

Comment l'Urne de Carter et Leslie se Déverse dans celle de Hempel

Paul FRANCESCHI

Université de Corse
p.franceschi@univ-corse.fr
<http://www.univ-corse.fr/~franceschi>

publié dans the *Canadian Journal of Philosophy*, Vol.29, July 1999, pp. 139-56

Le philosophe mit le pied sur la première marche du *futurotron*. C'était la première fois qu'il utilisait cet appareil pour ses recherches. Bien qu'il vienne seulement d'être mis au point et qu'il ne soit encore qu'à l'état de prototype, ce futurotron pouvait décidément rendre de grands services. De nombreux chercheurs de différentes disciplines l'avaient d'ailleurs déjà utilisé de manière très fructueuse.

Le philosophe prit place aux côtés du pilote sur le siège avant de la machine.

- Quel est le principe de fonctionnement de la machine? demanda-t-il.
- Ce serait un peu long à vous expliquer. C'est basé sur la mécanique quantique.

Le pilote demanda, pressé:

- Quelle époque avez-vous choisie?
- Je souhaite effectuer une investigation en l'an 2150.

Le pilote plaça le sélecteur sur l'année 2150.

- Bien. Nous allons commencer. Etes-vous prêt?
- Oui.

Le futurotron se mit à ronronner. Un flash lumineux intense s'ensuivit. Au bout de quelques minutes, une image floue apparut sur l'écran géant du prototype. Peu à peu, l'image devint plus nette, et puis finalement parfaitement précise. Le philosophe se mit à observer l'écran avec une extrême acuité. Il distinguait nettement des êtres, d'apparence humaine, qui étaient assis dans ce qui semblait être une salle de conférences. En les regardant attentivement, il constata que ceux-ci présentaient les caractéristiques de notre espèce actuelle. Leur physionomie était humaine, mais pourtant il y avait chez eux quelque chose de différent... Dans ce qui paraissait être une projection publique, le philosophe distinguait nettement un faisceau lumineux et coloré qui provenait du front de l'un des êtres placé au centre de la salle, et s'élargissait pour former ce qui constituait l'image projetée d'une véritable scène tridimensionnelle. Les autres individus, assis autour, regardaient attentivement la scène. La projection était d'une netteté parfaite, et d'un réalisme saisissant. On y distinguait nettement deux personnes qui déambulaient en conversant, dans un paysage représentant une plage superbe, battue par les vagues.

Le pilote interrompit soudain:

- Voilà, c'est fini. Nous ne disposons que d'une minute par séance. Le coût de l'appareil est très élevé, vous savez.

Il ajouta:

- Il reste juste une petite formalité à accomplir. Voulez-vous avoir l'amabilité de remplir avec soin ce formulaire. Vous devez indiquer le motif de l'utilisation de la machine. C'est pour nos statistiques.

Le passager parcourut du regard les feuillets tendus par le pilote. Il s'attarda sur la rubrique "Philosophie". Il sauta quelques paragraphes et parvint à la catégorie intitulée "Paradoxes". Celle-ci comportait plusieurs noms, qu'il convenait de cocher de manière adéquate: Eubulide/Menteur, Eubulide/Sorites, Russell, Cantor, Burali-Forti, Carter-Leslie, Goodman, Hempel... Le philosophe prit tranquillement un stylo et cocha la case Carter-Leslie. Puis sans hésiter, il plaça le stylo deux cases en-dessous, et cocha également: Hempel.

- Voilà, c'est fait.
- Vous avez coché deux cases, commenta le pilote.
- Oui, en effet, répondit le passager.
- En principe, on ne doit cocher qu'une seule rubrique.
- Ce serait un peu long à vous expliquer... Mais n'ayez crainte, cela ne faussera pas vos statistiques.

I Le problème de Hempel

Le problème de Hempel (*Hempel's problem*, soit HP) est basé sur le fait que les deux assertions suivantes:

- (H) Tous les corbeaux sont noirs
- (H') Tout ce qui est non-noir est un non-corbeau

sont logiquement équivalentes. La structure logique de (H) est de la forme:

- (H1) Tous les X sont Y

soit $\forall x (Xx \rightarrow Yx)$, alors que celle de (H') est la suivante:

- (H1') Tous les non-Y sont non-X

soit $\forall x (\sim Yx \rightarrow \sim Xx)$. De fait, la structure de la forme contraposée (H1') est clairement équivalente à celle de (H1). Il en résulte que la découverte d'un corbeau noir confirme (H) et également (H'), mais aussi que la découverte d'une chose non-noire qui n'est pas un corbeau telle qu'un flamand rose ou même un parapluie gris, confirme (H') et donc (H). Cette dernière conclusion apparaît comme paradoxale.

Les propositions (H1) et (H1') sont basées sur quatre propriétés X, $\sim X$, Y et $\sim Y$, correspondant respectivement à *corbeau*, *non-corbeau*, *noir*, et *non-noir*, dans la version originale de HP. Ces quatre propriétés déterminent elles-mêmes quatre catégories d'objets: XY, X \sim Y, \sim XY et \sim X \sim Y, qui correspondent aux *corbeaux noirs*, *corbeaux non-noirs*, *non-corbeaux noirs* et *non-corbeaux non-noirs*.

On peut observer ici qu'un *corbeau* est défini de manière précise dans la taxinomie au sein de laquelle il s'insère. Une catégorie comme celle des corbeaux peut être considérée comme bien définie, car elle est basée sur un ensemble de critères précis définissant l'espèce *corvus corax* et permettant l'identification de ses instances.

Il apparaît aussi qu'on peut construire sans difficulté des versions de HP où une variation par rapport à la classe des X est opérée. Si l'on assimile la classe des X à celle des tulipes, ou bien des dauphins, etc. en adaptant de manière corrélative la propriété Y, on obtient ainsi autant de versions valables de HP. Il apparaît ainsi que des changements peuvent être opérés au niveau de la classe X sans que cela n'affecte la problématique inhérente à HP.

