



HAL
open science

Une application des n-univers à l'argument de l'Apocalypse et au paradoxe de Goodman

Paul Franceschi

► **To cite this version:**

Paul Franceschi. Une application des n-univers à l'argument de l'Apocalypse et au paradoxe de Goodman. Sciences de l'Homme et Société. Université Pascal Paoli, 2002. Français. NNT: . tel-00002811

HAL Id: tel-00002811

<https://theses.hal.science/tel-00002811>

Submitted on 8 May 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**UNE APPLICATION DES N-UNIVERS
A L'ARGUMENT DE L'APOCALYPSE
ET AU PARADOXE DE GOODMAN**

Paul Franceschi

Université de Corse
p.franceschi@univ-corse.fr

<http://www.univ-corse.fr/~franceschi>

Thèse de doctorat soutenue à Corté le 6 novembre 2002
pour l'obtention du grade de docteur

sous la co-direction de:

Pr. Dominique Salini (Université de Corse)
Pr. Claude Panaccio (Université du Québec Trois Rivières)

rapporteurs:

Pr. Jean-Paul Delahaye (Université de Lille)
Pr. Daniel Andler (Ecole Normale Supérieure, Rue d'Ulm)

jury:

Pr. Claude Panaccio, Président (Université du Québec Trois Rivières)
Pr. Dominique Salini (Université de Corse)
Pr. Jean-Paul Delahaye (Université de Lille)
Pr. Jean-François Santucci (Université de Corse)

Discipline: Philosophie
Spécialité: Philosophie Analytique

Résumé

Un certain nombre de problèmes philosophiques présentent la particularité d'être basés sur une analogie entre une situation réelle et un modèle probabiliste. De tels problèmes reposent ainsi sur une analogie avec une urne. La présente thèse se propose de décrire et de mettre en application une méthodologie destinée à la résolution des problèmes philosophiques basés sur une analogie avec une urne. Cette méthodologie est basée sur l'utilisation des n-univers. A cet effet, je décris tout d'abord les n-univers de manière détaillée. Je discute également les difficultés de la théorie des n-univers liés à la démultiplication des critères ainsi qu'à la relation un/plusieurs entre les objets et un critère donné.

En premier lieu, je présente ensuite une application du modèle des n-univers à *l'argument de l'Apocalypse (Doomsday Argument)* et aux problèmes récemment apparus dans la littérature en relation avec celui-ci. Je m'attache également à montrer comment l'application du modèle des n-univers à plusieurs problèmes et expériences de pensée liés à *l'argument de l'Apocalypse* permet d'en clarifier les données et de dissiper l'ambiguïté qui préside à la description des situations correspondantes. Je développe ainsi une analyse des problèmes suivants liés à l'Argument de l'Apocalypse: le *modèle des deux urnes (two urn case)*, l'expérience du *pile ou face divin (God's Coin Toss)*, le *problème de la Belle au bois dormant (Sleeping Beauty Problem)*, le *Philosophe présomptueux (Presumptuous Philosopher)*, *Adam paresseux (Lazy Adam)*, le *paradoxe de la chambre d'exécution (Shooting-Room Paradox)*. Je présente enfin une solution pour l'Argument de l'Apocalypse, qui repose sur une troisième voie par rapport aux deux types principaux de solutions classiquement décrites.

En second lieu, je développe une application de la théorie des n-univers au paradoxe de Goodman (*Goodman's Paradox*). Je replace tout d'abord l'énoncé de Goodman dans le cadre des n-univers. Puis je propose une solution pour le paradoxe, sur la base d'une distinction entre deux modélisations de l'énoncé de Goodman dans des n-univers distincts.

Remerciements

Je remercie le Professeur Dominique Salini l'ensemble de ses conseils, ainsi que pour son soutien constant et sa confiance tout au long de la démarche.

Je suis également très reconnaissant envers le Professeur Claude Panaccio pour l'ensemble de ses commentaires, formulés à chacune des étapes de la présente thèse. C'est grâce à ses remarques précises et détaillées, son souci constant d'unité et ses orientations que cette dissertation a pu prendre sa structure finale et sa forme définitive.

Je remercie également les Professeurs Daniel Andler et Jean-Paul Delahaye pour des commentaires détaillés et très utiles portant sur l'ensemble de la présente thèse.

Je suis reconnaissant envers plusieurs experts anonymes du *Canadian Journal of Philosophy* pour leurs commentaires concernant mes articles "Une Solution pour l'Argument de l'Apocalypse" et "Comment l'urne de Carter et Leslie se déverse dans celle de Hempel" dont le matériau constitue une partie de l'analyse du problème de la définition de la classe de référence dans l'Argument de l'Apocalypse présentée ici.

Je remercie également plusieurs experts anonymes de *Dialogue* pour leurs commentaires relatifs à mon article "Une solution pour le paradoxe de Goodman", qui constitue le matériau du dernier chapitre.

Je suis enfin reconnaissant envers Nick Bostrom, Taner Edis, Bradley Monton pour des discussions très utiles concernant le sujet de la présente thèse.

Introduction

La présente dissertation se propose de présenter et de mettre en oeuvre une méthodologie destinée à l'étude de certains types de paradoxes et d'arguments philosophiques. Cette approche méthodologique repose sur l'utilisation des n-univers. J'ai décrit pour la première fois le modèle des n-univers dans mon article *Une solution pour le paradoxe de Goodman* publié dans *Dialogue* (2001). Cet article présentait le modèle des n-univers et le mettait ensuite en oeuvre afin de proposer une solution pour le paradoxe de Goodman. Dans la présente dissertation, je développe en premier lieu, de manière approfondie le modèle des n-univers. A cet égard, j'aborde les difficultés de la théorie des n-univers liées à la démultiplication des critères. De plus, j'y discute également les problèmes posés par la relation un/plusieurs entre les objets et un critère donné. En second lieu, je m'attache à décrire l'application de la théorie des n-univers à d'autres problèmes ou paradoxes contemporains. J'étends ainsi l'utilisation des n-univers à plusieurs autres problèmes en montrant comment leur utilisation permet de dissiper l'ambiguïté qui préside à l'énoncé de ces problèmes philosophiques et de formuler une solution. Ainsi, le problème de la Belle au bois dormant (*Sleeping Beauty Problem*), le *Paradoxe de la chambre d'exécution* (*Shooting-Room Paradox*) et enfin l'Argument de l'Apocalypse (*Doomsday Argument*) ainsi que le paradoxe de Goodman (*Goodman's Paradox*) sont finalement analysés à la lumière du modèle des n-univers. Il s'agit là à la fois des problèmes récents (l'Argument de l'Apocalypse et les problèmes qui lui sont liés) mais aussi d'un paradoxe classique (le paradoxe de Goodman). Il s'avère en effet que ces différents problèmes présentent la propriété commune d'être basés sur une *analogie avec une urne* (*urn analogy*). Et l'analyse qui sera développée plus loin tend à montrer que chacun de ces problèmes peut être analysé à la lumière d'une méthodologie commune, fondée sur l'utilisation des n-univers.

Dans le chapitre 1, je m'attache à décrire l'approche méthodologique basée sur le modèle des n-univers. Je décris tout d'abord les principes généraux régissant les n-univers, en présentant également un certain nombre de n-univers remarquables. Enfin, je m'attache à cerner plusieurs difficultés relatives aux n-univers: en premier lieu, la notion d'instances et d'objets multiples; en second lieu, les critères à démultiplication; en troisième lieu, les relations un/plusieurs entre les objets et un critère donné.

Dans le chapitre 2, je présente l'argument de l'Apocalypse en soulignant également plusieurs points importants: la notion d'univers déterministe ou indéterministe, mais aussi le problème lié à la définition de la classe de référence.

Dans le chapitre 3, je développe une analyse détaillée du *modèle des deux urnes*, une expérience de pensée qui joue un rôle crucial dans l'étude de l'Argument de l'Apocalypse. J'y modélise également le *modèle des deux urnes* en termes de n-univers. Je présente en outre une modélisation du *modèle des deux urnes* dans d'autres n-univers que celui dans lequel prend place l'expérience classique, où on constate cependant un effet équivalent à celui du *modèle des deux urnes* classique.

Dans le chapitre 4, je m'attache à étudier la situation humaine correspondant à l'Argument de l'Apocalypse. Je traduis ensuite la situation correspondante en termes de n-univers. Puis je m'attache à modéliser cette situation dans d'autres n-univers où on observe toujours un effet de même nature.

Je décris dans le chapitre 5 deux expériences importantes qui ont été exposées par John Leslie: l'*expérience des émeraudes* et l'*expérience Petit-Bled/Londres*. Je m'attache ensuite à modéliser les situations correspondantes en termes de n-univers.

Je décris dans le chapitre 6 le *pile ou face divin*, une expérience qui est également en relation avec l'Argument de l'Apocalypse. Je modélise ensuite cette expérience en termes de n-univers. Enfin, je donne une analyse du problème posé par le *pile ou face divin*.

Dans le chapitre 7, je présente le *problème de la Belle au bois dormant*. Je modélise ensuite ce problème dans le modèle des n-univers. Enfin, je présente une analyse du *problème de la Belle au bois dormant*.

Dans le chapitre 8, je présente une étude de l'*expérience des segments de temps*. J'y décris la situation correspondante en termes de n-univers, en montrant également comment la problématique

correspondante peut également prendre place dans d'autres n-univers que celui correspondant à la situation qui résulte de l'expérience classique. J'analyse ensuite l'expérience des segments de temps afin de présenter une solution au problème qui lui est propre.

Je développe dans le chapitre 9 une analyse de l'Argument de l'Apocalypse, bâtie à partir des résultats des chapitres précédents. Je mets l'accent sur le problème lié aux hypothèses qui sous-tendent DA. Je présente également plusieurs expériences qui permettent d'éclairer le problème posé par l'Argument de l'Apocalypse et notamment le *modèle des deux urnes* à décalage, ainsi que l'expérience de l'ange gardien. Finalement, je présente une réponse au problème de la définition de la classe de référence (*reference class problem*), puis une solution à l'Argument de l'Apocalypse.

Dans le chapitre 10, je présente une étude du *paradoxe de la chambre d'exécution*. J'analyse le paradoxe à partir d'une modélisation de l'expérience dans le n-univers. Je m'attache également à présenter une modélisation de l'expérience dans d'autres n-univers. Enfin, je propose une solution pour le paradoxe de la chambre d'exécution.

Enfin, j'expose dans le chapitre 11 une solution au *paradoxe de Goodman*, également fondée sur la modélisation du problème dans le cadre des n-univers. Je modélise tout d'abord l'énoncé de Goodman dans le cadre des n-univers. Après avoir défini les conditions de l'induction, je présente une solution au paradoxe, basée sur la distinction entre deux modélisations de l'énoncé de Goodman dans des n-univers différents.

Avant toute chose, il convient de considérer ici l'œuvre accomplie par John Leslie. Dans cette œuvre, deux points saillants me paraissent émerger. En premier lieu, la tâche qui consiste à répondre point par point à chacun des multiples objections opposées à l'Argument de l'Apocalypse apparaît comme une tâche immense. En second lieu, Leslie présente un style d'argumentation remarquablement original, basé sur la mise en relation de situations concrètes avec des expériences de pensée de type probabiliste. Leslie possède une habileté exceptionnelle pour illustrer ses propos à l'aide d'analogies probabilistes saisissantes. Les arguments développés par Leslie tirent ainsi leur force d'un nombre important d'expériences de pensée convaincantes, basées notamment sur des analogies avec des urnes.

D'une manière générale, le fait d'utiliser des analogies avec des urnes et donc de mettre en adéquation des situations concrètes et des modèles probabilistes constitue une démarche rationnelle et pertinente. Plus encore, cette démarche paraît constituer un secteur de la philosophie qui ouvre la voie à nombre d'applications concrètes. Ces applications répondent à une forme d'argument qui est la suivante. Je me trouve dans telle ou telle situation concrète donnée qui est analogue à un modèle probabiliste donné, et je dois donc raisonner de la même manière que je le ferais avec ce modèle probabiliste. Il y a là une réelle utilité à pouvoir disposer, de manière fiable et sûre, de ce type de raisonnement. Dans ce contexte, la philosophie apparaît comme une discipline utile, susceptible d'aider à résoudre les problèmes qui se présentent dans un certain nombre de situations concrètes. De manière générale, ceci conduit à poser le problème de savoir dans quelle mesure une analogie avec une urne est fondée. Dans quelle mesure une analogie avec une urne est-elle applicable à une situation réelle? L'Argument de l'Apocalypse constitue ainsi une instance de ce type de problème, car il s'agit d'un argument basé sur une analogie avec une urne. Le problème posé par l'Argument de l'Apocalypse est bien celui de savoir si le modèle de l'urne est applicable ou non à la situation humaine correspondant à l'Argument de l'Apocalypse. Mais comment vérifier cela? Comment avoir la certitude que l'analogie avec l'urne est fondée dans le cas de l'Argument de l'Apocalypse? On rencontre là un problème méthodologique. Et ce problème se pose d'une manière générale. Comment vérifier qu'une analogie avec une urne est fondée? On rencontre là une lacune, un vide méthodologique. Selon quel(s) critère(s) peut-on décider si une analogie avec une urne est valable? L'Argument de l'Apocalypse a conduit à un certain nombre d'expériences de pensée, différentes de l'Argument de l'Apocalypse lui-même. Certaines de ces expériences de pensée ont été développées par Leslie, mais d'autres ont notamment été exposées par Nick Bostrom. Là encore, ces expériences sont basées sur la mise en relation de situations concrètes avec des modèles probabilistes, c'est-à-dire une analogie avec une urne. Ainsi, l'examen de la littérature relative à l'Argument de l'Apocalypse rend plus nécessaire que jamais le fait de s'attacher à évaluer les arguments basés sur une analogie avec une urne et à développer une méthode qui permette de déterminer s'ils sont valides ou non.

Intéressons-nous au problème général de la validité d'une analogie avec une urne. Telle analogie entre une situation concrète donnée et tel modèle probabiliste est-elle valable? Ce problème concerne directement des arguments philosophiques qui présentent une telle structure. Dans ce type d'argument, on commence par considérer un modèle probabiliste donné et la conclusion correspondante. Puis¹ on considère qu'une situation concrète donnée est analogue à ce dernier modèle probabiliste. Compte tenu de l'analogie existant entre les deux situations, on en déduit que la conclusion qui prévaut dans le modèle probabiliste est également applicable à la situation concrète correspondante. La structure ce type d'arguments est ainsi la suivante:

- (1) dans le modèle probabiliste M , il s'ensuit la conclusion $C1(M)$
- (A2) la situation concrète m est analogue au modèle probabiliste M
- (3) \therefore dans la situation concrète m , il s'ensuit la conclusion $C1(m)$

Le problème qui émerge dans le cas d'une analogie avec une urne conduit à poser directement la question suivante: dans quelle mesure des arguments présentant une telle structure sont-ils valides? Dans ce type d'argument, il apparaît tout d'abord que la prémisse probabiliste peut être aisément vérifiée. Une telle prémisse fait habituellement l'objet d'un consensus. Aussi, ce n'est pas la prémisse probabiliste (1) qui est en général à mettre en doute. De même, il apparaît que la structure de la conclusion (3) constitue un élément solide de ce type d'argument. Si en effet la prémisse probabiliste et l'étape d'analogie sont vraies, alors la conclusion est nécessairement vraie. Dans ce type d'argument, il apparaît ainsi que l'étape douteuse est l'étape d'analogie (A2). L'étape du raisonnement qui consiste à déclarer qu'une situation concrète donnée est analogue au modèle probabiliste considéré, nécessite ainsi d'être prouvée. En conséquence, il convient de pouvoir disposer d'un moyen permettant de vérifier la valeur de vérité de l'étape d'analogie dans les arguments basés sur une analogie avec une urne. Dans ce qui suit, je me propose de présenter un tel outil.

¹ Il s'agit de l'étape d'*analogie*.

Chapitre 1. Approche méthodologique: les n-univers

1.1. Méthodologie

Alors que les paradoxes philosophiques contemporains font l'objet d'analyses et de tentatives de résolution répétées et multiples, on peut remarquer que l'approche méthodologique employée pour résoudre ces paradoxes est rarement discutée. Il est peu fréquent en effet que la méthodologie utilisée dans le but de résoudre les paradoxes ou les arguments philosophiques dans leur ensemble fasse l'objet d'une discussion. Habituellement, la méthode utilisée pour résoudre les paradoxes est basée sur l'application des principes généraux qui sont ceux de la philosophie analytique: clarté, rigueur, précision. L'application de ces principes constitue en soi une méthodologie. Cependant, il s'agit là davantage de la mise en oeuvre de principes généraux que d'une méthodologie détaillée et précise, spécialisée et destinée à la résolution de tel ou tel type de problème philosophique. On peut penser ainsi qu'il pourrait être utile de disposer d'une approche méthodologique plus précise, adaptée notamment au traitement de certains types de problèmes. Le présent chapitre se propose de décrire une telle approche méthodologique, destinée en particulier à clarifier les données résultant des problèmes présentant la structure d'une analogie avec une urne. Les chapitres suivants constitueront l'application de cette méthode à la recherche d'une solution concernant plusieurs paradoxes ou problèmes philosophiques contemporains: en premier lieu, l'*argument de l'apocalypse* ainsi que plusieurs problèmes qui lui sont liés tels que le *problème de la Belle au bois dormant* et le *paradoxe de la chambre d'exécution*, et en second lieu, le paradoxe de Goodman.

La classe de paradoxes ou d'arguments philosophiques concernés par la présente approche méthodologique présentent la structure d'une analogie avec une urne. Bien souvent, avec ce type d'arguments, le fond du problème consiste à déterminer si la situation réelle correspondant à l'exposé du paradoxe est bien en adéquation avec le modèle probabiliste correspondant. Ainsi, on peut penser que la modélisation de la situation réelle en termes précis constitue ici une exigence fondamentale. Dans ce but, j'introduirai ici un outil méthodologique qui permet de modéliser les situations correspondant aux énoncés de plusieurs problèmes philosophiques de manière précise, permettant ainsi de dissiper souvent l'ambiguïté inhérente à l'énoncé. Dans ce contexte, on peut observer que les expériences physiques réalisées dans les sciences expérimentales font l'objet de la description d'un protocole précis et rigoureux. On peut penser également que les expériences de pensée utilisées en philosophie doivent faire l'objet d'un protocole tout aussi précis. Dans ce but, je présenterai ainsi l'outil méthodologique que constituent les n-univers. Le principe méthodologique qui y est associé est le suivant: il s'agit de modéliser en termes de n -univers la situation réelle correspondant à l'énoncé du problème philosophique considéré. Il s'avère en effet souvent que les problèmes philosophiques étudiés comportent une ambiguïté dans la définition des conditions qui sont celles dans lesquelles les expériences présentées qui les sous-tendent prennent place. Et le modèle des n -univers a précisément pour finalité de dissiper cette ambiguïté. Une fois les expériences de pensée concernées décrites avec précision dans un n -univers dont les propriétés sont clairement définies, l'ambiguïté disparaît, permettant ainsi une clarification de la situation considérée. A cette fin, il convient tout d'abord de décrire les caractéristiques du n -univers correspondant à l'énoncé dans lequel le problème philosophique prend place.

Un second principe méthodologique qui accompagne la modélisation du problème considéré dans le n -univers correspondant est le suivant: une fois la situation correspondant au problème posé décrite en termes de n -univers, il convient de s'attacher à modéliser cette situation dans le n -univers le plus simple possible. Pour ce faire, on s'attache à choisir le n -univers le plus *simple* correspondant à la situation donnée.

Enfin, un troisième principe méthodologique que l'on peut mentionner est le suivant. Une fois la situation correspondant au problème posé modélisée dans un n -univers donné, il peut souvent se révéler utile de considérer des variations de la situation correspondant au problème dans d'autres n -univers, dont la structure se révèle légèrement différente. Il apparaît ainsi utile de décrire éventuellement des versions équivalentes de la situation que l'on cherche à modéliser, dans des n -univers structurellement différents. L'intérêt de cette approche est de distinguer différentes situations, afin de faciliter d'éventuelles comparaisons. Ceci a trait aux expériences de pensée et à la

méthodologie utilisée. En général, une expérience de pensée consiste dans le fait d'imaginer une situation la plus souvent réelle et d'en tirer des conséquences. Souvent, on se pose la question de savoir si l'expérience de pensée considérée peut être transcrite par d'autres situations réelles, tout en en préservant la teneur et la portée. Une des techniques utilisées dans la présente dissertation permet précisément d'apporter une réponse à de telles questions. La méthodologie utilisée est ainsi la suivante. On considère tout d'abord une expérience de pensée E_1 qui prend place dans une situation réelle m . On modélise ensuite E_1 dans un n -univers donné, qui est défini par un certain nombre de critères. On remplace alors l'un de ces critères par un critère différent et on obtient alors un second n -univers. On teste alors si l'expérience de pensée E_2 qui prend désormais place dans le second n -univers préserve l'impact et la portée de l'expérience originale modélisée dans le premier n -univers. Si tel est le cas, il est alors intéressant de s'intéresser à la modélisation de la situation m dans le second n -univers. Cette technique sera largement utilisée tout au long de la présente dissertation.

1.2. Définition

Considérons² la loi (L1) selon laquelle "le diamant raye les autres solides". A priori, (L1) s'impose à nous comme une vérité incontestable. Pourtant, il s'avère qu'à une température supérieure à 3550°C, le diamant fond. Aussi en dernière analyse, la loi (L1) se vérifie-t-elle à une température normale et en tout état de cause, lorsque la température est inférieure à 3550°C. Mais une telle loi ne s'applique pas au-delà de 3550°C. Ceci illustre combien l'énoncé des conditions dans lesquelles la loi (L1) est vérifiée est important, notamment en ce qui concerne les conditions de température. Ainsi, lorsqu'on énonce (L1), s'avère-t-il nécessaire de préciser les conditions de température dans lesquelles elle trouve à s'appliquer. Ceci revient à décrire le type d'univers dans lequel la loi est vérifiée.

Soit également la proposition (P1) suivante: "le volume de l'univers visible est supérieur à 1000 fois celui du système solaire". Une telle proposition s'impose à nous comme évidente. Mais là aussi, il apparaît que (P1) est vérifiée à l'époque moderne, mais qu'elle se révèle fautive dans les premiers instants de l'univers. En effet, lorsque l'âge de notre univers était de 10^{-6} seconde après le big bang, son volume était à peu près égal à celui de notre système solaire. Ici également, il apparaît donc nécessaire de spécifier, en même temps que la proposition (P1) les conditions de l'univers dans lequel elle s'applique. Une formulation non ambiguë de (P1) comporte donc une clause temporelle plus restrictive, telle que : "à notre époque, le volume de l'univers visible est supérieur à 1000 fois celui du système solaire". Ainsi, d'une manière générale, on peut penser que lorsqu'on énonce une généralisation, il est nécessaire de préciser les conditions de l'univers dans lequel celle-ci s'applique. La description précise de l'*univers de référence* est fondamentale, car selon les conditions de l'univers dans lequel on se place, la loi énoncée peut se révéler vraie ou fautive.

On observe dans notre univers la présence à la fois de constantes et de variables. On a ainsi des constantes, qui constituent les constantes fondamentales de l'univers : la vitesse de la lumière : $c = 2,998 \times 10^8$ m/s; la constante de Planck : $h = 6,626 \times 10^{-34}$ J.s; la charge de l'électron : $e = 1,602 \times 10^{-19}$ C; etc. On a d'autre part des variables. Parmi celles-ci, on peut citer notamment : la température, la pression, l'altitude, la localisation, le temps, la présence d'un rayonnement laser, la présence d'atomes de titanium, etc.

On a souvent tendance, lorsqu'on énonce une généralisation, à ne pas prendre en compte les constantes et les variables qui sont celles de notre univers envisagé dans sa totalité. Tel est le cas par exemple lorsqu'on considère la situation de notre univers le 1er janvier de l'an 2000, à 0h. On se place alors explicitement dans ce qui constitue une tranche, une coupe de notre univers. En effet, le temps n'est pas considéré alors comme une variable, mais bien comme une constante. Soit également la généralisation: "les dinosaures avaient le sang chaud"³. Ici, on se place explicitement dans un sous-univers du notre où les paramètres du temps et de l'espace ont une portée restreinte. La variable temporelle se réduit à l'époque particulière de l'histoire de la Terre qui a connu l'apparition des dinosaures : le Trias et le Crétacé. Et de même, le paramètre spatial se limite à notre planète : la Terre. De manière identique, les conditions de température sont changeantes au sein de notre univers, selon

² Les paragraphes qui suivent sont essentiellement tirés de mon *Une Solution pour le Paradoxe de Goodman* (2001).

³ Cette affirmation est controversée.

que l'on se situe à un emplacement ou à un autre de ce dernier : à l'Equateur terrestre, à la surface de Pluton, au cœur d'Alpha du Centaure, etc. Mais si l'on s'intéresse exclusivement au ballon servant à l'expérimentation au sein du laboratoire de physique, où la température est maintenue invariablement à 12°C, on peut considérer alors la température comme une constante. Car lorsqu'on exprime de telles généralisations, on se place non pas dans notre univers envisagé dans sa totalité, mais seulement dans ce qui constitue véritablement une partie spécifique, une restriction de ce dernier. On peut alors assimiler l'univers de référence dans lequel on se place à un sous-univers du notre. Il est ainsi fréquent d'exprimer des généralisations qui ne valent que pour l'époque présente ou pour nos conditions terrestres habituelles. Explicitement ou non, l'énoncé d'une loi comporte un univers de référence. Mais dans la plupart des cas, les variables et les constantes du sous-univers considéré sont distinctes de celles permettant de décrire notre univers envisagé dans sa totalité. Car les conditions sont extrêmement variées au sein de notre univers : les conditions sont très différentes selon que l'on se place à la 1ère seconde après le big bang, sur Terre à l'époque précambrienne, sur notre planète en l'an 2000, à l'intérieur de l'accélérateur de particules du CERN, au cœur de notre Soleil, à proximité d'une naine blanche, ou bien à l'intérieur d'un trou noir, etc.

On peut penser également qu'il est intéressant de pouvoir modéliser des univers dont même les constantes sont différentes des constantes fondamentales de notre univers. On peut ainsi souhaiter étudier par exemple un univers où la masse de l'électron est égale à $9,325 \times 10^{-31}$ kg, ou bien un univers où la charge de l'électron est égale à $1,598 \times 10^{-19}$ C. Et de fait, les univers-jouets, qui prennent en compte des constantes fondamentales différentes de celles de notre univers familier, sont étudiés par les astrophysiciens.

Enfin, lorsqu'on décrit les conditions d'une expérience de pensée, on se place, de manière explicite ou non, dans les conditions qui s'apparentent à celles d'un sous-univers. Lorsqu'on considère par exemple 100 boules extraites d'une urne durant 100 jours consécutifs, on se place alors dans une restriction de notre univers où la variable temporelle est limitée à une période de 100 jours et, où la localisation spatiale est extrêmement réduite, correspondant par exemple à un volume à peu près égal à 5 dm^3 . Par contre, le nombre d'atomes de zirconium ou de titane éventuellement présents dans l'urne, l'existence éventuelle d'un rayonnement laser, la présence ou l'absence d'une source sonore de 10 db, etc. peuvent être omis et ignorés. Dans ce contexte, il n'est pas nécessaire de prendre en compte l'existence de telles variables. Dans cette situation, il suffit de mentionner les variables et les constantes *effectivement* utilisées dans l'expérience de pensée. Car on peut penser en effet que le nombre de variables dans notre univers est si grand qu'il est impossible de les énumérer toutes. Et dès lors, il ne paraît pas possible de caractériser notre univers en fonction de toutes ses variables, car on peut en fournir une énumération infinie. Il apparaît suffisant de décrire le sous-univers considéré, en mentionnant uniquement les constantes et les variables qui jouent un rôle effectif dans l'expérience. Ainsi, dans de telles situations, on décrira le sous-univers considéré en ne mentionnant que les critères effectifs nécessaires à la description de l'expérience.

Ce qui précède incite à penser que d'une manière générale, afin de modéliser le contexte dans lequel prennent place des problèmes tels que l'Argument de l'Apocalypse ou bien le paradoxe de Goodman, il est opportun de décrire un univers donné en termes de variables et de constantes. On est amené ainsi à définir un n -univers ($n \geq 0$) comme un univers dont les critères comportent m constantes, et n variables, où les m constantes et les n variables constituent les *critères* de l'univers considéré. Dans ce cadre particulier, on définit un *1-univers temporel* (ΩT) comme un univers comportant un seul critère-variable : le temps. De même, on définit un *1-univers coloré* (ΩC) comme un univers comportant un seul critère-variable : la couleur. On définira aussi un *2-univers coloré et temporel* (ΩCT) comme un univers comportant deux critères-variables : le temps et la couleur. Etc. De même, un univers où tous les objets sont rouges, mais se caractérisent par une localisation différente sera modélisé par un *1-univers localisé* (ΩL) dont la couleur est un critère-constante (rouge).

On notera incidemment que le modèle à n -univers permet notamment de modéliser plusieurs situations intéressantes. Ainsi, un univers *temporel* peut être considéré comme un n -univers dont l'une des variables est un critère temporel. De plus, un univers où on considère un moment unique T_0 , dépourvu du phénomène de succession du temps, peut être considéré comme un n -univers dont le temps ne constitue pas une des variables, mais où il existe une constante-temps. De même, un univers *atemporel* correspond à un n -univers dont aucune variable ne correspond à un critère temporel, et où il n'existe aucune constante-temps.

1.2.1. Démultiplication

Il est intéressant de mettre l'accent sur la propriété de *démultiplication* d'un objet donné par rapport à un critère-variable du n-univers considéré. Lorsqu'un critère-variable χ d'un n-univers est démultiplié, il sera dénoté par χ^* . N'importe quel critère-variable d'un n-univers peut ainsi être démultiplié. Le fait pour un objet donné d'être démultiplié par rapport à un critère χ est la propriété pour cet objet de présenter plusieurs taxons du critère χ .

Prenons tout d'abord l'exemple du critère du temps. Le fait pour un objet donné d'être démultiplié par rapport au temps est le fait de se trouver à plusieurs positions temporelles. Dans notre univers, un objet o peut se trouver à *plusieurs* positions temporelles (successives) à la fois. Ainsi, dans notre univers dont l'un des critères-variables est celui du temps, il est usuel de constater qu'un objet o qui existe en T_1 existe également en T_2, \dots, T_n . Un tel objet a une durée de vie qui s'établit sur la période T_1-T_n . Ainsi, o est démultiplié par rapport au temps en T_1, \dots, T_n . Un tel n-univers présente la structure ΩT^* . De manière plus précise, on peut observer que nos objets usuels possèdent une propriété de *persistance* temporelle, qui constitue un cas particulier de démultiplication temporelle.

Intéressons-nous maintenant à la démultiplication par rapport au critère du lieu. Qu'est-ce pour un objet donné o que d'être démultiplié par rapport à la position locale? Une telle propriété caractérise un n-univers où un objet donné o peut se trouver au même moment à plusieurs positions spatiales à la fois. Un tel univers est ubiquiste. Il présente la structure ΩL^* . Un tel univers ne correspond pas à notre univers réel. Le modèle des n-univers permet également de modéliser des expériences de pensée concernant d'autres univers que le notre et notamment des n-univers à démultiplication locale⁴. Cependant, de tels n-univers ne seront pas décrits ni étudiés ici, car dans l'objet de la présente dissertation se limite à l'étude des n-univers qui permettent de modéliser des situations réelles.

1.2.2. Relation un ou plusieurs des objets multiples avec un critère donné

A ce stade, il convient d'opérer une distinction importante. Il est en effet nécessaire de distinguer deux types de situations. *Un* objet peut en effet se trouver à *plusieurs* taxons d'un critère-variable donné. Ceci correspond au cas qui vient d'être décrit de la démultiplication par rapport à un critère-variable donné. Si l'on prend l'exemple du critère-variable du temps, on a alors un objet qui occupe plusieurs positions spatiales (démultiplication par rapport au critère du lieu). Dans ce cas, cet objet peut se trouver au même moment par exemple en L_1, L_2, L_3 et L_4 . Un tel univers est ubiquiste et ne correspond pas à la situation observée dans notre univers réel. De même, dans le cas où un objet occupe plusieurs positions temporelles (démultiplication par rapport au critère du temps), cet objet peut exister par exemple aux positions temporelles T_1, T_2, T_3, T_4 .

Mais il est important de considérer un autre type de situation. En effet, *plusieurs* objets peuvent se trouver à *un* même taxon d'un critère⁵ donné. Considérons tout d'abord le critère temporel. Plaçons-nous, à titre d'exemple, dans un n-univers temporel à objets multiples, un $\Omega \alpha T$. Ici, on a globalement deux types de n-univers. Dans le premier type de n-univers, il existe un seul objet par position temporelle. A un moment donné, on ne peut donc avoir qu'un objet unique. On considère dans ce cas que chaque objet de ce n-univers est en relation *un* avec les taxons de temps. On dénote par $\Omega \alpha T \dots$ un tel n-univers. Considérons à l'inverse un n-univers temporel à objets multiples où plusieurs objets o_1, o_2, o_3 peuvent exister au même moment, à une même position temporelle donnée T_0 . Dans ce cas, la situation est fondamentalement différente du $\Omega \alpha T$, car plusieurs objets peuvent se trouver à une même position temporelle donnée. En d'autres termes, les objets peuvent coexister à un moment donné. Dans ce cas, on considère que les objets sont en relation *plusieurs* avec les taxons temporels. On dénote alors par $\Omega \alpha^* T \dots$ un tel n-univers. On notera ici qu'une telle propriété d'un n-univers correspond aux caractéristiques de notre univers réel.

Considérons également le critère du lieu. Compte tenu de ce qui précède, on est amené à distinguer deux types de n-univers. Dans le premier type de n-univers, *un* seul objet peut se trouver à *un* même

⁴ Il est à noter que l'on peut construire de même des n-univers à démultiplication par couleur, où les objets présentent à un même moment et en un même lieu plusieurs couleurs. De tels univers présentent la structure ΩC^* . Ces n-univers à démultiplication par couleur, de même que les n-univers à démultiplication locale, ne correspondent pas à notre univers réel.

⁵ Il peut s'agir d'un critère-variable ou d'un critère-constante.

taxon d'un critère de lieu. Il existe alors un seul objet par position spatiale. Ceci modélise par exemple la situation qui est celle des pièces d'un jeu d'échecs. On dénote par $\Omega\alpha L\dots$ un tel n-univers. Dans ce cas, les objets sont en relation *un* avec le critère spatial. A l'inverse, dans le second type de n-univers, *plusieurs* objets peuvent se trouver à *un* même taxon d'un critère de lieu. Ainsi, par exemple, les objets o_1, o_2, o_3 se trouvent en L_1 . Une telle situation correspond par exemple à une urne⁶ où se trouvent plusieurs boules. On dénote par $\Omega\alpha*L\dots$ un tel n-univers, où les objets sont en relation *plusieurs* avec les taxons de localisation.

Considérons enfin la relation des objets multiples par rapport à un critère-variable de couleur. On est alors conduit à distinguer deux types de situations. Dans le premier cas, les objets sont en relation *un* avec le critère de couleur si un taxon de couleur donné ne peut être possédé que par un seul objet. Dans ce type de n-univers, aucun objet ne possède la même couleur. On dénote $\Omega\alpha C\dots$ ce type de n-univers. A l'inverse, dans le second type de n-univers, les objets sont en relation *plusieurs* avec le critère de couleur si un taxon de couleur donné peut être possédé par plusieurs objets. Dans ce type de n-univers, plusieurs objets peuvent présenter la même couleur. Et on dénote $\Omega\alpha*C\dots$ ce type de n-univers. Ici, il apparaît que notre univers réel est modélisé par un $\Omega\alpha*C\dots$, car plusieurs objets peuvent y présenter une couleur identique.

Enfin, il convient de préciser un point important qui concerne la notation utilisée. Ceci a trait au cas où le n-univers considéré comporte des objets multiples et un critère-constante de temps ou/et un critère-constante de localisation. Considérons tout d'abord le cas où on se situe dans un n-univers comportant de multiples objets et un critère-constante de temps. Dans ce cas, il existe donc nécessairement de multiples objets multiples en T_0 . Par conséquent, dans le n-univers considéré, les objets multiples sont donc en relation *plusieurs* avec le critère-constante de temps. On se place donc dans la situation qui est celle d'un $\Omega\alpha*T_0\dots$. Considérons maintenant le cas où le n-univers considéré comporte un critère-constante de localisation. Dans cette hypothèse, il existe nécessairement plusieurs objets à la position spatiale L_0 . En conséquence, les objets multiples se trouvent en relation *plusieurs* avec le critère-constante de localisation, dans le n-univers considéré. On se situe donc dans un $\Omega\alpha*T_0\dots$.

Les remarques qui précèdent suggèrent ainsi une simplification en matière de notation utilisée. En effet, dès lors qu'un n-univers comporte des objets multiples et qu'il comporte un critère-constante de temps, les objets multiples sont nécessairement en relation *plusieurs* avec le critère-constante de temps. Le n-univers est ainsi un $\Omega\alpha*T_0\dots$. Mais on peut donc simplifier la notation correspondante en $\Omega\alpha\dots$. De même, si un n-univers comporte des objets multiples et un critère-constante de localisation, les objets multiples se trouvent nécessairement en relation *plusieurs* avec le critère-constante de localisation. Le n-univers concerné est donc un $\Omega\alpha*L_0\dots$. Et on peut ainsi simplifier la notation du n-univers considéré en $\Omega\alpha\dots$. Par conséquent, on peut simplifier les notations $\Omega\alpha*L_0\alpha*T_0$ par $\Omega\alpha$, $\Omega\alpha*L_0\alpha T$ par $\Omega\alpha T$, $\Omega\alpha*L_0\alpha*T$ par $\Omega\alpha*T$, $\Omega\alpha*L_0\alpha*T*$ par $\Omega\alpha*T*$, etc.

1.2.3. Sous-univers

A ce stade, il convient de s'attacher à décrire la notion de *sous-univers*, qui se révélera également utile par la suite. Commençons par en donner une définition informelle. Un n-univers Ω_2 est un sous-univers d'un autre n-univers Ω_1 s'il présente les mêmes critères-variables en nombre inférieur ou les mêmes critères sous forme simplifiée. Ainsi, un sous-univers d'un n-univers comportant n critères est un n-univers comportant un nombre de critères inférieur à n . De plus, la forme simplifiée d'un critère-variable à démultiplication est le même critère-variable sans démultiplication. Une autre forme simplifiée d'un critère-variable concerne les n-univers à objets multiples. Soit un n-univers où les objets multiples sont en relation *plusieurs* avec un autre critère-variable donné. La forme simplifiée de ce n-univers comporte alors les mêmes critères, mais les objets multiples se trouvent alors en relation *un* avec le même critère-variable.

De manière plus formelle, on a la définition suivante d'un sous-univers. Un n-univers Ω_2 est un sous-univers d'un autre n-univers Ω_1 s'il satisfait au moins l'une des trois conditions suivantes:

- (a) le n-univers Ω_2 présente au moins un critère-variable de moins que le n-univers Ω_1 . A titre d'exemple, un ΩCT est un sous-univers d'un ΩCTL . De même, un ΩT est également un sous-univers

⁶ L'urne est assimilée ici à une position spatiale unique donnée.

d'un ΩCTL . On notera ici qu'un cas particulier qui correspond à cette condition est que le critère-variable concerné du n-univers Ω_1 est constitué par la propriété de posséder de multiples objets, alors que Ω_2 ne présente pas un tel critère. Ainsi, un ΩCT est un sous-univers d'un $\Omega\alpha\text{CT}$.

- (b) si l'un au moins des critères-variables de Ω_1 est un critère à démultiplication alors le même critère dans Ω_2 n'est pas à démultiplication. Ainsi, par exemple, un ΩCT est un sous-univers d'un ΩCT^* .
- (c) si Ω_1 est un n-univers à objets multiples où les objets multiples sont en relation *plusieurs* avec un critère-variable donné, alors au niveau de Ω_2 les objets multiples se trouvent en relation *un* avec ce critère-variable. Soit l'exemple suivant: un $\Omega\alpha\text{T}^*\text{C}$ est un sous-univers d'un $\Omega\alpha^*\text{T}^*\text{C}$. De même, un $\Omega\alpha\text{LT}^*$ constitue un sous-univers d'un $\Omega\alpha^*\text{LT}^*$.

Formellement, il en résulte la définition suivante. Un n-univers Ω_1 dont les critères-variables sont $\Omega a_1 a_2 a_3 \dots a_n$ est un sous-univers d'un n-univers Ω_2 dont les critères-variables sont $\Omega a_1 a_2 a_3 \dots a_m$ si l'une au moins des trois conditions suivantes est vérifiée:

- (a) $m < n$
- (b) $m = n$ et $\exists i (1 \leq i \leq n)$ tel que $a_i(\Omega_1) = a_i^*$ et $a_i(\Omega_2) = a_i$
- (c) $m = n$ et $\exists i (1 \leq i \leq n)$ tel que $a_i(\Omega_1) = \alpha^*$ et $a_i(\Omega_2) = \alpha$

1.2.4. n-univers spécialisés

Le présent chapitre n'a pas pour objet l'étude des n-univers *per se*. Il vise en effet à décrire de manière précise certains n-univers qui permettent de modéliser les problèmes qui seront étudiés plus loin, à savoir *l'argument de l'apocalypse* et plusieurs problèmes qui lui sont liés, ainsi que le paradoxe de Goodman. Ainsi, les n-univers étudiés ici sont ceux qui permettent de modéliser des situations concrètes. Par conséquent, deux catégories de n-univers ne seront pas évoquées ici. Il s'agit d'une part des n-univers de nature concrète dont les propriétés sont fondamentalement différentes du notre. Car le modèle des n-univers permet également de décrire des n-univers concrets dont les propriétés ne sont pas celles du notre, tels que des univers ubiquistes, où les objets peuvent se trouver simultanément à plusieurs positions spatiales, etc. De même, le modèle des n-univers permet de décrire des univers abstraits. Mais étant donné que la modélisation des objets abstraits ne présente pas d'intérêt dans le cadre de la présente étude, ils ne seront pas mentionnés non plus. Ainsi, le cas des univers concrets aux propriétés différentes du notre ainsi que celui des univers abstraits ne seront pas évoqués ici. Car la présente étude se propose de montrer comment le modèle des n-univers peut se révéler utile pour la clarification des problèmes posés par certains problèmes philosophiques, essentiellement basés sur une analogie avec une urne. Mais ces derniers étant basés sur la description d'une situation concrète, on ne s'intéressera donc ici qu'aux n-univers concrets dont les propriétés correspondent à celles de notre univers réel. Leur étude se révèle suffisante pour les besoins de la présente étude. Compte tenu du fait qu'ils permettent de modéliser les situations réelles, les n-univers qui seront étudiés ici comportent un critère temporel et un critère de localisation⁷. Ainsi de tels n-univers comportent à la fois une constante ou une variable temporelle, ainsi qu'une constante ou une variable spatiale. Ces n-univers modélisant des situations réelles appartiennent donc à quatre catégories différentes, selon qu'ils présentent: (a) une constante temporelle et une constante spatiale; (b) une constante temporelle et une variable spatiale; (c) une variable temporelle et une constante spatiale; (d) une variable temporelle et une variable spatiale. Ils appartiennent donc à l'un des quatre types suivants: $\Omega L_0 T_0 \dots$, $\Omega L T_0 \dots$, $\Omega L_0 T \dots$, $\Omega L T \dots$.

1.3. n-univers caractéristiques

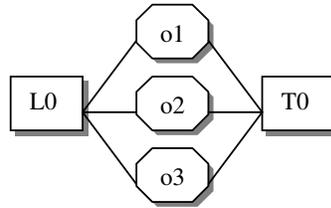
Il convient maintenant de s'attacher à décrire un certain nombre de n-univers caractéristiques, correspondant à des situations concrètes courantes. Ces n-univers serviront de cadre pour les expériences qui seront décrites plus loin.

1.3.1. Le n-univers alocalisé et atemporel à objets multiples: $\Omega\alpha$

Il s'agit d'un n-univers comportant de multiples objets: o_1, \dots, o_n et présentant une constante de localisation (un lieu unique L_0) ainsi qu'une constante temporelle (une position temporelle unique T_0).

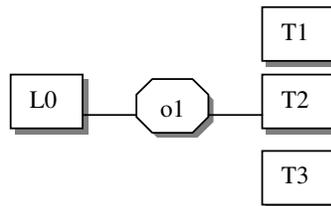
⁷ A l'inverse, les n-univers abstraits ne comportent ni critère temporel ni critère de localisation.

Dans le $\Omega\alpha$ (en notation étendue $\Omega\alpha*L_0\alpha*T_0$)⁸ les n objets existent en L_0 et en T_0 . Dans un tel n-univers, les objets sont nécessairement en relation *plusieurs* avec les constantes de localisation et de temps. Un tel univers correspond par exemple à une urne contenant plusieurs boules en un lieu unique et à un temps unique. Et le nombre de boules n'est pas susceptible de changer en fonction du temps et du lieu⁹.



1.3.2. Le n -univers alocalisé et temporel à objet unique: ΩT

Il s'agit d'un n-univers comportant un objet unique et présentant une constante de localisation (un lieu unique L_0) ainsi qu'une variable temporelle. Un tel n-univers comporte au moins une ou plusieurs positions temporelles: T_1, T_2, \dots, T_n . Dans le ΩT (en notation étendue $\Omega L_0 T$), l'objet unique n'est pas persistant par rapport au temps: il existe en T_1 , mais n'existe pas aux autres positions temporelles.

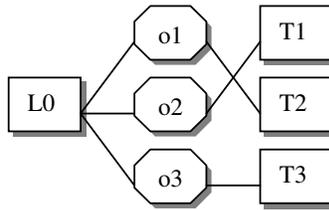


1.3.3. Le n -univers alocalisé et temporel à objets multiples: $\Omega\alpha T$

Il s'agit d'un n-univers comportant de multiples objets: o_1, \dots, o_n et comportant une constante de localisation (un lieu unique L_0) ainsi qu'une variable temporelle et donc plusieurs positions temporelles: T_1, T_2, \dots, T_n . Dans le $\Omega\alpha T$ (en notation étendue $\Omega\alpha*L_0\alpha T$), un objet donné n'est pas persistant par rapport au temps: il existe en L_0 à une position temporelle donnée, mais n'existe pas aux autres positions temporelles. Il est à noter que dans ce type de n-univers, plusieurs objets peuvent se trouver en L_0 à un moment donné. Ainsi, les objets sont en relation *plusieurs* avec la constante de localisation. A l'inverse toutefois, plusieurs objets ne peuvent se trouver en L_0 à une position temporelle donnée. Les objets se trouvent donc en relation *un* avec le critère-variable temporel.

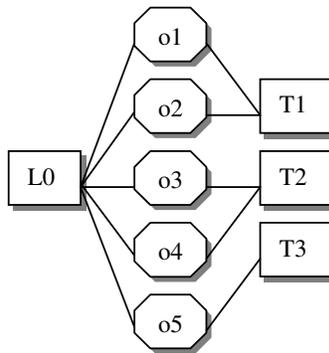
⁸ Ici, on pourrait noter plus simplement: $\Omega L_0\alpha*T_0$. En effet, dès lors que de multiples objets se existent dans un n-univers comportant une constante de localisation, ils se trouvent nécessairement en relation *plusieurs* avec la constante spatiale. De manière générale, il apparaît que tout n-univers à objets multiples comportant un critère-constante de localisation présente la structure $\Omega\alpha*L_0$ et peut donc être dénoté plus simplement $\Omega\alpha$. De même, tout n-univers à objets multiples comportant un critère-constante de temps présente la structure $\Omega\alpha*T_0$ peut être dénoté de manière plus simple $\Omega\alpha$. Ainsi, de manière générale, il suffit de mentionner qu'un n-univers est à objets multiples dès lors qu'il comporte un critère-constante spatial ou temporel.

⁹ Compte tenu du fait que cet univers est immuable, et que le nombre d'objets n'est pas susceptible de changer en fonction du lieu ou du temps, un tel univers pourrait sembler a priori inintéressant. Tel n'est cependant pas le cas. En effet, un tel univers s'avère intéressant pour modéliser des situations où on ne connaît pas, par exemple, le nombre d'objets n qui y sont contenus et où on fait des hypothèses sur ce dernier nombre.



1.3.4. Le n -univers alocalisé et temporel à objets multiples*: $\Omega\alpha^*T$

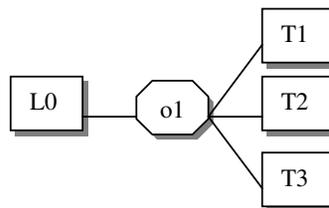
Il s'agit du même type de n -univers que le précédent, avec cette différence que plusieurs objets peuvent se trouver en L_0 à une position temporelle donnée. Dans le $\Omega\alpha^*T$ (en notation étendue $\Omega\alpha^*L_0\alpha^*T$), les objets sont ainsi en relation *plusieurs* avec le critère-variable temporel.



1.3.5. Le n -univers alocalisé et temporel* à objet unique: ΩT^*

Il s'agit d'un n -univers comportant un objet unique et présentant une constante de localisation ainsi qu'une variable temporelle. Le ΩT^* (en notation étendue $\Omega L_0 T^*$) comprend donc un lieu unique L_0 et plusieurs positions temporelles: T_1, T_2, \dots, T_n . L'objet est persistant par rapport au temps: s'il existe en T_5 , c'est qu'il a existé en T_1, T_2, T_3, T_4 . En termes de modèle probabiliste, un tel univers correspond par exemple à une urne contenant une seule boule en un lieu unique. La boule est susceptible d'exister de T_i à T_j , avec $i \leq j$. Ainsi la boule possède une durée d'existence variable Δ , telle que $\Delta = (j - i) + 1$.

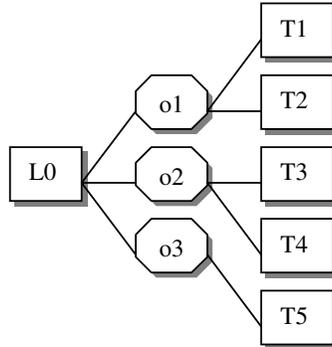
On peut noter qu'une autre façon d'appréhender un tel n -univers est de considérer que ce n -univers correspond à la durée d'existence de la boule. Un tel univers n'existe que pendant la durée d'existence de l'unique objet qu'y s'y trouve. On considère alors l'univers qui correspond à la durée pendant laquelle la boule existe. Ainsi, le n -univers correspond comment en T_i et finit en T_j . On observera également que dans ce type de n -univers, des notions telles que "avant T_i " ou "après T_j " ne sont pas définies. Elles ne possèdent donc pas de sens dans le $\Omega L_0 T^*$.



1.3.6. Le n -univers alocalisé et temporel* à objets multiples: $\Omega\alpha T^*$

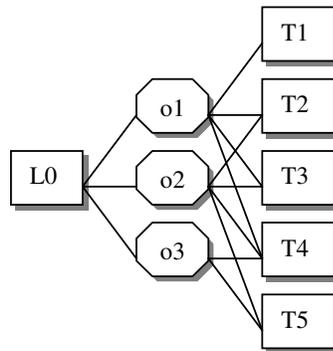
Il s'agit d'un n -univers comportant de multiples objets o_1, \dots, o_n et présentant une constante de localisation ainsi qu'une variable temporelle. Un tel n -univers comporte donc un lieu unique L_0 et n

positions temporelles T_1, T_2, \dots, T_n . Ici, chaque objet est persistant par rapport au temps: s'il existe en T_5 , c'est qu'il a existé en T_1, T_2, T_3, T_4 . La variable temporelle est donc à démultiplication. De plus, on peut avoir plusieurs objets à la position unique L_0 . Dans le $\Omega\alpha T^*$ (en notation étendue $\Omega\alpha^*L_0\alpha^*T^*$), les objets sont ainsi en relation *plusieurs* avec le critère-constante de localisation. Enfin, il existe un seul objet en L_0 à une position temporelle donnée: les objets sont donc en relation *un* avec le critère temporel.



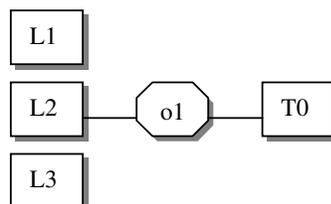
1.3.7. Le n -univers alocalisé et temporel* à objets multiples*: $\Omega\alpha^*T^*$

Il s'agit du même n -univers que le précédent, à cette différence que plusieurs objets peuvent se trouver en L_0 à une position temporelle donnée dans le $\Omega\alpha^*T^*$ (en notation étendue $\Omega\alpha^*L_0\alpha^*T^*$). Les objets sont donc en relation *plusieurs* avec le critère-constante de localisation. Le $\Omega\alpha^*L_0\alpha^*T^*$ correspond par exemple à une urne contenant plusieurs boules o_1, \dots, o_n en un lieu unique L_0 qui sont telles que chaque boule existe de T_i à T_j , et possède une durée d'existence variable $\Delta_1, \dots, \Delta_n$. Une caractéristique intéressante du $\Omega\alpha^*L_0\alpha^*T^*$ est que le nombre d'objets existant à une position temporelle donnée T_i est susceptible de varier. Une propriété importante du $\Omega\alpha^*L_0\alpha^*T^*$ est donc que le nombre d'objets est susceptible d'y varier en fonction du temps.



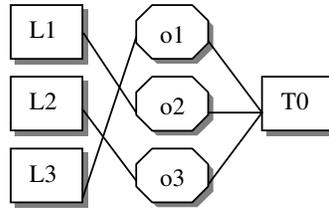
1.3.8. Le n -univers localisé et temporel à objet unique: ΩL

Il s'agit d'un n -univers comportant un objet unique et présentant une constante temporelle ainsi qu'une variable de localisation. Un tel n -univers présente donc une position temporelle unique T_0 ainsi que n positions locales L_1, L_2, \dots, L_n . Dans le ΩL (en notation étendue $\Omega L T_0$), l'objet unique existe en T_0 à une position temporelle donnée L_i .



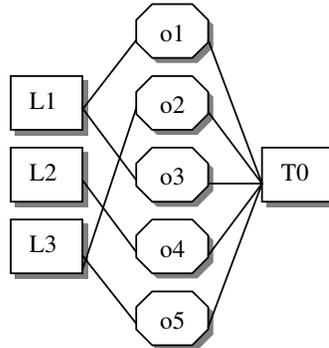
1.3.9. Le n -univers localisé et atemporel à objets multiples: $\Omega\alpha L$

Il s'agit d'un n -univers comportant de multiples objets o_1, \dots, o_n et présentant une constante temporelle ainsi qu'une variable de localisation. Un tel n -univers présente donc une position temporelle unique T_0 et n positions locales L_1, L_2, \dots, L_n . Dans le $\Omega\alpha L$ (en notation étendue $\Omega\alpha L\alpha^*T_0$), un objet donné ne peut se trouver qu'à une seule position spatiale en T_0 . En effet, un tel n -univers n'est pas à démultiplication locale (ubiquiste). Il existe donc dans ce n -univers plusieurs objets, tels que chaque objet se trouve placé à une position spatiale distincte. Dans le $\Omega\alpha L$, chaque objet se trouve ainsi à une position spatiale L_i en T_0 . De plus, de multiples objets existent en T_0 , de sorte que les objets sont en relation *plusieurs* avec la constante temporelle.



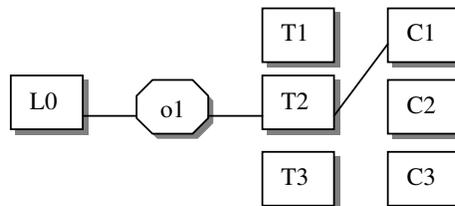
1.3.10. Le n -univers localisé et atemporel à objets multiples*: $\Omega\alpha^*L$

Il s'agit du même type de n -univers que précédemment, avec cette différence que l'on n'a pas un objet par position spatiale en T_0 . A l'inverse, plusieurs objets peuvent se trouver à une position spatiale distincte en T_0 . Ainsi, dans le $\Omega\alpha^*L$ (en notation étendue $\Omega\alpha^*L\alpha^*T_0$), les objets se trouvent en relation *plusieurs* avec la variable de localisation.



1.3.11. Le n -univers alocalisé, coloré et temporel à objet unique: ΩTC

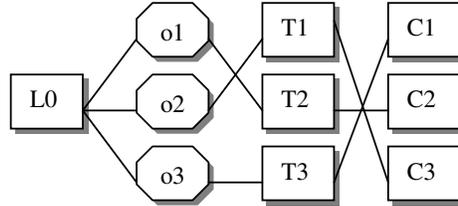
Il s'agit d'un n -univers comportant un objet unique et présentant à la fois une constante de localisation, une variable temporelle et une variable de couleur. Un tel n -univers présente donc un lieu unique L_0 , plusieurs positions temporelles T_1, T_2, \dots, T_n , ainsi que plusieurs taxons de couleur parmi C_1, C_2, \dots, C_p . Dans le ΩTC (en notation étendue $\Omega L_0 TC$), l'objet unique n'est pas persistant par rapport au temps: il existe en T_i avec une couleur donnée C_j mais n'existe pas aux autres positions temporelles.



1.3.12. Le n -univers alocalisé, coloré et temporel à objets multiples: $\Omega\alpha TC$

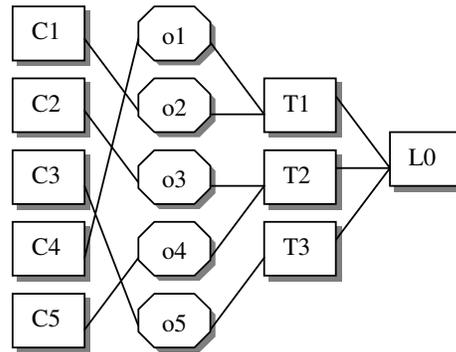
Il s'agit d'un n -univers comportant de multiples objets o_1, \dots, o_n , une constante de localisation, une variable temporelle et une variable de couleur. Un tel n -univers présente donc un lieu unique L_0 ,

plusieurs positions temporelles T_1, T_2, \dots, T_n , et plusieurs taxons de couleur C_1, C_2, \dots, C_p . Dans le $\Omega\alpha TC$ (en notation étendue $\Omega\alpha^*L_0\alpha TC$) les objets ne sont pas persistants par rapport au temps: les objets existent à une position temporelle donnée avec une couleur donnée mais n'existent pas aux autres positions temporelles. De plus, plusieurs objets peuvent se trouver en L_0 : les objets sont en relation plusieurs avec le critère-constante de localisation. Enfin, plusieurs objets ne peuvent se trouver à une position temporelle donnée: les objets sont donc en relation *un* avec le critère de temps.



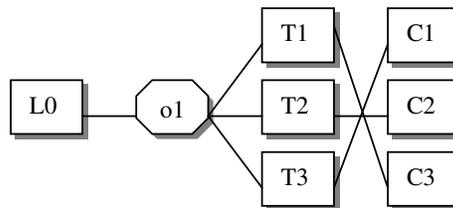
1.3.13. Le n -univers alocalisé, coloré et temporel à objets multiples*: $\Omega\alpha^*TC$

Il s'agit du même type de n -univers que le précédent, avec cette différence que plusieurs objets ne peuvent se trouver à une position temporelle donnée. Dans le $\Omega\alpha^*TC$ (en notation étendue $\Omega\alpha^*L_0\alpha^*TC$), les objets sont donc en relation *plusieurs* avec le critère de temps.



1.3.14. Le n -univers alocalisé, coloré et temporel* à objet unique: ΩT^*C

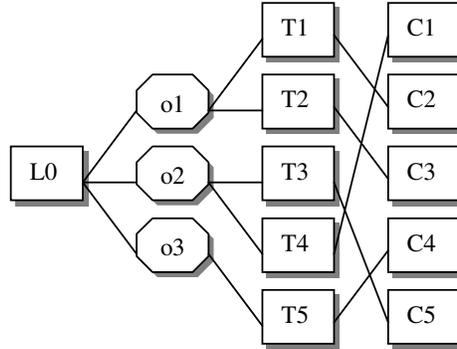
Il s'agit d'un n -univers comportant un objet unique, une constante de localisation, une variable temporelle à démultiplication et une variable de couleur. Un tel n -univers présente donc un lieu unique L_0 , plusieurs positions temporelles T_1, T_2, \dots, T_n et plusieurs taxons de couleur C_1, C_2, \dots, C_p . Dans le ΩT^*C (en notation étendue $\Omega L_0 T^*C$), l'objet unique est persistant par rapport au temps: s'il existe en L_0 en T_5 , c'est qu'il a également existé en L_0 en T_1, T_2, T_3, T_4 . Il est à noter également que l'objet unique est susceptible de changer de couleur au cours du temps. Dans le ΩT^*C , l'objet qui existe en L_0 durant la période T_3 - T_5 par exemple, peut présenter la couleur C_1 en T_3 et T_4 , mais posséder la couleur C_2 en T_5 .



1.3.15. Le n -univers alocalisé, coloré et temporel* à objets multiples: $\Omega\alpha T^*C$

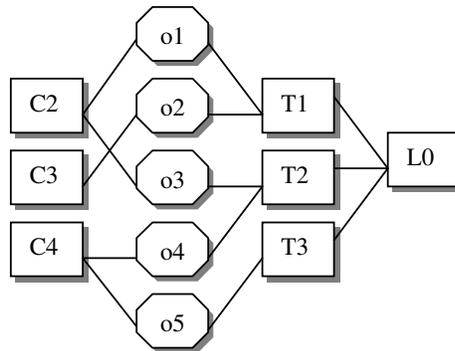
Un tel n -univers comporte de multiples objets o_1, \dots, o_n et présente une constante de localisation, une variable temporelle à démultiplication et une variable de couleur. Un tel n -univers présente donc un

lieu unique L_0 , plusieurs positions temporelles T_1, T_2, \dots, T_n et plusieurs taxons de couleur C_1, C_2, \dots, C_p . Dans le $\Omega\alpha T^*C$ (en notation étendue $\Omega\alpha L_0\alpha T^*C$), les objets sont persistants par rapport au temps: s'ils existent en L_0 en T_5 , c'est qu'ils ont existé en T_1, T_2, T_3, T_4 . Les objets sont donc à démultiplication temporelle. Il est à noter également que les objets sont susceptibles de changer de couleur au cours du temps. De plus, on peut avoir plusieurs objets en L_0 , à un moment donné. Enfin, on ne peut avoir qu'un objet à une position temporelle donné: les objets sont donc en relation *un* avec le critère temporel.



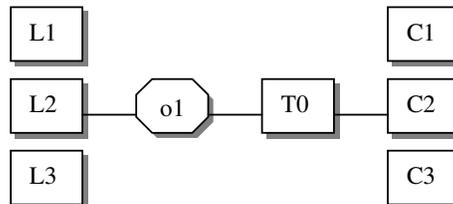
1.3.16. Le n -univers alocalisé, coloré et temporel* à objets multiples*: $\Omega\alpha^*T^*C$

Il s'agit du même n -univers que le précédent, avec cette différence que plusieurs objets peuvent se trouver à une position temporelle donnée. Dans le $\Omega\alpha^*T^*C$, (en notation étendue $\Omega\alpha^*L_0\alpha^*T^*C$) les objets sont donc en relation *plusieurs* avec le critère temporel.



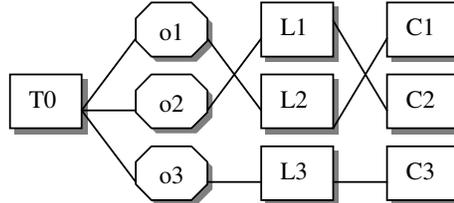
1.3.17. Le n -univers atemporel, localisé et coloré à objet unique: ΩLC

Il s'agit d'un n -univers comportant un objet unique et présentant une constante temporelle, une variable de localisation et une variable de couleur. Un tel n -univers présente donc une position temporelle unique T_0 , n positions locales L_1, L_2, \dots, L_n et n taxons de couleur C_1, C_2, \dots, C_p . Dans le ΩLC (en notation étendue ΩLT_0C) l'objet unique existe en T_0 à une position spatiale L_i avec une couleur C_j .



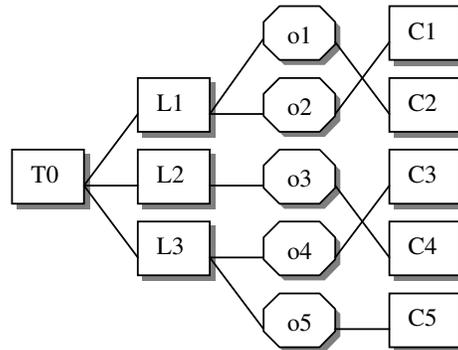
1.3.18. Le n -univers atemporel, localisé et coloré à objets multiples: $\Omega\alpha LC$

Ce n -univers comporte de multiples objets o_1, \dots, o_n , une constante temporelle, une variable de localisation et une variable de couleur. Un tel n -univers présente donc une position temporelle unique T_0 , n positions locales L_1, L_2, \dots, L_n et p taxons de couleur C_1, C_2, \dots, C_p . Dans le $\Omega\alpha LC$ (en notation étendue $\Omega\alpha L\alpha^*T_0C$), les objets existent en T_0 à une position spatiale donnée avec une couleur donnée. De plus, il plusieurs objets peuvent exister en T_0 : les objets sont en relation *plusieurs* avec le critère-constante de temps. Enfin, il existe une seul objet en T_0 à une position spatiale donnée: les objets sont en relation *un* avec le critère-variable spatial.



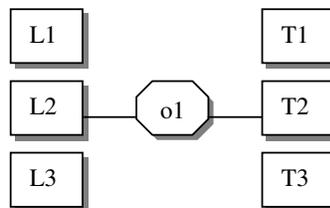
1.3.19. Le n -univers atemporel, localisé et coloré à objets multiples*: $\Omega\alpha^*LC$

Ce n -univers est identique au précédent, à cette différence toutefois que plusieurs objets peuvent se trouver en T_0 à une position spatiale donnée: dans le $\Omega\alpha^*LC$ (en notation étendue $\Omega\alpha^*L\alpha^*T_0C$), les objets sont en relation *plusieurs* avec le critère-variable de localisation.



1.3.20. Le n -univers localisé et temporel à objet unique: ΩLT

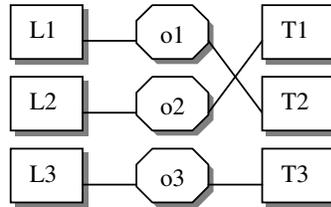
Il s'agit d'un n -univers comportant un objet unique, une variable de localisation et une variable temporelle. Un tel n -univers présente donc n positions spatiales L_1, L_2, \dots, L_n et plusieurs positions temporelles T_1, T_2, \dots, T_p . Dans le ΩLT , l'objet unique n'est pas persistant par rapport au temps: il existe en T_i à une position temporelle donnée L_i mais n'existe pas aux autres positions temporelles.



