

Correspondances d'agrégation de jugements basées sur le nombre de votes

Patricia Everaere¹
patricia.everaere@univ-lille1.fr

Sébastien Konieczny²
konieczny@cril.fr

Pierre Marquis²
marquis@cril.fr

¹ LIFL-CNRS, Université Lille 1, France

² CRIL-CNRS, Université d'Artois, France

Résumé :

L'agrégation de jugements vise à prendre des décisions collectives sur des questions (fermées) à partir d'un ensemble de jugements exprimés par des individus (agents) sur ces questions. De nombreux résultats importants en agrégation de jugements prennent la forme de théorèmes d'impossibilité, qui identifient des ensembles de conditions, individuellement rationnelles, ne pouvant être satisfaites conjointement. En conséquence, certaines conditions de rationalité doivent être abandonnées. Dans cet article, nous critiquons la condition de systématisme souvent considérée dans ces théorèmes. Nous définissons la famille des correspondances d'agrégation de jugements basées sur le nombre de votes (abrégées en SAC), qui rassemble des méthodes pour lesquelles la sélection d'une réponse dans un ensemble de jugements sociaux dépend seulement de ses interactions logiques avec les autres réponses et du nombre de votes individuels que ces réponses ont obtenus. Nous nous intéressons particulièrement à un sous-ensemble de SAC constitué des méthodes majoritaires ordonnées, qui sélectionnent des ensembles cohérents de jugements basés sur les scores de leurs éléments. Nous étudions les propriétés de rationalité offertes par toutes les méthodes de cette famille, et plus particulièrement par certaines méthodes spécifiques parmi elles.

Abstract:

Judgment aggregation consists in making collective decisions on a set of issues from the judgments expressed by individual agents on these issues. The resulting decisions take the form of social judgements sets in which issues are accepted or rejected. Many important results in judgment aggregation take the form of impossibility theorems, identifying sets of rationality conditions cannot be jointly satisfied. As a consequence, some rationality conditions must be given up. In this paper we criticize the systematicity condition often considered in such theorems. We define the family of Support-based Aggregation Correspondences (SAC), gathering methods for which the selection of an issue in a social judgements set depends only on its logical interactions with the other issues and of the level of support it gets from the individuals. We specifically focus on the subset of SAC consisting of ranked majority methods, which select consistent judgements sets based on the scores of their elements. We investigate the rationality conditions offered by all the methods of the family and by some specific methods in it.

Keywords: Choix social, agrégation de jugements, support, systématisme

1 Introduction

L'agrégation de jugements a pour objectif de prendre des décisions collectives sur un en-

	1	2	3	4	5	6	7	<i>maj</i>
φ_1	1	1	1	1	1	0	0	1
φ_2	1	1	1	0	0	1	1	1
φ_3	0	0	0	1	1	1	1	1

TABLE 1 – Un paradoxe doctrinal

semble de questions fermées, à partir des opinions (jugements) exprimées par les individus sur ces questions. Le résultat d'un tel processus d'agrégation est un ou plusieurs ensembles de jugements sociaux, qui reflètent l'opinion du groupe d'individus sur ces questions. L'agrégation de jugements est un sujet de recherche très actif en théorie du choix social et en philosophie politique (Bovens and Rabinowicz [2006]; Gärdenfors [2006]; List and Pettit [2002]; List and Puppe [2009]; List [2004]; Pettit [2001]). La plupart des travaux en agrégation de jugements se sont intéressés aux propriétés de rationalité des méthodes d'agrégation, et les principaux résultats sont des théorèmes d'impossibilité, qui montrent qu'aucune méthode d'agrégation de jugements ne satisfait toutes ces propriétés (Dietrich [2006]; List and Pettit [2002]).

Les difficultés posées par l'agrégation de jugements individuels peuvent être illustrées en considérant un exemple simple. Considérons trois formules logiques φ_1 , φ_2 , φ_3 , telle que $\neg\varphi_1 \vee \neg\varphi_2 \vee \neg\varphi_3$ est une formule valide. La table 1 montre les opinions de chaque individu d'un groupe de sept personnes $\{1, 2, 3, 4, 5, 6, 7\}$ sur ces trois formules (1 signifie "pour" et 0 "contre"). Chaque opinion individuelle est contrainte uniquement par la condition de rationalité liée à la cohérence de son opinion. Ainsi, ici, chaque vote est nécessairement cohérent avec la contrainte $\neg\varphi_1 \vee \neg\varphi_2 \vee \neg\varphi_3$. Supposons que le vote majoritaire simple soit utilisée pour prendre la décision collective. Comme une majorité d'individus est pour φ_1 , pour φ_2 et pour φ_3 , la décision prise par la majorité devrait être $\varphi_1 \wedge \varphi_2 \wedge \varphi_3$, qui n'est pas cohérent. Les exemples de cette nature sont appelés *paradoxes doctrinaux* (voir List and Puppe [2009] pour plus de détails).

Bien que cet exemple montre que le vote majoritaire simple n'est pas satisfaisant comme méthode d'agrégation de jugements (JA), un enjeu capital est de déterminer des méthodes alternatives capables de le faire. Comme les théorèmes d'impossibilité en théorie du vote, les théorèmes d'impossibilité en agrégation de jugements montrent qu'il n'y a aucune méthode qui peut satisfaire l'ensemble des propriétés rationnelles attendues (Dietrich [2006]; List and Pettit [2002]). En conséquence, certaines de ces conditions de rationalité doivent être abandonnées.

La plupart des théorèmes d'impossibilité considèrent une propriété appelée *systematicité*¹, qui demande que les jugements collectifs sur une question donnée ne dépendent pas des votes sur les autres questions. Cette propriété est souvent présentée comme la traduction dans le cadre de l'agrégation de jugement de la propriété d'indépendance aux alternatives non pertinentes (IIA) de la théorie du vote (Arrow [1963]). Cet argument est cependant discutable puisque, alors que la propriété IIA demande que *deux* choix soient comparés indépendamment des autres, la *systematicité* demande de prendre une décision sur *une* question indépendamment des autres. De ce fait, la *systematicité* interdit de considérer l'agrégation de jugements comme un processus d'optimisation qui tente de trouver un meilleur compromis, ce qui est souvent attendu (l'exemple 1 illustre ce point de vue²).

