



HAL
open science

Épistémologie sociale et épistémologie bayésienne. La vie sociale des bayésiens

Denis Bonnay, Mikaël Cozic

► **To cite this version:**

Denis Bonnay, Mikaël Cozic. Épistémologie sociale et épistémologie bayésienne. La vie sociale des bayésiens. Drouet, I. Le bayésianisme aujourd'hui: Fondements et pratiques, Editions Matériologiques, pp.113-166, 2016. hal-01950806

HAL Id: hal-01950806

<https://hal.science/hal-01950806>

Submitted on 11 Dec 2018

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.

Epistémologie sociale et épistémologie bayésienne.

La vie sociale des bayésiens¹

Denis BONNAY² et Mikaël COZIC³

0 Introduction

0.1 Epistémologie bayésienne et épistémologie sociale

L'épistémologie bayésienne (EB) est une théorie normative qui porte sur la structure et la révision des croyances individuelles. Elle repose essentiellement sur trois thèses :

- (B1) *le gradualisme* : une épistémologie adéquate doit considérer les *degrés de croyance* et non pas seulement les croyances « catégoriques » ou « binaires ». L'attitude épistémique des agents vis-à-vis de propositions est affaire de degrés, qui reflètent la confiance qu'ils ont à l'égard de la vérité des propositions qu'ils considèrent.
- (B2) *le probabilisme* : les degrés de croyance d'un agent rationnel se laissent représenter par une distribution de probabilités, telle que la notion est définie dans la théorie mathématique des probabilités.
- (B3) *la révision par conditionnalisation* : quand un agent apprend une information, il doit réviser ses croyances par conditionnalisation, c'est-à-dire adopter comme nouveaux degrés de croyances les degrés de croyance initiaux conditionnels à l'information.

L'EB s'est principalement consacrée à l'analyse du comportement épistémique d'un agent isolé, comme en témoigne la vaste littérature bayésienne sur la confirmation et l'infirmité des hypothèses scientifiques⁴. On y considère typiquement l'attitude épistémique d'un agent qui envisage différentes hypothèses, et qui fait évoluer ses degrés de croyance concernant ces hypothèses en fonction des informations empiriques qu'il apprend. En revanche, jusque récemment, l'EB a beaucoup moins étudié les questions qui se posent quand on considère le

¹ Nous remercions Isabelle Drouet pour sa patience et ses encouragements, et Richard Bradley, Franz Dietrich, Olivier Roy et Jan-Willem Romeijn pour les discussions que nous avons pu avoir sur l'agrégation et la révision sociale des probabilités subjectives. Nous avons bénéficié du soutien du Labex Institut d'étude de la cognition ANR-10-LABX-0087 IEC et de l'Idex ANR-10-IDEX-0001-02-PSL*. Mikaël Cozic a aussi été soutenu par l'Institut universitaire de France

² Université Paris-Ouest Nanterre et IHPST.

³ Université Paris-Est Créteil, IUF et IHPST.

⁴ Pour un traitement approfondi, voir par exemple Colin Howson et Peter Urbach, *Scientific Reasoning : the Bayesian Approach*, La Salle, Open Court, 2006, et, pour une présentation synthétique, voir Mikaël Cozic, « Confirmation et Induction » in Anouk Barberousse, Denis Bonnay et Mikaël Cozic (éds) *Précis de Philosophie des Sciences*, Paris, Vuibert, 2011, p. 62-99.

fait qu'un individu i est rarement le seul à avoir des informations ou une opinion pertinentes concernant une certaine question Q : comment i doit-il alors tenir compte des opinions des autres individus concernant Q ? doit-il seulement tenir compte des *informations* que les autres ont obtenues à propos de Q , ou bien également de la manière dont les autres ont exploité leurs informations pour se forger une opinion concernant Q ? Et, si i délègue à un groupe d'experts le soin de se pencher sur Q , comment doit-il faire pour tirer parti des différentes opinions émises par ces experts ?

Toutes ces questions sont au cœur de l'*épistémologie sociale*. L'objectif de ce chapitre est de présenter et de discuter les contributions de l'épistémologie bayésienne à l'épistémologie sociale.

0.2 Agrégation, délibération et révision

Considérons un groupe d'agents $G = \{1, \dots, n\}$. A la suite de Richard Bradley⁵, on peut distinguer deux questions centrales. La première question est celle de l'*agrégation* : il s'agit de savoir de quelle manière il convient de dériver un état épistémique collectif ou agrégé à partir de l'état épistémique des différents membres de G . Si, conformément à l'épistémologie bayésienne, l'état épistémique d'un agent se laisse décrire par des degrés de croyance formant une distribution de probabilités, alors la question de l'agrégation prend la forme spécifique de savoir de quelle manière il convient de *dériver une distribution de probabilités collective ou agrégée, à partir des distributions de probabilités individuelles*. On peut schématiser le problème de l'agrégation ainsi : il s'agit de savoir comment doit procéder un *agrégateur*, qui est informé de l'état épistémique de chaque membre de G , et qui doit, sur cette base, former l'opinion collective de G . Il est important de souligner que l'opinion collective résultante n'a pas à être acceptée par tous les membres de G (ni par aucun en particulier) : l'agrégation, en tant que telle, ne vise pas au *consensus*. Elle vise plutôt une sorte de *compromis*, en particulier quand la méthode d'agrégation fait l'objet d'un accord entre les membres de G . Quand c'est le cas, les membres de G peuvent s'accorder à voir dans le résultat de la procédure l'opinion du groupe, même si les désaccords antérieurs à l'agrégation persistent.

L'agrégation ne suppose pas de communication entre les membres du groupe (ni même, à vrai dire, entre les membres du groupe et l'agrégateur : celui-ci peut très bien mesurer les opinions des membres sans les interroger directement). La seconde famille de questions, par contraste, trouve son origine dans des échanges d'informations entre individus. Dans bien des situations, en effet, les membres d'un groupe peuvent se donner des informations en exprimant leurs opinions. On est alors dans une situation de *révision sociale* : comment les membres de G doivent-ils modifier leurs opinions à la suite de telles interactions épistémiques ? Comment, en particulier, un membre de G doit-il réviser son opinion quand il apprend celle d'un autre membre ?

L'organisation du chapitre va suivre cette distinction fondamentale entre agrégation et révision sociale, que nous aborderons respectivement dans la première et dans la deuxième

⁵ Richard Bradley, « Taking Advantage of Difference in Opinion », *Episteme*, 3(3), 2006, p. 141-155.

Section. Nous reviendrons dans la troisième Section sur cette distinction elle-même et sur les relations qui peuvent exister entre agrégation et révision.

1 Théories de l'agrégation probabiliste

1.1 Préliminaires

La question de l'agrégation des croyances est la question de savoir de quelle manière il convient de former des croyances collectives à partir des croyances des membres d'un groupe. Une telle question est riche en applications, en particulier quand il s'agit d'agréger les opinions d'un groupe d'*experts*. On peut songer, par exemple, à la sûreté nucléaire (voir par exemple Cooke, 1991, chapitre 2, pour un aperçu des domaines où les avis des experts sont systématiquement utilisés⁶). Nous reviendrons dans la Section 3 sur l'interprétation de la question de l'agrégation. Remarquons toutefois dès maintenant qu'une bonne partie des modèles et des critères que nous allons envisager dans la suite de cette Section s'appliquent non seulement quand on considère un groupe d'agents dotés d'opinions (c'est l'interprétation que nous privilégions), mais aussi quand on considère *un* agent qui doit se forger une opinion probabiliste à partir de plusieurs *sources d'information prenant déjà la forme de distribution de probabilités*. L'agrégation probabiliste trouve ainsi des applications naturelles dans les domaines où l'on s'appuie sur une pluralité de méthodes ou de modèles comme en sciences du climat et de la Terre (voir Allard et alii, 2012)⁷.

En général, on suppose que les probabilités en jeu sont définies sur une algèbre d'événements Σ définie sur un espace d'états ou de mondes possibles Ω . Sauf mention contraire, on supposera dans ce qui suit que l'algèbre est assez « riche », au sens où elle contient au moins trois événements disjoints (une telle algèbre est parfois dite « tertiaire »), et qu'elle est au plus dénombrable (la plupart des résultats se généralisent cependant au cas continu). Notons P l'ensemble des distributions de probabilités sur Σ . Un **profil** de probabilités (p_1, \dots, p_n) est la donnée d'une distribution de probabilités par agent, la distribution p_i étant censée représenter les degrés de croyances de l'agent i . La théorie de l'agrégation probabiliste s'intéresse aux propriétés qu'on peut ou qu'on doit attendre d'une **fonction d'agrégation probabiliste**⁸, c'est-à-dire d'une fonction $F : P^n \rightarrow P$ qui associe à chaque profil de probabilités une probabilité censée représenter les croyances collectives.

1.2 Moyennes et axiomes

L'agrégateur confronté à la diversité des jugements des experts d'un panel trouvera sans doute naturel de les synthétiser en considérant la moyenne de ces jugements. Ce n'est d'ailleurs pas le moindre mérite de l'hypothèse d'une expression quantitative de ces

⁶ Roger M. Cooke, *Experts in Uncertainty*, Oxford University Press, USA, 1991.

⁷ Denis Allard, Alessandro Comunian, & Philippe Renard, « Probability Aggregation Methods in Geoscience », *Mathematical Geosciences*, vol. 44, 2012, p. 545-581.

⁸ Cette fonction est également appelée un « *pooling operator* ».

jugements que de permettre le calcul d'une telle moyenne. On pourra alors déterminer la fonction d'agrégation F de la manière suivante :

$$F(p_1, \dots, p_n) (E) = \frac{1}{n} \sum_{i=1}^n p_i(E) \quad \text{Moyenne Arithmétique Simple (MAS)}$$

La probabilité attribuée par le panel à un événement E est simplement la moyenne des probabilités attribuées à cet événement par les membres du panel. On peut facilement vérifier que le résultat de l'application d'une telle règle est nécessairement une distribution de probabilité. Mais en allant ainsi au plus simple, on a supposé que tous les experts avaient un poids égal au sein du groupe. Cela n'a rien d'évident ; l'agrégateur peut au contraire considérer que le groupe comprend certains experts particulièrement renommés dont l'opinion a plus de poids que celle d'autres experts moins reconnus. Il préférera alors une moyenne arithmétique *pondérée* :

$$F(p_1, \dots, p_n) (E) = \sum_{i=1}^n k_i \cdot p_i(E) \quad \text{Moyenne Arithmétique Pondérée (MAP)}$$

où k_i est un nombre positif ou nul qui représente le poids de l'expert i , avec $\sum_{i=1}^n k_i = 1$ pour garantir que F est une distribution de probabilité. Lorsque le poids accordé à l'expert i est plus grand que le poids accordé à l'expert j , c'est-à-dire $k_i > k_j$, la probabilité $p_i(E)$ que l'expert i attribue à l'événement E a davantage d'impact sur la formation de la probabilité de groupe $F(p_1, \dots, p_n) (E)$ que la probabilité $p_j(E)$ que l'expert j attribue au même événement E . MAS et MAP ont l'avantage de la simplicité. Simples à expliquer et à comprendre, elles rendent bien compte de l'idée selon laquelle l'agrégateur attribue l'opinion collective en faisant la synthèse des opinions individuelles. MAP a en outre une forme paramétrique qui permet de prendre en compte la diversité épistémique du groupe. Pour autant, il y a d'autres moyennes que la moyenne arithmétique. On pourrait par exemple attribuer l'opinion collective en utilisant une moyenne géométrique. Pour un monde w , correspondant à un événement atomique de l'algèbre, on demandera

$$F(p_1, \dots, p_n) (w) = \frac{1}{c} \left(\prod_{i=1}^n p_i(w) \right)^{1/n} \quad \text{Moyenne Géométrique Simple (MGS)}$$

où c est une constante de normalisation qui vaut $\sum_{w' \in \Omega} \left(\prod_{i=1}^n p_i(w') \right)^{1/n}$ et dont le rôle est de garantir que F soit une distribution de probabilités en assurant que $\sum_{w' \in \Omega} F(p_1, \dots, p_n) (w') = 1$. Le choix d'une moyenne géométrique est particulièrement naturel si l'on s'intéresse moins aux valeurs absolues des probabilités qu'aux ratios, car la moyenne géométrique commute avec les ratios. Enfin, comme pour la moyenne arithmétique, on peut prendre en compte le fait que l'avis des différents experts n'a pas la même importance en leur attribuant des poids. On utilisera alors une moyenne géométrique pondérée de la manière suivante :

$$F(p_1, \dots, p_n) (w) = \frac{1}{c} \prod_{i=1}^n p_i(w)^{k_i} \quad \text{Moyenne Géométrique Pondérée (MGP)}$$

où k_i est le poids de l'expert i avec $\sum_{i=1}^n k_i = 1$ et $c = \sum_{w' \in \Omega} \prod_{i=1}^n p_i(w')^{k_i}$ la constante de normalisation. F , ainsi définie sur les atomes de Ω , s'étend de manière unique en une distribution définie sur l'algèbre Σ tout entière.

Que de telles méthodes apparaissent simples et naturelles ne suffit pas à en justifier l'usage, d'autant que leur diversité même invite à se demander pourquoi l'on devrait considérer que l'opinion de groupe résulte de l'application de l'une plutôt que de l'autre. A l'instar d'autres théories agrégatives comme les théories de l'agrégation des préférences ou des utilités dans la tradition d'Arrow (voir par exemple, pour un état de l'art contemporain, l'ouvrage collectif édité par Arrow, Sen et Suzumura⁹) ou comme la théorie de l'agrégation des jugements (voir Mongin et Dietrich, 2010)¹⁰, la théorie de l'agrégation probabiliste adopte, pour répondre à la question de savoir ce qui justifie telle ou telle règle, une approche *axiomatique* : elle s'intéresse aux relations entre des fonctions d'agrégation et des conditions abstraites contraignant ces fonctions, que l'agrégateur peut juger désirables ou pas, et montre que l'on devra choisir telle classe de fonctions (par exemple les MAP) si et seulement si on considère que l'agrégation doit satisfaire les propriétés énoncées par tels et tels axiomes.

L'une des conditions les plus simples et les plus attractives est la Préservation de l'Unanimité. Selon cette condition, quand il y a consensus parmi les membres du groupe sur la probabilité à accorder à un événement E , alors la probabilité collective doit s'aligner sur le consensus :

(PU) Préservation de l'Unanimité

Pour tout profil $(p_1, \dots, p_n) \in P^n$ et tout événement $E \in \Sigma$,

$$\text{si } p_1(E) = \dots = p_n(E) = \alpha, \text{ alors } F(p_1, \dots, p_n)(E) = \alpha.$$

Les fonctions d'agrégation qui ont le plus retenu l'attention sont les moyennes arithmétiques pondérées, c'est-à-dire les fonctions de la classe MAP (on parle parfois aussi de « linear pooling »). Σ . On peut aussi vérifier que, si un agrégateur suit une telle fonction, alors il obéit à la Préservation de l'Unanimité. La question théorique qui s'est naturellement posée est celle de savoir quel ensemble de conditions sur les fonctions d'agrégation *caractérise* la classe (MAP). La réponse met en jeu la propriété suivante, familière aux théoriciens du choix social :

(N) Neutralité¹¹

Il existe une fonction $m : [0, 1]^n \rightarrow [0, 1]$ telle que pour tout profil $(p_1, \dots, p_n) \in P^n$ et pour tout événement $E \in \Sigma$,

$$F(p_1, \dots, p_n)(E) = m(p_1(E), \dots, p_n(E))$$

La condition de Neutralité exprime une double idée : d'abord celle selon laquelle la croyance collective en un événement ne dépend que de la croyance des membres du groupe en *cet* événement. Et ensuite celle selon laquelle la méthode par laquelle la croyance collective en un

⁹ Kenneth J. Arrow, Amartya Sen & Kotaro Suzumura, *Handbook of Social Choice and Welfare*, 2 volumes, Amsterdam, North-Holland, 2011 et 2012.

¹⁰ Philippe Mongin & Franz Dietrich « Un bilan interprétatif de la théorie de l'agrégation logique », *Revue d'économie politique*, 120(6), 2010, p. 929-972.

¹¹ Les spécialistes de l'agrégation probabiliste parlent souvent de « *strong setwise function property* ».

