont habituellement mobilisés pour rejeter l'utilitarisme au motif qu'il justifierait le sacrifice de la vie, des libertés ou du bien-être de certains pour le plus grand bien de tous. J'ai tenté de montrer que la *lettre* des principes rawlsiens de justice ne leur permettait pas d'échapper à cela même que la *Théorie* reproche à l'utilitarisme. Eux aussi, dans des cas semblables, justifient la solution sacrificielle. Il m'a alors été répliqué que la *Théorie* n'avait pas été conçue pour s'appliquer à des cas de ce genre. De fait, elle est censée s'appliquer, en une circularité que Rawls revendique aujourd'hui explicitement, à une société « bien ordonnée », c'est-à-dire déjà en gros gouvernée par les principes rawlsiens de justice. J'ai alors souligné que la méthode que Rawls emploie pour comparer le principe d'utilité et ses propres principes de justice est extrêmement injuste (*unfair*). L'utilitarisme est accusé de favoriser le sacrifice dans des contextes que la « justice comme équité » exclut de son propre champ – alors même que, si elle devait y être appliquée, ses principes justifieraient le sacrifice tout autant que l'utilitarisme –, tout cela au motif que l'utilitarisme,

contrairement à la justice comme équité, se donne comme une doctrine morale globale, d'application universelle¹.

De quel droit peut-on exclure du champ d'une théorie des cas difficiles comme ceux-ci ? Au motif qu'ils seraient exceptionnels, extra-ordinaires ? On peut arguer au contraire que les situations sacrificielles en cause sont non seulement « ordinaires » – non pas au sens qu'elles seraient fréquentes, mais au sens qu'elles font intégralement partie de la réflexion morale, sans doute depuis que celle-ci existe –, mais de plus qu'elles se situent au fondement même de tout ordre culturel, social et politique. Cas difficile ne veut donc pas nécessairement dire cas extraordinaire. Faut-il alors les exclure au motif que la philosophie pratique doit se situer au plus près d'une philosophie du sens commun ? Mais qui nous dit que le sens commun ne s'est pas forgé à l'épreuve de cas ou dilemmes de ce genre ? Quelle démission pour la pensée, de toute façon, que de fuir devant la difficulté ! Un élément d'ordre esthétique se glisse peut-être ici. En ce qui me concerne, la philosophie morale me ferait bayer aux corneilles si elle se contentait de délivrer des conseils de sagesse « ordinaire ». Je ne sache pas d'ailleurs qu'elle ait jamais aidé quiconque à mener une vie sage. Si elle m'intéresse, c'est par les vérités d'ordre anthropologique qu'elle dévoile, souvent sans s'en apercevoir, lorsqu'elle consent, précisément, à se confronter à ces cas difficiles.

Je voudrais mettre ici à l'épreuve de deux cas difficiles l'assimilation reçue entre délibération morale ou prudentielle et mise en œuvre du syllogisme pratique (le schéma croyances – désirs/décision/action). Ce syllogisme implique que les croyances et les désirs sont déterminés antérieurement à, et indépendamment de, la décision qui va être prise à la suite d'une délibération mobilisant ces mêmes croyances et désirs. Sur l'indépendance et l'antériorité des désirs, beaucoup de choses ont déjà été écrites, qui amènent à douter de la validité universelle du schéma. Les croyances semblent plus résistantes. À propos d'un thème crucial pour la philosophie morale, à savoir le sacrifice de l'innocent, la première étude de cas remet pourtant en cause l'indépendance et l'antériorité des croyances par rapport aux décisions. Il apparaît, en conséquence, que les croyances ne sont pas (toujours) des raisons. Sur un sujet d'importance majeure pour l'avenir de l'humanité, la seconde étude de cas poursuit cette mise en question du syllogisme pratique. Elle montre

1. Sur cette controverse, on peut consulter Dupuy 1992b.

que l'«équilibre de la terreur» ne peut se justifier au regard d'un critère de prudence qu'à condition de renoncer à l'un des présupposés métaphysiques du syllogisme pratique, à savoir l'«ouverture» de l'avenir – en termes techniques, sa dépendance contrefactuelle par rapport à l'action présente.

L'universalisabilité et l'objectivité du jugement moral : le cas *Billy Budd*

Les philosophes de la morale, au moins depuis Kant, nous mettent en garde contre la tentation qui nous est trop familière de faire, en matière de jugement moral, une exception en ce qui nous concerne. Pour nombre de doctrines morales, le jugement moral est soumis à la règle d'universalisabilité. Le point de vue de l'agent en tant que tel ne doit bénéficier d'aucun privilège. Lorsqu'un sujet porte un jugement à propos d'un acte qu'il envisage d'accomplir, le fait qu'il soit le sujet de cet acte ne devrait en rien intervenir dans ce qui peut être dit à propos du caractère correct ou incorrect de ce jugement. La version que donne Henry Sidgwick du principe d'universalisabilité s'exprime ainsi :

Nous ne pouvons pas juger qu'une même action est bonne pour le sujet *A* et mauvaise pour le sujet *B* s'il n'y a pas dans leurs natures ou dans les circonstances dans lesquelles ils se trouvent des différences que nous puissions tenir pour justifiant de façon raisonnable que leurs devoirs soient différents. Si, en conséquence, je juge qu'une certaine action est bonne pour moi, implicitement je juge qu'elle est bonne pour toute autre personne dont la nature ou les circonstances ne diffèrent pas des miennes significativement. S'il est bien (resp. mal) pour moi de faire une certaine chose et s'il n'est pas bien (resp. mal) pour quelqu'un d'autre de faire la même chose, ce ne peut être que sur la base d'une différence entre les deux cas autre que le fait que cette personne et moi soyons distinctes².

Récemment, plusieurs philosophes britanniques ont défendu la thèse que l'on ne pouvait aussi facilement évacuer ce qu'ils appellent «the primacy of the agent's perspective» ou encore «the primacy of first-person practical judgements». Peter Winch a lancé le débat avec son article «The Universalizability of Moral Judgments»³. Cette

2. Sidgwick 1907, 384-385.

3. Winch 1972.

publication a suscité de très nombreuses réactions, parmi lesquelles je retiendrai les articles de David Wiggins⁴ et de Joseph Raz⁵. Winch remarque un glissement significatif dans la citation de Sidgwick. La première phrase énonce un principe avec lequel il dit se trouver en plein accord. Il n'y aurait pas de jugement moral s'il ne respectait pas ces conditions minimales de cohérence, d'intelligibilité et de rationalité. Cependant, il faut noter que la première phrase concerne des jugements faits par la même personne à propos d'actes par rapport auxquels elle est en situation de spectateur. La conclusion, elle, fait disparaître l'identité de celui qui porte le jugement. Ce saut, selon Winch, n'est pas légitime.

Winch défend la thèse que dans certains « cas difficiles », en particulier celui des dilemmes moraux, un jugement fait à la première personne (le spectateur qui porte le jugement et l'agent sont une seule et même personne) peut échapper à la législation du principe d'universalisabilité. Wiggins et Raz, qui acceptent dans une large mesure la thèse de Winch, ont examiné l'étendue des dommages qu'elle fait subir à ces positions méta-éthiques que sont le cognitivisme, l'objectivisme et le réalisme moral.