De même, la propriété *noir* peut être spécifiée avec précision, à partir d'une taxinomie des couleurs établie par rapport aux longueurs d'onde de la lumière¹. De plus, il est permis d'envisager des variations par rapport à la propriété Y. On pourra choisir ainsi des propriétés telles que *d'une longueur inférieure à 50 cm*, *vivant moins de 10 ans*, etc. De telles variations conduisent également à des versions acceptables de HP.

Enfin, on peut constater que la propriété *non-noir* peut faire l'objet d'une définition qui ne souffre pas d'ambiguïté, à partir notamment de la taxinomie précise des couleurs qui vient d'être mentionnée. De même, si on prend en considération des variations de la propriété Y telles que *inférieur à 40 cm*, ou bien *d'un diamètre supérieur à 25 cm*, etc., on parvient à des définitions de la propriété non-Y qui de même que *non-noir*, sont établies avec précision, et conduisent par ailleurs à des versions de HP comportant la même problématique que la version originale.

Ainsi, la classe des X, de même que les propriétés Y et non-Y pour HP peuvent faire l'objet d'une définition précise et non ambiguë. De plus, des variations opérées sur ces dernières conduisent à des versions acceptables de HP. Il n'en est pas de même pour la classe des non-X.

II La classe de référence des Z

¹ On sait qu'une lumière monochromatique, de longueur d'onde unique, ne se rencontre pratiquement qu'en laboratoire. Mais les couleurs naturelles peuvent être modélisées en termes de soustraction de lumières de certaines longueurs d'onde, à partir de la lumière blanche du Soleil.

La notion de *non-corbeau* présente dans la version originale de HP conduit à mettre en évidence un important problème. Qu'est-ce qui constitue une *instance* d'un non-corbeau? Intuitivement, un merle bleu, un flamand rose, un parapluie gris, voire même un entier naturel, constituent des non-corbeaux.

On se trouve ainsi confronté à la définition d'une nouvelle classe de référence - appelons-la Z - incluant les X et les non-X. La classe Z permet de définir de manière complémentaire celle des non-X, et dans la version originale de Hempel, celle des *non-corbeaux*. Ainsi Z est la classe de référence implicite par rapport à laquelle la définition de la classe X permet celle des non-X.

Doit-on alors envisager une classe Z qui aille jusqu'à inclure les objets abstraits? Faut-il considérer une notion de *non-corbeau* qui englobe des entités abstraites tels que les entiers naturels et les nombres complexes? Ou bien faut-il se limiter à une classe Z qui n'embrasse que des choses concrètes? Une telle discussion a son importance, car les objets abstraits sont en nombre infini, alors que les objets concrets individualisés ne sont présents qu'en nombre fini. Ce fait est de nature à influencer ultérieurement de manière déterminante sur l'application éventuelle d'un raisonnement bayésien.

On pourrait ainsi avoir une classe de référence Z comprenant à la fois les objets abstraits (les entiers naturels, les nombres réels et complexes, etc.) et les objets concrets tels que les artefacts, mais aussi les entités naturelles telles que les humains, les animaux, les végétaux, les météorites, les astres, etc. Une telle classe de référence est définie de manière très extensive. Et un tel choix a pour conséquence que la découverte de n'importe quel objet² confirme (H') et donc (H). A ce stade, n'importe quoi confirme (H).

Notons qu'on peut avoir aussi une conception de la classe Z incluant tous les objets de nature concrète qui viennent d'être cités, mais excluant cette fois les objets abstraits. Les instances de cette classe sont en nombre fini, de même que le cardinal de l'ensemble correspondant: la classe de référence Z inclut alors les animaux, les végétaux, les étoiles, etc.

Mais de manière alternative, on pourrait encore considérer la classe Z associant les corbeaux (*corvus corax*) et les goélands d'Audouin³ (*larus audouinii*). Dans ce cas, les instances de la classe des X (*corvus corax*) sont en nombre supérieur à celles de la classe des non-X (*larus audouinii*). Et on a toujours la version correspondante de HP⁴.

Enfin, rien ne semble interdire, à un niveau très restrictif, de choisir une classe Z composée de la classe X, additionnée seulement d'un unique élément tel qu'une tulipe rouge. Avec cette définition de Z, on a encore une version minimale de HP. Bien sûr, n'importe quel objet, ajouté à la classe des X et constituant celle des non-X conviendra, et confirmera alors à la fois (H') et (H). Ainsi, n'importe quel objet $\sim X \sim Y$ pourra confirmer (H).

Les remarques qui viennent d'être exposées appellent toutefois une objection immédiate. A des degrés divers, il est permis de penser que le choix de chacune des classes de référence Z venant d'être évoquées, relève de l'*arbitraire*.

Car il est permis de rejeter sur ces bases des définitions extrêmes de Z telles que celle définie plus haut et incluant tous les objets abstraits. De même, une classe Z englobant les entiers naturels ou les nombres complexes peut aussi être éliminée. La classe X est définie par rapport aux objets concrets que sont les *corbeaux*, et on n'a pas de raison particulière de choisir une classe Z qui englobe les entités abstraites.

De même, on pourra rejeter la définition de Z basée sur une restriction purement artificielle, associant simplement à X un objet déterminé tel qu'une tulipe rouge. Car je peux choisir, de manière arbitraire, l'objet qui constitue le complément de X, c'est-à-dire définir Z comme je l'entends. Une telle conception extrême apparaît comme sans rapport avec la définition initiale de X. Une classe Z ainsi définie n'est pas homogène. Et on n'a pas de justification pour légitimer l'association d'une *tulipe rouge* à la classe des *corbeaux* pour construire celle des Z.