A l'opposé, de nombreuses méthodes d'agrégation de jugement existantes ne satisfont pas la condition de neutralité, qui nous semble cependant une demande minimale. En effet, la neutralité établit que toutes les questions doivent a priori être considérées de la même façon, sans aucune priorité entre elles. La méthode basée sur les prémisses (*premise-based method* Bovens and Rabinowicz [2006]; Dietrich and Mongin [2010]; Pettit [2001]), la méthode basée sur les conclusions (*conclusion-based method* Bovens and Rabinowicz [2006]; Pettit [2001]; Pigozzi *et al.* [2009]), la règle de priorité séquentielle (*sequential priority rule* List [2004]), et les règles par quota (*quota rules* Dietrich and List [2007]) contredisent toutes la neutralité, en utilisant des informations complémentaires pour prendre les décisions (par exemple l'identification de l'ensemble des prémisses). Bien entendu, lorsque des informations additionnelles sont disponibles (comme la distinction entre

prémisses et conclusions, ou certaines priorités entre les questions), il est naturel de les utiliser et donc la neutralité n'a plus à être respectée. Cependant, dans le cas contraire, la neutralité est un moyen d'éviter l'arbitraire dans le processus d'agrégation. D'ailleurs, les méthodes basées sur des distances (Pigozzi [2006]) et plus généralement les méthodes basées sur la minimisation proposées récemment dans Lang *et al.* [2011] satisfont la neutralité.

Dans cet article, nous commençons par critiquer la propriété de *systematicité* lorsque la neutralité est attendue. Nous proposons à la fois un exemple naturel qui met en lumière les limites de la *systematicité*, et un théorème d'impossibilité montrant que cette propriété est incompatible avec d'autres propriétés rationnelles attendues. Ensuite nous définissons et étudions des méthodes d'agrégation de jugements qui sélectionnent des ensembles de réponses cohérents comme ensembles de jugements sociaux, en se basant sur les supports de leur éléments. Une raison pour s'intéresser à ces méthodes est qu'un ensemble collectif de jugements, s'il est cohérent, doit seulement être choisi par rapport aux votes obtenus. A titre d'illustration, sur l'exemple précédent, on voit que φ_1 et φ_2 reçoivent 5 votes sur 7 votants, alors que φ_3 en reçoit seulement 4. Bien que les trois formules soient acceptées par une majorité d'individus, les supports de φ_1 et de φ_2 sont strictement plus grands que celui de φ_3 . Cela peut être considéré comme un argument essentiel pour préférer l'ensemble de jugements $\{\varphi_1, \varphi_2\}$ (et donc $\neg\varphi_3$), aux ensembles de jugements $\{\varphi_2, \varphi_3\}$ et $\{\varphi_1, \varphi_3\}$. Cependant, de façon étonnante, la plupart des méthodes JA existantes ne prennent pas cette information en compte. Cela est d'autant plus surprenant que si l'on considère ces agrégations selon un point de vue épistémique³, plus une réponse reçoit de votes, plus elle est probable (voir section 5 pour une discussion sur ce point). Afin de remédier à cette constatation, nous introduisons la famille des correspondances basées sur le nombre de votes (alias SAC). Cette famille rassemble des méthodes pour lesquelles la sélection d'une question dans un ensemble de jugements social dépend seulement des ses interactions logiques avec les autres questions et du nombre de votes qu'elle reçoit. Nous nous intéressons plus particulièrement à un sous-ensemble de SAC constitué de méthodes majoritaires ordonnées. De telles méthodes choisissent des ensembles de jugements

1. ou une variante appelée *indépendance*.

2. Voir l'explication dans la suite de cette introduction.

3. Sous les hypothèses usuelles de fiabilité et d'indépendance (Condorcet [1785]; Everaere *et al.* [2010]; List and Goodin [2001]).

cohérents basés sur les scores majoritaires de leur éléments. Nous étudions les propriétés de rationalité de ces méthodes à la fois dans le cas général et pour certaines méthodes spécifiques. Nous présentons une correspondance qui satisfait toutes les propriétés intéressantes, ce qui nous conduit à un « théorème de possibilité ». Nous discutons également l'importance de certaines de ces méthodes majoritaires ordonnées, en considérant le problème de la recherche de la vérité dans l'agrégation.

La suite de cet article est organisée comme suit. Après des préliminaires techniques (section 2), des propriétés de rationalité pour les méthodes d'agrégation sont rappelées et discutées à la section 3. Les SAC et les méthodes majoritaires ordonnées sont définies à la section 4. Nous discutons section 5 le problème de la recherche de la vérité (truth tracking) correspondant à certaines méthodes majoritaires ordonnées. Enfin, la section 6 conclut l'article.

2 Préliminaires formels

Un *agenda* est un ensemble fini, non vide et totalement ordonné de formules propositionnelles non triviales (i.e., non contradictoires et non tautologiques) $X = \{\varphi_1, \dots, \varphi_m\}$.

Un *jugement* sur une formule φ_k de X est un élément de $D = \{1, 0, \star\}$, où 1 signifie que φ_k est supporté, 0 que $\neg\varphi_k$ est supporté, \star que ni φ_k ni $\neg\varphi_k$ n'est supporté. Un *ensemble de jugements* sur X est une application γ de X dans D , qui peut être vue aussi comme l'ensemble des formules $\{\varphi_k \in X \mid \gamma(\varphi_k) = 1\} \cup \{\neg\varphi_k \mid \varphi_k \in X \text{ et } \gamma(\varphi_k) = 0\}$. Pour chaque φ_k de X , γ satisfait $\gamma(\neg\varphi_k) = \neg\gamma(\varphi_k)$, où $\neg\gamma$ est donné par $\neg\gamma(\varphi_k) = \star$ ssi $\gamma(\varphi_k) = \star$, $\neg\gamma(\varphi_k) = 1$ ssi $\gamma(\varphi_k) = 0$, et $\neg\gamma(\varphi_k) = 0$ ssi $\gamma(\varphi_k) = 1$.

On attend souvent des ensembles de jugements qu'ils soient cohérents et résolus : Un ensemble de jugements γ sur X est *cohérent* ssi $\bigwedge_{\{\varphi_k \in X \mid \gamma(\varphi_k) = 1\}} \varphi_k \wedge \bigwedge_{\{\varphi_k \in X \mid \gamma(\varphi_k) = 0\}} \neg\varphi_k$ est cohérent. Il est *résolu* ssi $\forall \varphi_k \in X, \gamma(\varphi_k) = 0$ ou $\gamma(\varphi_k) = 1$.