1.3.21. Le n -univers localisé et temporel à objets multiples: $\Omega\alpha L\alpha T$

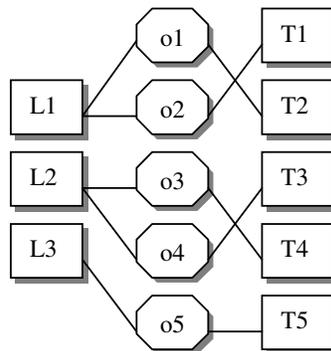
Il s'agit d'un n -univers comportant de multiples objets o_1, \dots, o_n , une variable de localisation et une variable temporelle. Ce n -univers présente n positions locales L_1, L_2, \dots, L_n et plusieurs positions

temporelles T_1, T_2, \dots, T_p . Dans le $\Omega\alpha L\alpha T$, les objets ne sont pas persistants par rapport au temps: ils existent en T_i à une position temporelle donnée L_i mais n'existent pas aux autres positions temporelles. En outre, un seul objet peut se trouver à une position spatiale donnée à un moment donné: les objets sont en relation *un* avec le critère-variable de localisation et avec le critère-variable temporel.



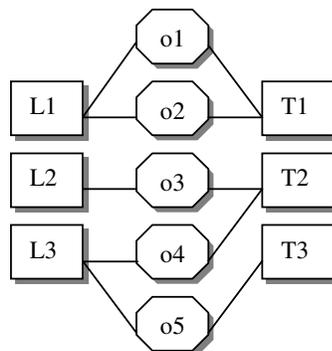
1.3.22. Le n -univers localisé et temporel à objets multiples*: $\Omega\alpha^*L\alpha T$

Il s'agit du même type de n -univers que le précédent, à cette différence que plusieurs objets peuvent se trouver à une position spatiale donnée à un moment donné: les objets sont en relation *plusieurs* avec le critère-variable de localisation. Cependant, plusieurs objets ne peuvent se trouver à une même position temporelle donnée. Les objets sont en relation *un* avec le avec le critère-variable temporel.



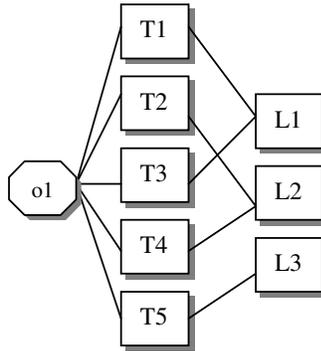
1.3.23. Le n -univers localisé et temporel à objets multiples*: $\Omega\alpha^*L\alpha^*T$

Il s'agit du même type de n -univers que le précédent, à cette différence que plusieurs objets peuvent se trouver à une position spatiale donnée à un moment donné: les objets sont en relation *plusieurs* avec le critère-variable de localisation et avec le critère-variable temporel.



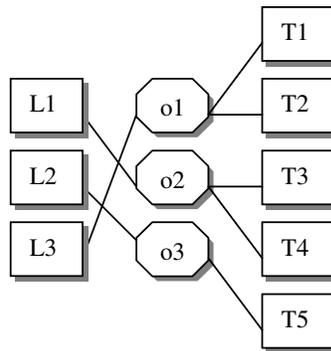
1.3.24. Le n -univers localisé et temporel* à objet unique: ΩLT^*

Un tel n -univers comporte un objet unique, une variable de localisation ainsi qu'une variable temporelle à démultiplication. Ce n -univers présente n positions locales L_1, L_2, \dots, L_n et plusieurs positions temporelles T_1, T_2, \dots, T_p . Dans le ΩLT^* , l'objet unique est persistant par rapport au temps: s'il existe en L_i en T_5 , c'est qu'il a existé en T_1, T_2, T_3, T_4 . Il est à noter également que l'objet unique est susceptible de changer de position spatiale au cours du temps. Dans le ΩLT^* , l'objet qui existe durant la période T_3 - T_5 par exemple, peut occuper la position spatiale L_1 en T_3 et T_5 , mais se trouver à la position spatiale L_2 en T_4 .



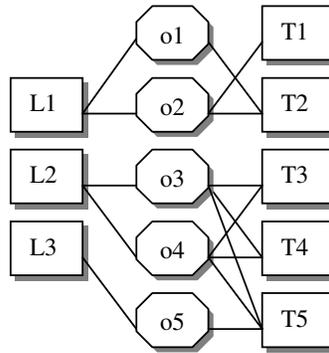
1.3.25. Le n -univers temporel* et localisé à objets multiples: $\Omega \alpha L \alpha T^*$

Il s'agit d'un n -univers comportant de multiples objets o_1, \dots, o_n , une variable temporelle et une variable de localisation. Un tel n -univers présente n positions locales L_1, L_2, \dots, L_n et plusieurs positions temporelles T_1, T_2, \dots, T_p . De plus, chaque objet est persistant par rapport au temps: s'il existe en T_5 , c'est qu'il a existé en T_1, T_2, T_3, T_4 . Dans le $\Omega \alpha T^* L$, le nombre d'objets est susceptible de varier en fonction du temps et de la position spatiale. En outre, un seul objet peut se trouver à une position spatiale donnée à un moment donné: les objets sont en relation *un* avec le critère-variable de localisation et avec le critère-variable temporel.



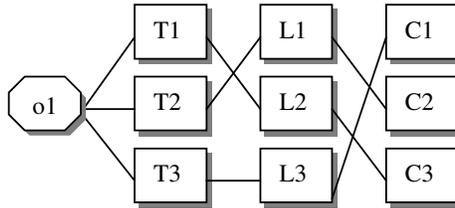
1.3.26. Le n -univers temporel* et localisé à objets multiples*: $\Omega \alpha^* L \alpha^* T^*$

Il s'agit du même type de n -univers que le précédent, avec cette différence que plusieurs objets peuvent se trouver à une position spatiale donnée à un moment donné: les objets sont en relation *plusieurs* avec le critère-variable de localisation et avec le critère-variable temporel.



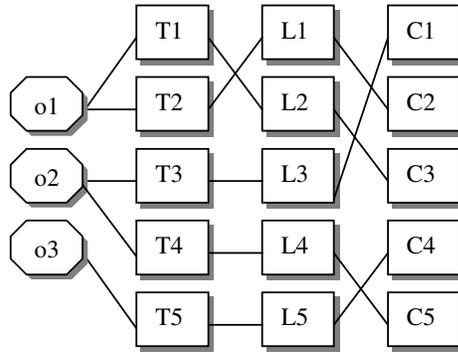
1.3.27. Le n -univers localisé et temporel* à objet unique: ΩLT^*C

Un tel n -univers comporte un objet unique, une variable de localisation, une variable temporelle à démultiplication et une variable de couleur. Un tel n -univers présente n positions spatiales L_1, L_2, \dots, L_n , plusieurs positions temporelles: T_1, T_2, \dots, T_p et plusieurs taxons de couleur C_1, C_2, \dots, C_q . Dans le ΩLT^*C , l'objet unique est persistant par rapport au temps: s'il existe en T_5 , c'est qu'il a existé en T_1, T_2, T_3, T_4 . Il est à noter également que l'objet unique est susceptible de changer de position spatiale et de couleur au cours du temps. Dans le ΩLT^*C , l'objet qui existe durant la période T_3 - T_5 par exemple, peut présenter la position spatiale L_1 et la couleur C_2 en T_3 et T_5 , mais posséder la position spatiale L_2 et la couleur C_5 en T_4 .



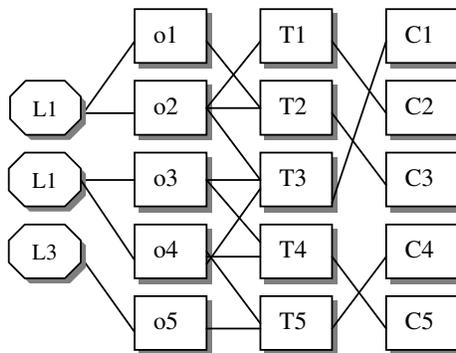
1.3.28. Le n -univers localisé et temporel* à objets multiples: $\Omega \alpha L \alpha T^*C$

Il s'agit d'un n -univers comportant de multiples objets o_1, \dots, o_n , une variable de localisation, une variable temporelle à démultiplication et une variable de couleur. Un tel n -univers présente n positions spatiales L_1, L_2, \dots, L_n , plusieurs positions temporelles: T_1, T_2, \dots, T_p et plusieurs taxons de couleur C_1, C_2, \dots, C_q . Dans le $\Omega \alpha L \alpha T^*C$, les objets sont persistants par rapport au temps: s'ils existent en T_5 , c'est qu'ils ont existé en T_1, T_2, T_3, T_4 . Il est à noter également que les objets sont susceptibles de changer de position spatiale et de couleur au cours du temps. Dans le $\Omega \alpha L \alpha T^*C$, un objet qui existe durant la période T_3 - T_5 par exemple, peut présenter la position spatiale L_1 et la couleur C_2 en T_3 et T_5 , mais avoir la position spatiale L_2 et la couleur C_5 en T_4 . Enfin, un seul objet peut se trouver à une position spatiale donnée à un moment donné: les objets sont en relation *un* avec le critère-variable de localisation et avec le critère-variable temporel.



1.3.29. Le n -univers localisé et temporel* à objets multiples*: $\Omega\alpha^*L\alpha^*T^*C$

Ce type de n -univers est le que le précédent, à cette différence que plusieurs objets peuvent se trouver à une position spatiale donnée à un moment donné: les objets sont en relation *plusieurs* avec le critère-variable de localisation et avec le critère-variable temporel.



Chapitre 2. L'Argument de l'Apocalypse

2.1. Le problème

On peut tout d'abord présenter ainsi, de manière informelle, l'*argument de l'apocalypse* (*Doomsday Argument*, DA, dans ce qui suit). Formulons tout d'abord les deux hypothèses suivantes concernant le futur de l'humanité:

- (H1₄) l'humanité comptera au total 100 milliards d'humains
- (H2₅) l'humanité comptera au total 10000 milliards d'humains

Maintenant, il s'avère que je suis moi-même l'humain n° 50×10^9 . Compte tenu de cela et du fait que je peux considérer mon rang au sein de l'humanité comme aléatoire, il s'ensuit une révision à la hausse de la probabilité de l'hypothèse (H1), car elle est rendue plus probable par la donnée concrète constituée par mon rang de naissance.

L'Argument de l'Apocalypse est précisément basé sur une analogie entre d'une part la situation concernant les humains ayant existé depuis le commencement de l'humanité et d'autre part une expérience concernant des boules contenues dans une urne. DA tire sa force du fait que le raisonnement qui prévaut dans la situation applicable à l'urne, vaut également pour la situation concernant l'humanité. Cette analogie qui est à la base de DA est clairement décrite par Nick Bostrom dans son introduction (1997) à l'Argument de l'Apocalypse. Cette description est la suivante:

The core idea is this. Imagine that two big urns are put in front of you, and you know that one of them contains ten balls and the other a million, but you are ignorant as to which is which. You know the balls in each urn are numbered 1, 2, 3, 4 ... etc. Now you take a ball at random from the left urn, and it is number 7. Clearly, this is a strong indication that that urn contains only ten balls. If originally the odds were fifty-fifty, a swift application of Bayes' theorem gives you the posterior probability that the left urn is the one with only ten balls. (Pposterior (L=10) = 0.999990). But now consider the case where instead of the urns you have two possible human races, and instead of balls you have individuals, ranked according to birth order. As a matter of fact, you happen to find that your rank is about sixty billion. Now, say Carter and Leslie, we should reason in the same way as we did with the urns. That you should have a rank of sixty billion or so is much more likely if only 100 billion persons will ever have lived than if there will be many trillion persons. Therefore, by Bayes' theorem, you should update your beliefs about mankind's prospects and realise that an impending doomsday is much more probable than you have hitherto thought¹⁰.

¹⁰ Trad.: L'idée centrale est celle-ci. Imaginez que deux grandes urnes se trouvent placées devant vous, et que vous savez que l'une d'elles contient dix boules et l'autre un million, mais vous êtes ignorant de laquelle il s'agit. Vous savez que les boules dans chaque urne sont numérotées 1, 2, 3, 4... etc. Maintenant vous prenez une boule au hasard de l'urne gauche, et il s'agit de la boule numéro 7. Clairement, il s'agit là d'une indication forte que cette urne contient seulement dix boules. Si à l'origine les probabilités étaient cinquante-cinquante, une application rapide du théorème de Bayes conduit à une probabilité a posteriori que l'urne gauche est celle qui contient seulement dix boules. (Pposterior (L=10) = 0,999990). Mais considérez maintenant le cas où au lieu des urnes il s'agit de deux races humaines possibles, et au lieu des boules il s'agit d'individus, rangés selon l'ordre de naissance. En fait, vous vous apercevez que votre rang de naissance est environ soixante milliards. Maintenant, disent Carter et Leslie, nous devrions raisonner comme nous l'avons fait avec les urnes. Que vous devriez avoir un rang de naissance de soixante milliards environ est beaucoup plus probable si seulement 100 milliards de personnes auront jamais vécu que s'il y en aura beaucoup de nombreux milliards de milliards. Par conséquent, par le théorème de Bayes, vous devriez mettre à jour votre croyance au sujet des perspectives de l'humanité et vous rendez compte qu'une apocalypse imminente est beaucoup plus probable que vous ne l'avez jusqu'ici pensé.

Bostrom met ainsi explicitement en parallèle deux types de situations: d'une part, le modèle probabiliste concernant des boules contenues dans une urne, et d'autre part la situation correspondant à l'humanité. L'expérience concernant les boules de l'urne conduit à un résultat non controversé, à savoir un décalage bayésien en faveur de l'hypothèse selon laquelle l'urne seulement 10 boules. A l'inverse, le résultat de même nature appliqué à notre humanité se révèle contraire à l'intuition. Aussi le problème posé par DA est-il de savoir si le raisonnement qui prévaut dans le modèle de l'urne s'applique valablement à la situation qui est celle de l'humanité. D'une manière générale, le problème posé est ainsi celui de savoir si l'analogie avec l'urne qui est à la base de DA se révèle valable. Afin de mettre en œuvre les conditions permettant de vérifier avec précision si cette analogie est fondée, il conviendra de s'attacher tout d'abord à modéliser les deux situations en termes de n-univers. Ceci sera réalisé dans les deux chapitres suivants. Avant toutefois de procéder à une telle modélisation, il convient préalablement de prendre en considération un certain nombre de points particuliers concernant DA, dont la clarification s'avère nécessaire afin de préparer l'analyse qui va suivre.

2.2. Le principe SSA

Le principe SSA (*Self-Sampling Assumption*) est le principe selon lequel des observateurs doivent raisonner comme s'ils étaient des observateurs choisis au hasard parmi l'ensemble des observateurs. Un tel principe constitue une prémisse de DA, car dans le modèle probabiliste de l'urne, la boule tirée est choisie au hasard. De même, on considère dans DA, que chaque humain qui applique DA est choisi au hasard parmi l'ensemble de tous les humains composant la classe de référence. SSA a été formalisé ainsi par Bostrom (2001):

(SSA) Observers should reason as if they were a random sample from the set of all observers in their reference class¹¹

On peut noter qu'une version légèrement différente du principe SSA est également donnée par William Eckhardt¹² sous la forme du principe HR (*Human Randomness Assumption*):

(HR) We can validly consider our birth rank as generated by random or equiprobable sampling from the collection of all persons who ever live¹³

La formulation d'Eckhardt est légèrement différente de celle utilisée par Bostrom. En effet, le principe SSA utilisé par Bostrom se réfère au fait que ce sont les observateurs qui sont choisis au hasard, alors que la formulation de HR par Eckhardt fait référence au fait que c'est le *rang de naissance*¹⁴ de chaque individu qui s'avère choisi au hasard. Il apparaît ainsi que SSA vaut pour les observateurs et que HR s'applique aux rangs de naissance des observateurs. Les deux principes doivent donc être distingués. Il apparaît en effet qu'une classe de référence composée de 999 humains et d'un dauphin constitue une violation du principe SSA parce que le dauphin n'est pas représentatif (*typical*) de la classe de référence. Mais une classe de référence composée de 1000 humains possédant chacun un numéro de naissance qui ne ferait pas l'objet d'un tirage aléatoire constituerait une violation de HR. Cette différence entre SSA et HR étant mentionnée, il conviendra de revenir ultérieurement sur cette dernière.

¹¹ Trad.: Les observateurs devraient raisonner comme s'ils étaient un échantillon aléatoire de l'ensemble de tous les observateurs au sein de leur classe de référence

¹² Cf. 1997, p. 248.

¹³ Trad.: Nous pouvons bien considérer notre rang de naissance comme résultant d'un tirage aléatoire ou équiprobable parmi l'ensemble de toutes les personnes qui auront jamais vécu.

¹⁴ *Birth rank*.

2.3. Version déterministe ou indéterministe de DA

Dans son analyse de DA, Leslie distingue à plusieurs reprises selon que DA prend place dans un univers déterministe ou bien dans un univers indéterministe. De manière générale, Leslie considère que DA se manifeste avec force dans un univers déterministe mais trouve également à s'appliquer, bien que de manière affaiblie, dans un univers indéterministe. Ainsi, conclut Leslie, DA se trouve seulement affaibli si notre univers est de nature indéterministe et si le futur de l'humanité n'est pas complètement scellé. Leslie est ainsi amené à distinguer deux types de situations. Dans un premier cas (a) qui est celui d'un univers déterministe, le futur de l'humanité est déjà fixé et le nombre d'humains est d'ores et déjà déterminé avec précision. Quels que soient les événements susceptibles de se produire dans le futur, le nombre d'humains qui composeront finalement l'humanité se révèle inexorablement fixé. Dans ce type de situation, Leslie considère que DA s'applique avec toute sa force. Dans un second cas (b), qui est celui d'un univers indéterministe, le nombre d'humains n'est pas encore déterminé et ne le sera¹⁵ qu'au temps T_1 , avec $T_1 > T_0$ (le raisonnement inhérent à DA prenant place en T_0). Dans cette dernière hypothèse, le futur de l'humanité n'est pas encore tracé de manière définitive et le nombre d'humains qui existeront finalement n'est pas encore déterminé. Ce nombre dépend d'événements futurs, qui ne sont pas encore produits. Dans ce second type de situation, Leslie considère que DA trouve encore à s'appliquer, même s'il s'en trouve affaibli. Ainsi, le fait de considérer une situation de type indéterministe a seulement pour effet de restreindre la portée de DA, sans en annihiler toutefois l'effet.

Ma façon d'appréhender le problème posé par la possibilité que notre univers soit indéterministe est la suivante. Ici, le but est de parvenir à apporter une solution au problème posé par DA. Pour cela, il est nécessaire de considérer une version de DA où le problème qui lui est inhérent se manifeste avec toute sa force. Dans l'hypothèse où le futur de l'humanité est rendu imprécis à cause de facteurs indéterministes, il en résulte (rejoignant ici Leslie) une version plus faible de DA. A l'inverse, si le futur de l'humanité est entièrement régi par des facteurs déterministes et que le nombre d'humains qui composeront l'humanité est d'ores et déjà fixé, alors DA se manifeste avec toute sa force. Par conséquent, il convient de prendre en considération la version la plus forte de DA. Et cette dernière version se manifeste dans le cadre d'un univers entièrement déterministe. Dans ce cas, DA se pose avec toute son acuité. Car il s'agit ici non pas de considérer notre situation humaine en particulier, mais de s'attacher véritablement à appréhender le problème général qui est posé par DA. Ainsi, dans ce qui suit, on se placera résolument dans le cas d'un univers déterministe, dans lequel le nombre d'humains qui existeront finalement se trouve d'ores et déjà fixé.

2.4. Le problème de la définition de la classe de référence

Un important problème qui apparaît lors de l'analyse de DA est celui de la définition de la *classe de référence* qui sous-tend le raisonnement. Un tel problème, qui a été notamment soulevé par Eckhardt (1993) n'a pas été éludé par Leslie, qui y apporte notamment¹⁶ une réponse détaillée dans *The End of the World*. Dans *Une Solution pour l'Argument de l'Apocalypse* (1998) et *Comment l'Urne de Carter et Leslie se Déverse dans celle de Hempel* (1999), j'ai présenté une analyse de DA fondée sur le problème posé par la définition de la classe de référence. Cette analyse peut être résumée brièvement comme suit¹⁷.

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

¹⁵ Pour simplifier.

¹⁶ Cf. également Leslie (1993).

¹⁷ Les paragraphes qui suivent sont tirés de mon *Comment l'Urne de Carter et Leslie se Déverse dans celle de Hempel* (1999), avec quelques modifications de détail.

¹⁸ Leslie utilise les termes de *human race*.

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 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, a 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 DA.

A ce stade, on peut s'interroger si une telle analyse a pour effet de détruire le problème posé par DA. En ce sens, constitue-t-elle une solution satisfaisante à DA? On peut penser toutefois que non. En effet, un défenseur de DA pourrait répliquer que cette analyse du problème lié à la classe de référence n'a pas pour effet véritablement de détruire DA, mais seulement d'en réduire la portée. Car l'effet de cette analyse, pourrait-on dire, est simplement que DA fonctionne sans la terreur habituellement associée à la disparition de la classe de référence. Il résulte en effet de cette analyse que la terreur associée à la notion d'Apocalypse et d'extinction de la classe de référence n'est plus une conclusion qui peut être légitimement associée à DA. Cependant, il en résulte seulement une version plus faible de DA, car le cœur du raisonnement inhérent à DA se révèle toujours présent. En effet, si on choisit la classe de référence de manière à faire disparaître l'effet de terreur lié à DA, on obtient toujours un décalage bayésien. Et ce décalage bayésien se révèle toujours contraire à l'intuition.

Considérons ainsi, pourrait poursuivre un défenseur de DA, une classe de référence assimilée de manière précise à la sous-espèce *Homo sapiens sapiens*. Et supposons que notre sous-espèce *Homo sapiens sapiens* soit remplacée, par le processus naturel de l'évolution, par la sous-espèce *Homo sapiens supersapiens*. Ici, on s'intéresse seulement à la classe de référence constituée d'*Homo sapiens sapiens*, sans se préoccuper du devenir de notre sous-espèce, qui évoluera ultérieurement en *Homo sapiens supersapiens*. Bien sûr, le fait de restreindre la classe de référence à *Homo sapiens sapiens* a pour effet d'éliminer la terreur habituellement inhérente à DA. Car le fait qu'*Homo sapiens sapiens* disparaisse n'apparaît plus comme un événement tragique mais comme un phénomène naturel lié à l'évolution. Cependant, pourrait-il être rétorqué, DA fonctionne toujours dans cette situation. Car l'argument entraîne toujours une augmentation de la probabilité de l'extinction de notre sous-espèce *Homo sapiens sapiens*. Ainsi, l'argument se révèle toujours valide, entraînant toujours un décalage bayésien contraire à l'intuition. Dans ce contexte en effet, DA s'avère bien toujours présent. Bien sûr, une telle situation ne s'accompagne pas de l'effet terrifiant qui lui est habituellement associé. Mais ici, on ne se préoccupe pas alors de savoir si d'autres individus existent ou non encore à l'extérieur de la classe de référence, une fois que cette dernière est éteinte. Car l'effet de DA ne porte que sur la classe de référence sur laquelle il opère. Et comme le mentionne Bostrom (2000), DA peut signifier

simplement que l'espèce humaine sera remplacée par une autre race plus rapidement qu'on n'aurait pu le penser de prime abord. Ainsi, en tout état de cause, pourrait poursuivre un défenseur de DA, il s'avère bien que l'argument fonctionne toujours, même si ce n'est que sous une forme atténuée.

Une fois dépourvu de son effet terrifiant, DA apparaît ainsi toujours valide. Ceci invite à s'intéresser aux autres problèmes soulevés par DA qui sont indépendants du problème de la définition de la classe de référence. Ce dernier problème est certainement important pour DA. Nous aurons l'occasion plus loin de revenir sur ce problème de la définition de la classe de référence et de présenter dans un chapitre ultérieur une analyse plus approfondie. Car il apparaît à ce stade qu'une véritable solution à DA se doit également d'apporter une réponse au problème posé par la définition de la classe de référence.

Chapitre 3. Le modèle des deux urnes

3.1. Description

DA constitue, on vient de le voir, un argument qui est basé sur une analogie entre d'une part une situation mettant en scène l'espèce humaine et d'autre part une situation concernant des boules dans une urne. Il s'avère ainsi nécessaire d'analyser tour à tour chacune de ces situations. Commençons par analyser tout d'abord de manière précise le modèle qui est celui de l'urne. Rappelons tout d'abord les termes dans lesquels l'expérience concernant des boules dans une urne est notamment décrite par Bostrom¹⁹:

Imagine that two big urns are put in front of you, and you know that one of them contains ten balls and the other a million, but you are ignorant as to which is which. You know the balls in each urn are numbered 1, 2, 3, 4 ... etc. Now you take a ball at random from the left urn, and it is number 7. Clearly, this is a strong indication that that urn contains only ten balls. If originally the odds were fifty-fifty, a swift application of Bayes' theorem gives you the posterior probability that the left urn is the one with only ten balls. ($P_{\text{posterior}}(L=10) = 0.999990$)²⁰.

Bostrom présente ainsi une situation qui a été dénommée ensuite par Korb et Oliver (1999) le *modèle des deux urnes* (*two urn case*). La situation qui est celle du *modèle des deux urnes* a été décrite par Bostrom afin de traduire de manière explicite l'analogie sur laquelle l'Argument de l'Apocalypse est basé. Le modèle probabiliste associé au *modèle des deux urnes* est ainsi le suivant. On considère une urne qui contient soit 10 boules, soit 1000 boules. On fait ainsi les deux hypothèses suivantes:

(H1₆) l'urne contient 10 boules

(H2₇) l'urne contient 1000 boules

Ces hypothèses sont telles que les probabilités a priori sont les suivantes: $P(H1) = P(H2) = 0,5$.²¹ Toutes les boules par ailleurs sont numérotées. Soit ensuite l'événement E: je tire une boule dans l'urne dont le numéro est 5. Il s'ensuit alors les probabilités conditionnelles suivantes: $P(E|H1) = 1/10$ et $P(E|H2) = 1/1000$. Ceci conduit à réviser à la hausse la probabilité selon laquelle l'urne contient 10 boules, par application de la formule de Bayes: $P'(H1) = [P(H1)*P(E|H1)] / [P(H1)*P(E|H1) + P(H2)*P(E|H2)] = P'(H1) = [0,5*0,1] / [0,5*0,1 + 0,5*0,001] = 0,05/0,0505 \approx 0,99$. D'où $P'(H2) = 1 - P'(H1) \approx 0,01$. Le *modèle des deux urnes* constitue simplement une illustration du principe général selon lequel il s'avère rationnel de réviser une probabilité initiale en fonction des éléments supplémentaires qui peuvent apparaître par la suite.²²

¹⁹ Cf. Bostrom (1997).

²⁰ Trad.: Imaginez que deux grandes urnes se trouvent placées devant vous, et que vous savez que l'une d'elles contient dix boules et l'autre un million, mais vous êtes ignorant de laquelle il s'agit. Vous savez que les boules dans chaque urne sont numérotées 1, 2, 3, 4... etc. Maintenant vous prenez une boule au hasard de l'urne gauche, et il s'agit de la boule numéro 7. Clairement, il s'agit là d'une indication forte que cette urne contient seulement dix boules. Si à l'origine les probabilités étaient cinquante-cinquante, une application rapide du théorème de Bayes conduit à une probabilité a posteriori que l'urne gauche est celle qui contient seulement dix boules. ($P_{\text{posterior}}(L=10) = 0,999990$).

²¹ A cet égard, on peut choisir d'autres probabilités initiales sans que cela n'affecte la portée de l'expérience. Quelles que soient en effet les probabilités a priori qui sont choisies, il s'ensuit en effet un décalage bayésien en faveur de l'hypothèse (H1).

²² Ainsi que l'illustre Leslie en mentionnant l'expérience suivante (*the story of the cat*, 1996, p. 187): 'Waking up in the night, you form two theories. Each has a half-chance of being right, you estimate. The first, that you left the back door open, gives the chances as 10 per cent that the neighbour's cat is in your bedroom. The second, that you shut the door, puts those chances at 0.01 per cent. You switch on the light and see the cat. You should now much prefer the first theory.' Trad.: 'Vous réveillant la nuit, vous formez deux théories. Chacun a une chance sur deux d'être exacte, estimez-vous. La première,

On peut observer par ailleurs que le *modèle des deux urnes* est tout à fait équivalent à une situation de même nature décrite par Leslie²³. Un tel modèle est également dénommé *l'expérience de la loterie* (*lottery case*) par Bartha & Hitchcock (1999).

Suppose that you know that your name is in a lottery name, but not how many other names the urn contains. You estimate, however, that there's a half chance that it contains a thousand names, and a half chance of its containing only ten. Your name then appears among the first three drawn from the urn. Don't you have rather strong grounds for revising your estimate? Shouldn't you now think it very improbable that there are another 997 names waiting to be drawn?²⁴

On peut noter que cette dernière version décrite par Leslie (*expérience de la loterie*) présente une différence par rapport à celle présentée par Bostrom (*modèle des deux urnes*). Soit en effet n le nombre de boules présentes dans l'urne. Bostrom considère ainsi une urne dont les boules sont numérotées de 1 à n . En revanche, Leslie considère des boules dont l'une des boules est individualisée et comporte le nom de l'observateur. Cependant dans leur effet, ces versions se révèlent tout à fait équivalentes. Car elles ont toutes deux pour conséquence que le calcul de la probabilité conditionnelle $P(E|H1)$ est égal à l'inverse du nombre de boules présentes dans l'urne: $P(E|H1) = 1/n$. Plus le nombre de boules présentes dans l'urne est faible et plus cette probabilité conditionnelle s'avère forte.

On peut noter ici que l'on peut également prendre en considération une variation équivalente du *modèle des deux urnes* qui prend en compte des boules de deux couleurs différentes: rouge et vert. Appelons une telle version la version *bicolore* du *modèle des deux urnes*. Dans ce cas, il s'ensuit les deux hypothèses suivantes:

(H1₈) l'urne contient 10 boules rouges

(H2₉) l'urne contient 1000 boules dont 10 boules rouges et 990 boules vertes

Et on a également l'événement E: je tire une boule dans l'urne dont la couleur est rouge. On vérifie ici aisément à ce stade que compte tenu de cette dernière formulation, on a également l'effet du *modèle des deux urnes*.

D'une manière générale, le *modèle des deux urnes* constitue un modèle qui conduit à réviser à la hausse une probabilité initiale selon laquelle l'urne ne contient qu'un petit nombre de boules. Il s'ensuit que lorsqu'une situation réelle correspond à ce modèle, il est rationnel de réviser à la hausse une

selon laquelle vous avez laissé la porte arrière ouverte, conduit à une chance de 10 pour cent que le chat des voisins se trouve dans votre chambre à coucher. La seconde, selon laquelle vous avez fermé la porte, place ces mêmes chances à 0,01 pour cent. Vous allumez la lumière et vous voyez le chat. Vous devriez maintenant nettement préférer la première théorie.'

²³ Cf. (1996, p. 22). Leslie mentionne également le même modèle probabiliste (1996, pp. 68-69): 'One of the balls in an urn has your name on it. Balls are drawn without replacement and the seventh ball drawn is 'yours.' This result favors the hypothesis that the urn contained just twenty balls over the hypothesis that it contained a thousand.' Trad.: 'Une des boules dans une urne comporte votre nom. Des boules sont tirées sans être remplacées et la septième boule dessinée est la votre. Ce résultat conduit à favoriser l'hypothèse selon laquelle l'urne a contenu seulement vingt boules par rapport à l'hypothèse selon laquelle elle en a contenu mille.'

²⁴ Trad.: Supposez que vous savez qu'un bulletin contenant votre nom se trouve parmi d'autres bulletins dans une loterie, mais que vous ne savez pas combien d'autres bulletins l'urne contient. Vous estimez, cependant, qu'il y a une chance sur deux qu'elle contient mille noms, et une chance sur deux qu'elle en contienne seulement dix. Votre nom apparaît alors parmi les trois premiers tirés de l'urne. N'avez-vous pas des raisons plutôt fortes pour mettre à jour votre évaluation initiale? Ne devriez-vous pas maintenant penser qu'il est très improbable qu'il reste encore dans l'urne 997 noms attendant d'être tirés?

estimation initiale selon laquelle le nombre d'objets est faible, par rapport à une hypothèse alternative selon laquelle le nombre d'objets est élevé. Pour cette raison, on peut l'appeler *modèle à minoration*.

Plusieurs caractéristiques du *modèle à minoration* méritent à ce stade, d'être soulignées. Il apparaît tout d'abord que dans le *modèle à minoration*, les probabilités a priori $P(H1)$ et $P(H2)$ sont données. En second lieu, il s'avère que l'on peut choisir n'importe quelles probabilités a priori. Dans l'exemple ci-dessus, on utilise $P(H1) = P(H2) = 0,5$. Cependant, toute autre probabilité initiale conviendrait. Ce qu'indique le *modèle à minoration*, c'est un décalage bayésien, une augmentation de la probabilité initiale selon laquelle le nombre d'objets est faible.

Il convient à ce stade de noter la comptabilité de SSA et du *modèle à minoration*. Il apparaît en effet que le protocole du *modèle à minoration*²⁵ inclut le fait que la boule est tirée au hasard. Ainsi, SSA et le *modèle à minoration* s'avèrent tout à fait compatibles. Plus encore, le tirage aléatoire de la boule et donc l'utilisation de SSA se révèle être une des conditions de l'application du *modèle à minoration*. Ainsi, les conclusions dérivées du *modèle à minoration* ne sont possibles que dès lors que la boule est tirée au hasard.

3.2. Modélisation dans le n-univers coloré à objets multiples ($\Omega\alpha C$ ou $\Omega\alpha^*C$)

Il convient maintenant de s'attacher à décrire le *modèle des deux urnes* en termes de n-univers. Compte tenu du principe méthodologique qui a été défini plus haut, on s'attachera ici à modéliser le *modèle des deux urnes* dans le n-univers le plus simple. Dans quel type de n-univers le *modèle des deux urnes* prend-il donc place? Compte tenu de la description qui en est fournie, il convient ici de s'attacher à déterminer les critères du n-univers correspondant. Il apparaît opportun de distinguer ici selon que l'on prend en compte le *modèle des deux urnes* de Bostrom (ou l'*expérience de la loterie* de Leslie) ou bien le *modèle des deux urnes* bicolore qui vient d'être mentionné.

3.2.1. Le modèle des deux urnes de Bostrom et le $\Omega\alpha C$

Considérons tout d'abord le *modèle des deux urnes* de Bostrom ou l'*expérience de la loterie* décrit par Leslie. Il s'avère en premier lieu que la situation du *modèle des deux urnes* se caractérise par la présence d'objets multiples: les boules. On se situe donc dans un n-univers à objets multiples ($\Omega\alpha\dots$)²⁶. En second lieu, on peut considérer que les boules se trouvent dans un lieu unique: l'urne. Le n-univers correspondant présente donc un critère-constante de localisation (L_0).

A ce stade, on peut s'interroger sur la temporalité du n-univers correspondant au *modèle des deux urnes*. Le *modèle des deux urnes* est en effet basé, comme un grand nombre d'expériences de pensée portant sur des boules contenues dans une urne, sur une probabilité *a priori* et une probabilité *a posteriori*. Ceci incite à penser que le n-univers qui en résulte est nécessairement temporel et comporte donc au moins deux positions temporelles: T_0 (temps de la probabilité a priori) et T_1 (temps de la probabilité a posteriori). Ceci conduit à modéliser le *modèle des deux urnes* dans un n-univers comportant de multiples objets: o_1, \dots, o_n situés un lieu unique (L_0) et possédant deux positions temporelles uniques T_0 et T_1 . On s'intéresse alors exclusivement au nombre de boules existant en T_1 . Le nombre de boules existant en T_0 est indifférent. Cependant, compte tenu du principe méthodologique de *simplicité* qui a été exposé, on peut se demander s'il n'est pas possible de se dispenser de la variable temporelle. Bien que le *modèle des deux urnes* soit basé sur une probabilité *a priori* et une probabilité *a posteriori*, il apparaît bien en effet que l'on peut se dispenser de recourir à un n-univers comportant une variable *temporelle*. Il suffit en effet de considérer que les probabilités initiales $P(H1)$ et $P(H2)$ ainsi que l'événement E (une boule a été tirée possédant le n° 5) font partie des données du problème. On considère alors, pour simplifier, que le raisonnement qui conduit au calcul bayésien prend place au moment précis où celui qui raisonne prend connaissance des données du problème. Dans ce contexte on le voit, il n'est pas nécessaire de recourir à une distinction temporelle basée sur la probabilité *a priori* (en T_0) et *a posteriori* (en T_1). Une telle version étant plus simple que la précédente, on la

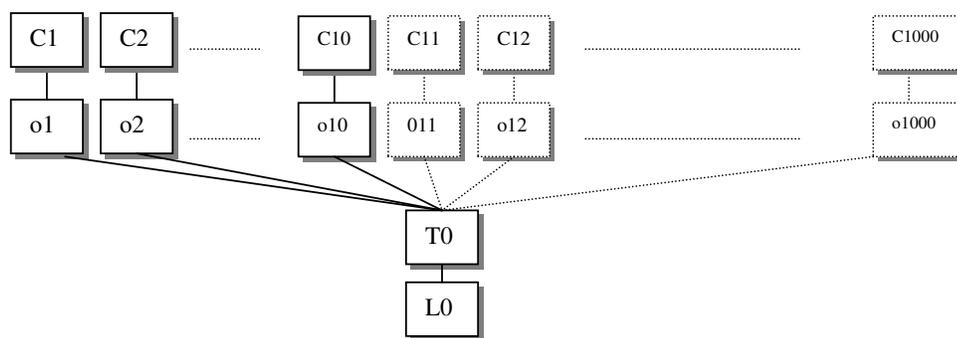
²⁵ Ceci vaut à la fois pour le *modèle des deux urnes* de Bostrom, l'*expérience de la loterie* de Leslie ou bien le *modèle des deux urnes* bicolore.

²⁶ On sait qu'un n-univers comportant plusieurs objets est un $\Omega\alpha$, car si les objets sont multiples, c'est qu'ils sont différents et donc qu'un *critère* donné α les différencie.

préfèrera donc afin de modéliser la situation correspondante. Et cela permet de modéliser la situation précédente dans un n-univers plus simple dépourvu de variable temporelle. On considèrera donc que la situation du *modèle des deux urnes* peut être modélisée dans un n-univers présentant une constante temporelle (T_0).

A ce stade, il s'avère nécessaire de prendre en compte le fait que les boules sont numérotées dans l'urne de Bostrom, alors que l'une des boules se trouve marquée afin de la distinguer des autres dans l'*expérience de la loterie* de Leslie. Il convient ainsi de déterminer quel critère permet de restituer ces caractéristiques. Il apparaît ici que l'on peut prendre en compte ces deux situations en considérant que la numérotation ou le marquage des boules s'effectue par l'utilisation d'un critère de couleur. Ainsi le n-univers correspondant est-il coloré ($\Omega C \dots$).

Considérons tout d'abord le *modèle des deux urnes* de Bostrom, où les boules sont numérotées. Une telle situation peut être modélisée dans un n-univers coloré, en considérant que le critère de couleur présente autant de taxons que de numéros de boules. Ceci correspond par exemple à un n-univers coloré qui comporte 1000 taxons de couleur. Il suffit de considérer par exemple 1000 nuances de gris. Dans ce type de situation, il apparaît que plusieurs objets différents ne peuvent avoir une même couleur. Ainsi, les multiples objets sont dans une relation *un* avec la couleur. On le voit, la situation correspondante est modélisée dans un $\Omega\alpha C$. La situation du *modèle des deux urnes* est alors modélisée dans un n-univers comportant de multiples objets colorés o_1, \dots, o_n , dans un lieu unique L_0 , à une position temporelle unique T_0 et où les objets sont dans une relation *un* avec la couleur. Les caractéristiques du n-univers où prend place le modèle de l'urne sont ainsi celles d'un $\Omega\alpha C$.



Considérons maintenant l'*expérience de la loterie* de Leslie. Leslie considère des boules dont l'une est individualisée et comporte le nom de l'observateur. Dans le *modèle des deux urnes* décrit par Bostrom, la boule tirée dans l'urne possède le numéro 5, alors que toutes les boules sont numérotées de 1 à n . Dans l'*expérience de la loterie* de Leslie, il suffit de considérer une seule boule dont le numéro est différent de celui des autres boules. Par commodité, on peut assimiler cette situation à l'urne de Bostrom, et modéliser de même cette situation dans le $\Omega\alpha C$, un n-univers à objets multiples atemporel, alocalisé et coloré, où les objets multiples sont en relation *un* avec la couleur²⁷.

3.2.2. Le modèle des deux urnes bicolore et le $\Omega\alpha^*C$

Une situation quelque peu différente est celle du *modèle des deux urnes bicolore*. Dans ce dernier cas, la situation correspondant au *modèle des deux urnes* est celle d'un n-univers à objets multiples atemporel, alocalisé et coloré comprenant seulement deux taxons de couleur: {rouge, vert}²⁸. Les

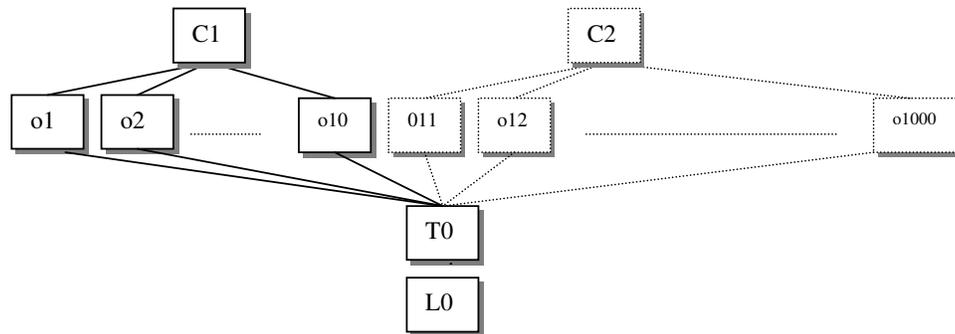
²⁷ Il y a place ici pour une autre interprétation de l'expérience de l'*expérience de la loterie* décrite par Leslie. On considère alors, de manière alternative, qu'il n'existe que deux catégories de boules: une seule boule comporte notre numéro gagnant, et les autres boules possédant un autre numéro perdant. Ainsi, on n'a que deux numéros portés sur les boules. Et cette variation peut alors être assimilée au *modèle des deux urnes* bicolore, qui fait l'objet de la section suivante.

²⁸ On peut bien sûr considérer un nombre de taxons plus important. Mais deux taxons sont suffisants pour modéliser la situation correspondant au *modèle des deux urnes*. Ici encore, on privilégie les n-univers les plus simples.

boules sont donc soit rouges, soit vertes. Dans cette hypothèse, il apparaît que plusieurs boules peuvent avoir une même couleur. On se situe donc dans un n-univers où les objets multiples sont dans une relation *plusieurs* avec le critère de couleur. On se place ainsi dans un $\Omega\alpha^*C$. On fait alors les hypothèses suivantes:

- (H1₁₀) l'urne contient 10 boules rouges
- (H2₁₁) l'urne contient 1000 boules dont 10 boules rouges et 990 boules vertes

avec les probabilités a priori $P(H1) = P(H2) = 0,5$. On considère l'événement E: une boule a été tirée et possède la couleur rouge. Il s'ensuit un décalage bayésien en faveur de l'hypothèse (H1).



Ici, une caractéristique intéressante de cette variation mérite d'être mentionnée. En effet, l'existence des 10 boules rouges est certaine, alors que l'existence des 990 boules vertes est hypothétique. Ainsi, on peut formuler les deux hypothèses, de manière alternative, en fonction de la présence éventuelle des boules vertes dans l'urne:

- (H1₁₂) l'urne ne contient aucune boule verte
- (H2₁₃) l'urne contient 990 boules vertes

A ce stade, il apparaît enfin intéressant de modéliser la situation correspondante en termes de n-univers purs. Il s'ensuit alors les deux hypothèses suivantes:

- (H1₁₄) le $\Omega\alpha^*C$ contient 10 objets rouges
- (H2₁₅) le $\Omega\alpha^*C$ contient 1000 objets dont 10 objets rouges et 990 objets verts

3.3. Modélisation dans d'autres n-univers

A ce stade, on peut se poser la question de savoir si le *modèle des deux urnes* peut être modélisé dans d'autres n-univers, présentant une structure différente du $\Omega\alpha C$ ou du $\Omega\alpha^*C$. De manière préliminaire, on peut s'interroger ici pour savoir si l'on peut modéliser le *modèle des deux urnes* dans un n-univers plus simple que le $\Omega\alpha C$, tel qu'un $\Omega\alpha$ où la variable-couleur se révèle absente. Il apparaît toutefois que non. En effet, dans le *modèle des deux urnes*, il est nécessaire de prendre en compte la couleur de la boule tirée (qui représente soit son numéro soit la marque qui la distingue des autres boules). Car la couleur joue ici le rôle d'un dispositif d'affichage, qui permet de modéliser l'événement E: une boule est extraite de l'urne qui possède le numéro 5 (ou la couleur n° 5). Et en l'absence de taxons de couleur, on ne peut restituer valablement l'événement E, et donc les conditions qui sont celles de l'expérience. Ainsi, le *modèle des deux urnes* ne peut valablement prendre place dans un $\Omega\alpha$.

Le *modèle des deux urnes*, on l'a vu, prend place dans un $\Omega\alpha C$ ou un $\Omega\alpha^*C$, c'est-à-dire dans un n-univers dont la structure est celle d'un $\Omega\alpha\beta$ ou d'un $\Omega\alpha^*\beta$. Ceci suggère que le *modèle des deux urnes*

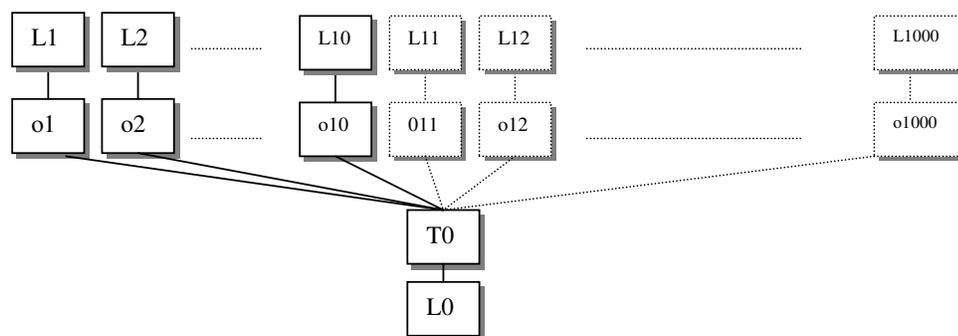
pourrait prendre place dans un autre n-univers où le critère β est différent du critère de couleur²⁹. On peut se demander ainsi si le *modèle des deux urnes* peut notamment prendre place dans un $\Omega\alpha L$ ou un $\Omega\alpha^*L$. Il convient de s'intéresser tour à tour à ces deux types de situations.

3.3.1. Modélisation du *modèle des deux urnes* dans le $\Omega\alpha L$

Considérons tout d'abord le cas d'un $\Omega\alpha L$. Une version du *modèle des deux urnes* peut-elle prendre place dans un $\Omega\alpha L$? Un tel n-univers comporte des objets multiples o_1, \dots, o_n et présente une variable de localisation. Le $\Omega\alpha L$ comporte de plus une constante de temps, de sorte que les objets existent à une position temporelle unique (T_0). En outre, dans un $\Omega\alpha L$, les objets sont en relation *un* avec le critère spatial. Il s'ensuit qu'un seul objet peut se trouver à une position spatiale donnée au moment unique T_0 . Enfin, on n'observe pas de démultiplication par rapport au lieu dans un tel univers, de sorte qu'un objet donné o_i ne peut occuper en T_0 qu'une position spatiale unique L_i . Pour cette raison, il existe dans le $\Omega\alpha L$ autant de positions spatiales³⁰ que d'objets. Soient donc L_1, L_2, \dots, L_n les positions spatiales correspondantes. Il apparaît ici que le modèle probabiliste correspondant est par exemple celui d'une roue qui comprend soit 10 soit 1000 cases et dont chacune des cases comprend une boule. Les cases correspondant à chaque boule sont masquées à l'exception d'une case qui constitue le numéro de la case correspondant à la boule qui a été tirée. On fait alors tourner la roue et lorsque celle-ci s'arrête, on regarde le numéro de la case correspondant à la boule qui est affichée. On fait alors les hypothèses suivantes:

- (H1₁₆) la roue contient 10 cases
- (H2₁₇) la roue contient 1000 cases

avec les probabilités a priori $P(H1) = P(H2) = 0,5$. Soit également l'événement E: la roue s'arrête à la position spatiale n° 5. On le voit, la position spatiale de la boule joue ici le rôle de dispositif d'affichage qui permet de connaître le numéro de la boule. Et cette fonction se révèle tout à fait équivalente à celle exercée par la couleur dans la version du *modèle des deux urnes* dans le $\Omega\alpha C$.



Il est à noter ici que l'on également la variation suivante formulée en termes de n-univers purs:

- (H1₁₈) le $\Omega\alpha L$ contient 10 objets en L_1-L_{10}
- (H2₁₉) le $\Omega\alpha L$ contient 1000 objets en L_1-L_{1000}

²⁹ Le raisonnement utilisé ici constitue une instance du principe méthodologique suivant, décrit dans l'introduction. On considère un n-univers Ω_1 dont la structure est $\Omega\alpha_1\alpha_2\dots\alpha_n$. On remplace alors l'un des critères α_i par un autre critère et il en résulte un n-univers Ω_2 . On teste alors si l'expérience de pensée E_2 qui est modélisée dans le Ω_2 produit un effet équivalent à l'expérience originale E_1 . Dans l'affirmative, on s'intéresse alors à la modélisation de la situation m dans le n-univers Ω_2 .

³⁰ On s'intéresse ici aux positions spatiales utiles. S'il existait d'autres positions spatiales non occupées par un objet, elles pourraient être purement et simplement ignorées.

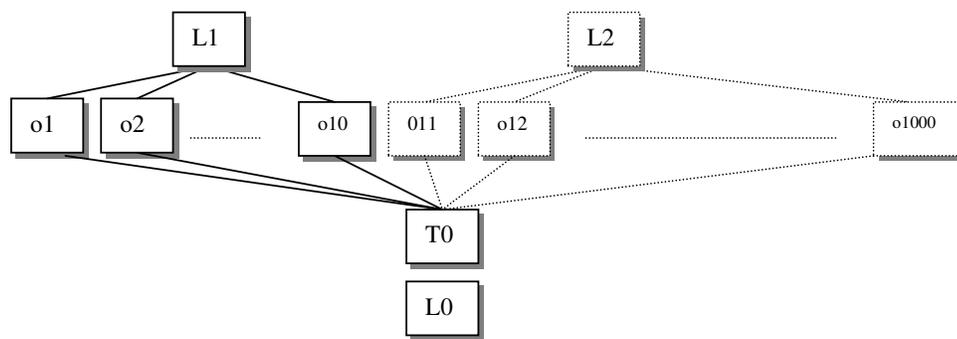
3.3.2. Modélisation du *modèle des deux urnes* dans le $\Omega\alpha^*L$

Peut-on maintenant modéliser la situation correspondant au *modèle des deux urnes* dans un $\Omega\alpha^*L$? Quelle serait donc la version correspondante du *modèle des deux urnes* dans un tel n-univers? Un $\Omega\alpha^*L$ comporte des objets multiples o_1, \dots, o_n et présente une constante de temps (T_0) ainsi qu'une variable de localisation déterminant plusieurs positions spatiales. Pour simplifier, on peut considérer un cas similaire au *modèle des deux urnes* bicolore où le $\Omega\alpha^*L$ ne comprend que deux positions spatiales: L_1 et L_2 . De plus, dans le $\Omega\alpha^*L$, les objets sont en relation *plusieurs* avec le critère de localisation. Ceci a pour conséquence que plusieurs objets peuvent se trouver à une position spatiale donnée au temps constant T_0 . Enfin, on n'observe pas de démultiplication par rapport au lieu dans le $\Omega\alpha^*L$, de sorte qu'un objet donné o_i ne peut occuper en T_0 qu'une position spatiale unique. Dans ce contexte, les hypothèses correspondantes sont les suivantes:

- (H1₂₀) l'urne contient 10 boules en L_1
- (H2₂₁) l'urne contient 1000 boules dont 10 boules en L_1 et 990 boules en L_2

On a de même l'événement E: je tire une boule dans l'urne et celle-ci se trouve en L_1 . On constate alors que l'on a bien dans cette situation un effet tout à fait équivalent à celui du *modèle des deux urnes* bicolore. Et on le voit, les hypothèses peuvent également être traduites en termes d'existence éventuelle de boules à la position spatiale L_2 :

- (H1₂₂) l'urne ne contient aucune boule en L_2
- (H2₂₃) l'urne contient 990 boules en L_2



A ce stade, on est à même de constater que l'expérience suivante (le *double amnésique*), également décrite par Leslie (1996, p. 209-10), met en scène une situation réelle qui présente tout à fait la structure du *modèle des deux urnes* dans le $\Omega\alpha^*L$:

Look again at your amnesiac self as you try to say where you are more likely to be, in London or in Little Puddle. You have nothing but population figures as a guide. To simplify things suppose that, much as all human observations must occur either before or after the infinitely brief instant of AD 2150, so also all humans must find themselves in Little Puddle or else in London., and you know it. You know, too, that the populations are fifty and ten millions, respectively. An appropriate model of the situation is an urn containing fifty balls marked 'Little Puddle' and ten millions marked 'London'. But now, what if you knew instead that your only possible location was either fifty-bodied Little Puddle, which definitely existed, or else ten-million-bodied London, which might be fictitious? Suppose you were markedly uncertain whether London was real. This, Carter and I suggest, can be interestingly comparable to knowing there have been humans before AD 2150, while being unsure that there will be any afterwards. Your situation now invites comparison with that of a man with an urn which he looks on as having, for instance, a 63 per cent probability of containing ten million 'London' balls in addition to the fifty marked 'Little Puddle' which he know it to contain. If you next discover that you are in Little Puddle then you should react as if you'd drawn a 'Little Puddle'

ball in such circumstances, which give excellent grounds for believing that the 'London' balls were imaginary. You have powerful new reasons for thinking London unreal³¹.

Leslie décrit ici une analogie avec une urne. Cette analogie met en relation une situation humaine concrète et un modèle probabiliste. Dans l'expérience du *double amnésique*, Leslie expose une situation concrète où une personne amnésique sait qu'elle se trouve soit à Petit-Bled, qui contient 50 habitants, soit à Londres, qui comprend 10 millions d'habitants. Le modèle probabiliste correspondant est celui d'une urne contenant soit 50 boules marquées Petit-Bled, soit 10 millions de boules marquées Londres. On a alors les hypothèses suivantes:

(H1₂₄) l'urne contient 50 boules en L₁ (Petit-Bled)

(H2₂₅) l'urne contient 10000050 boules dont 50 boules en L₁ (Petit-Bled) et 10000000 boules en L₂ (Londres)

qui peut être traduit par:

(H1₂₆) l'urne ne contient aucune boule en L₂

(H2₂₇) l'urne contient 10000000 boules en L₂

Ici, le fait de considérer que Londres n'a pas d'existence réelle³² correspond au fait que c'est l'hypothèse (H1) qui prévaut et que les boules marquées Londres n'existent pas dans l'urne. On le voit, cette dernière expérience décrite par Leslie correspond tout à fait à la version du *modèle des deux urnes* dans le $\Omega\alpha^*L$ qui vient d'être décrite.

Il est à noter que l'on également la variation suivante formulée en termes de n-univers purs:

(H1₂₈) le $\Omega\alpha^*L$ contient 10 objets en L₁

(H2₂₉) le $\Omega\alpha^*L$ contient 1000 objets dont 10 en L₁ et 990 en L₂

3.3.3. Modélisation du *modèle des deux urnes* dans le ΩL et l'expérience de la flèche

On l'a vu, le *modèle des deux urnes* prend place de manière standard dans le $\Omega\alpha C$. Bien qu'on ne puisse modéliser le *modèle des deux urnes* dans un $\Omega\alpha$, on a vu que l'expérience pouvait être

³¹ Trad.: Imaginez encore un double de vous amnésique alors que vous essayez de déterminer s'il est plus probable que vous vous trouvez à Londres ou bien à Petit-Bled. Vous ne possédez comme indice que les chiffres relatifs à la population. Pour simplifier les choses, supposez que, malgré tout toutes les observations humaines doivent se produire avant ou bien lors d'un l'instant infiniment bref après l'an 2150, de façon à ce que tous les humains doivent se trouver soit à Petit-Bled soit à Londres, et vous savez cela. Vous savez aussi que leurs populations sont respectivement de cinquante et dix millions. Un modèle convenable de la situation est ainsi une urne contenant cinquante boules marquées 'Petit-Bled' et dix millions marquées 'Londres'. Mais maintenant, que se passerait-il si vous saviez à la place que votre seul emplacement possible est soit à 50% de chances Petit-Bled, qui a certainement existé, ou Londres avec dix millions d'âmes, mais qui pourrait ne pas exister? Supposez que vous estimez très incertain que Londres existe. Ceci, ainsi que Carter et moi le suggérons, pourrait être ainsi comparable au fait de savoir qu'il y a eu des humains avant l'an 2150, tout en n'étant pas sûr qu'il y en aura après. Votre situation incite maintenant à la comparaison avec celle d'un homme avec une urne qui il regarde dessus comme ayant, par exemple, la probabilité de 63 pour cent de contenir dix millions de boules marquées 'Londres' en plus des cinquante marquées 'Petit-Bled' qu'il sait avec certitude que l'urne contient. Si vous découvrez après que vous êtes à Petit-Bled, alors vous devriez réagir comme si vous aviez tiré une boule marquées 'Petit-Bled' dans de telles circonstances, vous donnant ainsi d'excellentes raisons de croire que les boules marquées 'Londres' étaient imaginaires. Vous avez de fortes nouvelles raisons de considérer que Londres n'existe pas.

³² Cf. "thinking London unreal".

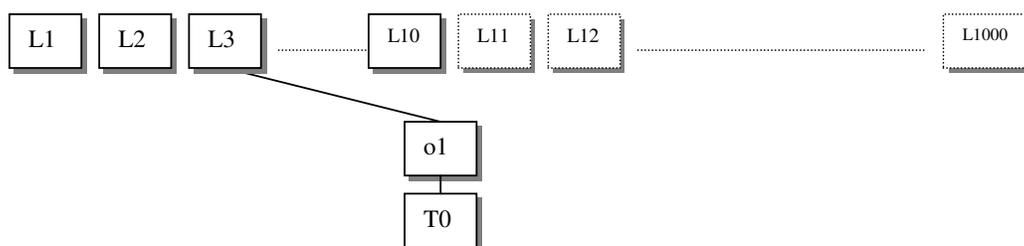
modélisée valablement dans un $\Omega\alpha L$. Le modèle probabiliste correspondant, on l'a vu, est alors celui d'une roue comportant plusieurs boules à différentes cases, dont l'une est finalement sélectionnée par un curseur. Une telle situation met en œuvre plusieurs boules et implique un univers à objets multiples. Mais cette situation apparaît très proche du modèle probabiliste qui est celui d'une roulette, où une boule unique est lancée. Intuitivement, une roulette prend place dans un n-univers à objet unique (la boule) et à positions spatiales multiples (les cases). Ceci suggère ainsi de vérifier si le *modèle des deux urnes* ne peut également être modélisé dans un ΩL . Considérons donc pour cela le modèle probabiliste suivant³³:

La roulette On considère une roulette qui comprend, de manière a priori équiprobable 10 cases ou bien 1000 cases. On lance la boule et la boule se trouve sur la case n° 5. A partir de là, je suis amené à réviser à la hausse ma probabilité initiale selon laquelle la roulette comprend seulement 10 cases.

On a ainsi les deux hypothèses suivantes:

- (H1₃₀) la roulette comporte 10 cases
- (H2₃₁) la roulette comporte 1000 cases

Ici, l'expérience met en scène un objet unique: la boule. De plus, l'expérience prend place à un moment unique T_0 . Le n-univers correspondant présente donc une constante temporelle. Et la boule apparaît comme l'objet qui détermine le tirage au sort de l'une des cases. Il s'ensuit que l'une des cases se trouve marquée (par la présence de la boule) alors que 9 cases ou bien 999 cases³⁴ ne sont pas marquées. On le voit, cette modélisation dans le ΩL préserve tout à fait l'effet qui est celui de la modélisation classique du *modèle des deux urnes* dans le $\Omega\alpha C$.



Il est frappant de constater ici à quel point la version du *modèle des deux urnes* qui prend place dans le ΩL est très proche de la variation du *modèle des deux urnes* qui est modélisée dans le $\Omega\alpha L$. Il apparaît ainsi que les 10 ou 1000 boules de la roue du $\Omega\alpha L$ correspondent aux 10 ou 1000 cases de la roulette dans le ΩL . Dans chaque cas on effectue un tirage aléatoire. Et l'événement E est alors matérialisé par la position spatiale de la boule sélectionnée dans le $\Omega\alpha L$ ou bien par la position spatiale sur laquelle s'arrête la boule dans le ΩL .

Pour éclairer davantage cette situation, il est également intéressant de considérer l'expérience de la flèche décrite par Leslie (1996, p. 209)³⁵:

³³ Ce modèle est directement suggéré par la modélisation du *modèle des deux urnes* dans le $\Omega\alpha L$ avec une roue. Les expériences modélisées en termes de roue ($\Omega\alpha L$) ou bien de roulette (ΩT) sont très proches.

³⁴ Selon l'hypothèse considérée.

³⁵ Cette expérience est présentée par Leslie dans un contexte où il répond aux arguments du type "urn analogies are inappropriate. We weren't given our birth times by a deity who pulled our souls from an urn at successive seconds and put them into human bodies.". Trad.: les "analogies avec l'urne sont inadéquates. Nos rangs de naissance ne nous ont pas été donnés par un Dieu qui a extrait nos âmes d'une urne seconde après seconde et les a incorporées dans des corps humains.".

You are hit by an arrow while walking around on a small island. Was this bad luck, or was the arrow aimed at you? If only luck was involved then this would be (as a rough approximation) as if the grid references of every square foot of the island had been put on slips of paper in an urn, your name being written on just one of them, etc³⁶.

Attachons-nous à décrire avec précision le n-univers dans lequel prend place l'expérience de la flèche. En premier lieu, il s'agit d'un univers présentant une variable spatiale, car la situation est basée sur le fait que l'on reçoit une flèche à un emplacement donné de l'île. Leslie se réfère explicitement à un univers localisé discret, en considérant que les différentes positions spatiales correspondent à une grille, où la structure des positions spatiales est celle d'une matrice, à la manière d'un jeu d'échecs. On se situe donc dans un ΩL ... D'autre part, on s'intéresse uniquement à la situation dans l'île à un moment donné. Aussi l'aspect temporel est-il absent. Ceci revient à considérer un n-univers comportant une constante temporelle. Enfin, une telle situation met en scène un objet unique: la flèche. A ce stade, on peut s'interroger afin de savoir à quoi correspondent les deux hypothèses décrites par Leslie: (a) la flèche a été envoyée au hasard; (b) la flèche a été envoyée à dessein. La situation présuppose un univers complexe, mais on peut en simplifier les données tout en s'attachant à en conserver l'essentiel. Considérons tout d'abord le cas où la flèche a été envoyée au hasard. Dans ce cas, le fait de recevoir la flèche s'analyse comme de la malchance, mais peut aussi bien se produire sur une île dont le territoire est petit. Par contre, si l'île possède un vaste territoire, le fait de recevoir la flèche suggère ici qu'elle a été envoyée à dessein. Ainsi, ces deux situations peuvent être traduites par les deux hypothèses suivantes:

(H1₃₂) une flèche a été tirée sur l'île qui comprend 10 kilomètres carrés

(H2₃₃) une flèche a été tirée sur l'île qui comprend 1000 kilomètres carrés

On le voit, une telle situation peut être valablement modélisée dans un ΩL , un n-univers à objet unique comportant une constante temporelle et une variable spatiale.

Le modèle probabiliste correspondant à la situation de l'expérience de la flèche est d'autre part suggéré par Leslie. Leslie suggère en effet un modèle qui est celui d'une urne contenant des morceaux de papier dont l'un comporte notre nom³⁷. Pour reprendre le modèle de l'urne utilisé jusqu'à présent, on peut modéliser la situation correspondante en considérant que l'une des boules porte la marque d'une flèche (ou notre nom comme Leslie le suggère) alors que toutes les autres boules ne comportent aucune marque. Il s'agit donc d'un n-univers à objets multiples. Ici, cette situation peut être modélisée en considérant que la présence de la flèche ou son absence sur une boule correspond à deux taxons de couleur dans un n-univers coloré. Il apparaît ainsi que le modèle correspondant prend place dans un $\Omega\alpha^*C$. Et cette situation est alors tout à fait celle du *modèle des deux urnes bicolore*³⁸.

L'expérience de la flèche est particulièrement intéressante et illustre mieux que toute autre l'expertise que possède Leslie pour mettre en œuvre des situations concrètes et déterminer le modèle probabiliste correspondant. L'expérience de la flèche est en effet basée sur une analogie avec une urne. Dans les analogies classiques avec une urne, une situation donnée, modélisée dans un n-univers donné, est assimilée à un modèle probabiliste qui prend place dans le même n-univers. Mais dans l'expérience de la flèche, il en va autrement. En effet, la situation correspondante est celle d'un n-univers à objet

³⁶ Trad.: Vous êtes heurté par une flèche tout en marchant autour d'une petite île. S'agit-il de malchance, ou bien la flèche était-elle dirigée intentionnellement vers vous? S'il s'agissait seulement de chance alors ce serait (de manière approximative) comme si les références dans une grille de chaque mètre carré de l'île avaient été inscrites sur des bouts de papier placés dans une urne, votre nom étant écrit sur seulement l'un d'eux, etc...