Winch a recours à un exemple littéraire, le roman de Herman Melville, *Billy Budd*⁶. Le contenu précis de l'histoire n'est peut-être pas d'une importance majeure, et tout autre dilemme moral aurait sans doute convenu. Je tiens cependant que ce n'est pas le fruit du hasard si, une fois de plus, après – ou plutôt avant – le choix de Sophie, le choix de Jim et bien d'autres cas que la littérature, le mythe ou la philosophie nous présentent, il s'agit de savoir s'il est bien de mettre l'innocent à mort. L'histoire se passe sur la frégate britannique *L'Indomptable*, au temps des guerres napoléoniennes, dans une période de grande tension pour la marine anglaise, suite à la fameuse mutinerie de la Nore. Le gabier de misaine Billy Budd est un marin d'une beauté exceptionnelle, au caractère angélique. Ce serait à vrai dire un être parfait, s'il n'était affligé d'une tendance à bégayer (« le bégaiement dans la parole divine » dont parle E.M. Forster dans le livret qu'il écrivit pour l'opéra du même nom de Benjamin Britten⁷) qui s'accroît dans les moments de forte émotion. Le maître d'armes Claggart hait Billy d'une haine faite d'envie, de jalousie, de désir inassouvi. Claggart fait tomber

4. Wiggins 1991.

5. Raz 1999, chap. 10, « The Truth in Particularism ».

6. Melville 1924.

7. Britten 1994.

Billy dans un piège et l'accuse devant le commandant Vere, faux témoignage à l'appui, de fomenter une mutinerie. Fou de rage et d'incompréhension, Billy se montre incapable de prononcer le moindre mot pour sa défense et frappe Claggart, qui tombe raide mort. Vere, après délibération avec ses officiers supérieurs, condamne Billy à mort et le fait exécuter sur le champ.

Selon nos trois auteurs (qui prennent quelques libertés avec le texte de Melville, mais laissons ici cette question de côté), le choix du commandant Vere serait pris en tenaille entre deux systèmes de raisons morales incommensurables et irréductibles. D'un côté, l'attachement à la loi martiale et au *Mutiny Act*, lui-même soutenu par un raisonnement moral de type conséquentialiste (il faut punir Billy Budd pour l'exemple, pour éviter la panique à bord). De l'autre, la loi de nature qui interdit absolument de mettre à mort un innocent.

Selon Winch, l'histoire montre que Vere agit bien en condamnant Billy à mort. Et cependant, ajoute Winch, si moi, Peter Winch, je m'étais trouvé à la place de Vere, exactement dans les mêmes circonstances, il aurait été mal pour moi de faire la même chose. En fait, condamner Billy à mort et le faire exécuter est quelque chose que je n'aurais tout simplement pas *pu* faire – et ce, non pas par manque de courage ou faiblesse de la volonté, mais bien par pure *impossibilité morale*.

Il existerait donc des situations dans lesquelles l'on pourrait tenir simultanément pour moralement fondés les deux jugements pratiques suivants, jugements, répétons-le, qu'il est impossible de séparer de ceux qui les portent :

- a) dans les circonstances *C* l'agent *A* décide qu'il est bien (*right*) pour lui de faire *Y*,
- b) dans les mêmes circonstances *C* l'agent *B* décide qu'il est mal (*wrong*) pour lui de faire *Y*,

alors que *A* et *B* ont exactement les mêmes caractéristiques parmi celles que la délibération peut légitimement tenir pour fixes⁸ : les capacités, les obligations, les engagements, etc.

Nombre de solutions tentantes à cet apparent paradoxe sont à écarter. La première consiste à donner de l'expression « Est bien pour l'agent » une interprétation purement subjectiviste : serait « bien

8. Wiggins 1991, 181, n. 43 : « deliberation can treat as fixed ». Je rassemble ici dans une même formule les démarches de Winch 1972 et de Wiggins 1991.

pour lui » ce qu'il croirait, penserait être bien pour lui. À l'opposé, la deuxième sauverait trop facilement le réalisme et l'objectivisme en affirmant que le caractère des agents *A* et *B* fait partie des caractéristiques qui interviennent légitimement dans la délibération ; or jamais deux agents n'auront exactement les mêmes traits de caractère⁹.

Winch, Wiggins et Raz essaient de dire quelque chose de beaucoup plus subtil et ambitieux. Parmi les considérations qui interviennent dans le jugement pratique, on trouve un facteur que nos auteurs ont du mal à nommer ou à décrire précisément – appelons-le le facteur *x* –, mais dont ils dessinent assez bien les caractéristiques formelles. Le facteur *x* tient du caractère de l'agent, de sa personnalité, de ses dispositions et sensibilités morales, mais seulement en tant que ces éléments ne constituent pas des données déterminées une fois pour toutes préalablement à l'action, mais s'actualisent en permanence dans la série des actions qui composent une vie¹⁰. Le facteur *x* est relatif à ce que l'agent découvre au sujet de lui-même en décidant et en agissant, mais en tant que cette découverte n'est pas celle d'une essence dont ses décisions et ses actions seraient le reflet, mais constitue une *self-discovery* et, en fait, une *self-determination*, une découverte et une création de soi simultanées à la prise de décision et à l'action¹¹ – un sartrien dirait : en tant que l'existence précède l'essence. Le facteur *x* ne peut faire partie des raisons qui conduisent l'agent à agir dans un certain sens et cependant

9. C'est l'interprétation que donne Samuel Guttenplan de Winch, citée par Wiggins 1991, 179-180.

10. Cf. Raz 1999, 245 : « The question is not, given my moral character what shall I do? To put it thus is to foreclose the possibility of a change in one's personal perspective, and to deny the self-determining, the self-creating aspect of decision and action ».

11. Winch 1972, 245 : « [Vere] is not merely concerned to decide to *do* something, but also to *find out* what is the right thing for him to do [...]. The deciding what to do is, in a situation like this, itself a sort of finding out what is the right thing to do... deciding is an integral part of what we call "finding out what I ought to do" [...] it seems [...] that what one finds out is something about oneself, rather than anything one can speak of as holding universally [...]. What a man finds out about himself is something that can be expressed only in terms of the moral ideas by consideration of which he arrives at his decision ». Raz 1999, 246 : « [The question what should *I* do? What is the right thing *for me* to do?] allows a continuous process of self-determination, for it leaves open the possibility, however unlikely, that my response will surprise me, that it will not confirm my own previously formed idea of myself [...] decision is part of discovery in the sense that it is open to development, and is part of the process in which we are part authors of our own character ».

il joue un rôle crucial dans la détermination de ce qui fait qu'il est bien pour l'agent d'agir comme il le fait¹².

