

INTERACTIVE, STRATEGIC,
AND EXPRESSIVE SOCIAL CHOICE

UMBERTO GRANDI

Umberto Grandi: *Interactive, Strategic, and Expressive Social Choice*, Université Toulouse 1 Capitole, 2020

ABSTRACT

This thesis presents research I conducted since obtaining my PhD in computational social choice, an interdisciplinary field of research which lies between artificial intelligence (in particular multi-agent systems, knowledge representation and reasoning), economics, and political science, and whose aim is the analysis of collective decision-making processes with tools from theoretical computer science.

My work aims at enriching social choice models under three aspects, corresponding to the three main chapters of this document: (a) take into account the interactions between voters with the introduction of an influence network or a communication network, (b) model the strategic behaviour of voters, (c) increase the expressiveness of the voting ballots whilst keeping the computational complexity of the decision process tractable. The introductory chapter presents the main lines of research of my work, gives a summary of the chapters, and defines the basic concepts for understanding the results that follow. The technical sections are supplemented with examples to quickly grasp the interest and scope of the results presented.

In Chapter 2 I augment the classic model of social choice with a network that connects voters and that can represent a network of influence or of communication between the agents. A large part of the chapter is dedicated to study models of opinion diffusion on a network which are adapted to voting situations and to proposing a model of strategic opinion diffusion. Two problems of algorithmic design are then presented. The first which enriches the system of liquid democracy with multi-agent and ranked delegations, and the second which proposes a notion of collective evaluation which is more resistant to bribery thanks to the use of the underlying network.

Chapter 3 introduces the notion of a voting game, in which the strategic dimension of voters' behaviour is taken into account. First, I define and study a class of games named after the famous Gibbard-Satterthwaite theorem, which allow to zoom in on a subset of strategic agents with restricted actions. Next, in a voting situation on several binary questions, I show how the use of pre-vote negotiations can be used as a selector of desirable Nash equilibria. The chapter concludes with the presentation of several results on iterative voting, using techniques from multi-agent simulation, reinforcement learning, and experimental economics.

Chapter 4 focuses on the expressiveness of voters, especially in the case of combinatorial voting, when the set of alternatives is of exponential size. Several approaches are presented, developing on a motivating example of planning a group activity. I begin with a detailed

comparison of several judgment aggregation frameworks in terms of succinctness (i.e., the size of the representation of a problem). Then, I present a voting system based on the elicitation of propositional goals. Finally, I analyse the axiomatic properties and the computational complexity of three voting rules inspired by the majority rule. Increasing further in expressiveness, I then study the aggregation of incomplete conditional preferences and I define a voting rule for the analysis of a collective sentiment enriched with the extraction of comparative preferences.

Each chapter is concluded by a presentation of the main open problems and of future directions of research specific to the materials presented in the chapter. In the conclusion (Chapter 5) I sketch a research project motivated by the development of autonomous artificial agents which can act as intermediaries in a collective decision.

RÉSUMÉ EN FRANÇAIS

Dans ce document, je présente les résultats de mes recherches depuis l'obtention de mon doctorat en suivant un fil conducteur qui en montre les évolutions. Mes recherches se situent en choix social computationnel, c'est-à-dire, l'analyse avec des outils d'informatique théorique de processus de décisions collectives prises par un ensemble d'agents autonomes. Il s'agit d'un champ de recherche multidisciplinaire qui se situe entre l'intelligence artificielle (notamment les systèmes multi-agents, la représentation et le raisonnement sur des connaissances), l'économie, et les sciences politiques. Pour ce qui concerne les travaux présentés ici, les techniques d'analyse utilisées et les applications visées, ils sont tous pour la plupart dans le domaine de l'intelligence artificielle.

Mes travaux ont pour but d'enrichir les études de choix social sous trois aspects, correspondants aux trois chapitres principaux de ce document: (a) prendre en compte les interactions entre votants avec l'introduction d'un réseau d'influence ou de communication, (b) modéliser le comportement stratégique des votants et, (c) augmenter l'expressivité des bulletins de vote (qu'il ne faut pas imaginer en papier!) tout en restant dans des classes de complexité computationnelle acceptables. Le chapitre introductif présente les grandes lignes de recherche dans lequel mon travail s'inscrit, donne un résumé du contenu des chapitres et définit les notions de bases pour la compréhension des résultats qui suivent. La plupart des sections techniques sont agrémentées d'exemples pour pouvoir rapidement saisir l'intérêt et la portée des résultats présentés.

Dans le Chapitre 2, j'augmente le modèle classique du choix social avec un réseau qui connecte les votants et qui peut représenter un réseau d'influence ou de communication entre agents. Une grande partie du chapitre est dédiée à étudier des modèles de diffusion d'opinion sur un réseau qui sont adaptés à des situations de vote et à ébaucher un modèle de diffusion stratégique. Deux études de conception algorithmique sont ensuite présentées. La première qui enrichit le système de démocratie liquide avec des délégations multi-agents et classées. La seconde définit une notion d'évaluation collective qui est plus résistante à la subornation grâce à l'utilisation d'un réseau.

Le Chapitre 3 introduit la notion d'un jeu de vote, dans lequel la dimension stratégique des votants est prise en compte. Premièrement, je définis et étudie la classe des jeux de Gibbard-Satterthwaite, qui permet de zoomer sur un sous-ensemble d'agents stratégiques avec actions restreintes. Ensuite, dans une situation de vote sur plusieurs questions binaires, je montre comment l'utilisation de négociations

précédant le vote peut être utilisée comme sélecteur d'équilibres de Nash désirables. Le chapitre se conclut avec la présentation de plusieurs travaux autour du vote itératif, avec des techniques de simulation multiagents, d'apprentissage par renforcement et d'économie expérimentale.

Le Chapitre 4 se concentre sur l'expressivité des votants, notamment en cas de vote combinatoire, quand l'ensemble des alternatives décisionnelles est de taille exponentielle. Plusieurs approches y sont présentées, s'appuyant sur un exemple de planification d'activités en groupe. Je commence par une comparaison détaillée des cadres d'agrégations des jugements sous l'angle de la concision (la taille de représentation d'un problème). Ensuite, je présente un système de vote fondé sur l'élicitation des buts propositionnels. Enfin, j'analyse les propriétés axiomatiques et la complexité computationnelle de trois règles de vote inspirées de la règle de majorité. Tout en montant en complexité d'élicitation, j'étudie ensuite l'agrégation des préférences conditionnelles incomplètes et je définis une règle de vote pour l'analyse du sentiment collectif enrichi avec l'extraction de préférences comparatives.

Tous les chapitres se terminent avec une présentation des problèmes ouverts et des futures directions de recherches spécifiques aux matériels présentés. Dans la conclusion (Chapitre 5) j'expose un projet de recherche fédérateur autour du développement d'agents artificiels autonomes qui puissent agir en tant qu'intermédiaires dans un processus de décision collective.

*Facesti come quei che va di notte
che porta il lume dietro a sè e non giova
Ma dopo sè fa le persone dotte*
— *Purgatorio, XXII 67–69*

ACKNOWLEDGEMENTS

I was very lucky in my professional life to meet a number of excellent teachers, supervisors, and colleagues who taught me how to do the complex job of a researcher (and it's certainly not over). This document should prove that I am now able to supervise the work of others, and I could not find a better image of this mission than the citation above by Dante: using one's lamp not to light up one's path, but the path of others.

Without all colleagues, PhD students, and Master students with whom I collaborated this thesis will be empty. Their names are mentioned in the relevant chapters (apologies if I forgot anyone), and they deserve one big thank for the great time we spent together.

This document was mostly written at home during the *COVID-19* pandemic. Many people did not share the same luck of staying safely at home and kept vital aspects of our societies running. This thesis is dedicated to them.

Finding the motivation to write this thesis was possible thanks to the support (at times the pressure) of Sophie, who was probably preoccupied of giving me an occupation while she was busy nurturing our child in her growing belly. I am now done writing and I am excited to welcome him!

Finally, I am grateful to the creators of the *classicthesis* style-file in LaTeX, to Flaticons for some of the pictures,¹ and to all musicians whose recordings helped me getting through the writing.

*Toulouse, August 2020
Revised in May 2021*

¹ In particular: *Photozidea-studio*, *Freepik*, and *Prosymbols* from www.flaticon.com.

CONTENTS

1	INTRODUCTION	1
1.1	Motivating example: automated personal assistants . . .	2
1.2	The architecture of collective decisions	4
1.3	Structure of the document	5
2	SOCIAL CHOICE AND SOCIAL NETWORKS	7
2.1	Opinion diffusion as aggregation	8
2.1.1	Related work on opinion diffusion	9
2.1.2	Basic definitions	10
2.1.3	Propositional opinion diffusion	11
2.1.4	Diffusion of preferences	14
2.1.5	Turn-taking and opinion diffusion in voting . .	15
2.1.6	Strategic opinion diffusion	16
2.2	Social networks as information filters	17
2.3	Smart voting	19
2.4	Personalised ratings	24
2.5	Conclusions and perspectives	28
3	VOTING GAMES AND ITERATIVE VOTING	29
3.1	Context: The game-theoretic structure of voting	30
3.2	Gibbard-Satterthwaite games	31
3.3	Equilibrium selection via pre-vote negotiations	33
3.4	Iterative voting	34
3.4.1	Restricted convergent dynamics	35
3.4.2	Reinforcement learning agents	36
3.4.3	Behavioural experiments in multiple referenda	37
3.5	Conclusions and perspectives	38
3.5.1	The <i>Itero</i> platform	38
4	EXPRESSIVE SOCIAL CHOICE	41
4.1	Judgments and binary views	43
4.2	Goal-based voting	45
4.3	Conditional preferences	49
4.4	Preference and polarity	51
4.5	Conclusions and perspectives	53
5	CONCLUSIONS AND PERSPECTIVES	55
5.1	Going further	56
	BIBLIOGRAPHY	61

INTRODUCTION

The study of social choice was deeply transformed when computer scientists started analysing the algorithmic aspects of collective decision making, leading to the thriving field of research of computational social choice (for introductions to the field see, e.g., [18] and [138]). In more recent years a second transformation is ongoing, less algorithmic in nature. Researchers are moving closer to applications and real-world situations by, e.g., simulating collective decisions in a lab (see, e.g., [134]), developing online applications for voting or fair division such as *Whale*¹ and *Spliddit*,² collecting datasets of voters preferences,³ or augmenting the formal models of social choice with more complex agents' architectures.⁴

Research presented in this thesis belongs to this second wave, and aims at enriching social choice studies focussing on three aspects of collective decisions: the interaction among voters, their strategic behaviour, and the expressive power of the voting ballots.

But what is social choice?

Social choice studies situations in which a set of autonomous agents need to take a collective decision, i.e., choosing among a set of alternatives on which each agent has individual preferences. Classical examples include the election of candidates, the allocation of resources, or the fair division of goods. It is a problem of a cooperative nature, as the alternative is chosen for the entire group of decision-makers, but the competitive aspect is dominant, as each agent has different and possibly conflicting preferences over the alternatives. The axiomatic method is one of the main techniques used by researchers in social choice theory to advance their knowledge, specifying a number of desiderata for a given collective decision rule, and then exploring the space so-defined, eventually characterising all such rules or prove that there is none. Cornerstone results using this method are Arrow's theorem and the Gibbard-Satterthwaite theorem, which are both impossibilities, as well as the axiomatic characterisation of the majority rule proven by May's theorem. Computational and algorithmic properties have been added to classical fairness desiderata, such as requir-

¹ <https://whale.imag.fr>

² www.spliddit.org

³ www.preflib.org

⁴ Some of these topics are discussed at length in a recent book edited by Ulle Endriss [53]. The most recent advances are summarised in the proceedings of the Dagstuhl seminar on Application-Oriented Computational Social Choice [83].

ing the existence of tractable algorithms for the computation of the vote outcome, or, symmetrically, that manipulating the result be computationally intractable as a barrier against malevolent behaviour by the agents involved. A complete introduction to the field can be found in the Handbook of Social Choice and Welfare [4] for the economics and political science approaches, and in the Handbook of Computational Social Choice [18] for a recent survey of the computational aspects.

And why is it relevant for artificial intelligence?

Research in multiagent systems, the area of artificial intelligence (AI) that studies the interaction of autonomous agents, received modern foundations from mathematical economics, and game theory in particular (see, e.g., [146, 154]). Collective decisions are examples of the possible interactions in a multiagent system, but one that requires a vast number of techniques in AI and that touches most aspects of the autonomous agents architecture. Moreover, even in applications that are far from traditional problems in AI, any computational study of a social choice problem would require knowledge of both economics and computer science techniques that are in the toolbox of an AI researcher. The fruitfulness of marrying economics models in social choice with AI and computational techniques is testified by applications such as the recent interest in e-democracy, the advent of group recommender systems, and the development of matching procedures at large scale.

1.1 MOTIVATING EXAMPLE: AUTOMATED PERSONAL ASSISTANTS

Let us consider an idealised situation, which will provide examples for the main issues that this thesis will touch upon:

Example 1.1.1 (Personal assistants fixing a meal). ⁵ *Consider five automated personal assistants which are given instructions by their owners to organise a meal at a restaurant for the five of them. Each owner specified their preference to her assistant, and gave access to her calendar. After a brief automated discussion it becomes clear that the conflicting issues to be decided upon are:*

- *whether the restaurant should be in the center or in the suburbs,*
- *lunch or dinner,*
- *fancy or casual restaurant.*

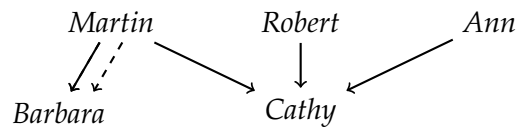
The automated personal assistants thus form a multiagent system who has to take a collective decision on three issues.

⁵ This example is based on Example 6.1 in the thesis of Novaro [126].

The example above considers a collective decision of low stake such as organising a meal, but similar examples could be constructed about higher-stake decisions such as the selection of candidates in a hiring process, the choice of which products a company should invest on, or policy features.

There are three aspects of collective decisions on which this thesis focuses on, corresponding to the three main chapters. First, let us consider a network of trust that relates the five personal assistants of Example 1.1.1, modelling the interactions among the decision-makers. This will be the focus of Chapter 2:

Example 1.1.2 (Interactive social choice). *Barbara told her assistant to do whatever Martin said on whether the restaurant should be fancy, or whether it should be in the center or not, but that she would prefer lunch over dinner. Cathy told her assistant that she would conform to the majority of the others about fancy or casual (except for Barbara), and that she would prefer to have dinner in the center. The other three owners gave direct preferences on the three issues to their assistants. The situation induces an influence network among the agents which can be depicted in the figure below, with dashed arrows corresponding to the first issue, and thick ones to the third issue:*



The five automated assistants know which other agents they trust, and their preference for the final vote will depend on the opinions of the other assistants. Depending on the discussion protocol they chose they might end up in an infinite loop of influence constantly changing their mind. Some of them might actually have this situation as a meta-goal, to prevent a decision to be taken and gain time...

Let us now focus on the strategic abilities of the automated assistants, who can reason about each others' preferences and act accordingly. Chapter 3 will study situations such as the following one:

Example 1.1.3 (Strategic social choice). *Let us focus on three automated assistants only. Barbara asked her assistant to organise the meal either in the center or for lunch (or both). Martin wishes to have either lunch or to be in a fancy restaurant (or both). And Ann would like a restaurant in the center or a fancy one, or both. Clearly, having lunch in the center in a fancy restaurant would satisfy the three of them. Their personal assistants decide to resolve the issue by a quick majority vote on each of the three questions. The result could be the following:*

Barbara	Center	Dinner	Casual
Martin	Suburbs	Lunch	Casual
Ann	Suburbs	Dinner	Fancy
Majority	Suburbs	Dinner	Casual

The result that the assistants will propose as the outcome of the majority vote is completely unsatisfactory as it does not satisfy any of the specified preferences. However, they all voted according to the preferences they received. Suppose now that the assistants are able to struck deals among each others before the final vote takes place. Barbara's assistant could ask Martin's one to vote for a restaurant in the center instead of the suburbs, as he is indifferent about this issue. And Martin's assistant can in turn ask Ann's one to vote in favour of lunch over dinner, as she is indifferent on this issue. Thus, by a pre-vote negotiation the three agents are able to find a voting pattern that satisfies the preferences of the three owners.

The decision that the assistants have to take has three features, making the space of possible alternative plans a combinatorial one. Expressing preferences in such domains is the subject of Chapter 4:

Example 1.1.4 (Expressive social choice). *As in Example 1.1.2, Robert specifies a clear preference for a casual lunch in the city center. Ann instead has the same preference as in Example 1.1.3, that is she is indifferent about lunch or dinner, but would like a restaurant in the center or a fancy one (or both). Barbara realises that if the restaurant is in the center, then she would prefer a casual one over a fancy one, however if the restaurant is in the suburbs then she would prefer a fancy one. Robert's preferences can be easily represented as a complete judgment on the three issues at stake, while Ann's preferences rather specify a set of possible answers to the three issues, which are logically correlated. Barbara's preferences are conditional, and can be represented by means of suitably defined logical statements. Once the problem of representation of preferences is solved, the automated assistants still face the issue of how to properly aggregate such compactly represented preferences into a choice for a meal, with the usual limitations in space and time resources.*

To sum up, autonomous agents involved in a collective decision can face situations which require adequate representations for the influence among the voters, for their strategic abilities, and for the expressive variety of their preferences. This document delve into each of these questions, providing models and algorithms for interactive, strategic, and expressive social choice.

1.2 THE ARCHITECTURE OF COLLECTIVE DECISIONS

A different angle to present the content of this thesis is to consider the improvements that each of the three main chapters bring to the overall architecture of a collective choice problem. Each individual agent in a multiagent system is typically characterised by three features: a set of actions at his disposal, a (typically partial) knowledge of its environment, and a preference or a goal on the possible configurations of the multiagent system. In the simplest model of a collective decision, that of an election over alternatives, agents can express their vote

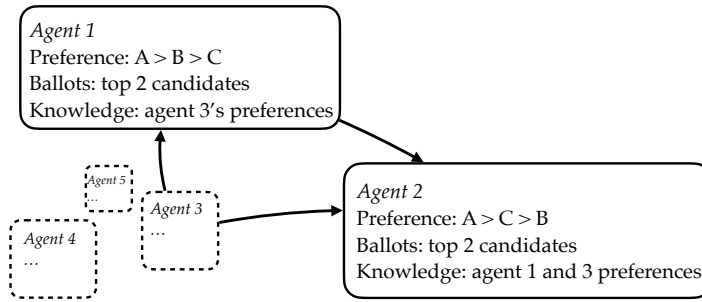


Figure 1: A multiagent system composed of voting agents taking a collective decision on candidate alternatives A, B, C . Each agent has preferences over the outcome of the vote, can express its opinion with voting ballots, and has partial knowledge about its environment. The voting rule is not represented, but in this case associates a winning candidate with every configuration of voting ballots. The arrow between agents are visualisations of the private knowledge of the agents, and in Chapter 2 they will also be used for representing social influence.

on ballots, such as communicating their most preferred candidate or providing an ordering of the preferred three candidates. They have full or partial knowledge about the preferences or the votes of other voters. And they have preferences on the outcome of the election—outcome that is decided by applying a voting rule on the ballots of all the voters. Figure 1 provides a visual explanation.

When studying the interactions among voters defined by their mutual influence in Chapter 2, we will work on the knowledge that each agent has of other agents, specifically the ones influencing it, and its effect on the resulting actions. An agent can thus be influenced by updating its preferences or its choice of action. In Chapter 3 we will focus on the action that the agents take depending on their knowledge of what actions the other agents will play, in this way taking a full game-theoretic perspective. In doing so, we will provide the agents with computational tools to analyse the strategic structure of collective decisions. The focus of Chapter 4 is the set of actions at the agents' disposal, i.e., the voting ballots, which are a crucial aspect of a collective decision when the set of alternatives has a combinatorial structure.

1.3 STRUCTURE OF THE DOCUMENT

This thesis presents the research I conducted after obtaining my PhD. It is organised in three chapters, each presenting a coherent set of results around three themes: modelling interactive, strategic, and expressive social choice.

Chapter 2 augments the classical model of social choice with a network relating the voters, which can represent the trust among the agents or a communication network. A large part of the chapter is dedicated to the study of models of opinion diffusion that are adapted to social choice situations, and to sketch a first model of strategic opinion diffusion. The last part of the chapter takes a mechanism design perspective, presenting two studies of algorithmic design. The first is a generalisation of liquid democracy accounting for multiagent ranked delegations, and the second is a study of bribery of two systems for collective rating.

Chapter 3 starts with the notion of a voting game, and thus takes the strategic aspects of collective decisions into account. The chapter begins with the introduction of a novel class of games called Gibbard-Satterthwaite games, which focuses on players with restricted sets of actions. Then, in a situation of binary vote on multiple issues, we show how the use of negotiations preceding the vote can be used as selectors of Nash equilibria. The chapter is then concluded by the presentation of a number of studies around the problem of iterative voting, with techniques from multiagent simulations, reinforcement learning, and experimental economics.

Chapter 4 focuses on the expressivity of the voters, more specifically in combinatorial vote, when the set of alternatives is of exponential size. Several approaches are presented, building on a motivating example of group activity selection. We start from a comparison of judgment aggregation languages from the point of view of succinctness, moving on to presenting a full-fledged analysis of voting based on propositional goals. Increasing the expressive power of individual ballots, we then consider the aggregation of conditional preferences that are probabilistic and incomplete, and, finally, present a voting rule for collective sentiment analysis in the presence of both polarities and comparative statements.

The introduction of each chapter presents the research problem, and details all collaborations concerning the results presented. Related work is surveyed to put my research in context. Each following section is complemented with examples aimed at quickly grasping the significance of the results and the interest of the problem considered. The last section of each chapter details open problems and future research directions.