événement est déterminée à partir des croyances individuelles en cet événement est la même, quel que soit l'événement envisagé. La première de ces idées peut se formuler, séparément, de la manière suivante :

(I) Indépendance¹²

Pour tout événement $E \in \Sigma$, il existe une fonction $m_E : [0, 1]^n \rightarrow [0, 1]$ telle que pour tout profil $(p_1, \dots, p_n) \in P^n$,

$$F(p_1, \dots, p_n)(E) = m_E(p_1(E), \dots, p_n(E))$$

La condition (I) est logiquement plus faible que la condition (N). Mais il s'avère qu'un cas particulier de Préservation de l'Unanimité suffit à renforcer la première au point de la rendre équivalente à la seconde. L'affaiblissement de (PU) en question est celui-ci :

(PU-0) Préservation de l'Unanimité-0

Pour tout profil $(p_1, \dots, p_n) \in P^n$, et tout événement $E \in \Sigma$,

$$\text{si } p_1(E) = \dots = p_n(E) = 0, \text{ alors } F(p_1, \dots, p_n)(E) = 0.$$

L'équivalence prend la forme de la Proposition suivante :

Proposition 1

Une fonction d'agrégation probabiliste satisfait (N) ssi elle satisfait (I) et (PU-0).

Nous sommes maintenant en mesure d'énoncer le théorème de caractérisation de la moyenne pondérée.

Théorème 2 (McConway¹³, Wagner¹⁴)

Une fonction d'agrégation probabiliste F satisfait (N) ssi F est dans la classe MAP.

Autrement dit, (N) suffit à caractériser l'agrégation par moyenne arithmétique pondérée¹⁵.

Le calcul « événement par événement » et la préservation de l'unanimité, en particulier de l'unanimité-0, sont des propriétés naturelles pour une moyenne, dont il est très facile de montrer que les moyennes arithmétiques pondérées les satisfont. La Proposition 1 et le Théorème 2 nous disent qu'elles suffisent à caractériser les fonctions de la classe MAP, ce qui était loin d'être évident. On peut également ajouter une condition d'invariance par permutation sur les agents, qui dit que F ne doit dépendre que des distributions, et pas de quel

¹² Le terme d'indépendance vient plutôt des théories du choix social et de l'agrégation des jugements. Les spécialistes de l'agrégation probabiliste parlent volontiers de « *weak setwise function property* ».

¹³ Kevin J. McConway, « Marginalization and linear opinion pools », *Journal of the American Statistical Association*, 76, 1981, p. 410-15.

¹⁴ Carl Wagner « Allocation, Lehrer Models, and the Consensus of Probabilities », *Theory and Decision* 14(2), 1982, p. 207-220.

¹⁵ Pour être précis, il faut ajouter que Wagner traite le cas d'une fonction de masse probabiliste $p : \Omega \rightarrow \mathbb{R}$ avec Ω fini tandis que McConway considère des probabilités définies sur des σ -algèbres..

agent a quelle distribution. Cette condition, ajoutée à (N), permet de caractériser la classe MAS.

La condition de Neutralité fait face à deux familles de critiques : d'une part, celles qui portent spécifiquement sur la condition d'Indépendance, et d'autre part, celles qui portent sur ce que la Neutralité ajoute à l'Indépendance – à savoir l'idée que la même méthode doit s'appliquer à tous les événements.

Cette seconde famille de critiques est aisée à comprendre, et nous permet d'introduire une dimension importante de l'agrégation épistémique. Supposons en effet que le groupe G soit *hétérogène* du point de vue de la compétence épistémique 'inter-événementielle' de ses membres: certains agents sont particulièrement aptes à juger si E , tandis que d'autres sont plus aptes à juger si E' . L'hétérogénéité inter-événementielle se distingue de l'hétérogénéité épistémique *globale*, qui tient dans le fait qu'un membre de G peut être plus compétent en général, et qui peut se traduire naturellement dans les poids w_i que l'agrégateur accorde aux différents agents. Dans le cas inter-événementiel, l'agrégateur voudra certainement accorder un poids différent aux agents quand il déterminera les croyances collectives vis-à-vis de E et de E' . Ces cas d'hétérogénéité épistémique inter-événementielle suggèrent une généralisation simple de la moyenne pondérée : on pourrait vouloir *relativiser les poids individuels aux événements en jeu*. Si w_i^E dénote le poids accordé à l'agent i quand il s'agit de prendre en compte la probabilité qu'il assigne à l'événement E , alors la généralisation suggérée se formule de la manière suivante :

$$F(p_1, \dots, p_n) (E) = \sum_{i=1}^n w_i^E \cdot p_i (E) \quad \text{(MAP-rel)}$$

La difficulté que rencontre immédiatement cette suggestion, c'est que, dans le cas général, la fonction collective engendrée ne sera pas une distribution de probabilités, de sorte que cette règle ne définit pas une classe de fonctions d'agrégation probabiliste.

1.3 Indépendance probabiliste, conditionnalisation et moyenne géométrique

Un autre réquisit naturel sur les fonctions d'agrégation probabiliste consiste à exiger que, si tous les membres du groupe considèrent les événements E et E' comme indépendants du point de vue probabiliste, alors les croyances collectives doivent elles-mêmes considérer E et E' comme indépendants. Formellement,

(PIP) Préservation de l'indépendance probabiliste¹⁶

Pour tout profil (p_1, \dots, p_n) , et tous événements $E, E' \in \Sigma$,

$$\text{si pour tout } i \in G, \quad p_i(E \cap E') = p_i(E) \times p_i(E'),$$

$$\text{alors } F(p_1, \dots, p_n) (E \cap E') = F(p_1, \dots, p_n) (E) \times F(p_1, \dots, p_n) (E')$$

¹⁶ Pour une étude récente, qui comporte un rappel des résultats principaux concernant (PIP), voir Carl Wagner « Peer Disagreement and Independence Preservation », *Erkenntnis*, 74, p. 277-288, 2011.

La condition (PIP) est déjà attractive quand on l'envisage de manière *statique*, puisqu'elle affirme que lorsqu'il y a consensus parmi les membres du groupe sur l'indépendance de deux événements, celle-ci est également endossée par les croyances agrégées. Du point de vue bayésien, elle prend encore davantage de relief quand on l'envisage du point de vue *dynamique*. En effet, d'après l'épistémologie bayésienne, un agent rationnel change ses degrés de croyance(s) par conditionnalisation. Cela implique que si deux événements sont indépendants pour l'agent i , apprendre que l'un des deux est advenu ne changera pas les croyances de i à propos de l'autre. Si tous les agents considèrent que E et E' sont indépendants, alors aucun d'entre eux n'est prêt à changer ses croyances sur l'un s'il apprend que l'autre est survenu. La condition (PIP) signifie alors que, si aucun des agents n'est prêt à changer ses croyances sur l'un des événements s'il apprend l'autre, alors le groupe non plus n'est pas prêt à changer ses croyances sur l'un s'il apprend l'autre.

En général, la moyenne pondérée *ne respecte pas* la préservation de l'indépendance probabiliste. C'est ce que met en évidence l'*Exemple* suivant.

Exemple 1

$$\Omega = \{w_1, w_2, w_3, w_4\}, E = \{w_1, w_2\} \text{ et } E' = \{w_2, w_3\}$$

Croyances de l'agent 1 :

$$p_1(w_1) = p_1(w_2) = p_1(w_3) = p_1(w_4) = 1/4$$

$$p_1(E) = p_1(E') = 1/2$$

$$p_1(E \cap E') = 1/4 = p_1(E) \times p_1(E')$$

Croyances de l'agent 2 :

$$p_2(w_1) = 4/7 ; p_2(w_2) = 1/7$$

$$p_2(w_3) = 2/35 ; p_2(w_4) = 8/35$$

$$p_2(E) = 5/7 ; p_2(E') = 1/5$$

$$p_2(E \cap E') = 1/7 = p_2(E) \times p_2(E')$$

Croyances agrégées : supposons que l'on agrège avec la fonction F définie par comme la moyenne arithmétique pondérée avec $k_1 = 4/10 = 2/5$ et $k_2 = 6/10 = 3/5$

$$F(p_1, p_2)(E) = 22/35 ; F(p_1, p_2)(E') = 8/25$$

$$F(p_1, p_2)(E \cap E') = 13/70 \approx 0,186 \neq F(p_1, p_2)(E) \times F(p_1, p_2)(E') = 176/875 \approx 0,201$$

Ce non-respect de (PIP) est l'un des arguments invoqués contre cette famille de fonctions d'agrégation probabiliste (Laddaga, 1977¹⁷). En particulier, si l'on considère que la question de savoir quels événements sont indépendants doit être traitée *a priori* et faire l'objet d'un consensus précédant l'attribution des probabilités aux événements individuels, alors on doit demander que l'agrégation probabiliste respecte ce consensus et satisfasse (PIP). Lehrer et Wagner (1983, Théorème 2¹⁸) ont montré en substance que la seule possibilité pour la règle

¹⁷ Robert Laddaga, « Lehrer and the consensus proposal », *Synthese*, 36, 1977, p. 473-477.

¹⁸ Keith Lehrer & Carl Wagner « Amalgamation and the Independence Issue: A Reply to Laddaga », *Synthese*, 55, 1983, p. 339-346

de moyenne pondérée de respecter (PIP) est d'être dictatoriale¹⁹, c'est-à-dire d'accorder un poids maximal $w_d = 1$ à l'un des agents, le dictateur.²⁰

Un autre réquisit qu'il est naturel de considérer est le suivant : qu'il ne fasse pas de différence pour le résultat qu'une information soit d'abord reçue par les membres du groupe puis qu'on agrège leurs opinions («réviser-puis-agrégé») ou qu'on agrège les opinions initiales et que l'information soit ensuite fournie à la probabilité agrégée («agréger-puis-réviser»). Sous l'hypothèse usuelle que la révision s'opère par conditionnalisation, ce réquisit s'exprime ainsi :

(CC) Commutativité avec la conditionnalisation²¹

Pour tout profil (p_1, \dots, p_n) , et tout événement $E \in \Sigma$ tel que pour tout $i \in G$, $p_i(E) > 0$

$$F(p_1(\cdot|E), \dots, p_n(\cdot|E)) = F(p_1, \dots, p_n)(\cdot|E)$$

L'*Exemple 1* montre que la moyenne arithmétique pondérée n'obéit pas non plus à la condition (CC). Puisque les agents 1 et 2 jugent les événements E et E' indépendants, s'ils apprennent que E' est le cas, leurs croyances en E ne changeront pas, et par conséquent la croyance agrégée non plus. En revanche, nous avons vu que les deux événements devenaient dépendants (en probabilité) quand on agrégeait directement les probabilités initiales. Par conséquent, $F(p_2(\cdot|E'), p_2(\cdot|E')) \neq F(p_1, p_2)(\cdot|E')$.

On peut renforcer la Commutativité avec la conditionnalisation, en exigeant la commutativité avec une famille plus vaste de méthodes de révision des croyances. En effet, pour un événement $E \subseteq \Omega$, la probabilité conditionnelle a la forme

$$p_i(\omega|E) = \frac{\mathbf{1}_E(\omega) \cdot p_i(\omega)}{\sum_{\omega \in \Omega} \mathbf{1}_E(\omega) \cdot p_i(\omega)}$$

où $\mathbf{1}_E : \Omega \rightarrow \{0, 1\}$ est la fonction caractéristique de l'événement E , qui assigne 1 à un état s'il appartient à E et 0 sinon. De manière plus générale, on peut considérer une fonction révisée qui a la forme

$$p_i^L(\omega) = \frac{L(\omega) \cdot p_i(\omega)}{\sum_{\omega \in \Omega} L(\omega) \cdot p_i(\omega)}$$

où $L(\cdot)$ est une fonction qui indique, pour chaque état, le support qui lui est apporté par une certaine information. On appelle une telle fonction une **fonction de vraisemblance**. On la définit comme suit dans le cas discret (Wagner, 2009). Soit un profil (p_1, \dots, p_n) de probabilités. $L: \Omega \rightarrow [0, \infty)$ est une **fonction de vraisemblance** pour (p_1, \dots, p_n) si pour tout $i \in G$,

¹⁹ De manière générale, une fonction d'agrégation probabiliste est dictatoriale s'il existe un agent $d \in \{1, \dots, n\}$ tel que pour tout profil (p_1, \dots, p_n) , $F(p_1, \dots, p_n) = p_d$.

²⁰ Ce résultat figure dans un argumentaire destiné à montrer, premièrement, que la moyenne pondérée et (PIP) ne sont pas conciliables, et, deuxièmement, qu'il faut rejeter (PIP) plutôt que la moyenne pondérée.

²¹ Carl Wagner, « Jeffrey conditioning and external Bayesianity », *Logic Journal of the IGPL*, 18(2), p. 336-345, 2010.

$$0 < \sum_{\omega \in \Omega} L(\omega) \cdot p_i(\omega)$$

et si (p_1^L, \dots, p_n^L) est une suite de distributions de probabilités où

$$p_i^L(\omega) = \frac{L(\omega) \cdot p_i(\omega)}{\sum_{\omega \in \Omega} L(\omega) \cdot p_i(\omega)} (*)$$

Dans ce cadre, la commutation de l'apprentissage et de l'agrégation s'exprime par la propriété dite de « Bayésianité externe » :

(BE) Bayésianité externe²²

F satisfait la propriété de Bayésianité externe si pour tout profil (p_1, \dots, p_n) et pour toute fonction de vraisemblance L pour (p_1, \dots, p_n) , alors L est aussi une fonction de vraisemblance pour $F(p_1, \dots, p_n)$ et la propriété de commutativité suivante vaut :

$$F(p_1^L, \dots, p_n^L) = [F(p_1, \dots, p_n)]^L$$

La Commutativité avec la conditionnalisation n'étant qu'un cas particulier de la Bayésianité externe, il suit que la Moyenne Arithmétique Pondérée viole la Bayésianité externe. En cela, elle diffère notablement de la Moyenne Géométrique Pondérée.

Proposition 3

Si une fonction d'agrégation probabiliste appartient à la classe des Moyennes Géométriques Pondérées, elle satisfait la Bayésianité externe.

La classe des fonctions d'agrégation qui satisfont la Bayésianité externe est plus large que la classe des fonctions qui se laissent décrire par une moyenne géométrique pondérée. Elle a été caractérisée par Genest, McConway & Schervish (*ibid.*). Genest (1984)²³ a également donné une caractérisation de la moyenne géométrique en ajoutant une condition à la propriété de Bayésianité externe.

Pour compléter la comparaison entre moyennes arithmétique et géométrique, il faut ajouter que si la moyenne géométrique satisfait également l'affaiblissement de la Préservation de l'Unanimité (PU) selon lequel, si toutes les probabilités sont identiques à p , alors $F(p_1, \dots, p_n)^L = p$, elle ne satisfait pas l'Indépendance. Cela est dû au fait que la probabilité agrégée d'un monde est affectée non seulement par les probabilités individuelles de ce monde, mais aussi, via la constante de normalisation, par les probabilités individuelles

²² La propriété de Bayésianité externe est en général formulée et étudiée dans le cas continu. Voir par exemple Christian Genest, Kevin J. McConway et Mark J. Schervish « Characterization of externally Bayesian pooling operators », *Annals of Statistics*, 14(2), p. 487-501, 1986.

²³ Christian Genest, « A characterization theorem for externally Bayesian group », *The Annals of Statistics*, 12, p. 1100-1105, 1984.

des autres mondes. Par ailleurs, on peut penser que la Préservation de l'Indépendance Probabiliste (PIP) marque une autre différence avec la moyenne arithmétique. Ainsi, l'utilisation d'une moyenne pondérée géométrique plutôt qu'arithmétique permet la préservation de l'indépendance dans le cas de l'*Exemple* précédent :

Exemple 1 (suite)

Ω, E, E', p_1 et p_2 sont définis comme précédemment

Croyances agrégées : supposons que l'on agrège par la fonction F définie comme la moyenne géométrique pondérée avec $k_1 = 4/10 = 2/5$ et $k_2 = 6/10 = 3/5$

$$F(p_1, p_2)(E) \approx 0,634$$

$$F(p_1, p_2)(E') \approx 0,303$$

$$F(p_1, p_2)(E \cap E') = F(p_1, p_2)(E) \times F(p_1, p_2)(E') \approx 0.192$$

De fait, on peut ainsi prouver que, pour toutes les algèbres comprenant au plus quatre atomes, comme dans l'*Exemple*, les fonctions de la classe MGP préservent l'indépendance. Malheureusement, ce résultat ne vaut pas pour les algèbres avec 5 atomes ou plus, ni a fortiori pour les algèbres infinies : les seules fonctions de la classe MGP, et plus généralement, les seules fonctions qui, comme la moyenne géométrique, satisfont (I) à une renormalisation près et qui satisfont (PIP), sont dictatoriales²⁴.

