Mémoire de licence sur l’agrégation du jugement

Noah, Saintonge

26 juin 2020

Résumé

On étudiera ici différents aspects de l’aggrégation du jugement. On s’informera tout d’abord du cadre formel de travail pour ensuite étudier la manière dont on peut orienter une décision en sa faveur, comment s’en protéger et enfin les limites et critiques de cette théorie.

Table des matières

Introduction : Un peu d'histoire de l'aggrégation du Jugement

L'agrégation du jugement est une science qui mêle la théorie du choix social à la logique pour donner une nouvelle approche du sujet.

La théorie du choix sociale est l'étude de la manière dont les préférences et intérêts individuels sont combinés en décisions collectives. L'exemple le plus étudié en théorie du choix social est celui des élections, que nous étudierons tout au long de ce mémoire.

L'étude systématique et formelle du vote a réellement commencée durant la révolution française, grâce à des mathématiciens français comme De Borda, Condorcet et Laplace.

De Borda

De Borda a montré que le scrutin majoritaire pouvait élire le mauvais candidat. En effet chaque individu vote pour un unique candidat et celui qui reçoit le plus de votes est alors élu. Ce type de scrutin ne tient cependant pas compte de l'ordre des préférences des individus.

50	x	y	z
30	y	z	x
30	z	y	x

Observons le tableau précédent, qui représente les préférences de vote des individus lors de l'élection de 3 candidats. On remarque que pour 50 des 110 électeurs x est le préféré, tandis que 30 électeurs préfères y et enfin 30 autres préfèrent z.Selon le principe du scrutin majoritaire x doit être élu, cependant en regardant de plus près on remarque que chez 60 électeurs, x est le pire candidat possible. Le scrutin de De Borda a pour visée de répondre à cette problématique.

Dans la méthode de De Borda, les électeurs rangent les n candidats par ordre de préférence. Chaque candidat à la première place obtient alors n points, chaque candidat à la seconde place obtient n-1 points et ainsi de suite. Cette méthode, bien que meilleure que le scrutin majoritaire est toutefois perméable aux manipulations. Un électeur peut par exemple placer l'opposant le plus dangereux à son propre candidat à la fin dans le but de lui faire perdre le plus de points possibles.

Condorcet

Comme De Borda, Condorcet a remarqué que le scrutin majoritaire est loin d'être la meilleure alternative. Ceci le mena alors à la formulation du critère de Condorcet : le candidat élu est celui qui gagne la majorité de ses "matchs" contre les autres candidats. Contrairement à De Borda, le vainqueur de Condorcet est celui qui bat la majorité des candidats en les comparant par paires. Cependant ce type de scrutin peut amener un groupe ayant des préférences cycliques (c.f tableau ci-dessous) à ne pas pouvoir élire de vainqueur.

Electeur 1	x	y	z
Electeur 2	y	z	x
Electeur 3	z	x	y

Condorcet a alors utilisé les probabilités pour régler ce problème en s'inspirant de celui des jurés dans les tribunaux. Les électeurs sont alors vus comme des jurés qui doivent voter pour la meilleure alternative. Il s'est appuyé sur la théorie de Rousseau selon laquelle le groupe prend des meilleures décisions que l'individu. Il a alors cherché la procédure d'agrégation qui permettrait de maximiser la probabilité du groupe de prendre la bonne décision. Il a alors énoncé le théorème suivant résumé ici dans les grandes lignes :

Sous certaines conditions, le groupe prend des meilleurs décisions que l'individu et la probabilité de voir le groupe prendre la bonne décision s'approche de 1 à mesure que la taille du groupe augmente.

Voilà les deux apports majeurs de Condorcet à la théorie du choix social.

De la théorie du choix social à l'agrégation du jugement

Les modèles de la théorie du choix social classique se concentrent sur des problèmes d'agrégation de préférences en vue d'un résultat collectif comme une élection, une politique donnée ou une action. Cependant ces modèles sont incapables de résoudre des problèmes de décisions dans lesquels des individus ont à se baser sur des décisions logiquement connectées entre-elles, comme lors d'assemblées générales. C'est ce manque que l'agrégation du jugement vise à combler.

Paradoxe Doctrinal

L'agrégation du jugement prend racine dans la jurisprudence. Le paradoxe d'un groupe d'agents rationnels prenant ensembles une décision irationnelle est apparu pour la première fois dans la littérature légale, où les jurys à la cour devaient justifier leur décisions.

Le problème connu sous le nom de paradoxe doctrinal a été redécouvert en 1921 par le théoricien légal Vacca en 1921 qui a soulevé des critiques sur le fait que la décision du groupe puisse s'écarter des préférences individuelles de chacun. Cependant le paradoxe doctrinal tel que connu actuellement est attribué au papier de 1986 de Kornhauser et Sager

Prenons un exemple donné par Kornhauser et Sager pour illuster le paradoxe doctrinal :

	Contrat valide p	Violation q	Accusé coupable r
Juge 1	1	1	1
Juge 2	1	0	0
Juge 3	0	1	0
Majorité	1	1	0

Un jury composé de trois personnes doit régler un cas de violation de contrat entre une accusation et une défense. Le contrat stipule que la défense est responsable seulement si le contrat est valide et qu'il a été violé (les deux prémises notées ici p et q).

