

Rapport sur le manuscrit de thèse présenté par François Durand

Jérôme Lang
LAMSADE — CNRS & Université Paris-Dauphine

François Durand (FD) a soumis un manuscrit de thèse intitulé «Vers des modes de scrutin moins manipulables». Le sujet de cette thèse se situe dans une large mesure en *choix social computationnel*, une discipline de recherche en plein essor qui est à la croisée entre l'économie mathématique (et plus précisément, le choix social) et l'informatique. Le fil conducteur de la thèse est l'étude qualitative et quantitative de la manipulabilité des règles de vote. FD définit pour cela un cadre original et général permettant d'inclure (entre autres) des règles de vote dont l'input est cardinal. Il établit ensuite un nombre de résultats théoriques, dont l'aboutissement est que sous une hypothèse réaliste sur la distribution des votes (plus faible que l'indépendance probabiliste), la recherche d'une règle de vote satisfaisante et de manipulabilité minimale peut se faire à l'intérieur des règles Condorcet-cohérentes et dont l'input est ordinal. FD décrit ensuite une plate-forme informatique, qu'il a développée, qui permet d'étudier de nombreux aspects de la manipulabilité des règles de vote. Il présente un nombre de résultats expérimentaux, d'abord en utilisant des données générées aléatoirement, selon plusieurs classes de distribution, puis en utilisant des données provenant d'élections réelles. L'aboutissement de cette partie est la découverte qu'en moyenne, certaines des règles de vote usuelles (ou des variantes de celles-ci), sont bien moins manipulables que d'autres, pourtant plus largement utilisées. La dernière partie de la thèse définit un cadre algorithmique général pour la recherche de règles de manipulabilité minimale pour des nombres de votants (n) et de candidats (m) fixés, et donne des résultats concrets pour certaines (petites) valeurs de n et m . Avant de présenter les résultats dans le détail, je ne peux m'empêcher de dire que chacune des deux moitiés de cette thèse aurait largement suffi à l'obtention d'un doctorat.

Le manuscrit de thèse de FD se compose d'une introduction, de dix chapitres regroupés en deux parties, d'une conclusion, et d'annexes.¹

L'introduction place le travail à venir dans le cadre du choix social. Pour cela, FD rappelle les principaux objectifs et difficultés de la théorie du vote, en insistant sur le théorème de Gibbard et Satterthwaite. Ce chapitre n'est pas

1. Ces annexes présentent des résultats significatifs, mais comme la manipulabilité n'y joue pas un rôle central, FD n'a pas souhaité les inclure dans la thèse. Cependant, l'annexe C, qui analyse l'application de systèmes de vote aux réseaux de télécommunications, qui est une contribution intéressante et pertinente en regard du thème central de la thèse, aurait pu figurer dans le document principal.

seulement un état de l’art objectif : FD y fait rentrer une vision personnelle très pertinente, défendant en partie la manipulation (c’est-à-dire l’attitude stratégique des votants), tout en condamnant la manipulabilité, qui peut aboutir à des résultats socialement indésirables.

La première partie de la thèse (chapitres 1 à 5) rassemble un ensemble de contributions de nature théorique.

Dans le chapitre 1, FD présente un formalisme très général pour les “systèmes” de vote (qu’il distingue des règles de vote, ou fonctions de choix social, par le fait que leur input est plus général). Il définit la notion d’espace électoral, où un état consiste en la donnée, pour chaque votant, d’une relation binaire sur les candidats (qui peut être une relation de préférence, mais pas nécessairement) et d’informations supplémentaires (qui peuvent être des évaluations cardinales ou binaires des candidats, entre autres). Cette notion est illustrée par plusieurs exemples bien choisis. Il donne ensuite la définition d’un espace électoral probabilisé. Il définit ensuite les systèmes de vote basés sur l’état, et plus généralement, les systèmes de vote (généraux), basés sur la notion plus générale de stratégie, et qui permettent par exemple de définir des protocoles à tours réels, où les votants sont appelés à s’exprimer en plusieurs étapes et pas nécessairement de façon cohérente. Après un détour identifiant les conditions nécessaires et suffisantes sur les nombres de votants et de candidats pour qu’il existe un système de vote anonyme et neutre dans l’espace électoral des ordres stricts totaux (généralisant un résultat de Moulin concernant les systèmes anonymes, neutres et Pareto-efficaces), FD définit la notion de manipulabilité dans les espaces électoraux, et montre que si l’objectif est de chercher des systèmes peu manipulables, on peut sans perte de généralité se concentrer aux systèmes de vote basés sur l’état. Il termine le chapitre en rappelant des notions habituelles (notamment autour de la notion de vainqueur de Condorcet et ses versions affaiblies) et introduit les règles de vote dont il aura besoin dans le reste du document (dont certaines ont été peu, voire pas, étudiées jusqu’à présent).²

