Binary Aggregation with Integrity Constraints

Umberto Grandi

Binary Aggregation with Integrity Constraints

ILLC Dissertation Series DS-2012-08

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 904
1098 XH Amsterdam
phone: +31-20-525 6051
fax: +31-20-525 5206
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl

These investigations were supported by the Netherlands Organisation for Scientific Research (NWO) Vidi project 639.022.706 on "Collective Decision Making in Combinatorial Domains".

Cover design by Inés Crespo. Cover art: detail from *Città ideale*, unknown painter, Galleria Nazionale delle Marche, Urbino. Source: Wikimedia Commons. Printed and bound by GVO | Ponsen & Looijen, Ede.

ISBN: 978-90-6464-575-4

Binary Aggregation with Integrity Constraints

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. D.C. van den Boom
ten overstaan van een door het college voor
promoties ingestelde commissie, in het openbaar
te verdedigen in de Agnietenkapel
op dinsdag 25 september 2012, te 10:00 uur

door

Umberto Grandi

geboren te Milaan, Italië.

Promotor: Prof. dr. Y. Venema
Co-promotor: Dr. U. Endriss

Overige leden: Prof. dr. K. R. Apt
Prof. dr. J. F. A. K. van Benthem
Prof. dr. C. Boutilier
Prof. dr. J. Lang
Prof. dr. C. List
Prof. dr. F. Rossi

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Contents

Acknowledgments

This dissertation is the product of four years during which many people have helped, encouraged and influenced me in many different ways. I want to express my gratitude to those who have shared part of this journey with me, and acknowledge the help and support of several among them.

First and foremost, I would like to thank Ulle Endriss for being one of the best supervisors a student can ask for, leaving me enough freedom to develop my own ideas, while constraining them to become a solid production of intellectual material which I am now able to present in this dissertation. I am grateful for his successful efforts in creating a lively research group during the past four years, and I thank him for his patience, for listening to my endless motivational doubts and for allowing me, perhaps too frequently, to travel and visit Italy.

Most of the results presented in this thesis would have not seen the light without several fruitful discussions with Daniele Porello. This reasearch has also been greatly influenced by the interaction with Stéphane Airiau, Davide Grossi, Jérôme Lang, Gabriella Pigozzi, Paolo Turrini and Joel Uckelman.

This dissertation was developed in an exciting research environment provided by the Institute for Logic, Language and Computation in Amsterdam, and I would like to thank all the people who have contributed to create it. I would like to thank Karin Gigengack, Tanja Kassenaar and Peter van Ormondt from the ILLC office, as well as Jenny Batson and Marco Vervoort. Particular thanks to Yde Venema for accepting to be my promotor. Special thanks also go to those who supported us so enthusiastically during the creation of the PhD council. I am deeply indebted to Hykel Hosni for having encouraged me to cross the Alps and apply for a PhD in Amsterdam.

Warm thanks goes to Inés Crespo for, among other things, having designed the cover of this dissertation and having carefully proofread some parts. I owe much to her wise suggestions and to her determination. I am grateful to Alessandro Porta for our endless Skype calls and for always being a good friend. I also want to warmly thank: Federico Sangati, Nina Gierasimczuk, Jakub Szymanik, Lena Kurzen, Cédric Dégremont, all the members of the PhD council, the organisers of the Logic Tea, the numerous office mates in Science Park C3.119, Dapper Brouwerij, Michael, Gabor for having satisfied my dream of knowing a true contemporary artist, Britta for being such a positive person, Yvon for the help with the samenvatting, Mape and Fabri for having rescued me during a hard winter, all the people who participated to the Amsterdam Summer School in the summer of 2011, and all those people I forgot to mention (there is at least one).

My family has kept on encouraging me during these years, and without their support I would have neither started nor completed this adventure. They will never cease to be my biggest source of inspiration.

Amsterdam
July, 2012.

Umberto Grandi

Chapter 1

Introduction

Every organised society of individuals calls for procedures to stage collective decisions. Since the first democracies to the advent of social networks and interactive software agents, the interest in the development of efficient procedures for collective decision making has only been increasing.

This dissertation provides a systematic study of a particular class of collective decision making problems, in which several individuals each need to make a yes/no choice regarding a number of issues and these choices then need to be aggregated into a collective choice. Inspired by potential applications in Artificial Intelligence, we put forward a systematic and flexible framework that aims to account for the wide variety of situations that can be encountered when dealing with the problem of collective choice.

1.1 Background

The literature on Economic Theory, in particular its branches of Welfare Economics, Public Choice and Social Choice Theory, comprises a centuries-old tradition of studies of the problem of collective decision making. Dating back as far as the 18th century, the work of Condorcet initiated a line of research in which renowned scholars such as Charles Dodgson (also known as Lewis Carroll) have contributed to the design and the analysis of voting procedures to be used in public elections or committee decisions (McLean and Urken, 1995). In more recent times, starting from the seminal work of Arrow (1963), Social Choice Theory has become a well-established formalism for the study of collective decision making.

Today's world is not quite similar to the one in which Condorcet and his colleagues carried out their research. The rise of new information technologies has endowed society with the possibility of taking collective decisions between large groups of people in a network, and novel theoretical problems have originated from the design of systems of autonomous software agents. Researchers in Artificial Intelligence, particularly from the field of Multiagent Systems (Sandholm, 1999;

Shoham and Leyton-Brown, 2009; Wooldridge, 2009), soon became interested in the work of social choice theorists, and started to borrow techniques from the literature on Economic Theory to analyse and study problems of collective choice in a new light.

Notable examples include the analysis of ranking systems carried out by Altman and Tennenholtz (2008, 2010), in which problems related to the design of a search engine are given a formal axiomatic treatment using tools from Social Choice Theory. Similar techniques have also been used to compare and evaluate the design of online recommender systems (Pennock et al., 2000), and to formalise the problem of aggregating the result of different search engines (Dwork et al., 2001). This line of research has proved useful not only for the study of the interaction of automatic software agents, but also for the implementation and the enhancement of existing procedures for collective decision making. As an example, Duke University in the U.S. has implemented a complex ranking procedure known as the Kemeny rule (Kemeny, 1959) to rank Ph.D. applicants, exploiting efficient heuristics developed by computer scientists (Conitzer, 2010).

Growing collaboration between Artificial Intelligence and Social Choice Theory has led to the creation of an entirely new research agenda under the name of Computational Social Choice (Chevaleyre et al., 2007; Procaccia, 2011; Brandt et al., Forthcoming). One particular problem of interest for this new community is the case of *social choice in combinatorial domains*, in which the space of alternatives from which individuals have to choose has a multi-attribute structure (Chevaleyre et al., 2008). Classical examples include voting in multiple referenda, in which individuals are asked to decide which propositions in a given set they accept; or electing a committee, in which a number of seats need to be filled with a set of possible candidates. The problem of decision making in combinatorial domains was first pointed out by political scientists (Brams et al., 1998; Lacy and Niou, 2000) and is now also receiving attention from researchers in Economic Theory (Ahn and Oliveros, 2012). In Artificial Intelligence such questions have been the subject of numerous publications. Starting from the work of Lang (2004), to a series of more recent developments (Lang, 2007; Xia et al., 2010; Xia, 2011), there have been several attempts to tackle the high complexity that arises in this context by using tools from Artificial Intelligence, such as methods for modelling preferences inspired by knowledge representation (Rossi et al., 2004; Lang and Xia, 2009; Li et al., 2011; Airiau et al., 2011).

1.2 Research Question

A central problem in Social Choice Theory, and, in view of our previous discussion, in all of its applications to Artificial Intelligence, is the problem of *aggregation*: Suppose a group of agents each supply a particular piece of information regarding a common problem and we want to aggregate this information into a collective view to obtain a summary of the individual views provided. A classical example is that of preferences (Arrow, 1963): each agent declares their individual preferences over a set of alternatives by providing an ordering over this set, and we are asked to amalgamate this information into a collective ranking that represents the individual preferences provided. The same methodology has also been applied more recently to a number of other types of information, among others beliefs (Maynard-Zhang and Lehmann, 2003; Konieczny and Pino Pérez, 2002, 2011) and judgments (List and Pettit, 2002).

One of the main features of the study of aggregation is the problem of *collective rationality*: given a rationality assumption that binds the choices of individuals, we ask whether the output of an aggregator still satisfies the same rationality assumption. Consider the following example:

Example 1.2.1. Three autonomous agents need to decide on whether to perform a collective action. This action is performed if two parameters are estimated to exceed a certain threshold. We can model the choice situation with a multi-attribute domain in which there are three issues at stake: "the first parameter is above the threshold" (T_1), "the second parameter is above the threshold" (T_2), and "the action should be performed" (A). The rationality assumption that links the three issues together can be modelled using a simple propositional formula, namely $T_1 \wedge T_2 \rightarrow A$. Consider now the following situation, in which the individual views on the three issues are aggregated using the *majority rule*, accepting an issue if a majority of the individual agents do:

	T_1	T_2	A
Agent 1	Yes	Yes	Yes
Agent 2	No	Yes	No
Agent 3	Yes	No	No
Majority	Yes	Yes	No

In this situation the collective action A is not performed, even though a majority of the individuals think that the first parameter exceeds the threshold and a (different) majority agree that also the second parameter exceeds the threshold. Situations like the one above are often considered paradoxical: even if each individual agent is rational (i.e., each of them satisfies the rationality assumption), the collective view derived using the majority rule is not. That is, the majority rule fails to *lift* the integrity constraint $T_1 \wedge T_2 \rightarrow A$ from the individual to the collective level. This example shows that the majority rule violates collective rationality in certain specific cases.

In this dissertation we put forward a general framework that encompasses most of the classical studies of collective rationality in Social Choice Theory, and that can prove useful to diverse research areas in Artificial Intelligence. We base our framework on binary aggregation, in which individuals are required to choose from a multi-issue domain in which issues represent different binary choices. We model rationality assumptions using a simple propositional language, and we give a precise definition of collective rationality with respect to a given rationality assumption. Classical work in Social Choice Theory has studied aggregation procedures with the axiomatic method, using axioms to express desirable properties of a procedure. For example, the principle that all individuals should be given equal weight is formalised in the axiom of *anonymity*, and the axiom of *neutrality* expresses a similar requirement of impartiality between issues. We classify rationality assumptions with respect to their syntactic properties, and we give a systematic treatment of the question of how we can relate collective rationality with respect to a syntactically defined sublanguage on the one hand, to classical axiomatic properties from Social Choice Theory on the other. For instance, Example 1.2.1 shows that the majority rule is not collectively rational with respect to the integrity constraint $T_1 \wedge T_2 \rightarrow A$, which formalises the rationality assumption in the example. A similar phenomenon can be observed when considering the 3-clause $T_1 \vee T_2 \vee A$ as rationality assumption: to see this, consider a scenario in which each of three agents accepts exactly one issue, and no two agents accept the same issue. On the other hand, any 2-clause (i.e., disjunctions of size 2) will always be lifted, i.e., the majority rule is collectively rational with respect to the language of 2-clauses. We will then be able to describe the majority rule in terms of classical axioms from Social Choice Theory or in terms of the languages for integrity constraints it lifts. It is results of this kind that we shall explore in depth in this dissertation.

Research in Computational Social Choice have mainly focused on the study of voting procedures (Brandt et al., Forthcoming), i.e., mechanisms for the selection of candidates depending on the preferences of a set of individuals (Brams and Fishburn, 2002). The study of voting procedures is strongly related to the problem of aggregation, since the selection of candidates can take place by aggregating individual preferences into a collective one. However, we shall not treat the problem of voting in this dissertation, referring to our conclusions for a discussion of the impact of our results on voting theory.

Nevertheless, two frameworks for the study of aggregation have been considered in Computational Social Choice, namely preference and judgment aggregation. Of the two, the former has received the most attention, being the subject of a growing number of papers (Conitzer, 2006; Pini, 2007; Gonzales et al., 2008; Endriss et al., 2009; Betzler et al., 2009; Pini et al., 2009; Hudry, 2010; Pini et al., 2011; Rossi et al., 2011). On the other hand, the framework of judgment aggregation, which was recently established as a central topic in Social Choice Theory (List and Pettit, 2002; List and Puppe, 2009), has to date given rise to only a small amount of publications in Artificial Intelligence, most of which focus on investigating the the computational complexity of the framework (Endriss et al., 2010a,b; Baumeister et al., 2011; Lang et al., 2011).

Both the framework of preference and the framework of judgment aggregation can be embedded into binary aggregation by devising suitable integrity constraints. An ordering over three alternatives, for instance, can be represented in binary aggregation with a binary ballot over three issues, one for each pair of alternatives. If $P_{a>b}$, $P_{b>c}$ and $P_{a>c}$ are three binary issues, with their natural interpretation, then we can represent the ordering $a > c > b$ with ballot $(1, 0, 1)$, signifying that both issues $P_{a>b}$ and $P_{a>c}$ are accepted and issue $P_{b>c}$ is rejected. The rationality assumption of transitivity can be represented with formulas like the following: $P_{a>b} \wedge P_{b>c} \rightarrow P_{a>c}$. The embedding is less straightforward for the case of judgment aggregation, and can be achieved by explicitly representing the logical correlations between the propositional formulas that constitute the object of judgment.

The problem of collective rationality is central to both preference and judgment aggregation, and theoretical results in these frameworks can thus be compared with our findings in binary aggregation. Inspired by situations like the one presented in the introductory Example 1.2.1, we provide a general definition of *paradox* in binary aggregation to account for situations in which the collective outcome does not fulfill the integrity constraint which is satisfied by all individuals. Making use of the embeddings of aggregation frameworks into binary aggregation, we are able to show that most paradoxes in aggregation theory, such as the Condorcet paradox (1785) and the discursive dilemma (List and Pettit, 2002), can be seen as instances of our general definition. Moreover, we analyse in depth the relation between our characterisation results and known impossibility theorems in both preference and judgment aggregation, putting forward a new proof method which attempts to identify the source of impossibilities in a clash between axiomatic properties and particular requirements of collective rationality.

All the results achieved in this dissertation are of a theoretical kind, and their presentation aims at proposing a theory of collective rationality in binary aggregation rather than developing solutions which are specific to a certain class of applications. This dissertation aims at providing sound foundations to more domain specific research, building a framework that takes into account the variety of new problems that may be encountered by researchers in Artificial Intelligence.

1.3 Chapter Overview

The structure of this dissertation is summarised in Figure 1.1. In Chapter 2 we give the basic definitions of the framework of binary aggregation with integrity constraints, which is the principal object of study of this dissertation. The two crucial definitions of paradox and of collective rationality are presented in the same chapter, as well as several axiomatic properties for the study of aggregation procedures. Chapter 4 constitutes the mathematical core of the dissertation, providing a number of characterisation results in binary aggregation that link classical axiomatic properties from Social Choice Theory with collective rationality. The generality of the framework of binary aggregation with integrity constraints is investigated along two lines of argument. First, by concentrating on the study of paradoxical situations, in Chapter 3 we show that our definition of paradox accounts for many of the classical occurrences of paradoxes in aggregation theory. Second, we show how characterisation results in binary aggregation can serve as a starting point for the investigation of new results in other frameworks of aggregation. Chapter 5 focuses on preference aggregation and Chapter 6 on judgment aggregation. In Chapter 7 we bring together the two lines of work by defining and analysing practical aggregation procedures for collectively rational aggregation. Chapter 8 concludes and contains a list of directions for future research.

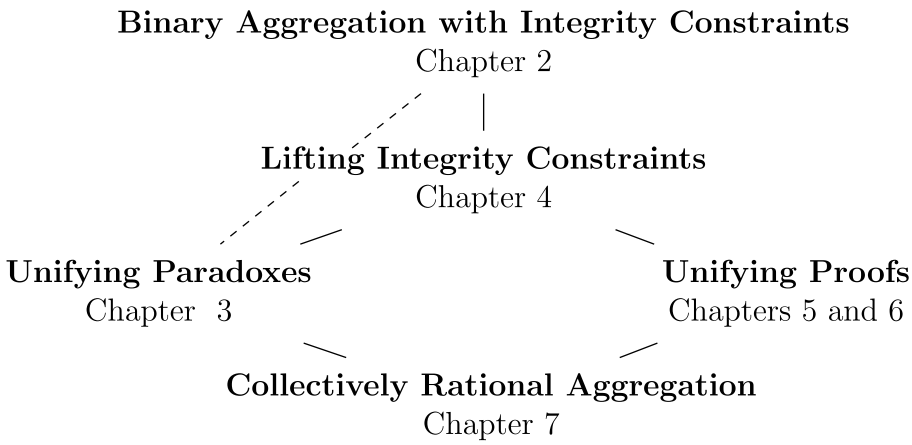

Figure 1.1: Structure of the dissertation.

The remaining part of this introduction provides a brief overview of the results presented in the dissertation following the structure in Figure 1.1. As shown in Figure 1.1, Chapters 2 and 4 constitute the core of the dissertation. However, Chapter 4 may be skipped by following the dashed line in Figure 1.1, forming a coherent presentation of aggregation paradoxes and possible escape routes towards collectively rational aggregation.

1.3.1 Binary Aggregation with Integrity Constraints

Chapter 2 is devoted to introducing the framework of *binary aggregation with integrity constraints*, which we put forward as a general framework for the study of aggregation problems. The chapter provides basic definitions for this setting, including the two crucial notions of *paradox* and of *collective rationality*, as well as a list of axiomatic properties that shall be used to study aggregation procedures.

Binary Aggregation

The ingredients of a decision problem are a set of individuals (possibly one) and a set of alternatives from which to make a choice. In this work, we concentrate on decision problems in which there are at least two individuals (a collective choice problem), and where the set of alternatives has a binary combinatorial structure, i.e., it is a product space of several binary domains associated with a set of issues, or attributes. We assume that each individual submits a yes/no choice for each of the issues and these choices are then aggregated into a collective one.

In Section 2.1.3 we provide several motivating examples showing the generality of this setting. The most natural example is that of collective decisions over multiple issues, e.g., multiple referenda and situations such as the one presented in Example 1.2.1. More complex objects such as preferences and judgments can also be modelled as elements of specific binary combinatorial domains.

At a very abstract level, virtually every individual expression has the potential to be described using a finite number of binary parameters. This is a common assumption when, for instance, the focus is on describing the diversity of elements in a set of alternatives (Nehring and Puppe, 2002), or distinguishing between possible worlds in an epistemic framework (Hintikka, 1962). **Binary aggregation can therefore be summarised as the study of the aggregation of individual expressions described by means of binary variables.**

Rationality Assumptions/Integrity Constraints

Individuals can be rational in many different ways. When they express preferences over a set of alternatives, like in the case of preference aggregation (Gaertner, 2006), a common assumption is to assume the transitivity of such preferences. Thus, if an alternative a is preferred to a second one b, and this is in turn preferred to a third alternative c, then the individual is also assumed to prefer a to c. Different assumptions are made in the field of judgment aggregation (List and Puppe, 2009), in which individuals express judgments over a set of correlated propositions. In that case, the rationality of a judging agent relates to the logical consistency of the set of propositions she accepted.

As shown by our initial Example 1.2.1, rationality assumptions in binary aggregation can be expressed by means of formulas in a simple propositional language. Rationality assumptions characteristic for other settings can also be formalised in this language, exploiting the embedding of the different frameworks into binary aggregation. We call a propositional formula enforcing a rationality assumption in binary aggregation an *integrity constraint*. An individual expression, i.e., a binary ballot, is called rational if it satisfies the formula in question.

This fact represents our first crucial observation: **rationality assumptions can be represented as propositional formulas, and can thus be classified and analysed in terms of their syntactic properties.** This is where mathematical logic will play a small but very important role.

Collective Rationality

Given a set of individuals each expressing a rational ballot, the natural question that arises is whether the collective outcome will be rational as well. As testified by the introductory Example 1.2.1, this is not always the case, even when one of the most natural aggregation procedures like the majority rule is being used.

We call a situation in which all individual ballots satisfy a given rationality assumption, but the aggregation results in an irrational outcome a *paradox* (see Definition 2.1.9). Chapter 3 is devoted to showing how most paradoxes of aggregation that are traditionally studied in the literature on Social Choice Theory can be seen as instances of our general definition of paradox in binary aggregation.

We call an aggregation procedure *collectively rational* for a given rationality assumption if, whenever all the ballots submitted by the individuals are rational, so is the outcome of aggregation (see Definition 2.1.8). The majority rule, for instance, is not collectively rational with respect to the integrity constraint $T_1 \wedge T_2 \rightarrow A$, as shown by our Example 1.2.1. Thus, **an aggregation procedure is collectively rational with respect to an integrity constraint if it lifts the rationality assumption given by the integrity constraint from the individual to the collective level.** In Chapter 4 we analyse how the notion of collective rationality varies depending on the syntactic structure of the integrity constraint at hand, and we look for axiomatic conditions that guarantee collective rationality of a given procedure.

1.3.2 Unifying Paradoxes

The observation of paradoxical situations has traditionally been the starting point of most theoretical work in Social Choice Theory. One of the most striking example was observed by Condorcet (1785) when analysing the use of majority aggregation for preferences. Consider for instance the following toy example, in which three colleagues are helping in choosing a colour for the cover of this dissertation:

Joel	Orange	$\succ$	Red	$\succ$	Green
Daniele	Red	$\succ$	Green	$\succ$	Orange
Stéphane	Green	$\succ$	Orange	$\succ$	Red
Majority	Orange $\succ$ Red $\succ$ Green $\succ$ Orange				

Table 1.1: A cyclical majority outcome.

In spite of the fact that all colleagues have rational (in this case, transitive) preferences, in the situation described by Table 1.1 the conclusion of the majority rule is that orange is both the best *and* the worst colour for the cover of this dissertation! Thus, we obtain an irrational majoritarian outcome starting from a profile of rational ballots.

In Chapter 3 we analyse the most important paradoxes arising from the use of the majority rule in different settings. Our analysis focuses on the Condorcet paradox (1785), the discursive dilemma in judgment aggregation (List and Pettit, 2002), the Ostrogorski paradox (1902) and the more recent work of Brams et al. (1998) on multiple election paradoxes. The purpose of Chapter 3 is to show that most paradoxes in aggregation theory can be seen as instances of our definition of paradox in binary aggregation (see Definition 2.1.9). Hence, we provide a unified treatment of aggregation paradoxes that enables us to analyse the syntactical properties of paradoxical rationality assumptions.

We can thus make in Section 3.4 our second important observation: **when the majority rule is concerned, all paradoxical integrity constraints feature a clause (i.e., a disjunction) of size at least 3**. For instance, our introductory Example 1.2.1 describes a paradoxical situation with respect to the integrity constraint $T_1 \wedge T_2 \rightarrow A$, which is equivalent to the 3-clause $\neg T_1 \vee \neg T_2 \vee A$.

This observation can be formalised into a general result: in Theorem 4.4.8 we show that the majority rule is collectively rational (i.e., it does not generate a paradox) if and only if the integrity constraint under consideration is equivalent to a conjunction of clauses of size at most 2.

1.3.3 Lifting Integrity Constraints

The observation of paradoxical situations is usually generalised into impossibility theorems, proving that aggregation is unfeasible under certain axiomatic conditions. Classical work in Social Choice Theory was restricted to particular studies of collective rationality in a given aggregation situation, and for given classes of aggregation procedures. The aim was to identify *the* appropriate set of axiomatic properties (e.g., to model real-word economies, specific moral ideals, etc.) and then to prove a characterisation (or impossibility) result for those axioms. Given the wide variety of potential applications in Artificial Intelligence, on the other hand, in this context we require a systematic study that, depending on the situation at hand, can give answers to the problem of collective rationality. With every new application the principles underlying a system may change, so we may be more interested in devising languages for expressing a range of different axiomatic properties rather than identifying the "right" set of axioms. Furthermore, we may be more interested in developing methods that will help us to understand the dynamics of a range of different social choice scenarios rather than in technical results for a specific such scenario.

We group integrity constraints into syntactically defined fragments of the propositional language, e.g., the set of conjunctions, or the set of disjunctions of limited size, and we study the class of procedures that are collectively rational with respect to all integrity constraints in a given language. We discover that **requiring collective rationality with respect to certain natural syntactically defined languages corresponds to known classical axiomatic properties from Social Choice Theory.**

Formally, we define classes of aggregation procedures in two ways. On the one hand, given a language $\mathcal{L}$, we define the class $\mathcal{CR}[\mathcal{L}]$ as the set of procedures that are collectively rational with respect to all integrity constraints in $\mathcal{L}$. On the other hand, given a set of axiomatic properties AX and a language $\mathcal{L}$, we define the class $\mathcal{F}_{\mathcal{L}}[\text{AX}]$ as the set of procedures satisfying axioms AX on domains defined by $\mathcal{L}$. What we seek are characterisation results of the following form, providing necessary and sufficient axiomatic conditions for an aggregation to be collectively rational with respect to a given language $\mathcal{L}$:

$$\mathcal{CR}[\mathcal{L}] \quad = \quad \mathcal{F}_{\mathcal{L}}[\text{AX}].$$

In Section 4.2 we prove a series of characterisation results for several fragments of the propositional language. A simple example can be obtained by considering the language of literals, i.e., propositional atoms together with their negation. We prove that a necessary and sufficient condition for an aggregation procedure to be collectively rational with respect to any literal is that the procedure be *unanimous*, i.e., it accepts/rejects an issue when all individuals agree to accept/reject it (Theorem 4.2.1).

Results of this form can also be interpreted as characterising classical axiomatic properties in terms of collective rationality. While providing a characterisation for many standard axioms from the literature, in Section 4.3 we also show that for some other natural properties such a characterisation is not possible.

A very interesting case is given by the class of aggregation procedures that are collectively rational with respect to *all* possible integrity constraints. We prove that each such procedure copies the ballot of a (possibly different) individual in every situation (Theorem 4.2.8), and we call these procedures *generalised dictatorships*. In Chapter 7 we argue that a meaningful choice of the individual ballot that best represents all the other ballots submitted by the individuals may generate new interesting aggregation procedures. We present one such rule, called the *average voter rule*, and we evaluate its axiomatic and computational properties.

1.3.4 Unifying Proofs

Classical frameworks in Social Choice Theory like preference aggregation (Gaertner, 2006) and judgment aggregation (List and Puppe, 2009) can be seen as instances of binary aggregation by devising suitable integrity constraints. Having shown in Chapter 3 that paradoxes in these settings can be seen as instances of a general definition in binary aggregation, in Chapter 5 and 6 we turn to the analysis of theoretical results.

We are able to obtain new (im)possibility theorems in both preference and judgment aggregation, by employing the characterisation results in binary aggregation presented in Chapter 4. More importantly, we devise a new uniform proof method for theoretical results in aggregation theory that sheds new light on the problems that lie behind impossibilities. The method consists of three basic steps:

(*i*) Given an aggregation problem, translate it into binary aggregation, obtaining, first, an integrity constraint that describes the domain of aggregation and, second, a set of axiomatic properties.

(*ii*) Use a characterisation result from binary aggregation to check whether collective rationality with respect to the given integrity constraint *clashes* with the axiomatic requirements.

(*iii*) Translate the result back into the original setting to obtain a possibility or an impossibility result.

Using this method, we look for clashes between the syntactic shape of the integrity constraints defining an aggregation problem on the one hand, and a given combination of axiomatic postulates on the other. The results that can be obtained by using this proof method may share similarities or may be weaker than known results from the literature on Social Choice Theory. However, the focus is not on the novelty or strength of single results, but rather on the generality and flexibility of the proof method we put forward. By unifying proofs in aggregation theory we gain a deeper understanding of the common problem behind many classical results: **impossibilities arise from clashes between axiomatic properties and requirements of collective rationality**.

We employ this methodology in Chapter 5 for the case of preference aggregation, proving both possibility and impossibility results for various combinations of axioms and different representations of preferences. We also present an alternative proof of Arrow's Theorem (Arrow, 1963), which focuses on the effect of collective rationality with respect to preferential integrity constraints on the set of winning coalitions for an aggregation procedure.

Chapter 6 is devoted to a study of the framework of judgment aggregation (List and Pettit, 2002). In particular we focus on the new problem of the *safety of the agenda* (Endriss et al., 2010a). An agenda, i.e., a set of formulas, is called safe with respect to a given class of judgment aggregation procedures if all aggregators in the class output consistent judgments on all profiles of consistent judgment sets. For several classes of procedures defined in axiomatic terms, we provide necessary and sufficient conditions for an agenda to be safe. The resemblance with the characterisation results presented in Chapter 4 is immediate, and in Section 6.3.4 we compare these findings. We conclude the chapter by analysing in Section 6.4 the computational complexity of recognising safe agendas, proving that it is Π_2^p-complete for all classes we considered (Theorem 6.4.7). Our findings thus suggest that this problem is highly intractable for all the classes of procedures under consideration.

We can therefore conclude, as pictured in Figure 1.2, that binary aggregation with integrity constraints constitutes a general framework for the analysis of collective rationality. It provides a unifying definition of paradox and general characterisation results that encompass the other frameworks of aggregation present in the literature.

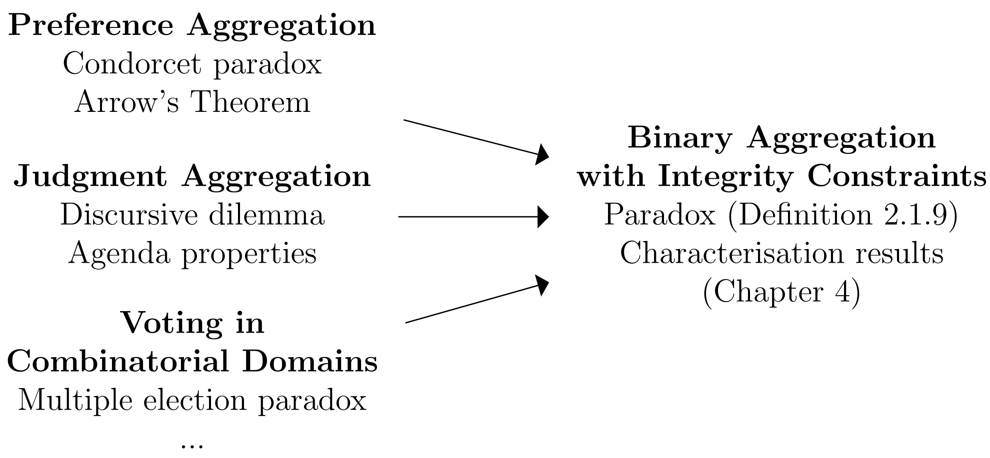

Figure 1.2: A general framework for aggregation theory.

1.3.5 Collectively Rational Aggregation

Having established the importance of the notion of collective rationality, the dissertation is completed with the analysis of some concrete aggregation procedures that are especially designed to be collectively rational. We propose in Chapter 7 the definition of three collectively rational rules, and we investigate the computational complexity of two classical problems: winner determination (WinDet) and strategic manipulation (Manip).

The former problem of winner determination for a given aggregation procedure demands to assess how difficult it is to compute the outcome in a given situation. The latter problem, Manip, focuses on the incentives that individuals may have in reporting their vote truthfully. A celebrated theorem by Gibbard (1973) and Satterthwaite (1975) shows that every reasonable voting procedure can be manipulated, i.e., individuals always have the opportunity to change the outcome of an election in their favour. The problem Manip asks how difficult it is to recognise whether an agent has incentives to deviate from her truthful ballot in a given situation.

The first rule we analyse is a generalised dictatorship which selects in each situation the ballots of those individuals that minimise the amount of disagreement with the other individual ballots. We call this rule the *average voter rule* (AVR) and we show that both the problems WinDet and Manip can be solved in polynomial time for this rule (Proposition 7.2.4 and Theorem 7.2.5).

The second rule we study is the *premise-based procedure* (PBP) for judgment aggregation, in which the judgment over a set of independent formulas called premises is aggregated by using the majority rule, and this collective judgment is then used to infer the acceptance or rejection of a a set of complex propositions defined over the premises (see, e.g., List and Puppe, 2009). We prove that WinDet for the PBP can be solved in polynomial time, while Manip is NP-complete (Proposition 7.3.2 and Theorem 7.3.3), thus showing the "jump" in computational complexity between winner determination and manipulability that is a good indicator of an aggregation rule which resists manipulation.

We end by analysing a well-known rule called the *distance-based rule* (DBP) (see, e.g., Konieczny and Pino Pérez, 2002; Pigozzi, 2006; Miller and Osherson, 2009). We limit our analysis to the problem of winner determination, showing that it is already highly unfeasible. We prove that WinDet for the DBP is complete for the class Θ_2^p, which contains those problems that can be solved in polynomial time using a logarithmic number of queries to an NP oracle (Theorem 7.4.5).

The results we obtain can be summarised in the following table:

	WinDet	Manip
AVR	P	P
PBP	P	NP-complete
DBP	Θ_2^p-complete	–

Table 1.2: Complexity of collectively rational aggregation.

1.3.6 Summary

Collective decision making in multi-issues domains is a problem of high interest to the Artificial Intelligence community, and has recently received considerable attention in the literature on Computational Social Choice. This dissertation provides a systematic study of aggregation in binary combinatorial domains, with particular attention to the problem of collective rationality.

This work is based on the following publications:

U. Grandi and U. Endriss. Lifting rationality assumptions in binary aggregation. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI-2010)*, 2010.

U. Endriss, U. Grandi, and D. Porello. Complexity of judgment aggregation: Safety of the agenda. In *Proceedings of the 9th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2010)*, 2010a.

U. Endriss, U. Grandi, and D. Porello. Complexity of winner determination and strategic manipulation in judgment aggregation. In *Proceedings of the 3rd International Workshop on Computational Social Choice (COMSOC-2010)*, 2010b.

U. Grandi and U. Endriss. Binary aggregation with integrity constraints. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, 2011.

U. Grandi and G. Pigozzi. On compatible multi-issue group decisions. In *Proceedings of the 10th Conference on Logic and the Foundations of Game and Decision Theory (LOFT-2012)*, 2012.

U. Endriss and U. Grandi. Graph aggregation. In *Proceedings of 4th International Workshop on Computational Social Choice (COMSOC-2012)*, 2012

U. Grandi. The common structure of paradoxes in aggregation theory. In *Proceedings of 4th International Workshop on Computational Social Choice (COMSOC-2012)*, 2012.

Chapter 2

Binary Aggregation with Integrity Constraints

In this chapter we provide the basic definitions of the framework of binary aggregation with integrity constraints, which constitutes the main object of study of this dissertation. In this setting, several individuals each need to make a yes/no choice regarding a number of issues and these choices then need to be aggregated into a collective choice. Depending on the application at hand, different combinations of yes/no may be considered *rational* and we describe such assumptions with an integrity constraint expressed in a simple logical language. The question then arises whether or not a given aggregation procedure will *lift* the rationality assumptions from the individual to the collective level, i.e., whether the collective choice will be rational whenever all individual choices are. We name this problem *collective rationality*, and we give it central status throughout this dissertation.

We provide formal definitions for the framework of binary aggregation with integrity constraints in Section 2.1, including the two crucial definitions of collective rationality and of paradox. In Section 2.2, we provide a list of desirable properties for aggregation procedures in the form of axioms. For some classes of procedures defined axiomatically we provide a mathematical representation in Section 2.3, and in Section 2.4 we compare our framework to the existing literature on binary aggregation.

2.1 Basic Definitions

Many aggregation problems can be modelled using a finite set of binary issues, whose combinations describe the set of alternatives on which a finite set of individuals need to make a choice. In this section, we give the basic definitions of the framework of binary aggregation with integrity constraints, and we define the two crucial concepts of paradox and of collective rationality. We present several practical examples of binary aggregation problems, taken from the literature on Social Choice Theory or inspired by practical cases of collective decision making.

2.1.1 Binary Aggregation

Let $\mathcal{I} = \{1, \ldots, m\}$ be a finite set of *issues*, and let $\mathcal{D} = D_1 \times \cdots \times D_m$ be a *boolean combinatorial domain*, i.e., $|D_i| = 2$ for all $i \in \mathcal{I}$. Without loss of generality we assume that $D_j = \{0, 1\}$ for all j. Thus, given a set of issues $\mathcal{I}$, the domain associated with it is $\mathcal{D} = \{0, 1\}^{\mathcal{I}}$. A *ballot* B is an element of $\mathcal{D}$.

Let $\mathcal{N} = \{1, \ldots, n\}$ be a finite set of *individuals*. Each individual submits a ballot $B_i \in \mathcal{D}$ to form a *profile* $\boldsymbol{B} = (B_1, \ldots, B_n)$. Thus, a profile consists of a binary matrix of size $n \times m$. We write b_j for the jth element of a ballot B, and $b_{i,j}$ for the jth element of ballot B_i within a profile $\boldsymbol{B} = (B_1, \ldots, B_n)$.

Definition 2.1.1. Given a finite set of issues $\mathcal{I}$ and a finite set of individuals $\mathcal{N}$, an *aggregation procedure* is a function $F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D}$, mapping each profile of binary ballots to an element of $\mathcal{D}$. Let $F(\boldsymbol{B})_j$ denote the result of the aggregation of profile $\boldsymbol{B}$ on issue j.

Aggregation procedures are defined for all possible profiles of binary ballots, a condition that takes the name of *universal domain* in the literature on Social Choice Theory. Aggregation procedures that are defined on a specific restricted domain, by making use of particular characteristics of the domain at hand, can always be extended to cover the full boolean combinatorial domain (for instance, by mapping all remaining profiles to a constant value).

2.1.2 Integrity Constraints

In many applications it is necessary to specify which elements of the domain are rational and which should not be taken into consideration. Since the domain of aggregation is a binary combinatorial domain, propositional logic provides a suitable formal language to express possible restrictions of rationality. In the sequel we shall assume acquaintance with the basic concepts of propositional logic. A list of the basic notions of propositional logic that we make use of in this dissertation can be found in Appendix A.

If $\mathcal{I}$ is a set of m issues, let $PS = \{p_1, \ldots, p_m\}$ be a set of propositional symbols, one for each issue, and let $\mathcal{L}_{PS}$ be the propositional language constructed by closing PS under propositional connectives. For any formula $\varphi \in \mathcal{L}_{PS}$, let $\text{Mod}(\varphi)$ be the set of assignments that satisfy φ.

Definition 2.1.2. An *integrity constraint* is any formula $\text{IC} \in \mathcal{L}_{PS}$.

Integrity constraints can be used to define what tuples in $\mathcal{D}$ we consider *rational* choices. Any ballot $B \in \mathcal{D}$ is an assignment to the variables $p_1, \ldots, p_m$, and we call B a *rational ballot* if it satisfies the integrity constraint IC, i.e., if B is an element of Mod(IC). A rational profile is an element of $\text{Mod(IC)}^{\mathcal{N}}$. In the sequel we shall use the terms "integrity constraints" and "rationality assumptions" interchangeably.

2.1.3 Examples

Let us now consider several examples of aggregation problems that can be modelled in binary aggregation by devising a suitable integrity constraint:

Example 2.1.3. (Multi-issue elections under constraints) A committee $\mathcal{N}$ has to decide on each of the three following issues: (U) financing a new university building, (S) financing a sports centre, (C) increasing catering facilities. As an approval of both a new university building and a sports centre would bring an unsustainable demand on current catering facilities, it is considered irrational to approve both the first two issues and to reject the third one. We can model this situation with a set of three issues $\mathcal{I} = \{U, S, C\}$. The integrity constraint representing this rationality assumption is the following formula: $p_U \wedge p_S \rightarrow p_C$. To see an example of a rational profile, consider the situation described in Table 2.1 for the case of a committee with three members. All individuals are rational, the only irrational ballot being $B = (1, 1, 0)$.

$\boldsymbol{B}$:		U	S	C
	i_1	0	1	0
	i_2	1	0	0
	i_3	1	1	1

Table 2.1: A rational profile for $p_U \wedge p_S \rightarrow p_C$.

The two examples that follow are classical settings from the literature on Social Choice Theory and will be studied in more detail in later chapters.

Example 2.1.4. (Preference aggregation) A set $\mathcal{N}$ of individuals has to agree on a ranking of three alternatives a, b and c. Each individual submits its own ranking of the alternatives from the most preferred to the least preferred, e.g., $b > a > c$. We can model this situation using a binary issue for every pair of alternatives: issue ab stands for "alternative a is preferred to alternative b". The set of issues is therefore $\mathcal{I} = \{ab, ba, bc, cb, ac, ca\}$. However, not every binary evaluation over this set of issues corresponds to a preference order. An integrity constraint needs to be devised to encode the properties of a strict preference relation: transitivity, completeness and anti-symmetry. This can be done by considering, for each combination of pairs of issues, the following integrity constraints: $p_{ab} \wedge p_{bc} \rightarrow p_{ac}$ standing for transitivity, and $p_{ab} \leftrightarrow \neg p_{ba}$, encoding the remaining two conditions of completeness and anti-symmetry. The correspondence between preference aggregation and binary aggregation is spelled out in detail in Section 3.1.2 and Chapter 5.

Example 2.1.5. (Judgment aggregation) A court composed of three judges has to decide on the liability of a defendant under the charge of breach of contract.

According to the law, the individual is liable if there was a valid contract and her behaviour was such as to be considered a breach of the contract. The court takes three majority decisions on the following issues: there was a valid contract (α), the individual broke the contract (β), the defendant is liable ($\alpha \wedge \beta$). We can model this situation using a set of six issues $\mathcal{I} = \{\alpha, \neg\alpha, \beta, \neg\beta, \alpha \wedge \beta, \neg(\alpha \wedge \beta)\}$ to model the decision of a judge on the three issues at stake, and a set of integrity constraints that reflect the consistency of a possible verdict. To do so, we need to rule out explicitly every inconsistent set that can be created using issues in $\mathcal{I}$:

Inconsistent sets of size 2: $\neg(p_x \wedge p_{\neg x})$ for all $x \in \{\alpha, \beta, \alpha \wedge \beta\}$,
$\neg(p_{\alpha\wedge\beta} \wedge p_{\neg\alpha})$ and $\neg(p_{\alpha\wedge\beta} \wedge p_{\neg\beta})$
Inconsistent set of size 3: $\neg(p_{\neg(\alpha\wedge\beta)} \wedge p_\alpha \wedge p_\beta)$

Situations like the one described in this example are the subject of a wide literature in Social Choice Theory under the name of *judgment aggregation* (List and Puppe, 2009). See Section 3.2.1 and Chapter 6 to see the correspondence between judgment aggregation and binary aggregation in more detail.

We conclude with two classical examples from voting theory:

Example 2.1.6. (Voting for candidates) A winning candidate has to be chosen from a set $C = \{1, \dots, m\}$ by an electorate $\mathcal{N}$. Let the set of issues be $\mathcal{I}{=}C$. Assume that we are using *approval voting* as voting procedure, in which individuals are submitting a set of candidates they approve (Brams and Fishburn, 2007). Then, we can model the situation without any integrity constraint, since every binary ballot over $\mathcal{I}$ corresponds to a set of candidates. Instead, if we consider more restrictive ballots like in the case of the *plurality rule*, in which each individual submits only its favourite candidate, we need to devise an integrity constraint that forces each individual to approve a single candidate in the list. This can only be done by taking the disjunction of all possible ballots:

$$(p_1 \wedge \neg p_2 \wedge \cdots \wedge \neg p_m) \vee (\neg p_1 \wedge p_2 \wedge \cdots \wedge \neg p_m) \cdots \vee (\neg p_1 \wedge \dots \neg p_{m-1} \wedge p_m)$$

The voting rule known as k-approval voting, in which individuals submit a set of k approved candidates, can be modelled in a similar fashion.

Example 2.1.7. (Voting for a committee) An electorate $\mathcal{N}$ needs to decide on a steering committee composed of a director, a secretary and a treasurer. Candidates can be chosen between c_1 and c_2, proposed by party F, and c_3 and c_4, proposed by party P. For political reasons, if the chosen director belongs to a certain party, then the remaining vacancies must be filled with candidates belonging to the other party. Let the set of issues be $\mathcal{I} = \{D{=}c_j, T{=}c_j, S{=}c_j \mid j = 1, \dots, 4\}$. In order for each ballot to correspond to a committee we need to add the following integrity constraints:[1] $D{=}c_j \rightarrow \bigwedge_{k \neq j} \neg D{=}c_k$ and $\bigvee_{j=1,\dots,4} D{=}c_j$ and similarly for

[1]As a shorthand for $p_{D=c_j}$, which stands for the propositional variable associated to issue $D{=}c_j$, we directly use the name of the issue.

T and S. Finally, to encode the requirement of political balance, we add the following formulas:

$$(D{=}c_1 \vee D{=}c_2) \rightarrow (T{=}c_3 \vee T{=}c_4) \wedge (S{=}c_3 \vee S{=}c_4)$$
$$(D{=}c_3 \vee D{=}c_4) \rightarrow (T{=}c_1 \vee T{=}c_2) \wedge (S{=}c_1 \vee S{=}c_2)$$

2.1.4 Paradoxes and Collective Rationality

Consider the situation introduced in Example 2.1.3: There are three issues at stake, and the integrity constraint is represented by the formula $\mathrm{IC} = p_U \wedge p_S \rightarrow p_F$. Suppose there are three individuals, choosing ballots $(0,1,0)$, $(1,0,0)$ and $(1,1,1)$, as in Table 2.1. Their choices are rational (they all satisfy IC). Assume now that we accept an issue j if and only if a majority of individuals do, employing what we will call the *majority rule*. Then, we would obtain the ballot $(1,1,0)$ as collective outcome, which fails to be rational. This kind of observation is often referred to as a paradox.

In the literature on Social Choice Theory, situations like the one above are ruled out by requiring aggregation procedures to satisfy a property called *collective rationality*, which forces the output of an aggregation procedure to be of the same form as the input, i.e., a rational ballot. In preference aggregation, for instance, the output of an aggregation procedure is often required to be a linear (or weak) order over a set of alternatives (Gaertner, 2006). In judgment aggregation the output is required to be a complete and consistent judgment over a set of propositional formulas (List and Puppe, 2009). In view of our general perspective on aggregation problems, we give here a definition of collective rationality that depends on the integrity constraint at hand:

Definition 2.1.8. Given an integrity constraint $\mathrm{IC} \in \mathcal{L}_{PS}$, an aggregation procedure $F : \mathcal{D}^{\mathcal{N}} \rightarrow \mathcal{D}$ is called *collectively rational* (CR) with respect to IC, if for all rational profiles $\boldsymbol{B} \in \mathrm{Mod}(\mathrm{IC})^{\mathcal{N}}$ we have that $F(\boldsymbol{B}) \in \mathrm{Mod}(\mathrm{IC})$.

Thus, F is CR with respect to IC if it *lifts* the rationality assumption given by IC from the individual to the collective level, i.e., if $F(\boldsymbol{B}) \models \mathrm{IC}$ whenever $B_i \models \mathrm{IC}$ for all $i \in \mathcal{N}$. An aggregation procedure that is CR with respect to IC cannot generate a paradoxical situation with IC as integrity constraint. From Definition 2.1.8 we can obtain a general definition of paradoxical behaviour of an aggregation procedure:

Definition 2.1.9. A *paradox* is a triple $(F, \boldsymbol{B}, \mathrm{IC})$, where $F : \mathcal{D}^{\mathcal{N}} \rightarrow \mathcal{D}$ is an aggregation procedure, $\boldsymbol{B}$ is a profile in $\mathcal{D}^{\mathcal{N}}$, IC is an integrity constraint in $\mathcal{L}_{PS}$, and $B_i \in \mathrm{Mod}(\mathrm{IC})$ for all $i \in \mathcal{N}$ but $F(\boldsymbol{B}) \notin \mathrm{Mod}(\mathrm{IC})$.