La cour peut statuer *directement*, en prenant en compte le vote de la majorité sur la conclusion r sans prendre en compte ce que les jurés pensent des prémisses (procédure basée sur la conclusion) ou *indirectement* en demandant l'avis des juges sur les prémisses et en donnant la décision de la cour sur r en appliquant la règle $(p \wedge q) \leftrightarrow r$ (procédure basée sur les prémisses). Dans cet exemple précis ces deux procédures donnent des résultats différents. En effet la défense est innocente selon la première et coupable selon la deuxième.

Les différences de jugement entre ces deux procédures ont longtemps été débattues, et certains comme Nash on même proposé des procédures alternatives. Kornhauser et Sager proposent ici de voter au cas par cas la procédure à utiliser lors d'un jugement.

Le dilemme discursif

Le philosophe Pettit a reconnu le premier que le paradoxe doctrinal illustre un problème bien plus général qu'une simple décision de cour. Il introduit alors le terme de *dilemme discursif* pour désigner toute décision de groupe dont l'issue dépend de la méthode d'agrégation choisie.

List a ensuite reconstruit l'exemple de Kornhauser et Sager comme ci-dessous. La différence est ici que la méthode d'agrégation choisie a été ajoutée à l'ensemble de notions que les juges doivent voter. Le dilemme discursif est caractérisé par le fait que le groupe aboutit à une décision incohérente, comme $\{p, q, (p \wedge q) \leftrightarrow r, \neg r\}$. La cour accepterait la méthode d'agrégation, donnerait un jugement positif sur les prémisses p et q, et rejeterait la conclusion r. Une position de ce type est intenable, elle reviendrait à relâcher l'accusé tout en admettant en même temps que les conditions nécessaires à sa culpabilité sont remplies.

	Contrat valide	Violation	Méthode d'agrégation	Accusé coupable
	p	q	$(p \wedge q) \leftrightarrow r$	r
Juge 1	1	1	1	1
Juge 2	1	0	1	0
Juge 3	0	1	1	0
Majorité	1	1	1	0

List et Petit on étudié le problème en regardant les ensembles de propositions choisis individuellement par les électeurs. La théorie de l'agrégation du jugement devient alors une enquête formelle sur les conditions dans lesquelles des jugements individuels cohérents peuvent aboutir à une décision de groupe incohérente.

I Notation de base

Nous allons rappeler au lecteur les principales notions de l'agrégation du jugement.

Langage : Un langage L est un ensemble de phrases, appelées *propositions* et doté d'un *opérateur de négation* $\neg$ et d'une notion de cohérence. Le langage est clos pour l'opération de négation ($\varphi \in \mathrm{L} \implies \neg\varphi \in \mathrm{L}$ et chaque ensemble de propoistions est cohérent ou incohérent (mais pas les deux à la fois).

Agenda : Ensemble fini de formules de logique propositionnelle de la forme $\Phi{=}\Phi^+ \bigcup \{\neg\varphi | \varphi \in \Phi^+\}$ telles que le *pre-agenda* Φ^+ ne contient que des formules non-niées.

Ensemble de décision : Un sous-ensemble $\mathrm{J} \subseteq \Phi$. J est *complete* si $\forall\varphi \in \Phi^+, \varphi \in \mathrm{J}$ *or* $\neg\varphi \in \mathrm{J}$. $\mathcal{J}(\Phi)$ représente l'ensemble de tous les ensembles de décision complets et cohérents sur Φ.

Soit $\mathcal{N} = [\![1, n]\!]$ un ensemble fini d'agents. Un *profile* $\mathbf{J} = (J_1, ..., J_n)$ est un vecteur d'ensembles de décision, un pour chaque agent. L'agent *i accepte* la formule $\varphi \in \Phi$ si $\varphi \in J_i$. La *coalition des supporters* de φ dans le profil J est notée $\mathcal{N}_\varphi^{\mathbf{J}} = \{i \in \mathcal{N} \mid \varphi \in J_i\}$.

Un profil est appelé *unanime* s'il est de la forme $\mathbf{J} = (J, ..., J)$.

Règle d'aggrégation : C'est une fonction $\mathcal{F} : \mathcal{J}(\Phi)^n \to 2^\Phi$ qui attribue un ensemble de décision collectif $\mathcal{F}(\mathbf{J})$ à chaque profil complet et cohérent (2^Φ représente le *powerset* de Φ). Si $\varphi \in \mathcal{F}(\mathbf{J})$ alors le groupe accepte φ.

Une règle du quotat est une règle d'aggrégation induite par un nombre $q \in \{0, ..., n+1\}$ telle que :

$$\mathcal{F}_q(\mathbf{J}) = \left\{\varphi \in \Phi \mid N_\varphi^{\mathbf{J}} \geq q\right\}$$

Pour $q = \left\lceil \frac{n}{2} \right\rceil$ on a une règle de la majorité, tandis que pour $q = n$ on a une règle de l'unanimité.