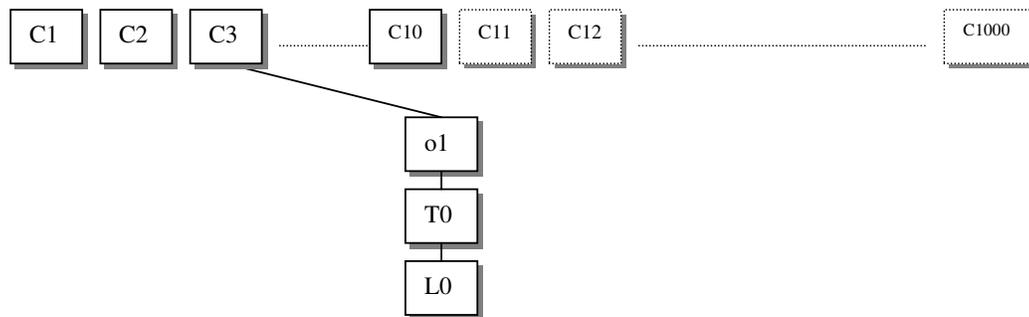
³⁷ Cf. "this would be (...) as if the grid references of every square foot of the island had been put on slips of paper in an urn, your name being written on just one of them, etc.". Trad.: "ce serait (...) comme si les références dans une grille de chaque mètre carré de l'île avaient été inscrites sur des bouts de papier placés dans une urne, votre nom étant écrit sur seulement l'un d'eux, etc..."

³⁸ On peut bien sûr modéliser également cette situation dans un *modèle des deux urnes* classique dans un $\Omega\alpha C$.

unique atemporel et localisé, un Ω_L . Mais Leslie suggère que le modèle probabiliste qui correspond à cette situation concrète est celui d'un $\Omega\alpha^*C$. On le voit, l'expérience comporte ce qui s'analyse comme une transposition d'un n-univers, le Ω_L , à un autre, le $\Omega\alpha^*C$. Cette transposition est basée sur la transformation des taxons de localisation en objets multiples. Et on constate de même l'équivalence entre d'une part la présence ou l'absence de la boule à une position spatiale donnée dans le Ω_L , et d'autre part le fait que la boule présente l'un ou l'autre des deux taxons de couleur dans le $\Omega\alpha^*C$. Il apparaît ainsi finalement que l'expérience de la flèche s'analyse comme une analogie avec une urne qui comporte une transposition d'un n-univers à un autre.

3.3.4. Modélisation du modèle des deux urnes dans un ΩC

On vient de le voir, le *modèle des deux urnes* peut être modélisé valablement dans un Ω_L , un n-univers à objet unique et comportant une variable spatiale. Un tel n-univers présente la structure $\Omega\beta$. Ceci suggère que le *modèle des deux urnes* peut également être modélisé notamment dans un ΩC . Dans ce cas, la variation correspondante est la suivante. On considère une boule unique, qui est susceptible de changer de couleur. La boule peut en effet prendre 10 couleurs différentes ou bien 1000, en fonction d'un tirage aléatoire. On a un code de couleurs correspondant à 1000 taxons de couleurs. Le tirage aléatoire a lieu et la boule prend la couleur n° 3. Compte tenu de cela, je révisé à la hausse ma probabilité initiale selon laquelle la boule est susceptible de prendre seulement 10 couleurs différentes.



Ainsi, on vérifie que l'expérience du *modèle des deux urnes* peut être valablement modélisée dans un ΩC . Par conséquent, les modélisations du *modèle des deux urnes* dans le Ω_L ou le ΩC confirment que le *modèle des deux urnes* peut être modélisée de manière classique dans un n-univers à objets multiples, mais également dans un univers à objet unique.

3.4. La version intégrale du modèle des deux urnes

Afin de permettre ultérieurement une meilleure comparaison entre le *modèle des deux urnes* et la situation correspondant à DA, il convient de préciser davantage les détails de la description du *modèle des deux urnes*. L'expérience comporte en effet la mention de l'événement E, qui consiste dans le fait que la boule n° 5 est extraite de l'urne. Ce tirage entraîne l'application du modèle à minoration, c'est-à-dire un décalage bayésien en faveur de l'hypothèse (H1). Mais l'événement E est-il le seul type d'événement susceptible de se produire dans la situation qui est celle du *modèle des deux urnes*? Il apparaît que non. En effet, un autre type d'événement – appelons-le F - d'une nature fondamentalement différente est susceptible de se produire. Il est ainsi possible de tirer dans l'urne la boule n° 400. Et l'événement F a alors pour effet de confirmer définitivement l'hypothèse (H2) et de falsifier l'hypothèse (H1). Ainsi, on est à même de décrire de manière plus précise une version plus complète du *modèle des deux urnes* où deux types d'événements, E et F sont susceptibles de se produire. Ceci

conduit à décrire ce qu'on peut appeler une version intégrale du *modèle des deux urnes*, où on distingue nettement entre les deux types d'événements:

Version intégrale du modèle des deux urnes Une urne se trouve en face de moi. Je sais qu'elle contient 10 boules ou bien 1000. Les boules sont numérotées 1, 2, 3, 4 ... etc.

Etape (a) Je tire une boule au hasard dans l'urne.

Etape (b) Je découvre le numéro de la boule. Ici, deux types d'événements sont susceptibles de se présenter. Soit l'événement E: je tire dans l'urne la boule n° 5. Dans ce premier cas, je conclus à un décalage bayésien en faveur de l'hypothèse selon laquelle l'urne ne contient que 10 boules. Soit également l'événement F: je tire dans l'urne la boule n° 400. Dans ce second cas, l'hypothèse selon laquelle l'urne ne contient que 10 boules se trouve falsifiée, et l'hypothèse selon laquelle elle contient 1000 boules est confirmée³⁹.

Cette dernière version apparaît ainsi comme une version qui décrit la situation correspondante de manière plus complète que la version classique du *modèle des deux urnes*. Dans ce contexte, la version classique du *modèle des deux urnes* apparaît comme une version tronquée. En distinguant entre les deux types d'événements susceptibles de survenir, la version intégrale du *modèle des deux urnes* met en lumière les conséquences de nature différente qui résultent des deux types d'événements susceptibles de se produire. Dans un cas, il s'ensuit une révision à la hausse d'une probabilité initiale. Dans l'autre cas, il en résulte la confirmation radicale d'une hypothèse et l'infirmité de l'hypothèse alternative.

3.5. Le modèle des deux urnes à décalage (version intégrale)

Considérons maintenant l'expérience suivante, qui constitue une variation de la version intégrale du *modèle des deux urnes*:

Modèle des deux urnes à formule d'incrémentatation (version intégrale) Soient les deux hypothèses:

(H1₃₄) une urne contient 10 boules numérotées

(H2₃₅) une urne contient 1000 boules numérotées

pour lesquelles les probabilités initiales P(H1) et P(H2) sont égales à 0,5. Avant de procéder au tirage d'une boule dans l'urne, on modifie les numéros de chacune des boules de la façon suivante. Soient n le nombre de boules dans l'urne, m le numéro original de la boule et m^* son numéro après incrémentatation. Dans ce dernier cas $m^* = m + 1 \pmod{n}$. Ainsi, si (H1) alors $m^* = m + 1 \pmod{10}$. Et de même si (H2) alors $m^* = m + 1 \pmod{1000}$. Il s'ensuit une incrémentatation (modulo n) des numéros de boules. Maintenant, est-ce que je me trouve dans une situation de même nature que dans la version intégrale du *modèle des deux urnes* et puis-je accepter de la même façon les conclusions qui résultent des deux types d'événements E* et F*, susceptibles de survenir?

Il apparaît en effet ici que deux types d'événements peuvent survenir lors du tirage d'une boule dans l'urne. D'une part un événement de type E: je tire une boule dont le numéro est 5. Dans ce cas, si l'événement E* survient, je tire alors la boule n° 6. Et d'autre part un événement de type F: je tire une boule dont le numéro est 300. Ainsi, si l'événement F* survient, je tire alors la boule n° 301. Ici, il apparaît que cette expérience est tout à fait identique à la version intégrale du *modèle des deux urnes*, à cette différence près que le numéro de chaque boule se trouve incrémenté (modulo n). On peut penser raisonnablement dans ce cas que malgré l'utilisation de la formule à incrémentatation, quiconque accepte la version intégrale du *modèle des deux urnes* dans sa formulation classique accepte également les conclusions qui résultent de cette version modifiée.

³⁹ La version intégrale présentée ici distingue de plus entre les étapes (a) et (b), de la même manière que l'expérience du *pile ou face divin* décrite par Bostrom.

Considérons maintenant l'expérience suivante, qui est basée sur les mêmes hypothèses et les mêmes probabilités initiales que précédemment, mais où le décalage s'effectue différemment:

Modèle des deux urnes à décalage (version intégrale) On procède de même que dans l'expérience précédente, mais avant de procéder au tirage d'une boule dans l'urne, on crée un décalage en choisissant une valeur quelconque de décalage. Pour cela, on modifie ainsi les numéros de boule de la façon suivante: $m^* = m + p \pmod{n}$ où p est un entier naturel non nul. Il s'ensuit alors. Si (H1) alors $m^* = m + p \pmod{10}$. Et si (H2) alors $m^* = m + p \pmod{1000}$. On procède alors au tirage de la boule dans l'urne. Suis-je alors prêt à en accepter les conclusions qui résultent des deux types d'événements E^* et F^* , susceptibles de survenir, de la même façon que j'accepte celles du *modèle des deux urnes* dans sa version intégrale?

Ici, il apparaît que cette expérience fait appel à une formule de décalage plus générale que la précédente. Il s'agit ainsi d'une version plus générale que le *modèle des deux urnes* à incrémentation, où le pas est constitué par une valeur entière quelconque. Dans cette situation, il apparaît que je peux raisonner ainsi. Avec la numérotation originale des boules, les boules sont numérotées 1, 2, 3, 4, Après décalage, même si toutes les boules ont changé de numéro, les boules sont toujours numérotées 1, 2, 3, 4, Ceci vaut quelle que soit la valeur de p . Par conséquent, je me trouve dans la même situation que dans la version intégrale classique du *modèle des deux urnes*. Les deux types d'événements E^* et F^* peuvent survenir, de même que E et F peuvent survenir dans la version intégrale du *modèle des deux urnes*. Ainsi dans ce cas, malgré le décalage des numéros de boule dans l'urne préalablement au tirage, il apparaît toujours rationnel, pour quiconque accepte le *modèle des deux urnes*, d'accepter les conclusions qui résultent du *modèle des deux urnes* à décalage.

Les versions du *modèle des deux urnes* à incrémentation et à décalage qui viennent d'être mentionnées font appel à des boules numérotées. Ces variations prennent donc place dans le $\Omega\alpha C$, un n -univers où les objets sont en relation *un* avec la couleur. Il s'avère toutefois intéressant de décrire la version d'une telle expérience dans le contexte qui est celui du *modèle des deux urnes* bicolore. Il apparaît en effet que l'on peut également mettre en évidence une version du *modèle des deux urnes* à décalage en considérant des boules de couleur à la place de boules numérotées. Une telle variation prend place dans le $\Omega\alpha^*C$, un n -univers où les objets sont en relation *plusieurs* avec la couleur. Dans cette variation, on considère une taxinomie de couleurs qui ne comporte que deux taxons: {rouge, vert}. On a alors les deux hypothèses:

(H1₃₆) l'urne contient 10 boules (rouges)

(H2₃₇) l'urne contient 1000 boules (dont 10 boules rouges et 990 boules vertes)

Considérons maintenant le mécanisme du décalage. Imaginons que les boules possèdent un numéro et une couleur initiale correspondant à ce numéro. Ainsi à l'origine, les boules n° 1 à 10 sont rouges, et les boules n° 11 à 1000 (dans l'hypothèse où l'urne contient 1000 boules) sont vertes. Chacun des numéros permet ensuite de transformer la couleur initiale en un nouveau taxon de couleur. La transformation s'effectue de la manière suivante. Soit m le numéro initial d'une boule et m^* son numéro après décalage. Soit p un entier naturel non nul, qui constitue la valeur de décalage. On modifie ainsi les couleurs des boules de la manière suivante: $m^* = m + p \pmod{n}$. Il s'ensuit ainsi. Si (H1) alors $m^* = m + p \pmod{10}$. Et si (H2) alors $m^* = m + p \pmod{1000}$. Les nouveaux taxons de couleurs sont alors attribués de la manière suivante. Après décalage, dans l'hypothèse (H1), la couleur des 10 boules est inchangée et l'urne contient toujours 10 boules rouges. Dans l'hypothèse (H2), l'urne contient toujours 10 boules rouges et 990 boules vertes. En revanche, il apparaît que si $p \geq 10$ ⁴⁰, les 10 boules qui étaient rouges avant décalage sont désormais vertes. Et réciproquement, 10 boules parmi celles qui étaient vertes avant décalage sont désormais rouges. Ainsi, après décalage, on a toujours 10 boules rouges et 990 boules vertes (dans le cas où l'hypothèse (H2) est vérifiée) dans l'urne. Compte

⁴⁰ Si $p < 10$, alors ce sont p boules initialement rouges qui deviennent vertes, et p boules initialement vertes qui deviennent rouges.

tenu de cette dernière propriété, il apparaît donc également ici que quiconque accepte les conclusions du *modèle des deux urnes* bicolore (version intégrale) accepte également les conclusions qui résultent du *modèle des deux urnes* à décalage (version intégrale) dans le $\Omega\alpha^*C$.

Allons plus loin et considérons maintenant la variation suivante de l'expérience précédente:

Modèle des deux urnes à formule de décalage (version intégrale) De même que dans les deux expériences précédentes, les boules qui se trouvent dans l'urne sont initialement numérotées. Avant de procéder au tirage d'une boule dans l'urne, on remplace les numéros initiaux par une formule de calcul, de la manière suivante.⁴¹ On modifie les numéros des boules en inscrivant sur chacune des boules à la place du numéro initial m la formule de calcul suivante: " $n - m$ ". De la sorte, les boules comportent désormais les inscriptions: $n - 1, n - 2, n - 3, \dots, n - 10$ dans l'hypothèse (H1). Et elles comportent de même les inscriptions $n - 1, n - 2, n - 3, \dots, n - 1000$ dans l'hypothèse (H2). On crée ainsi une formule de décalage en recalculant le numéro initial m^* de la manière suivante: $m^* = n - m$. Maintenant, avant de procéder au tirage de la boule dans l'urne et de connaître la valeur de n , avez-vous une confiance absolue dans les conclusions qui résulteront des deux types d'événements E^* et F^* susceptibles de survenir dans cette expérience, de même que vous l'aviez dans les conclusions de la version intégrale du *modèle des deux urnes*?

Dans cette nouvelle expérience, on observe une différence importante avec les protocoles précédents., Il apparaît en effet ici que l'on tire dans l'urne une boule qui comporte une formule de calcul, et non un numéro.

Avec ce nouveau protocole, il apparaît que je peux raisonner ainsi. Je sais que les boules comportent les formules $n - 1, n - 2, n - 3, n - 4, \dots$. Avoir un jeu de boules qui comporte les valeurs 1, 2, 3, 4, ... ou bien alors $n - 1, n - 2, n - 3, n - 4, \dots$ se révèle tout à fait équivalent. La seule différence est finalement dans les symboles " $n -$ " qui précèdent les nombres 1, 2, 3, 4, ... dans la seconde numérotation. Or j'ai une confiance absolue dans les conclusions qui s'ensuivent dans la version intégrale du *modèle des deux urnes*, avec les boules numérotées 1, 2, 3, 4, De plus, une fois la valeur de n connue, les nouveaux numéros de boules seront calculés et je me retrouverai dans la même situation qu'à l'origine, avec une urne comportant des boules numérotées 1, 2, 3, 4.... J'ai donc également une confiance absolue dans les conclusions qui résulteront des deux types d'événements E^* et F^* avec les boules numérotées $n - 1, n - 2, n - 3, n - 4, \dots$. Par conséquent, dans une telle hypothèse, j'accepte en totalité les deux types de raisonnements possibles, qui seront déterminés par la connaissance du numéro de la boule: soit (a) l'invalidation de (H1) si le numéro calculé de la boule est supérieur à 100; soit (b) le décalage bayésien et l'application du modèle à minoration si le numéro calculé de la boule est inférieur ou égal à 10. Ainsi, il apparaît que quiconque accepte les résultats de la version intégrale du *modèle des deux urnes* accepte également les résultats de la version intégrale du *modèle des deux urnes* à formule de décalage.

⁴¹ On a une version plus générale de cette expérience en considérant la variation suivante. On choisit un entier naturel p tel que $1 \leq p \leq n$. Puis on modifie les numéros de boule en inscrivant sur chacune des boules à la place du numéro initial la formule de calcul suivante: $m^* = p - m \pmod{n}$. L'utilisation de la formule $m^* = m + p \pmod{n}$ conduit bien sûr à des résultats équivalents.

Chapitre 4. La situation correspondant à DA

DA, on l'a vu, est basé classiquement sur une analogie entre le *modèle des deux urnes* et la situation humaine correspondante. Il convient donc à ce stade de s'attacher à modéliser en termes de n -univers la situation de l'humanité qui correspond à DA.

4.1. Modélisation de la situation correspondant à DA dans le n -univers temporel à objets multiples ($\Omega\alpha^*T^*C$)

Considérons tout d'abord la description de la situation correspondant à DA qui est fournie par Bostrom (1997):

(...) consider the case where (...) you have two possible human races, and (...) you have individuals, ranked according to birth order. As a matter of fact, you happen to find that your rank is about sixty billion. Now, say Carter and Leslie, we should reason in the same way as we did with the urns. That you should have a rank of sixty billion or so is much more likely if only 100 billion persons will ever have lived than if there will be many trillion persons. Therefore, by Bayes' theorem, you should update your beliefs about mankind's prospects and realise that an impending doomsday is much more probable than you have hitherto thought⁴².

Chaque humain possède ainsi un numéro qui lui est attribué en fonction de son rang de naissance. Et on a de plus les deux hypothèses suivantes:

(H1₃₈) l'humanité comptera au total 100 milliards d'humains

(H2₃₉) l'humanité comptera au total 10000 milliards d'humains

Enfin, l'événement E est le suivant: mon numéro de naissance individuel est 60×10^9 . Il s'ensuit un décalage bayésien en faveur de l'hypothèse (H1).

Il convient maintenant de modéliser la situation de l'humanité correspondant à DA en termes de n -univers. Quelles sont donc les caractéristiques du n -univers dans lequel prend place la situation humaine correspondant à DA? En premier lieu, cette situation se caractérise par la présence multiple d'humains. On se situe donc dans un n -univers à objets multiples ($\Omega\alpha\dots$). En second lieu, on peut considérer que les humains occupent un lieu unique: la Terre⁴³. Le n -univers correspondant présente donc une constante spatiale (L_0). En troisième lieu, il convient ici également de prendre en considération le fait que les humains possèdent un rang de naissance (*birth rank*) et que chaque humain possède un numéro unique. Ceci peut être pris en compte, comme on l'a vu précédemment avec le *modèle des deux urnes*, en considérant que la numérotation des humains s'effectue par l'utilisation de taxons de couleur: on considère ainsi une taxinomie de couleurs dont les taxons sont numérotés de 1 à n . Ainsi, le n -univers correspondant est-il également coloré. Et le fait que chaque humain possède un numéro de naissance unique se traduit par le fait que, dans le n -univers correspondant, les objets sont en relation *un* avec la couleur. A ce stade, on se situe donc dans un $\Omega\alpha C\dots$. Enfin, l'univers correspondant apparaît comme temporel ($\Omega T\dots$), puisque les humains vivent et meurent à une époque donnée. On peut donc modéliser la situation correspondante dans un n -univers qui comporte n positions temporelles: T_1, T_2, \dots, T_n . De plus, un humain donné occupe

⁴² Trad.: (...) considérez le cas où (...) l'on a deux races humaines possibles, et (...) où les individus sont rangés selon l'ordre de leur naissance. En fait, vous constatez que votre rang de naissance est environ soixante milliards. Maintenant, disent Carter et Leslie, nous devrions raisonner comme nous l'avons fait avec les urnes. Que vous ayez un rang de naissance de soixante milliards environ est beaucoup plus probable si seulement 100 milliards de personnes auront jamais vécu, par rapport au cas où il y en aura de nombreux milliards de milliards. Par conséquent, à l'aide du théorème de Bayes, vous devriez réévaluer votre croyance au sujet des perspectives d'avenir de l'humanité et vous vous rendez compte qu'une apocalypse imminente est beaucoup plus probable que vous ne l'avez jusqu'ici envisagé.

⁴³ Le fait d'attribuer, de manière alternative, une position spatiale unique à chaque humain n'entraîne pas de modification, dans ce qui suit.

plusieurs positions temporelles successives: il naît à une position temporelle donnée T_i , vit pendant les positions temporelles suivantes et meurt finalement à une position temporelle T_j (avec $i < j$). La durée de vie de cet humain s'établit ainsi sur la période T_i - T_j . Et le n-univers correspondant est donc à démultiplication temporelle⁴⁴ (ΩT^* ...). En outre, il convient d'observer que plusieurs humains peuvent se trouver à une position temporelle T_i donnée. Les objets sont donc en relation *plusieurs* avec le critère temporel. Ainsi, le n-univers correspondant est un $\Omega\alpha^*T^*$ En conséquence, il apparaît que la situation de l'humanité correspondant à DA peut être modélisée dans un $\Omega\alpha^*T^*C$. Une telle situation prend ainsi place dans un n-univers à objets multiples, comportant à la fois une constante de localisation, une variable temporelle et une variable de couleur, où les objets sont à démultiplication temporelle et sont en relation *plusieurs* avec le critère du temps. On peut observer ici qu'une caractéristique importante de cette situation est que le nombre d'humains vivant à un moment donné est variable. Ainsi, à une position temporelle T_i donnée, on a l'existence de n_i humains. On a ainsi: n_1 humains en T_1 , n_2 humains en T_2 , ..., n_n humains en T_n ⁴⁵.

4.2. Variations de la situation humaine correspondant à DA dans d'autres n-univers

Il est intéressant, à ce stade, de s'intéresser à des variations de la situation humaine correspondant à DA. Dans ce contexte, on étudiera en particulier des variations de cette dernière situation où la situation de l'humanité dans DA se trouve légèrement simplifiée, tout en préservant l'essentiel de l'analogie qui préside à DA. On s'intéressera ainsi en particulier à des sous-univers du $\Omega\alpha^*T^*C$ et aux situations humaines qui leur sont associées.

4.2.1. Variation dans un $\Omega\alpha^*TC$

On l'a vu, le n-univers décrivant la situation humaine correspondant à DA est un $\Omega\alpha^*T^*C$. Ainsi, le n-univers correspondant est-il à démultiplication temporelle (ΩT^* ...). Est-il cependant possible de modéliser une variation de la situation correspondant à DA dans un n-univers qui ne présente pas une telle propriété de démultiplication temporelle, tout en conservant un effet équivalent, c'est-à-dire l'analogie intuitive avec le *modèle des deux urnes* sur laquelle est basé DA? Dans ce cas, le n-univers correspondant serait un $\Omega\alpha^*TC$, un n-univers à objets multiples, comportant à la fois une constante de localisation, une variable temporelle et une variable de couleur, où les objets ne sont pas à démultiplication temporelle et sont en relation *plusieurs* avec le critère du temps. Quelle serait alors la situation humaine correspondante? Il apparaît ici qu'une telle situation correspond à celle où chaque humain occupe une position temporelle unique, correspondant par exemple au moment précis de sa naissance. Dans ce cas, on occulte le fait que chaque humain possède une durée de vie qui s'établit sur l'intervalle T_i - T_j . Et on considère pour simplifier, que chaque humain occupe une position temporelle donnée T_i . On le voit, une telle variation permet de conserver l'intuition qui préside à DA. La situation humaine correspondante se révèle suffisante pour permettre au problème inhérent à DA de prendre place. Ainsi, il apparaît que si on prend en compte une situation humaine correspondant à un $\Omega\alpha^*TC$, on constate que l'analogie intuitive avec l'urne qui préside à DA est toujours présente.

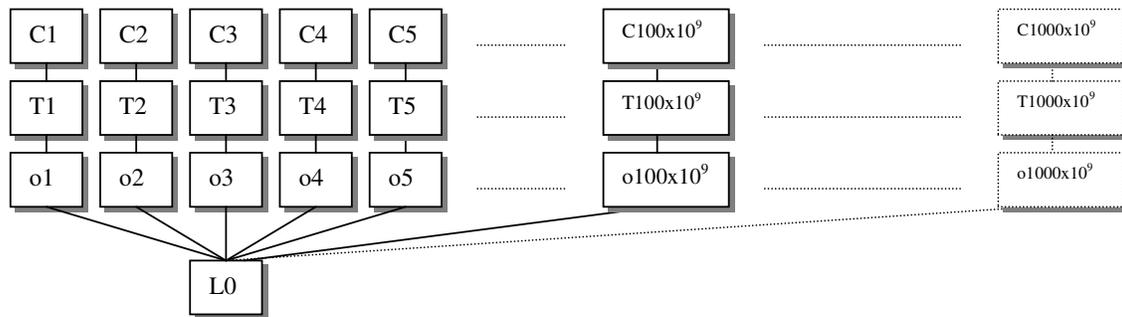
4.2.2. Variation dans un $\Omega\alpha TC$

Considérons maintenant une autre variation. La situation humaine correspondant à DA prend place dans un $\Omega\alpha^*T^*C$. Et on peut de même, on vient de le voir, considérer une variation qui en préserve l'effet, dans un $\Omega\alpha^*TC$. Mais est-il possible à ce stade de considérer, sans perte d'analogie, une variation de la situation humaine correspondante à DA qui prenne place dans un $\Omega\alpha TC$, où les objets sont cette fois en relation *un* avec le critère temporel. En effet, il apparaît ici que l'on peut imaginer une situation humaine où plusieurs humains ne peuvent occuper une même position temporelle. Il

⁴⁴ La persistance constitue ici un cas particulier de démultiplication temporelle.

⁴⁵ Ici, il est utile de rappeler que l'on se place dans l'hypothèse d'un univers déterministe. Soit T_m le moment présent. Dans un univers déterministe, les n_j en T_j ($j > m$) sont déterminés à l'avance, c'est-à-dire que la population d'humains future est d'ores et déjà fixée. A l'inverse, dans un univers indéterministe, les n_j en T_j ($j > m$) ne sont pas déterminés à l'avance.

suffit ainsi de considérer que le moment de la naissance de chaque être humain est mesuré avec une précision telle⁴⁶ que deux humains ne peuvent naître au même moment. Il s'ensuit alors qu'un moment unique est associé à la naissance de chaque être humain. On constate alors qu'une telle simplification permet toujours de capturer l'intuition qui préside à notre compréhension de la situation de l'humanité, dans des conditions telles que l'analogie avec l'urne est préservée. Il résulte de ce qui précède que la situation correspondant à DA peut être simplifiée en considérant une situation où un seul humain occupe une position temporelle donnée, qui correspond par exemple au moment extrêmement précis de sa naissance. Le numéro de naissance attribué à chaque humain correspond à ce moment très précis. Dès lors, le n-univers correspondant est simplement temporel ($\Omega T \dots$). Et une telle situation peut alors être modélisée, en lieu et place d'un $\Omega \alpha * T * C$, dans un n-univers à objets multiples, comportant une constante de localisation ainsi qu'une variable temporelle et une variable de couleur, c'est-à-dire un $\Omega \alpha TC$ ⁴⁷.



⁴⁶ Par exemple avec une précision de 10^{-10} s.

⁴⁷ On peut être tenté à ce stade de se débarrasser de la variable temporelle. Cela est-il possible? Il apparaît toutefois que non. En effet, dans la situation correspondant à DA, il est clair que les humains possèdent un numéro correspondant à leur rang de naissance. Si l'on perdait la temporalité, on perdrait donc le rang de naissance des humains et une caractéristique importante de la situation humaine. Ainsi, on ne peut réduire la situation correspondant à DA à un n-univers à objets multiples et coloré. La variable temporelle dans la situation correspondant à DA ne saurait être évitée.

Chapitre 5. L'expérience des émeraudes et l'expérience Petit-Bled/Londres

5.1. Description

Ainsi que cela a été mentionné plus haut, la portée des écrits de Leslie relatifs à DA s'étend bien au-delà de la seule analyse de DA. Leslie met en effet en oeuvre un principe méthodologique qui consiste à fonder un raisonnement concernant des situations réelles sur des modèles probabilistes, en se basant sur l'analogie existant entre les deux situations. Il en résulte souvent une argumentation très convaincante. Il s'agit là de l'exploration d'un champ spécifique de la philosophie qui se révèle très prometteur, puisqu'il ouvre la voie à des applications concrètes. Leslie mentionne par ailleurs d'autres types d'expériences, dont la nature se révèle fondamentalement différente de l'*expérience de la loterie*⁴⁸. Parmi ces dernières, on peut notamment citer l'*expérience des émeraudes* et l'*expérience Petit-Bled/Londres*. L'analyse révèle que ces deux types d'expériences présentent finalement une structure identique, qu'il convient maintenant de s'attacher à mettre en évidence. Considérons tout d'abord l'expérience suivante décrite par Leslie⁴⁹:

Imagine an experiment planned as follows. At some point in time, three humans would each be given an emerald. Several centuries afterwards, when a completely different set of humans was alive, five thousands humans would again each be given an emerald in the experiment. You have no knowledge, however, of whether your century is the earlier century in which just three people were to be in this situation, or the later century in which five thousand were to be in it. Do you say to yourself that if yours were the earlier century then the five thousand people *wouldn't be alive yet*, and that therefore you'd have no chance of being among them? On this basis, do you conclude that you might just as well bet that you lived in the earlier century?

Suppose you in fact betted that you lived there. If every emerald-getter in the experiment betted in this way, there would be five thousand losers and only three winners. The sensible bet, therefore, is that yours is instead the later century of the two⁵⁰.

Appelons cette expérience l'*expérience des émeraudes*. Leslie met ainsi en parallèle une situation réelle concernant des émeraudes et un modèle concernant des boules dans une urne. La structure de l'*expérience des émeraudes* est donc également celle d'une analogie avec une urne. L'*expérience des émeraudes* décrit ainsi une situation réelle, concrète, qu'il conviendra de modéliser en termes de n-univers. Dans l'énoncé de l'expérience, le modèle probabiliste correspondant n'est toutefois pas décrit de manière explicite, mais s'avère simplement suggéré. Il s'agit là d'une différence avec DA. Dans l'*expérience des émeraudes*, l'analogie avec l'urne est seulement implicite. On pourra toutefois rendre explicite le modèle probabiliste sur lequel est basé l'*expérience des émeraudes*. A partir de la description détaillée du n-univers correspondant à la situation des émeraudes, il conviendra donc de décrire explicitement le modèle probabiliste correspondant à l'*expérience des émeraudes*.

⁴⁸ La version de Leslie du *modèle des deux urnes*.

⁴⁹ Cf. (1996, p. 20).

⁵⁰ Trad.: Imaginez une expérience qui se déroule de la façon suivante. À un certain moment, on donne à trois humains chacun une émeraude. Plusieurs siècles après, alors qu'un ensemble d'humains complètement différent se trouvent vivants, on donne également dans l'expérience à cinq milliers d'humains chacun une émeraude. Vous ne savez en aucun cas, toutefois, si votre siècle est le siècle précédent où juste trois personnes devaient être dans cette situation, ou bien le siècle suivant où cinq mille personnes devaient être dans ce cas. Vous dites-vous que si le vôtre était le siècle précédent alors les cinq mille personnes ne seraient pas encore vivantes, et que donc vous n'auriez aucune chance d'être parmi eux? Sur cette base, concluez-vous que vous pourriez aussi bien parier que vous avez vécu en siècle plus tôt?

Supposez en fait que vous avez parié que vous avez vécu là. Si chacun de ceux qui ont reçu une émeraude dans l'expérience avait parié de cette manière, il y aurait donc cinq mille perdants et seulement trois gagnants. Le pari valable est donc que le vôtre est, parmi les deux siècles, celui qui est postérieur.

A ce stade, on peut observer qu'une autre expérience mentionnée par Leslie⁵¹ correspond au même type de raisonnement que celui qui prévaut au niveau de l'*expérience des émeraudes*. Cette expérience comporte toutefois des caractéristiques légèrement différentes. Considérons tout d'abord la description correspondante:

Compare the case of geographical position. You develop amnesia in a windowless room. Where should you think yourself more likely to be: in Little Puddle with a tiny situation, or in London? Suppose you remember that Little Puddle's population is fifty while London's is ten million, and suppose you have nothing but those figures to guide you. (...) Then you should prefer to think yourself in London. For what if you instead saw no reason for favouring the belief that you were in the larger of the two places? Forced to bet on the one or on the other, suppose you betted you were in Little Puddle. If everybody in the two places developed amnesia and betted as you had done, there would be ten million losers and only fifty winners. So, it would seem, betting on London is far more rational. The right estimate of your chances of being there rather than in Little Puddle, on the evidence on your possession, could well be reckoned as ten million to fifty⁵².

Appelons cela l'*expérience Petit-Bled/Londres*. Cette dernière est basée sur une situation réelle, concrète, à mettre en relation avec un modèle probabiliste sous-jacent. Il conviendra donc également d'analyser l'*expérience Petit-Bled/Londres* en termes de n -univers, puis de s'attacher à décrire le modèle probabiliste qui lui correspond, de manière à rendre explicite l'analogie avec l'urne correspondante. Ceci permettra de mettre en évidence une structure identique à celle de l'*expérience des émeraudes*. La structure commune aux deux expériences, toutefois, n'apparaîtra de manière nette que lorsque les deux expériences auront été modélisées dans le n -univers correspondant.

5.2. Modélisation en termes de n -univers

A ce stade, il convient de modéliser l'*expérience des émeraudes* et l'*expérience Petit-Bled/Londres* dans le cadre des n -univers.

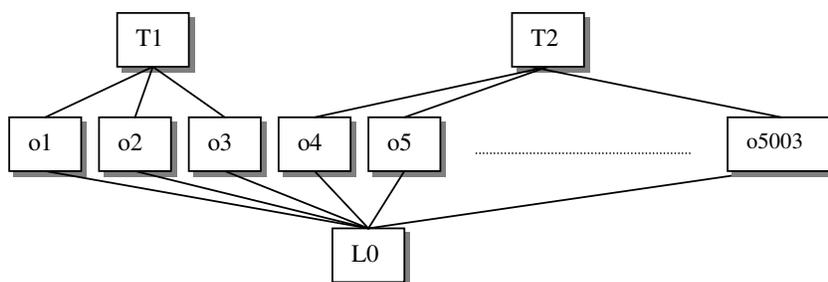
5.2.1. Modélisation de l'*expérience des émeraudes* dans le n -univers temporel à objets multiples ($\Omega\alpha^*T$)

Attachons-nous tout d'abord à modéliser la situation de l'*expérience des émeraudes* en termes de n -univers. Quelles sont donc les caractéristiques du n -univers correspondant à l'*expérience des émeraudes*? Ici, il apparaît en premier lieu, que la situation correspondante se caractérise par la présence de multiples objets: les émeraudes. En termes de n -univers, on se situe donc dans un n -univers à objets multiples ($\Omega\alpha\dots$). En second lieu, on peut considérer que les émeraudes se trouvent placées en un lieu unique: la Terre. Ainsi, le n -univers correspondant possède une constante-localisation (L_0). De plus, Leslie distingue explicitement deux moments dans l'expérience: une

⁵¹ Cf. (1996, p. 191).

⁵² Trad.: Comparez avec le cas de la position géographique. Vous développez une amnésie dans une salle sans fenêtres. Là où si vous vous pensez plus probablement pour être: dans peu de magma avec une situation minuscule, ou à Londres? Supposez-vous se rappellent que la population de peu de magma est cinquante tandis que Londres est dix millions, et vous suppose n'ont rien mais ces figures pour vous guider (.) alors que vous devrait préférer se penser à Londres. Pour ce qui si vous ne voyiez à la place aucune raison pour favoriser la croyance que vous étiez dans le plus grand des deux endroits? Forcé à parier sur celui ou de l'autre, supposez-vous vous parient étaient dans peu de magma. Si tout le monde dans les deux endroits développait l'amnésie et pariait comme vous aviez fait, il y aurait dix millions de perdants et seulement cinquante gagnants. Ainsi, elle semblerait, pariant sur Londres est plus raisonnable bien. La bonne évaluation de vos chances d'être là plutôt que dans peu de magma, sur l'évidence sur votre possession, a pu jaillir soit comptée en tant que dix millions à cinquante.

position temporelle correspondant à une époque donnée⁵³ et une autre position temporelle qui se situe plusieurs siècles plus tard⁵⁴. On peut donc modéliser la situation correspondante à l'aide d'un n-univers comportant deux positions temporelles: T_1 et T_2 . Le n-univers correspondant apparaît donc temporel ($\Omega T...$). En outre, il s'avère qu'une émeraude donnée ne peut occuper plusieurs positions temporelles successives. En effet, il résulte implicitement des conditions de l'expérience qu'une émeraude particulière ne peut occuper qu'une position temporelle donnée et que les émeraudes existant en T_1 n'existent pas en T_2 (et réciproquement). Ainsi le n-univers correspondant à l'expérience des émeraudes est-il un n-univers qui n'est pas à démultiplication temporelle ($\Omega T...$). Enfin, on peut observer que plusieurs émeraudes peuvent se trouver à une même position temporelle T_1 donnée: trois émeraudes existent ainsi en T_1 et cinq mille émeraudes en T_2 . Le nombre d'émeraudes existant à un moment donné se révèle donc variable. Par conséquent, les objets se trouvent dans une relation *plusieurs* avec la variable temporelle, dans le n-univers correspondant. Compte tenu de ce qui précède, il apparaît ainsi que l'expérience des émeraudes prend place dans un $\Omega\alpha^*T...$, un n-univers à objets multiples, comportant une constante de localisation et présentant une variable temporelle avec laquelle les objets se trouvent en relation *plusieurs*⁵⁵.



5.2.2. Modélisation de l'expérience Petit-Bled/Londres dans le n-univers localisé à objets multiples ($\Omega\alpha^*L$)

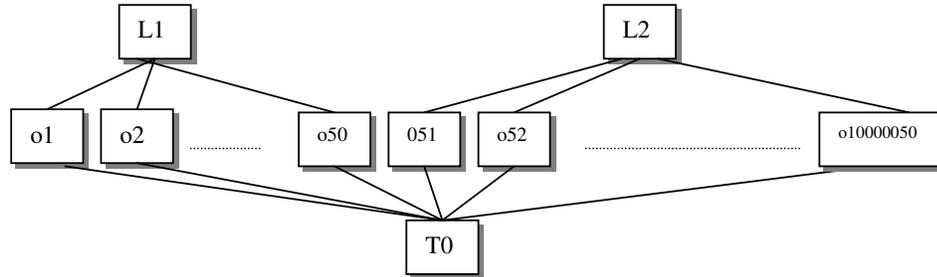
Attachons-nous maintenant à modéliser la situation de l'expérience *Petit-Bled/Londres* en termes de n-univers. Quelles sont donc les critères du n-univers correspondant à la situation concrète qui prévaut dans l'expérience *Petit-Bled/Londres*? Il apparaît ici en premier lieu, que la situation correspondante se caractérise par la présence de multiples habitants: 50 à Petit-Bled et 10 millions à Londres. Le n-univers correspondant est donc un n-univers à objets multiples ($\Omega\alpha...$). En second lieu, il apparaît que cette expérience prend place à un moment unique. On peut ainsi considérer que le n-univers correspondant possède une constante-temps (T_0). De plus, Leslie distingue explicitement deux positions spatiales dans l'expérience: Petit-Bled et Londres. On peut donc modéliser la situation correspondante à l'aide d'un n-univers comportant deux positions spatiales: L_1 et L_2 . Le n-univers correspondant apparaît ainsi comme un n-univers qui comporte une variable locale ($\Omega L...$). En outre, chacun des habitants se trouve soit à Petit-Bled soit à Londres. Ainsi, un habitant donné ne peut occuper plusieurs positions spatiales à la fois. Il s'ensuit que ceux qui se trouvent en L_1 ne peuvent se trouver en L_2 (et réciproquement). Ainsi le n-univers correspondant à l'expérience *Petit-Bled/Londres* est-il un n-univers qui n'est pas à démultiplication locale ($\Omega L...$). Enfin, on peut observer que

⁵³ Cf.: 'At some point in time'. Trad.: 'A un moment particulier donné'.

⁵⁴ Cf. 'several centuries afterwards'. Trad.: "plusieurs siècles après".

⁵⁵ On peut se poser la question ici de savoir si le n-univers correspondant à l'expérience des émeraudes n'est pas un n-univers coloré ($\Omega C...$). En ce sens, la couleur des émeraudes n'est elle pas un élément dominant? Il apparaît cependant que non, car la couleur mentionnée dans l'expérience est unique. En effet, on peut aussi bien concevoir cette expérience avec des rubis ou encore des saphirs. Il apparaît ainsi que l'on peut se dispenser de recourir à une variable de couleur, car tous les objets (les émeraudes) présentent une couleur identique. Dès lors, le n-univers correspondant n'est pas un n-univers coloré.

plusieurs personnes peuvent se trouver à une position spatiale L_i donnée: on a ainsi 50 habitants à Petit-Bled (L_1) et 10 millions à Londres (L_2). Par conséquent, les objets se trouvent dans une relation *plusieurs* avec la variable spatiale dans le n-univers correspondant. Compte tenu de ce qui précède, il apparaît que la situation de *l'expérience Petit-Bled/Londres* prend place dans un $\Omega\alpha^*L$, un n-univers à objets multiples, comportant une constante temporelle, et présentant une variable de localisation, avec laquelle les objets se trouvent en relation *plusieurs*.



5.3. Modèles probabilistes correspondants

On vient de le voir, les situations concrètes correspondant à *l'expérience des émeraudes* et à *l'expérience Petit-Bled/Londres* peuvent être modélisées respectivement dans le $\Omega\alpha^*T$ et le $\Omega\alpha^*L$. Il convient maintenant de s'attacher à décrire les modèles probabilistes correspondants. Considérons tout d'abord le $\Omega\alpha^*T$ dans lequel prend place l'expérience des émeraudes. Il apparaît ici que le modèle probabiliste correspondant est le suivant. On considère une urne (en un lieu L_0) qui contient 10 boules en T_1 et 990 autres boules en T_2 . On considère maintenant l'événement E qui consiste dans le fait qu'une boule a été tirée. Ayant tiré une boule⁵⁶, on se pose la question de savoir quelle est la probabilité qu'elle ait été tirée en T_1 ⁵⁷? On a ainsi les deux hypothèses:

- (H1₄₀) la boule provient des 10 boules en T_1
- (H2₄₁) la boule provient des 990 boules en T_2

On raisonne alors: $P(H1) = 10/(10+990) = 10/1000 = 0,01$ et $P(H2) = 990/1000 = 0,99$. Et on conclut que la probabilité que la boule provienne des 10 boules existant en T_1 est égale à 0,01, alors que la probabilité que la boule provienne des 990 boules en T_2 est égale à 0,99.

Considérons maintenant le $\Omega\alpha^*L$, le n-univers dans lequel se trouve modélisée *l'expérience Petit-Bled/Londres*. Dans ce cas, il apparaît que le modèle probabiliste est le suivant. On considère, au temps constant T_0 , deux urnes telles que celle qui est placée en L_1 contient 10 boules et celle qui se trouve en L_2 comporte 990 boules. Ayant tiré une boule, quelle est la probabilité qu'elle provienne de l'urne placée en L_1 ? On a alors les deux hypothèses:

- (H1₄₂) la boule est l'une des 10 boules provenant de l'urne placée en L_1
- (H2₄₃) la boule est l'une des 990 boules provenant de l'urne placée en L_2

Et on est conduit à raisonner de la même manière que précédemment. Ainsi: $P(H1) = 10/1000 = 0,01$ et $P(H2) = 990/1000 = 0,99$.

Comme on l'a vu plus haut, *l'expérience des émeraudes* peut être modélisé dans le $\Omega\alpha^*T$ alors que *l'expérience Petit-Bled/Londres* peut être modélisée dans le $\Omega\alpha^*L$. On le voit les deux n-univers

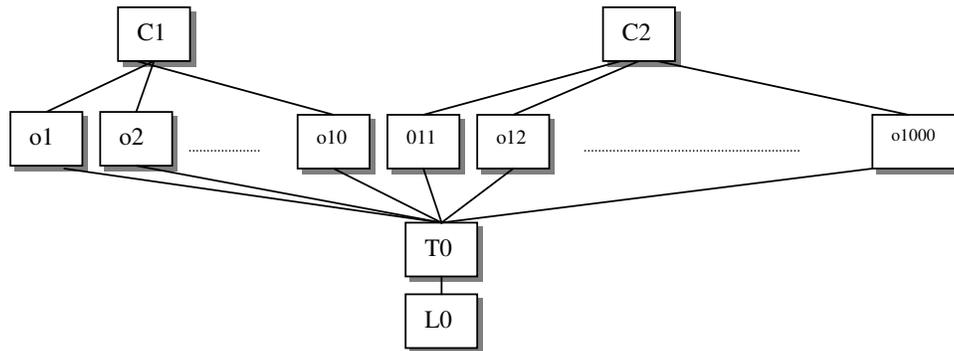
⁵⁶ Ici, on peut remarquer qu'il n'est pas nécessaire que les boules soient marquées ou numérotées d'une manière ou d'une autre.

⁵⁷ Il est à noter ici que l'on ne donne pas de probabilité initiale telle que $P(H1) = P(H2) = 0,5$.

correspondant à ces deux expériences présentent la structure commune $\Omega\alpha*\beta$. Ceci suggère que l'on pourrait également construire des versions analogues à l'expérience des émeraudes et à l'expérience Petit-Bled/Londres dans des n-univers où le critère n'est ni temporel ni spatial, mais correspond à un autre critère tel que la couleur. Attachons-nous donc à mettre en évidence un modèle analogue à l'expérience des émeraudes dans le $\Omega\alpha*C$. Soit donc un n-univers à objets multiples et présentant une variable de couleur. On considère pour simplifier que le critère-variable de couleur comprend deux taxons: C_1 (rouge) et C_2 (vert). De même, plusieurs objets donnés peuvent posséder une même couleur: les objets sont en relation *plusieurs* avec le critère de couleur. Ainsi, le modèle correspondant à l'expérience des émeraudes et à l'expérience Petit-Bled/Londres peut être ainsi décrit dans le $\Omega\alpha*C$. On considère, au temps constant T_0 , en un lieu L_0 une urne contenant 10 boules rouges et 990 boules vertes (soient mille boules au total). Ayant tiré une boule, quelle est donc la probabilité qu'elle soit rouge? On a les deux hypothèses:

- (H1₄₄) la boule tirée est l'une des 10 boules rouges
- (H2₄₅) la boule tirée est l'une des 990 boules vertes

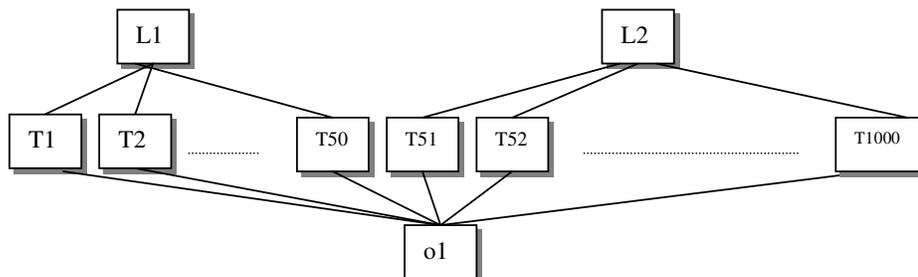
Et l'on est conduit au même raisonnement que précédemment. Il s'ensuit: $P(H1) = 10/1000$ et $P(H2) = 990/1000$.



Le modèle probabiliste correspondant à l'expérience des émeraudes et à l'expérience Petit-Bled/Londres qui prend place dans un $\Omega\alpha*C$ suggère également que l'on peut également modéliser des deux types d'expériences, sans perte de contenu, dans un n-univers à objet unique. Il apparaît en effet que l'on peut également modéliser les deux expériences dans un $\Omega T*L$, un n-univers à objet unique et comportant des variables temporelles et locales, où l'objet unique possède une propriété de démultiplication temporelle. Le modèle correspondant aux deux expériences est alors le suivant. Une boule unique existe à plusieurs positions temporelles: de T_1 à T_{50} , puis de T_{51} à T_{1000} . De T_1 à T_{50} , elle se trouve à la position spatiale L_1 , puis de T_{51} à T_{1000} , elle se trouve à la position spatiale L_2 . Considérant maintenant l'existence de la boule à une position temporelle donnée. Etant donné les deux hypothèses:

- (H1₄₆) la position temporelle de la boule est de T_1 à T_{50}
- (H2₄₇) la position temporelle de la boule est de T_{51} à T_{1000}

on est alors conduit à raisonner de la même manière que précédemment.



5.4. Leçons des deux expériences

Plusieurs points importants méritent d'être signalés concernant le modèle probabiliste qui est celui de l'expérience des émeraudes et de l'expérience *Petit-Bled/Londres*. Il convient tout d'abord ainsi de mettre notamment l'accent sur les différences existant entre ces deux dernières expériences et le modèle des deux urnes. Tout d'abord, à la différence du modèle des deux urnes, il ne sert à rien dans l'expérience des émeraudes et l'expérience *Petit-Bled/Londres* de considérer la probabilité conditionnelle de l'événement E par rapport aux hypothèses (H1) ou (H2). Car l'événement E est certain, que l'on se place dans l'une ou l'autre des deux hypothèses. Ainsi, $P(E|H1) = P(E|H2) = 1$.

En second lieu, le calcul de $P(H1)$ et de $P(H2)$ qui est effectué dans le modèle de l'expérience des émeraudes et l'expérience *Petit-Bled/Londres* se révèle indépendant des probabilités initiales $P(H1)$ et $P(H2)$ qui pourraient être éventuellement données (par exemple, $P(H1) = P(H2) = 0,5$). Il ne s'agit pas en effet dans le cas de l'expérience des émeraudes et l'expérience *Petit-Bled/Londres* de calcul de probabilités a posteriori, et on ne fait pas application, comme dans le modèle des deux urnes, du théorème de Bayes. En effet, le modèle de l'expérience des émeraudes et de l'expérience *Petit-Bled/Londres* permet de calculer directement une probabilité initiale. Aussi le modèle correspondant peut-il également être décrit sans que les probabilités initiales $P(H1)$ et $P(H2)$ soient initialement données.

L'expérience des émeraudes de même que l'expérience *Petit-Bled/Londres* apparaissent finalement basées sur un modèle qui conduit à privilégier une hypothèse selon laquelle le nombre d'objets est élevé par rapport à une hypothèse selon laquelle le nombre d'objets est faible. Pour cette raison, le modèle correspondant à ces deux expériences peut être dénommé *modèle à majoration*.

On peut observer enfin que le modèle à majoration se révèle tout à fait compatible avec SSA. Plus encore, les conclusions dérivées du modèle à majoration ne sont possibles que dès lors que la boule est tirée au hasard parmi les 1000 boules. Ainsi, SSA apparaît bien comme une condition pour une application valide du modèle à majoration. En ce sens, il s'agit d'une propriété commune avec le modèle des deux urnes dont SSA constitue également une des conditions d'application.

Chapitre 6. Le pile ou face divin et SIA

6.1. Le principe SIA

Le principe SIA (*Self-Indication Assumption*) a été formalisé ainsi par Bostrom (2000, *Observational Selection Effects and Probability*):

(SIA) Given the fact that you exist, you should (other things equal) favor hypotheses according to which many observers exist over hypotheses on which few observers exist⁵⁸

De manière intuitive, ce principe énonce que lorsque les deux types d'hypothèses suivantes sont formulées:

(H1₄₈) il existe peu d'observateurs

(H2₄₉) il existe beaucoup d'observateurs

et que l'on est soi-même un observateur, on doit favoriser l'hypothèse (H2). SIA est un principe controversé, qui a notamment fait l'objet de discussions par Olum (2000) ou Ćirković (2001). De plus, un certain nombre de solutions pour DA apparues dans la littérature sont elles-mêmes basées sur l'application de SIA⁵⁹.

6.2. Le pile ou face divin

L'expérience du *pile ou face divin* a été décrite par Bostrom (2000) de la manière suivante:

Stage (a): God first creates hundred cubicles. Each cubicle has a unique number painted on it on the outside (which can't be seen from the inside); the numbers are the integers between 1 and 100. God creates one observer in cubicle #1. Then God tosses a fair coin. If the coin falls tails, He does nothing more. If the coin falls heads, He creates one observer in each of cubicle #2-#100. Apart from this, the world is empty. It is now a time well after the coin has been tossed and any resulting observers have been created. Everyone knows all the above.

Stage (b): A little later, you have just stepped out of your cubicle and discovered that it is #1.

Question: What should your credence of the coin having fallen tails be at stages (a) and (b)?⁶⁰

Une telle expérience est basée sur les deux hypothèses suivantes:

(H1₅₀) il existe 1 observateur

(H2₅₁) il existe 100 observateurs

⁵⁸ Trad.: Etant donné le fait que vous existez, vous devez (toutes choses étant par ailleurs égales) favoriser les hypothèses selon lesquelles beaucoup d'observateurs existent par rapport aux hypothèses selon lesquelles peu d'observateurs existent.

⁵⁹ Cf. notamment la solution proposée par Olum (2000).

⁶⁰ Trad.: Étape (a): Dieu crée d'abord cent compartiments. Chaque compartiment comporte un nombre unique peint à l'extérieur (qui ne peut pas être vu de l'intérieur); les nombres sont les nombres entiers entre 1 et 100. Dieu crée un observateur dans le compartiment n°1. Dieu lance alors une pièce de monnaie équilibrée en l'air. Si la pièce de monnaie tombe sur pile, il ne fait rien d'autre. Si la pièce de monnaie tombe sur face, il crée alors un observateur dans chacun de compartiment n°2 à n°100. A part cela, le monde correspondant est vide. On se trouve maintenant bien après que la pièce de monnaie ait été jetée en l'air et tous les observateurs correspondants ont été créés. Ce qui précède est connu de tous. Étape (b): un peu plus tard, vous êtes juste sorti de votre compartiment et vous avez découvert qu'il s'agit du n°1. Question: Que devrait être votre croyance que la pièce de monnaie est tombée sur pile aux étapes (a) et (b)?

D'une manière générale, le problème qui résulte de l'expérience du *pile ou face divin* est de déterminer correctement les probabilités concernant les hypothèses (H1) et (H2) aux étapes (a) et (b) de l'expérience.

Par ailleurs, la description d'une expérience de ce type est également effectuée de la manière suivante par Leslie⁶¹:

Suppose all humans had to exist at the same time. Suppose you had no knowledge of whether there are any humans apart from yourself. You know merely that God had decided to toss his coin *just once*, and that if it had landed heads then he would have created ninety million humans, while if it had landed tails the he would have created only a single human. Would you therefore think the odds ninety million to one that God's coin had landed heads?⁶²

L'expérience décrite par Leslie et celle décrite par Bostrom apparaissent identiques, à une différence près toutefois. En effet, il s'avère que la description de Bostrom distingue les étapes (a) et (b), alors que ce n'est pas le cas avec la version de l'expérience décrite par Leslie. L'étape (b) de Bostrom n'est en effet pas incluse dans l'expérience de Leslie. Leslie considère ainsi simplement l'événement: "je sais que j'existe", qui correspond à l'étape (a) dans la version de Bostrom. A l'inverse, Bostrom considère deux événements successifs: E_a ("je sais que j'existe") et E_b ("je possède la numéro 1"). Ainsi, l'expérience décrite par Leslie inclut seulement les données de l'étape (a) dans l'expérience de Bostrom.

Le *pile ou face divin* se révèle problématique en ce sens que deux raisonnements a priori valides et conduisant à des conclusions différentes permettent de calculer les probabilités aux deux étapes de l'expérience. Ainsi, selon Bostrom (2000) et Olum (2000)⁶³, on peut utiliser a priori valablement les

⁶¹ Cf. 1996, p. 227.

⁶² Trad.: On suppose que tous les humains ont existé au même moment. Supposez-vous n'a eu aucune connaissance du fait qu'il existe ou non des humains à part vous-même. Vous savez simplement que Dieu avait décidé de lancer sa pièce de monnaie en l'air une seule fois, et que si elle était tombée sur face il aurait créé quatre-vingt-dix millions d'humains, alors que si elle était tombée sur pile il aurait créé seulement un seul humain. Penseriez-vous alors qu'il existe une chance de quatre-vingt-dix millions contre un que la pièce de monnaie lancée par Dieu soit tombée sur face?

⁶³ Cf. la présentation du pile ou face divin faite par Olum (2000): "Leslie and Bostrom argue as follows. Before you look at your room number, you should think that since the coin was fair the chance of heads was 1/2. Now if the coin was heads, then of course you would be in one of the first ten rooms. However, if the coin was tails, the chance to be in one of the first ten rooms is 1/100. Thus, according to Eq. (2) [Bayes' formula], you should now believe that the coin was heads with probability 0.99.

The alternative argument runs as follows. Before you look at your room number, you should think that the probability of heads is 0.99. There are one thousand possible people who would be right with that belief, whereas only ten would be right with the belief in heads. When you look at your room number, you should then update your probabilities using Eq. (2) [Bayes' formula]. The result is that in the end you think the chance is 1/2 that the coin was heads. Another way to say the same thing is that there are ten ways to have the coin heads and you in a room in the first ten, and ten ways to have the coin tails and you in a room in the first ten, and thus the chances for heads and tails are equal.". Trad.: "Leslie et Bostrom argumentent de la façon suivante. Avant que vous regardiez le numéro de votre pièce, vous devriez penser que puisque la pièce de monnaie était équilibrée la probabilité de tomber sur face était de 1/2. Maintenant si la pièce de monnaie était tombée sur face, alors naturellement vous vous trouveriez dans l'une des dix premières salles. Cependant, si la pièce de monnaie est tombée sur pile, la chance de se trouver dans l'une des dix premières salles est de 1/100. Ainsi, selon l'Eq. (2) [la formule de Bayes], vous devriez maintenant croire que la pièce de monnaie est tombée sur face avec une probabilité de 0,99. L'argument alternatif est le suivant. Avant que vous ne regardiez le numéro de votre pièce, vous devriez penser que la probabilité que la pièce tombe sur face est de 0,99. Il y a mille personnes qui auraient virtuellement raison avec cette croyance, tandis que seulement dix auraient raisonné de manière exacte avec la croyance que la pièce tombe sur face. Quand vous regardez le numéro de votre pièce, vous devriez alors mettre à jour vos probabilités en utilisant l'Eq. (2) [la

deux raisonnements suivants. Le premier (GCT-) conduit à raisonner ainsi. Puisque la pièce est équilibrée, la probabilité que celle-ci soit tombée sur pile est de $1/2$. Par conséquent, à l'étape (a), $P(H1) = 0,5$. Si l'expérience est répétée un grand nombre de fois, on aura donc bien $P(H1) = P(H2) = 0,5$. On se trouve alors replacé dans le cadre du *modèle à minoration*. Et par conséquent, à l'étape (b), par application directe du décalage bayésien résultant du modèle à minoration, il s'ensuit: $P'(H1) \approx 0,99$ et $P'(H2) \approx 0,01$.

Le raisonnement alternatif (GCT+) est le suivant. A l'étape (a), je raisonne ainsi. La boule qui a été tirée provient soit de la petite urne, soit de la grande. Maintenant, je me trouve donc dans la situation qui est celle du *modèle à majoration*. La probabilité⁶⁴ de tirer une boule dans la petite urne est de $1/101$ alors que la probabilité de tirer une boule dans la grande urne est de $100/101$. Par conséquent, $P(H1) \approx 0,01$ et $P(H2) \approx 0,99$. Ainsi, à cette étape du raisonnement, les probabilités a priori sont déterminées et je me trouve dans les conditions qui sont celles du *modèle à minoration*. Il s'ensuit, à l'étape (b), par application du théorème de Bayes: $P'(H1) = P'(H2) = 0,5$.

Prima facie, chacun des deux raisonnements apparaît valide. Pourtant, ils conduisent à des conclusions radicalement différentes, et l'un des deux seulement doit prévaloir. Ainsi, le problème posé par le *pile ou face divin* est le suivant: lequel des deux raisonnements (GCT-) ou (GCT+) est valide?

On peut remarquer, de manière préliminaire, que la formulation du problème place en concurrence deux types de raisonnements. Ainsi formulé, le problème conduit en effet à un choix exclusif entre deux raisonnements distincts. Il est cependant permis de penser que d'autres raisonnements pourraient également être considérés. On peut penser en effet que les raisonnements (GCT-) et (GCT+) ne décrivent pas de manière exhaustive l'ensemble des raisonnements qui peuvent être appliqués au *pile ou face divin*. Bostrom décrit d'ailleurs un troisième type de raisonnement, mais son caractère inconsistant conduit à un rejet immédiat. Il apparaît ainsi que la formulation en termes de choix exclusif est simplificatrice. On pourrait reformuler ainsi le problème en des termes plus généraux: quel est le raisonnement correct qui permet de calculer valablement les probabilités a priori $P(H1)$ et $P(H2)$ et les probabilités a posteriori $P'(H1)$ et $P'(H2)$ correspondant à l'expérience du *pile ou face divin*? Bien sûr, les deux raisonnements précédemment cités sont ceux qui retiennent prima facie l'analyse. Mais il s'avère toutefois que la formulation en termes de choix entre les raisonnements (GCT-) et (GCT+) se révèle réductrice.

6.3. Modélisation en termes de n-univers

Afin de permettre une analyse du *pile ou face divin*, il est nécessaire de modéliser l'expérience dans le cadre des n-univers. Avant de procéder toutefois à une telle modélisation, il est nécessaire de considérer plusieurs variations qui trouvent leur justification dans le fait qu'elles permettront de meilleures comparaisons avec plusieurs modèles⁶⁵ qui ont été décrits plus haut.

6.3.1. Variations du *pile ou face divin*

Afin de faciliter les comparaisons, l'expérience du *pile ou face divin* peut être tout d'abord simplifiée de manière à ce que sa structure et ses données soient analogues à celles du *modèle des deux urnes*. Une telle version est d'ailleurs décrite par Ken Olum (2000)⁶⁶:

formule de Bayes]. Le résultat est à la fin que vous pensez que la probabilité que la pièce de monnaie est tombée sur face est de $1/2$. Une autre manière de dire la même chose est qu'il y a dix manières différentes que la pièce de monnaie tombe sur face et que vous vous trouviez dans une chambre parmi les dix premières, et dix manières que la pièce de monnaie tombe sur pile et que vous vous trouviez dans une chambre parmi les dix premières, et ainsi les probabilités que la pièce tombe sur pile ou face sont égales."

⁶⁴ Avec les données qui sont celles de la description de Bostrom.

⁶⁵ Il s'agit des modèles à minoration et à majoration.

⁶⁶ Olum dans cette version ne distingue pas les étapes (a) et (b) dans l'expérience, mais elles peuvent être aisément rétablies.

Suppose that God tosses a fair coin. If it comes up heads, he creates ten people, each in their own room. If tails, he creates one thousand people, each in their own room. The rooms are numbered 1-10 or 1-1000. The people cannot see or communicate with the other rooms. Suppose that you know all this, and you discover that you are in one of the first ten rooms. How should you reason that the coin fell?⁶⁷

Cette version se révèle ainsi basée sur les deux hypothèses suivantes:

- (H1₅₂) il existe 10 observateurs
- (H2₅₃) il existe 1000 observateurs

A ce stade, on peut penser qu'il est également utile de modifier légèrement l'expérience du *pile ou face divin* de manière à la rendre compatible avec les données qui sont celles des modèles à *minoration* et à *majoration* précédemment décrits. La version décrite par Bostrom et Olum peut ainsi être adaptée de manière afin de faciliter les comparaisons. Il en résulte alors la version suivante du *pile ou face divin*:

Le contenu d'une urne dépend du lancer d'une pièce équilibrée. Si la pièce tombe sur pile, l'urne contient 10 boules rouges. Si la pièce tombe sur face, l'urne contient 1000 boules dont 10 boules rouges et 990 boules vertes. A l'étape (a), une boule est extraite de l'urne. A l'étape (b), on constate que la couleur de la boule extraite de l'urne est rouge. Le problème est d'évaluer la probabilité selon laquelle la pièce est tombée sur pile ou face, aux étapes (a) et (b) de l'expérience.

On le voit, une telle version est basée sur les deux hypothèses suivantes:

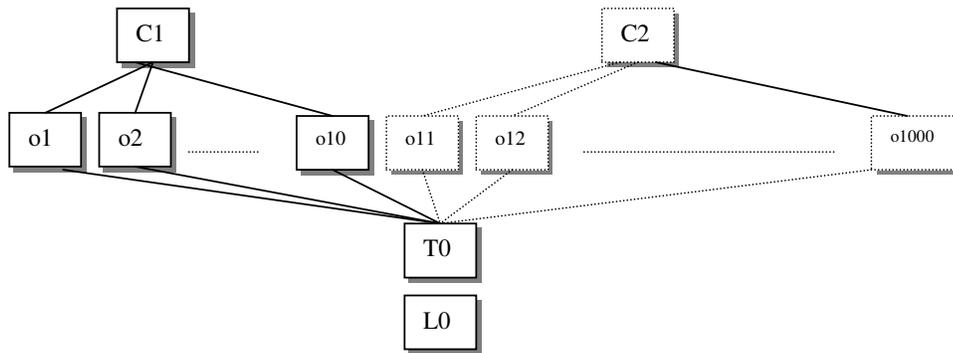
- (H1₅₄) l'urne contient 10 boules rouges
- (H2₅₅) l'urne contient 1000 boules dont 10 boules rouges et 990 boules vertes

qui sont également celles de la version bicolore du *modèle des deux urnes*. Le problème posé par le *pile ou face divin* est ici aussi de déterminer les probabilités P(H1) et P(H2) à l'étape (a) et les probabilités P'(H1) et P'(H2) à l'étape (b).

6.3.2. Modélisation dans le n-univers coloré à objets multiples ($\Omega\alpha^*C$)

A ce stade, on est à même de modéliser la situation du *pile ou face divin* en termes de n-univers. A cet égard, il apparaît que la situation du *pile ou face divin* s'avère tout à fait similaire à celle dans laquelle le *modèle des deux urnes* prend place. Le *pile ou face divin* prend ainsi place dans le cadre d'un n-univers qui est en tous points identique à celui du *modèle des deux urnes*. Etant donné que nous avons retenu ici la version du *pile ou face divin* appliquée à une urne bicolore, la version correspondante est donc celle du *modèle des deux urnes bicolore*. Le *pile ou face divin* prend ainsi place de la même manière que le *modèle des deux urnes bicolore* dans un $\Omega\alpha^*C$, un-univers à objets multiples atemporel, alocalisé et coloré. Un tel n-univers comporte de multiples objets colorés o_1, \dots, o_n , en un lieu unique L_0 et au temps constant T_0 , les taxons de couleur étant: {rouge, vert}. Il s'agit d'un n-univers où plusieurs boules peuvent présenter une même couleur. Par conséquent, les objets multiples y sont en relation *plusieurs* avec le critère de couleur.