L'agrégation de jugements vise à associer des ensembles de jugements collectifs à un profil d'ensembles de jugements individuels : Un *profil* $P = (\gamma_1, \dots, \gamma_n)$ sur X est un vecteur d'ensembles de jugements sur X . Il est *cohérent* (resp. *résolu*) lorsque chaque ensemble de jugements du profil est cohérent (resp. résolu).

Pour chaque agenda X , une *méthode d'agrégation de jugements* δ associée à un profil P sur X un ensemble non vide $\delta(P)$ d'ensembles de jugements γ_P sur X . Lorsque $\delta(P)$ est un singleton pour chaque P , la méthode d'agrégation de jugements est appelée une règle (déterministe) d'agrégation de jugements, et elle est appelée une correspondance d'agrégation de jugements sinon (Lang *et al.* [2011]).

3 Rationalité de l'agrégation de jugements

Quelles sont les propriétés que les méthodes d'agrégation de jugements doivent satisfaire ? De nombreuses intuitions différentes sont capturées par les propriétés suivantes (voir Lang *et al.* [2011]; Slavkovic [2012]). La première propriété établit que tous les profils cohérents doivent être considérés :

Universalité. Le domaine de δ est l'ensemble de tous les profils cohérents.

Cette propriété est souvent exprimée (List and Pettit [2002]) comme :

R-universalité. Le domaine de δ est l'ensemble de tous les profils cohérents et résolus.

D'autres propriétés établissent que les résultats doivent également être cohérents et résolus :

Rationalité collective. Pour tout profil P du domaine de δ , $\delta(P)$ est un ensemble de jugements collectifs cohérent.

Complétude collective. Pour tout profil P du domaine de δ , $\delta(P)$ est un ensemble de jugements collectifs résolu.

Il est usuellement attendu que les agents jouent des rôles symétriques :

Anonymat. Pour deux profils $P = (\gamma_1, \dots, \gamma_n)$ et $P' = (\gamma'_1, \dots, \gamma'_n)$ du domaine de δ qui sont des permutations l'un de l'autre, on a $\delta(P) = \delta(P')$.

Systématicité. Pour deux profils $P = (\gamma_1, \dots, \gamma_n)$ et $P' = (\gamma'_1, \dots, \gamma'_n)$ du domaine de δ , et deux formules φ_k et φ_l de X , telles que $\gamma_i(\varphi_k) = \gamma'_i(\varphi_l) \forall i$, si $\gamma_P(\varphi_k) = x$ pour tout $\gamma_P \in \delta_P$, alors $\gamma'_{P'}(\varphi_l) = x$ pour tout $\gamma'_{P'} \in \delta'_{P'}$.

Pour pouvoir définir la prochaine propriété qui exprime une forme de respect de la majorité, nous devons d'abord donner quelques définitions préliminaires : soit g_P^+ , g_P^- et g_P les fonctions d'évaluation de la majorité, définies de X dans \mathbb{Z} et données par $g_P^+(\varphi_k) = |\{\gamma_i \in P \mid \gamma_i(\varphi_k) = 1\}|$, $g_P^-(\varphi_k) = |\{\gamma_i \in P \mid \gamma_i(\varphi_k) = 0\}|$, et $g_P(\varphi_k) = g_P^+(\varphi_k) - g_P^-(\varphi_k)$. Ces fonctions peuvent être naturellement étendues à des fonctions définies de $X \cup \{\neg\varphi_k \mid \varphi_k \in X\}$ dans \mathbb{Z} telles que pour tout $\varphi_k \in X$, on a $g_P(\neg\varphi_k) = -g_P(\varphi_k)$, $g_P^+(\neg\varphi_k) = g_P^-(\varphi_k)$, et $g_P^-(\neg\varphi_k) = g_P^+(\varphi_k)$.

Ces notations nous permettent de définir formellement la règle d'agrégation majoritaire δ^{maj} , qui est la règle d'agrégation de jugements considérée dans l'exemple donné en introduction.

Définition 1 La règle d'agrégation majoritaire δ^{maj} est définie comme suit. Pour tout agenda X et tout profil P sur X , on a $\delta^{maj}(P) = \{\gamma_P^{maj}\}$, où pour tout $\varphi_k \in X$:

$$\gamma_P^{maj}(\varphi_k) = \begin{cases} 1 & \text{si } g_P(\varphi_k) > 0 \\ 0 & \text{si } g_P(\varphi_k) < 0 \\ \star & \text{si } g_P(\varphi_k) = 0 \end{cases}$$

On peut à présent définir la propriété de préservation de la majorité :

Préservation de la majorité. Si γ_P^{maj} est cohérent et résolu, alors $\delta(P) = \delta^{maj}(P)$.

La préservation de la majorité⁴ (Slavkovic [2012]) est une propriété très naturelle, qui établit que si le vote majoritaire simple sur chaque question conduit à un ensemble de jugements cohérent, alors la correspondance d'agrégation de jugements doit contenir cet ensemble, et seulement lui. En effet, il semble raisonnable d'attendre que le résultat soit celui du vote majoritaire simple si aucun paradoxe doctrinal n'existe.

D'autres propriétés logiques semblent aussi raisonnables pour les correspondances d'agrégation de jugements. Ainsi, lorsque toutes les informations disponibles sont données par le profil initial, on peut attendre que tous les jugements jouent des rôles symétriques :

Neutralité. Si $X = \{\varphi_1, \dots, \varphi_m\}$ et $X' = \{\varphi'_1, \dots, \varphi'_m\}$ sont deux agendas tel qu'il existe

4. Appelée préservation de la majorité forte dans Slavkovic [2012].

une permutation σ sur $\{1, \dots, m\}$ satisfaisant $\varphi_k = \varphi'_{\sigma(k)}$ pour tout $k \in \{1, \dots, m\}$, alors pour tous profils $P = (\gamma_1, \dots, \gamma_n)$ sur X et $P' = (\gamma'_1, \dots, \gamma'_n)$ sur X' tels que pour tout $i \in \{1, \dots, n\}$, pour tout $k \in \{1, \dots, m\}$, $\gamma_i(\varphi_k) = \gamma'_i(\varphi'_{\sigma(k)})$, on a $\delta(P) = \delta(P')$.