Une chose est certaine : ce facteur x est corrélé à la décision au point qu'il en hérite une propriété formelle essentielle : puisque l'agent est libre, sa décision n'est pas déterminée avant qu'il la prenne. Il en va de même pour le facteur x : il n'est pas déterminé avant que l'agent prenne sa décision. Et pourtant, il fait partie des éléments qui font de la décision la décision qu'il est bien pour l'agent de prendre. C'est là le paradoxe autour duquel nos auteurs tournent sans vraiment jamais arriver à le présenter ainsi. La décision est ce qui doit être expliqué, justifié et cependant cette même décision fait partie, par l'intermédiaire du facteur x qu'elle contribue à déterminer, des éléments qui expliquent ou justifient la décision. C'est comme si la décision s'expliquait ou se justifiait elle-même.

Cette structure formelle est certes inhabituelle, mais ses propriétés sont parfaitement cernables. C'est une structure de « point fixe » ou d'« équilibre réfléchi », bien en phase avec la démarche cohérentiste en philosophie morale. Dire que la décision fait partie des éléments qui justifient la décision ne veut pas dire que n'importe quelle décision fait l'affaire et qu'il aurait été tout aussi bien *pour Vere* de ne pas condamner Billy à mort que de le condamner¹³. Mais une telle structure peut admettre plusieurs points fixes et la manière (souvent contingente) d'entrer dans le système détermine le point fixe qui finalement prévaudra. C'est dire que le jugement pratique est dépendant par rapport au chemin qui y mène (notion de *path-dependency*, dont l'équivalent en philosophie pratique est le concept de procédure pure). La contrainte rappelée par Wiggins, à savoir que la justesse d'un jugement pratique est indépendante du fait que son auteur parvient ou non à ce jugement¹⁴, est donc invalidée par ce type de structure à point fixe. Il en résulte que c'est en situation et dans la dynamique même de la délibération pratique que l'on

12. Raz 1999, 241 : « It was not that Vere's character or personality, or moral sensibilities, or anything like that figures among his reasons. At least it does not figure as such, under that description. It was merely that it is part of what makes the decision right for him ». Et aussi : « How can the decision be right, or even right for me, because of something about me, when my condition was not part of my reasons » (*ibid.*, 240).

13. Cf. Wiggins 1991, 182 : « The position may seem self-contradictory. But it is not. There is no question of one agent's coming to both conclusions, and our endorsing each conclusion ».

14. *Ibid.*, 171 : « the correctness of a judgment should be independent of its author's actually arriving at it ». Cf. aussi la citation de Nagel à la même page, n. 29.

peut arriver à un jugement justifié ; et que donc le même agent, ou bien un autre agent aux caractéristiques identiques, placé dans les mêmes circonstances mais entrant dans la délibération par un autre bout, eût pu ou pourrait arriver à un jugement différent¹⁵.

On comprend pourquoi la coexistence de deux jugements opposés du type (a) et (b) ne menace pas vraiment le caractère objectif et universel du jugement pratique. Est objective et universelle la structure de point fixe avec l'ensemble de ses attracteurs (ou « points fixes » ou « équilibres réfléchis »). Relève simplement de la subjectivité, ou mieux de la dépendance du jugement par rapport à l'agent¹⁶, la dynamique singulière, avec son point d'entrée contingent, de la délibération. Ce que nous avons appelé le facteur x (le caractère ou les dispositions morales de l'agent) intervient dans ce qui fait que telle décision est celle qu'il est bien pour l'agent de prendre (une autre valeur de x , pour un autre agent aux caractéristiques pertinentes identiques ou pour le même agent, mais entrant dans la délibération autrement, expliquerait qu'une autre décision apparaisse comme celle qu'il serait bien de prendre). Et cependant, x ne peut constituer une raison pour l'agent, puisque x n'est pas déterminé indépendamment de et antérieurement à la détermination de la décision.

Nos auteurs nous disent deux choses qui peuvent paraître contradictoires. D'une part, dans des dilemmes moraux de ce type, la décision *exprime* le caractère, la personnalité de l'agent, etc., tous éléments qui ne font pas partie des *raisons* qu'a l'agent d'agir, mais qui le poussent à agir comme il le fait. Ainsi Raz :

Dans de tels cas, il est bien pour l'agent d'agir comme son caractère moral lui dicte d'agir. Mais la raison qu'il a d'agir comme il le fait n'est pas que cela correspond à ses dispositions ou que cela

15. Cf. Raz 1999, 245 : « In that sense but for the fact that Vere has already decided, Winch could not have known what was right for him. The answer must come from the person whose decision it is. Only they, by going through the process of confronting the issue in life or in their imagination, can discover/determine what is right for them ». Wiggins 1991 ne me semble pas à la hauteur du défi lorsqu'il écrit : « [...] it seems that what we the non-agents have to ask may be different from what the agent Captain Vere has to ask. By Winch's account, one part of what it may turn on for us is Vere's decision. But that cannot be anything it turns on *for Vere!* » (*ibid.*, 170). Pourquoi pas ? Parce que la décision n'est pas encore prise ? Wiggins ne voit pas ici que la délibération implique la résolution d'un problème de point fixe.

16. Cf. Wiggins 1991, 170 : « agent-oriented view of practical judgment ».

est plus cohérent avec ses décisions passées. C'est qu'il ne peut pas agir autrement¹⁷.

D'autre part, il paraît clair que l'agent délibère, donc qu'il fait varier dans son esprit sa décision, la traitant comme un paramètre, en appréciant les conséquences mais aussi en inférant quelque chose au sujet de son *x*, *i.e.* qui il est, son caractère moral, sa personnalité, etc.

Non seulement ce schéma de raisonnement n'est pas contradictoire, il révèle l'existence d'un schéma de raison pratique fort différent du syllogisme habituel, et donc il révèle que notre raison pratique n'est pas une, mais double. Encore faut-il s'en apercevoir. Cette dualité repose en dernière analyse sur les fictions que nous inventons pour nous penser libres dans un monde déterministe. Il se trouve que nous ne disposons pas d'une seule telle fiction, mais (d'au moins) deux.

D'un côté (le côté dominant), nous avons la tradition d'Aristote, de Leibniz, de Kant, etc. Christine Korsgaard la caractérise ainsi :

Afin de pouvoir faire quoi que ce soit, il nous faut tout simplement faire semblant d'ignorer [*ignore*] le fait que nous sommes déterminés, et décider ce que nous devons faire – exactement comme si [*als ob*] nous étions libres.

Selon cette fiction, nous sommes capables d'agir dans l'exacte mesure où nous sommes capables de démarrer, par l'effet de notre volonté, de nouvelles chaînes causales. Faire comme si nous étions libres nous amène à considérer des contrefactuels du type : « Si j'agissais autrement que je le fais, alors telles conséquences s'en suivraient ».