⁶⁷ Trad.: Supposez que Dieu jette une fois une pièce de monnaie en l'air. Si elle tombe sur face, il crée dix personnes, chacune dans leur propre pièce. Si elle tombe sur pile, il crée mille personnes, chacun également dans leur propre pièce. Les pièces sont numérotées de n°1 à n°10 ou bien n°1 à n°1000. Les gens ne peuvent pas voir ni communiquer avec les autres pièces. Supposez que vous savez tout cela, et que vous découvrez que vous vous trouvez dans l'une des dix premières salles. Quel raisonnement devez-vous tenir afin de savoir comment la pièce de monnaie est tombée?



6.4. Analyse

Avant de présenter une analyse du problème posé par le *pile ou face divin*, il est intéressant de s'attacher préalablement à déterminer les différences existant entre le *modèle des deux urnes* et le *pile ou face divin*. A cet égard, il apparaît qu'une première différence réside dans le fait que dans le *modèle des deux urnes*, les étapes (a) et (b) ne sont pas distinguées, alors qu'elles le sont notamment au niveau de la version du *pile ou face divin* décrite par Bostrom. Ainsi, le *modèle des deux urnes* prend en compte l'événement "la boule tirée possède le n° 5", alors que le *pile ou face divin* distingue deux types d'événements: à l'étape (a), l'événement E_a "une boule est extraite de l'urne" et à l'étape (b), l'événement E_b "la boule tirée possède le n° 5".

En second lieu, l'expérience du *pile ou face divin* se différencie du *modèle des deux urnes* par le fait que les probabilités a priori $P(H1)$ et $P(H2)$ sont déterminées à l'avance dans le *modèle des deux urnes*, alors qu'elles ne le sont pas dans l'expérience du *pile ou face divin*. Les probabilités initiales dans le *modèle des deux urnes* sont en effet telles que $P(H1) = P(H2) = 0,5$ ⁶⁸. A l'inverse, en ce qui concerne le *pile ou face divin*, les probabilités a priori ne sont pas définies dans les conditions initiales de l'expérience. Et ceci permet aux deux raisonnements concurrents (GCT-) et (GCT+) de prendre place. En particulier, (GCT-) tente de recalculer les probabilités initiales manquantes à partir du fait que la pièce est équilibrée, en posant ainsi $P(H1) = P(H2) = 0,5$. Ainsi, les deux situations se révèlent donc également distinctes, compte tenu du fait que les probabilités initiales sont fixées à l'avance dans le *modèle des deux urnes* alors qu'elles ne sont pas définies dans l'énoncé du *pile ou face divin*.

A ce stade, il convient de s'attacher à donner une analyse du problème posé par le *pile ou face divin*. Mon analyse conduira à assigner des probabilités à (H1) et (H2) aux deux étapes de l'expérience, c'est-à-dire après chacun des événements E_a ("une boule est extraite de l'urne") et E_b ("la couleur de la boule est rouge"). Il convient ainsi de s'attacher tout d'abord à déterminer $P(H1)$ et $P(H2)$ à l'étape (a), alors qu'une boule vient d'être extraite de l'urne. A cette fin, intéressons-nous tout d'abord à la structure du raisonnement qui sous-tend (GCT+). Attachons-nous donc à analyser en détail le point de vue qui est celui de (GCT+). Il convient ici de se placer du point de vue qui est celui du *modèle à majoration*. Rappelons tout d'abord les conditions de ce dernier modèle. Considérons ici les conditions qui sont celles du *modèle à majoration*⁶⁹ lorsque la situation correspondante est modélisée dans un $\Omega\alpha^*C$. Ceci permettra d'assurer la compatibilité avec la version du *pile ou face divin* modélisée dans le $\Omega\alpha^*C$. Le *modèle à majoration* met ainsi en scène, au temps constant T_0 et à la position spatiale L_0 une urne contenant 10 boules rouges et 990 boules vertes (soient mille boules au total). Ayant tiré une boule, quelle est donc la probabilité qu'elle soit rouge? Ici, on sait de manière certaine que:

(H2₅₆) l'urne contient 1000 boules dont 10 boules rouges et 990 boules vertes

et on a les deux hypothèses:

⁶⁸ On peut bien sûr considérer, de manière alternative, d'autres probabilités a priori pour le *modèle des deux urnes*, sans que l'effet de l'expérience (un décalage bayésien) n'en soit modifié.

⁶⁹ Qui correspond à l'expérience des émeraudes et à l'expérience Petit-Bled/Londres.

- (H3₅₇) la boule tirée est l'une des 10 boules rouges
- (H4₅₈) la boule tirée est l'une des 990 boules vertes

D'où il s'ensuit: $P(H3) = 10/1000$ et $P(H4) = 990/1000$.

Il est utile de rappeler ici que dans les conditions qui sont celles du *modèle à majoration*, l'existence des 10 boules rouges et des 990 boules vertes est certaine. Ainsi, les 10 boules rouges et les 990 boules vertes se trouvent physiquement dans l'urne correspondante. Ceci est à mettre en parallèle avec les deux hypothèses qui sont celles du *pile ou face divin* dans le $\Omega\alpha^*C$:

- (H1₅₉) l'urne contient 10 boules rouges
- (H2₆₀) l'urne contient 1000 boules dont 10 boules rouges et 990 boules vertes

qui peuvent être ainsi reformulées, si l'on met l'accent sur l'éventuelle existence des 990 boules vertes:

- (H1₆₁) l'urne ne contient aucune boule verte
- (H2₆₂) l'urne contient 990 boules vertes

A ce stade, il apparaît qu'il existe une différence fondamentale entre la situation du *pile ou face divin* et celle du *modèle à majoration*. En effet, au niveau du *pile ou face divin*, l'existence des 990 boules vertes est hypothétique et on ne possède pas de certitude quant à leur existence. En revanche, dans le *modèle à majoration*, l'existence des 990 boules vertes est certaine, alors qu'elle n'est qu'hypothétique au niveau du *pile ou face divin*. Il s'agit là d'une différence fondamentale entre les conditions du *modèle à majoration* et la situation qui est celle du *pile ou face divin*.

Olum (2000) mentionne d'ailleurs explicitement le fondement du type de raisonnement inhérent à (GCT+) qui est de considérer les observateurs potentiels comme des observateurs réels⁷⁰. Dans le contexte du *pile ou face divin*, des boules vertes hypothétiques sont considérées dans (GCT+) comme des boules réelles. Ceci apparaît bien comme le cœur de la forme de raisonnement inhérente à (GCT+) qui conduit à l'application du modèle à majoration dans le *pile ou face divin*. Mais il s'avère que l'on se trouve au niveau du *pile ou face divin*, dans la situation du *modèle à majoration*, d'une manière qui n'est qu'hypothétique. En définitive, (GCT+) apparaît comme une application inconditionnelle du modèle à majoration, alors que les conditions de ce dernier ne sont satisfaites que de manière hypothétique. En conséquence, (GCT+) s'analyse comme une application illicite du modèle à majoration.

Ceci a pour effet de mettre en lumière la nature de SIA. En effet, (GCT+) est sous-tendu par SIA. Par conséquent, l'analyse qui précède met en lumière la nature véritable de SIA qui n'est autre qu'une application du *modèle à majoration* alors que les conditions de ce dernier ne sont satisfaites que de manière hypothétique. SIA se révèle ainsi être une application illicite du *modèle à majoration*.

Les considérations qui précèdent renvoient au problème plus général de savoir quelle est la valeur du *modèle à majoration* lorsque l'existence des boules de couleur verte s'avère hypothétique. Les conditions du *modèle à majoration* ne sont en effet satisfaites que si (H2) est vérifiée, c'est-à-dire dans

⁷⁰ Cf. Olum (2000): 'The difference here hinges on whether one considers possible people in the same ways that one considers actual people. If instead of flipping a coin, God creates both sets of rooms, then Leslie and Bostrom and I all agree that you should think it much more probable that you are in the large set before you look at your room number, and equally probable afterward. Treating the two possibilities in the same way as two sets of actual observers implies the Self-Indication Assumption: the existence of a large number of observers in a possible universe increases the chance to find oneself in that universe.' Trad.: 'La différence tient ici au fait que l'on considère les personnes virtuelles de la même manière que l'on considère les personnes réelles. Si au lieu de lancer une pièce de monnaie, Dieu crée les deux séries de pièces, alors Leslie et Bostrom et moi-même sommes tous d'accord que vous devriez penser qu'il est beaucoup plus probable que vous vous trouvez dans la plus grande série avant de regarder le numéro de votre pièce, et également probable après l'avoir regardé. Le fait de traiter les deux possibilités comme deux séries d'observateurs réels implique la Self-Indication Assumption: l'existence d'un grand nombre d'observateurs dans un univers possible augmente la chance de se trouver dans ce même univers.'

notre exemple si les 990 boules vertes existent réellement. Or dans le *pile ou face divin*, le fait de savoir si (H2) est vrai ou non est précisément indéterminé. Ceci suggère de s'attacher à rechercher une réponse au problème plus général de savoir quelle la valeur du *modèle à majoration* lorsque (H2) est hypothétique et donc lorsque l'existence des 990 boules vertes n'est pas certaine. Dans une telle hypothèse, on a donc $P(H2) < 1$, alors que l'on a $P(H2) = 1$ dans le *modèle à majoration* standard. Il apparaît que ceci revient à se placer dans la situation qui est celle du *modèle à majoration*, mais avec l'information supplémentaire que $P(H2) < 1$. Quelle conclusion le *modèle à majoration* permet-il donc lorsque l'on sait que $P(H2) < 1$? Intuitivement toutefois, il semble que le *modèle à majoration* devrait trouver à s'appliquer d'une certaine façon, par exemple si la probabilité de l'existence des 990 boules vertes devait être très élevée et telle que $P(H2) = 0,99$. Ceci suggère de reconsidérer les conditions d'application du *modèle à majoration* de manière plus générale, en redéfinissant ces dernières quelle que soit la probabilité de l'hypothèse (H2). Dans ce contexte, l'application classique du *modèle à majoration* apparaît comme un cas particulier où l'hypothèse (H2) est confirmée, c'est-à-dire $P(H2) = 1$.

Afin de fixer les idées, plaçons-nous d'abord dans la situation spécifique qui est celle du *pile ou face divin*. Dans cette dernière hypothèse, on pose que l'urne contient 10 boules rouges de manière certaine mais aussi 990 boules vertes avec une probabilité égale à 0,5. Posons également, en suivant ici (GCT-) que la probabilité que les 990 boules vertes existent est égale à 0,5, puisque la pièce est équilibrée. On a ainsi $P(H1) = P(H2) = 0,5$. Et on a également les deux hypothèses suivantes:

- (H3₆₃) la boule tirée est l'une des 10 boules rouges
- (H4₆₄) la boule tirée est l'une des 990 boules vertes

Ici, on ne peut conclure, comme dans le modèle à majoration classique où $P(H2) = 1$ que $P(H3) = 10/1000$ et $P(H4) = 990/1000$. Car on se trouve dans une situation où $P(H2) = 0,5$. Ne devrait-on pas raisonner de la manière suivante, dans la situation qui est celle du *pile ou face divin*, en vertu d'une application conditionnelle du *modèle à majoration*? En premier lieu, quelle est la probabilité pour que la boule tirée soit l'une des 10 boules rouges? On envisage successivement les deux hypothèses, (H1) et (H2). Si (H1) est vraie, la probabilité de tirer une boule rouge est 1. En revanche, si (H2) est vraie, cette probabilité est égale à 1/100. Quelle est maintenant la probabilité de tirer une boule verte? Si (H1) est vraie, cette probabilité est égale à 0. En revanche, si (H2) est vraie, alors cette probabilité est égale à 99/100. Etant donné que la probabilité de (H1) et (H2) est égale à 0,5, on calcule ainsi que la probabilité de tirer une boule rouge $P(H3)$ est égale à $1*0,5+0,01*0,5 = 0,505$. Il s'agit là d'une application de la formule plus générale, qui est la suivante:

$$(65) P(H3) = P(H3|H1)*P(H1)+P(H3|H2)*P(H2)$$

On calcule de même que la probabilité de tirer une boule verte $P(H4)$ est égale à $0*0,5+0,99*0,5 = 0,495$. Là aussi, il s'agit d'une application de la formule générale:

$$(66) P(H4) = P(H4|H1)*P(H1)+P(H4|H2)*P(H2)$$

Il s'ensuit qu'il est légèrement plus probable de tirer une boule rouge plutôt qu'une boule verte. Ceci conduit ainsi, dans le cas où $P(H1) = P(H2) = 0,5$ à une probabilité légèrement supérieure de tirer une boule rouge. Ainsi, l'utilisation conditionnelle du *modèle à majoration* conduit à privilégier légèrement l'hypothèse selon laquelle la boule tirée est rouge. Il apparaît donc légèrement plus probable à l'étape (a) du *pile ou face divin* de tirer une boule rouge plutôt qu'une boule verte. Ceci illustre qu'il est tout à fait possible de raisonner à l'aide du *modèle à majoration*, même dans l'hypothèse où le nombre de boules n'est connu qu'avec une probabilité inférieure à 1. Cependant, dans ce cas, il est indispensable de faire intervenir cette dernière probabilité dans le calcul.

Ici, il apparaît que c'est précisément là que réside l'erreur dans le raisonnement qui sous-tend (GCT+): le *modèle à majoration* y est appliqué de manière inconditionnelle, sans prendre en compte le fait que $P(H2) = 0,5$ et d'une manière générale, $P(H2) < 1$. (GCT+) fait application du modèle à majoration alors que dans ce dernier modèle la condition $P(H2) = 1$ est exigée.

On l'a vu, on calcule que $P(H3) = 0,505$ et $P(H4) = 0,495$ (en vertu du modèle à majoration conditionnel) à l'étape (a) du *pile ou face divin*. Les probabilités initiales $P(H1)$ et $P(H2)$ demeurent cependant inchangées et telles que $P(H1) = P(H2) = 0,5$. Il reste à calculer $P'(H1)$ et $P'(H2)$ à l'étape (b) de l'expérience, lorsqu'on constate que la boule extraite de l'urne est rouge. A ce stade, il apparaît alors que l'on se trouve dans les conditions qui sont celles du *modèle des deux urnes*. On peut donc faire valablement application du *modèle à minoration*. Il s'ensuit, par application directe du théorème de Bayes: $P'(H1) \approx 0,99$ et $P'(H2) \approx 0,01$.

On le voit, la présente analyse appliquée au *pile ou face divin* conduit à retenir la conclusion qui résulte de (GCT-) et à rejeter le raisonnement qui sous-tend (GCT+). La présentation classique du *pile ou face divin* conduit à distinguer en effet deux raisonnements alternatifs: l'un (GCT-) basé sur le modèle à minoration, et l'autre (GCT+) fondé sur le modèle à majoration. A l'étape (a), il convient donc de calculer $P(H1)$ et $P(H2)$ en vertu de (GCT-). Et à l'étape (b), on calcule $P'(H1)$ et $P'(H2)$ à l'aide d'un raisonnement basé sur le *modèle à minoration*. La présente analyse se fonde sur (GCT-) et conduit au rejet de (GCT+), un raisonnement basé sur une application inconditionnelle du *modèle à majoration*. On le voit, l'analyse qui précède a ici permis de définir les conditions dans lesquelles il peut être fait usage d'un modèle à majoration différent de SIA, qui peut être appelé *modèle à majoration conditionnel*.

A ce stade, il convient de terminer par quelques remarques qui ont trait à la répétabilité de l'expérience du *pile ou face divin*. Olum (2000) souligne ainsi avec lucidité que SIA qui sous-tend (GCT+) souffre d'un défaut, à savoir que les résultats qui résultent de l'expérience du *pile ou face divin* ne sont pas répétables⁷¹. En effet, lorsque l'expérience est répétée, les probabilités $P(H1)$ et $P(H2)$

⁷¹ Cf. Olum (2000): 'Suppose that God flips his coin many times and creates many batches of people, some in sets of ten and some in sets of a thousand. Then even Leslie's argument doesn't yield a doomsday prediction. All sequences of coin flips are equally probable, in this interpretation, but the great majority of the time there are many small sets and many large sets. Thus before looking at your room number, you expect to be in a large set. After finding that you are in the first ten, you now reduce your estimate of the coin flip relevant to you to nearly equal probabilities of heads and tails. The same result could be seen in a repeated version of the "guarantee" protocol in which the experiment is done many times and the guarantee is that you will be one of the subjects in one of the runs, with equal probability to be any one of those subjects. Thus Leslie's argumentation depends on there being only a single universe. He agrees with this and says "In cases like this we must reject the intuition... that to estimate probabilities we ought to ask what bets would maximize winnings when the experiment was repeated infinitely many times" [3, page 228]. This does not seem to bother Leslie, but it is a strange claim. In some theories of probability, probability just means the bet that would maximize winnings, but even if one doesn't accept this as a definition, it still seems clear there is something wrong with a system for computing probabilities that yields odds vastly different from how a bettor should bet.' Trad.: 'Supposez que Dieu lance sa pièce de monnaie un grand nombre de fois et crée beaucoup de groupes de personnes, certains dans des groupes de dix et certains dans des groupes de mille. Dans ce cas même l'argument de Leslie ne conduit pas une prévision de nature apocalyptique. Toutes les séquences de lancers de pièces de monnaie sont équiprobables, selon cette interprétation, mais la plupart du temps, on aura beaucoup de petits groupes et beaucoup de grands groupes. Ainsi avant de regarder votre numéro de pièce, vous supposez que vous êtes dans un grand groupe. Après avoir constaté que vous êtes dans les dix premiers, vous réduisez maintenant votre estimation de lancer de pièce de monnaie vous concernant à des probabilités pratiquement égales concernant pile ou face. Le même résultat pourrait être obtenu avec une version répétée du protocole "garanti" où l'expérience est réalisée un grand nombre de fois et la garantie réside dans le fait que vous serez un des sujets dans un des lancers, avec une probabilité égale d'être n'importe lequel de ces sujets. Ainsi l'argumentation de Leslie dépend du fait qu'il existe seulement un seul univers. Il est d'accord avec ceci et dit "dans les cas comme celui-ci nous devons rejeter l'intuition... selon laquelle pour estimer des probabilités, nous devrions demander quelles façons de parier maximiseraient les gains obtenus lorsque l'expérience est répétée sans arrêt un grand nombre de fois" [3, page 228]. Ceci ne semble pas gêner Leslie, mais c'est une affirmation quelque peu étrange. Dans quelques théories probabilistes la probabilité signifie juste le pari qui maximiserait les gains, mais même si on n'accepte

tendent vers 0,5, puisque la pièce est équilibrée. Or SIA est basé sur le *modèle à majoration*, un modèle qui tire lui-même sa force du fait que ses conclusions se trouvent confirmés par la répétition de l'expérience ("in the long run..."). Il y a ainsi une contradiction interne dans le fait d'appliquer SIA, un principe basé sur la répétabilité à une expérience qui n'est elle-même pas répétable.

A cet égard, il est intéressant de considérer la description du *pile ou face divin* qui est faite par Leslie. En effet, l'auteur y met notamment l'accent sur le fait que le dé est lancé une seule fois (*just once*), ainsi que Leslie le souligne). La description de Bostrom ne fait pas mention d'une telle particularité. Cette dernière précision de Leslie mérite d'être analysée. En effet, l'énoncé du *pile ou face divin* comporte la mention selon laquelle la pièce lancée qui décide du choix est équilibrée. Ceci entraîne que si l'expérience est répétée, la probabilité que la pièce tombe sur pile et donc que l'hypothèse (H1) soit vraie tend vers 0,5. Dans ces circonstances, on peut alors poser $P(H1) = 0,5$ et $P(H2) = 0,5$. Et ceci a pour effet de replacer l'expérience du *pile ou face divin* dans les conditions exactes qui sont celles du *modèle des deux urnes*. Ici, il semble que la mention selon laquelle l'expérience n'a lieu qu'une fois ait pour objet de neutraliser un tel raisonnement, basé sur la répétabilité de l'expérience. On peut toutefois s'interroger sur le fondement d'une telle précision. Lorsqu'on effectue une expérience, il est essentiel que les conclusions qui résultent de cette expérience puissent être répétées et généralisées. Il s'agit là d'une exigence rationnelle de répétabilité. Un tel souci prévaut pour les expériences physiques. Mais on peut penser que ce qui prévaut pour les expériences concrètes vaut également pour les expériences de pensée. Par conséquent, on peut penser raisonnablement que l'exigence de répétabilité prévaut également pour les expériences de pensée.

En second lieu, cette exigence de non-répétabilité apparaît comme impossible à mettre en oeuvre. En effet, une expérience telle que le *pile ou face divin* est destinée à être exécutée mentalement. Mais par nature, une telle expérience est mise en oeuvre par chaque lecteur. Et elle tire précisément sa force du fait que chacun des lecteurs qui imagine la situation correspondante parvient à des conclusions identiques. Ainsi, une telle expérience, dès lors qu'elle est mentalement exécutée par plusieurs personnes, se révèle nécessairement répétée. Et l'exigence de non-répétabilité ne peut donc être satisfaite.

Enfin, dans la littérature relative à DA, Leslie fait un large usage des modèles probabilistes pour les appliquer à des situations réelles. Leslie possède pour cela une expertise qui fait précisément la force de son argumentation. Mais les modèles qu'il met en oeuvre possèdent habituellement une propriété de répétabilité, d'où son argumentation tire précisément son caractère convaincant. De nombreuses situations sont étudiées par Leslie dans des situations où elles sont répétées ("in the long run..."). Il n'y a là rien qu'une démarche rationnelle. Mais ceci contraste avec le fait d'exiger, pour l'expérience du *pile ou face divin* en particulier, qu'elle ne soit pas répétable.

6.5. Analyse des autres problèmes posés par SIA

L'application de SIA, on vient de le voir, entraîne un certain nombre de problèmes. Un de ces problèmes est notamment le raisonnement (GCT+) décrit avec l'expérience du *pile ou face divin*. On a vu qu'il convenait de rejeter l'application de SIA dans le cas de l'expérience du *pile ou face divin*, car il s'agissait d'une application du modèle à majoration alors que les conditions n'en étaient pas réunies. Cependant, il est apparu que SIA pouvait également conduire à d'autres types de problèmes. Bostrom (2001) décrit ainsi un autre problème qui résulte de l'utilisation du SIA: le *Philosophe présomptueux*. Il convient maintenant de s'y intéresser. Le *philosophe présomptueux* (the *Presumptuous Philosopher*) est décrit par Bostrom (2001, note 1) dans les termes suivants:

It is the year 2100 and physicists have narrowed down the search for a theory of everything to only two remaining plausible candidate theories, T_1 and T_2 (using considerations from super-duper symmetry). According to T_1 the world is very, very big but finite, and there are a total of a trillion trillion observers in the cosmos. According to T_2 , the world is very, very, very big but finite, and there are a trillion trillion trillion observers. The super-duper symmetry

pas ceci comme définition, il semble toujours clair ici que quelque chose va mal dans un système pour calculer des probabilités qui entraîne une chance extrêmement différente par rapport à la manière dont un parieur devrait parier.'

considerations are indifferent between these two theories. Physicists are preparing a simple experiment that will falsify one of the theories. Enter the presumptuous philosopher: "Hey guys, it is completely unnecessary for you to do the experiment, because I can already show to you that T_2 is about a trillion times more likely to be true than T_1 (whereupon the philosopher runs the *Incubator* thought experiment and appeals to SIA)!"⁷²

L'expérience est ici basée sur deux hypothèses:

- (H1₆₇) il existe 10^{24} êtres dans l'univers
- (H2₆₈) il existe 10^{36} êtres dans l'univers

Et on a également l'événement E_a : "je suis un observateur" (c'est-à-dire "une boule est extraite de l'urne"). On peut noter cependant ici que l'événement E_b , c'est-à-dire "je suis l'observateur n° 5" (qui correspond à: "la boule extraite de l'urne est rouge") du *pile ou face divin* n'est pas mentionné dans le *philosophe présomptueux*. Le raisonnement du *philosophe présomptueux* est le suivant. Compte tenu du fait que j'existe en tant qu'être dans l'univers et par application directe de SIA, il s'ensuit que la seconde hypothèse doit être privilégiée. On l'a vu en effet, SIA peut prendre place à partir de la constatation de l'existence d'un seul être dans le cosmos. SIA opère seulement à partir du simple constat de mon existence ("une boule est extraite de l'urne"). Car SIA est basé sur l'application du *modèle à majoration*. Contrairement au raisonnement basé sur le *modèle à minoration*, qui nécessite la connaissance du numéro correspondant au rang de l'observateur, SIA peut prendre place à partir de la seule connaissance de l'existence d'un être dans le cosmos.

De manière préliminaire, on peut observer ici que les considérations relatives à l'expérience⁷³ que préparent les physiciens ne sont pas indispensables ici. En effet, elles visent simplement à mettre en évidence la situation contraire à l'intuition qui résulte de l'application de SIA.

Une autre remarque préliminaire que l'on peut faire concernant le *philosophe présomptueux* est que l'expérience que s'apprêtent à réaliser les physiciens va entraîner une certitude, puisqu'elle s'apprête à falsifier l'une des deux théories, T_1 ou T_2 . Après l'expérience, T_1 ou T_2 se révélera fausse et par conséquent l'une des deux théories sera prouvée vraie. Ainsi, l'expérience conduit à une certitude alors que l'argument du *philosophe présomptueux* n'entraîne que l'augmentation, fût-elle importante, d'une probabilité. Le raisonnement qui est celui du *philosophe présomptueux* conduit ainsi à une probabilité forte, mais en aucun cas à une certitude telle que $P(T_1)$ ou $P(T_2) = 0$. De ce point de vue, la conclusion du *philosophe présomptueux* selon laquelle l'expérience est inutile se révèle fausse. L'expérience que s'apprêtent à mener les physiciens se révèle très utile, car elle va entraîner une certitude. A l'inverse, l'argument du *philosophe présomptueux* ne peut conduire qu'à une forte probabilité. Cependant, on peut observer qu'une telle remarque n'est pas décisive et ne suffit pas à bloquer le raisonnement qui est celui du *philosophe présomptueux*. En effet, il apparaît que l'on peut faire resurgir le problème en considérant un énoncé de l'expérience du *philosophe présomptueux* quelque peu différent où les physiciens s'apprêtent à réaliser une expérience qui va conduire à augmenter de manière très significative la probabilité de l'une des deux théories. Il suffit de faire disparaître la clause selon laquelle l'expérience entraînera une certitude. Et dans ce cas, le problème resurgit avec toute son acuité.

⁷² Trad.: On se trouve en l'an 2100 et les physiciens ont réussi à limiter la recherche d'une théorie de tout à seulement deux théories qui demeurent les deux seules plausibles candidates, T_1 et T_2 (selon des considérations ayant trait à la super-trop-super symétrie). Selon T_1 le monde est très, très grand mais fini, et il y a un total d'observateurs dans le cosmos qui s'élève à un trillion de trillions. Selon T_2 , le monde est très, très, très grand mais fini, et il y a un trillion de trillions de trillions d'observateurs. Les considérations de super-trop-super symétrie sont indifférentes par rapport à chacune de ces deux théories. Les physiciens préparent une expérience simple qui falsifiera l'une des théories. Mais voici qu'arrive le philosophe présomptueux: "Salut les gars, il est complètement inutile que vous fassiez l'expérience, parce que je peux déjà vous prouver que T_2 est à peu près un milliard de milliards de fois plus probablement vrai que T_1 (sur quoi le philosophe lance l'expérience de l'Incubateur et fait appel à SIA)!"

⁷³ Cf. from super-duper symmetry. Trad.: la super-trop-super symétrie.

A ce stade, il convient de s'attacher à modéliser l'expérience du *philosophe présomptueux* dans le cadre des n-univers. On peut traduire ainsi les deux hypothèses qui résultent de l'expérience:

(H1₆₉) l'univers comprend 10^{24} êtres

(H2₇₀) l'univers comprend 10^{36} êtres

et la situation correspondante peut être modélisée dans un $\Omega\alpha$, un n-univers à objets multiples, où les objets sont en relation *plusieurs* avec les constantes temporelle et spatiale. En termes de modèle probabiliste, on a alors les hypothèses correspondantes:

(H1₇₁) l'urne comprend 10^{24} boules

(H2₇₂) l'urne comprend 10^{36} boules

qui peuvent être traduites ainsi, avec les données quantitatives inhérentes au *modèle des deux urnes*:

(H1₇₃) l'urne comprend 10^1 boules

(H2₇₄) l'urne comprend 10^3 boules

La situation correspondante peut alors être transposée dans le $\Omega\alpha^*C$, avec les données qui sont celles du *modèle des deux urnes*:

(H1₇₅) l'urne contient 10 boules rouges

(H2₇₆) l'urne contient 1000 boules dont 10 boules rouges et 990 boules vertes

On le voit, la structure du *philosophe présomptueux* est clairement celle du raisonnement (GCT+) appliqué au *pile ou face divin*. En effet, la structure du *philosophe présomptueux* qui vient d'être mise en évidence laisse clairement apparaître que l'existence des 10^{12} êtres supplémentaires (ou les 990 boules vertes) est hypothétique. La structure du *philosophe présomptueux* apparaît clairement comme une variation du raisonnement (GCT+) appliqué au *pile ou face divin*. Par conséquent, le *philosophe présomptueux* s'analyse, de même que le raisonnement (GCT+) dans le *pile ou face divin*, comme un raisonnement non valide, car il est basé sur SIA, une application inconditionnelle du *modèle à majoration*. Or le *modèle à majoration* requiert pour s'appliquer valablement que (H2) soit vérifiée. Le *modèle à majoration* pourrait ainsi s'appliquer valablement si les hypothèses étaient les suivantes:

(H1₇₇) je proviens d'une population de 10^{24} êtres dans l'univers en T_1

(H2₇₈) je proviens d'une population de 10^{36} êtres dans l'univers en T_2

Dans un tel cas en effet, si les propositions suivantes:

(H1₇₉) il existe une population de 10^{24} êtres dans l'univers en T_1

(H2₈₀) il existe une population de 10^{36} êtres dans l'univers en T_2

étaient vraies, alors les conditions du *modèle à majoration* seraient satisfaites. On pourrait alors conclure valablement à une plus forte probabilité de l'hypothèse (H2), en vertu du *modèle à majoration*. En revanche, dans les conditions qui sont celles du *philosophe présomptueux*, on ne peut appliquer valablement le *modèle à majoration*, car l'existence d'une population supplémentaire de 10^{12} êtres est hypothétique. Et ceci interdit d'appliquer le *modèle à majoration* dans ces circonstances.

Ainsi qu'on vient de le voir, on peut résoudre les expériences basées sur SIA que constituent le *pile ou face divin* et le *philosophe présomptueux* décrites par Bostrom en montrant que leur structure est celle de l'application du *modèle à majoration* alors que les conditions n'en sont pas réunies. En particulier, l'existence des 990 boules vertes n'est pas établie. L'erreur de raisonnement dans SIA et les expériences en découlant est constituée par une application inconditionnelle du *modèle à majoration*. Car on ne peut faire légitimement dans ces conditions qu'une application conditionnelle du *modèle à majoration*. SIA fait application du *modèle à majoration* alors que les conditions n'en sont pas

satisfaites. Et le problème avec SIA est que ce dernier principe assimile une existence hypothétique à une existence certaine, ainsi que cela est mentionné de manière explicite par Olum (2000).

Par ailleurs, l'application de SIA a pu être présentée dans la littérature comme une solution à DA. Plusieurs analyses récentes⁷⁴ de DA sont en effet basées sur l'application de SIA dans la situation qui est celle de DA. Mais les considérations qui précèdent concernant SIA conduisent également à rejeter l'application de SIA en tant que solution à DA. Dans la situation correspondant à DA, on formule en effet deux hypothèses:

(H1₈₁) le nombre d'humains ayant jamais existé atteindra 100 milliards

(H2₈₂) le nombre d'humains ayant jamais existé atteindra 10000 milliards

Mais il apparaît ici que l'existence des 9900 milliards d'humains supplémentaires de l'hypothèse (H2) est hypothétique. Ici, l'analyse de SIA qui précède montre que l'application de SIA n'est pas valable dans la situation de DA, car on n'a de certitude, de même que pour le *pile ou face divin*, que l'effectif de l'humanité correspond à la grande urne. Et on se trouve donc dans une situation tout à fait analogue à celle du *pile ou face divin* où SIA, on l'a vu, ne peut être appliqué valablement, car il s'agit d'une application illicite du modèle à majoration. Il en va ainsi de même pour l'application de SIA dans la situation correspondant à DA.

Enfin, il convient de mentionner à ce stade qu'il serait erroné de présenter SSA et le *modèle à majoration* (dont SIA constitue une application irrégulière) comme deux principes alternatifs et exclusifs l'un de l'autre. En effet, comme on l'a vu plus haut, le *modèle à majoration* nécessite SSA pour fonctionner véritablement. Car pour appliquer le *modèle à majoration*, il s'avère nécessaire de tirer une boule au hasard, en vertu donc de SSA. Ainsi, SSA se révèle tout à fait compatible avec le modèle à majoration.

⁷⁴ Cf. Olum (2000).

Chapitre 7. Le problème de la Belle au bois dormant

7.1. Le problème

Le problème de la *Belle au bois dormant* (*Sleeping Beauty Problem*) a donné lieu récemment à des discussions par Adam Elga (2000) et David Lewis (2001)⁷⁵. Le problème de la *Belle au bois dormant* est en particulier décrit ainsi par Adam Elga (2000):

Some researchers are going to put you to sleep. During the two days that your sleep will last, they will briefly wake you up either once or twice, depending on the toss of a fair coin (Heads: once; Tails: twice). After each waking, they will put you to back to sleep with a drug that makes you forget that waking. When you are first awakened, to what degree ought you believe that the outcome of the coin toss is Heads?

First answer: 1/2, of course! Initially you were certain that the coin was fair, and so initially your credence in the coin's landing Heads was 1/2. Upon being awakened, you receive no new information (you knew all along that you would be awakened). So your credence in the coin's landing Heads ought to remain 1/2.

Second answer: 1/3, of course! Imagine the experiment repeated many times. Then in the long run, about 1/3 of the wakings would be *Heads-wakings* - wakings that happen on trials in which the coin lands Heads. So on any particular waking, you should have credence 1/3 that that waking is a Heads-waking, and hence have credence 1/3 in the coin's landing Heads on that trial. This consideration remains in force in the present circumstance, in which the experiment is performed just once.⁷⁶

Il apparaît que le problème de la *Belle au bois dormant* comporte les éléments suivants⁷⁷. Soient tout d'abord les deux hypothèses suivantes:

(H1₈₃) la *Belle au bois dormant* sera réveillée une seule fois

(H2₈₄) la *Belle au bois dormant* sera réveillée deux fois

⁷⁵ Cf. également Arntzenius (2002) et Monton (2002).

⁷⁶ Trad.: Des chercheurs se préparent à vous endormir. Pendant les deux jours que durera votre sommeil, ils vous réveilleront brièvement soit une soit deux fois, selon le résultat du lancer d'une pièce de monnaie équilibrée (face: une fois; pile: deux fois). Après chaque réveil, ils vous endormiront à nouveau à l'aide d'une drogue qui vous fera oublier ce dernier réveil. Lorsque vous êtes réveillé pour la première fois, à quel degré devez-vous croire que le résultat du lancer de la monnaie est face? *Première réponse:* 1/2, naturellement! Au début, vous étiez certain que la pièce de monnaie était équilibrée, de telle manière qu'au début votre croyance dans le fait que le résultat du lancer de la pièce était face était de 1/2. Etant réveillé, vous ne recevez aucune nouvelle information (vous saviez depuis toujours que vous seriez réveillé). Ainsi votre croyance dans le fait que le résultat du lancer de la pièce est devrait demeurer à 1/2. *Seconde réponse:* 1/3, naturellement! Imaginez l'expérience répétée un grand nombre de fois. A la longue, environ 1/3 des réveils seraient des *réveils/face* - les réveils qui se produisent dans les cas où la pièce de monnaie tombe sur face. Ainsi, à l'occasion d'un réveil donné, vous devriez avoir une croyance égale à 1/3 qu'il s'agit d'un réveil/face, et par conséquent avoir une croyance de 1/3 dans le fait que la pièce de monnaie est tombée sur face dans ce cas particulier. Cette constatation demeure avec force dans la circonstance présente, où l'expérience n'est exécutée qu'une seule fois.

⁷⁷ On peut mettre en correspondance les notations qui sont celles d'Elga et de Lewis, avec celles utilisée ici. On a ainsi les correspondances suivantes. Monday et Tuesday dans Elga correspondent respectivement ici à T_1 et T_2 . De même, $t1$ et $t2$ dans Elga correspondent respectivement à $\circ T_1|H2$ et $\circ T_2|H2$. Enfin, Elga considère l'hypothèse (H) selon laquelle (H1) est vraie si j'existe en T_1 , ce qui équivaut avec la présente notation à: $H1|\circ T_1$.

Soit également l'événement E: la *Belle au bois dormant* est réveillée. Les probabilités initiales sont les suivantes: $P(H1) = P(FACE) = 0,5$ et $P(H2) = P(PILE) = 0,5$. Ici, on cherche à connaître les probabilités a posteriori, compte tenu de la survenance de l'événement E: $P'(H1)$ c'est-à-dire $P(H1|E)$ et $P'(H2)$ c'est-à-dire $P(H2|E)$. Ici, la probabilité que la *Belle au bois dormant* soit réveillée si la pièce est tombée sur face c'est-à-dire $P(E|H1)$ est telle que $P(E|H1) = 1$. Et de même la probabilité que la *Belle au bois dormant* soit réveillée si la pièce est tombée sur pile c'est-à-dire $P(E|H2)$ est telle que $P(E|H2) = 1$. Le problème de la *Belle au bois dormant* peut être ainsi défini: quelle est la manière correcte de calculer $P'(H1)$ et $P'(H2)$ en prenant en compte le fait que la *Belle au bois dormant* vient d'être réveillée? Le problème de la *Belle au bois dormant* apparaît plus nettement lorsqu'on considère notamment deux types de raisonnements qui semblent valides et peuvent apparemment être tenus de manière concurrente dans ce type de situation. Un premier type de raisonnement (I) est le suivant. Aucune donnée nouvelle, aucun élément nouveau n'est venu modifier les probabilités initiales. Par conséquent, on doit conserver les probabilités initiales qui résultent de la donnée de l'énoncé selon laquelle la pièce est équilibrée. Ainsi, $P'(H1) = P'(H2) = 0,5$. Car il serait rationnel de modifier des probabilités initiales en fonction de données nouvelles. Mais étant donné qu'aucune donnée nouvelle n'est apparue, il n'y a pas lieu de modifier en conséquence les probabilités a priori. Un tel raisonnement correspond, de manière simplifiée, à celui qui est mis en oeuvre par Lewis (2001).

Mais de manière alternative, il semble que l'on puisse également tenir un raisonnement de type (II). Un tel raisonnement est notamment décrit par Elga (2000). Il s'établit comme suit. Formalisons par σT_i le fait que la *Belle au bois dormant* existe en T_i . Elga considère ainsi que lorsque la *Belle au bois dormant* se réveille, elle se trouve dans l'un des trois cas suivants:

- (E1₈₅) (H1) & σT_1
- (E2₈₆) (H2) & σT_1
- (E3₈₇) (H2) & σT_2

Elga pose montre tout d'abord que $P(E2) = P(E3)$. Elga se place dans la situation où la *Belle au bois dormant* apprend à son réveil que le jour est lundi. Elga s'attache ensuite à montrer que $P(E1) = P(E2)$. On a: $P(H1) = P(H1|E1 \vee E2)$ et d'autre part $P(H1|E1 \vee E2) = P(E1|E1 \vee E2) = P(E1)/[P(E1)+P(E2)]$. Or $P(H1) = 1/2$. Par conséquent: $P(E1)/[P(E1)+P(E2)] = 1/2$. D'où il s'ensuit $P(E1) = P(E2)$. Par conséquent, $P(E1) = P(E2) = P(E3)$. Et étant donné que $P(E1) + P(E2) + P(E3) = 1$, il s'ensuit que $P(E1) = P(E2) = P(E3) = 1/3$. Il en résulte que $P'(H1) = 1/3$ et $P'(PILE) = P'(H2) = 1 - 1/3 = 2/3$.

Ainsi, le *problème de la Belle au bois dormant* réside précisément dans le fait de déterminer lequel des raisonnements (I) ou (II) est valide. Car les conclusions qui résultent des deux types de raisonnements sont contradictoires. Le raisonnement (I) conduit à $P'(H1) = P'(H2) = 0,5$ alors que le raisonnement (II) conclut à $P'(H1) = 1/3$ et $P'(H2) = 2/3$.

7.2. Modélisation dans le n-univers temporel à objet unique (ΩT^*)

A ce stade, il convient de s'attacher à modéliser la situation concrète correspondant au problème de la *Belle au bois dormant* en termes de n-univers. Il apparaît tout d'abord que le problème de la *Belle au bois dormant* concerne un être unique: la Belle au bois dormant. On se situe donc dans un n-univers à objet unique. De plus, la situation prend place en un lieu donné. Le n-univers correspondant comporte donc une constante locale (L_0). De plus, on distingue dans l'expérience les jours 1 (lundi) et 2 (mardi). Le n-univers correspondant comporte donc deux positions temporelles: T_1 et T_2 . On se place donc dans un n-univers temporel, soit un $\Omega T \dots$. Enfin, la *Belle au bois dormant* peut exister à la fois aux positions temporelles T_1 et T_2 . Plus encore, l'objet unique apparaît ainsi comme persistant: s'il existe en T_2 , c'est qu'il a également existé en T_1 . Le n-univers correspondant est donc à démultiplication temporelle. Par conséquent, le *problème de la Belle au bois dormant* prend place dans un ΩT^* , un n-univers à objet unique, alocalisé et temporel, où l'objet unique possède une propriété de démultiplication par rapport au critère temporel. Dans ce contexte, l'événement E correspond au fait que la *Belle au bois dormant* existe (est réveillée) à l'une des deux positions temporelles, T_1 ou T_2 .

Attachons-nous maintenant à modéliser la situation concrète correspondant au problème de la *Belle au bois dormant* en termes d'urne. On considère une boule unique dans une urne. Cette boule existe en T_1 et peut de même exister également en T_2 . On a alors les deux hypothèses suivantes:

- (H1₈₈) la boule existe uniquement en T_1
- (H2₈₉) la boule existe à la fois en T_1 et en T_2

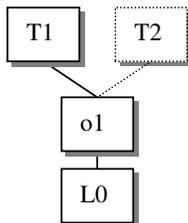
Et il est ici commode de considérer que (H1) et (H2) sont respectivement équivalentes à:

- (H1₉₀) la boule existe en T_1 et n'existe pas en T_2
- (H2₉₁) la boule existe en T_1 et existe en T_2

c'est-à-dire deux propositions dont la structure est:

- (H1₉₂) $oT_1 \wedge \bar{o}T_2$
- (H2₉₃) $oT_1 \wedge oT_2$

en formalisant par oT_i le fait que la boule existe en T_i et par $\bar{o}T_i$ le fait que la boule n'existe pas en T_i . On a également l'événement E: une boule est tirée dans l'urne (la *Belle au bois dormant* est réveillée).



Enfin, il apparaît utile de modéliser le problème de la *Belle au bois dormant* en considérant la situation en termes de n-univers purs. Ici, on considère que le n-univers correspondant n'existe que pendant la durée d'existence de l'objet unique qui s'y trouve. Le ΩT^* n'existe donc que pendant la durée d'existence de la boule unique. Ainsi, le ΩT^* possède soit une position temporelle (T_1), soit deux positions temporelles (T_1 et T_2). On a alors deux hypothèses:

- (H1₉₄) le ΩT^* comporte 1 position temporelle
- (H2₉₅) le ΩT^* comporte 2 positions temporelles

7.3. Modélisation dans d'autres n-univers

La modélisation du problème de la *Belle au bois dormant* dans le ΩT^* étant établie, on peut s'interroger maintenant pour savoir si problème de la *Belle au bois dormant* ne peut pas être modélisé également dans d'autres n-univers. Car la structure du ΩT^* est un $\Omega\beta^*$. Ceci suggère en effet que le problème de la *Belle au bois dormant* pourrait prendre place dans un n-univers tel qu'un ΩL^* ou un ΩC^* .

Il apparaît cependant ici que l'on peut écarter d'emblée le ΩL^* . En effet, un tel n-univers est à démultiplication locale, et l'objet unique peut s'y trouver en plusieurs lieux au même moment. Or une telle propriété ubiquiste ne correspond pas aux caractéristiques qui sont celles d'un univers concret tel que le notre dans lequel prend place le problème de la *Belle au bois dormant*.

De même, on peut éliminer également le ΩC^* , un n-univers comportant un objet unique à démultiplication par couleur. Dans un tel n-univers, un objet peut présenter plusieurs couleurs en un même lieu et au même moment. Or un tel n-univers ne correspond pas non plus aux caractéristiques qui sont celles de notre univers réel.

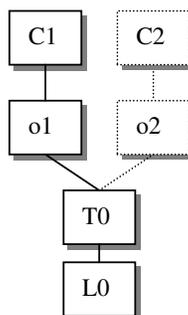
Enfin, on peut se demander si le problème de la *Belle au bois dormant* ne peut pas être modélisé dans un $\Omega\alpha$. Un tel n-univers comporte de multiples objets⁷⁸ en un lieu donné (L_0) et à un moment donné (T_0). De plus, les objets dans un tel n-univers sont en relation *plusieurs* avec les constantes de temps et de lieu. Plusieurs objets peuvent en effet se trouver au même moment T_0 en un même lieu L_0 . Dans un $\Omega\alpha$, le nombre d'objets ne peut varier par rapport à un critère-variable donné, puisqu'il n'existe que des critères-constants (de lieu et de temps). En conséquence, le nombre d'objets est constant dans le $\Omega\alpha$. Il en résulte alors la variation qui correspond au modèle probabiliste suivant. On tire une boule dans une urne à un moment T_0 donné. On a alors les deux hypothèses suivantes:

- (H1₉₆) l'urne contient une boule
- (H2₉₇) l'urne contient deux boules

On a aussi les probabilités initiales $P(H1) = P(H2) = 0,5$. Soit également l'événement E: je tire une boule dans l'urne. Ici, on a $P(E|H1) = P(E|H2) = 1$. Cependant ici, il apparaît que ce modèle ne permet pas de rendre compte du fait que les positions temporelles sont individualisées dans le problème de la *Belle au bois dormant*, où T_1 et T_2 sont clairement distinguées. Dans le $\Omega\alpha$ en effet, on ne peut distinguer les boules entre elles. Il s'avère ainsi nécessaire de recourir à un critère supplémentaire. Considérons donc un $\Omega\alpha^*C$, un n-univers à objets multiples et colorés où chaque objet possède une couleur différente. Il suffit ici de considérer deux taxons de couleur: rouge et vert. De plus, plusieurs objets peuvent posséder une même couleur. Il s'agit ainsi d'un n-univers où les objets multiples sont en relation *plusieurs* avec le critère de couleur. La situation correspondant au problème de la *Belle au bois dormant* est alors celle d'une urne contenant soit une boule rouge, soit une boule rouge et une boule verte. On a ainsi les deux hypothèses suivantes:

- (H1₉₈) l'urne contient une boule rouge
- (H2₉₉) l'urne contient deux boules dont une boule rouge et une boule verte

et l'événement E correspond ici au fait que l'on extrait une boule de l'urne. A ce stade, il apparaît qu'une telle modélisation du problème de la *Belle au bois dormant* dans le $\Omega\alpha^*C$ préserve les conditions qui sont celles de l'expérience originale.



On peut mentionner enfin la version correspondante, qui est définie en termes de n-univers purs. Un $\Omega\alpha^*C$ contient en T_0 et en L_0 soit un objet unique soit deux objets. On a les deux hypothèses:

- (H1₁₀₀) le $\Omega\alpha$ contient un objet rouge
- (H2₁₀₁) le $\Omega\alpha$ contient deux objets dont un objet rouge et un objet vert

⁷⁸ Que distingue un critère α .

7.4. Analyse

La modélisation du problème de la *Belle au bois dormant* dans le $\Omega\alpha^*C$ qui vient d'être décrite présente une particularité intéressante. En effet, le $\Omega\alpha^*C$ constitue également le n-univers où le *modèle des deux urnes* et l'expérience du *pile ou face divin* prennent place. Cette caractéristique permet ainsi de comparer aisément ces différentes situations. A ce stade, il apparaît nécessaire de s'attacher à souligner à la fois les points communs et les différences entre d'une part le problème de la *Belle au bois dormant* et d'autre part le *modèle des deux urnes* et le *pile ou face divin*.

Attachons-nous, en premier lieu, à mettre en évidence les points communs entre le problème de la *Belle au bois dormant* et le *modèle des deux urnes*. On a vu précédemment que le problème de la *Belle au bois dormant* pouvait être modélisé dans un $\Omega\alpha^*C$. Considérons donc la version du problème de la *Belle au bois dormant* définie dans un $\Omega\alpha^*C$, un n-univers à objets multiples et coloré, où les objets se trouvent en relation *plusieurs* avec le critère de couleur. On a ici deux taxons de couleur et les boules y sont rouges ou vertes. Cette version du problème de la *Belle au bois dormant* prend ainsi place dans le même n-univers que le *modèle des deux urnes* bicolore. Considérons le modèle probabiliste correspondant au problème de la *Belle au bois dormant*. On a ainsi soit une urne comportant une boule rouge, soit une urne comportant deux boules respectivement rouge et verte. Les deux hypothèses correspondantes sont les suivantes:

(H1₁₀₂) l'urne comprend 1 boule rouge

(H2₁₀₃) l'urne comprend 2 boules dont une boule rouge et une boule verte

Le problème de la *Belle au bois dormant* est rendu ici par le fait que l'urne contient une boule rouge si la pièce tombe sur face, et deux boules - respectivement rouge et verte - si la pièce tombe sur pile. Ici, l'existence de la boule rouge est certaine, alors que l'existence de la boule verte est hypothétique. La probabilité initiale P(FACE) c'est-à-dire P(H1) est égale à 0,5. Ainsi, P(H1) est la probabilité qu'une urne comportant une seule boule (rouge) ait été créée. Rappelons à cet égard les hypothèses qui sont celles du *modèle des deux urnes* bicolore dans le $\Omega\alpha^*C$:

(H1₁₀₄) l'urne contient 10 boules rouges

(H2₁₀₅) l'urne contient 1000 boules dont 10 boules rouges et 990 boules vertes

A ce stade, il apparaît clairement que les hypothèses (H1) et (H2) qui sont celles du problème de la *Belle au bois dormant* et du *modèle des deux urnes* présentent une structure identique.

En second lieu, il est intéressant de souligner les différences existant entre la situation où prend place le problème de la *Belle au bois dormant* et celle du *modèle des deux urnes*. De ce point de vue, il apparaît que dans le problème de la *Belle au bois dormant*, l'événement E est le suivant: la *Belle au bois dormant* se réveille. Et au niveau de la modélisation du problème de la *Belle au bois dormant* dans le $\Omega\alpha^*C$, l'événement E correspond à: "je tire une boule dans l'urne". Or le *modèle des deux urnes* qui conduit à l'application du modèle à minoration est basé sur l'événement E: je tire une boule dans l'urne dont le numéro est 5⁷⁹. Il s'agit là d'une différence importante entre les deux situations. Ici, un événement E analogue dans le problème de la *Belle au bois dormant* aurait été: je tire une boule dans l'urne dont la couleur est rouge (lundi, Monday). Mais l'événement E qui sert de support aux types de raisonnements concurrents dans le problème de la *Belle au bois dormant* est: "je tire une boule dans l'urne", sans mention de la couleur de la boule⁸⁰. Ainsi, les événements qui sous-tendent les deux situations se révèlent-ils être d'une nature distincte. Or cet événement est celui qui détermine le calcul des probabilités P(H1) et P(H2). Il s'agit là par conséquent d'une différence importante entre les deux problèmes.

⁷⁹ Plus précisément, dans le $\Omega\alpha^*C$, l'événement E est le suivant: je tire une boule dans l'urne dont le numéro de couleur est 5.

⁸⁰ Dans certaines variations du *problème de la beauté endormie*, l'événement E est le suivant: "*Beauté endormie* se réveille et apprend que le jour est lundi". Une telle version est notamment mentionnée par Monton (2002). Cf. "what probability should she assign to *H* on Monday, when she wakes up" (p. 47). Trad.: "quelle probabilité doit-elle assigner à *H* le lundi, lorsqu'elle se réveille".

A ce stade, il convient d'effectuer une comparaison entre la situation du problème de la *Belle au bois dormant* et celle du *pile ou face divin*. Commençons par décrire les points communs entre les deux expériences. On compare ici la version du problème de la *Belle au bois dormant* qui prend place dans un $\Omega\alpha^*C$ et la version du *pile ou face divin* qui est modélisée dans ce même n-univers. Ainsi, les deux versions prennent place dans un n-univers à objets multiples et coloré, où les objets sont en relation *plusieurs* avec le critère de couleur. Un tel n-univers comprend deux taxons de couleur (rouge et vert). D'autre part, plusieurs objets peuvent présenter une même couleur. Ainsi, les objets multiples sont en relation *plusieurs* avec le critère de couleur. Rappelons également comment les différentes situations qui sont celles du problème de la *Belle au bois dormant* prennent place. Comment modélise-t-on ainsi le fait que la *Belle au bois dormant* soit réveillée un lundi? Ceci correspond au fait que l'on tire dans l'urne une boule rouge. De même, à quoi correspond le fait que la *Belle au bois dormant* soit réveillée un mardi? Cela correspond au fait que l'on tire dans l'urne une boule verte. De même, comment modélise-t-on le fait que la *Belle au bois dormant* soit réveillée un lundi, consécutivement à un tirage face? Il apparaît ici que cela correspond au fait que l'on tire dans l'urne une boule rouge alors que la pièce est tombée sur face. On a en effet trois possibilités:

- (E1₁₀₆) (H1) & une boule rouge est tirée dans l'urne
- (E2₁₀₇) (H2) & une boule rouge est tirée dans l'urne
- (E3₁₀₈) (H2) & une boule verte est tirée dans l'urne

Enfin, les hypothèses respectives sur lesquelles portent les deux situations sont les suivantes. Dans le problème de la *Belle au bois dormant*, les hypothèses sont:

- (H1₁₀₉) l'urne comprend 1 boule rouge
- (H2₁₁₀) l'urne comprend 2 boules dont une boule rouge et une boule verte

alors que les hypothèses du *pile ou face divin* sont⁸¹:

- (H1₁₁₁) l'urne contient 10 boules rouges
- (H2₁₁₂) l'urne contient 1000 boules dont 10 boules rouges et 990 boules vertes

On le voit ainsi de manière générale, les caractéristiques du n-univers dans lequel le problème de la *Belle au bois dormant* et le *pile ou face divin*, ainsi que les hypothèses correspondantes présentent une structure tout à fait identique.

Attachons-nous maintenant à mettre en évidence les différences existant entre la version du problème de la *Belle au bois dormant* dans le $\Omega\alpha^*C$ et le *pile ou face divin*. Il apparaît ainsi que le problème de la *Belle au bois dormant* ne comporte qu'un seul événement E, alors que le *pile ou face divin* comporte deux événements E_a et E_b. L'événement E dans le problème de la *Belle au bois dormant* est ainsi le suivant: la *Belle au bois dormant* est réveillée. En termes d'urne, l'événement E correspond au fait qu'une boule est extraite de l'urne. Dans l'expérience du *pile ou face divin* décrite par Bostrom en revanche, les étapes (a) et (b) sont distinguées. L'étape (a) correspond au fait que l'événement E_a survient: une boule est extraite dans l'urne. Et l'étape (b) correspond au fait que l'événement E_b se produit: la boule possède la couleur n° 5⁸². On le voit, le problème de la *Belle au bois dormant* et le *pile ou face divin* diffèrent seulement par le fait que le problème de la *Belle au bois dormant* ne comporte que l'événement E_a correspondant à l'étape (a) alors que l'expérience du *pile ou face divin* décrite par Bostrom⁸³ comprend à la fois les événements E_a et E_b, correspondant aux étapes (a) et (b).

L'analyse qui précède montre que la situation qui est celle du problème de la *Belle au bois dormant* apparaît finalement tout à fait identique à une version de l'expérience du *pile ou face divin* décrite par Bostrom, qui ne comporterait que l'étape (a). Elle est également structurellement identique à la version

⁸¹ Il s'agit des mêmes hypothèses que celles du *modèle des deux urnes* bicolore dans le $\Omega\alpha^*C$.

⁸² Ou de manière équivalente: la boule n° 5 a été tirée.

⁸³ La version de l'expérience Le pile ou face divin décrite par Leslie décrite plus haut ne comporte que l'étape (a).

du *pile ou face divin* décrite par Leslie. Ainsi, il apparaît que le problème de la *Belle au bois dormant* constitue simplement une variation du *pile ou face divin*. Alors que l'expérience du *pile ou face divin* prend place dans un $\Omega\alpha^*C$, un n-univers à objets multiples et coloré où les objets sont en relation *plusieurs* avec la couleur, le problème de la *Belle au bois dormant* se situe dans un ΩT^* , un n-univers à objet unique, alocalisé et à démultiplication temporelle. Mais on l'a vu, on a une modélisation équivalente de problème de la *Belle au bois dormant* dans un $\Omega\alpha^*C$.

Compte tenu de l'équivalence entre le problème de la *Belle au bois dormant* et l'étape (a) de l'expérience du *pile ou face divin*, il convient donc d'apporter au problème de la *Belle au bois dormant* une réponse de même nature que celle fournie à l'étape (a) du *pile ou face divin*. Cette réponse est la suivante. A l'étape (a), l'événement E survient et la *Belle au bois dormant* se réveille (une boule est extraite de l'urne). A ce stade, la *Belle au bois dormant* peut faire une application conditionnelle du modèle à majoration. On calcule alors de même que dans *pile ou face divin*, la probabilité que la *Belle au bois dormant* soit réveillée le lundi $P(H3)$ est égale à $1*0,5+0,5*0,5 = 0,75$. Et de même la probabilité que la *Belle au bois dormant* soit réveillée le mardi $P(H4)$ est égale à $1 - P(H3) = 0,25$. Il s'agit là d'une application conditionnelle du modèle à majoration qui se révèle correcte. Ceci éclaire en revanche comment le raisonnement (II) est incorrect car il conduit à une application inconditionnelle du modèle à majoration. Le raisonnement (I) prévaut ainsi et de manière identique qu'à l'étape (a) du *pile ou face divin*, on a ainsi $P(H1) = P(H2) = 0,5$, étant donné que la pièce est équilibrée.

Finalement, quelle aurait été la situation si le problème de la *Belle au bois dormant* avait comporté une étape (b) analogue à celle du *pile ou face divin* dans sa version décrite par Bostrom? On pourrait alors reformuler le problème de la *Belle au bois dormant* de manière plus précise en distinguant deux types d'événements correspondant aux deux étapes (a) et (b). L'énoncé du problème de la *Belle au bois dormant* aurait alors comporté deux événements: un événement E_a , consistant dans le fait que la *Belle au bois dormant* est réveillée, et un événement E_b , consistant dans le fait que la *Belle au bois dormant* apprend que l'on est lundi (jour 1). Dans le modèle probabiliste correspondant, les événements E_a et E_b correspondent respectivement à "une boule est extraite de l'urne" et "la couleur de la boule est rouge". Une telle variation plus complète présenterait alors une structure tout à fait identique à celle de la version du *pile ou face divin* décrite par Bostrom. A l'étape (a), on aurait raisonné ainsi que cela vient d'être mentionné. Et de même, à l'étape (b), on aurait raisonné de la manière suivante, tout à fait analogue à celle qui prévaut pour la même étape du *pile ou face divin*, avec des données légèrement différentes. Ainsi, à l'étape (b), l'événement E_b survient et la *Belle au bois dormant* est informée qu'elle se trouve le lundi (la couleur de la boule extraite de l'urne est rouge). Dans ces circonstances, la *Belle au bois dormant* peut faire application du modèle à minoration. Il s'ensuit alors, par application de la formule de Bayes: $P'(H1) = 2/3$. D'où $P'(H2) = 1 - P'(H1) = 1/3$.

Chapitre 8. L'expérience des segments de temps

8.1. Description

Soit l'expérience suivante, qui a été exposée par Oliver et Korb (1998)⁸⁴:

if you number the minutes of your life starting from the first minute you were aware of the applicability of the Argument to your life span to the last such minute and if you then attempt to estimate the number of the last minute using your current sample of, say, one minute, then according to the Doomsday Argument, you should expect to die before you finish reading this article.⁸⁵

Une telle expérience montre comment un raisonnement de même nature que DA peut être appliqué au nombre de minutes de la vie de chaque être humain. Dans l'expérience de Oliver et Korb, l'événement E consiste dans le fait de tirer la première minute où nous avons été informé de la nature d'un argument du type de DA. La conclusion qui en résulte est une augmentation de la probabilité initiale selon laquelle le nombre de minutes qui nous reste à vivre est faible.

Une expérience du même type est également décrite par Jean-Paul Delahaye (1996) dans un manuscrit non publié. Delahaye expose ainsi une situation paradoxale fondée sur un raisonnement qui présente la structure de DA qu'il dénomme le *paradoxe des bébés*. Il distingue ainsi les bébés de type A et les bébés de type B. Les bébés de type A vivent moins d'un an. Les bébés de type B dépassent un an, vivant jusqu'à 70 ans en moyenne⁸⁶. Les données réelles concernant les bébés sont les suivantes: 1% des bébés sont des bébés de type A, leur durée de vie étant inférieure à un an. En revanche, 99% des bébés sont des bébés de type B, allant jusqu'à vivre environ 70 ans. Dans ce contexte, le raisonnement inhérent à DA peut prendre place. Considérons le fait que je suis un bébé qui est dans la première année de sa vie. La probabilité initiale que je meure dans ma première année est égale à 0,01. Mais la probabilité que je sois dans la première année de ma vie si je suis un bébé de type A est égale à 1, alors que la probabilité que je sois dans la première année de ma vie si je suis un bébé de type B est égale à 1/70. Ainsi, je dois réviser à la hausse la probabilité initiale selon laquelle je vais mourir dans ma première année. Le décalage bayésien conduit à une probabilité a posteriori qui est à peu près égale à 0,41.

Une particularité intéressante du paradoxe des bébés est que la conclusion qui en résulte est manifestement absurde. En effet, elle contredit les données qui sont vérifiées par l'expérience: les bébés de type A constituent 1% des bébés, alors que les bébés de type B en représentent 99%. Il s'ensuit une certitude que le décalage bayésien conduit à une conclusion qui contredit les faits, à la différence de DA où la conclusion qui en résulte est simplement contraire à l'intuition. Cette dernière particularité a ainsi conduit plusieurs auteurs à accepter purement et simplement les conclusions de DA. Mais dans le paradoxe des bébés, le décalage bayésien conduit à une conclusion paradoxale. Et la même attitude qui consiste à accepter la conclusion qui résulte de l'argument ne peut être adoptée ici.

Les deux expériences qui viennent d'être décrites présentent une structure commune, malgré quelques différences. En ce qui concerne tout d'abord les points communs entre les deux expériences, il apparaît en premier lieu qu'elles concernent un être unique (à la différence de DA qui concerne des êtres multiples), respectivement un être humain ou un bébé. En second lieu, les deux expériences se révèlent basées sur des probabilités relatives à la durée d'une vie.

⁸⁴ Cf. p. 405, n. 2.

⁸⁵ Trad.: si vous numérotez les minutes de votre vie à partir de la première minute où vous vous êtes rendu compte de l'applicabilité de l'argument à votre durée de vie, jusqu'à la dernière minute et si vous essayez alors d'estimer le nombre correspondant à cette dernière minute en utilisant votre échantillon actuel qui est, par exemple, d'une minute, puis en vertu de l'argument de l'apocalypse, vous devriez vous attendre à mourir avant que vous finissiez de lire cet article.

⁸⁶ Cf. (...) "1 % des bébés vivent moins d'un an (bébés de type A). 99 % des bébés dépassent un an et vivent alors - en moyenne - 70 ans (bébés de type B)."

A cet égard, les différences entre les deux expériences apparaissent mineures. Elles portent en effet tout d'abord sur le fait que la segmentation de la durée de vie s'effectue en minutes dans l'expérience de Oliver et Korb et en années dans celle de Delahaye. Cependant, on peut bien sûr produire une version de l'expérience de Oliver et Korb basée sur une segmentation en années ou bien encore une version de l'expérience de Delahaye basée sur une segmentation en minutes.

Une seconde différence entre les deux expériences réside dans la nature de l'événement E. En effet, dans l'expérience de Oliver et Korb, l'événement E est ainsi défini: étant donné que la minute 1 de mon existence est celle où j'ai été au courant pour la première fois de DA, je me trouve actuellement à la minute m (avec $m \geq 1$). En revanche, dans le paradoxe des bébés, l'événement E est le suivant: je suis à la minute 1 de mon existence. Une telle différence se révèle toutefois également secondaire. En effet, on pourrait considérer une version de l'expérience de Oliver et Korb basée sur un événement E défini de la manière suivante: mon existence est définie à partir du moment où j'ai été au courant pour la première fois de DA, et je me trouve actuellement à la minute 1 de mon existence. De même, on pourrait prendre en compte une version de l'expérience de Delahaye basée sur l'événement E: je suis actuellement à la minute m (avec $m > 1$) de mon existence. Il apparaît ici que de telles variations n'apportent pas de changement fondamental, et conduisent bien à des versions dans lesquelles le décalage bayésien de même nature que DA prend également place.

A partir des deux expériences précédentes, on peut en outre construire aisément d'autres variations qui conduisent à la même problématique. Soit par exemple la version suivante, qui présente sensiblement la même nature que les deux précédentes. Je suis au départ d'une course de 100 m. Je m'apprête à franchir le premier mètre. Je considère les deux hypothèses concurrentes:

(H1₁₁₃) je ne vais parcourir qu'un mètre

(H2₁₁₄) je vais parcourir cent mètres

pour lesquelles les probabilités initiales sont de 0,5. Par un raisonnement bayésien, je préfère l'hypothèse (H1) par rapport à (H2). Je conclus que je ne parviendrai pas au bout des 100 m et qu'avec une probabilité d'environ 0,99, je ne parviendrai pas à dépasser le premier mètre.