*Notes in the margin
highlight specific
results or features*

This chapter presents research on enriching the structure of a social choice problem by taking into consideration a network relating the decision-makers. The importance of network phenomena in collective decisions is difficult to underestimate. The effects of social media on real elections are under constant scrutiny, and social network analysis provides by now standard tools for political campaigners. Moreover, autonomous agents are already quite present on social media, and interact with decision-makers by filtering our access to information or directly as content producers. From the perspective of mechanism design as well, social networks can play an important role in the creation of novel collective-decision methods.

Social network analysis [47, 95] provides plenty of tools for the social choice theorist who wants to embark on such an enterprise. These tools, as we shall see, need however to be adapted to fit the (often combinatorial) structure of social choice problems, and important modelling considerations are required to clarify the role of networks in different scenarios. A network among voters can be seen as a compact representation of the mutual influence or trust among the decision-makers, but also as a map of the information channels that are available to them.

This chapter focuses on two social choice problems: voting, both on binary issues and on a set of alternative candidates, and collective rating of service-providers. Most of the techniques we use come from theoretical computer science and game-theory, complemented by initial results with multiagent simulations. The first two sections are concerned with the effect that social networks can have on a collective decision. They deal with two different aspects of this problem: social influence (Section 2.1), and non-uniform access to information (Section 2.2). The two sections that follow them take the perspective of mechanism design to exploit the presence of a social network and obtain decision mechanisms—delegative voting rules in Section 2.3 and a rating system in Section 2.4—that satisfy desirable properties.

I started this line of research at my arrival at IRIT in Toulouse, beginning a fruitful collaboration with Emiliano Lorini and Laurent Perrussel that is still ongoing. The results presented in this chapter include the work of two PhD students (Arianna Novaro and Rachael Colley) and three Master students (Sirin Botan, Claire Pillet, and James Stewart) whom I co-supervised in Toulouse. This research is also the result of international collaborations with Markus Brill from *TU Berlin*, Edith Elkind from *Oxford University*, Ulle Endriss from the

*Outline of the
chapter*

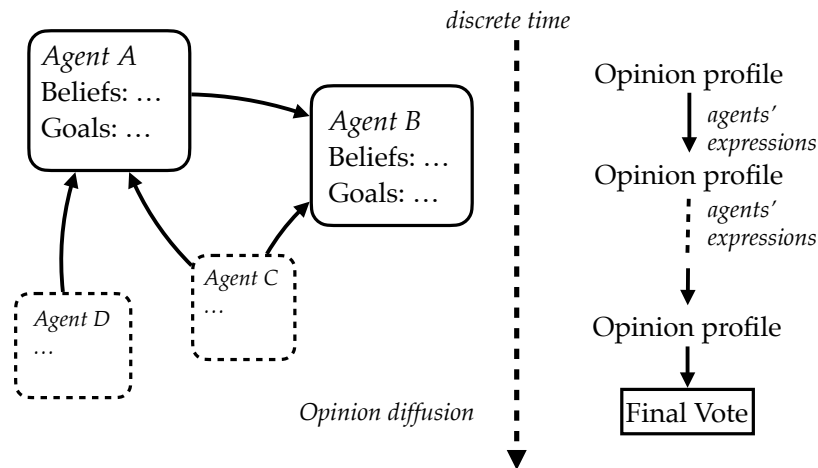


Figure 2: A set of agents is linked by a network of trust, or mutual influence. Each agent has a private and partial knowledge of the situation, and a private goal or preference. Agents exert their influence by expressing opinions, entering a process of opinion diffusion which results in a final vote.

*Survey on social
choice and social
networks*

University of Amsterdam, and Paolo Turrini from Warwick University. I wrote a survey in 2017 on social choice and social networks [72], which introduces the reader to various approaches on this problem. Research is however moving very fast on this topic, and many further publications appeared since then on this topic. Part of the work presented in this chapter is being conducted in the ANR-18-CE23-0009-01 SCONE project on social choice and social networks.¹

2.1 OPINION DIFFUSION AS AGGREGATION

In this section we focus on the role that social influence can play when preceding a vote. A high-level description of this process is depicted in Figure 2.

This section presents results on adapting and studying models of opinion diffusion for voting situations, in which the expressed opinion of the agents take the form of binary views over interconnected issues (Section 2.1.3) and linear orders (Section 2.1.4). Results studying the effects of the presented diffusion models on voting rules are summarised in Section 2.1.5, and a full-fledged epistemic model of strategic diffusion is described in Section 2.1.6.

¹ <https://www.irit.fr/scone>

2.1.1 Related work on opinion diffusion

Accurate introductions to the field of social network analysis can be found in the classical books of Jackson [95] and Easley and Kleinberg [47]. How information, diseases, or opinions diffuse on a given network is a classical problem studied in this field. Models of opinion diffusion can be classified depending on the mathematical representation that is chosen for the agents' opinions. Continuous models, in which nodes' opinions have values $x \in [0, 1]$ or equivalent, are widely studied. Perhaps the most well-known example is the DeGroot's model [88] (also known as the Lehrer-Wagner model [109]) in which agents update their opinions depending on the weight they attach to their neighbours opinions, defining a linear model based on iterated weighted sums. A wide literature developed around a generalisation of this model known as the Friedkin and Johnsen model (see, e.g., [61]), including recent developments which incorporate constraints on the multiple continuous beliefs held by the agents [62]. A large number of results in this literature characterise networks and model parameters that guarantee some forms of consensus at the end of the opinion diffusion process. Discrete models, instead, typically assume that agents can only have binary (two-state) opinions such as yes-no, for-against, or a list of finite opinions, such as colours. Granovetter [87] and Schelling [140] were the first in proposing and studying threshold models, with each agent in the population being assigned a numerical value that indicates the proportion of neighbours who must have a certain opinion before the agent adopts it. A generalisation of these models that is widely studied in computer science are linear threshold models [99, 100], which also include weights on agents' links, to model the strength of the influence relation. Finally, diffusion processes known as voter models study the probabilistic propagation of opinions, deciding at each time step if an individual takes the opinion of a random neighbour [33, 93], or, more closely to threshold models, the most approved opinion in her neighbourhood [133].

More recently, a stream of papers in artificial intelligence proposed generalisations of discrete opinion diffusion models with more complex representations of opinions. What is common among these models is the view that each agent, in updating her opinion, aggregates the opinions of her neighbours. Schwind et al. [141, 142] model opinions as knowledge bases, i.e., sets of propositional formulas, importing aggregators from the literature on belief merging. The work I will summarise in Section 2.1.4 studies how preferences represented as linear orders diffuse following majoritarian opinions in one agent's neighbourhood [21]. Finally, previous work which will be discussed in Section 2.1.3, as well the work of Christoff and Grossi [30, 32], focused on possibly interconnected binary views, importing aggrega-

Most research is on models with continuous opinions

Most results focus on finding consensus

The celebrated segregation game is a threshold model (see, e.g., ncase.me/polygons/)

Opinion diffusion as aggregation

tors from the field of judgment aggregation. Related to this last line of research is the more general treatment of opinion transformation proposed by List [112], which however did not explicitly consider a network relating the individuals. Models of binary opinion diffusion are closely related to the literature on boolean networks (see, e.g. [28, 98]), which stems from the modelling of biological regulatory networks [145].

Power indexes on social networks

My approach at opinion diffusion models was widely influenced by the work of Grabisch and Rusinowska on power indexes and influence functions [68–71]. Their work can in turn be traced back at a formal model of influencing power introduced by Isbell [94] and developed more recently by Hoede and Bakker [92]. Their first work focused on binary decisions, with an influence function determining the dynamics from a profile of individual opinions over n agents $x \in \{0, 1\}^n$ to the next [68, 70]. The key concept is the followers function, which specifies for each coalition of individuals the set of individuals that would follow their (unanimous) decision. The model was generalised to non-binary settings, with an ordered set of actions or a continuum of actions available to individuals [69, 71].

2.1.2 Basic definitions

The basic bricks of the model are a finite non-empty set of agents $\mathcal{N} = \{1, \dots, n\}$ and a directed graph $E \subseteq \mathcal{N} \times \mathcal{N}$ which is taken to represent a trust or influence network among the agents. While most social networks are often assumed to be undirected (e.g., modelling friendship) we take here a more general perspective, to account for situations in which influence is not necessarily mutual. We denote the set of influencers of an agent i in network E with $Inf(i) = \{j \mid (j, i) \in E\}$, reading $(j, i) \in E$ as “ j influences i ”. If E is reflexive, then for all $i \in \mathcal{N}$ we have that $i \in Inf(i)$, which is a realistic assumption in most applications of opinion diffusion.

We assume that each agent $i \in \mathcal{N}$ has an individual opinion, which we denote B_i for agent i , which in further sections will take different forms: a binary opinion, a ranking over a number of alternatives, a model of a propositional formula. We call *opinion profile*, denoted with $\mathbf{B} = (B_1, \dots, B_n)$, the ordered list of individual opinions of the agents.

The opinion of the influencers are aggregated

The model of social influence proposed here is based on aggregation: each individual i uses an aggregation function to update her opinion from those received by her influencers (and her own one). Formally, if \mathcal{X} is the space of all possible individual opinions, we consider a collection of functions $F_i : \mathcal{X}^n \rightarrow \mathcal{X}$ for each $n \in \mathbb{N}$ and for each $i \in \mathcal{N}$. The dependency on n accounts for different possible sizes of the influencers set. Notable examples of aggregation functions are the majority rule over binary opinions, a judgment aggregation func-

tion over models of a propositional formula, or a belief merging operator.

Opinion diffusion is then modelled as an iterative process in discrete time. Let $turn : \mathbb{N} \rightarrow 2^{\mathcal{N}}$ indicate at each point in time the set of agents updating. Let B_i^t be the opinion of agent i at time $t \in \mathbb{N}$, and:

A discrete-time iterative process

$$B_i^{t+1} = \begin{cases} F_i(B_i^t, \mathbf{B}^t \upharpoonright_{Inf(i)}) & \text{if } i \in turn(t) \\ B_i^t & \text{otherwise.} \end{cases}$$

Where $\mathbf{B}^t \upharpoonright_{Inf(i)}$ denotes the restriction of profile \mathbf{B} to the agents in $Inf(i)$. If $turn(t) = \mathcal{N}$ for all t we name the process *synchronous*. When the *turn* function instead selects one single individual at each turn the process is called *asynchronous*.² When individual opinions are complex, e.g., multiple binary issues or preferences, the *turn* function needs to be coupled with a further scheduler identifying which individual updates on which part of the opinion.

The first problem we study is the termination of the iterative process. Two notions can be defined. If we call a transition from a profile \mathbf{B} to a second profile \mathbf{B}' *effective* if $\mathbf{B}' \neq \mathbf{B}$, then we say that an opinion diffusion process *terminates universally* if there exists no infinite sequence of effective transitions starting at any initial profile. Instead, we say that an opinion diffusion process *terminates asymptotically* if from any initial profile there exists a sequence of transitions that reaches a stable profile, where a profile \mathbf{B} is said to be *stable* if $F_i(\mathbf{B}) = B_i$ for all $i \in \mathcal{N}$. For an asynchronous model the latter notion corresponds to requiring the model to be an absorbing Markov chain.

Universal and asymptotic termination

While previous research mostly focused on identifying initial conditions to reach a consensual opinion among the agents, our perspective of studying social influence before a vote leads us to consider a wider set of situations. Let a *termination profile* for initial profile \mathbf{B}^0 be any stable profile reachable from \mathbf{B}^0 . We can then investigate when a process *converges to a unique profile*, if all termination profiles reachable from an initial one coincide. For the specific case of preferences over alternatives, we can investigate when all termination profiles are aligned, e.g., single-peaked, or if they allow for a Condorcet winner more often than the initial one.

We do not only focus on consensus

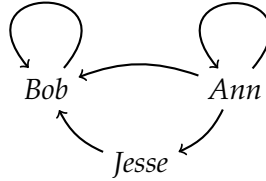
2.1.3 Propositional opinion diffusion

In this section we assume that individual opinions are binary views over a set of binary issues: $B_i \in \{0, 1\}^{\mathcal{J}}$ where \mathcal{J} is a finite non-empty set of issues, with 1 standing for a “yes” vote and 0 for a “no” vote.

² An anonymous reviewer correctly pointed out that the existence of a clock would still make the process synchronous. For the sake of uniformity with previous publications I am still using the old but imprecise terminology.

Issues can be correlated (see Example 2.1.2), in which case we will assume that the set of feasible individual opinions is restricted by an integrity constraint. Formally, if IC is compactly represented as a propositional formula over variables \mathcal{J} then we assume that $B_i \models IC$ for all i . Consider the following example:³

Example 2.1.1. Consider the following influence network among three agents, Ann, Bob and Jesse:



Suppose the three agents need to decide whether to approve the building of a swimming pool (first issue) and a tennis court (second issue) in the residence where they live. The following table presents their initial opinions and the evolution of the synchronous opinion diffusion process assuming that each agent uses the majority rule to update her opinion:

Initial opinions	Profile \mathbf{B}^1	Profile \mathbf{B}^2
$B_A^0 = (0, 1)$	$B_A^1 = (0, 1)$	$B_A^2 = (0, 1)$
$B_B^0 = (0, 0)$	$B_B^1 = (0, 0)$	$B_B^2 = (0, 1)$
$B_J^0 = (1, 0)$	$B_J^1 = (0, 1)$	$B_J^2 = (0, 1)$

Study termination
of IC-consistent
aggregators

There are two perspectives we can take when facing possibly correlated multiple issues. The first option is to use aggregation functions that always return IC-consistent opinions (or equivalently assume that issues are independent, which corresponds to $IC = \top$). Under these assumptions we can show the following:

Theorem 2.1.1. [81] If F_i satisfies ballot-monotonicity for each $i \in \mathcal{N}$, then propositional opinion diffusion universally terminates on the class of directed acyclic graphs with loops after at most $diam(E) + 1$ number of steps.

The assumption of ballot-monotonicity (which is tight) requires that for each profile $\mathbf{B} = (B_1, \dots, B_m)$, if $F(\mathbf{B}) = B^*$ then for any $1 \leq i \leq m$ we have that $F(\mathbf{B}_{-i}, B^*) = B^*$. That is, reinforcing a given opinion in the input does not change the result of the aggregation. $diam(E)$ is the length of the longest path in E .

We obtained more results in this line for specific aggregators such as the majority rule or the unanimity rule, also assuming independent issues, in joint work with Emiliano Lorini and Laurent Perrussel [81].

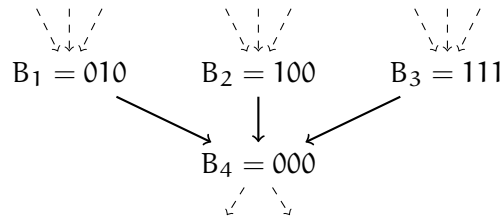
³ Revisited from a classical example in multi-issue voting [29].

There is considerable related work on opinion diffusion on a single binary issue. Most notably, we mention the classical work of Goles and Olivos [67] on threshold rules, and the recent work of Christoff and Grossi [32] on finding sufficient conditions for universal termination of arbitrary aggregation procedures, as well as the work of Bredereck and Elkind [19] focusing on asynchronous processes for the majority rule.

The second approach to deal with logical dependencies between the issues consists in using simpler aggregation functions, such as the issue-by-issue majority rule, with influence percolating on the network one issue at the time. In the presence of integrity constraints this approach is not straightforward to apply, as exemplified in the following situation:

Propositionwise diffusion

Example 2.1.2. *Four individuals are part of a collective decision-making process on whether to build a skyscraper (S), a hospital (H), or a new road (R). Law says that if S and H are built then R also should be built (corresponding to integrity constraint $(S \wedge H) \rightarrow R$). The four agents are part of the following influence network:*



Suppose that agent 4, which is influenced by the other three, asks her influencers their opinion on the 3 issues at the same time. Then she would face an inconsistent issue-by-issue majority result of 110, and it would be unclear how to update her opinion. However, agent 4 can also ask questions on only two issues at the time, in which case the result can be 100 or 001 depending on the order of questions. The same result can be reached by asking two 1-issue questions in sequence.

Generalising on the previous example, it seems that by limiting the size of queries that agents can ask we can circumvent the problem of updating on an inconsistent aggregated view. However, when the aggregated opinion of the influencers is consistent, we would still like to be able to reach it by asking proposition-wise queries. Borrowing concepts from the theory of boolean functions, we are able to prove the following result:

Theorem 2.1.2 ([11]). *Let IC be an integrity constraint, and let B' be reachable from an IC-consistent initial profile B by asking queries on all issues at the same time. Then, B' is reachable from B by asking queries on k issues at the time iff IC is k -geodetic.*

Transitivity or budget constraints work well with propositionwise diffusion

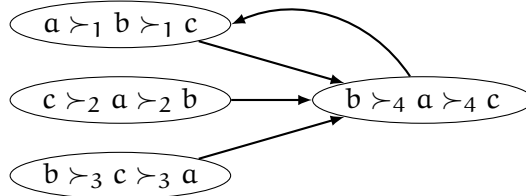
The notion of k -geodeticness is inspired by the work of Ekin, Hammer, and Kogan [48]. Intuitively, a constraint is 1-geodetic (or geodetic *tout court*) if any two of its models are connected by a shortest path on the boolean cube. To give an example, a simple pXORq is not geodetic, since any shortest path between the two models 10 and 01 passes through a non-model (00 or 11). On a more positive note, we can show that constraints used in decision-theoretic applications, such as transitivity constraints or budget constraints, are geodetic. In general, it is co-NP-complete to test if a given constraint is k -geodetic.

Further investigations of propositionwise opinion diffusion processes under integrity constraints can be found in joint work with Sirin Botan and Laurent Perrussel [11]. To the best of our knowledge, the only related work investigating constraints in discrete opinion diffusion is the work of Christoff and Grossi [30] on delegation graphs, where each node has at most one influencer.

2.1.4 Diffusion of preferences

In voting applications agents are typically assumed to express a preference over a set $A = \{a, b, c \dots\}$ of alternative candidates, with the most common assumption being a linear order (aka strict ranking). In this setting, the idea of issue-wise diffusion can be applied to pairs of alternatives that are *adjacent* in an individual preference order, as explained in the following example:

Example 2.1.3. Consider a simple influence network with $n = 4$ agents and $|A| = 3$ alternatives in the figure below.



The preferences of agents 1, 2, and 3 form a Condorcet cycle, i.e., the majority relation of their preferences is cyclic. In one possible asynchronous sequence of updates, agent 4 updates on pair ab , moving to preference $a >_4 b >_4 c$. After that, no further updates are possible: even though agent 4 disagrees with its influencers on pair ac , this pair cannot be swapped since it is no longer adjacent in \succ_4 . The agents reached a stable state, and the pairwise diffusion process terminated. If we instead consider synchronous updates, with all agents updating at the same time, possibly on different pairs of alternatives, we observe that if agents 1 and 4 update repeatedly on pair ab we obtain an infinite update sequence and the diffusion process does not terminate. Agents 2 and 3 are sources, i.e., they have no influencer, and thus never update their preferences.

Recall that asymptotic termination requires that in each opinion profile (denoted as \succ for linear orders over candidates) we can construct a path of updates that leads to a termination profile. Whilst the example above shows that asymptotic termination with synchronous pairwise majority diffusion is not always possible, we show the following result for asynchronous processes:

Theorem 2.1.3 ([21]). *Let \succ^0 denote the initial preference profile. For $|A| = 2$, asynchronous pairwise majority diffusion asymptotically terminates on any graph E . For $|A| \geq 3$, asynchronous pairwise majority diffusion asymptotically terminates on E when (\succ^0, E) satisfies linear local majority.*

The linear local majority assumes that all local majorities calculated on sets of influencers in any profile that is reachable by pairwise diffusion from \succ^0 on the network E are transitive. Basically, it requires that no Condorcet cycle is encountered in the diffusion process. It is a rather strong assumption, but observe that the termination result is shown for any graph E . We conjecture that Theorem 2.1.3 still holds without this restrictive assumption, but the proof is still an open problem.

Relaxing the linear local majority assumption is an open problem

We also initiated a qualitative study of the termination profiles. We obtained only limited results, the following one being of particular interest:

Proposition 2.1.1. *Let E be a DAG and let \succ^0 be a profile such that (\succ^0, E) satisfies the no-tie property (i.e., the diffusion process never leads to ties). If $\succ^0|_{\text{sources}}$ is aligned (single-peaked, single-crossing, Sen value-restriction), then all profiles at termination are also aligned.*

Opinion diffusion propagates the alignment of preferences

The proposition shows that pairwise diffusion propagates the preferential alignment of the sources of a directed acyclic graph to all other agents. Further results in this line can be found in our joint paper with Markus Brill, Edith Elkind, and Ulle Endriss [21].

2.1.5 Turn-taking and opinion diffusion in voting

Once models of opinion diffusion are adapted to voting-like situations, the next question to tackle is whether opinion diffusion is beneficial to the voting process. Ideally, an election organiser would like to know if promoting discussion (hence social influence) among voters is beneficial to the election itself, or if uninfluenced voters lead to better collective decision processes. The Master thesis of Claire Pillet [132] which I recently supervised goes in this direction, using the tool of multi-agent simulations.

In particular, her work aims at investigating the role of the turn-taking function that decides the order of the agents' expressions. For the simplest case of one binary decision and majoritarian opinion diffusion, the results mostly show the superiority of asynchronous

turn-taking in reaching a stable state, and highlight the role of the underlying influence graph in determining the number of switches in the result of a majority vote (we used Erdős-Rényi and Barabási-Albert graphs, with varying initial parameters).