II Les différentes méthodes d'aggrégation

II - 1 Rappels sur la notion d'agenda

Implication : Une proposition φ (ou un ensemble de propositions J) entraîne une proposition φ' (ou un ensemble de propositions J') si φ (resp. $\bigcap_{q \in J} q$) est un sous-ensemble de φ' (resp. $\bigcap_{q \in J'} q$).

Résolution : Un ensemble de propositions $\mathcal{J}$ résoud une proposition φ s'il entraîne φ ou $\neg\varphi$.

Etendue : L'étendue d'un agenda Φ est l'ensemble $\overline{\Phi}$ de propositions résolues par chaque ensemble de décision rationnel $J \in \mathcal{J}_{\Phi}$. C'est de manière équivalente la fermeture de X par union (ou intersection) et négation, c'est donc l'algèbre générée par Φ. L'étendue d'un agenda peut être très large, on a toutes les unions et intersections de propositions de l'agenda, toutes les intersections de négations d'unions de propositions de l'agenda et ainsi de suite. L'algèbre $\overline{\Phi}$ engendrée par Φ est elle-même un agenda où $\Phi \in \overline{\Phi}$. On a de plus que $\Phi et \Phi'$ sont deux agendas équivalents si $\overline{\Phi} = \overline{\Phi'}$

II - 2 Résultats préliminaires

Implication conditionnelle : $\varphi \in \Phi$ implique conditionnellement une autre proposition $\psi \in \Phi$, que l'on écrira $\varphi \models_c \psi$ si :

$\varphi \bigcup Y \models \psi$ pour un ensemble $\Psi \subseteq \Phi$ cohérent avec φ et $\neg\psi$.

On écrira de plus, pour $\varphi, \psi \in \Phi$, que $\varphi \models\!\models_c \psi$ si :

Il existe une suite de propositions $\varphi_1, ..., \varphi_k \in \Phi$ telle que $\varphi = \varphi_1 \models_c \varphi_2 \models_c ... \models_c \varphi_k = \psi$. $\models\!\models_c$ est donc la fermeture transitive de $\models_c$.

Indépendance par rapport à une proposition : Pour tout $\varphi \in \Phi$ et tous les profils $(A_1, ..., A_n), (A'_1, ..., A'_n)$, si $\varphi \in A_i \iff \varphi \in A'_i$ pour tous les individus i alors $\varphi \in \mathcal{F}(A_1, ..., A_n) \iff \varphi \in \mathcal{F}(A'_1, ..., A'_n)$

Préservation de l'unanimité : Pour tous les profils unanimes admissibles $(A, ..., A)$ on a $\mathcal{F}(A, ..., A) = A$.

Non-dictatoriale : Il n'existe pas d'individu $i \in \mathcal{N}$ (le dictateur) telle que $\mathcal{F}(A_1, ..., A_n) = A_i$ pour chaque profil admissible $(A_1, ..., A_n)$

Verouillage Complet : Un agenda Φ est complètement verouillé si, pour toutes les propositions $\varphi, \psi \in \Phi$ on a $\varphi \models\!\models_c \psi$

Incohérence minimale : Un ensemble $S \subseteq$ L est minimalement incohérent s'il est incohérent mais que tout ses sous-ensembles distincts de lui-même sont cohérents.

Niable un nombre pair de fois : Un agenda Φ admet un nombre pair de négations s'il existe un ensemble minimalement incohérent $\Psi \subseteq \Phi$ avec un sous-ensemble $\Omega \subseteq \Psi$ de taille paire tel que $(\Psi \backslash \Omega) \bigcup \{\neg\varphi : \varphi \in \Omega\}$ est cohérent.

On utilisera plus fréquemment la définition équivalente : Il existe un ensemble minimalement incohérent $\Psi \subseteq \Phi$ avec deux éléments distincts $\varphi, \psi \in \Psi$ tel que $(\Psi \backslash \{\varphi, \psi\} \bigcup \{\neg\varphi, \neg\psi\}$ est cohérent.

II - 3 L'aggrégation des préférences individuelles

Théorème 1. *Si un agenda est complètement verouillé et niable un nombre pair de fois alors il n'existe pas de règle d'aggrégation $\mathcal{F} : \mathcal{J}(\Phi)^n \to \mathcal{J}(\Phi)$ indépendante, non-dictatoriale et qui préserve l'unanimité. Dans tout autre cas, on peut trouver une telle règle.*