1.4 Discussion

1.4.1 La question de l'allocation des poids

Nous venons de présenter les deux méthodes d'agrégation probabiliste les plus discutées, par moyenne arithmétique et par moyenne géométrique. Elles sont toutes deux paramétriques : le résultat de l'agrégation dépend de l'assignation de poids aux différents agents. Cette flexibilité semble bienvenue : elle permet de prendre en compte le fait que tous les agents n'ont « pas le même poids ». Pour autant, parler ainsi en termes de poids reste un peu mystérieux, tout comme restent mystérieux les principes ayant vocation à guider l'attribution de ces poids. Intuitivement, les poids représentent l'importance des différents agents. Mais qu'entend-on ici par importance ? On peut imaginer au moins deux réponses, fort différentes. Premièrement, il peut s'agir de l'importance au sein du groupe de chacun des agents. Dans un groupe, tout le monde n'a pas le même statut, la voix de certains compte plus que la voix d'autres. La notion d'importance ainsi visée est une notion primitive, attachée au groupe en tant qu'entité collective. Il s'agit de sanctionner une réalité de l'organisation collective du groupe, sans préjuger des justifications de cet état de fait – un membre peut

²⁴ Christian Genest et Carl Wagner (1987) « Further Evidence against Independence Preservation in Expert Judgement Synthesis », *Aequationes Mathematicae*, 32 (1), 1987, p. 74-86.

avoir plus d'importance parce qu'il a plus de chances d'avoir raison, parce que les autres membres considèrent qu'il a plus de chances d'avoir raison, ou simplement parce qu'il est « le chef » du groupe en question. Quant à l'attribution des poids de manière à capturer cette importance, plusieurs façons de faire sont envisageables. La répartition des poids peut se faire par les agents eux-mêmes, s'ils parviennent à un consensus concernant leur importance respective au sein du groupe, ou par l'agrégateur, sur la base de son observation du fonctionnement du groupe.

Une autre interprétation des poids est possible, en termes directement épistémiques. Si l'on considère que l'agrégateur vise à former la meilleure opinion possible sur la base des avis du groupe d'experts, on peut interpréter les poids comme le reflet de la compétence que l'agrégateur attribue aux différents experts : plus un expert est considéré comme compétent en comparaison des autres experts du groupe, plus son opinion doit être décisive dans la formation de l'opinion collective, et donc plus le poids qui lui est accordé doit être élevé. Cette interprétation s'impose si l'on considère que les méthodes d'agrégation sont des méthodes destinées à nous permettre d'atteindre la vérité, en tout cas de nous en approcher ou de produire les meilleures estimations possibles. Dans cette perspective, on remarque d'emblée que les méthodes d'agrégation telles que considérées jusqu'ici ne nous disent pas comment faire. Elles ne disent rien quant à la manière de dériver, de justifier ou d'assigner les poids individuels w_1, \dots, w_n , alors même que ces poids ont un impact crucial sur le résultat de l'agrégation. Les poids sont censés refléter la compétence épistémique que l'agrégateur accorde aux différents membres du groupe, mais les résultats de caractérisation n'indiquent pas de méthode qui permettrait de fonder une mesure de la compétence épistémique des agents. Cette difficulté a été repérée depuis longtemps dans la littérature probabiliste. Genest & McConway²⁵ fournissent un panorama des grands types de réponses qu'on a tenté de lui apporter. Ces tentatives peuvent être intéressantes dans des contextes particuliers, mais, à ce jour, il n'existe pas de méthode conceptuellement satisfaisante permettant de fournir une solution générale au problème²⁶. On peut donner une idée des difficultés rencontrées en considérant une interprétation naturelle du poids w_i comme reflétant la probabilité pour que la distribution de probabilités de l'agent i soit « correcte ». On appelle parfois ces probabilités, qui ont des probabilités subjectives pour objet, des probabilités « véridiques »²⁷. La difficulté première tient au fait que l'usage de ces probabilités suppose que l'agrégateur considère qu'un et exactement un des membres du groupe aura une probabilité subjective « correcte », puisque la somme des poids vaut 1. C'est une hypothèse souvent héroïque et qui par ailleurs s'accommode difficilement du cadre de l'agrégation où les poids sont alloués *ex ante* et ne sont pas censés être modifiés une fois que l'agrégateur est informé des probabilités des différents membres du groupe. Il se peut donc que deux agents se voient attribuer deux poids

²⁵ Christian Genest et Kevin J. McConway, « Allocating the weights in the linear opinion pool », *Journal of Forecasting*, 9 (1), p. 53-73, 1990.

²⁶ Ce constat pessimiste est partagé par Robert T. Clemen et Robert L. Winkler, « Combining Probability Distributions From Experts in Risk Analysis », *Risk Analysis*, 19 (2), p. 187-203, 2007, qui écrivent : « En dépit des travaux de Genest et McConway (*ibid.*), nous ne disposons d'aucune méthode fondée en principe afin de déterminer les poids [des moyennes arithmétiques et géométriques pondérées] » (p. 207).

²⁷ L'expression est de D. Bunn, « Two Methodologies for the Linear Combination of Forecasts », *The Journal of Operational Research Society*, 32(3), pp. 213-222, 1981.

différents (donc deux probabilités différentes d'émettre des probabilités subjectives correctes, d'après l'interprétation que nous considérons) mais aient pourtant les mêmes distributions de probabilités. Pire encore, des règles comme les moyennes peuvent parfaitement aboutir à des distributions qui ne coïncident avec aucune des distributions individuelles, et donc qui a une probabilité nulle d'être la probabilité subjective « correcte ». Enfin, l'idée même d'une probabilité subjective « correcte » ne va pas de soi, en particulier quand il ne s'agit pas d'une probabilité subjective qui cherche à épouser une probabilité objective supposée. Quelle est, par exemple, la probabilité correcte pour que tel cheval, participant à telle course de chevaux le week-end prochain, gagne la course en question²⁸ ? La probabilité dégénérée qui accorde 0 ou 1 au cheval selon le résultat de la course ? La probabilité qui semble la mieux justifiée au vu des informations disponibles ?

1.4.2 Indépendance et localité

Dans la section précédente, nous avons discuté des points aveugles de l'axiomatisation, étant entendu que l'on allait utiliser une des deux méthodes présentées. Revenons maintenant sur les raisons que l'axiomatique nous donne de choisir l'une de ces deux méthodes. Le travail d'axiomatisation a montré le lien qui unit agrégation par moyenne arithmétique et indépendance. Y a-t-il de bonnes raisons d'imposer la condition d'Indépendance, qui force la fonction d'agrégation à être *locale*, c'est-à-dire à déterminer la distribution de probabilité collective « événement-par-événement » ?

Notons tout d'abord que les contreparties de l'Indépendance en théorie du choix social²⁹ et en théorie de l'agrégation des jugements, qui jouent un rôle essentiel dans les résultats d'impossibilité comme celui d'Arrow, ont été abondamment étudiées et discutées dans ces contextes. En général, deux arguments *pragmatiques* sont avancés en faveur des conditions d'Indépendance³⁰. Le premier est un argument de parcimonie : l'Indépendance garantit que, quand on veut déterminer l'attitude collective (préférence, utilité, jugement, probabilité) à propos d'un item (paire d'options, option isolée, proposition, événement), il suffit de s'informer des attitudes individuelles à l'égard de l'item en question. Le second argument concerne la *manipulabilité* de la procédure d'agrégation. Jusqu'à présent, nous avons fait abstraction de la manière dont l'agrégateur était informé des attitudes des individus. Il se peut que l'agrégateur doive se reposer sur les déclarations des individus. Mais la procédure d'agrégation peut ne pas inciter les individus à rapporter leurs vraies attitudes. On dit alors qu'elle est (individuellement) *manipulable*. Or, les conditions d'Indépendance peuvent jouer

²⁸ Nous considérons cet exemple comme un cas *potentiel* où une probabilité subjective existe sans contrepartie objective. Mais nous ne voulons pas prendre parti sur la question de savoir s'il y a réellement des cas de ce genre, ni sur celle de savoir quelle est la meilleure conception des probabilités objectives.

²⁹ On l'appelle, en théorie de l'agrégation des préférences, la condition d' « Indépendance des alternatives non-pertinentes ».

³⁰ Ici et dans ce qui suit, nous empruntons à Franz Dietrich et Christian List, « Probabilistic Opinion Pooling », Manuscrit, personal.lse.ac.uk/list/PDF-fileshttp://personal.lse.ac.uk/list/PDF-files/OpinionPoolingReview.pdf/OpinionPoolingReview.pdf, 2014. A paraître dans Alan Hajek et Christopher Hitchcock (eds.) *The Oxford Handbook of Probability and Philosophy*, Oxford: Oxford University Press.

un rôle crucial dans la non-manipulabilité d'une procédure d'agrégation³¹. Par ailleurs, on a supposé que les individus avaient à se prononcer sur un ensemble fixe et déterminé d'items. Or, il se peut qu'un « architecte » ait un contrôle sur la liste d'items à propos desquels les individus sont sollicités – ce qu'on appelle, en théorie de l'agrégation des jugements, un *agenda*. Et dans certains cas, l'architecte peut manipuler l'issue de la procédure d'agrégation en manipulant l'agenda. Ici encore, les conditions d'Indépendance s'avèrent cruciales pour empêcher la manipulation³². Il est aisé de comprendre pourquoi : si l'attitude collective concernant l'item x dépend des attitudes individuelles concernant un item $y \neq x$, alors en mettant ou en ne mettant pas y dans l'agenda, l'architecte peut influencer l'attitude collective concernant x .

Pour mieux saisir le lien entre indépendance et manipulabilité dans le cas probabiliste, faisons un détour par une propriété introduite par McConway, qu'il appelle la propriété de Marginalisation (Marg). Celle-ci se formule dans un cadre légèrement différent du cadre habituel de l'agrégation des probabilités. Au lieu d'avoir une fonction d'agrégation probabiliste définie sur un n -uplet de mesures de probabilités définies sur une algèbre fixe S , McConway considère une **classe F de fonctions d'agrégation probabiliste**, chaque fonction F_S de la classe étant indexée sur une σ -algèbre de Ω . Considérons une σ -algèbre Σ et l'une de ses sous- σ -algèbres T . La **marginalisation** d'une probabilité p sur Σ est la probabilité p^T sur T qui assigne aux événements de Σ compris dans T la probabilité que p leur assignait déjà. Une classe **F** de fonctions d'agrégation probabilistes obéit à la propriété de marginalisation si les opérations de marginalisation et d'agrégation commutent. Formellement,

Marginalisation (Marg)

Une classe **F** a la propriété de marginalisation si pour toute σ -algèbre Σ ,

$$[F_S(p_1, \dots, p_n)]^T = F_T(p_1^T, \dots, p_n^T)$$

pour toute sous- σ -algèbre T de Σ et toutes probabilités p_1, \dots, p_n sur Σ .

McConway³³ a montré que la propriété de marginalisation est équivalente à la condition d'Indépendance. Comme le suggèrent Dietrich & List, ce résultat éclaire, dans le cadre spécifique de l'agrégation probabiliste, le lien entre Indépendance et non-manipulabilité de l'agenda. En effet, on peut comprendre la propriété de marginalisation (et donc l'Indépendance dans un cadre probabiliste) comme affirmant qu'un architecte ne peut influencer l'opinion collective sur les événements de T en faisant varier la sur- σ -algèbre à laquelle s'applique la fonction d'agrégation probabiliste, c'est-à-dire en faisant varier l'agenda.

³¹ En théorie de l'agrégation des jugements, Dietrich & List ont montré que l'Indépendance, conjointement avec une condition de Monotonie, caractérise les procédures non-manipulables (Franz Dietrich et Christian List « Strategy-proof judgment aggregation », *Economics and Philosophy*, 2, 2007, p. 269-300).

³² Voir Franz Dietrich « Judgment Aggregation: (Im)possibility Theorems », *Journal of Economic Theory*, vol. 126(1), 2006, p. 286-298.

³³ Kevin J. McConway, « Marginalization and Linear Opinion Pools », *Journal of the American Statistical Association*, 76 (374), 1981, p. 410-414.

Y a-t-il, à côté de ces deux familles d'arguments pragmatiques, des arguments proprement épistémiques en faveur de la condition d'Indépendance ? On peut proposer un argument par l'ignorance. Si les distributions de probabilités des membres du groupe, et donc la probabilité qu'ils attachent à chacun des événements, sont la seule information dont dispose l'agrégateur, de sorte que, en particulier, il n'a aucune information particulière sur les relations de dépendance entre les différents événements, alors il semble raisonnable d'agréger « événement par événement ». En effet, par hypothèse, l'agrégateur ne sait pas quoi conclure pour un événement donné des informations dont il dispose concernant d'autres événements. A contrario, la condition (PIP) semble demander plus de la part de l'agrégateur : celui-ci doit non seulement considérer les probabilités individuelles, mais aussi préserver les relations d'indépendance, ce qui semble supposer des connaissances supplémentaires de sa part (pour justifier de la pertinence de cette préservation). Néanmoins, cette lecture est contestable : un partisan de (PIP) pourrait rétorquer que les distributions de probabilités des agents contiennent deux types d'information, des informations de premier niveau sur les probabilités attribuées aux différents événements, des informations de second niveau sur les relations entre ces événements (dépendance ou indépendance). Respecter l'Indépendance, c'est choisir une certaine manière de prendre en compte les informations du premier type et surtout ignorer les informations du second type, respecter (PIP), c'est subordonner la prise en compte des informations du premier type à la prise en compte de certaines informations du second type. Autrement dit, les caractérisations axiomatiques reposant d'un côté sur (I) et de l'autre sur (PIP) et ses généralisations montreraient la nécessité de choisir, dans les distributions de probabilité, le genre d'information qu'on souhaite privilégier.

On peut être surpris par la force de la condition d'Indépendance, et par le fait qu'elle suffit, presque à elle seule, modulo (PU-0), à forcer la moyenne arithmétique. Après tout, si les fonctions de la classe MAP opèrent point par point, n'en va-t-il pas de même des fonctions de la classe MGP ? Le diable est ici dans les détails : les fonctions de la classe MGP ne sont pas de pures moyennes géométriques, mais des moyennes géométriques normalisées. Le coefficient de normalisation dépend des moyennes obtenues pour tous les atomes de l'algèbre, et c'est donc à cause de lui que les fonctions de la classe MGP ne satisfont pas la condition d'Indépendance. On pourrait cependant protester qu'il s'agit là d'un artifice du cadre de représentation, plutôt que d'une propriété substantielle des méthodes considérées. L'esprit de la méthode géométrique est bien « événement par événement », même si la lettre de son application ne l'est pas, pour garantir que les probabilités atomiques somment à 1. Un résultat de Gilardoni³⁴ éclaire la compatibilité de principe entre indépendance et moyenne géométrique. Au lieu de spécifier les probabilités attachées à l'algèbre Σ en donnant les probabilités absolues associées aux événements atomiques (et, partant, à l'ensemble des événements), on pourrait tout aussi bien le faire en donnant les ratios de probabilité associés aux différentes paires d'événements atomiques. De même, la fonction d'agrégation ne calculerait pas directement les probabilités absolues des événements, mais les probabilités relatives des événements atomiques. Dans un tel cadre, la condition d'indépendance s'énonce de la manière suivante

³⁴ Gustavo L. Gilardoni, « On irrelevance of alternatives and opinion pooling », *Brazilian Journal of Probability and Statistics*, 16, 2002, p. 87-98.

(IR) Indépendance des ratios

Pour toute paire d'atomes $w_i, w_j \in \Omega$, il existe une fonction $m_{ij} : [0, 1]^n \rightarrow [0, 1]$ telle que pour tout profil $(p_1, \dots, p_n) \in P^n$,

$$\frac{F(p_1, \dots, p_n)(w_i)}{F(p_1, \dots, p_n)(w_j)} = m_{ij} \left(p_1(w_i)/p_1(w_j), \dots, p_n(w_i)/p_n(w_j) \right)$$

Gilardoni prouve que (IR), (PU) et la neutralité des labels (uniformité de la méthode pour tous les événements de l'algèbre), caractérisent, moyennant certaines conditions de régularité supplémentaires, les fonctions de la classe MGP sur les ratios. La leçon à tirer de ce résultat, qui reste étonnamment peu cité dans la littérature récente, nous semble importante.

