

Le dossier Ummo, ensemble de lettres envoyées par une civilisation extraterrestre présente sur Terre, a suscité depuis des décennies un intérêt persistant dans les milieux ufologiques. Présentant notamment une cosmologie et une nouvelle physique détaillées, il se distingue par son volume (plusieurs milliers de pages), sa cohérence interne et sa prétention à dépasser la science humaine de l'époque. Pourtant, depuis les années 1980, des doutes sérieux ont émergé, culminant avec l'aveu de José Luis Jordán Peña en 1993 d'avoir initié l'affaire comme un canular psychologique élaboré.

L'application du théorème de Bayes à l'évaluation d'un dossier ufologique comme Ummo constitue une démarche méthodologique visant à structurer l'incertitude plutôt qu'à produire une certitude.

Le théorème de Bayes est un résultat mathématique permettant de calculer la probabilité d'une hypothèse à la lumière de nouvelles observations. Il repose sur la mise à jour d'une probabilité initiale (dite *a priori*) en fonction de la vraisemblance des données observées sous différentes hypothèses.

Dans le cadre de l'ufologie, cela revient à comparer plusieurs explications possibles d'un phénomène observé et à ajuster leur plausibilité relative en fonction des éléments du dossier. La validité du calcul bayésien dépend entièrement des probabilités utilisées en entrée, notamment la probabilité *a priori* des hypothèses et les probabilités conditionnelles (vraisemblances).

Lorsqu'on applique la formule de Bayes, même des données relativement inhabituelles peuvent ne pas suffire à rendre une hypothèse initialement très improbable dominante. Cela s'explique par le poids des probabilités initiales dans le raisonnement bayésien.

Dans le domaine ufologique, il n'existe pas de base de données consensuelle permettant d'estimer objectivement les probabilités nécessaires. Les valeurs utilisées sont donc souvent discutables et dépendantes du cadre d'analyse choisi.

L'intérêt principal de l'approche bayésienne dans ce contexte n'est pas de trancher définitivement la nature d'un phénomène, mais de rendre explicites les hypothèses, les présupposés et leur influence sur la conclusion.

Le recours au théorème de Bayes en ufologie permet d'introduire une rigueur formelle dans l'analyse, mais ne supprime pas l'incertitude fondamentale liée à la qualité et à l'interprétation des données disponibles.

Ce long article est divisé en trois chapitres permettant de rentrer dans le détail de la méthode Bayésienne, du choix du prior et du choix des données à analyser.

Bayes 1 Le théorème de Bayes et la notion de parcimonie

Ce premier document explicite la méthode Bayésienne qui permettra ultérieurement une évaluation objective et probabiliste de l'origine du dossier, sans considération des aveux de Peña.

L'objectif sera de s'appuyer sur des analyses textuelles et factuelles de lettres choisies ou d'un ensemble de lettres traitant d'un thème donné, analyses effectuées par un ou plusieurs « experts » ; partant de là, nous appliquerons une approche bayésienne simplifiée pour estimer la probabilité que le dossier Ummo soit d'origine terrestre plutôt qu'extraterrestre.

Partant d'un prior que l'on explicitera dans le chapitre 2 suivant, nous quantifierons la sensibilité ($P(D|S)$) et la spécificité ($P(\neg D|\neg S)$) des évidences (explications de tous ces termes dans le texte), en montrant comment chaque type d'analyse textuelle (narrative/conceptuelle, épistémologique, empirique) va se traduire en probabilités à travers le théorème de Bayes. Cette méthode visera non pas à "démonter" un mythe, mais à illustrer comment des outils probabilistes peuvent éclairer des questions controversées en ufologie.

Introduction au Calcul Bayésien Formules et définitions

On rappelle ici la formule de Bayes et ses notations. On trouvera sur internet de très nombreuses références pour se rafraîchir la mémoire sur ce sujet.

Le lecteur intéressé pourra consulter l'ouvrage en français suivant très pédagogique.

Lê Nguyễn Hoàng et al. (2018). La formule du savoir: une philosophie unifiée du savoir fondée sur le théorème de Bayes. EDP Sciences.

Dans la suite, P est une probabilité, S est une hypothèse à vérifier, D sont des données nouvelles permettant de mieux estimer S.

$$P(S|D) = \frac{P(D|S)}{P(D)} P(S) = \frac{P(D|S)}{P(D|S)P(S)+P(D|\neg S)(1-P(S))} P(S)$$

Pour évaluer la probabilité que le dossier Ummo soit d'origine terrestre (hypothèse notée S) plutôt qu'extraterrestre (hypothèse $\neg S$), j'ai utilisé cette formule bayésienne .

Cette méthode a besoin d'une certaine probabilité initiale que l'on nomme "prior » P(S), mais aussi de données que l'on nomme **diagnostic** ou **indicateurs** ou **évidences** (d'où la lettre D qui vient de la médecine) et dont il faut chiffrer la probabilité à partir de faits vérifiés. Personnellement, je préfère le terme **indicateur**, mais je conserverai la lettre D utilisée par la communauté Bayésienne.

Quelques précisions sur le terme indicateur

Dans le cadre de l'analyse probabiliste d'un dossier, le terme *indicateur* désigne tout élément factuel, thématique ou structurel susceptible d'entrer en relation différentielle avec les hypothèses envisagées (notamment ici l'hypothèse d'une origine humaine ou non humaine). Les indicateurs ne constituent ni des preuves, ni des faits explicatifs en eux-mêmes : ils sont des variables d'observation intégrées dans un cadre interprétatif explicitement défini.

L'ensemble des indicateurs est noté D. Ces indicateurs/diagnostics notés (D) proviendront de différentes analyses comme expliqué en introduction. On peut aussi, plutôt que de détailler l'influence de chaque diagnostic D1, D2, D3, etc..., travailler avec une évidence composite globale. Dans tous les cas, D représentera soit une évidence unitaire, soit une évidence composite globale. Dans le cas d'une évidence composite globale, il faut bien comprendre que D regroupera un ensemble de traits corrélés (conceptuels, discursifs, structurels et empiriques), indicateurs interprétés comme les manifestations d'un même processus de production. Les composantes de D comme évidence globale ne sont pas considérées comme indépendantes et ne sont pas probabilisées séparément. L'analyse bayésienne peut porter alors exclusivement sur D pris comme un tout.

Dans le cadre bayésien standard, P(D) désigne la probabilité marginale des données. Elle ne mesure ni la cohérence d'un modèle ni la plausibilité d'une hypothèse ; elle intervient uniquement comme terme de normalisation dans la formule de Bayes. L'analyse doit se concentrer sur les vraisemblances $P(D|S)$ et $P(D|\neg S)$, qui seules portent une signification interprétative.

Par complément, **on notera -D l'ensemble des contre-indicateurs**, c'est-à-dire des observations qui seraient incompatibles avec le modèle considéré s'il était pris au pied de la lettre. Les contre-indicateurs ne doivent pas être confondus avec de simples absences de données ou des lacunes observationnelles : ils correspondent à des phénomènes dont l'existence est explicitement requise par le modèle, et dont la non-observation constitue une incompatibilité empirique.

Un exemple clair de contre-indicateur dans le dossier Ummo est fourni par la question du nuage occultant censé entourer l'étoile d'origine des Ummites. L'existence d'un tel écran est explicitement invoquée dans les lettres pour expliquer l'impossibilité d'observations astronomiques directes. Or, l'analyse astrophysique montre qu'aucune étoile candidate de type approprié ne présente un dispositif de ce genre, alors même qu'un tel nuage serait détectable par les moyens observationnels modernes. Cette non-détection ne relève pas d'un manque d'information, mais d'une prédiction empirique non vérifiée ; elle constitue donc un contre-indicateur au sens strict.

Il est essentiel de souligner que ni $P[D]$ ni la présence de contre-indicateurs ne permettent, pris isolément, de conclure à la validité ou à l'invalidité définitive d'une hypothèse. Deux cadres interprétatifs concurrents peuvent produire des valeurs comparables de $P[D]$, tout en divergeant profondément sur l'interprétation des contre-indicateurs. L'analyse bayésienne ne vise pas à établir une certitude, mais à structurer rationnellement le raisonnement et à évaluer la parcimonie relative des hypothèses. Nous reviendrons plus loin sur ce qu'est la parcimonie.

Ainsi, dans l'étude du dossier Ummo, les indicateurs et les contre-indicateurs doivent être compris comme des outils de clarification méthodologique. Ils permettent d'évaluer la cohérence interne d'un cadre explicatif et d'orienter l'analyse vers l'hypothèse la plus économique, sans jamais transformer cette orientation en démonstration ou en verdict définitif.

Le terme « indicateur » étant précisé, poursuivons la définition des différents paramètres.

Dans le contexte des tests diagnostiques ou des classifications binaires (par exemple, en médecine, en épistémologie bayésienne des sciences, ou en évaluation d'hypothèses comme on voudra le faire pour le dossier Ummo), les deux paramètres classiques sont :

La spécificité : $P(\neg D | \neg S)$

Où :

$\neg D$ signifie : "pas de données/évidence positive" (négatif au test) : ce sont les contre-indicateurs

$\neg S$ signifie : "pas de maladie" ou "hypothèse alternative vraie ».

Ce paramètre spécificité mesure la capacité du test (ou de l'indicateur D) à identifier correctement les cas **négatifs vrais**.

Le dénominateur de la formule de Bayes s'écrit : $P(D | S) P(S) + P(D | \neg S) (1 - P(S))$.

On y trouve le terme $P(D | \neg S)$ qui vaut $1 - P(\neg D | \neg S)$, c'est à dire le complément à la spécificité. Il représente la probabilité conditionnelle d'observer l'évidence positive D sous l'hypothèse alternative $\neg S$ est vraie. C'est le taux de **faux positifs**.

C'est pourquoi en général, on évalue plutôt le terme $P(D | \neg S)$ qui vaut $1 - P(\neg D | \neg S)$, c'est à dire le complément à la spécificité.

L'analyse qualitative de la spécificité ne constitue pas une estimation probabiliste autonome, mais fournit le cadre conceptuel dans lequel sont définies les fourchettes plausibles des probabilités conditionnelles utilisées dans l'analyse bayésienne de l'évidence composite D.

Ceci sera illustré graphiquement dans la suite.

La sensibilité : $P(D|S)$. Dans le contexte diagnostique et dans notre utilisation bayésienne, c'est la probabilité d'obtenir un résultat positif (évidence D observée) lorsque l'hypothèse S est vraie : par exemple, la maladie est présente, ou dans notre cas compte tenu du choix de S, le dossier Ummo est d'origine terrestre. Une sensibilité élevée signifie que le "test « détecte très bien les cas où S est vrai. C'est le **vrai positif**.