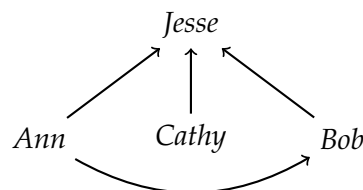
More interesting results are obtained for pairwise preference diffusion, when agents have preferences over multiple candidates. Assessing the quality of a preference profile is not an easy task, and a basic measure is the existence of a Condorcet winner. We then investigated whether pairwise majoritarian diffusion can make a Condorcet winner emerge in profiles that do not admit one. Initial results with 31 voters and 5 alternatives show a 10–15% increase in the number of profiles that have a Condorcet winner after the opinion diffusion process with respect to the initial profiles generated with impartial culture (uniform distribution). These results are irrespective of the turn function and the graph structure used. Moreover, the profiles obtained at termination of opinion diffusion have the further property that the Plurality rule elects the Condorcet winner more often than in uniformly drawn ones, while this property is not verified for the Borda rule. These results suggest that pairwise majoritarian diffusion has a beneficial effect on Plurality elections.

*Preference diffusion
increases the
number of profiles
with a Condorcet
winner*

2.1.6 Strategic opinion diffusion

From a cognitive perspective, the agents in all models described so far are extremely simple: they have no knowledge of their environment, no goal, nor strategy to attain it (recall Figure 2). In collaboration with Emiliano Lorini, Arianna Novaro, and Laurent Perrussel [79, 80] we developed two models of strategic opinion diffusion based on a possible-world semantic. Our models are best exemplified by yet one more example:

Example 2.1.4. *The are four agents: Jesse (the mother), Ann, Bob, and Cathy (the children). The family has to decide whether to go to the cinema or to make an excursion to the countryside. Jesse asks the kids what they would like to do. If they unanimously agree on the same option, then this option will be chosen. Otherwise, in case an agreement is not reached, the default option will be to go to the cinema. Each kid can decide either to abstain or to participate in the vote by declaring her/his preference. Members of the family influence each other following the influence network depicted below:*



Our model aims at representing the strategic action of exerting one's influence on another agent, or abstaining from doing so. We first assume common knowledge of the private opinions. Assume in the situation of Example 2.1.4 that it is common knowledge in the family that Jesse and Bob tacitly prefer the cinema, while Ann and Cathy tacitly prefer the excursion. Nonetheless, according to the rule of the game, if a kid does not explicitly says what she/he prefers, her/his preference will not count. In this situation, Ann should explicitly declare her preference to be sure that her goal of making an excursion will be achieved. Otherwise, Jesse's preference of going to the cinema will possibly prevail. In game-theoretic terms, Ann has a winning strategy to achieve her goal of making an excursion.

Modelling the strategic action of exerting one's influence

We then focus on the epistemic structure of opinion diffusion, dropping the assumption that the individual preferences of the agents are common knowledge among them. Therefore, if an agent does not reveal her opinions, the others will be uncertain about them. If we consider again the situation in Example 2.1.4, Ann would be uncertain about Cathy's preference before Cathy has publicly revealed it. Consequently, Ann does not know whether, by choosing to explicitly declare her preference, an unanimous agreement to make an excursion will necessarily be reached. While Ann does not have a uniform strategy (i.e., one that would achieve her goal in any configuration), she knows that the choice of explicitly declaring her opinions (weakly) dominates the choice of abstaining.

Full-fledged epistemic model of opinion diffusion on multiple issues

We assume that agents' goals are expressed compactly using a fragment of linear temporal logic (LTL) [148], which allows them to express their preferences over future opinion configurations, and to devise sequences of strategic actions to attain them. While we only obtain limited results on characterising such strategies in specific graph structures, we show that the problem of recognising if a given strategy profile is a (subjective) Nash equilibrium can be solved in polynomial space (PSPACE)—to be interpreted as a positive result since satisfaction of an LTL formula is already PSPACE-complete [148]. All other tasks we consider, such as the existence and uniqueness of a Nash equilibria, are significantly harder and we give upper bounds above exponential time.

2.2 SOCIAL NETWORKS AS INFORMATION FILTERS

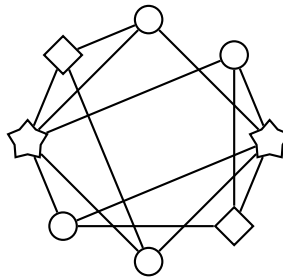
Let us consider once more an election with networked voters: a set of agents \mathcal{N} , connected on a network $E \subseteq \mathcal{N} \times \mathcal{N}$, having to decide on a set of alternatives $A = \{a, b, c \dots\}$. This time, however, the network is not meant to represent mutual trust or social influence between the agents, but maps the agents' exchanges of information. Think of an idealised Facebook wall (based on an undirected graph) or a Twitter page (based on a directed graph). The information access is there-

fore personalised for each voter. This is clearly an idealised situation, as typically voters do also have access to a global information poll, but leaving this aside allows us to focus on the consequences of the asymmetric access to information induced by the network.

In a recent paper which gained popular attention,⁴ Lerman, Yan, and Wu [110] defined and studied a network phenomenon whom they named the majority illusion: the existence of social networks in which the majority of the nodes belong to a given type, but for each of those nodes the majority of their neighbours belong to a different type, therefore creating the wrong perception (the illusion), that the majority type is different from the actual one. In ongoing work with Grzegorz Lisowski and Paolo Turrini we adapt the majority illusion to a voting situation, starting from the following motivating example:

The majority illusion in social networks

Example 2.2.1. *A group of 8 voters participate in a three-candidate election decided by a plurality vote (also known as first-past-the-post). Call the candidates a, b, c, and assume that ties are broken following the alphabetic order. The 8 voters are of three kinds, depicted as circles, diamonds, and stars. The figure below shows the undirected network of information exchange among the voters:*



The circle voters have a preference over candidates of $a > b > c$, denoting that they prefer candidate a to candidate b, and candidate b to candidate c. The star voters have a preference of $c > a > b$, and the diamond voters $b > a > c$. If a vote is taken, the plurality winner would be candidate a with 4 votes in favour, followed by c and b with 2 votes. Observe that candidate a is also a Condorcet winner, and that no voter would have incentive to manipulate if the sincere preferences of voters were common knowledge among them.

<i>Sincere vote</i>	
a	4 votes
b	2 votes
c	2 votes

We assume that voters access information on the election solely via the undirected network depicted in the figure above, and that their strategic behaviour

⁴ www.technologyreview.com/s/538866/the-social-network-illusion-that-tricks-your-mind/

will respond accordingly. Consider the circle agents voting for candidate *a*: each see 2 other voters supporting *c*, and 1 voting for *b*. If they voted for *a* (sincere vote) then *c* would win, which is their worst option. Circle voters then have an incentive to vote for *b* and have their second best alternative winning (by tie-breaking). As for the star and diamond voters, they expect candidate *a* to be the winning candidate with a margin of 2, and thus do not have incentives to manipulate since *a* is their second-best candidate. The new result of the election, with voters responding to the information they received, would thus be *b* winning with 6 votes, no vote for candidate *a*, and 2 for candidate *c*.

Asymmetric access to information leads to a worse election outcome

<i>Strategic response to network</i>	
a	<i>0 votes</i>
b	<i>6 votes</i>
c	<i>2 votes</i>

In our ongoing collaboration we conduct a computational study of illusion phenomena like the one described in the previous example. We generalise the majority illusion of Lerman, Yan, and Wu [110] to allow for more than two types, and show that recognising those networks for which a labeling of nodes would constitute an illusion is an NP-complete problem. We then turn to design polynomial algorithms for repairing a network to ensure that no illusion can occur, mostly by increasing the connectivity under a budget constraint. In doing so we take a similar approach to a recent stream of papers by Auletta, Ferraioli, and Greco [5] and Castiglioni, Ferraioli, and Gatti [26], in which the authors identify networks and initial distributions of opinions (binary case) where a minority vote can become a majority vote by updates due to social influence. We however focus on a non-binary case, and consider strategic agents that *respond* to their neighbours' opinions rather than *conform* to it.

2.3 SMART VOTING

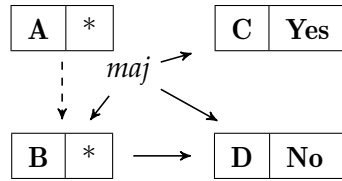
In contrast with the two previous sections, in which we proposed models of social influence and other network phenomena and investigated their effects on collective decision-making, we take in this section a mechanism design perspective. In particular, we are interested in how can social influence be exploited (and, first, elicited) to design novel voting procedures. Building on work on liquid democracy and proxy voting, in our collaboration with Rachael Colley and Arianna Novaro [35, 36] we designed a delegative voting method that allows voters to express rankings and complex delegations.

Let us first provide some context. Liquid democracy sits in the middle ground between a purely representative and a purely direct democratic system, allowing for transitive delegations among the voters.

A number of recent papers in AI investigate liquid democracy

Liquid democracy has been the subject of numerous recent investigations in AI, ranging from analysing its truth-tracking power [9, 34, 97] to its adaptation to multiple issues and alternatives [22, 31]. Our work takes inspiration from the work of Christoff and Grossi [31], which were the first authors to make a connection between computing the result of a liquid democracy instance and models of opinion diffusion such as those described in Section 2.1. Also, in line with recent work of Kotsialou and Riley [103] and Escoffier, Gilbert, and Pass-Lanneau [55, 56], we allow voters to express a ranking of possible delegations, to address potential delegation cycles.

While most implementations of liquid democracy introduce a platform for the elicitation of voters’ delegations, we aim at a decentralised voting system in which voters’ ballots and delegations can remain private. Our first novelty is the definition of a general language for voters’ ballots to express a direct vote, or a delegation to a single other agent, or a combination of the votes of multiple other agents. Second, to tackle the issue of delegation cycles we allow voters to express a number of prioritized delegations, with a final backup vote with the lowest priority. An example of what we call a smart ballot profile can be seen on the following figure:



Our “smart voting” system allows for multiagent and ranked delegations

In the profile described above four agents A, B, C, and D, are taking a binary collective decision, with the possible direct votes being Yes, No or abstain, denoted with *. Voter A expresses a multiagent and ranked delegation: she wants her vote to coincide with the majority of B, C, and D’s votes, and in case this leads to a delegation cycle (i.e. such majority cannot be computed) she gives a single-agent delegation to B, denoted by a dashed arrow. Voter B delegates to D, expressing a single-agent delegation. Voter D casts a direct vote against, while C votes in favour. Voters A and B abstain (*) as their final backup option.

Formally, we start in full generality by assuming that a set of \mathcal{N} voters needs to decide over \mathcal{J} issues, with each issue $i \in \mathcal{J}$ having values in a domain $D(i)$, including absention (*).

Definition 2.3.1. A smart ballot of agent a for an issue $i \in \mathcal{J}$ is an ordering $((S^1, F^1) > \dots > (S^k, F^k) > x)$ where $k \geq 0$ and for $1 \leq h \leq k$ we have that $S^h \subseteq \mathcal{N}$ is a set of agents, $F^h : D(i)^{S^h} \rightarrow D(i)$ is a resolute aggregation function and $x \in D(i)$ is an alternative.

We moreover assume that a smart ballot is *valid*, in the following sense: for each agent a and issue $i \in \mathcal{J}$, for all $1 \leq s \neq t \leq k$ we have

that (i) if $S^s \cap S^t \neq \emptyset$ then F^s is not equivalent to F^t , and (ii) $a \notin S^s$. A smart ballot can therefore be seen as a preference ordering over the agent’s desired delegations, with a backup vote at the bottom of the ordering. The validity requirement ensures the non-redundancy of the preference ordering and forbids delegation loops.

Once smart ballots are collected—by a central authority or on a decentralised ledger—they need to be “unravalled” into a standard voting profile, one in which each voter expresses either a direct vote or an abstention, to finally apply a standard voting rule and compute the result of the election. Thus, our object of study is the following:

Definition 2.3.2. An unravelling procedure \mathcal{U} for issue $i \in \mathcal{I}$ and agents in \mathcal{N} is any function associating with each valid smart ballot profile a profile of a direct votes: $\mathcal{U} : (B_{1i} \times \dots \times B_{ni}) \rightarrow D(i)^n$.

Smart ballot profiles are unravalled into profiles of direct votes to apply a voting rule

Note that while most of the related work in liquid democracy aims at assigning a final delegate (often named the “guru”) to each voter, our unravelling procedures aim at assigning a direct voting ballot to each delegating agent.

We followed two principles in defining four greedy unravelling procedures: use the highest preference level of voters when breaking delegation cycles, and keep the unraveling procedure polynomial. We propose four procedures (details are omitted in the interest of space and can be found in our paper [35]): basic update (**U**), update with direct vote priority (**DU**), update with random voter selection (**RU**), update with direct vote priority and random voter selection (**DRU**). We illustrate the first one in the following:

Example 2.3.1. Consider one binary issue i with $D(i) = \{0, 1\}$ and agents $\mathcal{N} = \{A, \dots, E\}$. Their valid ballots in profile B_i , stating their preferences for delegations or direct votes, are shown schematically in the table below:

	1 st	2 nd	3 rd
A	({B, C}, B ∧ C)	({D}, D)	1
B	1	-	-
C	({D}, D)	0	-
D	({E}, E)	1	-
E	({A}, A)	({B}, B)	0

We spell out here the unravelling $\text{UNRAVEL}(\mathbf{U})$ on B_i , a procedure which looks for direct votes at the highest preference level possible, propagating them via the delegations specified by the agents at the same preference level:

- Starting at the first preference level, the direct vote of B is stored in a partial direct vote vector $X = (\Delta, 1, \Delta, \Delta, \Delta)$, with Δ as a symbol denoting agents whose vote has not yet been computed.

- As there are no direct or computable votes at level 1 (no agent has a single-agent delegation to B), we move to the second level.
- The direct votes of C and D are thus added to X, and the vote of E is computed by copying the vote of B, obtaining a new $X = (\Delta, 1, 0, 1, 1)$.
- The preference level is set back to 1.
- A's vote can now be computed from her first preference level, and vector $X = (0, 1, 0, 1, 1)$ is thus returned as the unravelled profile of direct votes of the agents.

We show that the four greedy unravelling procedures we propose do terminate on any profile of valid ballots, and that there exists one such profile on which they output four different results. In order to prove that the computational complexity of running our unravelling procedures is polynomial, we need to fix a compact representation for the functions F^s formalising multiagent delegations (an explicit representation would render the rules trivially polynomial). We therefore restrict our attention to binary-valued issues, and define **BOOL** to be the set of smart ballot profiles where for each agent a and issue i we have that every F_{ai}^h in ballot B_{ai} is a contingent propositional formula in DNF. We show the following:

Our four greedy unravellings terminate in polynomial time

Proposition 2.3.1 ([35]). $\text{UNRAVEL}(\#)^{\text{BOOL}}$ terminates in at most $\mathcal{O}(n^2 \cdot \max_p(\mathbf{B}) \cdot 2 \max_\varphi(\mathbf{B}))$ time steps, for $\# \in \{\mathbf{U}, \mathbf{DU}, \mathbf{RU}, \mathbf{DRU}\}$.

Where n is the number of agents, $\max_p(\mathbf{B})$ is the highest preference level of any ballot in \mathbf{B} and $\max_\varphi(\mathbf{B})$ is the maximum length of any formula from an agent in \mathbf{B} . The figure in Proposition 2.3.1 is thus polynomial in the input of the procedures.

To distinguish further our four unravelling procedures from each other, and to be able to compare them with similar procedures proposed in the literature, we initiated an axiomatic study of two properties of participation, in line with the work of Kotsialou and Riley [103]. All results we show are restricted to smart ballots with single-agent ranked delegations only. We show that all four unravelling procedure we proposed satisfy *cast-participation*, provided that the voting rule used on the unravelled profile is monotonic. That is, agents who vote directly have an incentive to do so, rather than to express any other ballot. We also show that this property ceases to hold if we allow for multi-agent delegations. On the other hand, none of our unravelling procedures satisfy the property of *guru-participation*, which focuses on the incentive that an agent has to receive and accept a delegation. Indirectly, this result shows that our unravelling procedures differ from the depth-first and breadth first procedures proposed by Kotsialou and Riley [103].

We also studied two unravelling procedures based on minimisation. We call a vector \mathbf{c} a consistent certificate for a smart ballot profile \mathbf{B}

if each component corresponds to a delegation of the corresponding agent such that no delegation cycles is created. For instance, the certificate $(1, 3, 1)$ implies that agent 1 uses her first delegation, agent 2 her third, and agent 3 her first (with delegation cycles to be checked reading profile \mathbf{B}). The rule `MinSum` finds the consistent certificate that minimises the sum of its components, corresponding to the rule that minimises the overall ranks of delegations used to unravel a profile. Similarly, the rule `MinMax` finds the consistent certificate that minimises the maximum component, corresponding to a classical fairness criteria of minimising the delegation rank of the worst-off voter. We show that there are profiles of multiagent delegations on which the six rules considered all have different results, but that they all coincide on standard liquid democracy ballots, i.e., ballots where each agent only expresses a single-agent delegation and a backup vote, with the exception of `MinMax`.

We then focused on determining the computational complexity of `MinSum` and `MinMax`, showing that they are both NP-complete, and identifying a class of smart ballots on which they are polynomial. Let `BOOL` be the class of all contingent propositional formulas in DNF, and let `LIQUID` be the language of ranked single-agent delegations. We call `BOUNDEDMINSUM` the decision problem associated with the search problem of finding a solutions to `MinSum`. The input of `BOUNDEDMINSUM` is a smart profile \mathbf{B} , such that every ballot is restricted to `BOOL`, and a constant $M \in \mathbb{N}$. The problem then asks if there is a consistent certificate \mathbf{c} that unravels \mathbf{B} such that $\sum_{i \in \mathcal{N}} c_i \leq M$. An equivalent definition is given for `BOUNDEDMINMAX`. By reducing from the NP-complete problem `SAT-DNF` we obtain the following result.

Theorem 2.3.1. *`BOUNDEDMINSUM` and `BOUNDEDMINMAX` are both NP-complete.*

Investigating further the source of the computational complexity of these two rules, we were able to devise two polynomial algorithms for the computation of the corresponding unravelling, assuming that the agent ballots are in the `LIQUID` language.

Theorem 2.3.2. *Finding an outcome of `MinSum` and `MinMax` can be done in polynomial time when individual ballots are restricted to `LIQUID` (i.e., single-agent ranked delegations).*

The proof of Theorem 2.3.2 is obtained by a non-trivial adaptation of the Edmonds algorithm for minimum spanning arborescence trees for `MinSum` and a novel algorithm for `MinSum`.

Work on multiagent ranked delegations in voting is just at its beginning, and plenty of problems remain open. First, our initial study of participation should be complemented by more general results including multiagent delegations, as well as the analysis of further axiomatic properties that are typically studied in voting. Second, as the

Unravelling by minimising the sum or the max of the delegation ranks

Unravelling based on minimisation is NP-complete

Unravelling based on minimisation is polynomial on single-agent delegations

*Smart voting should
be implemented on a
distributed ledger*

name “smart voting” suggests, its intended application is the development of collective decisions on a decentralised ledger, in line with recent initial work by Kotsialou et al. [104] and Riley et al. [136] on the design of electronic institutions. The report by Dhillon et al. [45] is a starting point in this direction, presenting a detailed analysis of smart contracts for electronic voting and liquid democracy.

2.4 PERSONALISED RATINGS

*We study the
bribery-proofness of
two rating systems*

In this section we change the type of collective decision problem under scrutiny, moving from a classical voting situation to the realm of ratings and recommendations (see, e.g., Bobadilla et al., 2013). This is the problem of a mechanism that has to provide to its users an assessment of the value of a given service-provider (we will use a restaurant as running example), based on the evaluations received by a subset of its users (which we will call customers). This problem is typically solved—and commercially exploited by well-known rating systems such as Tripadvisor—by taking the average of the evaluations received from the customers. In collaboration with Paolo Turrini and James Stewart [84–86] we compared this system with a personalised version, in which the rating provided to a user is the average of the evaluations of the user’s neighbours in a given social network. More precisely, we evaluate the possibility of bribery by the restaurant, who can invest some of its utility to obtain fake ratings from the customers to improve its performance. We show that personalised ratings are superior to what we call objective ratings under two aspects. First, from an algorithmic point of view we are able to devise an algorithm for finding optimal bribing strategies for objective ratings that runs in polynomial time, while we show that the same problem is NP-complete for the case of personalised ratings. Second, we show that the revenue of a bribing strategy for objective ratings increases with the number of users of the rating system, and is thus potentially unbounded, while we identify a bound in the personalised case that is dependent on the structure of the underlying social network.

We first need to introduce some notation and definitions to present our model. Call r the *restaurant* (or any service-provider), which is evaluated by a finite non-empty set of individuals $C = \{c_1, \dots, c_n\}$, called *customers*. Customers are connected in a network $E \subseteq C \times C$, which we assume to be symmetric and reflexive. $N(c)$ denotes the neighbourhood of customer c in E . The evaluations of the customers are represented as a function $eval : C \rightarrow Val \cup \{*\}$, where $Val \subseteq [0, 1]$ is a set of evaluation values (e.g., a discrete scale or the whole $[0, 1]$ interval) and $*$ is the symbol for no evaluation. Let $V \subseteq C$ be the set of customers that expressed an evaluation different than $*$, and call

them *voters*. The objective rating of the restaurant ignores the network and averages over all evaluations received:

$$\text{O-RATING}(eval) = \text{avg}_{c \in V} eval(c),$$

Personalised ratings, on the other hand, average only over one customer's neighbours:

$$\text{P-RATING}(c, eval) = \text{avg}_{k \in N(c) \cap V} eval(k)$$

The first important assumption we make is that the rating viewed by a customer corresponds to the propensity of a customer to visit the restaurant. Assuming also for simplicity that each customer is worth 1 utility point to the restaurant, we obtain the following formulas for the initial restaurant utilities under *eval* for O-RATING (u_O^0) and P-RATING (u_P^0):

The rating corresponds to the probability of visiting the restaurant

$$u_O^0 = |C| \text{O-RATING}(eval) \quad u_P^0 = \sum_{c \in C} \text{P-RATING}(c, eval)$$

If we now define a bribing strategy of a restaurant as a function $\sigma : C \rightarrow Val$ that specifies for each customer the amount of utility received from the restaurant, we can formulate our second important assumption as follows: the effect of a bribe of $x \in Val$ on a customer c is an increase of x on her evaluation (capped at 1). Formally, if $eval^\sigma$ represents the customers' evaluations after bribing strategy σ , then $eval^\sigma(c) = \min\{1, eval(c) + \sigma(c)\}$. Thus, we can define the utility of a bribing strategy σ for O-RATING (u_O^σ) and P-RATING (u_P^σ) as follows:

The effect of a bribe is linear on the user evaluation

$$u_O^\sigma = |C| \text{O-RATING}(eval^\sigma) - \sum_{c \in C} \sigma(c)$$

$$u_P^\sigma = \sum_{c \in C} \text{P-RATING}(c, eval^\sigma) - \sum_{c \in C} \sigma(c)$$

Finally, the revenue of a bribing strategy is the utility obtained from it minus the initial utility of the restaurant. Formally, $r_P(\sigma) = u_P^\sigma - u_P^0$ (an equivalent definition is given for r_O).