Contrairement à ce que pourrait laisser penser une lecture trop rapide des résultats classiques de caractérisation de Wagner, la condition d'Indépendance ne fait pas choisir la classe MAP contre la classe MGP. C'est uniquement sur le fond d'un certain choix de représentation (valeurs absolues des probabilités) que l'indépendance caractérise la classe MAP. Un autre choix de représentation (valeurs relatives des probabilités) fait obtenir la classe MGP à partir de la condition d'indépendance correspondante, (IR) – augmentée de quelques hypothèses auxiliaires. Autrement dit, si l'on fait l'hypothèse d'indépendance, faire une moyenne arithmétique est le choix qui s'impose si l'on raisonne en termes de valeurs absolues des probabilités tandis que faire une moyenne géométrique est le choix qui s'impose si l'on raisonne en termes de valeurs relatives.

1.4.3 Agrégation probabiliste et agrégation des jugements

Pour terminer cette discussion, nous voudrions revenir sur les liens qui unissent les travaux axiomatiques précédents et la théorie de l'agrégation logique, qui a transposé aux jugements catégoriques la riche tradition de la théorie du choix social.

Il est courant de distinguer un concept catégorique de croyance et un concept gradué. Le concept catégorique est celui qu'on utilise quand on attribue à quelqu'un la « simple » croyance que p . Le concept de croyance gradué, celui qui est retenu par l'épistémologie bayésienne, est celui qu'on utilise quand on dit de quelqu'un qu'il croit fortement ou faiblement que p , ou quand on dit qu'il croit que p plus qu'il ne croit que q .

L'agrégation des croyances catégoriques a fait l'objet de nombreuses recherches depuis une dizaine d'années et constitue un domaine que l'on appelle la *théorie de l'agrégation des jugements* ou *théorie de l'agrégation logique*. La représentation des croyances catégoriques s'appuie généralement sur les langages formels élaborés par la logique mathématique. Les agents sont censés s'exprimer sur un ensemble de propositions représentées par un sous-ensemble (l'*agenda*) de l'ensemble de formules du langage formel de référence. L'ensemble des croyances (catégoriques) d'un agent i est alors représenté par l'ensemble des formules exprimant les propositions qu'il croit (l'*ensemble de jugements* de i), qui est un sous-ensemble de l'*agenda*. Souvent, on fait l'hypothèse que les ensembles de jugement sont (i) logiquement cohérents et (ii) complets (au sens où pour toute paire formée d'une formule et

de sa négation, ils contiennent au moins l'une des deux formules). La fonction d'agrégation prend la forme d'une fonction qui associe un ensemble de jugements (les jugements collectifs) à un profil d'ensembles de jugements (les jugements individuels).

Dans ce cadre, les théoriciens de l'agrégation des jugements ont montré que, lorsque l'on considère des conditions axiomatiques attractives sur la fonction d'agrégation, on se heurte rapidement à des théorèmes d'impossibilité analogues à ceux, bien connus, que l'on trouve en théorie du choix social concernant l'agrégation des préférences. Ces conditions axiomatiques sont pour la plupart des transpositions de celles que l'on a formulées pour l'agrégation des probabilités (Préservation de l'Unanimité³⁵, Indépendance³⁶, Neutralité³⁷, Anonymat³⁸ ou, plus faiblement, Absence de dictature³⁹), mais elles conduisent à des résultats fort différents. Essentiellement les mêmes propriétés qui étaient conciliables dans le cadre probabiliste deviennent inconciliables dans le cadre catégorique. Plus précisément, l'un des théorèmes d'impossibilité centraux pour l'agrégation des jugements catégoriques, qui dit qu'il n'existe pas de fonction d'agrégation respectant les conditions d'Unanimité, d'Indépendance et d'Anonymat peut être vu comme un corollaire du théorème 2⁴⁰. En effet, les croyances catégoriques peuvent être vues comme le cas particulier de degrés de croyance égaux à 0 ou à 1. Les conditions d'Unanimité et d'Indépendance imposent alors que la fonction d'agrégation soit une moyenne arithmétique pondérée de 0 et de 1, tandis que l'Anonymat impose que les poids soient équitablement partagés entre les agents, et, partant, que la moyenne arithmétique pondérée soit une moyenne arithmétique simple. Mais cette moyenne arithmétique simple ne sera pas en général à valeur dans $\{0,1\}$, rendant l'agrégation impossible.

Le contraste entre l'impossibilité d'agrégérer les jugements catégoriques et, sous les mêmes hypothèses, la possibilité d'agrégérer les degrés de croyance révèle la difficulté qu'il y a à passer des uns aux autres. En effet, s'il existait une procédure respectant les conditions sous lesquelles on dérive l'impossibilité qui permettrait de définir les croyances catégoriques à partir des degrés de croyance, il serait possible d'agrégérer les croyances catégoriques en appliquant cette procédure à la suite d'une agrégation probabiliste. Voyons cela d'un peu plus près, en considérant un exemple d'une telle procédure. Selon la thèse de Locke⁴¹, la croyance catégorique se réduit à un degré de croyance suffisamment fort. Ainsi, on définirait la croyance catégorique que p par la croyance que p au degré au moins r pour un certain r

³⁵ Selon la Préservation de l'unanimité, les formules qui sont acceptées par tous les membres du groupe sont acceptées par le groupe lui-même.

³⁶ Selon la condition d'Indépendance, l'attitude du groupe à l'égard d'une formule dépend uniquement de l'attitude de ses membres à l'égard de cette formule - la manière dont l'attitude des membres détermine l'attitude du groupe pouvant changer d'une formule à l'autre.

³⁷ Qu'on parle plutôt de « Neutralité » que de « Systématicité » dans le contexte de la théorie de l'agrégation logique. Selon la condition de Systématicité, l'attitude du groupe à l'égard d'une formule dépend uniquement de l'attitude de ses membres à l'égard de cette formule, et la manière dont l'attitude des membres détermine l'attitude du groupe est la même quelle que soit la formule.

³⁸ La fonction d'agrégation doit traiter tous les individus à la même enseigne.

³⁹ La fonction d'agrégation ne sélectionne pas purement et simplement un l'ensemble de jugements de l'un des membres du groupe.

⁴⁰ Voir Xavier Gilard, « Comment parvient-on à un consensus ? », mémoire de master, Université Paris Ouest Nanterre, août 2013.

⁴¹ Voir Richard Foley, « Belief, Degrees of Belief, and the Lockean Thesis », in F. Huber & C. Schmidt-Petri (eds.), *Degrees of Belief*, Dordrecht: Springer, 2009.

suffisamment grand compris dans l'intervalle $[1/2 ; 1]$. En combinant moyenne arithmétique simple et thèse de Locke, on obtient alors un mécanisme d'agrégation des croyances catégoriques qui est équivalent à une procédure d'agrégation à la majorité qualifiée, le seuil r correspondant au quota exigé. Un tel mécanisme ne permet bien sûr pas d'obtenir des jugements complets : il se peut que ni p ni $non-p$ ne soient soutenus par une majorité qualifiée. Mais l'on pourrait espérer qu'il garantisse la cohérence. Il n'en est rien : on peut montrer facilement que pour toute règle de majorité qualifiée, il existe un agenda pour lequel la cohérence échoue pour certains profils (Dietrich et List⁴²). On peut voir cet échec comme un nouvel échec pour l'agrégation des jugements catégoriques (même en l'absence de réquisit de complétude), ou, comme l'agrégation probabiliste est quant à elle toujours possible, comme un échec pour la transformation des jugements probabilistes en jugements catégoriques. De fait, les scénarii utilisés pour montrer la possibilité d'incohérence pour les règles de majorité qualifiée sont similaires aux scénarii que l'on retrouve dans le paradoxe de la loterie ou le paradoxe de la préface et qui visent à mettre en échec la thèse de Locke. Le lien entre les théorèmes d'impossibilité et ces paradoxes a été étudié de manière détaillée par Douven et Romeijn⁴³ ainsi que par List et Dietrich⁴⁴.

Ce contraste entre résultats d'impossibilité pour des jugements discrets et résultats de possibilité pour des variables continues fait écho à celui que l'on trouve en théorie du choix social où les résultats d'impossibilité laissent place à des résultats de possibilité quand on passe de l'agrégation des *préférences* à l'agrégation des *fonctions d'utilité*⁴⁵. Le parallèle a toutefois ses limites, parce que le rapport entre les ensembles de jugements et les probabilités subjectives est bien différent de celui entre les préférences et les fonctions d'utilité. Si l'on voulait un parallèle plus exact, il faudrait comparer l'agrégation des distributions de probabilités non pas à l'agrégation des jugements, mais à l'agrégation des *probabilités comparatives*. Une probabilité comparative est une relation \geq définie sur l'espace des événements et dont la signification est la suivante : $E \geq E'$ si l'événement E est au moins aussi probable que l'événement E' . Une telle relation est une probabilité comparative si elle peut être représentée par une distribution de probabilité. Lainé, LeBreton et Trannoy⁴⁶ ainsi que Weymark⁴⁷ (1997) se sont précisément intéressés à l'agrégation des probabilités comparatives. Ils montrent que, pour un espace d'états fini et l'algèbre des parties de cet espace, des conditions de Préservation de l'Unanimité⁴⁸ et d'Indépendance conduisent en effet à des issues dictatoriales ou quasi-dictatoriales.

⁴² Franz Dietrich et Christian List, « Judgment Aggregation by Quota Rules », *Journal of Theoretical Politics*, vol. 19 (4), 2007, p. 391-424.

⁴³ Igor Douven et Jan-Willem Romeijn, « The Discursive Dilemma as a Lottery Paradox », *Economics and Philosophy*, 23 (3), 2007, pp 301-319.

⁴⁴ Franz Dietrich et Christian List, « From Degrees of Beliefs to Beliefs: Lessons from Judgment-Aggregation Theory », *manuscript*, 2014.

⁴⁵ Voir Amartya Sen, *Collective Choice and Social Welfare*, San Francisco: Holden-Day, 1970.

⁴⁶ Jean Lainé, Michel LeBreton et Alain Trannoy « Group Decision Making Under Uncertainty. A Note on the Aggregation of 'Ordinal Probabilities' », *Theory and Decision*, 21, pp. 155-161

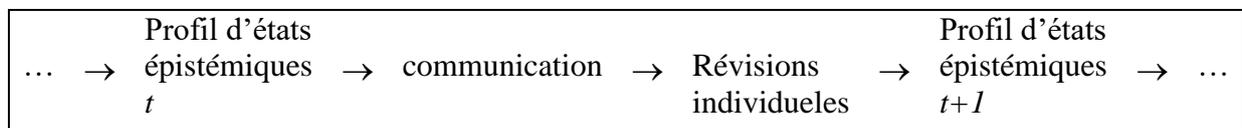
⁴⁷ John Weymark, « Aggregating Ordinal Probabilities on Finite Sets », *Journal of Economic Theory*, 75, 1997, p. 407-432.

⁴⁸ Une première condition dite de Pareto Faible, affirme que si tous les agents jugent E strictement plus probable que E' , le groupe doit juger de même. Elle conduit à une quasi-dictature. Si on ajoute une condition de

En théorie de l'agrégation des jugements, les résultats de possibilité et d'impossibilité dépendent fortement des propriétés de l'agenda. L'enquête théorique porte donc non seulement sur les propriétés de la fonction d'agrégation, mais aussi sur celles de l'agenda et sur leurs interactions. Par exemple, certains des résultats fondamentaux caractérisent les conditions sur l'agenda qui rendent incompatibles un certain nombre d'axiomes sur la fonction d'agrégation. Récemment, Dietrich et List⁴⁹ ont exporté cette manière de procéder en étudiant l'agrégation de degrés de croyances définis sur des agendas qui ne sont plus nécessairement des σ -algèbres. Ils prouvent alors des généralisations de la Proposition 1 et du Théorème 2 en variant corrélativement les propriétés structurelles de l'agenda et la condition de Préservation de l'Unanimité de la fonction d'agrégation.

2 Théories de la révision sociale

Dans cette section, nous nous intéressons non plus à une situation où les membres du groupe transmettent leurs croyances à un agrégateur en charge de déterminer une opinion collective, mais à une situation où il y a interaction épistémique entre eux, i.e. où les agents se communiquent leurs opinions. La séquence élémentaire d'un tel processus de *révision sociale* se laisse concevoir comme la transition d'un profil d'états épistémiques t à un profil d'états épistémiques $t+1$:



Elaborer un modèle de révision sociale consiste, pour l'essentiel, à spécifier chacun des ingrédients de cette séquence élémentaire. Tous les modèles qui nous concernent incluent une représentation probabiliste des degrés de croyances dans la représentation de l'état épistémique des agents. En ce sens, ils sont tous bayésiens. Mais dans un modèle *purement* bayésien, (i) les états épistémiques individuels se réduisent à des distributions de probabilité, (ii) les messages communiqués prennent la forme d'événements (au sens technique) et (iii) les révisions sont effectuées par conditionnalisation sur ces événements. Nous présenterons un tel modèle ultérieurement. Nous allons commencer par un modèle qui s'écarte significativement de ce modèle purement bayésien, le modèle DLW, qui réutilise, à d'autres fins, les moyennes pondérées présentées dans la section précédente.

Pareto Indifférence selon laquelle si tous les agents jugent E et E' aussi probables, alors le groupe doit juger de même, alors on obtient une dictature.

⁴⁹ Franz Dietrich et Christian List, « Probabilistic opinion pooling generalized », manuscript, 2013.

2.1 Le modèle DLW

Le modèle le plus influent de révision sociale probabiliste est le modèle DLW, qui a été proposé indépendamment par De Groot⁵⁰ et Lehrer & Wagner⁵¹. Dans la version de Lehrer et Wagner, les agents se communiquent les uns aux autres leurs degrés de croyances concernant un événement spécifique E. Dans la version plus générale de De Groot, ils se communiquent la totalité de leurs distributions de probabilité.

Les deux ingrédients fondamentaux du modèle DLW sont les suivants. D'une part, chaque agent i révisé sa probabilité initiale P_i conformément à la règle de la moyenne pondérée, en attribuant à chaque agent k (y compris lui-même) un poids w_{ik} (les poids sommant à 1). On note W la matrice $n \times n$ des poids épistémiques que les agents s'attribuent les uns aux autres – notons qu'on suppose pour l'instant que ces poids sont donnés une fois pour toutes et valent à toutes les étapes du processus de révision. D'autre part, les agents communiquent et révisent leurs probabilités de manière itérée : à l'étape 1, ils se communiquent et prennent en compte leurs probabilités initiales. À l'étape 2, ils se communiquent et prennent en compte les probabilités obtenues à l'étape 1, etc. Le caractère itératif du modèle DLW le rapproche des méthodes empiriques du type Delphi. Ces méthodes, développée à partir des années 1950, visent à faire converger les prédictions d'un groupe d'experts à l'aide d'un processus itératif analogue : les experts échangent leurs prédictions sur une question donnée, commentent les résultats obtenus et sont ensuite invités à réviser leurs prédictions, le processus ayant vocation à être répété un petit nombre de fois, dans l'espoir de produire un consensus.

Revenons à la formalisation proposée par le modèle DLW. Le processus part de la matrice colonne P^0 qui représente les probabilités initiales des agents, et à chaque étape, la matrice colonne P^k est remplacée par la matrice colonne $P^{k+1} = W \times P^k = W^{k+1} \times P^0$.

Probabilités à l'étape k
P_1^k
...
P_n^k

Poids épistémiques attribués à 1, ..., n		
w_{11}	...	w_{1n}
...
w_{n1}	...	w_{nn}

 \Rightarrow

Probabilités à l'étape $k+1$
$P_1^{k+1} = \sum_{i=1}^n w_{1i} \times P_i^k$
...
$P_n^{k+1} = \sum_{i=1}^n w_{ni} \times P_i^k$

⁵⁰ Morris H. DeGroot « Reaching a consensus », *Journal of the American Statistical Association*, 69(345), pp. 118-121.

⁵¹ Keith Lehrer et Carl Wagner, *Consensus in Science and Society*, Dordrecht, Reidel, 1981. En accord avec la littérature qui a suivi, nous présentons le modèle DLW comme un modèle de révision sociale et non d'agrégation. Notons que l'interprétation qu'en donnent Lehrer et Wagner du modèle DLW est plus ambiguë. Ils affirment en effet que « dans notre modèle normatif, le processus d'agrégation doit être vu comme une fiction heuristique [...] ce modèle en tant que modèle de rationalité est un modèle synchronique plutôt que diachronique. Il nous dit ce que l'information totale contenue dans un groupe à l'état initial ajoute à une étape spécifique de l'acquisition de l'information. » (p. 26).

Le résultat fondamental obtenu dans ce cadre est que, sous certaines conditions assez peu restrictives, les opinions des membres du groupe vont converger. Les conditions en question mettent en jeu le concept de **chaîne de respect épistémique**. On dit qu'il y a une chaîne de respect épistémique entre l'agent i et l'agent j s'il existe une suite d'agents dont le premier est i , le dernier est j , et telle que chaque agent attribue à l'agent suivant un poids épistémique non-nul.