L'unanimité établit que si tous les agents s'accordent sur une réponse, alors les jugements collectifs sur cette réponse doivent être l'avis unanime :

Unanimité.

- Pour tout $\varphi_k \in X$, si $\gamma_i(\varphi_k) = 1 \forall \gamma_i \in P$, alors pour tout $\gamma_P \in \delta(P)$, on a $\gamma_P(\varphi_k) = 1$.
- Pour tout $\varphi_k \in X$, si $\gamma_i(\varphi_k) = 0 \forall \gamma_i \in P$, alors pour tout $\gamma_P \in \delta(P)$, on a $\gamma_P(\varphi_k) = 0$.

Enfin, nous définissons le swap $\gamma_P|_S$ d'un ensemble de jugements collectifs γ_P par rapport à un ensemble $S \subseteq X$ comme $\gamma_P|_S(\varphi_k) = \gamma_P(\varphi_k)$ si $\varphi_k \notin S$, et $\gamma_P|_S(\varphi_k) = \neg\gamma_P(\varphi_k)$ si $\varphi_k \in S$. Nous pouvons alors considérer la propriété suivante :

Swap Optimalité. Si $\gamma_P \in \delta(P)$, alors il n'existe pas $\varphi_k, \varphi_l \in X$ tel que $\gamma_P|_{\{\varphi_k, \varphi_l\}}$ est cohérent et $\max(g_P^+(\neg\varphi_k), g_P^+(\neg\varphi_l)) > \max(g_P^+(\varphi_k), g_P^+(\varphi_l))$.

La swap optimalité demande que la sélection d'ensembles de jugements collectifs soit basée sur les votes reçus par chaque réponse. Donc si on peut remplacer dans un ensemble de jugements collectifs un couple de formules de l'agenda par leur négation, on ne peut trouver une meilleure fonction d'évaluation pour aucune de ces deux négations de formules. La **R-swap optimalité** est une relaxation de la **swap optimalité** quand seuls des profils résolus sont considérés.

Il apparaît que toutes ces propriétés ne sont pas compatibles ensemble. D'abord, nous présentons le théorème d'impossibilité de List and Pettit [2002] pour les règles d'agrégation de jugements seulement⁵.

Proposition 1 (List and Pettit [2002]) *Il n'existe aucune règle d'agrégation de jugements qui satisfait R-universalité, rationalité collective, complétude collective, systématisme et anonymat.*

5. C'est-à-dire lorsque l'on force le résultat à être un singleton.

P	φ_1	φ_2	φ_3	φ_4	φ_5	φ_6
1	1	1	0	1	1	1
2	1	1	1	0	1	1
3	1	1	1	1	0	1
4	1	1	1	1	1	0
5	1	0	1	1	1	1
6	1	0	1	1	1	1

P'	φ_1	φ_2	φ_3	φ_4	φ_5	φ_6
1'	0	1	1	1	1	1
2'	0	1	1	1	1	1
3'	0	1	1	1	1	1
4'	0	1	1	1	1	1
5'	1	0	0	0	0	0
6'	1	0	0	0	0	0

TABLE 2 – Profils P et P'

Ce théorème est très négatif, mais il repose sur des hypothèses très fortes. La première de ces hypothèses est l'hypothèse de complétude sur les individus (R-universalité), qui peut être critiquée, parce qu'on ne peut pas forcément attendre de tous les agents d'avoir une opinion sur toutes les questions possibles. On peut critiquer également la propriété de complétude collective, qui, bien qu'utile pour prendre des décisions, oblige à faire des choix même s'il n'y a aucune raison de choisir une solution plutôt qu'une autre. Ainsi, pour respecter la complétude collective, on va imposer de discriminer des ensembles de jugements, selon des informations supplémentaires qui ne sont pas présentes dans le profil initial, ce qui entre en conflit avec l'anonymat et la neutralité. Supposons par exemple une égalité parfaite (disons sur une question unique φ dans l'agenda, avec 4 votes pour et 4 votes contre), pourquoi et comment faire une distinction entre φ et $\neg\varphi$? On peut se référer à Gärdenfors [2006] pour une critique de la propriété de complétude collective.

La propriété de systématicité est également extrêmement criticable. En effet, elle empêche de considérer l'agrégation de jugements comme un processus d'optimisation permettant d'aboutir à un meilleur compromis, ce qui est souvent attendu. L'exemple suivant illustre cette limitation :

Exemple 1 Soit un agenda X composé des six formules $\varphi_1 = (\neg a \vee \neg b \vee \neg c \vee \neg d \vee \neg e)$, $\varphi_2 = a$, $\varphi_3 = b$, $\varphi_4 = c$, $\varphi_5 = d$, $\varphi_6 = e$. Soient les profils P et P' sur cet agenda, représentés dans la table 2.

Dans le profil (resolu) P , toutes les formules reçoivent une majorité de votes, donc en utilisant un vote majoritaire simple toutes les formules doivent être sélectionnées, ce qui conduit à un ensemble de jugements collectifs incohérent. Donc (au moins) l'une des six formules doit être rejetée par la correspondance d'agrégation de jugements. Il y a une unanimité pour φ_1 , donc il semble raisonnable de sélectionner φ_1 dans le

résultat. Toutes les autres formules sauf φ_2 sont quasi-unanimes (elles reçoivent tous les votes sauf un). La formule la moins supportée est φ_2 , donc le résultat le plus raisonnable semble être $\gamma_P = \{\varphi_1, \neg\varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6\}$.

Soit à présent le profil P' . Le vote majoritaire simple conduit à un ensemble de jugements collectifs cohérent $\gamma_{P'} = \{\neg\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6\}$, qui apparaît donc comme la solution à retenir. Ainsi, bien que les jugements individuels soient les mêmes dans les deux profils pour φ_2 , dans le résultat attendu pour P $\neg\varphi_2$ est sélectionné, alors que pour P' φ_2 est sélectionné. Comme φ_2 reçoit les mêmes votes pour et contre dans les deux profils, aucune méthode d'agrégation de jugements satisfaisant la systématicité ne peut faire une telle distinction.