We illustrate the definitions of the model in the following examples:

Example 2.4.1 (O-RATING). Consider the three situations in Figure 3, illustrating the effect of a single bribe of 0.1 on each of the customers. The top-left (σ_A), top-right (σ_B) and bottom (σ_C) strategies, although costing the same (0.1) to the restaurant induce different returns. In particular $u_O^{\sigma_A} = u_O^{\sigma_B} = 3 \times 0.65 - 0.1 = 1.85$ while $u_O^{\sigma_C} = 3 \times 0.4\bar{3} - 0.1 = 1.2$. Observe how the first two give a positive change to the O-RATING while the third one does not. Hence, σ_A and σ_B are profitable strategies while σ_C is not.

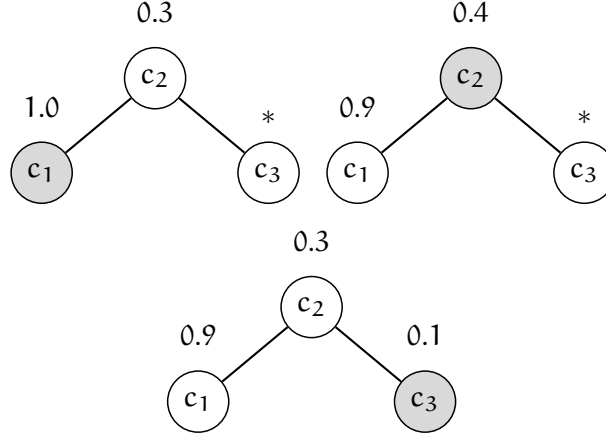


Figure 3: The effect of bribing different customers with the same amount of 0.1 under O-RATING and P-RATING. The bribed customer is in grey. Customers' evaluations are reported above each node.

Example 2.4.2 (P-RATING). Consider now the effect of the three strategies on P-RATING, starting from σ_A . We have that $\text{P-RATING}^{\sigma_A}(c_1) = \text{P-RATING}^{\sigma_A}(c_2) = 0.65$ while $\text{P-RATING}^{\sigma_A}(c_3) = 0.3$. Note already how bribing c_1 does not affect c_3 , whose P-RATING stays the same. We thus have that $u_p^{\sigma_A} = 1.5 = u_p^0$, which means $r_P(\sigma_A) = 0$ and thus σ_A is not a profitable strategy. Observe also that any strategy that behaves like σ_A , bribing only customer c_1 , gives either a zero or a negative revenue. Conversely, σ_B is different. In this case, we also have that $\text{P-RATING}^{\sigma_B}(c_1) = \text{P-RATING}^{\sigma_B}(c_2) = 0.65$. However, $\text{P-RATING}^{\sigma_B}(c_3) = 0.4$, which means that $u_p^{\sigma_B} = 1.6$ and thus $r_P(\sigma_B) = 0.1$. So, the strategy σ_B yields a positive revenue to the restaurant. Finally, σ_C is such that $\text{P-RATING}^{\sigma_C}(c_1) = 0.6$, $\text{P-RATING}^{\sigma_C}(c_2) = 0.4\bar{3}$ and $\text{P-RATING}^{\sigma_C}(c_3) = 0.2$, which means $u_p^{\sigma_C} = 1.1\bar{3}$ and thus $r_P(\sigma_C) = -0.3\bar{6}$. So, the strategy σ_C yields a negative revenue to the restaurant.

Our comparison of the two rating systems under the perspective of bribery presents a number of results depending on the information available to the restaurant about the customer network and the presence of non-voters. Our main results are obtained assuming that not all customers have expressed their opinion, and that the network relating the customers is known to the restaurant.

Let an *O-greedy strategy* be any efficient strategy that redistributes all the initial resources u_0^0 of the restaurant among the voters (without wasting utility). The following proposition shows that these are all and only the optimal bribing strategies for O-RATING:

Proposition 2.4.1 ([86]). Let $V \neq C$ and σ a strategy. The following three statements are equivalent:

- (i) σ is optimal for O-RATING,

- (ii) σ is an O-greedy strategy,
- (iii) σ yields a payoff of $\left(\frac{|C|}{|V|} - 1\right) \times \min\{u_O^0, \sum_{c \in V} (1 - \text{eval}(c))\}$.

Thus, a polynomial-time algorithm to find an optimal bribing strategy for O-RATING is the following: starting from an evaluation vector eval , distribute all available resources u_O^0 to the voters, without exceeding the maximal evaluation of 1. By either exhausting the available budget or distributing it all, we are guaranteed the maximum gain by Proposition 2.4.1.

A polynomial algorithm for finding optimal bribes under O-RATING

On the other hand, we show that for P-RATING the same problem is computationally intractable. Let us first formalise bribery as a decision problem (NVKL stands for non-voters and known network):

BRIBE-NVKL

Instance: Network (C, E) , evaluation eval_0 , $\rho \in \mathbb{Q}$

Yes-Instance: An instance of BRIBE-NVKL s.t. there exists a strategy σ with $r(\sigma) \geq \rho$

By providing a polynomial reduction from the known NP-complete problem of finding an independent set on 3-regular graphs, aka IS-REG(3) [63], we are able to show the following:

Optimal bribing under P-RATING is computationally intractable

Theorem 2.4.1 ([84]). BRIBE-NVKL is NP-complete.

We are also able to obtain simple quantitative results comparing the revenue that can be obtained under the two studied rating systems. We start from the following:

Proposition 2.4.2. Let σ be an efficient strategy s.t. $|B(\sigma)| = 1$, and let \bar{c} be such that $\sigma(\bar{c}) \neq 0$. Then $r_P(\sigma) < |N(\bar{c})|$.

Proposition 2.4.2 shows that increasing the number of individuals that are not connected to an agent that is being bribed does not increase the upper bound on the revenue of the bribing strategy for P-RATING. This is not true when we use O-RATING:

The revenue under P-RATING is bounded by the number of neighbours

Proposition 2.4.3. Let σ be an efficient strategy bribing voters only. The revenue $r_O(\sigma)$ of σ is monotonically increasing in the number of non-voters.

Thus, under realistic assumptions on the network structure, such as a the presence of a very large proportion of non-voters and with customers having a small number of connections, these results show that bribing under O-RATING is increasingly rewarding, while under P-RATING this is no longer the case.

While our model is built on simplifying assumptions that may limit its relevance in practical applications—some of which can be easily generalised to, e.g., weighted networks, or weighted averages for the computation of ratings—our results show that taking into account the social network that relates the users of a rating system can be the key to devise mechanisms that are resistant to bribery.

2.5 CONCLUSIONS AND PERSPECTIVES

This chapter presented research on the multiple consequences of taking into consideration the social network connecting the agents involved in a collective decision-making process. A big effort was directed towards adapting and designing diffusion models to fit voting situations, where the opinion of the agents can take a complex form (Section 2.1). An important problem that remained open is the conjecture that asynchronous issue-wise diffusion with arbitrary integrity constraints (or at least with the constraints induced by preference orders) asymptotically terminates *on any network*. I also presented initial studies on the effects of social network phenomena on collective decisions, be it social influence (Section 2.1.5) or information filtering (Section 2.2). Recent papers in the AI literature also perform similar studies [46, 152], and there is ample ground for further investigations. A mechanism design perspective is however more likely to lead to impactful and lasting results. The study of delegative voting systems such as liquid democracy is a thriving topic in AI at the moment, and our proposed model of multiagent ranked delegations seems to have the right generality to encompass and analyse the various approaches proposed in the literature (Section 2.3). Also, while the study of personalised ratings presented in Section 2.4 is rather complete, the model should be tested under more realistic information assumptions (such as the restaurant knowing a distribution of the customers' evaluations) as well as with tools from the recommender systems literature.

Going further in research on social choice and social networks would require, among other things, better models of strategic reasoning in social influence. Most work focused on strategic actions by an external agent to control an election, such as adding or deleting edges of a network [26], or manipulating individuals' opinions [58, 153] and opinion polls [7]. What is missing however is a model for the strategic actions of an agent involved in a social influence process preceding an election. The model described in Section 2.1.6 is an initial attempt in this direction that simply allows agents to disclose or retain their otherwise sincere opinion. Recent work by Shepherd and Goldsmith [144] builds on the same assumptions as our model, but uses reinforcement learning for modelling the agents' strategic abilities. An interesting connection could also be drawn with voting games, and in particular with the results presented in Chapter 3, and Section 3.3 in particular: substituting the agents' monetary transfers with actions of social influence might lead to an alternative mechanism for equilibrium selection in aggregation games.

The theory of multiagent systems received modern foundations from economics, and more precisely game theory (see, e.g., [146, 154]). An autonomous agent interacting with other peers is endowed with actions, has (typically partial) knowledge about its environment and the other agents' situations, and guide its choice of action towards the satisfaction of a goal, often represented as the maximisation of a utility function.

Voting is an interaction protocol to take collective decisions among multiple agents (typically at least 3), and as such it does not escape game-theoretical and strategic considerations. From early studies in economics to the computational complexity of manipulation, a wide number of papers have investigated the strategic behaviour of voters, of the election organiser, or of the candidates.

This chapter summarises my work on strategic aspects of voting, which focused on strategic actions by the voters (in contrast to strategic actions by the candidates, or by an election organiser). Section 3.1 introduces the strategic structure of voting and discusses some relevant related work. In Section 3.2 we introduce Gibbard-Satterthwaite games, in which multiple manipulators are faced with strategic actions in a one-shot game. Section 3.3 studies collective decisions on multiple binary issues, proposing pre-vote negotiations as a mean of equilibrium selection. Section 3.4 reports on several research directions on the framework of iterative voting, which studies agents dynamics on a voting game, presenting simulations on restricted dynamics, experiments with reinforcement learning agents, and behavioral experiments in the lab.

Most of the work presented in this chapter was initiated while I was a postdoc in Francesca Rossi's group at the *University of Padova*. The work on GS games was initiated during a visit of Arkadii Slinko from the *University of Auckland* in this university, and continued during his visits to the *University of Toulouse*. The interest of our research group on this topic was formalised by the organisation of the Workshop on Voting Games and Iterative Voting,¹ financed by the *COST IC1205* action on computational social choice. The experimental work on iterative voting resulted from the collaborative project *ANR-14-CE24-0007 Cocorico-Codec*, which successfully brought together the French communities in computer science, political science, and economics, around various issues in contemporary social choice research. The use of reinforcement learning in iterative voting is a fascinating col-

¹ www.irit.fr/~Umberto.Grandi/iterative

laboration that aims at posing the first bricks towards the desing of AIs for voters' support.²

3.1 CONTEXT: THE GAME-THEORETIC STRUCTURE OF VOTING

One of the cornerstones of social choice theory, the theorem by Gibbard [64] and Satterthwaite [139], states that there exists no onto, non-dictatorial, and strategy-proof voting rule for more than three candidates. Strategy-proofness of a voting rule requires that in all preference profiles no voter has an incentive to vote insincerely, provided that all other voters keep their vote fixed. Equivalently, in a sincere voting profile no voter regrets having voted sincerely, once the outcome and each others' votes is made public. In game-theoretic terms, strategy-proofness can be expressed as the following requirement. Construct a game in which voters are the players, the set of actions available to players are linear orders over the candidates, and their utility is defined by their sincere preference order over the results of a voting rule, which associates a winning candidate with each profile of rankings. If the profile of sincere votes is a Nash equilibria in any such game then the rule is strategy-proof.

Voting can be studied as a game

Similar game-theoretic constructions were used extensively in the literature on economics, and more recently computer-science, to analyse the strategic structure of voting. A first observation—already done by Farquharson [59]—is the existence of a multitude of meaningless Nash equilibria in which voters submit irrational votes. If we take a simple example of a plurality vote, and we assume all voters have the same sincere preference over the candidates $a > b > c$, then the sincere vote profile (a, a, a) is a Nash equilibria, but so are (b, b, b) and (c, c, c) which are clearly Pareto-dominated, but still do not allow for individual deviations. This observation calls for refinements of the basic game-theoretic model of voting or of the solution concepts considered. Much work has been published in economics, starting from the seminal work of Farquharson [59], with one of the most popular models being that of Myerson and Weber [123], which enriches voting games using utilities. Recent work in computer science have focused on equilibrium refinement, e.g., the models of truth-biased voters [129] and lazy voters [44] which introduce micro-incentives to eliminate undesirable equilibria. The study of best-response dynamics on a voting game is still a popular subject in computer science under the name of iterative voting (for a survey see, e.g., [118]). The report of the Dagstuhl seminar on Computation and Incentives in Social Choice provides a good overview of recent research on the context of this chapter [51].

Refinements of voting games to avoid irrational equilibria

² A vision I sketched in a position paper [73], also shared by Cesar Hidalgo in his recent TED talk (www.ted.com/talks/cesar_hidalgo_a_bold_idea_to_replace_politicians).

3.2 GIBBARD-SATTERTHWAITE GAMES FOR k -APPROVAL

The Gibbard-Satterthwaite theorem implies that strategic voting is unavoidable, but focuses on the specific perspective of a single voter who has complete knowledge of the other voters' preferences. The following example illustrates what happens when more than one such strategic voter or manipulator is assumed to vote strategically:

Example 3.2.1 ([76]). *The table below describes a voting situation, in which four voters rank five candidates a, b, c, d, w from the most to the least preferred, with the most preferred candidates on the left of the lists. We assume that the voting rule is the Plurality rule, i.e., the winner is the candidate with the highest number of first positions, and that ties are resolved following a predetermined ordering over the candidates $w > a > b > c > d$.*

Voters	Preferences
Voter 1	a b c d w
Voter 2	b a c d w
Voter 3	w c a d b
Voter 4	d w a b c

In the situation described the plurality winner is w , thanks to tie-breaking. This result is the worst possible for the first two voters, who each have the opportunity to manipulate. Voters 1 and 2 are the only Gibbard-Satterthwaite manipulators (GS-manipulators) at this profile: voter 1's insincere strategy consists in voting for b instead of a , thus making b the winner, and voter 2 can vote insincerely in favour of a instead of b , thus making candidate a the winner. However, if both manipulate at the same time, their efforts will cancel out. If we zoom in on the situation when voter 1 and voter 2 are strategically choosing between their sincere vote and their manipulative vote, while other voters are not strategic, we see that voters 1 and 2 are playing an anti-coordination game.

Two GS-manipulators can face an anti-coordination game

Let us now extend our analysis to all the voters. We observe that voter 3 is happy and does not have reasons for strategising. Voter 4 does not have any incentive to manipulate: the current winner w is in his second position, so giving her more support will not change the outcome. For instance, the voter does not have any incentive to vote for w instead of d . However, this move is a very strong countermanipulation if voter 4 fears a strategic move from any of the first three players: giving additional support to w makes any manipulation impossible, ending all strategic considerations. Moreover, in game-theoretic terms this move weakly dominates his sincere vote.

In collaboration with Edith Elkind, Francesca Rossi, and Arkadii Slinko [49, 50, 76] we conducted an extensive analysis of the games that Gibbard-Satterthwaite manipulators can play, adding one more

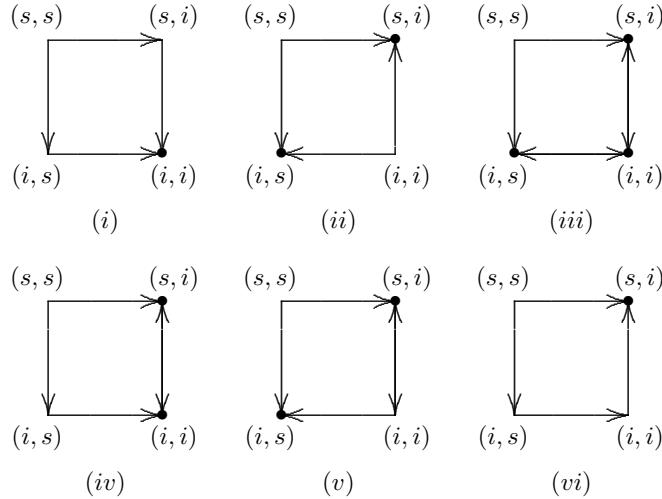


Figure 4: Diagrams for 2-by-2 GS-Games. Each vertex is a strategy profile for the two players, with s being the sincere strategy and i the insincere one. Vertical arrows indicate the preference of the first player, and horizontal ones those of the second player on the vote outcome. NE are denoted by thicker dots.

tile to the analysis of the cognitive and computational complexity of strategic voting.

To showcase some of our results, consider the diagrams in Figure 4. These are representations of all kind of games that two GS-manipulators can face, assuming that each of them have only one manipulation strategy (named i in the diagrams for insincere) other than their sincere one (named s). First, we observe that all GS-games so defined have at least one Nash equilibrium. However, some of them have two (they are anti-coordination games). For the plurality rule, we characterise which diagrams can be realised, i.e., we construct a voting profile in which there are two GS-manipulators with one manipulative strategy each, playing a game depicted in our diagrams, or we prove that there exists no such profile:

Theorem 3.2.1 ([76]). *The only diagrams of Figure 4 realizable by Plurality are (ii), (iii), (iv) and (v).*

In a companion paper, we adapted the cognitive hierarchy framework of Camerer, Ho, and Chong [25] to a voting situation, in an attempt to refine our model to bounded rational voters. The main difference is at the first level of the hierarchy, since we assumed that level-0 (non-strategic) voters play their sincere strategy rather than playing a random one. At level 1 voters are strategic voters who assume that all other voters will be at level 0, and hence will vote sincerely. At level 2 voters will assume that all other voters are either at level 1, which gives them the ability to countermanipulate. We do not go further than this level in our analysis, since there is evidence that

*GS-games for 2
players and 2
strategies have a
simple graphical
representation*

very few voters would be capable of higher-level reasoning (see, e.g. the work of Crawford, Costa-Gomes, and Iriberry, 2013).

In our paper we focus on the basic question that a voter at the level 2 of our hierarchy would face: knowing if one of her manipulation strategies weakly dominates truth-telling (or any other manipulation strategy). We analyse it from the perspective of computational complexity, and we find that under mild assumptions on the set of manipulation strategies this problem is easy for 1-approval (plurality) and 2-approval:

Proposition 3.2.1 ([50]). *Given a GS-game, where for each voter the set of strategies A_i consists of i 's truthful vote and a subset of i 's minimal manipulations, and two strategies v'_1, v''_1 of voter 1, we can decide in polynomial time whether v'_1 weakly dominates v''_1 .*

Manipulation weakly dominates truth-telling is polynomial for 2-approval

This result does not extend to 3-approval. Determining the complexity for this problem is still open, but we are able to show that from $k \geq 4$ the problem of determining whether a manipulation strategy is at the second level of the hierarchy or whether it is an improving strategy is computationally intractable:

Proposition 3.2.2 ([50]). *For every fixed $k \geq 4$, given a GS-game and a strategy v of voter 1, it is NP-hard to decide whether v is a level-2 strategy, and it is coNP-hard to decide whether v is an improving strategy.*

From $k \geq 4$ reasoning about GS-games is computationally intractable

The variety of the results we obtained show how rich the framework of GS-games is, and how complex the process of strategic voting can be to voters, from both a game-theoretic and a computational perspective. A recent position paper by Slinko [149] highlights a number of open problems and directions for future work on this topic.

3.3 EQUILIBRIUM SELECTION VIA PRE-VOTE NEGOTIATIONS

As observed in Section 3.1, the simplest construction of a voting game allows for undesirable Nash equilibria in which players vote irrationally. In collaboration with Davide Grossi and Paolo Turrini [75], we investigate this phenomenon in games where voters decide over multiple binary issues and have quasi-dichotomous preferences. Borrowing from the work of Jackson and Wilkie [96], we propose the introduction of pre-vote negotiations as a mean of equilibrium selection, in an attempt to model the complex process of influence that precedes a collective decision.

Our object of study are aggregation games, in which a set of voters have to decide on multiple binary issues, with an aggregation procedure deciding the collective outcome. We focus on non-manipulable aggregation rules such as quota rules and oligarchies, characterised by their sets of winning and losing coalitions. Agents are endowed with simple goals represented by conjunctions of literals from the set

We focus on multi-issue binary voting with quasi-dichotomous utilities

of issues, and a payoff function that associates a utility with each outcome of the aggregation rule. The strategic actions of the agents are thus guided by so-called quasi-dichotomous preferences, in which the goal represents the uncompromising positions of the voters, and the payoff allows to compare profiles that both satisfy or both do not satisfy the goal.