In Chapter 3 we explore the generality of Definition 2.1.9 by showing that classical paradoxes introduced in several frameworks of aggregation are instances of this definition. In Chapter 4 we study, for several fragments of the propositional language the class of procedures that are CR with respect to all integrity constraints in a given language, characterising it in axiomatic terms.

2.1.5 Rationality Constraints vs. Feasibility Constraints

An important remark needs to be made about the nature of integrity constraints. In the previous section, we have introduced the concept of integrity constraint as formalising a rationality assumption that separates ballots that are rational from those to which a meaning cannot even be attached. In our framework, both the individuals and the collective outcome should comply with the constraints that define a problem. This, however, is not the only way in which integrity constraints can be used. Consider the following example.

Example 2.1.10. A committee $\mathcal{N}$ has to decide on whether to accept or reject three bills A, B and C. For budgetary reasons, only two of the three bills can be financed. This can be modelled in binary aggregation using a set of three issues $\mathcal{I} = \{A, B, C\}$ and integrity constraint $\neg(p_A \wedge p_B \wedge p_C)$.

In this example, the integrity constraint expresses a condition of feasibility rather than rationality, partitioning the set of ballots into those that are feasible and those that are not. Similar constraints are usually enforced on the outcome, but may not be imposed on individuals. This is because we may be interested in knowing the sincere evaluations of individuals, even if unfeasible, rather than having them misrepresent their judgment to satisfy the integrity constraint.

In this dissertation we use integrity constraints in their first interpretation, i.e., as rationality assumptions that need to be satisfied by both the collective and the individuals. Feasibility constraints in aggregation theory have been studied extensively using the framework of logic-based belief merging (Konieczny and Pino Pérez, 2002, 2011). A combination of the two approaches constitutes a highly promising direction for future work.

2.2 The Axiomatic Method

Aggregation procedures are traditionally studied using the axiomatic method. Axioms are used to express desirable properties of an aggregation procedure, and these axioms are then combined in an attempt to find the most desirable aggregation system. This methodology is widespread in the whole literature on Economic Theory, as testified by several important results which were proven using the axiomatic method in a number of disciplines: notable examples are the definition of the Nash solution for bargaining problems (Nash, 1950), the treatment by von Neumann and Morgenstern (1947) of decision making under uncertainty and, finally, Arrow's Theorem in preference aggregation (Arrow, 1963). In this section, we adapt the most important axioms familiar from standard Social Choice Theory, and more specifically from judgment aggregation (List and Puppe, 2009) and binary aggregation (Dokow and Holzman, 2010a), to our framework.

Let $\mathcal{X} \subseteq \mathcal{D}^{\mathcal{N}}$ be a subset of the set of profiles. The first axiomatic property we take into consideration is called unanimity:

Unanimity (U): For any profile $\boldsymbol{B} \in \mathcal{X}$ and any $x \in \{0,1\}$, if $b_{i,j} = x$ for all $i \in \mathcal{N}$, then $F(\boldsymbol{B})_j = x$.

Unanimity postulates that, if all individuals agree on issue j, then the aggregation procedure should implement that choice for j. This axiom stems from a reformulation of the Paretian requirement, which is traditionally assumed in preference aggregation. In Chapter 5 we discuss this correspondence in more detail. Several weaker versions of this axiom have been proposed. The most common assumption in the literature on judgment aggregation (List and Puppe, 2009) requires the individuals to agree on *all* issues in order for the collective outcome to agree with the individual ballot. This notion is considerably weaker than our axiom of unanimity, but the two formulations are equivalent for procedures satisfying the additional axiom of independence (which we shall see soon).

Another common property is the requirement that an aggregation procedure should treat all issues in the same way. We call this axiom issue-neutrality:

Issue-Neutrality ($\mathrm{N}^{\mathcal{I}}$): For any two issues $j, j' \in \mathcal{I}$ and any profile $\boldsymbol{B} \in \mathcal{X}$, if for all $i \in \mathcal{N}$ we have that $b_{i,j} = b_{i,j'}$, then $F(\boldsymbol{B})_j = F(\boldsymbol{B})_{j'}$.

The axiom of issue-neutrality often comes paired with another requirement of symmetry between issues, that focuses on the possible values that issues can take.[2] We propose this axiom under the name of domain-neutrality:

Domain-Neutrality ($\mathrm{N}^{\mathcal{D}}$): For any two issues $j, j' \in \mathcal{I}$ and any profile $\boldsymbol{B} \in \mathcal{X}$, if $b_{i,j} = 1 - b_{i,j'}$ for all $i \in \mathcal{N}$, then $F(\boldsymbol{B})_j = 1 - F(\boldsymbol{B})_{j'}$.

This axiom is a generalisation to the case of multiple issues of the axiom of neutrality introduced by May (1952). The two notions of neutrality above are independent from each other but dual: issue-neutrality requires the outcome on two issues to be the same if all individuals agree on these issues; domain-neutrality requires them to be reversed if all the individuals make opposed choices on the two issues.

The following property requires the aggregation to be a symmetric function of its arguments, and it is traditionally called anonymity.

[2] Sometimes the two conditions are paired together in a single requirement of neutrality (see, e.g., Riker, 1982, Chapter 3).

Anonymity (A): For any profile $\boldsymbol{B} \in \mathcal{X}$ and any permutation $\sigma : \mathcal{N} \to \mathcal{N}$, we have that $F(B_1, \ldots, B_n) = F(B_{\sigma(1)}, \ldots, B_{\sigma(n)})$.

The next property we introduce has played a crucial role in several studies in Social Choice Theory, and comes under the name of independence:

Independence (I): For any issue $j \in \mathcal{I}$ and any two profiles $\boldsymbol{B}, \boldsymbol{B}' \in \mathcal{X}$, if $b_{i,j} = b'_{i,j}$ for all $i \in \mathcal{N}$, then $F(\boldsymbol{B})_j = F(\boldsymbol{B}')_j$.

This axiom requires the outcome of aggregation on a certain issue j to depend only on the individual choices regarding that issue. In preference aggregation the corresponding axiom is called *independence of irrelevant alternatives.* In the literature on judgment aggregation, the combination of independence and issue-neutrality takes the name of *systematicity.* This axiom is at the basis of the "welfaristic view" for ordinal utility in Social Choice Theory (see Roemer, 1996, p. 28). This research assumption states that a society, in making its choices, should only be concerned with the well-being of its constituents, discarding all "non-utility information"; in particular, past behaviour or hypothetical situations other than the one a society is facing (independence), and particular characteristics or correlations between the issues at hand (issue-neutrality). In the same spirit, the axiom of anonymity requires that the collective decision should disregard names, weights or importance of the individuals in a society.

We now introduce two axioms of monotonicity. The first, which we call I-monotonicity, is often called *positive responsiveness* and is formulated as an (inter-profile) axiom for independent aggregation procedures. The second version of monotonicity is designed for neutral procedures, and it was introduced by Endriss et al. (2010a):

I-Monotonicity (M^{I}): For any issue $j \in \mathcal{I}$ and any two profiles $\boldsymbol{B}, \boldsymbol{B}' \in \mathcal{X}$, if $b_{i,j} = 1$ entails $b'_{i,j} = 1$ for all $i \in \mathcal{N}$, and for some $s \in \mathcal{N}$ we have that $b_{s,j} = 0$ and $b'_{s,j} = 1$, then $F(\boldsymbol{B})_j = 1$ entails $F(\boldsymbol{B}')_j = 1$

N-Monotonicity (M^{N}): For any two issues $j, j' \in \mathcal{I}$ and any profile $\boldsymbol{B} \in \mathcal{X}$, if for all $i \in \mathcal{N}$ we have that $b_{i,j} = 1$ entails $b_{i,j'} = 1$ and for some $s \in \mathcal{N}$ we have that $b_{s,j} = 0$ and $b_{s,j'} = 1$, then $F(\boldsymbol{B})_j = 1$ entails $F(\boldsymbol{B})_{j'} = 1$.

M^{I} expresses that, if an issue j is collectively accepted and receives additional support (from an individual s), then it should continue to be collectively accepted. On the other hand, axiom M^{N} says that, if issue j is collectively accepted and issue j' is accepted by a strict superset of the individuals accepting j, then j' should also be collectively accepted. Under the assumption of systematicity the two versions of monotonicity are equivalent.

Not all aggregation procedures satisfy each of these axioms. The literature on Social Choice Theory is plagued with impossibility results showing that there is no aggregator satisfying certain combinations of axioms. To see an example, consider a simple constant procedure that outputs the same ballot in every profile. This procedure is independent and issue-neutral, since it treats all issues in the same way in every profile. On the other hand, it is neither domain-neutral nor unanimous.

The last property for aggregation procedures that we are going to introduce is traditionally considered a negative one. We choose not to state it as an axiom, but rather as a property defining a class of functions.

Definition 2.2.1. An aggregation procedure F is called a *dictatorship* if there exists an individual $i \in \mathcal{N}$ such that $F(\boldsymbol{B}) = B_i$ for all profiles $\boldsymbol{B}$.

A dictatorship copies the ballot of the same individual in every profile. This notion is in clear conflict with the axiom of anonymity previously introduced. In Definition 4.2.7 we will generalise this notion by defining the class of *generalised dictatorships* as those procedures that copy the ballot of a (possibly different) individual in every profile.

We conclude with an important remark. It is crucial to observe that all axioms are *domain-dependent*: It is possible that an aggregation procedure satisfies an axiom only on a subdomain $\mathcal{X} \subseteq \mathcal{D}$ in which individuals can choose their ballots. For instance, consider the following example. With two issues, let $\mathrm{IC} = (p_2 \to p_1)$ and let F accept the first issue if a majority of the individuals accept it, and accept the second issue only if the first one was accepted and the second one has the support of a majority of individuals. This procedure is clearly not independent on the full domain, but it is easy to see that it satisfies independence when restricted to $\mathcal{X} = \mathrm{Mod}(\mathrm{IC})^{\mathcal{N}}$. We will make extensive use of this fact in Chapter 4.

One last observation: Note that the notion of integrity constraint never occurs in the definition of an axiom. This choice reflects the view that axiomatic properties should be separate from the behaviour of an aggregation procedure with respect to domain specific problems like that of collective rationality.

2.3 Representation Results

Axiomatic properties like the ones we introduced in Section 2.2 can be used to define classes of procedures, and in Chapter 4 we are going to make use of this construction extensively. In this section, we start investigating some of these classes providing, for some of them, a useful mathematical representation. All the results proved in this section are adaptations of known results from the literature, even if those results are rarely stated explicitly.

Let us first introduce some further notation: denote with $N_j^{\boldsymbol{B}} = \{i \in \mathcal{N} \mid b_{i,j} = 1\}$ the set of individuals accepting issue j in profile $\boldsymbol{B}$. We begin by considering aggregation procedures that satisfy the axiom of independence, proving a representation result in terms of *winning coalitions*:[3]

Proposition 2.3.1. *An aggregation procedure F satisfies* I *if and only if for every issue j there exists a collection of subsets $\mathcal{W}_j \subseteq \mathcal{P}(\mathcal{N})$ such that $F(\boldsymbol{B})_j = 1$ if and only $N_j^{\boldsymbol{B}} \in \mathcal{W}_j$. Let $\mathcal{W}_j$ be the set of winning coalitions of F for issue j.*

Proof. Let F be an independent procedure, and let j be an issue in $\mathcal{I}$. Define $\mathcal{W}_j$ as the set of all sets $A \subseteq \mathcal{N}$ such that there exists a profile $\boldsymbol{B}$ with $N_j^{\boldsymbol{B}} = A$ and $F(\boldsymbol{B})_j = 1$. As F is independent, for every profile $\boldsymbol{B}'$ with $N_j^{\boldsymbol{B}'} = A$ we have that $F(\boldsymbol{B}')_j = 1$. Thus, F is defined by the set of winning coalitions $\mathcal{W}_j$. On the other hand, given a set of winning coalitions $\mathcal{W}_j$ for F, let $\boldsymbol{B}$ and $\boldsymbol{B}'$ be two distinct profiles such that $b_{i,j} = b'_{i,j}$. It is straightforward to observe that this implies that $N_j^{\boldsymbol{B}} = N_j^{\boldsymbol{B}'}$, and hence that F has the same outputs on j in the two profiles. Thus, F is independent. □

When combined with issue-neutrality, independence generates procedures that are defined by a single set of winning coalitions, the same for every issue:

Corollary 2.3.2. *An aggregation procedure F satisfies* I *and* $\mathrm{N}^{\mathcal{I}}$ *if and only if there exists a collection of subsets $\mathcal{W} \subseteq \mathcal{P}(\mathcal{N})$ such that $F(\boldsymbol{B})_j = 1$ if and only if $N_j^{\boldsymbol{B}} \in \mathcal{W}$.*

Let us now consider the case of procedures that satisfy the axioms of anonymity, independence and issue-neutrality. We prove that for these procedures the acceptance of an issue depends solely on the number of individuals accepting it.[4]

Proposition 2.3.3. *An aggregation procedure F satisfies* A, I *and* $\mathrm{N}^{\mathcal{I}}$ *on the full domain $\mathcal{D}$ if and only if there exists a function $h : \{0, \ldots, |\mathcal{N}|\} \to \{0, 1\}$ such that $F(\boldsymbol{B})_j = 1 \Leftrightarrow h(|N_j^{\boldsymbol{B}}|) = 1$.*

Proof. We prove that if F satisfies A, I and $\mathrm{N}^{\mathcal{I}}$ over the full domain $\mathcal{D}$, then $|N_j^{\boldsymbol{B}}| = |N_{j'}^{\boldsymbol{B}'}|$ for profiles $\boldsymbol{B}, \boldsymbol{B}'$ and issues j, j' implies $F(\boldsymbol{B})_j = F(\boldsymbol{B}')_{j'}$. Thus, the fact that a set of individuals is a winning coalition depends only on the cardinality of the set, which can be specified using a function $h : \{0, \ldots, |\mathcal{N}|\} \to \{0, 1\}$ as in the statement of Proposition 2.3.3.

Let F be an anonymous, independent and issue-neutral procedure. Since $|N_j^{\boldsymbol{B}}| = |N_{j'}^{\boldsymbol{B}'}|$, we can rearrange the individuals in profile $\boldsymbol{B}'$ obtaining profile $\boldsymbol{C}$ such that $c_{i,j'} = b_{i,j}$. By anonymity the result of F on issue j does not change moving from profile $\boldsymbol{B}'$ to $\boldsymbol{C}$. Let us now construct a fourth profile $\boldsymbol{D}$ such that $d_{i,j} = b_{i,j}$ and $d_{i,j'} = c_{i,j'}$ for all $i \in \mathcal{N}$. By independence, we have that $F(\boldsymbol{B})_j = F(\boldsymbol{D})_j$. By issue-neutrality, we have that $F(\boldsymbol{D})_j = F(\boldsymbol{D})_{j'}$. It is now sufficient to apply independence one more time to obtain that $F(\boldsymbol{D})_{j'} = F(\boldsymbol{C})_{j'}$ and conclude the chain of equalities by using the axiom of anonymity. □

[3]Rules defined in terms of winning coalitions are sometimes referred to as "voting by committee" in the literature on Social Choice Theory (Barberà et al., 1991).

[4]This is a known result. List and Pettit (2002), for instance, use this insight (adapted to the case of judgment aggregation) in the proof of their impossibility theorem.

Representation results along the lines of Proposition 2.3.3 can easily be obtained for various classes of procedures. Examples can be found, for instance, in our previous work (Endriss et al., 2010a).

Note that a (somewhat surprising) consequence of Proposition 2.3.3 is the following corollary:[5]

Corollary 2.3.4. *If the number of individuals is even and there are at least two issues, then there exists no aggregation procedure that satisfies* A, I, $\mathrm{N}^{\mathcal{I}}$ *and* $\mathrm{N}^{\mathcal{D}}$ *on the full domain.*

Proof. Let j and j' be two issues, and let $\boldsymbol{B}$ be a profile such that exactly half of the individuals accept j and reject j', and the other half reject j and accept j'. By $\mathrm{N}^{\mathcal{D}}$ we have that $F(\boldsymbol{B})_j = 1 - F(\boldsymbol{B})_{j'}$. Since the number of individuals accepting the two issues is exactly the same, by Proposition 2.3.3 we obtain the additional requirement that $F(\boldsymbol{B})_j = F(\boldsymbol{B})_{j'}$, in contradiction with the requirement given by domain-neutrality. □

2.3.1 Quota Rules

An aggregation procedure F for n individuals is a *quota rule* if for every issue j there exists a quota $0 \leqslant q_j \leqslant n+1$ such that $F(\boldsymbol{B})_j = 1$ if and only if $|N_j^{\boldsymbol{B}}| \geqslant q_j$. The class of quota rules, which we denote as $\mathcal{QR}$, was introduced by Dietrich and List (2007a) in the framework of judgment aggregation. Quota rules are axiomatised as the class of procedures satisfying the axioms of anonymity, independence and I-monotonicity, as we prove in the following result:[6]

Proposition 2.3.5. *An aggregation procedure F satisfies* A, I, *and* M^{I} *on the full domain $\mathcal{D}$ if and only if it is a quota rule.*

Proof. First, observe that if we add the assumption of anonymity to the statement of Proposition 2.3.1, we obtain a representation for independent and anonymous aggregation procedures. These functions can be characterised in terms of a set of acceptance functions indicating, for each issue j, the size of possible winning coalitions, i.e., functions $h_j : |\mathcal{N}| \to \{0,1\}$ for each j such that $F(\boldsymbol{B})_j = 1 \Leftrightarrow h_j(|N_j^{\boldsymbol{B}}|) = 1$. If we now add the assumption of monotonicity in its independence version, and we apply it to issue j, we can infer that whenever $h_j(m) = 1$ for a certain m, then for all $t \geqslant m$ it must be the case that $h_j(t) = 1$. Thus, we can define an acceptance quota for each issue by letting q_j be the minimal $m \leqslant |\mathcal{N}|$ such that $h_j(m) = 1$. In case such an m does not exist (i.e., in case $h_j(t) = 0$ for all t), it is sufficient to fix $q_j = n+1$. □

[5] Observe that Corollary 2.3.4 is an impossibility result which does not feature any requirement of collective rationality. This result is related to the known fact that there exists no resolute voting procedure for 2 alternatives and an even number of individuals which is anonymous and neutral (see, e.g., Moulin, 1983, for a generalisation of this result).

[6] An analogous version of Proposition 2.3.5 for the framework of judgment aggregation can be found in the work of Dietrich and List (2007a).

A quota rule is called *uniform* if the quota is the same for all issues. We can obtain an axiomatisation of this class by adding the axiom of issue-neutrality to Proposition 2.3.5:

Corollary 2.3.6. *An aggregation procedure F satisfies* A, I, $\mathrm{N}^{\mathcal{I}}$ *and* M^{I} *on the full domain $\mathcal{D}$ if and only if it is a uniform quota rule.*

2.3.2 The Majority Rule

A particular quota rule, which we study in detail in Section 4.4.2, is the *majority rule*. The majority rule is the uniform quota rule that accepts an issue if and only if a majority of individuals accept it. In case the number of individuals is odd, the majority rule has a unique definition by setting the quota to $q = \frac{n+1}{2}$. In case the number of individuals is even, the majority rule does not have a unique definition, to account for ties between acceptances and rejections. One possibility is to favour rejection, defining the *strict majority rule* with quota $q = \frac{n+2}{2}$, or to favour acceptance, defining the *weak majority rule* with quota $q = \frac{n}{2}$. We study these rules in more detail in Section 4.4.2. We now make the assumption that the number of individuals is odd and we provide an axiomatisation of the majority rule in this case.

May (1952) provided an axiomatisation of the majority rule in the case of preference aggregation over two alternatives. We can obtain a more general version of his result, which accounts for the case of multiple issues, by adding the axioms of issue-neutrality and domain-neutrality to Proposition 2.3.5:

Proposition 2.3.7. *If the number of individuals is odd, an aggregation procedure F satisfies* A, $\mathrm{N}^{\mathcal{I}}$, $\mathrm{N}^{\mathcal{D}}$, I *and* M^{I} *if and only if it is the majority rule.*

Proof. By Corollary 2.3.6, we know that F is a uniform quota rule. The axiom of domain-neutrality then forces us to treat the two sets $N_j^{\boldsymbol{B}}$ and $\mathcal{N} \setminus N_j^{\boldsymbol{B}}$ symmetrically. Hence, the only possibility is to fix the quota at $\frac{n+1}{2}$. □

2.4 Previous Work on Binary Aggregation

Wilson (1975) has been the first to define and study the framework of binary aggregation. His seminal paper contains several impossibility results for independent aggregation procedures, including a generalisation of the famous impossibility result by Arrow (1963). Wilson starts from an attribute space A of properties that can be assigned to individuals, and studies the aggregation of such assignments. This setting corresponds to aggregating binary ballots using A as the set of issues. Given a collection $\mathcal{B}$ of subsets of the space of all assignments $\mathcal{D}$, Wilson calls an aggregation procedure "responsive" with respect to $\mathcal{B}$ if, whenever all individual assignments belong to a subset $\mathcal{X} \in \mathcal{B}$, then the outcome also belongs to $\mathcal{X}$. He then proves several representation results for independent and responsive procedures depending on the structure of the collection of subsets $\mathcal{B}$. Arrow's Theorem is obtained as a corollary of one of his results. Wilson's notion of responsive aggregator corresponds to our notion of collective rationality with respect to a *family* of integrity constraints. We will follow the same approach in Chapter 4, when we concentrate on the study of collectively rational procedures with respect to a *language* of integrity constraints.

A similar setting has been investigated more recently by Dokow and Holzman (2009, 2010a). Their definition of collective rationality is the same as Wilson's, although they consider a single subdomain $\mathcal{X} \subseteq \mathcal{D}$ of rational ballots at a time rather than a family of such subsets. As previously remarked, propositional logic is fully expressive with respect to subsets of $\mathcal{D}$, hence our approach is equivalent to that of Dokow and Holzman. Our choice of using formulas rather than sets is motivated by the possibility of classifying integrity constraints by means of syntactic properties and by the the compactness of this representation. Consider for instance the subset defined by any non-complete cube, i.e., a conjunction in which not all literals occur. The representation of this set by means of a formula is exponentially more concise than the full list of elements of the same set.

Another framework for binary aggregation is adopted by Nehring and Puppe (2007, 2010). Although their aim is more general, they also concentrate on the study of aggregation procedures over property spaces, a setting that is closer to the original framework of Wilson (1975). We refer to Section 6.2.3 for a more detailed discussion of this framework.

In several papers (see, e.g., List and Puppe, 2009; Nehring and Puppe, 2010; Dokow and Holzman, 2010a) it has been observed that the framework of judgment aggregation for propositional logic is equivalent to that of binary aggregation (see also our Example 2.1.5). In Chapter 6, in particular Section 6.2.3, we discuss in detail the relation between these diverse frameworks for aggregation.

Rubinstein and Fishburn (1986) generalised Wilson's framework allowing individuals to choose elements of certain vector spaces. The case of binary aggregation is subsumed by considering the vector space $\mathcal{D} = \{0, 1\}^{\mathcal{I}}$.

An important, although not substantial, difference between our framework and classical approaches to binary aggregation resides in our definition of aggregation procedure. Both Dokow and Holzman (2010a) and Nehring and Puppe (2007) define an aggregation procedure on a specific domain $\mathcal{X} \subseteq \{0, 1\}^m$, including in this definition the notion of collective rationality with respect to the integrity constraint that defines $\mathcal{X}$. The same approach is also used in the literature on judgment aggregation (List and Puppe, 2009). Instead, we define aggregation procedures on all possible profiles, studying collective rationality as an additional property. As already pointed out at the end of Section 2.2, our choice is motivated by an attempt to separate the definition of an aggregation procedure and its axiomatic properties from the notion of collective rationality, which depends on the domain of rational ballots on which the aggregation is performed.

Chapter 3

Paradoxes of Aggregation

Most work in Social Choice Theory started with the observation of paradoxical situations. From the Marquis de Condorcet (1785) and Jean-Charles de Borda (1781) to more recent American court cases (Kornhauser and Sager, 1986), a wide collection of paradoxes have been analysed and studied in the literature on Social Choice Theory (see, e.g., Nurmi, 1999). In this chapter we present some of the most well-known paradoxes that arise from the use of the majority rule in different contexts, and we show how they can be expressed in binary aggregation as instances of our Definition 2.1.9. Such a uniform representation of the most important paradoxes in Social Choice Theory enables us to make a crucial observation concerning the syntactic structure of paradoxical integrity constraints: they all feature a disjunction of literals of size at least 3. This observation will give rise to one of the main theorems of this dissertation (Theorem 4.4.8).

In Section 3.1, we introduce one of the most notable paradoxes in Social Choice Theory, the Condorcet paradox, and we show how settings of preference aggregation can be seen as instances of binary aggregation by devising a suitable integrity constraint. Section 3.2 repeats this construction for the framework of judgment aggregation and for the paradoxical example which gave rise to this area of research, namely the doctrinal paradox. In Section 3.3 we deal with the Ostrogorski paradox, in which a paradoxical feature of representative majoritarian systems is analysed. In Section 3.5 we introduce two further paradoxes generated by the use of the majority rule on multiple issues: the paradox of divided government and the paradox of multiple elections. In Section 3.6 we conclude.

3.1 The Condorcet Paradox and Preference Aggregation

During the Enlightment period in France, several active scholars dedicated themselves to the problem of collective choice, and in particular to the creation of new procedures for the election of candidates. Although these are not the first documented studies of the problem of social choice (McLean and Urken, 1995), Marie Jean Antoine Nicolas de Caritat, the Marquis de Condorcet, was the first to point out a crucial problem of the most basic voting rule that was being used, the majority rule (Condorcet, 1785). The paradox he discovered, that now comes under his name, is explained in the following paragraphs:

> **Condorcet Paradox**. Three individuals need to decide on the ranking of three alternatives $\{\triangle, \bigcirc, \square\}$. Each individual expresses her own ranking in the form of a linear order, i.e., an irreflexive, transitive and complete binary relation over the set of alternatives. The collective outcome is then aggregated by pairwise majority: an alternative is preferred to a second one if and only if a majority of the individuals prefer the first alternative to the second. Consider the profile described in Table 3.1.

$$\begin{array}{c} \triangle <_1 \bigcirc <_1 \square \\ \square <_2 \triangle <_2 \bigcirc \\ \bigcirc <_3 \square <_3 \triangle \\ \hline \triangle < \bigcirc < \square < \triangle \end{array}$$

Table 3.1: The Condorcet paradox.

> When we compute the outcome of the pairwise majority rule on this profile, we notice that there is a majority of individuals preferring the circle to the triangle ($\triangle < \bigcirc$); that there is a majority of individuals preferring the square to the circle ($\bigcirc < \square$); and, finally, that there is a majority of individuals preferring the triangle to the square ($\square < \triangle$). The resulting outcome fails to be a linear order, giving rise to a circular collective preference between the alternatives.

Condorcet's paradox was rediscovered in the second half of the XXth century while a whole theory of *preference aggregation* was being developed, starting with the work of Black (1958) and Arrow's celebrated result (Arrow, 1963). In this section, we review the framework of preference aggregation, we show how this setting can be embedded into the framework of binary aggregation with integrity constraints, and we show how the Condorcet paradox can be seen as an instance of our general definition of paradox (Definition 2.1.9).

3.1.1 Preference Aggregation

The framework of *preference aggregation* (see, e.g., Gaertner, 2006) considers a finite set of individuals $\mathcal{N}$ expressing preferences over a finite set of alternatives $\mathcal{X}$. A preference relation is represented by a binary relation over $\mathcal{X}$. Preference relations are traditionally assumed to be *weak orders*, i.e., reflexive, transitive and complete binary relations. In some cases, in order to simplify the framework, preferences are assumed to be *linear orders*, i.e., irreflexive, transitive and complete binary relations. In the first case, we write aRb for "alternative a is preferred to alternative b or it is equally preferred as b", while in the second case aPb stands for "alternative a is strictly preferred to b". In the sequel we shall assume that preferences are represented as linear orders. We refer to Chapter 5 for a more detailed presentation of other assumptions in preference aggregation.

Each individual submits a linear order P_i, forming a profile $\boldsymbol{P} = (P_1, \ldots, P_{|\mathcal{N}|})$. Let $\mathcal{L}(\mathcal{X})$ denote the set of all linear orders on $\mathcal{X}$. Aggregation procedures in this framework are called *social welfare functions* (SWFs):

Definition 3.1.1. Given a finite set of individuals $\mathcal{N}$ and a finite set of alternatives $\mathcal{X}$, a *social welfare function* is a function $F : \mathcal{L}(\mathcal{X})^{\mathcal{N}} \to \mathcal{L}(\mathcal{X})$.

Note that a SWF is defined for every logically possible profile of linear orders, a condition that traditionally goes under the name of universal domain, and that it always outputs a linear order. This last condition was given the name of "collective rationality" by Arrow (1963). As we have seen in Table 3.1, the Condorcet paradox proves that the pairwise majority rule is not a SWF because, in Arrow's words, it fails to be "collectively rational". In the following section we will formalise this observation by devising an integrity constraint that encodes the assumptions underlying Arrow's framework of preference aggregation.

3.1.2 Translation

Given a preference aggregation problem defined by a set of individuals $\mathcal{N}$ and a set of alternatives $\mathcal{X}$, let us now consider the following setting for binary aggregation. Define a set of issues $\mathcal{I}_{\mathcal{X}}$ as the set of all pairs (a, b) in $\mathcal{X}$. The domain $\mathcal{D}_{\mathcal{X}}$ of aggregation is therefore $\{0, 1\}^{|\mathcal{X}|^2}$. In this setting, a binary ballot B corresponds to a binary relation P over $\mathcal{X}$: $B_{(a,b)} = 1$ if and only if a is in relation to b (aPb). Given this representation, we can associate with every SWF for $\mathcal{X}$ and $\mathcal{N}$ an aggregation procedure that is defined on a subdomain of $\mathcal{D}_{\mathcal{X}}^{\mathcal{N}}$. We now characterise this domain as the set of models of a suitable integrity constraint.

Using the propositional language $\mathcal{L}_{PS}$ constructed over the set $\mathcal{I}_{\mathcal{X}}$, we can express properties of binary ballots in $\mathcal{D}_{\mathcal{X}}$. In this case the language consists of $|\mathcal{X}|^2$ propositional symbols, which we shall call p_{ab} for every issue (a, b). As we already anticipated in Example 2.1.4, the properties of linear orders can be enforced on binary ballots using the following set of integrity constraints, which we shall call $\mathrm{IC}_<$:[1]

Irreflexivity: $\neg p_{aa}$ for all $a \in \mathcal{X}$

Completeness: $p_{ab} \vee p_{ba}$ for all $a \neq b \in \mathcal{X}$

Transitivity: $p_{ab} \wedge p_{bc} {\rightarrow} p_{ac}$ for $a, b, c \in \mathcal{X}$ pairwise distinct

Note that the size of this set of integrity constraints is polynomial in the number of alternatives in $\mathcal{X}$. It is now straightforward to see that every SWF corresponds to an aggregation procedure that is collectively rational with respect to $\mathrm{IC}_<$ and *vice versa*.[2]

In case preferences are expressed using weak orders rather than linear orders, it is sufficient to modify the integrity constraint $\mathrm{IC}_<$ to obtain a similar correspondence between SWFs and aggregation procedures. Recall that a weak order is a reflexive, transitive and complete binary relation over $\mathcal{X}$. Let therefore $\mathrm{IC}_\leqslant$ be the following set of integrity constraints:

Reflexivity: p_{aa} for all $a \in \mathcal{X}$

Completeness: $p_{ab} \vee p_{ba}$ for all $a \neq b \in \mathcal{X}$

Transitivity: $p_{ab} \wedge p_{bc} {\rightarrow} p_{ac}$ for $a, b, c \in \mathcal{X}$ pairwise distinct

3.1.3 The Condorcet Paradox in Binary Aggregation

The translation presented in the previous section enables us to express the Condorcet paradox in terms of Definition 2.1.9. Let $\mathcal{X} = \{\triangle, \bigcirc, \square\}$ and let $\mathcal{N}$ contain three individuals. Consider the profile $\boldsymbol{B}$ for $\mathcal{I}_\mathcal{X}$ described in Table 3.2, where we have omitted the values of the reflexive issues $(\triangle, \triangle)$ (always 0 by $\mathrm{IC}_<$), and specified the value of only one of $(\triangle, \bigcirc)$ and $(\bigcirc, \triangle)$ (the other can be obtained by taking the opposite of the value of the first), and accordingly for the other alternatives. Every individual ballot in Table 3.2 satisfies $\mathrm{IC}_<$, but the outcome obtained using the majority rule *Maj* (which corresponds to pairwise majority in preference aggregation) does not satisfy $\mathrm{IC}_<$: the formula $p_{\triangle\bigcirc} \wedge p_{\bigcirc\square} \rightarrow p_{\triangle\square}$ is falsified by the outcome. Therefore, $(Maj, \boldsymbol{B}, \mathrm{IC}_<)$ is a paradox by Definition 2.1.9.

The integrity constraint $\mathrm{IC}_<$ can be further simplified for the case of 3 alternatives $\{a, b, c\}$. The formulas encoding the transitivity of binary relations are equivalent to just two positive clauses: The first one, $p_{ba} \vee p_{cb} \vee p_{ac}$, rules out the cycle $a{<}b{<}c{<}a$, and the second one, $p_{ab} \vee p_{bc} \vee p_{ca}$, rules out the opposite cycle $c{<}b{<}a{<}c$. That is, these constraints correspond exactly to the two Condorcet cycles that can be created from three alternatives.

[1] We will use the notation IC both for a single integrity constraint and for a set of formulas—in the latter case considering as the actual constraint the conjunction of all the formulas in IC.

[2] A technicality: to every SWF correspond *many* binary aggregation procedures, depending on how we extend the procedure outside of $\mathrm{Mod}(\mathrm{IC}_<)^\mathcal{N}$.

	△○	○□	△□
Agent 1	1	1	1
Agent 2	1	0	0
Agent 3	0	1	0
Maj	1	1	0

Table 3.2: Condorcet paradox in binary aggregation.

3.2 The Discursive Dilemma and Judgment Aggregation

The discursive dilemma emerged from the formal study of court cases that was carried out in recent years in the literature on law and economics, generalising the observation of a paradoxical situation known as the "doctrinal paradox" (Kornhauser and Sager, 1986, 1993). Such a setting was first given mathematical treatment by List and Pettit (2002), giving rise to an entirely new research area in Social Choice Theory known as *judgment aggregation*. Earlier versions of this paradox can be found in work by Guilbaud (1952) and Vacca (1922). We now describe one of the most common versions of the discursive dilemma:

> **Discursive Dilemma**. A court composed of three judges has to decide on the liability of a defendant under the charge of breach of contract. According to the law, the individual is liable if there was a valid contract and her behaviour was such as to be considered a breach of the contract. The court takes three majority decisions on the following issues: there was a valid contract (α), the individual broke the contract (β), the defendant is liable ($\alpha \wedge \beta$). Consider a situation like the one described in Table 3.3.

	α	β	$\alpha \wedge \beta$
Judge 1	yes	yes	yes
Judge 2	no	yes	no
Judge 3	yes	no	no
Majority	yes	yes	no

Table 3.3: The discursive dilemma.

> All judges are expressing consistent judgments: they accept the third proposition if and only if the first two are accepted. However, when aggregating the judgments using the majority rule we obtain an inconsistent outcome: even if there is a majority of judges who believe that

> there was a valid contract, and even if there is a majority of judges who believe that the individual broke the contract, the individual is considered *not liable* by a majority of the individuals.

In this section we review the framework of judgment aggregation (List and Puppe, 2009), and we provide a characterisation of judgment aggregation procedures as collectively rational procedures with respect to a suitable set of integrity constraints. This in turn enables us to show that the discursive dilemma is also an instance of our general definition of paradox. For a more detailed introduction to the framework of judgment aggregation we refer to Chapter 6.

3.2.1 Judgment Aggregation

Judgement aggregation (JA) considers problems in which a finite set of individuals $\mathcal{N}$ has to generate a collective judgment over a set of interconnected propositional formulas[3] (List and Puppe, 2009). Formally, given a finite propositional language $\mathcal{L}$, an *agenda* is a finite nonempty subset $\Phi \subseteq \mathcal{L}$ that does not contain any doubly-negated formulas and that is closed under complementation (i.e, $\alpha \in \Phi$ whenever $\neg\alpha \in \Phi$, and $\neg\alpha \in \Phi$ for every non-negated $\alpha \in \Phi$).

Each individual in $\mathcal{N}$ expresses a *judgment set* $J \subseteq \Phi$, as the set of those formulas in the agenda that she judges to be true. Every individual judgment set J is assumed to be *complete* (i.e., for each $\alpha \in \Phi$ either α or its complement are in J) and *consistent* (i.e., there exists an assignment that makes all formulas in J true). If we denote by $\mathcal{J}(\Phi)$ the set of all complete and consistent subsets of Φ, we can give the following definition:

Definition 3.2.1. Given a finite agenda Φ and a finite set of individuals $\mathcal{N}$, a *JA procedure* for Φ and $\mathcal{N}$ is a function $F : \mathcal{J}(\Phi)^{\mathcal{N}} \to 2^{\Phi}$.

Note that no additional requirement is imposed on the collective judgment set. A JA procedure is called *complete* if the judgment set it returns is complete on every profile. A JA procedure is called *consistent* if, for every profile, the outcome is a consistent judgment set.

3.2.2 Translation

Given a judgment aggregation framework defined by an agenda Φ and a set of individuals $\mathcal{N}$, let us now construct a setting for binary aggregation with integrity constraints that interprets it, generalising from our previous Example 2.1.5. Let the set of issues $\mathcal{I}_\Phi$ be equal to the set of formulas in Φ. The domain $\mathcal{D}_\Phi$ of aggregation is therefore $\{0,1\}^{|\Phi|}$. In this setting, a binary ballot B corresponds

[3] We shall not treat here the case of judgment aggregation in more general logics (Dietrich, 2007). We refer to Appendix A for a brief introduction of propositional logic.

to a judgment set: $B_\alpha = 1$ if and only if $\alpha \in J$. Given this representation, we can associate with every JA procedure for Φ and $\mathcal{N}$ a binary aggregation procedure on a subdomain of $\mathcal{D}_\Phi^\mathcal{N}$.

It is important to remark that is not exactly the standard way of interpreting JA in binary aggregation. The embedding that is given, for instance, by Dokow and Holzman (2009, 2010a), associates with every judgment set a binary ballot over a set of issues representing only the *positive* formulas in Φ, considering a rejection of the issue associated with a formula φ as an acceptance of its negation $\neg\varphi$. The same embedding is given by List and Puppe (2009, Section 2.3). In our translation we made the choice of introducing both an issue for φ and one for $\neg\varphi$, adding an additional integrity constraint to enforce the completeness of a judgment set. This allow us to easily generalise the framework to the case of incomplete ballots (see, e.g., Dietrich and List, 2008a), without having to resort to an additional symbol for abstention (see, e.g., Dokow and Holzman, 2010b)

As we did for the case of preference aggregation, we now define a set of integrity constraints for $\mathcal{D}_\Phi$ to enforce the properties of consistency and completeness of individual judgment sets. Recall that the propositional language is constructed in this case on $|\Phi|$ propositional symbols p_α, one for every $\alpha \in \Phi$. Call an inconsistent set of formulas each proper subset of which is consistent *minimally inconsistent set* (mi-set). Let IC_Φ be the following set of integrity constraints:

Completeness: $p_\alpha \vee p_{\neg\alpha}$ for all $\alpha \in \Phi$

Consistency: $\neg(\bigwedge_{\alpha \in S} p_\alpha)$ for every mi-set $S \subseteq \Phi$

While the interpretation of the first formula is straightforward, we provide some further explanation for the second one. If a judgment set J is inconsistent, then it contains a minimally inconsistent set, obtained by sequentially deleting one formula at the time from J until it becomes consistent. This implies that the constraint previously introduced is falsified by the binary ballot that represents J, as all issues associated with formulas in a mi-set are accepted. *Vice versa*, if all formulas in a mi-set are accepted by a given binary ballot, then clearly the judgment set associated with it is inconsistent.

Note that the size of IC_Φ might be exponential in the size of the agenda. This is in agreement with considerations of computational complexity (see, e.g., Papadimitriou, 1994): Since checking the consistency of a judgment set is NP-hard, while model checking on binary ballots is polynomial, the translation from JA to binary aggregation must contain a superpolynomial step (unless P=NP). A more detailed discussion of the computational complexity of these two frameworks can be found in Section 7.5.

In conclusion, the same kind of correspondence we have shown for SWFs holds between complete and consistent JA procedures and binary aggregation procedures that are collectively rational with respect to IC_Φ.

3.2.3 The Discursive Dilemma in Binary Aggregation

The same procedure that we have used to show that the Condorcet paradox is an instance of our general definition of paradox applies here for the case of the discursive dilemma. Let Φ be the agenda $\{\alpha, \beta, \alpha \wedge \beta\}$, in which we have omitted negated formulas, as for any $J \in \mathcal{J}(\Phi)$ their acceptance can be inferred from the acceptance of their positive counterparts. Consider the profile $\boldsymbol{B}$ for $\mathcal{I}_\Phi$ described in Table 3.4.

	α	β	$\alpha \wedge \beta$
Judge 1	1	1	1
Judge 2	0	1	0
Judge 3	1	0	0
Maj	1	1	0

Table 3.4: The discursive dilemma in binary aggregation.

Every individual ballot satisfies IC_Φ, while the outcome obtained by using the majority rule contradicts one of the constraints of consistency, namely $\neg(p_\alpha \wedge p_\beta \wedge p_{\neg(\alpha\wedge\beta)})$. Hence, $(Maj, \boldsymbol{B}, \text{IC}_\Phi)$ constitutes a paradox by Definition 2.1.9.

3.3 The Ostrogorski Paradox

Another paradox listed by Nurmi (1999) as one of the main paradoxes of the majority rule on multiple issues is the Ostrogorski paradox. Ostrogorski (1902) published a treaty in support of procedures inspired by direct democracy, pointing out several fallacies that a representative system based on party structures can encounter. Rae and Daudt (1976) later focused on one such situation, presenting it as a paradox or a dilemma between two equivalently desirable procedures (the direct and the representative one), giving it the name of "Ostrogorski paradox". This paradox, in its simplest form, occurs when a majority of individuals are supporting a party that does not represent the view of a majority of individuals on a majority of issues.

> **Ostrogorski Paradox**. Consider the following situation: there is a two party contest between the Mountain Party (MP) and the Plain Party (PP); three individuals (or, equivalently, three equally big groups in an electorate) will vote for one of the two parties if their view agrees with that party on a majority of the three following issues: economic policy (E), social policy (S), and foreign affairs policy (F). Consider the situation described in Table 3.5.

	E	S	F	Party supported
Voter 1	MP	PP	PP	PP
Voter 2	PP	PP	MP	PP
Voter 3	MP	PP	MP	MP
Maj	MP	PP	MP	PP

Table 3.5: The Ostrogorski paradox.

> The result of the two party contest, assuming that the party that has the support of a majority of the voters wins, declares the Plain Party the winner. However, we notice that a majority of individuals support the Mountain Party both on the economic policy E *and* on the foreign policy F. Thus, the elected party, the PP, is in disagreement with a majority of the individuals on a majority of the issues.

Bezembinder and van Acker (1985) generalised this paradox, defining two different rules for compound majority decisions. The first, the representative outcome, outputs as a winner the party that receives support by a majority of the individuals. The second, the direct outcome, outputs the party that receives support on a majority of issues by a majority of the individuals. An instance of the Ostrogorski paradox occurs whenever the outcome of these two procedures differ.

Stronger versions of the paradox can be devised, in which the losing party represents the view of a majority on *all* the issues involved (see, e.g., Rae and Daudt, 1976; see also our Table 3.7). Further studies of the "Ostrogorski phenomenon" have been carried out by Deb and Kelsey (1987) as well as by Eckert and Klamler (2009). The relation between the Ostrogorski paradox and the Condorcet paradox has been investigated in several papers (Kelly, 1989; Rae and Daudt, 1976), while a comparison with the discursive dilemma was carried out by Pigozzi (2005).

3.3.1 The Ostrogorski Paradox in Binary Aggregation

In this section, we provide a binary aggregation setting that represents the Ostrogorski paradox as a failure of collective rationality with respect to a suitable integrity constraint.

Let $\{E, S, F\}$ be the set of issues at stake, and let the set of issues $\mathcal{I}_O = \{E, S, F, A\}$ consist of the same issues plus an extra issue A to encode the support for the first party (MP).[4] A binary ballot over these issues represents the

[4]We hereby propose a model that can be used for instances of the Ostrogorski paradox concerning at most two parties. In case the number of parties is bigger than two, the framework can be extended adding one extra issue for every party.

individual view on the three issues E, S and F: if, for instance, $b_E = 1$, then the individual supports the first party MP on the first issue E. Moreover, it also represents the overall support for party MP (in case issue A is accepted) or PP (in case A is rejected). In the Ostrogorski paradox, an individual votes for a party if and only if she agrees with that party on a majority of the issues. This rule can be represented as a rationality assumption by means of the following integrity constraint IC_O:

$$p_A \leftrightarrow [(p_E \wedge p_S) \vee (p_E \wedge p_F) \vee (p_S \wedge p_F)]$$

An instance of the Ostrogorski paradox can therefore be represented by the profile $\boldsymbol{B}$ described in Table 3.6.

	E	S	F	A
Voter 1	1	0	0	0
Voter 2	0	0	1	0
Voter 3	1	0	1	1
Maj	1	0	1	0

Table 3.6: The Ostrogorski paradox in binary aggregation.

Each individual in Table 3.6 accepts issue A if and only if she accepts a majority of the other issues. However, the outcome of the majority rule is a rejection of issue A, even if a majority of the issues gets accepted by the same rule. Therefore, the triple $(Maj, \boldsymbol{B}, \text{IC}_O)$ constitutes a paradox by Definition 2.1.9.

Using this formalism we can easily devise a stronger version of the Ostrogorski paradox, in which the winning party disagrees with a majority of the individuals on *all* issues. Such a profile is described in Table 3.7.

	E	S	F	A
Voter 1	1	0	0	0
Voter 2	0	1	0	0
Voter 3	0	0	1	0
Voter 4	1	1	1	1
Voter 5	1	1	1	1
Maj	1	1	1	0

Table 3.7: Strong version of the Ostrogorski paradox.

3.4 The Common Structure of Paradoxical Integrity Constraints

Let us make an important remark concerning the syntactic structure of the integrity constraints that formalise the three paradoxes we have presented so far. The first formula, encoding the transitivity of a preference relation in the Condorcet paradox, is the implication $p_{ab} \wedge p_{bc} \rightarrow p_{ac}$. This formula is equivalent to $\neg p_{ab} \vee \neg p_{bc} \vee p_{ac}$, which is a clause of size 3, i.e., it is a disjunction of three different literals. The second formula, presented in Section 3.2 to represent the discursive dilemma, is also equivalent to a clause of size 3, namely $\neg p_{\alpha} \vee \neg p_{\beta} \vee \neg p_{\neg(\alpha \wedge \beta)}$. The last formula, which formalises the majoritarian constraint underlying the Ostrogorski paradox, is equivalent to the following conjunction of clauses of size 3:

$$(p_A \vee \neg p_E \vee \neg p_F) \wedge (p_A \vee \neg p_E \vee \neg p_S) \wedge (p_A \vee \neg p_S \vee \neg p_F) \wedge$$
$$\wedge (\neg p_A \vee p_E \vee p_F) \wedge (\neg p_A \vee p_E \vee p_S) \wedge (\neg p_A \vee p_S \vee p_F)$$

The observation that the integrity constraints formalising the most classical paradoxes in aggregation theory all feature a clause of size at least 3 is not a coincidence. In Section 4.4.2 we will formalise this observation with a theorem that characterises the class of integrity constraints that are lifted by the majority rule as those and only those that can be expressed as a conjunction of clauses of maximal size 2 (see Theorem 4.4.8).[5]

3.5 Further Paradoxes on Multiple Issues

In this section we describe two further paradoxes that can be analysed using our framework of binary aggregation with integrity constraints: the paradox of divided government and the paradox of multiple elections. Both situations concern a paradoxical outcome obtained by using the majority rule on an aggregation problem defined on multiple issues. The first paradox can be seen as an instance of a more general behaviour described by the second paradox.