Cet exemple montre clairement que les jugements individuels sur cette question ne peuvent être considérés indépendamment de ceux des autres. Cette demande est d'ailleurs en conflit avec d'autres propriétés que nous avons donné, ce qui nous donne le théorème d'impossibilité suivant :

Proposition 2 Il n'y a pas de correspondance d'agrégation de jugements qui satisfait la R-universalité, l'unanimité, la préservation de la majorité, la R-swap optimalité, et la systématicité.

Comme l'universalité implique la R-universalité et que la swap optimalité implique la R-swap optimalité, nous obtenons immédiatement le corollaire :

Corollaire 1 Il n'y a pas de correspondance d'agrégation de jugements qui satisfait l'universalité, l'unanimité, la préservation de la majorité, la swap optimalité, et la systématicité.

Pour cette raison, la systématique est laissée de côté dans la suite, et nous cherchons des correspondances d'agrégation de jugements qui satisfont l'universalité, la rationalité collective, l'anonymat, la neutralité, la préservation de la majorité, l'unanimité, et la swap optimalité (ou R-swap). Dans la section suivante, nous montrons que de telles correspondances existent.

4 Agrégation basée sur le nombre de votes

Nous présentons à présent une famille générale de correspondances basées sur le nombre de votes pour laquelle la sélection d'ensembles cohérents de jugements collectifs est basée seulement sur le nombre de votes obtenu par leurs éléments. Dans ce but, nous introduisons d'abord quelques définitions :

Définition 2 *Etant donné un agenda $X = \{\varphi_1, \dots, \varphi_m\}$, une solution potentielle $M = \langle \alpha_1, \dots, \alpha_m \rangle$ est un vecteur tel que chaque α_k est soit φ_k ou $\neg\varphi_k$, et $\bigwedge_{i=1}^m \alpha_i$ est cohérent. M_X est l'ensemble de toutes les solutions potentielles de X .*

Clairement, il existe une équivalence directe entre les solutions potentielles et les ensembles cohérents et résolus de jugements collectifs. On peut remarquer que M_X n'est jamais vide.

Définition 3 *Etant donné un profil P sur X , l'ensemble support de P est l'ensemble $ss(P) = \{\langle \varphi_i, g_P^+(\varphi_i), g_P^-(\varphi_i) \rangle \mid \varphi_i \in X\}$.*

Ainsi, l'ensemble support du profil P contient les fonctions d'évaluation pour et contre chaque formule de X .

Définition 4 *Une correspondance d'agrégation de jugements basée sur le nombre de votes (SAC) δ est une correspondance d'agrégation de jugements tel qu'il existe une fonction f telle que : $\delta(P) = f(ss(P))$.*

La définition des SAC demande donc simplement que le résultat de l'agrégation ne soit déterminé que par la valeur des votes pour ou contre chaque proposition, à l'exclusion de toute autre information (comme par exemple une classification des formules en prémisses/conclusions).

Clairement, chaque SAC satisfait un certain nombre de propriétés intéressantes :

Proposition 3 *Chaque SAC satisfait l'universalité, la rationalité collective, l'anonymat et la neutralité.*

Bien que la règle d'agrégation majoritaire δ^{maj} définie auparavant n'est pas une SAC puisqu'elle ne satisfait pas la condition de rationalité collective, on peut aisément définir une SAC basée sur elle, et plus généralement, sur les méthodes majoritaires qualifiées (méthodes à quota⁶) :

Définition 5 *Soit q un entier. La règle d'agrégation q -quota δ^{Q_q} est définie comme suit. Pour tout agenda X , et tout profil P sur X , on a $\delta^{Q_q}(P) = \{\gamma_P^{Q_q}\}$, où pour tout $\varphi \in X$:*

$$\gamma_P^{Q_q}(\varphi) = \begin{cases} 1 & \text{si } g_P(\varphi) > q \\ 0 & \text{si } g_P(\varphi) < q \\ \star & \text{sinon} \end{cases}$$

La SAC $\delta_{sac}^{Q_q}$ associée à δ^{Q_q} est définie comme $\delta_{sac}^{Q_q}(P) = \delta^{Q_q}(P)$ si $\gamma_P^{Q_q}$ est cohérent, et $\delta_{sac}^{Q_q}(P) = M_X$ sinon.

Les méthodes $\delta_{sac}^{Q_q}$ posent clairement un problème lorsque $\gamma_P^{Q_q}$ est incohérent.

Nous introduisons maintenant une famille spécifique de SAC qui évite ce problème, la famille des méthodes majoritaires ordonnées. Pour ces correspondances, la sélection des solutions potentielles est basée sur leur score agrégé.

Définition 6 *Le vecteur de score d'une solution potentielle $M = \langle \alpha_1, \dots, \alpha_m \rangle$ étant donné P est le vecteur des votes dans P en faveur des questions sélectionnées par cette solution potentielle :*

$$s(M) = (g_P^+(\alpha_1), \dots, g_P^+(\alpha_m))$$

Définition 7 *Une fonction d'agrégation est une application⁷ \oplus de \mathbb{R}^m dans \mathbb{R} , qui satisfait :*

- Si $x_i \geq x'_i$, alors $\oplus(x_1, \dots, x_i, \dots, x_m) \geq \oplus(x_1, \dots, x'_i, \dots, x_m)$ (**monotonie**)

6. δ^{maj} est la méthode d'agrégation 0-quota δ^{Q_0} .

7. Rigoureusement, une famille d'applications, une pour chaque m .

- $\oplus(0, \dots, 0) = 0$ **(minimalité)**
- Si σ est une permutation sur $\{1, \dots, m\}$,
alors $\oplus(x_1, \dots, x_m) = \oplus(x_{\sigma(1)}, \dots, x_{\sigma(m)})$ **(symétrie)**

Des propriétés supplémentaires peuvent aussi être considérées pour la fonction \oplus :

- Si $x_i > x'_i$, alors
 $\oplus(x_1, \dots, x_i, \dots, x_m) > \oplus(x_1, \dots, x'_i, \dots, x_m)$ **(monotonie stricte)**
- Si $\oplus(x_1, \dots, x_m, z) \geq \oplus(y_1, \dots, y_m, z)$
alors $\oplus(x_1, \dots, x_m) \geq \oplus(y_1, \dots, y_m)$ **(décomposition)**
- Si $\forall i, z > y_i$ alors
 $\oplus(z, x_1, \dots, x_k) > \oplus(y_1, \dots, y_{k+1})$ **(préférence stricte)**

Définition 8 Une méthode d'agrégation de jugements majoritaire ordonnée $\delta^{RM\oplus}$ associée à chaque profil P sur l'agenda X l'ensemble des solutions potentielles ayant le plus grand vecteur de score selon la fonction d'agrégation \oplus , i.e., $\delta^{RM\oplus}(P) = \{M \in M_X \mid \nexists M' \in M_X \text{ t.q. } \oplus(s(M')) > \oplus(s(M))\}$.