*Truthful equilibria
can be totally
inefficient*

We conduct a detailed study of equilibria in aggregation games. We are able to show that truthful equilibria, in which voters submit a ballot that satisfy their goals, can be totally inefficient, i.e., they can satisfy none of the agents' goals, even if goals are mutually consistent. However, if we allow for endogenous transfers, i.e., binding transfers of utility among players should a given strategy be played, we obtain two positive results. If we define an endogenous aggregation game as an aggregation game in which voters are allowed to sign binding agreements transferring utility before the game is played, and we call surviving NE any Nash equilibria of the underlying aggregation game that can be sustained by a transfer of utility in the resulting two-phase game, we are able to show the following:

*Efficient equilibria
can be sustained by
pre-play
negotiations*

Theorem 3.3.1 ([75]). *Let A be an endogenous aggregation game for a non-manipulable aggregator F . Every efficient and truthful NE of A is a surviving NE.*

Theorem 3.3.2 ([75]). *Let A be an endogenous aggregation game for a non-manipulable aggregator F in which a coalition C of voters have consistent goals, and such that C is a winning coalition of F for the conjunction of their goals. Then, every surviving NE of A is C -efficient.*

In words, the good Nash equilibria of an aggregation game (those that satisfy the goals of all the agents) are surviving NE and, conversely, all surviving NE of an aggregation game satisfy all the goals of a winning coalition of the underlying aggregation function (provided that these goals are consistent with each other).

*Boolean games with
shared control of
variables reduce to
exclusive control*

Endogenous transfers aside, the analysis of aggregation games could be pushed further, as they augment the well-studied setting of boolean games by adding variables that are under the shared control of a set of agents. In a related collaboration with Belardinelli et al. [8] we analysed aggregation games and related mathematical structures from the perspective of the verification of logical formulas expressed in the language of alternated-time temporal logic (ATL), presenting a polynomial reduction from shared control to exclusive control in boolean games.

3.4 ITERATIVE VOTING: CONVERGENT DYNAMICS, REINFORCEMENT LEARNING, AND LAB EXPERIMENTS

In iterative voting a discrete-time process is defined starting from a classical voting situation (i.e., a profile of preferences and a voting

rule), by providing information to the voters about the current state of the vote such as a poll or the current winner, and let them respond to this information by eventually changing their vote. This process is repeated until an equilibrium or a deadline is reached. Iterative voting has become an established research agenda in computational social choice in recent years, and three streams of work can be identified. First, iterative voting can model the iterative response to polls that precedes and election, with the main investigations being the convergence to an equilibrium (see among others, [111, 121], and [118] for a survey) and the possible manipulations by a central authority [7]. A second, normative, interpretation uses iterative voting to define novel voting rules, interpreting the iterative process preceding the decision as a pre-election deliberative step [2, 78, 130]. The *Itero*³ platform we developed can be seen as an attempt to popularise and study iterative voting rules. Third, each step in an iterative vote can be used to investigate the strategic voting abilities of each agent, typically by means of experiments in the lab [1, 119].

3.4.1 Restricted convergent dynamics

Starting from the observation that many of the dynamics initially proposed in this literature do not converge at a stable state (i.e., a Nash equilibria of the corresponding voting game), in a collaboration with Andrea Loreggia, Francesca Rossi, K. Brent Venable, and Toby Walsh [78], we focused on designing restrictions on the voters' actions that guarantee the convergence of the induced dynamics (and still have some descriptive value). Formally, we start from a set of voters each associated with a linear order P_i^0 over a set of candidates, and a voting rule F which associates a single winner with each profile \mathbf{P} of linear orders. Iterative voting is a discrete-time process, with one or possibly more voters strategically changing their vote at each time point $k = 0, 1, \dots$, thus defining a sequence of profiles \mathbf{P}^k for each time point k . The two restricted dynamics we considered are the following:⁴

Second-chance (SC): the manipulator i moves the second-best candidate in P_i^0 to the top of her reported ballot P_i^{k+1} , unless the current winner $w = F(\mathbf{P}^k)$ is already her best or second-best candidate in P_i^0 .

Best-upgrade (BU): the manipulator i moves the most preferred candidate in P_i^0 which is above $w = F(\mathbf{P}^k)$ in P_i^k to the top of her reported ballot P_i^{k+1} , among those that can become the new winner of the election.

³ <https://itero.irit.fr/>

⁴ The names used here are different than in the published version [78].

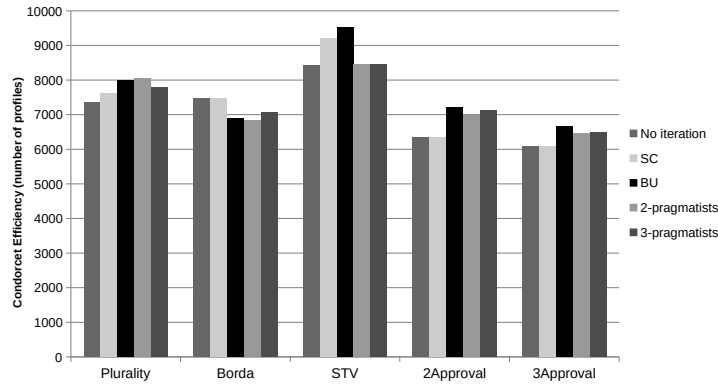


Figure 5: Increase in Condorcet efficiency for different voting rules, comparing the non-iterated profiles with the second-choice, best update, 2 and 3-pragmatists dynamics. The results are obtained with 10 voters and 25 candidates—simulating a Doodle vote—averaging over 10.000 profiles obtained with the Polya-Eggenberger urn model with 10% correlation.

We analysed theoretically the convergence of the two dynamics and the axiomatic properties of the iterative voting process, and evaluated experimentally the properties of the attractor profiles. To showcase one of our results, we tested the Condorcet efficiency (i.e., the percentage of Condorcet winners that are elected when one exists) of a voting rule in the set of initial profiles \mathbf{P}^0 compared with the corresponding ones obtained at convergence of the iterative process. Figure 5 shows that this parameter is increased by iterative voting under both restricted dynamics (for comparison we also plotted the results of the 2 and 3-pragmatists dynamics of Reijngoud and Endriss [135]). The improvement is limited, especially under correlated preferences, but still statistically significant.

Restricted dynamics in iterative voting increase the Condorcet efficiency

3.4.2 Reinforcement learning agents

Most dynamics defined in the literature are myopic and memory-less

Further convergent dynamics have been proposed in the literature, but all lacked some form of descriptive power: agents are assumed to be myopic since they do not look further than one step ahead in the iterative process, and memory-less since they forget all previous actions by themselves or by the other agents. While a full-fledged theoretical analysis of a more realistic agent seemed out of reach,⁵ in collaboration with Stéphane Airiau and Filippo Studszinski Perotto [2] we tested experimentally the performance of autonomous agents designed with reinforcement learning. Figure 6 showcases our main result. We again focused on the Condorcet-efficiency, showing that in a very simple setting of iterated plurality elections, with minimal

Reinforcement learning agents are able to reach good collective decisions in iterative voting

⁵ With the notable exception of the local dominance model of strategic voting [120].

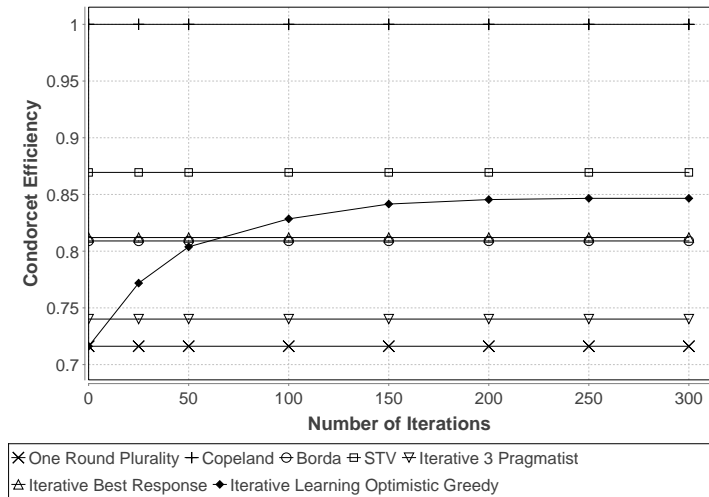


Figure 6: Performance of the reinforcement learning agents in terms of Condorcet efficiency, with 9 voters and 7 candidates. Averaged over 7331 profiles with a Condorcet winner generated with the Polya-Eggenberger urn model with 10% correlation.

information given to the voters (the winner of the previous election), our learning agents are able to “find” a Condorcet winner in approximately 85% of the cases. While the number of iterations required is high (about 300), this performance is superior to standard iterative voting dynamics as well as to many classical voting rules.

Further results in this line have been obtained in the Master thesis of Loujain Layka [108], which I supervised in 2019, who adapted our learning agent simulation to the case of multi-issue voting.

3.4.3 Behavioural experiments in multiple referenda

The model of multi-issue binary voting, e.g. multiple simultaneous referenda, is particularly interesting. It has been observed that in some cases the resulting combination of accepted and rejected issues can be the worst outcome for a majority or even all the voters (the so-called paradox of multiple elections [15]), and in recent years Bowman, Hodge, and Yu [14] proposed to use iterative voting to inform voters of the possible outcome and give them a chance to steer the result towards more favourable outcomes. In collaboration with Stéphane Airiau, Jérôme Lang, and Ali Ozkes [1] we designed a behavioural experiment to test such a protocol. A set of subjects each faced a different 2×2 -matrix with each cell containing their final potential gain should the cell be elected, and had to vote in a 2-issue referenda on which column (left-right) and which row (top-bottom) should be chosen. The winner would be made public and the election repeated until the same winning cell was elected 3 times in a row (or a

We tested the framework of iterative voting in multiple referenda in lab experiments

cap of 15 repeated elections was reached). The experiments were conducted at the Laboratoire d’Economie Experimentale de Paris (LEEP) in 2019, and our analysis of the results showed two main results. First, we analysed the behaviour of voters at the first step of the election, discovering that they tend to make optimistic choices when casting a vote on non-separable variables. Second, we observed an increase in social welfare at the end of iterative voting with respect to truthful voting, confirming the results of [14] which were obtained via computer simulations.

3.5 CONCLUSIONS AND PERSPECTIVES

In this chapter I presented several lines of research investigating the strategic behaviour of voters involved in a collective decision. The underlying model is that of a voting game, in which voters are players, actions are the possible voting ballots, and the players’ preferences are defined by combining a voting rule with a profile of sincere preferences given to the voters. The Gibbard-Satterthwaite games presented in Section 3.2 act as magnifiers that allow the researcher to restrict the number of manipulators and their action sets to analyse in depth a specific situation of strategic voting. There are a number of problems left open in our analysis, most notably closing the gap between the computational complexity of reasoning on level-2 strategies for 3-approval. In Section 3.3 we proposed pre-vote negotiations as equilibrium selectors in voting on multiple binary issues. An interesting link could be drawn here with the research line presented in Chapter 2 by assuming that such pre-vote negotiations need to be consistent with a predetermined trust network connecting the individuals. Section 3.4 is perhaps the most exploratory of the three main research directions presented. Iterative voting is under scrutiny by researchers for a number reasons. First, as a model of repeated interactions it can be rather close to what is experienced by voters who participate in multiple elections, view polls, and learn how to respond and behave in these situations. Second, it can suggest novel voting procedures which exploit the iteration phase to inform the voters of each others’ preferences. In particular, the use of reinforcement learning agents can be expanded by allowing agents to access a poll summarising the agents’ vote (in line with the recent work of Baumeister, Selker, and Wilczynski [7]), and their information structure can be refined by the introduction of a communication network.

3.5.1 *The Itero platform*

An important direction in my future research is the use of the *Itero*⁶ voting platform, which was developed in 2021 by Joseph Boudou dur-

⁶ <https://itero.irit.fr/>

ing his postdoc in the *ANR SCONE* research project. This platform allows the organisation of iterative elections using the plurality rule. Election organisers can decide the final deadline for the election as well as the duration of each iteration round. At each round, after the initial one, all voters can see the plurality score of the candidates at the previous rounds, and have the opportunity to update their vote. Figure 7 and 8 show screenshots of the voting interface and of the information available to the voters at each round.

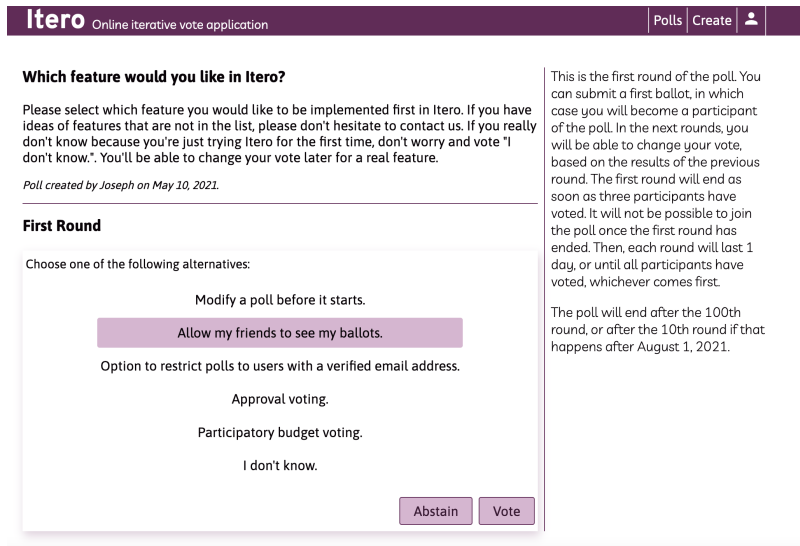


Figure 7: Screenshot of the voting interface on the *Itero* platform.

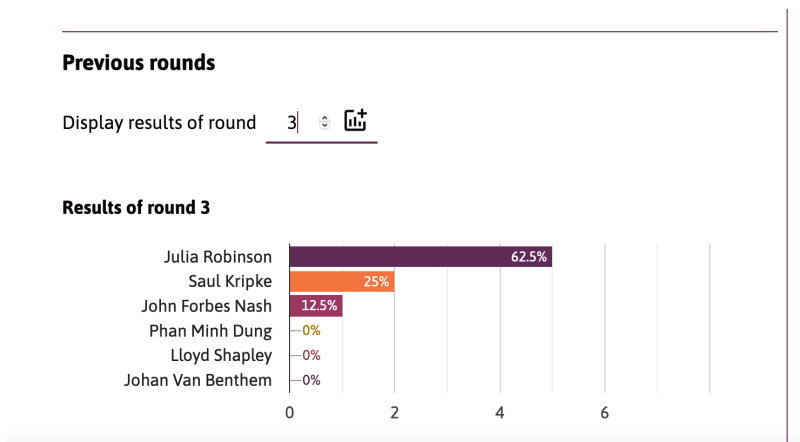


Figure 8: Screenshot of the results of previous rounds of an iterative election on the *Itero* platform.

The platform was developed with two purposes in mind. First, as a tool for outreach workshops (such as the *EU Researchers Night*), building on the successful experience of the *Whale*⁷ platform developed in

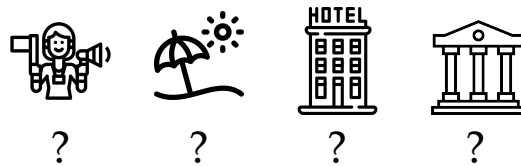
⁷ <https://whale.imag.fr/>

Grenoble. The underlying idea is to make voters aware that reliable information about other voters' preferences can be useful to avoid undesirable voting outcomes, formalising a common experience in countries that use plurality voting in their political elections (plurality with runoff for the case of France). Going further, iterative voting can be proposed as a simple deliberative platform, with the aim of reaching consensus, or at least an equilibrium outcome. Second, specific voting experiments could be run online (taking inspiration for instance from the *Voter Autrement* online experiments [13]) in order to obtain data about iterative voting and strategic voting.

EXPRESSIVE SOCIAL CHOICE

A crucial aspect of any collective decision process is the elicitation and the representation of the voters' preferences. To give a simple example, a vote over candidates with the plurality rule only elicits the most preferred candidate from voters, while a vote with STV or Borda asks for the entire ranking over candidates. Typically, the more information is elicited the better. But in many applications the combinatorial structure of the problem can significantly bind this process, adding computational or cognitive limitations.

Example 4.0.1 (Running example: automated travel planner). ¹ *An automated travel planner application collects the preferences of a group of travellers and creates a custom travel plan for the group. The list of possible activities for the day are the following:*



The group can do a guided tour of the city, visit the beach, have a stopover at the hotel during the day, and visit a museum. A travel plan is any subset of these activities.² What does the group want to do?

We first observe that in our running example there are $2^4 = 16$ possible travel plans. While it is not inconceivable to ask for a complete ranking of the 16 travel plans to each traveller, we can already expect some cognitive limitations, especially if we do not allow for ties. For instance, travellers can rank just their top three activities and then copy the order of presentation on the app for the remaining ones, with the risk that the 4th activity in the presentation order comes very high in the collective ranking without any grounding in the agents' real preferences. Moreover, as soon as the number of possible activities increases, the number of possible travel plans grows exponentially.

This chapter presents a number of compact languages for preference representation that can be implemented by the travel planner to elicit and then aggregate the preferences of the group members:

The combinatorial nature of the problem makes complete preference elicitation infeasible

¹ This example is inspired by the travelling group problem studied by Klamler and Pferschy [101].

² In contrast with the planning literature, the order in which the activities are executed is not relevant in our model. We assume that activities can be done in any order or will follow a predetermined order.

Judgments: the automated planner can ask yes/no questions on possibly complex propositions such as “do you want to take the guided tour and visit the museum?”, or “do you require a stopover at the hotel if we visit the beach?”. Section 4.1 presents a unified view on several frameworks for the aggregation of judgments, analysing the relative succinctness of the frameworks.

Goals: the members of the group can express any propositional formula over the activities, such as “I want to either visit the museum or do the guided tour, and I do not want to go to the beach”. Section 4.2 presents extensive work on goal-based voting, presenting a number of rules that associate travel plans with each collection of goal-formulas.

Conditional preferences: some clever members of the group might realise that they have preferential dependencies, such as “If we visit the beach I prefer the guided tour to the museum (as I’ll be dressed informally), otherwise I prefer the museum to the guided tour”. In Section 4.3 we use the framework of conditional preference networks (CP-nets) to compactly express such statements, expanding it with constraints and allowing for incomplete preferential statements.

Collective sentiment analysis: the automated travel planner is complemented with a bot that collects the travellers’ internet data and proposes a plan without having to ask for the group members’ preferences directly. Section 4.4 presents a proposal to expand classical sentiment analysis of text with the extraction and the aggregation of individual comparative judgments.

Work presented in this chapter is the result of numerous collaborations. I started working on a unified framework for judgment aggregation languages towards the end of my doctoral studies in Amsterdam, and the project advanced considerably thanks to a course I gave at ESSLI 2014 in Tübingen, during which Ronald de Haan (one of the co-authors) answered an open question I presented. The work on goal-based voting is the bulk of the PhD thesis of Arianna Novaro [126], which I supervised in Toulouse from 2016 to 2019. Work on conditional preferences started at the *University of Padova* with the PhD thesis of Cristina Cornelio [38] (probabilistic extension) and a collaboration with Nicolas Maudet and Hang Luo (constraints), and continued in Toulouse during the visit of Adrian Haret from *TU Wien* in our research group (incomplete CP-nets). Research on sentiment analysis also started at *University of Padova*, during the visit of Vijay Saraswat from *IBM TJ Watson laboratory*.

4.1 JUDGMENTS AND BINARY VIEWS

Consider once more the situation described in Example 4.0.1. Suppose that the preferences of the travellers over the four possible activities are elicited by means of the following questions:

- Q1:** Do you want to include a stop-by at the hotel?
Q2: Do you want to take the tour of the city centre or visit the museum (or both)?
Q3: Do you require a stop-by at the hotel if a visit to the beach is scheduled?
Q4: Do you want to do at least one of the three activities (tour, museum, beach)?

Observe that the questions are not logically independent. There are (at least) two possible choices in the design of such an automated planner. The first is to represent each question as a propositional formula, with the travellers' answers thus being consistent sets of such formulas. If we denote with propositional variable T the guided tour, B the beach, H the hotel, and M the museum, then question $Q2$ corresponds to formula $T \vee M$, and $Q3$ to formula $B \rightarrow H$.

The second possibility is to pre-process the questions in a constraint, and represent the travellers' answers as true/false assignments to propositional variables Q_1, Q_2, Q_3, Q_4 that are required to satisfy the constraint. In the example defined above such constraint would correspond to formula:

$$(Q_1 \rightarrow Q_3) \wedge [(Q_4 \wedge Q_3) \rightarrow Q_1] \wedge (Q_4 \rightarrow Q_2)$$

In collaboration with Ulle Endriss, Ronald de Haan, and Jérôme Lang [54], we compared four possible specification languages for such applications in terms of their succinctness, i.e., the ability of a language to compactly represent the specification of a problem such as the one described in the previous example.

We begin by introducing the basic language which represents sets of feasible combinations over m items by explicitly listing them:

Definition 4.1.1. *The basic language (BASIC) for the specification of collective decision problems over logically structured domains is $\mathcal{L}_0 = \{X \mid X \subseteq \{0, 1\}^m, X \neq \emptyset, m \in \mathbb{N}\}$.*

We then define four compact *specification languages*, each composed by a set \mathcal{L} of specifications and an interpretation function $\tau : \mathcal{L} \rightarrow \mathcal{L}_0$. In the interest of space, we give only two definitions corresponding to the well-studied settings of formula-based judgment aggregation (see, e.g., [6] or [52]) and binary aggregation with integrity constraints [74]:

*We focus here on
formula-based
judgment
aggregation and
binary aggregation
with constraints*

Definition 4.1.2. *The language of judgment aggregation (JA) is defined as follows:*

$$\begin{aligned}\mathcal{L}_{JA} &= \{\{\phi_1, \neg\phi_1, \dots, \phi_m, \neg\phi_m\} \mid \phi_j \in \mathcal{P}\mathcal{L}, m \in \mathbb{N}\} \\ \tau_{JA}(\Phi) &= \mathcal{J}(\Phi),\end{aligned}$$

where $\mathcal{P}\mathcal{L}$ is the language of propositional logic and $\mathcal{J}(\Phi)$ is the set of complete and consistent subsets of Φ .