Le chapitre 2 consiste en un ensemble de variations autour de l’un des résultats-clé de la thèse. FD commence par définir la *condorcification* f^* d’un système de vote f , qui retourne le vainqueur de Condorcet s’il en existe un, et le résultat selon f sinon, puis le critère de *coalition majoritaire informée*, qui est satisfait par un système de vote si toute coalition comprenant une majorité stricte de votants est capable, en connaissant les votes des autres votants, de décider du résultat. Puis il montre le résultat-clé du chapitre : si un système de vote satisfait le critère de la coalition majoritaire informée, alors sa condorcification est au plus aussi manipulable (au sens ensembliste) que lui. Il examine la possibilité de rendre un système encore moins manipulable en remplaçant le test d’existence d’un vainqueur de Condorcet par celle d’un vainqueur de Condorcet faible ou relatif³, et montre que cette voie mène en général à l’échec, sauf dans

2. Il manque au «zoo» présenté en fin de chapitre quelques systèmes communs, notamment la règle de Copeland, dont on peut se demander pourquoi ils ont été laissées en-dehors du champ de la thèse.

3. Les relations de préférences considérées dans cette thèse peuvent contenir des indifférences.

le cas particulier de certains systèmes de vote (notamment la pluralité à deux tours). FD définit ensuite la notion de *vainqueur de Condorcet résistant* comme (je simplifie) un candidat c pour lequel, pour toute paire d'autres candidats d et e , il existe une majorité absolue de votants qui préfère c simultanément à d et e ; il définit ensuite un système de vote Condorcet-résistant comme un qui élit le vainqueur de Condorcet résistant lorsqu'il en existe un (il remarque que si ce critère est plus faible que la Condorcet-cohérence, celles des règles usuelles qui ne sont pas Condorcet-cohérentes ne sont pas non plus Condorcet-résistantes). Cette notion mène à deux résultats très intéressants : (informellement, et sous certaines conditions simples) un système de vote Condorcet-cohérent est manipulable dans un état seulement s'il n'existe pas de vainqueur de Condorcet résistant dans cet état; puis vient la version forte du résultat-clé du début du chapitre : si un système de vote n'est pas Condorcet-résistant, alors sa condorcification est *strictement* moins manipulable (au sens ensembliste) que lui. Une conclusion de ce chapitre est que lorsqu'on cherche un système de vote de manipulabilité minimale, il suffit de restreindre la recherche au sein des systèmes Condorcet-cohérents.

Le chapitre 3 introduit une famille de «critères majoritaires» qu'un système de vote peut satisfaire ou non. Outre le critère de la coalition majoritaire informée, dont j'ai parlé *supra*, celui de la *coalition majoritaire ignorante* exprime que toute coalition majoritaire peut forcer le résultat du vote, sans la connaissance des autres votes, et celui du *bulletin majoritaire* est une version faible du critère (plus simple) exprimant que toute coalition majoritaire peut faire élire un candidat en exprimant des bulletins identiques. FD établit ensuite des liens entre ces critères et l'existence et la nature «condorcetienne» des équilibres de Nash forts du jeu correspondant à un profil de préférences. Par exemple, il montre que sous certaines conditions raisonnables, l'ensemble des systèmes de vote qui garantissent l'existence d'un équilibre de Nash fort pour tout profil possédant un vainqueur de Condorcet est inclus dans l'ensemble des systèmes satisfaisant le critère de la coalition majoritaire ignorante et contient la classe des systèmes vérifiant le critère du bulletin majoritaire. Les preuves sont claires, et les contre-exemples utilisés pour montrer que certaines des inclusions sont strictes sont élégants. Puis FD étudie systématiquement, pour de nombreux systèmes de vote classiques, lesquels de ces critères majoritaires ils satisfont. Certains de ces résultats étaient connus, d'autres sont originaux. Les conditions nécessaires et/ou suffisantes établies pour les règles positionnelles à score ou les règles à élimination successive sont particulièrement intéressantes.