Rappelons la formule de Bayes développée pour poursuivre l'explication des notations :

$$P(S|D) = \frac{P(D|S)P(S)}{P(D|S)P(S) + P(D|\neg S)(1 - P(S))}P(S)$$

* $P(S|D)$: Probabilité postérieure (posterior) de S étant donné D : **c'est ce que l'on cherche. C'est à dire réviser notre probabilité S a priori à la lumière de nouvelles données factuelles D.**

* $P(S)$: Prior = probabilité initiale de S (avant les corrections par les évidences D).

J'utiliserai en général pour mes illustrations un prior neutre de 50 % (0.5) : ce choix n'a ici qu'une valeur pédagogique et ne constitue pas un prior épistémiquement justifié.

Car prendre un prior de 0.5 n'est pas une attitude bayésienne et on verra pourquoi plus loin.

Afin de s'y retrouver dans tout ce vocabulaire et ces notations, j'illustre graphiquement ce que tout cela signifie.

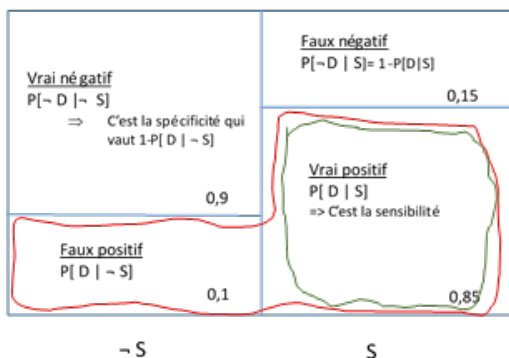
Pour faciliter la lecture du graphique, le monde est divisé ici graphiquement en deux hypothèses égales $P[S]=P[\neg S]=1/2$.

D représente les indicateurs et $\neg D$ les contre-indicateurs.

Les valeurs choisies pour cette illustration sont $\{P(S)=0.5, P(\neg D|\neg S)=0.9, P(D|S)=0.85\}$.

Ces valeurs sont celles que l'on trouvera plus loin dans l'exemple de fonctionnement du notebook Mathematica consacré à la formule de Bayes, et ici simplement calculées à la main.

Deux hypothèses : S origine humaine et $\neg S$ origine ET sont examinées à partir d'indicateurs D et de contre indicateurs $\neg D$. S et $\neg S$ se partagent l'univers selon les probabilités $P[S]$ et $P[\neg S]=1-P[S]$. Pour ce graphique, on choisit $P[S]=P[\neg S]=0,5$. Chaque cas est ensuite scruté selon les indicateurs et les contre indicateurs.



Poids des indicateurs $P[D] = \text{Aire de la zone D}$ (sans oublier les pondérations)
 Soit $P[D] = P[D|S]*P[S]+P[D|\neg S]*P[\neg S]$

La surface pondérée des vrais positifs donne $P[D|S]*P[S]$

Si on la compare à $P[D]$, le rapport des surfaces pondérées donne $P[D|S]*P[S]/P[D]$

\Rightarrow C'est la probabilité conditionnelle $P[S|D]$ car : $P[S|D]=P[D|S] * P[S]/P[D]$

Qui n'est autre que la formule de Bayes.

Exemple avec $P[S]=0,5 : P[S|D]=0,85/(0,85+0,1)=0,89$

Simulation graphique du théorème de Bayes

Pour illustrer de manière plus interactive le théorème de Bayes, j'ai utilisé un notebook Mathematica disponible à l'adresse suivante :

<https://demonstrations.wolfram.com/BayessTheoremAndInverseProbability/>

J'ai choisi les paramètres suivants dans la simulation :

- Un prior $P[S]=0.5$
- La sensibilité $P(D|S)$ varie de 0.75 à 0.9
- La spécificité $P(\neg D|\neg S)$ varie de 0.8 à 0.95
- Le tout pour estimer $P(S|D)$.

Les fourchettes numériques proposées n'ont pas en général pas vocation à être optimales, mais à représenter un intervalle raisonnable de probabilités couvrant des appréciations divergentes d'experts. Chacun pourra les faire varier sur la base de sa propre expertise.

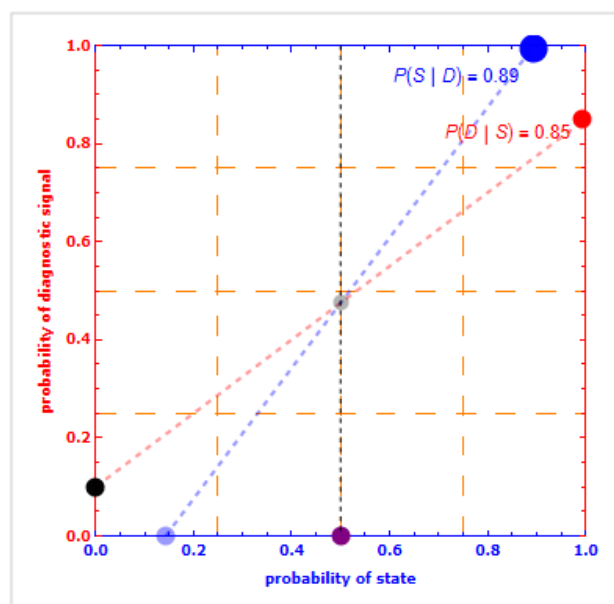
On peut également dans ce notebook faire varier $P(S)$ sur la base d'autres arguments. Nous y reviendrons. On peut ainsi manipuler en ligne les curseurs pour obtenir la valeur $P(S|D)$ affichée sur le graphique. J'ai reporté en Annexe 1 les explications de l'auteur traduites en français concernant le fonctionnement de son notebook.

A titre d'exemple, voici un graphique obtenu en réglant les différents paramètres à des valeurs non extrêmes utilisant les paramètres du graphique ci-dessus.

$\{P(S)=0.5, P(\neg D|\neg S) = 0.9, P(D|S) = 0.85\}$.

=> $P(S|D) = 0.89$ soit une probabilité conditionnelle de 0.89.

Le poids des évidences conduirait ici à une probabilité a posteriori d'une origine humaine du dossier largement dominante par rapport à l'hypothèse d'une origine non terrestre. Mais bien entendu ce n'est qu'une illustration, les évidences D n'ayant pas été énoncées ni démontrées dans ce travail par des travaux d'experts.



Dans la suite, je vais utiliser à titre de démonstrations ce graphique interactif en faisant varier les paramètres $\{P(S), P(-D|-S), P(D|S)\}$, puis je relèverai à chaque fois le résultat $P(D|S)$ obtenu. Le lecteur peut faire de même directement sur internet.

Encore une fois, il ne s'agit ici que d'illustrer le comportement de la formule de Bayes.

Avertissement : Les valeurs probabilistes obtenues ne doivent pas être interprétées comme des fréquences objectives ni comme des degrés de vérité, mais exclusivement comme le résultat probabiliste conditionnel d'un modèle bayésien exploratoire reposant sur des hypothèses explicitement énoncées.

Cas 1 On va commencer par des valeurs de $P(S)$ favorable à l'approche sceptique (pas d'ET sur Terre pour écrire des lettres), valeur que l'on peut illustrer par certains sondages officiels à titre purement informatif. Au niveau mondial, un sondage Reuters du 8 avril 2010 (Annexe 2) indiquait qu'un adulte sur cinq (20 %) parmi les personnes interrogées dans 22 pays (représentant 75 % du PIB mondial) affirme croire que des êtres extraterrestres sont descendus sur Terre et vivent parmi nous dans nos communautés déguisées en « nous ». En France, 10 % des adultes seulement adhèrent à cette croyance, et 21 % en Espagne.

Ces données issues de sondages d'opinion ne sont utilisées ici qu'à titre illustratif pour explorer la sensibilité du modèle à différents choix de prior, et ne constituent en aucun cas une estimation empirique de la probabilité réelle de l'hypothèse considérée.

Les résultats numériques présentés dans la suite n'ont donc pas vocation à conclure, mais uniquement à illustrer la dynamique du théorème de Bayes.

Nous allons considérer le cas où $P(S)$ est de l'ordre de 80 %.

- En prenant par exemple des valeurs telles que $P(-D|-S) = 0.8$ et $P(D|S) = 0.75$, on obtient $P(S|D) = 0.94$.

Cette valeur représente le résultat conditionnel d'un choix de paramètres dans un cadre exploratoire.

Cas 2 Si on part d'une position neutre $P(S) = 0.5$, et en réglant les différents paramètres à leur valeur suivante $\{P(S) = 0.5, P(-D|-S) = 0.88, P(D|S) = 0.83\}$, alors on obtient

$P(S|D) = 0.87$ soit une probabilité conditionnelle de 0.87.

Cas 3 $P(S) = 0.2$, c'est-à-dire que la probabilité qu'il y ait des ET sur Terre est de l'ordre de 0.2. Si j'utilise les valeurs suivantes de la sensibilité et de la spécificité :

$\{P(S) = 0.2, P(-D|-S) = 0.95, P(D|S) = 0.9\}$, alors on obtient $P(S|D) = 0.82$

Si j'utilise les valeurs suivantes de la sensibilité et de la spécificité, on obtient :

$\{P(S) = 0.2, P(-D|-S) = 0.8, P(D|S) = 0.75\}$, soit $P(S|D) = 0.48$

Tout ceci illustre la dynamique de la formule de Bayes et la nécessité de justifier clairement via une expertise soignée le choix des différents paramètres, à commencer par le choix du prior.

Un prior bayésien doit s'appuyer sur des faits

J'ai attiré ci-dessus l'attention du lecteur sur le fait que le choix du prior ne peut se faire n'importe comment. Dans un raisonnement bayésien, le choix du prior constitue une étape essentielle, car il

encode l'état des connaissances antérieures à l'examen détaillé des indicateurs analysés par la suite. **Contrairement à une idée répandue, le choix d'un prior égal à 0,5 ne correspond pas à une position de neutralité : il suppose implicitement une symétrie a priori entre deux hypothèses dont le statut épistémique peut être profondément asymétrique.**

Concernant le dossier Ummo par exemple, l'hypothèse d'une origine humaine du corpus Ummo repose sur des mécanismes bien documentés (productions pseudonymes, canulars complexes, récits pseudo-scientifiques élaborés), tandis que l'hypothèse d'une origine non humaine ne bénéficie d'aucune confirmation indépendante avérée. Un prior égal à 0,5 reviendrait donc à postuler, sans justification empirique, une équivalence initiale entre une hypothèse historiquement fréquente et une hypothèse extraordinairement exceptionnelle.

Le prior à retenir doit provenir d'une donnée antérieure indépendante de l'analyse principale. On comprend bien à ce stade que le choix du prior est fondamental dans le cadre de l'analyse Bayésienne d'une part, et qu'une fois un prior choisi, cette donnée antérieure ne pourra plus être invoquée dans l'évaluation de la formule de Bayes.