De même, l'association au sein d'une même classe Z des corbeaux et des goélands d'Audouin, apparaît comme un choix que rien ne vient légitimer. Pourquoi pas en effet l'association des corbeaux et des chardonnerets? De telles associations sont symptomatiques d'une sélection purement artificielle.

Ainsi, les choix des classes de référence Z évoquées plus haut révèlent-ils une nature arbitraire et artificielle. En effet, ne doit-on pas s'attacher à rechercher une classe Z qui soit la plus naturelle et la plus

² Tout objet $\sim X \sim Y$ dans la classe Z ainsi extensivement définie.

³ L'effectif total des goélands d'Audouin est évalué à environ 3000 couples (cf. Thibault 1983, 132).

⁴ Ceci permet incidemment de vérifier que HP ne trouve pas son origine dans une disproportion de la classe des X par rapport à celle des non-X. Que les instances de la classe des X soient en nombre supérieur à celles des non-X n'empêche pas l'émergence d'une version de HP.

homogène possible, compte tenu de la définition des X? On peut penser qu'on doit s'efforcer d'opérer une détermination de la classe des Z qui soit la plus objective possible. Dans la version originale de HP, le choix pour la classe X des corbeaux ne détermine-t-il pas implicitement une classe Z qui soit directement en rapport avec celle des corbeaux? Une classe Z englobant naturellement celle des corbeaux telle que celle des *corvidés*, ou bien celle des *oiseaux*, apparaît comme une bonne candidate. Car une telle classe est au moins déterminée implicitement par le contenu de la classe X.

Mais avant d'analyser des versions construites selon ce fondement, il convient de s'intéresser préalablement à des versions non paradoxales de HP.

III L'analogie avec l'urne

Il est admis de manière notoire que certaines versions⁵ de HP ne conduisent à aucun paradoxe. Tel est le cas notamment si on envisage une classe de référence Z matérialisée par des boîtes, ou bien un jeu de cartes. On peut considérer également une version de HP associée à une urne. On considère ainsi une classe X où les objets sont en nombre fini, et qui ne comprend que des *boules* et des *tétraèdres*. La classe Y est elle-même réduite à deux couleurs: *rouge* et *vert*. On a ainsi quatre types d'objets: des boules rouges, des boules vertes, des tétraèdres rouges et des tétraèdres verts. Dans ce contexte, on a la version suivante de HP:

- (H2) Toutes les boules sont rouges
- (H2') Tous les objets non-rouges sont des non-boules

Il apparaît ici que le cas des tétraèdres rouges peut être ignoré. En effet, leur rôle est indifférent et on peut ainsi se désintéresser de leur présence dans l'urne. Ils peuvent être assimilés à des objets parasites, dont la présence éventuelle dans l'urne ne possède pas d'importance. On est ainsi amené à prendre en considération une urne contenant des objets significatifs constitués par des boules rouges, des boules vertes, et des tétraèdres verts. Et le fait que les objets non-rouges ne puissent être que verts, et que les non-boules ne puissent être que des tétraèdres conduit à considérer, de manière équivalente:

- (H3) Toutes les boules sont rouges
- (H3') Tous les objets verts sont des tétraèdres

qui constitue clairement une version non paradoxale de HP. En effet, le tirage d'une boule rouge confirme (H3) et (H3'), alors que le tirage d'un tétraèdre vert confirme (H3') et (H3).

Considérons maintenant le cas où l'urne contient six objets significatifs⁶. On vient de tirer trois boules rouges et un tétraèdre vert (le tirage est 3-0-1⁷) et on fait alors l'hypothèse (H3). A ce stade, la probabilité que toutes les boules soient rouges, correspond à trois tirages (3-0-3, 4-0-2 et 5-0-1) parmi six possibles (3-0-3, 3-1-2, 3-2-1, 4-0-2, 4-1-1, 5-0-1). De même, la probabilité que tous les objets verts soient des tétraèdres, est identique. Ainsi, $P(H3) = P(H3') = 1/2$ et de même, $P(\sim H3) = P(\sim H3') = 1/2$.

Ces probabilités initiales étant posées, considérons maintenant le cas où l'on vient d'effectuer un nouveau tirage dans l'urne. On tire une autre boule rouge (le tirage est 4-0-1). Ceci correspond à trois compositions possibles de l'urne (4-0-2, 4-1-1, 5-0-1). Soit E l'événement consistant dans le *tirage d'une boule rouge* dans l'urne. On a alors la probabilité de tirer une boule rouge si toutes les boules de l'urne sont rouges, c'est-à-dire $P(E, H3)$ telle que $P(E, H3) = 2/3$, puisque deux cas (4-0-2, 5-0-1) correspondent au fait que toutes les boules soient rouges. De même, $P(E, \sim H3) = 1/3$. La situation est identique si on considère $P(E, H3')$ et $P(E, \sim H3')$. On est alors à même de calculer la probabilité *a posteriori* que toutes les boules soient rouges à l'aide de la formule de Bayes: $P'(H3) = [P(H3) \times P(E, H3)] / [P(H3) \times P(E, H3) + P(\sim H3) \times P(E, \sim H3)] = (0,5 \times 2/3) / (0,5 \times 2/3 + 0,5 \times 1/3) = 2/3$. Et $P'(\sim H3) = 1/3$. On a des résultats identiques pour

⁵ A proprement parler, il ne s'agit donc pas de versions de HP, puisqu'elles sont non paradoxales. Mais les propositions concernées possèdent la structure logique de (H) et (H').

⁶ Les tétraèdres rouges éventuellement trouvés dans l'urne sont considérés comme des objets non significatifs.