Les chapitres 2 et 3 ne considéraient que des règles anonymes; dans le chapitre 4, FD généralise les notions et résultats obtenus aux situations où les votants ont une importance et des rôles différents. Il reprend de la littérature des jeux simples la notion de famille de coalition gagnante; étant donné une famille de coalitions gagnantes \mathcal{M}_x pour tout candidat x , FD définit un \mathcal{M} -vainqueur de Condorcet (unique sous certaines hypothèses sur \mathcal{M}) comme un candidat c tel que pour tout d , l'ensemble des votants qui préfère c à d est dans \mathcal{M}_c et l'ensemble des votants qui préfère d à c n'est pas dans \mathcal{M}_d . Il définit ensuite naturellement la \mathcal{M} -condorcification de façon similaire à la condorcifi-

cation, remplaçant la condition d’existence d’un vainqueur de Condorcet par celle d’un \mathcal{M} -vainqueur de Condorcet. Les différents critères majoritaires des systèmes de vote sont eux aussi généralisés dans cette direction, et FD généralise le résultat du chapitre 3 en montrant que sous certaines conditions sur \mathcal{M} et f , la \mathcal{M} -condorcification de f est au plus aussi manipulable que f . Ce résultat étant d’autant plus puissant que \mathcal{M} est grande, FD définit la famille maximale d’un système de vote comme la famille des coalitions qui peuvent manipuler de manière informée en faveur d’un candidat, et montre que, sous certaines hypothèses, la condorcification utilisant cette famille maximale est la moins manipulable des condorcifications de ce type. Il termine le chapitre en appliquant la condorcification généralisée à des systèmes de vote anonymes ne vérifiant pas le critère de la coalition majoritaire informée, et puis à des systèmes de vote non anonymes et/ou non neutres, pour lesquels il donne des exemples concrets pertinents.

Dans le chapitre 5, FD formalise et prouve le résultat intuitif qui énonce que les systèmes de vote cardinaux, en laissant plus de liberté d’expression aux votants que les systèmes de vote ordinaux, sont de ce fait plus manipulables. La notion de manipulabilité considérée ici est quantitative : le taux de manipulabilité est la probabilité de l’ensemble des états manipulables, pour une distribution donnée. FD commence par définir une notion de décomposabilité sur les espaces électoraux probabilisés, qui est vérifiée en particulier si les votes sont indépendants (mais qui est bien plus générale que cela) ; puis il montre le résultat-clé de ce chapitre : pour tout espace électoral probabilisé qui vérifie la condition de décomposabilité, et tout système de vote non-ordinal, il existe un système de vote ordinal qui est cohérent avec lui (qui en est une «tranche», selon la terminologie de FD) et dont le taux de manipulabilité est inférieur ou égal au sien. Par exemple, le vote par notation (où chaque votant donne une évaluation numérique à chaque candidat, le candidat maximisant la somme des notes étant vainqueur) est cohérent avec toutes les règles de vote positionnelles à score, et il en existe donc une qui est au plus aussi manipulable que lui. (FD montre par ailleurs qu’il est difficile de renforcer ce résultat : la condition de décomposabilité est indispensable, et remplacer la notion quantitative de manipulabilité par la manipulabilité relative au sens ensembliste ne fonctionne pas non plus.) La combinaison de ce résultat avec le résultat-clé du chapitre 2 est que (sous certaines conditions) la recherche d’un système de vote de manipulabilité minimale peut se faire au sein des systèmes Condorcet-cohérents ordinaux.

La seconde partie de la thèse (chapitres 6 à 9) rassemble un ensemble de contributions essentiellement algorithmiques et expérimentales. FD a déjà identifié des conditions garantissant que certains systèmes de vote sont moins manipulables (qualitativement ou quantitativement) que d’autres. Pour obtenir des arguments plus conclusifs sur des systèmes de vote spécifiques, il propose de recourir à des études empiriques reposant, d’une part, sur la génération aléatoire, et d’autre part sur des votes issues de bases de données réelles. C’est l’objet des chapitres 6 à 9.