Le rasoir d'Ockham comme principe méthodologique et la parcimonie

Dans le cadre bayésien, le principe du rasoir d'Ockham n'est pas une règle philosophique extérieure, mais une conséquence directe du calcul des probabilités. En effet, les hypothèses complexes, qui mobilisent davantage de paramètres ou de mécanismes, répartissent leur probabilité a priori sur un plus grand nombre de configurations possibles. Elles sont donc, toutes choses égales par ailleurs, moins probables que des hypothèses plus simples.

Ainsi, à pouvoir explicatif comparable, une hypothèse plus simple bénéficie naturellement d'un avantage probabiliste dans la mise à jour bayésienne. Le rasoir d'Ockham apparaît alors non comme une préférence subjective pour la simplicité, mais comme une propriété structurelle du raisonnement probabiliste.

Ce principe ne stipule pas que l'hypothèse la plus simple est nécessairement vraie, mais qu'à pouvoir explicatif équivalent, une hypothèse plus complexe est pénalisée tant qu'elle n'apporte pas un gain explicatif décisif. En termes modernes, et notamment dans le cadre de l'analyse bayésienne, la parcimonie n'est pas une préférence esthétique pour la simplicité, mais une conséquence mathématique du fait que les hypothèses complexes répartissent leur probabilité a priori sur un plus grand nombre de paramètres et de configurations possibles (Annexe 3).

Ce principe invite à comparer non pas ce qui est imaginable, mais ce qui est le plus probable compte tenu des données disponibles.

Ainsi, dans le cadre bayésien, la parcimonie résulte conjointement du choix du prior, qui pénalise les scénarios nécessitant des hypothèses nombreuses ou spécifiques, et du terme de vraisemblance $P(D|S)$, qui favorise les scénarios capables de rendre compte de l'ensemble des évidences sans ajustement fin.

Ce dernier point très important est porté en Annexe 4 car il nécessite un développement particulier pour être bien compris.

Conclusion du chapitre 1

Le théorème de Bayes fournit un cadre formel permettant de structurer rationnellement l'évaluation d'hypothèses à partir de données observées. Il ne produit pas de certitudes, mais organise de manière explicite l'articulation entre hypothèses, données et plausibilités initiales.

L'introduction des notions de sensibilité, de spécificité, d'indicateurs et de contre-indicateurs permet de traduire des analyses qualitatives en un cadre probabiliste cohérent, tout en conservant une transparence sur les hypothèses implicites du raisonnement.

Il apparaît également que le choix du prior joue un rôle central, non pas comme un paramètre arbitraire, mais comme une synthèse des connaissances antérieures. Ce choix doit être explicité et discuté, sans jamais être confondu avec les données utilisées dans la mise à jour bayésienne.

Enfin, le lien entre le théorème de Bayes et le principe du rasoir d'Ockham montre que la préférence pour des hypothèses parcimonieuses n'est pas une simple règle heuristique, mais une conséquence du cadre probabiliste lui-même : à pouvoir explicatif équivalent, les hypothèses les plus simples sont naturellement favorisées.

Annexes du chapitre 1

Annexe 1 Fonctionnement du notebook Mathematica

Cette démonstration permet d'explorer la relation quantitative entre deux évaluations de probabilité conditionnelle, $P(S | D)$ et $P(D | S)$, inverses l'une de l'autre, où P représente la probabilité, D est une proposition concernant un « signal diagnostique » et S une proposition concernant une variable d'état d'intérêt. Généralement, la relation entre ces probabilités inverses est comprise grâce au théorème de Bayes.

Vous pouvez faire varier :

(1) le point rouge, $P(D | S)$, la **sensibilité** du test diagnostique ;

(2) le point bleu, $P(D | [\text{Non}]S) = 1 - P([\text{Non}]D | [\text{Non}]S)$, où $P([\text{Non}]D | [\text{Non}]S)$ est la **spécificité** du test diagnostique ;

(3) le point violet, $P(S)$, le taux de base de l'état.

Dans cette notation, $[\text{Non}]S$ et $[\text{Non}]D$ sont les symboles de la négation des propositions S et D , que l'on lit respectivement « non S » et « non D ».

La démonstration illustre également le théorème de Bayes. Ce théorème compare la probabilité inconditionnelle de l'état S , $P(S)$, également appelée probabilité a priori, à la probabilité conditionnelle $P(S | D)$, communément appelée probabilité a posteriori. L'idée de base est simple : si un signal D est plus probable étant donné l'état S qu'en moyenne ($P(D | S) > P(D)$), alors la probabilité a posteriori que l'état S ait observé le signal D est supérieure à sa probabilité a priori ($P(S | D) > P(S)$), et inversement si le signal est moins probable étant donné l'état. Cela ressort logiquement de l'expression du théorème de Bayes, puisque la moyenne pondérée au dénominateur du membre de droite correspond à $P(D)$. Le graphique le met clairement en évidence en observant la différence horizontale entre le point bleu et la ligne verticale passant par le point violet lorsque vous modifiez les paramètres sous-jacents spécifiés par les curseurs.

Il est difficile de suivre simultanément toutes les probabilités conditionnelles et inconditionnelles utilisées dans le raisonnement par probabilité inverse, surtout lorsqu'une ou plusieurs de ces probabilités varient. Les lignes pointillées rouges et bleues de l'animation visent à alléger cette charge cognitive, en utilisant la notion de moyenne pondérée comme « contrainte » sur le calcul des probabilités d'états et de signaux.

Par exemple, lorsque le taux de base ou la probabilité a priori $P(S)$ varie, tandis que $P(D | S)$ et $P(D | \neg S)$ restent constants, $P(S)$ doit demeurer une moyenne pondérée des deux probabilités

conditionnelles inverses $P(S|D)$ et $P(S|\neg D)$, même si les poids $\{P(D), 1-P(D)\}$ et les probabilités conditionnelles inverses utilisées pour déterminer cette moyenne changent tous avec $P(S)$! Cette contrainte particulière de « moyenne pondérée » sur $P(S)$ est illustrée par la ligne pointillée bleue. Symétriquement, la ligne pointillée rouge représente une autre contrainte : toutes les moyennes pondérées possibles $P(D)$ des deux extrémités $P(D|S)$ et $P(D|\neg S)$, avec des pondérations variables $\{P(S), 1-P(S)\}$. Le point essentiel est que ces deux contraintes ne sont pas indépendantes. Premièrement, leur intersection doit se produire précisément aux probabilités inconditionnelles $P(S)$, $P(D)$.

Deuxièmement, étant donné la ligne pointillée rouge, c'est-à-dire la spécification de la sensibilité et de la spécificité du test via les paramètres de l'animation, et le taux de base, également spécifié paramétriquement dans l'animation, la ligne pointillée bleue est entièrement déterminée. Les extrémités de cette ligne, les deux probabilités conditionnelles inverses $P(D|S)$ et $P(D|\neg S)$, sont particulièrement intéressantes, car c'est leur différence qui indique réellement à quel point le signal diagnostique est « informatif » pour l'état. Il est pertinent d'utiliser la démonstration pour observer à quel point un bon test diagnostique (c'est-à-dire sensible et spécifique) peut être peu informatif pour des valeurs très élevées ou très faibles du taux de base $P(S)$.

Annexe 2 Sondage Ipsos pour Reuters

One in Five (20%) Global Citizens Believe That Alien Beings Have Come Down To Earth and Walk Amongst Us In Our Communities Disguised As Humans

India (45%) And China (42%) Most Likely To Believe Aliens Are In Our Midst

New York, NY— A new Reuters News poll conducted by Ipsos and released today indicates that one in five (20%) of presumably human adults surveyed in 22 countries (representing 75% of the world's GDP) say they believe that alien beings have come down to earth and walk amongst us in our communities disguised as "us".

In the alternate, the survey of over 23,000 adults—1000+ respondents per country -- shows that 80% indicate they don't believe that disguised aliens walk within our midst.

Those countries with citizens most likely to believe that the disguised visitors are walk and close by are from India (45%) and China (42%). Those countries where inhabitants are least likely to believe that aliens are walking about disguised as humans are from Belgium (8%), Sweden (8%) and the Netherlands (8%). The following list of findings **begins with** the countries where citizens are **most likely to agree** that "alien beings have come to earth and walk amongst us in our communities disguised as us" and ascends to those countries where citizens are **least likely to agree** with the proposition:

India 45% agree/55% disagree

China 42% agree/50% disagree

Japan 29% agree/71% disagree

South Korea 27% agree/73% disagree

Italy 25% agree/75% disagree

United States 24% agree/76% disagree

Brazil 24% agree/76% disagree

Australia 23% agree/77% disagree
Russia 21% agree/79% disagree
Spain 21% agree/79% disagree
Poland 19% agree/81% disagree
Czech Republic 18% agree/82% disagree
Great Britain 16% agree/84% disagree
Canada 16% agree/84% disagree
Hungary 14% agree/86% disagree
Argentina 13% agree/87% disagree
Germany 11% agree/89% disagree
Mexico 10% agree/90% disagree
France 9% agree/91% disagree
Sweden 8% agree/92% disagree
Belgium 8% agree/92% disagree
Netherlands 8% agree/92% disagree

Those who believe that "alien beings have come to earth and walk amongst us in our communities disguised as us" (20%) are more likely to be men (22%) compared to women (17%), under the age of 35 (25%) compared to those aged 35-54 (16%) and those over the age of 55 (11%) and those more likely to be higher educated (22%) compared to those who are lower or middle educated (19%). There is virtually no difference in terms of income level (lower 22%, higher 20% or middle 19%) and marital status (married 19% versus other 21%)

Those who do not believe that "alien beings have come to earth and walk amongst us in our communities disguised as us" (80%) are more likely to be women (83%) compared to men (78%), those over the age of 55 (89%) compared with those aged 35-54 (84%) and those under the age of 35 (75%), and those who are least educated (low and middle at 81% both) compared with those who are highest educated at 78%. There is virtually no difference in terms of income level (middle 81%, high 80% and low 78%) and marital status (married 81% versus other 79%).

These are the findings of an Ipsos poll conducted between November 4th, 2009 and January 13th, 2010, on behalf of Thompson Reuters News Service. For this survey an international sample of 24,077 adults aged 18-64 were interviewed in a total of 22 countries representing 75% of the world's GDP. The countries included Argentina, Australia, Belgium, Brazil, Canada, China, France, Germany, Hungary, India, Italy, Japan, Mexico, Poland, Russia, and South Korea, Spain, Sweden, the Czech Republic, the Netherlands, Great Britain and the United States. Approximately 1000+ individuals participated on a country by country basis via the Ipsos online panel. Weighting was then employed to balance demographics and ensure that the sample's composition reflects that of the adult population according to the most recent country Census data and to provide results intended to approximate the sample universe. A survey with an unweighted probability sample of this size and a 100% response rate would have an estimated margin of error of +/-3.1 percentage points 19 times out of 20 per country of what the results would have been had the entire population of adults in that

country had been polled. All sample surveys and polls may be subject to other sources of error, including, but not limited to coverage error, and measurement error.