8.2. Modélisation dans le n-univers temporel et coloré à objet unique (ΩT^*C)

Il convient maintenant de s'attacher à modéliser l'expérience des segments de temps dans le cadre des n-univers. Dans quel type de n-univers l'expérience des segments de temps prend-elle place? Il apparaît ici que l'expérience concerne en premier lieu un être humain unique. Il s'agit là d'une importante différence avec DA qui concerne de multiples êtres humains. Ainsi, une première caractéristique du n-univers considéré est qu'il comporte un objet unique. En second lieu, on peut considérer que l'objet unique se trouve en un lieu constant. Le n-univers correspondant présente donc une constante locale (L_0). En troisième lieu, le n-univers considéré apparaît comme temporel⁸⁷, puisque l'objet unique existe à différentes positions temporelles. L'expérience de Oliver et Korb est en effet basée sur un temps discret correspondant à une segmentation en minutes, alors que celle de Delahaye est basée sur une segmentation en années. Il apparaît ainsi que l'on se place dans un ΩT En outre, l'objet unique est susceptible d'exister à différentes positions temporelles. Ainsi, le n-univers en question est-il à démultiplication temporelle. Plus encore, l'objet unique est persistant: s'il existe en T_i et en T_j (avec $i < j$) alors il a existé en $T_i, T_{i+1}, \dots, T_{j-1}, T_j$. En conséquence, le n-univers dans lequel prend place l'expérience des segments de temps présente les caractéristiques suivantes: il s'agit d'un n-univers présentant un objet unique, comportant une constante de localisation et une variable temporelle à démultiplication. Ainsi, l'expérience des segments de temps prend-elle place dans un ΩT^* .

A ce stade, on est à même de modéliser l'expérience des segments de temps en termes d'urne. On considère ainsi une boule unique, qui existe à différentes positions temporelles successives: la boule

⁸⁷ A la différence notamment du *modèle des deux urnes*.

existe en T_1, T_2, \dots, T_n (pour $n \geq 1$). La boule possède une propriété de persistance temporelle: si elle existe en T_i , alors elle a existé en T_1, T_2, \dots, T_{i-1} . On formule alors les deux hypothèses suivantes⁸⁸:

(H1₁₁₅) la boule existe jusqu'en T_{10}

(H2₁₁₆) la boule existe jusqu'en T_{1000}

qui sont équivalentes à:

(H1₁₁₇) la durée d'existence de la boule est T_1-T_{10}

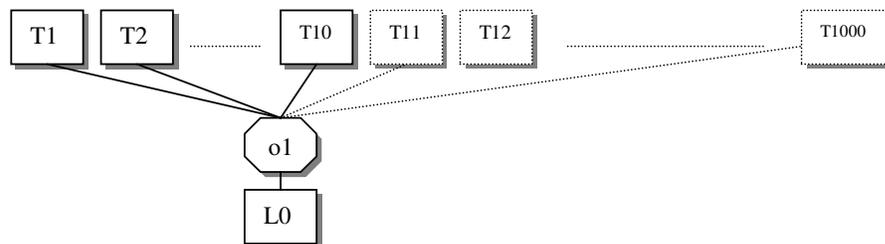
(H2₁₁₈) la durée d'existence de la boule est T_1-T_{1000}

Ici, l'événement E est le suivant: la boule se trouve en T_1 . On a également les probabilités a priori $P(H1) = P(H2) = 0,5$. On s'intéresse aux probabilités a posteriori $P'(H1)$ et $P'(H2)$. Ici, $P'(H1)$ correspond à la probabilité conditionnelle $P(H1|E)$. On calcule alors $P'(H1) \approx 0,99$ à l'aide de la formule de Bayes et de même $P'(H2) = 1-P'(H1) \approx 0,01$. De manière informelle, l'expérience des segments de temps correspond ici à une situation où on cherche à connaître la durée de vie de l'objet unique présent dans l'urne. Il en résulte alors des variations analogues où on considère une situation concrète correspondant à la durée d'une course, la durée d'une vie humaine, la durée d'existence d'une association, etc.

Il est également intéressant de s'attacher à exprimer l'expérience en termes de n-univers purs. On considère un n-univers temporel qui comporte un objet unique en un lieu donné L_0 . L'objet est susceptible d'exister à différentes positions temporelles successives. Ici, on considère que le n-univers n'existe que pendant la durée de vie de l'objet unique qui s'y trouve. On a alors deux hypothèses:

(H1₁₁₉) le ΩT^* comprend 10 positions temporelles

(H2₁₂₀) le ΩT^* comprend 1000 positions temporelles



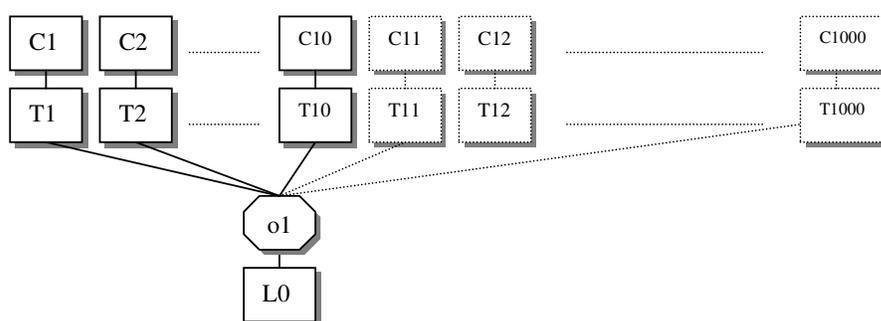
De manière informelle, cette variation de l'expérience des segments de temps correspond à une situation où l'observateur cherche à connaître une caractéristique de l'univers dans lequel il se trouve: je me trouve au temps T_1 dans un univers dont je sais qu'il est temporel, mais dont j'ignore le nombre exact de positions temporelles. Soit T_1-T_n la durée de vie du ΩT^* . Avant T_1 , le ΩT^* correspondant n'existe pas. Après T_n , le ΩT^* n'existe plus. Immégré en T_1 dans le ΩT^* , je cherche à connaître la durée de vie du n-univers dans lequel je me trouve. Par un raisonnement bayésien, je suis conduit à privilégier l'hypothèse selon laquelle l'univers contient peu de positions temporelles.

A ce stade, on peut s'interroger pour savoir si la modélisation de l'expérience des segments de temps dans le ΩT^* est complètement satisfaisante. En effet, considérons le modèle de l'urne correspondant. On a une boule unique, qui existe soit de T_1 à T_{10} , soit de T_1 à T_{1000} . L'expérience est basée sur le fait que l'événement E est le suivant: la boule existe en T_1 . Mais comment, à ce stade, sait-on que l'on est en T_1 ? Dans notre univers réel, un dispositif d'affichage⁸⁹ se révèle nécessaire, afin que l'observateur

⁸⁸ Ici, T_{10} et T_{1000} sont choisis de manière à rendre l'expérience compatible avec les données qui sont celles du *modèle des deux urnes*. On aurait pu bien sûr choisir des données correspondant exactement aux expériences d'Oliver et Korb, et Delahaye.

⁸⁹ Ou de manière plus générale, un dispositif permettant à l'observateur, à l'aide d'une perception sensorielle, de prendre connaissance de la position temporelle correspondante.

puisse connaître la position temporelle dans laquelle il se trouve. Il peut s'avérer ici nécessaire, afin de mieux restituer le contenu de l'expérience des segments de temps, de faire appel à un critère-variable supplémentaire, permettant de faire connaître à l'observateur la position temporelle correspondante. Ceci peut s'effectuer de plusieurs manières. On peut utiliser par exemple une boule numérotée, dont la numérotation change en fonction de la position temporelle: la boule possède le numéro 1 en T_1 , le numéro 2 en T_2 , ..., le numéro n en T_n . En termes de critères-variables, on peut faire appel à un critère-variable de couleur, correspondant par exemple à 1000 nuances de gris. Ainsi, chaque taxon de couleur correspond à un numéro de 1 à 1000. La boule présente alors la couleur 1 en T_1 , la couleur 2 en T_2 , ..., la couleur n en T_n . Compte tenu de ces éléments, on est amené à restituer l'expérience des segments de temps dans le ΩT^*C , un n -univers coloré à démultiplication temporelle. Une particularité de cet univers est que la couleur est indexée sur la position temporelle correspondante. Ici, le critère de couleur supplémentaire apparaît essentiellement comme un critère qui permet de procurer les taxons permettant d'afficher et de rendre visibles les positions temporelles.



8.3. Modélisation dans d'autres n -univers

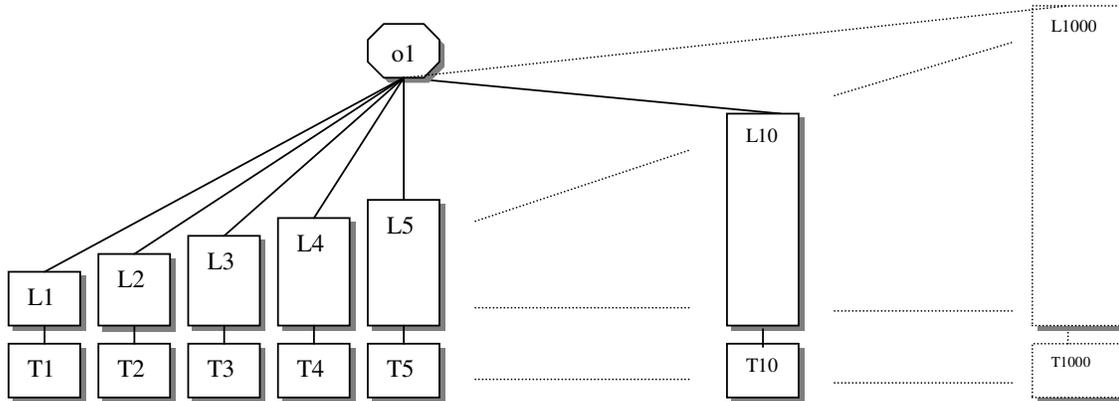
On peut s'interroger pour savoir si l'expérience des segments de temps ne peut pas également être modélisée dans un autre n -univers que le ΩT^*C . Intéressons-nous en premier lieu aux sous-univers du ΩT^*C . Pourrait-on ainsi par exemple modéliser l'expérience des segments de temps dans un sous-univers du ΩT^*C ? On peut tout d'abord se poser ainsi la question de savoir si l'on peut se dispenser du critère temporel. Ne peut-on ainsi transposer l'expérience des segments de temps dans un ΩC ? Dans un ΩC , l'objet unique possède une couleur donnée au temps T_0 et au lieu L_0 . Mais il apparaît qu'on ne peut pas alors restituer fidèlement une caractéristique de l'expérience qui est la persistance temporelle, qui constitue elle-même un cas particulier de démultiplication temporelle. Il faudrait pour cela recourir à un critère-variable de couleur à démultiplication, qui ne constitue pas une caractéristique de notre univers réel. Ainsi, il apparaît alors que l'expérience des segments de temps ne peut être modélisée dans un ΩC .

Peut-on de même restituer l'expérience des segments de temps dans un ΩT^* , un n -univers d'où le critère de couleur est absent? Mais on l'a vu plus haut, la couleur est utilisée comme dispositif d'affichage afin de faire connaître la position temporelle. Car en l'absence du critère de couleur, on ne peut plus rendre compte de l'événement E , à savoir du tirage d'un numéro de segment de temps. Ainsi, l'expérience des segments de temps ne peut pas non plus être modélisée dans le ΩT^* .

Il convient d'envisager en second lieu si l'expérience des segments de temps ne peut pas être rendue dans un n -univers présentant une structure identique à celle du ΩT^*C . Le ΩT^*C possède en effet la structure $\Omega\beta^*C$. Pourrait-on ainsi replacer l'expérience des segments de temps dans un ΩL^*C ? Cependant, il apparaît qu'on se heurte ici au fait qu'un tel n -univers est à démultiplication locale, c'est-à-dire qu'un objet peut s'y trouver en plusieurs endroits à un même moment T_0 . Et une telle propriété ubiquiste ne correspond pas aux caractéristiques de notre univers réel.

De manière alternative, l'expérience des segments de temps ne peut-elle être valablement restituée dans un ΩT^*L , qui présente également la structure $\Omega T^*\beta$. Considérons pour cela l'expérience suivante:

L'escalier Un escalier possède soit 10 soit 1000 marches. Et ces hypothèses sont a priori équiprobables. Je gravis la marche L_1 en T_1 , la marche L_2 en T_2 , ..., la marche L_n en T_n . Si je suis parvenu à la marche n , c'est donc que j'ai gravi les marches 1, 2, ..., $n-1$. Maintenant, je me trouve à la marche n° 5 de l'escalier. Quelle est maintenant la probabilité a posteriori que l'escalier possède seulement 10 marches?



Les caractéristiques du n-univers correspondant à une telle expérience sont les suivantes. Ici, je constitue, en tant qu'humain, l'objet unique dans le n-univers. De ce point de vue, l'expérience est tout à fait similaire à la version décrite par Oliver et Korb ou bien au paradoxe des bébés. En second lieu, le n-univers considéré comporte une variable temporelle ainsi qu'une variable spatiale qui correspond aux différentes marches. On a ainsi n positions temporelles T_1, T_2, \dots, T_n et n positions spatiales L_1, L_2, \dots, L_n . En outre, je peux exister à plusieurs positions distinctes. Le n-univers correspondant est donc à démultiplication temporelle. Par contre, je ne peux me trouver qu'à une position spatiale unique à un moment donné T_i . Ainsi, il apparaît qu'une telle expérience constitue une situation qui peut être modélisée dans un $\Omega T * L$. Ici, l'objet unique possède une position spatiale unique L_i à une position temporelle unique T_i . Ainsi, le numéro de position temporelle est indexé sur la position spatiale. On le voit, le fait de considérer ici des positions spatiales en plus du critère temporel à démultiplication permet de visualiser la position temporelle correspondante. Le critère spatial, indexé sur la position temporelle, joue ici un rôle tout à fait analogue à celui joué par la couleur dans le $\Omega T * C$. Les taxons de localisation jouent donc ici aussi le rôle du dispositif d'affichage de la position temporelle correspondante. On le voit, la modélisation de l'expérience des segments de temps dans le $\Omega T * L$ restitue parfaitement les conditions qui sont celles de l'expérience des segments de temps.

On peut noter de même que l'on obtient une variation analogue de l'expérience des segments de temps en considérant un autre n-univers qui présente la structure $\Omega T * \beta$, c'est-à-dire un $\Omega T * S$. Il s'agit d'un n-univers à objet unique comportant une variable temporelle et une variable sonore avec n notes de musique: S_1, S_2, \dots, S_n . Il en résulte la version suivante:

La harpe incrémentale Une harpe comporte soit 10 soit 1000 cordes. A priori, ces deux hypothèses sont équiprobables. La harpe produit des notes de musique dans les conditions suivantes: la note S_1 est jouée en T_1 , la note S_2 est jouée en T_2 , ..., la note S_n est jouée en T_n . Si la note de musique n a été jouée, c'est donc que les notes 1, 2, ..., $n-1$ ont été jouées préalablement. Maintenant, j'entends que la note S_1 est jouée sur la harpe. Quelle est maintenant la probabilité a posteriori que la harpe comprenne seulement 10 cordes?

On le voit, la variable sonore qui est indexée sur la position temporelle, joue ici de même que précédemment le rôle du dispositif qui permet de connaître la position temporelle en la rendant

audible. Ici également, la modélisation de l'expérience des segments de temps dans le ΩT^*S restitue tout à fait les conditions de l'expérience originale.

Les variations précédentes de l'expérience des segments de temps suggèrent maintenant que l'on peut construire une version de l'expérience qui prend place dans un n-univers à objets multiples. Il suffit de considérer que les objets multiples jouent le rôle du dispositif d'affichage qui était respectivement exercé dans le ΩT^*C , le ΩT^*L et le ΩT^*S par les critères de couleur, de localisation et de son. On devrait ainsi être à même de reconstruire l'expérience des segments de temps dans $\Omega\alpha^*T^*$ ou bien un $\Omega\alpha T^*$. Dans ce cas, c'est le nombre de boules qui jouera le rôle du dispositif d'affichage, en indiquant ainsi la position temporelle.

Il convient de distinguer ici le cas du $\Omega\alpha T^*$ ou du $\Omega\alpha^*T^*$. Considérons tout d'abord si l'expérience peut prendre place dans un $\Omega\alpha^*T^*$. Un tel n-univers comporte des objets multiples et présente une variable temporelle. De plus, les objets y sont en relation *plusieurs* avec le critère de temps. Ainsi, plusieurs objets peuvent se trouver à une même position temporelle. On a alors une urne dont le fonctionnement est le suivant: au temps T_1 la boule o_1 est placée dans l'urne, au temps T_2 la boule o_2 est ajoutée dans l'urne, ..., au temps T_n la boule o_n est ajoutée dans l'urne. Les hypothèses correspondantes sont alors les suivantes:

- (H1₁₂₁) l'urne contiendra 10 boules en T_{10}
- (H2₁₂₂) l'urne contiendra 1000 boules en T_{1000}

Soit également l'événement E: une boule se trouve dans l'urne en T_5 . Compte tenu de cela, je mets alors à jour mes probabilités initiales concernant le nombre de boules qui seront finalement contenues dans l'urne. Ainsi, il s'avère que l'expérience des segments de temps peut être valablement restituée dans un $\Omega\alpha^*T^*$.

D'autre part, on peut se poser la question de savoir si une version de l'expérience des segments de temps peut prendre place dans un $\Omega\alpha T^*$? Un tel n-univers qui comporte des objets multiples est à démultiplication par rapport au critère temporel, et les objets y sont en relation *un* avec le temps. Il s'ensuit qu'il existe un seul objet par position temporelle. Une seule boule se trouve donc à une position temporelle donnée. Mais cette dernière caractéristique interdit de modéliser l'expérience des segments de temps dans le $\Omega\alpha T^*$. En effet, on a vu précédemment qu'il était nécessaire de pouvoir disposer de plusieurs boules à une position temporelle donnée, afin de réaliser ainsi un dispositif d'affichage de la position temporelle. En l'absence d'une telle propriété, la position temporelle de la boule ne peut être affichée, et l'expérience des segments de temps ne peut être valablement restituée.

8.4. L'expérience des segments de temps appliquée à l'objet humanité

A ce stade, il convient de mentionner une conséquence importante de l'expérience des segments de temps. En effet, il apparaît que l'on peut construire une version de l'expérience des segments de temps appliquée à l'humanité. On considère ici l'humanité en tant qu'objet unique, prise dans son ensemble et sans considération du nombre de ses membres à un moment donné. Etant donné ce qui précède, il apparaît que l'on peut mettre en évidence une variation de l'expérience des segments de temps qui correspond à la situation de l'humanité. L'effet de cette dernière version se révèle sensiblement équivalent à celui de DA classique. On a vu en effet plus haut que l'expérience des segments de temps prenait place dans le ΩT^*C . Considérons donc la variante de l'expérience des segments de temps appliquée à l'objet humanité. La variation correspondante peut alors être décrite de la manière suivante:

L'expérience des segments de temps appliquée à l'objet humanité Considérons l'humanité prise dans son ensemble. L'humanité possède une durée de vie, comprise entre le moment de sa

naissance (l'apparition d'*homo sapiens*⁹⁰) et sa disparition. Considérons le nombre de millions d'années⁹¹ écoulés depuis l'apparition de l'humanité. On fait alors les deux hypothèses suivantes:

(H1₁₂₃) l'existence de l'humanité durera 10 millions d'années (*apocalypse proche*)

(H2₁₂₄) l'existence de l'humanité durera 1000 millions d'années (*apocalypse lointaine*)

Les probabilités initiales sont telles que $P(H1) = P(H2) = 0,5$. L'événement E consiste ici dans le fait que l'humanité se trouve dans le 3ème million d'années de son existence. Comment calcule-t-on alors les probabilités a posteriori $P'(H1)$ et $P'(H2)$? Etant donné cela, il s'ensuit ici, de la même manière que pour l'expérience des segments de temps classique, un décalage bayésien en faveur de l'hypothèse selon laquelle l'Apocalypse interviendra bientôt (*apocalypse proche*).

Et il s'ensuit une situation tout à fait analogue à celle qui prévaut pour un être humain unique. La variation de l'expérience des segments de temps appliquée à l'objet humanité prend place dans le ΩT^*C , un n-univers à objet unique, temporel et coloré, où l'objet unique peut se trouver à plusieurs temporelles différentes. Ici, la segmentation au niveau de l'existence de l'humanité s'établit au niveau du million d'années. Et on constate qu'une telle variation de l'expérience des segments de temps produit un effet tout à fait similaire à l'expérience classique des segments de temps, en s'appliquant cette fois à l'humanité prise dans son ensemble. On note ici qu'il convient de raisonner par rapport à l'humanité prise dans son ensemble, considérée en tant qu'entité unique. On considère ici l'existence de l'humanité dans le premier million d'années, puis dans le second million d'années, ..., puis dans le e -ième million d'années. Une telle variation se révèle intéressante, car intuitivement, son effet est similaire à celui de DA classique. Elle concerne l'humanité dans sa globalité, et existant à différentes positions temporelles. L'argument peut ainsi être mis en œuvre dans le premier million d'années d'existence de l'humanité, puis dans le second million d'années, ..., puis dans le e -ième million d'années d'existence de l'humanité. Remarquons dans cette variation que peu importe quel membre de l'espèce humaine met en œuvre l'argument. En effet, il suffit ici qu'il soit mis en œuvre, à l'époque considérée par un ou plusieurs humains. Il s'agit ici en quelque sorte d'une mise en œuvre collective de l'argument. Car on considère ici l'objet humanité pris dans son ensemble, et une segmentation du temps basée sur le million d'années. Une telle situation conduit à une augmentation de la probabilité initiale relative à l'hypothèse selon laquelle le nombre de segments de temps dans l'existence de l'humanité sera faible. Ce que signifie ainsi le décalage bayésien qui résulte de DA-temps, c'est que la probabilité selon laquelle l'humanité comprendra un faible nombre de segments de temps doit être revue à la hausse.

8.5. L'expérience des segments de temps en tant qu'objection à DA

L'expérience des segments de temps est apparue dans la littérature comme une objection à DA. La conséquence de l'expérience des segments de temps se révèle en effet contraire à l'intuition et a pu ainsi être soulevée en tant qu'objection à DA. Tel est en particulier le point soulevé par Oliver et Korb (1998) lorsqu'ils décrivent leur version de l'expérience des segments de temps.

On peut noter ici qu'une telle objection ne constitue pas en soi une réfutation de DA. En effet, il apparaît que la conclusion qui résulte de l'expérience des segments de temps est effectivement contraire à l'intuition, comme le soulignent Oliver et Korb. Cependant, tel est également le cas pour DA dans sa version classique. Mais le problème est précisément de savoir pourquoi DA est contraire à l'intuition. Bien sûr, l'expérience des segments de temps apparaît encore plus fortement contraire à l'intuition que DA et la variation appliquée aux bébés se révèle de surcroît paradoxale. Finalement, ce qui résulte de l'expérience des segments de temps en tant qu'objection à DA, c'est que cette dernière expérience constitue une conséquence contraire à l'intuition supplémentaire qui résulte de la structure

⁹⁰ On peut choisir ici, si on le souhaite, toute autre classe de référence, telle qu'une classe de référence correspondant au genre *homo*, à la sous-espèce *homo sapiens sapiens*, etc. Car on obtient alors une version équivalente.

⁹¹ On peut choisir, de manière équivalente, toute autre unité de temps que le million d'années.

du raisonnement qui sous-tend DA. Ceci, toutefois, ne résout pas DA, car il s'avère nécessaire de préciser pour quelle raison sa conclusion est contraire à l'intuition.

Il existe également une autre raison en vertu de laquelle l'expérience des segments de temps ne peut constituer une objection valable à DA. Cette raison est la suivante. On vient de le voir, une variation de l'expérience des segments de temps est constituée par l'expérience des segments de temps appliquée à l'humanité considérée en tant qu'objet unique. Une telle variation constitue l'expérience des segments de temps appliquée à l'objet humanité. Mais la conclusion de cette dernière expérience est de même nature que celle résultant de DA: il convient de réviser à la hausse une probabilité initiale selon laquelle l'extinction de l'humanité est proche. Ainsi, les conclusions qui résultent de l'expérience des segments de temps appliquée à l'objet humanité et de la version classique de DA présentent une nature convergente et se renforcent finalement. Par conséquent, loin de contrecarrer DA, l'expérience des segments de temps, dans sa variante que constitue l'expérience des segments de temps appliquée à l'objet humanité, renforce finalement la conclusion contraire à l'intuition de DA.

8.6. Analyse

Compte tenu de ce qui précède, il convient maintenant de s'attacher à déterminer pour quelle raison la conclusion qui résulte de l'expérience des segments de temps est contraire à l'intuition (et même paradoxale dans le cas du paradoxe des bébés). Le raisonnement qui a pour support l'expérience des segments de temps est basé sur une analogie avec le *modèle des deux urnes*. Et cette dernière analogie conduit à l'application du modèle à minoration. Intuitivement toutefois, l'application du modèle à minoration n'est pas valable dans le cas de l'expérience des segments de temps. Il convient donc de s'attacher à déterminer pourquoi l'analogie entre la situation concrète correspondant à l'expérience des segments de temps et le *modèle des deux urnes* n'est pas valable.

Attachons-nous pour cela à déterminer quel est le modèle probabiliste qui correspond véritablement à l'expérience des segments de temps. Considérons ainsi la modélisation de l'expérience des segments de temps dans le ΩT^* qui vient d'être décrite. Cette dernière prend place dans un n-univers à objet unique à démultiplication temporelle. Il en résulte, de manière directe, le modèle probabiliste suivant:

L'objet à incrémentation-temps (dans le ΩT^*) Une boule unique existe pendant une période de temps donnée. On formule ainsi les deux hypothèses suivantes:

(H1₁₂₅) la boule existe jusqu'en T_{10}

(H2₁₂₆) la boule existe jusqu'en T_{1000}

La boule possède une propriété de persistance temporelle: si elle existe en T_i , alors elle a existé en T_1, T_2, \dots, T_{i-1} . Les intervalles qui séparent les différentes positions temporelles sont égaux. Ici, l'événement E est le suivant: la boule existe en T_1 . On a également les probabilités a priori $P(H1) = P(H2) = 0,5$. On cherche à calculer les probabilités a posteriori $P'(H1) = P(H1|E)$ et $P'(H2) = P(H2|E)$.

On peut de même modéliser de manière plus saisissante la situation concrète correspondant à l'expérience des segments de temps dans un ΩT^*C . Le critère de couleur supplémentaire sert ici de dispositif de visualisation et permet de restituer la couleur de la boule unique. On se situe alors dans un n-univers à objet unique, coloré et à démultiplication temporelle. Ici, la boule possède une couleur, choisie parmi n taxons de couleur différents. On a ainsi la couleur C_1 , la couleur C_2 , ..., la couleur C_n . Pour simplifier, on considère ici que le fait que la boule présente la couleur C_m est rendu ici par le fait que la boule présente le n° m . Le modèle probabiliste correspondant est alors le suivant:

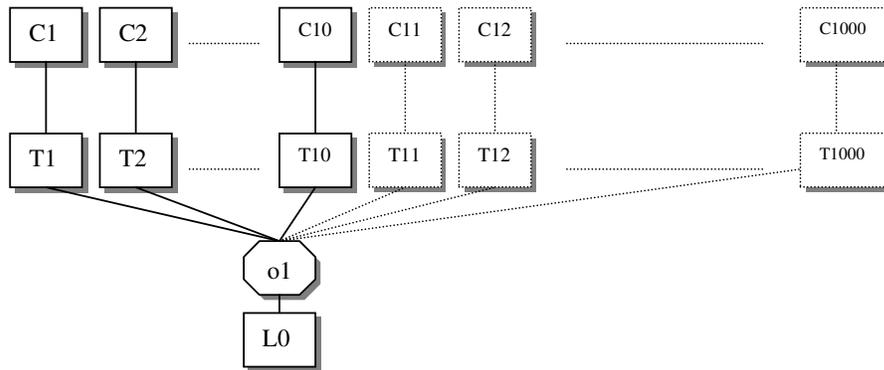
L'objet à incrémentation-temps (dans le ΩT^*C) Une boule unique existe pendant une période de temps donnée. La boule possède une propriété de persistance temporelle: si elle existe en T_i , alors elle a existé en T_1, T_2, \dots, T_{i-1} . Les intervalles qui séparent les différentes positions temporelles sont égaux. De plus, les C_1 - C_n taxons de couleur sont indexés sur la position

temporelle: la boule présente la couleur C_1 en T_1 , la couleur C_2 en T_2 , ..., la couleur C_n en T_n . On formule ainsi les deux hypothèses suivantes:

- (H1₁₂₇) la boule présentera les numéros 1 à 10
- (H2₁₂₈) la boule présentera les numéros 1 à 1000

Ici, l'événement E est le suivant: la boule présente le n° 1. On a également les probabilités a priori $P(H1) = P(H2) = 0,5$. On cherche à calculer les probabilités a posteriori $P'(H1) = P(H1|E)$ et $P'(H2) = P(H2|E)$.

Ici, la couleur de la boule sert de dispositif de visualisation et est assimilée à un numéro porté sur la boule. On considère donc de manière équivalente que la boule présente le numéro 1 en T_1 , le numéro 2 en T_2 , ..., le numéro n en T_n . L'expérience peut être restituée en considérant une boule sur laquelle s'affiche un numéro en fonction de la position temporelle correspondante.

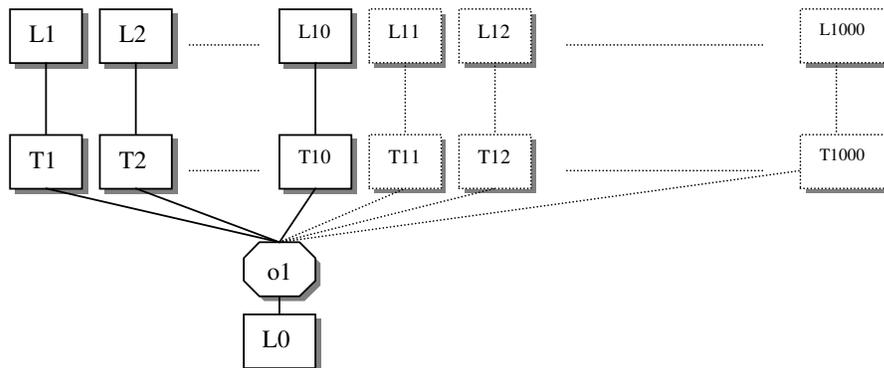


On a de même la variation qui est la suivante:

L'objet à incrémentation-espace (dans le $\Omega T * L$) Une boule unique existe pendant une période de temps donnée. La boule possède une propriété de persistance temporelle: si elle existe en T_i , alors elle a existé en T_1, T_2, \dots, T_{i-1} . Les intervalles qui séparent les différentes positions temporelles sont égaux. De plus, la boule unique occupe successivement les L_1-L_n taxons de localisation qui sont indexés sur la position temporelle: la boule se trouve placée en L_1 en T_1 , en L_2 en T_2 , ..., en L_n en T_n . On formule ainsi les deux hypothèses suivantes:

- (H1₁₂₉) la boule occupera les positions spatiales de 1 à 10
- (H2₁₃₀) la boule occupera les positions spatiales de 1 à 1000

Ici, l'événement E est le suivant: la boule se trouve à la position spatiale n° 1. On a également les probabilités a priori $P(H1) = P(H2) = 0,5$. On cherche à calculer les probabilités a posteriori $P'(H1) = P(H1|E)$ et $P'(H2) = P(H2|E)$.



Dans cette dernière expérience, on le voit, la position spatiale de la boule sert de dispositif de visualisation de la position temporelle correspondante.

A ce stade, on peut analyser de manière détaillée le modèle probabiliste correspondant à l'expérience des segments de temps. Il apparaît en effet de manière nette, dans ce type de situation que l'événement E consiste dans le fait que la boule présente le numéro 1 qui correspond à la position temporelle T_1 lors du tirage. De manière générale, il s'avère que la boule tirée présente le n° i en T_i . Le numéro de la boule n'est donc pas choisi au hasard, mais se révèle *indexé* sur la position temporelle correspondante. Il apparaît ainsi qu'un tel tirage s'effectue ici en violation de HR. Le fait que l'on observe dans l'expérience des segments de temps une violation manifeste de HR entraîne une conséquence importante. Il s'ensuit en effet que l'analogie existant entre l'expérience des segments de temps et le *modèle des deux urnes* n'est pas fondée. Et par conséquent, le raisonnement qui conduit à un décalage bayésien en faveur de l'hypothèse (H1) en vertu du modèle à minoration, n'apparaît désormais plus valide.

Quel est donc le raisonnement qui doit être valablement tenu dans cette situation, en remplacement de celui basé sur l'application du modèle à minoration? Compte tenu ici du fait que le numéro 1 de la boule est celui de la position temporelle, il s'ensuit que la probabilité que la boule comporte le numéro 1 en T_1 si l'urne contient 10 boules, est égale à 1. Et de même, la probabilité que la boule comporte le numéro 1 en T_1 si l'urne contient 1000 boules, est égale à 1. On a ainsi $P(E|H1) = P(E|H2) = 1$. Une telle situation apparaît ainsi fondamentalement différente de celle du *modèle des deux urnes*, où $P(E|H1) = 1/10$ et $P(E|H2) = 1/1000$. Dans le modèle de *l'objet à incrémentation-temps*, les probabilités que la boule présente le numéro 1 en T_1 si l'urne contient 10 ou 1000 boules sont donc identiques et égales à 1. Et ceci ne permet pas une quelconque modification des probabilités initiales.

On peut observer ici que le modèle de *l'objet à incrémentation-temps* présente une première propriété qui est la suivante: l'expérience se déroule sur une période qui s'étend de T_1 à T_n . Cette période est strictement délimitée et aucun événement particulier en rapport avec l'expérience qui se déroulerait préalablement à T_1 c'est-à-dire en T_0 , n'est mentionnée ici. Ainsi, la période de temps sur laquelle la totalité de l'expérience prend place se trouve bien circonscrite à T_1 - T_n .

A ce stade, une autre particularité intéressante du modèle de *l'objet à incrémentation-temps* mérite également d'être détaillée. Il s'agit du fait que l'expérience correspondante est susceptible d'être itérée, sans que sa nature ne s'en trouve affectée ou modifiée. On peut donc considérer une variation où la situation correspondante à l'expérience de *l'objet à incrémentation-temps* est itérée. Considérons pour cela la situation suivante:

L'objet à incrémentation-temps itéré (dans le ΩT^*C) Une boule unique existe pendant une période de temps donnée. La boule possède une propriété de persistance temporelle: si elle existe en T_i , alors elle a existé en T_1, T_2, \dots, T_{i-1} . Les intervalles qui séparent les différentes positions temporelles sont égaux. De plus, les C_1 - C_n taxons de couleur sont indexés sur la position temporelle: la boule présente la couleur C_1 en T_1 , la couleur C_2 en T_2 , ..., la couleur C_n en T_n . On formule les deux hypothèses suivantes:

- (H1₁₃₁) la boule présentera les numéros de 1 à 10
- (H2₁₃₂) la boule présentera les numéros de 1 à 1000

Ici, les événements successifs E_1, E_2, \dots, E_i ($1 \leq i \leq 10$) sont respectivement les suivants: la boule possède le n° 1 en T_1 , le n° 2 en T_2 , ..., le n° i en T_i . On a également les probabilités a priori $P(H1) = P(H2) = 0,5$. On cherche à calculer les probabilités a posteriori $P(H1|E_1)$ et $P(H2|E_1)$, $P(H1|E_2)$ et $P(H2|E_2)$, ..., $P(H1|E_i)$ et $P(H2|E_i)$.

Comme précédemment, la boule présente ici le numéro i en T_i . Dans une telle situation, on peut prévoir avec précision le prochain numéro que comportera la boule. Ainsi, si l'on considère l'événement E_1 , on se trouve replacé dans la situation qui est celle du modèle de *l'objet à incrémentation-temps* (non itéré). De même, si on se place dans le cas de l'événement E_2 , il apparaît alors que l'on rencontre une situation tout à fait analogue. La seule différence est que les hypothèses peuvent être ainsi formulées:

- (H1₁₃₃) la boule présentera les numéros de 2 à 10
- (H2₁₃₄) la boule présentera les numéros de 2 à 1000

En effet, l'hypothèse où la boule présente le numéro 1 peut désormais être écartée, car la position temporelle T_1 est écoulée. Si l'on considère maintenant que la boule possédera le n° 2 en T_2 , le n° 3 en T_3 , ..., le n° i en T_i ($1 \leq i \leq 10$), il apparaît à ce stade que l'expérience correspondante est structurellement analogue à celle du modèle de *l'objet à incrémentation-temps* (non itéré) dont les hypothèses seraient les suivantes:

- (H1₁₃₅) la boule présentera les numéros de 1 à 9
- (H2₁₃₆) la boule présentera les numéros de 1 à 9

Ceci montre finalement que la situation qui est celle de l'expérience itérée de *l'objet à incrémentation-temps* se compose finalement d'une série d'expériences successives de *l'objet à incrémentation-temps*.

On le voit, ces propriétés (temporalité limitée à T_1 - T_n et itérabilité) du modèle de l'expérience de *l'objet à incrémentation-temps* sont présentées à la fois par la version de l'expérience des segments de temps décrite par Korb et Oliver, mais aussi par la variation présentée par Delahaye.

Ce qui précède permet également de clarifier une différence existant entre la version de l'expérience des segments de temps de Korb et Oliver et celle de Delahaye. Rappelons, que dans l'expérience de Oliver et Korb, l'événement E est le suivant: étant donné que la minute 1 de mon existence est celle où j'ai été au courant pour la première fois de DA, je me trouve actuellement à la minute m (avec $m \geq 1$). En revanche, dans la version de Delahaye, l'événement E est: je suis à la minute 1 de mon existence. Le modèle de *l'objet à incrémentation-temps* s'applique donc de manière adéquate à la version décrite par Delahaye. En ce qui concerne la variation de Korb et Oliver, on peut considérer, pour fixer les idées, un cas où $m > 1$ tel que par exemple $m = 10000$. Dans ce cas, il résulte de l'expérience itérée de *l'objet à incrémentation-temps* qu'une telle version se réduit également au modèle de *l'objet à incrémentation-temps*. Car au début de l'expérience ($m = 1$), les hypothèses sont les suivantes:

- (H1₁₃₇) le nombre de minutes restantes de mon existence sera 4000000
- (H2₁₃₈) le nombre de minutes restantes de mon existence sera 40000000

Cependant, lorsque $m = 10000$, les hypothèses sont les suivantes:

- (H1₁₃₉) le nombre de minutes restantes de mon existence sera 3999000
- (H2₁₄₀) le nombre de minutes restantes de mon existence sera 39900000

Ceci permet ainsi de conclure que l'expérience des segments de temps décrite par Korb et Oliver est bien structurellement identique à celle décrite par Delahaye.

Finalement, il résulte de ce qui précède que l'expérience des segments de temps n'est pas isomorphe⁹² au modèle des deux urnes, mais bien au modèle de *l'objet à incrémentation-temps*. Dans ce contexte, le raisonnement qui doit être valablement tenu n'est pas fondé sur l'analogie entre la situation de l'expérience des segments de temps et le modèle des deux urnes, mais bien sur l'analogie entre la situation de l'expérience des segments de temps et le modèle de *l'objet à incrémentation-temps*. Et l'application de ce dernier modèle conduit à la conclusion que les probabilités a priori demeurent inchangées, une conclusion qui se révèle cette fois tout à fait conforme à l'intuition.

A ce stade, il apparaît également qu'une telle analyse vaut également pour la variation que constitue l'expérience des segments de temps appliquée à l'objet humanité. Une telle situation est en effet en tous points similaire à l'expérience des segments de temps appliquée à un être humain donné. Ainsi, la situation correspondant à l'expérience des segments de temps appliquée à l'objet humanité présente-t-elle une propriété d'itérabilité. Car l'argument peut être mis en œuvre dans le premier million d'années d'existence de l'humanité, puis dans le second million d'années, ..., puis dans le e -ième million

⁹² Selon la terminologie utilisée à propos de DA par Timothy Chambers (2001). Elle peut de fait être étendue à tout argument basé sur une analogie avec un modèle probabiliste.

d'années d'existence de l'humanité. En conséquence, le critère d'itérabilité de l'objet à incrémentation-temps se trouve bien vérifié au niveau de la situation correspondant à l'expérience des segments de temps appliquée à l'objet humanité. Ainsi, dans la situation correspondant à l'expérience des segments de temps appliquée à l'objet humanité, il en va de même que pour l'expérience des segments de temps: on ne peut raisonner valablement en appliquant le modèle à minoration, car la situation est isomorphe au modèle de l'objet à incrémentation-temps. Dans le cas de l'expérience des segments de temps appliquée à l'objet humanité, il ne s'ensuit donc aucun décalage bayésien concernant les probabilités initiales, car la situation est isomorphe au modèle de l'objet à incrémentation-temps, où les probabilités initiales demeurent inchangées. Il apparaît en outre qu'il existe une position temporelle (T_0) dans la situation correspondant à l'expérience des segments de temps appliquée à l'objet humanité mais non dans l'objet à incrémentation-temps. Cependant, aucun événement susceptible d'influer sur la situation se déroulant de T_1 à T_n n'apparaît susceptible de s'y dérouler. Par conséquent, on n'observe en T_0 aucun événement substantiel. En conclusion, la situation correspondant à l'expérience des segments de temps appliquée à l'objet humanité vérifie bien les critères qui sont ceux de l'*objet à incrémentation-temps*.

Chapitre 9. Analyse de l'Argument de l'Apocalypse

9.1. La version bi-dimensionnelle de DA

Avant de procéder à une analyse de DA, il convient d'analyser les conséquences de l'existence de la variation de l'expérience des segments de temps exposée plus haut et que nous avons dénommée l'expérience des segments de temps appliquée à l'objet humanité. La structure de cette dernière variation, on l'a vu, est isomorphe à l'objet à incrémentation-temps. Mais l'existence de l'expérience des segments de temps appliquée à l'objet humanité suggère également qu'il est possible de mettre en évidence une variation de DA appliquée au critère temporel. Considérons en effet la variation suivante:

DA-temps à objets multiples On considère les deux hypothèses suivantes:

(H1₁₄₁) l'existence de l'humanité durera 10 millions d'années (*apocalypse proche*)

(H2₁₄₂) l'existence de l'humanité durera 1000 millions d'années (*apocalypse lointaine*)

Les probabilités initiales sont telles que $P(H1) = P(H2) = 0,5$. L'événement E consiste ici dans le fait que je me trouve dans le 3ème million d'années d'existence de l'humanité. Comment calcule-t-on alors les probabilités a posteriori $P'(H1)$ et $P'(H2)$?

Intuitivement en effet, *DA-temps à objets multiples* entraîne une révision à la hausse de l'hypothèse selon laquelle l'humanité connaîtra une extinction proche. Une telle conclusion apparaît très proche de celle qui résulte de la version classique de DA. On le voit, *DA-temps à objets multiples* prend place dans un n-univers à objets multiples, temporel et coloré, où les objets se trouvent en relation *plusieurs* avec le critère temporel, soit un $\Omega\alpha^*T^*C$.

Une conséquence immédiate de la mise en évidence de *DA-temps à objets multiples* est qu'il en résulte véritablement une version double, bi-dimensionnelle de DA. Cette version double de DA repose à la fois sur DA classique et sur DA-temps à objets multiples. Les deux variations prennent place dans le $\Omega\alpha^*T^*C$. Une telle version bi-dimensionnelle de DA conduit à une révision à la hausse d'une part (a) de l'hypothèse selon laquelle le nombre d'humains ayant jamais existé sera faible, et d'autre part (b) de l'hypothèse en vertu de laquelle le nombre de segments de temps dans l'existence de l'humanité sera faible. Elle repose en effet en premier lieu (a) sur DA classique (DA-objets) qui conduit à une révision à la hausse de la probabilité selon laquelle le nombre d'humains ayant jamais existé sera faible. Elle concerne les humains en général. Le raisonnement est basé ici sur le rang de naissance de chaque être humain, lui-même membre de la classe de référence. Et la version bi-dimensionnelle de DA repose en second lieu (b) sur DA-temps à objets multiples. Cette dernière version concerne l'humanité dans sa globalité, envisagée en tant qu'objet unique, et existant à différentes positions temporelles. Elle conduit à une révision à la hausse de la probabilité selon laquelle le nombre de segments de temps dans l'existence de l'humanité sera faible.

Ainsi, la version bi-dimensionnelle de DA résulte de la conjonction de deux variations de DA: DA-objets et DA-temps à objets multiples. Les deux versions conduisent à des conclusions convergentes, car elles sont toutes deux basées sur le modèle à minoration. Elles conduisent ainsi toutes deux à une conclusion de même nature qui consiste en un double décalage bayésien. Ces deux versions se renforcent mutuellement, de sorte que leurs conclusions convergentes en retirent une force supplémentaire. Compte tenu du fait qu'elle inclut à la fois DA-objets et DA-temps à objets multiples, la version bi-dimensionnelle de DA se révèle plus forte que la version classique de DA, dont la portée se limite à DA-objets.

A cet égard, on peut remarquer ici qu'au sens strict, la version classique de DA (DA-objets) ne permet pas de faire des hypothèses sur l'époque de l'extinction de l'humanité. La version classique de DA ne permet légitimement de faire des hypothèses que sur le nombre total d'humains ayant jamais existé. Dans la présentation classique de l'argument, on en déduit habituellement une date

correspondant à l'extinction de l'humanité⁹³. Mais cette conclusion va toutefois au-delà de la portée de la version classique de DA. Rappelons en effet les hypothèses qui correspondent précisément à la situation qui est celle de DA:

(H1₁₄₃) l'humanité atteindra au total 100 milliards d'humains

(H2₁₄₄) l'humanité atteindra au total 10000 milliards d'humains

Ces deux hypothèses portent, de manière stricte, sur le nombre d'humains ayant composé l'humanité. Mais on ne peut pas en déduire une conclusion relative à la durée d'existence de l'humanité considérée dans son ensemble. Intuitivement, le nombre global d'humains et la durée d'existence de l'humanité apparaissent grossièrement corrélées. Cependant, une telle corrélation ne s'ensuit pas nécessairement. Car on peut imaginer plusieurs hypothèses dans lesquelles l'hypothèse (H1) est vérifiée mais où une extinction précoce de l'humanité ne s'ensuit pas. Ceci peut correspondre par exemple à une situation où les derniers humains deviennent quasiment immortels. Dans ce cas, DA s'interprète comme le fait que l'humanité découvre finalement le secret d'une quasi immortalité. Une autre situation, plus classique, correspond au fait que le nombre d'humains vivant à un moment donné pourrait devenir très faible. Dans ce cas, l'humanité pourrait très bien ne compter que très peu d'individus sur une période de temps extrêmement longue. L'hypothèse (H1) serait alors vérifiée sans qu'il s'ensuive nécessairement une extinction prématurée de l'humanité. On peut noter qu'une interprétation de cette nature est notamment mentionnée par Bostrom. Bostrom⁹⁴ (2000) souligne en effet que DA peut être interprété comme le fait que le nombre d'humains vivant sur Terre va diminuer très fortement dans le futur. Et ce fait se révèle tout à fait compatible avec le fait que l'humanité survivra sur une période très longue.

Parallèlement, DA-temps à objets multiples ne doit pas non plus être interprété de manière erronée. En effet, DA-temps autorise seulement à réviser à la hausse une probabilité initiale selon laquelle le nombre de millions d'années dans l'existence de l'humanité sera faible. Mais DA-temps à objets multiples ne permet pas de conclure que le nombre d'humains ayant jamais existé sera faible plutôt qu'important. Rappelons en effet les hypothèses précises sur lesquelles DA-temps est basé:

(H1₁₄₅) l'existence de l'humanité durera 10 millions d'années (*apocalypse prochaine*)

(H2₁₄₆) l'existence de l'humanité durera 1000 millions d'années (*apocalypse lointaine*)

Ce que signifie ainsi le décalage bayésien qui résulte de DA-temps à objets multiples, c'est seulement que la probabilité selon laquelle l'humanité comprendra un faible nombre de segments de temps doit être revue à la hausse. Il ne s'ensuit pas nécessairement que le nombre d'humains ayant jamais existé se révélera finalement faible. En effet, une disparition plus rapide que prévue de l'humanité pourrait bien se produire sans qu'il s'ensuive nécessairement que le nombre d'humains ayant jamais existé se révèle faible. Car on pourrait bien observer une progression exponentielle des humains dans les dernières années⁹⁵. Ceci pourrait par exemple être dû au fait que la durée de vie des humains serait devenue beaucoup plus courte au fil des siècles. A la suite d'importants changements génétiques intervenus au niveau de l'espèce humaine, la durée de vie de chaque humain serait ainsi devenue très brève. Dans cette hypothèse, on observera bien une durée de vie réduite de l'humanité, mais en revanche un nombre extrêmement élevé d'humains. Dans ce cas, on aura bien un nombre de segments de temps faible dans l'existence de l'humanité, mais un nombre d'humains ayant composé l'humanité

⁹³ Ou de manière plus restrictive de la classe de référence correspondante.

⁹⁴ Cf. p. xxx: "It should be pointed out that even if DA is basically correct, there is still vast scope for alternative interpretations of the result other than that humankind is likely to go extinct soon. For example, one may think that: (...) The size of the human population will decrease in the future; this reconciles DA with even extremely long survival of the human species". Trad.: "Il convient de remarquer que même si DA est fondamentalement correct, il y a encore une large place pour des interprétations alternatives du résultat autres que le fait que l'humanité est susceptible de connaître une prochaine extinction. Par exemple, on peut penser que: (...) la taille de la population humaine diminuera à l'avenir; ceci réconcilie DA avec une survie extrêmement longue de l'espèce humaine".

⁹⁵ Ce scénario est celui de l'expérience de la *shooting-room*, qui sera étudiée au chapitre suivant.

corrélativement très élevé. Ainsi, de même que pour DA-objets, on ne doit pas interpréter DA-temps à objets multiples au-delà de ce qui se révèle strictement contenu dans les hypothèses.

A ce stade, on peut maintenant s'interroger sur les conséquences de l'existence d'une version bi-dimensionnelle de DA. Quel est donc l'effet de la version bi-dimensionnelle de DA par rapport à la seule version classique de DA? Ici, il faut imaginer que l'humanité comporte deux dimensions: une dimension *objets* (ou composante) qui correspond au nombre d'humains qui font partie de l'humanité; et une seconde dimension *temporelle*, qui traduit le fait que l'humanité possède une période de naissance, une période d'existence et une position temporelle correspondant à son extinction. Ces deux dimensions objets et temporelle correspondent aux critères-variables qui sont ceux du n -univers dans lequel se trouve modélisée la situation correspondante: le $\Omega\alpha * T * C$. Car la couleur sert ici de critère de visualisation. En particulier, dans DA-objets, la couleur, on l'a vu, sert de critère de visualisation du rang de naissance. Au niveau de DA-temps à objets multiples, la couleur permet la visualisation de la position temporelle correspondante. En termes de n -univers, la version bi-dimensionnelle de DA porte à la fois sur le nombre d'objets et le critère temporel. La version bi-dimensionnelle de DA porte à la fois sur la dimension *objets* et la dimension *temporelle*. La version de DA qui résulte de la conjonction de ces deux versions est plus forte que la seule version classique de DA, car cette dernière ne porte que sur une seule dimension de l'humanité: la dimension *objets* liée au nombre d'humains composant l'humanité. La version bi-dimensionnelle de DA est également plus forte car elle permet d'écarter certaines interprétations exclusivement liées à DA-objets, qui ont été mentionnées plus haut et basées sur un scénario où (a) les derniers humains deviendront immortels ou bien (b) où l'humanité vivra très longtemps avec un petit nombre d'humains, etc. La version temporelle de DA permet d'écarter ces interprétations liées à DA-objets. Mais cette version bi-dimensionnelle permet également d'éliminer les interprétations exclusivement liées à DA-temps, qui viennent d'être mentionnées, fondées sur une progression exponentielle des humains. De telles interprétations sont à écarter car elles sont rendues impossibles à cause de la présence conjointe de la version classique de DA et de DA-temps à objets multiples.

A ce stade, on peut s'interroger pour savoir s'il n'existe pas une réduction de DA-temps (à objets multiples) à DA-objets. Les deux versions de DA qui composent la version bi-dimensionnelle de DA ne se réduisent-elles pas à une seule et même version? Considérons donc DA-temps à objets multiples. A chaque segment de temps est associé un intervalle $p-q$ qui correspond au nombre d'humains ayant existé jusqu'en p ou en q . Leslie (1996, p. 202) mentionne d'ailleurs dans son analyse du problème de la *définition de la classe de référence* une telle simplification. Il considère des étiquettes associées à intervalles de nombres:

The point survives even when one's mathematics becomes covered with approximation signs, 'at least' signs, signs standing for ranges of figures rather than single figures, and so on. Compare how the report that your name has come out of a lottery urn 'in something like the tenth draw', or 'between the sixth and the sixteenth draw', can act powerfully against the suggestion that there are 'at least a thousand other names' still waiting to be drawn, or 'from one to two thousand other names'. (...) the doomsday argument doesn't demand accurate figures. Approximations and ranges can fit its purposes admirably.⁹⁶

et conclut que le décalage bayésien inhérent à DA s'applique de manière identique dans une telle situation. Cette remarque de Leslie conduit ainsi à la version suivante de DA: considérons, de manière

⁹⁶ Trad.: Ce point particulier survit de même lorsqu'on utilise en guise de mathématiques des signes correspondant à des approximations, 'au moins' des signes, c'est-à-dire des signes représentant des plages de nombres plutôt que de simples nombres, et ainsi de suite. Voyez comment l'information selon laquelle votre nom est sorti d'une urne de loterie 'approximativement au dixième coup', ou 'entre le sixième et le seizième coup', peut agir puissamment contre l'idée qu'il y a 'au moins mille autres noms attendant toujours pour être tirés, ou bien 'un à deux mille autres noms'. (...) L'argument de l'apocalypse n'exige pas des nombres précis. Les approximations et les plages de nombres peuvent servir ses buts de manière admirable.

simplifiée l'objet humanité, correspondant par exemple à l'espèce *homo sapiens*. Intéressons-nous seulement au nombre de milliards d'individus composant l'humanité. On s'intéresse au nombre de milliards d'*homo sapiens* composant de l'objet humanité et on raisonne ainsi: actuellement je suis dans le 50ème milliard. Par un raisonnement bayésien, je conclus à un décalage en faveur de l'hypothèse selon laquelle l'humanité ne comptera que 60 milliards d'âmes de préférence à l'hypothèse alternative selon laquelle elle en comptera finalement 1000 milliards.

Mais rien n'oblige toutefois à utiliser des intervalles égaux. Ne peut-on ici choisir des intervalles différents? Et dès lors que l'on choisit des intervalles différents, on peut utiliser les intervalles qui correspondent précisément à chaque segment de temps. Ici, il apparaît que l'on n'a pas toutefois une réduction directe de DA-objets classique à DA-temps. On a toutefois une corrélation grossière, qui est due au fait que le nombre de segments de temps et le nombre d'*Homo sapiens* sont grossièrement corrélés (puisque grosso modo, le nombre d'humains augmente lorsque le numéro de segment de temps augmente). Cependant, une telle corrélation ne se produisant pas nécessairement, on ne peut donc considérer que l'on a une réduction de DA-objets classique à DA-temps à objets multiples.

9.2. Le problème de la classe de référence au niveau de DA-objets et de DA-temps

Le problème de la définition de la classe de référence dans DA a été exposé plus haut. Ce problème a été étudié sous l'angle de la définition de la classe de référence par restriction ou par extension. Ce dernier aspect a ainsi été envisagé par rapport à la version classique de DA, c'est-à-dire DA-objets. Cependant, compte tenu de la version bi-dimensionnelle de DA qui vient d'être décrite, il convient également, dans un souci d'exhaustivité, de s'intéresser au problème de la définition de la classe de référence au niveau de DA-temps à objets multiples.

Au niveau de DA-objets, le problème de la définition de la classe de référence consiste dans la difficulté de définir de manière appropriée les humains qui constituent la classe de référence. Le problème de définir avec précision les individus qui composent l'humanité se trouve ainsi posé. En l'absence d'un critère objectif, comment définir la classe de référence par restriction ou par extension, d'une manière qui ne soit pas arbitraire.

Intéressons-nous maintenant à DA-temps à objets multiples. DA-temps à objets multiples prend place dans un n-univers à objets multiples, temporel et coloré, où les objets se trouvent en relation *plusieurs* avec le critère temporel, un $\Omega\alpha^*T^*C$. La durée de vie de l'humanité est divisée en 10 ou 1000 segments de temps. Maintenant, étant donné que ma position actuelle se trouve dans le 3ème segment de temps, je conclus alors à un décalage bayésien en faveur de l'hypothèse selon laquelle l'existence de l'humanité durera 10 millions d'années (*apocalypse prochaine*). Dans cette dernière version, il apparaît que le problème de la définition de la classe de référence se pose également. Car doit-on assimiler la classe de référence dont on dénombre les segments de temps au genre *homo*, à l'espèce *homo sapiens*, à la sous-espèce *homo sapiens sapiens*, etc.? Ici aussi, la classe de référence est la même que pour DA-objets, et sa définition est ambiguë. Que l'on s'intéresse au nombre d'humains qui composent la classe de référence ou bien à la durée de vie de cette dernière, on rencontre le même problème au niveau de la définition de la classe de référence. Ainsi, ce dernier problème se pose à la fois pour DA-objets et DA-temps à objets multiples. En conséquence, le problème de la définition de la classe de référence apparaît bien comme un problème qui concerne directement la version bi-dimensionnelle de DA, car il en concerne les deux dimensions: les objets et le temps.

Cependant, il apparaît que le problème de la définition de la classe de référence dans DA, que ce soit au niveau de DA-objets ou de DA-temps à objets multiples, peut être neutralisé en considérant une classe de référence définie avec précision, en tolérant toutefois un certain degré d'arbitraire. Dans ce cas, on s'intéresse exclusivement au sort d'une certaine classe de référence telle que *homo sapiens sapiens*. On ne se préoccupe pas alors du sort des autres sous-espèces, espèces ou genres, etc. qui ont précédé ou qui succéderont à *homo sapiens sapiens*. Dans ce cas, il en résulte une version de DA moins forte, car dépourvue de la terreur associée à sa version originale, mais qui produit toujours une conclusion contraire à l'intuition. Ainsi, il apparaît ici que quelle que soit la classe de référence choisie, avec un certain degré d'arbitraire, on observe toujours le décalage bayésien caractéristique de la problématique inhérente à DA. DA est donc toujours présent, même si c'est toutefois sous une forme plus faible. Ainsi, la réponse apportée au problème de la définition de la classe de référence dans DA est-elle la suivante: le décalage bayésien contraire à l'intuition se produit quelle que soit la classe de

référence choisie. Par conséquent, ceci autorise à choisir, avec un certain degré d'arbitraire, l'une ou l'autre classe de référence.

Il existe toutefois des situations proches de DA au niveau desquelles n'apparaît pas le problème lié à la définition de la classe de référence. Considérons pour cela les variations de l'expérience des segments de temps décrites par Oliver et Korb et par Delahaye. La variante de l'expérience des segments de temps décrite par Oliver et Korb met en scène un seul être humain. Dans ce cas, la variation correspondante de l'expérience des segments de temps est 1-dimensionnelle et s'applique exclusivement au critère temporel. Ici, la classe de référence unique est composée des segments de temps composant l'existence de l'être humain unique concerné par l'expérience. Dans ce contexte, il apparaît que la définition de la classe de référence ne souffre pas d'ambiguïté. La notion de durée d'existence d'un être humain unique peut être définie avec précision. Ainsi, dans cette dernière hypothèse, il apparaît que le problème de la définition de la classe de référence a disparu. De manière identique, lorsqu'on considère les données qui sont celles du paradoxe des bébés, on n'a pas d'ambiguïté pour définir la classe de référence, composée de l'ensemble des segments de temps composant l'existence du bébé. Ainsi de manière générale, lorsque l'on considère l'existence d'un être *unique* comme c'est le cas pour la version décrite par Oliver et Korb ou le paradoxe des bébés, la définition de la classe de référence s'effectue de manière précise et non ambiguë. Dans ces situations, le problème de la définition de la classe de référence n'émerge pas. Pourtant, on observe bien dans ces versions un problème qui est celui d'une utilisation contraire à l'intuition du modèle à minoration. On peut noter ainsi qu'il existe des versions proches de DA basées sur la structure de l'expérience des segments de temps, pour lesquelles le problème de la définition de la classe de référence ne se pose pas. En effet, dans ces variations, la classe de référence est définie dans l'énoncé de manière non ambiguë et le choix de cette dernière s'effectue d'une manière qui n'est pas arbitraire.

Cependant, on peut observer à ce stade que le problème de la définition de la classe de référence se manifeste également au niveau de la variation de l'expérience des segments de temps appliquée à l'objet humanité. Dans cette dernière variation, on constate bien en effet l'émergence du problème de la définition de la classe de référence. Un tel problème se manifeste au niveau de la définition précise des membres qui composent l'objet unique humanité.

Ce qui précède conduit finalement à la réponse suivante au problème de la définition de la classe de référence dans DA. Il apparaît ainsi, en ce qui concerne la version bi-dimensionnelle de DA, on peut tolérer un certain niveau d'arbitraire au niveau du choix de la classe de référence, car l'effet contraire à l'intuition de l'argument se manifeste quelle que soit la classe de référence choisie, parmi les formulations "moyennes" de cette dernière, c'est-à-dire celles qui évitent les choix trop extrêmes par restriction ou par extension⁹⁷. Un choix de la classe de référence assimilée par exemple à la sous-espèce *homo sapiens sapiens*, ou bien à l'espèce *homo sapiens*⁹⁸, satisfait un tel critère de choix. Dans cette hypothèse, il en résulte une version plus faible de DA, car dépourvue de la terreur qui lui était initialement associée.

9.3. L'urne à incrémentation

Il apparaît ici opportun d'introduire deux modèles probabilistes, correspondant à deux variations différentes, qui prennent place dans un $\Omega\alpha T^*C$: l'urne à incrémentation-temps et l'urne à incrémentation-objet. Ces deux modèles trouveront un rôle important à jouer dans la discussion qui va suivre. Considérons tout d'abord la situation suivante:

L'urne à incrémentation-temps Des boules numérotées sont expulsées d'un dispositif, à des intervalles de temps réguliers, puis placées dans une urne. Les boules sont numérotées ainsi selon leur ordre d'expulsion du dispositif. Ainsi, l'urne contient 1 boule au temps T_1 , 2 boules au temps T_2 , ..., n boules au temps T_n . Ici, l'intervalle qui sépare deux positions temporelles est

⁹⁷ Il s'agit essentiellement de la solution au problème de la classe de référence retenue par Leslie (1996).

⁹⁸ Toute classe de référence intermédiaire, c'est-à-dire comprise entre *homo sapiens sapiens* et *homo sapiens* convient également.

constant, de sorte que $T_{i+1} - T_i = \Delta$ (pour tout $i \geq 1$). Cependant, on ne connaît pas le nombre total de boules que l'urne contiendra finalement. Au départ, deux hypothèses sont envisagées:

- (H1₁₄₇) l'urne contiendra finalement 10 boules
- (H2₁₄₈) l'urne contiendra finalement 1000 boules

Les probabilités initiales sont $P(H1) = P(H2) = 0,5$ ⁹⁹. Soit maintenant l'événement E: l'urne contient 1 boule en T_1 . Quelles sont maintenant les probabilités a posteriori $P'(H1)$ et $P'(H2)$?

Ici, il apparaît que l'on peut raisonner de la même manière que dans le cas de *l'objet à incrémentation-temps*. En effet, dans la situation de *l'urne à incrémentation-temps*, la probabilité que la boule n° 1 soit expulsée du dispositif en T_1 si l'urne contient 10 boules est égale à 1. Et de même, la probabilité que la boule n° 1 soit expulsée du dispositif en T_1 si l'urne contient 1000 boules est égale à 1. Ainsi, $P(H1|E) = P(H2|E) = 1$. Les deux hypothèses sont donc équiprobables lorsque survient l'événement E. Ainsi, dans la situation qui est celle de *l'urne à incrémentation-temps*, le tirage de la boule n° 1 n'entraîne pas une révision à la hausse de la probabilité de l'une ou l'autre des hypothèses. Dès lors, il est rationnel de ne pas appliquer le théorème de Bayes dans cette situation. Une telle situation apparaît ainsi, dans ses conséquences, fondamentalement différente de celle du *modèle des deux urnes* où la probabilité que j'extraie du dispositif la boule n° 1 en T_1 est égale à 1/10 si l'urne contient 10 boules et à 1/1000 si l'urne en contient 1000. Les deux situations sont donc fondamentalement différentes.

Ici, on peut remarquer que le modèle de *l'urne à incrémentation-temps* correspond exactement à celui du *distributeur séquentiel d'objets (consecutive token dispenser)* décrit par Eckhardt (1997)¹⁰⁰. La description faite par Eckhardt¹⁰¹ est en effet la suivante:

(...) suppose on each trial the *consecutive token dispenser* expels either 50 (early doom) or 100 (late doom) consecutively numbered tokens at the rate of one per minute.¹⁰²

⁹⁹ On pourrait bien sûr considérer également une version de *l'urne incrémentale à objets multiples* où les probabilités $P(H1)$ et $P(H2)$ initiales ne sont pas données.

¹⁰⁰ Cf. 1997, p. 251.

¹⁰¹ Eckhardt utilise le *distributeur séquentiel d'objets* afin de montrer comment dans ces conditions, il n'est pas possible de déterminer un nombre au hasard, sans connaître préalablement le nombre de boules qui sera, au total, expulsé par la machine. La citation complète est ainsi la suivante: "(...) suppose on each trial the *consecutive token dispenser* expels either 50 (early doom) or 100 (late doom) consecutively numbered tokens at the rate of one per minute. The sampler's task is to select a randomly ranked token as it is expelled. A moment's reflection reveals that this cannot be done without prior knowledge of the pool size for each trial; the sampler must contrive to choose randomly from a domain of 1 through 50 when the machine is going to expel only 50, and randomly from 1 through 100 in other cases, without knowing which cases are which. It makes not the least difference whether the pool size for each trial is decided by a deterministic or indeterministic process; the obstacle is ignorance, not indeterminateness. Without knowledge of the trial's termination, it is impossible to select random ranks." Trad.: "(...) supposons qu'à chaque essai le *distributeur séquentiel d'objets* expulse 50 (apocalypse prochaine) ou bien 100 (apocalypse lointaine) objets numérotés de manière consécutive au taux d'un par minute. La tâche de l'échantillonneur est de choisir un objet comportant un numéro aléatoire au moment où il est expulsé. Une rapide réflexion révèle que ceci ne peut être réalisé sans la connaissance préalable de la taille de l'échantillon à l'occasion de chaque essai; l'échantillonneur doit s'arranger pour choisir de manière aléatoire dans un domaine entre 1 et 50 quand la machine va en expulser seulement 50, et aléatoirement entre 1 et 100 dans les autres cas, sans connaître de quel cas il s'agit à chaque fois. Et il n'y a aucune différence si la taille de l'échantillon à chaque essai est déterminée par un processus déterministe ou indeterministe; le problème réside dans l'ignorance, non dans l'indétermination. Sans connaissance du moment de l'arrêt de l'essai, il est impossible de choisir des rangs aléatoires."