Cependant, il existe une autre fiction qu'on peut rattacher aux philosophes suivants : Spinoza, pour qui être libre c'est acquiescer à la nécessité ; Bergson, pour qui les possibles ne préexistent pas au choix : c'est le choix qui les crée rétroactivement ; Heidegger et sa notion paradoxale d'*Entschlossenheit*, c'est-à-dire la faculté que nous avons de « choisir notre destin » ; et bien sûr Sartre et sa reprise de cette notion sous la forme :

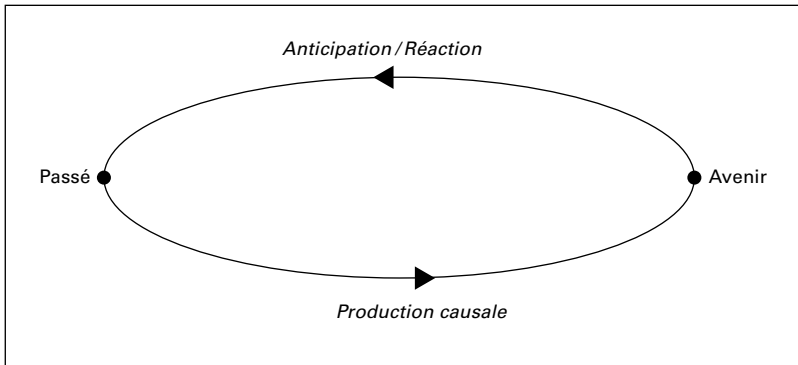
Être fini, c'est se choisir, c'est-à-dire se faire annoncer ce qu'on est en se projetant vers un possible, à l'exclusion des autres¹⁸.

17. Raz 1999, 243.

18. Sartre 1943, 604.

Dans cette autre fiction, je me considère comme déterminé par un x , disons, dans le vocabulaire sartrien, par une essence : je ne suis libre, dans aucun monde possible, d'agir contrairement à cette essence. Comment donc puis-je croire à la fiction que je suis libre ? En faisant comme si j'étais libre de choisir cette essence en choisissant mon existence (« Choisir son destin », « se déterminer », etc.). Les contrefactuels que je vais utiliser dans mes raisonnements pratiques sont du type rétrograde (*backtracking conditionals*) : « Si je faisais ceci au lieu de cela, c'est que mon essence serait celle-ci plutôt que celle-là, et donc... »

J'ai montré, dans mon travail de ces dernières années¹⁹, que deux conceptions du temps de l'action correspondent à ces deux types de raison pratique. Au type classique correspond le temps comme « jardin aux sentiers qui bifurquent », pour reprendre, après David K. Lewis, l'image de Borges. Dans ce temps, que j'ai appelé le temps de l'histoire, nos raisonnements font intervenir des arbres de décision. L'avenir y est « ouvert » au sens qu'il dépend contre-factuellement de l'action présente. Au type autre, correspond un temps que j'ai appelé le temps du projet, lequel se traduit formellement par un cercle. La détermination simultanée du passé (le x de l'agent, son essence, etc.) et de l'avenir (la décision) pose un problème de point fixe. J'ai appelé « équilibre projeté » la ou les solutions de ce problème.



Définition d'un équilibre projeté

19. Voir par exemple : Dupuy 2000, 323-345 ; Dupuy 1992b, 191-220 ; Dupuy 1989, 37-62 ; Dupuy 1998.

L'équilibre projeté est un point fixe du cercle qui mène de l'avenir à l'avenir en passant par le passé, au moyen d'une fonction d'anticipation/réaction et d'une fonction de production causale.

C'est sur le cas de la dissuasion nucléaire que je vais illustrer ce concept.

Le cas de la dissuasion nucléaire ou le salut comme fatalité improbable

Nous sommes à l'automne 1995. Le président de la République française tient une conférence de presse à l'Élysée. La reprise des essais nucléaires dans le Pacifique durant l'été a provoqué un tollé international. Jacques Chirac a décidé de les interrompre définitivement et aussi de réduire la capacité de la force de frappe française au niveau permettant l'anéantissement d'une population de soixante millions d'âmes, soit l'équivalent de la population française. Un membre américain de Greenpeace l'interpelle.

Le représentant de Greenpeace (G) : « Monsieur le Président, comment justifiez-vous le fait que la France soit le seul pays nucléaire à afficher une politique de dissuasion exclusivement axée sur des représailles qui visent les villes et les populations civiles, sans manifester le moindre souci d'épargner les innocents ? »

Jacques Chirac (P) : « La politique nucléaire de la France, telle qu'elle a été réaffirmée dans le *Livre blanc* de 1994, se définit par la volonté et la capacité de faire redouter à un adversaire, quel qu'il soit et quels que soient ses moyens, des dommages inacceptables, hors de proportion avec l'enjeu d'un conflit, s'il cherche à s'en prendre à nos intérêts vitaux. Son objectif est par là même de "rendre inconcevable le recours à la guerre totale comme moyen actif de la politique" »²⁰.

(G) « Votre prédécesseur, François Mitterrand, eut un jour ce mot : "La dissuasion, c'est moi". Monsieur le Président, si la dissuasion nucléaire française, c'est vous, vous sentez-vous capable, le moment venu, de donner l'ordre qui amènera la liquidation physique de soixante millions d'innocents ? Ne trouvez-vous pas cette perspective monstrueuse ? Derrière les abstractions de

20. Institut des hautes études de Défense nationale 1996a, 11.

la stratégie, n'êtes-vous pas sensible aux horribles souffrances des enfants que vous condamnerez ainsi à la mort la plus barbare ? »

- (P) « La France n'a pas, Monsieur, de leçon de morale à recevoir de votre organisation. La France ne sera jamais l'agresseur, donc elle n'encourt pas de responsabilité morale pour ce qu'elle pourrait être amenée à faire en riposte à l'agression ou au chantage. Notre arsenal nucléaire est destiné à prévenir la guerre, non à la livrer²¹. Ne confondons pas ce qui n'est qu'une menace en cas d'atteinte avérée à notre propre existence et l'acte lui-même »²².
- (G) « Comment expliquez-vous, Monsieur le Président, que la France, contrairement aux deux autres membres nucléaires de l'Alliance atlantique, les États-Unis et la Grande-Bretagne, ait systématiquement refusé de faire entrer l'arme atomique dans le domaine d'application du droit des conflits ? Ce droit reprend pourtant les critères classiques de la guerre juste, à savoir la discrimination entre combattants et non-combattants et la proportionnalité entre la riposte et l'enjeu ? »
- (P) « La politique de la France repose sur la conviction que l'arme atomique “n'est pas un moyen de guerre, car elle va outre les limites du raisonnable”²³. Je le répète, la stratégie française est strictement défensive et rejette toute confusion entre dissuasion et emploi »²⁴.
- (G) « Mais enfin, Monsieur le Président, si vous êtes si certain que cela que votre dissuasion est efficace et que jamais vous n'appuierez sur le bouton, quelle est en fin de compte son efficacité ? Puisque l'ennemi sait que vous n'appuierez pas sur le bouton parce que vous croyez que votre dissuasion est efficace, il vous attaquera et vous ne ferez rien en retour. Comment expliquer d'ailleurs que la France ait toujours refusé toute doctrine de riposte graduée et n'ait aucune politique de conduite d'une guerre nucléaire limitée²⁵ ? Vous ne voulez pas considérer le

21. Hassner 1995, 203.

22. Institut des hautes études de Défense nationale 1996a, 26, n. 2 ; et Institut des hautes études de Défense nationale 1996b, 5 - 11.