Definition 4.1.3. *The language of binary aggregation with integrity constraints (IC) is defined as follows:*

$$\begin{aligned}\mathcal{L}_{IC} &= \{\Gamma \mid \Gamma \in \mathcal{P}\mathcal{L}_{\{p_1, \dots, p_m\}} \text{ satisfiable}, m \in \mathbb{N}\} \\ \tau_{IC}(\Gamma) &= \text{Mod}(\Gamma),\end{aligned}$$

where $\text{Mod}(\Gamma)$ is the set of models of the formula Γ , and $\mathcal{P}\mathcal{L}_{\{p_1, \dots, p_m\}}$ is the language of propositional logic over m variables.

In words, the language of judgment aggregation requires voters to submit consistent subsets of m propositional formulas built over an arbitrary set of variables, while binary aggregation with constraints asks voters to submit yes/no answers to m propositional variables subject to an integrity constraint. An example of the two languages was provided at the beginning of this section.

All frameworks we study are equally expressive, i.e., the interpretation functions τ_{JA} and τ_{IC} are surjective. We then focus on the relative succinctness of the two languages, inspired by the work of Cadoli et al. [23] and Gogic et al. [65]:

Definition 4.1.4. *Given two specification languages \mathcal{L}_1 and \mathcal{L}_2 , we say that \mathcal{L}_1 is **at least as succinct as** \mathcal{L}_2 , and write $\mathcal{L}_1 \preceq \mathcal{L}_2$, if there exists a function $f: \mathcal{L}_2 \rightarrow \mathcal{L}_1$ and a polynomial p s.t., for all $X \in \mathcal{L}_2$, we have:*

- $f(X) \equiv X$, and
- $\text{size}(f(X)) \leq p(\text{size}(X))$.

A straightforward result can be obtained by observing that propositional formulas are strictly more succinct than lists of feasible models:

Proposition 4.1.1. *IC is strictly more succinct than BASIC.*

We then show the following less straightforward result:

Theorem 4.1.1 ([54]). *JA is strictly more succinct than IC, unless the Polynomial Hierarchy collapses.*

That is, representing a logically structured domain for a collective decision by means of propositional formulas leads to more compact representations than a pre-processing of the problem in a constraint. This gain in the size of the representation comes with the computational cost of checking if an individual judgment is valid: it is NP-hard to check if a JA set of formulas is consistent, while it is polynomial to check if a IC specification is a model of the constraints.

However, in our paper we compared the computational complexity of higher-level tasks related to the aggregation of the individual judgments, most notably the computational complexity of winner determination for a number of aggregation rules. We showed that most of the time the computational complexity of winner determination does not vary between the two specification languages, with few notable exceptions.

The computational complexity of aggregation tasks does not differ in the two frameworks

4.2 GOAL-BASED VOTING

The travelling group problem (recall the running Example 4.0.1) can be approached as a multiple referendum, asking travellers a yes/no vote on each of the four activities. Such a choice would cause well-known problematic issues, such as those pointed out by Brams, Kilgour, and Zwicker [16], but we want to focus here on the voters' problem of deciding how to vote. Suppose for instance that there are three agents: Ann who wants to do all activities, Barbara who only wants to visit the museum, and Camille who wants to do one activity among the guided tour, museum, or beach, but no more than one. In a multiple referenda they might vote as follows:

	Guided Tour	Beach	Hotel	Museum
Ann	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Barbara	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Camille	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

By taking a majority vote on each of the issues, the result would see Camille particularly unhappy, since the group will visit both the beach and the museum. Had she known the other travellers' votes she would have supported the museum, or no activity at all.

In collaboration with Arianna Novaro, Dominique Longin, and Emiliano Lorini [126–128], we studied extensively the possibility of letting voters express a set of possible plans by means of propositional formulas over atoms representing the issues at stake. Such a setting corresponds to social choice with dichotomous preferences, with the combinatorial aspect of the problem making the aggregation non-trivial. It is also the same setting as the one studied in belief merging (see, e.g., [102]), while the normative properties we wish to impose on our aggregators are different. In particular, we observed a tension between fairness properties of a goal-based voting rule and its decisiveness or resoluteness, i.e., its ability to output a unique decision in most situations. One would indeed want to avoid an automated travel planner to output a large set of equally preferred travel plans, requiring for instance a second vote to choose which of the equally preferred travel plans to implement.

Each voter submits a propositional goal over the issues

We observe a tension between fairness and resoluteness

Let us identify this tension formally. The setting of goal-based voting sees a non-empty finite set of agents $\mathcal{N} = \{1, \dots, n\}$ taking a collective decision over a set of binary issues $\mathcal{J} = \{1, \dots, m\}$. We assume all issues to be independent from one another. Each agent expresses an individual goal as a propositional formula over atoms in \mathcal{J} , which we shall call γ_i . For instance, in the example described above the set of issues is $\{T, B, H, M\}$ and Camille's goal is $T \vee B \vee M$. A goal-profile $\Gamma = (\gamma_1, \dots, \gamma_n)$ is the choice of a goal by each of the voters, and a goal-based voting rule F is a collection of functions for all $n, m \in \mathbb{N}$ that associates with each goal profile a non-empty set of issues: $F : (\mathcal{L}_{\mathcal{J}})^n \rightarrow \mathcal{P}(\{0, 1\}^m) \setminus \emptyset$. The strongest definition we can give to formalise the resoluteness of a rule is the following:

Resoluteness (R) A rule F is resolute if and only if on all goal profiles Γ we have that $|F(\Gamma)| = 1$.

This condition can be relaxed to account for cases in which the electorate is equally split on an issue, allowing the rule to abstain:

Weak-resoluteness (WR) A rule F is weakly resolute if and only if for all profiles Γ and issues $j \in \mathcal{J}$ there exists a partial conjunction φ of literals in \mathcal{J} such that $F(\Gamma) = \text{Mod}(\varphi)$.

Clearly, if a rule F is resolute then it is also weakly resolute. Let us now introduce two axiomatic properties pertaining to fairness, the first requiring that all agents are to be treated equally, and the second that no issue is biased towards acceptance or rejection:

Anonymity (A): A rule F is anonymous if for any profile Γ and any permutation $\sigma : \mathcal{N} \rightarrow \mathcal{N}$, we have that $F(\gamma_1, \dots, \gamma_n) = F(\gamma_{\sigma(1)}, \dots, \gamma_{\sigma(n)})$.

Duality (D): A rule F satisfies duality if for all profiles Γ we have that $F(\bar{\gamma}_1, \dots, \bar{\gamma}_n) = \{(1 - v(1), \dots, 1 - v(m)) \mid v \in F(\Gamma)\}$, where $\bar{\gamma} = \gamma[\neg 1 \mapsto 1, \dots, \neg m \mapsto m]$.

In line with a well-known impossibility in classical social choice theory [122], we prove the following impossibility:

Proposition 4.2.1 ([126]). *There is no resolute rule F satisfying anonymity and duality.*

Let us now consider a different fairness axiom, already considered in the literature on belief merging (known as IC2 in this literature):

Model-unanimity (IC2): if the conjunction of the individual goals $\bigwedge_{i \in \mathcal{N}} \gamma_i$ is consistent, then $F(\Gamma) = \text{Mod}(\bigwedge_{i \in \mathcal{N}} \gamma_i)$.

While model-unanimity seems to be a harmless condition to assume, we show the following simple impossibility result, formalising a deeper tension between resoluteness and fairness:

Proposition 4.2.2 ([126]). *There is no rule F that satisfies model-unanimity and is weakly resolute.*

Let us now consider the adaptation of a classical—and heavily criticised—axiom from social choice theory. Let $D_m = \{(a, b) \mid a, b \in \mathbb{N} \text{ and } a + b \leq 2^m\}$ and $C = \{\{0\}, \{1\}, \{0, 1\}\}$. Let also $m_i(j)$ be a two-component vector counting the total number of zeros and the total number of ones in the models of goal γ_i :

Independence (I): A rule F is independent if there are functions $f_j : D_m^n \rightarrow C$ for $j \in \mathcal{J}$ and $n, m \in \mathbb{N}^+$ such that for all profiles Γ we have $F(\Gamma) = \prod_{j \in \mathcal{J}} f_j(m_1(j), \dots, m_n(j))$.

This adaptation of the axiom of independence states the existence of functions that decide the outcome of a rule F on each issue independently from the result on other issues. Such property is particularly interesting for goal-based voting, thanks to the following proposition:

Proposition 4.2.3. *Every independent rule F is weakly resolute.*

We can therefore look into the class of independent rules to seek for goal-based voting rules that satisfy both resoluteness and fairness conditions. In our work we present three adaptations of the most classical of the independent rules: the majority rule. While the formal definitions are left out in the interest of space, as well as the formal definitions of the axioms of equality, neutrality, and monotonicity (for the details consult the PhD thesis of Arianna Novaro [126]), we are able to characterise axiomatically one of our rules as follows:

Theorem 4.2.1 ([126]). *A rule satisfies equality, independence, neutrality, anonymity, monotonicity, unanimity, and duality iff it is TrueMaj.*

Intuitively, the goal-based rule *TrueMaj* accepts a given issue if the overall sum of acceptances in the individual models, discounted by the number of models submitted by each agent, is greater than the same figure for the number of rejections. By Theorem 4.2.1 this rule seems to realise the sought trade-off between resoluteness and fairness in goal-based voting.

Given that goal-based voting rules need take a compactly represented input, their computational complexity is necessarily higher than their classical voting counterparts. We focused on the problem of winner determination, i.e., constructing the winning outcome from the voters' propositional formulas. We prove that the three adaptations of the majority rule we propose are PP-hard, where PP is the class of problems that are solvable by a non-deterministic Turing machine that accepts in strictly more than half of all non-deterministic choices if and only if the answer to the problem is yes:

Theorem 4.2.2 ([126]). *Winner determination for TrueMaj, EMaj, and 2sMaj is PP-hard.*

We adapt the classical independence axiom to goal-based voting

We define three variations of the majority rule: TrueMaj, EMaj, and 2sMaj

Winner determination for TrueMaj, EMaj, and 2sMaj is PP-hard

	\mathcal{L}^\wedge		\mathcal{L}^\vee		\mathcal{L}^\oplus	
	E	D	E	D	E	D
<i>EMaj</i>	SP	SP	M	SP	M	M
<i>TrueMaj</i>	SP	SP	M	SP	M	M
<i>2sMaj</i>	SP	SP	SP	SP	M	M

Table 1: Manipulation and strategy-proofness results for the three adaptations of the majority rule to goal-based voting, and the language restrictions \mathcal{L}^\wedge , \mathcal{L}^\vee and \mathcal{L}^\oplus . In particular E stands for erosion manipulation, D for dilatation, SP for strategy-proof, and M for manipulable.

The typical PP-complete problem is MAJSAT, asking for a given propositional formula ϕ if $|\text{Mod}(\phi)| > |\text{Mod}(\neg\phi)|$, and the theorem is proven by a reduction from a suitable modification of MAJSAT. Problems in this class are particularly hard, the class PP containing both NP and co-NP, and P^{PP} containing the entire polynomial hierarchy. We also provide the following upper bounds:

Theorem 4.2.3 ([126]). *Winner determination for TrueMaj and EMaj is in PSPACE, and for 2sMaj is in P^{PP} .*

In line with the strategic view described in Chapter 3 we also investigated the possible manipulation strategies of the voters when using propositional goals. We considered three main manipulative strategies: erosion, by submitting a more restrictive goal than the sincere one, dilatation, by submitting a less restrictive goal than the sincere one, and unrestricted manipulation. A common example of manipulation by erosion is the common behaviour on Doodle polls to report a less available agenda in order to steer the decision towards more preferred slots. As the outcome of a goal-based voting rule can contain more than one subset of the issues, we also need to consider several measures of satisfaction for assessing an agent’s manipulation move. Our findings show that manipulation is ubiquitous, albeit computationally hard (but no harder than computing the result of the rules themselves). Strategy-proofness can be guaranteed by restricting the language of possible goals to conjunctions (corresponding to judgment aggregation with abstentions) or to disjunctions, in which case we obtain strategy-proofness only with respect to dilatation. Our results are summarised in Table 1.

We consider three manipulative strategies: erosion, dilatation, unrestricted

Characterising a goal language for tractable winner determination remains open

There are two interesting problems that remain open. First, given the high computational complexity of computing the result of the three adaptations of majority we proposed, it would be interesting to identify a non-trivial subclass of propositional formulas over which winner determination is tractable (or easier than PP). Second, in a

similar way it would be interesting to identify the boundary between manipulable and non-manipulable instances of goal-based voting, by suitably restricting the language for goals.

4.3 CONDITIONAL PREFERENCES

In the previous two sections we dealt with languages to compactly represent dichotomous preferences over a combinatorial domain, by either expressing our best option or a set of options in the form of a propositional goal. In this section we consider a more complex form of preference named *ceteris-paribus*: an agent prefers the value of one of the features of the domain to another value, *all other things being equal*. Such statements can be represented compactly in a much-studied language called CP-nets [12] to express partial orders over a combinatorial domain.

Ceteris paribus preferences compare outcomes that differ on one variable only

Example 4.3.1. Consider the following CP-net, built on our running Example 4.0.1. Ann, one of the users of the automated travel planner, has the following preferences: she prefers going to the beach than not going, but if she goes to the beach she prefers a stop-by at the hotel and not going to the museum. She would also like to do the guided tour of the city. This can be represented compactly as follows: let \bar{T} indicates not doing the guided tour, and accordingly for three other binary variables B, H, M. The two variables B and T are the parent variables for Ann, since her preference on these two activities do not depend on the other activities. We can therefore write $T > \bar{T}$ and $B > \bar{B}$. The two variables H and M depend on the value of B, thus: $B : H > \bar{H}$, $B : \bar{M} > M$. Recall that preferences are interpreted all other things being equal, so the previous statements induce the following partial order over travel plans:

$$\begin{array}{rcl}
 TBH\bar{M} & > & TB\bar{H}\bar{M} & > & \bar{T}B\bar{H}\bar{M} \\
 & > & & > & \dots \\
 TBHM & > & TB\bar{H}M & \dots & \\
 & > & & > & \dots \\
 \bar{T}BHM & > & \bar{T}B\bar{H}M & \dots &
 \end{array}$$

A vast amount of papers considered generalisations of CP-nets and designed algorithms for their use. More recently, this line of research has been under the scrutiny of behavioural experiments, in an attempt to measure the practical use of conditional preferences by humans. The results are mostly negative, with lexicographic preferences being more common than conditional ones (preliminary results have been presented by [3]). Still, CP-nets have been an influential language in preference representation, and are still being applied in normative studies of preferences such as the recent literature on endowing artificial intelligences with ethical principles (see, e.g., [24, 114]).

My contributions in this area pertain to voting with conditional preferences, designing algorithms or analysing the computational complexity of computing the outcome of a known voting rule when the preferences of the voters is represented compactly using (generalisation of) CP-nets.

*Voting with CP-nets
under feasibility
constraints*

In collaboration with Hang Luo, Nicolas Maudet, and Francesca Rossi we tackled the problem of voting under constraints with CP-nets [82]. When the dependencies of the voters' CP-nets are aligned (technically, there exists an acyclic topological ordering of the union of the CP-nets dependency graphs) a sequential vote can be taken using any voting rule [105]. In our paper we considered a binary setting with the addition of a constraint that binds the possible outcomes of the vote. We characterise settings in which sequential vote by majority is guaranteed to satisfy the constraint, under very restrictive assumptions of CP-nets being separable and individuals' top outcomes satisfying the constraints. A more interesting result can be obtained when the constraint is binary, in which case sequential majority voting can be used with any profile of CP-nets that satisfies some conditions of coherence with respect to the constraint. Finally, we design an algorithm that adapts the sequential procedure of majority voting to satisfy the constraint when none of the above-mentioned restrictions hold. We show that the algorithm is correct and that it runs in polynomial time in the size of the compact input.

*We represent
multi-agent
preferences as a
probabilistic CP-net*

We take a different stance in a collaboration with Cristina Cornelio, Judy Goldsmith, Nick Mattei, and Francesca Rossi [39]. We first introduce and study a generalisation of CP-nets that can quantify the uncertainty of the individual preferences (called PCP-nets) and then use it as a representation of a profile of CP-nets. We define two methods to build a PCP-net from the users' preferences, and combine them with two algorithms we designed to extract the top outcome of the aggregated PCP-net. We analyse the two methods both axiomatically and experimentally, concluding that one of the two methods is superior as a voting procedure.

*We study the
aggregation of
gCP-nets*

Finally, building on the work of Rossi, Venable, and Walsh [137] that defined voting semantics over collections of CP-nets, and on the work of Goldsmith et al. [66] that generalised CP-nets to account for incomplete and more general logical CP-statements, in collaboration with Adrian Haret and Arianna Novaro we studied the aggregation of incomplete CP-nets [90]. The assumption of completeness of a CP-net needs to be relaxed with non-trivial dependency structures or in applications where the preference statements of the voters are elicited in an online or iterative fashion. In our paper we use gCP-nets, which allow for conditional statements involving more than one variable, such as "if we go to the museum and to the beach then I prefer not having a guided tour", which would be represented with formula $M \wedge B : \bar{T} > T$. We study the four semantics defined in

the literature for voting with multiple CP-nets, and show that most algorithmic tasks are PSPACE-complete. Albeit being mostly negative results from the point of view of computational complexity, our results show that the aggregation of gCP-nets is not harder than reasoning with a single gCP-net. However, the assumption of working with gCP-nets do not allow us to clearly separate hard from easy aggregation problems. The recent work of Lukasiewicz and Malizia [116] presents a clearer picture of the computational complexity of aggregation tasks by working with *complete* CP-nets, obtaining results ranging from polynomial algorithms to Σ_2^P -completeness.

4.4 PREFERENCE AND POLARITY

The abundance of user data on the Internet stimulated the development of applications such as recommender systems that are directly able to interact with a user or a group of users without having to elicit her preferences. A technique that is often at the basis of such applications is sentiment analysis, a collection of algorithms from natural language processing that can elicit a polarity (typically: positive, negative, or neutral) from textual data (for an introduction see, e.g., [113]). One can imagine the use of similar techniques in group decision-making applications such as group recommender systems, and in collaboration with Francesca Rossi and Vijay Saraswat we proposed a method inspired from social choice theory for collective sentiment analysis [77]. We started from the observation that classical sentiment analysis falls short as soon as the number of alternatives to be considered is larger than one:

Example 4.4.1 (Adapted from [77]). *Let us take a simplified example of our running Example 4.0.1. A fully automated travel planner has access to a group of users' textual data on previous trips as well as their social media textual activity. The problem is simple: it has to decide if the group prefers a visit to the museum to a day at the beach (only one of the two activities is possible). Of the 35 group participants, sentiment analysis techniques concluded that 20 people are talking positively about going to museums while 15 are talking negatively about it, and 30 people are talking positively about going to the beach while 5 negatively. However, what the participants express on social media is just the tip of their preferences, which we can assume being a ranking of which of the two activities they would choose if they were confronted with the choice. We assume that the preferences of the voters are represented in the following table, with activities to the left being preferred to those to the right, and the bar in the center signalling the threshold of positive vs negative polarity:*

20 people:	Museum	Beach	
10 people:		Beach	Museum
5 people:			Museum Beach
Sentiment analysis			Beach
Majority rule			Museum

The most popular alternative is not necessarily the best choice

With 30 participants out of 35 talking positively about going to the beach, against 20 for the museum, a sentiment analysis application would conclude that a trip to the beach is the right option to recommend to this group. However, when the group is presented with this choice, or maybe when they are already laying under the sun discussing about the possible travel options for the day, they will realise that 25 of them would have preferred going to the museum, with only 10 of them choosing a day at the beach.

A generalisation of the Borda rule to deal with both polarity and comparisons

The classical Borda rule—and all scoring rules in general—have been shown to be particularly flexible in adapting to non-conventional input from voters (see, e.g., the adaptations of the Borda rule to partial orders [41] and bucket orders [57]). In our paper we chose to adapt the Borda rule to the kind of input that can be extracted from textual information via sentiment analysis techniques. First, we define a unified structure called SP-structure which can represent both polarity and comparative statements over the set of alternatives. Then, we assign a score to each alternative depending on its polarity as well as on the number of alternatives that are preferred to it. We name such rule $B_{\underline{\alpha}}^*$ -rule, with $\underline{\alpha}$ being the vector of scores. This rule generalises both classical sentiment analysis aggregation (which would correspond to Approval voting over the alternatives marked with positive polarity) and the Borda rule:

Theorem 4.4.1 ([77]). *If a profile of SP-structures \mathbf{S} is purely preferential, then for all $\underline{\alpha}$ we have that $B_{\underline{\alpha}}^*(\mathbf{S}) = \text{Borda}(\mathbf{S})$. If a profile \mathbf{S} is purely sentimental, then for all $\underline{\alpha}$ such that $\alpha_2 = \alpha_3$ we have that $B_{\underline{\alpha}}^*(\mathbf{S}) = \text{Approval}(\mathbf{S})$.*

We also analyse the axiomatic properties of the $B_{\underline{\alpha}}^*$ -rule, we show that it runs in polynomial time in the size of the input (a profile of SP-structures), and that its communication complexity is $\Theta(nm^2)$ where n is the number of agents and m the number of alternatives (for comparison, the communication complexity of the Borda rule is $\Theta(nm \log(m))$). The social-choice specialist will be reminded of the Fallback voting rule [17], in which voters express a set of ranked and approved candidates. This rule can give different results than our $B_{\underline{\alpha}}^*$ -rule, which is tailored to handle incomplete extracted data rather than an elicited voting ballot.