II - 3.1 Aggrégation non-oligarchique

On dit d'une règle d'aggrégation $\mathcal{F}$ qu'elle est *oligarchique* s'il existe un ensemble non-vide $\mathcal{M} \subseteq \mathcal{N}$ *(les oligarques)* et un ensemble de décision $\Psi \in \mathcal{J}(\Phi)$ *(l'origine)* tels que pour tout $\varphi \in \Phi$ et tous les profils admissibles $(J_1, ..., J_n)$,

$$\varphi \in \mathcal{F}(J_1, ..., J_n) \iff \begin{cases} \varphi \in J_i \text{ pour } \textit{tous} \text{ les oligarques } i \in M \text{ si } \varphi \in \Phi \backslash \Psi \\ \varphi \in J_i \text{ pour } \textit{certains} \text{ oligarques } i \in M \text{ si } \varphi \in \Psi. \end{cases}$$

Sous cette définition d'oligarchie on constate qu'on groupe d'oligarques possède lde pouvoir de déterminer la décision collective sur une proposition φ dès qu'ils sont unanimes sur cette proposition. Les oligarques possèdent également le pouvoir de forcer le groupe à changer sa décision sur toute proposition φ dès qu'ils ne sont pas en accord avec le reste du groupe.

Semi-verouillé : Un agenda Φ est *semi-verouillé* si, pour toutes propositions $\varphi, \psi \in \Phi$, $[\varphi \models\models_c \psi \textit{ et } \psi \models\models_c \varphi]$ ou $[\varphi \models\models_c \neg\psi \textit{ et } \neg\psi \models\models_c \varphi]$.

Théorème 2. *Si un agenda est semi-verouillé et niable un nombre pair de fois, il n'existe pas de règle d'aggrégation $\mathcal{F} : \mathcal{J}(\Phi)^n \to \mathcal{J}(\Phi)$ non-oligarchique, indépendante et qui préserve l'unanimité. Autrement, une telle règle existe.*

Lemme 1. *Chaque agenda non-trivial qui est semi-verouillé admet un nombre pair de négations.*

Corollaire 2.1. *Toute règle d'aggrégation $\mathcal{F} : \mathcal{J}(\Phi)^n \to \mathcal{J}(\Phi)$ indépendante et qui préserve l'unanimité est oligarchique mais toutes ne sons pas dictatoriales si et seulement si l'agenda est semi-verouillé et non trivial.*

II - 3.2 Agrégation Anonyme

L'anonymité requiert un traitement égal entre tous les individus

Anonymité : Pour tous les profils admissibles $(J_1, ..., J_n), (J'_1, ..., J'_n)$ qui sont des permutations les uns des autres, on a $\mathcal{F}(J_1, ..., J_n) = \mathcal{F}(J'_1, ..., J'_n)$

Verouillé : Un agenda Φ est *verouillé* s'il contient une proposition $\varphi \in \Phi$ telle que $\varphi \models\models_c \neg\varphi$ et $\neg\varphi \models\models_c \varphi$

On a par exemple que l'agenda constitué de $a, a \wedge b, a \wedge \neg b, a \wedge c$ et leurs négations. On a bien $a \models\models_c \neg a$ et $\neg a \models\models_c a$ (mais aucune porposition n'implique conditionnellement $a \wedge c$. On a de plus le théorème suivant :

Théorème 3. *Soit $n \in 2\mathbb{N}$. Si l'agenda est vérouillé, il n'exite pas de règle d'aggrégation $\mathcal{F} : \mathcal{J}(\Phi)^n \to \mathcal{J}(\Phi)$ indépendante, anonyme et qui préserve l'unanimité. Autrement, une telle règle existe.*

Corollaire 3.1. *Il existe des règles d'aggrégation $\mathcal{F} : \mathcal{J}(\Phi)^n \to \mathcal{J}(\Phi)$ indépendantes, qui préservent l'unanimité et l'anonymat pour tout groupe de taille n si et seulement si l'agenda est verouillé.*

II - 3.3 Aggregation sans pouvoir de veto

Dans une règle d'aggrégation oligarchique, on remarque que les oligarues on tous le pouvoir de véto. Même les règles d'aggrégation anonymes n'évitent pas la présence d'un tel pouvoir. En effet elles peuvent donner un pouvoir de véto à chaque individu. Pour des questions de démocratie on peut donc souhaiter l'absence de pouvoir de véto.

Absence de pouvoir de véto individuel : Pour tous les profiles admissibles $(J_1, ..., J_n)$ dans lesquels *n-1* ensembles de décision individuels coïncident (ils sont tous égaux à J) alors on a $\mathcal{F}(J_1, ..., J_n) = J$

Verouillage minimal : Un agenda Φ est *verouillé de manière minimale* s'il contient au moins deux propositions non équivalentes $\varphi, \psi \in \Phi$ telles que $\varphi \models\models_c \psi$ et $\psi \models\models_c \varphi$.

On obtient alors le théorème suivant :

Théorème 4. *Si un agenda est verouillé de manière minimale alors il n'existe pas de règle d'aggrégation* $\mathcal{F} : \mathcal{J}(\Phi)^n \to \mathcal{J}(\Phi)$ *indépendante sans pouvoir de véto individuel. Autrement, une telle règle existe si* $n \geq 2^{\frac{|\Phi|}{2}-1}$ *(pour les grands groupes) et une telle règle n'existe pas si* $n \leq k_\Phi$ *où* k_Φ *est la taille du plus grand sous-ensemble minimalement incohérent de* Φ *(pour les petits groupes).*

Le théorème est toujours valide si on impose de plus les conditions d'anonymité, de monotonicité et de préservation de l'unanimité sur $\mathcal{F}$.