23. *Ibid.*, 9 - 11 ; et Hassner 1995, 202.

24. Institut des hautes études de Défense nationale 1996a, 11.

25. *Ibid.*, 9 ; Ministère de la Défense 1994, 56. Cf. aussi Institut des hautes études de Défense nationale 1996b : « Toute arme nucléaire de théâtre, à usage tactique,

cas où votre dissuasion échouerait, comme si cela suffisait à rendre la chose impossible ? »²⁶

- (P) « La définition de notre politique est liée à la “liberté d’appréciation des plus hautes autorités de l’État” et donc fondée sur l’incertitude, garantie fondamentale de la dissuasion²⁷. Vous souffrirez, Monsieur, que je n’en discute pas plus avant en public ».

Cet échange²⁸ illustre le *double bind* dans lequel le défenseur de la dissuasion nucléaire semble condamné à tomber. Pour éviter une condamnation éthique évidente, il lui faut insister sur le fait qu’il ne fait que menacer l’adversaire de représailles monstrueuses sans que l’acte lui-même fasse partie de ses plans (au regard de l’éthique déontologique, encore lui faut-il montrer que cette menace n’est pas en soi monstrueuse, même si elle n’est pas mise à exécution). Cela n’est pas suffisant et il lui faut ajouter que la menace ne sera pas mise à exécution. De là on glisse à : « il est inconcevable qu’elle soit mise à exécution » et pour finir à : « il est impossible qu’elle soit mise à exécution ». Mais cela même rend la dissuasion inefficace. Pour que la dissuasion fût à la fois efficace et morale, il faudrait que ce que promet la menace, à savoir l’apocalypse nucléaire, fût un événement à la fois possible et non actuel (*i.e.* non réalisé). Il faudrait que cet événement fût possible, pour qu’il pût avoir, même non réalisé, un impact sur le monde réel – car, ainsi que l’écrit Clausewitz dans *De la guerre* :

En raison de leurs conséquences, les événements possibles doivent être jugés comme réels.

entretient chez autrui l’idée que l’on se ménage une possibilité de recours au nucléaire “intermédiaire”, dans des situations où l’intégrité de notre territoire national n’est pas en danger *et qui relèvent d’autres impératifs que ceux qui justifient la dissuasion*. Nous devrions donc renoncer de façon démonstrative à celles que nous avons et nous interdire d’en mettre d’autres au point. L’opinion internationale s’inquiétera toujours davantage de l’existence de ce type d’armes, et ce d’autant plus que les pays proliférateurs chercheront, quant à eux, à les développer » (Je souligne).

26. Hassner 1995, 202.

27. Institut des hautes études de Défense nationale 1996a, 11.

28. Dont je n’ai pas besoin de dire qu’il est apocryphe, tant il est évident qu’un président de la République française n’aurait pas la patience de se soumettre à un interrogatoire serré sur le sujet. J’ai rassemblé des extraits de documents officiels, en formulant les questions auxquels ils répondent implicitement.

Mais il faudrait évidemment qu'il restât non réalisé, pour que l'aventure nucléaire ne débouchât pas sur le mal absolu. Or je crois possible de montrer que l'acceptabilité de la dissuasion au regard de la prudence implique non seulement que l'apocalypse soit possible, mais aussi qu'elle soit actuelle (elle est inscrite dans l'avenir). On soupçonne que la discussion de ces questions fait intervenir des considérations métaphysiques. Est ici en jeu la question de savoir si l'un des axiomes de l'argument dominateur de Diodore Kronos doit être maintenu en dépit de son incompatibilité avec d'autres axiomes tout autant conformes au bon sens. Cet axiome dit : « Il y a un possible dont la réalisation n'a jamais lieu, ni dans le présent, ni dans le futur ». Si l'on nie cet axiome, donc si l'on pose que le possible est soit ce qui est vrai, soit le sera, alors la dissuasion ne peut être à la fois efficace et morale. Et pourtant, je crois possible de montrer que la dissuasion n'est acceptable au regard de la prudence que si l'on se place dans un cadre métaphysique qui nie l'axiome en question.

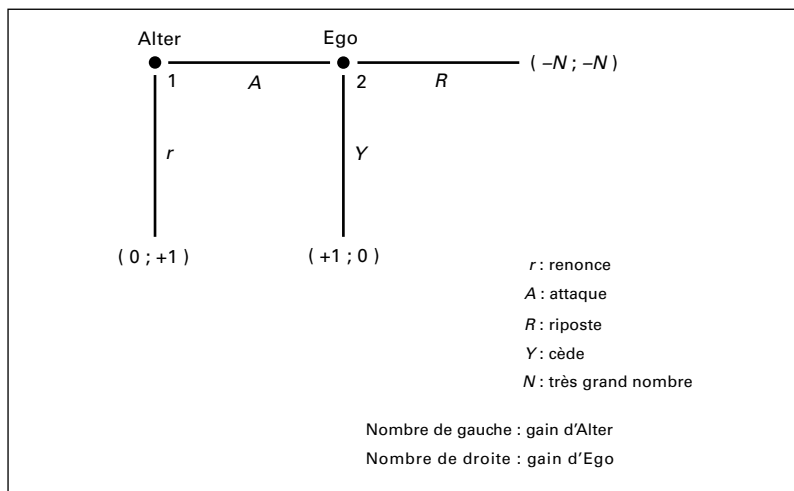
L'un des plus intelligents défenseurs américains de la politique de vulnérabilité mutuelle connue sous le nom de MAD (*Mutually Assured Destruction*) est Bernard Brodie. La citation suivante résume parfaitement sa pensée :

C'est un curieux paradoxe de notre temps que l'un des facteurs essentiels qui font que la dissuasion [nucléaire] marche vraiment, et marche bien, est la peur sous-jacente qu'elle pourrait échouer en cas de crise très grave. Dans ces circonstances, *on ne tente pas le destin*. Si nous étions absolument certains que la dissuasion nucléaire était efficace à cent pour cent dans son rôle de protection contre une attaque nucléaire, alors sa valeur dissuasive contre une guerre conventionnelle tomberait à peu de choses ou même à zéro²⁹.

On comprend la thèse qui est ici énoncée : la dissuasion parfaite est autoréfutante, mais la dissuasion imparfaite peut être efficace et justifiée au regard de la prudence (et même au regard de l'éthique). Je prétends, cependant, que cette thèse ne peut être démontrée (et même exprimée d'une manière qui fasse sens) dans le cadre du raisonnement pratique habituel, en termes de croyances, désirs, intentions et décisions – disons : en termes de stratégie.

Raisonnons sur l'arbre de décision suivant :

29. Brodie 1973, 430-431.



La dissuasion nucléaire et la structure MAD en théorie des jeux

La structure MAD est la réunion de ce jeu et du jeu obtenu en échangeant les positions d'Alter et d'Ego. L'attaque A est une agression conventionnelle, c'est-à-dire non nucléaire. Les joueurs sont tous deux vulnérables : c'est ce que représente la perte N (considérable, équivalente à l'anéantissement de leur société) pour chacun d'entre eux ; ils sont tous deux invulnérables : c'est ce que représente la perte N pour leur adversaire si celui-ci choisit R . Victimes d'une frappe nucléaire, ils ne périront pas seuls.