We also conduct an empirical analysis to assess the performance of the $B_{\underline{\alpha}}^*$ -rule in incomplete domains. To do so, we generated profiles of SP-structures, computed the result of the $B_{\underline{\alpha}}^*$ -rule, and then deleted

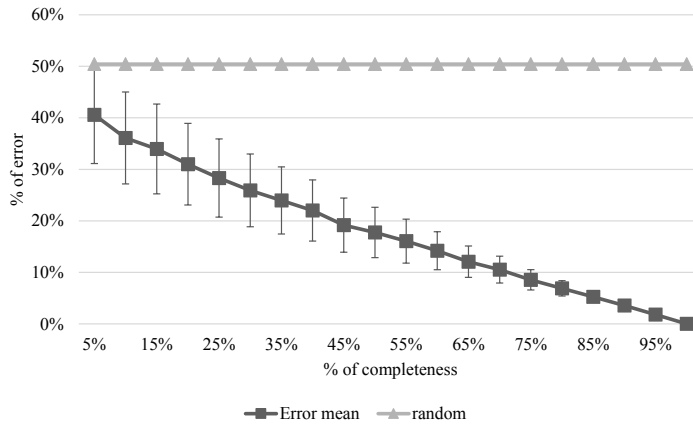


Figure 9: Mean error of the B_{α}^* -rule on incomplete profiles in terms of score.

a number of voters from the same profile, computing at each time the new result of the rule and its distance from the winner in the complete profile, using the same notion of score as for the computation of the rule. Our results are summarised in Figure 9, which shows that the rule always behave better than a random procedure, and that its accuracy grows quickly when completeness in increased.

4.5 CONCLUSIONS AND PERSPECTIVES

This chapter begun with a running example of social choice over a combinatorial domain, in which complete elicitation or representation of the agents' preferences is computationally intractable. We therefore introduced and studied compact languages for the representation of the agents' preferences, and rules for the computation of the collective outcome. We started in Section 4.1 with a comparison of four languages for the aggregation of judgments in terms of their relative succinctness. Section 4.2 instead assumed that the voters' input comes in the form of propositional goals, and proposed adaptations of the majority rule to this setting. Conditional preferences were introduced in Section 4.3, which presented three lines of research in voting with CP-nets: under constraints, using probabilistic CP-nets, and dealing with incomplete input. Finally, Section 4.4 considered input extracted from users' textual data, combining a polarity with comparative statements over the alternatives.

Overall, the languages we discussed in this chapter deal with applications that are purely discrete. In ongoing work with Linus Boes, Rachael Colley, Jérôme Lang, and Arianna Novaro we are generalising known aggregation rules to settings in which alternatives are associated to a numerical weight, such as participatory budgeting or collective scheduling, showing yet one more bridge from knowledge representation to collective decision-making. Building on the

framework of judgment aggregation with weights, first introduced by Nehring and Pivato [124], we are able to show that a number of previous problems—which we name *collective combinatorial optimisation problems*—and the corresponding aggregation rules can be viewed as instances of a more general definition. This class of problems includes participatory budgeting, collective scheduling, and collective spanning trees.

The variety of decision problems that can be modelled with a combinatorial approach, and the vast number of techniques developed in AI to deal with the resulting exponential domains, allow for an incredible number of bridges between knowledge representation, reasoning, and social choice. A road map and a number of open problems are sketched in an invited talk by Lang [106].

*Bridges between
knowledge
representation,
reasoning, and social
choice*

One of the main critiques to these approaches is the prohibitive computational complexity of the resulting representations or algorithms (e.g., in this chapter we saw computational problems in PP or in PSPACE). This is a fair point, and research in this area should have as a first priority to minimise complexity and maximise applicability. On the one hand, applications abound in which the scale of the problem is not central, and SAT or ILP solvers can be used efficiently to solve certain NP-hard problems. For instance, experiments on NP-hard participatory budgeting rules are often run using ILP solvers (see, e.g., [151]). On the other hand, depending on the importance of the collective decision being taken, it can be crucial to obtain guarantees on the time execution of collective decision methods on realistic instances. Moreover, human voters can often profit from more expressivity (as many of our examples showed), and goal-oriented autonomous agents have already a hard (computational) time to reason qualitatively about their options.³

³ A simple example: an autonomous agent who needs to decide about the most probable of two events described by two propositional formulas has to solve an instance of MAJ-SAT which is a PP-complete problem.

CONCLUSIONS AND PERSPECTIVES

This thesis presented an overview of my past research, in which I focused on enriching models of collective decision-making to take into account the interaction among decision-makers, their strategic behaviour, and to provide expressive yet compact means for preference elicitation and aggregation.

Interactive social choice

Chapter 2 presented a number of research lines arising from the addition of a social network—represented as a graph—connecting the participants of a collective decision. It is the longest chapter and my most active research line. I started by presenting a number of attempts at adapting diffusion models to social choice settings, in which individual opinions are represented as multiple binary views over possibly interconnected issues, or linear orders over a set of alternatives, and social networks are seen as representing trust or influence among agents. Such models were analysed theoretically, mostly focusing on the problem of termination of the diffusion process or on characterising the profiles at termination, as well as with multiagent simulations, analysing whether opinion diffusion can help aligning the preferences of the decision-makers. I also presented a model of strategic opinion diffusion in which agents can choose whether to exert their influence or retain from it towards the satisfaction of given individual goals. By taking a different interpretation of networks, seen as communication channels for information flow among the agents, I presented initial results on designing algorithms to identify phenomena generalising the majority illusion. In a final change towards a mechanism design perspective, I presented work on conceiving algorithms for eliciting and exploiting the trust networks among individuals towards better decision-making. I described two models, one that proposed a personalised version of collective ratings, and the other that enhances liquid democracy with multi-agent ranked delegations.

Strategic social choice

A long-lasting research problem in social choice theory analyses the strategic behaviour of voters, and in Chapter 3 I presented my contributions in this area. First, in the classical setting of k -approval voting, I introduced and studied Gibbard-Satterthwaite games, which are magnifying lenses on a strategic vote involving multiple voters.

Our results show, both analytically and with algorithmic techniques, that manipulation is complex, perhaps too complex for the voters involved in the game, with some exceptions in the easiest settings. Second, starting from a vote on multiple binary issues with agents provided with quasi-dichotomous utilities—a setting we called aggregation game—I presented the use of pre-vote negotiations and showed that they exclude undesirable equilibria and sustain efficient equilibria of the original aggregation game. Third, iterative voting considers best or better response dynamics in voting, either as a model of a repeated response to polls or as a voting procedure in itself. I presented an experimental study directed at designing best-response dynamics that are guaranteed to terminate, a behavioural experiment in the lab aimed at investigating real voters' behaviour in iterative voting, and work on the development of reinforcement learning agents that adapt their vote in multiple repeated elections and can reach collective decisions comparable to those of well-known voting rules.

Expressive social choice

Chapter 4 linked classical social choice with knowledge representation techniques, opening the possibility of designing and studying social choice procedures over combinatorial domains. Four settings were considered, in increasing order of expressive power of the voters' ballots. I presented a comparison of judgment aggregation frameworks in terms of their succinctness in representing combinatorial domains over multiple binary issues. In this setting agents express binary views over multiple issues, or a set of propositional formulas from an agenda. This assumption can be relaxed in two ways. First, I presented and analysed rules for the aggregation of propositional goals, which represent compactly dichotomous preferences over a binary combinatorial domain. Second, I studied two generalisations of conditional preference networks and the corresponding aggregation problems, one with probabilistic statements and the second one allowing incomplete specifications. I also presented work proposing rules for collective sentiment analysis, in which the agents' preferences are not elicited but extracted with natural language processing techniques from their textual expressions.

5.1 GOING FURTHER

In this section I summarize the main avenues research directions that I intend to pursue in the coming years. Open problems and detailed future research directions were described in the closing sections of each chapter.

Specific open questions and future work are discussed at the end of each chapter

Interactive democracy

Research in computational social choice has considerably expanded the kind of social choice problems considered in classical studies of collective decision making, as well as the set of techniques employed in their analysis, and this thesis provided several such examples. One of the factors behind the success of this contemporary view on social choice is the shift from almost-universal normative principles applied to political elections or to the construction of social welfare measures, to a problem-based approach on arguably lower-stake decisions such as participatory budgeting, committee elections, or school matching.

My future research will sit in this latter line of work, more precisely in the study of collective decision mechanisms which may involve a number of back-and-forth interactions between the voters and the mechanism, a research agenda that has been termed “interactive democracy” [20]. In particular, I plan to focus on three such interactive mechanisms:

- Starting from the research presented in Section 2.3 on **liquid democracy**, I plan to investigate the relation between delegative voting and opinion diffusion. The two models can be seen as being one the inverse of the other, e.g., the fact that agent *a* delegates to agent *b* can be seen as a formalisation that agent *b* influences agent *a*. Termination states of opinion diffusion models can therefore be used to characterise the possible results of delegative elections. Also, other collective decision-making mechanisms can be made delegative,¹ and a full-fledged model of delegations in multi-issue voting under constraints could be proposed as an encompassing framework.
- In **participatory budgeting**, a number of projects have to be funded by a collectivity with an overall budget to be respected. On the theoretical side, I plan to investigate further the relations between judgment aggregation and participatory budgeting, building on initial previous work developing an encompassing framework for collective discrete optimisation problem. On a more practical side, I started a collaboration with the *Mairie de Toulouse* which resulted in a student project analysing their data from their 2019 participatory budgeting campaign. In the long term I plan to extend and develop this collaboration in an attempt to popularise the numerous research results on participatory budgeting and at the same time obtain datasets for testing new rules, in line with recent research on this topic [107].
- A tension can be identified in many real-world situations involving a social choice: on the one hand the use of simple decision

¹ See e.g. initial work by Talmon et al. presented at the *Workshop of Liquid Democracy* in Toulouse, 2021 (www.irit.fr/scone/events/workshop-on-liquid-democracy/).

rules with bad properties is widespread, and on the other hand rules with good theoretical properties are often difficult to explain to voters and, most importantly, to election organisers. A classical example is the use of the plurality rule for political elections. **Iterative voting** proposes a possible solution to similar situations: the repetition of a collective decision with a bad rule can actually lead to results that have good theoretical properties. In future work I plan to investigate the use of iterative voting rules as simple deliberative processes. The *Itero* platform presented at the end of Chapter 3 is part of this agenda, and will be exploited in both outreach events and behavioural experiments.

Agent-mediated social choice

If we focus on those high-stake decisions that were at the core of classical research in social choice, such as political elections, there seems to be a growing consensus that democracy is experiencing serious problems entering in the XXIst century, but there is disagreement on how to cure them. Scientific and technological advancements in information technology are disrupting centuries-old institutions such as the banking and finance sector, and transportation: democratic institutions are even older, and there are repeated proposals to "redesign our political system" taking advantage of XXIst century innovations (see, e.g., [117]). Moreover, well-known scholars in computer science are beginning to propose encompassing visions for eDemocracy and its specifications [27, 143], complemented with more focused position papers in multi-agent systems [20] and media studies [42].

*Can solutions for
democracy's
problems come from
innovations in IT?*

When it comes to democracy, the use of AI techniques has rather bad connotations [91], the most notorious application having been the use of automated micro-targeting in political campaigns. In my opinion there are two reasons why current artificial intelligence techniques fall short in aiming at solving democracy's current problems: the restricted scope of widespread machine learning techniques, and the mainstream game-theoretic design of agents. Machine learning's most successful techniques (classifiers, deep learning...) seem to be less adapted to help improving low-frequency but high-stake decision-making such as elections. Corpora of past decisions comes in the form of history books, which are hard leaning material for a machine learning algorithm. Moreover, assessing the fitness or the quality of an election result—the basic requirement to feedback a learning automated agent—is a research question in itself, and heavily relies on the values that are being (again, collectively) adopted. If we then turn to research in multi-agent systems, we see that its modern foundations lie in rational decision theory and game theory [146, 154], i.e., on self-interested agents. Democracy certainly features many "egoistic"

aspects, but this is not the first word that we would use to characterise a democratic process.

The ideal Athenian dream of direct democracy is often idealised and aimed at by many techno-friendly citizens, including a number of researchers. Such a system seems however to suffer foundational problems such as a very low voter participation, often skewed towards the extremes of the political spectrum, eroding the legitimacy of its decisions. My proposal is to assess the feasibility of voting avatars, autonomous agents that take the role of citizens in a direct democracy, voting on their behalf on the hundreds of questions on the daily political agenda [73]. In the same way as autonomous agents are taking the lead in finance, trading, and eCommerce in general [43, 60, 131, 147], the time might be ripe for the development of autonomous agents acting as proxies of human voters in collective decision-making, and the design of voting rules specific to this new user set, making high-stake collective decisions with high-frequency a reality.

Leaving aside the practical problems of privacy and security, such a project requires the construction of autonomous agents that are able to understand the reasons behind a citizen's political vote or view, to be able to take part in collective decisions on her behalf, the development of voting and debating procedures to be used by autonomous artificial proxies, and the design of computational-friendly mechanisms for modelling social influence and the diffusion of opinions. Some such techniques have been already discussed at length in this thesis, with the main challenge in my opinion being the development of voting agents that are not purely game-theoretical. An initial step in this direction could be the study of goal-based agents in voting, i.e., agents whose rationality is based on satisfaction rather than maximisation, as in our research on goal aggregation presented in Chapter 4. To move further, a compact language for principles should be devised, together with inference mechanisms to decide which alternative is closer to such principles and decide the voting behaviour of the autonomous agents. Research in ethics and moral decisions in artificial intelligence is quite close and can suggest interesting starting points [37, 115]. A related approach has been taken by Stirling [150], who defines individual preferences as statements that are conditional on the choices of some of the other voters. His results assume a numerical representation of preferences, but his model of collective satisficing provides one of the few examples of collective decision-making that is not based on optimisation. One further idea is not to abandon the game-theoretical view at all, but let the voting avatars play an assistance game, or inverse cooperative reinforcement learning game [89, 125], in which the voting avatar would be able to learn which fitness function to use in its implementation from observing the behaviour of the voter.

Direct democracy suffers of low participation and cognitive load limitations

Autonomous agents acting as voters' proxies in high-stakes/high-frequency collective decisions

There are multiple ethical questions that this research agenda will indirectly raise. First, the actual use of voting avatars in collective decisions such as elections requires a level of security, data protection, and privacy which current techniques are not yet able to handle (but research is advancing fast in this area, as witnessed by recent applications of cryptography to collective decision-making [27] or the governance of electronic institutions [104]). Second, the problem of vote-selling would need to be carefully checked at the moment of prototyping. Last but not least, societies will have to decide whether foundational aspects of democracy should be left to autonomous agents which, albeit acting on behalf of a human, will still have a large degree of autonomy. Ongoing debates on ethical and societal consequences of artificial intelligence are preparing the ground for these decisions.

BIBLIOGRAPHY

- [1] Stéphane Airiau, Umberto Grandi, Jérôme Lang, and Ali Ozkes. “One-shot and iterative voting for multiple referenda: an experimental study”. In: *Under submission* (2020). URL: <https://ssrn.com/abstract=3747713>.
- [2] Stéphane Airiau, Umberto Grandi, and Filippo Studzinski Perotto. “Learning Agents for Iterative Voting”. In: *Proceedings of the 5th International Conference on Algorithmic Decision Theory (ADT)*. 2017.
- [3] Thomas E. Allen, Muye Chen, Judy Goldsmith, Nicholas Mattei, Anna Popova, Michel Regenwetter, Francesca Rossi, and Christopher E. Zwillig. “Beyond Theory and Data in Preference Modeling: Bringing Humans into the Loop”. In: *Proceedings of the 4th International Conference on Algorithmic Decision Theory (ADT)*. 2015.
- [4] Kenneth J. Arrow, A. K. Sen, and Kotaro Suzumura, eds. *Handbook of Social Choice and Welfare*. Vol. 1. Elsevier, 2002.
- [5] Vincenzo Auletta, Diodato Ferraioli, and Gianluigi Greco. “On the complexity of reasoning about opinion diffusion under majority dynamics”. In: *Artificial Intelligence* 284 (2020), pp. 103–288.
- [6] Dorothea Baumeister, Gábor Erdélyi, and Jörg Rothe. “Judgment Aggregation”. In: *Economics and Computation, An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*. Ed. by Jörg Rothe. Springer, 2016.
- [7] Dorothea Baumeister, Ann-Kathrin Selker, and Anaëlle Wilczynski. “Manipulation of Opinion Polls to Influence Iterative Elections”. In: *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2020.
- [8] Francesco Belardinelli, Umberto Grandi, Andreas Herzig, Dominique Longin, Emiliano Lorini, Arianna Novaro, and Laurent Perrussel. “Relaxing Exclusive Control in Boolean Games”. In: *Proceedings of the Sixteenth conference on Theoretical Aspects of Rationality and Knowledge (TARK)*. 2017.
- [9] Daan Bloembergen, Davide Grossi, and Martin Lackner. “On Rational Delegations in Liquid Democracy”. In: *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. 2019.
- [10] Jesús Bobadilla, Antonio Hernando, Fernando Ortega, and Abraham Gutiérrez. “Recommender systems survey”. In: *Knowledge-Based Systems* 46 (2013), pp. 109–132.

- [11] Sirin Botan, Umberto Grandi, and Laurent Perrussel. “Multi-Issue Opinion Diffusion under Constraints”. In: *Proceedings of the 18th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2019.
- [12] Craig Boutilier, Ronen I. Brafman, Carmel Domshlak, Holger H. Hoos, and David Poole. “CP-nets: A Tool for Representing and Reasoning with Conditional Ceteris Paribus Preference Statements”. In: *Journal of Artificial Intelligence Research* 21 (2004), pp. 135–191.
- [13] Sylvain Bouveret, Renaud Blanch, Antoinette Baujard, François Durand, Herrade Igersheim, Jérôme Lang, Annick Laruelle, Jean-François Laslier, Isabelle Lebon, and Vincent Merlin. *Voter Autrement 2017 for the French Presidential Election*. Tech. rep. HAL, 2019.
- [14] C. Bowman, J. K. Hodge, and A. Yu. “The potential of iterative voting to solve the separability problem in referendum elections”. In: *Theory and Decision* 77 (2014), 111–124.
- [15] S. Brams and W. Zwicker D. Kilgour. “The paradox of multiple elections”. In: *Social Choice Welfare* 15 (1998), 211–236.
- [16] Steven J. Brams, D. Marc Kilgour, and William S Zwicker. “The paradox of multiple elections”. In: *Social Choice and Welfare* 15.2 (1998), pp. 211–236.
- [17] Steven Brams and M. Remzi Sanver. “Voting Systems that Combine Approval and Preference”. In: *The Mathematics of Preference, Choice and Order*. Ed. by Steven J. Brams, William V. Gehrlein, and Fred S. Roberts. Springer, 2009.
- [18] Felix Brandt, Vincent Conitzer, Ulle Endriss, Ariel D Procaccia, and Jérôme Lang. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.
- [19] Robert Bredereck and Edith Elkind. “Manipulating Opinion Diffusion in Social Networks”. In: *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*. 2017.
- [20] Markus Brill. “Interactive Democracy”. In: *Proceedings of the 17th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2018.
- [21] Markus Brill, Edith Elkind, Ulle Endriss, and Umberto Grandi. “Pairwise Diffusion of Preference Rankings in Social Networks”. In: *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*. 2016.
- [22] Markus Brill and Nimrod Talmon. “Pairwise Liquid Democracy”. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. 2018.

- [23] Marco Cadoli, Francesco M. Donini, Paolo Liberatore, and Marco Schaerf. "Space Efficiency of Propositional Knowledge Representation Formalisms". In: *Journal of Artificial Intelligence Research (JAIR)* 13 (2000), pp. 1–31.
- [24] Roberta Calegari, Andrea Loreggia, Emiliano Lorini, Francesca Rossi, and Giovanni Sartor. "Modeling Contrary-to-Duty with CP-nets". In: *CoRR abs/2003.10480* (2020). arXiv: 2003.10480. URL: <https://arxiv.org/abs/2003.10480>.
- [25] Colin F. Camerer, Teck-Hua Ho, and Juin-Kuan Chong. "A cognitive hierarchy model of games". In: *The Quarterly Journal of Economics* 119.3 (2004), pp. 861–898.
- [26] Matteo Castiglioni, Diodato Ferraioli, and Nicola Gatti. "Election Control in Social Networks via Edge Addition or Removal". In: *The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI)*. 2020.
- [27] David Chaum. "Random-Sample Voting". In: (2016). URL: rsvoting.org/whitepaper/white\textunderscorepaper.pdf.
- [28] Daizhan Cheng, Hongsheng Qi, Zhiqiang Li, and Jiang B. Liu. "Stability and stabilization of Boolean networks". In: *International Journal of Robust and Nonlinear Control* 21 (2010), pp. 134–156.
- [29] Yann Chevaleyre, Ulle Endriss, Jérôme Lang, and Nicolas Maudet. "Preference Handling in Combinatorial Domains: From AI to Social Choice". In: *AI Magazine, Special Issue on Preferences* 29.4 (2008), pp. 37–46.
- [30] Zoé Christoff and Davide Grossi. "Binary Aggregation with Delegable proxy: An Analysis of Liquid Democracy". In: *Proceedings of the 16th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*. 2017.
- [31] Zoé Christoff and Davide Grossi. "Binary Voting with Delegable Proxy: An Analysis of Liquid Democracy". In: *Proceedings of the 16th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*. 2017.
- [32] Zoé Christoff and Davide Grossi. "Stability in Binary Opinion Diffusion". In: *Proceedings of the 6th International Conference on Logic, Rationality and Interaction (LORI)*. 2017.
- [33] Peter Clifford and Aidan Sudbury. "A model for spatial conflict". In: *Biometrika* 60.3 (1973), pp. 581–588.
- [34] Gal Cohensius, Shie Mannor, Reshef Meir, Eli A. Meir, and Ariel Orda. "Proxy Voting for Better Outcomes". In: *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. 2017.