On peut également simplifier ce théorème avec le corollaire suivant, qui ne requiert aucune borne.

Corollaire 4.1. *Il existe une règle d'agrégation* $\mathcal{F} : \mathcal{J}(\Phi)^n \to \mathcal{J}(\Phi)$ *indépendante et sans pouvoir de véto individuel pour tout groupe assez grand de taille n si et seulement si l'agenda n'est pas minimalement verouillé.*

III Les comportements stratégiques

Après avoir vu différentes méthodes d'aggrégation des préférences individuelles en préférences de groupe, voyons maintenant dans quelle mesure il est possible de manipuler ces méthodes d'aggrégation pour changer l'issue en sa faveur.

III - 1 Préférences sur les ensembles de décision

Les stratégies que nous étudierons ici impliquerons toutes l'envie d'obtenir une "meilleure" issue. Ainsi pour étudier la notion de comportement stratégiue il est essentiel de définir la notion de préférences selon les opinions. Autrement dit quand une opinion est préférée à une autre

Dans le pire des cas, le nombre d'opinions pouvant potentiellement jouer un rôle est une exponentielle du nombre d'issues. Par exemple, pour m issues on aura 2^m opinions possibles. Il est donc déraisonnable d'attendre de chaque agent qu'il définisse une relation de préférence sur chaue opinion réalisable. Il est plus logique d'utiliser un langage plus spécialisé et compact pour représenter cette relation de préférence.

III - 1.1 Préférences basées sur les sous-ensembles

La première méthode compacte pour préciser les préférences des agents que nous allons étudier est la méthode des préférences basées sur les *sous-ensembles*. Pour cette relation de préférences selon les décisions, un agent avec une opinion sincère $J \in \mathcal{J}(\Phi)$ donne un sous-ensemble $L \in J$ d'issues importantes. Dès lors pour tout ensembles $J_1, J_2 \in \mathcal{J}(\Phi)$ l'ensemble J_1 est préféré à l'ensemble J_2 si $L \subseteq J_1$ et $L \nsubseteq J_2$. En clair, tout ensemble de décision incluant L est toujours prévéré à un ensemble de décision n'incluant pas L.

III - 1.2 Autres relations de préférences

On peut également avoir d'autres relations de préférences. Par exemple :

1. Quand un ensemble de décision est préféré à un autre s'il est en accord avec un ensemble de décision optimal fixé à l'avance sur au moins une issue sur laquelle l'autre ensemble de décision est en désaccord.

2. Quand un ensemble de décision est préféré à un second ensemble de décision s'il est en accord avec un ensemble de décision optimal fixé à l'avance sur au moins une issue sur laquelle l'autre ensemble de décision est en désaccord et si le premier

ensemble est en accord avec l'ensemble de décision optimal sur toutes les issues sur lesquelles le second ensemble est en acord avec l'ensemble de décision optimal.

3. On y trouve également toutes les préférences qui préférent un unique ensemble de décision "préféré" et classent ensuite de manière arbitraire les autres ensembles.

III - 2 La Manipulation : Comment arriver au résultat désiré

III - 2.1 Le mensonge

Le mensonge implique de vouloir influencer le résultat de la procédure d'aggrégation en sa faveur en donnant une décision non sincère. Ceci peut par exemple s'apparenter au dilemme du vote utile dan une élection, qui mettra l'électeur certain de voir son candidat perdre face à un dilemme cornélien. Doit-il voter selon ses préférences et risquer alors de voir le candidat le moins proche de ses idéaux remporter l'élection ou doit-il faire taire ses préférences et voter pour le candidat ayant les idées les plus proches de son favori mais qui a des chances bien plus élevées de l'emporter ?

III - 2.2 La corruption

Une autre forme de comportement stratégique dans l'aggrégation du jugement est la corruption. Dans ces conditions, un agent externe souhaite influencer l'issue d'un scénario en corrompant un certain nombre d'individus.

Le corrupteur a un ensemble $L \in \Phi$ de conclusions désirées qu'il veut que l'opinion collective atteigne. De plus, le corrupteur a un budget qui lui permet de corrompre jusqu'à k individus. Pour tous les individus corrompus, le corrupteur peut spécifier un ensemble de décision arbitraire (complet et cohérent). On pourra chercher à déterminer si le corrupteur peut sélectionner jusqu'à k individus et spécifier un ensemble de décision pour chaque individu tel que l'issue de la procédure d'aggrégation est meilleure (par rapport à L) que sans corruption.