3.5.1 The Paradox of Divided Government

The paradox of divided government is a failure of collective rationality that was pointed out for the first time by Brams et al. (1993). Here we follow the presentation of Nurmi (1997).

[5] This observation is strongly related to a result proven by Nehring and Puppe (2007) in the framework of judgment aggregation, which characterises the set of paradoxical agendas for the majority rule as those agendas containing a minimal inconsistent subset of size at least 3. See also our previous work (Grandi, 2012).

The paradox of divided government. Suppose that 13 voters (equivalently, groups of voters) can choose for Democratic (D) or Republican (R) candidate for the following three offices: House of Representatives (H), Senate (S) and the governor (G). It is a common assumption that in case the House of Representatives gets a Republican candidate, then at least one of the remaining offices should go to Republicans as well. Consider now the profile in Table 3.8.

	H	S	G
Voters 1-3	D	D	D
Voter 4	D	D	R
Voter 5	D	R	D
Voter 6	D	R	R
Voters 7-9	R	D	R
Voters 10-12	R	R	D
Voter 13	R	R	R
Maj	R	D	D

Table 3.8: The paradox of divided government.

As shown in Table 3.8, it is exactly the combination that had to be avoided (i.e., RDD) that is elected, even if no individual voted for it.

This paradox can be easily seen as a failure of collective rationality: it is sufficient to replace the letters D and R with 0 and 1, and to formulate the integrity constraint as $\neg(p_H \wedge \neg p_S \wedge \neg p_G)$. The binary ballot $(1, 0, 0)$ is therefore ruled out as irrational, encoding the combination (R,D,D) that needs to be avoided.

This type of paradox can be observed in cases like the elections of a committee, such as in our Example 2.1.7. Even if it is recognised by every individual that a certain committee structure is unfeasible (i.e., it will not work well together), this may be the outcome of aggregation if the majority rule is being used.

In view of our discussion in Section 2.1.5, we may consider the constraint underlying the paradox of divided government as a feasibility constraint, rather than a constraint of rationality. Under such an interpretation this situation would cease to be paradoxical, while still showing the failure of the majority rule to output a feasible outcome.

3.5.2 The Paradox of Multiple Elections

Whilst the Ostrogorski paradox was devised to stage an attack against representative systems of collective choice based on parties, the paradox of multiple elections (MEP) is based on the observation that when voting directly on multiple issues, a combination that was not supported nor liked by any of the voters can be the winner of the election (Brams et al., 1998; Lacy and Niou, 2000). While the original model takes into account the full preferences of individuals over combinations of issues, if we focus on only those ballots that are submitted by the individuals, then an instance of the MEP can be represented as a paradox of collective rationality. Let us consider a simple example described in Table 3.9.

> **Multiple election paradox**. Suppose three voters need to take a decision over three binary issues A, B and C. Their ballots are described in Table 3.9.

	A	B	C
Voter 1	1	0	1
Voter 2	0	1	1
Voter 3	1	1	0
Maj	1	1	1

Table 3.9: The multiple election paradox (MEP).

> The outcome of the majority rule in Table 3.9 is the acceptance of all three issues, even if this combination was not voted for by any of the individuals.

While there seems to be no integrity constraint directly causing this paradox, we may represent the profile in Table 3.9 as a situation in which the three individual ballots are bound by a budget constraint $\neg(p_A \wedge p_B \wedge p_C)$ (like in our Example 2.1.10). Even if all individuals are giving acceptance to two issues each, the result of the aggregation is the unfeasible acceptance of all three issues.

As can be deduced from our previous discussion, every instance of the MEP gives rise to several instances of a binary aggregation paradox for Definition 2.1.9. To see this, it is sufficient to find an integrity constraint that is satisfied by all individuals and not by the outcome of the aggregation.[6] On the other hand, every instance of Definition 2.1.9 in binary aggregation represents an instance of the MEP, as the irrational outcome cannot have been voted for by any of the individuals. In Section 7.2 we define an interesting aggregation procedure, called the average voter rule, which avoids both the MEP and any other failure of collective rationality.

[6]Such a formula always exists. Consider for instance the disjunction of the formulas specifying each of the individual ballots. This integrity constraint forces the result of the aggregation to be equal to one of the individual ballots on the given profile, thus generating a binary aggregation paradox from a MEP.

The multiple election paradox gives rise to a different problem than that of consistency, to which this dissertation is dedicated, as it is not directly linked to an integrity constraint established in advance. The problem formalised by the MEP is rather the *compatibility* of the outcome of aggregation with the individual ballots. Individuals in such a situation may be forced to adhere to a collective choice which, despite it being rational, they do not perceive as representing their views (Grandi and Pigozzi, 2012).

In their paper, Brams et al. (1998) provide many versions of the multiple election paradox, varying the number of issues and the presence of ties. Lacy and Niou (2000) enrich the model by assuming that individuals have a preference order over combinations of issues and submit just their top candidate for the election. They present situations in which, e.g., the winning combination is a Condorcet loser (i.e., it loses in pairwise comparison with all other combinations). Some answers to the problem raised by the MEP have already been proposed in the literature on Artificial Intelligence. For instance, a sequence of papers have studied the problem of devising sequential elections to avoid the MEP in case the preferences of the individuals over combinations of multiple issues are expressed in a suitable preference representation language (Lang, 2007; Lang and Xia, 2009; Xia et al., 2011; Conitzer and Xia, 2012).

3.6 Conclusions

The first lesson that can be drawn from this chapter dedicated to paradoxes of aggregation is that the majority rule is not a good aggregation procedure to be employed when dealing with collective choices over multiple issues. This fact stands out as a counterpart to May's Theorem (1952), which proves that the majority rule is the only aggregation rule *for a single binary issue* that satisfies a set of highly desirable conditions. The sequence of paradoxes we have analysed in this chapter shows that this is not the case when multiple issues are involved. While this fact may not add anything substantially new to the existing literature, the wide variety of paradoxical situations encountered in this chapter stresses even further the negative features of the majority rule for multi-issue domains.

A second conclusion is that most paradoxes of Social Choice Theory share a common structure, and that this structure is formalised by our Definition 2.1.9, which stands out as a truly general definition of paradox in aggregation theory. Moreover, by analysing the integrity constraints that underlie some of the most classical paradoxes, we were able to identify a common syntactic feature of paradoxical constraints (cf. Section 3.4). This observation is the starting point of the following chapter, in which we build a systematic theory of collective rationality depending on the syntactic properties of integrity constraints.

The paradoxical situations presented in this chapter constitute a fragment of the problems that can be encountered in the formalisation of collective choice problems. First, all the paradoxes we presented feature the majority rule as the procedure used for aggregation. Paradoxical situations can be encountered in the study of many other aggregation procedures, e.g., in the case of the Borda paradox (McLean and Urken, 1995). Second, all paradoxes concern problems of aggregation in which the input given by the individuals is of the same form as the output. Paradoxical situations concerning voting procedures (Nurmi, 1999), which take as input a set of preferences and output a set of winning candidates, are therefore not included in our analysis.

We close this section with an observation regarding the interpretation of some of the paradoxes presented in this chapter. We have already remarked how some of these examples have been employed in the literature to show weaknesses and advantages of either the direct approach to democratic choice (represented by issue-by-issue aggregation) or the representative one. The last two paradoxes especially (the paradox of divided government and the MEP) seem to suggest that direct decisions over multiple issues should be avoided, at least when issues are not completely independent from one another. In our view, elections over multi-issue domains cannot be escaped: not only do they represent a model for the aggregation of more complex objects like preferences and judgments, as seen in Section 3.1 and 3.2, but they also stand out as one of the biggest challenges to the design of more complex automated systems for collective decision making.

A crucial problem in the modelling of real-world situations of collective choice is that of identifying the set of issues that best represent a given domain of aggregation, and devising an integrity constraint that models correctly the correlations between those issues. This problem obviously represents a serious obstacle to a mechanism designer, and is moreover open to manipulation. However, we believe that structuring collective decision problems with more detailed models *before* the aggregation takes place, e.g., by discovering a shared order of preferential dependencies between issues (Lang and Xia, 2009; Airiau et al., 2011), facilitates the definition of collective choice procedures on complex domains without having to elicit the full preferences of individuals. Such models can be employed in the design and the implementation of automated decision systems, in which a safe aggregation, i.e., one that avoids paradoxical situations, is of the utmost necessity. One of the main aims of this dissertation is exactly to provide tools allowing to stage direct elections on correlated issues in a safe way.

Chapter 4

Lifting Individual Rationality

Individual agents may be considered rational in many different ways, and for most cases, as exposed in Chapter 3, it is possible to devise paradoxical situations leading to an irrational collective outcome. The purpose of this chapter is to develop a theoretical analysis of the relation between axiomatic properties and collective rationality with respect to given integrity constraints, generalising what observed in the previous chapter to a full-fledged theory of collective rationality for aggregation procedures in binary aggregation.

In Section 4.1 we introduce two definition schemas for classes of aggregation procedures: the first in terms of collective rationality with respect to a given language for integrity constraints, and the second by using classical axiomatic requirements. The relation between these two definitions is studied in detail in Section 4.2. For several fragments of the propositional language we provide necessary and sufficient axiomatic conditions for an aggregation procedure to be collectively rational with respect to all integrity constraints in the given fragment. The analysis is continued in Section 4.3, this time focusing on axioms rather than on languages for integrity constraints. For several axiomatic requirements we provide negative results that rule out possible characterisations of such properties in terms of collective rationality. In Section 4.4 we concentrate on the class of quota rules, providing precise bounds on quotas to ensure collective rationality with respect to languages of clauses. We also characterise the set of integrity constraints that are lifted by the majority rule, thus formalising our observations from Chapter 3. Section 4.5 discusses related work and concludes.

4.1 Classes of Aggregation Procedures

Recall that a binary aggregation problem is given by a set of agents $\mathcal{N}$ having to take a decision on which combination of binary issues $\mathcal{I}$ to choose. Depending on the situation at hand, a subset of such combinations is designated as the set of rational choices, and is specified by means of a propositional formula in the language $\mathcal{L}_{PS}$ associated to $\mathcal{I}$ (cf. Section 2.1).

Let therefore $\mathcal{I}$ be a finite set of issues and let $\mathcal{L}_{PS}$ be the propositional language associated with it. We call any subset $\mathcal{L}$ of $\mathcal{L}_{PS}$ a *language*. Examples include the set of atoms PS, or the set of formulas of a given size, as well as more classical fragments obtained by restricting the set of connectives that can be employed in the construction of formulas, like the set of clauses, obtained from the set of literals using only disjunctions. In Section 2.1.4 we called an aggregation procedure collectively rational with respect to a formula IC $\in \mathcal{L}_{PS}$ if the outcome of aggregation satisfies the same integrity constraint IC as the individuals on every rational profile. We now extend this definition to collectively rational procedures with respect to a given language $\mathcal{L}$:

Definition 4.1.1. Given a language $\mathcal{L} \subseteq \mathcal{L}_{PS}$, define $\mathcal{CR}[\mathcal{L}]$ to be the class of aggregation procedures that lift all integrity constraints IC $\in \mathcal{L}$:

$$\mathcal{CR}[\mathcal{L}] \;:=\; \{F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D} \mid \mathcal{N} \text{ is finite and } F \text{ is CR for all IC} \in \mathcal{L}\}.$$

Note that in this definition we do not fix the number of individuals, making $\mathcal{I}$ the only parameter that is fixed in advance. This choice is arguably a natural one, as a decision *problem* is usually defined before specifying the number of individuals that are going to take part in the decision *process*. However, its appeal does not only reside in its practical use; rather is it a mathematical assumption that allows us to gain clarity in some of the results that follows. Many of our results, e.g. Theorems 4.2.1 and Corollary 4.3.3, still hold if we fix the number of individuals in Definition 4.1.1, as shown in our previous work (Grandi and Endriss, 2010).

The next step is to introduce notation for defining classes of aggregation procedures in terms of classical axioms as the ones we listed in Section 2.2. Recall that an axiom may be satisfied on a subdomain of interest $\mathcal{X} \subseteq \mathcal{D}$, but not on the full domain $\mathcal{D}^{\mathcal{N}}$ (see the observation on page 23). Here, we are interested in domains defined by means of integrity constraints (i.e., propositional formulas), as this is interpreted as the domain of rational ballots. We therefore need a notation to identify procedures that satisfy an axiom on the subdomain $\text{Mod(IC)}^{\mathcal{N}}$ induced by a given integrity constraint IC.

Let $F_{\restriction \text{Mod(IC)}^{\mathcal{N}}}$ denote the restriction of the aggregation procedure F to the subdomain of rational ballots $\text{Mod(IC)}^{\mathcal{N}}$. We give the following definition:

Definition 4.1.2. An aggregation procedure F satisfies a set of axioms AX with respect to a language $\mathcal{L} \subseteq \mathcal{L}_{PS}$, if for all constraints IC $\in \mathcal{L}$ the restriction $F_{\restriction \text{Mod(IC)}^{\mathcal{N}}}$ satisfies the axioms in AX. This defines the following class:

$$\mathcal{F}_{\mathcal{L}}[\text{AX}] \;:=\; \{F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D} \mid \mathcal{N} \text{ is finite and } F_{\restriction \text{Mod(IC)}^{\mathcal{N}}} \text{ sat. AX for all IC} \in \mathcal{L}\}$$

In particular, $\mathcal{F} := \{F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D} \mid \mathcal{N} \text{ is finite}\}$ is the class of all aggregation procedures for a given $\mathcal{I}$. In the sequel we shall omit mentioning explicitly that $\mathcal{N}$ is finite, keeping it as a general underlying assumption.

4.1.1 Collective Rationality and Languages

In this section we study the behaviour of the classes defined in the previous section with respect to set-theoretic and logical operations performed on the languages and on the axioms. In particular, we give a definition of languages for integrity constraints that is specific to the study of collectively rational procedures.

Let $\mathcal{L}$ be a language. Define $\mathcal{L}^\wedge$ to be the closure of $\mathcal{L}$ under conjunction, i.e., the set of finite conjunctions of formulas in $\mathcal{L}$. We now prove that the class of collectively rational procedures is invariant under closing the language under conjunction, i.e., that the set of collectively rational procedures for $\mathcal{L}$ and for $\mathcal{L}^\wedge$ coincide:

Lemma 4.1.3. *$\mathcal{CR}[\mathcal{L}^\wedge] = \mathcal{CR}[\mathcal{L}]$ for all $\mathcal{L} \subseteq \mathcal{L}_{PS}$.*

Proof. $\mathcal{CR}[\mathcal{L}^\wedge]$ is clearly included in $\mathcal{CR}[\mathcal{L}]$, since $\mathcal{L} \subseteq \mathcal{L}^\wedge$. It remains to be shown that, if an aggregation procedure F lifts every constraint in $\mathcal{L}$, then it lifts any conjunction of formulas in $\mathcal{L}$. This fact is rather straightforward; however, we now prove it in detail to get acquainted with the definition of $\mathcal{CR}[\mathcal{L}]$. Let $\bigwedge_k \mathrm{IC}_k$ with $\mathrm{IC}_k \in \mathcal{L}$ be a conjunction of formulas in $\mathcal{L}$, and let $\boldsymbol{B} \in \mathrm{Mod}(\bigwedge_k \mathrm{IC}_k)^\mathcal{N}$ be a profile satisfying this integrity constraint. Note that $\mathrm{Mod}(\bigwedge_k \mathrm{IC}_k) = \bigcap_k \mathrm{Mod}(\mathrm{IC}_k)$, thus $\boldsymbol{B} \in \mathrm{Mod}(\mathrm{IC}_k)^\mathcal{N}$ for every k. Now suppose that $F \in \mathcal{CR}[\mathcal{L}]$, then when we apply F to profile $\boldsymbol{B}$ we have that $F(\boldsymbol{B}) \in \mathrm{Mod}(\mathrm{IC}_k)$ for every k by collective rationality of F. This in turn implies $F(\boldsymbol{B}) \in \mathrm{Mod}(\bigwedge_k \mathrm{IC}_k)$, thus proving that F is CR with respect to $\bigwedge_k \mathrm{IC}_k$. □

This lemma entails that different languages for integrity constraints can define the same class of CR procedures. For instance, we have that the language of cubes (conjunctions of literals) generates the same class as the language of literals, i.e., $\mathcal{CR}[\mathit{cubes}] = \mathcal{CR}[\mathit{literals}]$, since the former is obtained from the latter by closing it under conjunction. A more interesting fact is that procedures that are CR with respect to clauses (disjunctions of literals) are CR with respect to *any* integrity constraint in $\mathcal{L}_{PS}$, i.e., $\mathcal{CR}[\mathit{clauses}] = \mathcal{CR}[\mathcal{L}_{PS}]$. This holds because every propositional formula is equivalent to a formula in conjunctive normal form (CNF), where it is expressed precisely as a conjunction of clauses.

One last remark about this lemma. In Section 2.1 we defined collective rationality with respect to a single formula, rather than with respect to a set of integrity constraints. Lemma 4.1.3 provides a formal underpinning for this choice: an aggregation procedure is CR with respect to a set of formulas if and only if it is CR with respect to a single formula given by the conjunction of all integrity constrains in the set.

We have just proven that the class $\mathcal{CR}[\mathcal{L}]$ is invariant under closing the language under conjunction. Another such property is the closure under logical equivalence.[1] Recall that two formulas are logically equivalent when they share

[1]It is important to stress the fact that we consider logical equivalence *inside* the language $\mathcal{L}_{PS}$, not allowing the use of additional propositional variables.

the same set of models. Let us indicate with $\mathcal{L}^{\equiv}$ the set of formulas in $\mathcal{L}_{PS}$ that are equivalent to a formula in $\mathcal{L}$. We have the following lemma:

Lemma 4.1.4. $\mathcal{CR}[\mathcal{L}^{\equiv}] = \mathcal{CR}[\mathcal{L}]$ *for all* $\mathcal{L} \subseteq \mathcal{L}_{PS}$.

The proof of the lemma is straightforward from our definitions. It is sufficient to observe that an equivalent formulation of our definition of collective rationality can be given by substituting formulas with the set of rational ballots given by their models. Two formulas that are logically equivalent have the same set of models, giving rise to the same requirement of collective rationality.[2]

Bringing together the results of Lemma 4.1.3 and of Lemma 4.1.4, we can now give the following definition:

Definition 4.1.5. A *language for integrity constraints* $\mathcal{L}$ is a subset of $\mathcal{L}_{PS}$ that is closed under conjunction and logical equivalence.

In the following sections we often characterise languages by means of syntactic properties, e.g., *cubes* or *clauses*, denoting the language for integrity constraints generated by these formulas, i.e., the subset of $\mathcal{L}_{PS}$ obtained by closing the original language under conjunction and logical equivalence. For instance, the language of 2-*clauses* (i.e., disjunctions of size at most two) indicates the language of formulas that are equivalent to a conjunction of clauses of size at most two[3]. The language of *literals* and that of *cubes* coincide, as well as the language of *clauses* and the full language $\mathcal{L}_{PS}$, as we have previously remarked.

Tautologies and contradictions play a special role in languages for integrity constraints. First, observe that if a language $\mathcal{L}$ includes a tautology (or a contradiction, respectively), then by closure under logical equivalence $\mathcal{L}$ contains *all* tautologies (all contradictions, respectively). Thus, we indicate with $\top \in \mathcal{L}$ the fact that $\mathcal{L}$ contains all tautologies, and with $\bot \in \mathcal{L}$ the fact that $\mathcal{L}$ contains all contradictions. Second, not all languages for integrity constraints include both tautologies and contradictions, or either of them. For instance, the language of literals includes the contradiction $p \wedge \neg p$ but it does not contain any tautology. On the other hand, the language of positive clauses, composed by clauses in which all literals occur positively, does not include neither tautologies nor contradictions.

Nevertheless, it is easy to see that collective rationality with respect to tautologies and contradictions corresponds to a vacuous requirement: In the first case, the outcome of a procedure will always satisfy a tautology, and in the second case the set of rational ballots is empty. These remarks constitute a proof of the following lemma.

[2]This is the standard approach in the literature on binary aggregation (cf. Dokow and Holzman, 2010a). Our choice of using formulas rather than sets is motivated by the compactness of this representation (see the observation at page 27) and by the possibility of using syntax to classify rationality assumptions.

[3]The language of *2-clauses* can be equivalently defined by closing the set of 2-CNF under logical equivalence.

Lemma 4.1.6. $\mathcal{CR}[\mathcal{L} \cup \{\top\}] = \mathcal{CR}[\mathcal{L} \cup \{\bot\}] = \mathcal{CR}[\mathcal{L}]$ *for all* $\mathcal{L} \subseteq \mathcal{L}_{PS}$.

We now move to answering the question of whether the operations that we have included in Definition 4.1.5 are all the operations that we can perform on $\mathcal{L}$ leaving the set $\mathcal{CR}[\mathcal{L}]$ invariant. The following result provides a positive answer to this question, provided that a language include tautologies and contradictions:

Lemma 4.1.7. *Given two languages for integrity constraints* $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *containing* $\top$ *and* $\bot$, *if it is the case that* $\mathcal{L}_1 \neq \mathcal{L}_2$, *then* $\mathcal{CR}[\mathcal{L}_1] \neq \mathcal{CR}[\mathcal{L}_2]$.

Proof. As the two languages both contain tautologies and contradictions, they must differ on a contingent formula φ. Without loss of generality we can consider a formula $\varphi \in \mathcal{L}_2$ such that $\varphi \notin \mathcal{L}_1$. We want to prove that there exists an aggregation procedure $F \in \mathcal{CR}[\mathcal{L}_1]$ that is not CR with respect to φ. This in turn implies that F is not in $\mathcal{CR}[\mathcal{L}_2]$, and that the two classes $\mathcal{CR}[\mathcal{L}_1]$ and $\mathcal{CR}[\mathcal{L}_2]$ are different.

Let $|\mathcal{N}| = n$ where $n = |\text{Mod}(\varphi)|$ and let F be a procedure in $\mathcal{CR}[\mathcal{L}_1]$ defined for $\mathcal{N}$.[4] Observe that $n \neq 0, 2^{|\mathcal{I}|}$, since φ is a contingent formula. We claim that it is possible to modify the behaviour of F on a single profile $\boldsymbol{B}$ in order to create another procedure F' that is still CR with respect to $\mathcal{L}_1$ but sends the profile $\boldsymbol{B}$ of φ-rational ballots to an outcome that does not satisfy φ. To do so it is sufficient to find a profile $\boldsymbol{B} = (B_1, \ldots, B_n)$ of models of φ and a ballot B_c outside $\text{Mod}(\varphi)$ such that whenever each of $B_1, \ldots, B_n$ satisfy the same formula $\psi \in \mathcal{L}_1$ then also $B_c \models \psi$. If we can find such a $\boldsymbol{B}$ and B_c, then by setting $F'(\boldsymbol{B}) = B_c$ and $F'(\boldsymbol{B}') = F(\boldsymbol{B}')$ for all remaining $\boldsymbol{B}' \neq \boldsymbol{B}$ we obtain an aggregator procedure that is in $\mathcal{CR}[\mathcal{L}_1]$ but not in $\mathcal{CR}[\mathcal{L}_2]$.

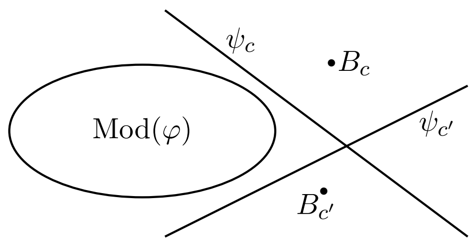

Figure 4.1: ψ_c separates $\boldsymbol{B}_c$ from $\text{Mod}(\varphi)$.

Suppose for the sake of contradiction that such a profile does not exist, i.e., that for every choice of n ballots in $\text{Mod}(\varphi)$ and ballot B_c outside $\text{Mod}(\varphi)$, there is a formula $\psi \in \mathcal{L}_1$ that separates them: for all i we have that $B_i \models \psi$ but $B_c \not\models \psi$. Note that $|\text{Mod}(\varphi)| = n$, as well as the size of the profile $\boldsymbol{B}$ we are looking for.

[4] In Section 4.2 we will prove that $\mathcal{CR}[\mathcal{L}]$ can never be empty for any language $\mathcal{L}$ and any set of agents $\mathcal{N}$ (see Theorem 4.2.8), i.e., that it is always possible to find an F meeting our requirement. For the sake of this proof it is sufficient to consider a dictatorship of the first individual, i.e., a procedure that outputs the first coordinate of a profile.

This entails that we can construct a profile $\boldsymbol{B}_\varphi$ which contains all distinct models of φ, and that for every $B_c \not\models \varphi$ there is a formula in $\mathcal{L}_1$, that we shall call ψ_c, that separates the set $\text{Mod}(\varphi)$ from B_c.

We have assumed that $\varphi \notin \mathcal{L}_1$, i.e., φ is not equivalent to a conjunction of formulas in $\mathcal{L}_1$. Let us now consider the conjunction $\bigwedge \psi_c$ for $B_c \not\models \varphi$. We claim that $\varphi \equiv \bigwedge \psi_c$, in contradiction with our assumption, since all ψ_c are in $\mathcal{L}_1$ and $\mathcal{L}_1$ is closed under conjunction. By construction we know that $\text{Mod}(\varphi) \subseteq \text{Mod}(\bigwedge \psi_c)$, as all models of φ are individual ballots in the profile $\boldsymbol{B}_\varphi$. We need to prove the other inclusion. Assume for the sake of contradiction that there exists a B_* in $\text{Mod}(\bigwedge \psi_c) \setminus \text{Mod}(\varphi)$. By construction, there exists a formula $\psi_* \in \mathcal{L}_1$ that separates B_* from $\text{Mod}(\varphi)$, and this formula is included in $\bigwedge \psi_c$. But by construction $B_* \not\models \psi_*$, therefore it cannot be included in $\text{Mod}(\bigwedge \psi_c)$.

By reaching this contradiction we have concluded the proof: it is possible to modify F on the profile $\boldsymbol{B}$ of all models of φ to output a ballot that is not a model of φ but respects all integrity constraints in $\mathcal{L}_1$. That is, F' so defined is in $\mathcal{CR}[\mathcal{L}_1]$ but not in $\mathcal{CR}[\varphi]$. □

A crucial assumption used in the previous proof is that $\mathcal{CR}[\mathcal{L}]$ contains procedures defined for arbitrarily large sets of individuals.[5] The construction we used in the previous proof would not be possible if we had fixed in advance the number of individuals. An example can be obtained in the special case of $|\mathcal{N}| = 1$, considering the language $\mathcal{L}$ obtained from $\{p, \neg p, q, \neg q\}$. This language cannot express the formula $\varphi = p \leftrightarrow q$, even if the class $\mathcal{CR}[\mathcal{L} \cup \{\varphi\}] = \mathcal{CR}[\mathcal{L}]$, as can be seen with some easy calculation. Similar counterexamples involving bigger sets of individuals are hard to obtain.

We conclude this section by establishing some easy properties of $\mathcal{CR}[\mathcal{L}]$ and of $\mathcal{F}_\mathcal{L}[\text{AX}]$ that shall be useful in the next sections.

Lemma 4.1.8. *The following facts hold:*

(i) *If* $\mathcal{L}_1 \subseteq \mathcal{L}_2$, *then* $\mathcal{CR}[\mathcal{L}_1] \supseteq \mathcal{CR}[\mathcal{L}_2]$;

(ii) $\mathcal{CR}[\mathcal{L}_1 \cup \mathcal{L}_2] = \mathcal{CR}[\mathcal{L}_1] \cap \mathcal{CR}[\mathcal{L}_2]$ *for all* $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{L}_{PS}$;

(iii) $\mathcal{CR}[\mathcal{L}_1 \cap \mathcal{L}_2] = \mathcal{CR}[\mathcal{L}_1] \cup \mathcal{CR}[\mathcal{L}_2]$ *for all* $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{L}_{PS}$.

The proof is straightforward from our definitions. Similar properties can be proven for classes of procedures defined in terms of axioms. We write $\mathcal{F}[\text{AX}]$ as a shorthand for $\mathcal{F}_{\{\top\}}[\text{AX}]$, the class of procedures that satisfy the axioms in AX over the full domain $\mathcal{D}$. It is easy to see that the following lemma holds:

Lemma 4.1.9. *The following facts hold:*

(i) *if* $\mathcal{L}_1 \subseteq \mathcal{L}_2$ *then* $\mathcal{F}_{\mathcal{L}_1}[\text{AX}] \supseteq \mathcal{F}_{\mathcal{L}_2}[\text{AX}]$;

(ii) *in particular, if* $\top \in \mathcal{L}$, *then* $\mathcal{F}[\text{AX}] \supseteq \mathcal{F}_\mathcal{L}[\text{AX}]$;

(iii) $\mathcal{F}_\mathcal{L}[\text{AX}_1, \text{AX}_2] = \mathcal{F}_\mathcal{L}[\text{AX}_1] \cap \mathcal{F}_\mathcal{L}[\text{AX}_2]$.

[5] This is the only result in this chapter the proof of which hinges on this assumption.

Observe that for most axioms an additional fact holds: if the axiomatic property AX is satisfied on the full domain $\mathcal{D}$, then AX is also satisfied on every subdomain of $\mathcal{D}$. This is true in particular for all the axioms we considered.[6] Thus, for most axioms AX it holds that $\mathcal{F}[\text{AX}] \subseteq \mathcal{F}_{\mathcal{L}}[\text{AX}]$ for all $\mathcal{L} \subseteq \mathcal{L}_{PS}$. Moreover, by (*ii*) of Lemma 4.1.9, if $\top \in \mathcal{L}$ then $\mathcal{F}[\text{AX}] = \mathcal{F}_{\mathcal{L}}[\text{AX}]$ for all $\mathcal{L} \subseteq \mathcal{L}_{PS}$.

4.1.2 From Classes of Aggregation Procedures to Integrity Constraints and Back

In the first part of Section 4.1 we have associated with any language for integrity constraints $\mathcal{L}$ a class of aggregation procedures $\mathcal{CR}[\mathcal{L}]$ that are collectively rational with respect to all formulas in $\mathcal{L}$. Once a set of issues $\mathcal{I}$ is fixed, $\mathcal{CR}[-]$ can therefore be viewed as an operator from the set of languages for integrity constraints (i.e., subsets of $\mathcal{L}_{PS}$ closed under conjunction and logical equivalence) to subsets of the class $\mathcal{F}$ of all aggregation procedures for $\mathcal{I}$. In this section we introduce an inverse operation, that we shall call $\mathcal{LF}[-]$, which, given a class of procedures, outputs the set of integrity constraints that are *lifted* by all procedures in that class. As we will see, $\mathcal{LF}$ is the left inverse of $\mathcal{CR}$, but on the other side the two operators do not commute.

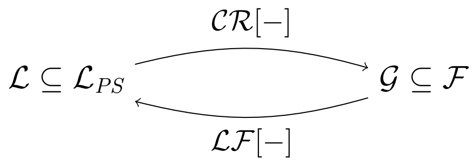

Figure 4.2: The operators $\mathcal{CR}[-]$ and $\mathcal{LF}[-]$.

Definition 4.1.10. Given a class of aggregation procedures $\mathcal{G} \subseteq \mathcal{F}$, let $\mathcal{LF}[\mathcal{G}]$ be the set of integrity constraints that are lifted by all $F \in \mathcal{G}$:

$$\mathcal{LF}[\mathcal{G}] = \{\varphi \in \mathcal{L}_{PS} \mid F \text{ is CR with respect to } \varphi \text{ for all } F \in \mathcal{G}\}$$

$\mathcal{LF}[\mathcal{G}]$ is the intersection of all $\mathcal{LF}[\{F\}]$ for $F \in \mathcal{G}$. We now prove the following:

Proposition 4.1.11. *Let $\mathcal{I}$ be a set of issues, $\mathcal{L}$ a language for integrity constraints containing $\top$ and $\bot$, and $\mathcal{G} \subseteq \mathcal{F}$ a class of aggregation procedures on $\mathcal{I}$. Then the following facts are true:*
(*i*) $\mathcal{LF}[\mathcal{CR}[\mathcal{L}]] = \mathcal{L}$
(*ii*) $\mathcal{CR}[\mathcal{LF}[\mathcal{G}]] \supseteq \mathcal{G}$ *and this inclusion is strict for some classes.*

[6]See the observation at the end of Section 4.3.2 for an example of an axiomatic property that can be satisfied on the full domain but falsified on a proper subset of it.

Proof. (*i*) We start by proving that $\mathcal{LF}$ is a left inverse of $\mathcal{CR}$. A direct consequence of our definitions is that $\mathcal{L} \subseteq \mathcal{LF}[\mathcal{CR}[\mathcal{L}]]$, and we now prove the other inclusion. We want to show that if an integrity constraint φ is lifted by all procedures that are CR with respect to $\mathcal{L}$, then φ belongs to $\mathcal{L}$. This is a straightforward consequence of Lemma 4.1.7. Assume for the sake of contradiction that $\varphi \notin \mathcal{L}$ (and thus that φ is neither a tautology nor a contradiction). By Lemma 4.1.7, there exists a procedure F which is collectively rational for $\mathcal{L}$ but not for φ, against our assumption that all procedures in $\mathcal{CR}[\mathcal{L}]$ are CR with respect to φ. Therefore φ is in $\mathcal{L}$.

(*ii*) It is straightforward from our definitions that $\mathcal{CR}[\mathcal{LF}[\mathcal{G}]] \supseteq \mathcal{G}$. Recall that a dictatorship is a procedure that copies the ballot of a given individual in every profile. It can be easily observed that such a procedure is collectively rational for every integrity constraint, and in Section 4.2 this fact will be given a formal proof (see Theorem 4.2.8). Let us therefore consider a class of procedures $\mathcal{G}$ not containing any dictatorship. In view of our previous observation we know that all dictatorships are contained in $\mathcal{CR}[\mathcal{LF}[\mathcal{G}]]$, as they are collectively rational for any integrity constraint. As we assumed that $\mathcal{G}$ does not contain any dictatorship, we infer that $\mathcal{CR}[\mathcal{LF}[\mathcal{G}]] \supsetneq \mathcal{G}$. □

In the following section we will prove several characterisation results for the class $\mathcal{CR}[\mathcal{L}]$ in terms of classical axioms. These results define unambigously the class of integrity constraints lifted by the class of procedures under consideration. To see this, suppose we can prove that $\mathcal{CR}[\mathcal{L}] = \mathcal{F}_{\mathcal{L}}[\text{AX}]$ for given $\mathcal{L}$ and AX; then, by part (*i*) of Proposition 4.1.11, we have that $\mathcal{LF}[\mathcal{F}_{\mathcal{L}}[\text{AX}]] = \mathcal{L}$.[7] On the other hand, a characterisation of the class $\mathcal{LF}[\mathcal{G}]$ cannot be turned into a characterisation of $\mathcal{CR}[\mathcal{L}]$. By part (*ii*) of Proposition 4.1.11 we can infer that if $\mathcal{LF}[\mathcal{G}] = \mathcal{L}$ then $\mathcal{G} \subseteq \mathcal{CR}[\mathcal{L}]$, but there is no guarantee that $\mathcal{G}$ will be equal to $\mathcal{CR}[\mathcal{L}]$.

Let us spend one last paragraph concerning this asymmetry between the two definitions. As we will see in the next section, the two operators commute for every class of procedures defined in terms of classical axiomatic properties like those described in Section 2.2 (provided that they contain the set of generalised dictatorships, see Definition 4.2.7). This is because classical axioms define sufficiently big classes of procedures. The conditions we have imposed on languages for integrity constraints guarantee that languages are closed under the "right" set of operations when talking about collective rationality. This is not so for classes of aggregation procedures, but the results that we prove in the following section suggest that defining classes of procedures by means of axiomatic properties might be a good candidate to obtain the commutativity of the two operators.

[7]The behaviour is slightly different in case $\mathcal{L}$ does not contain $\top$ or $\bot$, in which case $\mathcal{LF}[\mathcal{F}_{\mathcal{L}}[\text{AX}]] = \mathcal{L} \cup \{\top, \bot\}$.

4.2 Characterisation Results for Propositional Languages

The aim of this section is to explore the relationship between the two definitions of classes of aggregation procedures introduced in Section 4.1: collectively rational procedures on one side, and procedures defined by axiomatic requirements on the other. In particular, we look for results of the following form:

$$\mathcal{CR}[\mathcal{L}] \quad = \quad \mathcal{F}_{\mathcal{L}}[\text{AX}],$$

for languages $\mathcal{L}$ and axioms AX. We call such findings *characterisation results*: they provide necessary and sufficient axiomatic conditions for an aggregation procedure to be collectively rational with respect to a language for integrity constraints. The focus of this section is on languages: We provide complete characterisation for some basic classes of languages defined in a syntactic fashion, proving the correspondence with some of the main classical axioms from the literature on Social Choice Theory. We shift the focus to axioms in Section 4.3.

Definitions of all the axiomatic properties we refer to in this section can be found in Section 2.2. Axioms will be denoted with the capital letter associated with it, e.g., we will write U for unanimity and I for independence. Recall that a language for integrity constraints is a set of propositional formulas closed under conjunction and logical equivalence.

4.2.1 Full Characterisations

Recall that a procedure is unanimous if it shares the view of the individuals in case they all agree, either all accepting or rejecting a certain issue. The first characterisation result shows that the set of aggregation procedures that lift all rationality constraints that can be expressed as conjunctions of literals is precisely the class of unanimous procedures:

Theorem 4.2.1. $\mathcal{CR}[\mathit{literals}] = \mathcal{F}_{\mathit{literals}}[\text{U}]$.

Proof. One direction is easy: If $X := \text{Mod}(\ell)$ is a domain defined by a literal ℓ, then every individual ballot must agree with it, either positively or negatively depending on its sign. This entails, by unanimity, that the collective outcome agrees with the individual ballots. Thus, F is collectively rational with respect to ℓ, and by Lemma 4.1.3 F is CR with respect to the full language of literals.

For the other direction, suppose that $F \in \mathcal{CR}[\mathit{literals}]$. Fix an issue $j \in \mathcal{I}$. Pick a profile $\boldsymbol{B} \in \mathcal{D}^n$ such that $b_{i,j} = 1$ (or 0) for all $i \in \mathcal{N}$. That is, $\boldsymbol{B} \in \text{Mod}(p_j)^{\mathcal{N}}$ (or $\neg p_j$, respectively). Since F is collectively rational for every literal, including p_j and $\neg p_j$, it must be the case that $F(\boldsymbol{B})_j = 1$ (or 0, respectively), proving unanimity of the aggregator. □

As remarked in Section 4.1.1, the language generated from literals is the same as the language of *cubes*, i.e., finite conjunctions of literals. We can therefore state the following corollary.

Corollary 4.2.2. $\mathcal{CR}[cubes] = \mathcal{F}_{cubes}[\mathrm{U}]$.

An *equivalence* is a bi-implication of literals where the literals are both positive (or both negative, which amounts to the same thing). Call the language for integrity constraints generated by equivalences $\mathcal{L}_{\leftrightarrow}$, i.e., the set $\{p_j \leftrightarrow p_k \mid p_j, p_k \in PS\}$ closed under conjunction and logical equivalence. This language allows us to characterise issue-neutral aggregators, i.e., procedures that treat distinct issues in the same way:

Theorem 4.2.3. $\mathcal{CR}[\mathcal{L}_{\leftrightarrow}] = \mathcal{F}_{\mathcal{L}_{\leftrightarrow}}[\mathrm{N}^{\mathcal{I}}]$.

Proof. To prove the first inclusion ($\supseteq$), pick an equivalence $p_j \leftrightarrow p_k$. This defines a domain in which issues j and k share the same pattern of acceptance/rejection, and since the procedure is neutral over issues, we get $F(\boldsymbol{B})_j = F(\boldsymbol{B})_k$. Therefore, the constraint given by the initial equivalence is lifted. Thus, we can conclude by Lemma 4.1.3 that the full langauge $\mathcal{L}_{\leftrightarrow}$ is lifted.

For the other direction ($\subseteq$), suppose that a profile $\boldsymbol{B}$ is such that $b_{i,j} = b_{i,k}$ for every $i \in \mathcal{N}$. This implies that $\boldsymbol{B} \in \mathrm{Mod}(p_j \leftrightarrow p_k)^{\mathcal{N}}$, and since F is in $\mathcal{CR}[\mathcal{L}_{\leftrightarrow}]$, $F(\boldsymbol{B})_j$ must be equal to $F(\boldsymbol{B})_k$. This holds for every such $\boldsymbol{B}$, proving that F is neutral over issues. □

With an analogous proof we can obtain a characterisation result involving the axiom of domain-neutrality. Recall that a procedure is domain-neutral if it symmetric with respect to any two issues. An XOR formula is a bi-implication of one negative and one positive literal. Let $\mathcal{L}_{XOR}$ be the language for integrity constraints generated from $\{p_j \leftrightarrow \neg p_k \mid p_j, p_k \in PS\}$.

Theorem 4.2.4. $\mathcal{CR}[\mathcal{L}_{XOR}] = \mathcal{F}_{\mathcal{L}_{XOR}}[\mathrm{N}^{\mathcal{D}}]$.

Proof. The first inclusion is straightforward: When every individual ballot in a profile satisfies the same XOR formula, then this means that there are two issues the behaviour of which is symmetrical. By domain-neutrality, the outcome of the aggregation is also symmetrical, and therefore the constraint is lifted.

To prove the remaining inclusion ($\subseteq$), suppose that a profile $\boldsymbol{B}$ is such that $b_{i,j} = 1 - b_{i,k}$ for every $i \in \mathcal{N}$. This implies that $\boldsymbol{B} \in \mathrm{Mod}(p_j \leftrightarrow \neg p_k)^{\mathcal{N}}$. As before, since F is in $\mathcal{CR}[\mathcal{L}_{XOR}]$, it must be the case that $F(\boldsymbol{B})_j = 1 - F(\boldsymbol{B})_k$ and F is domain-neutral. □

Consider now the language $\mathcal{L}^+_{\rightarrow}$ of positive implications, generated from formulas of the form $p_j \rightarrow p_k$, or equivalently $\neg p_j \rightarrow \neg p_k$, for $p_j, p_k \in PS$. Since $\mathcal{L}^+_{\rightarrow} \supseteq \mathcal{L}_{\leftrightarrow}$, we know that $\mathcal{CR}[\mathcal{L}^+_{\rightarrow}] \subseteq \mathcal{CR}[\mathcal{L}_{\leftrightarrow}] = \mathcal{F}_{\mathcal{L}_{\leftrightarrow}}[\mathrm{N}^{\mathcal{I}}]$. Therefore, a characterisation of the language of positive implications must involve the axiom of neutrality in combination with others. The right combination is the following.

Theorem 4.2.5. $\mathcal{CR}[\mathcal{L}^+_\rightarrow] = \mathcal{F}_{\mathcal{L}^+_\rightarrow}[\mathrm{N}^\mathcal{I}, \mathrm{M}^\mathrm{N}]$.

Proof. ($\supseteq$) Let us first consider the case of individual ballots all satisfying a certain positive implication IC $= p_j \rightarrow p_k$. We want to prove that if F satisfies I and M^I then F lifts the integrity constrait IC. Note that if an individual accepts issue j then she also accepts issue k. Therefore, the first part of the antecedent forming the axiom of N-monotonicity is satisfied. We now have to consider two cases: if for all $i \in \mathcal{N}$ we have that $b_{i,j} = b_{i,k} = 1$, then by issue-neutrality we have that $F(\boldsymbol{B})_j = F(\boldsymbol{B})_k$. The constraint IC is therefore satisfied, as the only way to falsify it is by accepting j and rejecting k. If on the other hand there is an individual i such that $b_{i,j} = 0$ while $b_{i,k} = 1$, then $\boldsymbol{B}$ fully satisfies the antecedent of M^N and therefore $F(\boldsymbol{B})_k = 1$ whenever $F(\boldsymbol{B})_j = 1$, again making it impossible to falsify the integrity constraint IC.

For the remaining inclusion ($\subseteq$), suppose that a profile $\boldsymbol{B}$ is such that whenever $b_{i,j} = 1$ then $b_{i,k} = 1$ for every $i \in \mathcal{N}$. This implies that $\boldsymbol{B} \in \mathrm{Mod}(p_j \rightarrow p_k)^\mathcal{N}$. Since we assumed F to be in $\mathcal{CR}[\mathcal{L}^+_\rightarrow]$, $F(\boldsymbol{B})_j = 1$ entails $F(\boldsymbol{B})_k = 1$, for the constraint $p_j \rightarrow p_k$ has to be lifted. Therefore F is N-monotonic. We have already remarked that, due to Theorem 4.2.3, all procedures in $\mathcal{CR}[\mathcal{L}^+_\rightarrow]$ are also issue-neutral. □

The last result would suggest that a characterisation of the language of negative implications (i.e., when exactly one of the two literals is negative) might be proven by considering the axiom of domain-neutrality combined with N-monotonicity. Unfortunately, in the absence of additional restrictions on the set of profiles, this characterisation does not hold. We prove a partial characterisation for this class in Section 4.2.3.

We conclude this section by characterising the classes of collectively rational procedures for languages at the extremes of the spectrum: the full language $\mathcal{L}_{PS}$, the language of tautologies, and that of contradictions. For the last two classes the characterisation is straightforward. Recall that $\mathcal{F} = \{F : \mathcal{D}^\mathcal{N} \rightarrow \mathcal{D}\}$ is the class of all aggregation procedures (for fixed $\mathcal{I}$). We have already stated in Lemma 4.1.6 that tautologies and contradictions are vacuous requirements for what concerns collective rationality, and here we use these arguments to give a characterisation result for this trivial class of formulas. Let $\{\top\}$ be the language of all tautologies, and $\{\bot\}$ be the language of all contradictions:

Proposition 4.2.6. $\mathcal{CR}[\{\top\}] = \mathcal{CR}[\{\bot\}] = \mathcal{F}$.

If on the other hand we turn to study the class of procedures that lift *any* integrity constraint in $\mathcal{L}_{PS}$ we discover an interesting class of procedures. Let us give the following definition,that generalises the notion of dictatorship.[8]

[8]This class was introduced by Cariani et al. (2008) in the context of judgment aggregation under the name of *rolling dictatorships*. A related (but different) notion is that of *positional dictatorships*, introduced by Roberts (1980a) and rather standard in Social Choice Theory

Definition 4.2.7. An aggregation procedure $F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D}$ is a *generalised dictatorship*, if there exists a map $g : \mathcal{D}^{\mathcal{N}} \to \mathcal{N}$ such that $F(\boldsymbol{B}) = B_{g(\boldsymbol{B})}$ for every $\boldsymbol{B} \in \mathcal{D}^{\mathcal{N}}$.