Comme attendu, on a :

Proposition 4 Toute méthode d'agrégation de jugements majoritaire ordonnée est une SAC.

De nombreuses méthodes d'agrégation de jugements de la littérature appartiennent à la famille des méthodes d'agrégation de jugements majoritaires ordonnées. Ainsi, la règle du sous-agenda de poids maximal (*maxweight subagenda rule*) introduite dans Lang *et al.* [2011] est une méthode majoritaire ordonnée avec $\oplus = \Sigma$. Une autre méthode définie dans Lang *et al.* [2011], l'opérateur δ^{RA} , n'appartient pas directement à cette famille mais $\delta^{RMleximax}$ (qui est membre de cette famille) raffine δ^{RA} . D'autres fonctions d'agrégation ont du sens pour définir de nouvelles méthodes majoritaires ordonnées. On peut citer par exemple *max*, *min*, *leximax*, *leximin*, Σ^n (la somme des puissances $n^{ièmes}$), etc.

Introduisons à présent une dernière fonction d'agrégation, qui va nous permettre de définir une méthode d'agrégation de jugements majoritaire ordonnée intéressante.

Définition 9 – Soit q un nombre positif,
 $maj_q(x_1, \dots, x_n) = |\{x_i \mid x_i > q \text{ et } i \in \{1, \dots, n\}\}|$
– On note $Maj = maj_{\lfloor \frac{P}{2} \rfloor}$.

Les méthodes majoritaires ordonnées δ^{RMmaj_q} sélectionnent les solutions potentielles ayant le score agrégé le plus élevé, où le score d'une solution est le nombre de formules qui sont supportées par une majorité (dépendante de la valeur de q). Ce qui est intéressant, c'est que δ^{RMMaj} est différent de $\delta^{RMleximax}$, $\delta^{RMleximin}$, et $\delta^{RM\Sigma}$:

Exemple 2 Soit $X = \{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$ un agenda avec $\varphi_1 = a$, $\varphi_2 = b$, $\varphi_3 = \{(a \vee b) \wedge c\}$ et $\varphi_4 = (\neg a \vee \neg b) \wedge d$, et soit le profil P sur X donné par la table 3.

	φ_1	φ_2	φ_3	φ_4
1	1	0	1	1
2	1	0	0	1
3	1	1	0	0
4	1	1	1	0
5	1	1	0	0
6	0	1	0	1
7	0	1	1	1

TABLE 3 – Profil P

M_1, M_2, M_3 ci-dessous sont des solutions potentielles étant donné P :

- $M_1 = \langle \varphi_1, \varphi_2, \varphi_3, \neg\varphi_4 \rangle$,
et $s(M_1) = (5, 5, 3, 3)$
- $M_2 = \langle \varphi_1, \neg\varphi_2, \neg\varphi_3, \varphi_4 \rangle$,
et $s(M_2) = (5, 2, 4, 4)$
- $M_3 = \langle \neg\varphi_1, \varphi_2, \neg\varphi_3, \varphi_4 \rangle$,
et $s(M_3) = (2, 5, 4, 4)$

M_2 et M_3 sont strictement préférées à M_1 lorsque *Maj* est la fonction d'agrégation puisqu'elles contiennent 3 formules majoritaires alors que M_1 n'en contient que 2. Une telle conclusion ne peut être obtenue si *leximax*, *leximin* ou Σ sont utilisés : $\delta^{RMleximax}$, $\delta^{RMleximin}$, et $\delta^{RM\Sigma}$ sélectionnent M_1 .

Intéressons nous maintenant aux propriétés de rationalité satisfaites par les méthodes majoritaires ordonnées $\delta_{RM\oplus}$:

Proposition 5

- $\delta_{RM_{\oplus}}$ satisfait l'universalité, la rationalité collective, l'anonymat et la neutralité.
- Si \oplus satisfait la monotonie stricte, alors $\delta_{RM_{\oplus}}$ satisfait la préservation de la majorité.
- Si \oplus satisfait la préférence stricte et la décomposition, alors $\delta_{RM_{\oplus}}$ satisfait l'unanimité et la swap optimalité.

Il n'est guère surprenant que la propriété de systématique ne soit pas satisfaite en général par $\delta_{RM_{\oplus}}$ (c'est une conséquence du théorème 1 et de la proposition 5). Cependant, les propriétés restantes peuvent être satisfaites conjointement, ce qui donne un « théorème de possibilité » :

Théorème 1 $\delta_{RM_{leximax}}$ satisfait l'universalité, la préservation de la majorité, l'anonymat, l'unanimité, la rationalité collective, la neutralité et la swap optimalité.

La proposition suivante illustre l'importance de la condition de complétude sur le profil initial :

Proposition 6 Si P est un profil résolu, alors $\delta_{RM_{leximax}}(P) = \delta_{RM_{leximin}}(P)$.

Pour des profils quelconques, $\delta_{RM_{leximax}}$ et $\delta_{RM_{leximin}}$ donnent (en général) des résultats différents. Au passage, cela montre encore que l'hypothèse d'absence d'abstention n'est pas une simplification technique sans conséquence ; en effet, elle entraîne que des méthodes différentes coïncident si des ensembles de jugements résolus sont considérés.

Des résultats additionnels sont obtenus pour des méthodes majoritaires ordonnées spécifiques :

Proposition 7

- $\delta_{RM_{\Sigma}}$ and $\delta_{RM_{Maj}}$ satisfait l'universalité, l'anonymat, la neutralité, la préservation de la majorité, la rationalité collective et la R-swap optimalité. Aucune d'entre elles ne satisfait l'unanimité ou la swap optimalité.
- $\delta_{RM_{leximin}}$ satisfait l'universalité, l'anonymat, la neutralité, la préservation de la majorité, la rationalité collective, l'unanimité et la R-swap optimalité. Elle ne satisfait pas la swap optimalité.