⁷ Avec la notation: *n-p-q* (boules rouges - boules vertes - tétraèdres verts).

$P'(H3')$ et $P'(\sim H3')$. Ainsi, $P'(H3) > P(H3)$, et $P'(H3') > P(H3')$, de sorte que l'hypothèse (H3) de même que l'hypothèse équivalente (H3') se trouvent confirmées par le tirage d'une nouvelle boule rouge.

Examinons enfin la situation où, en lieu et place d'une boule rouge, on tire dans l'urne un tétraèdre vert (le tirage est 3-0-2). Soit donc F l'événement consistant dans le *tirage d'un tétraèdre vert*. Dans ce cas, les combinaisons possibles sont au nombre de trois (3-0-3, 3-1-2, 4-0-2). Mais parmi celles-ci, deux (3-0-3, 4-0-2) correspondent à une situation où les hypothèses (H3) et (H3') sont confirmées. Ainsi, $P(F, H3) = P(F, H3') = 2/3$ et $P(F, \sim H3) = P(F, \sim H3') = 1/3$. Le calcul bayésien fournit les mêmes résultats que dans l'hypothèse précédente du tirage d'une boule rouge. Ainsi, dans l'hypothèse du tirage d'un tétraèdre vert, on calcule les probabilités a posteriori $P'(H3) = P'(H3') = 2/3$ et $P'(\sim H3) = P'(\sim H3') = 1/3$. Ainsi, le tirage d'un tétraèdre vert confirme à la fois (H3') et (H3).

Il est intéressant de constater que l'on peut construire aisément des versions de HP permettant d'établir de manière non paradoxale le raisonnement précédent. Considérons ainsi un bloc minéral cubique de 1m de côté. Un tel objet de 1m³ est divisé en 1000 blocs cubiques de 1 dm³, constitués soit de quartz, soit d'albite. On examine cinquante de ces blocs, et on constate que plusieurs d'entre eux sont constitués d'albite de qualité gemme. On est amené à faire l'hypothèse que tous les blocs d'albite sont de qualité gemme. On a alors la version suivante de HP:

- (H4) Tous les blocs d'albite sont de qualité gemme
- (H4') Tous les blocs de qualité non-gemme ne sont pas de l'albite

ce qui équivaut à:

- (H5) Tous les blocs d'albite sont de qualité gemme
- (H5') Tous les blocs de qualité non-gemme sont du quartz

où on a bien l'équivalence entre (H5) et (H5') et où un raisonnement bayésien correct peut s'établir. Un tel exemple (appelons-le *l'urne minérale*) peut également être transposé à d'autres propriétés X et Y, dès lors que des conditions identiques sont préservées.

IV Une solution au problème

On doit, compte tenu de ce qui a été exposé plus haut⁸, s'attacher à mettre en évidence une définition de la classe Z qui ne présente pas un caractère arbitraire et artificiel, mais s'avère au contraire la plus naturelle et la plus homogène possible, compte tenu de la définition de X qui est donnée. Considérons pour cela la version suivante⁹ de HP:

- (H6) Tous les autours des palombes cyrno-sardes possèdent une envergure inférieure à 3,50 m
- (H6') Tous les oiseaux possédant une envergure supérieure à 3,50 m ne sont pas des autours des palombes cyrno-sardes

Dans cette version particulière de (H'), la classe X est celle des *autours des palombes cyrno-sardes*¹⁰, et la classe de référence Z est celle des *oiseaux*. Cette dernière classe présente un rapport évident avec celle des autours des palombes cyrno-sardes. Il est permis de penser qu'une telle manière de définir Z en fonction de X constitue une manière naturelle. En effet, une telle définition ne présente pas un caractère aussi évidemment arbitraire que cela était le cas avec les exemples de classes Z mentionnés plus haut.

⁸ Cf. § II.

⁹ Cette version particulière de HP est choisie ici parce qu'elle est basée sur une classe X correspondant à la *sous-espèce* *accipiter gentilis arrigonii*. A l'inverse, la version originale de HP est fondée sur l'*espèce* *corvus corax*. Le choix d'une sous-espèce pour la classe des X permet simplement ici un niveau supplémentaire d'intégration.

¹⁰ Les autours des palombes cyrno-sardes (*accipiter gentilis arrigonii*) constituent une sous-espèce de l'autour des palombes, spécifique à la Corse et à la Sardaigne. Cette sous-espèce endémique se distingue de la forme nominale de l'autour des palombes par les caractéristiques suivantes (cf. Thibault 1983): la coloration de la tête est noirâtre au lieu de brun noirâtre; le dos est moins brun; le dessous est plus foncé.

Bien sûr, on peut observer qu'il est possible de choisir, de manière plus restreinte mais aussi naturelle, une classe Z correspondant au genre *accipiter*. Une telle classe présente un caractère homogène. Elle comprend notamment les espèces *accipiter gentilis* (autour des palombes) mais également *accipiter nisus* (épervier d'Europe), *accipiter novaehollandiae* (autour variable), *accipiter melanoleucus* (autour noir et blanc).

Toutefois, de manière alternative, et selon le même point de vue, on pourrait également étendre la classe Z aux instances de la famille des *accipitridés*¹¹ - plus étendue - comprenant à la fois le genre *accipiter* qui vient d'être mentionné, mais également les genres *milvus* (milan), *buteo* (buse), *aquila*, etc. Une telle classe inclut notamment les espèces *milvus migrans* (milan noir), *milvus milvus* (milan royal), *buteo buteo* (buse variable), *aquila chrysaetos* (aigle royal), etc.