Théorème 4⁵²

S'il existe un agent qui se respecte et s'il existe une chaîne de respect épistémique depuis chaque autre membre du groupe vers lui, alors les opinions des membres du groupe convergent.

L'une des caractéristiques remarquables de ce Théorème est que l'atteinte d'un consensus ne dépend que des poids épistémiques que les individus s'attribuent les uns aux autres : si ces poids ont la propriété appropriée, alors *quelles que soient les probabilités individuelles initiales* et donc même si ces probabilités sont très différentes, le modèle converge vers des probabilités consensuelles.

Pour discuter le modèle DLW, il faut bien distinguer ses deux ingrédients : (a) une règle de révision qui régit la manière dont les agents changent leurs croyances à chaque étape, et (b) l'itération de ce processus avec des poids épistémiques constants. Nous abordons tour à tour chacun de ces ingrédients.

(a) La règle de révision retenue est la moyenne arithmétique pondérée, qui est appliquée à chaque étape par chaque membre du groupe.

L'une des propriétés essentielles de la moyenne pondérée envisagée comme règle de révision est l'**Intermédierité** : quand le groupe se réduit à deux agents et que ces agents s'accordent un poids épistémique non-nul, alors la croyance révisée en E de l'un prend une valeur intermédiaire entre sa croyance initiale et la croyance initiale de l'autre. Cette propriété est directement liée au débat en cours, en épistémologie sociale, sur le désaccord entre « pairs épistémiques » (voir notamment les travaux de Kelly⁵³, Christensen⁵⁴ et Elga⁵⁵). La question qui est au cœur de ce débat est celle de savoir quelle attitude doit être adoptée par un agent qui apprend qu'un autre agent, qu'il tient pour aussi compétent que lui pour juger si E , n'entretient pas les mêmes croyances vis-à-vis de E . L'Intermédierité correspond à l'une des deux principales réponses que l'on peut donner à cette question, celle selon laquelle, dans les situations de ce genre, un agent rationnel fera « un pas vers l'autre ». De manière plus générale, le modèle DLW soulève la question de la *justification* de la MAP comme règle de

⁵²Ce théorème est appelé « probability consensus theorem » dans Keith Lehrer, « When Rational Disagreement is Impossible », *Noûs*, 10, 1976, p. 327-332. Dans l'ouvrage de Lehrer & Wagner cité dans la note précédente, il s'agit du théorème 7.2.

⁵³ Thomas Kelly, « The epistemic significance of disagreement », in John Hawthorne et Tamar Gendler (eds.), *Oxford Studies in Epistemology*, vol 1, Oxford : Oxford University Press, 2005, p. 167-196.

⁵⁴ David Christensen, « Epistemology of Disagreement: The Good News », *Philosophical Review*, 116(2), 2007, p. 187-217.

⁵⁵ David M. Elga, « Reflection and Disagreement », *Noûs*, 41 (3), 2007, p. 478-502.

révision⁵⁶. Dans leur ouvrage de 1981, Lehrer et Wagner considèrent que la caractérisation axiomatique de la moyenne pondérée, présentée dans la première partie, fournit une réponse adéquate à cette question. La discussion se ramène alors essentiellement aux raisons d'adopter la condition d'Indépendance (I). La position de Lehrer et Wagner est que « la charge de la preuve appartient à ceux qui proposent de ne pas [...] adopter [la condition d'Indépendance] » (1981, p. 112). Si l'argument peut sembler un peu court, il ne s'agit pas cependant d'une simple pétition de principe. Lehrer et Wagner interprètent (I) comme une condition d'invariance « relativement faible », stipulant que la détermination pour un agent de la nouvelle probabilité qu'il accorde à E ne doit pas varier quand varient les probabilités accordées par les autres agents aux autres événements. L'agent doit pouvoir les ignorer, en l'absence d'information spécifique quant à la manière dont elles devraient l'influencer dans sa révision pour E. Cependant, nous avons vu dans la première partie que cette idée d'indépendance à l'égard des autres événements n'était pas le monopole de la condition (I), puisqu'elle pouvait être capturée, dans un autre cadre de représentation, par la condition (IR), conduisant à une autre classe de fonctions de révision.

(b) L'autre ingrédient du modèle DLW est l'itération de la révision. Celle-ci soulève deux questions : pourquoi, tout d'abord, les membres du groupe devraient-ils continuer à réviser après la première révision ? Pourquoi, ensuite, ces révisions devraient-elles toujours se faire avec les mêmes poids épistémiques ?

La première question est à l'origine d'une famille de critiques du modèle DLW, qui lui reprochent de procéder à un « double comptage » des opinions⁵⁷ : si chaque agent a tenu compte de l'opinion de tous les autres agents (à l'étape 1), pourquoi devrait-il continuer à le faire ultérieurement ? On peut vouloir répondre que, à chaque nouvelle étape, les agents apprennent bien quelque chose : ils apprennent la manière dont les autres agents tiennent compte de l'opinion des autres membres du groupe. Par exemple, quand l'agent i apprend que l'agent j a révisé son opinion P_j^0 en P_j^1 à partir des mêmes informations que lui, il a une information (incomplète) sur les poids épistémiques attribués par j aux autres agents. Si l'on suit cette interprétation de l'itération, alors *la signification des poids épistémiques change d'une étape à l'autre de l'itération* : à l'étape 1, les poids épistémiques reflètent la confiance qu'ont les agents dans leurs opinions sur l'état de la nature (attitude d'ordre 1). À l'étape 2, en revanche, ils reflètent la confiance qu'ils ont dans leurs jugements révisés sur l'état de la nature, et par conséquent leur confiance dans la manière dont les membres du groupe évaluent les opinions des autres sur l'état de la nature (attitude d'ordre 2). Et l'on peut ainsi monter dans la hiérarchie des croyances. De manière générale, les poids épistémiques à l'étape $k + 1$ reflètent la confiance qu'ont les membres du groupe dans la manière dont ils se sont évalués à l'étape k (attitude d'ordre $k + 1$).

⁵⁶ L'hypothèse normative selon laquelle c'est effectivement la règle appropriée est par exemple explicitement soutenue par Lehrer dans son article de 1976 (p. 328).

⁵⁷ Voir Lehrer et Wagner (1981), p. 127.

A la lumière de cette analyse de l'itération, l'hypothèse selon laquelle les poids épistémiques sont constants d'une étape à l'autre semble excessivement restrictive⁵⁸ : après tout, i peut avoir confiance en j pour juger si E est le cas, mais pas pour juger si les membres du groupe jugent correctement si E est le cas. Il est donc naturel d'envisager une généralisation du modèle DLW, appelons-la « DLW+ »⁵⁹, dans laquelle les poids épistémiques sont autorisés à changer d'une étape à l'autre. Parmi les éléments primitifs du modèle, au lieu d'avoir la matrice W , on a désormais une suite infinie de matrices W^1, W^2, \dots . Les conditions du théorème précédent ne sont plus satisfaites, mais on peut tout de même formuler des conditions qui garantissent la convergence vers un consensus. Notons β^k le plus petit nombre de la matrice W^k . Le résultat s'énonce alors comme suit :

Théorème 5 (Lehrer & Wagner, 1981)

Si la suite des sommes $\beta^1, \beta^1 + \beta^2, \dots$ diverge vers $+\infty$, alors les opinions des membres du groupe convergent.

Les Théorèmes 4 et 5 utilisent les fonctions d'agrégation de la classe MAP, qui ont les faveurs de Lehrer et Wagner. Cependant, intuitivement, si la propriété qui amène le consensus est l'Intermédiarité, on devrait pouvoir obtenir des résultats analogues pour d'autres classes, pourvu qu'elles amènent également les agents à « faire un pas les uns vers les autres ». Gilardoni et Clayton⁶⁰ ont montré que les conditions de convergence pour les fonctions de la classe MGP étaient essentiellement les mêmes que pour la classe MAP considérée par Lehrer et Wagner : la moyenne géométrique conduit au consensus aussi sûrement que la moyenne arithmétique. Mais ce n'est pas le cas de toutes les moyennes, puisque Gilardoni et Clayton montrent également, résultat étonnant, que le fait de converger dans ces conditions caractérise la classe MAPUMGP.

2.2 Conditionnalisation et probabilités d'ordre supérieur

Quand on l'envisage comme un modèle de révision sociale, le modèle DLW s'écarte sensiblement de l'épistémologie bayésienne. Plus précisément, il introduit une distinction forte à l'intérieur de la *hiérarchie* des croyances d'un individu, concept qu'il nous faut désormais approfondir. Jusqu'ici, nous avons implicitement considéré que les individus avaient des croyances probabilistes sur la « nature ». L'espace d'états sous-jacent S est donc un espace des états de la nature. Et ces croyances sont des **croyances sur la nature** ou croyances d'ordre 1. Les individus peuvent aussi entretenir des croyances sur les croyances que les autres ont sur la nature, des croyances sur les croyances que les autres ont sur les croyances que lui ou les autres ont sur la nature, etc. On appellera ces croyances des

⁵⁸ Lehrer, dans son article de 1976, justifie cependant l'hypothèse de constance des poids par le fait que, durant le processus de révision, les individus n'apprennent pas de nouvelles informations et par conséquent n'ont pas de raison de remettre en question leur confiance initiale dans les membres du groupe. Cet argument est problématique à plusieurs égards. D'abord, il prête le flanc à l'objection du « double comptage ». Ensuite, on peut justement considérer que durant le processus d'itération, les membres du groupe apprennent quelque chose – ils ont des informations sur l'évolution des croyances des autres membres du groupe.

⁵⁹ Lehrer et Wagner l'appellent le « modèle étendu ».

⁶⁰ Gustavo Gilardoni et Murray K. Clayton, « On Reaching a Consensus using DeGroot's iterative pooling », *The Annals of Statistics*, 21 (1), 1993, p. 391-401.

croyances d'ordre supérieur. Notons qu'à chaque type de croyance correspond un type d'information : un individu peut apprendre quelque chose sur la nature mais aussi sur les croyances des autres à propos de la nature, ou sur leurs croyances à propos des croyances que lui ou les autres ont sur la nature, etc.

Considérons maintenant le modèle DLW du point de vue de la distinction entre ces deux types de croyance. Dans le modèle DLW, ce sont les croyances sur la nature qui font l'objet d'une représentation probabiliste. Les agents se présentent donc comme des agents bayésiens pour les croyances d'ordre 1. Mais ils apprennent des informations d'ordre supérieur (d'ordre 2, pour être précis). Et ces informations d'ordre supérieur ne sont pas traitées de manière bayésienne. D'une part, les agents n'ont pas de croyances initiales probabilisées d'ordre supérieur (sauf à interpréter les poids comme des probabilités, interprétation dont a vu les difficultés précédemment). Et, d'autre part, après avoir appris ces informations d'ordre supérieur, ils ne révisent pas leurs croyances par conditionnalisation. Les croyances sur la nature et les croyances d'ordre supérieur obéissent donc à un régime épistémique différent. Le traitement des secondes s'écarte sensiblement des hypothèses fondamentales de l'épistémologie bayésienne.

En quoi consisterait une approche purement bayésienne de la révision sociale ? Fondamentalement, à considérer que *le témoignage d'autrui est un événement au sein d'une algèbre à propos de laquelle les agents entretiennent des croyances probabilistes*. Quand il apprend un tel événement, l'agent révisera alors ses croyances par conditionnalisation, comme il le fait face aux événements 'habituels'. Notons $\lceil p_j(E) = \alpha \rceil$ l'événement selon lequel l'agent j croit au degré α que E est le cas. L'approche bayésienne revient à supposer que quand l'agent i apprend cette information (d'ordre supérieur), sa croyance en E prend la valeur $p_i(E \mid \lceil p_j(E) = \alpha \rceil)$.

Il existe deux principales manières de représenter les croyances probabilistes d'ordre supérieur. La première consiste à construire une hiérarchie explicite en partant d'un ensemble S d'états de la nature, en ajoutant ensuite, pour chaque agent, l'espace de ses croyances probabilistes possibles sur S , et en continuant la construction⁶¹. L'état épistémique d'un agent (on parle parfois de « type » dans ce contexte) est alors identifié à une suite de distributions de probabilités d'ordres croissants. La seconde option élabore la hiérarchie des croyances de manière implicite : on se donne immédiatement un ensemble Ω d'états du monde et, pour chaque agent i , une fonction $\pi_i(\cdot)$ qui à chaque état du monde $\omega \in \Omega$ fait correspondre une distribution de probabilités sur Ω . Dans ce cas, on peut définir l'événement « j croit au degré α que E » comme

$$\lceil p_j(E) = \alpha \rceil = \{ \omega \in \Omega : \pi_j(\omega)(E) = \alpha \}$$

Supposons qu'à chaque état du monde $\omega \in \Omega$ on fasse correspondre un état de la nature $s \in S$. On peut alors vérifier que chaque état du monde induit de surcroît une hiérarchie (infinie) de

⁶¹ La définition inductive exacte est complexe, notamment parce qu'elle met en jeu une contrainte de cohérence entre les niveaux. Pour une présentation rigoureuse et pédagogique, voir Michael Maschler, Eilon Zamir et Shmuel Solan, *Game Theory*, Cambridge, Cambridge University Press, 2013, Définition 11.6.

croyances probabilistes d'ordre supérieur, pour chaque agent. Un état du monde peut donc se concevoir comme une description complète de la réalité, qui inclut l'état de la nature et les croyances de tous les ordres des agents. Dans la suite, nous parlerons d'un **système de croyances probabilistes** pour désigner un tel modèle de croyances probabilistes d'ordre supérieur. (Un système de croyances probabilistes est l'analogie de la sémantique de Kripke multi-agent pour la logique épistémique.) C'est sur cette seconde manière de modéliser les croyances probabilistes d'ordre supérieur que nous allons nous appuyer pour présenter une famille de résultats initiés par un article fondateur d'Aumann en 1976⁶².

Considérons deux individus et faisons l'hypothèse de l'*a priori* commun (*common prior assumption*, également appelée « doctrine d'Harsanyi »). En chaque état du monde, les croyances d'un individu i sont obtenues par conditionnalisation à partir d'une information privée et d'une distribution de probabilité commune à tous les membres du groupe notée $p : \Omega \rightarrow [0, 1]$. Plus précisément, on suppose, pour chaque individu i , une *partition informationnelle* $I_i : \Omega \rightarrow \wp(\Omega)$. I_i représente l'information que reçoit i en chaque état du monde possible. L'hypothèse de l'*a priori* commun se formule donc ainsi : pour chaque individu i et événement E ,

$$\pi_i(\omega)(E) = p(E \mid I_i(\omega))$$

L'hypothèse de l'*a priori* commun est tout sauf triviale ; il est aisé de construire des systèmes de croyances probabilistes qui la violent. Si elle vaut, alors le système de croyances probabilistes correspondant se laisse représenter comme un triplet $(\Omega, p, (I_i)_{i \in G})$. Les partitions informationnelles permettent de définir un concept de **connaissance** : en un état $\omega \in \Omega$, un individu i sait que E est le cas si, et seulement si, dans tous les états compatibles avec ses informations, E est le cas. Autrement dit,

$$K_i E = \{\omega \in \Omega : I_i(\omega) \subseteq E\}$$

A partir du concept de connaissance, on peut définir celui de **connaissance commune** : il est de connaissance commune entre i et j que E si tous les deux savent que E , tous les deux savent que l'autre sait que E , tous les deux savent que tous les deux savent que l'autre sait que E , etc. Définissons I^* comme la borne inférieure des deux partitions I_1 et I_2 , c'est-à-dire comme la partition la plus fine parmi les partitions qui sont plus grossières que I_1 et I_2 . Alors l'opérateur de connaissance commune est défini ainsi :

$$K^* E = \{\omega \in \Omega : I^*(\omega) \subseteq E\}$$

On peut maintenant énoncer le résultat d'Aumann (1976) :

Théorème 6

⁶² Robert Aumann, « Agreeing to Disagree », *The Annals of Statistics*, 4 (6), 1976, p. 1236-1239.

Soit $(\Omega, p, (I_1, I_2))$ un système de croyances probabilistes qui satisfait l'hypothèse de l'*a priori* commun⁶³. Si, en un état du monde $\omega \in \Omega$, il est de connaissance commune que l'agent 1 croit E au degré α_1 et que l'agent 2 croit E au degré α_2 , alors $\alpha_1 = \alpha_2$.