That is, a generalised dictatorship copies the ballot of a (possibly different) individual in every profile. Call this class GDIC. This class fully characterises the class of collectively rational aggregators for the full propositional language $\mathcal{L}_{PS}$:

Theorem 4.2.8. $\mathcal{CR}[\mathcal{L}_{PS}] = \text{GDIC}$.

Proof. Clearly, every generalised dictatorship lifts any arbitrary integrity constraint IC $\in \mathcal{L}_{PS}$. To prove the other direction, suppose that $F \notin \text{GDIC}$. Hence, there exists a profile $\boldsymbol{B} \in \mathcal{D}^{\mathcal{N}}$ such that $F(\boldsymbol{B}) \neq B_i$ for all $i \in \mathcal{N}$. This means that for every i there exists an issue j_i such that $F(\boldsymbol{B})_{j_i} \neq b_{i,j_i}$. We now want to build a propositional formula that is satisfied by all individuals and not by the collective outcome, proving that F is not CR with respect to the full propositional language. Define a literal ℓ_{j_i} to be equal to p_{j_i} if $b_{i,j_i} = 1$, and to $\neg p_{j_i}$ otherwise. Consider as integrity constraint IC the following formula: $\bigvee_i \ell_{j_i}$. Clearly, $B_i \models \text{IC}$ for every $i \in \mathcal{N}$, i.e., $\boldsymbol{B}$ is a rational profile for the integrity constraint IC. But by construction, $F(\boldsymbol{B}) \not\models \text{IC}$, as $F(\boldsymbol{B})$ differs from the individual ballots on all literals in IC. Therefore, F is not collectively rational for IC and does not belong to the class $\mathcal{CR}[\mathcal{L}_{PS}]$. □

As a concluding remark, recall that the language generated by clauses coincides with the full propositional language, as every propositional formula is equivalent to a conjunction of clauses by taking its conjunctive normal form. We therefore obtain the following:

Corollary 4.2.9. $\mathcal{CR}[\textit{clauses}] = \text{GDIC}$.

We analyse restricted languages of clauses in Section 4.4.

4.2.2 How to Combine Characterisation Results

Most of the characterisation results presented thus far characterise a class of procedures determined by a *single* axiom and by a *uniform* description of the language. In this section we briefly explain to what extent such results can be combined to allow us to make predictions regarding the collective rationality of procedures satisfying several such axioms, or in the case where the integrity constraints can be chosen from a more complex language.

(Roemer, 1996). It denotes social choice functions that follow the choice of the individual having a certain position in society (e.g., egalitarian maximin). The same term is also used to indicate a generalisation of the median rule in single-peaked domains (Moulin, 1988).

Consider the case of two characterisation results $\mathcal{CR}[\mathcal{L}_1] = \mathcal{F}_{\mathcal{L}_1}[\mathrm{AX}_1]$ and $\mathcal{CR}[\mathcal{L}_2] = \mathcal{F}_{\mathcal{L}_2}[\mathrm{AX}_2]$. By part (*ii*) of Lemma 4.1.8 and by $\mathcal{F}_{\mathcal{L}_1 \cup \mathcal{L}_2}[\mathrm{AX}_1, \mathrm{AX}_2] \subseteq \mathcal{F}_{\mathcal{L}_1}[\mathrm{AX}_1] \cap \mathcal{F}_{\mathcal{L}_2}[\mathrm{AX}_2]$ we can infer that

$$\mathcal{F}_{\mathcal{L}_1 \cup \mathcal{L}_2}[\mathrm{AX}_1, \mathrm{AX}_2] \subseteq \mathcal{CR}[\mathcal{L}_1 \cup \mathcal{L}_2].$$

(But note that the other inclusion is not always true.) This entails that if we express constraints in the language $\mathcal{L}_1 \cup \mathcal{L}_2$ or in any of its sublanguages, then picking procedures from $\mathcal{F}_{\mathcal{L}_1 \cup \mathcal{L}_2}[\mathrm{AX}_1, \mathrm{AX}_2]$ is a sufficient condition for collective rationality. If, instead, we start from a class of procedures satisfying axioms AX_1 and AX_2 on the language $\mathcal{L}_1 \cup \mathcal{L}_2$ then we can infer that these procedures lift any language $\mathcal{L} \subseteq \mathcal{L}_1 \cup \mathcal{L}_2$, since as we previously observed $\mathcal{F}_{\mathcal{L}_1 \cup \mathcal{L}_2}[\mathrm{AX}_1, \mathrm{AX}_2]$ is included in $\mathcal{CR}[\mathcal{L}_1 \cup \mathcal{L}_2]$, which in turn is included in $\mathcal{CR}[\mathcal{L}]$.

Suppose, for instance, that we are considering an aggregation procedure F that satisfies both the axiom of unanimity and the axiom of domain-neutrality, i.e., a function in $\mathcal{F}[\mathrm{U}, \mathrm{N}^{\mathcal{D}}]$. Then we are certain that any integrity constraint that can be expressed as a conjunction of literals and XOR-formulas is lifted by F, i.e., any formula in the language for integrity constraint generated by $\mathit{cubes} \cup \mathcal{L}_{XOR}$. On the other hand, if we can describe an integrity constraint as a conjunction of an XOR-formula with an equivalence, then by Theorems 4.2.3 and 4.2.4 it is sufficient to pick an aggregator that is both neutral over issues and domain-neutral to guarantee collective rationality.

4.2.3 Partial Characterisation Results

In this section we prove two partial characterisations for languages of implications, i.e., we give only sufficient axiomatic conditions for an aggregator to be CR with respect to the language under consideration. Recall from Theorem 4.2.5 that the language of positive implications $\mathcal{L}^+_{\rightarrow}$ is the language for integrity constraints generated by formulas of the form $p_j \rightarrow p_k$ for $p_j, p_k \in PS$.

Proposition 4.2.10. $\mathcal{CR}[\mathcal{L}^+_{\rightarrow}] \supsetneq \mathcal{F}_{\mathcal{L}^+_{\rightarrow}}[\mathrm{N}^{\mathcal{I}}, \mathrm{M}^{\mathrm{I}}, \mathrm{I}]$.

Proof. It is sufficient to notice that $\mathcal{F}_{\mathcal{L}^+_{\rightarrow}}[\mathrm{N}^{\mathcal{I}}, \mathrm{M}^{\mathrm{I}}, \mathrm{I}] \subseteq \mathcal{F}_{\mathcal{L}^+_{\rightarrow}}[\mathrm{N}^{\mathcal{I}}, \mathrm{M}^{\mathrm{N}}]$ to obtain the inclusion from Theorem 4.2.5. To see this, observe that under the assumption of neutrality and independence the two versions of monotonicity coincide, as we already remarked in Section 2.2.

We give a direct counterexample to prove that the inclusion is strict. Let $\mathcal{I}$ consist of 2 issues j and k, and $\mathcal{N}$ of 2 individuals i_1 and i_2. Define F to accept both issues in all profiles except for the one where both individuals accept all issues, in which we fix F as rejecting both issues. F is clearly collectively rational for the language of positive implications over two propositional symbols, as the only ballots falsifying such constraints are never in the output of F. Consider now the following two profiles.

$\boldsymbol{B}$:

	j	k
i_1	0	1
i_2	1	1
F	1	1

$\boldsymbol{B}'$:

	j	k
i_1	1	1
i_2	1	1
F	0	0

The second profile $\boldsymbol{B}'$ is obtained from the first by increasing support on issue j, which is accepted by F in $\boldsymbol{B}$ while it is rejected in $\boldsymbol{B}'$, contradicting I-monotonicity. Note that F is also not independent, as the outcome on issue k changes even if the individual ballots concerning this issue coincide in the two profiles. Therefore, F is collectively rational with respect to positive implications but is neither I-monotonic nor independent. However, F has to satisfy the axiom of issue-neutrality. The reason is that this axiom corresponds to collective rationality with respect to equivalences, as we have seen in Theorem 4.2.3, and equivalences can be expressed as conjunctions of positive implications. Formally, as $\mathcal{L}^+_{\rightarrow} \supseteq \mathcal{L}_{\leftrightarrow}$, we have by Lemma 4.1.8 that $\mathcal{CR}[\mathcal{L}^+_{\rightarrow}] \subseteq \mathcal{CR}[\mathcal{L}_{\leftrightarrow}]$, and by Theorem 4.2.3 this last class is equal to $\mathcal{F}[\mathrm{N}^{\mathcal{I}}]$. □

With a similar proof we can obtain a partial characterisation result for the language of negative implications $\mathcal{L}^-_{\rightarrow}$, i.e., formulas of the form $p_j \rightarrow \neg p_k$ and $\neg p_j \rightarrow p_k$ for $p_j, p_k \in PS$:

Proposition 4.2.11. $\mathcal{CR}[\mathcal{L}^-_{\rightarrow}] \supsetneq \mathcal{F}_{\mathcal{L}^-_{\rightarrow}}[\mathrm{N}^{\mathcal{D}}, \mathrm{M}^{\mathrm{I}}, \mathrm{I}]$.

Proof. Let F be a domain-neutral, independent and I-monotonic procedure. Assume first that $p_j \rightarrow \neg p_k$ is the integrity constraint, and let $\boldsymbol{B}$ be a rational profile such that $F(\boldsymbol{B})_j = 1$. We want to show that $F(\boldsymbol{B})_k = 0$ to prove the collective rationality of F with respect to the integrity constraint $p_j \rightarrow \neg p_k$. By individual rationality we have that $\{i \mid b_{i,j} = 1\} \subseteq \{i \mid b_{i,k} = 0\}$. We now build a profile $\boldsymbol{B}'$ from $\boldsymbol{B}$ by modifying the individual ballots on issue j only, such that $b'_{i,j} = 1 - b_{i,k}$ for all i. Observe that $\boldsymbol{B}'$ is still a rational profile for the initial integrity constraint. By M^{I} we can infer that $F(\boldsymbol{B}')_j = 1$, as we increased the number of individuals accepting j in $\boldsymbol{B}'$ without changing the individual ballots on any other issue. Now by $\mathrm{N}^{\mathcal{D}}$ we have that $F(\boldsymbol{B}')_j = 1 - F(\boldsymbol{B}')_k$, since $b'_{i,j} = 1 - b'_{i,k}$ for all i, and by I this entails $F(\boldsymbol{B})_k = 0$, as $b_{i,k} = b'_{i,k}$ for all i.

Assume now that the integrity constraint is $\neg p_j \rightarrow p_k$, and let $\boldsymbol{B}$ be a rational profile in which $F(\boldsymbol{B})_j = 0$. As in the first part of the proof, we want to show that $F(\boldsymbol{B})_k = 1$ to prove collective rationality with respect to the integrity constraint. In this case, we know that $\{i \mid b_{i,j} = 0\} \subseteq \{i \mid b_{i,k} = 1\}$, and let $\boldsymbol{B}'$ be a second profile constructed from $\boldsymbol{B}$ such that $b'_{i,j} = 1 - b_{i,k}$ for all i. Assume now that $F(\boldsymbol{B}')_j = 1$. Since in the move from $\boldsymbol{B}$ to $\boldsymbol{B}'$ we increased the amount of *rejections* for j, we can apply M^{I} , this time from $\boldsymbol{B}'$ to $\boldsymbol{B}$, to conclude that also $F(\boldsymbol{B})_j = 1$, against our initial assumption. Thus, $F(\boldsymbol{B}')_j = 0$, and the proof can be concluded using the same arguments as in the previous case.

A counterexample similar to the one used in Proposition 4.2.10 can be devised to prove that the inclusion is strict. □

We can combine the two partial characterisations proved in this section using Lemmas 4.1.8 and 4.1.9 to obtain sufficient (but not necessary) axiomatic conditions for lifting the full language of implications $\mathcal{L}_{\to}$, i.e., the language for integrity constraints generated from the set $\{p \to q \mid p, q \in PS\}$:

$$\mathcal{CR}[\mathcal{L}_{\to}] \supsetneq \mathcal{F}_{\mathcal{L}_{\to}}[\mathrm{M^I}, \mathrm{I}, \mathrm{N}^{\mathcal{I}}, \mathrm{N}^{\mathcal{D}}].$$

Observe that $\mathcal{L}_{\to}$ = *2-clauses*, i.e., the class of formulas that are equivalent to a conjunction of clauses of maximal size two. This constitutes a very interesting language for integrity constraints and, as we will see in Section 4.4.2, the majority rule is one of the collectively rational procedures with respect to this language.

4.3 Characterisations Results for Classical Axiomatic Properties

In the previous section we proved several characterisation results for various simple fragments of the propositional language associated with an aggregation problem. In this section we shift our focus from syntactic descriptions of languages to axiomatic properties of aggregation procedures, having the axioms as variables when exploring the possibility for a characterisation result.

We first generalise some of the results proven in the previous section to more general characterisations of axioms, dropping the domain restriction given by the language. Then, we prove some negative results involving axioms like anonymity or independence, which are properties that constrain the aggregation on more than one profile. For these axioms a characterisation cannot be found.

At the end of the section we turn to the problem of determining which axioms can be characterised using collective rationality and which cannot. While we do prove a formal result for all the axioms listed in Section 2.2, we only provide a preliminary answer to this more general question.

4.3.1 Axioms Characterisation

Consider the class $\mathcal{F}$[AX], dropping the subscript $\mathcal{L}$, as representing the class of procedures that defines an axiom. As observed at the end of Section 4.1.1, for all the axiomatic properties considered we know that $\mathcal{F}[\mathrm{AX}] \subseteq \mathcal{F}_{\mathcal{L}}[\mathrm{AX}]$ for all $\mathcal{L} \subseteq \mathcal{L}_{PS}$. Some of the characterisation results proved in the previous section can be easily generalised to the class $\mathcal{F}$[AX], becoming therefore characterisations of classical axioms.

Corollary 4.3.1. *The following equivalences hold:*
(*i*) $\mathcal{F}[\mathrm{U}] = \mathcal{CR}[\mathit{literals}]$.
(*ii*) $\mathcal{F}[\mathrm{N}^{\mathcal{I}}] = \mathcal{CR}[\mathcal{L}_{\leftrightarrow}]$.
(*iii*) $\mathcal{F}[\mathrm{N}^{\mathcal{D}}] = \mathcal{CR}[\mathcal{L}_{XOR}]$.

Proof. Refer to Theorems 4.2.1, 4.2.3 and 4.2.4. For all three classes we prove that $\mathcal{F}_{\mathcal{L}}[\mathrm{AX}] = \mathcal{F}[\mathrm{AX}]$, for the relevant axiom and language. We do so by proving that the condition required by the axiom is a vacuous requirement *outside* domains defined by formulas in $\mathcal{L}$. In the first case, suppose that $\boldsymbol{B}$ is a profile in which all individuals unanimously accept (reject) a given issue j. This means that $\boldsymbol{B} \in \mathrm{Mod}(p_j)$ ($\boldsymbol{B} \in \mathrm{Mod}(\neg p_j)$, respectively). Thus, all profiles in the scope of the axiom of unanimity are models of a literal. The other cases are similar. In the second one, if two formulas share the same pattern of acceptance/rejection (issue-neutrality), then we are in a profile that is a model of a bi-implication. Similar remarks hold for the third case of domain-neutrality. Therefore, in the three cases under consideration, for a procedure to satisfy an axiom on domains defined by $\mathcal{L}$ is equivalent to satisfying the same axiom on the full domain. □

4.3.2 Negative Results

Results of this form cannot be proven for other important axioms, for which it is not even possible to obtain a characterisation result. We study four such classes in the following result. Recall that $\mathcal{LF}[\mathcal{G}]$ is the set of integrity constraints lifted by all the aggregation procedures in $\mathcal{G}$. Let $\{\top, \bot\}$ be the language for integrity constraints composed of all tautologies and contradictions.

Proposition 4.3.2. *The following equivalences hold:*

$$\mathcal{LF}[\mathcal{F}[\mathrm{I}]] = \mathcal{LF}[\mathcal{F}[\mathrm{A}]] = \mathcal{LF}[\mathcal{F}[\mathrm{M}^{\mathrm{I}}]] = \mathcal{LF}[\mathcal{F}[\mathrm{M}^{\mathrm{N}}]] = \{\top, \bot\}$$

Proof. We prove this proposition by constructing for any contingent formula φ (i.e., such that both φ and $\neg\varphi$ are satisfiable) an independent, anonymous and monotonic procedure that is not collectively rational with respect to φ. Let φ be such a formula, and let $B^\star \in \mathcal{D}$ be a ballot such that $B^\star \not\models \varphi$. Consider now the constant procedure $F(\boldsymbol{B}) = B^\star$ for all profiles $\boldsymbol{B}$: this procedure is independent, anonymous, I-monotonic and N-monotonic, but it is not collectively rational with respect to φ. □

An immediate corollary of this result is that it is not possible to obtain a characterisation of these axioms in terms of collective rationality:

Corollary 4.3.3. *There is no language $\mathcal{L} \subseteq \mathcal{L}_{PS}$ such that $\mathcal{CR}[\mathcal{L}] = \mathcal{F}[\mathrm{I}]$. The same holds for $\mathcal{F}[\mathrm{A}]$, $\mathcal{F}[\mathrm{M}^{\mathrm{I}}]$ and $\mathcal{F}[\mathrm{M}^{\mathrm{N}}]$.*

Proof. Recall from Section 4.1.2 that in the presence of a characterisation result, the set of integrity constraints lifted by a class of procedures is uniquely determined. Suppose then that $\mathcal{CR}[\mathcal{L}] = \mathcal{F}[\text{AX}]$ for $\text{AX} \in \{\text{I}, \text{A}, \text{M}^\text{I}, \text{M}^\text{N}\}$ and a given $\mathcal{L}$. Proposition 4.3.2 forces $\mathcal{L}$ to be equal to $\{\top, \bot\}$, but we have already proven that this class characterises the whole set of procedures $\mathcal{F}$ (cf. Proposition 4.2.6 combined with Lemma 4.1.8). Therefore, such a characterisation cannot exist. □

Note that this argument can be generalised to prove that the class $\mathcal{F}_\mathcal{L}[\text{I}]$ (and the same holds for A, M^I and M^N) cannot be characterised for any restriction given by a language $\mathcal{L}$. It is sufficient to note that the constant procedure employed in the proof of Proposition 4.3.2 can be defined regardless of the domain restriction.

Our interest toward these axioms does not cease here. On the contrary, Proposition 4.3.2 showed that the classes of monotone, independent and anonymous procedures behave in the same way as the full class $\mathcal{F}$ for what concerns collective rationality. This suggests that interesting characterisations can be studied *inside* those classes, replacing the set $\mathcal{F}$ of all procedures with, e.g., the class $\mathcal{F}[\text{I}]$. We are going to pursue this approach in the following sections.

Let us conclude the session with some remarks about the structure of the axiomatic properties seen so far. The results proven in this section are consistent with the intuition that assumptions regarding the collective rationality of an aggregator can only condition the outcome in view of a single profile at a time. Axioms like independence or monotonicity coordinate the behaviour of the aggregator on more than one profile, and for this reason cannot be characterised as collective rationality with respect to a particular language. This ideal line that can be drawn to separate "intra-profile axioms" from "inter-profile axioms" has been given the name of "single-profile" versus "multi-profile" approach in the literature on Social Choice Theory (Samuelson, 1967; Roberts, 1980a).

The multi-profile approach to Social Choice Theory is the rather standard study of aggregation procedures as functions defined on the domain of all profiles, while under the single-profile approach the object of study is a single profile at a time, together with its outcome. Classical axioms have first been formulated under the former approach, which led to the celebrated Arrow's Impossibility Theorem (Arrow, 1963). In the following decades several authors proposed the single-profile approach to Social Choice Theory as a possible escape from Arrow's impossibility (Samuelson, 1967). Unfortunately, as several theorems have shown (see, e.g., Roberts, 1980b; Pollak, 1979), the impossibility persists even after the classical axioms are transformed into their so-called "single-profile analogues".

It is not our purpose to give a formal treatment of such notions,[9] but we may put forward the following informal definition.

Definition 4.3.4. (*informal*) An axiom AX is an *intra-profile* property of aggregation procedures if it can be written in the form $\forall \boldsymbol{B} \;\; \Psi(\boldsymbol{B}, F(\boldsymbol{B}))$ where Ψ is a

[9]For a formalisation of similar concepts in first-order logic of predicates we refer to the work of Rubinstein (1984) and our previous work (Grandi and Endriss, 2012).

property of the profile $\boldsymbol{B}$ and of the outcome $F(\boldsymbol{B})$. An axiom is an *inter-profile* property otherwise.

It is easy to see that collective rationality with respect to a certain language is an intra-profile requirement, since it comes of the following form:

$$\forall \boldsymbol{B} \text{ if } B_i \models \text{IC for all } i \in \mathcal{N} \text{ then also } F(\boldsymbol{B}) \models \text{IC}$$

To prove that anonymity, independence and monotonicity are genuine inter-profile properties would require a more precise definition than the one we have provided, but for now it is sufficient to observe that their formulation involves a quantification on two distinct profiles, and that this cannot be easily translated into a single-profile statement. By summing up these remarks, we can conclude that if an axiom can be characterised as collective rationality with respect to a certain language, then it is necessarily intra-profile, and the universally quantified formula expressing this axiom is exactly that of collective rationality. Axioms that are instead genuinely inter-profile, like independence, monotonicity and anonymity, but also non-imposition, non-dictatorship and permutation-neutrality (see Riker, 1982), cannot be characterised in terms of collective rationality, and it is easy to prove a result like Proposition 4.3.2 for such classes.

On the other hand, not all intra-profile requirements can be expressed as collective rationality with respect to a certain language. A counterexample can be found in a property inspired from the condition of "unrestricted domain over triples" presented by Pollak (1979). This condition is the single-profile analogue of the axiom that classically goes under the name of universal domain. It is a condition of richness imposed on a single profile $\boldsymbol{B}$, and in binary aggregation it requires that for every combination that can be constructed with three individuals and three issues, i.e., for every profile over $\{0,1\}^3$, there exist three issues i_1, i_2 and i_3 and three individuals such that their ballots in $\boldsymbol{B}$ restricted to issues i_1, i_2 and i_3 coincide with the given subprofile. If this is the case, we set the outcome of the function to accept all the issues, i.e., $F(\boldsymbol{B}) = (1, \ldots, 1)$. The condition is genuinely intra-profile but in view of its syntactic structure (it contains an existential quantifier over three different individuals and issues) cannot be expressed as collective rationality with respect to a given language.

4.4 Quota Rules and Languages of Clauses

This section is devoted to a thorough exploration of the classes of collectively rational procedures for several languages of clauses inside the class of quota rules.

We introduced quota rules in Section 2.3.1 as procedures assigning a quota $0 \leqslant q_j \leqslant n+1$ to every issue j such that $F(\boldsymbol{B})_j = 1 \Leftrightarrow |\{i \mid b_{i,j} = 1\}| \geqslant q_j$. In the same section we also proved that quota rules are axiomatised as the class of independent, anonymous and monotone procedures. If we denote with $\mathcal{QR}$ the set of quota rules, then we can write $\mathcal{QR} = \mathcal{F}[\text{A}, \text{I}, \text{M}^{\text{I}}]$.

There are several reasons why the choice of quota rules and languages of clauses constitutes an interesting combination. First, since we have proven in Section 4.3 that no characterisation result is possible neither for independent, nor for anonymous, nor for monotonic procedures, it is important to explore classes of collectively rational procedures inside those classes. As remarked earlier, by combining all three axioms together we obtain the set of quota rules. Second, languages of clauses are the most expressive ones, ranging from literals, to implications, to the full expressivity of $\mathcal{L}_{PS}$.

By Corollary 4.2.9 we know that to obtain interesting results it is necessary to limit either the size or the shape of the clauses that build up a language for integrity constraints. In this section we concentrate on languages defined by bounding the size of clauses, and quota rules seem a perfect candidate to deal with such restrictions, as they allow us to put restrictions on quotas and constrain them with equations.

We will assume for the rest of this section that the number of issues is always strictly bigger than the limitation on the size of a clause. This is because in case the number of issues is smaller or equal than the bound on clauses this limitation is fictitious, and the language becomes fully expressive, thereby ruling out any non-trivial characterisation.

4.4.1 Positive and Negative Clauses

We start by studying the special case of positive and negative clauses of arbitrary size, obtaining necessary and sufficient conditions for quota rules to lift such constraints, and exploring characterisation results inside these classes.

For $k \geqslant 1$, we define an exact k-clause as a clause of length k, i.e., a clause in which exactly k propositional symbols occur either positively or negatively but not both.[10] A k-clause is a clause of size *at most* k. A k-pclause is a positive k-clause, i.e., a disjunction where all literals are positive, and a k-nclause is a negative k-clause, where all literals are negative. Given a clause $\varphi = \ell_1 \vee \cdots \vee \ell_k$, we say that an issue j *occurs in* φ if one and only one of p_j and $\neg p_j$ is one of the literals of φ. For instance, the following formula $p \vee q \vee \neg p \vee \neg r \vee \neg r$ is a 2-clause in which two propositional symbols q and r occur, while p occurs in a spurious way and does not add to the length of the clause.

Proposition 4.4.1. *A quota rule is CR for an exact k-pclause* IC *if and only if $\sum_j q_j < n + k$, with j ranging over all issues that occur in* IC *and n being the number of individuals, or $q_j = 0$ for at least one issue j that occurs in* IC.

Proof. Suppose IC $= p_1 \vee \cdots \vee p_k$ and let $i_1, \ldots, i_k$ be the corresponding issues. Given that IC is a positive clause, the only way to generate a paradox is by

[10] This is to exclude from the count redundant subformulas of a clause like $p_j \vee p_j$ or $p_j \vee \neg p_j$.

rejecting all issues $i_1, \dots, i_k$. It is easy to see that this cannot occur if the quota for one of these issues is 0. We can therefore assume that all quotas are positive.

Suppose now that we can create a paradoxical profile $\boldsymbol{B}$. Every individual ballot B_i must accept at least one issue to satisfy the integrity constraint; therefore the profile $\boldsymbol{B}$ contains at least n acceptances concerning issues $i_1, \dots, i_k$. On the other hand, since $F(\boldsymbol{B})_j = 0$ for all $j = 1, \dots k$, we have that the number of individuals accepting an issue j is strictly lower than q_j. As previously remarked, there are at least n acceptances on the profile $\boldsymbol{B}$ and the maximal number of acceptances that allows rejection on all issues is $\sum_j (q_j - 1)$. Hence we have that $n \leqslant \sum_j (q_j - 1)$. This is equivalent to $n + k \leqslant \sum_j q_j$, since all j are distinct, thus we can construct a paradox with our IC if and only if this inequality holds. By taking the contrapositive we obtain the statement of Proposition 4.4.1. □

With a similar proof we get the analogous result for negative clauses:

Proposition 4.4.2. *A quota rule is CR for an exact k-nclause* IC *if and only if* $\sum_j q_j > (k-1)n$*, with j ranging over all issues that occur in* IC *and n being the number of individuals, or $q_j = n + 1$ for at least one issue j that occurs in* IC*.*

Proof. As for the previous proof, we can start by assuming that all quotas are $\leqslant n$, for otherwise the constraint is trivially lifted. Suppose now that we can create a paradoxical profile $\boldsymbol{B}$. For the constraint to be lifted there must be at least n rejections on the profile $\boldsymbol{B}$, and the maximal amount of rejections that allows acceptance of all issues, in order to create such a paradox, is $\sum_j (n - q_j)$; hence $n \leqslant \sum_j (n - q_j)$. Therefore, since each quota refers to distinct issues, we can construct a paradox with this IC if and only if $\sum_j q_j \leqslant (k-1)n$, and by taking the contrapositive we obtain the statement of Proposition 4.4.2. □

In case $k = 1$, i.e., the case of a literal p_j or $\neg p_j$, we obtain $q_j < n+1$ from the first proposition and $q_j > 0$ from the second, thus forcing the rule to be unanimous on issue j (quota rules satisfying $1 \leqslant q_j \leqslant n$ for all j are unanimous). This is consistent with Proposition 4.2.1.

We now want to turn these results into characterisation results along the lines of those proved in Section 4.2. We first have to define languages of clauses from our definition of clauses of a limited size. Let *k-pclauses* and *k-nclauses* denote, respectively, the set of positive (negative) clauses of size $\leqslant k$. Denote with $\mathcal{QR}_{eq}$ the set of quota rules such that all quotas q_j for $j \in \mathcal{I}$ satisfy the equation eq in the subscript. The function $\lceil x \rceil$ is defined as the smallest integer bigger or equal than x. We prove the following corollary of Propositions 4.4.1 and 4.4.2:

Corollary 4.4.3. *The following inclusions are true:*

(i) $\mathcal{QR}_{q_j \leqslant \lceil \frac{n}{k} \rceil} \subseteq \mathcal{CR}[k\text{-}pclauses]$

(ii) $\mathcal{QR}_{q_j \geqslant n - \lceil \frac{n}{k} \rceil + 1} \subseteq \mathcal{CR}[k\text{-}nclauses]$

Proof. The first result is a consequence of Proposition 4.4.1 and the fact that $\lceil \frac{n}{k} \rceil < \frac{n}{k} + 1$. If all quotas $q_j \leqslant \lceil \frac{n}{k} \rceil$ then for any subset of k issues that might occur in a k-pclause we have that $\sum_j q_j \leqslant \sum_j \lceil \frac{n}{k} \rceil < \sum_j (\frac{n}{k} + 1) = n + k$.

Analogously, referring this time to Proposition 4.4.2, we have that for any k set of issues $\sum_j q_j \geqslant \sum_j (n - \lceil \frac{n}{k} \rceil + 1) > \sum_j (n - \frac{n}{k} - 1 + 1) = (k-1)n$. □

The significance of the previous result resides in the fact that the bounds given by those equations are the lowest uniform bounds we can give to quotas to guarantee collective rationality. This is what is proven in the following corollary, which focuses on uniform quota rules.

Corollary 4.4.4. *A uniform quota rule is CR with respect to:*
(*i*) *a k-pclause if and only if* $q \leqslant \lceil \frac{n}{k} \rceil$;
(*ii*) *a k-nclause if and only if* $q \geqslant n - \lceil \frac{n}{k} \rceil + 1$.

Proof. *i*) In the case of uniform quota rules the equation in Proposition 4.4.1 takes the following form $\sum_j q = kq < n + k$. This holds if and only if $q < \frac{n}{k} + 1$, that is equivalent, as remarked in the proof of the previous corollary, to $q \leqslant \lceil \frac{n}{k} \rceil$.

ii) In the same way, using a single quota in the equation of Proposition 4.4.2 we obtain $kq > (k-1)n$, i.e., $q > n - \frac{n}{k}$ which is equivalent to $q \geqslant n - \lceil \frac{n}{k} \rceil + 1$. □

Note that the two equations in the previous proposition are incompatible except for the case of $k = 2$ and n being odd, in which case $q = \frac{n+1}{2}$. This proves that no uniform quota rule is collectively rational on both positive *and* negative clauses of a given size, except for the case of n being odd and $q = \frac{n+1}{2}$. This quota rule is the majority rule, and it is now time to study this procedure in more detail.

4.4.2 The Majority Rule

The majority rule is the uniform quota rule that accepts an issue whenever there are more individuals accepting the issue than rejecting it. The majority rule is perhaps one of the most natural aggregation rules. It is arguably the one that is used the most in practical applications, but as we noted in Chapter 3, it also generates a plethora of paradoxical situations that have been widely studied in the literature. As we have seen in Section 2.3.2, the majority rule is axiomatised by A, I, $N^{\mathcal{I}}$, M^{I} and $N^{\mathcal{D}}$ (cf. Proposition 2.3.7).

In case the number of individuals is odd the majority rule has a unique definition by setting the quota to $q = \frac{n+1}{2}$. The case of an even number of individuals is more problematic, to account for profiles in which exactly half of the individuals accept an issue and exactly half reject it. We give two different definitions. The *weak majority rule* with quota $q = \frac{n}{2}$ accepts an issue if and only if at least half of the individuals accepts it. The *strict majority rule* accepts an issue if and only if a strict majority of the individuals accepts it, i.e., it is the uniform quota rule with quota $q = \frac{n+2}{2}$. The first rule favours acceptance, while the second favours rejection of an issue. In the following sections we will characterise the set of integrity constraints that are lifted by the majority rule, by the weak majority rule, and by the strict majority rule.

Odd Number of Individuals: The Majority Rule

In this section we make the assumption that the number of individuals is odd, and we indicate with *Maj* the uniform quota rule with quota $q = \frac{n+1}{2}$. We make the additional assumption that there are at least 3 individuals. The majority rule in the case of 1 individual is the identity function that outputs the ballot received by the only individual.

Let us begin with a base-line result that proves collective rationality of the majority rule in case the integrity constraint is equivalent to a conjunction of 2-clauses:

Proposition 4.4.5. *The majority rule is in $\mathcal{CR}$[2-clauses].*

Proof. Since the majority rule is a uniform quota rule, its quota has to satisfy both types of constraints from Corollary 4.4.3 to lift both positive and negative clauses. Thus $q \leqslant \lceil \frac{n}{k} \rceil$ and $q \geqslant n - \lceil \frac{n}{k} \rceil + 1$, and as remarked in the previous section these are incompatible unless $k = 2$ and n is odd, in which case $q = \frac{n+1}{2}$. It remains to check the case of mixed clauses $\text{IC} = p_i \vee \neg p_j$. A paradoxical profile for the majority rule with respect to this integrity constraint features a first majority of individuals rejecting issue i and a second majority of individuals supporting issue j. By the pigeonhole principle these two majorities must have a non-empty intersection, i.e., there exists one individual that rejects issue i and accepts issue j. This is incompatible with the requirement that all individual ballots satisfy IC, therefore the majority rule is collectively rational with respect to every 2-clause, and by Lemma 4.1.3 with respect to the full language of *2-clauses*. □

An easy corollary of this proposition covers the case of just 2 issues:

Corollary 4.4.6. *If $|\mathcal{I}| \leqslant 2$, then the majority rule is in $\mathcal{CR}[\mathcal{L}_{PS}]$.*

Proof. This follows immediately from Proposition 4.4.5 and from the observation that every given IC for two issues can be put in conjunctive normal form, and since there are only two propositional symbols this formula must consist of conjunctions of clauses of size at most 2 (recall that repetitions of the same atom in a clause do not count). □

As we have remarked in Section 3.4, all classical paradoxes involving the majority rule can be formalised in our framework by means of an integrity constraint that consists of (or is equivalent to) one or more clauses of size bigger than two. We now generalise this observation to a theorem that completes the characterisation of the integrity constraints lifted by the majority rule, proving that these are all and only those formulas that are expressible as conjunctions of 2-clauses. We need some preliminary definitions and a lemma.

Let *minimally falsifying partial assignment* (mifap-assignment) for an integrity constraint IC be an assignment to some of the propositional variables that cannot be extended to a satisfying assignment, although each of its proper subsets can. We first prove a a crucial lemma about mifap-assignments. Given a propositional formula φ, associate with each mifap-assignment ρ for φ a conjunction $C_\rho = \ell_1 \wedge \cdots \wedge \ell_k$, where $\ell_i = p_i$ if $\rho(p_i) = 1$ and $\ell_i = \neg p_i$ if $\rho(p_i) = 0$ for all propositional symbols p_i on which ρ is defined. The conjunction C_ρ represents the mifap-assignment ρ and it is clearly inconsistent with φ.[11]

Lemma 4.4.7. *Every non-tautological formula φ is equivalent to $(\bigwedge_\rho \neg C_\rho)$ with ρ ranging over all mifap-assignments of φ.*

Proof. Let A be a total assignment for φ. Suppose $A \not\models \varphi$, i.e., A is a falsifying assignment for φ. Since φ is not a tautology there exists at least one such A. By sequentially deleting propositional symbols from the domain of A we eventually find a mifap-assignment ρ_A for φ included in A. Hence, A falsifies the conjunct associated with ρ_A, and thus the whole formula $(\bigwedge_\rho \neg C_\rho)$.

Assume now $A \models \varphi$ but $A \not\models (\bigwedge_\rho \neg C_\rho)$. Then there exists a ρ such that $A \models C_\rho$. This implies that $\rho \subseteq A$, as C_ρ is a conjunction. Since ρ is a mifap-assignment for φ, i.e., it is a falsifying assignment for φ, this contradicts the assumption that $A \models \varphi$. □

We are now able to provide a full characterisation of the set of integrity constraints that are lifted by the majority rule in case the set of individuals is odd. Recall from Definition 4.1.10 that $\mathcal{LF}[F]$ is the set of integrity constraints lifted by F.

Theorem 4.4.8. $\mathcal{LF}[Maj] =$ *2-clauses.*

Proof. One direction is entailed by Proposition 4.4.5: the majority rule is CR with respect to conjunctions of 2-clauses.

For the opposite direction assume that IC $\notin$ *2-clauses*, i.e., that IC is not equivalent to a conjunction of 2-clauses. We now build a paradoxical profile for the majority rule. By Lemma 4.4.7 we know that IC is equivalent to the conjunction $\bigwedge_\rho \neg C_\rho$ of all mifap-assignments ρ for IC. We can therefore infer that at least one mifap-assignment ρ^* has size > 2, for otherwise IC would be equivalent to a conjunction of 2-clauses.

[11]The notion of mifap-assingment corresponds to what are called *minimally inconsistent sets* in the judgment aggregation literature (List and Puppe, 2009). For a detailed discussion of the relationship between binary aggregation and judgment aggregation refer to Chapter 6. Formulas $\neg C_\rho$ associated with mifap-assignments ρ for IC are also known as the *prime implicates* of IC (Marquis, 2000). Lemma 4.4.7 is a reformulation of the known result that a formula is equivalent to the conjunction of its prime implicates.

Consider now the following profile. Let y_1, y_2, y_3 be three propositional variables that are fixed by ρ^*. Assume that there are at least 3 individuals. Let the first individual i_1 accept the issue associated with y_1 if $\rho(y_1) = 0$, and reject it otherwise, i.e., let $b_{1,1} = 1 - \rho^*(y_1)$. Furthermore, let i_1 agree with ρ^* on the remaining propositional variables. By minimality of ρ^*, this partial assignment can be extended to a satisfying assignment for IC, and let B_{i_1} be such an assignment. Repeat the same construction for individual i_2, this time changing the value of ρ^* on y_2 and extending it to a satisfying assignment to obtain B_{i_2}. The same construction for i_3, changing the value of ρ^* on issue y_3 and extending it to a satisfying assignment B_{i_3}. If there are other individuals in $\mathcal{N}$, let individuals i_{3s+1} have the same ballot B_{i_1}, individuals i_{3s+2} ballot B_{i_2} and individuals i_{3s+3} ballot B_{i_3}. The basic profile for 3 issues and 3 individuals is shown in Table 4.1.

	y_1	y_2	y_3
i_1	1-$\rho^*(y_1)$	$\rho^*(y_2)$	$\rho^*(y_3)$
i_2	$\rho^*(y_1)$	1-$\rho^*(y_2)$	$\rho^*(y_3)$
i_3	$\rho^*(y_1)$	$\rho^*(y_2)$	1-$\rho^*(y_3)$
Maj	$\rho^*(y_1)$	$\rho^*(y_2)$	$\rho^*(y_3)$

Table 4.1: A paradoxical profile for the majority rule.

As can be seen from the table, and easily generalised to the case of more than 3 individuals, there is a majority supporting ρ^* on every variable on which ρ^* is defined. Since ρ^* is a mifap-assignment and therefore cannot be extended to an assignment satisfying IC, the majority rule in this profile is not collectively rational with respect to IC. □

This result may be considered a "syntactic counterpart" of a result by Nehring and Puppe (2007), in which it is proven that in the framework of judgment aggregation the majority rule will output a consistent outcome if and only if the set of formulas under consideration satisfies what is called the *median property*, i.e., that no minimally inconsistent subsets of size $\geqslant 3$ can be constructed from such formulas. A more detailed discussion about this correspondence can be found in Section 6.3.4.

Recall that the result of Theorem 4.4.8 does not give rise to a characterisation result, i.e., it does not imply that $\mathcal{CR}$[*2-clauses*] = *Maj* (cf. Proposition 4.1.11). On the contrary, this class includes all generalised dictatorships. We will provide a characterisation of this class inside the class of quota rules in Section 4.4.3.

Even Number of Individuals: Weak Majority and Strict Majority

If the aggregation problem features an even number of individuals the majority rule can take two forms. Recall that the weak majority rule (*W-Maj*) is the uniform quota rule with quota $q = \frac{n}{2}$ while the strict majority rule (*S-Maj*) has quota $q = \frac{n+2}{2}$.

The main difference to the case of an odd number of individuals is that both the weak and the strict majority rule do not satisfy the axioms of domain-neutrality $N^{\mathcal{D}}$.[12] To see this, consider the profile in Table 4.2. The profile features

	j_1	j_2
i_1	0	1
i_2	0	1
i_3	1	0
i_4	1	0
W-Maj	1	1
S-Maj	0	0

Table 4.2: Weak and strict majority.

two issues that are both accepted by two individuals and rejected by exactly the same number. The weak majority rule accepts both issues, since at least half of the individuals agree with them. On the other hand, the strict majority rule rejects both issues, as it requires at least 3 individuals accepting an issue to form a majority. Both cases are in violation of domain neutrality, as the individual ballots on the two issues are symmetric: the first rule favours acceptance, while the second favours rejection.

For what concerns the behaviour of the two procedures with respect to collective rationality, we can prove the following proposition.

Proposition 4.4.9. *W-Maj and S-Maj are CR with respect to $\mathcal{L}^+_\rightarrow$ (i.e., 2-clauses in which one literal is negative and one is positive). W-Maj is CR with respect to 2-pclauses. S-Maj is CR with respect to 2-nclauses.*

Proof. A closer inspection of the proof of Proposition 4.4.5 reveals that the case for mixed clauses is also applicable for an even number of individuals. The same result can also be obtained from Theorem 4.2.5, since both *W-Maj* and *S-Maj* satisfy issue-neutrality and N-monotonicity. The second part of the proof is a direct consequence of Corollary 4.4.4. □

Note that the profile in Table 4.2 is a counterexample to *W-Maj* being CR with respect to a negative 2-clause, i.e., $\neg p_1 \vee \neg p_2$, and a counterexample to *S-Maj* being CR with respect to the positive 2-clause $p_1 \vee p_2$.

Unfortunately, a result analogous to Theorem 4.4.8 for the case of an even number of individuals cannot be proven. We therefore refer to the more general

[12] Since they are both uniform quota rules, they still satisfy I, A, $N^{\mathcal{I}}$ and M^{I}.

result about uniform quota rules that is proven in Section 4.4.3 (Corollary 4.4.11). To have an idea of the behaviour of the weak and strict majority rule on clauses of size bigger than 2, let us consider a positive clause of size 3. While the majority rule generates a paradox for any set of individuals of odd size, when this set is of even cardinality the weak majority generates a paradox if an only if there are no more than 4 individuals. This is a consequence of Corollary 4.4.11 but it can be seen directly with some easy calculation.

4.4.3 General Clauses

In this section we present a general result for the collective rationality of an arbitrary quota rule with respect to an arbitrary k-clause. This result generalises our previous results concerning positive and negative clauses, and clauses of size 2. At the end of the section we prove some conclusive characterisations for collectively rational procedures inside the class of quota rules.

We prove the following general result for arbitrary k-clauses:[13]

Theorem 4.4.10. *A quota rule is CR with respect to an exact k-clause* IC *iff*

$$\sum_{j \text{ negative}} q_j + \sum_{j \text{ positive}} (n - q_j + 1) \quad > \quad n(k-1) \tag{4.1}$$

for issues j that occur positively or negatively in IC*, or $q_j = 0$ for some issue j that occurs positively in* IC*, or $q_j{=}n+1$ for issue j that occurs negatively in* IC.

Proof. The case of quota rules being constant on one of the issues (i.e., the case $q_j = 0, n+1$) is straightforward. We can therefore assume that all quotas are $0 < q_j < n+1$. Suppose now that we can generate a paradoxical profile $\boldsymbol{B}$ for the k-clause IC. The only way to falsify the integrity constraint is to output an assignment $F(\boldsymbol{B})$ that rejects all issues that occur positively in IC and accepts all those that are negative. We therefore concentrate our attention to the subprofile $\boldsymbol{B}^k$ defined by restricting the individual ballots to the k issues occurring in IC.

Since individual ballots are rational, there are at least n "correct" symbols in this subprofile, i.e., a 1 for a positive issue or a 0 for a negative one. We refer to such entries with a C in Table 4.3. We now want to count how many "wrong" symbols are present on this subprofile. As $\boldsymbol{B}$ is a paradoxical profile, all issues that occur negative in IC have to be accepted. Therefore, for each of those issues at least q_j individuals have the wrong symbol, in this case a 1. For the same reason, every issue that occurs positively in IC is rejected, so the profile $\boldsymbol{B}^k$ contains at least $n - q_j + 1$ individuals rejecting such issue. Summing up, since the number of cells in this subprofile is nk, we can generate a paradoxical profile if and only if there are enough cells in $\boldsymbol{B}^k$ to account for the minimal number of

[13]This proposition corresponds to a result proved by Dietrich and List (2007a, Theorem 2c) for quota rules in the framework of judgment aggregation.

	j_1	j_2	...	j_k
i_1	C	W	...	W
i_2	W	W	...	C
⋮	⋮	⋮		⋮
i_3	W	C	...	W
F	W	W	...	W

Table 4.3: A paradoxical profile for general clauses.

correct and wrong symbols to generate a paradox. This turns into the equation $n + \sum_{jpos}(n - q_j + 1) + \sum_{jneg} q_j \leqslant nk$. By taking the contrapositive of the last statement we get Equation 4.1. □

It is easy to see that the special cases of positive and negative clauses, i.e., our Propositions 4.4.1 and 4.4.2, can be obtained as corollaries of the previous result. The special case of *uniform* quota rules is particularly interesting:

Corollary 4.4.11. *A uniform quota rule with $q \neq 0, n+1$ is CR with respect to a k-clause* IC *if and only if*

$$(k_2 - k_1)q > n(k_2 - 1) - k_1 \tag{4.2}$$

where k_1 (k_2, respectively) is the number of positive (negative, respectively) issues in IC.

Our Corollary 4.4.4 can be proven as the special case of $k_1 = k$ and $k_1 = 0$. The other special case of $k_1 = k_2$ lead to a satisfiable equation only in case of $k = 2$, proving two important facts: First, *every* quota rule lifts a 2-clause in which one issue is positive and the other is negative, for the equation in this case is always true (it reduces to $0 > -1$). More importantly, it implies that $\mathcal{CR}[k\text{-}clauses]$ does not contain any uniform quota rule for $k > 3$ when the number of issues is even, since this language includes also k-clauses where exactly half of the issues are negative and half are positive, in which case the equation does not have any solutions.