Annexe 3 Le rasoir d'Ockham déduit du bayésianisme

Extrait du livre de Lê Nguyễn Hoàng et al (2018). *La formule du savoir: une philosophie unifiée du savoir fondée sur le théorème de Bayes*. EDP Sciences.

« Venons-en enfin à l'un de mes plus grands instants de jouissance dans mes méditations de la formule de Bayes. J'entrai dans le bureau de collègues à l'École Polytechnique Fédérale de Lausanne (EPFL) à l'heure du déjeuner. Deux de mes collègues discutaient du concept d'Ockham learning, qui est intimement lié au rasoir d'Ockham. Je me posai alors la question de l'interprétation bayésienne du rasoir d'Ockham. Se pourrait-il que la formule de Bayes implique le rasoir d'Ockham ?

Considérons un langage dans lequel nos théories seront décrites. Ce langage peut être le français, la logique formelle ou un langage de programmation informatique. Chaque théorie est alors décrite par une phrase (potentiellement très longue) dans ce langage, c'est-à-dire une suite finie de symboles du langage. Appelons T_n l'ensemble des théories décrites par une phrase à n symboles. Pour être conforme au bayésianisme, l'a priori sur ces théories doit être tel que la somme des crédences $P[T_n]$ en les théories à n symboles est égale à 1. Autrement dit, le bayésianisme impose la condition suivante :

$$P[T_1] + P[T_2] + P[T_3] + P[T_4] + \dots = 1.$$

Or chaque quantité $P[T_n]$ est positive, et il y a une infinité de telles quantités. La théorie des sommes infinies nous dit alors que, si la somme infinie de ces termes positifs est finie, c'est que, nécessairement, les termes $P[T_n]$ de la somme deviennent arbitrairement petits pour de grandes valeurs de n . Alors que cette pensée traversa tout à coup mon esprit, je me jetai au tableau et écrivit :

$$\sum_{n=1}^{\infty} P[T_n] < \infty \implies \lim_{n \rightarrow \infty} P[T_n] = 0.$$

Or, écrire cela, c'est exactement dire que les théories qui nécessitent plus de symboles pour être décrites sont les moins crédibles a priori. Incroyable ! La formule de Bayes implique le rasoir d'Ockham !

La formule de Bayes va même plus loin et nous précise à quel point les théories plus longues à décrire sont moins crédibles. En effet, le nombre de théories à n symboles est exponentiel en n . On en déduit que la crédence a priori d'une théorie à n symboles décroît exponentiellement en n ! Autrement dit, les théories plus sophistiquées ne sont donc pas juste moins crédibles ; elles sont exponentiellement moins crédibles !

Je fus vivement saisi par cette découverte délicieuse — d'autant que je n'avais pas encore rencontré le démon de Solomonoff à ce moment-là. Non seulement cette découverte confortait-elle encore plus la formule de Bayes, mais elle permettait aussi de lever le voile sur la mystérieuse acceptation commune du rasoir d'Ockham. Pour la pure bayésienne, le rasoir d'Ockham n'est pas un principe philosophique qu'il faut s'efforcer d'accepter ; le rasoir d'Ockham est un théorème mathématique du paradigme bayésien. »

Annexe 4 Parcimonie et théorème de Bayes Le principe de parcimonie, ou rasoir d'Ockham, consiste à privilégier les hypothèses qui rendent compte des observations avec le moins de mécanismes supplémentaires ou d'hypothèses auxiliaires.

Dans le cadre de cette étude, les données peuvent être constitués d'observations élémentaires indépendantes, mais aussi d'évidences composites globales notée D, elles-mêmes issue d'un ensemble d'indicateurs :

$D = \{D_1, D_2, \dots, D_n\}$ où chaque D_i représente un aspect observé du corpus (structure discursive, cohérence scientifique, erreurs factuelles, etc.).

L'objectif de cette annexe est de montrer que, dans un cadre bayésien, la parcimonie n'est pas une règle externe, mais une conséquence directe de la manière dont sont évaluées les probabilités.

Rappel de la formule de Bayes dans ce cadre

La probabilité d'un scénario S (par exemple origine humaine ou non humaine) conditionnellement à l'évidence composite D s'écrit :

$$P(S|D) = \frac{P(D|S)P(S)}{P(D|S)P(S) + P(D|\neg S)(1 - P(S))} P(S)$$

- $P(S)$ est le prior,
- $P(D|S)$ mesure la compatibilité globale des indicateurs avec le scénario,
- $P(S|D)$ est la probabilité a posteriori.

Parcimonie et prior $P(S)$

Dans ce cadre, un scénario peut être dit plus ou moins parcimonieux selon qu'il fait intervenir :

- Un nombre limité de mécanismes explicatifs (scénario simple),
- Ou au contraire des conditions spécifiques, des ajustements ad hoc ou des structures explicatives multiples (scénario complexe).

Un scénario complexe doit satisfaire simultanément davantage de contraintes implicites. Il correspond donc à un ensemble plus restreint de configurations possibles, ce qui conduit à lui attribuer un prior plus faible :

$$P(S \text{ complexe}) < P(S \text{ simple})$$

Dans le langage des indicateurs, cela signifie qu'un scénario nécessitant de nombreux ajustements pour expliquer les observations est a priori moins plausible.

Parcimonie et évidence composite D

L'évidence D est constituée d'un ensemble d'indicateurs et de contre-indicateurs :

- Certains éléments de D renforcent le scénario (indicateurs),
- D'autres le fragilisent (contre-indicateurs).

La quantité $P(D|S)$ mesure la capacité du scénario à rendre compte de l'ensemble de ces éléments simultanément.

Même si les composantes de D ne sont pas strictement indépendantes, cette probabilité peut être interprétée comme une mesure de la cohérence globale du corpus sous le scénario considéré.

Rôle de la sensibilité et de la spécificité

Dans le cadre adopté, chaque indicateur peut être caractérisé qualitativement par :

- Sa sensibilité : probabilité d'observer cet indicateur si le scénario est vrai,
- Sa spécificité : probabilité de ne pas observer cet indicateur si le scénario est faux.

Un scénario parcimonieux présente typiquement :

- Des indicateurs de sensibilité élevée (les éléments observés sont attendus),
- Des contre-indicateurs de faible probabilité (peu d'incohérences à expliquer).

À l'inverse, un scénario complexe peut nécessiter :

- D'expliquer certains indicateurs par des mécanismes spécifiques,
- De justifier des contre-indicateurs par des exceptions ou des ajustements.

Cela revient à diminuer la probabilité globale $P(D|S)$.

Interprétation en termes de cohérence globale

Dans ce cadre, la parcimonie peut être reformulée ainsi :

- Un scénario parcimonieux rend les indicateurs globalement attendus sans ajustement particulier.
- Un scénario complexe peut expliquer chaque élément séparément, mais au prix de justifications multiples et spécifiques.

La probabilité conjointe de l'ensemble des indicateurs devient alors plus faible, car elle suppose la réalisation simultanée de plusieurs conditions peu probables.

Interprétation en termes de « réglage fin »

On peut également interpréter cette situation en termes de réglage fin :

- Un scénario simple n'a pas besoin d'être ajusté pour produire les observations : celles-ci sont compatibles avec un large ensemble de situations.
- Un scénario complexe nécessite au contraire un ajustement précis pour reproduire les données observées.

Dans ce cas, la probabilité d'obtenir spontanément une telle configuration est plus faible.

Formulation générale

Si l'on introduit implicitement des paramètres θ associés à un scénario S , la vraisemblance s'écrit :

$$P(D|S) = \int P(D|\theta, S) P(\theta|S) d\theta$$

Un scénario comportant de nombreux paramètres (ou mécanismes implicites) répartit sa probabilité sur un grand nombre de configurations possibles.

Ainsi :

- Certaines configurations peuvent bien expliquer les indicateurs,

- Mais elles représentent une fraction réduite de l'ensemble des cas possibles.

La probabilité moyenne $P(D|S)$ en est diminuée.

Synthèse dans le cadre des indicateurs

Dans le cadre adopté, la parcimonie se traduit par deux effets combinés :

1. Effet sur le prior
Un scénario nécessitant de nombreux ajustements est pénalisé a priori.
2. Effet sur la compatibilité globale des indicateurs
Un scénario qui doit expliquer séparément indicateurs et contre-indicateurs attribue une probabilité globale plus faible à l'évidence composite.

Ainsi, un scénario parcimonieux bénéficie d'un avantage probabiliste global, car il rend les observations attendues sans multiplication de mécanismes explicatifs.

Conclusion de l'annexe 4

Dans le cadre bayésien appliqué à une évidence composite constituée d'indicateurs et de contre-indicateurs, le principe de parcimonie n'est pas une règle externe, mais une conséquence directe du calcul probabiliste.

Un scénario est favorisé non parce qu'il est « simple » au sens intuitif, mais parce qu'il attribue une probabilité plus élevée à l'ensemble des observations sans nécessiter d'ajustements multiples.

La parcimonie apparaît ainsi comme une propriété émergente du raisonnement bayésien, traduisant la cohérence globale entre un scénario et les indicateurs observés.

Bayes 2 Choix d'un prior dans un cadre bayésien appliqué à un scénario ufologique

Le chapitre 1 explique ce qu'est la méthode basée sur la formule de Bayes, ainsi que la notion de parcimonie, pour comparer la probabilité d'un scénario par rapport à un autre.

Il s'agit maintenant d'appliquer cette méthodologie au dossier Ummo, mais sans utiliser encore ce que l'on appelle les évidences ou indicateurs D disséminés tout au long de ce dossier.

Car une première étape importante, dans le cadre de l'approche bayésienne que nous venons de décrire consiste à choisir le prior $P(S)$: ***ce choix doit être effectué indépendamment des évidences D relevées dans le dossier, afin d'éviter tout double comptage de l'information.***

Le présent chapitre 2 vise à expliciter ce choix dans le cas de différents scénarii, en particulier en montrant que certains d'entre eux reposent implicitement sur une conjonction de conditions distinctes, ce qui a des conséquences directes sur leur probabilité a priori $P(S)$.