Le théorème d'Aumann affirme en substance que, sous l'hypothèse de l'*a priori* commun, s'il n'y a pas consensus entre les agents concernant E , il n'y a pas de transparence complète entre eux à propos de leurs opinions respectives concernant E , puisqu'elles ne peuvent pas être de connaissance commune. Pour le dire encore autrement : quand les différences de probabilités entre deux agents ne *peuvent* venir que du fait qu'ils ont reçu des informations différentes, alors la connaissance commune de ces probabilités empêche qu'elles soient différentes. Dans les commentaires qu'il fait de son résultat, Aumann mentionne le modèle DLW et va jusqu'à affirmer qu'on peut considérer son modèle comme un « fondement théorique pour la réconciliation de probabilités subjectives ». La proximité se fera sans doute mieux sentir si l'on considère la *version dynamique* du résultat d'Aumann démontrée par Geanakoplos & Polemarchakis en 1982⁶⁴. On suppose maintenant que les agents suivent un protocole de communication et de révision⁶⁵ particulier, qui ressemble à certains égards à celui de DLW : à chaque étape, l'un des deux agents annonce son degré de croyance en un événement fixé E . L'autre agent révisé ensuite sa probabilité pour E par conditionnalisation et annonce sa probabilité révisée.

Théorème 7 (Geanakoplos & Polemarchakis, 1982)

Soit $(\Omega, p, (I_1, I_2))$ un système de croyances probabilistes qui satisfait l'hypothèse de l'*a priori* commun. Si les agents communiquent et révisent leurs croyances à propos d'un événement E conformément au protocole décrit ci-dessus, alors, en un nombre fini d'étapes, ils aboutissent à un consensus concernant E .

Afin de mieux comprendre le contenu de ce Théorème, illustrons-le d'un exemple.

Exemple 2 (Geanakoplos & Polemarchakis, 1982)

$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9\}$, $p(\omega) = 1/9$ pour tout $\omega \in \Omega$

$I_1 = \{\{\omega_1, \omega_2, \omega_3\}, \{\omega_4, \omega_5, \omega_6\}, \{\omega_7, \omega_8, \omega_9\}\}$

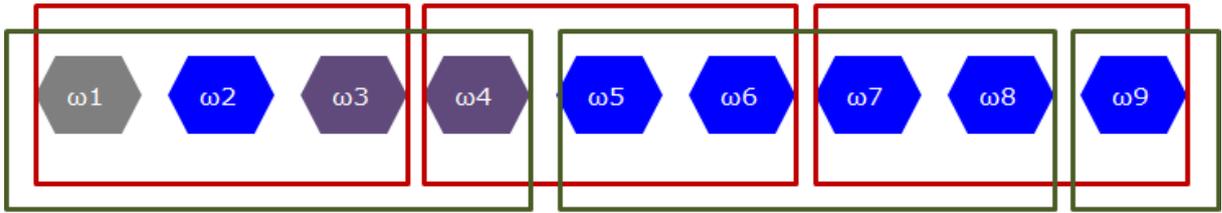
$I_2 = \{\{\omega_1, \omega_2, \omega_3, \omega_4\}, \{\omega_5, \omega_6, \omega_7, \omega_8\}, \{\omega_9\}\}$

Supposons que l'état réel soit ω_1 et $E = \{\omega_3, \omega_4\}$.

⁶³ On suppose que Ω est fini et que la probabilité *a priori* $p(\cdot)$ accorde une probabilité non-nulle à tout événement de la partition $I_1 \vee I_2$.

⁶⁴ John Geanakoplos et Hercule Polemarchakis, « We Can't Disagree Forever », *Journal of Economic Theory*, 28 (1), 1982, p. 192-200

⁶⁵ Ce que Bacharach appelle un « dialogue », Michael Bacharach, « Some extensions of a claim of Aumann in an axiomatic model of knowledge », *Journal of Economic Theory*, 37(1), 1985, p. 167-190.



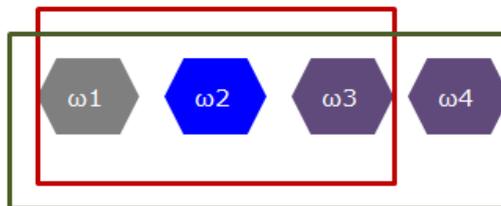
Dans ce cas, l'agent 1 apprend que l'état réel appartient à l'ensemble $I_1(\omega_1) = \{\omega_1, \omega_2, \omega_3\}$ et a pour probabilité initiale $p_1^0(E) = p(E \mid I_1(\omega_1)) = 1/3$. De son côté, l'agent 2 apprend que l'état réel appartient à l'ensemble $I_2(\omega_1) = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ et a pour probabilité initiale $p_2^0(E) = p(E \mid I_2(\omega_1)) = 1/2$. Les opinions des deux agents divergent donc.



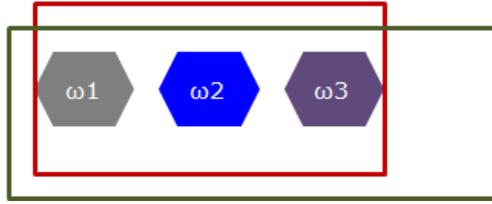
Supposons maintenant que l'agent 1 informe l'agent 2 de son degré de croyance. Celui-ci apprend alors qu'il est dans un des états du monde $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6$ où l'agent 1 croit E au degré $1/3$. Mais l'agent 2 a déjà eu l'information, plus forte, selon laquelle $I_2(\omega_1)$ est le cas. Donc il ne change pas sa croyance initiale.



Quand l'agent 2 annonce à son tour son degré de croyance $p_2^1(E) = 1/2$, l'agent 1 est informé de l'événement $\{\omega_1, \omega_2, \omega_3, \omega_4\}$ et ne change pas non plus sa croyance en E .



Mais quand il annonce à l'agent 2 que $p_1^1(E) = 1/3$, il lui apprend quelque chose de nouveau : si le monde réel était ω_4 , les croyances *révisées* de l'agent 1 seraient en effet $p_1^1(E) = 1$ (que l'on obtient en conditionnalisant $p(\cdot)$ sur $I_1(\omega_4)$ puis sur $\lceil p_2^1(E) = 1/2 \rceil$). Donc le monde réel n'est pas ω_4 .



Par conditionnalisation, on obtient $p_2^2(E) = 1/3$. Les croyances des deux agents ont donc convergé vers un consensus.

La version dynamique du théorème d'Aumann exhibe un processus de révision sociale, purement bayésien, au terme duquel les agents aboutissent au consensus et à la connaissance commune de ce consensus – cette dernière conséquence est explicitée par Bacharach (1985, Prop. 12), qui généralise les deux théorèmes précédents. Si l'on se concentre sur la version dynamique, il apparaît que, dans le cas typique, chaque agent accroît son information initiale sur l'état du monde à chaque fois que l'autre agent annonce son degré de croyance en E . Mais il est important de souligner que le processus ne conduit pas nécessairement à un partage intégral de l'information privée I_i que chaque agent i a reçue avant la phase de communication. Et cela peut d'ailleurs induire un manque de *robustesse* du consensus obtenu : il se peut que celui-ci ne soit pas celui que les agents atteindraient s'ils partageaient directement leurs informations initiales respectives⁶⁶.

L'hypothèse de l'*a priori* commun joue bien sûr un rôle fondamental dans le résultat d'Aumann (1976). Aumann y voit d'ailleurs une manière de tester l'hypothèse – un test qui délivrerait un verdict *défavorable* si l'on accepte l'idée qu'il est fréquent que deux individus soient en désaccord alors même que leurs opinions respectives sont de connaissance commune. L'hypothèse implique que toute différence dans les croyances *a posteriori* provient d'une différence d'information puisque, si elle vaut, deux agents qui reçoivent les mêmes informations ont les mêmes croyances probabilistes. Que devient le résultat d'Aumann si l'on renonce à l'hypothèse de l'*a priori* commun ? Comme on peut s'y attendre, sans cette hypothèse, la connaissance commune des croyances *a posteriori* n'implique plus de consensus entre les agents. En revanche, elle implique toujours que, *concernant l'événement E en question*, les différences de probabilités subjectives induites par les différences d'informations privées disparaissent⁶⁷. Plus précisément, pour chacun des deux agents i ,

$$p_i(E \mid I_i(\omega)) = p_i(E \mid I^*(\omega))$$

On retrouve bien sûr le résultat d'Aumann dans le cas particulier de l'hypothèse de l'*a priori* commun (i.e. où $p_1 = p_2$).

Les deux résultats que nous venons d'évoquer considèrent le cas où $|G| = 2$. Envisageons maintenant le cas plus général d'un groupe fini d'agents. Quand tous les agents communiquent leurs croyances probabilistes à tous les autres, la version dynamique du

⁶⁶ John Geanakoplos, « Common Knowledge » in R. Aumann et S. Hart, *Handbook of Game Theory*, vol. 2, chapitre 40, Amsterdam, Elsevier, 1994, p. 1450.

⁶⁷ *ibid.* p. 1454.

théorème d'Aumann (le Théorème 7) vaut toujours en substance⁶⁸. Plus précisément, le protocole de communication en question selon lequel, à chaque étape, chaque agent communique sa croyance révisée publiquement, aboutit à la connaissance commune des croyances des agents et au consensus si l'hypothèse de l'*a priori* commun vaut. Mais quand on considère un groupe plus large, il est naturel d'envisager d'autres familles de protocoles de communication. Tous ne conduisent évidemment pas à la connaissance commune des croyances probabilistes, ni au consensus. Plusieurs contributions ont étudié l'impact de différents protocoles de communication sur ces deux issues possibles⁶⁹. Parikh & Krasucki (1990)⁷⁰ ont par exemple étudié les protocoles de communication « équitables » (*fair*), c'est-à-dire les protocoles par lesquels, pour le dire rapidement, tout agent communique, directement ou pas, avec tout autre agent, un nombre indéfini de fois. C'est le cas des protocoles de type « tourniquet » (*round robin*) où les agents sont placés en cercle et où chaque agent communique son degré de croyance exclusivement à son voisin de droite⁷¹. Parikh & Krasucki montrent que de tels protocoles conduisent eux aussi au consensus quand l'hypothèse de l'*a priori* commun vaut, tandis que la question de savoir si, et en quel sens, il y a également connaissance commune est discutée par Heifetz (1996), qui envisage l'existence d'incertitude sur le protocole⁷². La demande que les protocoles soient équitables est analogue à la condition sur les chaînes de respect dans les Théorèmes 4 et 5 de Lehrer et Wagner : elle garantit que les révélations de n'importe quel agent aient un impact sur les croyances de tout autre agent.

2.3 Discussion des modèles de révision sociale

La littérature sur la révision sociale a donc élaboré deux familles distinctes de modèles, sur la base desquelles on peut par exemple explorer la question de savoir dans quelles conditions un groupe d'agents communiquant entre eux est susceptible d'atteindre un consensus. Cette dualité appelle naturellement une comparaison entre les deux types de modèles, et en particulier une analyse de leur compatibilité et de leurs justifications respectives. Une première chose qu'il convient de souligner est qu'il est aisé de concevoir des situations où la règle de moyenne arithmétique pondérée (MAP) entre en conflit avec nos intuitions.

Exemple 3

Les agents i et j ont des probabilités initiales p_i et p_j sur un espace Ω où chaque état de la nature rend vraies ou fausses les propositions A et B :

⁶⁸ Le théorème d'Aumann lui-même vaut pour $|G| \geq 2$. Pour la généralisation à $|G| > 2$ dans le cas dynamique, voir Jonathan Cave « Learning to Agree », *Economics Letters*, 12, 1983, pp. 147-152.

⁶⁹ Cette littérature s'est d'ailleurs rapidement abstraite du problème original de la communication des croyances probabilistes : elle a considéré, de manière plus générale, la communication de la valeur de fonctions f dont l'argument est un événement I . Pour un événement E , la fonction qui associe à chaque information possible I la probabilité de E conditionnelle à I est une fonction de ce genre.

⁷⁰ Rohit Parikh et Paul Krasucki, « Communication, Consensus and Knowledge », *Journal of Economic Theory*, 52(1), 1990, p. 178-189.

⁷¹ Pour d'autres illustrations, voir Paul Krasucki, « Protocols Forcing Consensus », *Journal of Economic Theory*, 70 (1), 1996, p. 266-272 ou Frédéric Koessler, « Common Knowledge and consensus with noisy communication », *Mathematical Social Science*, 42 (2), 2001, p. 139 – 159.

⁷² Avaidia Heifetz, « Comment on consensus without common knowledge », *Journal of Economic Theory*, 70, 1996, p. 273-277.

	AB	$A \neg B$	$\neg AB$	$\neg A \neg B$
p_i	0,9	0,1	0	0
p_j	0,1	0	0,9	0

Supposons que l'agent i croit que l'agent j sait si B est le cas, et que j communique à i sa probabilité initiale à l'égard de AB – soit $p_j(AB) = 0,1$. Etant donné les croyances d'ordre supérieur de i , i va en inférer que j croit que B est le cas – soit $p_j(B) = 1$. Puisque i se fie à j concernant la valeur de vérité de B , i va éliminer les états possibles où B n'est pas le cas. Il serait donc naturel que i adopte comme probabilité révisée :

	AB	$A \neg B$	$\neg AB$	$\neg A \neg B$
p_i	1	0	0	0

Cet épisode viole la propriété d'Intermédierité et ne peut donc pas être reconstruit comme l'application d'une fonction de révision appartenant aux classes MAP ou MGP. Notons d'ailleurs que si l'on suppose les mêmes croyances d'ordre supérieur pour i mais les croyances d'ordre 1 suivantes :

	AB	$A \neg B$	$\neg AB$	$\neg A \neg B$
p_i	0,9	0,05	0,05	0
p_j	0,1	0	0,9	0

alors il deviendra naturel d'endosser comme probabilité révisée (à la suite du même message de j) :

	AB	$A \neg B$	$\neg AB$	$\neg A \neg B$
p_i	90/95	0	5/95	0

On est alors face à une violation non plus seulement de l'Intermédierité, mais de la propriété de Neutralité (N) et même de la propriété d'Indépendance (I), si l'on suppose que les croyances révisées de i sont engendrées par application d'une fonction d'agrégation probabiliste aux probabilités initiales de i et j . En variant encore une fois légèrement les probabilités initiales, on peut également violer l'Unanimité (U), puisque les probabilités suivantes :

	AB	$A \neg B$	$\neg AB$	$\neg A \neg B$
p_i	0.9	0.1	0	0
p_j	0.9	0	0.1	0

aboutissent, toujours par le même raisonnement, à cette probabilité révisée :

	AB	$A \neg B$	$\neg AB$	$\neg A \neg B$
p_i	1	0	0	0

Alors que i et j sont initialement d'accord sur le degré de croyance à accorder à la proposition AB , ce consensus se perd quand i apprend que j accorde à AB le même degré de croyance que lui. Ces remarques n'ont rien d'étonnant ni de paradoxal : les règles de moyenne pondérée, qui sont initialement des règles d'agrégation, ont une capacité limitée à tenir compte des croyances d'ordre supérieur des agents, alors que celles-ci jouent évidemment un rôle crucial dans toute révision sociale. En revanche, elles ont l'avantage d'être *frugales* en ce sens qu'elles ne requièrent que peu d'informations. Par contraste, l'approche purement bayésienne, qui permet de rendre compte des exemples qui précèdent, suppose ou exige que i détermine la probabilité conditionnelle $p_i(E \mid [p_j(E) = \alpha])$ et ait donc une opinion (probabiliste) sur la manière dont l'événement E et le témoignage de j sur E sont liés. C'est précisément ce que nous avons supposé dans les exemples en faisant l'hypothèse que i pensait que j savait si B était le cas ou non.

Quelle leçon en tirer quant aux rôles respectifs du modèle bayésien et du modèle DLW ? On peut les voir comme des modèles *concurrents*, destinés à modéliser les mêmes situations de mise en commun des opinions individuelles, avec chacun leurs avantages et leurs inconvénients. Dans une situation où les croyances de niveau supérieur sont complexes et jouent un rôle important, et où l'on peut disposer d'informations détaillées sur les croyances conditionnelles des agents, le modèle bayésien sera le plus adéquat. Dans une situation où les croyances de niveau supérieur sont plus grossières et peuvent se laisser résumer à l'attribution de poids différents aux opinions des agents, lorsque les croyances conditionnelles des agents ne sont pas accessibles ou lorsqu'elles ne sont pas déterminées, le modèle DLW pourra s'imposer. Si l'on considère que le modèle Bayésien reste le modèle de référence en termes de normes de rationalité, et que le modèle DLW vaut surtout par sa frugalité, il faut se demander dans quelle mesure le modèle DLW est compatible avec le modèle Bayésien, c'est-à-dire dans quelle mesure une application du modèle DLW donnant un certain résultat pourrait être « rationalisée » comme le résultat d'une application du modèle Bayésien donnant le même résultat, moyennant des hypothèses supplémentaires sur les croyances conditionnelles des individus. Dans une situation à deux agents, une telle rationalisation est possible exactement lorsque les probabilités initiales de l'agent sont identiques à son espérance concernant les probabilités initiales de l'autre agent⁷³.