Ces différentes définitions acceptables de la classe Z trouvent leur justification dans la taxinomie au sein de laquelle s'insère l'autour des palombes cyrno-sarde. De manière plus systématique, ce dernier appartient à la sous-espèce *accipiter gentilis arrigonii*, à l'espèce *accipiter gentilis*, au genre *accipiter*, à la famille des *accipitridés*, à l'ordre des *falconiformes*, à la classe des *oiseaux*, au sous-embranchement des *vertébrés*, à l'embranchement des *chordés*¹², au règne *animal*, etc. Il en résulte que les variations suivantes de (H') sont acceptables, dans le sens qui vient d'être défini:

- (H7') Tous les *autours des palombes* possédant une envergure supérieure à 3,50 m ne sont pas des *autours des palombes cyrno-sardes*
- (H8') Tous les *autours* possédant une envergure supérieure à 3,50 m ne sont pas des *autours des palombes cyrno-sardes*
- (H9') Tous les *accipitridés* possédant une envergure supérieure à 3,50 m ne sont pas des *autours des palombes cyrno-sardes*
- (H10') Tous les *falconiformes* possédant une envergure supérieure à 3,50 m ne sont pas des *autours des palombes cyrno-sardes*
- (H11') Tous les *oiseaux* possédant une envergure supérieure à 3,50 m ne sont pas des *autours des palombes cyrno-sardes*
- (H12') Tous les *vertébrés* possédant une envergure supérieure à 3,50 m ne sont pas des *autours des palombes cyrno-sardes*
- (H13') Tous les *chordés* possédant une envergure supérieure à 3,50 m ne sont pas des *autours des palombes cyrno-sardes*
- (H14') Tous les *animaux* possédant une envergure supérieure à 3,50 m ne sont pas des *autours des palombes cyrno-sardes*

On a ainsi plusieurs versions de (H'), correspondant à des variations de la classe Z qui sont elles-mêmes rendues possibles par le fait que l'autour des palombes cyrno-sarde appartient à *n* catégories, déterminées par la taxinomie à laquelle il appartient.

Et de fait, lorsque je rencontre un autour des palombes appartenant à la forme nominale (*accipiter gentilis gentilis*), il s'agit à la fois d'un autour des palombes (*accipiter gentilis*) non-cyrno-sarde (*non-accipiter gentilis arrigonii*), d'un autour (*accipiter*) non-autour des palombes cyrno-sarde, d'un accipitridé non-autour des palombes cyrno-sarde, d'un falconiforme non-autour des palombes cyrno-sarde, d'un oiseau (*aves*) non-autour des palombes cyrno-sarde, mais aussi d'un vertébré non-autour des palombes cyrno-sarde, d'un chordé non-autour des palombes cyrno-sarde, d'un animal non-autour des palombes cyrno-sarde. Ainsi, l'instance d'*accipiter gentilis gentilis* que je viens d'observer, appartient à la fois à toutes ces catégories.

Et lorsque je rencontre une baleine grise, ce n'est pas un oiseau non-autour des palombes cyrno-sarde, mais c'est bien un vertébré non-autour des palombes cyrno-sarde, ainsi qu'un chordé non-autour des palombes cyrno-sarde, de même qu'un animal non-autour des palombes cyrno-sarde.

¹¹ Les ornithologues distinguent encore la classe des *accipitriformes*, correspondant à l'ensemble des accipitridés, auxquels s'ajoutent les *pandlionidés*, telles que *pandlion haliaetus* (balbuzard pêcheur), etc.

¹² L'embranchement des *chordés* comprend l'ensemble des vertébrés et quelques invertébrés qui présentent la propriété de posséder une corde dorsale, au moins à une période donnée de leur vie.

D'une manière générale, un objet x découvert appartient à n niveaux dans la taxinomie au sein de laquelle il s'insère. Il appartient ainsi à une sous-espèce¹³, à une espèce, à un sous-genre, à un genre, à un super-genre, à une sous-famille, à une famille, à une super-famille, à un sous-embranchement, à un embranchement, à un règne... On peut assigner à la sous-espèce le niveau¹⁴ 1 de la taxinomie, à l'espèce le niveau 2, ..., à la super-famille le niveau 8, etc. Et si au sein de (H), la classe X se situe au niveau p , il est clair que Z doit se situer à un niveau q tel que $q > p$. Mais comment fixer Z à un niveau q qui ne soit pas arbitraire? Car la classe de référence Z correspond à un niveau d'intégration. Mais où doit-on s'arrêter? Doit-on fixer Z au niveau de l'espèce, du sous-genre, du genre, ..., du règne? On ne possède pas de *critère objectif* permettant de choisir un niveau q parmi les possibilités qui sont offertes. Je peux choisir q proche de p en opérant par *restriction*; mais de manière aussi concluante, je suis autorisé à choisir q éloigné de p , en appliquant un principe d'*extension*. Alors pourquoi choisir telle classe de référence définie de manière restrictive plutôt que telle autre définie de façon extensive? On ne possède pas en réalité de *critère* pour légitimer le choix, selon que l'on procède par *restriction* ou par *extension*, de la classe Z . Dès lors, il apparaît que celle-ci ne peut être définie qu'*arbitrairement*. Et il s'ensuit ici nettement que la détermination de la classe des Z et donc des non- X relève de l'*arbitraire*.

Mais le choix de la classe de référence Z se révèle fondamental. Car selon que je choisirai telle ou telle classe de référence Z , il en résultera qu'un objet x donné confirmera ou non (H). Pour tout objet x , je peux construire une classe Z telle que x appartient à non- X , comme je peux choisir une classe Z telle que x n'appartient pas à non- X . Ainsi, ce choix est laissé à mon arbitraire. Pour un objet x donné, je peux construire une classe Z telle que cet objet confirme (H) et une autre classe Z' telle que cet objet ne confirme pas (H).