Décomposition d'un scénario ufologique

Dans le cadre des différents dossiers OVNI, supposer que les observations sont dues à des vaisseaux extraterrestres induit déjà un ensemble d'hypothèses que l'on peut décomposer comme suit :

- S1: existence d'intelligences extraterrestres
- S2 : capacité de déplacement interstellaire de ces intelligences

- S3: possibilité pour elles de contourner les contraintes relativistes

Dans le cadre général du dossier Ummo, on peut ajouter les hypothèses particulières suivantes :

- S4: présence effective discrète au sol dans l'environnement terrestre
- S5: capacité à produire un corpus structuré pour nous instruire

Le scénario Ummo d'une origine non humaine peut donc être reformulé comme la réalisation simultanée de plusieurs sous-hypothèses :

$$S_{\text{non humain}} = S1 \cap S2 \cap S3 \cap S4 \cap S5$$

⇒ Cela implique la satisfaction simultanée de plusieurs conditions de nature hétérogène.

Du point de vue épistémique, il s'agira d'attribuer une certaine valeur de probabilité à chacune de ces sous hypothèses. De manière très qualitative, voici ce que l'on peut en dire :

- S1: **ignorance sur comment la vie naît et se développe dans le temps sur une exoplanète habitable**
- S2: **ignorance sur le temps nécessaire à une ethnie intelligente pour atteindre la capacité technologique à sortir de son berceau stellaire sans auparavant s'être autodétruite**
- S3: **ignorance sur la capacité d'une ethnie à atteindre un niveau technologique permettant de contourner la barrière lumineuse**
- S4: **non observé de manière robuste même s'il existe une certaine croyance**
- S5: **non réellement discriminant** (également compatible avec le scénario noté Shumain)

Qualitativement, aucune de ces conditions ne peut être considérée comme empiriquement établie.

Conséquence sur la structure probabiliste

Dans le cadre bayésien : $P(S_{\text{non humain}}) = P(S1 \cap S2 \cap \dots \cap S5)$

Si on fait une hypothèse d'indépendance des propositions, on pourrait écrire :

$$P(S_{\text{non humain}}) = P(S1) * P(S2) * \dots * P(S5)$$

Cette dernière factorisation n'est introduite qu'à des fins illustratives et n'est pas nécessaire à l'argument principal, qui repose uniquement sur la nature conjonctive du scénario.

Dans tous les cas, la probabilité d'une conjonction d'événements est inférieure ou égale à celle de chacun des événements pris isolément, puisque pour chaque i , $P(Si) \leq 1$.

Ainsi, plus un scénario requiert un nombre élevé de conditions simultanées, plus il va être contraint. Cette propriété ne dépend pas d'une hypothèse d'indépendance stricte, mais de la structure même du scénario.

Illustration (ordre de grandeur)

a Si l'on considère, à titre purement illustratif, que chacune des sous-hypothèses S_i possède une probabilité élevée : $P(S_i) = 0,9$, alors, sous hypothèse d'indépendance pour faciliter l'illustration :

$$P(S_{\text{non humain}}) \approx 0,9^5 \approx 0,59.$$

Dans cette illustration :

- Chaque condition prise isolément apparaît très plausible ($P(S_i)=0.9$).
- **Mais leur combinaison** conduit à limiter l'asymétrie avec l'autre hypothèse.

b Si on décide que $P(S_i)=1$ sauf pour S_3 (dépassement de la vitesse de la lumière), et que l'on choisisse $P(S_3)=0.1$ de manière prudente, cette hypothèse devient fortement influente dans ce cadre multiplicatif illustratif.

J'insiste ici sur le fait que ces illustrations sont basées sur une factorisation non justifiée ; la contrainte principale provient de la nature conjonctive du scénario, indépendamment d'une hypothèse d'indépendance stricte.

Lien avec la parcimonie

Dans le cadre général exposé dans le chapitre 1, la parcimonie se traduit par une pénalisation des scénarios nécessitant :

- un grand nombre de conditions spécifiques,
- des mécanismes explicatifs multiples,
- une coordination implicite entre plusieurs hypothèses.

Un scénario S_{non} humain apparaît, de ce point de vue, comme structurellement plus contraint que le scénario S_{humain} , qui repose sur des mécanismes connus que nous évoquerons plus loin.

Cette différence de structure peut justifier l'introduction d'une asymétrie a priori :

$$P(S_{humain}) > P(S_{non\ human})$$

Cette asymétrie ne dépend pas des évidences, mais de la complexité relative des scénarios que l'on souhaite comparer entre eux.

Un scénario humain

Le scénario d'une origine humaine peut également être explicité sous forme d'une conjonction de conditions, bien que celles-ci correspondent à des mécanismes empiriquement observés [1].

On peut ainsi considérer que $S_{humain}=H_1 \cap H_2 \cap H_3$ avec :

- H_1 désigne la production intentionnelle d'un corpus structuré par un ou plusieurs auteurs humains,
- H_2 la capacité à mobiliser et combiner des connaissances hétérogènes issues de différents domaines,
- H_3 la présence possible d'erreurs, d'incohérences ou d'éléments construits ad hoc.

Ces trois types de mécanismes sont largement documentés dans les productions humaines, qu'il s'agisse d'œuvres littéraires, d'essais spéculatifs ou de corpus pseudo-scientifiques. [1]

On peut citer, à titre illustratif, le cas de George Adamski [2], dont les récits d'origine extraterrestre, produits dans les années 1950, présentent une structuration élaborée, une mobilisation de références scientifiques de l'époque, ainsi que diverses incohérences relevées dans une partie de la

littérature critique. Ce type de corpus illustre la capacité d'un auteur humain à produire des constructions discursives complexes intégrant des éléments scientifiques, spéculatifs et narratifs. Plus généralement, de nombreux exemples de productions pseudo-scientifiques ou spéculatives montrent que la combinaison de connaissances hétérogènes et la présence d'éléments ad hoc ou erronés constituent des caractéristiques fréquentes de ce type de documents. Ces exemples ne constituent pas des preuves au sens strict, mais des illustrations empiriques de mécanismes effectivement observés.

À titre d'ordre de grandeur, et sans prétention d'estimation empirique, on peut associer aux composantes du scénario Shumain des probabilités élevées, par exemple :

$P(H1) \approx 1$, $P(H2) \in [0,7;1]$ et $P(H3) \in [0,8;1]$: la combinaison de ces valeurs (proposées à titre indicatif car elles ne reposent pas vraiment sur une calibration statistique), conduirait dans cet exemple à une probabilité globale $P(\text{Shumain})$ de l'ordre de $[0,6;1]$.

Le scénario humain apparaît ainsi comme moins contraint, dans la mesure où il ne requiert pas de conditions exceptionnelles ou non observées, mais seulement la mise en œuvre de capacités attestées. Ces valeurs doivent toutefois être comprises comme des indications qualitatives destinées à refléter la fréquence et la robustesse de tels mécanismes, et non comme des mesures quantitatives au sens strict.

Le choix du prior et sa contestabilité

Comme déjà évoqué, le choix du prior ne peut être dérivé des données que l'on veut analyser sans introduire un double comptage. Il repose donc sur des considérations externes comme nous venons de le faire. L'importance de la parcimonie des hypothèses appliquée aux scénarios dits non humains » et « humains » précédents conduit à retenir : $P(\text{Shumain}) > P(\text{Snon humain})$.

De plus, l'expérience générale des productions humaines comparables comme nous venons également de l'expliquer suggère qualitativement une probabilité globale $P(\text{Shumain})$ élevée.

Dans ce cadre, il est proposé d'adopter une fourchette qui ne constitue pas une estimation empirique démontrée, mais une hypothèse de travail destinée à refléter une asymétrie structurelle entre scénarios, tout en laissant une marge (en n'allant pas jusqu'à $P(S) = 1$) :

Soit $P(\text{Shumain}) \in [0,6;0,9]$

Le choix de cette fourchette peut légitimement faire l'objet de désaccords, en particulier selon les présupposés épistémologiques du lecteur. Rappelons que dans un cadre bayésien, le prior n'est pas unique. Un croyant pourrait donc privilégier $P(S) = 0,6$. Un sceptique plutôt $P(S) = 0,9$.

L'objectif n'est pas d'imposer une valeur ou une plage de valeurs, mais de rendre explicite cette hypothèse et d'examiner la stabilité des conclusions que l'on en tirera lors de l'examen des données sur une plage de valeurs raisonnables.

Sensibilité au prior et rôle des données

La formule de Bayes peut s'écrire sous forme de rapports de vraisemblance :

$(P(S|D) / P(\neg S|D)) = (P(S) / P(\neg S)) \times (P(D|S) / P(D|\neg S))$. Le posterior dépend du prior **et** du poids des données.

Le terme $P(D|S) / P(D|\neg S)$ constitue le facteur de Bayes, qui mesure la force discriminante des données, et permet de transformer un prior en posterior (Voir Annexe de ce chapitre 2).

Si ce facteur est proche de 1, les données sont peu informatives et le posterior reste proche du prior ; s'il est très élevé ou très faible, les données dominent et le posterior devient peu sensible au prior. Autrement dit, un prior raisonnable fixe un point de départ, mais ce sont in fine les données qui détermineront l'ampleur de la mise à jour.

Ainsi, la sensibilité du résultat au prior dépend directement de la force discriminante des données.

Si les indicateurs et contre-indicateurs conduisent à un facteur de Bayes élevé, la probabilité a posteriori devient largement indépendante du prior initial. À l'inverse, en présence de données peu discriminantes, le prior conserve une influence significative.

L'analyse doit donc être comprise comme une évaluation conjointe du prior et de la capacité des données à le modifier.

En conséquence, la robustesse d'une conclusion bayésienne ne repose pas sur le choix d'un prior unique, mais sur sa stabilité vis-à-vis d'une plage de valeurs plausibles.

La prise en compte explicite du facteur de Bayes $P(D|S) / P(D|\bar{S})$, qui mesure la force discriminante des données, montre que les conclusions dépendront non seulement du prior, mais surtout de la capacité des données à le faire évoluer dans un sens ou dans l'autre, ce qui constitue le cœur du raisonnement bayésien.

Conclusion du chapitre 2

Ce chapitre avait pour objectif de définir un cadre bayésien explicite permettant d'introduire un prior indépendant des données que l'on va examiner ultérieurement.

L'analyse montre que le choix d'un prior ne peut être réduit à une valeur arbitraire, mais doit prendre en compte la structure des scénarios considérés. En particulier, la décomposition du scénario non humain en un ensemble de conditions distinctes met en évidence son caractère conjonctif, tandis que le scénario humain repose sur des mécanismes empiriquement observés. Cette différence de structure peut justifier l'introduction d'une asymétrie a priori.

Dans ce cadre, il est proposé d'adopter une fourchette large : $P(\text{Shumain}) \in [0,6;0,9]$.

Cette fourchette ne constitue pas une estimation empirique, mais une hypothèse méthodologique visant à refléter une asymétrie de complexité entre scénarios. Cette fourchette de priors a vocation à être utilisée dans des analyses ultérieures qui seront fondées sur différentes évidences D.