¹⁰² Trad.: "(...) supposons qu'à chaque essai le distributeur d'objets séquentiel expulse 50 (apocalypse prochaine) ou bien 100 (apocalypse lointaine) objets numérotés de manière consécutive au taux d'un par minute.

De manière identique, il s'agit également du dispositif décrit récemment par George Sowers (2002).¹⁰³

There are two urns populated with balls as before, but now the balls are not numbered. Suppose you obtain your sample with the following procedure. You are equipped with a stopwatch and a marker. You first choose one of the urns as your subject. It doesn't matter which urn is chosen. You start the stopwatch. Each minute you reach into the urn and withdraw a ball. The first ball withdrawn you mark with the number one and set aside. The second ball you mark with the number two. In general, the n^{th} ball withdrawn you mark with the number n . After an arbitrary amount of time has elapsed, you stop the watch and the experiment. In parallel with the original scenario, suppose the last ball withdrawn is marked with a seven. Will there be a probability shift? An examination of the relative likelihoods reveals no.¹⁰⁴

Considérons maintenant la variation suivante du modèle probabiliste précédent:

L'urne à incrémentation-objet Une boule numérotée est expulsée, à des intervalles de temps irréguliers par un dispositif, puis ajoutée dans une urne. Les boules sont numérotées ainsi selon leur ordre d'expulsion du dispositif. L'urne contient donc 1 boule au temps T_1 , 2 boules au temps T_2 , ..., n boules au temps T_n . L'intervalle qui sépare deux positions temporelles est *variable*, de sorte que $T_{i+1} - T_i = \Delta_i$ (pour tout $i \geq 1$). Cependant, on ne connaît pas le nombre total de boules que l'urne contiendra finalement. Au départ, deux hypothèses sont formulées:

- (H1₁₄₉) l'urne contiendra finalement 10 boules
- (H2₁₅₀) l'urne contiendra finalement 1000 boules

Les probabilités initiales sont $P(H1) = P(H2) = 0,5$ ¹⁰⁵. Soit maintenant l'événement E: l'urne contient 1 boule en T_1 . Quelles sont maintenant les probabilités a posteriori $P'(H1)$ et $P'(H2)$?

Il apparaît ici qu'une telle expérience est identique à la précédente, à la différence de l'intervalle de temps qui est ici variable. Il apparaît toutefois qu'une telle caractéristique n'a pas d'incidence sur le raisonnement correspondant. On est donc fondé à raisonner de manière identique.

Cette dernière expérience suggère également de considérer, de manière alternative, la variation suivante où des tirages multiples sont effectués:

L'urne à incrémentation-objet itérée Une boule numérotée est expulsée, à des intervalles de temps irréguliers par un dispositif, puis ajoutée dans une urne. Les boules sont numérotées ainsi selon leur ordre d'expulsion du dispositif. L'urne contient donc 1 boule en T_1 , 2 boules en T_2 , ..., n boules en T_n . L'intervalle qui sépare deux positions temporelles est *variable*. Cependant, on ne connaît pas le nombre total de boules que l'urne contiendra finalement. Au départ, on formule deux hypothèses:

¹⁰³ Cf. Sowers (2002, p. 39).

¹⁰⁴ Trad.: Il y a deux urnes comportant des boules de même que précédemment, mais maintenant les boules ne sont pas numérotées. Supposez-vous que vous obtenez votre échantillon avec le procédé suivant. Vous êtes équipés d'un chronomètre et d'un marqueur. Vous choisissez d'abord une des urnes comme sujet d'étude. Peu importe ici quelle urne est choisie. Vous lancez le chronomètre. A chaque minute, vous puisez dans l'urne et retirez une boule. La première boule retirée présente le numéro un et est mise de côté. La deuxième boule possède le numéro deux. De manière générale, la $n^{\text{ième}}$ boule retirée possède le numéro n . Après qu'une durée arbitraire se soit écoulée, vous arrêtez la montre ainsi que l'expérience. De manière similaire au scénario original, supposez que la dernière boule retirée présente un sept. Cela va-t-il conduire à un décalage de la probabilité initiale? Un examen des probabilités correspondantes indique que non.

¹⁰⁵ On pourrait bien sûr considérer également une version de l'urne *incrémentale à objets multiples* où les probabilités $P(H1)$ et $P(H2)$ initiales ne sont pas données.

- (H1₁₅₁) l'urne contiendra finalement 10 boules
- (H2₁₅₂) l'urne contiendra finalement 1000 boules

Les probabilités initiales sont $P(H1) = P(H2) = 0,5$. Soit maintenant les événements E_1, E_2, E_3, E_4, E_5 ¹⁰⁶ qui correspondent respectivement à: je tire la boule 1 en T_1 , la boule 2 en T_2 , la boule 3 en T_3 , la boule 4 en T_5 . Compte tenu des événements E_1, E_2, E_3, E_4, E_5 , quelles sont maintenant les probabilités a posteriori $P'(H1)$ et $P'(H2)$?

Une telle expérience apporte-t-elle une modification fondamentale par rapport à l'expérience non itérée? Il apparaît toutefois que non. Il s'avère en effet qu'à l'étape E_1 , on se trouve exactement dans la situation de l'expérience précédente, c'est-à-dire de l'urne à *incrémentat*ion-objet. Mais de même, à chacune des étapes E_2, E_3, E_4, E_5 , on peut se ramener à une situation tout à fait équivalente à celle de l'étape E_1 et donc de l'urne à *incrémentat*ion-objet. Il suffit en effet de considérer une variation de l'expérience précédente où les hypothèses et les probabilités initiales sont les mêmes que précédemment. Les boules dans l'urne sont numérotées 1, 2, 3, ..., 10 ou bien 1, 2, 3, ..., 1000. A l'étape E_1 , je tire une première boule dans l'urne au temps T_1 et elle porte le numéro 1. Maintenant, cette boule est retirée de l'urne et les boules qui vont être ajoutées dans l'urne sont renumérotées par *d*écémentation. Les boules se trouvant dans l'urne sont donc numérotées 1, 2, 3, ..., 9 ou bien 1, 2, 3, ..., 999. De même, les hypothèses sont ajustées de manière à ne prendre en compte que les boules qui peuvent être ajoutées dans l'urne:

- (H1₁₅₃) l'urne contiendra finalement 9 boules
- (H2₁₅₄) l'urne contiendra finalement 999 boules

Là, j'effectue un second tirage en T_2 et j'extrait la boule n° 1. A nouveau, les boules sont renumérotées par *d*écémentation et les hypothèses sont ajustées de même que précédemment en *d*écémentant les valeurs maximales. Et ainsi de suite.... Dans cette dernière situation, il apparaît ainsi que je tire toujours la boule n° 1. Cette dernière variation met en évidence le fait que le fait de considérer des événements itérés $E_1, E_2, E_3, \dots, E_n$, se ramène en fait au cas de l'urne à *incrémentat*ion-objet. Par conséquent, il n'est pas nécessaire de recourir à une urne *incrémentat*ionnelle à tirages multiples et il suffit de prendre en considération le modèle qui est celui de l'urne à *incrémentat*ion-objet.

9.4. Le modèle des deux urnes à *incrémentat*ion

Considérons maintenant l'expérience suivante:

*Le modèle des deux urnes à incrémentat*ion-temps (objets multiples) On réalise tout d'abord l'expérimentation correspondant au *modèle des deux urnes* au temps T_0 , avec une urne qui contient n boules. A ce stade, la boule n° e est tirée. Une seconde expérience est ensuite mise en œuvre, qui se déroule à des intervalles de temps *constants*, de T_1 à T_e . Un second dispositif expulse en T_1 une boule n° 1, en T_2 une boule n° 2, ..., en T_n une boule n° n . Maintenant, selon le résultat du *modèle des deux urnes* réalisé en T_0 , le dispositif s'arrête lorsque la boule n° e a été expulsée. Ici, le sujet dans la seconde expérience cherche à évaluer le nombre de boules qui étaient contenues en T_0 dans la première urne. Il formule les deux hypothèses suivantes:

- (H1₁₅₅) le nombre de boules contenues dans la première urne est 10
- (H2₁₅₆) le nombre de boules contenues dans la première urne est 1000

Les probabilités initiales sont telles que $P(H1) = P(H2) = 0,5$. Se trouvant devant la seconde urne et prenant connaissance du numéro e de la boule expulsée de l'urne, quelles probabilités le sujet attribue-t-il aux deux hypothèses?

¹⁰⁶ Ici, on pourrait considérer une variation de cette expérience avec $E_1, E_2, E_3, \dots, E_n$, pour $n \leq 10$.

Il apparaît ici que le *modèle des deux urnes à incrémentation-temps* constitue une combinaison de deux expériences: d'une part le *modèle des deux urnes* et d'autre part l'*urne à incrémentation-temps*. D'un point de vue temporel, les deux expériences se révèlent compatibles, le *modèle des deux urnes* se déroulant en T_0 , et l'*urne à incrémentation-temps* prenant place de T_1 à T_n . Il s'avère ici que les deux expériences sont liées et qu'en particulier, la seconde expérience est indexée sur les résultats de la première. Ainsi, la boule n° e est expulsée de la seconde urne en T_e . A ce stade, si je suis informé des conditions de l'expérience et que je me trouve face à la seconde urne en tirant la boule n° e en T_e , ne suis-je pas autorisé à raisonner de la même manière que je le fais dans la situation classique qui est celle du *modèle des deux urnes*? Car la position temporelle T_e , lorsque la seconde expérience s'arrête, est directement liée au fait que la boule n° e a été choisie au hasard dans la première expérience. Dans ce cas, informé des conditions de l'expérience et tirant la boule n° e en T_e , je mets valablement en oeuvre le raisonnement qui sous-tend le *modèle des deux urnes*, et je conclus à un décalage bayésien en faveur de l'hypothèse (H1). Ainsi, il s'avère que quiconque accepte les conclusions qui résultent du *modèle des deux urnes*, accepte également les conclusions qui sont celles du *modèle des deux urnes à incrémentation-temps*.

Il est intéressant également de considérer une autre variation. Celle-ci est constituée d'une combinaison du *modèle des deux urnes* et de l'*objet à incrémentation-temps*. Cette expérience est la suivante:

*Le modèle des deux urnes à incrémentation-temps (objet unique)*¹⁰⁷ On réalise en premier lieu l'expérimentation correspondant au *modèle des deux urnes* au temps T_0 , avec une urne qui contient n boules. A ce stade, la boule n° e est tirée. En second lieu, on a également une seconde expérimentation qui se déroule à des intervalles de temps *constants*, de T_1 à T_e . Cette seconde expérimentation se déroule de la manière suivante. Une boule unique, qui possède une propriété de persistance temporelle présente un numéro indexé sur la position temporelle correspondante: la boule présente le numéro 1 en T_1 , le numéro 2 en T_2 , ..., le numéro 3 en T_n . La seconde expérience s'arrête en T_e alors que la boule présente le numéro e . Le sujet dans la seconde expérience cherche à évaluer le nombre de boules qui étaient contenues en T_0 dans la première urne. Il formule les deux hypothèses suivantes:

(H1₁₅₇) le nombre de boules contenues dans la première urne est 10

(H2₁₅₈) le nombre de boules contenues dans la première urne est 1000

Soit également la variation suivante, légèrement différente, qui constitue la même expérience que le *modèle des deux urnes à incrémentation-temps (objets multiples)*, à la seule différence que les intervalles entre les positions temporelles T_1 à T_n sont variables:

Le modèle des deux urnes à incrémentation-objet On réalise tout d'abord l'expérimentation correspondant au *modèle des deux urnes* au temps T_0 , avec une urne qui contient n boules. A ce stade, la boule n° e est tirée. Une seconde expérience est ensuite mise en oeuvre, qui se déroule à des intervalles de temps *variables*, de T_1 à T_e . Un second dispositif expulse en T_1 une boule n° 1, en T_2 une boule n° 2, ..., en T_n une boule n° n . Maintenant, selon le résultat du *modèle des deux urnes* réalisé en T_0 , le dispositif s'arrête lorsque la boule n° e a été expulsée. Ici, le sujet dans la seconde expérience cherche à évaluer le nombre de boules qui étaient contenues en T_0 dans la première urne. Il formule les deux hypothèses suivantes:

(H1₁₅₉) le nombre de boules contenues dans la première urne est 10

(H2₁₆₀) le nombre de boules contenues dans la première urne est 1000

¹⁰⁷ On pourrait considérer également l'expérience du *modèle des deux urnes à incrémentation-espace (objet unique)*, constituée d'une combinaison du *modèle des deux urnes* et de l'*objet à incrémentation-espace*.

et conclut à un décalage bayésien en faveur de l'hypothèse (H1).

Ici aussi, il apparaît que de même que précédemment, quiconque est confiant dans les résultats du *modèle des deux urnes* est également prêt à accepter les conclusions qui résultent du *modèle des deux urnes à incrémentation-objet*. A ce stade, il apparaît clairement que le *modèle des deux urnes à incrémentation-objet* constitue une combinaison de deux modèles précédemment étudiés: d'une part le *modèle des deux urnes* et d'autre part, l'urne à l'urne à *incrémentation-objet*. De tels modèles sont tout à fait compatibles sur le plan temporel, car le *modèle des deux urnes* prend place en T_0 , alors que celui de l'urne à *incrémentation-objet* se déroule de T_1 à T_c .

9.5. *Objections par rapport à l'analyse de Sowers*

George Sowers (2002) a présenté récemment une solution à DA, qui se révèle, à plusieurs égards, intéressante. Bien qu'elle développe plusieurs points pertinents et contribue à clarifier le problème, il me semble cependant que l'analyse de Sowers s'avère insuffisante, car elle n'atteint pas véritablement le coeur du problème posé par DA.

Sowers décrit ainsi DA comme un argument basé sur une analogie existant entre la situation humaine correspondant à DA et le *modèle des deux urnes*. Sowers considère que le rang de naissance de chaque humain n'est pas aléatoire, car il est indexé sur la position temporelle correspondante. Par conséquent, l'analogie avec le *modèle des deux urnes* n'est pas valable. Son analyse consiste à montrer comment la situation humaine correspondant à DA est en fait analogue à la situation suivante¹⁰⁸:

There are two urns populated with balls as before, but now the balls are not numbered. Suppose you obtain your sample with the following procedure. You are equipped with a stopwatch and a marker. You first choose one of the urns as your subject. It doesn't matter which urn is chosen. You start the stopwatch. Each minute you reach into the urn and withdraw a ball. The first ball withdrawn you mark with the number one and set aside. The second ball you mark with the number two. In general, the n^{th} ball withdrawn you mark with the number n . After an arbitrary amount of time has elapsed, you stop the watch and the experiment. In parallel with the original scenario, suppose the last ball withdrawn is marked with a seven. Will there be a probability shift? An examination of the relative likelihoods reveals no.¹⁰⁹

L'expérience ici décrite par Sowers est, on l'a vu plus haut, celle que nous avons dénommée *l'urne à incrémentation-temps*. La conclusion à laquelle elle conduit, étant donné votre rang de naissance, est que les probabilités initiales concernant les deux hypothèses concurrentes (H1) et (H2) sur lesquelles est basé DA demeurent inchangées.

¹⁰⁸ Cf. Sowers (2002, p. 39). Ce dispositif est identique à celui du *distributeur séquentiel d'objets* décrit par William Eckhardt (1997, p. 251): '(...) suppose on each trial the *consecutive token dispenser* expels either 50 (early doom) or 100 (late doom) consecutively numbered tokens at the rate of one per minute.'. Trad.: '(...) supposons qu'à chaque essai le *distributeur séquentiel d'objets* expulse 50 (apocalypse prochaine) ou bien 100 (apocalypse lointaine) objets numérotés de manière consécutive au taux d'un par minute'.

¹⁰⁹ Trad.: Il y a deux urnes comportant des boules de même que précédemment, mais maintenant les boules ne sont pas numérotées. Supposez-vous que vous obtenez votre échantillon avec le procédé suivant. Vous êtes équipés d'un chronomètre et d'un marqueur. Vous choisissez d'abord une des urnes comme sujet d'étude. Peu importe ici quelle urne est choisie. Vous lancez le chronomètre. A chaque minute, vous puisez dans l'urne et retirez une boule. La première boule retirée présente le numéro un et est mise de côté. La deuxième boule possède le numéro deux. De manière générale, la $n^{\text{ième}}$ boule retirée possède le numéro n . Après qu'une durée arbitraire se soit écoulée, vous arrêtez la montre ainsi que l'expérience. De manière similaire au scénario original, supposez que la dernière boule retirée présente un sept. Cela va-t-il conduire à un décalage de la probabilité initiale? Un examen des probabilités correspondantes indique que non.

Le point fondamental qui résulte de l'analyse de Sowers est que le rang de naissance de chaque humain n'est pas aléatoire, mais se révèle indexé sur la position temporelle correspondante¹¹⁰. Par conséquent, selon Sowers, DA comporte une violation manifeste du principe HR ou SSA. Ainsi, DA constitue un argument basé sur une fausse analogie. On peut remarquer ici qu'une telle conclusion est tout à fait similaire au point de vue développé à la fois par William Eckhardt et Elliott Sober.

En effet, Eckhardt en effet met l'accent sur le fait que le rang de naissance ne peut faire l'objet d'un tirage aléatoire sans la connaissance du nombre d'objets total figurant dans l'urne. Il considère de même que l'analogie qui sous-tend DA se révèle infondée car notre rang de naissance ne présente pas un caractère aléatoire¹¹¹.

De même, Sober (2001) considère de même que notre rang de naissance ne présente pas un caractère aléatoire. Il fait valoir que le mécanisme qui génère un numéro aléatoire est absent et par conséquent, le principe HR ou SSA se trouve violé¹¹².

Ma première observation par rapport à l'analyse de Sowers est la suivante. Sowers décrit l'expérience de l'urne à incrémentation-temps, où les boules sont expulsées d'une urne à des intervalles de temps constants (chaque minute). Les boules sont numérotées dans l'ordre de leur expulsion de l'urne. Cependant, une telle situation ne correspond pas véritablement à la situation humaine correspondant à DA où les humains apparaissent sur Terre à des intervalles *variables*. Sowers considère ainsi que la cause de DA réside dans le caractère non-aléatoire du numéro et son indexation par rapport au temps¹¹³. Or cela ne s'avère pas tout à fait précis, car le rang de naissance de chaque humain se trouve

¹¹⁰ Cf. également Sowers (2002, p. 44): "The doomsday argument has been shown to be fallacious due to the incorrect assumption that you are a random sample from the set of all humans that have ever existed." Trad.: "l'argument de l'apocalypse s'est donc révélé fallacieux, à cause de l'hypothèse incorrecte selon laquelle vous êtes un échantillon aléatoire de l'ensemble de tous les humains qui ont jamais existé."

¹¹¹ Cf. (1993, p. 486): "A procedure that invariably yields an existent item cannot be equiprobable sampling, since in that case nonexistent members of the pool could not be receiving appropriate weight. Yet the sampling procedure employed in the Doomsday Argument invariably yields a result – a human rank current at the time of the argument's discovery. Hence, this cannot be equiprobable sampling from an ensemble only part of which currently exists." et aussi (1997, p. 257): "Examples such as the *consecutive token machine* illustrate the impossibility of choosing a random rank without knowledge of how many ranks exist overall (or without being able to interact with the population as in drawing from an urn)." Trad.: "un procédé qui renvoie invariablement un objet existant ne peut pas être considéré véritablement comme équiprobable, puisque dans ce cas les membres non-existants de l'échantillon ne pourraient pas être pris en considération. Pourtant le procédé de prélèvement utilisé dans l'argument de l'apocalypse donne invariablement un résultat - un rang humain actuel correspondant au moment de la découverte de l'argument. Par conséquent, ceci ne peut pas être considéré comme un prélèvement équiprobable d'un objet au sein d'un ensemble dont seulement une partie existe actuellement." Et aussi (1997, p. 257): les "exemples tels que le distributeur séquentiel d'objets illustrent l'impossibilité de choisir un rang aléatoire sans la connaissance préalable de combien de rangs existent au total (ou sans pouvoir interagir avec la population de la même manière que lorsque l'on tire dans une urne)."

¹¹² Cf. (2001): "I think that Leslie's sampling assumption (S) is false. My present temporal position is *not* the result of random sampling from the temporal locations of all human beings any more than I am writing this paper in July 2000 because a time was selected at random for me to do so. (...) But who or what has the propensity to randomly assign me a temporal location in the duration of the human race? There is no such mechanism." Trad.: "Je pense que l'hypothèse d'équiprobabilité (s) de Leslie est fausse. Ma position temporelle actuelle n'est *pas* le résultat de l'échantillonnage aléatoire des positions temporelles de tous les êtres humains, pas plus que j'écris cet article en juillet 2000 parce qu'une période de temps a été choisi au hasard pour moi pour ce faire (...) Mais qui ou quoi a donc la propension à m'assigner au hasard une position temporelle par rapport à la durée de la race humaine? Aucun mécanisme de ce type n'existe."

¹¹³ Cf. Sowers (2002, p. 40): "My claim is that by assigning a rank to each person based on birth order, a time correlation is established in essentially the same way that the stopwatch process established a

en effet indexé directement sur le numéro d'ordre d'apparition sur Terre de chaque humain, mais pas sur la position temporelle correspondante. Dans la situation correspondant à DA, les humains apparaissent à des intervalles de temps variables. Toutefois, on peut penser qu'il ne s'agit pas là d'un point fondamental. Car on pourrait considérer, de manière alternative, une variation de l'urne à incrémentation-temps décrite par Sowers où les boules sont expulsées de l'urne à des intervalles de temps variables et où les numéros des boules correspondent à leur ordre d'expulsion de l'urne.

J'en viens maintenant au point qui me paraît essentiel, et que je considère comme une importante objection par rapport à l'analyse de Sowers. Cette dernière est en effet basée sur le fait que le rang de naissance de chaque humain n'est pas aléatoire, mais se révèle indexé sur la position temporelle correspondante. Et ceci, selon Sowers, interdit l'analogie avec le modèle des deux urnes, dont l'une des conditions est l'existence d'un tirage aléatoire. Cependant, je m'attacherai ici à démontrer qu'il est tout à fait possible d'avoir à la fois un tirage qui présente un caractère aléatoire *et* une indexation du rang de naissance de chaque humain sur la position temporelle correspondante.

Tout d'abord, il est certainement exact, comme l'indique Sowers, que le raisonnement fondé sur une analogie entre la situation humaine correspondant à DA et le *modèle des deux urnes* n'est pas valable. Car une analyse sommaire montre en effet que la situation humaine correspondant à DA est temporelle, alors que le *modèle des deux urnes* est atemporel. On a pu vérifier en effet que la situation humaine correspondant à DA pouvait être modélisée dans un n-univers temporel, alors que le *modèle des deux urnes* prenait place dans un n-univers atemporel. En termes de n-univers, la situation humaine correspondant à DA prend place dans un n-univers temporel, alors que le *modèle des deux urnes* prend place dans un n-univers atemporel. Aussi les deux situations ne peuvent pas être considérées comme isomorphes.

Dans ce contexte, on peut penser que l'analogie qu'ont à l'esprit Carter et Leslie, et également Bostrom, est davantage celle existant entre d'une part la situation humaine correspondant à DA et d'autre part le *modèle des deux urnes à incrémentation-objet*. Il est permis de penser ainsi que c'est cette dernière analogie qui constitue véritablement le cœur du raisonnement inhérent à DA. Par conséquent, le décalage bayésien auquel conduit DA est davantage basé sur une analogie entre la situation humaine correspondant à DA et l'expérience du *modèle des deux urnes à incrémentation-objet* décrite plus haut. Dans cette dernière situation en effet, le numéro de chaque boule expulsée du second dispositif est indexé sur l'ordre de son apparition dans l'urne. Dans ce cas, je tire la boule n° 6000000000. Mais je sais également que la boule précédente est la boule n° 5999999999 et que la boule suivante sera la boule n° 6000000001, etc. Pourtant, cela ne m'empêche pas de raisonner de la même manière que dans le *modèle des deux urnes* classique et de conclure à un décalage bayésien en faveur de l'hypothèse (H1). On le voit, dans le *modèle des deux urnes à incrémentation-objet*, on observe bien la *conjonction* du tirage aléatoire et de l'indexation du rang de naissance sur la position temporelle.

Les remarques qui précèdent m'amènent à considérer que l'analyse concrète développée par Sowers ne se révèle cependant pas suffisante pour résoudre DA. Car l'analyse de Sowers ne s'applique pas véritablement à ce qui constitue le cœur de DA, c'est-à-dire une analogie entre la situation humaine correspondant à DA et le *modèle des deux urnes à incrémentation-objet*. Cette dernière situation constitue en fait une combinaison du *modèle des deux urnes* et de l'*urne à incrémentation-temps*. Car on observe bien dans cette dernière situation l'existence conjointe d'un tirage aléatoire et d'une indexation du rang de naissance de chaque humain sur la position temporelle correspondante. Dans ce contexte, la conclusion de Sowers selon laquelle son analyse conduit à la fin de DA apparaît trop forte. Il a certainement apporté des étapes supplémentaires conduisant vers la résolution de DA et clarifié des points importants. Mais son analyse n'atteint pas toutefois pas véritablement les formulations les plus fortes de DA.

correlation with the balls.". Trad.: "Ce que j'affirme ici, c'est qu'en affectant un rang à chaque personne basé sur l'ordre de naissance, une corrélation avec le temps est établie d'une manière qui est essentiellement la même que le processus du chronomètre a établi une corrélation avec les boules.".

9.6. Analyse

A ce stade, il convient désormais de s'attacher à présenter une solution à DA. Les éléments qui précèdent conduisent tout d'abord à reformuler le problème posé par DA. Il convient ainsi de prendre en considération le fait que l'analogie qui sous-tend véritablement DA ne s'établit pas entre la situation correspondant à DA et le *modèle des deux urnes*, mais bien entre la situation correspondant à DA et le *modèle des deux urnes à incrémentation*. On vient de le voir en effet en discutant l'analyse développée par Sowers, la situation humaine correspondant à DA est temporelle, alors que le *modèle des deux urnes* est atemporel. C'est ce que montrent les modélisations successives de ces deux situations dans le cadre des n-univers. Car la situation humaine correspondant à DA se trouve modélisée dans un $\Omega\alpha T^*C$ ¹¹⁴, alors que le modèle des deux urnes prend place dans un $\Omega\alpha^*C$ ou un $\Omega\alpha C$. Par conséquent, l'analogie sur laquelle est fondée DA ne s'établit pas véritablement entre la situation humaine correspondant à DA et le *modèle des deux urnes*, mais bien entre la situation humaine correspondant à DA et le *modèle des deux urnes à incrémentation*. Ici, les deux situations apparaissent bien comme isomorphes, car tant la situation humaine correspondant à DA que le modèle des deux urnes à incrémentation prennent place dans le $\Omega\alpha T^*C$. Par conséquent, c'est à la version de DA basée sur cette dernière analogie qu'il convient de s'attacher à donner une solution.

Dans ce contexte, le problème posé par DA apparaît comme le suivant: la situation correspondant à DA est-elle analogue au *modèle des deux urnes à incrémentation*? Mais comme le souligne Sowers et ainsi que cela résulte également de l'analyse développée par Eckhardt, une approche alternative est la suivante: la situation correspondant à DA n'est-elle pas analogue au modèle de *l'urne à incrémentation*? Ceci conduit finalement à formuler le problème posé par DA de la manière suivante: la situation correspondant à DA est-elle analogue au *modèle des deux urnes à incrémentation* ou bien à celui de *l'urne à incrémentation*? Il s'avère ainsi nécessaire de s'attacher à apporter une réponse précise à cette question. Pour cela, il conviendra de s'attacher ici à déterminer quels sont les éléments objectifs qui permettent de privilégier l'une ou l'autre des deux analogies. Quels sont donc les éléments objectifs qui permettent de choisir l'un ou l'autre des deux modèles probabilistes? Pour cela, il conviendra d'étudier tour à tour chacune des deux situations.

A ce stade, il s'avère également nécessaire de prendre en considération la version bi-dimensionnelle de DA. Dans ce contexte, DA-objets apparaît d'une part comme un raisonnement basé sur une analogie entre la situation correspondant à DA-objets et le *modèle des deux urnes à incrémentation-objet*. Et de même, DA-temps constitue d'autre part un raisonnement basé sur une analogie entre la situation correspondant à DA-temps et le *modèle des deux urnes à incrémentation-temps (objet unique)*. Ainsi le problème posé par DA-objets est-il celui de savoir si la situation correspondant à DA-objets est analogue au *modèle des deux urnes à incrémentation-objet* ou bien à *l'urne à incrémentation-objet*. Parallèlement, le problème posé par DA-temps est de savoir si la situation correspondant à DA-temps est analogue au *modèle des deux urnes à incrémentation-temps (objet unique)* ou bien à *l'objet à incrémentation-temps*. Il se révèle nécessaire d'analyser ici chacune de ces deux situations.

Considérons tout d'abord DA-objets. Le problème posé par DA-objets est le suivant: la situation correspondant à DA-objets est-elle analogue au *modèle des deux urnes à incrémentation-objet* ou bien à *l'urne à incrémentation-objet*? Intuitivement, l'analogie entre la situation correspondant à DA-objets et le *modèle des deux urnes à incrémentation-objet* est basée sur le fait que la situation suivante est susceptible de correspondre à DA-objets:

Dieu et l'humanité-objet Avant la naissance de l'humanité, Dieu appelle l'un après l'autre chacun des futurs humains. Lorsque vient votre tour, il vous annonce: "Durant l'existence future de l'humanité, un très grand nombre d'humains seront appelés à vivre. Toi aussi, tu seras appelé à participer à l'aventure de l'humanité. Pour des raisons qui t'échappent encore, le nombre d'humains qui composeront la future humanité est d'ores et déjà fixé. Dans cette urne, j'ai placé autant de boules numérotées que l'humanité comptera d'humains pendant toute la durée de son existence. Les numéros de boules correspondent à l'ordre d'apparition de chaque humain durant

¹¹⁴ De manière générale, on a pu observer que la modélisation pouvait s'effectuer dans le $\Omega\alpha^*T^*C$, mais aussi dans le $\Omega\alpha T^*C$ et le $\Omega\alpha TC$.

l'existence de l'humanité. Et je vais maintenant procéder à un tirage au sort afin de déterminer quel sera ton rang de naissance. Tu vas donc tirer au hasard une boule. Maintenant, vous vous avancez et tirez la boule n° 60×10^9 .

Quelques millions d'années plus tard, vous vivez sur la planète Terre et vous avez perdu la mémoire de tous les événements précédents. Vous formulez alors les deux hypothèses suivantes:

(H1₁₆₁) le nombre d'*homo sapiens*¹¹⁵ sera égal à 100 milliards

(H2₁₆₂) le nombre d'*homo sapiens* sera égal à 10000 milliards

auxquelles vous attribuez respectivement une probabilité initiale de 0,5. Considérant maintenant votre rang de naissance au sein de l'espèce *homo sapiens* qui est n° 60×10^9 , vous êtes ramené à réviser à la hausse la probabilité initiale selon laquelle le nombre total d'*homo sapiens* atteindra seulement 100 milliards.

Dans une telle situation, n'est il pas rationnel de raisonner de la même manière que lorsqu'on est en présence d'un *modèle des deux urnes*? N'est-on pas fondé à conclure ici à un décalage bayésien en faveur de l'hypothèse (H1). On le voit, le raisonnement qui sous-tend l'expérience *Dieu et l'humanité-objet* se révèle de même nature que celui qui prévaut dans le *modèle des deux urnes* classique.

Considérons ici, par commodité, que le tirage au sort de l'expérience *Dieu et l'humanité-objet* se déroule au temps T_0 . A la suite de ce tirage, la situation correspondant à DA-objets prend place. Survient alors en effet l'événement E: votre rang de naissance est le n° 60×10^9 . On le voit, l'ensemble constitué par le tirage au sort de l'expérience *Dieu et l'humanité-objet* et la situation correspondant à DA-objets présente une structure isomorphe à celle du modèle probabiliste du *modèle des deux urnes à incrémentation-objet*. Dans ce contexte, le décalage bayésien final constitue une conclusion qui apparaît tout à fait rationnelle.

Les développements qui précèdent permettent de mettre en lumière le problème posé par l'analogie entre DA-objets et le *modèle des deux urnes à incrémentation-objet*. Car il s'agit en effet de savoir si la situation correspondant à DA-objets a été précédée ou non par un tirage au sort du type de celui qui intervient dans l'expérience *Dieu et l'humanité-objet*. Car cette dernière situation correspond précisément au *modèle des deux urnes*. Et une telle situation repose sur l'hypothèse¹¹⁶ d'un tirage aléatoire parmi les n numéros correspondant aux n humains qui composeront l'humanité. En effet, on ne peut se prononcer avec certitude sur l'hypothèse déterministe. Compte tenu de nos connaissances actuelles, le statut de cette dernière hypothèse apparaît indéterminé. Et étant donné l'ignorance dans laquelle nous sommes des conditions ayant précédé notre existence, nous ne sommes pas en mesure non plus de confirmer ou d'infirmer l'existence préalable à l'existence de l'humanité d'un tirage aléatoire. Il n'en demeure pas moins que si l'hypothèse d'un tirage aléatoire se trouve vérifiée, alors l'application du *modèle des deux urnes à incrémentation-objet* est légitime. Par contre, si cette dernière hypothèse est infirmée, alors le modèle du *modèle des deux urnes à incrémentation-objet* ne peut être valablement appliqué. A ce stade, il s'avère que l'existence ou non avant l'existence de l'humanité d'un tirage aléatoire du type de celui mentionné dans l'expérience *Dieu et l'humanité-objet* ne peut être déterminée avec certitude. Ainsi, le statut d'une telle hypothèse est-il *indéterminé*. On pourra arguer qu'un mode opérationnel tel que celui qui préside à *Dieu et l'humanité-objet* est très improbable. Cependant, on ne peut conclure ici avec certitude, sur le fondement d'éléments objectifs, qu'une

¹¹⁵ On considère ici une version de DA-objets qui tient compte des remarques précédentes concernant le problème de la définition de la classe de référence. On considère ainsi une version de DA-objets basée sur une classe de référence correspondant, avec un certain degré d'arbitraire, à l'espèce *homo sapiens*.

¹¹⁶ Cette situation repose également sur l'hypothèse d'un univers déterministe, c'est-à-dire d'une détermination préalable à l'existence de l'humanité du nombre d'humains qui la composeront finalement. On a vu cependant plus haut que l'on se plaçait délibérément, dans la présente discussion, dans l'hypothèse d'une situation déterministe, une situation qui avantage DA.

situation de nature analogue ne s'est pas produite¹¹⁷. A ce stade, il apparaît que l'on ne possède pas d'éléments objectifs pour considérer que notre existence est précédée ou non d'une situation telle que *Dieu et l'humanité-objet*. En l'absence d'éléments objectifs confirmant ou infirmant l'hypothèse d'une telle situation, il apparaît naturel de considérer cette dernière comme hypothétique et indéterminée. La possibilité d'une telle situation existe mais aucun élément de preuve n'est disponible. L'incertitude prévaut ici, et par conséquent, il convient de retenir, concernant cette dernière hypothèse, une nature indéterminée. Une première conséquence de cette analyse est ainsi la suivante: on n'est pas fondé à faire légitimement application du *modèle des deux urnes à incrémentation-objet*, car l'isomorphisme de la situation correspondant à DA-objet avec ce dernier modèle n'a pu être établi. En effet, dans le *modèle des deux urnes à incrémentation-objet*, l'existence au temps T_0 d'une expérience du type *modèle des deux urnes* est établi avec certitude. A l'inverse, dans la situation correspondant à DA, l'incertitude demeure quant à l'existence préalablement à l'apparition de l'humanité d'une situation de ce type. Par conséquent, il est rationnel de ne pas admettre l'analogie entre la situation correspondant à DA et le *modèle des deux urnes* à incrémentation. Cette dernière analogie n'est donc pas fondée. Il s'ensuit à ce stade que le décalage bayésien ne peut pas être légitimement appliqué. Par conséquent, on n'est pas fondé ici à mettre en oeuvre le décalage bayésien qui résulte du modèle à minoration. Ceci montre finalement que l'analogie qui préside à DA se révèle infondée. On le voit, on n'a pu mettre en évidence ici l'existence certaine d'un tirage aléatoire au temps T_0 . Dans ce contexte, mon numéro d'humain¹¹⁸ ne peut être considéré, de manière certaine, comme aléatoire. Le fait que le numéro correspondant à l'événement E dans DA n'est pas aléatoire, en violation directe de HR¹¹⁹, a été

¹¹⁷ En lieu et place de l'hypothèse divine prise en compte dans l'expérience *Dieu et l'humanité-objet*, on peut envisager si nécessaire une expérience de même nature où le tirage aléatoire s'effectue sur des bases purement physiques.

¹¹⁸ *Birth rank*.

¹¹⁹ Ici, il convient de reconsidérer la formulation des principes HR (Eckhardt) ou SSA (Bostrom). On peut penser en effet que la formulation de ces principes doit éventuellement être modifiée pour capturer l'intuition qui préside à DA et à la notion de tirage aléatoire. Compte tenu de ce qui précède, il apparaît que la formulation correcte du principe est la suivante:

(HR*) l'attribution des numéros aux membres de la classe de référence doit être effectuée de manière aléatoire

Car on doit bien considérer que c'est l'attribution des numéros aux membres de la classe de référence qui doit être effectuée de manière aléatoire. Mon numéro dans l'urne doit être attribué au hasard. Un tel principe apparaît bien analogue à celui formulé par Eckhardt:

(HR) We can validly consider our birth rank as generated by random or equiprobable sampling from the collection of all persons who ever live. Trad.: Nous pouvons bien valablement considérer notre rang de naissance comme résultant d'un prélèvement aléatoire ou équiprobable parmi l'ensemble de toutes les personnes ayant jamais existé.

Par contre, il se distingue du principe SSA défini par Bostrom:

(SSA) Observers should reason as if they were a random sample from the set of all observers in their reference class. Trad.: Les observateurs devraient raisonner comme s'ils étaient un échantillon aléatoire de l'ensemble de tous les observateurs au sein de leur classe de référence

En effet, compte tenu de ce qui précède, ce n'est pas la fait que je ne suis pas un humain aléatoire qui se révèle problématique ici, mais le fait que mon *numéro* ne m'est pas attribué de manière aléatoire. La violation de HR à l'origine de DA qui vient d'être décrite doit ainsi être distinguée d'une objection qui a pu être proposé, en particulier par Eckhardt, en tant que solution à DA (Cf. Eckhardt (1997, p. 257-8): "The mere fact of being a likely candidate for random member in one class, however, often precludes a similar status in some other class. A random human is surely an exceptional vertebrate,

souligné par plusieurs auteurs. Mais la *cause* du fait que mon rang de naissance n'est pas aléatoire n'est pas, comme l'indiquent les analyses de Sowers, Eckhardt ou Sobers, le fait que ce numéro de naissance soit indexé sur la position temporelle ou le numéro d'ordre correspondant, mais bien le fait que l'on n'est pas en mesure d'établir l'existence d'un mode opératoire aléatoire préalable à l'existence de l'humanité.

Ce qui précède a permis de mettre en évidence le fait que l'analogie qui sous-tend DA se révèle infondée. En effet, la présence d'un modèle des deux urnes en T_0 se révèle certaine dans le *modèle des deux urnes à incrémentation-objet*, alors qu'elle ne présente qu'un caractère hypothétique dans la situation correspondant à DA. Cependant, puisque DA constitue ainsi un raisonnement de nature fallacieuse, quel est donc le raisonnement qui doit être valablement tenu dans une telle situation? Dès lors qu'on n'est pas fondé à mettre en oeuvre valablement le *modèle des deux urnes à incrémentation-objet*, est-on donc pour cela autorisé à faire usage légitimement du *modèle de l'urne à incrémentation-objet*? Le raisonnement valide qui doit être tenu, en présence d'une situation correspondant à DA-objet, en lieu et place de la version classique de DA, est-il celui qui prévaut dans le cas de l'*urne à incrémentation-objet*? Dois-je ainsi conclure que $P(E|H1) = P(E|H2) = 1$ et qu'il s'ensuit que les probabilités initiales sont inchangées? Il convient d'analyser cela en détail¹²⁰. La question primordiale est alors celle-ci. Suis-je fondé de manière certaine à faire application du modèle de l'*urne à incrémentation-objet*? Or il apparaît que l'application du modèle de l'*urne à incrémentation-objet* se révèle également hasardeuse, pour les raisons suivantes. Le modèle de l'*urne à incrémentation-objet* se révèle en effet susceptible d'être englobé dans un modèle plus sophistiqué, dont il constitue alors

and so on. (...) Whether or not we are likely candidates for the status of random human, we cannot be counted random mammals, random vertebrates, or random solid objects; we are too atypical as members of the last three classes." Trad.: "Le simple fait d'être un candidat probable pour être un membre choisi au hasard dans une classe donnée, cependant, exclut souvent un statut semblable dans une autre classe. Un humain aléatoire est sûrement un vertébré exceptionnel, et ainsi de suite (...) Si nous sommes des candidats probables pour le statut d'humain aléatoire, nous ne pouvons pas être des mammifères choisis de manière aléatoire, des vertébrés choisis au hasard, ou des objets solides aléatoirement choisis; nous sommes trop atypiques comme membres des trois dernières classes." La violation de HR qui vient d'être décrite doit en effet être distinguée de l'objection à DA basée sur le fait que chacun d'entre nous n'est pas un humain représentatif de la classe de référence. Cette objection a pu en effet être opposée à DA. Mais une telle objection a pu être rejetée (à mon sens avec succès) par Leslie. En effet, il s'avère que l'on peut mettre en évidence des variations où la problématique inhérente à DA est présente, quelle que soit la composition de la classe de référence. Car ce qui fait l'objet des hypothèses dans DA et le *modèle des deux urnes*, c'est le nombre d'humains ou bien le nombre d'objets contenus dans l'urne. Et on peut aussi bien considérer une version du *modèle des deux urnes* où les objets placés dans l'urne sont hétéroclites. Soit donc une urne qui contient 10 objets consistant en deux tulipes, un stylo, trois tétraèdres, un dé à coudre et trois escargots, ou bien 1000 objets comprenant les mêmes dix objets que précédemment, mais aussi six roses, deux cubes, cinq sphères, ..., huit coccinelles. Chacun des objets est numéroté. Je tire dans l'urne un tétraèdre comportant le numéro 5 et j'applique alors le raisonnement basé sur l'analogie avec le *modèle des deux urnes* qui conduit à un décalage bayésien en faveur de l'hypothèse selon laquelle l'urne ne contient que 10 objets. Ici, il apparaît bien que je peux mettre en oeuvre valablement le raisonnement qui est celui du *modèle des deux urnes*. En conséquence, même si je ne suis pas un humain représentatif de la classe de référence, je peux valablement mettre en oeuvre le raisonnement inhérent à DA dès lors que mon numéro se trouve déterminé de manière aléatoire.

¹²⁰ Il apparaît que si l'analogie entre la situation correspondant à DA-objets et l'*urne à incrémentation-objet* prévaut, le fait que la situation relative au nombre d'humains ayant jamais existé soit de nature déterministe ou indéterministe est indifférent. En effet, le modèle de l'*urne à incrémentation-objet* s'applique dans le cas d'un univers déterministe mais également dans le cas d'un univers indéterministe où le nombre de boules qui seront finalement expulsées de l'urne n'est pas connu au moment où débute l'expérience. Eu égard à la nature indéterministe ou non de cette situation, le modèle de l'*urne à incrémentation-objet* se révèle donc indifférent. En comparaison, le modèle alternatif qui est celui des *deux urnes à incrémentation-objet* ne vaut que pour un univers déterministe.

seulement l'un des composants. Car le modèle de l'urne à *incrément-objet* prend place de T_1 à T_n . Et ce dernier modèle est susceptible d'être englobé dans un modèle plus général qui est celui des *deux urnes à incrément-objet* et qui se déroule aux positions temporelles T_0, T_1 à T_n . Dans ce cas, un modèle des deux urnes prend alors place au temps T_0 . Or précisément, dans la situation correspondant à DA-objet, il existe une période de temps, celle qui précède la naissance de l'humanité, qui correspond à T_0 . Et à ce moment précis, une expérience du type du modèle des deux urnes est susceptible de prendre place. Dans ces conditions, il est tout à fait possible que la situation correspondant à DA-objet soit isomorphe au modèle des *deux urnes à incrément-objet*. Cependant, on n'est pas en mesure, ainsi qu'on l'a vu plus haut, de confirmer ou d'infirmer cela sur un fondement objectif. En toute hypothèse cependant, l'incertitude pèse sur ce type de situation. En conséquence, il apparaît ici que l'on ne peut faire valablement application du modèle de l'urne à *incrément-objet*. Ainsi, on n'est pas fondé à conclure, de manière certaine, comme le font Sowers, Eckhardt ou Sobers, que la situation correspondant à DA-objet est isomorphe au modèle de l'urne à *incrément-objet*¹²¹.

La situation qui régit DA-temps se révèle tout à fait parallèle. Le problème posé par DA-temps est en effet le suivant: la situation correspondant à DA-temps est-elle analogue au *modèle des deux urnes à incrément-temps* ou bien à l'urne à *incrément-temps*? Intuitivement, l'analogie entre la situation correspondant à DA-temps et le *modèle des deux urnes à incrément-temps* est basée sur le fait que la situation suivante est susceptible de correspondre à DA-temps:

Dieu et l'humanité-temps Avant la naissance de l'humanité, Dieu appelle l'un après l'autre, chacun des futurs humains. Lorsque vient votre tour, il vous annonce: "L'existence future de l'humanité s'étendra sur un certain nombre de millions d'années. Pour des raisons qui t'échappent encore, la durée d'existence de la future humanité est d'ores et déjà fixée. Toi aussi, tu seras appelé à participer à l'aventure de l'humanité. J'ai découpé la période pendant laquelle existera l'humanité en autant de millions d'années. Dans cette urne, j'ai placé autant de boules numérotées que l'existence de l'humanité comportera de millions d'années. Je vais maintenant procéder à un tirage au sort afin de déterminer pendant quel million d'années tu vivras sur Terre. Tu vas donc tirer au hasard une boule. Maintenant, vous vous avancez et tirez la boule n° 7.

Quelques millions d'années plus tard, vous vivez sur la planète Terre et vous avez perdu la mémoire de tous les événements précédents. Vous formulez alors les deux hypothèses suivantes:

- (H1₁₆₃) l'existence d'*homo sapiens* durera 10 millions d'années (*apocalypse prochaine*)
- (H2₁₆₄) l'existence d'*homo sapiens* durera 1000 millions d'années (*apocalypse lointaine*)

auxquelles vous attribuez respectivement une probabilité initiale de 0,5. Considérant maintenant que le million d'années d'existence de l'espèce *homo sapiens* dans lequel vous vous trouvez est le n° 7, vous êtes ramené à réviser à la hausse la probabilité initiale selon laquelle l'existence d'*homo sapiens* ne s'étendra que sur 10 millions d'années.

Une telle situation ne présente-t-elle pas des caractéristiques qui conduisent à un décalage bayésien en faveur de l'hypothèse (H1), de manière identique au raisonnement qui prévaut dans le cadre d'un *modèle des deux urnes* classique. On peut considérer ici que l'expérience *Dieu et l'humanité-temps* proprement dite se déroule en T_0 . A la suite de cette expérience, la situation correspondant à DA-temps prend alors place. L'ensemble constitué par l'expérience *Dieu et l'humanité-temps* et la situation correspondant à DA-temps se révèle ainsi analogue au *modèle des deux urnes à incrément-temps*. Ici aussi, le problème posé par DA est celui de savoir si la situation correspondant à DA-temps a été précédée ou non par une situation du type *Dieu et l'humanité-temps*. Car cette dernière situation correspond au modèle probabiliste du *modèle des deux urnes*. Ici aussi, il apparaît que l'analogie entre la situation correspondant à DA-temps et le *modèle des deux urnes à incrément-temps* repose sur deux hypothèses: d'une part (a) celle d'un univers déterministe où le nombre de millions d'années qui

¹²¹ Ainsi qu'on l'a vu, Sowers utilise plus précisément une analogie avec le modèle de l'urne à *incrément-temps*.

composeront l'existence de l'humanité est d'ores et déjà fixé. Et d'autre part (b), l'existence d'un tirage aléatoire précédant notre existence portant sur le numéro de million d'années correspondant. On le voit, les conditions de validité de l'analogie entre la situation correspondant à DA-temps et le *modèle des deux urnes à incrémentation-temps* sont de la même nature que celles qui prévalent entre la situation correspondant à DA-objets et le *modèle des deux urnes à incrémentation-objet*.

Finalement, il apparaît que l'on peut classer plusieurs solutions à DA rencontrées dans la littérature en deux catégories. Une première catégorie de solutions considère que l'analogie entre la situation correspondant à DA et le modèle des deux urnes est valablement établie, et par conséquent, le décalage bayésien qui en résulte doit être accepté. Les analyses développées par Carter et Leslie appartiennent ainsi à cette première catégorie. Une seconde série de solutions présentes dans la littérature rejettent l'analogie entre la situation correspondant à DA et le modèle des deux urnes. Les solutions qui appartiennent à cette catégorie sont notamment celles développées par Eckhardt, Sowers et Sobers. Ces solutions mettent en avant le fait que l'analogie qui prévaut véritablement est celle existant entre la situation correspondant à DA et l'urne à incrémentation-temps (objets multiples). Ces deux types de solutions consistent donc soit à accepter, soit à rejeter le raisonnement analogique sur lequel est basé DA.

Ainsi, la présente analyse se distingue de ces deux derniers types de solutions en choisissant une troisième voie: chacune des deux analogies avec l'urne précédentes n'est ni acceptée, ni rejetée. Car l'existence ou non d'une situation analogue au modèle des deux urnes en T_0 c'est-à-dire avant la naissance de l'humanité, se révèle indéterminée. L'existence éventuelle d'une telle situation n'est ni confirmée, ni infirmée, sur la base des éléments objectifs dont nous disposons. Etant donné que cette existence n'est pas confirmée, on ne peut raisonner de manière certaine selon l'analogie avec le *modèle des deux urnes à incrémentation-objet*. Mais étant donné que cette existence n'est pas infirmée, on ne peut non plus raisonner avec certitude selon l'analogie avec le *modèle de l'urne à incrémentation-temps*.

De plus, la solution décrite ici est également basée sur une étape préliminaire. Il s'agit ainsi de montrer en effet que l'analogie sur laquelle est basé DA ne correspond pas exactement aux modèles qui sont décrits dans l'une ou l'autre de ces solutions. En effet, il résulte tout d'abord des développements précédents que l'analogie sur laquelle est susceptible d'être basé DA prévaut entre la situation correspondant à DA et le modèle des deux urnes à incrémentation-objets, mais non (comme dans la solution de Carter-Leslie) avec le modèle des deux urnes. En second lieu, une seconde analogie sur laquelle est susceptible d'être basé DA prévaut entre la situation correspondant à DA et le modèle des deux urnes à incrémentation-objets, et non (comme dans la solution de Eckhardt-Sobers-Sowers) avec le modèle de l'urne à incrémentation-temps (objets multiples).

La solution basée sur une *troisième voie* qui a été choisie ici repose sur le fait que l'on n'est pas en mesure de déterminer si l'analogie qui prévaut est celle existant entre la situation correspondant à DA et le modèle des deux urnes à incrémentation-objets, ou bien le modèle de l'urne à incrémentation-objet. Ainsi que les développements qui précèdent l'ont montré, il convient en réalité de reformuler ces deux types de solutions en considérant l'alternative suivante: soit (a) l'analogie qui prévaut est celle existant entre la situation correspondant à DA et le modèle des deux urnes à incrémentation-objets; soit (b) l'analogie est celle existant entre la situation correspondant à DA et le modèle de l'urne à incrémentation-objet. Les deux catégories de solutions précédemment mentionnées considèrent que l'une ou l'autre analogie prévaut, de manière *déterminée*. En revanche, la présente solution considère que la situation est *indéterminée*. Et ceci a une importante conséquence. Compte tenu de cette indétermination, on ne peut faire valablement application du modèle des deux urnes à incrémentation-objets. Il ne s'ensuit donc pas un décalage bayésien. Mais on ne peut non plus faire légitimement application du raisonnement qui prévaut dans le modèle de l'urne à incrémentation-objet. Ceci pose la question de savoir quel est finalement le modèle probabiliste qui est véritablement analogue à la situation correspondant à DA. A ce stade, il convient de considérer l'expérience suivante:

Le modèle des deux urnes à incrémentation-objet hypothétique On réalise éventuellement, avec une probabilité $P(\alpha)$, l'expérimentation α correspondant au *modèle des deux urnes* au temps T_0 , avec une urne qui contient n boules. A ce stade, la boule $n^\circ e$ est éventuellement tirée. Une seconde expérience est ensuite mise en œuvre, de manière certaine, et se déroule à des

intervalles de temps *variables*, de T_1 à T_c . Un second dispositif expulse en T_1 une boule n° 1, en T_2 une boule n° 2, ..., en T_n une boule n° n . Maintenant, selon le résultat du *modèle des deux urnes* réalisé en T_0 , le dispositif s'arrête lorsque la boule n° e a été expulsée. Ici, le sujet dans la seconde expérience cherche à évaluer le nombre de boules qui étaient contenues en T_0 dans la première urne. Il formule les deux hypothèses suivantes:

(H1₁₆₅) le nombre de boules contenues dans la première urne est 10

(H2₁₆₆) le nombre de boules contenues dans la première urne est 1000

et s'interroge à l'issue de la seconde expérience sur la révision éventuelle de la probabilité de l'hypothèse (H1).

Ici, à la différence du *modèle des deux urnes à incrémentation-objet*, le déroulement au temps T_0 d'un *modèle des deux urnes* présente un caractère hypothétique. En ce sens, le *modèle des deux urnes à incrémentation-objet* peut être considéré comme un cas particulier du *modèle des deux urnes à incrémentation-objet hypothétique* dans le cas où $P(\alpha) = 1$. De même, le *modèle de l'urne à incrémentation-objet* peut être considéré comme un cas particulier du *modèle des deux urnes à incrémentation-objet hypothétique* avec $P(\alpha) = 0$. On le voit, dans le *modèle des deux urnes à incrémentation-objet hypothétique*, la probabilité $P(\alpha)$ que prenne place au temps T_0 un *modèle des deux urnes* se révèle indéterminée. Et cette propriété se révèle tout à fait analogue à la situation qui prévaut dans la situation humaine correspondant à DA. Car dans la situation humaine, l'existence ou non avant la naissance de l'humanité d'un tirage au sort du type de celui décrit dans l'expérience Dieu et l'humanité-objet demeure indéterminée. On ne dispose pas de certitude dans un sens ou dans l'autre et toute probabilité telle que $P(\alpha) = 1$ ou $P(\alpha) = 0$ ne peut être retenue qu'en l'absence de preuve. On le voit, on a bien là une analogie entre les deux situations. Compte tenu de cette dernière analogie, comment doit-on alors modifier les probabilités a priori? Supposons pour cela que nous ayons la preuve de l'existence avant la naissance de l'humanité d'un tirage au sort tel que celui qui prévaut dans l'expérience Dieu et l'humanité-objet. Dans ce cas, on peut raisonner avec $P(\alpha) = 1$ et la situation est analogue au *modèle des deux urnes à incrémentation-objet*. On calcule alors un décalage bayésien qui correspond à la différence entre les probabilités a posteriori et a priori: $(P'(H1) - P(H1)) * 1$. Supposons à l'inverse que nous disposions de la preuve contraire de l'absence avant la naissance de l'humanité d'une situation du type Dieu et l'humanité-objet. Dans ce cas, on est autorisé à raisonner avec $P(\alpha) = 0$ et la situation se révèle alors analogue au *modèle de l'urne à incrémentation-objet*. On calcule alors un décalage bayésien qui correspond à la différence entre les probabilités a posteriori et a priori, multiplié par la valeur 0: $(P'(H1) - P(H1)) * 0$. On le voit, la formule générale qui permet de calculer le décalage bayésien dans la situation du *modèle des deux urnes à incrémentation-objet hypothétique* est la suivante: $(P'(H1) - P(H1)) * P(\alpha)$. Dans le cas réel qui est celui de la situation humaine correspondant à DA, la probabilité que se déroule au temps T_0 un *modèle des deux urnes* présente un caractère très improbable. Il paraît raisonnable de retenir par exemple à cet égard une probabilité du type: $P(\alpha) = 0,00001$. Dans ce cas, on calcule donc de la manière suivante le décalage bayésien: $P'(H1) - P(H1) * 0,00001$. Il s'ensuit un décalage bayésien dont la valeur est tout à fait négligeable. Et ce dernier résultat doit encore être pondéré par le fait que la classe de référence est choisie arbitrairement et que l'univers pourrait bien être indéterministe. Finalement, il apparaît que cette façon de raisonner se révèle tout à fait différente de celle qui prévaut d'une part dans l'argument de l'Apocalypse, et d'autre part, dans la ligne de raisonnement développée par Eckhardt-Sobers-Sowers. Ainsi, la présente solution constitue une troisième voie par rapport à ces deux types de solutions.

9.7. Variations de DA

A ce stade, il convient de s'intéresser également à un certain nombre de variations de DA et de montrer comment la solution précédente s'applique également à ces dernières. Plusieurs expériences du même type que la version classique de DA et entraînant des conséquences contraires à l'intuition ont notamment été décrites par Bostrom (2001). En particulier, Bostrom introduit les expériences

suivantes: *le Conseil du serpent (Serpent's Advice)* et *Adam paresseux (Lazy Adam)*¹²². Bostrom attribue les conséquences contraires à l'intuition qui résultent de ces expériences à SSA. Bostrom (2001) introduit ainsi l'expérience *le conseil du serpent*:

Eve and Adam, the first two humans, knew that if they gratified their flesh, Eve might bear a child, and if she did, they would be expelled from Eden and would go on to spawn billions of progeny that would cover the Earth with misery. One day a serpent approached the couple and spoke thus: "Pssst! If you embrace each other, then either Eve will have a child or she won't. If she has a child then you will have been among the first two out of billions of people. Your conditional probability of having such early positions in the human species given this hypothesis is extremely small. If, on the other hand, Eve doesn't become pregnant then the conditional probability, given this, of you being among the first two humans is equal to one. By Bayes's theorem, the risk that she will have a child is less than one in a billion. Go forth, indulge, and worry not about the consequences!"¹²³

L'expérience du *Conseil du serpent* est intéressante en ce sens qu'elle met en scène les premiers humains: Eve et Adam. Leur rang de naissance est donc respectivement 1 et 2. Dans cette situation, Eve et Adam peuvent mettre en œuvre un raisonnement de type DA. Il est d'autre part implicite dans la présentation de Bostrom, que les hypothèses sous-jacentes à la situation correspondant à DA sont les suivantes:

(H1₁₆₇) l'humanité comptera au total 2 humains

(H2₁₆₈) l'humanité comptera au total 1000 milliards d'humains

Et il s'ensuit, de manière analogue à la version classique de DA, un décalage bayésien en faveur de l'hypothèse (H1).

L'expérience du *Conseil du serpent* présente quelques différences avec la version classique de DA. En premier lieu, comme on l'a vu, la situation correspondante met en scène les premiers humains: Adam et Eve. On peut considérer ici que les événements E_1 et E_2 consistent respectivement dans le tirage des numéros 1 et 2. Pour simplifier, on pourrait ne considérer ici qu'un seul l'événement E. Cet événement E pourrait être ainsi assimilé au fait qu'Adam effectue le tirage du numéro 1.

En second lieu, Bostrom insiste également sur un autre élément propre à la situation du *Conseil du serpent*. Il s'agit du fait qu'Eve et Adam sont sur le point de s'unir charnellement, entraînant ainsi une forte probabilité d'engendrer une descendance, favorisant ainsi intuitivement l'hypothèse (H2). Eve et Adam se préparent ainsi à agir d'une manière qui rend plus plausible l'hypothèse (H2) par rapport à l'hypothèse (H1). Cependant, une telle particularité ne distingue pas véritablement la situation du *Conseil du serpent* par rapport à la version classique de DA, dans la mesure où dans notre situation actuelle, nous nous trouvons également dans une situation de même nature. A notre époque, pourrait-on dire, ce sont la plupart de ceux qui raisonnent selon le DA qui agissent de manière à engendrer une descendance, alors que dans le *Conseil du serpent*, cela ne concerne qu'Adam et Eve. Ainsi, une telle particularité de la situation du *Conseil du serpent* se révèle-t-elle finalement d'importance minime.

¹²² Bostrom décrit également l'expérience du *Tour de cartes d'Eve (Eve's Card Trick)*, qui s'analyse comme une variation encore plus forte que *Adam paresseux*.

¹²³ Trad.: Eve et Adam, les deux premiers humains, ont su que s'ils s'unissaient, Eve pourrait porter un enfant, et dans ce cas, ils seraient expulsés de l'Eden et se mettraient à engendrer une progéniture comportant des milliards d'individus qui se répandraient sur Terre dans la misère. Un jour, un serpent s'approcha ainsi du couple et lui tint ce discours: Pssst! Si vous vous unissez, alors soit Eve aura un enfant soit elle n'en aura pas. Si elle a un enfant alors vous aurez été dans les deux premiers parmi des milliards de personnes. Votre probabilité conditionnelle de posséder de telles positions temporelles si proches au niveau de l'espèce humaine dans cette hypothèse est extrêmement faible. Si, d'un autre côté, Eve ne devient pas enceinte alors la probabilité conditionnelle, étant donné ceci, que vous soyez parmi les deux premiers humains est égale à un. Par application du théorème de Bayes, le risque qu'elle ait un enfant est de moins d'un sur un milliard. Allez-y donc, ne vous retenez pas, et ne vous inquiétez pas des conséquences!.

On le voit, la différence essentielle du *Conseil du serpent* avec la version classique de DA réside simplement dans le fait que Adam et Eve sont les deux premiers humains. L'événement E consiste ici dans le tirage du numéro 1. On le voit, malgré cette différence, la solution qui vient d'être décrite s'applique également à l'expérience du *Conseil du serpent*. En ce sens, la situation du *Conseil du serpent* s'analyse simplement comme une variation de DA appliquée aux premiers humains. Compte tenu de ces éléments, il apparaît que la solution à DA qui vient d'être décrite s'applique également au *Conseil du serpent*.

Bostrom¹²⁴ présente également une autre variation de DA, plus forte que le *Conseil du serpent*, qu'il nomme *Adam paresseux*:

Assume as before that Adam and Eve were once the only people and that they know for certain that if they have a child they will be driven out of Eden and will have billions of descendants. But this time they have a foolproof way of generating a child, perhaps using advanced *in vitro* fertilization. Adam is tired of getting up every morning to go hunting. Together with Eve, he devises the following scheme: *They form the firm intention that unless a wounded deer limps by their cave, they will have a child.* Adam can then put his feet up and rationally expect with near certainty that a wounded deer – an easy target for his spear – will soon stroll by.¹²⁵

Adam paresseux présente un argument dont la structure est globalement de la même nature que DA. Ici, l'expérience a pour finalité de mettre en évidence des conséquences encore plus contraires à l'intuition que dans la version classique de DA.

Il apparaît ici que l'expérience *Adam paresseux* présente un certain nombre de caractéristiques qui la distinguent de DA. En premier lieu, la situation est de même nature que dans le *Conseil du serpent*, où Adam et Eve constituent les humains n° 1 et 2. Il s'ensuit, comme on l'a vu, une situation où l'événement E s'assimile au fait qu'Adam tire le numéro 1. Ici toutefois, Adam et Eve agissent d'une manière encore plus forte que dans le *Conseil du serpent*, compte tenu des propriétés du mode de reproduction humain qui est utilisé. En effet, le mode de reproduction choisi est encore plus sûr que le mode naturel, car on peut alors considérer comme certain que si Adam et Eve ont un enfant, il en résultera nécessairement une descendance composée de milliards d'humains. Il s'agit donc plus, comme dans l'expérience du *Conseil du serpent*, d'une forte probabilité, mais d'une certitude. Mais ceci s'analyse en définitive comme une variation quantitative, par degré, de l'expérience précédente.

En second lieu, Adam et Eve ont la ferme intention d'avoir un enfant, à moins qu'un événement extrêmement improbable, à savoir l'apparition dans la caverne d'un cerf blessé, ne survienne. La

¹²⁴ Bostrom donne encore une autre variation, également plus forte que DA, qui est à *Eve's Card Trick*: "One morning, Adam shuffles a deck of cards. Later that morning, Eve, having had no contact with the cards, decides to use her willpower to retroactively choose what card lies top. She decides that it shall have been the dame of spades. In order to ordain this outcome, Eve and Adam form the firm intention to have a child unless the dame of spades is top. They can then be virtually certain that when they look at the first card they will indeed find the dame of spades." Trad.: "Un matin, Adam mélange un jeu de cartes. Un peu plus tard ce même matin, Eve, qui n'a eu aucun contact avec les cartes, décide d'employer son intellect pour déterminer rétroactivement quelle carte se trouve au-dessus. Elle décide qu'il s'agit de la dame de Pique. Afin de rendre solennel ce résultat, Eve et Adam forment l'intention ferme d'avoir un enfant à moins que la dame de Pique se trouve au-dessus. Ils peuvent alors être pratiquement certains que lorsqu'ils regarderont la première carte, ils trouveront en effet la dame de Pique."

¹²⁵ Trad.: Supposez comme précédemment qu'Adam et Eve étaient par le passé les seules personnes qui existaient et qu'ils savaient de manière certaine que s'ils avaient un enfant, ils seraient conduits hors de l'Eden et auraient des milliards de descendants. Mais cette fois ils disposent d'un moyen infaillible de faire un enfant, comme d'utiliser une technique de fertilisation *in vitro* avancée. Adam est fatigué de se lever chaque matin pour aller à la chasse. Avec Eve, il conçoivent l'arrangement suivant: *Ils forment la ferme intention, qu'à moins qu'un cerf blessé arrive en boitant dans leur caverne, ils auront un enfant.* Adam peut maintenant se détendre les jambes et prévoir de manière rationnelle avec une quasi-certitude qu'un cerf blessé - une cible facile pour sa lance - va bientôt venir déambuler.

conclusion qui résulte de l'expérience est que compte tenu du décalage bayésien en faveur de l'hypothèse (H1) consécutif à l'événement E, un cerf blessé va apparaître dans la caverne. Une telle conclusion est encore plus forte que celle qui résulte de DA.