We are now ready to prove a characterisation analogous to that of Theorem 4.4.8 for uniform (non-constant) quota rules:

Proposition 4.4.12. *Let q be different from $0, n+1$, and let F_q be the corresponding uniform quota rule. Then, $\mathcal{LF}[F_q]$ is the language for integrity constraints generated from all k-clauses that satisfy Equation 4.2.*

Proof. One direction is Corollary 4.4.11. For the other direction we resort to Lemma 4.4.7 again. An integrity constrain IC is equivalent to the conjunction of the negation of its mifap-assignments $\bigwedge_\rho \neg C_\rho$. Therefore, if F_q generates a paradox with IC then there is at least one of the mifap-assignments that is satisfied by the outcome of F_q, whilst being falsified by every individual. Since every $\neg C_\rho$ is a clause, by Corollary 4.4.11 it must not satisfy equation 4.2. Thus, if F_q is not collectively rational with respect to IC then IC is not in the language generated by k-clauses satisfying Equation 4.2. □

In exactly the same way we can prove a characterisation for the set of integrity constraints lifted by general quota rules:

Proposition 4.4.13. *Let q_j be different from $0, n+1$, and let F be a quota rule. Then, $\mathcal{LF}[F]$ is the language for integrity constraints generated from k-clauses satisfying Equation 4.1.*

Using the equations introduced in this section we are able to prove some interesting results about the characterisation of collectively rational procedures *inside* the class of quota rules. We have already seen some partial inclusion in Section 4.4.1 for the language of positive and negative clauses in case the set of individuals is odd, and we can now prove the following:

Proposition 4.4.14. $\mathcal{CR}[\text{2-clauses}] \cap \mathcal{QR} = \{Maj\}$

Proof. Let q_1 and q_2 be two quotas relative to two distinct issues. Since we assume that every 2-clause is lifted, these two quotas satisfy the following system of equations, obtained by instantiating Equation 4.1 to the case of positive, negative, and mixed 2-clauses:

$$\begin{aligned} q_1 + q_2 &< n + 2 \\ q_1 + q_2 &> n \\ q_1 + n - q_2 + 1 &> n \\ q_2 + n - q_1 + 1 &> n \end{aligned}$$

From the first two equations we obtain $q_1 + q_2 = n + 1$, since quotas are integers. From the other equations we obtain $|q_1 - q_2| < 1$, which for integer values entails $q_1 = q_2$. We can then conclude that $q_1 = q_2 = \lceil \frac{n+1}{2} \rceil$, thus obtaining the majority rule. □

We end this section by proving an expected negative result for the characterisation of general languages of clauses inside the class of quota rules:

Proposition 4.4.15. $\mathcal{CR}[\text{k-clauses}] \cap \mathcal{QR} = \emptyset$ *for all $k > 2$.*

Proof. A quota rule in $\mathcal{CR}[\text{k-clauses}]$ must be CR with respect to both positive and negative clauses, therefore both equations in Propositions 4.4.1 and 4.4.2 have to be satisfied. But these are unsolvable for $k > 2$. To see this fact consider the first equation, which forces $\sum q_j < n + k$ for any subset of issues of size k, and the second equation requiring $\sum q_j > (k-1)n$ on the same subsets. These two equations are compatible only if $(k-1)n < n+k$, from which we obtain $n < \frac{k}{k-2}$. This in turn is true only if $n < 3$, against our general assumption that there are at least 3 individuals. □

Observe that the equations involved in this proof do not assume that clauses have size strictly smaller than k, hence this result can be strengthened to the language of clauses of size *exactly* k.

4.5 Conclusions and Related Work

In this chapter we developed a general theory of collective rationality in binary aggregation. We classified integrity constraints into syntactically defined languages, and we investigated the problem of how to guarantee collective rationality with respect to all integrity constraints in a given language by means of classical axiomatic properties (see, e.g., Theorems 4.2.1, 4.2.3 and 4.2.4). We also investigated the related problem of characterising, given a classical axiomatic requirement for aggregation procedures, the set of integrity constraints that are lifted by all procedures satisfying that property (see, e.g., Proposition 4.3.2). In the last part of the chapter we concentrated on quota rules, i.e., procedures defined by means of acceptance quotas for every issue. In particular, we obtained a complete characterisation of the set of integrity constraints lifted by the majority rule and by general quota rules (Theorems 4.4.8 and 4.4.10).

While the framework of binary aggregation is well-known in the literature on Social Choice Theory (cf. Section 2.4), the results presented in this chapter constitute the first systematic study of collective rationality with respect to languages for integrity constraints in this setting. Classical approaches usually concentrate on *Arrovian* aggregation procedures, i.e., procedures that are both unanimous and independent, while most of our characterisation results do not make such restrictive assumptions. While the restriction to Arrovian aggregators is in line with standard assumptions in economics, those assumptions are not always justified in AI applications. In this section we review some of the classical approaches to the problem of collective rationality from the literature on Social Choice Theory and we make some conclusive remarks.

Wilson (1975) has been the first to define and study the framework of binary aggregation, obtaining general characterisation results for independent aggregation procedures that generalise the more famous impossibility theorem by Arrow (1963). As pointed out in Section 2.4, Wilson's notion of responsive aggregator for a family of subsets corresponds to our notion of collective rationality with respect to a language for integrity constraints. Being focused on independent procedures, Wilson characterised classes of collectively rational procedures in terms of the structure of winning coalitions defining those procedures (cf. Proposition 2.3.1).

In Chapter 5 we follow a similar approach by providing a proof of Arrow's Theorem which focuses on linking requirements of collective rationality for preference relations with the structure of winning coalitions that defines an independent procedure.

Dokow and Holzman (2009, 2010a) focused on the similar problem of characterising "impossibility domains", i.e., subsets of the full set of binary ballots $\{0,1\}^{\mathcal{I}}$ on which every independent, unanimous and collectively rational procedure is dictatorial. They represented rationality assumptions directly as sets of feasible binary ballots, and they provided graph-theoretic conditions for such a subset to be an impossibility domain. As we have remarked at the beginning of this section, our work differs in that we do not concentrate on independent aggregation procedures. However, an interesting line for future research is to investigate whether a syntactic counterpart of the properties characterising impossibility domains can be devised.

A similar approach has been taken by Nehring and Puppe (2010), who focused on the study of monotonic and independent procedures. As remarked earlier, our Theorem 4.4.8 can be considered as a syntactic analogue of a result proved by the same authors in earlier work (Nehring and Puppe, 2007), which deals with the characterisation of impossibility domains for a class of procedures including the majority rule.

As we have already remarked in several places throughout the chapter, many of the results proven in Section 4.4 are analogous to those proven by Dietrich and List (2007a) in the framework of judgment aggregation. The use of integrity constraints, however, enables us to obtain results that are more applicable both in a theoretical and a practical way. First, as we argue in Chapter 5 and 6, we can easily reduce possibility and impossibility results in other settings of aggregation to some of the characterisation results presented in this chapter. Second, from a computational perspective the use of formulas to model rationality assumptions, rather than referring to the consistency of judgment sets, leads to problems that are substantially easier to compute (cf. Section 7.5).

All results in this chapter can be generalised to cover the case of an infinite number of issues and an infinite number of individuals (except for those concerning quota rules and the majority rule, whose definitions hinge on the finiteness of the set of individuals). Related work on this topic has been carried out by Herzberg and Eckert (2012), focusing on the study of independent aggregation procedures for infinite electorates.

Together with Chapter 2, this chapter constitutes a first step towards a general application-oriented theory of aggregation, a topic that is crucial to the development of several AI applications and, above all, to the design of multiagent systems. The achievements of the chapter are twofold: First, this work constitutes the first systematic study of collective rationality for non-independent aggregation procedures in binary aggregation. Second, to achieve these results we introduced the novel concept of languages for integrity constraints. By developing a complex theoretical machinery around this notion we were able to build a link between classical axiomatic properties and collective rationality, proving interesting characterisations and laying the basis for further investigations of aggregation theory.

The results proven in the present chapter should not be interpreted as limiting the possibility of consistent aggregation, but rather as specifying for each application at hand the right conditions that make it possible. Perhaps the most intriguing direction for future research is to employ the machinery developed in this chapter to design collectively rational aggregation procedures to tackle complex problems of aggregation that occur in many AI applications. We make some preliminary steps in this direction in Chapter 7.

Chapter 5

Preference Aggregation

Preference aggregation (PA) is one of the central topics of Social Choice Theory. It started from the seminal work of Black (1958) and Arrow (1963), and it rapidly evolved to a full-fledged theory (Arrow et al., 2002). PA studies the problem of how to aggregate the preferences of a number of individuals into a collective preference over a set of alternatives. In Chapter 3 we showed how PA can be embedded into binary aggregation with integrity constraints (BA with IC), and how the Condorcet paradox can be seen as an instance of our general definition of paradox. The focus of this chapter is on theoretical results in PA. We recall the basic definitions of PA and we show how classical and new results can be obtained with a new proof method that makes use of the characterisation results presented in Chapter 4. By translating aggregation problems from PA into BA with IC, we are able to identify the source of impossibilities in a clash between the integrity constraints defining a preference domain, and a list of axiomatic properties.

In Section 5.1 we review the framework of PA, and we recall its translation into BA with IC. We illustrate our proof method in Section 5.2, by proving an impossibility result for quota rules and a possibility result for non-independent procedures. In Section 5.3 we provide an alternative proof of Arrow's Theorem. Although inspired by our new proof method, this proof does not hinge on any characterisation result from Chapter 4, but rather develops its own argument. As corollaries we obtain a characterisation of oligarchies and a version of Arrow's Theorem for complete and transitive preferences. We conclude in Section 5.4 by discussing different approaches that have been put forward in the literature to relate preference aggregation with binary and judgment aggregation.

5.1 The Framework of Preference Aggregation

In this section we review the basic definitions of the framework of PA (Gaertner, 2006; Taylor, 2005) and we recall its translation into BA with IC that was introduced in Section 3.1.2, extending the correspondence to the level of axioms.

5.1.1 Basic Definitions

Let $\mathcal{N}$ be a set of *individuals* expressing preferences over a set $\mathcal{X}$ of *alternatives*. We represent such preferences with a binary relation on $\mathcal{X}$. In this section we concentrate on two ways of representing preferences, linear orders and weak orders. Other options are possible, and we make use of different assumptions in Sections 5.2.2 and 5.3. Recall that a binary relation is a *linear order* if it is irreflexive, transitive and complete. The term aP_ib stands for "individual i strictly prefers alternative a to alternative b". The choice of a linear order P_i for each individual constitutes a *preference profile* $\boldsymbol{P} = (P_1, \ldots, P_n)$. A *weak order* is a binary relation that is reflexive, transitive and complete. We denote weak orders with the letter R, thus aR_ib stands for "individual i weakly prefers a to b" and call $\boldsymbol{R} = (R_1, \ldots, R_n)$ a profile of weak orders. Note that every weak order R induces an irreflexive and transitive binary relation, usually referred to as the strict part of R, and denoted with $R^<$, namely the relation that holds between a and b whenever aRb holds but bRa does not.

If we denote with $\mathcal{L}(\mathcal{X})$ the set of all linear orders on $\mathcal{X}$, then the set of all *profiles* of (linear) preference orders is the set $\mathcal{L}(\mathcal{X})^{\mathcal{N}}$.

Definition 5.1.1. A *social welfare function* (SWF) for $\mathcal{X}$ and $\mathcal{N}$ defined on linear orders is a function $w : \mathcal{L}(\mathcal{X})^{\mathcal{N}} \to \mathcal{L}(\mathcal{X})$.

A SWF associates with every *preference profile* $\boldsymbol{P} = (P_1, \ldots, P_n) \in \mathcal{L}(\mathcal{X})^{\mathcal{N}}$ a linear order $w(\boldsymbol{P})$, which in most interpretations is taken to represent the aggregation of the preferences of the individuals into a "social preference order" over $\mathcal{X}$. The same definition can be given using the set $\mathcal{R}(\mathcal{X})$ of all weak orders over $\mathcal{X}$ as the domain of aggregation, defining a SWF for $\mathcal{N}$ and $\mathcal{X}$ *defined on weak orders* as a function $w : \mathcal{R}(\mathcal{X})^{\mathcal{N}} \to \mathcal{R}(\mathcal{X})$.

It is important to note that in our definition of SWFs there are two hidden conditions that could be stated as axioms, but that we have instead included as an integral part of the formal framework of preference aggregation. The first is usually called *unrestricted* or *universal domain*: it requires a SWF to be defined over *all* preference profiles in $\mathcal{L}(\mathcal{X})^{\mathcal{N}}$. Domain restrictions, such as *single-peaked* preferences (Black, 1958), are the most common escape from Arrow's impossibility theorem (see, e.g., Gaertner, 2001). The second hidden condition is called *collective rationality* by Arrow (1963, Chapter VIII, Section V). It requires the outcome of the aggregation to be a linear (weak, respectively) order, i.e., it requires the outcome to conform to the same rationality constraints as the input received from the individuals.

5.1.2 Axioms

Since the seminal work of Arrow (1963), the literature on preference aggregation has made extensive use of the axiomatic method to classify and study SWFs.

There are several properties that an aggregation mechanism may satisfy, and some of them have been argued to be natural requirements for a SWF. In this section we list some of the most important axioms presented in the literature. For some of the axioms, we use the same terminology as in Section 2.2, since we will prove in Section 5.1.3 that there is a direct correspondence between the two formulations. We start with the three properties that led to the proof of Arrow's Theorem:

Pareto Condition (P): For all profiles $\boldsymbol{P} \in \mathcal{L}(\mathcal{X})^{\mathcal{N}}$, if aP_ib for every individual $i \in \mathcal{N}$, then $a\,w(\boldsymbol{P})\,b$.

Independence of Irrelevant Alternatives (IIA): For all profiles $\boldsymbol{P}$ and $\boldsymbol{P}'$ in $\mathcal{L}(\mathcal{X})^{\mathcal{N}}$, if $aP_ib \Leftrightarrow aP_i'b$ for all $i \in \mathcal{N}$, then $a\,w(\boldsymbol{P})\,b \Leftrightarrow a\,w(\boldsymbol{P}')\,b$.

Non-dictatorship (NDIC): There is no individual $i \in \mathcal{N}$ such that $w(\boldsymbol{P}) = P_i$ for every profile $\boldsymbol{P} \in \mathcal{L}(\mathcal{X})^{\mathcal{N}}$.

The (weak) Pareto condition (also known as unanimity) states that, whenever every individual strictly prefers alternative a to alternative b, so does society. IIA forces the social ranking of two alternatives a and b to depend only on their relative ranking by the individuals. A formulation of these axioms for the case of weak orders can be easily obtained by considering profiles in $\mathcal{R}(\mathcal{X})^{\mathcal{N}}$ rather than in $\mathcal{L}(\mathcal{X})^{\mathcal{N}}$. The only exception is the weak Pareto condition, which is usually stated for the strict order $R^<$ induced by a weak order R.

Weak Pareto Condition (WP): For all profiles $\mathcal{R} \in \mathcal{R}(\mathcal{X})^{\mathcal{N}}$, if $a\,R_i^<\,b$ for every individual $i \in \mathcal{N}$, then $a\,w(\mathcal{R})^<\,b$.

The axioms WP, IIA and NDIC are the most classical set of impossible requirements for SWFs: Arrow's Theorem (1963) states that there is no SWF defined on weak (linear, respectively) orders that satisfies WP (P, respectively), I and NDIC in case there are at least 3 alternatives.

Other axiomatic properties have been proposed in the literature. We state here their formulation for the case of linear orders. We refer to the literature (Gaertner, 2006) for a formulation of these properties in the case of weak orders, in case it cannot be obtained directly from our version. The first one we consider is the axiom of *anonymity* (also known as *equality*, cf. Arrow, 1963):

Anonymity (A): For any profile $\boldsymbol{P} \in \mathcal{L}(\mathcal{X})^{\mathcal{N}}$ and any permutation $\sigma : \mathcal{N} \to \mathcal{N}$, we have that $F(P_1, \ldots, P_n) = F(P_{\sigma(1)}, \ldots, P_{\sigma(n)})$.

Another property is the principle of *neutrality*, i.e., that all alternatives should be treated the same way. This axiom takes different forms in the literature. It is often stated for independent procedures, or it employs permutations, in line with the axiom of anonymity (Arrow, 1963; Taylor, 2005). Here, we state a formulation of this axiom for linear orders which we adapted from Gaertner (2006).

Neutrality: For any four alternatives $a, b, c, d \in \mathcal{X}$ and profile $\boldsymbol{P} \in \mathcal{L}(\mathcal{X})^{\mathcal{N}}$, if for all $i \in \mathcal{N}$ we have that $aP_ib \Leftrightarrow cP_id$, then $a\,w(\boldsymbol{P})\,b \Leftrightarrow c\,w(\boldsymbol{P})\,d$.

We conclude by listing the two following axioms, adapted from the work of May (1952). The first axiom can be clearly recognised as a condition of monotonicity, while the second is a generalisation of the original axiom of neutrality proposed by May (1952). This formulation differs from the previous axiom N, as it requires a SWF to treat preferences on different alternatives *symmetrically*:

Positive Responsiveness: For all $a, b \in \mathcal{X}$ and any two profiles $\boldsymbol{P}$ and $\boldsymbol{P}'$ in $\mathcal{L}(\mathcal{X})^{\mathcal{N}}$, if aP_ib entails aP'_ib for all $i \in \mathcal{N}$, and for some $s \in \mathcal{N}$ we have that bP_sa and aP'_sb, then $a\,w(\boldsymbol{P})\,b$ entails $a\,w(\boldsymbol{P}')\,b$.

May's Neutrality: For all alternatives $a, b, c, d \in \mathcal{X}$ and profile $\boldsymbol{P} \in \mathcal{L}(\mathcal{X})^{\mathcal{N}}$, if for all $i \in \mathcal{N}$ we have that $aP_ib \Leftrightarrow \neg(cP_id)$ then $a\,w(\boldsymbol{P})\,b \Leftrightarrow \neg(d\,w(\boldsymbol{P})\,c)$.

5.1.3 Translation Revisited

We now recall the embedding of PA into BA with IC which was presented in Section 3.1.2. Given a set of alternatives $\mathcal{X}$, we can construct a set of issues $\mathcal{I}_{\mathcal{X}}$ given by all pairs of alternatives and integrity constraints $\text{IC}_<$ ($\text{IC}_\leqslant$, respectively) for the case of linear orders (weak orders), to encode the rationality constraints of preference aggregation. This enabled us to obtain a correspondence between SWFs defined on linear orders and aggregation procedures that are CR with respect to $\text{IC}_<$, and, in the same way, between SWFs defined on weak orders and aggregation procedures that are CR with respect to $\text{IC}_\leqslant$. This correspondence is not a bijection, since every SWF is associated with a set of aggregation procedures, depending on how the function is extended outside the domain defined by the integrity constraint of preferences.

The correspondence extends to axiomatic properties. By substituting the expressions aP_ib or aR_ib with $b_{i,ab} = 1$, $a\,w(\boldsymbol{P})\,b$ with $F(\boldsymbol{B})_{ab} = 1$, and preference profile $\boldsymbol{P}$ with binary profile $\boldsymbol{B}$, we obtain for most of the axioms presented in Section 5.1.2 their equivalent formulation for binary aggregation. The axiom of independence of irrelevant alternatives (IIA) corresponds to the axiom of independence (I) defined in Section 2.2. The Pareto condition corresponds to a weaker version of the axiom of unanimity (U), restricted to the case of individuals agreeing on the acceptance of an issue. For the case of weak orders the resulting axiom is even weaker (but sufficient to obtain deep impossibilities, see Section 5.3). The axiom of non-dictatorship corresponds to the negation of our Definition 2.2.1. The two axioms of anonymity are in direct correspondence, as well as that of neutrality and the axiom $\text{N}^{\mathcal{I}}$. Finally, the axiom of positive responsiveness corresponds to the independence version of M^{I}, and May's neutrality axiom corresponds to the axiom $\text{N}^{\mathcal{D}}$.

This direct correspondence between axiomatic properties enables us to study classes of procedures and to translate results from one framework to the other. For instance, to every anonymous, positively responsive SWF defined on weak orders that satisfies IIA corresponds an aggregation procedure that is CR with respect to $\mathrm{IC}_{\leqslant}$ and that satisfies I, A and M^{I} on $\mathrm{Mod}(\mathrm{IC}_{\leqslant})$. The restriction on $\mathrm{Mod}(\mathrm{IC}_{\leqslant})$ is of crucial importance, as we have no information on the behaviour of F outside the domain defined by $\mathrm{IC}_{\leqslant}$ or $\mathrm{IC}_{<}$.

5.2 (Im)possibility Results From Lifting Results

In this section we prove two results concerning the aggregation of preferences by making use of our characterisation results of Chapter 4, illustrating the proof method we outlined in the introduction: translate a problem from PA to BA with IC, and look for clashes between integrity constraints and axiomatic properties. The first result is a reformulation of a known result by Wilson (1975), while the second result is novel.

5.2.1 An Impossibility Result

It has been argued in several places that independence represents the crucial source of impossibilities in preference aggregation, identifying the problem in a clash between this axiom and the transitivity of collective preference (Saari, 2008). In this section we prove an impossibility result in this spirit, showing how a combination of axioms, including independence, clashes with the requirements of collective rationality for preferences.

Recall that the integrity constraint of transitivity in $\mathrm{IC}_{<}$ can be simplified for 3 alternatives to the conjunction of two positive clauses $p_{ba} \vee p_{cb} \vee p_{ac}$ and $p_{ab} \vee p_{bc} \vee p_{ca}$ (cf. the observation on page 32). Call a SWF *imposed* if for some pair of distinct alternatives a and b we have that a is always collectively preferred to b in every profile.[1] We now prove the following result, which is a weaker version of a result by Wilson (1975):

Proposition 5.2.1. *If $|\mathcal{X}| \geqslant 3$ and $|\mathcal{N}| \geqslant 2$, then any anonymous, independent and positively responsive SWF for $\mathcal{X}$ and $\mathcal{N}$ is imposed.*

Proof. The first step is to move to BA with IC, using the correspondence outlined in Section 5.1.3: To every anonymous, independent and positively responsive SWF corresponds a binary aggregation procedure that is collectively rational for $\mathrm{IC}_{<}$ and that satisfies A, I and M^{I} . Recall that by Proposition 2.3.5, every A, I, M^{I} aggregation procedure is a quota rule.

The second step consists in checking whether the axiomatic requirements clash with the integrity constraint under consideration. In this case the answer is

[1] The negation of this property is known as *citizen sovereignity* (cf. Arrow, 1963).

positive: by exploiting some of the characterisation results proven in Chapter 4 we are able to prove that, if a quota rule is collectively rational for $\mathrm{IC}_<$, then it is imposed, i.e., at least one of the quotas q_{ab} is equal to 0.

Suppose, for the sake of contradiction, that every quota $q_{ab} > 0$. In view of our previous discussion, for any three alternatives $a, b, c \in \mathcal{X}$ the integrity constraints corresponding to transitivity are $p_{ba} \vee p_{cb} \vee p_{ac}$ and $p_{ab} \vee p_{bc} \vee p_{ca}$. These are positive clauses of size 3; thus, by Proposition 4.4.1 we obtain the following inequalities on quotas:

$$q_{ba} + q_{cb} + q_{ac} < n + 3$$
$$q_{ab} + q_{bc} + q_{ca} < n + 3$$

Furthermore, it is easy to see that the requirements of completeness and antisymmetry of a linear order force the quotas to satisfy the following:

$$q_{ab} + q_{ba} = n + 1$$
$$q_{bc} + q_{cb} = n + 1$$
$$q_{ac} + q_{ca} = n + 1$$

Now, adding the two inequalities we obtain that $\sum_{a,b \in \mathcal{X}} q_{ab} < 2n + 6$ and adding the three equalities we obtain $\sum_{a,b \in \mathcal{X}} q_{ab} = 3n + 3$. The two constraints together admit a solution only if $n < 3$. Thus, it remains to analyse the case of 2 individuals; but it is easy to see that our constraints do not admit a solution in positive integers for $n = 2$. This shows that there must be a quota $q_{ab} = 0$ for certain distinct a and b as soon as $n \geqslant 2$; hence, the SWF is imposed. □

5.2.2 A Possibility Result

Let a *pair judgment* be a binary relation over $\mathcal{X}$ that is antisymmetric and complete, i.e., it requires each individual to express a (strict) preference on each pair of alternatives, without assuming any further property. This representation of preferences may be useful in situations in which transitivity constitutes a too strong requirement. Moreover, it may be employed to express preferential dependencies in multi-issue domains (Rossi et al., 2011; Airiau et al., 2011) or, more generally, edges in a directed graph (Endriss and Grandi, 2012).

In line with our definitions of Section 5.1, we can define a SWF on pair judgments as a function that assigns a collective pair judgment to every profile of pair judgments. Let IC_{pair} be the following integrity constraint.

Completeness and antisymmetry: $p_{ab} \leftrightarrow \neg p_{ba}$ for all $a \neq b \in \mathcal{X}$

To every SWF defined on pair judgments corresponds an aggregation procedure that is CR with respect to $\mathrm{IC}_{\mathrm{pair}}$. As for the standard case of linear and weak orders, the correspondence extends to the level of axioms. To the best of our knowledge SWFs defined on pair judgments have not yet been studied in the literature, and we now prove a possibility result concerning this class of procedures.

Proposition 5.2.2. *There exists a SWF defined on pair judgments that satisfies the Pareto condition, May's neutrality, independence and positive responsiveness.*

Proof. Once more, we make use of the proof method outlined in the introduction to the chapter. The first step is to move to the more general framework of BA with IC. By our previous discussion, it is sufficient to show that there exists an aggregation procedure that is CR with respect to IC_{pair} and that satisfies U, $N^{\mathcal{D}}$, I and M^{I} to obtain the required result.

The second step consists in checking whether the axiomatic properties required for the aggregation procedure clash with the integrity constraint that defines the domain, making use of our characterisation results from Chapter 4. In our case we can observe that IC_{pair} is expressed in what we called the XOR-language $\mathcal{L}_{XOR}$ (cf. Section 4.2). By our Theorem 4.2.4, we know that every aggregation procedure that satisfies $N^{\mathcal{D}}$ is CR with respect to any integrity constraint expressed in $\mathcal{L}_{XOR}$. We can therefore conclude that any aggregation procedure which satisfies U, $N^{\mathcal{D}}$, I and M^{I} (e.g., the majority rule) is CR with respect to IC_{pairs}.

The third step consists in translating everything back to our initial framework. Notice that we have actually proven a stronger result: As long as a SWF satisfies May's neutrality (which corresponds to $N^{\mathcal{D}}$), we can focus on the remaining axioms to obtain a SWF defined on pair judgments. We can use in this case the pair-wise majority rule to prove the existence of a SWF defined on pair judgments which satisfies U, $N^{\mathcal{D}}$, I and M^{I} . □

Not only have we obtained a proof of a relatively interesting statement in Proposition 5.2.2, but a closer inspection of the proof revealed that, as long as we include May's neutrality, it is possible to find a SWF defined on pair judgments for many other combinations of axioms. This allows, for instance, for SWFs that give different weights to individuals or alternatives.

5.3 Arrow's Theorem

Arrow's Theorem (1963) is considered one of the cornerstones of Social Choice Theory, with which every new result needs to be compared. In this section, we pick up the challenge providing an alternative proof of Arrow's result in line with the proof method presented in the previous section. While our proof does not stand out in terms of succintness when compared with proofs based on combinatorics (see, e.g., Geanakoplos, 2005), and while it employs known techniques based on the study of winning coalitions (see, e.g., Kirman and Sondermann, 1972), its contribution to the literature can be assessed in two important aspects: First, by referring to a more general framework (BA with IC), it sheds new light on the "source" of Arrow's impossibility, identifying it in a clash between axiomatic requirements and collective rationality with respect to the integrity constraints of preference. Second, the flexibility of our proof method enables us to obtain different versions of Arrow's result, including a characterisation of oligarchies usually attributed to Gibbard (1969), with minor adjustments from the original proof.

We begin by proving the following lemma:

Lemma 5.3.1. *If* $|X| \geqslant 3$*, every unanimous and independent aggregation procedure* F *for* $\mathcal{I}_\mathcal{X}$ *that is CR with respect to transitivity is issue-neutral with respect to non-reflexive issues.*[2]

Proof. Let F be an aggregation procedure for $\mathcal{I}_\mathcal{X}$ that satisfies both U and I. By the representation result in Proposition 2.3.1, F is characterised by a set of winning coalitions $\mathcal{W}_{ab}$ for every issue $ab \in \mathcal{I}_\mathcal{X}$, such that $F(\boldsymbol{B}) = 1$ if and only if $N_j^{\boldsymbol{B}} \in \mathcal{W}_{ab}$. We now prove that the collection of winning coalitions is the same for all (non-reflexive) issues, hence F satisfies $\mathrm{N}^{\mathcal{I}}$.

Note that the $\mathcal{W}_{ab}$ are not empty (due to unanimity). Consider any three alternatives a, b and c, ad let $C \in \mathcal{W}_{ab}$. We will employ collective rationality to show that C must also be a winning coalition for each of the other five issues associated with the three alternatives, namely ba, ac, ca, bc and cb. A simple inductive argument then suffices to show that C will in fact have to be a winning coalition for all (non-reflexive) issues.

Now suppose F is CR with respect to transitivity. Let us first see how to prove that $C \in \mathcal{W}_{ac}$: Consider a scenario in which ab and ac are accepted by the agents in C and only those, and in which bc is accepted by all agents, as described in Figure 5.1.

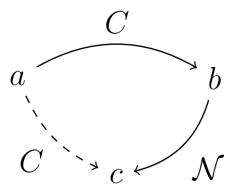

Figure 5.1: Collective transitivity entails issue-neutrality.

By definition of C, ab is collectively accepted, and by unanimity bc is also collectively accepted. Then, by collective transitivity, ac must be collectively accepted. Hence, C is a winning coalition for ac, i.e., $C \in \mathcal{W}_{ac}$. We can use a similar argument for the other edges: e.g., to show $C \in W_{cb}$ consider the case with C

[2]This lemma ceases to hold if we lift the restriction to non-reflexive issues, i.e., issues different from bb with $b \in \mathcal{X}$. However, this restriction suits well to our problem, since we do not want to differentiate the proof between irreflexive and reflexive preferences.

accepting all of ca, ab and cb; then to show $C \in \mathcal{W}_{ba}$ consider the case with C accepting all of bc, ca and ba; and so forth. □

Since transitivity is included in both integrity constraints of preferences $\mathrm{IC}_<$ and $\mathrm{IC}_\leqslant$, it is a straightforward consequence of the previous proof that Lemma 5.3.1 can be extended to these more restrictive constraints, obtaining a proof of what is known in PA as the *"contagion lemma"*:

Lemma 5.3.2. *If* $|X| \geqslant 3$*, every unanimous and independent aggregation procedure* F *for* $\mathcal{I}_\mathcal{X}$ *that is CR with respect to* $\mathrm{IC}_<$ *or* $\mathrm{IC}_\leqslant$ *is issue-neutral with respect to non-reflexive issues.*

We are now ready to state and prove Arrow's Theorem:

Theorem 5.3.3 (Arrow, 1963, weak orders)**.** *Given a finite set of individuals* $\mathcal{N}$ *and a finite set of alternatives* $\mathcal{X}$ *such that* $|\mathcal{X}| \geqslant 3$*, every independent and weakly Paretian SWF for* $\mathcal{N}$ *and* $\mathcal{X}$ *defined on weak orders is dictatorial.*

Proof. Let w be an independent and unanimous SWF for $\mathcal{N}$ and $\mathcal{X}$. By the translation of PA into BA with IC of Section 5.1.3, w corresponds to an aggregation procedure F_w on issues $\mathcal{I}_\mathcal{X}$ that is CR with respect to $\mathrm{IC}_\leqslant$ and satisfies axioms I and a weaker version of U. Closer inspection of the proof of Lemma 5.3.2 shows that it can be proved by weakening the assumption of unanimity to the property corresponding to weak Pareto. Thus, we can assume that F_w is also issue-neutral with respect to non-reflexive issues. Combining this observation with our representation result in Proposition 2.3.1, we can characterise F_w in terms of the set of winning coalitions $\mathcal{W}$. We now prove that $\mathcal{W}$ is an *ultrafilter* (Davey and Priestley, 2002), i.e., a collection of subsets such that: (i) $\emptyset \notin \mathcal{W}$; (ii) if $C_1 \in \mathcal{W}$ and $C_2 \in \mathcal{W}$ then $C_1 \cap C_2 \in \mathcal{W}$ (closed under intersection); (iii) for all $C \subseteq \mathcal{N}$, either C or $\mathcal{N} \setminus C$ is in $\mathcal{W}$ (maximality). The proof is then concluded by observing that an ultrafilter over a finite set is *principal*, i.e., that is defined as those subsets of $\mathcal{N}$ containing a given individual i^*, which is therefore the dictator. Thus, moving back to preference aggregation, we obtain the desired conclusion that every weakly Paretian and independent SWF defined on at least three alternatives is dictatorial.

(i) It is straightforward to observe that the empty set being a winning coalition is in direct contradiction with the weak Pareto condition, therefore $\emptyset \notin \mathcal{W}$.

(ii) In order to prove that $\mathcal{W}$ is closed under intersection, let C_1 and C_2 be two winning coalitions in $\mathcal{W}$ and consider the following profile over three distinct alternatives $a, b, c \in \mathcal{X}$ (recall that we assumed $|\mathcal{X}| \geqslant 3$). Let exactly the individuals in C_1 accept issues ab, exactly the individuals in C_2 accept issue bc, and exactly the individuals in $C_1 \cap C_2$ accept issue ac, as described in the left part of Figure 5.2. By independence, we can ignore the judgments of the individuals on the remaining issues. Since both C_1 and C_2 are winning coalitions, both issues

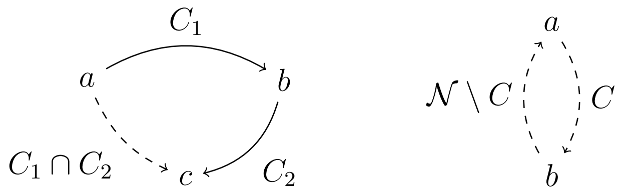

Figure 5.2: The set of winning coalitions is an ultrafilter.

ab and bc are accepted. By collective rationality with respect to transitivity, issue ac also has to be accepted. Therefore, $C_1 \cap C_2$ is a winning coalition in $\mathcal{W}$.

(*iii*) We conclude by proving maximality for $\mathcal{W}$. Let $C \subseteq \mathcal{N}$, and consider a profile in which exactly the individuals in C accept issue ab and exactly the individuals in $\mathcal{N} \setminus C$ accept issue ba, as described in the right part of Figure 5.2. By collective rationality with respect to completeness, either issue ab or issue ba has to be accepted. Thus, at least one of C or its complement $\mathcal{N} \setminus C$ is a winning coalition. □

Notice that in the proof of Theorem 5.3.3 we have not used the assumption of reflexivity of weak orders, and that Lemma 5.3.2 holds for both weak and linear orders. This implies that the same proof holds for the usual statement of Arrow's Theorem for linear orders. Moreover, by using Lemma 5.3.1 in place of Lemma 5.3.2, the same proof shows the following result:

Theorem 5.3.4. *Given a finite set of individuals $\mathcal{N}$ and a finite set of alternatives $\mathcal{X}$ such that $|\mathcal{X}| \geqslant 3$, every independent and unanimous SWF for $\mathcal{N}$ and $\mathcal{X}$ defined on complete and transitive binary relations is dictatorial.*

To illustrate further the flexibility that is brought about by our new proof method, let us prove a result that drops the assumption of completeness from the statement of Theorem 5.3.3. Define a *preorder* as a reflexive and transitive relation. Define $N_{ab} = \{i \in \mathcal{N} \mid aP_ib\}$. A SWF w is called an *oligarchy* if there exists a subset of individuals $A \subseteq \mathcal{N}$ such that for all profiles $\boldsymbol{P}$ we have that $a\, w(\boldsymbol{P})\, b$ if and only if $A \subseteq N^{\boldsymbol{P}}_{ab}$. We now provide an alternative proof of the following result, usually attributed to Gibbard (1969):

Theorem 5.3.5. *Given a finite set of individuals $\mathcal{N}$ and a finite set of alternatives $\mathcal{X}$ such that $|\mathcal{X}| \geqslant 3$, every independent and unanimous SWF for $\mathcal{N}$ and $\mathcal{X}$ defined on preorders is an oligarchy.*

Proof. In the proof of Theorem 5.3.3 we have used the assumption of completeness of a weak order to obtain the proof of maximality of the set of winning coalitions $\mathcal{W}$. Therefore, the first two conditions on the set of winning coalitions (i.e., $\emptyset \notin \mathcal{W}$ and closure under finite intersections) are still satisfied. We need to prove two simple properties to obtain our conclusion. First, we observe that $\mathcal{W}$ is non-empty, since by weak Pareto it contains the full set $\mathcal{N}$. Second, we prove that $\mathcal{W}$ is closed under supersets, i.e., if $C \in \mathcal{W}$ then $C \subseteq D$ implies $D \in \mathcal{W}$. In this case $\mathcal{W}$ is called a *filter* (Davey and Priestley, 2002). Let therefore $C \in \mathcal{W}$ and $C \subseteq D$. Construct a profile $\boldsymbol{B}$ in which issue ab is accepted by exactly the individuals in C, issue bc by all individuals, and issue ac by exactly the individuals in D, as described in Figure 5.3. Since $C \in \mathcal{W}$ and F is unanimous, both issues ab and bc are accepted, and by CR with respect to transitivity we obtain that also issue ac is accepted, and thus $D \in \mathcal{W}$.

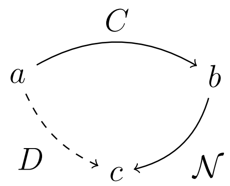

Figure 5.3: The set of winning coalitions is a filter.

We can now conclude the proof by observing that every filter over a finite set is defined as the set of $C \subseteq \mathcal{N}$ such that $A \subseteq C$ for a certain $A \subseteq \mathcal{N}$. To see this, it is sufficient to take the intersection of all winning coalitions (which is non-empty by closure under intersection): w is an oligarchy of the individuals in this set. □

A more detailed study of the aggregation of partially ordered preferences has been carried out in recent work by Pini et al. (2009) and by Xia and Conitzer (2011). For a more detailed study of the use of ultrafilters in PA we refer to Daniëls and Pacuit (2009) and Herzberg and Eckert (2012). A reformulation of Arrow's Theorem along the lines of the one presented in this section can be found in our previous work on graph aggregation (Endriss and Grandi, 2012).

5.4 Conclusions and Related Work

There exists a considerable amount of work in the literature exploring the relation between PA and other frameworks of aggregation (List and Pettit, 2004; Dietrich and List, 2007b; Grossi, 2009, 2010; Porello, 2010). Our approach is similar to that of Dietrich and List (2007b). In their work, the authors embed the framework of PA into the framework of judgment aggregation in general logics Dietrich (2007), by using a simple first-order logic of orders. They then obtain Arrow's Theorem as a corollary of a more general result that is proven in judgment aggregation. In a similar way, we have shown in this chapter how to obtain new theorems or new proofs of known results in PA by embedding PA into BA with IC, using a simple propositional language to express the integrity constraints of preference.

This does not entail that the characterisation results proven in our general framework are stronger than the results we obtain as corollaries in PA. This is the argument of the work of Porello (2010), who showed, for the case of judgment aggregation, how Arrow's Theorem and its correspondent in judgment aggregation are in fact equivalent when compared *at the level of PA*. It is likely that a similar result can be proven for our characterisation results in BA with IC. A similar conceptual perspective is the one taken by Grossi (2009, 2010). In these papers the author obtains an embedding of PA into judgment aggregation based on multi-valued logic and, *vice versa*, an embedding of judgment aggregation into PA by considering a particular structure of preferences. The aim of Grossi is more conceptual: what is sought is, on the one hand, a notion of logical consequence that can alone encode the logic behind preferences, and, on the other hand, a representation of the preferences induced by consistent judgment sets of propositional formulas.

Our purpose is different. In this chapter we showed how preference aggregation can be interpreted in BA with IC for several representations of preferences, and we have put forward a new proof method for PA problems that refers to our characterisation results proved for BA with IC. The results we have obtained may share many similarities or even be weaker than known results from the literature on PA, especially for the case of independent aggregation rules. However, the focus is not on the novelty or strength of single results, but on the generality and flexibility of the proof method we put forward. By unifying proofs in aggregation theory we gain a deeper understanding of the common problem behind impossibility results: impossibilities arise from clashes between axiomatic properties and requirements of collective rationality.

Chapter 6

Judgment Aggregation

In this chapter we study the framework of judgment aggregation (JA), a topic that in recent years has received considerable attention both in Social Choice Theory and in Artificial Intelligence. JA studies situations in which a set of agents express their judgments over a set of correlated propositions, and this needs to be aggregated into a collective judgment. Such a process may lead to paradoxical situations, as seen in Section 3.2, and in this chapter we investigate the source of such paradoxes proving a series of results in JA, inspired by our findings in binary aggregation with integrity constraints (BA with IC).

We define a new problem in the study of JA procedures called the *safety of the agenda*: Given a class of procedures defined axiomatically, we seek necessary and sufficient conditions on the set of propositions under consideration (i.e., the agenda) to avoid paradoxical situations. Once we have studied this problem for some classes of procedures defined axiomatically, we turn to identify the computational complexity of recognising safe agendas for such classes. We show that this problem is highly intractable, proving completeness for the complexity class Π_2^p.

In Section 6.1 we give the basic definitions of the framework of JA we shall be working with, and we list axiomatic properties for the study of judgment aggregation procedures. The relation between the framework of JA and BA with IC is investigated in detail in Section 6.2, showing that the two frameworks are equivalent in terms of expressive power. Section 6.3 presents the problem of safety of the agenda, and contains several safety results for classes of procedures defined axiomatically. We also compare our findings with characterisation results from Chapter 4, providing another example of the generality of the proof method for (im)possibility results described in the introduction and used in Chapter 5 to obtain several results in preference aggregation. In Section 6.4 we study the computational complexity of recognising safe agendas for such classes for which a safety result was proven. Section 6.5 concludes.

6.1 The Framework of Judgment Aggregation

In this section, we give precise definitions for the framework of JA we shall be working with for the rest of the chapter, giving particular attention to computational problems arising from the use of logical formulas as objects of aggregation. We introduce some new terminology to shed light on the difference between the "syntactic" and "logical" properties of a judgment set, a difference that we believe is worth stressing. We also introduce a list of axioms specifying desirable properties for a judgment aggregation procedure. All our definitions are closely related to existing ones, resulting in a framework that is essentially equivalent to the version given by List and Puppe (2009). We shall refer to the setting defined in this section as the *formula-based framework* for JA.

6.1.1 Basic Definitions

Let PS be a set of propositional variables, and $\mathcal{L}_{PS}$ the set of propositional formulas built from PS using the usual connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$ (see Appendix A). If α is a propositional formula, define $\sim\alpha$ as the *complement* of α, i.e., $\neg\alpha$ if α is not negated, and as β if $\alpha = \neg\beta$.

Definition 6.1.1. An *agenda* is a finite nonempty set $\Phi \subseteq \mathcal{L}_{PS}$ that does not contain any doubly-negated formulas and that is closed under complementation (i.e., if $\alpha \in \Phi$ then $\sim\alpha \in \Phi$).

In a slight departure from the common definition in the literature (List and Puppe, 2009), note that we *do* allow for tautologies and contradictions in the agenda. Our reason for relaxing the framework in this manner is that one of our interests is to study the computational complexity of JA, and recognising a tautology or a contradiction is itself a computationally intractable problem. The complexity results we will prove in Section 6.4 would however remain unchanged if we added the additional requirement that agendas should not contain tautologies and contradictions.

Definition 6.1.2. A *judgment set* J on agenda Φ is a subset of the agenda $J \subseteq \Phi$.

We call a judgment set J: *complete* if $\alpha \in J$ or $\sim\alpha \in J$ for all $\alpha \in \Phi$; *complement-free*[1] if for all $\alpha \in \Phi$ it is not the case that both α and its complement are in J; and *consistent* if there exists an assignment that makes all formulas in J true.

Denote with $\mathcal{J}(\Phi)$ the set of all complete and consistent subsets of Φ. Given a set $\mathcal{N} = \{1, \dots, n\}$ of n *individuals*, denote with $\boldsymbol{J} = (J_1, \dots, J_n)$ a *profile* of judgment sets, one for each individual.

[1] This property is called *weak consistency* by Dietrich (2007), and *consistency* by List and Pettit (2002). Our choice of terminology is intended to stress the fact that it is a purely syntactic notion, not involving any model-theoretic concept.

Definition 6.1.3. A *judgment aggregation procedure* (JA procedure) for agenda Φ and a set of individuals $\mathcal{N}$ is a function $F : \mathcal{J}(\Phi)^{\mathcal{N}} \to 2^{\Phi}$.

That is, a JA procedure maps any profile of individual judgment sets to a single collective judgment set (an element of the powerset of Φ). Since F is defined on the set of all profiles of consistent and complete judgment sets, we are already assuming a *universal domain*, which is sometimes stated as a separate property (List and Pettit, 2002). The definition also includes a condition of *individual rationality:* all individual judgment sets are complete and consistent.

6.1.2 Axiomatic Properties

In analogy with Section 2.2, we now list the most important axioms and properties that have been introduced in the literature for JA procedures. We will use the same terminology and letters of Section 2.2 to denote axiomatic properties for JA procedures. This should create no confusion since, as we will prove in Section 6.2, there is direct correspondence between the two formulations.

In Definition 6.1.3 we did not put any constraints on the collective judgment set, the outcome of aggregation. This is the role of the following definition:

Definition 6.1.4. A JA procedure F, defined on an agenda Φ, is said to be:
(i) *complete* if $F(\boldsymbol{J})$ is complete for every $\boldsymbol{J} \in \mathcal{J}(\Phi)^{\mathcal{N}}$;
(ii) *complement-free* if $F(\boldsymbol{J})$ is complement-free for every $\boldsymbol{J} \in \mathcal{J}(\Phi)^{\mathcal{N}}$;
(iii) *consistent* if $F(\boldsymbol{J})$ is consistent for every $\boldsymbol{J} \in \mathcal{J}(\Phi)^{\mathcal{N}}$.

We now present several *axioms* to provide a normative framework in which to state what the desirable properties of an acceptable JA procedure should be. The first axiom is a very basic requirement, restricting possible aggregators F in terms of fundamental properties of the outcomes they produce.

Weak Rationality (WR): F is complete and complement-free.

This condition differs from what is called "collective rationality" in the literature on JA (List and Puppe, 2009), as we do not require the collective judgment set to be consistent. However, we will see that WR corresponds to an instance of our notion of collective rationality for binary aggregation with a suitable integrity constraint (cf. Section 6.3.4). The first reason to separate the notion of consistency from the other conditions is that the requirements of WR are purely syntactic notions that can be checked automatically in an easy way. The second is that the notion of consistency is not intrinsic to the aggregation function, but depends more on the properties of the agenda. This will be made more precise in Section 6.3, where we will study the consistency of a class of aggregators depending on the agenda.

The following are the most important axioms discussed in the literature on JA (List and Pettit, 2002; List and Puppe, 2009; Nehring and Puppe, 2010).

Unanimity (U): If $\varphi \in J_i$ for all $i \in \mathcal{N}$ then $\varphi \in F(\boldsymbol{J})$.

Anonymity (A): For any profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^{\mathcal{N}}$ and any permutation $\sigma : \mathcal{N} \to \mathcal{N}$ we have $F(J_1, \ldots, J_n) = F(J_{\sigma(1)}, \ldots, J_{\sigma(n)})$.

Neutrality (N): For any φ, ψ in the agenda Φ and profile $\boldsymbol{J} \in \mathcal{J}(\Phi)$, if for all $i \in \mathcal{N}$ we have that $\varphi \in J_i \Leftrightarrow \psi \in J_i$, then $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \psi \in F(\boldsymbol{J})$.

Independence (I): For any φ in the agenda Φ and profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$ in $\mathcal{J}(\Phi)$, if $\varphi \in J_i \Leftrightarrow \varphi \in J'_i$ for all $i \in \mathcal{N}$, then $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \varphi \in F(\boldsymbol{J}')$.