- [35] Rachael Colley, Umberto Grandi, and Arianna Novaro. “Smart Voting”. In: *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*. 2020.
- [36] Rachael Colley, Umberto Grandi, and Arianna Novaro. “Smart Voting”. In: *Under submission* (2021).
- [37] Vincent Conitzer, Markus Brill, and Rupert Freeman. “Crowdsourcing Societal Tradeoffs”. In: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2015.
- [38] Cristina Cornelio. “Preference reasoning and aggregation over combinatorial domains in uncertain and multi-agent scenarios”. PhD thesis. Italy: Università di Padova, 2016.
- [39] Cristina Cornelio, Umberto Grandi, Judy Goldsmith, Francesca Rossi Nicholas Mattei, and K. Brent Venable. “Reasoning with PCP-nets in a Multiagent Context”. In: *Proceedings of the 14th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2015.
- [40] Vincent P Crawford, Miguel A Costa-Gomes, and Nagore Iriberrri. “Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications”. In: *Journal of Economic Literature* 51.1 (2013), pp. 5–62.
- [41] John Cullinan, Samuel K. Hsiao, and David Polett. “A Borda count for partially ordered ballots”. In: *Social Choice and Welfare* (2013), pp. 1–14.
- [42] Lincoln Dalberg. “Democracy via Cyberspace”. In: *New Media* 3.2 (2001), pp. 157–177.
- [43] Rajdeep K. Dash, Nicholas R. Jennings, and David C. Parkes. “Computational-Mechanism Design: A Call to Arms”. In: *IEEE Intelligent Systems* 18.6 (2003), pp. 40–47.
- [44] Yvo Desmedt and Edith Elkind. “Equilibria of plurality voting with abstentions”. In: *Proceedings of the 11th ACM conference on Electronic Commerce (EC)*. 2010.
- [45] Amrita Dhillon, Grammateia Kotsialou, Peter McBurney, Luke Riley, et al. *Introduction to Voting and the Blockchain: some open questions for economists*. Tech. rep. Competitive Advantage in the Global Economy (CAGE), 2019.
- [46] John A. Doucette, Alan Tsang, Hadi Hosseini, Kate Larson, and Robin Cohen. “Inferring true voting outcomes in homophilic social networks”. In: *Autonomous Agents and Multi-Agent Systems* 33.3 (2019), pp. 298–329.
- [47] David Easley and Jon Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. New York, NY, USA: Cambridge University Press, 2010.

- [48] Oya Ekin, Peter L. Hammer, and Alexander Kogan. "On Connected Boolean Functions". In: *Discrete Applied Mathematics* 96-97 (1999), pp. 337-362.
- [49] Edith Elkind, Umberto Grandi, Francesca Rossi, and Arkadii Slinko. "Gibbard-Satterthwaite Games". In: *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*. 2015.
- [50] Edith Elkind, Umberto Grandi, Francesca Rossi, and Arkadii Slinko. "Cognitive hierarchy and voting manipulation in k-approval voting". In: *Mathematical Social Sciences* 108 (2020), pp. 193-205.
- [51] Edith Elkind, Christian Klamler, Jeffrey S. Rosenschein, and M. Remzi Sanver. "Computation and Incentives in Social Choice (Dagstuhl Seminar 12101)". In: *Dagstuhl Reports* 2.3 (2012). Ed. by Edith Elkind, Christian Klamler, Jeffrey S. Rosenschein, and M. Remzi Sanver, pp. 1-22.
- [52] Ulle Endriss. "Judgment Aggregation". In: *Handbook of Computational Social Choice*. Ed. by Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. Cambridge University Press, 2016.
- [53] Ulle Endriss, ed. *Trends in Computational Social Choice*. AI Access, 2017.
- [54] Ulle Endriss, Umberto Grandi, Ronald de Haan, and Jérôme Lang. "Succinctness of Languages for Judgment Aggregation". In: *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR)*. 2016.
- [55] Bruno Escoffier, Hugo Gilbert, and Adèle Pass-Lanneau. "The Convergence of Iterative Delegations in Liquid Democracy in a Social Network". In: *Proceedings of the 12th International Symposium on Algorithmic Game Theory (SAGT)*. 2019.
- [56] Bruno Escoffier, Hugo Gilbert, and Adèle Pass-Lanneau. "Iterative Delegations in Liquid Democracy with Restricted Preferences". In: *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*. 2020.
- [57] Ronald Fagin, Ravi Kumar, Mohammad Mahdian, D. Sivakumar, and Erik Vee. "Comparing partial rankings". In: *SIAM Journal on Discrete Mathematics* 20 (2006), pp. 47-58.
- [58] Piotr Faliszewski, Rica Gonen, Martin Koutecký, and Nimrod Talmon. "Opinion Diffusion and Campaigning on Society Graphs". In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. 2018.
- [59] Robert Farquharson. *Theory of Voting*. Yale University Press, 1969.

- [60] Joan Feigenbaum, David C. Parkes, and David M. Pennock. "Computational challenges in e-commerce". In: *Communications of the ACM* 52 (2009), 70–74.
- [61] Noah E. Friedkin and Eugene C. Johnsen. *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*. Structural Analysis in the Social Sciences. Cambridge University Press, 2011.
- [62] Noah E. Friedkin, Anton V. Proskurnikov, Roberto Tempo, and Sergey Parsegov. "Network Science on Belief System Dynamics under Logical Constraints". In: *Science* 354.6310 (2016), pp. 321–326.
- [63] Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co., 1990.
- [64] A. Gibbard. "Manipulation of Voting Schemes: A General Result". In: *Econometrica* 41.4 (1973), pp. 587–601.
- [65] Goran Gogic, Henry A. Kautz, Christos H. Papadimitriou, and Bart Selman. "The Comparative Linguistics of Knowledge Representation". In: *Proc. 14th International Joint Conference on Artificial Intelligence (IJCAI)*. 1995.
- [66] Judy Goldsmith, Jérôme Lang, Mirosław Truszczyński, and Nic Wilson. "The Computational Complexity of Dominance and Consistency in CP-Nets". In: *Journal of Artificial Intelligence Research* 33 (2008), pp. 403–432.
- [67] E. Goles and J. Olivos. "Periodic behaviour of generalized threshold functions". In: *Discrete Mathematics* 30.2 (1980), pp. 187–189.
- [68] Michel Grabisch and Agnieszka Rusinowska. "A model of influence in a social network". In: *Theory and Decision* 69.1 (2010), pp. 69–96.
- [69] Michel Grabisch and Agnieszka Rusinowska. "A model of influence with an ordered set of possible actions". In: *Theory and Decision* 69.4 (2010), pp. 635–656.
- [70] Michel Grabisch and Agnieszka Rusinowska. "Different Approaches to Influence Based on Social Networks and Simple Games". In: *Collective Decision Making*. Ed. by Adrian Van Deemen and Agnieszka Rusinowska. Vol. 43. Theory and Decision Library C. Springer, 2010, pp. 185–209.
- [71] Michel Grabisch and Agnieszka Rusinowska. "A model of influence with a continuum of actions". In: *Journal of Mathematical Economics* 47.4–5 (2011), pp. 576–587.

- [72] Umberto Grandi. “Social Choice and Social Networks”. In: *Trends in Computational Social Choice*. Ed. by Ulle Endriss. AI Access, 2017. Chap. 9, pp. 169–184.
- [73] Umberto Grandi. “Agent-Mediated Social Choice”. In: *The Future of Economic Design*. Ed. by Laslier JF., Moulin H., Sanver M., and Zwicker W. Studies in Economic Design. Springer, 2019.
- [74] Umberto Grandi and Ulle Endriss. “Lifting Integrity Constraints in Binary Aggregation”. In: *Artificial Intelligence* 199-200 (2013), pp. 45–66.
- [75] Umberto Grandi, Davide Grossi, and Paolo Turrini. “Negotiable Votes”. In: *Journal of Artificial Intelligence Research* 64 (2019), pp. 895–929.
- [76] Umberto Grandi, Daniel Hughes, Francesca Rossi, and Arkadii Slinko. “Gibbard–Satterthwaite games for k-approval voting rules”. In: *Mathematical Social Sciences* 99 (2019), pp. 24–35.
- [77] Umberto Grandi, Andrea Loreggia, Francesca Rossi, and Vijay Saraswat. “A Borda Count for Collective Sentiment Analysis”. In: *Annals of Mathematics and Artificial Intelligence* 77.3 (2016), pp. 281–302.
- [78] Umberto Grandi, Andrea Loreggia, Francesca Rossi, Kristen Brent Venable, and Toby Walsh. “Restricted Manipulation in Iterative Voting: Condorcet Efficiency and Borda Score”. In: *Proceeding of the 3rd International Conference on Algorithmic Decision Theory (ADT)*. 2013.
- [79] Umberto Grandi, Emiliano Lorini, Arianna Novaro, and Laurent Perrussel. “Strategic Disclosure of Opinions on a Social Network”. In: *Proceedings of the 16th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2017.
- [80] Umberto Grandi, Emiliano Lorini, Arianna Novaro, and Laurent Perrussel. “Games of Influence”. In: *Journal of Logic and Computation* (2021). Published online.
- [81] Umberto Grandi, Emiliano Lorini, and Laurent Perrussel. “Propositional Opinion Diffusion”. In: *Proceedings of the 14th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2015.
- [82] Umberto Grandi, Hang Luo, Nicolas Maudet, and Francesca Rossi. “Aggregating CP-nets with unfeasible outcomes”. In: *Proceedings of the 20th International Conference on Principles and Practice of Constraint Programming (CP)*. 2014.

- [83] Umberto Grandi, Stefan Napel, Rolf Niedermeier, and Kristen Brent Venable. "Application-Oriented Computational Social Choice (Dagstuhl Seminar 19381)". In: *Dagstuhl Reports* 9.9 (2020), pp. 45–65.
- [84] Umberto Grandi, James Stewart, and Paolo Turrini. "The Complexity of Bribery in Network-Based Rating Systems". In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 2018.
- [85] Umberto Grandi, James Stewart, and Paolo Turrini. "Personalised Rating". In: *Autonomous Agents and Multiagent Systems* 34.55 (2020).
- [86] Umberto Grandi and Paolo Turrini. "A Network-Based Rating System and its Resistance to Bribery". In: *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*. 2016.
- [87] M. Granovetter. "Threshold models of collective behavior". In: *American Journal of Sociology* 83.6 (1978), pp. 1420–1443.
- [88] Morris H. de Groot. "Reaching a Consensus". In: *Journal of the American Statistical Association* 69.345 (1974), pp. 118–121.
- [89] Dylan Hadfield-Menell, Stuart J Russell, Pieter Abbeel, and Anca Dragan. "Cooperative Inverse Reinforcement Learning". In: *Proceedings of the 30th Conference on Advances in Neural Information Processing Systems (NIPS)*. 2016.
- [90] Adrian Haret, Arianna Novaro, and Umberto Grandi. "Preference aggregation with incomplete CP-nets". In: *Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning (KR)*. 2018.
- [91] Dirk Helbing, Bruno S. Frey, Gerd Gigerenzer, Ernst Hafen, Michael Hagner, Yvonne Hofstetter, Jeroen van den Hoven, Roberto V. Zicari, and Andrej Zwitter. "Will Democracy Survive Big Data and Artificial Intelligence". In: *Scientific American* (February 2017).
- [92] C. Hoede and R. R. Bakker. "A theory of decisional power". In: *The Journal of Mathematical Sociology* 8.2 (1982), pp. 309–322.
- [93] Richard A. Holley and Thomas M. Liggett. "Ergodic Theorems for Weakly Interacting Infinite Systems and the Voter Model". In: *The Annals of Probability* 3.4 (1975), pp. 643–663.
- [94] J. R. Isbell. "A class of simple games". In: *Duke Mathematical Journal* 25.3 (1958), pp. 423–439.
- [95] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

- [96] Matthew O. Jackson and Simon Wilkie. "Endogenous Games and Mechanisms: Side Payments among Players". In: *Review of Economic Studies* 72(2) (2005), pp. 543–566.
- [97] Anson Kahng, Simon Mackenzie, and Ariel D Procaccia. "Liquid democracy: An algorithmic perspective". In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 2018.
- [98] Stuart Kaufmann. "Homeostasis and Differentiation in Random Genetic Control Networks". In: *Nature* 224 (1969), pp. 177–178.
- [99] David Kempe, Jon Kleinberg, and Éva Tardos. "Maximizing the spread of influence through a social network". In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2003.
- [100] David Kempe, Jon Kleinberg, and Éva Tardos. "Influential Nodes in a Diffusion Model for Social Networks". In: *Proceedings of the 32nd International Colloquium on Automata, Languages and Programming (ICALP-2005)*. 2005.
- [101] Christian Klamler and Ulrich Pferschy. "The traveling group problem". In: *Social Choice and Welfare* 29 (2007), pp. 429–452.
- [102] Sébastien Konieczny and Ramón Pino Pérez. "Merging Information Under Constraints: A Logical Framework". In: *Journal of Logic and Computation* 12.5 (2002), pp. 773–808.
- [103] Grammateia Kotsialou and Luke Riley. "Incentivising Participation in Liquid Democracy with Breadth-First Delegation". In: *Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 2020.
- [104] Grammateia Kotsialou, Luke Riley, Amrita Dhillon, Toktam Mahmoodi, Peter McBurney, Paul Massey, and Richard Pearce. "Using Distributed Ledger Technology for Shareholder Rights Management". In: *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. 2018.
- [105] J. Lang and L. Xia. "Sequential composition of voting rules in multi-issue domains". In: *Mathematical social sciences* 57 (2009), pp. 304–324.
- [106] Jérôme Lang. "Knowledge Representation, Reasoning, and Computational Social Choice - - Invited Talk -". In: *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR)*. 2016.
- [107] Annick Laruelle. "Voting to select projects in participatory budgeting". In: *European Journal of Operational Research* 288.2 (2021), pp. 598–604.

- [108] Loujayn Layka. "Multiagent Reinforcement Learning for Iterative Voting". MA thesis. France: Université Jean Monnet Saint-Etienne, 2019.
- [109] Keith Lehrer and Carl Wagner. *Rational Consensus in Science and Society*. Springer, 1981.
- [110] Kristina Lerman, Xiaoran Yan, and Xin-Zeng Wu. "The "majority illusion" in social networks". In: *PloS one* 11.2 (2016).
- [111] Omer Lev and Jeffrey S. Rosenschein. "Convergence of Iterative Scoring Rules". In: *Journal of Artificial Intelligence Research* 57 (2016), 573–591.
- [112] Christian List. "Group Communication and the Transformation of Judgments: An Impossibility Result". In: *Journal of Political Philosophy* 19.1 (2011), pp. 1–27.
- [113] Bing Liu. *Sentiment Analysis and Opinion Mining*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers, 2012.
- [114] A. Loreggia, N. Mattei, F. Rossi, and K. B. Venable. "Value Alignment Via Tractable Preference Distance". In: *Artificial Intelligence Safety and Security*. Ed. by R. V. Yampolskiy. CRC Press, 2018.
- [115] Andrea Loreggia, Francesca Rossi, and Kristen Brent Venable. "Modelling Ethical Theories Compactly". In: *Proceedings of the AAAI Workshop on AI, Ethics and Society*. 2017.
- [116] Thomas Lukasiewicz and Enrico Malizia. "Complexity results for preference aggregation over (m)CP-nets: Pareto and majority voting". In: *Artificial Intelligence* 272 (2019), pp. 101–142.
- [117] Pia Mancini. "Why it is time to redesign our political system". In: *European View* 14.69 (2015), pp. 69–75.
- [118] Reshef Meir. "Iterative Voting". In: *Trends in Computational Social Choice*. Ed. by Ulle Endriss. AI Access, 2017. Chap. 4, pp. 69–86.
- [119] Reshef Meir, Kobi Gal, and Maor Tal. "Strategic voting in the lab: compromise and leader bias behavior". In: *Autonomous Agents and Multi Agent Systems* 34.1 (2020), p. 31.
- [120] Reshef Meir, Omer Lev, and Jeffrey S. Rosenschein. "A local-dominance theory of voting equilibria". In: *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*. 2014.
- [121] Reshef Meir, Maria Polukarov, Jeffrey S. Rosenschein, and Nicholas R. Jennings. "Convergence to Equilibria in Plurality Voting." In: *Proceedings of the Twenty-fourth conference on Artificial Intelligence (AAAI)*. 2010.
- [122] H. Moulin. *The strategy of social choice*. North-Holland, 1983.

- [123] R. Myerson and R. Weber. “A Theory of Voting Equilibria”. In: *American Political Science Review* 87.1 (1993), pp. 102–114.
- [124] Klaus Nehring and Marcus Pivato. “The median rule in judgement aggregation”. In: *Economic Theory* (2021).
- [125] Andrew Y. Ng and Stuart Russell. “Algorithms for Inverse Reinforcement Learning”. In: *Proceedings of the 17th International Conference on Machine Learning (ICML)*. 2000.
- [126] Arianna Novaro. “Goal-based voting”. PhD thesis. France: Université Toulouse III Paul Sabatier, 2019.
- [127] Arianna Novaro, Umberto Grandi, Dominique Longin, and Emiliano Lorini. “Goal-Based Collective Decisions: Axiomatics and Computational Complexity.” In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. 2018.
- [128] Arianna Novaro, Umberto Grandi, Dominique Longin, and Emiliano Lorini. “Strategic Majoritarian Voting with Propositional Goals (Extended Abstract).” In: *Proceedings of the 18th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2019.
- [129] S. Obraztsova, E. Markakis, and D. R. M. Thompson. “Plurality Voting with Truth-biased Agents”. In: *Proceedings of the 6th International Symposium on Algorithmic Game Theory*. 2013.
- [130] Svetlana Obraztsova, Evangelos Markakis, Maria Polukarov, Zinovi Rabinovich, and Nicholas R. Jennings. “On the Convergence of Iterative Voting: How Restrictive Should Restricted Dynamics Be?” In: *Proceedings of the 19th AAAI Conference on Artificial Intelligence (AAAI)*. 2015.
- [131] David C. Parkes. “On AI, Markets and Machine Learning”. In: *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. 2017.
- [132] Claire Pillet. “The influence of turn-taking on the diffusion of opinions”. MA thesis. France: Université Claude Bernard Lyon 1, 2020.
- [133] Usha Nandini Raghavan, Réka Albert, and Soundar Kumara. “Near linear time algorithm to detect community structures in large-scale networks”. In: *Physical Reviews E* 76 (3 2007), p. 036106.
- [134] Michel Regenwetter, Bernard Grofman, A. A. J. Marley, and Ilia Tsetlin. *Behavioral Social Choice: Probabilistic Models, Statistical Inference, and Applications*. Cambridge University Press, 2006.

- [135] Annemieke Reijngoud and Ulle Endriss. “Voter Response to Iterated Poll Information”. In: *Proceedings of the 11th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2012.
- [136] Luke Riley, Grammateia Kotsialou, Amrita Dhillon, Toktam Mahmoodi, Peter McBurney, and Richard Pearce. “Deploying a Shareholder Rights Management System onto a Distributed Ledger”. In: *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. 2019.
- [137] Francesca Rossi, Kristen Brent Venable, and Toby Walsh. “mCP Nets: Representing and Reasoning with Preferences of Multiple Agents”. In: *Proceedings of the 19th National Conference on Artificial Intelligence, 16th Conference on Innovative Applications of Artificial Intelligence (AAAI)*. 2004.
- [138] Jörg Rothe, ed. *Economics and Computation, An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*. Springer, 2016.
- [139] M. A. Satterthwaite. “Strategy-proofness and Arrow’s Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions”. In: *Journal of Economic Theory* 10.2 (1975), pp. 187–217.
- [140] Thomas Schelling. *Micromotives and macrobehavior*. Norton, 1978.
- [141] Nicolas Schwind, Katsumi Inoue, Gauvain Bourgne, Sébastien Konieczny, and Pierre Marquis. “Is promoting beliefs useful to make them accepted in networks of agents?” In: *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*. 2016.
- [142] Nicolas Schwind, Katsumi Inoue, Gauvain Bourgne, Sébastien Konieczny, and Pierre Marquis. “Belief Revision Games”. In: *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-2015)*. 2015.
- [143] Ehud Y. Shapiro. “Foundations of Internet-enabled Democracy”. In: CoRR abs/1710.02873 (2017). arXiv: [1710.02873](https://arxiv.org/abs/1710.02873). URL: <http://arxiv.org/abs/1710.02873>.
- [144] Patrick Shepherd and Judy Goldsmith. “A Reinforcement Learning Approach to Strategic Belief Revelation with Social Influence”. In: *The 34th AAAI Conference on Artificial Intelligence (AAAI)*. 2020.
- [145] Ilya Shmulevich, Edward R. Dougherty, Seungchan Kim, and Wei Zhang. “Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks”. In: *Bioinformatics* 18 (2002), pp. 261–274.

- [146] Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2009.
- [147] Carles Sierra. “Agent-Mediated Electronic Commerce”. In: *Autonomous Agents and Multi-Agent Systems* 9.3 (2004), pp. 285–301.
- [148] A. P. Sistla and E. M. Clarke. “The complexity of propositional linear temporal logics”. In: *Journal of the Association of Computing Machinery* 32.3 (1985), pp. 733–749.
- [149] Arkadii Slinko. “Beyond Gibbard and Satterthwaite: Voting Manipulation Games”. In: *The Future of Economic Design*. Ed. by Laslier JF., Moulin H., Sanver M., and Zwicker W. Studies in Economic Design. Springer, 2019.
- [150] Wynn C. Stirling. *Theory of Social Choice on Networks*. Cambridge University Press, 2016.
- [151] Nimrod Talmon and Piotr Faliszewski. “A Framework for Approval-Based Budgeting Methods”. In: *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. 2019.
- [152] Alan Tsang and Kate Larson. “The Echo Chamber: Strategic Voting and Homophily in Social Networks”. In: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2016.
- [153] Bryan Wilder and Yevgeniy Vorobeychik. “Controlling Elections through Social Influence”. In: *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. 2018.
- [154] M. Wooldridge. *An introduction to multiagent systems*. Second edition. John Wiley & Sons, 2009.