La thèse énoncée par Brodie nous dit : c'est une incertitude sur son efficacité qui rend la dissuasion efficace. Supposons que cette incertitude soit une variable stratégique. Appelons e la (très faible) probabilité que l'agression (conventionnelle) A d'Alter débouche sur une conflagration nucléaire et le résultat $(-N)$ pour les deux adversaires. Ego pourrait raisonner ainsi : il est rationnel pour moi de choisir de ménager une incertitude e telle que Ne reste supérieur à 1. En effet, s'il attaquait, l'espérance de gain d'Alter serait : 1 fois $(1 - e)$ moins N fois e , soit environ : $1 - Ne$; contre 0 s'il renonçait. Étant rationnel, il devrait donc renoncer. Admettons même que, dans ces circonstances, Alter, irrationnellement, m'attaque avec une probabilité i . À jouer de l'incertitude e , mon gain probable serait : 1 fois $(1 - i)$ plus i fois [0 fois $(1 - e)$ moins N fois e], soit, approximativement : $1 - i - eiN$. Comme eN est de l'ordre de grandeur de 1, ce résultat est de l'ordre de grandeur de $1 - mi$, (m entier positif), il est donc proche de 1. Or si moi, Ego, je renonçais à

jouer de l'incertitude (ce qui rendrait ma dissuasion inefficace, parce qu'absolument non crédible), j'obtiendrais 0. Il est donc rationnel pour moi, Ego, de jouer de l'incertitude³⁰. Mais ce raisonnement est évidemment faux. L'incertitude n'a ici aucune valeur stratégique et elle ne rend la dissuasion aucunement crédible, car, s'il s'agissait véritablement d'une variable stratégique, Alter saurait trop bien que, le moment venu, en cas d'échec de la dissuasion, Ego ne confierait certes pas à une loterie ou à une roulette russe – même affectée des probabilités e et $1 - e$ – le soin de décider pour lui : il céderait. Dans la situation MAD, l'incertitude ne joue aucun rôle dans le temps de l'histoire. C'est cela, très précisément, que l'on nomme l'*effet de boule de cristal*.

Aussi bien n'est-ce pas là l'argument des défenseurs de MAD. Ils voient bien que pour rendre la dissuasion efficace, il faut renoncer à la stratégie, il faut choisir de ne pas choisir. Ils pourraient certes dire, à la manière rendue célèbre par Thomas Schelling et sur le mode de la mal fameée *Doomsday Machine* : lions-nous les mains, rendons la riposte nucléaire automatique. Mais ce serait encore de la stratégie. Ils prennent bien soin de dire autre chose, même si c'est au prix d'une énigme, qu'ils n'éclairent pas. Ils invoquent, comme Bernard Brodie, le destin, la fatalité. Non pas une *fatalité conditionnelle*, qui se traduirait par un conditionnel du type : « si tu m'agresses, il est inévitable que je fasse tout sauter ». On serait encore dans l'intentionnalité. De plus – et cet oxymore reste incompréhensible –, il s'agit d'une *fatalité improbable*. L'apocalypse est comme inscrite dans l'avenir, mais sa probabilité d'occurrence est – Dieu merci – extrêmement faible. Ces notions semblent faire sens, mais elles n'ont de toute évidence pas trouvé une conceptualisation adéquate.

Considérons la thèse selon laquelle la dissuasion nucléaire s'autoréfute si elle est parfaite. Supposons donc que la dissuasion nucléaire soit efficace à cent pour cent face au risque d'une conflagration nucléaire. Le segment R , celui qui mène à la catastrophe, ne fait pas partie du monde actuel. La raison en est, non pas que Ego, en 2, ne joue pas R , mais bien qu'Alter, en 1, jouant r , ne lui donne pas l'occasion de jouer en 2. En d'autres termes, la proposition : « Si

30. Nixon à Haldeman : « I call it the Madman Theory, Bob. I want the North Vietnamese to believe I've reached the point where I might do *anything* to stop the war. We'll just slip the word to them that, "for God sake, you know Nixon is obsessed about Communism. We can't restrain him when he's angry – and he has his hand on the nuclear button" – and Ho Chi Minh himself will be in Paris in two days begging for peace ».

toi, Alter, tu m'agresses en jouant A , moi, Ego, je riposte en jouant R » est prise au sérieux par Alter. Celui-ci renonce. Pour justifier l'autoréfutation, il faudrait maintenant montrer que les hypothèses faites conduisent à une absurdité, au sens où, de la non-actualité de R , on devrait en vérité conclure au fait que Alter choisit d'attaquer. Nous sommes dans le temps de l'histoire, celui qui convient au raisonnement stratégique. On ne pourrait conclure ainsi qu'au prix d'une grave confusion, celle qui consiste à inférer l'impossibilité de la non-actualité. Sous les hypothèses faites, R , bien que non actuel, reste possible, et le fait que R est possible a des effets sur le monde actuel. Le paradoxe de l'autoréfutation ne trouve pas à s'exprimer dans le temps de l'histoire.

Les deux lignes argumentatives en question peuvent être défendues, mais elles appellent la problématique du temps du projet.

Demandons-nous quel est l'*équilibre projeté* du jeu en question. Il ne peut inclure (Ego, R) : la réaction d'Alter anticipant cela serait : (Alter, r), lequel interdit causalement la réalisation de (Ego, R). En revanche, (Ego, Y) donne, par anticipation/réaction : (Alter, A), lequel boucle causalement avec (Ego, Y).

L'équilibre projeté est donc (Alter, A ; Ego, Y). Il est identique à l'équilibre parfait obtenu par *backward induction*. Lui aussi conclut donc à l'inefficacité de la dissuasion nucléaire. Mais la raison en est fondamentalement différente. L'équilibre parfait conclut à l'inefficacité de la dissuasion par manque de crédibilité ; l'équilibre projeté, pour cause d'autoréfutation. La séquence (Alter, r ; Ego, R) se lit, dans le temps du projet : (Alter renonce à attaquer ; Ego riposte nucléairement), ce qui montre assez son impossibilité pour cause d'absurde. Dans le temps de l'histoire, l'interprétation de la même séquence serait : (Alter renonce à attaquer parce que, s'il attaquait, alors Ego riposterait nucléairement). La riposte nucléaire d'Ego, bien que non actuelle puisque Alter renonce, serait un possible, dont l'effet se fait sentir dans le monde actuel. Mais, dans le temps du projet, il n'y a pas de possible qui ne soit actualisé. L'absurdité de (Alter renonce à attaquer ; Ego riposte nucléairement) est, dans le temps du projet, la trace du paradoxe de l'autoréfutation, lequel, précisément, prend alors la forme d'un raisonnement par l'absurde dont la conclusion est l'inefficacité de la dissuasion nucléaire.