Bien sûr, dans le cas où Z est choisi arbitrairement, le raisonnement bayésien propre à HP "fonctionne", mais correspond à un point de vue arbitraire et artificiel: ayant trouvé l'objet x , (H) est confirmée. Mais on peut aussi bien choisir, de manière aussi artificielle et plus restrictive, une classe Z dont x est absent et où x ne confirme pas (H). Ainsi, on n'est pas autorisé à conclure, de manière *objective*, que la découverte de l'objet x confirme (H). Car raisonner ainsi reviendrait à conférer une valeur universelle et générale à un point de vue qui n'est que l'expression d'un choix arbitraire.

Comment ce résultat peut-il se concilier avec les faits mentionnés plus haut¹⁵, concernant l'existence de versions non paradoxales de HP? Il convient d'observer ici que le raisonnement bayésien peut s'établir dans chaque cas où la classe des Z est finie, et où ce fait est connu préalablement à l'expérience¹⁶. On peut alors conclure à un décalage bayésien. Mais à ce stade, il convient de distinguer les cas où la classe Z est déterminée, préalablement à l'expérience, par un critère objectif, et les cas où elle ne l'est pas.

Dans le premier cas, le contenu de la classe Z est donné préalablement à l'expérience, et la classe Z n'est donc pas choisie de manière arbitraire, mais selon un critère objectif. Dès lors, le raisonnement bayésien est correct et fournit des informations pertinentes. Tel est notamment le cas lorsqu'on considère une version de HP appliquée à une urne, ou encore une version telle que l'urne minérale. Dans cette dernière hypothèse, la composition de la classe Z est fixée à l'avance. On a alors une différence importante avec le critère de Nicod¹⁷: un objet $\sim X \sim Y$ confirme (H) et un objet XY confirme (H').

A l'inverse, lorsque la classe Z n'est pas fixée et déterminée préalablement à l'expérience par un critère objectif, on peut choisir subjectivement cette dernière à n'importe quel niveau d'extension ou de

¹³ Il est possible de considérer, si on le souhaite, de manière alternative une autre taxinomie que notre actuelle taxinomie scientifique. Cela n'entache pas le point de vue qui est exposé, puisque les conclusions sont identiques, dès lors que les principes de la classification sont respectés.

¹⁴ Il est évidemment possible de prendre en considération des taxinomies plus fines et comprenant des subdivisions supplémentaires à partir des différentes sous-espèces. De manière évidente, cela n'affecte pas le présent raisonnement.

¹⁵ Cf. § III.

¹⁶ Comme on l'a vu, le raisonnement bayésien n'a pas sa place lorsqu'on considère une classe Z incluant les ensembles infinis tels que les entiers naturels, les réels, etc.

¹⁷ Le critère de Nicod est ainsi défini (Hempel 1945, 11), pour $S_1 = (H)$ et $S_2 = (H')$: "(...) let a, b, c, d be four objects such that a is a raven and black, b is a raven but not black, c not a raven but black and d neither a raven nor black. Then, according to Nicod's criterion, a would confirm S_1 , but be neutral with respect to S_2 ; b would disconfirm both S_1 and S_2 ; c would be neutral with respect to both S_1 and S_2 , and d would confirm S_1 , but be neutral with respect to S_2 ."

restriction, mais les conclusions issues du raisonnement bayésien doivent être considérées comme purement arbitraires, et ne présentent donc pas de valeur objective. Car on ne possède pas alors de *fondement* et de *justification* pour choisir tel ou tel niveau de *restriction* ou d'*extension*. Ainsi, dans ce cas, le critère de Nicod selon lequel tout objet $\sim X \sim Y$ est neutre vis-à-vis de (H) et tout objet XY est neutre vis-à-vis de (H') trouve à s'appliquer.

On le voit, la présente solution a pour effet de préserver l'équivalence d'une proposition et de sa contraposition. Et de même, le principe de la confirmation d'une généralisation par chacune de ses instances s'en trouve également conservé.

V Une solution commune pour le problème de Hempel et l'Argument de l'Apocalypse

L'Argument de l'Apocalypse (*Doomsday Argument*, soit DA) attribué à Brandon Carter, a été décrit par John Leslie (1992). DA peut être exposé de la manière suivante. Considérons l'événement A: *l'extinction définitive de l'espèce humaine se produira avant l'an 2150*. On peut estimer à une chance sur 100 la probabilité que cette disparition survienne: $P(A) = 0,01$. Soit également $\sim A$ l'événement: *l'extinction définitive de l'espèce humaine ne se produira pas avant 2150*. Soit encore E l'événement: *je vis durant les années 1990*. On peut par ailleurs estimer aujourd'hui à 50 milliards le nombre d'humains ayant existé depuis la naissance de l'humanité: soit H_{1997} un tel nombre. De même, la population actuelle peut être évaluée à 5 milliards: $P_{1997} = 5 \times 10^9$. On calcule ainsi qu'un humain sur dix, si l'événement A survient, aura connu les années 1990. On évalue alors la probabilité que l'humanité soit éteinte avant 2150, si j'ai connu les années 1990: $P(E, A) = 5 \times 10^9 / 5 \times 10^{10} = 0,1$. Par contre, si l'humanité passe le cap des années 2150, on peut penser qu'elle sera appelée à une expansion beaucoup plus importante, et que le nombre des humains pourra s'élever par exemple à 5×10^{12} . Dans ce cas, la probabilité que l'humanité ne soit pas éteinte après 2150, si j'ai connu les années 1990 s'évalue ainsi: $P(E, \sim A) = 5 \times 10^9 / 5 \times 10^{12} = 0,001$. Ceci permet maintenant de calculer la probabilité a posteriori de l'extinction de l'espèce humaine avant 2150, à l'aide de la formule de Bayes: $P'(A) = [P(A) \times P(E, A)] / [P(A) \times P(E, A) + P(\sim A) \times P(E, \sim A)] = (0,01 \times 0,1) / (0,01 \times 0,1 + 0,99 \times 0,001) \approx 0,5025$. Ainsi, la prise en compte du fait que je vis actuellement fait passer la probabilité de l'extinction de l'espèce humaine avant 2150 de 0,01 à 50,25.