Il convient de souligner que ce choix reste discutable et dépend des présupposés épistémologiques adoptés. L'intérêt de l'approche proposée réside précisément dans le fait de rendre ces hypothèses explicites, et de permettre d'en évaluer l'impact sur les résultats.

L'étape suivante consistera à choisir des indicateurs issus du corpus et à examiner, à l'aide du théorème de Bayes, dans quelle mesure ceux-ci modifient les probabilités a priori ainsi définies. La robustesse des conclusions sera alors appréciée au regard de leur sensibilité à cette plage de priors.

Bibliographie du chapitre 2

[1] « Hoaxes » dans Ufo Encyclopedia de Jérôme Clark (1998)

[2] « George Adamski » dans Ufo Encyclopedia de Jérôme Clark (1998)

[3] Navarro D.J., Foxcroft, D.R. (2020). Apprentissage des statistiques avec Jamovi : un tutoriel pour les étudiants en psychologie et autres débutants. Chapitre 16 sur les statistiques bayésiennes.

Annexe du chapitre 2 Facteur de Bayes et son interprétation [3]

Dans le cadre bayésien adopté, la mise à jour des probabilités peut s'écrire sous forme de rapports :

$$(P(S|D)/P(-S|D)) = (P(S)/P(-S)) \times (P(D|S)/P(D|-S))$$

Le terme $BF = P(D|S)/P(D|-S)$ est appelé **facteur de Bayes**.

Il traduit dans quelle mesure les données D favorisent le scénario S par rapport à son alternative -S

Interprétation générale

- $BF > 1$: les données favorisent S
- $BF < 1$: les données favorisent -S
- $BF = 1$: les données sont non informatives

Le facteur de Bayes agit comme un **multiplicateur du prior exprimé en cote** $P(S)/P(-S)$. Le chapitre 16 de la référence [3] propose une échelle d'interprétation graduée du facteur de Bayes.

Facteur de Bayes BF	Interprétation en faveur de S	Facteur de Bayes BF	Interprétation en faveur de -S
1 à 3	Indicateur faible	1/3 à 1	Indicateur faible
3 à 20	Indicateur positif	1/20 à 1/3	Indicateur positif
20 à 150	Indicateur fort	1/150 à 1/20	Indicateur fort
>150	Indicateur très fort	<1/150	Indicateur très fort

Lien avec le prior

La relation complète s'écrit : **cotes a posteriori = cotes a priori × BF**

- Le prior fixe le point de départ (la cote a priori).
- Le facteur de Bayes mesure la contribution des données (BF)

Ainsi, un facteur de Bayes proche de 1 indique que les données modifient peu le prior. Et un facteur de Bayes élevé ou très faible indique que les données dominent le résultat final.

La robustesse d'un résultat dépend directement de la valeur du facteur de Bayes et de la plage de priors considérée.

Le facteur de Bayes constitue le mécanisme central de la mise à jour bayésienne.

Il permet de quantifier la force des indicateurs D indépendamment du prior, et donc d'évaluer dans quelle mesure les données sont capables de modifier une hypothèse initiale.

Son interprétation graduée fournit un cadre opérationnel pour apprécier la portée réelle des résultats obtenus dans l'analyse principale.

Bayes 3 Le choix des données dans la période 1966-1969

Dans les deux chapitres précédents, nous avons expliqué en détail le fonctionnement du théorème de Bayes, puis le choix d'un prior argumenté sur un tel dossier ufologique, et enfin la question du facteur de Bayes, dont l'estimation repose sur une expertise d'un certain jeu de données : ce facteur de Bayes BF devra être évalué de la manière la plus impartiale possible, afin au final de faire la part entre la probabilité d'un dossier d'origine humaine ou non humaine.

Il faut donc répondre à plusieurs questions :

- Quel jeu de données choisir sans biaiser volontairement ou involontairement l'analyse ?
- Comment choisir les différents sujets afin d'échantillonner l'ensemble du dossier de manière non biaisée ?
- Comment éviter la répétition de thèmes trop dépendants les uns des autres, ce qui peut augmenter ou diminuer artificiellement le BF ? Autrement dit, il faudrait arriver à sélectionner des thèmes indépendants les uns des autres.
- Chacun des choix précédents doit être clair afin de pouvoir ajuster le BF les concernant à la lumière d'autres expertises.

Répondre à ces questions, et sélectionner des thèmes et des sujets précis est l'objet de ce chapitre : une fois l'ensemble des sujets clarifié, il faudra ensuite développer, pour chacun d'entre eux, les arguments permettant d'établir pour chacun un facteur de Bayes.

Ce dernier point n'est pas l'objectif de ce chapitre et fera l'objet d'articles futurs.

Le dossier Ummo et le corpus fondateur des années 1966 à 1969

Cela fait maintenant plusieurs années que je m'intéresse à ce dossier, et j'ai d'ailleurs publié de nombreux articles sur le site Ummo-sciences.org.

En prenant un peu de recul, je me rends compte que j'ai parfois utilisé des informations qui venaient bien après les aveux de Jordan Peña de 1993. Mon souci est qu'après cette date, et la diffusion très élargie de l'affaire Ummo, je considère de manière plus que circonspecte l'authenticité des lettres et autres tweets des années 2000, dans le sens où rien ne garantit qu'elles soient l'œuvre de la même source.

Dans le but de préserver une certaine pureté du dossier à analyser, j'ai donc décidé de me limiter aux lettres classées D du siècle dernier sur le site ummo-sciences.org, suivant en cela la classification dite de Darnaude et de son catalogue Ummocat.

Si on examine maintenant la production des lettres indicées D-xx, on se rend compte que celles-ci sont émises par « salves » :

- Une grosse salve de lettres qui couvrent les années 1966 à 1969, respectivement par année {36,32,15,16} soit 99 lettres, ou encore 430 pages.
- Puis quasiment rien sur les années 1970 à 1986, hormis deux salves de lettres en 1970 (10), 1973 (10). Cette période représente à peine 40 pages.
- Et enfin une nouvelle salve de lettres en 1987 {11} et 1988 {15} qui représente environ 150 pages.

La première salve ressemble fortement à une sorte d'initialisation pour amorcer un dispositif ou une expérience ou un premier contact via une centaine de documents.

Les phases réduites peuvent correspondre à des interventions nécessaires pour perpétuer ou relancer l'expérience ou répondre à un contexte précis pour les contactés (séminaire, réunion, autre...).

La dernière salve, 20 ans après les années madrilènes, est beaucoup plus métaphysique et religieuse et synthétise clairement beaucoup d'informations déjà disponibles.

Puis les lettres se raréfient très fortement fin 1988.

La phase dite d'initialisation est évidemment fondamentale, et ce qu'elle contient doit constituer à la fois un socle de référence et une méthode éprouvée permettant d'entrer dans un processus d'échanges pérennes. Comme on l'a vu, cette phase représente environ 70 % des lettres classées D.

C'est pour cela que j'ai décidé de m'intéresser à cette première salve, qui représente **le corpus fondateur**, produit pendant la période madrilène (1966 à 1969).

Ce corpus représente un ensemble de 430 pages et une centaine de lettres, ce qui est considérable. On fait l'hypothèse que sur cette période, les lettres de la source n'ont pas été « polluées » par d'autres sources, et que la période contient réellement les idées fondatrices que la source voulait transmettre aux contactés espagnols. Il conviendra a posteriori de vérifier tout cela.

Les grands thèmes développés dans la période madrilène

Par où commencer l'étude de ces 430 pages ? Car cet ensemble forme un corpus dense, techniquement hétérogène, mobilisant aussi bien des sciences dites « dures » que des domaines métaphysiques et psychologiques. Face à un tel ensemble, l'analyse peut rapidement devenir coûteuse si elle n'est pas structurée méthodologiquement.

Le tableau suivant présente une répartition des domaines abordés dans les lettres des années madrilènes. Cette répartition est à ce stade très schématique mais cela permet d'avoir une première vue d'ensemble globale.

Cette répartition par thèmes ne permet pas d'inférer directement les intentions des auteurs, mais elle met en évidence la place occupée par certains domaines dans le discours.

Notamment, on observe dans le tableau suivant que les thèmes à caractère métaphysiques, théologique et psychologique représentent $\approx 50\%$ du corpus.

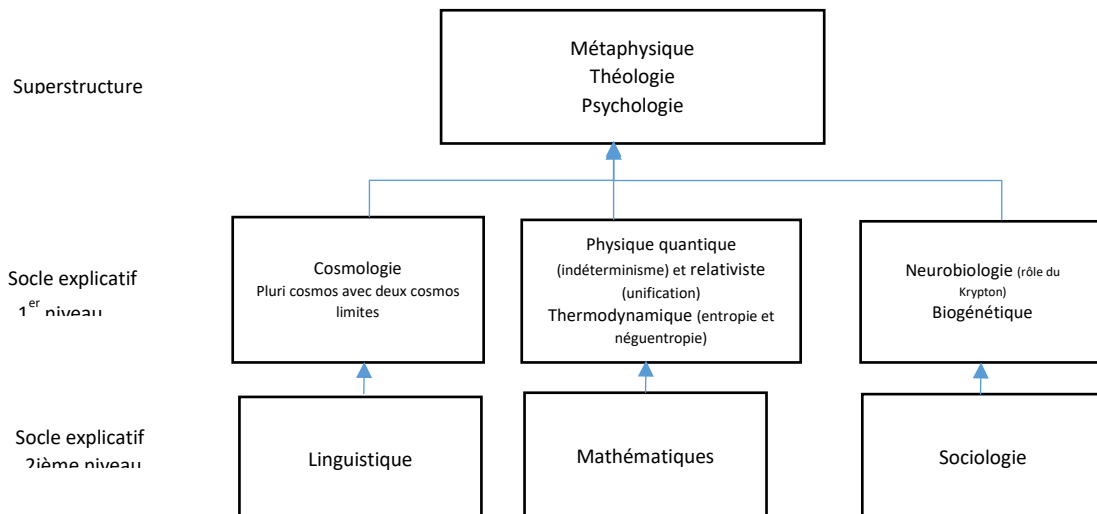
Pour être bien compris et acceptés, ces messages métaphysiques et psychologiques s'appuient sur un ensemble de concepts scientifiques cohérents : c'est là qu'interviennent les sciences dites dures (3 à 7), qui représentent un socle d'environ 45 % du corpus. Le reste des thèmes est plus réduit mais demeure pertinent, notamment l'aspect linguistique.