La situation qui est celle d'*Adam paresseux* apparaît essentiellement en définitive comme une variation de DA-objet. En effet, dans la version classique de DA-objet, on se situe dans un univers déterministe et le nombre total d'humains qui composeront finalement l'humanité se trouve fixé au moment où survient l'événement E. Dans le cas d'*Adam paresseux*, la situation apparaît différente puisque le nombre total d'humains est fixé à des milliards, à moins qu'un cerf blessé ne surgisse dans la caverne. Ceci s'assimile à une situation où le nombre total d'humains qui composera l'humanité est fixé, avec une quasi certitude. Toutefois, le fait qu'il existe une chance infime qu'un cerf blessé apparaisse dans la caverne suffit à transformer la situation correspondante en une situation qui n'est pas entièrement déterministe. Dans ce contexte, la situation correspondant à *Adam paresseux* s'assimile à une variation de DA qui prend place dans un univers partiellement indéterministe.

Chapitre 10. Le paradoxe de la chambre d'exécution

10.1. Le problème

L'expérience de la chambre d'exécution a été décrite à l'origine par Leslie (1992, p.538) en tant qu'objection à l'Argument de l'Apocalypse. Leslie indique que cette expérience a trouvé son origine dans une correspondance avec David Lewis¹²⁶. Leslie présente ainsi l'objection à DA fondée sur l'expérience de la *chambre d'exécution*¹²⁷:

Thrust into a room, you are assured that 90% of those who enter it will be shot. Panic! But you then learn that you will leave the room alive unless double-six is thrown, first time, with two dice. How is this compatible with the assurance that 90% will be shot? Successive batches of people thrust into the room are successively larger, exponentially, so that the forecast "90% will be shot" will be confirmed when double-six is eventually thrown. If knowing this, and knowing also that the dice falls would be utterly unpredictable event to a Laplacean demon who knew everything about the situation when the dice were thrown, then shouldn't one's panic vanish?¹²⁸

D'autre part, Leslie mentionne également dans *The End of the World* (p.251-6) l'expérience de la chambre d'exécution en répondant une nouvelle fois aux objections à DA basées sur l'expérience de la chambre d'exécution. Ainsi, le contexte dans lequel Leslie mentionne l'expérience de la chambre d'exécution est celui de la réfutation d'une objection à DA. L'expérience de la chambre d'exécution est en effet basée sur une croissance exponentielle de la population et peut ainsi suggérer une situation où DA se trouve réfuté.

Cependant, le problème inhérent à l'expérience de la chambre d'exécution a ensuite été repris et présenté par Bartha et Hitchcock (1999) en tant que paradoxe indépendant. Bartha et Hitchcock en effet ne décrivent plus l'expérience de la chambre d'exécution en tant qu'objection à DA, mais en tant que situation intrinsèquement paradoxale. Les auteurs décrivent ainsi le *paradoxe de la chambre d'exécution* comme un problème en relation avec DA, mais qui s'en distingue toutefois. En effet, DA est un problème qui conduit à une conclusion contraire à l'intuition, alors que le *paradoxe de la chambre d'exécution* entraîne une conclusion véritablement paradoxale. De fait, il s'avère que certains auteurs ont finalement accepté les conclusions qui résultent de DA. Mais une telle démarche se révèle impossible à suivre avec le *paradoxe de la chambre d'exécution*, qui présente une nature paradoxale. Bartha et Hitchcock décrivent ainsi le *paradoxe de la chambre d'exécution* comme un paradoxe autonome, qui peut éventuellement être disjoint de l'Argument de l'Apocalypse. Bartha et Hitchcock décrivent ainsi le paradoxe. Une pièce (la chambre d'exécution, *shooting-room*) contient successivement 1, 9, 90, 900, 9000, ..., $9 \cdot 10^n$ personnes. On lance alors deux dés comportant six faces. Si le résultat n'est pas un double six, toutes les personnes qui se trouvaient dans la chambre d'exécution peuvent en sortir, vivantes. Par contre, si le résultat est un double six, alors toutes les personnes se trouvant dans la chambre d'exécution sont éliminées. Le paradoxe réside ici dans le fait que deux raisonnements peuvent semble-t-il être valablement tenus, alors qu'il en résulte des conclusions contradictoires. On peut en effet raisonner tout d'abord de la manière suivante (I): si je me trouve dans l'une des pièces, je pense que la probabilité que je sois éliminé est égale à la probabilité de

¹²⁶ Cf. n. 6 p. 539.

¹²⁷ *Shooting-room objection*.

¹²⁸ Trad.: Poussé dans une salle, on vous assure que 90% de ceux qui y sont entrés seront tués. Panique! Mais vous apprenez alors que vous quitterez la salle vivant, à moins qu'un double-six ne sorte, la première fois, lorsque deux dés seront lancés. Comment cela est-il compatible avec l'assurance que 90% seront tués? Les séries successives de personnes poussées dans la salle sont successivement plus grandes, de manière exponentielle, de sorte que la prévision "90% seront tués" sera confirmée lorsqu'un double-six sera par la suite tiré. Si vous savez cela, et vous savez également que le résultat du lancer des dés seraient un événement tout à fait imprévisible même pour un démon de Laplace qui saurait tout au sujet de la situation lorsque les dés seraient jetés, alors la panique ne devrait-elle pas disparaître?

tirer un double six, c'est-à-dire 1/36. On peut observer ici que le raisonnement choisi par Leslie est le raisonnement (I), puisqu'il conclut (1992, p. 538¹²⁹) que la panique initiale de celui qui va participer à l'expérience de la chambre d'exécution devrait finalement disparaître. Mais de manière alternative, on peut tenir le raisonnement (II) suivant: la probabilité que je sois éliminé est égale 0,9 (plus précisément cette probabilité est $\geq 0,9$ pour tenir compte du cas 1) puisque 90/100 des personnes entrées dans la chambre d'exécution y périront finalement. Le paradoxe naît ici du fait que deux raisonnements apparemment valides concluent respectivement aux conclusions que la probabilité d'être éliminé dans la chambre d'exécution est respectivement de 1/36 et de 9/10.

10.2. Modélisation dans le n-univers spatio-temporel à objets multiples ($\Omega\alpha^*T^*L$)

A ce stade, il convient de s'attacher à modéliser l'expérience de la chambre d'exécution dans le cadre des n-univers. Quel est donc le n-univers qui correspond à la situation de l'expérience de la chambre d'exécution? En premier lieu, il s'avère que la situation correspondant à l'expérience comporte $n+1$ positions temporelles T_0, T_1, \dots, T_n , en considérant ici que T_0 correspond au début de l'expérimentation. Par conséquent, l'expérience de la chambre d'exécution prend place dans un univers temporel. Le n-univers correspondant est donc un $\Omega T \dots$. En second lieu, l'expérience comporte trois positions spatiales: L_0 (la chambre d'exécution), L_1 (l'emplacement ou l'antichambre où se trouvent les objets au début de l'expérience et L_2 (l'emplacement où se trouvent les objets après être sortis de la chambre d'exécution). Elle prend donc place dans un n-univers comportant une variable spatiale. On se situe donc dans un $\Omega L \dots$. En troisième lieu, il apparaît que l'expérience concerne de nombreux humains. Elle correspond donc à un n-univers à objets multiples, à savoir un $\Omega\alpha \dots$. En outre, chaque humain peut exister à plusieurs positions temporelles, et peut en particulier se révéler persistant par rapport au temps: s'il existe en T_5 en L_0 (la chambre d'exécution), c'est qu'il a existé en T_1, T_2, T_3, T_4 en L_1 . La variable temporelle est donc à démultiplication. Enfin, plusieurs humains peuvent se trouver à un emplacement donné (L_0, L_1 ou L_2) à une position temporelle donnée. Ainsi, les objets sont en relation *plusieurs* avec le critère du temps. Par conséquent, l'expérience de la chambre d'exécution prend place dans un $\Omega\alpha^*T^*L$, un n-univers à objets multiples, temporel et localisé, où les objets peuvent exister à plusieurs positions temporelles et sont en relation *plusieurs* avec le critère temporel. En reprenant ainsi les données qui sont celles de la version du *paradoxe de la chambre d'exécution* décrite par Bartha et Hitchcock (1999)¹³⁰, il s'ensuit la situation suivante:

- en T_0 , tous les objets se trouvent en L_1
- en T_1 , 1 objet (o_1) se trouve en L_0
- en T_2 , 9 objets (o_2 à o_{10}) se trouvent en L_0 et 1 objet (o_1) en L_2
- en T_3 , 90 objets (o_{11} à o_{100}) se trouvent en L_0 et 10 objets (o_1 à o_{10}) en L_2
- en T_4 , 900 objets (o_{101} à o_{1000}) se trouvent en L_0 et 100 objets (o_1 à o_{100}) en L_2
- ...
- en T_i , $9 \times 10^{i-2}$ objets ($o_{10^{i-2}}$ à $o_{10^{i-1}}$) se trouvent en L_0 et 10^{i-2} objets (o_1 à $o_{10^{i-2}}$) en L_2

10.3. Modélisation dans d'autres n-univers ($\Omega\alpha^*T$)

On peut s'interroger pour savoir si le *paradoxe de la chambre d'exécution* ne peut pas être modélisé dans d'autres n-univers, en particulier dans des sous-univers du $\Omega\alpha^*T^*L$. A ce stade, on peut se demander si l'expérience de la chambre d'exécution ne peut pas prendre place dans un n-univers dépourvu de variable spatiale. Ceci permettrait de modéliser plus simplement l'expérience de la chambre d'exécution dans un $\Omega\alpha^*T^*$. Ici, la question est de savoir si l'on peut considérer valablement

¹²⁹ Cf. p. 538: "I agree that one's panic ought to vanish in the situation described, despite the fact that most people thrust into the room would get bullets." Trad.: "Je conviens que la panique qui habite chacun devrait disparaître dans la situation qui est décrite, malgré le fait que la plupart des personnes poussées dans la salle recevraient des balles."

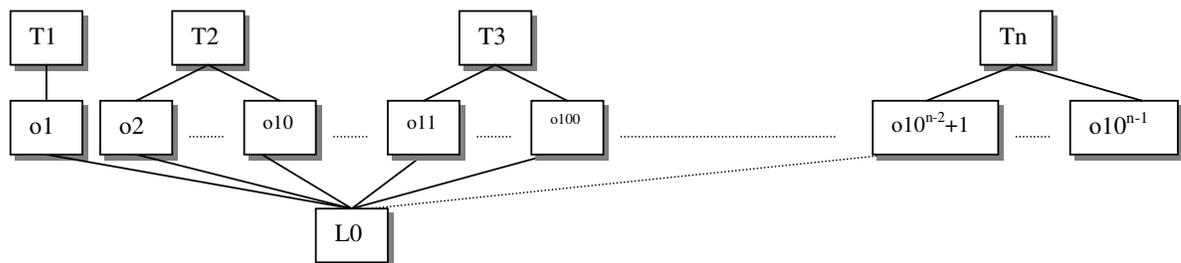
¹³⁰ Cf. §2: "(...) the groups are summoned to the room in increasingly large sizes: 1, 9, 90, 900, 9000, 90000, ...". Trad.: "(...) les groupes sont appelés à la salle avec des tailles de plus en plus grandes: 1, 9, 90, 900, 9000, 90000."

une seule position spatiale (L_0) au lieu de trois. Il apparaît en effet que l'on peut aisément se dispenser de L_2 , qui correspond à l'emplacement où se trouvent les objets après être sortis de la chambre d'exécution. En effet, on pourrait aussi bien concevoir que les objets, une fois sortis de la chambre d'exécution, sont ensuite replacés en L_1 .

La question reste toutefois posée de savoir si l'on peut se dispenser de la position spatiale L_1 , qui correspond au lieu où se trouvent les objets au début de l'expérience. Essayons donc de modéliser l'expérience de la chambre d'exécution dans un $\Omega\alpha^*T^*$. Ici, on considère toujours comme précédemment un n-univers à objets multiples, comportant une variable temporelle. Mais ici, l'expérience prend place dans un lieu unique, la chambre d'exécution. On se situe donc dans un n-univers présentant une constante locale (L_0). De plus, les objets sont en relation *plusieurs* avec le critère du temps. Dans la modélisation précédente de l'expérience de la chambre d'exécution dans le $\Omega\alpha^*T^*L$, chaque objet était également persistant par rapport au temps. Cependant, il apparaît ci que compte tenu des conditions nouvelles de l'expérience, à savoir que la modélisation s'effectue par rapport à un lieu unique L_0 , une telle caractéristique n'est plus indispensable. Car ici, le scénario correspondant est le fait que des objets sont créés, à un moment donné T_i en L_0 et qu'ils existent uniquement en T_i . L'expérience de la chambre d'exécution prend ainsi place dans un $\Omega\alpha^*T$, un n-univers à objets multiples comportant une constante spatiale ainsi qu'une variable temporelle sans démultiplication, où les objets sont en relation *plusieurs* avec le critère temporel. Ici, le nombre d'objets est variable en fonction du temps. On considère donc que les objets soit existent en T_i dans la chambre d'exécution soit n'existent pas. En reprenant les données qui sont celles de la version du *paradoxe de la chambre d'exécution* décrite par Bartha et Hitchcock, il s'ensuit un scénario où les objets sont créés, au fur et à mesure, dans l'ordre suivant:

- en T_1 , 1 objet (o_1) se trouve en L_0
- en T_2 , 9 objets (o_2 à o_{10}) se trouvent en L_0
- en T_3 , 90 objets (o_{11} à o_{100}) se trouvent en L_0
- en T_4 , 900 objets (o_{101} à o_{1000}) se trouvent en L_0
- en T_5 , 9000 objets (o_{1001} à o_{10000}) se trouvent en L_0
- ...
- en T_i , $9 \times 10^{i-2}$ objets ($o_{10^{i-2}}$ à $o_{10^{i-1}}$) se trouvent en L_0

jusqu'à ce qu'un double-six soit lancé. Dans ce dernier cas, l'expérience s'arrête. On s'intéresse alors au nombre d'objets n_i qui se trouvent dans la chambre d'exécution. Le nombre total n d'objets est: $n = 10^{\max-1}$ et les objets sont numérotés de 1 à $10^{\max-1}$. On le voit finalement, on peut également modéliser l'expérience de la chambre d'exécution dans le cadre d'un n-univers plus simple que le $\Omega\alpha^*T^*L$, à savoir un $\Omega\alpha^*T$.



10.4. L'expérience de la chambre d'exécution en tant qu'objection à DA

Dans ce qui suit, ma préoccupation principale sera de m'attacher à proposer une solution permettant de résoudre le *paradoxe de la chambre d'exécution*. Cependant, je m'attacherai préalablement à

considérer l'expérience de la chambre d'exécution en tant qu'objection à DA¹³¹. Si l'on se place exclusivement du point de vue de DA, en considérant que l'expérience de la chambre d'exécution constitue une objection à DA, ainsi que l'appréhendait Leslie à l'origine, on peut observer que cette objection ne suffit pas à bloquer DA. L'objection fondée sur l'expérience de la chambre d'exécution échoue en effet car DA doit être considéré d'un point de vue plus général que celui de la version classique basée sur le nombre d'humains composant l'humanité qui est DA-objets. On l'a vu en effet, il convient de prendre en considération la version bi-dimensionnelle de DA puisque le raisonnement qui sous-tend DA classique s'applique également au critère temporel sous la forme de DA-temps. Et lorsqu'on prend en compte la version bi-dimensionnelle de DA, il apparaît que l'objection à DA basée sur l'expérience de la chambre d'exécution ne vaut que pour DA-objets. En effet, elle ne s'applique pas à DA-temps qui conserve alors toute sa force. Ceci est illustré par le scénario suivant, qui permet de mettre en œuvre à la fois DA-objets et DA-temps, avec les données qui sont celles d'une progression exponentielle de la population que comporte la *chambre d'exécution*. On considère ainsi que la population humaine évolue dans les proportions suivantes: en T_1 , il existe 1 humain; en T_2 , 9 humains; en T_3 , 90 humains; en T_4 , 900 humains; ...; en T_n , $9 \times 10^{n-2}$ humains. On considère dans l'expérience que l'on connaît cette règle de progression. Considérons également que l'on se trouve au temps T_{11} , avec 9×10^9 humains. On est donc en présence des deux versions de DA: DA-objets et DA-temps. Considérons en premier lieu DA-objets. DA-objets (DA- $\Omega\alpha T * C$) prend place dans le n -univers à objets multiples, coloré et temporel. On raisonne alors ainsi. Je suis le 50 milliardième humain et je fais les deux hypothèses suivantes:

- (H1₁₆₉) l'humanité comptera au total 9×10^{10} humains au cours de son histoire
- (H2₁₇₀) l'humanité comptera au total 9×10^{18} humains au cours de son histoire

Compte tenu du fait que je me situe dans la tranche comprise entre 9 milliards et 90 milliards d'humains, je conclus donc, en vertu du modèle à minoration, à un décalage bayésien en faveur de l'hypothèse (H1).

Considérons en second lieu DA-temps. DA-temps (DA- $\Omega T * C$) prend place dans le n -univers à objets unique, coloré et temporel. On raisonne alors de la manière suivante. Je me trouve dans la 7ème période temporelle de l'histoire de l'humanité. Je fais les deux hypothèses suivantes:

- (H1₁₇₁) l'humanité connaîtra 8 périodes temporelles au cours de son histoire
- (H2₁₇₂) l'humanité connaîtra 20 périodes temporelles au cours de son histoire

Compte tenu du fait que je me situe actuellement dans la 7ème période temporelle de l'histoire de l'humanité, je conclus alors à un décalage bayésien en faveur de l'hypothèse (H1). On le voit, les hypothèses qui résultent des deux versions de DA se révèlent tout à fait compatibles. Il en résulte en effet la version conjointe suivante:

- (H1₁₇₃) l'humanité connaîtra 8 périodes temporelles et comptera au total 9×10^7 humains au cours de son histoire
- (H2₁₇₄) l'humanité connaîtra 20 périodes temporelles et comptera au total 9×10^{18} humains au cours de son histoire

On le voit, on se trouve bien ici dans une situation qui permet à la fois à DA-objets et DA-temps de prendre place. Ici, DA-objets peut effectivement conduire à une situation où l'objection basée sur l'expérience de la chambre d'exécution peut être opposée. Mais même si cette objection devait s'avérer effective, il s'avère que DA resurgit alors sous la forme de DA-temps. Et ceci montre que la progression exponentielle de la population ne constitue pas une objection qui suffit à bloquer DA. Ainsi, l'expérience de la chambre d'exécution ne constitue pas une objection qui peut être opposée valablement à DA.

¹³¹ Leslie présente également de solides arguments conduisant au rejet de l'expérience de la chambre d'exécution en tant qu'objection à DA.

10.5. Analyse

Il convient maintenant de s'attacher à présenter une analyse du *paradoxe de la chambre d'exécution*. Avant cela, il s'avère toutefois nécessaire de considérer préalablement plusieurs variations de ce dernier. Dans l'énoncé du *paradoxe de la chambre d'exécution*, il n'est pas précisé notamment si, me trouvant dans la chambre d'exécution, je connais ou non ma position temporelle T_i . De même, il n'est pas mentionné dans l'énoncé si j'ai ou non connaissance du nombre de personnes n_i qui sont dans la chambre d'exécution au moment T_i où je m'y trouve. Ainsi, on peut s'interroger pour savoir si le fait de connaître ma position temporelle T_i ou bien le nombre de personnes n_i qui sont dans la chambre d'exécution au moment T_i où je m'y trouve, a une incidence sur le problème posé par le *paradoxe de la chambre d'exécution*. S'agit-il alors de variations du paradoxe ou bien ce dernier disparaît-il si on prend en compte ces variations de l'énoncé? Considérons ainsi le fait que je me trouve dans la chambre d'exécution en T_i . A ce stade, je suis à même de déterminer avec précision le nombre de personnes qui se trouvent dans la chambre d'exécution. Car je sais que la progression de la population qui occupe la chambre d'exécution est exponentielle. On dénombre en effet en $T_1, T_2, T_3, T_4, \dots, T_n$ respectivement 1, 9, 90, 900, ..., $9 \times 10^{n-2}$ personnes dans la chambre d'exécution. Dès lors, sachant que ma position temporelle est T_i , je peux en déduire que le nombre de personnes qui se trouvent dans la chambre d'exécution est $9 \times 10^{i-2}$. Et de même, connaissant le nombre de personnes n_i qui se trouvent dans la chambre d'exécution, je peux déduire i et donc T_i . Ainsi, dans ces variations, si je connais ma position temporelle T_i alors je connais n_i et d'autre part si je connais n_i alors je connais ma position temporelle T_i . Ainsi, le fait de connaître ma position temporelle T_i ou bien le nombre de personnes n_i qui sont dans la chambre d'exécution au moment où je m'y trouve, a pour conséquence que je connais à la fois T_i et n_i . En outre, dans chacune de ces hypothèses, il apparaît que le problème posé par le *paradoxe de la chambre d'exécution* est toujours présent. On le voit ainsi, le fait que je connaisse ou non ma position temporelle T_i ou bien le nombre de personnes n_i qui se trouvent dans la chambre d'exécution, se révèle indifférent, car il en résulte toujours une variation correspondante du *paradoxe de la chambre d'exécution*.

Considérons maintenant la modélisation de l'expérience de la chambre d'exécution dans le $\Omega \alpha * T * L$. Dans ce cas, rappelons que tous les objets qui disparaîtront finalement dans la chambre d'exécution se trouvent réunis en L_1 . Dans le *paradoxe de la chambre d'exécution*, la contradiction apparaît entre d'une part un raisonnement de type (I) qui conclut à une probabilité de 1/36 et un raisonnement de type (II) qui conduit à une probabilité de 9/10. Si l'on effectue un parallèle avec la version bi-dimensionnelle de DA, on est amené aux constatations suivantes. En premier lieu, la probabilité 1/36 est liée à la position temporelle de l'humanité et donc au critère *temporel*. Cette dernière probabilité apparaît donc bien liée au tirage d'un numéro de segment de temps. En second lieu, la probabilité 9/10 est liée au rang parmi les humains et donc au critère *objets*.

Quel est donc le raisonnement correct qui peut être tenu dans la situation qui est celle de l'expérience de la chambre d'exécution? Il apparaît ici que le raisonnement correct est le suivant. Je me trouve dans la chambre d'exécution en T_i . Là, je sais que le double dé sera lancé et que je disparaîtrai uniquement dans l'hypothèse où un double six est tiré. Par conséquent, la probabilité que je disparaisse est égale à 1/36. Je sais également que 90/100 de ceux qui sont entrés dans la chambre d'exécution auront finalement disparu. Mais ceci se révèle tout à fait compatible avec le fait que la probabilité que je disparaisse est égale à 1/36. En effet, une telle situation est vérifiée dans le contexte d'une progression exponentielle du nombre d'occupants de la chambre d'exécution. On le voit finalement, le raisonnement retenu est le raisonnement de type (I). Ainsi, on peut conclure que c'est la probabilité de 1/36 qui prévaut dans l'expérience de la chambre d'exécution, et donc le raisonnement (I)¹³².

Dans ce contexte, qu'est-ce qui est donc fallacieux dans le raisonnement de type (II)? Ce dernier raisonnement est basé sur le fait que l'argument suivant est utilisé:

- (175) 0,90 des gens vont disparaître
- (176) ∴ la probabilité que je disparaisse est de 0,90

¹³² Une telle solution n'est pas nouvelle. Il s'agit essentiellement de l'analyse présentée par Leslie (1992, p. 538).

Mais l'inférence de (175) à (176) n'est pas valide (il est à noter que Bartha et Hitchcock concluent de manière identique, mais sur un fondement toutefois différent¹³³) car elle méconnaît la progression exponentielle en fonction de la position temporelle de la population qui se trouve dans la *chambre d'exécution*.

¹³³ Cf. (1999, §2): "It is instructive to see what is wrong with the inference from '90% of those who enter the room will die' to 'my probability of dying (given that I have entered the room) is .9.' This inference rests on the crucial assumption that we are not special with respect to the order in which we are summoned to the room." Trad.: "Il est instructif de voir ce qui est erroné avec l'inférence selon laquelle '90% de ceux qui entrent dans la salle mourront' à 'ma probabilité de mourir (étant donné que je suis entré dans la salle) est de 0,9.' Cette inférence repose sur l'hypothèse cruciale selon laquelle nous ne sommes pas spéciaux en ce qui concerne l'ordre dans lequel nous sommes appelés à la salle."

Chapitre 11. Le paradoxe de Goodman

11.1. Le problème

Le paradoxe de Goodman¹³⁴ (*Goodman's Paradox*, soit GP dans ce qui suit) a été énoncé par Nelson Goodman (1946). Goodman expose son paradoxe de la manière suivante¹³⁵. Soit une urne contenant 100 boules. Une boule est tirée chaque jour dans l'urne, durant 99 jours, jusqu'à aujourd'hui. A chaque fois, la boule extraite de l'urne est rouge. Intuitivement, on s'attend à ce que la 100ème boule tirée soit également rouge. Cette prédiction est basée sur la généralisation selon laquelle toutes les boules dans l'urne sont rouges. Cependant, si on considère la propriété S "tiré avant aujourd'hui et rouge ou tiré après aujourd'hui et non rouge", on constate que cette propriété est également vérifiée par les 99 instances déjà observées. Mais la prédiction qui en résulte cette fois, basée sur la généralisation selon laquelle toutes les boules sont S, est que la 100ème boule sera non rouge. Et ceci est contraire à la conclusion précédente, qui est elle-même pourtant conforme à notre intuition¹³⁶.

Goodman exprime ainsi GP à l'aide d'une induction énumérative. Et on peut modéliser GP en termes de SR (*straight rule*). Si l'on prend (D) pour la définition du prédicat "rouge", (I) pour l'énumération des instances, (H) pour la généralisation en résultant, et (P) pour la prédiction correspondante, on a alors :

- (D₁₇₇) R = rouge
- (I₁₇₈) Rb₁·Rb₂·Rb₃·...·Rb₉₉
- (H₁₇₉) Rb₁·Rb₂·Rb₃·...·Rb₉₉·Rb₁₀₀
- (P₁₈₀) ∴ Rb₁₀₀

Et de même, avec le prédicat S :

- (D*₁₈₁) S = rouge et tiré avant T ou non rouge et tiré après T
- (I*₁₈₂) Sb₁·Sb₂·Sb₃·...·Sb₉₉
- (H*₁₈₃) Sb₁·Sb₂·Sb₃·...·Sb₉₉·Sb₁₀₀ qui équivaut à :
- (H*₁₈₄) Rb₁·Rb₂·Rb₃·...·Rb₉₉·~Rb₁₀₀

¹³⁴ Ce chapitre est tiré de mon 'Une solution pour le paradoxe de Goodman' publié dans *Dialogue* (2001), avec quelques modifications mineures.

¹³⁵ Avec quelques adaptations mineures.

¹³⁶ Cf. Goodman "A Query On Confirmation", p. 383: "Suppose we had drawn a marble from a certain bowl on each of the ninety-nine days up to and including VE day and each marble drawn was red. We would expect that the marble drawn on the following day would also be red. So far all is well. Our evidence may be expressed by the conjunction " $Ra_1 \cdot Ra_2 \cdot \dots \cdot Ra_{99}$ " which well confirms the prediction Ra_{100} ." But increase of credibility, projection, "confirmation" in any intuitive sense, does not occur in the case of every predicate under similar circumstances. Let "S" be the predicate "is drawn by VE day and is red, or is drawn later and is non-red." The evidence of the same drawings above assumed may be expressed by the conjunction " $Sa_1 \cdot Sa_2 \cdot \dots \cdot Sa_{99}$ ". By the theories of confirmation in question this well confirms the prediction " Sa_{100} "; but actually we do not expect that the hundredth marble will be non-red. " Sa_{100} " gains no whit of credibility from the evidence offered." Trad.: "Supposez que nous avons tiré une bille d'une cuvette donnée durant chacun des quatre-vingt-dix-neuf jours qui précèdent et qui comprennent le jour de VE et que chaque bille tirée était rouge. Nous nous attendons alors à ce que la bille tirée le jour suivant soit également rouge. Jusqu'ici tout se présente bien. Les faits matériels sur lesquels nous nous basons peuvent être exprimés par la conjonction " $Ra_1 \cdot Ra_2 \cdot \dots \cdot Ra_{99}$ " qui confirme bien la prévision Ra_{100} ." Mais l'augmentation de la crédibilité, la projection, la "confirmation", quelque intuition qu'on en ait, ne se produisent pas pour tout prédicat dans des circonstances tout à fait semblables. Soit "S" le prédicat "est tiré avant le jour de VE et est rouge, ou est tiré plus tard et est non-rouge." Les faits matériels qui résultent des mêmes tirages que précédemment peuvent être exprimés par la conjonction " $Sa_1 \cdot Sa_2 \cdot \dots \cdot Sa_{99}$ ". Et de par les théories de la confirmation en question ceci confirme bien la prévision " Sa_{100} "; mais en réalité, nous n'escomptons pas que la centième bille sera non-rouge. " Sa_{100} " ne gagne aucun brin de crédibilité supplémentaire avec les faits matériels dont on dispose."

(P*₁₈₅) ∴ Sb₁₀₀ c'est-à-dire finalement :

(P'*₁₈₆) ∴ ~Rb₁₀₀

Le paradoxe réside ici dans le fait que les deux généralisations (H) et (H*) conduisent respectivement à des prédictions (P) et (P*) qui sont contradictoires. Intuitivement, l'application de SR à (H*) paraît erronée.

Goodman donne aussi dans *Fact, Fiction and Forecast* une version légèrement différente de son paradoxe, appliquée cette fois aux émeraudes¹³⁷. Cette forme est très bien connue et est basée sur le prédicat "grue" = vert et observé avant T ou non vert et observé après T.

Le prédicat S utilisé dans Goodman (1946) présente avec "grue", une *structure* commune. Soient P et Q deux prédicats, cette structure correspond à la définition : (P et Q) ou (~P et ~Q). Dans ce qui suit, on désignera par *grue* un prédicat présentant cette structure particulière, sans distinguer selon que la forme spécifique utilisée est celle de Goodman (1946) ou (1954).

11.2. La dualité unification/différenciation

J'ai devant moi des instances. Dois-je les décrire en mettant l'accent sur leurs différences ? Ou bien dois-je les décrire en insistant sur leurs propriétés communes ? Je peux procéder d'une manière ou de l'autre. Mettre l'accent sur les différences entre les instances, c'est opérer par *différenciation*. A l'inverse, mettre en évidence leurs propriétés communes, c'est procéder par *unification*. Il convient de s'intéresser tout à tour à chacun de ces deux modes d'opérer.

Soient les 100 boules composant l'urne de Goodman (1946). Considérons tout d'abord le cas où mon intention est de mettre l'accent sur les différences entre les instances. Là, une option est d'appréhender le moment particulier et unique, où chacune d'elles est extraite de l'urne. On considère alors les prédicats : *rouge et tiré le jour 1, rouge et tiré le jour 2, ..., rouge et tiré le jour 99*. On a ainsi 99 prédicats différents. Mais ceci interdit d'appliquer SR, qui nécessite un seul et même prédicat. Qu'est-ce donc que distinguer selon le moment où chaque boule est tirée ? C'est mettre l'accent sur une différence essentielle entre chacune des boules, fondée sur le critère du temps. On individualise ainsi chaque boule, et il en résulte autant de prédicats différents : tiré en T₁, tiré en T₂, ..., tiré en T₉₉. Ceci empêche ensuite tout mouvement inductif par application de SR. En effet, on ne dispose pas alors d'une propriété commune pour permettre l'induction et appliquer SR. Ici, la cause du problème réside dans le fait d'avoir réalisé une *différenciation extrême*.

De manière alternative, je peux également procéder par *différenciation* en opérant une mesure extrêmement précise¹³⁸ de la longueur d'onde de la lumière définissant la couleur de chacune des boules. J'obtiendrai alors une mesure de longueur d'onde unique pour chacune des boules de l'urne. Ainsi, j'ai 100 boules devant moi, et je connais avec précision la longueur d'onde de la lumière de 99 d'entre elles. Les boules ont respectivement une longueur d'onde de 722,3551 nm, 722,3643 nm, 722,3342 nm, 722,3781 nm, etc. Je dispose dès lors de 99 prédicats distincts P3551, P3643, P3342, P3781, etc. Mais je suis alors dans l'impossibilité d'appliquer SR, qui exige un seul prédicat. Ici aussi,

¹³⁷ *Ibid.*, p. 73-4 : "Suppose that all emeralds examined before a certain time *t* are green. At time *t*, then, our observations support the hypothesis that all emeralds are green; and this is in accord with our definition of confirmation. [...] Now let me introduce another predicate less familiar than "green". It is the predicate "grue" and it applies to all things examined before *t* just in case they are green but to other things just in case they are blue. Then at time *t* we have, for each evidence statement asserting that a given emerald is green, a parallel evidence statement asserting that that emerald is grue." Trad.: "Supposez que toutes les émeraudes examinées avant un certain temps *t* soient vertes. Au temps *t*, alors, nos observations soutiennent l'hypothèse que toutes les émeraudes sont vertes; et ceci est en accord avec notre définition de la confirmation [...] Maintenant laissez-moi introduire un autre prédicat moins familier que "vert". C'est le prédicat "grue" et il s'applique à toutes les choses examinées avant *t* dans le cas où elles sont vertes mais aux autres choses dans le cas où elles sont bleues. Alors au temps *t* nous avons, pour chaque fait matériel rapportant qu'une émeraude donnée est verte, un fait matériel rapportant parallèlement que l'émeraude est grue."

¹³⁸ Par exemple avec une précision de 10⁻⁴ nm.

les propriétés communes font défaut pour pouvoir mettre en oeuvre le processus inductif. De la même manière que précédemment, il s'avère ici que j'ai réalisé une *différenciation extrême*.

Que se passe-t-il maintenant si je procède exclusivement par *unification* ? Considérons le prédicat R correspondant à "rouge ou non rouge". On tire 99 boules rouges avant le temps T. Elles sont toutes R. On prédit que la 100ème boule sera R après T, c'est-à-dire rouge ou non rouge. Mais cette forme d'induction n'apporte ici aucune information. La conclusion produite est vide d'information. On appellera *induction vide* ce type de situation. Dans ce cas, on observe que le processus d'unification des instances par la couleur a été réalisé de manière radicale, en annihilant à cet égard, toute démarche de différenciation. La cause du problème réside ainsi dans la mise en oeuvre d'un processus d'*unification extrême*.

Si l'on se place du point de vue de la *couleur*, il apparaît que chacun des cas envisagés précédemment fait appel à une taxinomie différente des couleurs. Ainsi, il est fait usage successivement :

- de notre taxinomie usuelle des couleurs basée sur 9 prédicats : violet, indigo, bleu, vert, jaune, orangé, rouge, blanc, noir
- d'une taxinomie fondée sur une mise en relation des longueurs d'onde des couleurs avec l'ensemble des nombres réels (*taxinomie réelle*)
- d'une taxinomie fondée sur un prédicat unique (*taxinomie à taxon unique*) : rouge ou non rouge

Or il s'avère que chacun de ces trois cas peut être replacé dans une perspective plus générale. En effet, de multiples taxinomies des couleurs sont susceptibles d'être utilisées. Et celles-ci peuvent être ordonnées de la plus grossière (taxinomie à taxon unique) à la plus fine (taxinomie réelle), de la plus *unifiée* à la plus *différenciée*. On a notamment la hiérarchie suivante des taxinomies :

- TAX₁ = {rouge ou non rouge} (taxinomie à taxon unique)
- TAX₂ = {rouge, non rouge} (taxinomie binaire)
- ...
- TAX₉ = {violet, indigo, bleu, vert, jaune, orangé, rouge, blanc, noir} (taxinomie basée sur les couleurs spectrales, ainsi que blanc et noir)
- ...
- TAX₁₆₇₇₇₂₁₆ = {(0, 0, 0), ..., (255, 255, 255)} (taxinomie utilisée en informatique et distinguant 256 nuances de rouge, vert et bleu)
- ...
- TAX_R = {370, ..., 750} (taxinomie réelle basée sur la longueur d'onde de la lumière)

Au sein de cette hiérarchie, il apparaît que l'usage de taxinomies extrêmes telles que celle basée sur un taxon unique, ou bien la taxinomie réelle, conduisent à des problèmes (respectivement *unification extrême* et *différenciation extrême*). Ainsi, les problèmes mentionnés plus haut lors de l'application d'un raisonnement inductif basé sur SR surviennent lorsque le choix dans la dualité unification/différenciation s'effectue de manière trop radicale. De tels problèmes concernent l'induction en général. Ceci incite à penser que l'on doit plutôt raisonner ainsi : je ne dois privilégier ni l'*unification*, ni la *différenciation*. Un prédicat tel que "rouge", associé à notre taxinomie usuelle des couleurs¹³⁹ (TAX₉), correspond précisément à un tel critère. Il correspond à un choix équilibré dans la dualité unification/différenciation. Ceci permet d'éviter les problèmes précédents. Cela n'empêche pas toutefois l'émergence de nouveaux problèmes, dès lors que l'on cherche à mettre en oeuvre un raisonnement inductif, dans certaines situations. Et un de ces problèmes est naturellement GP.

Ainsi, il apparaît que l'enjeu du choix dans la dualité unification/différenciation est capital du point de vue de l'induction, car selon que je choisirai une manière ou bien l'autre, je pourrai ou non utiliser SR et produire des inférences inductives valables. Confronté à plusieurs instances, on peut mettre en oeuvre soit un processus de différenciation, soit un processus d'unification. Mais le choix effectué conditionne largement le succès ultérieur du raisonnement inductif réalisé sur ce fondement. Je dois décrire à la fois les propriétés communes et les différences. A partir de là, un raisonnement inductif correct peut prendre place. Mais d'ores et déjà, il apparaît que le rôle de la dualité unification/différenciation s'avère crucial pour l'induction. Plus précisément, il apparaît à ce stade

¹³⁹ Ou toute taxinomie qui en est proche.

qu'un choix correct dans la dualité unification/différenciation constitue une des conditions de l'induction.

11.3. Plusieurs problèmes concernant l'induction

Les problèmes qui viennent d'être évoqués constituent l'illustration de plusieurs difficultés inhérentes à la mise en oeuvre du processus inductif. Cependant, à la différence de GP, ils n'engendrent pas véritablement une contradiction. De ce point de vue, ils se distinguent de GP.

Considérons maintenant la situation suivante. Je tire 99 boules respectivement aux temps T_1, T_2, \dots, T_{99} . La 100ème boule sera tirée en T_{100} . On constate que les 99 boules tirées sont rouges. Elles sont donc à la fois rouges et tirées avant T_{100} . Soit R le prédicat "rouge" et T le prédicat "tiré avant T_{100} ". On a alors :

- (I₁₈₇) $RTb_1, RTb_2, \dots, RTb_{99}$
- (H₁₈₈) $RTb_1, RTb_2, \dots, RTb_{99}, RTb_{100}$
- (P₁₈₉) $\therefore RTb_{100}$

Par application directe de SR, il s'ensuit la prédiction : "la 100ème boule est rouge et tirée avant T_{100} ". Mais ceci est en contradiction avec les données de l'expérience en vertu desquelles la 100ème boule est tirée en T_{100} . Là aussi, le raisonnement inductif est basé sur une formalisation qui est celle de SR. Et de même que pour GP, SR conduit ici à une contradiction. Appelons $\Delta 2$ ce problème, où deux prédicats sont utilisés.

Il apparaît que l'on peut construire aisément une forme de $\Delta 2$ basée sur un seul prédicat. Une manière de faire cela est de considérer le prédicat unique S défini comme "rouge et tiré avant T_{100} " en lieu et place des prédicats R et T utilisés précédemment. Il s'ensuit alors la même contradiction.

Plus encore, il apparaît que l'on peut mettre en évidence une autre version ($\Delta 1$) comportant un seul prédicat de ce problème, sans utiliser la propriété "rouge" qui se révèle ici inutile. Soit en effet T le prédicat tiré avant T_{100} . On a alors :

- (I₁₉₀) $Tb_1, Tb_2, \dots, Tb_{99}$
- (H₁₉₁) $Tb_1, Tb_2, \dots, Tb_{99}, Tb_{100}$
- (P₁₉₂) $\therefore Tb_{100}$

Ici aussi, la conclusion selon laquelle la 100ème boule est tirée avant T_{100} contredit les données de l'expérience selon lesquelles la 100ème boule est tirée en T_{100} . Et on a alors un effet contradictoire, à l'instar de GP, sans que la structure de "grue" ait été mise en oeuvre. Compte tenu du fait que seul le critère du temps est utilisé pour construire ce problème, il sera désigné dans ce qui suit par $\Delta 1$ -temps.

Il apparaît ici que les problèmes tels que $\Delta 1$ -temps et $\Delta 2$ conduisent de même que GP à une contradiction. Tel n'est pas le cas pour les autres problèmes relatifs à l'induction évoqués précédemment¹⁴⁰, qui entraînent soit l'impossibilité de réaliser l'induction, soit une conclusion vide d'information. Cependant, il s'avère que la contradiction rencontrée dans $\Delta 1$ -temps n'est pas de même nature que celle observée dans GP. En effet, dans GP, on a une contradiction entre les deux prédictions concurrentes (P) et (P*). En revanche, dans $\Delta 1$ -temps, la contradiction apparaît entre d'une part les conditions de l'expérience ($T \geq 100$) et d'autre part la prédiction résultant de la généralisation ($T < 100$).

En tout état de cause, les problèmes qui viennent d'être rencontrés suggèrent que le formalisme de SR ne capture pas l'ensemble de nos intuitions relatives à l'induction. Il convient donc de s'attacher à définir avec précision les conditions de l'induction, et d'adapter en conséquence le formalisme utilisé. Mais avant toutefois de procéder à une telle analyse, il est nécessaire de préciser davantage les différents éléments du contexte de GP.

¹⁴⁰ Cf. §2.

11.4. Modélisation en termes de n -univers

On modélise la situation inhérente à GP par rapport à des boules contenues dans une urne. On s'attache à décrire avec précision le n -univers correspondant. Dans le contexte des n -univers, qu'est-ce maintenant qu'être *rouge* ? Ici, être "rouge" correspond à deux types de situations, selon le type de n -univers dans lequel on se place. Il peut s'agir en premier lieu, d'un n -univers dont la couleur est l'une des constantes. Dans ce type d'univers, la couleur des objets n'est pas susceptible de varier, et tous les objets y sont invariablement rouges.

Le fait d'être "rouge" peut correspondre, en second lieu, à un n -univers dont la couleur constitue un des critères-variables. Là, un objet peut être rouge ou non rouge. Soit le cas d'un ΩC . Dans un tel univers, un objet est rouge ou non rouge *dans l'absolu*. Aucun changement de couleur n'y est possible, car aucun autre critère-variable n'existe, duquel puisse dépendre une telle variation. Et dans un ΩCT , être rouge, c'est être rouge au temps T . Au sein d'un tel univers, être rouge, c'est être rouge *relativement* au temps T . De même, dans un 3-univers coloré, temporel et localisé (ΩCTL), être rouge, c'est être rouge au temps T et au lieu L . Etc. Dans de tels univers, être rouge, c'est être rouge *relativement* aux autres critères-variables. Et il en va de même pour les n -univers qui modélisent un univers tel que le notre.

Ici se pose le problème du *statut des instances* d'un type d'objet donné. Qu'est-ce donc qu'être une instance, dans le présent cadre ? Ce problème a son importance, car les versions originales de GP sont basées sur des *instances* de boules (1946) et d'émeraudes (1954). Si l'on prend en compte le cas de Goodman (1946), les instances considérées sont 100 boules différentes. Pourtant, si on considère une boule unique, tirée aux temps T_1, T_2, \dots, T_{100} , on constate que la problématique de GP est toujours présente. Il suffit en effet de considérer une boule dont la couleur est susceptible de varier au cours du temps. On a tiré 99 fois la boule aux temps T_1, T_2, \dots, T_{99} , et on constaté chaque fois que la boule était rouge. Il en résulte la prédiction selon laquelle la boule sera rouge en T_{100} . Pourtant, cette prédiction s'avère contradictoire avec une prédiction concurrente basée sur les mêmes observations, et la projection du prédicat S "rouge et tiré avant T_{100} ou non rouge et tiré en T_{100} "¹⁴¹.

Le présent cadre doit être à même d'appréhender la diversité de ces situations. Peut-on parler ainsi d'un 1-univers instancié et temporel, ou bien d'un 1-univers instancié et coloré ? Ici, on doit observer que le fait d'être instancié, pour un univers donné, correspond à un critère-variable supplémentaire. Car sinon, qu'est-ce qui permet de différencier les instances entre elles ? Si aucun critère ne les distingue, il s'agit alors d'une seule et même chose. Et si elles sont distinctes, c'est qu'un critère permet de les différencier. Ainsi, un 1-univers instancié et temporel est en fait un 2-univers, dont le 2ème critère, qui permet de distinguer les instances entre elles, n'est en fait pas mentionné ni explicité. En rendant explicite ce second critère-variable, il est donc clair que l'on se situe dans un 2-univers. De même, un 1-univers instancié et coloré est en réalité un 2-univers dont l'un des critères est la couleur et le second critère existe mais n'est pas précisé.

Un autre aspect qui mérite d'être souligné ici, est la question de la réduction d'un n -univers donné à un autre. N'est il pas possible en effet, de réduire logiquement un n -univers à un système de critères différent ? Intéressons-nous par exemple à un ΩCTL . Pour caractériser l'univers correspondant, on a 3 critères-variables : couleur, temps, localisation. Il apparaît que l'on peut réduire ce 3-univers à un 2-univers. Cela peut s'effectuer en réduisant deux des critères du 3-univers à un seul. En particulier, on réduira les critères de couleur et de temps à un critère unique de *tcouleur** (*shmolor*¹⁴²). Et on ne conservera que deux taxons de *tcouleur** : G et $\sim G$. Soient donc un critère de couleur comportant deux taxons (rouge, non rouge) et un critère de temps comportant deux taxons (avant T , après T). Si l'on associe les taxons de couleur et de temps, on obtient quatre nouveaux prédicats : rouge avant T , rouge après T , non rouge avant T , non rouge après T , que l'on dénotera respectivement $RT, R\sim T, \sim RT$ et $\sim R\sim T$. Plusieurs de ces prédicats sont compatibles entre eux (RT et $R\sim T$, RT et $\sim R\sim T$, $\sim RT$ et $R\sim T$, $\sim RT$ et $\sim R\sim T$) alors que d'autres sont incompatibles (RT et $\sim RT$, $R\sim T$ et $\sim R\sim T$). A ce stade, on a

¹⁴¹ Une telle remarque s'applique également à l'énoncé de Goodman, *Fact, Fiction and Forecast*.

¹⁴² Ainsi que le mentionne J.S. Ullian (1961, p. 387).

plusieurs manières (16)¹⁴³ de grouper les prédicats compatibles, permettant d'obtenir deux nouveaux prédicats G et $\sim G$ de tcouleur* :

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RT \wedge R \sim T		X				X	X	X				X	X	X		X
RT \wedge \sim R \sim T			X			X			X	X		X	X		X	X
\sim RT \wedge R \sim T				X			X		X		X	X		X	X	X
\sim RT \wedge \sim R \sim T					X			X		X	X		X	X	X	X

Dans chacun de ces cas, il en résulte bien un nouveau critère unique de tcouleur* (Z), se substituant aux deux critères précédents de couleur et de temps. On dénotera Z_i ($0 \leq i \leq 15$) les taxons de tcouleur* ainsi obtenus. S'il est clair que Z_{15} conduit à l'induction vide, on observera que plusieurs cas correspondant à la situation où les instances sont RT conduisent à la problématique GP. On notera ainsi que Z_2 , c'est-à-dire *grue*₂ (en assimilant les Z_i à *grue*_i et les Z_{15-i} à *bleen*_i) est basé sur la définition : *grue*₂ = rouge avant T *et* non rouge après T. Il s'agit là d'une interprétation *conjonctive* de la définition de "grue". De même, *grue*₇ correspond à une définition de "grue" basée sur un *ou exclusif*. Enfin, *grue*₁₂ est basé sur la définition classique : *grue*₁₂ = rouge avant T *ou* non rouge après T, où la disjonction s'interprète comme un *ou inclusif*.

De la même manière, il s'avère également qu'un Ω CT peut se réduire à un 1-univers tcoloré* (ΩZ). Et de façon plus générale, un n -univers est ainsi *réductible* à un $(n-1)$ -univers (pour $n > 1$). Ainsi, si l'on considère un univers donné, plusieurs caractérisations en termes de n -univers peuvent être valablement utilisées. On peut notamment appréhender un même univers comme un Ω CTL, ou bien comme un ΩZL . De la même manière, on peut se représenter un Ω CT comme un ΩZ . A ce stade, aucune de ces vues n'apparaît fondamentalement meilleure que l'autre. Mais chacune de ces deux caractérisations constituent des façons alternatives de décrire une même réalité. Ceci montre finalement qu'un n -univers constitue en fait une caractérisation abstraite d'un univers réel ou imaginaire. Un n -univers constitue ainsi un *système de critères*, comportant des constantes et des variables. Et pour caractériser un même univers réel ou imaginaire donné, on peut recourir valablement à plusieurs n -univers. Chacun d'entre eux apparaît finalement comme une caractérisation distincte de l'univers considéré, faisant simplement appel à un jeu de primitives différent.

11.5. Conditions de l'induction

Le fait que le formalisme de SR entraîne l'effet de GP suggère que l'intuition qui préside à notre notion d'induction n'est pas entièrement capturée par SR. Il est ainsi permis de penser que si l'approche formelle est nécessaire et utile pour servir de support à l'induction, elle ne constitue pas toutefois une démarche suffisante. Car il paraît également essentiel de capturer l'intuition qui préside à notre raisonnement inductif. Aussi s'avère-t-il nécessaire de compléter l'approche formelle de l'induction par une approche sémantique. Goodman lui-même fait mention d'une définition de l'induction¹⁴⁴. Il définit l'induction comme la projection de caractéristiques du passé dans le futur, ou d'une manière plus générale, comme la projection de caractéristiques correspondant à un aspect donné d'un objet à travers un autre aspect. Cette définition correspond à notre intuition de l'induction. On peut penser toutefois qu'il convient de la compléter en prenant en compte les observations précédentes¹⁴⁵ relatives à la dualité différenciation/unification. En ce sens, on a pu observer que l'induction consiste en une inférence à partir d'instances présentant à la fois des propriétés communes, et des différences. Soient

¹⁴³ Soit $C(0, 4)+C(1, 4)+C(2, 4)+C(3, 4)+C(4, 4) = 2^4$, où $C(p, q)$ désigne le nombre de combinaisons de q éléments pris p à la fois.

¹⁴⁴ Cf. Goodman, "A Query On Confirmation", p. 383 : "Induction might roughly be described as the projection of characteristics of the past into the future, or more generally of characteristics of one realm of objects into another." Trad.: "L'induction pourrait être décrite de manière sommaire comme la projection des caractéristiques du passé dans le futur, ou plus généralement des caractéristiques d'un domaine d'objets dans un autre."

¹⁴⁵ Cf. §2 ci-dessus.

les instances-source (instances-S) les instances sur lesquelles portent (I) ou (I*) et l'instance-destination (instance-D) celle qui fait l'objet de (P) ou (P*). Les propriétés communes concernent les instances-S et les propriétés différenciées s'établissent entre les instances-S et l'instance-D. Il en résulte la définition suivante : l'induction consiste précisément dans le fait que l'instance-D¹⁴⁶ possède également la propriété commune aux instances-S, alors que l'on fait varier le(s) critère(s) sur le(s)quel(s) est (sont) basé(es) les différences entre les instances-S et l'instance-D. Le raisonnement inductif est ainsi fondé sur le caractère constant d'une propriété, alors que telle autre propriété est variable.

De cette définition de l'induction découlent directement plusieurs conditions de l'induction. Il convient de les examiner tour à tour. Les deux premières conditions sont ainsi les suivantes :

- (C1) les instances-S doivent présenter des propriétés communes
- (C2) les instances-S et l'instance-D doivent présenter des propriétés distinctives

Ceci a pour conséquence qu'on ne peut appliquer l'induction dans deux circonstances particulières : d'une part (a) lorsque les instances ne laissent apparaître aucune propriété commune. On appellera un tel cas une *différenciation totale* des instances. Les problèmes correspondant à cette circonstance particulière ont été évoqués plus haut¹⁴⁷. Et d'autre part (b) lorsque les instances ne présentent aucune propriété distinctive. On appellera une telle situation *unification totale*. Les problèmes rencontrés dans ce type de situation ont également été mentionnés précédemment¹⁴⁸.

On doit noter ici qu'il ne s'agit pas là de propriétés intrinsèques des instances, mais bien de l'analyse qui est effectuée par celui qui s'apprête à raisonner par induction.

Compte tenu de la définition de l'induction qui a été donnée, une troisième condition peut être ainsi énoncée :

- (C3) un critère-variable est nécessaire pour les propriétés communes des instances-S et un autre critère-variable pour les propriétés distinctives

Ceci se rapporte à la structure de l'univers de référence considéré. En conséquence, deux critères-variables sont au minimum nécessaires, dans la structure de l'univers de référence correspondant. On appellera cela la *condition minimale* de l'induction. Par conséquent, un 2-univers est au minimum nécessaire pour que les conditions de l'induction soient satisfaites. Ainsi, un Ω CT conviendra. De même, un 2-univers temporel et localisé (Ω TL) satisfera également les conditions qui viennent d'être définies, etc¹⁴⁹.

On peut noter qu'une autre façon d'énoncer cette condition est la suivante : le critère-variable pour les propriétés communes et le critère-variable pour les propriétés différenciées doivent être distincts. On ne doit pas avoir confusion entre les deux. On peut appeler cela la *condition de séparation* des propriétés communes et des propriétés distinctives. Un tel principe apparaît comme une conséquence de la condition minimale pour l'induction : on doit avoir deux critères pour réaliser l'induction, et ces critères doivent être différents. Si l'on choisit un même critère pour les propriétés communes et les propriétés différenciées, on se ramène de fait à un seul critère et au contexte d'un 1-univers, lui-même insuffisant pour réaliser l'induction.

Enfin, une quatrième condition de l'induction résulte de la définition précédente :

- (C4) on doit projeter les propriétés communes des instances-S (et non les propriétés distinctives)

Les conditions de l'induction qui viennent d'être énoncées permettent désormais de traiter les problèmes liés à l'utilisation de SR évoqués plus haut¹⁵⁰. Il s'ensuit en effet que les projections¹⁵¹

¹⁴⁶ On peut bien sûr prendre en considération, de manière alternative, plusieurs instances-D.

¹⁴⁷ Cf. §2 ci-dessus.

¹⁴⁸ *Ibid.*

¹⁴⁹ Pour l'application de cette condition, on doit tenir compte des remarques mentionnées plus haut concernant le problème du statut des instances. Ainsi, on doit en réalité assimiler un 1-univers instancié et temporel à un 2-univers dont l'un des critères est temporel, et le second critère n'est pas explicite. De même, un 1-univers instancié et coloré s'assimile en fait à un 2-univers dont l'un des critères est temporel, et le second critère n'est pas spécifié.

¹⁵⁰ Cf. §3 ci-dessus.

suivantes sont correctes : $C^\circ T$ dans un ΩCT , $C^\circ L$ dans un ΩCL , $Z^\circ L$ dans un ΩZL , etc. A l'inverse, les projections suivantes sont incorrectes : $T^\circ T$ dans un ΩT , $Z^\circ Z$ dans un ΩZ . En particulier, on notera ici que la projection $T^\circ T$ dans le ΩT est celle de $\Delta 1$ -temps. $\Delta 1$ -temps prend en effet place dans un ΩT , alors que l'induction exige à la fois des propriétés communes et des propriétés distinctives. Ainsi, un 2-univers est au minimum nécessaire. D'habitude, le critère du temps est utilisé pour la différenciation. Mais ici, il est utilisé pour l'unification ("tiré avant T"). Cela peut se faire, mais à condition qu'on utilise un critère distinct pour les propriétés différenciées. Cependant, alors qu'il en résulte ici des propriétés communes, on perd les propriétés différenciées. Il manque donc un second critère - correspondant aux propriétés différenciées - à l'univers considéré, pour réaliser valablement l'induction. Ainsi $\Delta 1$ -temps trouve-t-il son origine dans une violation de la *condition minimale* de l'induction. On peut formuler cette solution de manière équivalente, par rapport à la *condition de séparation*. En effet, dans $\Delta 1$ -temps, un même critère temporel (tiré avant T/tiré après T) est utilisé pour les propriétés communes et pour les propriétés différenciées, alors que deux critères distincts sont nécessaires. Il s'agit ainsi d'une violation manifeste de la condition de séparation.

Enfin, les conditions de l'induction définies plus haut conduisent à adapter le formalisme utilisé pour décrire GP. Il s'avère en effet nécessaire de distinguer entre la propriété commune et la(les) propriété(s) distinctive(s). On utilisera donc le formalisme suivant en lieu et place de celui utilisé plus haut :

$$(I_{193}) \quad RT_1 \cdot RT_2 \cdot RT_3 \cdot \dots \cdot RT_{99}$$

$$(H_{194}) \quad RT_1 \cdot RT_2 \cdot RT_3 \cdot \dots \cdot RT_{99} \cdot RT_{100}$$

où R désigne la propriété commune et les T_i une propriété distinctive. Il est à noter qu'ici, il peut s'agir au choix d'un objet unique, ou bien d'instances que distingue un critère donné (qui n'entre pas en jeu dans le processus inductif) selon le n -univers dans lequel on se place. Ainsi, on utilisera dans le cas d'une instance unique α , dont la couleur est susceptible de varier selon le temps :

$$(I_{195}) \quad RT_1 \alpha \cdot RT_2 \alpha \cdot RT_3 \alpha \cdot \dots \cdot RT_{99} \alpha$$

ou dans le cas où plusieurs instances $\alpha_1, \alpha_2, \dots, \alpha_{99}, \alpha_{100}$ existent¹⁵² :

$$(I_{196}) \quad RT_1 \alpha_1 \cdot RT_2 \alpha_2 \cdot RT_3 \alpha_3 \cdot \dots \cdot RT_{99} \alpha_{99}$$

11.6. Solution

Compte tenu des conditions de l'induction et du cadre des n -univers qui viennent d'être définis, on est désormais en mesure de s'attacher à déterminer l'origine de GP. Il convient pour cela tout d'abord de décrire avec précision les conditions de l'univers de référence dans lequel GP prend place. En effet, dans la version originale de GP, le choix de l'univers de référence n'est pas défini avec précision. Or on peut penser qu'il est essentiel, afin d'éviter toute ambiguïté, que ce dernier soit décrit précisément.

L'univers de référence dans lequel se place Goodman (1946) n'est pas défini explicitement, mais plusieurs éléments de l'énoncé permettent d'en préciser la nature. Goodman fait ainsi mention des couleurs "rouge" et "non rouge". Aussi la couleur constitue-t-elle un des critères-variables de l'univers de référence. De plus, Goodman distingue les boules qui sont tirées aux temps $T_1, T_2, T_3, \dots, T_{100}$. Ainsi, le temps est également un critère-variable de l'univers considéré. Par conséquent, on peut décrire l'univers minimal dans lequel se place Goodman (1946) comme un ΩCT . De même, dans Goodman (1954), les critères-variables de couleur (vert/non vert) et de temps (tiré avant T/tiré après T) sont expressément mentionnés. Dans les deux cas, on se place donc, implicitement dans le cadre *minimal* d'un ΩCT .

Goodman fait par ailleurs mention d'*instances* de boules ou d'émeraudes. Faut-il à ce stade recourir à un critère-variable supplémentaire permettant de distinguer les instances entre elles ? Il apparaît que

¹⁵¹ Avec les notations C (couleur), T (temps), L (localisation) et Z (tcouleur*).

¹⁵² Toutefois, dès lors que le fait qu'il existe une ou plusieurs instances n'est pas essentiel dans la formulation du problème considéré, on pourra évidemment s'abstenir d'en faire mention.

non. D'une part en effet, comme on l'a vu précédemment¹⁵³, il s'avère que l'on a bien une version de GP en considérant simplement un ΩCT et un objet unique dont la couleur est susceptible de varier au cours du temps. D'autre part, il apparaît que si le critère qui sert à distinguer les instances n'est pas utilisé dans le processus inductif, il ne sert alors ni en tant que critère commun, ni en tant que critère différencié. Il s'ensuit alors que l'on peut se dispenser de recourir à ce 3ème critère supplémentaire. Ainsi, il s'avère que le fait de prendre en compte une instance unique ou bien plusieurs instances, n'est pas essentiel dans la formulation de GP. Dans ce qui suit, on pourra donc considérer que l'énoncé s'applique, indifféremment, à un objet unique ou à plusieurs instances que distingue un critère qui n'est pas utilisé dans le processus inductif.

Désormais, nous sommes en mesure de replacer GP dans le cadre des n -univers. Compte tenu du fait que le contexte de GP est celui d'un ΩCT ... *minimal*, on envisagera successivement deux situations : celle d'un ΩCT , puis celle d'un $\Omega CT\alpha$ (où α désigne un 3ème critère).

11.6.1. "Grue" dans le 2-univers coloré et temporel

Envisageons tout d'abord l'hypothèse d'un ΩCT . Dans un tel univers, être "rouge", c'est être rouge au temps T. On dispose alors d'un critère de couleur pour les propriétés communes, et d'un critère de temps pour les propriétés différenciées. Dès lors, il apparaît tout à fait légitime de projeter la propriété commune de couleur ("rouge"), à travers le temps différencié. Une telle projection s'avère conforme aux conditions de l'induction énoncées plus haut.

Qu'en est-il maintenant de la projection de "grue" ? On a observé précédemment¹⁵⁴ que le ΩCT était réductible à un ΩZ . Ici, le fait d'utiliser "grue" (et "bleen") en tant que primitives, est caractéristique du fait que le système de critères utilisé est celui d'un ΩZ . Qu'en est-il alors lorsqu'on projette "grue", dans le ΩZ ? Dans un tel univers de référence, l'unique critère-variable est la tcouleur*. Un objet y est "grue" ou "bleen" dans l'*absolu*. Dès lors, si l'on dispose bien d'un critère commun (la tcouleur*), il apparaît que le critère différencié fait défaut, pour mettre en oeuvre valablement l'induction. Et la situation dans laquelle on se trouve est celle d'une indifférenciation extrême. Ainsi, une telle projection s'effectue en violation de la condition minimale de l'induction. Par conséquent, il s'avère que GP ne peut prendre place dans le ΩCT , et se trouve bloqué au stade de la projection de "grue".

Mais ces remarques préliminaires sont-elles suffisantes pour fournir, dans le contexte d'un ΩCT , une solution satisfaisante à GP ? On peut penser que non, car le paradoxe s'y présente également sous une autre forme, qui est celle de la projection de la tcouleur* à travers le temps. On peut formaliser ainsi cette projection $Z^{\circ}T$:

$$\begin{aligned} (I^*_{197}) & GT_1 \cdot GT_2 \cdot GT_3 \cdot \dots \cdot GT_{99} \\ (H^*_{198}) & GT_1 \cdot GT_2 \cdot GT_3 \cdot \dots \cdot GT_{99} \cdot GT_{100} \text{ qui équivaut à :} \\ (H^*_{199}) & RT_1 \cdot RT_2 \cdot RT_3 \cdot \dots \cdot RT_{99} \sim RT_{100} \\ (P^*_{200}) & GT_{100} \text{ qui équivaut à :} \\ (P^*_{201}) & \sim RT_{100} \end{aligned}$$

où il est manifeste que les éléments de la problématique GP sont encore présents.

Fondamentalement dans cette version, il apparaît que les propriétés communes sont empruntées au système de critères du ΩZ , alors que les propriétés différenciées proviennent du ΩCT . Une première analyse révèle donc que la projection de "grue" dans ces conditions comporte un défaut qui consiste dans le choix d'un système de critères donné pour les propriétés communes (la tcouleur*) et d'un système de critères différent pour les propriétés différenciées (le temps). Car la sélection de la tcouleur* est caractéristique du choix d'un ΩZ , alors que l'utilisation du temps est révélatrice du fait que l'on se place dans un ΩCT . Mais on se doit de choisir l'un ou l'autre des systèmes de critères réductibles pour réaliser l'induction. Dans les hypothèses envisagées précédemment, le choix des critères pour les propriétés communes et différenciées s'effectuait dans *un même système de critères*. Mais ici, le choix des critères pour les propriétés communes et les propriétés différenciées s'effectue dans *deux systèmes de critères* (réductibles) différents. Ainsi, les critères commun et différencié

¹⁵³ Cf. §4.

¹⁵⁴ *Ibid.*

choisis pour l'induction ne sont pas véritablement distincts. Et ceci apparaît comme une *violation de la condition de séparation*. Par conséquent, une des conditions de l'induction n'est pas respectée.

Cependant, la projection $Z^{\circ}T$ possède un certain support intuitif, car elle est basée sur le fait que les notions de "grue avant T" et "grue après T" se révèlent, intuitivement, porteuses de sens. Faisons donc abstraction de la violation des conditions de l'induction qui vient d'être mentionnée, et considérons donc cette situation. Dans ce contexte, GP est toujours présent, puisqu'on observe une contradiction entre (P) et (P*). C'est à cette contradiction qu'il convient désormais de s'intéresser. Soit l'étape d'équivalence entre (H*) et (H'*). On conçoit que "grue avant T" s'assimile ici à RT, car le fait que les instances-S sont rouges avant T résulte clairement des conditions de l'expérience. En revanche, il convient de s'intéresser à l'étape selon laquelle (P*) entraîne (P'). Selon la définition classique¹⁵⁵ : "grue" = $\{RT \wedge R\sim T, RT \wedge \sim R\sim T, \sim RT \wedge \sim R\sim T\}$. Qu'est-ce donc qu'être "grue après T" ? Là, il apparaît qu'un objet "grue" peut être $R\sim T$ (ceci correspond au cas $RT \wedge R\sim T$) ou bien $\sim R\sim T$ (ceci correspond aux cas $RT \wedge \sim R\sim T$ et $\sim RT \wedge \sim R\sim T$). En conclusion, l'objet peut être soit $R\sim T$ soit $\sim R\sim T$. Ainsi, le fait de savoir qu'un objet est "grue après T" ne permet pas de conclure que cet objet est $\sim R\sim T$, car ce dernier peut également être $R\sim T$. En conséquence, l'étape selon laquelle (P*) entraîne (P') se révèle finalement *fausse*. D'où il s'ensuit que la contradiction entre (P) et (P*) n'a plus de raison d'être.