5 Fiabilité des ensembles de jugements

Parallèlement aux propriétés de rationalité, un autre critère de sélection d'une méthode d'agrégation de jugements est sa capacité à déterminer la vérité. Cette capacité est évaluée comme la probabilité de choisir le « vrai » ensemble de jugements collectifs, en supposant que les agents sont fiables, i.e., chaque agent a plus de chance de voter la une réponse qui est la bonne que de s'abstenir ou voter la une mauvaise réponse. Formellement, supposons qu'il y a un vrai (juste) ensemble de jugements γ^* sur X , i.e., $\gamma^*(\varphi_k) = 1$ si φ_k est vrai, et $= 0$ sinon, un agent i est *fiable* si pour tout $\varphi_k \in X$, $P(\gamma_i(\varphi_k) = \gamma^*(\varphi_k)) > P(\gamma_i(\varphi_k) \neq \gamma^*(\varphi_k))$.

Il est d'usage dans le problème de la recherche de vérité (Condorcet [1785]; Everaere *et al.* [2010]; List and Goodin [2001]) de faire certaines hypothèses. On suppose que les n agents prennent leurs décisions indépendamment des autres (hypothèse d'indépendance individuelle) et indépendamment des décisions prises sur les autres questions (hypothèse d'indépendance sur les questions). De plus, pour simplifier les calculs, une hypothèse d'homogénéité est souvent faite, supposant que la probabilité qu'un agent fasse le bon choix sur une question est la même, quels que soient l'agent et la question, i.e., $\forall i \in \{1, \dots, n\}, \forall k \in \{1, \dots, m\}, P(\gamma_i(\varphi_k) = \gamma^*(\varphi_k)) = p$.

Ces hypothèses sont celles du théorème du jury de Condorcet (Condorcet [1785]; List and Goodin [2001]) pour le vote. Ce théorème établit que pour une question donnée φ_k , si les individus sont fiables, homogènes et indépendants, alors le jugement le plus probable sur φ_k est celui qui est supporté par une majorité d'individus ; et si la taille du profil tend vers l'infini, alors la probabilité de déterminer le vrai jugement sur la question tend vers 1. Le théorème du jury de Condorcet est l'une des justifications de nos méthodes majoritaires ordonnées ; en effet, les hypothèses qui sous-tendent ce théorème impliquent que plus une réponse est supportée, plus elle est sûre.

Cela établi, notre but est de déterminer dans quelle mesure les critères de recherche de la vérité peuvent discriminer les méthodes majoritaires ordonnées. Afin de répondre à cette question, on considère le problème comme une estimation du maximum de vraisemblance : le but est de déterminer les méthodes d'agrégation de

jugements qui expliquent au mieux les jugements individuels observés.

Proposition 8 *Sous les hypothèses d'indépendance individuelle, d'homogénéité, de fiabilité et d'indépendance entre questions, la méthode d'agrégation de jugements donnée par l'estimation du maximum de vraisemblance sur le profil P sur X est δ_{RM_Σ} .*

Ainsi δ_{RM_Σ} est la meilleure méthode parmi les méthodes d'agrégation de jugements (pas seulement parmi les méthodes majoritaires ordonnées) pour identifier la vraie solution.

Comme nous l'avons déjà dit, les hypothèses retenues sont tout à fait standard lorsque des méthodes de vote sont considérées (Young [2003]). L'hypothèse la plus discutable est sans doute celle sur l'indépendance entre les questions, qui n'est satisfaite que rarement :

Exemple 3 *Soit deux formules, $\varphi_1 = a \wedge b$ et $\varphi_2 = \neg a$. Si un agent vote pour φ_1 , on sait qu'il ne peut supporter φ_2 (puisque φ_2 et φ_1 sont incohérents). Donc $P(g_P^+(\varphi_1) = k_1 \wedge g_P^+(\varphi_2) = k_2)$ est différent de $P(g_P^+(\varphi_1) = k_1) \cdot P(g_P^+(\varphi_2) = k_2)$.*

Dans ce cas, l'hypothèse d'indépendance entre questions n'est pas satisfaite, donc δ_{RM_Σ} n'est plus nécessairement le meilleur choix comme méthode d'agrégation de jugements⁸. Ainsi, nous allons nous passer de cette hypothèse d'indépendance entre questions, tout en continuant de considérer l'estimation du maximum de vraisemblance pour discriminer les méthodes majoritaires ordonnées. On peut relever que même si cette hypothèse d'indépendance n'est pas faite, une réponse à une question donnée est d'autant plus sûre que le nombre de votes qu'elle a reçus est élevé. Cela signifie que la recherche de la vérité peut être faite question par question, le problème étant ensuite de prendre des décisions basées sur un vecteur de formules plus ou moins supportées (et donc plus ou moins probables).

Intuitivement, on peut imaginer trois politiques naturelles pour choisir des ensembles de jugements : ceux qui contiennent le plus possible de

formules supportées par une majorité, ou ceux qui contiennent le plus possible de formules les plus probables, ou enfin ceux qui contiennent le moins possible de formules les moins probables. Formellement :

Définition 10 *Une méthode d'agrégation de jugements δ satisfait la **Meilleure/Card/Pire** politique ssi, pour tout profil P , toute les solutions potentielles de $\delta(P)$ sont préférées à toutes les autres solutions potentielles par rapport au critère **Meilleure/Card/Pire**.*

Card *Une solution potentielle M est préférée à une solution potentielle M' pour le critère **Card** ssi $|\{\alpha_k \in M \mid g_P^+(\alpha_k) > n/2\}| > |\{\alpha_k \in M' \mid g_P^+(\alpha_k) > n/2\}|$.*

Meilleure *Une solution potentielle M est préférée à une solution potentielle M' pour le critère **Meilleure** ssi $\exists \alpha_k \in M$ s.t. $\forall \alpha_l \in M', g_P^+(\alpha_k) > g_P^+(\alpha_l)$.*

Pire *Une solution potentielle M est préférée à une solution potentielle M' pour le critère **Pire** ssi $\exists \alpha_l \in M'$ s.t. $\forall \alpha_k \in M, g_P^+(\alpha_k) > g_P^+(\alpha_l)$.*