Ainsi donc, la critique de la dissuasion nucléaire au regard de la prudence fait intervenir des arguments qui relèvent de deux théories très différentes. Il se trouve qu'elles concluent dans le même sens, en condamnant l'une et l'autre la dissuasion nucléaire. Mais c'est là un résultat contingent, comme on va le voir maintenant.

L'argument si confus des défenseurs de MAD en termes de fatalité improbable trouve une traduction rigoureuse dans le temps du projet. L'incertitude peut rendre la dissuasion nucléaire justifiée au regard de la prudence.

Dans le temps de l'histoire, l'incertitude ne peut affecter qu'un conditionnel contrefactuel du type : « Si toi, Alter, tu m'attaquais, moi, Ego, je riposterais par une frappe nucléaire ». Dans le temps du projet, c'est la fonction d'anticipation/réaction du passé par rapport à l'avenir qui peut être affectée d'incertitude. *Si Ego jouait R en 2, que ferait Alter en 1 ?* À une telle question, la réponse du temps de l'histoire est évidemment : Alter aurait joué *A* en 1, sinon Ego ne serait pas en mesure de jouer en 2. Cependant, dans le temps du projet, ce n'est pas le passé que l'avenir tient pour fixe, mais la fonction d'anticipation/réaction du passé. S'il n'y avait pas d'incertitude, il faudrait donc répondre à la question posée par : si Ego jouait *R* en 2, Alter jouerait *r* en 1³¹. C'est l'impossibilité causale de cette réponse qui permet de conclure que (Ego, *R*) ne fait pas partie de l'équilibre projeté, donc n'est pas un possible.

Si, cependant, la fonction d'anticipation/réaction du passé est affectée d'incertitude, on peut donner sens à la configuration, de prime abord étrange, suivante. (Ego, *R*) peut faire partie de l'équilibre projeté, donc acquérir le statut ontologique d'*actuel*, si l'une des valeurs prises par cette fonction permet la réalisation causale de (Ego, *R*). Il faut donc que la probabilité du conditionnel contrefactuel : (*Si Ego jouait R en 2, Alter jouerait A en 1*) soit strictement positive. Cette probabilité est celle de l'échec de la dissuasion. Mais son sens est radicalement différent du *e* que nous avons considéré dans le temps de l'histoire, comme probabilité de (*Si Alter jouait A en 1, Ego jouerait R en 2*). Dans le temps de l'histoire, l'incertitude s'attache à un acte intentionnel. La dissuasion a été un échec, c'est un fait du passé, et la question est de savoir si Ego va maintenant mettre sa menace à exécution, provoquant l'apocalypse. Dans le temps du projet, le destin apocalyptique fait « déjà » partie du monde actuel, comme fatalité, précisément. Il ne tient qu'aux acteurs du drame de la rendre improbable, en s'abstenant de jouer *A*.

Encore faut-il que (Ego, *R*) fasse vraiment partie de l'équilibre projeté. Pour déterminer si tel est le cas, il faut comparer ce que Ego

31. Il s'agit d'un *backtracking counterfactual conditional* (conditionnel contrefactuel rétrograde) du type : « Si j'achetais un appartement avenue Foch à Paris, (c'est que) je serais riche » ; comparer à : « Si j'achetais un appartement avenue Foch à Paris, je serais pauvre » (car ruiné).

(et Alter, dans le jeu parallèle) retire de son *choix de ne pas choisir* en inscrivant l'apocalypse dans l'avenir, et ce qu'il retirerait à y inscrire la soumission, soit (Ego, Y). Dans ce dernier cas, 0; dans le premier : si f est la probabilité de (Si Ego jouait R en 2, Alter jouerait A en 1), le calcul probabiliste qui, ici, reprend un sens, donne : 1 fois $(1 - f)$ moins N fois f , soit environ : $1 - Nf$. Si l'espérance de dommage, Nf , reste inférieure à la valeur de l'enjeu, 1, le jeu consistant à faire de l'anéantissement mutuel un destin est, bien qu'extrêmement risqué, prudemment justifié. La dissuasion n'est qu'imparfaitement efficace – cette imperfection étant mesurée par la probabilité du contrefactuel rétrograde (Si Ego jouait R en 2, Alter jouerait A en 1) –, mais c'est précisément cette imperfection qui sauve l'efficacité de la dissuasion du risque d'autoréfutation.

Résumons. Dans le *temps de l'histoire*, la menace de provoquer la catastrophe nucléaire n'a aucun effet dissuasif, parce qu'elle n'est pas crédible. Cette absence de crédibilité renvoie à la rationalité minimale d'un joueur préférant perdre un enjeu fini plutôt que de disparaître dans une apocalypse terminale. Dans le *temps du projet*, la catastrophe ne peut avoir d'effet dissuasif parfait, car si la dissuasion était parfaite, elle s'auto-anéantirait. Si la catastrophe est inscrite dans l'avenir, on est dissuadé ; mais le fait d'être dissuadé empêche causalement la catastrophe de se produire. Elle ne se produit donc pas ; mais c'est donc qu'elle est impossible : donc, on n'est pas dissuadé. Si on est dissuadé, on n'est pas dissuadé. Donc, on n'est pas dissuadé.

Il reste cependant, qui peut justifier la dissuasion au regard de la prudence, le rôle crucial de l'imperfection dans le temps du projet. Ce qui fait que la dissuasion ne s'autoréfute pas est son imperfection même. C'est parce qu'il y a une probabilité f que la dissuasion ne marche pas qu'elle marche avec une probabilité $1 - f$. Ce qui pourrait passer pour une tautologie (ce serait évidemment le cas dans la métaphysique du temps de l'histoire) n'en est absolument pas une ici, puisque la proposition précédente n'est pas vraie pour $f = 0$. Le fait que la dissuasion ne marche pas avec une probabilité f strictement positive est ce qui permet l'inscription de la catastrophe dans l'avenir, et c'est cette inscription qui rend la dissuasion efficace, à f près. Notons qu'il serait incorrect de dire que c'est la *possibilité* de l'erreur, avec la probabilité f , qui sauve l'efficacité de la dissuasion – comme si l'erreur et l'absence d'erreur constituaient les deux branches d'une bifurcation. Dans le temps du projet, il n'y a pas de sentiers qui bifurquent. L'erreur n'est pas seulement possible, elle est actuelle, inscrite dans le temps – comme un raté de

plume, en quelque sorte. En d'autres termes, ce qui a des chances de nous sauver est cela même qui nous menace. Pour rendre la perspective de la catastrophe dissuasive, il faut la voir comme un destin, mais un destin qui ne s'accomplirait que par le fait d'une erreur, dont nous serions seuls responsables.