Afin d’être en mesure de réaliser cette étude expérimentale, FD a implémenté une plateforme logicielle (programmée en Python) permettant de dé-

terminer le taux de manipulabilité (au sens probabiliste, et pour une culture donnée) de différents systèmes de vote. La description de cette plateforme, et des algorithmes et heuristiques qu'elle utilise, est l'objet du chapitre 6. La première brique consiste à implémenter, pour chaque système, un ou plusieurs algorithme(s) permettant de déterminer si, étant donné un état et un système de vote, il existe une manipulation coalitionnelle (au sens de la définition habituelle ou de certaines de ses variantes) à partir de cet état. L'objectif étant de savoir s'il existe au moins une manipulation, le programme (générique) utilise des heuristiques consistant à chercher d'abord une manipulation en faveur des candidats pour laquelle l'existence d'une telle manipulation est la plus plausible, et à appliquer d'abord des algorithmes incomplets, mais rapides, ce qui est particulièrement pertinent pour les – nombreux – systèmes de vote pour lesquels la manipulation coalitionnelle est NP-difficile ; pour ces problèmes difficiles, FD essaie d'abord les manipulations dites «triviales», puis en cas d'échec, lance un algorithme approché polynomial, puis, au besoin, un algorithme exact. Pour certains systèmes, en particulier certaines variantes de STV, FD a construit des algorithmes originaux ; pour d'autres, il a implémenté des algorithmes existants. Au-delà de son intérêt pour le reste de la thèse, ce chapitre a une grande valeur pour la communauté du choix social (computationnel ou non) : c'est la première plateforme (ouverte, de surcroît) de calcul de manipulations et de manipulabilité. La plateforme permet de générer aléatoirement des votes selon plusieurs familles de cultures, ou de les importer à partir de fichiers externes. Sous ses aspects modestes, le chapitre contient aussi des contributions originales (algorithmes de manipulation coalitionnelle pour certains systèmes de vote).

Dans les chapitres 7 à 9, FD entre dans le coeur des expérimentations. Les chapitres 7 et 8 sont fondés sur des votes générés aléatoirement, selon deux classes de culture : dans le chapitre 7, des cultures dite «sphéroïdales», qui généralisent la culture impartiale (en particulier à des systèmes non ordinaux), et dans le chapitre 8, des cultures dites «basées sur un spectre politique», qui généralisent la culture unimodale («single-peaked»). Dans les deux cas, FD examine l'effet, selon le système de vote, des variations du nombre de candidats, du nombre de votants et en particulier de sa parité, du type de manipulation, et de certains des paramètres de la distribution. (Ces expérimentations s'accompagnent de résultats analytiques de convergence du taux de manipulabilité de la règle veto). Puis il propose une analyse fine des résultats obtenus, accompagnée d'explications convaincantes, et dresse des conclusions. Alors qu'on aurait pu s'attendre à des résultats comparatifs peu conclusifs, c'est l'inverse qui se produit : ces conclusions relativement surprenantes plaident en faveur de certaines règles, et plus ou moins indépendamment des différents paramètres. Tandis que les chapitres 7 et 8 considéraient des données générées aléatoirement, le chapitre 9 considère des données issues d'expériences de vote réelles. Certaines de ces expériences proviennent de la base de données *PrefLib*, d'autres sont originales. Certaines contiennent des évaluations cardinales, d'autres (la plupart) des votes ordinaux. Dans certaines d'entre elles, il y a tout lieu de penser que les préférences exprimées sont sincères. Bien que les données soient de nature très différentes de celles des chapitres 7 et 8, il est assez surprenant de constater que les résultats concordent largement avec ceux des chapitres 7 et 8. (Ils

confirment aussi que l’existence d’un vainqueur de Condorcet est fréquente, et – sur une note moins optimiste – que la manipulabilité d’un profil est fréquente également.) En bref : le système STV⁴ et ses variantes se montrent peu manipulables ; en particulier, la condorcification de STV est moins manipulable que la version standard (cf. chapitre 2), elle-même moins manipulable que sa version avec tours indépendants (cf. chapitre 2), elle-même moins manipulable que les autres systèmes considérés. Comme on peut s’y attendre, les systèmes basées sur des inputs cardinaux, mais aussi la règle de Borda, ont des taux de manipulabilité particulièrement forts.

Les chapitres 1 à 9 nous ont appris beaucoup de choses (formellement et empiriquement) sur la nature des systèmes de vote les moins manipulables (au sens relatif ou quantitatif du terme). Dans le Chapitre 10, FD propose d’aller plus loin et d’identifier *précisément* le(s) système(e) de vote de taux de manipulabilité minimal, au sens de la culture impartiale, pour des petites valeurs du nombre de votants n et de candidats m . Pour cela, il montre que la recherche d’un système de manipulabilité minimale se ramène à la résolution d’un problème d’optimisation combinatoire, avec malheureusement un nombre de variables exponentiel en n et m (malgré l’utilisation de procédés fins limitant le nombre de variables), ce qui, en pratique, limite la résolution de problème à de très petites valeurs de n et m . Cependant, les résultats sont très instructifs, et corroborent en grande partie les conclusions des études empiriques des chapitres 7 à 9. Le travail présenté dans ce chapitre, qui est un véritable travail de choix social computationnel, est d’une grande originalité.