Domaine	Intitulé	Part estimée
1	Métaphysique / théologie	25 %
2	Psychologie / libre arbitre	25 %

Domaine	Intitulé	Part estimée
3	Physique fondamentale	14 %
4	Cosmologie	10 %
5	Astronomie	8 %
6	Mathématiques	7 %
7	Biologie / neurobiologie	6 %
8	Sociologie	3 %
9	Linguistique	2%

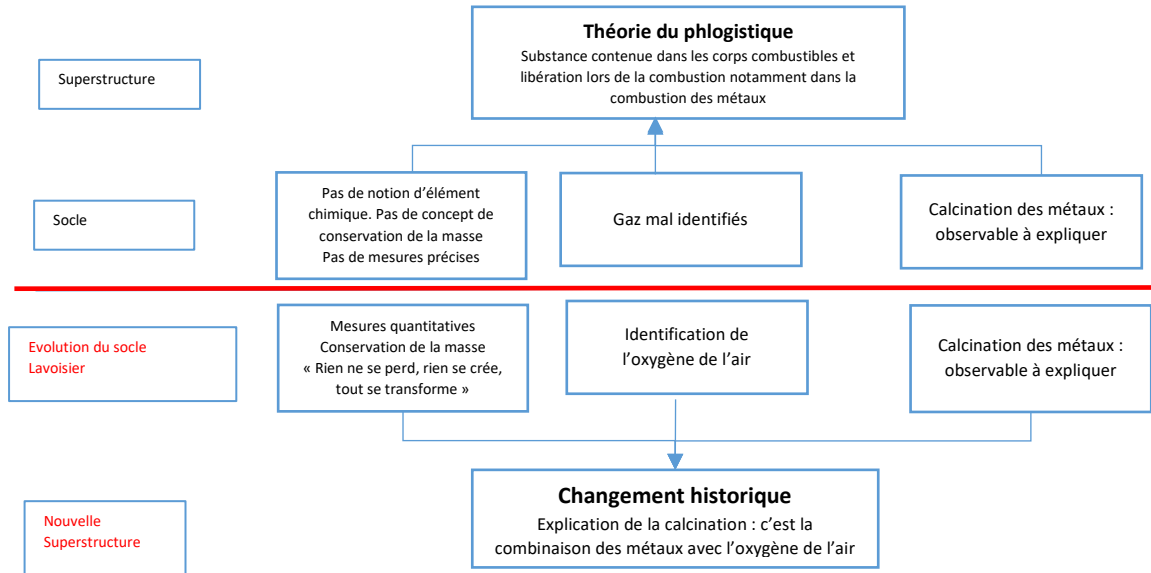
Avant de choisir les thèmes à étudier en profondeur, il convient en préliminaire de rappeler que pour les Ummites, les thèmes des domaines 1 et 2 s'appuient fortement sur un socle de sciences dures ; cependant, ces thèmes 1 et 2 ne renvoient pas en retour de nouveautés vers ces sciences dures : la dépendance est unidirectionnelle.

Le schéma ci-après illustre cette dépendance structurelle et asymétrique entre les domaines mobilisés dans le corpus. Les flèches montantes (des sciences dures vers les domaines 1 et 2) représentent une dépendance épistémique réelle : la métaphysique (1) et la psychologie (2) s'appuient explicitement sur des concepts ou cadres issus de la physique Ummite, de la cosmologie Ummite, de la biologie du cerveau Ummite. Comme expliqué en Annexe de ce chapitre 3, l'Univers deviendrait auto conscient capable de se corriger lui-même en une sorte de boucle cybernétique. Dans une perspective bayésienne stricte, ces boucles cybernétiques n'apportent aucune information indépendante susceptible d'augmenter la vraisemblance des hypothèses fondées sur les sciences dures. **Le schéma doit donc être lu comme la représentation d'un édifice où les niveaux supérieurs dépendent des niveaux inférieurs, sans rétroaction épistémique effective.**



Un exemple illustratif classique

Un exemple historique classique permet d'éclairer ce type de situation : celui de la théorie du phlogistique (XVII^e siècle). Cette théorie proposait un cadre explicatif cohérent et conceptuellement riche pour rendre compte de la combustion et de la calcination des métaux, et s'inscrivait dans une métaphysique de la matière largement acceptée à cette époque, en l'absence de notions opératoires telles que l'élément chimique, la conservation de la masse et l'absence de mesures quantitatives précises. C'est la partie haute du schéma.



L'émergence de la chimie Lavoisienne (partie basse du schéma) ne correspond pas à une simple correction locale de cette théorie, mais à un changement plus profond du socle conceptuel, fondé sur de nouveaux outils expérimentaux, un langage théorique renouvelé et l'identification de l'oxygène. Ce changement de cadre rend la théorie du phlogistique inutile plutôt que directement réfutée, entraînant l'effondrement de la superstructure explicative qui en dépendait.

L'exemple illustre ainsi comment une construction théorique rationnelle et historiquement située peut-être abandonnée à la suite d'une transformation du socle scientifique sous-jacent.

La théorie du phlogistique n'a pas été abandonnée parce que ses constructions conceptuelles auraient été naïves ou inconsistantes, mais parce que ses fondations empiriques n'apportaient aucun gain explicatif indépendant face à une hypothèse concurrente plus parcimonieuse.

Cet épisode, fréquemment enseigné en histoire et en philosophie des sciences, illustre un principe méthodologique général : une superstructure interprétative, même sophistiquée, ne peut compenser la faiblesse de ses bases empiriques, et ne modifie pas, à elle seule, l'évaluation probabiliste globale d'une hypothèse.

C'est pourquoi dans la suite je vais m'intéresser aux thèmes suivants qui constituent selon moi les vrais socles fondamentaux de toute la métaphysique Ummite, à savoir :

- ⇒ La physique fondamentale et ses applications, la cosmologie et l'astronomie, la neurobiologie, et la linguistique (qui permet de traduire ces concepts à l'aide d'un certain langage particulier).

Cet ensemble de thèmes couvre à mon sens correctement les bases du discours Ummite.

Le choix des sujets dans les différents thèmes

J'ai écrit depuis 2021 différents articles que l'on trouvera sur le site d'Ummo-sciences.org.

Pour rester cohérent avec la période 1966-1969, je ne vais conserver que les sujets que j'ai déjà traités et qui ne se basent pas sur des données supplémentaires hors de cette période « socle ».

Une fois le tri effectué, il resterait selon moi les articles suivants et les lettres correspondantes sur lesquels me baser :

ana37 - Analyse D57-1 et le JWST	2021	D 57-1	1967	D 74	1969		
ana38 - Retour sur février 1934	2021	D 57	1967				
art38-Les aurores boréales sur Ummo	2021	D 41-3	1966	D 74	1969		
art39-Le moteur à plasma	2021	D 41-6	1966	D 69-3	1968		
art40 - Ioumma étoile de type K	2021	D 74	1969				
art41 - Non Wolf 424 n'est pas l'étoile des Ummites	2021	D 21	1966	D 32	1966	D74	1969
art43 - La matière négative dans l'univers et sa localisation	2021	D 41 15	1966				
art50 - Comment calculer un angle	2025	D 59	1967				
art51 - Les étoiles proches de type K	2025	D 74	1969				
ana44 - La D63 et les propyls	2026	D 63	1967				
ana 45 Charroux et Ummo	2026	D 41	1966				
ana 46 Wheeler et Ummo	2026	D 41 D 45 D 52	1966	D 57 D59	1967		
ana 47 Rôle du krypton : analyse physique et épistémologique	2026	D21 D33-3	1966	D58-3	1967		
ana 48 Genèse épistémologique probable de la physique des particules	2026	D 59	1967				
Langue et langage Ummite (non publié)		(*)					
Les ordinateurs Ummites (non publié)		D69 D71	1967				

(*) Cet article ne repose pas exclusivement sur les lettres des années 1966-1969.

Tout ceci donne beaucoup de matière pour positionner chacun des sujets précédents dans les grands thèmes mis en évidence dans le paragraphe précédent.

Physique théorique et applications D1

art39-Le moteur à plasma

ana 46 Wheeler et Ummo

ana 48 Genèse épistémologique probable de la physique des particules

Les ordinateurs Ummites (non publié)

Cosmologie D2

ana 45 Charroux et Ummo

art43 - La matière négative dans l'univers et sa localisation

Linguistique D3

Langue et langage Ummite (non publié)

Neurobiologie D4

ana 47 Rôle du krypton : analyse physique et épistémologique

Astronomie et mathématiques D5

ana37 - Analyse D57-1 et le JWST

ana38 - Retour sur février 1934

art38-Les aurores boréales sur Ummo

art40 - Ioumma étoile de type K

art41 - Non Wolf 424 n'est pas l'étoile des Ummites

art50 - Comment calculer un angle

art51 - Les étoiles proches de type K

ana44 - La D63 et les proplyds

Le climat de l'exoplanète Ummo D6

Un lecteur attentif remarquera que mes différentes études sur le climat de l'exoplanète Ummo ne figurent pas dans le volet astronomie ci-dessus et il aura raison (**art 36** Habitabilité de la planète Ummo et **art 42** Zone habitable des exoplanètes et application à Ummo). Ces études ne figurent pas dans le tableau ci-dessus car l'estimation de l'habitabilité repose sur deux informations non disponibles dans la période madrilène considérée (1966 à 1969), à savoir :

- la métallicité de l'étoile Ioumma, fondamentale pour estimer sa luminosité (Tweet OAY de 2021)
- la pression atmosphérique de l'atmosphère d'Ummo, seulement connue en 2015 (OT 8)

De surcroît, l'art 42 nécessite une intervention supplémentaire du sulfureux 312 en 2021.

Mais pour ne pas supprimer ce thème particulier, je consacrerai un article complet sur ce sujet en faisant une analyse bayésienne appliquée au climat de Umho, et nous verrons bien sur ce qu'il en sortira.

Chacun des 6 thèmes précédents sera analysé comme constituant un indicateur global D_i , ceci afin d'éviter la répétition de sujets trop dépendants les uns des autres.

En effet, si je note $\{d_1, \dots, d_n\}$ les sous-indicateurs présents dans chaque thème D_i , l'objectif sera de travailler avec une évidence composite globale regroupant ces sous-indicateurs. Dans tous les cas, D_i représentera soit une évidence unitaire (par exemple en neurobiologie ou en linguistique, quand il n'y a qu'un seul sous-indicateur), soit une évidence composite globale comme en astrophysique.

Cela signifie que dans le cas d'une évidence composite globale, chaque D_i regroupera un ensemble de traits corrélés (conceptuels, discursifs, structurels et empiriques), c'est-à-dire des sous-indicateurs di interprétés comme les manifestations d'un même schéma. Les composantes d_i de chaque D_i comme évidence globale ne seront pas considérées comme indépendantes et ne seront pas probabilisées séparément, ce qu'il faudra bien entendu vérifier. L'analyse bayésienne peut porter alors exclusivement sur D_i pris comme un tout.

Conclusion du chapitre 3 La qualité et non la quantité des sujets choisis

Il ne s'agit pas dans l'application de la méthode bayésienne de traiter exhaustivement tous les sujets énumérés ci-dessus comme je viens de l'expliquer ; le nombre d'indicateurs retenus pour l'estimation du facteur de Bayes ne constitue pas en soi un gage de pertinence.