Systematicity (S): For any φ, ψ in the agenda Φ and profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$ in $\mathcal{J}(\Phi)$, if $\varphi \in J_i \Leftrightarrow \psi \in J'_i$ for all $i \in \mathcal{N}$, then $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \psi \in F(\boldsymbol{J}')$.

Unanimity expresses the idea that if all individuals accept a given judgment, then so should the collective.[2] Anonymity states that aggregation should be symmetric with respect to individuals, i.e., all individuals should be treated the same. Neutrality is a symmetry requirement for propositions: if the same subgroup accepts two propositions, then either both or neither should be collectively accepted. Independence says that if a proposition is accepted by the same subgroup under two otherwise distinct profiles, then that proposition should be accepted either under both or under neither profile. Systematicity is satisfied if and only if both neutrality and independence are. While all of these axioms are intuitively appealing, they are stronger than they may seem at first, and several impossibility theorems, establishing inconsistencies between certain combinations of axioms with other desiderata, have been proved in the literature. The original impossibility theorem of List and Pettit (2002), for instance, shows that there can be no collectively rational aggregation procedure satisfying A and S.

A further important property is monotonicity. We introduce two different axioms for monotonicity. The first is the one commonly used in the literature (Dietrich and List, 2007a; List and Puppe, 2009). It implicitly relies on the independence axiom. The second, introduced in our previous work (Endriss et al., 2010a), is designed to be applied to neutral procedures. For systematic procedures the two formulations are equivalent.

I-Monotonicity (M^I): For any φ in the agenda Φ and any two profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$ in $\mathcal{J}(\Phi)$, if $\varphi \in J_i$ entails $\varphi \in J'_i$ for all $i \in \mathcal{N}$, and for some $s \in \mathcal{N}$ we have that $\varphi \notin J_s$ and $\varphi \in J'_s$, then $\varphi \in F(\boldsymbol{J})$ entails $\varphi \in F(\boldsymbol{J}')$.

N-Monotonicity (M^N): For any φ, ψ in the agenda Φ and profile $\boldsymbol{J}$ in $\mathcal{J}(\Phi)$, if $\varphi \in J_i \Rightarrow \psi \in J_i$ for all $i \in \mathcal{N}$ and $\varphi \notin J_s$ and $\psi \in J_s$ for some $s \in \mathcal{N}$, then $\varphi \in F(\boldsymbol{J}) \Rightarrow \psi \in F(\boldsymbol{J})$.

[2] As already remarked in Section 2.2, this notion of unanimity is stronger than the standard formulation which can be found in the literature (List and Puppe, 2009), but the two formulations are equivalent under the assumption of independence.

That is, M^I expresses that if φ is collectively accepted and receives additional support (from s), then it should continue to be collectively accepted. Axiom M^N says that if φ is collectively accepted and ψ is accepted by a strict superset of the individuals accepting φ, then ψ should also be collectively accepted.

The axioms we have introduced can be used to define different *classes* of aggregation procedures: Given an agenda Φ and a list of desirable properties AX provided in the form of axioms, we define $\mathcal{F}_\Phi[\text{AX}]$ to be the set of all procedures $F : \mathcal{J}(\Phi)^n \to 2^\Phi$ that satisfy the axioms in AX.

Representation results, like those we proved in Section 2.3, can also be stated for classes of JA procedures. In Section 6.2 we shall extend the correspondence between JA and binary aggregation to the level of axiomatic properties, making it possible to translate representation results from one framework to the other.

6.2 From Judgments to Binary Ballots, and Back

In Section 6.1 we have presented the formula-based framework for JA, and we shall now recall its embedding into BA with IC introduced in Section 3.2.1, extending it to a correspondence between axiomatic properties. We then provide an inverse translation of BA with IC into JA, by allowing the use of constraints in the latter setting. We conclude by discussing the relations between various equivalent frameworks for JA that have been introduced in the literature.

6.2.1 Translation Revisited

Given an agenda Φ, we have shown in Section 3.2 that we can construct an integrity constraint IC_Φ over a set of issues $\mathcal{I} = \Phi$ that encodes the logical correlations between formulas of Φ. The first part of the formula IC_Φ encodes the completeness of a judgment set, while the second part encodes its consistency by ruling out explicitly the acceptance of all formulas in each minimally inconsistent subset of Φ. We were then able to show that there is a correspondence between every complete and consistent JA procedure for Φ and every aggregation procedure that is collectively rational with respect to IC_Φ.[3]

This correspondence can be easily extended to axiomatic properties. It is sufficient to substitute $\boldsymbol{J}$ for $\boldsymbol{B}$ and the expression "$\varphi \in \boldsymbol{J}_i$" for "$b_{i,\varphi} = 1$" in the axioms presented in Section 2.2 to obtain their formulation for JA given in Section 6.1. There is a direct correspondence between the two axioms of independence, between the axiom of neutrality and that of issue-neutrality ($N^\mathcal{I}$) for binary aggregation, between the axioms of anonymity, and between the two axioms of monotonicity M^I and M^N and their counterparts in binary aggregation. The axiom of unanimity in JA is slightly weaker than its BA version, involving

[3]The correspondence is not a bijection, as every JA procedure corresponds to several aggregation procedures, depending on how it is extended to cover the full boolean domain $\mathcal{D}$.

only the acceptance of a formula, but the two formulations are equivalent under the assumption of weak rationality. The axiom of domain-neutrality ($N^{\mathcal{D}}$) does not have any correspondent in the literature on JA.

This correspondence entails, for instance, that every JA procedure for Φ that satisfies systematicity and I-monotonicity corresponds to an aggregation procedure that is collectively rational with respect to IC_Φ and that satisfies I, $N^{\mathcal{I}}$ and M^I on $Mod(IC_\Phi)^{\mathcal{N}}$. Using the notation introduced in previous sections, every procedure in $\mathcal{F}_\Phi[AX]$ corresponds to an aggregation procedure in $\mathcal{F}_{IC_\Phi}[AX]$. The restriction to the domain defined by the integrity constraint IC_Φ is necessary since the binary aggregation procedure obtained by the translation can be extended arbitrarily outside the domain of consistent and complete judgment sets.

6.2.2 From Binary Ballots to Judgments

The formula-based framework for JA that we introduced in Section 6.1 can be expanded to cover the case of more general logical languages (Dietrich, 2007), and to the case of external propositional constraints (Dietrich and List, 2008b). This last case is particularly simple: given an agenda $\Phi \subseteq \mathcal{L}_{PS}$ and a set of *propositional constraints* $\Psi \subseteq \mathcal{L}_{PS}$, a judgment set $J \subseteq \Phi$ is said to be Ψ-consistent if it is consistent *with respect to* all formulas in Ψ (see Appendix A). The remaining definitions of Section 6.1 can be adapted by substituting every occurrence of consistency with Ψ-consistency. The initial JA setting is obtained by considering an empty set of constraints, or $\Psi = \{\top\}$.

Utilising these notions we are able to show that this slightly more general formula-based framework for JA is equivalent to BA with IC. Given a set of issues $\mathcal{I} = \{1, \ldots, m\}$ and an integrity constraint IC, define an agenda $\Phi = \{p_1, \ldots, p_m\}$ and a propositional constraint $\Psi = \{IC\}$. There is a direct correspondence between rational binary ballots over $\mathcal{I}$ and judgment sets over Φ that are Ψ-consistent. Therefore, every BA procedure that is CR with respect to IC corresponds to a complete and Ψ-consistent JA procedure for Φ. The correspondence extends to the level of axioms. We used these observations in our previous work to provide an alternative proof of Theorem 4.4.8 (Grandi and Endriss, 2011).

6.2.3 Related Work in Judgment Aggregation

Other frameworks have been proposed for the study of judgment aggregation from a more abstract perspective, and we already discussed some of them in Section 2.4. In this section, we give a more detailed overview of two such settings, comparing them with the formula-based framework defined so far.

Dokow and Holzman (2010a) study the aggregation of binary ballots on domains $\mathcal{X} \subseteq \{0,1\}^m$. Aggregation procedures are defined as functions $f : \mathcal{X}^{\mathcal{N}} \to \mathcal{X}$, directly including collective rationality in their definition. This framework is equivalent to formula-based JA, as explained in detail by the authors for both truth-functional agendas (i.e., agendas divided into a set of independent premises and a set of conclusions whose truth values can be inferred from the judgments on premises) and the general case (Dokow and Holzman, 2009, 2010a).

Nehring and Puppe (2010) propose another approach, based on their previous work on strategy-proof social choice (Nehring and Puppe, 2007). The authors consider as domain of aggregation a set X equipped with a set of properties $\mathcal{H} \subseteq 2^X$. Formally, they define a *property space* $(X, \mathcal{H})$ as a set X and a collection of subsets $\mathcal{H} \subseteq 2^X$ such that (i) if $H \in \mathcal{H}$ then $H \neq \emptyset$, (ii) if $H \in \mathcal{H}$ then also its complement $X \setminus H \in \mathcal{H}$, (iii) for all $x, y \in X$ there exists an $H \in \mathcal{H}$ such that $x \in H$ but $y \notin H$. There is a natural embedding of X into $\{0,1\}^{|\mathcal{H}|}$, listing for each element the properties that are satisfied by it (marked with a 1) and those that are not (marked with a 0). This observation is used to show the equivalence of this setting with that of binary aggregation and with that of formula-based JA (Nehring and Puppe, 2010, Section 2.1).

All these frameworks for judgment aggregation are equivalent: aggregation procedures, our main object of study, can be transferred from one setting to the other, keeping their axiomatic properties unchanged. We can identify the main difference between these frameworks in the way in which rationality assumptions are represented: by explicitly providing a set of rational ballots (Dokow and Holzman, 2010a); by introducing the concept of property space (Nehring and Puppe, 2007); by referring to the consistency of propositional logic (formula-based JA); by explicitly using a propositional formula (BA with IC).

From a computational perspective, the last two frameworks of BA with IC and formula-based JA are the closest to a possible implementation. The main reason supporting this claim is the compactness of the representation of rationality constraints.[4] The framework of BA with IC has moreover the advantage of being computationally more tractable than the formula-based framework for JA, at least for what concerns some basic problems. The easiest example is the problem of checking the rationality of a given binary ballot, which corresponds in JA to checking the consistency of a given judgment set. While for BA this problem can be solved with model checking in polynomial time, making it a tractable problem, in JA it corresponds to checking the satisfiability of a set of formulas, which is considered intractable (unless P is equal to NP). For a more detailed discussion of the computational complexity of similar problems in these two frameworks we refer to Section 7.5.

6.3 Safety of the Agenda

In this section, we introduce the concept of safety of the agenda (SoA): an agenda is safe for a class of JA procedures, if consistency is guaranteed for every procedure

[4]See also the observation on page 27, regarding the full expressivity of propositional logic with respect to characterising sets of ballots.

in that class. This concept captures a central problem in an application-driven study of the subject. In this section we characterise safe agendas for a number of classes of procedures defined axiomatically, and we relate such findings to the results we obtained in Chapter 4.

6.3.1 Problem Definition

Procedures for judgment aggregation are traditionally studied using the axiomatic method, and results are often negative. The usual finding is that dictatorships are the only JA procedures satisfying a number of appealing conditions (see e.g. List and Pettit, 2002; Pauly and van Hees, 2006; Gärdenfors, 2006; Nehring and Puppe, 2010; Dokow and Holzman, 2010a). An important set of results in the literature on JA are *possibility theorems*, sometimes called "characterisation results" (Nehring and Puppe, 2007; List and Puppe, 2009): Given some axioms as desiderata for the aggregation procedure (always including consistency), the aim of a possibility theorem is to characterise agendas on which such conditions are satisfiable. Despite their theoretical interest, results of this form are somewhat less relevant for applications. The reason is that actual users are more likely to want an assurance that aggregation *will* be safe (provided certain axioms are satisfied and the agenda has certain properties) rather than to learn that there *exists* a safe form of aggregation (satisfying certain axioms). Moreover, in view of the stress we have put on the distinction between "logical" and "syntactic" properties of an aggregation procedure and the collective judgment set it produces, a thorough study of the consistency of a class of procedures depending on the agenda is of immediate relevance. We therefore introduce the following concept:

Definition 6.3.1. An agenda Φ is *safe* with respect to a class of JA procedures $\mathcal{F}$, if every procedure in $\mathcal{F}$ is consistent when applied to judgment sets over Φ.

The example for a *discursive dilemma* presented in Section 3.2 demonstrates the *un*safety of the agenda $\{\alpha, \neg\alpha, \beta, \neg\beta, \alpha\wedge\beta, \neg(\alpha\wedge\beta)\}$ with respect to the majority rule.

6.3.2 Agenda Properties

While possibility theorems address a different issue than the one we are interested in here, some of the properties of agendas defined in that context are still potentially useful for our purposes (List and Puppe, 2009). One of these is the so-called *median property*. Later, we will use this property, and some of its variants, to characterise agendas that are safe for certain classes of aggregation procedures. We call an inconsistent set Δ *nontrivially inconsistent* if it does not contain any contradiction ($\perp \notin \Delta$).

Definition 6.3.2. We say that an agenda Φ satisfies the *median property* (MP), if every nontrivially inconsistent subset of Φ has itself an inconsistent subset of size 2.

The name of this property, introduced by Nehring and Puppe (2007), derives from a property of the set of all judgment sets $\mathcal{J}(\Phi)$ viewed as a subset of a particular vector space. The typical phrasing of this property in the literature is that an agenda satisfies the median property if all minimally inconsistent subsets of Φ have size 2. For agendas without tautologies the two formulations are equivalent. In our case, we have to include an additional check of nontriviality in case there is a contradictory formula in the agenda. We can generalise the median property as follows:

Definition 6.3.3. An agenda Φ satisfies the *k-median property* (kMP) for $k \geqslant 2$, if every inconsistent subset of Φ has itself an inconsistent subset of size at most k.

Observe that we have dropped the restriction to *nontrivially* inconsistent sets in Definition 6.3.3, because for trivially inconsistent sets it is always the case that there is an inconsistent subset of size at most k (namely one of size 1). The MP of Definition 6.3.2 and the 2MP are the same property.

Agendas satisfying the MP are already quite simple, but the restriction can be made tighter by requiring all inconsistent subsets to have a particular form:

Definition 6.3.4. An agenda Φ satisfies the *simplified median property* (SMP), if every nontrivially inconsistent subset of Φ has itself an inconsistent subset of the form $\{\varphi, \psi\}$ with $\models \varphi \leftrightarrow \neg\psi$.

A further simplification yields:

Definition 6.3.5. An agenda Φ satisfies the *syntactic simplified median property* (SSMP), if every nontrivially inconsistent subset of Φ has itself an inconsistent subset of the form $\{\varphi, \neg\varphi\}$.

Agendas satisfying the SSMP are composed of uncorrelated formulas, i.e., they are essentially equivalent to agendas composed of atoms alone. The SMP is less restrictive, allowing for logically equivalent but syntactically different formulas.

Observe that every agenda that satisfies the SMP also satisfies the MP. The converse is not true: $\Phi = \{p, \neg p, p \wedge q, \neg(p \wedge q)\}$ satisfies the MP, but not the SMP. Similarly, every agenda that satisfies the SSMP also satisfies the SMP. Again, the converse is not true: $\Phi = \{p, \neg p, p \wedge p, \neg(p \wedge p)\}$ satisfies the SMP, but not the SSMP.

6.3.3 Linking Agenda Properties and Axioms

We now prove several characterisation results for the safe aggregation of judgments. For several classes of aggregation procedures we will give necessary and sufficient conditions for an agenda to be safe on that class, i.e., all aggregation procedures in that class are consistent when applied to judgment sets over the agenda. We choose to concentrate on classes of procedures defined by weakening the axiomatisation of the majority rule (cf. Section 2.3). The first result is familiar from the literature (Nehring and Puppe, 2007), although it is presented there in a different formulation.

Proposition 6.3.6. *If the number of individuals is odd, then an agenda* Φ *is safe for the majority rule if and only if* Φ *satisfies the* MP.

This proposition is a direct consequence of a result proved by Nehring and Puppe (2007) (see Theorem 3 in the survey by List and Puppe (2009) for a formulation in the framework of JA), which in turn is equivalent to our Theorem 4.4.8. In their work, Nehring and Puppe show that if the number of individuals is odd, under the assumption of collective rationality (WR plus consistency), I-monotonicity, unanimity, systematicity, and anonymity, there exists an aggregation procedure on agenda Φ if and only if Φ satisfies the median property. The witness they give as a consistent aggregator is nothing other than the majority rule. Since Proposition 6.3.6 speaks of a "class" consisting only of a single procedure, namely the majority rule, the concept of safety of the agenda and the kind of concept inherent in a possibility theorem coincide and our result is a direct consequence of theirs. Unfortunately, the same kind of approach cannot be used to adapt other possibility theorems available in the literature, because the classes of procedures we consider in the sequel each contain more than just a single procedure.

The following result shows that the SMP is a necessary and sufficient condition for an agenda to be safe with respect to systematic and anonymous JA procedures. Recall that an agenda satisfies the SMP if all its inconsistent subsets contain two formulas of which the first is equivalent to the negation of the second. We show that this is a necessary and sufficient condition for a systematic and anonymous JA procedure to output a consistent outcome:

Proposition 6.3.7. *An agenda* Φ *is safe for* $\mathcal{F}_\Phi[\mathrm{WR}, \mathrm{A}, \mathrm{S}]$ *if and only if* Φ *satisfies the* SMP *and does not contain a contradictory formula.*[5]

Proof. ($\Leftarrow$) Suppose that Φ satisfies the SMP and suppose, for the sake of contradiction, that there exists a procedure $F \in \mathcal{F}_\Phi[\mathrm{WR}, \mathrm{A}, \mathrm{S}]$ and a profile $\boldsymbol{J}$ such that $F(\boldsymbol{J})$ is inconsistent. Note first that $F(\boldsymbol{J})$ cannot be trivially inconsistent, as we assumed that Φ does not contain a contradictory formula. Therefore, $F(\boldsymbol{J})$ contains a minimally inconsistent subset of size 2 of the form $\{\varphi, \psi\}$ with

[5]The additional requirement of Φ not containing a contradictory formula may be dropped by adding the axiom of unanimity U to the statement. To see this, observe that a unanimous JA procedure cannot accept a contradictory formula, since by individual rationality a contradiction is rejected by all the individuals. Thus, contradictory formulas do not play any role in the safety of an agenda when unanimity is assumed.

$\models \varphi \leftrightarrow \neg\psi$. Now, since every individual judgment set is consistent, we have that $\varphi \in J_i \Leftrightarrow {\sim}\psi \in J_i$ for all $i \in N$, which implies, by neutrality (which follows from systematicity), that $\varphi \in F(\boldsymbol{J}) \Leftrightarrow {\sim}\psi \in F(\boldsymbol{J})$. As we have $\{\varphi, \psi\} \subseteq F(\boldsymbol{J})$, this entails that ${\sim}\psi \in F(\boldsymbol{J})$, which is a contradiction, since the outcome must be complement-free.

($\Rightarrow$) For the other direction, suppose that Φ violates the SMP, i.e., there exists a nontrivially inconsistent subset that does not contain two formulas one of which is equivalent to the negation of the other. This set must contain a minimally inconsistent subset, which we shall call X. In case X has size $\geqslant 3$, we know by Proposition 6.3.6 that the majority rule will generate a discursive dilemma (since the MP is violated). Therefore the agenda is not safe on the class $\mathcal{F}_\Phi$[WR,A,S], which includes the majority rule. In case X has size 2, then it must be of the form $\{\varphi, \psi\}$ with $\varphi \models \neg\psi$ but $\neg\psi \not\models \varphi$. Consider then the following systematic and anonymous aggregation procedure for 3 individuals, defined with the notation used in Proposition 2.3.3, adapted to the JA framework: $h(0) = h(1) = 1$ and $h(2) = h(3) = 0$. That is, F_h accepts a proposition only if it is accepted by 0 or 1 individual. Consider the following profile, restricted to φ and ψ and their complements: $J_1 = \{{\sim}\varphi, {\sim}\psi\}$, $J_2 = \{\varphi, {\sim}\psi\}$, $J_3 = \{{\sim}\varphi, \psi\}$. Note that each of these sets is consistent. This profile (opportunely extended to a profile on the whole agenda) will generate an inconsistent outcome, since both φ and ψ are accepted by only one of the individuals. This proves that when the SMP is violated there always exists a function satisfying WR, S and A that generates an inconsistent outcome. □

With similar arguments we can prove that the SMP also characterises those agendas that are safe for the class of anonymous and neutral JA procedures:

Proposition 6.3.8. *An agenda Φ is safe for $\mathcal{F}_\Phi[\mathrm{WR}, \mathrm{A}, \mathrm{N}]$ if and only if Φ satisfies the* SMP *and does not contain a contradictory formula.*[6]

Proof. To prove this result it is sufficient to observe that the assumption of systematicity in the first part of the proof of Proposition 6.3.7 can be relaxed to neutrality, and, since the class of systematic rules is contained in that of neutral rules, the left-to-right direction of Proposition 6.3.7 entails the analogous direction for Proposition 6.3.8. □

We now prove a more restrictive characterisation result for the class of anonymous and independent JA procedures, showing that the SSMP is a necessary and sufficient condition for an agenda to be safe for this class, i.e., an anonymous and independent JA procedure is guaranteed to output a consistent judgment set if and only if the agenda is composed of uncorrelated formulas:

[6] See footnote 5.

Proposition 6.3.9. *An agenda* Φ *is safe for* $\mathcal{F}_\Phi[\text{WR}, \text{A}, \text{I}]$ *if and only if* Φ *satisfies the* SSMP *and does not contain a contradictory formula.*[7]

Proof. ($\Leftarrow$) Suppose Φ satisfies the SSMP. Let $\boldsymbol{J}$ be a profile such that $F(\boldsymbol{J})$ is inconsistent. Since Φ does not include any contradictory formula we can assume that $F(\boldsymbol{J})$ is nontrivially inconsistent. The SSMP now tells us that there must exist a formula $\varphi \in \Phi$ such that $\{\varphi, \neg\varphi\} \subseteq F(\boldsymbol{J})$, in contradiction with the property of being complement-free. (Note that if an agenda does not contain contradictions and satisfies the SSMP, then *any* weakly-rational procedure is consistent.)

($\Rightarrow$) The fact that Φ does not satisfy the SSMP is equivalent to the existence of two distinct formulas φ and ψ in Φ such that $\varphi \models \psi$. Consider then the constant function that accepts φ and rejects ψ in every profile: this is clearly a weakly-rational, independent, and anonymous function, and it generates for every profile an inconsistent outcome. □

Another class of procedures that has not yet been considered is $\mathcal{F}_\Phi[\text{A,S,M}^\text{I}]$, corresponding to the uniform quota rules (cf. Corollary 2.3.6, adapted to the JA framework). Here, a characterisation result of the kind we seek is available in the literature for certain subclasses of $\mathcal{F}_\Phi[\text{A,S,M}^\text{I}]$, namely uniform quota rules with a specific bound on the quota (Dietrich and List, 2007a). We state this interesting result as follows (recall that n is the number of individuals):

Proposition 6.3.10. *Let* $k \geqslant 2$*. An agenda* Φ *is safe for the class of uniform quota rules* F_q *for* n *individuals satisfying* $q > n - \frac{n}{k}$ *(where* n *is the number of individuals) if and only if* Φ *satisfies the* kMP*.*

Proposition 6.3.10 is a reformulation of Corollary 2(a) in the work of Dietrich and List (2007a), which is in turn equivalent to our Corollary 4.4.11.

6.3.4 Safety of the Agenda in BA with IC

In this section we translate the problem of safety of the agenda into BA with IC and we show how to obtain new proofs and strengthen some of the results proved in previous sections by making use of our characterisation results of Chapter 4. This section constitutes a further illustration of our proof method for (im)possibility results presented in Chapter 5: given an aggregation problem, in this case the safety of an agenda, translate it into BA with IC and look for clashes between the axiomatic requirements and collective rationality.

Let us first take a closer look at the integrity constraint associated with a given agenda Φ. Recall from Section 3.2.1 that IC_Φ consists of a first part given by the completeness requirements $p_\alpha \vee p_{\neg\alpha}$ for $\alpha \in \Phi$, and a second part made of formulas of the form $\neg \bigwedge_{\alpha \in S}(p_\alpha)$ for all $S \subseteq \Phi$ that are minimally inconsistent

[7] See footnote 5.

(mi-sets). Define IC^0_Φ to be the conjunction of the literals $p_\top$ and $\neg p_\bot$ associated with tautologies and contradictions (i.e., mi-sets of size 1) occurring as conjuncts of IC_Φ. Define IC^1_Φ to be the conjunction of $p_\varphi \leftrightarrow \neg p_{\neg\varphi}$ for $\varphi \in \Phi$, combining the formulas of completeness with the simplest mi-subsets of size 2. Finally, define IC^2_Φ to be the remaining part of IC_Φ (concerning mi-sets of size $\geqslant 3$).

As shown by the following lemma, all agenda properties defined in Section 6.3.2 can be characterised in terms of the syntactic properties of the formula IC_Φ associated with it. Recall from Section 4.2 that $\mathcal{L}_\leftrightarrow$ is the language of equivalences.

Lemma 6.3.11. *Given an agenda $\Phi \subseteq \mathcal{L}_{PS}$:*

(*i*) Φ *satisfies the* MP *if and only if* $\mathrm{IC}_\Phi \in$ *2-clauses.*
(*ii*) Φ *satisfies the* kMP *if and only if* $\mathrm{IC}_\Phi \in$ *k-clauses.*
(*iii*) Φ *satisfies the* SMP *if and only if* $\mathrm{IC}^2_\Phi \in \mathcal{L}_\leftrightarrow$.
(*iv*) Φ *satisfies the* SSMP *if and only if* IC^2_Φ *is empty.*

Proof. By definition of IC_Φ, the biggest clause occurring in it has the same size of the maximal mi-set included in Φ (if its size is $\geqslant 2$). Therefore, points (*i*) and (*ii*) are direct consequences of our definitions.

We first prove (*iv*). Recall that an agenda Φ satisfies the SSMP if every mi-subset of Φ is of the form $\{\alpha, \neg\alpha\}$. This corresponds to IC_Φ being equivalent to the conjunction of $p_\alpha \leftrightarrow \neg p_{\neg\alpha}$ for all positive $\alpha \in \Phi$, and to the conjunction of $p_\top$ and $\neg p_\bot$ in case Φ contains tautologies. All these formulas are included in IC^1_Φ and IC^0_Φ; therefore the SSMP is characterised by IC^2_Φ being empty.

For what concerns the characterisation of the SMP expressed in point (*iii*), it is sufficient to recall that this property holds if every mi-subset of Φ is of the form $\{\alpha, \sim\beta\}$ for α logically equivalent to β. Equivalences between formulas are expressed in IC^2_Φ using bi-implications. Thus, the SMP corresponds to adding to IC^0_Φ and IC^1_Φ a set of positive bi-implications $p_\alpha \leftrightarrow p_\beta$ for any equivalent α and β in Φ. □

Using this lemma we are able to obtain a new proof of some of the safety results presented in Section 6.3.3, inspired by our new proof method presented in the introduction. Let us start by giving a new proof of Proposition 6.3.6. We want to prove that, if the number of individuals is odd, an agenda Φ is safe for the majority rule if and only if Φ satisfies the MP. By Definition 6.3.1 and the translation of JA presented in Section 6.2, an agenda Φ is safe with respect to the majority rule if and only if the majority rule is collectively rational with respect to IC_Φ. Our Theorem 4.4.8 fully characterises the set of integrity constraints that are lifted by the majority rule as such formulas that are expressible as conjunctions of clauses of maximal size 2 (i.e., the language of *2-clauses*). It is therefore sufficient to use Lemma 6.3.11 (*i*) to conclude that this holds if and only if the agenda Φ satisfies the MP. With similar arguments we can obtain a proof of the safety result for quota rules proved in Proposition 6.3.10 by referring to our Corollary 4.4.11 for uniform quota rules and using Lemma 6.3.11 (*ii*).

Let us now make one step further, and strengthen our Proposition 6.3.8, making use of the same proof method:

Proposition 6.3.12. *An agenda Φ is safe for $\mathcal{F}_\Phi[\mathrm{WR}, \mathrm{N}]$ if and only if Φ satisfies the* SMP *and does not contain any contradictory formula.*[8]

Proof. By making use of the translation of JA into binary aggregation we know that Φ is safe with respect to complete, complement-free and neutral JA procedures if and only if IC_Φ does not generate a paradox with any issue-neutral procedure for binary aggregation. Let us focus on the structure of IC_Φ. The first part IC^0_Φ is empty, since Φ does not contain contradictions. For what concerns IC^1_Φ, it is easy to see that complete and complement-free procedures are characterised by collectively rational procedures with respect to IC^1_Φ. Therefore, we can concentrate on the remaining condition. We know by Theorem 4.2.3 that an issue-neutral procedure is collectively rational for IC if and only if IC belongs to $\mathcal{L}_\leftrightarrow$. Thus, we can conclude that Φ is safe for the class of neutral an weakly rational procedures if and only if IC^2_Φ is equivalent to a conjunction of equivalences, and by (*iii*) of Lemma 6.3.11, this is equivalent to Φ satisfying the SMP. $\square$

Proposition 6.3.9 can also be strengthened, dropping both the anonymity and the independence assumption, obtaining an interesting characterisation for the axiom of weak rationality.

Proposition 6.3.13. *An agenda Φ is safe for $\mathcal{F}_\Phi[\mathrm{WR}]$ if and only if Φ satisfies the* SSMP *and does not contain any contradictory formula.*[9]

Proof. As remarked in the previous proof, complete and complement-free JA procedures are associated with collectively rational procedures with respect to IC^1_Φ. Thus, using this time point (*iv*) in Lemma 6.3.11, an agenda Φ free from contradictions is safe with respect to complete and complement-free procedures if and only if Φ satisfies the SSMP. $\square$

6.4 Complexity of Safety of the Agenda

In this section, we establish the complexity of deciding whether an agenda satisfies the median property (or one of its variants), and we use these results to show that checking the safety of an agenda is Π^p_2-complete for several classes of aggregators, each characterised by a combination of the most important axioms for JA discussed in the literature.

[8]See footnote 5.
[9]See footnote 5.

6.4.1 Background: Complexity Theory

We shall assume familiarity with the basics of complexity theory up to the notion of NP-completeness. Helpful introductions include the textbooks by Papadimitriou (1994) and Arora and Barak (2009).

We will work with Π_2^p(also known as coNP^{NP} or "coNP with an NP oracle"), a complexity class located at the second level of the polynomial hierarchy. This is the class of decision problems for which a negative answer can be computed in polynomial time by a nondeterministic machine that has access to an oracle for answering queries to SAT (or any other NP-complete problem). To prove a problem Π_2^p-complete, we have to prove both membership in Π_2^p and Π_2^p-hardness. To prove membership, we need to provide an algorithm that, when provided with a *certificate* intended to establish a negative answer, can verify the correctness of that certificate in polynomial time, if we assume that the algorithm has access to a SAT-oracle.

The main challenge is typically to prove hardness. This can be done by giving a polynomial-time reduction from a problem that is already known to be Π_2^p-hard to the problem under investigation. For this purpose, we will make use of *quantified boolean formulas* (QBFs). While QSAT, the satisfiability problem for general QBFs, is PSPACE-complete, by imposing suitable syntactic restrictions we can generate complete problems for any level of the polynomial hierarchy. Consider a QBF of the following form:

$$\forall x_1 \cdots x_r \exists y_1 \cdots y_s . \varphi(x_1, \ldots, x_r, y_1, \ldots, y_s)$$

Here φ is an arbitrary propositional formula and $\{x_1, \ldots, x_r\} \cup \{y_1, \ldots, y_s\}$ is the set of all propositional variables occurring in φ (that is, above could be any QBF for which any existential quantifier occur inside the scope of all universal quantifiers). The problem of checking whether a formula of this form is *satisfiable* (i.e., *true*), which we shall denote $\forall\exists$SAT, is known to be Π_2^p-complete (Stockmeyer, 1976; Arora and Barak, 2009). In the sequel, we shall abbreviate formulas of the above type by writing $\forall \underline{x} \exists \underline{y} . \varphi(\underline{x}, \underline{y})$.

6.4.2 Membership

We shall write MP for the problem of deciding whether a given agenda Φ satisfies the MP, and similarly for the other properties defined in Section 6.3.2.

Lemma 6.4.1. MP*,* SMP*,* SSMP*, and* kMP *are all in* Π_2^p*.*

Proof. We shall present the proof for kMP, which is intuitively the most difficult of the four problems. The proofs for the other three problems are very similar.

We need to give an algorithm that decides the correctness of a certificate for the *violation* of the kMP in polynomial time, assuming it has access to a SAT-oracle. For a given agenda Φ (with $n = |\Phi|$), such a certificate is a set $\Delta \subseteq \Phi$

that (a) needs to be inconsistent and that (b) must *not* have any inconsistent subsets of size $\leqslant k$. Inconsistency of Δ can be checked with a single query to the SAT-oracle. If $n' = |\Delta|$, then there are $\sum_{i=1}^{k} \binom{n'}{i}$ nonempty subsets of Δ, which is polynomial in n' (and thus also in n).[10] Hence, the second condition can be checked by a further polynomial number of queries to the oracle. □

6.4.3 Hardness

To help intuition, observe that, similarly to ∀∃SAT, the median property and its variants ask questions beginning with a universal and ending in an existential quantification (roughly: "for *all* subsets ... there *exists* a subset ..."). To formally prove Π_2^p-hardness, we need to show that, although ∀∃SAT may seem a more general problem, it can be reduced to our seemingly more specific problems.

We first prove a technical lemma. Let ∀∃SAT^2 be the problem of checking whether a QBF of the following form is true, *given that we already know* that (*i*) φ is not a tautology, (*ii*) φ is not a contradiction, and (*iii*) φ is not logically equivalent to a literal:

$$\forall \underline{x} \exists \underline{y}. \varphi(\underline{x}, \underline{y}) \ \wedge \ \forall \underline{x} \exists \underline{y}. \neg\varphi(\underline{x}, \underline{y})$$

Lemma 6.4.2. ∀∃SAT^2 *is* Π_2^p*-hard.*

Proof. By reduction from ∀∃SAT: Given any QBF of the form $\forall \underline{x} \exists \underline{y}. \varphi(\underline{x}, \underline{y})$, we show that checking its satisfiability is equivalent to running ∀∃SAT^2 on $(\varphi \vee a) \wedge b$ with a being universally and b existentially quantified, for two new propositional variables a and b not occurring in φ, i.e., to checking the satisfiability of the formula

$$\forall \underline{x} \forall a \exists \underline{y} \exists b. [(\varphi(\underline{x}, \underline{y}) \vee a) \wedge b] \ \wedge \ \forall \underline{x} \forall a \exists \underline{y} \exists b. \neg[(\varphi(\underline{x}, \underline{y}) \vee a) \wedge b].$$

First, note that $(\varphi \vee a) \wedge b$ cannot be a tautology, a contradiction, or equivalent to a literal, therefore the side constraints specified in the definition of ∀∃SAT^2 are satisfied. Notice then that the first conjunct above is true exactly when the original formula $\forall \underline{x} \exists \underline{y}. \varphi(\underline{x}, \underline{y})$ is true. This is because b can always be set to true, and the original formula has to be true whenever a is set to false (a falls under the scope of a universal quantifier). Therefore, a positive answer to the ∀∃SAT^2 instance entails a positive answer to the original ∀∃SAT instance. The other direction is trivial, for the second of the above conjuncts is always satisfiable (by making b false). □

We are now able to prove Π_2^p-hardness for the SSMP:

Lemma 6.4.3. SSMP *is* Π_2^p*-hard.*

[10] This figure is not polynomial in k, but note that this does not affect the argument, because k is a constant.

Proof. We shall give a reduction from $\forall\exists\text{SAT}^2$ to SSMP; the claim then follows from Lemma 6.4.2. Take any instance of $\forall\exists\text{SAT}^2$, i.e., the question whether $\forall\underline{x}\exists\underline{y}.\varphi(\underline{x},\underline{y}) \wedge \forall\underline{x}\exists\underline{y}.\neg\varphi(\underline{x},\underline{y})$ is true for some φ with $\not\models \varphi$, $\varphi \not\models \bot$, and $\not\models \varphi \leftrightarrow \ell$ for literals ℓ. Suppose $\underline{x} = \langle x_1, \ldots, x_r\rangle$, and define an agenda as follows:[11]

$$\Phi \;=\; \{x_1, \neg x_1, x_2, \neg x_2, \ldots, x_r, \neg x_r, (\varphi \wedge \top), \neg(\varphi \wedge \top)\}$$

We now prove that Φ satisfies the SSMP if and only if the answer to our $\forall\exists\text{SAT}^2$-question is YES. To see this, consider the following facts. First, as φ is neither a tautology nor a contradiction, any inconsistent subset of Φ must be nontrivially inconsistent. Second, by construction of Φ (consisting largely of literals), any inconsistent subset of Φ not including a pair of syntactic complements must include either $(\varphi \wedge \top)$ or $\neg(\varphi \wedge \top)$, as well as a (complement-free) subset of $\{x_1, \neg x_1, \ldots, x_r, \neg x_r\}$. That is, the only way of violating the SSMP is to find a subset of literals from $\{x_1, \neg x_1, \ldots, x_r, \neg x_r\}$ to make true that forces either $(\varphi \wedge \top)$ or $\neg(\varphi \wedge \top)$ to be false. But this is precisely the situation in which our instance of $\forall\exists\text{SAT}^2$ requires a negative answer. For the same reason, suppose the answer to our initial $\forall\exists\text{SAT}^2$-question is NO. This means that we are able to find an assignment ρ for the variables in $\underline{x}$ that makes either φ or $\neg\varphi$ unsatisfiable. Suppose we are in this last case, then we can construct a subset of Φ containing $\neg(\varphi \wedge \top)$, and including a literal x_i if this is set true by the assignment ρ, and $\neg x_i$ otherwise. This is an inconsistent subset of Φ, and since φ is neither a tautology nor a contradiction, this falsifies the SSMP. □

Proving hardness for the SMP works similarly:

Lemma 6.4.4. SMP *is* Π_2^p*-hard.*

Proof. The construction used is the same as for the proof of Lemma 6.4.3. The only additional insight required is the observation that for the special kind of agenda constructed in that proof, the SMP and the SSMP coincide: By excluding formulas φ that are equivalent to literals, we ensure that the agenda Φ constructed in the previous proof does not contain equivalent formulas. □

Finally, for the MP and the kMP we give proofs using reductions from the SSMP:

Lemma 6.4.5. MP *is* Π_2^p*-hard.*

Proof. We shall give a polynomial-time reduction from SSMP, a Π_2^p-complete problem by Lemma 6.4.3, to MP.

Let Φ be an agenda on which we want to test the SSMP, and divide the formulas in Φ into a positive and a negative part: $\Phi = \Phi_+ \cup \{\neg\varphi \mid \varphi \in \Phi_+\}$. Let

[11] Using $(\varphi \wedge \top)$ rather than φ in Φ ensures that the agenda defined does not include doubly-negated formulas.

$\Phi_+ = \{\varphi_1, \ldots, \varphi_m\}$. Now build the set Φ'_+ in the following way: copy all formulas in Φ_+ m times, every time renaming the variables occurring in φ_i, obtaining the set of formulas φ_i^j for $1 \leqslant i, j \leqslant m$. For every i, substitute φ_i^i with $\varphi_i^i \vee p^i$, where p^i is a variable not occurring in any of the φ_i^j. Finally, add $p^1, \ldots, p^m$ to the agenda. We obtain the following set:

$$\begin{array}{rcl} \Phi'_+ & = & \{p^1, \varphi_1^1 \vee p^1, \ldots, \varphi_m^1, \\ & & p^2, \varphi_1^2, \varphi_2^2 \vee p^2, \ldots, \varphi_m^2, \\ & & \vdots \\ & & p^m, \varphi_1^m, \ldots, \varphi_m^m \vee p^m\} \end{array}$$

Define $\Phi' = \Phi'_+ \cup \{\neg\varphi \mid \varphi \in \Phi'_+\}$. We claim that Φ satisfies the SSMP if and only if Φ' satisfies the MP. One direction is easy: if Φ does not satisfy the SSMP, then there exists a minimally inconsistent subset of size $k \geqslant 2$ not containing both a formula and its complement. If this subset is $X = \{\varphi_{i_1}, \ldots, \varphi_{i_k}\}$, then the there exists a subset of Φ', namely $X' = \{\neg p^{i_1}, \varphi_{i_1}^{i_1} \vee p^{i_1}, \varphi_{i_2}^{i_1}, \ldots, \varphi_{i_k}^{i_1}\}$, that is a minimally inconsistent subset of size $k+1 \geqslant 3$, thereby falsifying the MP.

For the opposite direction, suppose that Φ' does not satisfy the MP. That is, there exists a minimally inconsistent subset of size $\geqslant 3$. By construction of Φ', we know that such a subset must only contain formulas with the same superscript or their complements (all other formulas having different variables). If this subset does not contain any p^i or $\neg p^i$, then we can find a copy of it in Φ, which then violates the SSMP. If instead either p^i or $\neg p^i$ is contained in this set for some i, then by minimality also $\varphi_i^i \vee p^i$ or its negation must be included. We can now reason by cases: if both p^i and $\varphi_i^i \vee p^i$ are in the set, then by dropping the disjunction we will still get an inconsistent subset, against the assumption of minimality; $\neg p^i$ and $\neg(\varphi_i^i \vee p^i)$ cannot be in the set for the same reason; finally, p^i together with the negation of $\varphi_i^i \vee p^i$ are already inconsistent. Therefore, we can conclude that all minimally inconsistent subsets that can be built from Φ' are of the form $\{\neg p^i, \varphi_i^i \vee p^i, (\neg)\underline{\varphi}_j^i\}$, where $\underline{\varphi}_j^i$ is a vector of formulas with the same superscript and the prefix $(\neg)$ is intended to indicate that any number of formulas in that vector can be negated. It is now easy to see that $\{\varphi_i, (\neg)\underline{\varphi}_j\}$ is a minimally inconsistent subset of Φ that falsifies the SSMP. □

A similar construction can be done to prove the following:

Lemma 6.4.6. kMP *is* Π_2^p*-hard for every* $k \geqslant 2$.

Proof. As for the previous lemma, we build a reduction from SSMP to kMP by devising a suitable agenda. Let therefore Φ be an agenda on which we want to test the SSMP, and let $\Phi_+ = \{\varphi_1, \ldots, \varphi_m\}$ be the positive formulas in Φ. As in the previous proof, construct Φ'_+ by copying all formulas in Φ_+ m times, and substituting for every i the formula φ_i^i with $\varphi_i^i \vee p^i$. Instead of adding $p^1, \ldots, p^m$ to the new agenda, we add a chain of length $k-1$ of the form $\{p^i_{k-1}, p^i_{k-1} \to p^i_{k-2}, \ldots, p^i_2 \to \neg p^i_1\}$ for every i. Define $\Phi' = \Phi'_+ \cup \{\neg\varphi \mid \varphi \in \Phi'_+\}$. We claim that Φ satisfies the SSMP if and only if Φ' satisfies the kMP. Again, one direction is easy: if Φ does not satisfy the SSMP, then there exists a minimally inconsistent subset of size $k \geqslant 2$ not containing both a formula and its complement. If this subset is $X = \{\varphi_{i_1}, \ldots, \varphi_{i_l}\}$, then construct the following subset $X' = \{p^{i_1}_{k-1}, p^{i_1}_{k-1} \to p^{i_1}_{k-2}, \ldots, p^{i_1}_2 \to \neg p^{i_1}_1, \varphi^{i_1}_{i_1} \vee p^{i_1}, \varphi^{i_1}_{i_2}, \ldots, \varphi^{i_1}_{i_l}\}$. This subset has size $k-1+l \geqslant k+1$ and it is minimally inconsistent. To prove minimality, notice that if we leave out any of the formulas belonging to the chain then by minimality of X there exists an assignment that makes p^{i_1} true as well as all formulas in X other than $\varphi^{i_1}_{i_1}$. Therefore Φ' fails the kMP.

Now for the opposite direction. Suppose that Φ' does not satisfy the kMP. That is, there exists a minimally inconsistent subset of size $\geqslant k+1$. By construction of Φ', we know that such a subset must only contain formulas with the same superscript or their complements. As in the previous proof, we can reason by cases to conclude that either this subset does not contain any p^i or $\neg p^i$, in which case we can find a copy of it in Φ, or it is of the form $X' = \{p^i_{k-1}, p^i_{k-1} \to p^i_{k-2}, \ldots, p^i_2 \to \neg p^i_1, \varphi^i_i \vee p^i, (\neg)\underline{\varphi}^i_j\}$, since including any partial chain of implications would contradict minimality. As in the previous proof, we can now use the set $\{\varphi_i, (\neg)\underline{\varphi}_j\}$ to prove that Φ falsifies the SSMP. □

6.4.4 Summary of Complexity Results

We have shown that deciding whether a given agenda Φ satisfies the MP, the SMP, the SSMP, or the kMP is both in Π^p_2 and Π^p_2-hard. Furthermore, in Section 6.3.3 we have linked these properties to the safety of Φ for various combinations of axioms. As an immediate corollary to all of these results, we obtain our theorem concerning the complexity of SoA:

Theorem 6.4.7. *Checking the safety of an agenda is Π^p_2-complete for any of these classes of aggregation procedures:*

- (i) *the majority rule, corresponding to* $\mathcal{F}_\Phi[\mathrm{WR}, \mathrm{A}, \mathrm{S}, \mathrm{M}^{\mathrm{I}}]$*;*
- (ii) *systematic rules:* $\mathcal{F}_\Phi[\mathrm{WR}, \mathrm{A}, \mathrm{S}]$*;*
- (iii) *neutral rules:* $\mathcal{F}_\Phi[\mathrm{WR}, \mathrm{A}, \mathrm{N}]$ *and* $\mathcal{F}_\Phi[\mathrm{WR}, \mathrm{N}]$*;*
- (iv) *independent rules:* $\mathcal{F}_\Phi[\mathrm{WR}, \mathrm{A}, \mathrm{I}]$*;*
- (v) *weakly rational rules:* $\mathcal{F}_\Phi[\mathrm{WR}]$
- (vi) *for any $k \geqslant 2$, the class of uniform quota rules F_m with $m > n - \frac{n}{k}$, where n is the number of individuals.*

Proof. We first prove Π^p_2-hardness. (i) is a direct consequence of Proposition 6.3.6 and Lemma 6.4.5. In the same way (ii) is derived from Proposition 6.3.7 and Lemma 6.4.4, (iii) from Proposition 6.3.8, Corollary 6.3.12 and Lemma 6.4.4, and (iv) and (v) from Proposition 6.3.9, Corollary 6.3.13 and Lemma 6.4.3. Finally, (*vi*) follows from Proposition 6.3.10 together with Lemma 6.4.6. Membership in Π_2^p follows from Lemma 6.4.1 in all six cases. □

6.5 Conclusions

In this chapter, we have given precise definitions for the formula-based framework for JA and we have given a thorough account of the new problem of safety of the agenda. We have shown how JA and its axiomatic properties can be embedded into BA with IC and, *vice versa*, how BA with IC can be interpreted in a slightly more general framework of JA which allows for integrity constraints (Dietrich and List, 2008a). We then focused on the problem of safety of the agenda. Unlike classical (im)possibility theorems, the problem at hand is how to guarantee that the collective judgment will be consistent, given some axiomatic properties enforced on aggregation procedures. These results share many similarities with our lifting results of Chapter 4, and we investigated this correspondence in detail. The new proof method that we outlined in Chapter 5 has proven useful also in this case, proving it a truly general approach for the understanding of impossibilities in aggregation theory. Finally, we showed how the formula-based framework can be used to study the computational complexity of JA. In this particular case, we have proven that the problem of checking the safety of an agenda is Π_2^p-complete for all the classes of JA procedures we considered.