On peut se persuader que cette analyse ne dépend pas du choix de la définition classique de "grue" (*grue*₁₂) qui est effectué, en considérant d'autres définitions. Soit par exemple la définition basée sur *grue*₉ : "grue" = $\{RT \wedge \sim R\sim T, \sim RT \wedge \sim R\sim T\}$ et "bleen" = $\{RT \wedge R\sim T, \sim RT \wedge R\sim T\}$. Mais dans cette version, on constate que l'on n'a pas l'émergence de GP, car les instances-S, qui sont RT, peuvent être à la fois "grue" et "bleen". Et il en va de même si on considère une définition conjonctive (*grue*₂) telle que "grue" = $\{RT \wedge \sim R\sim T\}$. Dans un tel cas en effet, les instances-S ne sont "grue" que si elles sont RT mais également $\sim R\sim T$. Or ceci ne correspond pas aux conditions initiales de GP dans le Ω^2CT où on ignore si les instances-S sont $\sim R\sim T$.

On pourrait penser également que le problème est lié à l'utilisation d'une taxinomie de tcouleur* basée sur deux taxons (G et $\sim G$). Considérons donc une taxinomie de la tcouleur* basée sur 4 taxons : $Z_0 = RT \wedge R\sim T, Z_1 = RT \wedge \sim R\sim T, Z_2 = \sim RT \wedge R\sim T, Z_3 = \sim RT \wedge \sim R\sim T$. Mais dans cette hypothèse, il apparaît clairement que dès lors que les instances-S sont par exemple Z_1 , on se trouve replacé dans la situation précédente.

Le fait de considérer "grue après T", "grue avant T", "bleen avant T", "bleen après T" s'assimile à une tentative d'exprimer "grue" et "bleen" par rapport à nos propres critères, et en particulier celui du temps. Il s'agit là d'une forme d'anthropocentrisme, sous-tendue par l'idée d'exprimer le ΩZ à l'aide des taxons du ΩCT . Dès lors que l'on connaît le code définissant les relations entre deux n -univers réductibles - le ΩZ et le ΩCT - et que l'on possède des données partielles, on peut être tenté d'élucider complètement les prédicats du n -univers étranger. Sachant que les instances sont GT, $G\sim T, \sim GT, \sim G\sim T$, je peux déduire qu'elles sont respectivement $\{RT, \sim RT\}, \{R\sim T, \sim R\sim T\}, \{\sim RT\}, \{R\sim T\}$. Mais comme on l'a vu, du fait que les instances sont GT et RT, je ne peux pas déduire qu'elles seront $\sim R\sim T$.

Le raisonnement dans cette version de GP est basé sur l'idée apparemment inductive que ce qui est "grue avant T" est également "grue après T". Mais dans le contexte qui est celui du ΩZ , lorsqu'un objet est "grue", il est "grue" dans l'absolu. Car aucun critère supplémentaire n'existe qui puisse faire varier sa tcouleur*. Ainsi, lorsqu'un objet est GT, il est *nécessairement* $G\sim T$. Et de l'information selon laquelle un objet est GT, on peut donc conclure, par *dédution*, qu'il est également $G\sim T$.

De ce qui précède, il s'ensuit que la version de GP liée à $Z^{\circ}T$ présente les caractères apparents de l'induction, mais il ne s'agit pas d'une forme authentique de ce type de raisonnement. $Z^{\circ}T$ constitue ainsi une forme déguisée de l'induction pour deux raisons principales: d'une part, il s'agit d'une projection à travers le critère différencié du temps, qui constitue le mode standard de notre pratique inductive. Et d'autre part, elle est basée sur le principe intuitif selon lequel tout ce qui est GT est également $G\sim T$. Mais comme on l'a vu, il s'agit en réalité là d'une forme déductive de raisonnement, dont la véritable nature se trouve masquée par un apparent mouvement inductif. Et ceci conduit à conclure que la forme de GP apparentée à $Z^{\circ}T$ s'analyse en fait véritablement comme une *pseudo-induction*.

¹⁵⁵ Il s'agit de celle basée sur le ou inclusif (*grue*₁₂).

11.6.2. "Grue" dans le 3-univers coloré, temporel et localisé

Envisageons maintenant le cas d'un $\Omega CT\alpha$. Ce type d'univers de référence correspond également à la définition d'un $\Omega CT\dots$ *minimal*, mais il comporte également un 3ème critère-variable¹⁵⁶. Choisissons pour ce dernier un critère tel que la localisation¹⁵⁷. Soit donc un ΩCTL . Considérons tout d'abord (H) dans un tel 3-univers. Etre "rouge" dans le ΩCTL , c'est être rouge au temps T et au lieu L. D'après les conditions de GP, la couleur correspond aux propriétés *communes*, et le temps aux propriétés *différenciées*. On a alors la projection C°TL suivante :

$$\begin{aligned} (I_{202}) & RT_1L_1 \cdot RT_2L_2 \cdot RT_3L_3 \cdot \dots \cdot RT_{99}L_{99} \\ (H_{203}) & RT_1L_1 \cdot RT_2L_2 \cdot RT_3L_3 \cdot \dots \cdot RT_{99}L_{99} \cdot RT_{100}L_{100} \\ (P_{204}) & \therefore RT_{100}L_{100} \end{aligned}$$

où compte tenu des conditions de l'induction, il s'avère légitime de projeter la propriété commune ("rouge") des instances-S, à travers le temps et le lieu différenciés, et de prédire que la 100ème boule sera rouge. Une telle projection apparaît tout à fait correcte, et s'avère en tous points conforme aux conditions de l'induction mentionnées plus haut.

Qu'en est-il maintenant de (H*) dans le ΩCTL ? On a pu observer que le ΩCTL pouvait se réduire à un ΩZL . Dans ce dernier n -univers, les critères-variables sont la couleur* et la localisation. Le fait d'être "grue" y est *relatif* au lieu : être "grue", c'est être "grue" au lieu L. Ce qui est alors projeté est la couleur*, c'est-à-dire le fait d'être "grue" ou "bleen". On a donc un critère commun de couleur* et un critère différencié de localisation. Dès lors, si on considère que les instances-S sont "grue", on peut fort bien projeter la propriété commune "grue" à travers un critère différencié de localisation. Soit donc la projection Z°L dans le ΩZL :

$$\begin{aligned} (I^*_{205}) & GL_1 \cdot GL_2 \cdot GL_3 \cdot \dots \cdot GL_{99} \\ (H^*_{206}) & GL_1 \cdot GL_2 \cdot GL_3 \cdot \dots \cdot GL_{99} \cdot GL_{100} \\ (P^*_{207}) & \therefore GL_{100} \end{aligned}$$

Une telle projection est conforme aux conditions mentionnées plus haut, et constitue par conséquent une forme valable de l'induction.

Dans ce contexte, on peut projeter valablement un prédicat ayant une structure identique à celle de "grue", dans le cas des émeraudes. Considérons la définition "grue" = vert avant T ou non vert après T, où T = 10 milliards d'années. On sait qu'à cette époque, notre Soleil sera éteint, et deviendra progressivement une naine blanche. Les conditions de notre atmosphère seront radicalement différentes de ce qu'elles sont actuellement. Et la température s'élèvera notamment dans des proportions considérables, pour atteindre 8000°. Dans ces conditions, la structure de nombreux minéraux se transformera radicalement. Il devrait normalement en être ainsi pour nos émeraudes actuelles, qui devraient voir leur couleur modifiée, à la suite de l'énorme élévation de température qui s'ensuivra. Ainsi, j'observe actuellement une émeraude: elle est "grue" (pour T = 10 milliards d'années). Si je projette cette propriété à travers un critère de *lieu*, j'en conclus légitimement que l'émeraude trouvée au cœur de la forêt amazonienne sera également "grue", de même également que l'émeraude qui vient d'être extraite d'une mine d'Afrique du Sud.

A ce stade, on pourrait s'interroger pour savoir si la projetabilité de "grue" n'est pas liée intrinsèquement au choix d'une définition de "grue" basée sur le ou inclusif (*grue*₁₂) ? Cependant, on vérifie aisément en utilisant une définition alternative de "grue" que sa projection demeure valide¹⁵⁸.

¹⁵⁶ Une même solution s'applique, bien sûr, si l'on considère un nombre de critères-variables supérieur à 3.

¹⁵⁷ Tout autre critère distinct de la couleur ou du temps, conviendrait également.

¹⁵⁸ En particulier, il apparaît que la projection d'une définition conjonctive (*grue*₂) nous est en fait familière. En effet, nous ne faisons pas autre chose lorsque nous projetons le prédicat "être vert avant maturité et rouge après maturité" applicable aux tomates, à travers un critère différencié de lieu : ceci est vrai des 99 instances-S observées en Corse et en Provence, et se projette valablement à une 100ème instance située en Sardaigne. On peut observer qu'un tel type de projection est notamment considéré comme non problématique par Franck Jackson (1975): "There seems no case for regarding 'grue' as nonprojectible if it is defined this way. An emerald is *grue*₁ just if it is green up to T and blue

On remarque qu'on a ici l'expression du fait que la taxinomie basée sur la tcouleur* est plus grossière que celle basée sur le temps et la couleur. En effet, la première ne comprend que 2 taxons (grue/bleen), alors que la seconde en comprend 4. En réduisant les critères de couleur et de temps à un critère unique de tcouleur*, on a remplacé 4 taxons ($RT \wedge R \sim T$, $RT \wedge \sim R \sim T$, $\sim RT \wedge R \sim T$, $\sim RT \wedge \sim R \sim T$) par 2. Ainsi, "grue" constitue de ce point de vue un prédicat plus grossier que "rouge". L'univers qui est décrit n'a pas changé, mais les n -univers qui sont des systèmes de critères décrivant ces univers sont différents. Avec la tcouleur* ainsi définie, on dispose de moins de prédicats pour décrire une même réalité. Les prédicats "grue" et "bleen" sont pour nous peu informatifs, et le sont moins, en tout état de cause que nos prédicats "rouge", "non rouge", "avant T", etc. Mais cela n'empêche pas toutefois "grue" et "bleen" d'être projetables.

Alors que la projection de "grue" se révèle valide dans le ΩZL , on remarquera cependant que l'on n'observe pas dans ce cas la contradiction entre (P) et (P*). Car ici (I*) équivaut bien à :

$$(I^*_{208}) RT_1L_1 \cdot RT_2L_2 \cdot RT_3L_3 \cdot \dots \cdot RT_{99}L_{99}$$

puisque, sachant d'après les données initiales de GP que les instances-S sont RT, on remplace valablement les GL_i par les RT_iL_i ($i < 100$). Mais il apparaît que dans cette hypothèse, (P*) n'entraîne pas :

$$(P^*_{209}) \sim RT_{100}L_{100}$$

car on ne possède pas d'indication relative à la temporalité de la 100ème instance, du fait que seule la localisation constitue ici le critère différencié. En conséquence, on a bien dans le cas du ΩCTL une version construite à partir des éléments de GP où la projection de "grue" s'effectue valablement, mais qui ne se révèle pas paradoxale.

11.7. Conclusion

Dans la solution à GP proposée par Goodman, un prédicat est projetable ou non projetable dans l'*absolu*. Et on a d'autre part une correspondance entre les prédicats implantés¹⁵⁹/ non implantés et les prédicats projetables / non projetables. Goodman par ailleurs ne fournit pas de justification à cette assimilation. Dans la présente approche, on n'a pas une telle dichotomie, car un prédicat donné P se révèle projetable dans un n -univers donné, et non projetable dans un autre n -univers. Ainsi, P est projetable *relativement* à tel univers de référence. On a donc la distinction projetable / non projetable *relativement à tel n-univers*. Et cette distinction est justifiée par les conditions de l'induction, et par le mécanisme fondamental de celle-ci par rapport à la dualité unification/différenciation. On a ainsi des n -univers où "vert" est projetable et d'autres où il ne l'est pas. De même, "grue" se révèle ici projetable relativement à certains n -univers. Ni *vert* ni *grue* ne sont projetables dans l'absolu, mais seulement relativement à tel univers donné. De même que d'autres prédicats, "grue" est projetable dans certains univers de référence, mais non projetable dans d'autres¹⁶⁰.

thereafter, and if we discovered that all emeralds so far examined had this property, then, other things being equal, we would probably accept that all emeralds, both examined and unexamined, have this property (...". Trad.: "Il ne semble pas y avoir de quoi considérer là 'grue' comme non projectible s'il est défini de cette manière. Une émeraude est grue₁ uniquement si elle est verte jusqu'à T et à bleue après, et si nous découvrons que toutes les émeraudes examinées jusqu'à présent ont eu cette propriété, puis, toutes choses étant par ailleurs égales, nous accepterions alors certainement que toutes les émeraudes, examinées ou non, ont probablement cette propriété (...". Si on devait replacer un tel prédicat dans la présente analyse, on devrait alors considérer que la projection s'effectue par exemple à travers un critère différencié de localisation (p. 115).

¹⁵⁹ C'est-à-dire "entrenched" (Goodman, *Fact, Fiction and Forecast*).

¹⁶⁰ La conception développée dans J. Holland, K. Holyoak, R. Nisbett et P. Thagard (*Induction*, Cambridge, MA; Londres, MIT Press, 1986) me paraît constituer une variation de la solution de Goodman, orientée vers le traitement informatique des données et basée sur la distinction intégré / non intégré dans la hiérarchie par défaut. Mais la solution de Holland présente les mêmes inconvénients que celle de Goodman : quelle justification sinon anthropocentrique, possède-t-on pour cette distinction ? Cf. p. 235 : "Concepts such as "grue", which are of no significance to the goals of the

Ainsi, il s'avère qu'une des causes de GP réside dans le fait que dans GP, on s'attache classiquement à opérer une dichotomie entre les prédicats projetables et les prédicats non projetables. Les solutions classiquement proposées pour résoudre GP sont respectivement basées sur la distinction temporel / non temporel, local / non local, qualitatif / non qualitatif, implanté / non implanté, etc. et une mise en correspondance avec la distinction projectable / non projectable. On s'interroge ainsi sur le caractère projectable ou non, dans l'*absolu*, de tel prédicat P* présentant la structure de "grue". Ceci résulte du fait que dans GP, on a une contradiction entre les deux prédictions concurrentes (P) et (P*). On en déduit classiquement qu'une des deux prédictions doit être rejetée, en même temps qu'une des deux généralisations (H) ou (H*) sur lesquelles ces prédictions sont respectivement basées. A l'inverse, dans la présente analyse, que l'on se place dans le cas de la projection *authentique* Z°L ou de la *pseudo-projection* Z°T, on n'a pas la contradiction entre (P) et (P*). Dès lors, on ne se trouve plus contraint de rejeter soit (H) soit (H*). Et la distinction entre prédicats projetables / non projetables ne se révèle plus indispensable¹⁶¹.

Comment s'effectue dans ce contexte le choix de nos *n*-univers usuels ? Des *n*-univers tels que le Ω CT, le Ω CTL, le Ω ZL etc. conviennent pour réaliser l'induction. Mais nous tendons naturellement à privilégier ceux qui sont basés sur des critères structurés assez finement pour permettre un maximum de *combinaisons* de projections. Si l'on opère à partir des critères Z et L dans le Ω ZL, on s'autorise un nombre de combinaisons restreint : Z°L et L°Z. A l'inverse, si l'on retient les critères C, T et L, on se place dans le Ω CTL et on a la possibilité des projections C°TL, T°CL, L°CT, CT°L¹⁶², CL°T, TL°C. On a ainsi un maximum de combinaisons. Ceci semble inciter à préférer le Ω CTL au Ω ZL. Bien sûr, le pragmatisme semble devoir jouer un rôle dans le choix optimal de nos critères. Mais il semble que ce ne soit qu'un des multiples facteurs qui interagissent pour permettre l'optimisation de nos critères pour effectuer les opérations primitives de regroupement et de différenciation, afin de pouvoir ensuite

learner, will never be generated and hence will not form part of the default hierarchy. (...) Generalization, like other sorts of inference in a processing system, must proceed from the knowledge that the system already has". Trad.: "Des concepts tels que "grue", qui sont sans signification par rapport aux buts poursuivis par le chercheur, ne seront jamais créés et par conséquent ne feront pas partie de la hiérarchie de défaut (...). La généralisation, comme d'autres sortes d'inférence dans un système de traitement des données, doivent opérer à partir des connaissances que le système possède déjà."

La présente analyse se distingue aussi de celle présentée par Susan Haack dans *Evidence and Inquiry* (1993), car l'existence d'espèces naturelles (*natural kinds*) ne constitue pas ici une condition pour l'induction. Cf. p. 134 : "There is a connection between induction and natural kinds. [...] the reality of kinds and laws is a necessary condition of successful inductions". Trad.: "Il existe une connexion entre l'induction et les espèces naturelles [...] la réalité des espèces et des lois est une condition nécessaire pour réussir l'induction". Dans le présent contexte, le fait que les conditions de l'induction (un critère commun, un critère différencié distinct, etc.) soient satisfaites convient pour réaliser l'induction.

¹⁶¹ Une remarque similaire est formulée par Franck Jackson en conclusion de son article ("Grue", p. 131) : "[...] the SR can be specified without invoking a partition of predicates, properties or hypotheses into the projectible and the nonprojectible". Trad.: "[...] la SR peut être posée sans qu'elle nécessite une partition entre les prédicats, les propriétés ou les hypothèses entre le projectible et le non projectible". Pour Jackson, tous les prédicats non contradictoires sont projectibles : "[...] all (consistent) predicates are projectible." (p. 114) Trad.: "[...] tous les prédicats (consistants) sont projectibles.". Une telle conclusion apparaît toutefois plus forte que celle qui résulte de la présente analyse. Car pour Jackson, tous les prédicats sont donc projectibles dans l'*absolu*. Mais dans le présent contexte, on n'a pas de prédicats projectibles ou non projectibles dans l'*absolu*. Ce n'est que *relativement* à un *n*-univers donné, qu'un prédicat P se révèle projectible ou non projectible.

De manière plus générale, la présente analyse se distingue essentiellement de celle de Jackson en ce sens que la solution proposée à GP ne repose pas sur la condition contrefactuelle (*counterfactual condition*). Cette dernière en effet apparaît trop liée à l'utilisation de certains prédicats (*examined, sampled, etc.*). En revanche, dans le présent contexte, on considère le problème d'un point de vue général, indépendamment de la nature particulière des prédicats composant la définition de *grue*.

¹⁶² Une telle projection correspond par exemple à la généralisation selon laquelle "Les statues-menhirs anthropomorphes sont de la couleur du granit et de l'Age du Bronze".

généraliser, classer, ordonner, faire des hypothèses ou prévoir¹⁶³. Parmi ces facteurs, on peut notamment citer : le pragmatisme, la simplicité, la souplesse de mise en oeuvre, la polyvalence¹⁶⁴, l'économie de moyens, la puissance¹⁶⁵, mais aussi la nature de notre univers réel, la structure de nos organes de perception, l'état de nos connaissances scientifiques, etc¹⁶⁶. Nos *n*-univers habituels sont optimisés par rapport à ces différents facteurs. Mais ceci laisse valablement la place au choix d'autres systèmes de critères, en fonction des variations de l'un ou l'autre de ces paramètres.

¹⁶³ Comme le souligne Ian Hacking (1993, p. 9): "Utiliser un nom pour une espèce, c'est (entre autres choses) vouloir réaliser des généralisations et former des anticipations concernant des individus de cette espèce. La classification ne se limite pas au tri : elle sert à prédire. C'est une des leçons de la curieuse "énigme" que Nelson Goodman publia il y a quarante ans."

¹⁶⁴ Le fait qu'un même critère puisse servir à la fois de critère commun et de critère différencié (en recourant éventuellement à des taxons différents).

¹⁶⁵ C'est-à-dire le nombre de combinaisons rendues possibles.

¹⁶⁶ Cette énumération ne prétend pas être exhaustive. Une étude plus approfondie de cette question serait bien entendu nécessaire.

Conclusion

Les développements qui précèdent ont été l'occasion d'étudier plusieurs problèmes liés à l'argument de l'Apocalypse ainsi que le paradoxe de Goodman, à l'aide d'une même méthodologie, qui est celle des n-univers. En particulier, l'argument de l'Apocalypse présentait une série de problèmes qui lui étaient étroitement liés: le problème de la Belle au bois dormant, le pile ou face divin, l'expérience des segments de temps, le paradoxe de la chambre d'exécution. Dans ce cas, le nombre de ces problèmes étant important, et la nécessité de faire appel à une méthodologie s'avérait donc pressante. En ce qui concerne le paradoxe de Goodman, les variations du problème décrites dans la littérature s'avéraient nombreuses, et le besoin d'une clarification s'avérait de même essentiel.

Il est intéressant de dresser, à ce stade, un bilan de l'apport qui a résulté de l'utilisation du cadre des n-univers lors de l'étude de ces différents problèmes. Pour chacun des problèmes étudiés en effet, l'intérêt de l'utilisation des n-univers s'est manifesté à plusieurs niveaux. On étudiera ainsi tour à tour quel a été l'apport particulier de la méthodologie des n-univers à l'occasion de l'étude de chacun des problèmes rencontrés.

Considérons en premier lieu le *modèle des deux urnes*. L'utilisation des n-univers a permis de distinguer deux variations fondamentales de ce dernier modèle. Une première variation qui correspond au *modèle des deux urnes* de Bostrom a été modélisée dans le $\Omega\alpha C$, un n-univers à objets multiples et coloré, où les objets sont en relation *un* avec le critère de couleur. Une seconde variation, correspondant au *modèle des deux urnes* bicolore a pris place dans le $\Omega\alpha^*C$, un n-univers à objets multiples et coloré, où les objets sont en relation *plusieurs* avec le critère de couleur. Par ailleurs, il est également apparu que d'autres modélisations présentant un effet équivalent au modèle des deux urnes, pouvaient être modélisées dans d'autres n-univers. Ceci illustre comment la méthodologie des n-univers s'est avérée utile par le mécanisme de la *transposition*. La connaissance de la structure précise d'un modèle probabiliste particulier permet en effet, à l'aide du mécanisme de la transposition, de déterminer rapidement des modèles probabilistes qui possèdent une portée équivalente. Ceci permet de donner un support à l'intuition et de comparer plus aisément les différentes situations. En ce qui concerne en particulier le modèle des deux urnes, les transpositions suggérées par la structure de la modélisation dans le n-univers correspondant ont permis de comparer aisément des situations qui semblaient, *prima facie*, tout à fait différentes. En particulier, le modèle des deux urnes a pu être modélisé dans le $\Omega\alpha L$ ou le $\Omega\alpha^*L$, des n-univers à objets multiples et présentant une variable de localisation. Ceci a notamment permis de montrer que la modélisation du modèle des deux urnes dans le $\Omega\alpha^*L$ correspondait à l'expérience du *double amnésique* décrite par Leslie. Finalement, il est également apparu que le modèle des deux urnes pouvait également être modélisé avec un effet équivalent dans un n-univers comportant un objet unique. La situation correspondante a pu ainsi être modélisée dans un ΩL , un n-univers à objet unique et comportant une variable spatiale. A ce stade, il est également apparu qu'une telle modélisation correspondait à *l'expérience de la flèche* décrite par Leslie.

Intéressons-nous, en second lieu, à la *situation humaine correspondant à DA*. Cette dernière a pu être modélisée dans un $\Omega\alpha^*T^*C$, c'est-à-dire un n-univers à objets multiples, comportant à la fois une constante de localisation, une variable temporelle et une variable de couleur, où les objets sont à démultiplication temporelle et en relation *plusieurs* avec le critère du temps. A ce stade, la démarche qui a été poursuivie a consisté à rechercher une modélisation de la situation correspondante dans un n-univers plus simple que le $\Omega\alpha^*T^*C$. L'analyse a ainsi montré que la situation humaine correspondant à DA pouvait également être modélisée, sans perte de contenu, dans un $\Omega\alpha^*TC$, un n-univers présentant les mêmes caractéristiques que le précédent, à l'exception de la propriété de démultiplication temporelle des objets. De même, on a pu observer également que la situation humaine correspondant à DA pouvait encore être modélisée avec un effet équivalent dans un n-univers encore plus simple, le $\Omega\alpha TC$, un n-univers aux mêmes propriétés que le précédent, mais où les objets sont cette fois en relation *un* avec le critère temporel.

En ce qui concerne en particulier l'expérience des émeraudes et l'expérience Petit-Bled/Londres, l'utilisation des n-univers a permis de montrer que l'expérience des émeraudes pouvait être modélisée dans un $\Omega\alpha^*T$ et que l'expérience Petit-Bled/Londres prenait place dans un $\Omega\alpha^*L$. Il est apparu ainsi que les deux expériences pouvaient être modélisées dans un n-univers présentant la structure commune d'un $\Omega\alpha^*\beta$. Cette dernière structure a ainsi permis de reconstruire le modèle probabiliste correspondant aux deux expériences dans un $\Omega\alpha^*C$. Cette dernière modélisation a présenté un intérêt particulier car elle prenait place dans le même type de n-univers que celui correspondant au modèle des deux urnes. Une telle propriété a finalement permis de comparer aisément les deux modèles, en mettant en évidence les points communs et les différences entre ces derniers.

S'agissant également de l'expérience du *pile ou face divin*, la démarche a consisté à modéliser cette dernière dans un n-univers dont les caractéristiques sont identiques à celui dans lequel se trouvent modélisés à la fois le modèle des deux urnes, l'expérience des émeraudes et l'expérience Petit-Bled/Londres. L'analyse a ainsi montré qu'une telle modélisation prenait place dans le $\Omega\alpha^*C$. La modélisation dans un n-univers identique de chacune de ces expériences a ainsi permis de faciliter les comparaisons, en rendant ainsi plus aisée la détermination des points communs et des différences entre les différentes expériences. La comparaison avec le modèle des deux urnes bicolore a notamment permis de montrer que le raisonnement (GCT+) était basé sur une application inconditionnelle du modèle à majoration, dans des circonstances où les conditions de ce dernier modèle n'étaient réunies que de manière hypothétique.

En ce qui concerne le problème de la *Belle au bois dormant*, la modélisation en termes de n-univers a permis de montrer que l'expérience prenait place dans un ΩT^* , un n-univers à objet unique où ce dernier possède une propriété de démultiplication par rapport au critère temporel. L'analyse a ensuite montré qu'il était possible de restituer le problème de la Belle au bois dormant dans le $\Omega\alpha^*C$, permettant ainsi une comparaison avec l'expérience des émeraudes, l'expérience Petit-Bled/Londres et le pile ou face divin. Il s'est ainsi avéré que la situation correspondant au problème de la Belle au bois dormant était finalement tout à fait identique à une version de l'expérience du pile ou face divin décrite par Bostrom, qui ne comporterait que l'étape (a). Le problème de la Belle au bois dormant s'est également révélé présenter une structure identique à la version du pile ou face divin décrite par Leslie.

S'agissant encore de l'expérience des *segments de temps*, cette dernière a pu être modélisée dans un ΩT^*C , un n-univers à objet unique, temporel et coloré, où l'objet unique présente une propriété de démultiplication par rapport au critère temporel. L'analyse a également montré que d'autres variations de l'expérience pouvaient prendre place respectivement dans un ΩT^*L et un ΩT^*S . Plus encore, il s'est avéré que l'expérience des segments de temps pouvait également être modélisée dans $\Omega\alpha^*T^*$, un n-univers temporel à objets multiples, où les objets présentent une propriété de démultiplication par rapport au critère temporel et se trouvent en relation *plusieurs* avec ce dernier. Une telle modélisation s'est montrée intéressante car il s'agit également d'un n-univers où la situation correspondant à DA peut prendre place. Enfin, l'analyse basée sur les n-univers a finalement montré que l'expérience des segments de temps était isomorphe au modèle de l'objet à incrémentation-temps.

L'apport des n-univers a également concerné l'analyse de *l'argument de l'Apocalypse*. Ceci a permis d'illustrer comment la méthodologie des n-univers se révèle utile pour l'analyse des arguments basés sur une analogie avec une urne. L'argument de l'Apocalypse appartient en effet à cette dernière catégorie d'arguments. Et plusieurs des problèmes étudiés présentaient, on l'a vu, la structure d'un argument basé sur une analogie entre une situation réelle et un modèle probabiliste donné. Tel était également le cas pour l'expérience des segments de temps. Et le problème posé par les arguments fondés sur une analogie avec une urne est précisément de déterminer si l'analogie se révèle valable. Dans le cas contraire, l'argument constitue un cas particulier d'argument fallacieux: la fausse analogie. Dans ce contexte, le cadre des n-univers a essentiellement constitué ici un outil méthodologique permettant de déterminer avec précision si l'analogie sur laquelle est fondée chacun des deux arguments que constituent l'argument de l'apocalypse et l'expérience des segments de temps, est valablement établie. Il s'est donc agi de définir tout d'abord des modèles probabilistes, pour lesquels le

raisonnement correspondant n'est pas controversé. Il s'est agi ensuite de déterminer avec précision, pour chacune des situations concrètes correspondantes, quel était le modèle probabiliste correspondant. C'est à ce stade particulier que le cadre des n-univers a trouvé à s'appliquer. Il est ainsi apparu comme un outil méthodologique permettant de définir le modèle probabiliste correspondant à une situation concrète donnée, par l'analyse du n-univers correspondant. Concernant en particulier l'argument de l'Apocalypse, l'analyse à l'aide des n-univers a finalement permis de montrer que la situation correspondant à DA-objets ne pouvait être considérée de manière certaine, comme isomorphe avec le modèle de l'urne à incrémentation-objet, ni avec celui des deux urnes à incrémentation-objet. Il en est ainsi résulté une solution à ma connaissance nouvelle et basée sur une troisième voie, conduisant à un raisonnement basé sur le fait qu'aucune des deux analogies précédentes n'est véritablement fondée. La solution finalement proposée repose sur une analogie alternative, qui s'établit entre la situation correspondant à DA-objets et le modèle des deux urnes à incrémentation-objet hypothétique.

Enfin, il convient de mentionner l'apport de l'outil méthodologique que constituent les n-univers dans le cadre de l'analyse du *paradoxe de Goodman*. Il est apparu en premier lieu que l'on pouvait modéliser la situation correspondante dans un n-univers à objets multiples ou bien dans un n-univers à objet unique. En privilégiant le n-univers le plus simple, c'est-à-dire le n-univers à objet unique, il est ensuite apparu que l'énoncé de Goodman pouvait être modélisé dans deux types de n-univers différents: d'une part dans un ΩT^*C , un n-univers à objet unique à démultiplication temporelle et coloré; et d'autre part dans un ΩT^*CL . Dans le ΩT^*C , l'analyse a ensuite montré que les conditions de l'induction n'étaient pas satisfaites et qu'il s'agissait en fait d'une pseudo-induction qui masquait véritablement un raisonnement de type déductif. Dans le ΩT^*CL en revanche, il est apparu que l'induction pouvait valablement prendre place et que la conclusion résultant du raisonnement inductif était fondée. En ce qui concerne le paradoxe de Goodman, le cadre des n-univers a ainsi permis de clarifier les données correspondant à la situation dans laquelle le paradoxe prend place. L'analyse a ainsi montré que le paradoxe de Goodman trouvait son origine dans une définition ambiguë de l'univers de référence correspondant à l'énoncé du problème. Il convenait donc de lever cette ambiguïté. Ceci a pu être réalisé en modélisant la situation correspondante dans deux n-univers structurellement différents. Une telle distinction a ainsi permis de mettre en évidence deux variations distinctes du problème de Goodman, conduisant alors à une résolution de nature fondamentalement différente pour chacune d'entre elles.

On le voit finalement, l'outil méthodologique que constituent les n-univers a permis ici de clarifier les données des différents problèmes étudiés tout au long de la présente thèse. Qu'il s'agisse de clarifier les problèmes essentiellement basés sur des analogies avec une urne, de déterminer à l'aide d'une transposition si une situation donnée est structurellement identique à une autre, ou bien enfin de déterminer les variations prenant place dans des univers de référence structurellement différents, leur rôle s'est finalement révélé essentiel à chacune des étapes de la présente discussion.

Références

- Arntzenius, F. (2002) *Reflections on Sleeping Beauty*, *Analysis*, 62-1, 53-62
- Barrow, J.D. & Tipler, F.J. (1986) *The anthropic cosmological principle*, New York: Oxford University Press
- Bartha, P. & Hitchcock, C. (1999) *No One Knows the Date of the Hour: An Unorthodox Application of Rev. Bayes's Theorem*, *Philosophy of Science (Proceedings)*, 66, 229-353
- Bartha, P. & Hitchcock, C. (2000) *The Shooting-Room Paradox and Conditionalizing on Measurably Challenged Sets*, *Synthese*, 108 (3), 403-437
- Bertolet, R. J. (1976) *On the Merits of Entrenchment*, *Analysis*, 37, 29-31
- Bostrom, N. (1997) *Investigations into the Doomsday argument*, Preprint, <http://www.anthropic-principles.com/preprints/inv/investigations.html>
- Bostrom, N. (1999) *The Doomsday Argument is Alive and Kicking*, *Mind*, 108 (431), 539-50
- Bostrom, N. (2000) *Observation Selection Effects and Probability*, Ph.D. Dissertation, London School of Economics, Preprint, <http://www.anthropic-principle.com/preprints>
- Bostrom, N. (2001) *The Epistemological Mystique of Self-locating Belief*, preprint, <http://www.anthropic-principles.com/preprints>
- Bostrom, N. (2001) *The Doomsday Argument, Adam & Eve, UN++, and Quantum Joe*. *Synthese*, 127, 359-387
- Bricker, P. (1996) *Isolation and Unification: The Realist Analysis of Possible Worlds*, *Philosophical Studies*, 84, 225-238
- Carter, B. (1974) *Large number coincidences and the anthropic principle in cosmology*, in *Confrontation of cosmological theories with data*, M.S. Longair, Ed., Reidel: Dordrecht, 291-298
- Carter, B. (1990) *Large Number Coincidences and the Anthropic Principle in Cosmology*, dans *Physical Cosmology and Philosophy*, J. Leslie, Ed., Macmillan Publishing Company
- Chambers, T. (2001) *Do Doomsday's Proponents Think We Were Born Yesterday?*, *Philosophy*, 76, 443-50
- Ćirković, M. (2001) *Is Many Likelier than Few? A Critical Assessment of the Self-Indication Assumption*, Preprint, <http://anthropic-principle.com/preprints/SIA1.pdf>
- Delahaye, J-P. (1996) *Recherche de modèles pour l'argument de l'apocalypse de Carter-Leslie*, Manuscrit
- Dieks, D. (1992) *Doomsday – Or: the Dangers of Statistics*, *Philosophical Quarterly*, 42 (166), 78-84.
- Eckhardt, W. (1992) *A Shooting-Room View of Doomsday*, *Journal of Philosophy*, 94 (5), 244-259
- Eckhardt, W. (1993) *Probability Theory and the Doomsday Argument*, *Mind*, 102 (407), 483-88
- Elga, A. (2000) *Self-locating Belief and the Sleeping Beauty Problem*, *Analysis*, 60, 143-147
- Fain, H. (1967) *The Very Thought of Grue*, *Philosophical Review*, 76, 61-73
- Forrest, P. (1986) *Ways Worlds Could Be*, *Australasian Journal of Philosophy*, 64, 15-24
- Foster, M. and Martin, M. (eds.) (1966) *Probability, Confirmation, and Simplicity*, Odyssey Press
- Franceschi, P. (1998) *Une solution pour l'argument de l'apocalypse*, *Canadian Journal of Philosophy*, 28, 227-46
- Franceschi, P. (1999) *Comment l'urne de Carter et Leslie se déverse dans celle de Hempel*, *Canadian Journal of Philosophy*, 29, 139-56
- Franceschi, P. (2001) *Une solution pour le paradoxe de Goodman*, *Dialogue*, 40, 99-123
- Gendler, T.S.(1998) *Galileo and the Indispensability of Scientific Thought Experiment*, *British Journal for the Philosophy of Science*, 49-3, 397-424
- Gott, R.J. (1993) *Implications of the Copernican Principle for our Future Prospects*, *Nature*, 363, 315-319
- Goodman, N. (1946) *A Query On Confirmation*, *Journal of Philosophy*, 43, 383-385, dans *Problems and Projects*, Indianapolis, Bobbs-Merrill, 1972, p. 363-366
- Goodman, N. (1954) *Fact, Fiction and Forecast*, Cambridge, MA, Harvard University Press
- Goodman, N. (1978) *Ways of Worldmaking*, Indianapolis: Hackett
- Gott, R. (1993) *Implications of the Copernican Principle for our Future Prospects*, *Nature*, 363, 315-319

- Gott, R.J. (1997) *Future prospects discussed*, Nature, 368, 108
- Grünbaum, A. (1973) *Philosophical Problems of Space and Time*, Dordrecht, Holland et Boston, MA: D. Reidel Publishing Company
- Haack, S. (1993) *Evidence and Inquiry*, Oxford, Cambridge, MA, Blackwell
- Hacking, I. (1993) *Le plus pur nominalisme*, Combas, L'éclat
- Hardin, C. L. (1993), *Color for Philosophers*, Indianapolis, Ind.: Hackett
- Horowitz, T. & G. Massey (eds.) (1991) *Thought Experiments in Science and Philosophy*, Savage MD: Rowman and Littlefield
- Horwich, P. (1982) *Probability and Evidence*, Cambridge: Cambridge University Press
- Hugues, C. (1997) *Same-Kind Coincidence and the Ship of Theseus*, Mind, 106, 53-67
- Jackson, F. (1975) 'Grue', *Journal of Philosophy*, 72, 113-131
- Kopf, T., Krtous, P, and Page, D. N. (1994) *Too Soon for Doom Gloom?*, Physics Preprints Archive, gr-qc/9407002
- Korb, K. & Oliver, J. (1998) *A Refutation of the Doomsday Argument*, Mind, 107, 403-410
- Leibniz, G. W. (1703-1716) *Principes de la nature et de la grâce, La Monadologie, et autres textes*, Flammarion: Paris 1996
- Leibniz, G. W. (1704-1765) *Nouveaux essais sur l'entendement humain*, Flammarion: Paris 1966
- Leslie, J. (1989) *Universes*, London: Routledge
- Leslie, J., (1990) ed. *Physical cosmology and philosophy*, Collier Macmillan: New York
- Leslie, J. (1992) *Time and the Anthropic Principle*, Mind, 101, 521-540
- Leslie, J. (1993) *Doom and Probabilities*, Mind, 102, 489-491
- Leslie, J. (1996) *The End of the World: the science and ethics of human extinction*, London, Routledge
- Lewis, D. (1986) *On the Plurality of Worlds*, Oxford: Blackwell
- Lewis, D. (1997), *Naming the Colors*, Australasian Journal of Philosophy, 75, 325-342
- Lewis, D. (2001) *Sleeping Beauty: Reply to Elga*, Analysis, 61, 171-176
- Locke, J. (1975) *An Essay Concerning Human Understanding*, ed. P. Niddich, Oxford: Clarendon Press
- McTaggart, J. M. E. (1908) *The Unreality of Time*, Mind, 68, 457-484
- Maudlin, T. (1990) *Substances and Spacetimes: What Aristotle Would have Said to Einstein*, Studies in the History and Philosophy of Science, 21, 531-561
- Maud, B. (1995) *Colors: Their Nature and Representation*, Cambridge: Camb.University Press
- Miscevic, N. (1992) *Mental Models and Thought Experiments*, International Studies in the Philosophy of Science, 3, 215-226
- Monton, B. (2002) *Sleeping Beauty and the Forgetful Bayesian*, Analysis, 62, 47-53
- Monton, B. (2002) *The Doomsday Argument without Knowledge of Birth Rank*, Preprint, <http://www.anthropic-principle.com/preprints>
- Nassau, K. (1983) *The Physics and Chemistry of Color*, New York: Wiley
- Olum, K. (2000) *The Doomsday Argument and the Number of Possible Observers*, Preprint, <http://xxx.lanl.gov/abs/gr-qc/0009081>
- Rosenkrantz, R. D. (1982) *Does the Philosophy of Induction Rest on a Mistake?*, Journal of Philosophy, 79, 78-97
- Salmon, W. (1966) *The Foundations of Scientific Inference*, University of Pittsburgh Press
- Skyrms, B. (1976) *Possible Worlds, Physics and Metaphysics*, Philosophical Studies, 30, 323-332
- Slots, M. A. (1974) *Entrenchment and Validity*, Analysis, 34, 204-207
- Smith, Q. (1994) *Anthropic explanations in cosmology*. Australasian Journal of Philosophy, 72, 371-382
- Sober, E. (2001) *A Critique of Two Versions of the Doomsday Argument - Gott's Line and Leslie's Wedge*, Preprint, <http://www.anthropic-principle.com/preprints>
- Sorensen, R. (1992) *Thought Experiments*, Oxford: Oxford University Press
- Sowers, G. F. (2002) *The Demise of the Doomsday Argument*, Mind, 111, 37-45
- Stalnaker, R. (1976) *Possible Worlds*, Noûs, 10, 65-75
- Tye, M. (2000) *Consciousness, Color, and Content*, Cambridge, Mass.: MIT/Bradford
- Ullian, J.S. (1961) *More one 'Grue' and Grue*, Philosophical Review, 70, 386-389
- Weingard, R. (1976) *On the Unity of Space*, Philosophical Studies, 29, 15-20
- Westphal, J. (1987) *Color: A Philosophical Introduction*, Oxford: Blackwell Publisher

Will, F. (1947) *Will the Future Be Like the Past?*, *Mind*, 56, 332-347
Zabludowski, A. (1974) *Concerning a Fiction About How Facts Are Forecast*, *Journal of Philosophy*,
71, 97-112

INDEX

- $\Omega\alpha^*L$, 16, 19, 35, 36, 49, 125, 135
 $\Omega\alpha^*LC$, 19
 $\Omega\alpha^*T$, 14, 15, 17, 18, 49, 83, 85, 95, 107, 109, 125
 $\Omega\alpha^*T^*$, 15, 18, 83, 85, 95, 107, 109, 125
 $\Omega\alpha^*T^*C$, 18, 83, 85, 95, 125
 $\Omega\alpha L$, 16, 19, 20, 22, 34, 37, 125, 135
 $\Omega\alpha L\alpha T$, 20, 22
 $\Omega\alpha L\alpha T^*$, 22
 $\Omega\alpha L\alpha T^*C$, 22
 $\Omega\alpha LC$, 19
 $\Omega\alpha T$, 13, 15, 17, 18, 21, 76, 95, 125
 $\Omega\alpha T^*$, 15, 18, 21, 76, 95
 $\Omega\alpha T^*C$, 18, 95
 $\Omega\alpha TC$, 17, 95, 125
 ΩL , 14, 15, 18, 19, 21, 22, 36, 37, 39, 66, 135
 ΩLC , 18
 ΩLT , 19, 21, 22
 ΩLT^* , 21, 22
 ΩLT^*C , 22
 ΩT , 13, 16, 17, 66, 73, 74, 75, 76, 77, 78, 79, 80, 118, 127
 ΩT^* , 17, 66, 73, 74, 75, 76, 77, 78, 79, 80, 127
 ΩT^*C , 17, 74, 75, 76, 77, 78, 80, 127
 ΩTC , 16
Adam paresseux, 2, 102, 103, 104
aléatoire, 24, 25, 31, 37, 39, 88, 92, 93, 94, 96, 97, 100
ambiguïté, 2, 4, 7, 87, 118, 127
analogie, 2, 4, 5, 6, 7, 12, 24, 25, 29, 36, 38, 43, 44, 46, 47, 78, 80, 81, 92, 93, 94, 95, 96, 98, 99, 100, 101, 126
analogie avec une urne, 2, 4, 5, 6, 7, 12, 36, 38, 46, 126
argument, 2, 4, 5, 6, 7, 12, 24, 27, 29, 46, 53, 59, 60, 61, 71, 77, 81, 83, 85, 87, 92, 93, 101, 103, 109, 125, 126, 128
argument de l'Apocalypse, 2, 4, 101, 125, 126
Bartha, Paul, 30, 105, 106, 107, 110, 128
Bostrom, Nick, 3, 5, 24, 25, 27, 29, 30, 31, 32, 40, 43, 52, 53, 54, 55, 56, 57, 60, 62, 69, 70, 84, 94, 97, 101, 102, 103, 125, 126, 128, 135
boule, 14, 15, 24, 25, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 49, 50, 51, 54, 55, 56, 57, 58, 59, 61, 63, 66, 67, 68, 69, 70, 72, 73, 76, 78, 79, 80, 81, 87, 88, 89, 90, 91, 92, 94, 96, 99, 100, 111, 112, 113, 114, 115, 121
Carter, Brandon, 3, 24, 26, 35, 36, 43, 94, 100, 128
Ćirković, Milan, 52, 128
classe de référence, 4, 25, 26, 27, 28, 77, 83, 84, 86, 87, 96, 97, 98, 101, 136
conclusion, 6, 27, 58, 59, 61, 71, 77, 78, 81, 82, 83, 84, 86, 92, 93, 94, 96, 104, 105, 111, 113, 114, 120, 123, 127
conditions de l'induction, 5, 114, 117, 118, 119, 120, 121, 122, 123, 127
Conseil du serpent, 102, 103
constante, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 32, 34, 35, 37, 38, 43, 44, 45, 47, 48, 65, 72, 107, 125
couleur, 9, 10, 11, 16, 17, 18, 19, 22, 30, 32, 33, 34, 38, 39, 41, 43, 44, 45, 48, 50, 55, 56, 57, 66, 67, 68, 69, 70, 74, 75, 76, 78, 79, 80, 85, 112, 113, 115, 116, 118, 119, 121, 122, 123, 125
critère, 2, 4, 5, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 27, 31, 32, 33, 34, 35, 44, 50, 55, 65, 67, 68, 69, 74, 75, 76, 78, 82, 83, 85, 86, 87, 106, 107, 108, 109, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 134
critère-constante, 9, 10, 11, 13, 15, 17, 19, 31
critère-variable, 9, 10, 11, 12, 13, 14, 19, 20, 21, 22, 23, 50, 67, 74, 115, 117, 118, 119, 121
DA-objet, 83, 85, 86, 95, 96, 98, 100, 104, 108, 127, 136
DA-temps, 77, 83, 84, 85, 86, 95, 99, 108, 136
décalage bayésien, 25, 27, 29, 31, 33, 39, 40, 42, 43, 54, 56, 71, 72, 77, 80, 82, 83, 84, 85, 86, 91, 92, 94, 96, 97, 98, 99, 100, 101, 102, 104, 108
Delahaye, Jean-Paul, 1, 3, 71, 72, 73, 81, 87, 128

démultiplication, 2, 4, 10, 11, 12, 15, 16, 17, 21, 22, 34, 35, 44, 48, 50, 65, 66, 70, 72, 74, 75, 76, 78, 106, 107, 125, 126, 127

déterministe, 4, 26, 44, 88, 96, 98, 99, 104, 134

Dieks, Dennis, 128

Dieu et l'humanité-temps, 99

durée, 10, 14, 15, 44, 66, 71, 72, 73, 76, 84, 86, 87, 89, 92, 93, 95, 99

Eckhardt, William, 25, 26, 88, 92, 93, 95, 97, 98, 99, 100, 101, 128

Elga, Adam, 64, 65, 128, 129

étape d'analogie, 6

expérience de la flèche, 36, 37, 38, 125, 135

expérience de la loterie, 30, 31, 32, 46

expérience de pensée, 4, 8, 9, 34

expérience des émeraudes, 4, 46, 47, 48, 49, 50, 51, 56, 126, 135

expérience des segments de temps, 4, 71, 72, 73, 74, 75, 76, 77, 78, 80, 81, 83, 87, 125, 126, 135

expérience Petit-Bled/Londres, 4, 46, 47, 48, 49, 50, 51, 56, 126, 135

généralisation, 8, 111, 114, 123

Goodman, Nelson, 2, 4, 5, 8, 111, 112, 115, 116, 118, 122, 124, 125, 127, 128

Gott, Richard, 128, 129

grue, 112, 114, 116, 119, 120, 121, 122, 123

Hacking, Ian, 124, 129

hasard, 24, 25, 29, 31, 38, 40, 51, 63, 80, 88, 91, 93, 96, 97, 98, 99

Hitchcock, Christopher, 30, 105, 106, 107, 110, 128

homo sapiens, 77, 86, 87, 96, 99

homo sapiens sapiens, 77, 86, 87

HR, 25, 80, 93, 97

indéterministe, 4, 26, 44, 98, 101, 104, 134

induction, 111, 112, 113, 114, 116, 117, 118, 119, 120, 121, 123, 127, 136

isomorphe, 81, 82, 83, 96, 99, 126, 127

Korb, Kevin, 29, 71, 72, 73, 75, 77, 81, 87, 129

Leslie, John, 3, 4, 5, 24, 26, 29, 30, 31, 32, 35, 36, 37, 38, 43, 46, 47, 48, 53, 57, 59, 60, 69, 70, 85, 87, 93, 94, 98, 100, 105, 106, 108, 109, 125, 126, 128, 129

lieu, 2, 4, 5, 7, 10, 12, 13, 14, 15, 16, 17, 18, 24, 31, 32, 34, 35, 38, 39, 43, 45, 47, 48, 49, 50, 51, 55, 56, 57, 58, 60, 64, 65, 66, 67, 68, 71, 72, 73, 74, 75, 83, 91, 97, 98, 100, 102, 103, 106, 107, 108, 109, 114, 115, 118, 121, 125, 127

localisation, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 31, 34, 35, 39, 44, 45, 47, 49, 72, 75, 76, 79, 115, 118, 121, 122, 125

loi, 8, 9

méthodologie, 2, 4, 7, 8, 125, 126

modèle de l'objet à incrémentation-espace, 79, 91

modèle de l'objet à incrémentation-temps, 78, 80, 81, 82, 83, 88, 91, 95, 126

modèle de l'urne à incrémentation, 87, 88, 89, 90, 91, 92, 93, 94, 95, 98, 99, 100, 101, 127, 136

modèle de l'urne à incrémentation-temps, 87, 88, 91, 92, 93, 94, 99, 100

modèle des deux urnes, 2, 4, 5, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 51, 54, 55, 56, 59, 60, 62, 68, 69, 72, 73, 78, 80, 81, 88, 90, 91, 92, 94, 95, 96, 98, 99, 100, 101, 125, 126, 127, 135, 136

modèle des deux urnes à décalage, 5, 40, 41, 42, 135

modèle des deux urnes à incrémentation, 41, 90, 91, 92, 94, 95, 96, 98, 99, 100, 101, 127, 136

modèle des deux urnes à incrémentation-objet, 91, 92, 94, 95, 96, 98, 100, 101, 127

modèle des deux urnes à incrémentation-objet hypothétique, 100, 101, 127

modèle des deux urnes à incrémentation-temps, 90, 91, 95, 99

modèle probabiliste, 2, 5, 6, 7, 14, 25, 29, 30, 34, 36, 37, 38, 46, 47, 49, 50, 51, 62, 67, 68, 70, 78, 80, 81, 89, 96, 99, 100, 125, 126

Monton, Bradley, 3, 64, 68, 129

objection à DA, 77, 98, 105, 107, 108, 135, 136

objet unique, 10, 13, 14, 15, 16, 17, 18, 19, 21, 22, 37, 38, 39, 50, 65, 66, 67, 70, 72, 73, 74, 75, 76, 77, 78, 83, 87, 91, 95, 118, 119, 125, 126, 127, 134, 135

objets multiples, 4, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 31, 32, 34, 35, 37, 38, 39, 43, 44, 45, 47, 48, 50, 55, 62, 67, 68, 69, 70, 76, 83, 84, 85, 86, 88, 89, 90, 91, 100, 106, 107, 108, 125, 126, 127, 134, 135, 136

observateur, 30, 32, 52, 61, 73

Oliver, Jonathan, 29, 71, 72, 73, 75, 77, 81, 87, 129

Olum, Ken, 52, 53, 54, 55, 57, 59, 63, 129
 paradoxe, 2, 3, 4, 5, 7, 9, 12, 71, 72, 75, 78, 87, 105, 106, 107, 109, 111, 112, 119, 125, 127, 128, 136
 paradoxe de Goodman, 2, 3, 4, 5, 7, 9, 12, 111, 125, 127, 128, 136
 paradoxe de la chambre d'exécution, 2, 5, 7, 105, 106, 107, 109, 125, 136
 philosophe présomptueux, 60, 61, 62
 pile ou face divin, 2, 4, 40, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 68, 69, 70, 125, 126, 135
 position, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 32, 34, 35, 37, 39, 43, 44, 47, 48, 49, 50, 56, 66, 73, 74, 75, 76, 78, 79, 80, 81, 82, 85, 86, 91, 92, 93, 94, 98, 106, 107, 109, 110
 position spatiale, 11, 16, 18, 19, 20, 21, 22, 23, 34, 35, 37, 39, 43, 49, 50, 56, 75, 79, 80, 107
 position temporelle, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 32, 34, 44, 48, 50, 66, 73, 74, 75, 76, 79, 80, 81, 82, 85, 91, 92, 93, 94, 98, 106, 109, 110
 prémisse, 6, 25
 probabilité, 24, 27, 29, 30, 31, 36, 37, 39, 40, 49, 50, 51, 53, 54, 55, 56, 58, 59, 60, 61, 62, 65, 68, 70, 71, 72, 73, 75, 77, 78, 80, 83, 84, 88, 89, 92, 96, 99, 100, 101, 102, 103, 105, 109, 110
 probabilité a posteriori, 24, 27, 29, 31, 71, 75
 probabilité a priori, 31
 problème de la Belle au bois dormant, 2, 4, 7, 64, 65, 66, 67, 68, 69, 70, 125, 126
 problème de la définition de la classe de référence, 3, 5, 26, 28, 85, 86, 87, 96, 134
 propriété, 4, 10, 12, 15, 42, 44, 50, 51, 60, 65, 66, 73, 74, 76, 78, 79, 80, 81, 91, 101, 111, 112, 114, 117, 118, 119, 121, 122, 125, 126
 protocole, 7, 31, 42, 59
 relation, 2, 4, 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 32, 33, 34, 35, 36, 41, 43, 44, 47, 48, 49, 50, 55, 62, 67, 68, 69, 70, 76, 83, 86, 105, 106, 107, 113, 125, 126
 relation plusieurs, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 33, 35, 41, 44, 48, 49, 50, 55, 62, 67, 68, 69, 70, 76, 83, 86, 106, 107, 125, 126
 relation un, 2, 4, 10, 11, 12, 13, 15, 17, 18, 19, 20, 21, 22, 32, 34, 36, 41, 43, 44, 76, 125
 SIA, 52, 57, 59, 60, 61, 62, 63, 135
 situation humaine correspondant à DA, 43, 44, 92, 93, 94, 95, 101, 125, 135
 situation réelle, 2, 5, 7, 8, 30, 35, 46, 47, 126
 Sober, Elliott, 93, 129
 sous-univers, 8, 9, 11, 12, 44, 74, 106
 Sowers, George, 89, 92, 93, 94, 95, 98, 99, 100, 101, 129, 136
 SSA, 25, 31, 51, 63, 93, 97, 102, 134
 taxon, 10, 11, 41, 74, 113
 temporel, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 31, 32, 38, 43, 44, 47, 48, 50, 55, 65, 72, 73, 74, 75, 76, 77, 83, 85, 86, 87, 91, 92, 94, 106, 107, 108, 109, 115, 117, 118, 119, 121, 123, 125, 126, 134, 135, 136
 temps, 5, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 26, 31, 34, 35, 44, 48, 49, 50, 55, 56, 59, 67, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 83, 84, 85, 86, 87, 88, 89, 90, 91, 93, 94, 95, 96, 97, 99, 100, 101, 106, 107, 108, 109, 112, 113, 114, 115, 116, 118, 119, 120, 121, 122, 123, 125, 126, 135
 tirage, 25, 31, 37, 39, 40, 41, 42, 69, 74, 80, 88, 90, 93, 94, 96, 97, 99, 100, 101, 102, 103, 109
 transposition, 39, 125, 127
 univers de référence, 8, 9, 117, 118, 119, 121, 122, 127
 urne, 2, 3, 5, 6, 9, 11, 13, 14, 15, 24, 25, 29, 30, 31, 32, 33, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 49, 50, 54, 55, 56, 57, 58, 59, 61, 62, 63, 66, 67, 68, 69, 70, 72, 73, 76, 80, 85, 87, 88, 89, 90, 91, 92, 93, 94, 95, 97, 98, 99, 100, 101, 111, 112, 115, 126, 127, 128, 136
 variable, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 31, 33, 34, 35, 38, 39, 44, 45, 48, 50, 72, 74, 75, 76, 89, 106, 107, 115, 117, 125

TABLE DES MATIERES

RESUME	2
REMERCIEMENTS	3
INTRODUCTION	4
CHAPITRE 1. APPROCHE METHODOLOGIQUE: LES N-UNIVERS.....	7
1.1. METHODOLOGIE	7
1.2. DEFINITION	8
1.2.1. Démultiplication.....	10
1.2.2. Relation un ou plusieurs des objets multiples avec un critère donné.....	10
1.2.3. Sous-univers	11
1.2.4. n-univers spécialisés.....	12
1.3. N-UNIVERS CARACTERISTIQUES	12
1.3.1. Le n-univers alocalisé et atemporel à objets multiples: $\Omega\alpha$	12
1.3.2. Le n-univers alocalisé et temporel à objet unique: ΩT	13
1.3.3. Le n-univers alocalisé et temporel à objets multiples: $\Omega\alpha T$	13
1.3.4. Le n-univers alocalisé et temporel à objets multiples*: $\Omega\alpha^*T$	14
1.3.5. Le n-univers alocalisé et temporel* à objet unique: ΩT^*	14
1.3.6. Le n-univers alocalisé et temporel* à objets multiples: $\Omega\alpha T^*$	14
1.3.7. Le n-univers alocalisé et temporel* à objets multiples*: $\Omega\alpha^*T^*$	15
1.3.8. Le n-univers localisé et temporel à objet unique: ΩL	15
1.3.9. Le n-univers localisé et atemporel à objets multiples: $\Omega\alpha L$	16
1.3.10. Le n-univers localisé et atemporel à objets multiples*: $\Omega\alpha^*L$	16
1.3.11. Le n-univers alocalisé, coloré et temporel à objet unique: ΩTC	16
1.3.12. Le n-univers alocalisé, coloré et temporel à objets multiples: $\Omega\alpha TC$	16
1.3.13. Le n-univers alocalisé, coloré et temporel à objets multiples*: $\Omega\alpha^*TC$	17
1.3.14. Le n-univers alocalisé, coloré et temporel* à objet unique: ΩT^*C	17
1.3.15. Le n-univers alocalisé, coloré et temporel* à objets multiples: $\Omega\alpha T^*C$	17
1.3.16. Le n-univers alocalisé, coloré et temporel* à objets multiples*: $\Omega\alpha^*T^*C$	18
1.3.17. Le n-univers atemporel, localisé et coloré à objet unique: ΩLC	18
1.3.18. Le n-univers atemporel, localisé et coloré à objets multiples: $\Omega\alpha LC$	19
1.3.19. Le n-univers atemporel, localisé et coloré à objets multiples*: $\Omega\alpha^*LC$	19
1.3.20. Le n-univers localisé et temporel à objet unique: ΩLT	19
1.3.21. Le n-univers localisé et temporel à objets multiples: $\Omega\alpha L\alpha T$	19
1.3.22. Le n-univers localisé et temporel à objets multiples*: $\Omega\alpha^*L\alpha T$	20
1.3.23. Le n-univers localisé et temporel à objets multiples*: $\Omega\alpha^*L\alpha^*T$	20
1.3.24. Le n-univers localisé et temporel* à objet unique: ΩLT^*	21
1.3.25. Le n-univers temporel* et localisé à objets multiples: $\Omega\alpha L\alpha T^*$	21
1.3.26. Le n-univers temporel* et localisé à objets multiples*: $\Omega\alpha^*L\alpha^*T^*$	21
1.3.27. Le n-univers localisé et temporel* à objet unique: ΩLT^*C	22
1.3.28. Le n-univers localisé et temporel* à objets multiples: $\Omega\alpha L\alpha T^*C$	22
1.3.29. Le n-univers localisé et temporel* à objets multiples*: $\Omega\alpha^*L\alpha^*T^*C$	23
CHAPITRE 2. L'ARGUMENT DE L'APOCALYPSE	24
2.1. LE PROBLEME	24
2.2. LE PRINCIPE SSA.....	25
2.3. VERSION DETERMINISTE OU INDETERMINISTE DE DA	26
2.4. LE PROBLEME DE LA DEFINITION DE LA CLASSE DE REFERENCE.....	26

CHAPITRE 3. LE MODELE DES DEUX URNES.....	29
3.1. DESCRIPTION	29
3.2. MODELISATION DANS LE N-UNIVERS COLORE A OBJETS MULTIPLES ($\Omega\alpha C$ OU $\Omega\alpha^*C$).....	31
3.2.1. <i>Le modèle des deux urnes de Bostrom et le $\Omega\alpha C$</i>	31
3.2.2. <i>Le modèle des deux urnes bicolore et le $\Omega\alpha^*C$</i>	32
3.3. MODELISATION DANS D'AUTRES N-UNIVERS.....	33
3.3.1. <i>Modélisation du modèle des deux urnes dans le $\Omega\alpha L$</i>	34
3.3.2. <i>Modélisation du modèle des deux urnes dans le $\Omega\alpha^*L$</i>	35
3.3.3. <i>Modélisation du modèle des deux urnes dans le ΩL et l'expérience de la flèche</i>	36
3.3.4. <i>Modélisation du modèle des deux urnes dans un ΩC</i>	39
3.4. LA VERSION INTEGRALE DU MODELE DES DEUX URNES.....	39
3.5. LE MODELE DES DEUX URNES A DECALAGE (VERSION INTEGRALE)	40
CHAPITRE 4. LA SITUATION CORRESPONDANT A DA	43
4.1. MODELISATION DE LA SITUATION CORRESPONDANT A DA DANS LE N-UNIVERS TEMPOREL A OBJETS MULTIPLES ($\Omega\alpha^*T^*C$)	43
4.2. VARIATIONS DE LA SITUATION HUMAINE CORRESPONDANT A DA DANS D'AUTRES N-UNIVERS	44
4.2.1. <i>Variation dans un $\Omega\alpha^*TC$</i>	44
4.2.2. <i>Variation dans un $\Omega\alpha TC$</i>	44
CHAPITRE 5. L'EXPERIENCE DES EMERAUDES ET L'EXPERIENCE PETIT- BLED/LONDRES	46
5.1. DESCRIPTION	46
5.2. MODELISATION EN TERMES DE N-UNIVERS	47
5.2.1. <i>Modélisation de l'expérience des émeraudes dans le n-univers temporel à objets multiples ($\Omega\alpha^*T$)</i>	47
5.2.2. <i>Modélisation de l'expérience Petit-Bled/Londres dans le n-univers localisé à objets multiples ($\Omega\alpha^*L$)</i>	48
5.3. MODELES PROBABILISTES CORRESPONDANTS	49
5.4. LEÇONS DES DEUX EXPERIENCES	51
CHAPITRE 6. LE PILE OU FACE DIVIN ET SIA	52
6.1. LE PRINCIPE SIA	52
6.2. LE PILE OU FACE DIVIN	52
6.3. MODELISATION EN TERMES DE N-UNIVERS	54
6.3.1. <i>Variations du pile ou face divin</i>	54
6.3.2. <i>Modélisation dans le n-univers coloré à objets multiples ($\Omega\alpha^*C$)</i>	55
6.4. ANALYSE	56
6.5. ANALYSE DES AUTRES PROBLEMES POSES PAR SIA.....	60
CHAPITRE 7. LE PROBLEME DE LA BELLE AU BOIS DORMANT	64
7.1. LE PROBLEME	64
7.2. MODELISATION DANS LE N-UNIVERS TEMPOREL A OBJET UNIQUE (ΩT^*)	65
7.3. MODELISATION DANS D'AUTRES N-UNIVERS.....	66
7.4. ANALYSE	68
CHAPITRE 8. L'EXPERIENCE DES SEGMENTS DE TEMPS	71
8.1. DESCRIPTION	71
8.2. MODELISATION DANS LE N-UNIVERS TEMPOREL ET COLORE A OBJET UNIQUE (ΩT^*C)	72
8.3. MODELISATION DANS D'AUTRES N-UNIVERS.....	74
8.4. L'EXPERIENCE DES SEGMENTS DE TEMPS APPLIQUEE A L'OBJET HUMANITE	76
8.5. L'EXPERIENCE DES SEGMENTS DE TEMPS EN TANT QU'OBJECTION A DA.....	77
8.6. ANALYSE	78

CHAPITRE 9. ANALYSE DE L'ARGUMENT DE L'APOCALYPSE	83
9.1. LA VERSION BI-DIMENSIONNELLE DE DA	83
9.2. LE PROBLEME DE LA CLASSE DE REFERENCE AU NIVEAU DE DA-OBJETS ET DE DA-TEMPS	86
9.3. L'URNE A INCREMENTATION.....	87
9.4. LE MODELE DES DEUX URNES A INCREMENTATION	90
9.5. OBJECTIONS PAR RAPPORT A L'ANALYSE DE SOWERS.....	92
9.6. ANALYSE	95
9.7. VARIATIONS DE DA.....	101
CHAPITRE 10. LE PARADOXE DE LA CHAMBRE D'EXECUTION	105
10.1. LE PROBLEME	105
10.2. MODELISATION DANS LE N-UNIVERS SPATIO-TEMPOREL A OBJETS MULTIPLES ($\Omega\alpha^*T^*L$) ...	106
10.3. MODELISATION DANS D'AUTRES N-UNIVERS ($\Omega\alpha^*T$)	106
10.4. L'EXPERIENCE DE LA CHAMBRE D'EXECUTION EN TANT QU'OBJECTION A DA.....	107
10.5. ANALYSE	109
CHAPITRE 11. LE PARADOXE DE GOODMAN	111
11.1. LE PROBLEME	111
11.2. LA DUALITE UNIFICATION/DIFFERENCIATION	112
11.3. PLUSIEURS PROBLEMES CONCERNANT L'INDUCTION	114
11.4. MODELISATION EN TERMES DE N-UNIVERS	115
11.5. CONDITIONS DE L'INDUCTION.....	116
11.6. SOLUTION	118
11.6.1. "Grue" dans le 2-univers coloré et temporel.....	119
11.6.2. "Grue" dans le 3-univers coloré, temporel et localisé	121
11.7. CONCLUSION	122
CONCLUSION	125
REFERENCES	128
INDEX	131
TABLE DES MATIERES.....	134