La thèse se termine par une conclusion très détaillée, qui synthétise les conclusions des différents chapitres, et propose de nouvelles pistes de recherche, chapitre par chapitre.

La thèse de François Durand est impressionnante à plusieurs égards : la quantité, la variété et l’originalité des résultats, leur importance et leur impact attendu, et la forme du document.

Quand on parle de la *variété* des résultats d’une thèse, on peut vouloir signifier deux choses différentes : certaines thèses, que j’appelle «technocentrées» appliquent une même technique à plusieurs problèmes ; d’autres, bien plus rares, que j’appelle «centrées sur un problème» s’attachent à étudier toutes les facettes d’un problème en utilisant toutes les techniques pertinentes. Les secondes sont bien plus intéressantes, ont un fil conducteur plus homogène, et sont aussi bien plus difficiles à réaliser, en raison du nombre de techniques différentes à maîtriser : *c’est exactement ce que la thèse de FD réalise* ; les contributions qui y sont présentées font appel aux mathématiques discrètes, aux probabilités, à l’optimisation combinatoire, à l’algorithmique, à la programmation, et à l’expérimentation sur des données réelles. Thématiquement, elles portent sur le choix social classique (chapitres 1 à 5), le choix social computationnel (chapitres 6 et 10), et le choix social expérimental (chapitres 7, 8, et surtout 9). Cette thèse est donc tout-à-fait novatrice et originale dans la façon d’aborder un problème

4. FD utilise la dénomination «VTI» issue de la traduction française de *Single Transferable Vote*.

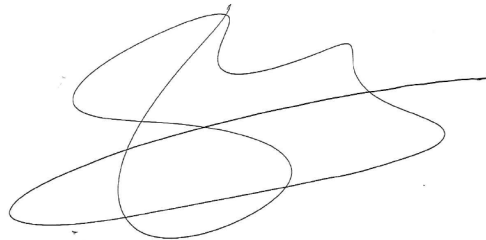
sous l'angle de différentes disciplines ; dans mon domaine de recherche, qui est celui du choix social computationnel, c'est la première thèse de ce type ; elle fera date, et j'espère qu'elle fera des émules.

L'importance des résultats est indiscutable. C'est une thèse qui propose et défend, justement, une *thèse* : on ne doit pas s'arrêter au constat que toutes les règles de vote raisonnables sont manipulables, mais faire une étude comparative et quantitative de leur manipulabilité, afin de voir si on peut dégager des conclusions claires ; et justement, FD dégage une conclusion claire : parmi toutes les règles étudiées, c'est la règle STV et certaines de ses variantes (en particulier, la version condorcifiée de STV, qui est le grand vainqueur de la thèse, si je puis dire) qui sont les moins manipulables ; le fait que STV soit déjà utilisé, sans que cela soulève de problème important, dans des élections politiques, achève de plaider en la faveur de cette classe de règles. Certes, la manipulabilité n'est pas tout, et je m'attendais peut-être à voir un peu plus de discussion sur le compromis à faire entre manipulabilité et d'autres propriétés, puisque les règles qui réussissent bien du point de vue du taux de manipulabilité sont aussi parmi celles qui violent un certain nombre de propriétés importantes (comme la monotonie). Une autre conclusion importante de la thèse est que les méthodes cardinales sont toujours plus manipulables que les méthodes ordinales, en raison de la plus grande liberté d'action laissée aux votants. Enfin, la thèse plaide en faveur du critère de la Condorcet-cohérence, et de la «condorcification» des règles non Condorcet-cohérentes. Toutes ces conclusions auront certainement un fort impact sur la communauté académique, et mon souhait est qu'elles puissent en avoir un au-delà.

Enfin, cette thèse est remarquablement bien construite et bien écrite. La présence d'exemples pertinents et judicieusement placés, la clarté des preuves et des explications, l'analyse des résultats, et la progression finement pensée et planifiée des arguments étayant les thèses principales du document, rend sa lecture passionnante.

Au vu de tout ce que je viens d'écrire : il ne fait absolument aucun doute que le manuscrit de thèse de François Durand contient bien plus qu'il ne faut pour recevoir le titre de docteur de l'Université de Paris-6.

Paris, le 3 septembre 2015
Jérôme Lang

A handwritten signature in black ink, consisting of several overlapping loops and a long horizontal stroke extending to the right.