JA has recently received considerable attention in Artificial Intelligence, especially from the research community of Multiagent Systems and Computational Social Choice (Klamler and Eckert, 2008; Grossi, 2009; Endriss et al., 2010a,b; Nehama, 2010; Lang et al., 2011; Slavkovik and Jamroga, 2012; Slavkovik, 2012). The reason for this trend is clear: in a multiagent system, autonomous agents may have different opinions on the same issues (maybe due to a difference in access to the relevant information, or due to different reasoning capabilities), and some joint course of action needs to be extracted from these diverse views. JA has the potential to provide a formal basis for this kind of collective decision making in multiagent systems. Moreover, JA shares strong links with the framework of abstract argumentation theory (Dung, 1995; Caminada and Pigozzi, 2011), an important research topic in Artificial Intelligence. One of the most important contributions of this chapter is that of providing precise and operational definitions of the formula-based framework for JA that can prepare the ground for a possible implementation. Our complexity results are also pointing in this direction, being one of the first studies of computational complexity in JA (the only other publications to date being the papers by Baumeister et al., 2011; Slavkovik and Jamroga, 2011).

JA has played an important role in several areas of philosophy, such as deliberative democracy (Dryzek and List, 2003), epistemic democracy (List and Goodin, 2001; Dietrich, 2010), and group agency (List and Pettit, 2011; Dietrich and List, 2010). BA with IC constitutes an equivalent framework in which the conditions behind the impossibilities are more explicit. Understanding the consequences of some of our findings (e.g., the interpretation in the debate on group agency of our Theorem 4.2.8, stating the necessity of a generalised dictatorship in order to ensure collective rationality with respect to all integrity constraints) constitutes an exciting direction for future research.

Chapter 7

Collectively Rational Procedures

In this chapter we concentrate on procedures that are especially designed to be collectively rational. In Section 7.1 we extend the basic definition of aggregation procedures to allow for ties in the outcome, and we adapt classical axioms to this case. We introduce the two problems that we shall use to analyse the computational complexity of aggregation procedures. The first problem (winner determination) formalises how difficult it is to compute the collective ballot on a given profile. The second problem is that of strategic manipulation, i.e., deciding whether individuals have incentives to report their truthful judgment or whether they can gain by voting strategically. Since the seminal work of Bartholdi et al. (1989a,b) and Bartholdi and Orlin (1991), both problems are widely employed in the analysis of voting rules (Faliszewski et al., 2009). While a "good" procedure should be easy to use (i.e., winner determination should be performed in polynomial time), it has been argued in the literature that *hardness* results concerning the problem of strategic manipulation may constitute a sufficient shield against its practical occurrence.[1]

In Sections 7.2, 7.3 and 7.4 we introduce three aggregation procedures: the *average voter rule*, the *premise-based procedure* and the *distance-based procedure*. For each of these rules we investigate both axiomatic properties and the computational complexity of the problems of winner determination and strategic manipulation. We prove that the first rule is both easy (polynomial) to use and to manipulate, that the second is hard (NP-complete) to manipulate, while for the last one the problem of winner determination is already very hard (Θ_2^p-complete). We conclude in Section 7.5 by comparing the computational complexity of using the framework of judgment aggregation (JA) and the framework of binary aggregation with integrity constraints (BA with IC).

[1]This opinion has been criticized in several papers, since it is based on a worst-case analysis of the problem (see, e.g., Faliszewski and Procaccia, 2010; Walsh, 2011a).

7.1 Basic Definitions

Throughout this dissertation we did not allow aggregation procedures to output ties in the collective outcome, be it for binary ballots, judgment sets or preferences (cf. Sections 2.1, 6.1, 5.1). However, in many practical cases this problem is hardly avoidable, and it is usually solved by introducing suitable tie-breaking rules. In this section, we provide a definition for irresolute aggregation procedures, i.e., rules that output a subset of collective outcomes, and we adapt the axiomatic properties introduced in Section 2.2 to this more general setting. We then define two classical problems that are central to the computational analysis of aggregation procedures. The first problem of *winner determination* asks how difficult it is to compute the winner of a given profile (*a* winner, in case of irresolute procedures). The second problem is that of *strategic manipulation*. In the context of voting, a player is said to be able to manipulate a voting rule when there exists a situation in which voting in a manner that does not truthfully reflect her preferences will result in an outcome that she prefers to the outcome that would be realised if she were to vote truthfully (Gaertner, 2006). What would constitute an appropriate definition of manipulation in the context of JA or BA with IC is not immediately clear, because in these frameworks there is no notion of preference. However, by designing a suitable notion of "closeness" on binary ballots or judgment sets, it is possible to build a preference ordering starting from the initial ballot or judgment set that an individual submitted. This is the approach followed by Dietrich and List (2007c) for judgment aggregation, and by Everaere et al. (2007) in the related setting of belief merging.

Here, we follow Dietrich and List (2007c) and assume that a player's individual judgment or ballot is also her most preferred outcome and amongst any two outcomes she will prefer the one that is "closer" to that most preferred outcome. We will measure "closeness" using the Hamming distance and we will call an aggregation procedure F *manipulable* if it permits a situation where an agent can change the outcome to get closer to her truthful judgment or ballot by reporting untruthfully.

7.1.1 Axioms for Irresolute Aggregation Procedures

Given a set of issues $\mathcal{I}$ and a set of individuals $\mathcal{N}$, an *irresolute* aggregation procedure is a function $F : \mathcal{D}^{\mathcal{N}} \to 2^{\mathcal{D}}$, associating a subset of $\mathcal{D}$ with every profile of binary ballots in $\mathcal{D} = \{0,1\}^{\mathcal{I}}$. The case of *resolute* aggregation procedures (i.e., Definition 2.1.1) can be viewed as the special case of irresolute procedures that output a set of size one for every profile.

Classical axiomatic properties (cf. Section 2.2) can be adapted to the case of irresolute aggregation procedures. Some axioms, like that of anonymity, do not need major changes in an irresolute framework, unlike those of unanimity and independence for which we provide novel versions.

Since most of the axiomatic properties we considered come in the form of implications, norming the behaviour of an aggregation procedure in case certain conditions occur, there are two natural ways to generalise classical axiomatic properties to deal with sets of possible outcomes. Starting from the same premises as the resolute version of an axiom, the first natural generalisation is to require *all* ballots in the winning set to comply with the conclusion of the resolute version of the axiom. Instead, the second possibility consists in requiring the *existence* of a winning ballot satisfying the conclusion of the resolute version of the axiom. We refer to the two versions as the "strong" and "weak" version of an axiom, even though for some properties (e.g., independence and monotonicity) the strong version does not actually imply the weak one.

In the list that follows, we will mainly choose the existential, i.e., weak, generalisation of resolute axioms, except for the axiom of unanimity for which we find the universal version more appealing. Lang et al. (2011) discuss both generalisations of some of the axioms that follow, for irresolute JA procedures. Recall that $\mathcal{X}$ indicates a subset of $\mathcal{D}^{\mathcal{N}}$.

Anonymity* (A*): For any profile $\boldsymbol{B} \in \mathcal{X}$ and any permutation $\sigma : \mathcal{N} \to \mathcal{N}$, we have that $F(B_1, \dots, B_n) = F(B_{\sigma(1)}, \dots, B_{\sigma(n)})$.

Strong Unanimity* (U*): For any profile $\boldsymbol{B} \in \mathcal{X}$ and any $x \in \{0, 1\}$, if $b_{i,j} = x$ for all $i \in \mathcal{N}$, then $b_j = x$ for all $B \in F(\boldsymbol{B})$.

Weak Independence* (I*): For any issue $j \in \mathcal{I}$, $x \in \{0, 1\}$ and profiles $\boldsymbol{B}, \boldsymbol{B}' \in \mathcal{X}$, if $b_{i,j} = b'_{i,j}$ for all $i \in \mathcal{N}$, then there exists a $B \in F(\boldsymbol{B})$ such that $b_j = x$ iff there exists $B' \in F(\boldsymbol{B}')$ such that $b'_j = x$.

Since all procedures we consider in this chapter will not satisfy the property of independence, we want to provide a formulation of monotonicity which, despite being considerably weaker than the natural adaptation of M^{I} for irresolute procedures, can be satisfied by non-independent procedures. Consider therefore the following notion: if there exists a ballot B in the collective outcome such that $b_j = 1$ and we increase acceptance for issue j while keeping individual judgments about *all other issues fixed*, then there still exists a winning ballot B' such that $b'_j = 1$. Note that in the presence of an integrity constraint, increasing acceptance of an issue keeping everything else fixed is not always possible, therefore the applicability of this property is quite limited. This requirement can be formalised in the following axiom:

Weak Monotonicity* (M*): For any $j \in \mathcal{I}$ and profile $\boldsymbol{B} \in \mathcal{X}$, if $F(\boldsymbol{B})$ contains a ballot B such that $b_j = 1$ and $\boldsymbol{B}'$ is obtained from $\boldsymbol{B}$ by increasing acceptance of issue j keeping everything else fixed, then there also exists a ballot B' in $F(\boldsymbol{B}')$ such that $b'_j = 1$.[2]

[2] A similar property is presented in the context of judgment aggregation by Lang et al. (2011),

Similar definitions can be devised for the framework of JA. An *irresolute JA procedure* is a function that associates with every profile of complete and consistent judgment sets over an agenda Φ a *set* of judgment sets over Φ. Axiomatic properties can be obtained from those we have previously listed for binary aggregation, using the translation between the two frameworks developed in Section 6.2. We state the axiom of unanimity as an example, the others being easily derived in a similar fashion:

Strong Unanimity* (U*): For any profile $\boldsymbol{J} \in \mathcal{J}(\Phi)$ and formula $\varphi \in \Phi$, if $\varphi \in J_i$ for all $i \in \mathcal{N}$, then $\varphi \in J$ for all $J \in F(\boldsymbol{J})$.

7.1.2 Winner Determination

The problem of winner determination in voting theory is that of deciding, given a profile and a candidate, whether the given candidate is one of the election winners. In our setting, binary issues take the role of candidates. The question that winner determination poses for resolute aggregation procedures is therefore whether a given issue is accepted by the collective outcome:

WinDet(F)
Instance: Integrity constraint IC, profile $\boldsymbol{B} \in \text{Mod}(\text{IC})^{\mathcal{N}}$, issue $j \in \mathcal{I}$.
Question: Is it the case that $F(\boldsymbol{B})_j = 1$?

By solving WinDet once for each issue in $\mathcal{I}$, we can compute the collective outcome from an input profile. Note that asking instead whether a given ballot $B^\star$ is equal to $F(\boldsymbol{B})$ does not lead to an appropriate formulation of the winner determination problem, because for actually computing the winner we would then have to solve our decision problem an exponential number of times (one for each possible $B^\star$).

For irresolute aggregation procedures, we need to resort to *partial* binary ballots to obtain the following formulation:

WinDet*(F)
Instance: Integrity constraint IC, profile $\boldsymbol{B} \in \text{Mod}(\text{IC})^{\mathcal{N}}$, subset $I \subseteq \mathcal{I}$, partial ballot $\rho : I \to \{0, 1\}$.
Question: Is there a $B^\star \in F(\boldsymbol{B})$ with $B^\star_j = \rho(j)$ for all $j \in I$?

To see that this is an appropriate formulation for a decision problem corresponding to the task of computing *some* winning set, note that we can compute a winner using a polynomial number of queries to WinDet* as follows. First, ask whether there exists a winning set that accepts a given issue $j_1 \in \mathcal{I}$. In case the

under the name of *insensitivity to reinforcement of collective judgments*. Our property is the existential version of it, for we do not require issue j to be accepted by all outcome ballots but rather by at least one of the collective outcomes.

answer is positive, consider a second issue j_2 and query the problem with a partial ballot ρ accepting both issues in $I = \{j_1, j_2\}$. In case of negative answer, use a different ρ which rejects the first issue and accepts the second one. Continue this process until all issues in $\mathcal{I}$ have been covered.[3]

The same formulation applies to resolute JA procedures, this time considering a formula in the agenda rather than a binary issue:

WinDet(F)
Instance: Agenda Φ, profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^{\mathcal{N}}$, formula $\varphi \in \Phi$.
Question: Is φ an element of $F(\boldsymbol{J})$?

For the case of irresolute JA procedures we can adapt the winner determination problem by using subsets of the agenda in place of partial ballots:

WinDet*(F)
Instance: Agenda Φ, profile $\boldsymbol{J} \in J(\Phi)^{\mathcal{N}}$, subset $L \subseteq \Phi$.
Question: Is there a $J^\star \in F(\boldsymbol{J})$ with $L \subseteq \mathcal{J}^\star$?

7.1.3 Strategic Manipulation

In this section, we define a notion of manipulation for aggregation procedures that is inspired by a more general definition proposed by Dietrich and List (2007c) in the context of judgment aggregation. Their definition is based on the idea that we can induce a preference relation over judgment sets by assuming that an agent's true judgment set J is her most preferred outcome, and between any two outcomes the one that is "closer" to J is preferred. One of the most appealing choices for such a notion of "closeness" is the Hamming distance, which we now define for the case of binary ballots:

Definition 7.1.1 (Hamming distance)**.** Given two binary ballots B and B' in $\mathcal{D}$, the Hamming distance $H(B, B')$ between B and B' is the number of issues on which they differ: $H(B, B') = \sum_{j \in \mathcal{I}} |b_j - b'_j|$.

That is, $H(B, B')$ is an integer between 0 (complete agreement) and $|\mathcal{I}|$ (complete disagreement). For example, if $\mathcal{I} = \{j_1, j_2, j_3\}$, then the Hamming distance between $B = (1, 1, 1)$ and $B' = (0, 1, 0)$ is $H(B, B') = 2$. Intuitively, if B_i is the true ballot of agent i, then i "prefers" B over B' if $H(B_i, B) < H(B_i, B')$.

An aggregation procedure F is said to be manipulable at a given profile $\boldsymbol{B}$, if there exists an alternative ballot $B'_i \in \mathrm{Mod}(\mathrm{IC})$ for some agent $i \in \mathcal{N}$ such that $H(B_i, F(B'_i, \boldsymbol{B}_{-i})) < H(B_i, F(\boldsymbol{B}))$, where $\boldsymbol{B}_{-i}$ is the partial profile obtained by removing B_i from $\boldsymbol{B}$. That is, by reporting B'_i rather than her truthful ballot

[3] In accordance with recent work by Hemaspaandra et al. (2012), we can argue that our formulation of WinDet is the correct decision problem associated with the *search* problem of actually computing a winning binary ballot.

B_i, agent i can achieve the outcome $F(B_i', \boldsymbol{B}_{-i})$ and that outcome is closer (in terms of the Hamming distance) to her truthful (and most preferred) set B_i than the outcome $F(\boldsymbol{B})$ that would get realised if she were to truthfully report B_i. A procedure that is not manipulable at any profile is called *strategy-proof*.

To study the computational complexity of strategic manipulation of a (resolute) aggregation procedure F, we formulate manipulation as a decision problem:

Manip(F)
Instance: Integrity constraint IC, profile $\boldsymbol{B} \in \text{Mod(IC)}^{\mathcal{N}}$, agent $i \in \mathcal{N}$.
Question: Is there a ballot $B_i' \in \text{Mod(IC)}$ such that $H(B_i, F(B_i', \boldsymbol{B}_{-i})) < H(B_i, F(B_i, \boldsymbol{B}_{-i}))$?

Notice that, differently from the previous question of winner determination, we are now asking *whether* an agent can manipulate successfully, rather than *how*. Therefore, this problem does not correspond to the practical (and harder) problem of computing an actual strategy for the manipulator. However, since we are only interested in studying *hardness* results for manipulation (see also the discussion at the beginning of the chapter), we can safely concentrate on this formulation, which provides a lower bound for the corresponding search problem.

The case of irresolute procedures is more complicated to address, since it would require a comparison between two different *sets* of winning ballots. The problem of extending preferences from alternatives to sets of alternatives is the subject of a wide literature in Social Choice Theory (Barberà et al., 2004). Given the high degree of arbitrariness already introduced in defining preferences from ballots, we will now reduce the case of irresolute procedures to the well-defined one for resolute procedures. Therefore, we introduce a given (polynomially computable) tie-breaking rule t and we study, for a given irresolute aggregation procedure F, the associated problem Manip(F^t) where F^t is the composition of F with the given tie breaking rule t.

The same definitions can be adapted to the framework of judgment aggregation. First, let us give a definition of the Hamming distance for judgment sets. Given a judgment set J, we write $J(\varphi) = 1$ if $\varphi \in J$ and $J(\varphi) = 0$ if $\varphi \notin J$. Let Φ^+ be the set of all positive formulas in Φ.

Definition 7.1.2 (Hamming distance)**.** Given an agenda Φ and two complete and complement-free judgment sets $J, J' \in 2^\Phi$, the Hamming distance $H(J, J')$ between J and J' is the number of positive formulas on which they differ: $H(J, J') = \sum_{\varphi \in \Phi^+} |J(\varphi) - J'(\varphi)|$.

That is, $H(J, J')$ is an integer between 0 (complete agreement) and $\frac{|\Phi|}{2}$ (complete disagreement). For example, if the agenda is $\Phi = \{p, \neg p, q, \neg q, p \wedge q, \neg(p \wedge q)\}$, then the Hamming distance between $J = \{\neg p, q, \neg(p \wedge q)\}$ and $J' = \{p, \neg q, \neg(p \wedge q)\}$ is $H(J, J') = 2$. Distances other than H can be employed, and recent work in judgment aggregation has focused on developing notions of distance specifically devised for judgment sets (Duddy and Piggins, 2012).

The problem of manipulability for resolute JA procedures can therefore be stated as follows (the case of irresolute JA procedures is dealt with by considering the composition of F with a given tie-breaking rule t):

Manip(F)
Instance: Agenda Φ, profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^{\mathcal{N}}$, agent $i \in \mathcal{N}$.
Question: Is there a $J_i' \in \mathcal{J}(\Phi)$ s.t. $H(J_i, F(J_i', \boldsymbol{J}_{-i})) < H(J_i, F(J_i, \boldsymbol{J}_{-i}))$?

7.2 The Average Voter Rule

As we argued in Chapter 4, the best aggregation procedures from the point of view of collective rationality are those belonging to the class of generalised dictatorships (Definition 4.2.7). Such procedures copy the ballot of a possibly different individual in every profile, and are thus collectively rational with respect to every integrity constraint. Despite including standard dictatorships, which copy the ballot of the same individual in all profiles, this class also contains interesting novel procedures, that can be obtained by selecting the most representative ballots among those submitted by the individuals.[4]

This section provides the definition of an irresolute generalised dictatorship that associates with every profile of binary ballots the set of those individual ballots that minimise the sum of disagreements with the remaining individuals in the profile. We call this aggregation procedure the *average voter rule* (AVR). The AVR satisfies good axiomatic properties, and we analyse its computational complexity, showing that both problems WinDet and Manip can be solved in polynomial time.

7.2.1 Definition of the Procedure

The average voter rule is defined as follows:[5]

Definition 7.2.1. The *average voter rule* (AVR) chooses those individual ballots that minimise the sum of the Hamming distance to all other individual ballots:

$$\text{AVR}(\boldsymbol{B}) = \underset{\{B_i \mid i \in \mathcal{N}\}}{\text{argmin}} \sum_{s \in \mathcal{N}} H(B_i, B_s),$$

[4]An example of one such rule is found in a science-fiction story by Isaac Asimov entitled "Franchise". In this story, an electronic democracy has been established in the United States, and an enormous computer called "MULTIVAC" is in charge to select a single person as "voter of the year", who would determine the result of the election by answering to a set of questions.

[5]This rule was introduced in our previous work under the name of distance-based generalised dictatorship (Grandi and Endriss, 2011).

Recall that the Hamming distance H has been defined in Definition 7.1.1. The AVR is an irresolute aggregation procedure based on minimisation and it draws inspiration from well-known *distance-based rules* (Miller and Osherson, 2009; Lang et al., 2011). However, to the best of our knowledge this rule has never been defined in the literature on aggregation theory. The crucial difference between the AVR and standard distance-based procedures is that the outcome will always be a set of ballots proposed by the individuals.

7.2.2 Axiomatic Properties

We now turn to the analysis of the axiomatic properties of the AVR. We begin by investigating the validity of those axioms for irresolute procedures which were presented in Section 7.1.1:

Proposition 7.2.2. *The* AVR *satisfies* $U^\star$*,* $A^\star$ *and* $M^\star$*, and it does not satisfy* $I^\star$*.*

Proof. The AVR is clearly anonymous and, as the outcome is composed of individual ballots, it is also unanimous. To see that the AVR satisfies the monotonicity condition $M^\star$, assume that B_i is the ballot of one of the average voters in profile $\boldsymbol{B}$ and that $b_{ij} = 1$, i.e., issue j is accepted. If we now increase acceptance of j by modifying other individual ballots, then the Hamming distance of B_i from other individual ballots decreases, hence leaving B_i in the collective outcome.

To see that the independence condition $I^\star$ does not hold, consider the profile in Table 7.1.

	j_1	j_2	j_3	j_4	j_5
B_1	1	1	0	1	1
B_2	0	1	1	0	1
B_3	1	0	1	1	0
Maj	1	1	1	1	1
AVR	1	1	0	1	1

Table 7.1: Majority differs from AVR

According to the AVR, the first individual ballot is selected as the group outcome: The first ballot differs on three issues from each of the remaining two ballots, having a sum of the Hamming distances of 6. The last two ballots differ on 4 issues from each other, having a total distance of 7. The first ballot is therefore selected by minimisation. If we call $\boldsymbol{B}$ the profile described in Table 7.1, then we have that $\text{AVR}(\boldsymbol{B})_{j_3} = 0$. We now construct a different profile $\boldsymbol{B}'$ such that $\text{AVR}(\boldsymbol{B}')_{j_3} = 1$, while keeping the individual judgments about the same issue j_3 fixed, contradicting the axiom of independence $I^\star$.

Let therefore $\boldsymbol{B}'$ be obtained from $\boldsymbol{B}$ by changing the third individual ballot to $B_3' = (1,1,1,1,1)$. We have that $H(B_1, B_2) = 3$, $H(B_1, B_3') = 1$ and $H(B_2, B_3') = 2$. The third ballot B_3' has now the lowest total Hamming distance with a value of 3, and is the only ballot that gets selected. This contradicts independence $\mathrm{I}^\star$, as there is no ballot in the outcome for $\boldsymbol{B}'$ which rejects the third issue j_3. □

The AVR, if composed with a suitable tie-breaking rule, satisfies other axiomatic properties for resolute aggregation procedures:

Proposition 7.2.3. *The* AVR *satisfies* U, $\mathrm{N}^{\mathcal{I}}$ *and* $\mathrm{N}^{\mathcal{D}}$ *for any choice of tie-breaking rule.*

Proof. Since the AVR is a generalised dictatorship for any choice of tie-breaking rule, it lifts all integrity constraints in $\mathcal{L}_{PS}$. In particular, it lifts the languages of literals, of equivalences and the XOR language $\mathcal{L}_{XOR}$. Therefore, by our characterisation results in Theorems 4.2.1, 4.2.3 and 4.2.4, the AVR belongs to the class $\mathcal{F}[\mathrm{U}, \mathrm{N}^{\mathcal{I}}, \mathrm{N}^{\mathcal{D}}]$. □

Let us move our analysis to the relation between the AVR and the majority rule. Independence is only satisfied by the latter, and even if they share many axiomatic properties, the AVR does not coincide with the majority rule. This can be seen by considering the profile in Table 7.1, in which the outcome of the majority rule is different from all individual ballots.

Despite such differences, it can be observed that if the outcome of the majority rule in a given profile coincides with one of the individual ballots composing such a profile, then the outcome of the AVR coincides with that of the majority rule. To see this, call the outcome of the majority B_{Maj}, and assume that B_{Maj} coincides with one of the individual ballots. Each other ballot that differs in at least one issue with B_{Maj} would differ with a number of individuals that is strictly larger than the majority. Therefore, the individual ballots coinciding with B_{Maj} has the best score for the AVR. This observation suggests another possibility for the definition of generalised dictatorships based on minimisation, and in Section 7.2.5 we define another rule which selects those ballots that minimise the distance from the outcome of the majority rule.

7.2.3 Winner Determination

Winner determination is a tractable problem for the AVR:

Proposition 7.2.4. $\textsc{WinDet}^\star$(AVR) *is in* P.

Proof. Given a profile $\boldsymbol{B}$ and a partial ballot $\rho : I \to \{0,1\}$, the problem of winner determination asks whether there exists a winning ballot $B^\star$ which extends ρ. Computing the Hamming distance between two binary ballots B and B' can be done in polynomial time, thus the score of an individual ballot $\sum_{s\in\mathcal{N}} H(B_i, B_s)$ can also be obtained in a polynomial number of steps. To obtain a polynomial algorithm for the problem of winner determination, it is therefore sufficient to compute the score of each individual ballot in $\boldsymbol{B}$, and then check whether there exists a B_i with minimal score (i.e., a winner) which extends ρ. □

7.2.4 Manipulation

Unfortunately, manipulating the average voting rule is also a tractable problem:

Theorem 7.2.5. Manip(AVR^t) *is in* P.

Proof. Recall the definition of Manip. Let $\boldsymbol{B}_{-i}$ be a partial profile, B_i the truthful ballot of individual i and IC the integrity constraint. We now present a polynomial algorithm to decide whether individual i can gain by reporting untruthfully. Let us first make some assumptions. Suppose that tie-breaking is done alphabetically, i.e., in case of tie B_j wins against B_k if and only if $j < k$. Assume moreover that the winner is the first ballot, i.e., $\text{AVR}^t(\boldsymbol{B}) = B_1$, and that the individual contemplating manipulation is the n-th individual, where $n = |\mathcal{N}|$.

There are two possible strategies for the manipulator. She can either try to change the outcome of the aggregation in favour of another individual, or attempt to win the election herself. In the first case, the best strategy for the manipulator is to copy the ballot of one of the other individuals, and then check whether this results in a better outcome. This can be done in polynomial time: it is sufficient to compute $H(B_n, \text{AVR}^t(\boldsymbol{B}_{-n}, B_j))$ for $j \neq n, 1$ and then confront it with $H(B_n, B_1)$. If for some j the former figure yields smaller number than the latter, then manipulation is possible and the outcome should be YES.

In case the first strategy does not succeed, i.e., either none of the individual ballots can be obtained as an outcome of a manipulated profile, or none of the resulting outcomes is preferred to the truthful winner B_1, then the manipulator's only strategy is to submit a ballot that (a) will be selected as the winner by the AVR^t and (b) is preferred by the manipulator to the truthful winner B_1. We now devise a greedy algorithm to find such a ballot.

Let $w_j = \sum_{i<n} |b_{ij} - b_{nj}|$, be the total disagreement on issue j from n's truthful ballot B_n, and assume that $H(B_1, B_n) = K$. Starting from the issue with the lowest $w_j > 0$, swap one issue at a time in B_n to construct ballot B'_n. At every step, compute the outcome of the aggregation of the new profile. If $\text{AVR}^t(\boldsymbol{B}_{-n}, B'_n) = B'_n$ then there is a manipulation strategy, and thus output YES, otherwise repeat this step swapping the issue with lowest w_j among the remaining ones. Continue for $K - 1$ steps, and if none of the $K - 1$ ballots succeeded in manipulation then output NO.

This greedy algorithm requires an iteration of K polynomial steps, where $K \leqslant |\mathcal{I}|$. Thus, the problem of manipulation of the AVR with tie-breaking can be solved in polynomial time.

Note that the first part of the algorithm, in which the manipulator's effort is to favour one of the other individual ballots, is not necessary since all possible strategies of manipulation will anyway be considered in the second part of the algorithm. The proof of correctness is straightforward. □

Observe that the proof of Theorem 7.2.5 provides an algorithm that also solves the search problem associated to MANIP in polynomial time. Thus, despite being easy to use and enjoying interesting axiomatic properties, the AVR has the drawback of also being easy to manipulate.

7.2.5 The Majority Voter Rule

Minimising the disagreement between individual ballots is not the only possibility of defining meaningful generalised dictatorships. Another rule can be obtained by using the outcome of the majority rule as point of reference, and choose those individual ballots that minimise the amount of disagreements with the outcome of the majority rule:[6]

Definition 7.2.6. The *majority voter rule* (MVR) chooses those individual ballots that minimise the Hamming distance to the outcome of the majority rule:

$$\text{MVR}(\boldsymbol{B}) = \underset{\{B_i | i \in \mathcal{N}\}}{\text{argmin}} \ H(B_i, Maj(\boldsymbol{B}))$$

Despite having a similar definition, the AVR and the MVR do not coincide, as can be shown by considering the profile in Table 7.2.

	j_1	j_2	j_3	j_4	j_5
B_1	1	1	0	0	0
B_2	0	1	1	0	0
B_3	1	0	0	1	0
B_4	0	0	1	1	0
B_5	0	0	0	0	1
Maj	0	0	0	0	0
MVR	0	0	0	0	1

Table 7.2: The AVR differs from the MVR.

In this example, the majority rule yields the outcome $(0, 0, 0, 0, 0)$. The MVR chooses the individual ballot which minimises the Hamming distance from the outcome of the majority rule, therefore $\text{MVR}(\boldsymbol{B}) = B_5$. Instead, the AVR chooses

[6]This definition is inspired by the Endpoint rule introduced by Miller and Osherson (2009), which however is not restricted to choosing the winners in the set of individual ballots.

as winners any of the first 4 individual ballots. To see this, observe that the Hamming distance of B_5 from all other ballots is 3, thus the total score of B_5 is 12, and each of the remaining ballots has a score of 11.

The MVR therefore constitutes another interesting rule that can be created by combining the definition of a distance-based aggregation rule with the notion of generalised dictatorship. Similar definitions can also be given in the framework of judgment aggregation, adding new definitions to the wide variety of rules based on distances that have been introduced in the literature (Pigozzi, 2006; Miller and Osherson, 2009; Lang et al., 2011). In addition to solving the problem of *consistency* by being collectively rational for every integrity constraints, generalised dictatorships also satisfy a strong notion of *compatibility* (Grandi and Pigozzi, 2012), avoiding situations in which the outcome was not voted for by any of the individuals (i.e., they avoid multiple election paradoxes, cf. Section 3.5).

7.3 The Premise-Based Procedure

In this section we move to the realm of JA, presenting one of the first rules that was devised to guarantee consistent aggregation of judgments. There are two basic types of JA procedures that (can be set up so as to) produce consistent outcomes that have been discussed in the JA literature from its very beginnings, namely the *premise-based* and the *conclusion-based* procedure (Kornhauser and Sager, 1993; Dietrich and Mongin, 2010). The basic idea is to divide the agenda into premises and conclusions. In the premise-based procedure, we apply the majority rule to the premises and then infer which conclusions to accept given the collective judgments regarding the premises;[7] under the conclusion-based procedure we directly ask the agents for their judgments on the conclusions and leave the premises unspecified in the collective judgment set. That is, the conclusion-based procedure does not result in complete outcomes, which is why we shall not consider it any further here. The premise-based procedure, on the other hand, can be set up in a way that guarantees consistent and complete outcomes, which provides a usable procedure of some practical interest.

In this section, we first formally introduce the precise variant of the premise-based procedure we shall analyse. We then study the complexity of the winner determination and manipulation problems for this procedure. We end this section by discussing an analogous definition of the premise based procedure in binary aggregation. For ease of exposition, throughout this section we shall assume that the number of agents n is odd.

[7] This is what is commonly understood by "premise-based procedure". Dietrich and Mongin (2010), who call this rule *premise-based majority voting*, have also investigated a more general class of premise-based procedures in which the procedure used to decide upon the premises need not be the majority rule.

7.3.1 Definition of the Procedure

For many JA problems, it may be natural to divide the agenda into premises and conclusions. Let $\Phi = \Phi^p \cup \Phi^c$ be an agenda divided into a set of premises Φ^p and a set of conclusions Φ^c, each of which is closed under complementation.

Definition 7.3.1. The *premise-based procedure* (PBP) for Φ^p and Φ^c is the function mapping each profile $\boldsymbol{J}=(J_1, \ldots, J_n) \in \mathcal{J}(\Phi)^{\mathcal{N}}$ to the following judgment set:

$$\begin{aligned} \text{PBP}(\boldsymbol{J}) &= \Delta \cup \{\varphi \in \Phi^c \mid \Delta \models \varphi\}, \\ &\quad \text{where } \Delta = \{\varphi \in \Phi^p \mid \#\{i \mid \varphi \in J_i\} > \frac{n}{2}\} \end{aligned}$$

If we want to ensure that the PBP always returns judgment sets that are consistent and complete, then we have to impose certain restrictions:

- If we want to guarantee *consistency*, we have to impose restrictions on the premises. In view of our discussion in Section 6.3, in particular our Proposition 6.3.6, the majority rule is guaranteed to be consistent if and only if the agenda Φ satisfies the *median property*, i.e., if every inconsistent subset of Φ has itself an inconsistent subset of size $\leqslant 2$ (cf. Definition 6.3.2). This result immediately transfers to the PBP: it is consistent if and only if the set of premises satisfies the median property.

- If we want to guarantee *completeness*, we have to impose restrictions on the conclusions: for any assignment of truth values to the premises, the truth value of each conclusion has to be fully determined.

Deciding whether a set of formulas satisfies the median property is Π_2^p-hard (see Lemma 6.4.5). That is, in its most general form, deciding whether the PBP can be applied correctly is a highly intractable problem (and, as we shall see, a problem that is most likely considerably harder than either using or manipulating the PBP). For a meaningful analysis, we therefore restrict attention to the following case. First, we assume that the agenda Φ is closed under propositional variables: $p \in \Phi$ for any propositional variable p occurring within any of the formulas in Φ. Second, we equate the set of premises with the set of literals. Clearly, the above-mentioned conditions for consistency and completeness are satisfied.

So, to summarise, the procedure we consider in this section is defined as follows: Under the assumption that the agenda is closed under propositional variables, the PBP accepts a literal ℓ if and only if more individual agents accept ℓ than do accept $\sim\ell$, and the PBP accepts a compound formula if and only if it is entailed by the accepted literals. For consistent and complete profiles, and under the assumption that n is odd, this leads to a resolute JA procedure that is consistent and complete.

7.3.2 Axiomatic Properties

When restricted to the set of premises, the PBP shares all the properties of the majority rule.[8] However, when the whole agenda is considered, the PBP does not satisfy most of the axiomatic properties we have so far encountered for JA procedures (cf. Section 6.1), except for anonymity. It certainly does not satisfy independence, as a change on the individual ballots over premises influences the outcome over the set of conclusions. Moreover, as can be inferred from the example presented in Table 7.3, the PBP is not even unanimous on conclusions.

	p_1	p_2	p_3	$p_1 \vee p_2 \vee p_3$
J_1	1	0	0	1
J_2	0	1	0	1
J_3	0	0	1	1
PBP($\boldsymbol{J}$)	0	0	0	0

Table 7.3: The PBP is not unanimous on conclusions.

Despite all these shortcomings, the PBP is a good example of a consistent and complete judgment aggregation procedure, and constitutes one of the few examples of such a rule that is not defined using a notion of distance.

7.3.3 Winner Determination

Winner determination is a tractable problem for the premise-based procedure:

Proposition 7.3.2. WinDet(PBP) *is in* P.

Proof. Counting the number of agents accepting each of the premises and checking for each premise whether the positive or the negative instance has the majority is easy. This determines the collective judgment set as far as the premises are concerned. Deciding whether a given conclusion should be accepted by the collective now amounts to a model checking problem (is the conclusion φ true in the model induced by the accepted premises/literals?), which can also be done in polynomial time. □

7.3.4 Strategic Manipulation

Manipulating the premise-based procedure, on the other hand, is intractable.

[8] A representation result analogous to Proposition 2.3.7 can be proven in the realm of judgment aggregation.

Theorem 7.3.3. MANIP(PBP) *is* NP-*complete.*

Proof. We first establish NP-membership. An untruthful judgment set J'_i yielding a preferred outcome can serve as a certificate. Checking the validity of such a certificate means checking that (a) J'_i is actually a complete and consistent judgment set and that (b) the outcome produced by J'_i is better than the outcome produced by the truthful set J_i. As for (a), checking completeness is easy. Consistency can also be decided in polynomial time: for every propositional variable p in the agenda, J'_i must include either p or $\neg p$; this admits only a single possible model; all that remains to be done is checking that all compound formulas in J'_i are satisfied by that model. As for (b), we need to compute the outcomes for J_i and J'_i (by Proposition 7.3.2, this is polynomial), compute their Hamming distances from J_i, and compare those two distances.

Next, we prove NP-hardness by reducing SAT to MANIP(PBP). Suppose we are given a propositional formula φ and want to check whether it is satisfiable. We will build a judgment profile for three agents such that the third agent can manipulate the aggregation if and only if φ is satisfiable. Let $p_1, \ldots, p_m$ be the propositional variables occurring in φ, and let q_1, q_2 be two additional propositional variables. Define an agenda Φ that contains all atoms $p_1, \ldots, p_m, q_1, q_2$ and their negation, as well as $m + 2$ syntactic variants of the formula $q_1 \vee (\varphi \wedge q_2)$ and their negation. For instance, if $\psi := q_1 \vee (\varphi \wedge q_2)$, we might use the syntactic variants ψ, $\psi \wedge \top$, $\psi \wedge \top \wedge \top$, and so forth. The judgment profile $\boldsymbol{J}$ is defined in Table 7.4 (the rightmost column has a "weight" of $m + 2$).

	p_1	p_2	$\cdots$	p_m	q_1	q_2	$q_1 \vee (\varphi \wedge q_2)$
J_1	1	1	$\cdots$	1	0	0	?
J_2	0	0	$\cdots$	0	0	1	?
J_3	1	1	$\cdots$	1	1	0	1
$F(\boldsymbol{J})$	1	1	$\cdots$	1	0	0	0

Table 7.4: Manipulation for the PBP is hard.

The judgments of agents 1 and 2 regarding $q_1 \vee (\varphi \wedge q_2)$ are irrelevant for our argument, so they are indicated as "?" in the table (but note that they can be determined in polynomial time; in particular, $J_1(q_1 \vee (\varphi \wedge q_2)) = 0$ for any φ).

If agent 3 reports her judgment set truthfully (as shown in the table), then the Hamming distance between J_3 and the collective judgment set will be $1 + (m + 2) = m + 3$. Note that agent 3 is decisive about all propositional variables (i.e., premises) except q_1 (which will certainly get rejected). Now:

- If φ is satisfiable, then agent 3 can report judgments regarding $p_1, \ldots, p_m$ that correspond to a satisfying assignment for φ. If she furthermore accepts

q_2, then all $m+2$ copies of $q_1 \vee (\varphi \wedge q_2)$ will get accepted in the collective judgment set. Thus, the Hamming distance from J_3 to this new outcome will be at most $m+2$, i.e., agent 3 will have manipulated successfully.

- If φ is not satisfiable, then there is no way to get any of the $m+2$ copies of $q_1 \vee (\varphi \wedge q_2)$ accepted (and q_1 will get rejected in any case). Thus, agent 3 has no means of improving over the Hamming distance of $m+3$ she can guarantee for herself by reporting truthfully.

Hence, φ is satisfiable if and only if agent 3 can manipulate successfully, and our reduction from SAT to MANIP(PBP) is complete. □

Thus, manipulating the PBP is significantly harder than using it, at least in terms of worst-case complexity (and under the assumption that P $\neq$ NP).

7.3.5 PBP in Binary Aggregation

The definition of the premise-based procedure we presented in this section can be adapted to the case of binary aggregation. Let $\mathcal{I}$ be a set of issues divided into a set of premises $\mathcal{I}^p$ and a set of conclusions $\mathcal{I}^c$. If B is a ballot over $\mathcal{I}$, indicate with B^p the restriction of B to $\mathcal{I}^p$ and with B^c the restriction to $\mathcal{I}^c$. Consider the following definition:

Definition 7.3.4. The *premise-based procedure* (PBP) for binary aggregation associates with every profile $\boldsymbol{B}$ over issues $\mathcal{I} = \mathcal{I}^p \cup \mathcal{I}^c$ and integrity constraint IC $\in \mathcal{L}_{PS}$ the following set of collective ballots B over $\mathcal{I}$:

- $\text{PBP}(\boldsymbol{B})^p = \text{AVR}(B_1^p, \ldots, B_n^p)$;
- $\text{PBP}(\boldsymbol{B})^c = \{B \in \{0,1\}^{\mathcal{I}^c} \mid (\text{PBP}(\boldsymbol{B})^p, B) \models \text{IC}\}$

The main difference from Definition 7.3.1 is the use of AVR in place of the majority rule for the aggregation on premises. This choice allows us to always obtain a rational outcome over premises, without having to limit the use of this rule to a particular class of integrity constraints.

There are several applications in which this procedure may constitutes a natural candidate for the aggregation of binary ballots. Consider for example the case of integrity constraints that come in the form of conclusion functions, selecting an acceptance/rejection value over conclusions depending on the outcome on the premises. This is the case, for instance, of the discursive dilemma as presented in Chapter 3 (cf. Table 3.4). Formally, let $PS^p = \{p_j \mid j \in \mathcal{I}_p\}$ be the set of propositional variables associated with premises, and $\mathcal{L}^p$ the propositional language obtained from PS^p. A set of *conclusion functions* for $\mathcal{I}$ is represented by a set of formulas $\{p_j \leftrightarrow \varphi_j \mid \varphi_j \in \mathcal{L}^p\}$ for every conclusion $j \in \mathcal{I}^c$. For instance, the integrity constraint representing the discursive dilemma in Table 3.4

is $p_{\alpha \wedge \beta} \leftrightarrow (p_\alpha \wedge p_\beta)$. In similar cases the PBP represents an interesting example of a collectively rational procedure, sharing the same axiomatic properties as its JA version.

7.4 The Distance-Based Procedure

Recent work in JA has shown that ideas from belief merging (Konieczny and Pino Pérez, 2011) can be imported into JA to yield practical aggregation procedures that are complete and consistent (Pigozzi, 2006; Miller and Osherson, 2009; Lang et al., 2011). Similar definitions can be provided in binary aggregation, and in this section we introduce one of the easiest examples of a distance-based rule, studying the complexity of its winner determination problem.

7.4.1 Definition of the Procedure

When defining the AVR in Section 7.2, we were interested in devising a method to choose those ballots that best represent the individual views in a profile. Once the integrity constraint is known, this restriction can be lifted, considering the whole set of rational ballots as the search space for winning ballots. This approach is inspired by the literature on belief merging (Konieczny and Pino Pérez, 2011), in which, however, the set of individual views to be aggregated is a set of propositional models rather than a single evaluation.

Definition 7.4.1. Given a set of issues $\mathcal{I}$ and an integrity constraint IC, the *distance-based procedure* DBP is the function mapping each profile $\boldsymbol{B} \in \text{Mod}(\text{IC})^{\mathcal{N}}$ to the following set of collective ballots:

$$\text{DBP}(\boldsymbol{B}) \quad = \quad \underset{B \in \text{Mod}(\text{IC})}{\text{argmin}} \sum_{i \in \mathcal{N}} H(B, B_i)$$

The DBP is an *irresolute* procedure, returning a (nonempty) set of collective ballots. A winning ballot under the DBP minimises the amount of disagreement with the individual binary ballots (i.e., it minimises the sum of the Hamming distances with all individual ballots). Note that in cases where the majority rule leads to a rational outcome, the outcome of the DBP coincides with that of the majority rule (making it a resolute procedure over these profiles). Also note that the DBP shares many features with the *Kemeny rule* for preference aggregation (Kemeny, 1959). We will elaborate more on this similarity in the proof of Lemma 7.4.4.

7.4.2 Axiomatic Properties

The DBP does not satisfy many axiomatic properties. Certainly it is not independent, as it builds on the idea that correlations between issues should be exploited rather than neglected. Moreover, examples showing that the DBP does not satisfy unanimity $\mathrm{U}^\star$ can be found in the literature on judgment aggregation (Lang et al., 2011, Proposition 24). By the same argument as the one employed in Section 7.2.2 for the AVR, it is straightforward to see that the DBP satisfies anonymity $\mathrm{A}^\star$ and monotonicity $\mathrm{M}^\star$.

7.4.3 Winner Determination

We now want to analyse the complexity of the winner determination problem for the DBP. As the DBP is not resolute, we study the decision problem $\textsc{WinDet}^\star$. Our findings reveal that $\textsc{WinDet}^\star$(DBP) is Θ_2^p-complete, thus showing that this rule is very hard to compute. The class Θ_2^p (also known as $\Delta_2^p(O(\log n))$ or $\mathrm{P}_{||}^{\mathrm{NP}}$) is the class of problems that can be solved in polynomial time asking a logarithmic number of queries to an NP-oracle (Wagner, 1987).

To obtain our result, we first have to devise an NP-oracle that will then be used in the proof of Θ_2^p-membership. We shall use the following problem:

$\textsc{WinDet}_K^\star(F)$
Instance: Integrity constraint IC, profile $\boldsymbol{B} \in \mathrm{Mod}(\mathrm{IC})^{\mathcal{N}}$, subset $I \subseteq \mathcal{I}$, partial ballot $\rho : I \to \{0,1\}$, $K \in \mathbb{N}$.
Question: Is there a $B^\star$ with $B_j^\star = \rho(j)$ for all $j \in I$ such that $\sum_{i\in\mathcal{N}} H(B^\star, B_i) \leqslant K$?

That is, we ask whether there exists a binary ballot $B^\star$ with Hamming distance at most K extending a partial ballot ρ. We now show that this problem lies in NP.

Lemma 7.4.2. $\textsc{WinDet}_K^\star$(DBP) *is in* NP.

Proof. We devise an algorithm that, given a certificate intended to provide a positive answer to $\textsc{WinDet}_K^\star$(DBP), can check the validity of the certificate in polynomial time. Such a certificate is given by a ballot $B^\star$, that can be guessed among the set of all ballots $\{0,1\}^{\mathcal{I}}$. It is then sufficient to check that: (*i*) the certificate is rational, i.e., $B^\star \models \mathrm{IC}$ (*ii*) the certificate extends ρ, i.e., $B_j^\star = \rho(j)$ for all $j \in I$ and (*iii*) $\sum_{i\in\mathcal{N}} H(B^\star, B_i) \leqslant K$. Each of the three steps can be done in polynomial time, thus resulting in a non-deterministic polynomial algorithm. □

Lemma 7.4.3. $\textsc{WinDet}^\star$(DBP) *is in* Θ_2^p.

Proof. The problem $\textsc{WinDet}^\star$(DBP) asks whether there exists a *winning* ballot that extends a given partial ballot ρ. Since the Hamming distance of a ballot from a profile is bounded from above by a polynomial figure, we can solve this problem by performing a binary search over K and asking a logarithmic number of queries to $\textsc{WinDet}_K^\star$.