Dans le cadre de la mise à jour bayésienne, l'information apportée par un ensemble d'évidences

$D = \{D_1, D_2, \dots, D_n\}$ dépend non seulement de leur contenu, mais surtout de leur indépendance conditionnelle relativement aux scénarios considérés.

En effet, la multiplication des facteurs de Bayes associés à chaque indicateur n'est rigoureusement justifiée que si ces derniers apportent des informations distinctes et non redondantes.

À l'inverse, la prise en compte de plusieurs éléments reposant sur un même mécanisme explicatif conduit à un double comptage de l'information et à une surestimation artificielle de l'évidence.

Il en résulte que la qualité des indicateurs (entendue comme leur capacité à discriminer entre les scénarios, ainsi que leur ancrage dans des mécanismes indépendants) prime largement sur leur quantité.

Le découpage en grands domaines disciplinaires retenu dans cette analyse apparaît globalement pertinent, dans la mesure où ces champs reposent sur des cadres théoriques et des mécanismes explicatifs distincts, autorisant une approximation raisonnable de leur indépendance, ce qui devra bien sûr être vérifié.

Une sélection restreinte d'indicateurs robustes et clairement distincts est ainsi préférable à une accumulation d'éléments partiellement corrélés, dont l'effet combiné peut induire un biais significatif dans l'évaluation probabiliste.

C'est pourquoi je pense que la meilleure façon d'éviter tout double comptage consiste à n'étudier que les 6 familles thématiques proposées $\{D_1, D_2, D_3, D_4, D_5, D_6\}$, sachant qu'au sein de chaque famille, on va probablement identifier un sujet englobant tous les autres. Ce choix (toujours discutable) sera de toute manière évalué lors de l'examen détaillé des éléments mobilisés au sein de chaque domaine. Car plusieurs arguments peuvent reposer sur un même mécanisme explicatif et ne

constitueront pas des sources d'information indépendantes. C'est par exemple le cas des différents sujets associées au système astrophysique, qui relèvent d'un modèle unique de localisation de l'étoile Ioumma.

Afin de préserver la validité de l'analyse bayésienne, il est donc préférable de regrouper ces éléments en un nombre restreint d'indicateurs correspondant à des mécanismes distincts, plutôt que de les traiter séparément.

Une telle approche permet d'éviter tout double comptage de l'information et de garantir une évaluation plus robuste des facteurs de Bayes associés.

Epilogue

L'approche retenue repose sur une délimitation explicite du corpus aux lettres de la période 1966–1969, suivie d'une structuration en grands thèmes disciplinaires (cosmologie, astrophysique, biologie, etc.), chacun étant associé à un indicateur D_i . Cette construction vise à garantir une sélection des données fondée sur des mécanismes explicatifs distincts, tout en évitant la multiplication d'éléments redondants, chaque indicateur pouvant regrouper plusieurs observations au sein d'un même thème et constituer ainsi un indicateur composite. Ce cadre permet de limiter les biais de sélection et d'assurer une certaine reproductibilité de l'analyse. Il repose toutefois sur l'hypothèse que ces grands thèmes peuvent être considérés comme relativement indépendants, hypothèse adoptée ici comme une approximation de travail et qui devra être appréciée de manière critique lors de l'interprétation des résultats.

Le travail restant à faire est considérable, car pour chaque indicateur global D_i , il va falloir en estimer avec expertise le facteur de Bayes, c'est-à-dire :

$$BF_i = \frac{P(D_i|S)}{P(D_i|\neg S)}$$

Une fois seulement ce travail accompli, on pourra évaluer le facteur de Bayes global :

$$BF = \frac{P(D_1, \dots, D_6|S)}{P(D_1, \dots, D_6|\neg S)}$$

Si nous avons bien choisi nos indicateurs de manière à ce qu'ils soient indépendants, ce que révéleront les études ultérieures, on pourra même écrire le facteur de Bayes global comme :

$$BF = BF_1 * \dots * BF_6$$

Dans tous les cas, la cote a posteriori s'écrirait selon la formulation générale :

$$Cote\ Posterior = \frac{P(S)}{P(\neg S)} BF$$

Si les indicateurs sont « bien choisis » et indépendants, on pourrait même écrire :

$$Cote\ Posterior = \frac{P(S)}{P(\neg S)} BF_1 * \dots * BF_6$$

Cette dernière formule montre que l'on pourrait introduire de nouveaux indicateurs correspondants à de nouveaux facteur de Bayes (par exemple BF_n), et que si ces nouveaux indicateurs sont

totalément indépendants des précédents, il suffira de multiplier le résultat précédent par BFn pour affiner la cote a posteriori.

Annexe 1 du chapitre 3 Synthèse de la lettre D 731 : la métaphysique Ummite, le rôle de l'homme, son libre arbitre (D33).

Avertissement : la cohérence interne du système conceptuel synthétisé ci-après, y compris dans ses formulations tardives, ne constitue en elle-même aucun indice de validité empirique. Elle témoigne uniquement de la stabilité d'un cadre interprétatif construit à partir d'hypothèses physiques préalables, dont la solidité demeure la question centrale.

La lettre D 731 a pour but de décrire la place de l'homme (OEMMI) dans l'Univers.

Pour ce faire, les Ummites décrivent leur métaphysique, elle-même basée sur une physique du pluri-cosmos qui nous est totalement inconnue. La description préliminaire du pluri cosmos a pour but de mettre en lumière deux Cosmos limites particuliers μ et $\sigma\mu$, en plus de notre propre Cosmos de masse M et de son Cosmos jumeau de masse -M. Il existerait de surcroît des masses imaginaires jouant un rôle essentiel dans la genèse du Pluri Cosmos, permettant également les interactions-perturbations entre les différents Cosmos. Celles-ci conduisent notamment à la formation de galaxies et cela rejoint dans certains cas notre modèle de Big-Bang. **Ces masses imaginaires constituent la seule nouveauté de la troisième salve de lettre (1987-1988), et vont permettre une synthèse des concepts des années madrilènes.**

Ce cadre physique étant posé, il s'agit de décrire les deux Univers limites μ et $\sigma\mu$, et leur lien privilégié avec les âmes μ des OEMMI et les êtres vivants. Le vivant se distingue par sa néguentropie, et le vivant évolue selon une sorte de Darwin accéléré et orienté par des lois bio génétiques bien précises : l'objectif est de guider l'évolution des espèces vers l'apparition tôt ou tard d'un OEMMI adapté à son environnement, et capable de communiquer avec les deux Univers limites. Ce guidage (effectué par les lois biogénétiques) se fait à partir de l'encéphale des OEMMI à travers des réseaux d'atomes de Krypton qui envoient (et reçoivent) instantanément des informations à la fois avec l'Univers μ via leur âme μ , mais également avec l'Univers $\sigma\mu$, véritable cerveau global de l'Univers : une rétroaction de l'Univers $\sigma\mu$ est alors effectuée pour orienter d'une part une civilisation planétaire via des patrons prédéfinis et d'autre part pour orienter le phylum de l'évolution sur la planète considérée : c'est l'orthogénèse.

Les masses imaginaires ne sont donc pas uniquement des éléments de la physique Ummite : elles jouent un rôle central dans la métaphysique Ummite, notamment dans la connexion entre l'encéphale de l'OEMMI et les univers μ et $\sigma\mu$. Grâce à l'effet frontière, ces masses permettent à l'être conscient d'accéder, même brièvement, à des plans de réalité non locaux et non temporels, où les règles habituelles de la causalité et de la continuité sont suspendues. La masse imaginaire devient le vecteur ontologique du lien entre matière, conscience et transcendance.

Ces descriptions relationnelles étant faites, on passe à l'évocation de la place de l'homme dans l'Univers, et à la notion d'un Dieu créateur intemporel. Dieu-WOA a ainsi créé un Univers auto conscient capable de se corriger lui-même en une sorte de boucle cybernétique. Ceci ne peut se réaliser sans l'apparition d'êtres néguentropiques, car galaxies et étoiles seules ne peuvent engendrer que de l'entropie, ce qui est contraire au dessein de WOA : car la fonction de l'OEMMI (être conscient) est d'augmenter la néguentropie (l'ordre) de l'Univers par son « témoignage » en servant de capteur-témoin au WAMM-WAAM. Ce dernier est la création de WOA, et aussi la création

de l'OEMMII d'une certaine façon puisqu'avec notre pensée nous contribuons avec les autres quintillions d'humains dans la pluralité des cosmos, à le remodeler.

Le libre arbitre n'est pas évoqué dans ce texte, mais il est écrit ailleurs (D33) que le libre arbitre provient de ce qu'on appelle l'indéterminisme en mécanique quantique. Sans indéterminisme, les liens relationnels décrits précédemment seraient alors déterministes, ce qui ne se passe pas du tout ainsi. Une orientation de ce libre arbitre est alors nécessaire, et Dieu a édicté des lois morales aux OEMMI afin que leurs actions favorisent la néguentropie au détriment de l'entropie. Le principe d'économie, ou entropie, nous pousse à être égoïstes et à négliger l'amour envers nos semblables. A contrario, aimer son prochain et participer à une meilleure cohésion sociale contribue à créer des ilots de néguentropie, favorables au plan de captation du WAAM. L'amour transcende alors les valeurs éthiques et humanistes et devient un concept intégré dans la science Ummite. Celui qui viole la loi de Dieu mérite la condamnation de tous les OEMMII du WAAM-WAAM car il porte gravement préjudice à tous les OEMMI du Pluri-Cosmos.

Les Ummites, peuple profondément religieux, notent une certaine analogie entre leur métaphysique et le christianisme, en nous alertant toutefois sur le fait que certains détails de cette religion ne sont que pure invention humaine a posteriori, car Jésus, UMMOWOA, a quant à lui réellement existé.

Le tableau suivant illustre que la lettre D731 ne constitue pas une refondation des concepts des années madrilènes, mais à l'aide du nouveau concept de masses imaginaires, une mise en cohérence tardive de thèmes déjà présents dans les textes initiaux.

Domaine	Volumes 1–2 (1966–1967)	Lettre D731 (1987)	Nature du lien
Physique fondamentale	Indéterminisme, structures non ponctuelles ,etc.	Formalisation plus systématique	Continuité
Cosmologie	WAAM, pluralité des univers, univers limites	Modélisation globale intégrée + masses imaginaires	Nouveau concept / puis synthèse
Psychologie	Libre arbitre fondé sur l'indéterminisme	Psyché collective, conscience intégrée	Extension
Métaphysique	Références à WOA et à la finalité	Ontologie explicite et structurée	Systematisation
Méthode	Appels aux sciences empiriques	Justification épistémique rétrospective	Rationalisation