III - 2.3 Contrôle en ajoutant ou en retirant des propositions

Une autre forme de comportement stratégique est le contrôle. Dans cette configuration, un agent externe souhaite influencer l'issue d'une procédure d'aggrégation du jugement en influençant ses conditions. Ainsi, comme par exemple dans une des réunions du conseil d'administration d'une grosse compagnie. La personne en charge de la réunion pourra énoncer un problème et décider de ne proposer que les suggestions qui l'intéressent au reste du conseil d'administration. Les autres membres auront alors l'impression d'avoir tout pouvoir de décision tandis que le manipulateur aura eu le résultat escompté.

III - 2.4 Echange de propositions

Lorsqu'un agent externe souhaite influencer l'issue d'une procédure d'aggrégation, il peut aussi le faire lorsque le choix des électeurs sur les issues données dépend des précédentes. Par exemple, soit deux issues $\varphi, \psi \in \Phi$ si on a que accepter φ implique accepter ψ et que accepter ψ ne donne aucune assurance sur l'acceptance ou le rejet de φ alors l'agent aura tout intérêt à s'assurer que φ soit présentée au reste du groupe avant ψ pour obtenir le résultat désiré.

III - 3 Comment empêcher la manipulation

Certaines conditions d'indépendance peuvent empêcher un type spécifique de manipulation par les votants et par les maîtres de l'agenda.

Indépendance sur Ψ : Pour chaque proposition $\varphi \in \Psi$ et chaque couple de profiles $(J_1, ..., J_n), (J'_1, ..., J'_n) \in \mathcal{J}(\Phi)$, F est indépendante sur Ψ si on a que $[\forall i, J_i \models \varphi \iff J'_i \models \varphi] \implies [\mathcal{F}(J_1, ..., J_n) \models \varphi \iff \mathcal{F}(J'_1, ..., J'_n) \models \varphi$

On obtient alors les affirmations suivantes :

- En imposant l'indépendance, toute décision sur une proposition $\varphi \in \Phi$ ne peut être modifiée en ajoutant ou retirant des propositions dans Φ autre que φ
- Si $\mathcal{F}$ viole la condition d'indépendance, alors il existe un profil $(J_1, ...J_n) \in \mathcal{J}(\Phi)$ tel que la décision sur une proposition $\varphi \in X$ peut être modifiée par une manipulation d'agenda du type de la remarque précédente.

Pour empêcher la manipulation d'agenda il nous faut donc garantir une condition d'indépendance sur la fonction d'aggrégation.

IV Les limites et critiques de l'aggrégation du jugement

IV - 1 Le rôle de la délibération

Une des objections les plus communes à l'approche de l'aggrégation du jugement est celle selon laquelle les procédures d'aggrégation vont remplacer la délibération. Pour concrétiser cette objection, on peut lire les résultats de l'aggrégation du jugement comme suggérant que les membres d'un conseil d'administration ou les scientifiques ne sont pas autorisés à communiquer entre eux. Au lieu de ça, ils sont contraints de soumettre leurs opinions à une boîte noire qui va les aggréger entre elles en une seule décision finale. Magnus écrira par exemple que *les procédures d'aggrégations traitent les scientifiques comme les entrées séparées d'un algorithme*.

Tous les scientifiques ne sont ccependant pas d'accord pour affirmer que l'aggrégation du jugement mettra un terme à la délibération. Miriam Solomon écrit en 2006 que les groupes en délibération sont soumis à la pression du groupe et qu'un individu pourrait être incité au consensus et amené à fare disparaître des preuves d'une incohérence lorsqu'un désagrément apparaît. Elle conclut alors que l'aggrégation du jugement, *sans la délibération*, évite la pression de groupe et permet une vision collective en accord avec toutes les données disponibles.

K.Brad Wray écrit dans une réponse à Solomon que cette drnière était trop rapide à écarter la délibération. Il affirme que le consensus est nécessaire dans une équipe en collaboration pour avoir une vue d'ensemble et que la délibération demeure donc un important procédé. Wray affirme de plus qu'il existe des stratégies pour écarter (ou tout du moins limiter) la pression de groupe et, sans toutefois rejeter complètement l'aggrégation du jugement, est persuadé qu'elle n'est pas essentielle dans toutes les décisions de groupe.

Je suis d'accord avec Wray pour dire que le consensus est nécessaire et que la délibération est un important procédé, de plus Wray à raison de dire qu'il existe des stratégies pour limiter la pression de groupe sans toutefois avoir besoin des méthodes d'aggrégation que nous avons étudié plus haut. Cependant il est à mon sens erroné de dire que l'aggrégation du jugement n'est pas essentielle dans chaque prise de décision, c'est là confondre le rôle de la délibération et de l'aggrégation du jugement, qui sont de manière complémentaires d'arriver à une décision de groupe. L'aggrégation du jugement fournit des fondations solides et fournit un objectif à la délibération, un résultat cohérent et irréfutable tandis que la délibération est la manière de mener les débat et leur déroulement du début au résultat. Ainsi l'un peut avoir lieu sans l'autre mais c'est ensemble que ces deux modes d'aggrégation des décisions font le plus sens.