Conclusion

L'examen de deux cas difficiles (l'un au regard de la morale, l'autre à celui de la prudence) nous a amenés à remettre en cause l'assimilation traditionnelle mais trop aisément tenue pour allant de soi entre raisonnement normatif et syllogisme pratique. Le point commun entre ces deux études de cas est qu'elles impliquent une autre métaphysique du temps que celle qui sert de support au syllogisme pratique. Ce dernier présuppose que l'avenir est ouvert, au sens où il dépend contrefactuellement de l'action présente, et, en conséquence, tient pour acquis le troisième axiome de l'argument dominateur : il existe (au moins) un possible dont la réalisation n'a jamais lieu, ni dans le présent, ni dans le futur. La négation de cet axiome ne menace pas le libre arbitre mais conduit à une métaphysique du temps dans laquelle l'avenir est tenu pour fixe, en quelque sorte « écrit à l'avance » : ce qui ne se produira pas est impossible. De là la catégorie d'« impossibilité morale » qui intervient dans la discussion du cas Billy Budd et l'*a priori* surprenante référence au destin dans l'évaluation prudentielle de la dissuasion nucléaire. Dans ce « temps du projet », ainsi que je l'ai appelé, l'avenir et le passé se déterminent mutuellement, le passé contribuant à la réalisation causale de l'avenir et celui-ci à la détermination du passé dans la mesure où le passé se détermine par rapport à l'anticipation de l'avenir. Le raisonnement pratique consiste à rechercher le ou les points fixes de cette structure circulaire, à la manière de la recherche d'un « équilibre réfléchi » dans la démarche méta-éthique dite « cohérentiste ».

Ce qui fait de ces cas des cas « difficiles » s'éclaire à la lumière de cette analyse. Qui dit structure de point fixe dit la possibilité d'une grande variété de « paysages » ou configurations de points fixes, ou plus généralement d'« attracteurs »³², pour un même problème pratique. Le « bon » cas, celui qui contribue à faire penser que la phi-

32. Les termes « paysages » et « attracteurs » (on dit aussi *eigenbehaviors*) sont propres à la théorie des systèmes dynamiques, la première discipline qui a eu à traiter de ces structures.

losophie pratique est une science qui sait déterminer « la » solution de nos problèmes, est celui où il existe un point fixe et où celui-ci est unique. Mais ce cas est probablement exceptionnel.

Le cas Billy Budd, tel que nous l'avons analysé à la suite de Winch, Wiggins et Raz, présente une configuration à deux points fixes. La pluralité des points fixes est sans doute le cas général. Cette configuration amène à considérablement complexifier, sans pour autant l'abandonner, l'analyse de ce que nous tenons pour être l'universalité et l'objectivité du jugement moral. La dissuasion nucléaire présente elle aussi une multiplicité d'attracteurs (correspondant, respectivement, à l'inefficacité de la dissuasion parfaite et au caractère prudemment justifié de la dissuasion imparfaite), mais ce qui fait son intérêt exceptionnel est la nature complexe du second de ces attracteurs (la « fatalité improbable »). Il s'agit en effet de la *superposition* de deux points fixes, correspondant l'un à l'erreur qui permet l'inscription dans l'avenir de la catastrophe, et l'autre à l'efficacité de la dissuasion.

À la lumière de cette analyse, le « dilemme », au sens usuel de ce terme, peut s'analyser ainsi. Il s'agit du cas où il n'existe aucun point fixe, mais où, en revanche, existe un attracteur prenant la forme d'un « cycle limite »³³ nous renvoyant constamment d'une solution à une autre³⁴.

D'autres configurations, plus « pathologiques » encore, sont possibles, qui contribuent à enrichir ce que Charles Larmore a appelé les « Patterns of Moral Complexity »³⁵, c'est-à-dire la substance dont est faite une vie humaine réfléchie³⁶. Une typologie des cas difficiles est donc envisageable sur la base des configurations d'attracteurs dans la métaphysique du temps du projet. La tâche reste pour l'essentiel à accomplir.

Jean-Pierre DUPUY

École Polytechnique et Université Stanford

33. Voir note ci-dessus.

34. J'ai montré, comme rappelé en introduction, que, sur la question de la rationalité du sacrifice qui oppose Rawls à l'utilitarisme, il n'y a pas d'équilibre réfléchi, mais un cycle limite qui nous renvoie constamment de la solution sacrificielle à la solution non sacrificielle. Appliqué au cas Billy Budd, cela donne : si Vere arrive à la conclusion qu'il faut condamner Billy, aussitôt il doit conclure de façon opposée, et inversement. Cf. Dupuy 1996.

35. Larmore 1987.

36. Cf. Canto-Sperber 2001.

Références

- BRITTEN B. (1994), *Billy Budd, Libretto by E.M. Forster and E. Crozier* [1951] (avec une traduction française par J. Bégaud), Paris, Éditions Premières Loges (L'Avant-Scène Opéra).
- BRODIE B. (1973), *War and Politics*, New York, Macmillan.
- CANTO-SPERBER M. (2001), *L'Inquiétude morale et la Vie humaine*, Paris, PUF.
- DUPUY J.-P. (1989), « Common Knowledge, Common Sense », *Theory and Decision*, 27, p. 37-62.
- DUPUY J.-P. (1992a), « Two Temporalities, Two Rationalities : A New Look at Newcomb's Paradox », in *Economics and Cognitive Science*, P. BOURGINE et B. WALLISER (éd.), Oxford – New York, Pergamon, p. 191-220.
- DUPUY J.-P. (1992b), *Le Sacrifice et l'Envie*, Paris, Calmann-Levy.
- DUPUY J.-P. (1996), « On the Rationality of Sacrifice », Document CREA n° 9628.
- DUPUY J.-P. (éd.) (1998) *Self-Deception and Paradoxes of Rationality*, Stanford, Center for the Study of Language and Information.
- DUPUY J.-P. (2000), « Philosophical Foundations of a New Concept of Equilibrium in the Social Sciences : Projected Equilibrium », *Philosophical Studies*, 100, p. 323-345.
- HASSNER P. (1995), *La Violence et la Paix*, Paris, Éditions Esprit.
- INSTITUT DES HAUTES ÉTUDES DE DÉFENSE NATIONALE (1996a), *La Dissuasion*, Comité 5, décembre.
- INSTITUT DES HAUTES ÉTUDES DE DÉFENSE NATIONALE (1996b), *L'Avenir de la défense nucléaire de l'Europe*, Groupe intercomités n° 8, décembre.
- LARMORE C. (1987), *Patterns of Moral Complexities*, Cambridge, Cambridge University Press.
- MELVILLE H. (1924), *Billy Budd, Sailor*, Chicago, University of Chicago Press ; trad. fr. par P. Leyris, *Billy Budd, marin*, Paris, Gallimard, 1980.
- MINISTÈRE DE LA DÉFENSE (1994), *Livre blanc sur la défense*, Paris, La Documentation française.
- RAZ J. (1999), *Engaging Reason. On the Theory of Value and Action*, Oxford, Oxford University Press.
- SARTRE, J.-P. (1943), *L'Être et le Néant*, Paris, Gallimard.

- SIDGWICK H. (1907), *The Methods of Ethics*, Londres, Macmillan, 7^e éd.
- WIGGINS D. (1991), « Truth, and Truth as Predicated of Moral Judgments » in *Needs, Values, Truth. Essays in the Philosophy of Value*, Oxford, Blackwell, p. 139-184.
- WINCH P. (1972), « The Universalisability of Moral Judgements », in *Ethics and Actions*, Londres, Routledge & Kegan Paul.