J'ai présenté dans mon article "Une Solution pour l'Argument de l'Apocalypse" une solution à DA, dont les lignes essentielles peuvent être ainsi décrites. L'argument de DA est basé sur une classe de référence unique qui est celle des *humains*¹⁸. Mais comment cette classe de référence doit-elle être définie? Faut-il la limiter aux seuls représentants de notre *sous-espèce* actuelle *Homo sapiens sapiens*? Ou bien doit-on l'étendre à tous les représentants de l'*espèce Homo sapiens*, en incluant cette fois, outre *Homo sapiens sapiens*, *Homo sapiens neandertalensis*...? Ou faut-il inclure dans la classe de référence l'ensemble du genre *Homo*, englobant alors tous les représentants successifs de *Homo erectus*, *Homo habilis*, *Homo sapiens*, etc.? Et n'est-il pas nécessaire d'aller jusqu'à prévoir une classe plus étendue encore, incluant tous les représentants d'un *super-genre S*, composé non seulement du genre *Homo*, mais aussi des nouveaux genres *Surhomo*, *Hyperhomo*, etc. qui résulteront des évolutions prévisibles de notre espèce actuelle?

Il apparaît ainsi qu'on peut considérer une classe de référence réduite en procédant par *restriction*, ou appréhender une classe plus grande en faisant le choix d'une classe de référence par *extension*. On peut donc opérer pour le choix de la classe de référence en appliquant soit un principe de restriction, soit un principe d'extension. Et selon que l'on applique l'un ou l'autre principe, différents niveaux de choix sont à chaque fois possibles.

Mais il apparaît qu'on ne dispose pas de *critère objectif* qui permette de légitimer le choix de telle ou telle classe de référence. Et même notre sous-espèce actuelle *Homo sapiens sapiens* ne peut être considérée comme un choix naturel et adéquat pour la classe de référence. Car n'est-il pas permis de penser que notre concept paradigmatique d'humain est appelé à subir des évolutions? Et d'autre part, le fait d'exclure des sous-espèces antérieures telles qu'*Homo sapiens neandertalensis*, ainsi que les futures évolutions de notre espèce, ne relève-t-il pas d'une démarche anthropocentrique?

Dès lors que l'on ne dispose pas d'un critère de choix objectif, on peut opter, arbitrairement, pour l'une ou l'autre des classes qui viennent d'être décrites. On peut par exemple identifier la classe de référence

¹⁸ Leslie utilise les termes de *human race*.

à l'espèce *Homo sapiens*, et conclure à un décalage bayésien. On a bien alors une augmentation de la probabilité a posteriori de l'extinction d'*Homo sapiens*.

Mais ce décalage bayésien vaut aussi bien pour une classe de référence plus restreinte encore, telle que notre sous-espèce *Homo sapiens sapiens*. Là aussi, l'application de la formule de Bayes met en évidence une sensible augmentation de la probabilité a posteriori de la fin prochaine d'*Homo sapiens sapiens*.

Cependant, de manière identique, le décalage bayésien s'applique aussi à une classe de référence plus réduite encore qui est celle des représentants d'*Homo sapiens sapiens n'ayant pas connu l'ordinateur*. Une telle classe de référence est certainement vouée à une extinction prochaine. Toutefois, il ne s'agit pas là d'une conclusion de nature à nous effrayer, car les potentialités évolutives de notre espèce sont telles que la succession des nouvelles espèces à celles qui les ont précédées, constitue une des caractéristiques de notre mode d'évolution.

On le voit, cette solution conduit ici à *accepter la conclusion* (le décalage bayésien) de Carter et Leslie pour une classe de référence donnée, tout en la plaçant sur un pied d'égalité avec des conclusions de même nature relatives à d'autres classes de référence, tout à fait inoffensives. La prise en compte de différents niveaux de restriction, rendue légitime par l'absence d'un critère objectif de choix, conduit finalement à l'innocuité de l'argument.

Ainsi, il apparaît que l'argument basé sur la classe de référence et son choix arbitraire par restriction ou par extension constitue une solution commune à HP et DA. HP et DA sont en définitive sous-tendus par un même problème inhérent à la définition de la classe Z de HP et de la classe de référence unique de DA. On a donc une solution de même nature pour les deux paradoxes.

Il convient ici de conclure en présentant un élément qui tend à confirmer l'origine commune des deux problèmes. On observera tout d'abord que l'on n'est pas en mesure de mettre en évidence une version de DA correspondant véritablement à la version originale de HP. En effet, une classe de référence telle que celle des *corbeaux* n'est pas transposable dans DA. L'argument inhérent à DA est en effet basé sur l'usage du principe anthropique, et requiert évidemment une classe de référence composée d'êtres intelligents. Lorsque Leslie¹⁹ envisage l'extension de la classe de référence, il précise expressément que la condition pour l'appartenance à la classe de référence est l'aptitude à produire un raisonnement anthropique.

Il est en revanche possible de décrire une version de HP constituée à partir des éléments de DA. Si on prend X pour notre actuelle sous-espèce *Homo sapiens sapiens*, et Y pour *ne seront vivants qu'avant 2150*, on obtient la version suivante de HP:

(H15) Tous les *Homo sapiens sapiens* ne seront vivants qu'avant l'an 2150

(H15') Tous ceux qui vivront après 2150 seront des non-*Homo sapiens sapiens*

Dans ce contexte, un être humain vivant en 1997 constitue une instance confirmant (H15). Parallèlement, la découverte d'un *Homo sapiens sapiens* après 2150 conduit à infirmer (H15). Enfin, la découverte d'un non-*Homo sapiens sapiens* vivant après 2150 constitue une confirmation de (H15') et donc de (H15). Compte tenu de cette formulation particulière, il est clair qu'on n'observe actuellement que des instances confirmant (H15). En revanche, après 2150, on pourra avoir des instances infirmant (H15), ou des instances confirmant (H15').