IV - 2 Le rôle des valeurs

Une autre objection à l'apprche de l'aggrégation du jugement est à propos du risque par induction. En particulier, Magnus s'inquiète que les procédures d'aggrégation ne prennent pas en compte les conclusions non vérifiées que les scientifiques peuvent faire, telles que le raisonnement par induction.

En effet, les scientifique induisent des conclusions a propos de données ou au moins décident s'ils vont publier un papier si des preuves suffisamment solides supportent leur hypotèses. De telles inductions peuvent être incertaines, et pour les réaliser un individu doit, comme le précise Richard Rudner dans son papier de 1953, les inductions de la forme $\varphi \rightarrow \psi$ ne peuvent être faite seulement si on connaît les risques d'induire ψ à partir de φ. Magnus soulève l'inquiétude suivante, qui est également celle soulevée par List et Petit dans le paradoxe doctrinal : "si l'on sonde les scientifiques sur [$\varphi \rightarrow \psi$] alors nous devrons accepter toute décision en accord avec leurs valeurs intrinsèques". Comme une fonction d'aggrégation du jugement "élude le rôle des valeurs" elles ne sont pas en adéquation avec le savoir collectif scientifique. En effet, on pourrait arriver comme dans le paradoxe à une situation où la majorité pense qu'il est risqué d'inférer ψ à partir de φ mais où une autre majorité décide d'accepter ψ.

Cette objection est liée à la manière dont la délibération se lie à l'aggrégation du jugement. En effet, les valeurs peuvent être éludées si la dlibération ne se fait pas et, comme il a déjà été dit plus haut, la délibération et l'aggrégation du jugement ne sont pas exclusif. En effet, il est inconcevable que l'aggrégation du jugement puisse remplacer un processus de délibération et d'analyse pour bien gérer le risque dans la recherche scientifique.

IV - 3 Impossibilité dans l'aggrégation du jugement

Rappelons tout d'abord quelques définitions :

La première condition certifie qu'une règle d'aggrégation doit accepter en entrée tout profil d'ensembles de décision individuels cohérents.

Domaine Universel : Le domaine de $\mathcal{F}$ est l'ensemble de tous les profils possibles d'ensembles de décision complets et cohérents sur Φ

La seconde condition restraint la sortie de la règle d'aggrégation, en imposant qu'elle produise un ensemble de décision collectf rationnel pour chaque profil admissible d'ensembles de décision individuel.

Rationnalité collective : Pour chaque profil $(J_1, ..., J_n)$ dans le domaine de $\mathcal{F}$, $\mathcal{F}(J_1, ..., J_n)$ est un ensemble de décision collectif complet et cohérent sur Φ

La troisième et la quatrième condition restreignent la manière dont les résultats sont générés à partir des entrées.

Systématicité : Pour chaque couple de profils $(J_1, ..., J_n), (J'_1, ..., J'_n)$ dans le domaine de $\mathcal{F}$ et chaque couple de proposition $\varphi, \psi \in \Phi$, $[\varphi \in J_i \iff \psi \in J'_i \forall i \in N] \implies [\varphi \in \mathcal{F}(J_1, ..., J_n) \iff \psi \in \mathcal{F}(J'_1, ..., J'_n)]$.

Cette condition peut être exprimée comme une obligation que :

- L'ensemble de décision collectif sur chaque propositino de Φ dépend seulement des décisions des individus sur cette proposition et non des décision sur les autres propositions (*indépendance*)
- Le critère pour déterminer la décision collective sur chaque proposition doit être le même sur chaque proposition de Φ (*neutralité*)

On rappelle de plus la condition d'anonymat.

Anonymat : Pour chaque couple de profils $(J_1, ..., J_n), (J'_1, ..., J'_n)$ dans le domaine de $\mathcal{F}$ qui sont des permutations les uns des autres, on a $\mathcal{F}(J_1, ..., J_n) = \mathcal{F}(J_1, .., J_n)$

Et enfin la monotonicité

Monotonicité : Si une proposition $\varphi \in \Phi$ est collectivement acceptée pour un profil donné $(J_1, ..., J_n)$ et si on considère un autre profil $(J_1, ..., J'_i, ..., J_n)$ dans lequel un autre individu i accepte φ (on a donc $\varphi \in J'_i$ tandis que $\varphi \notin J_i$ tandis que tous les autres individus ont le même ensemble de décision que précédemment, alors φ est également collectivement acceptée par le second profil.

On a alors le théorème suivant :

Théorème 5. *Soit $\{\varphi, \psi, \varphi \wedge \psi\}$ (où $\varphi \wedge \psi$ peut être remplacé par $\varphi \vee \psi$ ou $\varphi \to \psi$). Alors il n'existe pas de règle d'aggrégation qui satisfait les conditions de omaine universel, de rationnalité collective, de systématicité et d'anonymat. (List et Pettit 2002)*

Il peut être déçevant de voir qu'une telle fonction n'existe pas mais ce n'est pas un grand pas en arrière. En effet, la condition de domaine universel ne reflète pas les normes des publications scientifiques. Il est donc possible de l'abandonner pour éviter la situation d'impossibilité.