On peut également considérer que ces deux modèles sont *complémentaires* et n'ont pas vocation à s'appliquer dans les mêmes contextes informationnels. Le cadre d'Aumann, par exemple, suppose typiquement que les différents individus disposent d'informations privées qui sont différentes : la prise en compte des opinions individuelles des autres agents est alors un moyen indirect d'exploiter des informations auxquelles on n'a pas eu accès directement. A l'inverse, le modèle DLW s'apparente à une négociation, où chaque agent doit en quelque sorte faire la part entre la qualité de son opinion et la qualité de l'opinion des autres. Il peut alors s'appliquer à des situations d'« équilibre dialectique »⁷⁴ dans lesquelles un désaccord

⁷³ Denis Bonnay et Mikaël Cozic, « Weighted Averaging and Bayesian Conditionnalization », manuscript, 2015.

⁷⁴ Cf. par exemple Keith Lehrer, « Rationality as Weighted Averaging », 57, 1983, p. 290.

subsiste quand bien même les membres du groupe ont déjà échangé toutes leurs informations. Dans ce cas, la révision sociale n'est pas un moyen détourné d'intégrer des informations privées, mais plutôt une négociation quant à la manière dont les informations mises en commun doivent être interprétées. L'importance du contexte informationnel, et en particulier l'importance de la distinction entre les contextes où les agents ont des informations asymétriques et ceux où ils n'en ont pas, est défendue par Dietrich (2010)⁷⁵ et List et Dietrich (2014)⁷⁶. Il faut néanmoins préciser que Dietrich s'intéresse prioritairement à la question de l'agrégation plutôt qu'à la révision sociale. C'est pourquoi il applique cette distinction pour clarifier les domaines légitimes d'application de différentes règles d'agrégation, plutôt que, comme nous venons de le faire, pour clarifier ceux du modèle DLW et du modèle bayésien orthodoxe. En l'occurrence, il conclut qu'une famille de règles d'agrégation que nous n'avons pas examinées, les règles *multiplicatives* sont celles qui devraient s'appliquer en cas d'asymétrie informationnelle. Elles ne respectent ni la Préservation de l'unanimité, ni la Bayésianité externe, ce qui est bienvenu dans ce contexte d'après Dietrich.

3 Agréger et réviser : quand et pour quoi faire

Dans les précédentes Sections, nous avons passé en revue les deux grandes familles de théories qui prennent en charge la « vie sociale » d'agents bayésiens. La présente Section sera de nature plus interprétative. Nous allons revenir sur la distinction entre agrégation et révision sociale. D'une part, pour en approfondir chacun des termes ; et d'autre part, pour discuter des relations qui peuvent ou doivent exister entre les deux concepts.

3.1 Variété de l'agrégation

Nous avons caractérisé l'agrégation des croyances des membres d'un groupe comme un processus par lequel une croyance collective ou agrégée était formée à partir des croyances des membres de ce groupe. La simplicité de cette caractérisation est peut-être trompeuse. Si les théoriciens de l'agrégation prennent rarement le temps d'en dire plus sur ce en quoi consiste précisément l'agrégation, il n'est pas exclu que des phénomènes sensiblement hétérogènes soient visés par un seul et même formalisme. Il vaut donc la peine de se demander ce que sont les croyances agrégées obtenues à l'issue de ce processus et quels types de buts peuvent être poursuivis par un agrégateur. On peut se convaincre de l'intérêt potentiel de la question en considérant la théorie de l'agrégation des préférences. Amartya Sen a défendu à de nombreuses reprises⁷⁷ l'idée que les relations binaires individuelles définies sur l'ensemble des états sociaux possibles pouvaient être interprétées au moins de deux manières distinctes : comme représentant les jugements individuels comparatifs sur le sort ou l'« intérêt » du groupe dans les différents états sociaux (dans ce cas $x \prec_i y$ signifie que i juge que l'état x est collectivement meilleur que l'état y), ou comme représentant des

⁷⁵ Franz Dietrich, « Bayesian Group Belief », *Social Choice and Welfare*, 35(4), 2010, pp. 595-626.

⁷⁶ *Op.cit.*

⁷⁷ Voir notamment Amartya Sen, « Social Choice Theory: A Re-Examination », *Econometrica*; 45(1), 1977, p. 53-89.

comparaisons de bien-être individuel (dans ce cas $x \prec_i y$ signifie que l'état x est supérieur à l'état y du point de vue du bien-être de i). Cela conduit à deux interprétations distinctes de l'agrégation, comme *synthèse* des opinions sur le bien collectif, ou comme *arbitrage* entre le bien-être de différents individus, destiné à déterminer le bien collectif. Dans le cas de l'agrégation des croyances, les variations interprétatives ne sont pas les mêmes. Il n'y a pas en effet d'ambiguïté concernant la nature des *entrées* du mécanisme d'agrégation : il s'agit des degrés de croyances des membres du groupe. En revanche, on peut concevoir de manières différentes le *résultat* du processus – l'état épistémique « collectif » ou « agrégé » – et corrélativement le but du processus d'agrégation.

Un premier objectif que l'on peut poursuivre est de fournir un *résumé* du profil des croyances individuelles (p_1, \dots, p_n). Le résultat de l'agrégation est alors une sorte de statistique de ses entrées, si bien qu'il est naturel de parler d'une conception *statistique* de l'agrégation, au sens des statistiques descriptives. Selon cette perspective, l'agrégation a pour but de fournir une représentation concise des croyances des membres du groupe. Les moyennes sont les règles d'agrégation qui correspondent naturellement à cette interprétation, puisqu'une des fonctions des moyennes est précisément de fournir *une* valeur pour synthétiser un ensemble de valeurs. Notons cependant que les moyennes ne sont pas les seules opérations à remplir cette fonction – mode, médiane, écart-type, par exemple, concourent au même but, et un statisticien considèrera généralement que c'est la juxtaposition de ces valeurs qui permet de bien résumer l'ensemble de données considéré. Inversement, les moyennes peuvent avoir d'autres fonctions que celle que leur assignent les statistiques descriptives : on peut voir la moyenne, par exemple, comme le résultat d'une négociation entre les différents individus. Dans un article récent⁷⁸, Christian List propose de distinguer trois sortes de croyances collectives (et, plus généralement, d'attitudes collectives). L'interprétation que nous discutons correspond à ce qu'il appelle une attitude « agrégée » (*aggregate attitude*). Les exemples qu'il donne, le prix d'une action sur un marché de prédiction ou l'opinion politique d'une catégorie de population, relèvent bien de ce que nous appelons la conception statistique. Le prix de marché est une manière d'agréger les croyances des agents et reflète un certain équilibre entre les croyances individuelles. Les opinions politiques d'une catégorie de la population seront elle typiquement résumées à l'aide d'un pourcentage (70% des électeurs de droite pensent que ...). Comme List (*ibid.*) le soutient en substance, sous cette interprétation statistique de l'agrégation, il n'est pas clair qu'il faille exiger que le résultat satisfasse aux conditions de rationalité imposées aux attitudes individuelles – en l'occurrence, exiger qu'il constitue une distribution de probabilité – et qu'il y ait un sens dans lequel il s'agisse véritablement de *croyances* collectives.

Par contraste, l'agrégation peut avoir pour objectif de délivrer un authentique état épistémique qui puisse être attribué au groupe en tant que tel. Il ne s'agit pas alors de décrire de manière succincte l'ensemble des opinions individuelles, mais plutôt de décrire de manière complète l'opinion collective du groupe. On peut parler, dans ce cas, d'agrégation *attributive*. C'est typiquement le genre d'agrégation en jeu lorsqu'on s'intéresse à l'opinion d'un groupe constitué, comme un parti, une entreprise ou un syndicat. Sous cette interprétation,

⁷⁸ Christian List, « Three kinds of collective attitudes », *Erkenntnis*, 79 (9), 2014, p. 1601-1622.

l'agrégateur est le groupe lui-même ou, pour reprendre l'expression de Hogarth⁷⁹, la « personnalité synthétique du groupe ». L'agrégation attributive correspond à ce que List (*ibid.*) appelle une attitude de groupe (*corporate attitude*), qu'il définit comme une attitude possédée par un groupe qui a le statut d'agent intentionnel (une entité capable d'avoir des croyances, des désirs et d'agir sur la base de ces croyances et de ces désirs). List considère qu'il est constitutif de telles attitudes qu'elles satisfont à des contraintes de rationalité similaires à celles imposées aux attitudes individuelles, sans quoi ce ne serait pas d'authentiques attitudes. Nous acceptons ce point, en soulignant néanmoins le fait que les modèles auxquels nous nous intéressons sont très idéalisés par rapport aux conditions de rationalité, plus faibles, qui sont réellement exigées par l'attribution d'attitudes⁸⁰. Si l'on fait abstraction de cette considération, les modèles d'agrégation des probabilités auxquels nous nous sommes intéressés semblent particulièrement appropriés pour traiter de l'agrégation attributive, puisqu'ils imposent la même structure aux états épistémiques individuels et à l'état résultant de l'agrégation.

On pourrait vouloir distinguer cette deuxième conception (« attributive ») de l'agrégation d'une troisième, qui a en commun avec elle de viser à la formation d'authentiques états épistémiques⁸¹. La différence est qu'il ne s'agit plus de former un état épistémique susceptible d'être attribué *au groupe G*, mais, pour un tiers, de s'appuyer sur les états épistémiques des membres de *G*, en les traitant comme des informations à partir desquelles former son propre état épistémique. Rien n'empêche d'ailleurs que ce tiers soit un membre de *G*. Cette différence avec l'interprétation attributive semble en particulier se montrer dans la manière dont la fonction d'agrégation tient compte des caractéristiques épistémiques des différents membres de *G*. En effet, quand il s'agit de former l'état épistémique de *G*, on s'attend spontanément à ce que le poids accordé à chaque membre dépende de la place qu'il occupe au sein du groupe – dans un groupe 'égalitaire', par exemple, tous les membres seront traités de la même manière, une condition qu'on appelle traditionnellement l'Anonymat en théorie de l'agrégation des préférences. Alors que, quand un tiers s'appuie sur les opinions des membres de *G*, on s'attend à ce qu'il tienne compte de ce qu'il sait de leur hétérogénéité épistémique, et que les poids varient en fonction de l'expertise que lui-même attribue aux différents membres. L'objectif de l'agrégateur est en effet d'exploiter au mieux les opinions dont il est informé, et il se peut que cela passe par une attention différenciée à celles des uns et des autres : on peut parler d'agrégation *consultative*, au sens où il s'agit d'obtenir les meilleures croyances possibles sur la base des croyances de chacun des membres du groupe, comme on cherchera à le faire après avoir consulté les avis d'un panel d'experts.

⁷⁹ Robert M. Hogarth, « Methods for Aggregating Opinions », in H. Jungermann et G. De Zeeuw, *Decision Making and Change in Human Affairs*, Springer, 1975, p. 231-255.

⁸⁰ Sur cette question, voir Denis Bonnay et Mikaël Cozic, « Principe et charité et sciences de l'homme », in Thierry Martin (ed.) *Les sciences humaines sont-elles des sciences ?*, Paris, Vuibert, 2011, p. 119-157.

⁸¹ On trouve par exemple une trace de cette distinction dans l'article de Christian Genest & James Zidek, « Combining Probability Distributions: A Critique and an Annotated Bibliography », *Statistical Science*, vol.1(1), 1986, p. 114-148: « ...we reviewed a number of normative methods for amalgamating opinions which could be used either when the decision maker is a third person who is consulting a group of experts...or when the third person is simply the 'synthetic personality of the group' » (nous soulignons).

Afin de clarifier et de tester cette typologie tripartite, voyons les différentes dimensions qui permettent de distinguer chacun des types d'agrégation retenus. Un premier facteur concerne la nature du groupe dont les croyances sont agrégées. On peut distinguer groupes inorganisés et groupe organisés⁸². Un groupe inorganisé est une simple collection d'individu qu'on a décidé de considérer (pris) ensemble, mais qui ne se considèrent pas eux-mêmes comme un groupe. Les Français de plus de trente ans, les acheteurs d'Iphone sont de tels groupes inorganisés. Un groupe organisé est un groupe auquel ses membres ont conscience d'appartenir et auquel ils se rapportent en tant que tel : le parti socialiste, le conseil d'Etat, le gouvernement sont des groupes de ce genre. L'agrégation statistique et l'agrégation consultative peuvent opérer sur un groupe organisé comme sur un groupe inorganisé, en revanche l'agrégation attributive semble devoir opérer sur un groupe organisé – ou en tout cas, semble opérer de manière privilégiée sur de tels groupes, si l'on veut laisser ouverte la possibilité de groupes qui soient des agents intentionnels sans que leurs membres aient besoin pour autant de se rapporter à ce groupe en tant que tel.

Comment les opinions individuelles sont-elles prises en compte par la fonction d'agrégation ? Doivent-elles toutes être traitées sur un pied d'égalité, ou sont-elles hiérarchisées ? Si elles sont hiérarchisées, qu'est-ce qui détermine leur place ? La réponse à ces questions constitue un autre facteur déterminant pour le processus d'agrégation. Pour l'agrégation statistique, il semble que toutes les opinions se valent, dans la mesure où il s'agit de résumer l'ensemble de ces opinions. Dans le cas de l'agrégation attributive, les opinions individuelles comptent dans la mesure où elles comptent pour déterminer l'opinion du groupe. Ainsi, comme discuté dans la première section, pour peu que l'avis des différents membres du groupe n'ait pas la même importance s'agissant des questions considérées, il sera naturel d'accorder des poids différents aux opinions individuelles, possiblement en utilisant une moyenne pondérée plutôt qu'une moyenne simple. L'opinion du groupe ne survient pas uniquement sur la base des opinions individuelles correspondantes⁸³ (List, *ibid.*). Entrent également en jeu, par exemple, les relations de déférence entre les membres du groupe, telles que manifestées par l'attitude d'un membre qui s'en remet aux opinions d'un autre sur un certain sujet. A la limite, on peut imaginer des cas où l'opinion du groupe ne dépend pas des opinions individuelles, même pondérées, mais d'autres circonstances, par exemple des opinions passées du groupe (ainsi si le groupe est commis à une certaine continuité d'opinion à travers le temps) ou d'autres intérêts (si le groupe doit adopter des opinions en conformité avec sa raison d'être, par exemple). Apparaissent ici les limites des formalismes présentés s'agissant de l'agrégation attributive. L'étendue et la variabilité des déterminants de l'opinion du groupe rendent hautement douteuse la possibilité de la présenter comme fonction des opinions individuelles, fussent-elles pondérées ou amalgamées selon une règle spécifique. Le cas de l'attribution consultative est encore différent. Il s'agit cette fois d'utiliser au mieux les opinions individuelles pour atteindre, sinon la vérité, du moins les meilleures conjectures possibles. Les opinions individuelles doivent alors être prises en compte à la mesure de leur

⁸² Denis Bonnay, « A Clustering-Based Approach to Collective Beliefs », à paraître dans Thomas Boyer-Kassem, Conor Mayo-Wilson et Michael Weisberg (éds.), *Scientific Collaboration and Collective Knowledge*, Oxford University Press, New York.

⁸³ List (*ibid.*)

qualité : on leur accordera d'autant plus de poids qu'elles proviennent d'individus mieux informés, ou mieux à même d'exploiter les informations qui sont les leurs.

Un dernier facteur concerne le rôle de l'agrégateur : s'agit-il d'une simple fiction théorique, l'agrégation faisant sens indépendamment d'un agent l'opérant, ou le processus d'agrégation est-il au contraire commandé par l'agrégateur ? La visée de l'agrégation statistique est essentiellement pragmatique : on peut concevoir que la description à fournir dépende des intérêts ou des buts de l'agrégateur – dans certains cas, il sera pertinent de faire la moyenne, alors que dans d'autres cas, la moyenne n'aura pas d'intérêt pour l'agrégateur. L'agrégation attributive, au contraire, ne laisse pas de place à l'agrégateur. Si l'on admet que l'opinion de groupe existe indépendamment de celui qui l'observe, on dira qu'il ne s'agit pas pour l'agrégateur de faire naître une opinion collective, mais plutôt de l'attribuer correctement. S'agissant de l'opinion consultative, l'agrégateur apparaît comme le juge des compétences. C'est à lui qu'il appartient de décider de l'importance à accorder aux opinions individuelles selon leur valeur épistémique supposée. Pour cela, l'agrégateur peut s'en remettre aux opinions des membres du groupe quant à leur fiabilité respective, mais l'évaluation n'en restera pas moins *in fine* la sienne.