Les ensembles de jugements préférés selon les trois critères ne coïncident pas en général. Selon le critère **Card** les ensembles de jugements qui contiennent un nombre maximal de formules supportées par une majorité d'agents, et donc un nombre maximal de réponses probables, sont choisis. Cependant cela n'empêche pas de sélectionner des ensembles de jugements contenant des formules ayant un très faible support, et de ne pas sélectionner des réponses très supportées. **Pire**, lorsque le nombre d'abstentions est élevé, il est possible qu'aucune réponse ne soit supportée par une majorité d'agents. Dans ce cas, **Card** n'est pas suffisamment discriminant. Les choix alternatifs sont donnés par le critère **Meilleure** et le critère **Pire**, qui évaluent respectivement la vraisemblance d'un ensemble de jugements comme le meilleur (resp. le pire) support d'une formule de cet ensemble. Ainsi **Meilleure** (resp. **Pire**) correspond à une évaluation optimiste (resp. pessimiste) d'un ensemble de jugements. La proposition suivante montre que pour chaque politique parmi **Meilleure/Card/Pire**, il existe une correspondance majoritaire ordonnée qui la satisfait :

Proposition 9

- $\delta_{RM_{Maj}}$ satisfait la politique **Card**
- $\delta_{RM_{leximax}}$ satisfait la politique **Meilleure**
- $\delta_{RM_{leximin}}$ satisfait la politique **Pire**

8. On peut noter qu'une hypothèse similaire sur l'indépendance entre questions est faite pour utiliser l'estimation du maximum de vraisemblance afin de trouver le meilleur ordre par des méthodes d'agrégation en théorie du vote, en dépit du fait que des problèmes similaires se posent (voir par exemple Young [2003]).

6 Conclusion

Nous avons proposé un ensemble de résultats sur le problème de l'agrégation de jugements, en particulier :

- Une discussion des propriétés de rationalité de l'agrégation de jugements, en particulier une critique de la propriété de systématisme, un appui à la propriété de neutralité, et l'introduction d'une nouvelle propriété (swap optimalité).
- Un théorème d'impossibilité.
- Une nouvelle famille de méthodes d'agrégation de jugements, les correspondances basées sur le nombre de votes, pour qui la sélection d'ensembles cohérents de jugements collectifs est basée seulement sur le nombre de votes obtenus par leurs éléments. Nous nous sommes particulièrement intéressés à un sous-ensemble de ces méthodes, et avons montré qu'elles satisfont la plupart des propriétés de rationalité attendues.
- Une justification liée à la vraisemblance des résultats fournis par certains opérateurs de cette famille a également été donnée.

Ce travail ouvre de nombreuses perspectives pour des recherches à venir. L'une d'entre elles consiste à considérer d'autres postulats de rationalité pour l'agrégation de jugements (comme la réactivité et la monotonie, considérées dans Grossi and Pigozzi [2012]) et aussi d'identifier les ensembles maximaux de propriétés qui peuvent être satisfaites ensemble, afin de passer de théorèmes d'impossibilité à des « théorèmes de possibilité ». Il serait aussi intéressant d'étudier la complexité calculatoire des SAC, dans le cas général et pour certaines méthodes spécifiques.

Remerciements

Les auteurs remercient Srdjan Vesic pour ses nombreuses remarques sur la version préliminaire de ce travail qui ont permis d'améliorer cet article.

Références

- K. J. Arrow. *Social Choice and Individual Values*. Wiley, New York, second edition, 1963.
- L. Bovens and W. Rabinowicz. Democratic answers to complex questions. *Synthese*, 150 :131–153, 2006.

- Marquis de Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie Royale, Paris, 1785.
- F. Dietrich and C. List. Judgment aggregation by quota rules : Majority voting generalized. *Journal of Theoretical Politics*, 4(19) :391–424, 2007.
- F. Dietrich and P. Mongin. The premiss-based approach to judgment aggregation. *Journal of Economic Theory*, 145(2) :562–582, 2010.
- F. Dietrich. Judgment aggregation : (im)possibility theorems. *Economic Theory*, 126 :286–298, 2006.
- P. Everaere, S. Konieczny, and P. Marquis. The epistemic view of belief merging : Can we track the truth ? In *Proceedings of the Nineteenth European Conference on Artificial Intelligence (ECAI'10)*, pages 621–626, 2010.
- P. Gärdenfors. An Arrow-like theorem for voting with logical consequences. *Economics and Philosophy*, 22(2) :181–19, 2006.
- Davide Grossi and Gabriella Pigozzi. Introduction to judgment aggregation. In *Lectures on Logic and Computation (ESSLLI 2011)*, volume 7388 of *Lecture Notes in Computer Science*, pages 160–209. Springer, 2012.
- J. Lang, G. Pigozzi, M. Slavkovik, and L. Van Der Torre. Judgment aggregation rules based on minimization. In *Proceedings of the Thirteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK'11)*, pages 238–246, 2011.
- C. List and R. Goodin. Epistemic democracy : Generalizing the Condorcet jury theorem. *Journal of Political Philosophy*, 9(3) :277–306, 2001.
- C. List and P. Pettit. Aggregating sets of judgments. an impossibility result. *Economics and Philosophy*, 18 :89–110, 2002.
- C. List and C. Puppe. Judgment aggregation : A survey. In P. Anand, C. Paul, and P. Pattanaik, editors, *The Handbook of Rational and Social Choice*. Oxford University Press, 2009.
- C. List. A model of path-dependence in decisions over multiple propositions. *American Political Science Review*, 98(3) :495–513, 2004.

- P. Pettit. Deliberative democracy and the discursive dilemma. *Philosophical Issues*, 11 :268–299, 2001.
- G. Pigozzi, M. Slavkovic, and L. van der Torre. A complete conclusion-based procedure for judgment aggregation. In *Proceedings of the First International Conference on Algorithmic Decision Theory (ADT'09)*, volume 5783 of *Lecture Notes in Computer Science*, pages 1–13. Springer, 2009.
- G. Pigozzi. Belief merging and the discursive dilemma : an argument-based account to paradoxes of judgment aggregation. *Synthese*, 152(2) :285–298, 2006.
- M. Slavkovic. *Judgment Aggregation in Multiagent Systems*. PhD thesis, Université du Luxembourg, 2012.
- H. P. Young. Condorcet's theory of voting. *The American Political Science Review*, 82(4) :1231–1244, 2003.