Dietrich et List proposent également des moyens d'éviter l'impossibilité en affaiblissant la condition de domaine universel, mais alors on se retrouverait dans une situation où la condition d'anonymat et de domaine universel se retrouveraient violées. Or la condition d'anonymat est importante dans toute aggrégation de décisions dans le sens où elle permet à chacun de se sentir en sécurité pour donner son opinion en toute honnêteté.

La seconde alternative pour éviter l'impossibilité est de permettre des fonctions d'aggrégations incomplètes. On peut donc avoir des fonctions d'aggrégations qui remplissent les conditions demandées mais qui sont incapables d'affirmer ou infirmer certaines propositions considérées. Une manière naturelle d'implémenter ce fait serait d'avoir une fonction d'aggrégation qui valide toute proposition validée par une majorité absolue d'électeurs (si 51% des électeurs décident de valider φ alors la fonction d'aggrégation pourra valider φ sans avoir besoin de communiquer ses statistiques).

Cette alternative risquerait alors de nous ramener à la situation du paradoxe doctrinal et alors de nous conduire une fois de plus à une situation où une communauté (scientifique par exemple) va affirmer que φ et $\varphi \to \psi$ sont vraies mais que ψ est fausse.

Des commités comme le commité international des éditeurs de journaux médicaux sont donc en faveur de conditions comme celle de l'unanimité, dans les groupes de travail et affirme donc que c'est une des principales conditions qui devraient être satisfaites.

Cependant, la règle de l'unanimité est trop restrictive pour être recommandée comme une règle générale de l'aggrégation du jugement pour la collaboration scientifique. Les scientifiques ne voient pas cette règle comme une règle réaliste et ne la prennent donc pas en compte. Ainsi, on a que retirer la complétude amène à une règle de l'unanimité mal appliquée comme unique manière d'aggréger les préférences individuelles. Ainsi, on peut raisonnablement conclure que retirer la complétude n'est pas une manière valide d'éviter l'impossibilité.

Une autre solution a alors été trouvée par une équipe de chercheurs (Bright, Dang and Heesen) qui proposent alors de retirer la condition de domaine universel. On peut alors utiliser la condition de vote majoritaire proposition par proposition. Ainsi une proposition est acceptée si et seulement si la majorité l'accepte. Cette manière d'aggréger les préférences satisfait l'unanimité, l'anonymat, la systématicité et satisfait la complétude si jamais le nombre d'individus est impair.

V Conclusion

Nous avons donc ici étudié une vue d'ensemble de l'aggrégation du jugement. Tout d'abord, nous connaissons maintenant les conditions principales qui sont étudiées lorsqu'il est question de construire une fonction d'aggrégation des préférences individuelles.

Nous avons étudié différentes manières d'aggréger des préférences individuelles en préférences de groupes et quels sont les moyens possibles pour se prémunir de la manipulation du vote pour le garder aussi neutre que possible.

Nous avons ensuite observé les principales critiques autour de l'aggrégation du jugement et notamment comment certains scientifiques semblent déterminés à l'opposer au dialogue et à la délibération au sein de la communauté scientifique. Cependant ces notions ne sont pas opposées et il est bon de les mettre en lien pour pouvoir produire des délibérés justes avec un but précis.

Enfin, nous avons vu que l'aggrégation du jugement a ses limites et ne doit ainsi pas être considérée comme un outil tout puissant auquel on peut se fier sans ne plus devoir utiliser son bon sens mais comme un moyen de s'assurer que le résultat des délibérations satisfait des conditions prédéfinies au départ.

Il serait intéréssant de baser nos modèles politiques sur cette théorie et d'étudier comment éliminer le dilemme du vote utile et la dépendance aux alternatives non pertinentes dans une élection en appliquant des théories compréhensibles pour le grand public.

Références

[1] Davide Grossi et Gabriella Pigozzi. Judgment Aggregation *A Primer*. Morgan Claypool Publishers, 2000

[2] Sirin Botan, Arianna Novaro et Ulle Endriss. Group Manipulation in Judgment Aggregation

[3] Ronald de Haan. Complexity Results for Manipulation, Bribery and Control of the Kemeny Judgment Aggegation Procedure. Institute of Computer Graphics and Algorithms, March 2017

[4] Franz Dietrich. Judgment aggregation and agenda manipulation, December 2015

[5] Franz Dietrich. Judgment aggregation :(im)possibility theorems, December 2003

[6] Franz Dietrich et Christian List. Propositionwise judgment aggregation : the general case*. May 2009

[7] Christian List. The theory of judgment aggregation : An introductory review. August 2009

[8] Bright, L.K., Dang, H. Heesen, R. A Role for Judgment Aggregation in Coauthoring Scientific Papers. Erkenn 83, 231–252 (2018).

[9] Rudner, R. (1953). The scientist qua scientist makes value judgments. Philosophy of Science

[10] Solomon Miriam. Groupthink versus the wisom of crowds : The social epistemology of deliberation and dissent. The Southern Journal of Philosophy, 2006