On peut résumer la caractérisation ainsi obtenue des trois types d'agrégation dans le tableau suivant :

	Agrégation statistique	Agrégation attributive	Agrégation consultative
Visée	descriptive	reconnaissance ⁸⁴	épistémique
Nature du groupe	organisé ou non	organisé	organisé ou non
Mode de prise en compte des opinions individuelles	égalitaire	qualifiée (importance)	qualifiée (fiabilité)
Prise en compte d'autres éléments que les opinions individuelles	non	oui	non ⁸⁵
Rôle de l'agrégateur	pragmatique ⁸⁶	pas de véritable rôle	juge de la fiabilité

⁸⁴ On entend par 'visée reconnaissance' le fait qu'il s'agit pour l'agrégateur de reconnaître ce qu'est l'opinion du groupe, de lui attribuer l'opinion qui est la sienne.

⁸⁵ Ce 'non' doit être doublement précisé. D'une part, lorsque l'agrégateur pondère les opinions individuelles en jugeant la compétence des experts, il prend en compte un autre élément, la compétence de l'expert, mais cet élément n'est pas indépendant des opinions individuelles, on peut le voir comme une propriété de celles-ci. En revanche, dans le cas de l'agrégation consultative, la structure du groupe, par exemple, est un élément indépendant des opinions à agréger. D'autre part, l'agrégateur peut être amené à prendre en compte les opinions des individus sur la qualité des opinions des autres individus (voir la section suivante).

⁸⁶ Selon les buts de l'agrégateur, telle ou telle manière de résumer sera plus pertinente.

L'agrégation statistique est typiquement le fait d'un sondeur qui vise à décrire de manière synthétique les opinions au sein d'un groupe. Peu importe que ce groupe soit ou non organisé, seules les opinions individuelles sont prises en compte, et elles le sont de manière égalitaire. Le rôle de l'agrégateur est purement pragmatique, et porte sur la meilleure manière de synthétiser les opinions individuelles. L'agrégation attributive est typiquement le fait d'un observateur qui vise à reconnaître les attitudes qui sont celles d'un groupe organisé donné, par exemple un journaliste qui attribue des opinions à une entreprise ou à un gouvernement. Les opinions individuelles sur les questions considérées comptent, mais d'autres éléments sont à prendre en compte, comme l'histoire du groupe, et elles sont qualifiées selon le rôle au sein du groupe des individus qui les portent. L'agrégateur n'a pas de rôle substantiel, il ne s'agit que de rendre au groupe ce qui lui appartient. L'agrégation consultative est typiquement le fait d'un décideur s'appuyant sur un panel d'experts, elle a une visée épistémique et peut s'appliquer que le groupe soit organisé ou non. Seules les opinions individuelles sont prises en compte mais elles sont qualifiées en fonction de l'expertise de ceux qui les tiennent. Même s'il peut déléguer la chose, l'agrégateur est en dernière instance le juge de cette fiabilité.

Les distinctions conceptuelles que nous venons d'esquisser peuvent avoir un impact sur la plausibilité des conditions étudiées par l'approche axiomatique. Nous avons déjà évoqué celle selon laquelle le résultat de l'agrégation est lui-même une distribution de probabilités. Cet axiome vaut certainement pour les agrégations attributives et consultatives, sans doute pas pour une agrégation statistique. Autre exemple : l'axiome d'indépendance (I). (I) vaut le plus clairement pour l'agrégation statistique : la synthèse des opinions sur un événement donné ne semble dépendre, par définition, que des opinions sur cet événement. Pour l'agrégation consultative, c'est la justification par l'ignorance qui est pertinente, mais elle ne vaut pas à tout coup : l'agrégateur peut avoir des raisons de nature épistémique pour calculer les valeurs agrégées en faisant dépendre le verdict sur un événement des jugements sur d'autres événements. Pour l'agrégation attributive, on ne peut *a priori* s'attendre à ce que (I) soit satisfaite, mais les raisons pragmatiques présentées plus haut (sous-section 1.4.2.) d'accepter (I) sont pertinentes, afin par exemple de garantir que l'opinion du groupe ne puisse pas être manipulée par ses membres.

3.2 Agréger ou réviser ?

3.2.1 La contrainte par horizon de consensus

Comme le remarque Bradley (*op. cit.*), une situation d'agrégation est toujours en même temps une situation possible de révision, puisque les membres du groupe pourraient chacun décider de réviser leurs opinions individuelles à la lumière des opinions de celles des autres. Comment faut-il voir les relations entre agrégation et révision ? Les révisions que les individus sont susceptibles d'effectuer contraignent-elles la manière dont un agrégateur est susceptible de les agréger ? Bradley répond positivement à cette dernière question, en considérant le cas particulier où les révisions individuelles conduisent à un consensus, comme elles le feront si les individus révisent par moyenne arithmétique pondérée, dans les conditions décrites dans le Théorème 4. Il soutient que, si les processus individuels de révision devaient amener à un consensus, c'est ce consensus qui devrait *dès le départ*

constituer l'opinion agrégée des individus⁸⁷. Appelons cette thèse la thèse de la *contrainte par horizon de consensus* – le jugement collectif agrégé est contraint par les consensus susceptibles de se dégager à la suite d'opérations de révision.

De fait, il peut sembler tentant d'adopter pour valeur agrégée la valeur potentiellement consensuelle, lorsqu'elle existe. On résout alors les problèmes d'agrégation en les ramenant à des problèmes de révision. Mais qu'est-ce qui justifie de céder à cette tentation ? La justification envisagée par Bradley est épistémique : si l'agrégation ignore les révisions possibles, elle ignore des informations susceptibles d'améliorer le jugement collectif, en l'occurrence les informations que les individus possèdent sur la qualité des informations des autres individus, ou sur leur habileté à interpréter les informations qu'ils possèdent. Cela suggère que la thèse de la contrainte par horizon de consensus, si elle vaut, vaut pour l'agrégation consultative, qui vise à former la meilleure opinion possible sur la base de ce que savent les individus.

Dans le cas de l'agrégation statistique ou attributive, par contraste, il ne semble pas y avoir de bonne raison d'accepter la thèse de la contrainte par horizon de consensus. L'agrégation statistique vise à décrire de manière synthétique un ensemble d'opinions individuelles *à un moment donné* : ce qui pourrait arriver à ces opinions si les individus venaient à en changer d'une manière ou d'une autre ne semble pas un genre d'informations pertinent. L'agrégation attributive, de même, vise à correctement reconnaître ce que pense un groupe constitué. L'horizon de consensus concerne ce que ce groupe pourrait être amené à penser, mais le fait est que, si ce n'est pas là ce que pense le groupe, on ne voit pas pourquoi l'agrégateur devrait se fonder sur cet horizon plutôt que sur les opinions actuelles des membres du groupe.

Le cas de l'agrégation consultative est différent, puisque les opinions présentes des membres du groupe n'ont pas le même privilège : tous les moyens sont bons pour que l'agrégateur parvienne aux meilleurs jugements possibles. L'argument de Bradley repose sur le fait que les individus du groupe n'ont pas seulement des jugements de premier niveau sur les questions considérées, mais ont aussi des jugements de deuxième niveau (et de niveaux supérieurs) sur la qualité des jugements de premier niveau des autres membres. Si l'agrégateur n'utilise que les jugements de premier niveau, il décide de ne pas tenir de compte de jugements non dénués de valeur. Néanmoins, il ne suit pas pour autant que l'on doive accepter le principe de contrainte par horizon de consensus pour l'agrégation consultative. En effet, celui-ci suppose que l'agrégateur s'aligne sur la manière dont les individus du groupe évaluent leurs compétences respectives. Ce n'est pas évident : l'agrégateur peut avoir des raisons de ne pas faire confiance aux jugements d'ordre supérieur des individus, ou de les qualifier – par exemple, en se méfiant d'un individu qui accorde systématiquement beaucoup plus d'importance à son jugement qu'aux jugements d'autrui. S'il est naturel pour l'agrégateur de prendre en compte les jugements d'ordre supérieur des individus, lorsqu'il y a

⁸⁷ « supposons par exemple que notre théorie de la révision nous dise que les jugements individuels améliorés [après révision] seront d'accord quant à la valeur d'une issue particulière. Alors il serait naturel d'exiger que le jugement agrégé sur cette issue ne devrait être rien d'autre que le jugement consensuel auquel tout le monde finira par arriver en révisant son jugement à la lumière des jugements rapportés par les autres » (*ibid.*, p. 147).

accès, il est tout aussi naturel de considérer qu'il les prendra en compte en les qualifiant par la confiance qu'il leur accorde, tout comme il le fera pour les jugements de premier niveau. Il n'y a alors pas de raison de penser que le résultat auquel parviendra l'agrégateur sera identique au consensus qu'atteindrait le groupe sur la base de la façon dont chacun est disposé à prendre en compte les avis des autres.

Notons cependant qu'on peut envisager une interprétation différente de la contrainte par horizon de consensus, en voyant celle-ci comme une norme de rationalité pour les opinions de groupe : un groupe rationnel *devrait* adopter comme opinion le consensus auquel il aboutirait si chacun de ses membres révisait leurs croyances en fonction des opinions des autres. Dans ce cas, la contrainte par horizon de consensus vaudrait pour l'agrégation attributive, pour les groupes soumis à une telle norme.

3.2.2 *Agrégation consultative, révision et déférence.*

Le lien entre agrégation consultative et révision peut encore s'envisager d'une autre manière. Dans la mesure où il s'agit pour l'agrégateur de former les meilleurs jugements possibles sur la base des opinions du groupe, cette agrégation peut être vue comme une forme de révision, par laquelle l'agrégateur se forme sa nouvelle opinion sur les questions considérées. Dans une situation de révision standard, on a comme point de départ les opinions initiales de l'agent (ses « *priors* » dans un cadre probabiliste), et l'on cherche comment les mettre à jour pour tenir compte de certaines informations nouvelles. Mais, *si l'on suppose que l'agrégation consultative est correctement représentée par la théorie de l'agrégation conventionnelle*, alors on doit remarquer qu'elle ne postule pas de telles probabilités initiales. L'agrégateur adopte une attitude de *déférence épistémique* à l'égard des opinions des membres du groupe.

Il faut immédiatement préciser cette affirmation. Dans le cas très particulier où l'agrégateur est un membre du groupe, les probabilités initiales de l'agent font bien partie de la situation modélisée par la théorie conventionnelle, puisqu'elles font partie des données concernant le groupe. On peut alors vouloir comparer le résultat de l'agrégation consultative et le résultat d'une révision sociale où l'agrégateur réviserait à partir de ses probabilités initiales (telles que données dans le profil du groupe) et des opinions des autres membres du groupe. Si l'on utilise une même règle sociale dans les deux cas, par exemple une moyenne arithmétique pondérée avec les mêmes poids, les résultats seront les mêmes. Notons qu'il n'est cependant pas complètement évident que les résultats *doivent* être les mêmes. Lorsque je réviserai mes croyances sur la base des croyances d'autrui, on pourrait soutenir qu'il est rationnel pour moi d'accorder un privilège épistémique à mes croyances, en raison par exemple de l'accès privilégié dont je dispose aux raisons qui les justifient. A l'inverse, l'agrégation consultative semble relever d'une perspective en troisième personne : si l'agrégateur est un membre du groupe, il ne doit pas se traiter différemment des autres membres du groupe. Dans le cas général, l'agrégation consultative ne semble pas présupposer que l'agrégateur ait des probabilités initiales ou, s'il en a, qu'elles doivent être prises en compte. C'est une différence importante entre agrégation consultative et révision : l'agrégation consultative produit des croyances probabilistes à partir d'un type particulier d'informations, les croyances des membres du groupe, sans que cette formation s'analyse

comme une mise à jour de croyances initiales. En toute généralité, l'agrégation consultative ne semble donc pas réductible à la révision, même si elles sont comparables dans des cas particuliers.

C'est du moins la conclusion à laquelle on arrive si l'on maintient l'hypothèse selon laquelle le cadre conventionnel représente bien l'agrégation consultative, et si par conséquent on fait de la déférence épistémique une propriété constitutive de l'agrégation consultative. Mais la conclusion sera toute autre si l'on remet en cause cette hypothèse et si l'on s'en tient à la définition par laquelle nous avons introduit l'agrégation consultative : comme un processus par lequel un agent forme ses opinions sur la base des opinions d'autres agents. Alors, le cadre analytique de la théorie de l'agrégation ne s'appliquera à l'agrégation consultative que dans des cas très particuliers, ou par souci de simplicité, parce qu'il s'avère plus complexe de mobiliser un formalisme de révision des croyances. Ces cas seront précisément ceux où l'agrégation consultative prend place dans une situation de déférence épistémique. Ce qui rend la clarification difficile, c'est qu'il y a des espèces différentes de déférence épistémique. (a) L'agrégateur peut, sans se juger lui-même incompetent, mais peut-être par exemple parce que sa fonction l'exige, faire abstraction de ses propres jugements. On sera là dans un cas de *déférence épistémique qui n'est pas épistémiquement motivée*. Mais alors l'agrégation en question sera de l'ordre de l'*agrégation attributive*. (b) Ou alors, l'agrégateur peut se juger totalement incompetent et se fier exclusivement aux jugements des membres du groupe. On sera là dans un cas de *déférence épistémique épistémiquement motivée*. L'agrégation consultative est compatible avec ce genre de situations. Dans ce cas, s'appuyer sur le cadre analytique conventionnel de la théorie de l'agrégation est tout à fait plausible.

4 Conclusion.

Au terme de cette présentation, la vie sociale des bayésiens apparaît riche et variée. Elle mêle des extensions du cadre bayésien habituel et des modélisations spécifiques. Les extensions du cadre bayésien habituel traitent les informations sociales, concernant les croyances des agents, comme des informations « mondaines » habituelles, consistant en observations sur l'état de la nature. Le théorème d'Aumann exposé dans la deuxième section illustre la fécondité de cette approche : la caractérisation bayésienne de la rationalité individuelle permet de dériver des propriétés remarquables des échanges d'information. Les modélisations spécifiques tirent parti de ce que les informations sociales ont déjà une forme probabiliste pour proposer des procédures de transformation de ces informations, comme l'amalgamation par moyenne pondérée. Typiques de cette approche sont les caractérisations axiomatiques des moyennes pondérées exposées dans la première partie : elles permettent de comprendre le choix de telle ou telle procédure comme une conséquence du souci de transformer les opinions des agents en respectant certaines propriétés.

Un des principaux enjeux pour l'amélioration des conditions de vie sociales des bayésiens, ou en tout cas pour une meilleure compréhension de celles-ci, nous semble consister dans une meilleure articulation des analyses conceptuelles relevant de l'ontologie ou de l'épistémologie

sociale, telles qu'esquissées dans la troisième section, et des modèles visant à prendre en charge tel ou tel type de contexte social et informationnel.

Si par exemple l'agrégation statistique est une entreprise pragmatique *relative aux fins poursuivies par l'agrégateur*, cette approche devrait prendre en compte explicitement ces fins et montrer non pas que telle procédure d'agrégation est recommandée parce qu'elle satisfait telles propriétés axiomatiques désirables ou naturelles, mais plutôt que, lorsque telle fin est poursuivie, telle procédure d'agrégation est recommandée parce qu'elle satisfait des propriétés qui permettent d'atteindre le but fixé. Par exemple, un des buts de l'agrégation statistique peut être d'identifier les meilleurs messages à adresser à un ensemble d'individus pour l'amener à un certain état épistémique – c'est le cas lorsqu'un sondage d'opinion est réalisé par un homme politique soucieux d'optimiser son discours de campagne. Dans un tel scénario, les bonnes manières de synthétiser les opinions sont celles qui seront prédictives quant à la manière dont les membres du groupe recevront certains messages. Ou encore, dans le cas de l'agrégation consultative, on s'attend à ce que le choix de telle ou telle procédure d'agrégation dépende de la situation épistémique et varie par exemple selon que les membres du groupe ont des informations différentes ou au contraire partagent les mêmes informations. Comme envisagé dans les travaux récents de Dietrich et List (2014), on cherche alors la meilleure procédure d'agrégation *pour une situation épistémique donnée*.