Il convient ici d'observer que (H15) ne permet pas véritablement de servir de support à une version de DA. En effet, la classe de référence s'identifie ici de manière précise à *Homo sapiens sapiens*, alors que dans la version originale de DA, la classe de référence est constituée par l'espèce humaine (*human race*). En conséquence, on n'a pas à proprement parler une identité entre les événements sous-tendus par (H15) et A, de sorte que (H15)-(H15') ne constitue pas une version conjointe²⁰ de DA et HP.

Mais cette version de HP étant constituée à partir des éléments de DA, on doit être à même, à ce stade, de vérifier l'origine commune des deux problèmes, en montrant comment les arguments apportés en défense de DA par rapport à la classe de référence, peuvent également être utilisés en support de HP.

¹⁹ "How much widening of the reference class is appropriate when we look towards the future? There are strong grounds for widening it to include our evolutionarily much-altered descendants, three-armed or otherwise, as 'humans' for doomsday argument purposes - granted, that's to say, that their intelligence would remain well above the chimpanzee level." (1996, 262)

²⁰ C'est-à-dire comportant simultanément les deux problèmes.

On connaît la réponse apportée par Leslie à l'objection selon laquelle la classe de référence des *humains* pour DA est ambiguë, et au fait que la prise en considération des espèces ayant précédé ou issues des évolutions d'*Homo sapiens sapiens*, conduit à une classe de référence hétérogène, de nature composite. Elle est exposée dans la réponse faite à Eckhardt (1993):

How far should the reference class extend? (...) One can place the boundary more or less where one pleases, provided that one adjusts one's prior probability accordingly. Exclude, if you really want to, all future beings with intelligence quotients above five thousand, calling them demi-gods and not humans²¹.

et développée dans *The End of the World*²²:

The moral could seem to be that one's reference class might be made more or less what one liked. (...) What if we wanted to count our much-modified descendants, perhaps with three arms or with godlike intelligence, as 'genuinely human'? There would be nothing wrong in this. Yet if we were instead interested in the future only of two-armed humans, or of humans with intelligence much like that of humans today, then there would be nothing wrong in refusing to count any others²³.

Pour Leslie, on peut aller jusqu'à inclure dans la classe de référence, des descendants de l'humanité devenus très éloignés de notre espèce actuelle par le fait de l'évolution. Mais de manière libérale, Leslie accepte aussi qu'on limite la classe de référence aux seuls individus proches de notre humanité actuelle. On est ainsi libre de choisir la classe de référence que l'on souhaite, en opérant soit par extension, soit par restriction. Il suffira dans chaque cas d'ajuster en conséquence la probabilité initiale.

Il apparaît ici que ce mode de réponse peut être transposé, *littéralement*, à une objection à HP de même nature, fondée sur la classe de référence de (H15)-(H15'). On peut fixer - pourrait ainsi dire un défenseur de HP - la classe Z comme on le souhaite, et assigner à "tous ceux" le contenu désiré. On peut par exemple limiter Z à l'*espèce* *Homo sapiens*, ou bien l'assimiler à l'ensemble du *genre* *Homo*, incluant alors les évolutions de notre espèce telles que *Homo spatialis*, *Homo computeris*, etc. Ce qui importe - pourrait poursuivre ce défenseur - est de déterminer au préalable la classe de référence, et de s'en tenir à cette définition lorsque les différentes instances sont ensuite rencontrées.

Ainsi, il s'avère que les arguments avancés en support de la classe de référence de DA peuvent être transposés pour la défense de HP. Ceci constitue un élément supplémentaire, allant dans le sens de l'origine commune des deux problèmes, liée à la définition d'une classe de référence. DA et HP appellent par conséquent le même type de réponse. Ainsi, l'urne de Carter et Leslie se déverse dans celle de Hempel²⁴.

Références

- ECKHARDT, W. 1993. "Probability Theory and the Doomsday Argument." *Mind*, 102 (1993): 483-8
- FRANCESCHI, P. "Une Solution pour l'Argument de l'Apocalypse." *Canadian Journal of Philosophy* (à paraître)
- GOODMAN, N. 1955. *Fact, Fiction and Forecast*. Cambridge: Harvard University Press.
- HEMPEL, C. 1945. "Studies in the logic of confirmation." *Mind*, 54 (1945): 1-26 et 97-121
- LESLIE, J. 1992. "Time and the Anthropic Principle." *Mind*, 101 (1992): 521-40
- . 1993. "Doom and probabilities." *Mind*, 102 (1993): 489-91
- . 1996. *The End of the World: the science and ethics of human extinction*. London and New York: Routledge.
- PAPINEAU, D. 1995. "Methodology: the Elements of the Philosophy of Science." In *Philosophy A Guide Through the Subject*, ed. A.C. Grayling. Oxford: Oxford University Press.

²¹ Cf. Leslie (1993, 491).

²² Ce point de vue est détaillé par Leslie, dans la partie intitulée "Just who should count as being human?" (1996, 256-63).

²³ Cf. Leslie (1996, 260).

²⁴ Je remercie deux experts anonymes du *Canadian Journal of Philosophy* pour leurs commentaires, concernant une version précédente de cet article.

SAINSBURY, M. 1988. *Paradoxes*. New York: Cambridge University Press.
THIBAUT, J-C. 1983. *Les oiseaux de Corse*. Paris: De Gerfau.