More precisely, since $\sum_{B\in\boldsymbol{B}} H(B^\star, B) \leqslant K^\star := |\mathcal{I}| \times |\mathcal{N}|$, $K^\star$ is polynomial in the size of the input. We can then ask a first query to $\textsc{WinDet}_K^\star$(DBP) with

$K = \frac{K^\star}{2}$ and empty partial ballot ρ. In case of a positive answer we can continue the search with a new $K = \frac{K^\star}{4}$, otherwise we move to the higher half of the interval querying $\text{WinDet}^\star_K$(DBP) with $K = \frac{3}{4} \times K^\star$. This process eventually ends after a logarithmic number of steps, providing the exact Hamming distance K^w of a winning candidate from the profile $\boldsymbol{J}$ under consideration. It is now sufficient to run the problem $\text{WinDet}^\star_K$(DBP) with $K = K^w$ and partial ballot ρ as in the original instance of $\text{WinDet}^\star$(DBP) we wanted to solve. In case the answer is positive, since there cannot be a winning ballot with Hamming score strictly less than K^w, one of the winning ballots extends the partial ballot ρ. On the other hand, in case of a negative answer all ballots extending ρ have Hamming distance bigger than K^w, and therefore cannot belong to the winning set. □

Next, we show that the upper bound established by Lemma 7.4.3 is tight. We exploit the similarity of the DBP to the Kemeny rule in preference aggregation to build on a known Θ^p_2-hardness result by Hemaspaandra et al. (2005).

Lemma 7.4.4. $\text{WinDet}^\star$(DBP) *is* Θ^p_2*-hard.*

Proof. We build a reduction from the problem Kemeny Winner, as defined by Hemaspaandra et al. (2005). An instance of this problem consists of a set of candidates C, a profile of weak orders $\boldsymbol{R} = (R_1, \dots, R_n)$ over C and a designated candidate c. Define the *Kemeny score* of a given candidate c as the following expression:

$$KemenyScore(c, \boldsymbol{R}) = \min\{ \sum_{i=1}^{n} d(R_i, Q) \mid Q \text{ is a weak order and } c \in \text{top}(Q)\}$$

Where $d(R_i, Q)$ is the Hamming distance between two preference orders and $\text{top}(Q)$ is the set of most preferred candidates in Q. The problem asks whether the Kemeny score of c is less than or equal to the Kemeny score of all other candidates $d \in C$.

We now build an instance of $\text{WinDet}^\star$(DBP) to decide this problem, exploiting the (polynomial) embedding of preference aggregation into binary aggregation that we have presented in Sections 3.1.2 and 5.1.3. Let the set of issues be $\mathcal{I}_C = \{ab \mid a, b \in C\}$, and let $\text{IC}_\leqslant$ be the set of integrity constraints enforcing the properties of weak orders. Given a preference profile $\boldsymbol{R}$, let $\boldsymbol{B^R}$ be the binary profile obtained by encoding each weak order R_i over C in a ballot B^R_i over $\mathcal{I}_C$.

Note that, if R and Q are two weak orders, then $d(R, Q) = H(B^R, B^Q)$. Thus, to obtain an answer to the initial Kemeny Winner instance with designated candidate c, it is sufficient to ask a query to $\text{WinDet}^\star$(DBP) using $\mathcal{I}_C$ as a set of issues, $\boldsymbol{B^R}$ as a profile, and a partial ballot ρ such that $\rho(cd) = 1$ for all $d \in C$ with $d \neq c$. If the winning ballot features c as one of the top candidates (i.e., issues cd are accepted for all other candidates d), then its Kemeny score will be lower or equal than that of all other candidates, providing a positive answer to the original problem. □

Putting Lemma 7.4.3 and 7.4.4 together yields a complete characterisation of the complexity of winner determination under the distance-based procedure:

Theorem 7.4.5. WinDet*(DBP) *is* Θ_2^p*-complete.*

Theorem 7.4.5 shows that the problem of using the DBP is highly intractable. However, by adapting efficient heuristics developed for the Kemeny rule (which, as seen in the proof of Lemma 7.4.4, is the preference aggregation version of the DBP) it may be possible to obtain a tractable implementation of the DBP for both binary ballots and judgment sets (Davenport and Kalagnanam, 2004; Conitzer et al., 2006; Betzler et al., 2009).

Given that winner determination is a highly intractable problem, we shall not investigate the complexity of strategic manipulation for the DBP.

7.4.4 The DBP in Judgment Aggregation

The first example of a JA procedure based on distances was introduced by Pigozzi (2006) based on the work of Konieczny and Pino Pérez (2002) in belief merging. The procedure was defined under the restrictive assumptions that the agenda be closed under propositional variables and all compound formulas be unanimously accepted (or rejected) by all agents. Most importantly, the syntactic information contained in the agenda was discarded by moving the aggregation from the level of formulas to the level of models.

A syntactic variant of this procedure can be obtained by merging judgment sets rather than models corresponding to judgment sets. This procedure was first introduced by Miller and Osherson (2009) under the name Proptotype$_\mathrm{d}$, and studied independently in our previous work (Endriss et al., 2010b) for the case of d equal to the Hamming distance. The definition of the DBP we have provided for binary aggregation is a straightforward adaptation of this rule to the case of binary aggregation. Note that the DBP does not coincide with the procedure of Pigozzi (2006), even for agendas closed under propositional variables. The main reason is that the DBP is sensitive to logical correlations between formulas of the agenda: accepting an atom that is correlated with other formulas in the agenda "counts" more in our procedure than accepting an independent one.

As shown in previous work (Endriss et al., 2010b), winner determination for the judgment aggregation version of the DBP is also hard. An analogous version of Lemma 7.4.2 can be proven by encoding the requirements of consistency of a judgment set in an integer program, thus showing that WinDet$^\star_K$(DBP) for judgment aggregation also lies in NP. A proof of Θ_2^p-completeness for WinDet*(DBP) can also be obtained with a reduction from the problem of Kemeny winner. To obtain this result, it is sufficient to enforce the properties of weak orders on judgment sets by means of a polynomial number of spurious formulas in the agenda, representing the integrity constraints of transitivity and reflexivity.

7.5 Conclusions: Which Framework is Easier?

In this chapter we have studied the computational complexity of three procedures for collectively rational aggregation. The *average voter rule*, a generalised dictatorship that selects those individual ballots that minimise the amount of disagreements with the remaining ballots in the profile, resulted in a polynomially computable rule that is also easy to manipulate. The problem of computing the winner of the *distance-based procedure*, defined in a similar way but without restricting the outcome to those ballots submitted by the individuals, is considerably harder, and we proved a Θ_2^p-completeness result for this case. Finally, the *premise based procedure*, which we defined and studied in the framework of judgment aggregation, satisfies good computational properties, being easy to use and hard to manipulate (NP-complete).

In Section 6.2.2 we have observed that JA and BA with IC have the same expressive power, once we enrich the formula-based framework for JA with the use of constraints (Dietrich and List, 2008b). However, we have also observed that the embedding of JA into BA with IC may result in an exponential number of integrity constraints (cf. Section 3.2.2). In this conclusive section, we want to compare the computational complexity of the two frameworks on a number of basic problems.

From the point of view of an individual, a decision problem framed in BA with IC is substantially easier to deal with than one expressed in the JA formalism. The first problem we consider is the following:

DefCheck
Instance: Integrity constraint IC, ballot $B \in \{0,1\}^{\mathcal{I}}$
(agenda Φ, judgment set $J \in 2^{\Phi}$, respectively)
Question: Is B rational?
(is J consistent, respectively?)

While in BA deciding whether $B_i \models \text{IC}$ can be solved with a polynomial model checking, DefCheck in JA corresponds to solving the satisfiability of the set of formulas in J_i, a classical NP-complete problem. Another problem that can be considered is that of inferring knowledge from the result of aggregation:

WinInf
Instance: Winning ballot $F(\boldsymbol{B})$, formula $\varphi \in \mathcal{L}_{PS}$
(Winning set $F(\boldsymbol{J})$, formula $\varphi \in \mathcal{L}_{PS}$, respectively)
Question: Is it the case that $F(\boldsymbol{B}) \models \varphi$?
(is it the case that $F(\boldsymbol{J}) \models \varphi$, respectively?)

In this case also, the former instance can be solved in polynomial time with model checking while the latter is significantly harder. To see this, consider that the outcome of a JA procedure is a set of formulas, and that knowledge inference from a set of propositional formulas is coNP-hard.

From the point of view of the mechanism designer, the computational complexity of the two frameworks does not differ significantly. Consider for instance the problem of safety of the agenda, as presented in Section 6.3 for the formula-based framework of JA. We have seen that for most classes of aggregation procedures this problem is Π_2^p-complete, by investigating the complexity of checking a number of agenda properties. Let us concentrate on one such property, the median property (Definition 6.3.2). As remarked in Section 6.3.4, an equivalent version of this problem in BA with IC is given by the following statement: given an integrity constraint IC, how hard is to check that IC is equivalent to a conjunction of clauses of maximal size 2? Similar problems in the literature are all placed in the second level of the polynomial hierarchy (see, e.g., Schaefer and Umans, 2002, *L9: Short CNF), and this is a good indication that the computational complexity of checking the median property in the two frameworks should be comparable. A different result is obtained if we consider the SSMP (Definition 6.3.5), which requires an agenda to be "atomic", i.e., to allow for all possible judgment sets. Its analogue in BA is the case of IC being a tautology, thus admitting all possible ballots as rational. In this case, the complexity of the BA version of this problem is significantly easier (being a coNP-complete problem) than its JA version, which we proved to be Π_2^p-complete.

If we concentrate on the use of specific aggregation procedures, as we did in this chapter, then the two frameworks have comparable computational complexity, as we observed in Section 7.4.4.

In conclusion, the framework of BA with IC is computationally at most as hard as dealing with formulas in the classical JA framework, and it is significantly easier in some situations. While for many applications it may be more natural to directly use propositional formulas rather than structure the problem with binary issues and devise a suitable integrity constraint, we can conclude from the present section that this step decreases significantly the computational complexity of many basic problems.

Chapter 8

Conclusions and Perspectives

If we have succeeded in our purpose, then the reader will agree that collective rationality should take centre stage in the study of aggregation problems. Inspired by the wide spectrum of potential applications in Artificial Intelligence, we have put forward a systematic study of collective rationality in binary aggregation, focusing on the syntactic structure of the integrity constraint that defines an aggregation problem. Moreover, we have shown that a failure in collective rationality is at the basis of most of the classical paradoxes in aggregation theory, and that the source of many (im)possibility results in the literature on preference and judgment aggregation lies in a clash between a set of axiomatic properties and requirements of collective rationality. By providing a unifying view on aggregation problems, our framework of binary aggregation with integrity constraints proved to be a useful and flexible tool for both the analysis of theoretical questions as well as for the development of solutions for application-oriented problems.

Let us now look back at what has been achieved in this dissertation. **Chapter 2** defined the framework of binary aggregation with integrity constraints, in which individuals make yes/no choices over a finite set of issues and an aggregation procedure merges them into a collective choice. The choices of individuals are bound by an integrity constraint or rationality assumption, which we represented as a formula in a simple propositional language. We provided a definition of collective rationality which features integrity constraints as a parameter: an aggregation procedure is collectively rational with respect to an integrity constraint if the collective outcome satisfies the constraint whenever all individual ballots do. We called a counterexample to collective rationality a paradox.

In **Chapter 3** we explored the generality of our definition of paradox by showing that many classical paradoxes from the literature on Social Choice Theory can be seen as instances of our definition. In particular, we focused on the Condorcet paradox, the discursive dilemma, the Ostrogorski paradox, and, to a lesser extent, the paradox of multiple elections.

In **Chapter 4** we defined the class $\mathcal{CR}[\mathcal{L}]$ as the class of collectively rational procedures with respect to all integrity constraints in a given language $\mathcal{L}$. By studying those classes, we discovered that classical axiomatic properties from the literature on Social Choice Theory correspond to requirements of collective rationality with respect to some natural syntactically defined languages. For instance, the class of procedures that are collectively rational with respect to the language of literals corresponds to the class of unanimous procedures; the language of equivalences $\mathcal{L}_{\leftrightarrow}$ is associated with the requirement of issue-neutrality; and the XOR-language $\mathcal{L}_{XOR}$ is associated with the axiom of domain-neutrality. Similar correspondences cannot be proven for certain other axiomatic requirements. For the axioms of anonymity, independence, and two versions of monotonicity, we proved negative results which rule out possible characterisations in terms of collective rationality. We then moved to the study of procedures defined by means of acceptance quotas, a class that includes the majority rule. We characterised the set of integrity constraints lifted by the majority rule as the language of *2-clauses*, i.e., disjunctions of size at most 2, and we provided conditions about the size of a clause for other quota rules.

Classical frameworks like preference and judgment aggregation can be embedded into binary aggregation by devising suitable integrity constraints, and in Chapters 5 and 6 we compared theoretical results in these settings with our findings in binary aggregation.

Chapter 5 focused on preference aggregation. We obtained a possibility and an impossibility theorem for different representations of preferences by making use of our characterisation results from Chapter 4. We also put forward a new proof of Arrow's Theorem which aimed at reducing the impossibility to a clash between the Arrovian axiomatic requirements on the one hand, and collective rationality with respect to the preferential integrity constraint on the other.

In **Chapter 6** we presented in detail the framework of judgment aggregation. We focused on the problem of safety of the agenda, i.e., characterising the set of agendas on which a given class of procedures is guaranteed to output a consistent outcome. For several classes of procedures defined axiomatically, we identified the class of safe agendas by giving conditions on the structure of the inconsistent subsets that can be created from formulas in the agenda. For instance, we proved that a safe agenda for independent procedures only contains inconsistent subsets composed of a formula and its negation (we called this condition the syntactically simplified median property). We confronted these results with our findings in binary aggregation and we investigated the computational complexity of recognising safe agendas. For all the classes under consideration, we proved that the problem is Π_2^p-complete, i.e., highly intractable.

Chapter 7 is devoted to studying aggregation procedures that are designed to be collectively rational, and that are good candidates to be used in practice. We focus on three rules that can be employed in both binary and judgment aggregation: the average voter rule, the premise-based procedure and the distance-based procedure. For each rule, we investigated their axiomatic properties as well as the computational complexity of the problems of winner determination and strategic manipulation (except for the distance-based rule, for which determining a winner is already too hard a problem to justify a study of manipulation). We concluded the chapter by comparing the computational complexity of some simple problems in binary and judgment aggregation.

There are numerous directions in which this work can be extended. **Voting theory** represents a closely connected topic in which many of the results presented in this dissertation may find application. First, as shown by our Example 2.1.6, approval voting, k-approval voting and the plurality rule can be seen as binary aggregation procedures over suitably defined domains. The use of constraints is not common in the study of voting procedures, but constitutes an interesting direction for future investigations, as testified by recent work on the topic (Lu and Boutilier, 2011). Second, our framework can easily generalise to account for the case of full (rather than binary) combinatorial domains, i.e., product spaces $\mathcal{D} = D_1 \times \cdots \times D_m$ with $|D_i| \geqslant 2$. In this more general setting, integrity constraints could be expressed in the propositional language $\mathcal{L}'_{PS}$ built on a set of atomic formulas $PS = \{x_j{=}a_j \mid j = 1, \dots, m \text{ and } a_j \in D_j\}$ (cf. Example 2.1.7). Axioms for aggregation procedures on general combinatorial domains can be adapted from the literature (see, e.g., Lang, 2007). Preliminary results linking collective rationality with axiomatic properties can be obtained for the simpler case of voting for committees, in which the domains D_j are equal to a set of available candidates D and the combinatorial domain is the set $\mathcal{D} = D^m$. Moreover, a language to compactly express preferences over a combinatorial domain could be devised from $\mathcal{L}'_{PS}$ by adding a binary predicate $<$. A study of the expressivity and succintness of this language constitutes another interesting idea for future work, comparing it to existing languages for compact representation of preferences (Uckelman, 2009; Bienvenu et al., 2010).

Binary issues can also be used to represent **preferential dependencies**, for instance by encoding the graph underlying a given CP-net (Boutilier et al., 2004). Every edge in the graph could be associated with a binary issue, and integrity constraints could then be devised to encode rationality assumptions *about* preferential dependencies. Moreover, interesting aggregation procedures may be devised by first aggregating individual dependency structures, and then devising a sequential order of local elections based on the collective dependency graph obtained. A preliminary study along these lines has been carried out by Airiau et al. (2011). A similar approach has also been followed by Xia et al. (2008) and Conitzer et al. (2011).

The generality of the framework of binary aggregation with integrity constraints may also suggest that the **aggregation of logical structures** constitute an interesting theoretical setting for the study of collective agents. As models describing autonomous agents get more refined, our simple propositional language may not be sufficiently expressive or compact to encode more complex constraints.

The general framework for judgment aggregation developed by Dietrich (2007) goes in this direction. However, in this line of research the domain of aggregation is still composed of a sequence of yes/no choices over a set of (non-propositional) formulas, a framework that does not deviate significantly from binary aggregation. Instead, it would be interesting to study the aggregation of non-propositional logical *models* (e.g., first-order models, Kripke structures, etc.) using formulas in more complex logical languages to express properties of such structures, such as rationality assumptions. We have recently made an initial step in this direction by providing a first study of graph aggregation in which integrity constraints may be expressed in modal logic (Endriss and Grandi, 2012).

Our study of the **computational complexity** of aggregation procedures is based on a worst-case analysis. Despite constituting a solid first step in preparation for a possible implementation of a procedure, this approach has been criticised as failing to capture the actual distribution of the hard instances of a problem (Faliszewski and Procaccia, 2010; Walsh, 2011a). Thus, an important step to be taken in future work is a more refined analysis of those problems studied in the dissertation, by for instance using tools from average complexity (Faliszewski and Procaccia, 2010) or parametrized complexity (Betzler, 2010). Another interesting possibility is performing an evaluation of the distribution of the hard instances of some of our problems, an approach that may be useful to shed light on very hard questions such as our findings on the computational complexity of recognising safe agendas, as recently done in the domain of the manipulation of elections (Walsh, 2011b).

Perhaps the most intriguing direction for future research is to move **from theory to applications**. The flexibility of our framework has the potential of bringing interesting insights to the study and development of a diverse range of aggregation problems, like ranking and recommender systems, but it is in the design of mechanisms for complex collective decisions that lies its biggest potential. A recent position paper by Boutilier and Lu (2011) advocates a move from the study of high-stake and low-frequency situations like elections to low-stake and high-frequency situations such as those that are encountered in electronic commerce and similar applications. While the theoretical framework developed in this dissertation does not straightforwardly point in this direction, it is in its applicability and in its potential of suggesting novel solutions to real-world problems that it will find its main testing ground.

Appendix A

Propositional Logic

This appendix is devoted to introducing the basic terminology and definitions of propositional logic (for a more detailed presentation see, e.g., Shoenfield, 1967).

A.1 Formulas

Propositional logic is perhaps the simplest formal language that can be defined. The basic ingredients are a set of *atomic propositions* $PS = \{p_1, \dots, p_m\}$[1] and *propositional connectives*: negation ($\neg$), conjunction ($\wedge$) and disjunction ($\vee$). The set of *propositional formulas* built on *PS*, which we denote $\mathcal{L}_{PS}$, is defined recursively in the following way:

(*i*) every element of *PS* is a formula;
(*ii*) if φ is a formula, then $\neg\varphi$ is a formula;
(*iii*) if both φ and ψ are formulas, then $\varphi \vee \psi$ and $\varphi \wedge \psi$ are formulas.

Parentheses may be used to make the priorities between connectives clearer. We will make use of two additional connectives: implication ($\varphi \rightarrow \psi$) as a shorthand for $\neg\varphi \vee \psi$, and equivalence ($\varphi \leftrightarrow \psi$) as a shorthand for $(\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$. Here ends the syntactic definition of the propositional language $\mathcal{L}_{PS}$.

Definition A.1.1. We give the following definitions for types of formulas:
(*i*) an *atom* is an element of *PS*;
(*ii*) a *literal* is an atom or its negation;
(*iii*) a *cube* is a finite conjunction of literals;
(*iv*) a *clause* is a finite disjunction of literals.

[1]Here, we make the more restrictive assumption that the set of propositional symbols *PS* is finite, although it is usually only assumed to be countable.

A.2 Models

Formulas in the propositional language $\mathcal{L}_{PS}$ are used to express properties of binary product spaces $\{0,1\}^m$ (which is usually called its semantics). An *assigment* is a function $\rho : PS \to \{0,1\}$ (equivalently, an element of $\{0,1\}^{PS}$). We now define, in a recursive fashion, what it means for an assignment ρ to be a *model* of φ (and we write this $\rho \models \varphi$):

(*i*) if $\varphi \in PS$ then $\rho \models \varphi$ iff $\rho(\varphi) = 1$;
(*ii*) if $\varphi = \neg\psi$ then $\rho \models \varphi$ iff $\rho \not\models \psi$;
(*iii*) if $\varphi = \psi_1 \wedge \psi_2$ then $\rho \models \varphi$ iff $\rho \models \psi_1$ and $\rho \models \psi_2$;
(*iv*) if $\varphi = \psi_1 \vee \psi_2$ then $\rho \models \varphi$ iff $\rho \models \psi_1$ or $\rho \models \psi_2$.

Every formula $\varphi \in \mathcal{L}_{PS}$ corresponds to the subset of $\{0,1\}^{PS}$ of those assigments that are models of φ. We call this set $\mathrm{Mod}(\varphi)$. Since we have a finite number of propositional symbols, we can associate with every assignment ρ a formula φ that has ρ as its unique model. For instance, the formula associated with the assignment ρ that rejects all propositional symbols in PS is $\varphi_\rho = \neg p_1 \wedge \cdots \wedge \neg p_m$. Thus, every subset of $\{0,1\}^{PS}$ corresponds to the set of models of a certain formula, obtained as the disjunction of all formulas corresponding to each assignment in the given subset. $\mathcal{L}_{PS}$ is therefore fully expressive with respect to subsets of $\{0,1\}^{PS}$.

Definition A.2.1. We give the following definitions:

(*i*) A *tautology* is a formula that is true for every assignment to its propositional variables. We denote a tautology with the symbol $\top$.
(*ii*) A *contradiction* is a formula that is false for every assignment to its propositional variables. We denote a contradiction with the symbol $\bot$.
(*iii*) A formula is *satisfiable* (also, consistent) if it admits a model. A formula is *falsifiable* if its negation admits a model.
(*iv*) A formula is *contingent* if it is both satisfiable and falsifiable.
(*v*) A formula φ is *consistent with a formula* ψ if its conjunction $\varphi \wedge \psi$ admits a model.

Bibliography

The numbers in brackets following each entry in the bibliography indicate the page on which a reference has been cited.

D. S. Ahn and S. Oliveros. Combinatorial voting. *Econometrica*, 80(1):89–141, 2012. (p. 2)

S. Airiau, U. Endriss, U. Grandi, D. Porello, and J. Uckelman. Aggregating dependency graphs into voting agendas in multi-issue elections. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, 2011. (p. 2, 43, 82, 135)

A. Altman and M. Tennenholtz. Axiomatic foundations for ranking systems. *Journal of Artificial Intelligence Research (JAIR)*, 31:473–495, 2008. (p. 2)

A. Altman and M. Tennenholtz. An axiomatic approach to personalized ranking systems. *Journal of the ACM*, 57(4), 2010. (p. 2)

S. Arora and B. Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2009. (p. 103)

K. J. Arrow. *Social Choice and Individual Values*. John Wiley & Sons, second edition, 1963. (p. 1, 3, 11, 20, 26, 30, 31, 61, 73, 77, 78, 79, 81, 83)

K. J. Arrow, A. K. Sen, and K. Suzumura, editors. *Handbook of Social Choice and Welfare*, volume 1. Elsevier, 2002. (p. 77)

S. Barberà, H. Sonnenschein, and L. Zhou. Voting by committees. *Econometrica*, 59(3):595–609, 1991. (p. 24)

S. Barberà, W. Bossert, and P. K. Pattanaik. Ranking sets of objects. In S. Barberà, P. J. Hammond, and C. Seidl, editors, *Handbook of Utility Theory*. Kluwer Academic Publishers, Dordrecht, 2004. (p. 116)

J. J. Bartholdi and J. Orlin. Single transferable vote resists strategic voting. *Social Choice and Welfare*, 8(4):341–354, 1991. (p. 111)

J. J. Bartholdi, C. A. Tovey, and M. A. Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6(2):157–165, 1989a. (p. 111)

J. J. Bartholdi, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6(3):227241, 1989b. (p. 111)

D. Baumeister, G. Erdélyi, and J. Rothe. How hard is it to bribe the judges? A study of the complexity of bribery in judgment aggregation. In *Proceedings of the 2nd International Conference on Algorithmic Decision Theory (ADT-2011)*, 2011. (p. 4, 108)

N. Betzler. *A Multivariate Complexity Analysis of Voting Problems*. PhD thesis, Friedrich-Schiller-Universitt Jena, 2010. (p. 136)

N. Betzler, M. R. Fellows, J. Guo, R. Niedermeier, and F. A. Rosamond. Fixed-parameter algorithms for Kemeny rankings. *Theoretical Computer Science*, 410 (45):4554–4570, 2009. (p. 4, 130)

T. Bezembinder and P. van Acker. The Ostrogorski paradox and its relation to non-transitive choice. *Journal of Mathematical Sociology*, 11:131–158, 1985. (p. 37)

M. Bienvenu, J. Lang, and N. Wilson. From preference logics to preference languages, and back. In *Proceeding of the Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (KR-2010)*, 2010. (p. 135)

D. Black. *The Theory of Committees and Elections*. Cambridge University Press, 1958. (p. 30, 77, 78)

Borda, Jean-Charles de. *Memoire sur les Elections au Scrutin*. Histoire de l'Academie Royale des Sciences, Paris, 1781. (p. 29)

C. Boutilier and T. Lu. Probabilistic and utility-theoretic models in social choice: Challenges for learning, elicitation, and manipulation. In *Proceedings of the IJCAI-2011 Workshop on Social Choice and Artificial Intelligence*, 2011. (p. 136)

C. Boutilier, R. Brafman, C. Domshlak, H. Hoos, and D. Poole. CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. *Journal of Artifcial Intelligence Research*, 21:135–191, 2004. (p. 135)

S. J. Brams and P. C. Fishburn. Voting procedures. In K. J. Arrow, A. K. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*. Elsevier, 2002. (p. 4)

S. J. Brams and P. C. Fishburn. *Approval voting*. Springer, 2nd edition, 2007. (p. 18)

S. J. Brams, D. M. Kilgour, and W. S. Zwicker. A new paradox of vote aggregation. In *Annual Meeting of the American Political Science Association*, 1993. (p. 39)

S. J. Brams, D. M. Kilgour, and W. S. Zwicker. The paradox of multiple elections. *Social Choice and Welfare*, 15(2):211–236, 1998. (p. 2, 9, 41, 42)

F. Brandt, V. Conitzer, and U. Endriss. Computational social choice. In G. Weiss, editor, *Multiagent Systems*. MIT Press, Forthcoming. (p. 2, 4)

M. Caminada and G. Pigozzi. On judgment aggregation in abstract argumentation. *Autonomous Agents and Multi-Agent Systems*, 22(1):64–102, 2011. (p. 108)

F. Cariani, M. Pauly, and J. Snyder. Decision framing in judgment aggregation. *Synthese*, 163(1):1–24, 2008. (p. 55)

Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. A short introduction to computational social choice. In *Proceedings of the 33rd Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM-2007)*, 2007. (p. 2)

Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. Preference handling in combinatorial domains: From AI to social choice. *AI Magazine*, 29(4):37–46, 2008. (p. 2)

Condorcet, Marie Jean Antoine Nicolas de Caritat, Marquis de. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris, 1785. (p. 5, 8, 9, 29, 30)

V. Conitzer. *Computational aspects of preference aggregation*. PhD thesis, Computer Science Department, Carnagie Mellon University, 2006. (p. 4)

V. Conitzer. Making decisions based on the preferences of multiple agents. *Communications of the ACM*, 53(3):84–94, 2010. (p. 2)

V. Conitzer and L. Xia. Paradoxes of multiple elections: An approximation approach. In *Proceedings of the 13th International Conference on Principles of Knowledge Representation and Reasoning (KR-2012)*, 2012. (p. 42)

V. Conitzer, A. J. Davenport, and J. Kalagnanam. Improved bounds for computing Kemeny rankings. In *Proceedings of The Twenty-First National Conference on Artificial Intelligence (AAAI-2006)*, 2006. (p. 130)

V. Conitzer, J. Lang, and L. Xia. Hypercubewise preference aggregation in multi-issue domains. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, 2011. (p. 135)

T. R. Daniëls and E. Pacuit. A general approach to aggregation problems. *Journal of Logic and Computation*, 19(3):517–536, 2009. (p. 87)

A. J. Davenport and J. Kalagnanam. A computational study of the Kemeny rule for preference aggregation. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-2004)*, 2004. (p. 130)

B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, 2nd edition, 2002. (p. 85, 87)

R. Deb and D. Kelsey. On constructing a generalised Ostrogorski paradox: necessary and sufficient conditions. *Mathematical Social Sciences*, 14:161–174, 1987. (p. 37)

F. Dietrich. A generalised model of judgment aggregation. *Social Choice and Welfare*, 28(4):529–565, 2007. (p. 34, 87, 90, 94, 136)

F. Dietrich. General representation of epistemically optimal procedures. *Social Choice and Welfare*, 26(2):263–283, 2010. (p. 108)

F. Dietrich and C. List. Judgment aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics*, 19(4):391–424, 2007a. (p. 25, 70, 74, 92, 100)

F. Dietrich and C. List. Arrow's theorem in judgment aggregation. *Social Choice and Welfare*, 29(1):19–33, 2007b. (p. 87)

F. Dietrich and C. List. Strategy-proof judgment aggregation. *Economics and Philosophy*, 23(3):269–300, 2007c. (p. 112, 115)

F. Dietrich and C. List. Judgment aggregation without full rationality. *Social Choice and Welfare*, 31(1):15–39, 2008a. (p. 35, 108)

F. Dietrich and C. List. Judgment aggregation under constraints. In T. Boylan and R.Gekker, editors, *Economics, Rational Choice and Normative Philosophy*. Routledge, London, 2008b. (p. 94, 131)

F. Dietrich and C. List. The aggregation of propositional attitudes: towards a general theory. *Oxford Studies in Epistemology*, 3:215–234, 2010. (p. 108)

F. Dietrich and P. Mongin. The premiss-based approach to judgment aggregation. *Journal of Economic Theory*, 145:562–582, 2010. (p. 122)

E. Dokow and R. Holzman. Aggregation of binary evaluations for truth-functional agendas. *Social Choice and Welfare*, 32(2):221–241, 2009. (p. 27, 35, 74, 95)

E. Dokow and R. Holzman. Aggregation of binary evaluations. *Journal of Economic Theory*, 145(2):495–511, 2010a. (p. 21, 27, 35, 48, 74, 94, 95, 96)

E. Dokow and R. Holzman. Aggregation of binary evaluations with abstentions. *Journal of Economic Theory*, 145(2):544–561, 2010b. (p. 35)

J. S. Dryzek and C. List. Social choice theory and deliberative democracy: A reconciliation. *British Journal of Political Science*, 33(01):1–28, 2003. (p. 108)

C. Duddy and A. Piggins. A measure of distance between judgement sets. *Social Choice and Welfare*, 2012. In press. (p. 117)

P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–358, 1995. (p. 108)

C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank aggregation methods for the web. In *Proceedings of the 10th International Conference on World Wide Web (WWW-2001)*, 2001. (p. 2)

D. Eckert and C. Klamler. A geometric approach to paradoxes of majority voting in abstract aggregation theory. In *Proceedings of the First International Conference on Algorithmic Decision Theory (ADT-2009)*, 2009. (p. 37)

U. Endriss and U. Grandi. Graph aggregation. In *Proceedings of 4th International Workshop on Computational Social Choice (COMSOC-2012)*, 2012. (p. 82, 87, 136)

U. Endriss, M. S. Pini, F. Rossi, and K. B. Venable. Preference aggregation over restricted ballot languages: Sincerity and strategy-proofness. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI-2009)*, 2009. (p. 4)

U. Endriss, U. Grandi, and D. Porello. Complexity of judgment aggregation: Safety of the agenda. In *Proceedings of the 9th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2010)*, 2010a. (p. 4, 11, 22, 25, 92, 108)

U. Endriss, U. Grandi, and D. Porello. Complexity of winner determination and strategic manipulation in judgment aggregation. In *Proceedings of the 3rd International Workshop on Computational Social Choice (COMSOC-2010)*, 2010b. (p. 4, 108, 130)

P. Everaere, S. Konieczny, and P. Marquis. The strategy-proofness landscape of merging. *Journal of Artificial Intelligence Research (JAIR)*, 28:49–105, 2007. (p. 112)

P. Faliszewski and A. D. Procaccia. AI's war on manipulation: Are we winning? *AI Magazine*, 31(4):53–64, 2010. (p. 111, 136)

P. Faliszewski, E. Hemaspaandra, L. A. Hemaspaandra, and J. Rothe. A richer understanding of the complexity of election systems. In S. Ravi and S. Shukla, editors, *Fundamental Problems in Computing: Essays in Honor of Professor Daniel J. Rosenkrantz*. Springer, 2009. (p. 111)

W. Gaertner. *Domain Conditions in Social Choice Theory*. Cambridge University Press, 2001. (p. 78)

W. Gaertner. *A Primer in Social Choice Theory*. Oxford University Press, 2006. (p. 7, 10, 19, 31, 77, 79, 112)

P. Gärdenfors. A representation theorem for voting with logical consequences. *Economics and Philosophy*, 22(02):181–190, 2006. (p. 96)

J. Geanakoplos. Three brief proofs of Arrow's impossibility theorem. *Economic Theory*, 26(1):211–215, 2005. (p. 83)

A. Gibbard. Social choice and the Arrow conditions. Mimeo, 1969. (p. 84, 86)

A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973. (p. 12)

C. Gonzales, P. Perny, and S. Queiroz. Preference aggregation with graphical utility models. In *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, (AAAI-2008)*, 2008. (p. 4)

U. Grandi. The common structure of paradoxes in aggregation theory. In *Proceedings of 4th International Workshop on Computational Social Choice (COMSOC-2012)*, 2012. (p. 39)

U. Grandi and U. Endriss. Lifting rationality assumptions in binary aggregation. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI-2010)*, 2010. (p. 46)

U. Grandi and U. Endriss. Binary aggregation with integrity constraints. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, 2011. (p. 94, 117)

U. Grandi and U. Endriss. First-order logic formalisation of impossibility theorems in preference aggregation. *Journal of Philosophical Logic*, 2012. In press. (p. 61)

U. Grandi and G. Pigozzi. On compatible multi-issue group decisions. In *Proceedings of the 10th Conference on Logic and the Foundations of Game and Decision Theory (LOFT-2012)*, 2012. (p. 42, 122)

D. Grossi. Unifying preference and judgment aggregation. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2009)*, 2009. (p. 87, 88, 108)

D. Grossi. Correspondences in the theory of aggregation. In *Proceedings of the 9th Conference on Logic and the Foundations of Game and Decision Theory (LOFT-8)*, 2010. (p. 87, 88)

G. Guilbaud. Les théories de l'intérêt général et le problème logique de l'aggrégation. *Economie appliquée*, 5:501–505, 1952. (p. 33)

E. Hemaspaandra, H. Spakowski, and J. Vogel. The complexity of Kemeny elections. *Theoretical Computer Science*, 349(3):382–391, 2005. (p. 129)

E. Hemaspaandra, L. A. Hemaspaandra, and C. Menton. Search versus decision for election manipulation problems. *CoRR*, abs/1202.6641, 2012. (p. 115)

F. Herzberg and D. Eckert. Impossibility results for infinite-electorate abstract aggregation rules. *Journal of Philosophical Logic*, 41:273–286, 2012. (p. 74, 87)

J. Hintikka. *Knowledge and Belief*. Cornell University Press, 1962. (p. 7)

O. Hudry. On the complexity of Slater's problems. *European Journal of Operational Research*, 203(1):216–221, 2010. (p. 4)

J. S. Kelly. The Ostrogorski paradox. *Social Choice and Welfare*, 6(1):71–76, 1989. (p. 37)

J. Kemeny. Mathematics without numbers. *Daedalus*, 88:577–591, 1959. (p. 2, 127)

A. Kirman and D. Sondermann. Arrow's Theorem, many agents, and invisible dictators. *Journal of Economic Theory*, 5(2):267–277, 1972. (p. 83)

C. Klamler and D. Eckert. A geometric approach to judgment aggregation. In *Proceedings of the 2nd International Workshop on Computational Social Choice (COMSOC-2008)*, 2008. (p. 108)

S. Konieczny and R. Pino Pérez. Merging information under constraints: A logical framework. *Journal of Logic and Computation*, 12(5):773–808, 2002. (p. 3, 13, 20, 130)

S. Konieczny and R. Pino Pérez. Logic based merging. *Journal of Philosophical Logic*, 40(2):239–270, 2011. (p. 3, 20, 127)

L. A. Kornhauser and L. G. Sager. Unpacking the court. *Yale Law Journal*, 96 (1):82–117, 1986. (p. 29, 33)

L. A. Kornhauser and L. G. Sager. The one and the many: adjudication in collegial courts. *California Law Review*, 81(1):1–59, 1993. (p. 33, 122)

D. Lacy and E. M. S. Niou. A problem with referendums. *Journal of Theoretical Politics*, 12(1):5–31, 2000. (p. 2, 41, 42)

J. Lang. Logical preference representation and combinatorial vote. *Annals of Mathematics and Artificial Intelligence*, 42(1-3):37–71, 2004. (p. 2)

J. Lang. Vote and aggregation in combinatorial domains with structured preferences. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI-2007)*, 2007. (p. 2, 42, 135)

J. Lang and L. Xia. Sequential composition of voting rules in multi-issue domains. *Mathematical Social Sciences*, 57:304–324, 2009. (p. 2, 42, 43)

J. Lang, G. Pigozzi, M. Slavkovik, and L. van der Torre. Judgment aggregation rules based on minimization. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-XIII)*, 2011. (p. 4, 108, 113, 118, 122, 127, 128)

M. Li, B. Vo, and R. Kowalczyk. Majority-rule-based preference aggregation on multi-attribute domains with structured preferences. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-11)*, 2011. (p. 2)

C. List and R. E. Goodin. Epistemic democracy: Generalizing the Condorcet jury theorem. *Journal of Political Philosophy*, 9(3):277–306, 2001. (p. 108)

C. List and P. Pettit. Aggregating sets of judgments: An impossibility result. *Economics and Philosophy*, 18(1):89–110, 2002. (p. 3, 4, 5, 9, 11, 24, 33, 90, 91, 92, 96)

C. List and P. Pettit. Aggregating sets of judgments: Two impossibility results compared. *Synthese*, 140:207–235, 2004. (p. 87)

C. List and P. Pettit. *Group Agency: The Possibility, Design and Status of Corporate Agents*. Oxford University Press, 2011. (p. 108)

C. List and C. Puppe. Judgment aggregation: A survey. In *Handbook of Rational and Social Choice*. Oxford University Press, 2009. (p. 4, 7, 10, 13, 18, 19, 21, 27, 34, 35, 67, 90, 91, 92, 96, 98)

T. Lu and C. Boutilier. Budgeted social choice: From consensus to personalized decision making. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, 2011. (p. 135)

P. Marquis. Consequence finding algorithms. In D. Gabbay and P. Smets, editors, *Handbook on Defeasible Reasoning and Uncertainty Management Systems*. Kluwer Academic Publishers, Dordrecht, 2000. (p. 67)

K. O. May. A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica*, 20(4):680–684, 1952. (p. 21, 26, 42, 80)

P. Maynard-Zhang and D. J. Lehmann. Representing and aggregating conflicting beliefs. *Journal of Artificial Intelligence Research (JAIR)*, 19:155–203, 2003. (p. 3)

I. McLean and A. B. Urken. *Classics of social choice*. University of Michigan Press, 1995. (p. 1, 30, 43)

M. Miller and D. Osherson. Methods for distance-based judgment aggregation. *Social Choice and Welfare*, 32(4):575–601, 2009. (p. 13, 118, 121, 122, 127, 130)

H. Moulin. *The strategy of social choice*. North-Holland, 1983. (p. 25)

H. Moulin. *Axioms of Cooperative Decision Making*. Cambridge University Press, 1988. (p. 56)

J. F. Nash. The bargaining problem. *Econometrica*, 18:155–162, 1950. (p. 20)

I. Nehama. Approximate judgment aggregation. In *Proceedings of the 3rd International Workshop on Computational Social Choice (COMSOC-2010)*, 2010. (p. 108)

K. Nehring and C. Puppe. The structure of strategy-proof social choice. Part I: General characterization and possibility results on median spaces. *Journal of Economic Theory*, 135(1):269–305, 2007. (p. 27, 39, 68, 74, 95, 96, 97, 98)

K. D. Nehring and C. Puppe. A theory of diversity. *Econometrica*, 70(3):1155–1198, 2002. (p. 7)

K. D. Nehring and C. Puppe. Abstract Arrowian aggregation. *Journal of Economic Theory*, 145(2):467–494, 2010. (p. 27, 74, 91, 95, 96)

H. Nurmi. Compound majority paradoxes and proportional representation. *European Journal of Political Economy*, 13(3):443–454, 1997. (p. 39)

H. Nurmi. *Voting paradoxes and how to deal with them.* Springer, 1999. (p. 29, 36, 43)

M. Ostrogorski. *La démocratie et l'organisation de partis politiques.* Calmann-Levy, Paris, 1902. (p. 9, 36)

C. H. Papadimitriou. *Computational Complexity.* Addison Wesley, 1994. (p. 35, 103)

M. Pauly and M. van Hees. Logical constraints on judgement aggregation. *Journal of Philosophical Logic*, 35:569–585, 2006. (p. 96)

D. M. Pennock, E. Horvitz, and C. L. Giles. Social choice theory and recommender systems: Analysis of the axiomatic foundations of collaborative filtering. In *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-2000)*, 2000. (p. 2)

G. Pigozzi. Two aggregation paradoxes in social decision making: the Ostrogorski paradox and the discursive dilemma. *Episteme: A Journal of Social Epistemology*, 2(2):33–42, 2005. (p. 37)

G. Pigozzi. Belief merging and the discursive dilemma. *Synthese*, 152(2):285–298, 2006. (p. 13, 122, 127, 130)

M. S. Pini. *Reasoning with Preferences and Uncertainty.* PhD thesis, University of Padova, 2007. (p. 4)

M. S. Pini, F. Rossi, K. B. Venable, and T. Walsh. Aggregating partially ordered preferences. *Journal of Logic and Computation*, 19(3):475–502, 2009. (p. 4, 87)

M. S. Pini, F. Rossi, K. B. Venable, and T. Walsh. Incompleteness and incomparability in preference aggregation: Complexity results. *Artificial Intelligence*, 175(7-8):1272–1289, 2011. (p. 4)

R. A. Pollak. Bergson-Samuelson social welfare functions and the theory of social choice. *The Quarterly Journal of Economics*, 93(1):73–90, 1979. (p. 61, 62)

D. Porello. Ranking judgments in Arrow's setting. *Synthese*, 173:199–210, 2010. (p. 87, 88)

A. D. Procaccia. Computational social choice: The first four centuries. *ACM XRDS*, 18(2):31–34, 2011. (p. 2)

D. Rae and H. Daudt. The Ostrogorski paradox: A peculiarity of compound majority decision. *European Journal of Political research*, 4:391–398, 1976. (p. 36, 37)

W. Riker. *Liberalism against Populism.* W.H.Freeman, 1982. (p. 21, 62)

K. W. S. Roberts. Interpersonal comparability and social choice theory. *The Review of Economic Studies*, 47(2):421–439, 1980a. (p. 55, 61)

K. W. S. Roberts. Social choice theory: The single-profile and multi-profile approaches. *Review of Economic Studies*, 47(2):441–50, 1980b. (p. 61)

J. E. Roemer. *Theories of distributive justice.* Harvard University Press, 1996. (p. 22, 56)

F. Rossi, K. B. Venable, and T. Walsh. mCP-nets: Representing and reasoning with preferences of multiple agents. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-2006)*, pages 729–734, 2004. (p. 2)

F. Rossi, K. B. Venable, and T. Walsh. *A Short Introduction to Preferences: Between Artificial Intelligence and Social Choice.* Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2011. (p. 4, 82)

A. Rubinstein. The single profile analogues to multi profile theorems: Mathematical logic's approach. *International Economic Review*, 25(3):719–730, 1984. (p. 61)

A. Rubinstein and P. C. Fishburn. Algebraic aggregation theory. *J. Economic Theory*, 38(1):63–77, 1986. (p. 27)

D. G. Saari. *Disposing Dictators, Demystifying Voting Paradoxes.* Cambridge University Press, 2008. (p. 81)

P. Samuelson. Arrow's mathematical politics. In S. Hook, editor, *Human Values and Economic Policy: A Symposium.* New York University Press, 1967. (p. 61)

T. W. Sandholm. Distributed rational decision making. In G. Weiss, editor, *Multiagent systems.* MIT Press, 1999. (p. 1)

M. A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187 – 217, 1975. (p. 12)

M. Schaefer and C. Umans. Completeness in the polynomial-time hierarchy: A compendium. *SIGACT News*, 33(3):32–49, 2002. (p. 132)

J. R. Shoenfield. *Mathematical Logic.* Addison-Wesley, 1967. (p. 137)

Y. Shoham and K. Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations.* Cambridge University Press, 2009. (p. 2)

M. Slavkovik. *Judgment Aggregation for MultiAgent Systems.* PhD thesis, Université du Luxembourg, 2012. (p. 108)

M. Slavkovik and W. Jamroga. Distance-based judgment aggregation of three-valued judgments with weights. In *Proceedings of the IJCAI-2011 Workshop on Social Choice and Artificial Intelligence*, 2011. (p. 108)

M. Slavkovik and W. Jamroga. Computational and social theoretic properties of weighted distance-based aggregation of three-valued judgments. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2012)*, 2012. (p. 108)

L. J. Stockmeyer. The polynomial-time hierarchy. *Theoretical Computer Science*, 3(1):1 – 22, 1976. (p. 103)

A. D. Taylor. *Social Choice and the Mathematics of Manipulation.* Cambridge University Press, 2005. (p. 77, 79)

J. Uckelman. *More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains.* PhD thesis, University of Amsterdam, 2009. (p. 135)

R. Vacca. Opinioni individuali e deliberazioni collettive. *Rivista internazionale di filosofia del diritto*, 52:52–59, 1922. (p. 33)

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behaviour.* Princeton Univeristy Press, 1947. (p. 20)

K. W. Wagner. More complicated questions about maxima and minima, and some closures of NP. *Theoretical Computer Science*, 51(1-2):53–80, 1987. (p. 128)

T. Walsh. Is computational complexity a barrier to manipulation? *Annals of Mathematics and Artificial Intelligence*, 62(1-2):7–26, 2011a. (p. 111, 136)

T. Walsh. Where are the hard manipulation problems? *Journal of Artificial Intelligence Research (JAIR)*, 42:1–29, 2011b. (p. 136)

R. B. Wilson. On the theory of aggregation. *Journal of Economic Theory*, 10 (1):89–99, 1975. (p. 26, 27, 73, 81)

M. Wooldridge. *An introduction to multiagent systems.* John Wiley & Sons, second edition, 2009. (p. 2)

L. Xia. *Computational Voting Theory: Game-Theoretic and Combinatorial Aspects*. PhD thesis, Duke University, 2011. (p. 2)

L. Xia and V. Conitzer. A maximum likelihood approach towards aggregating partial orders. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, 2011. (p. 87)

L. Xia, V. Conitzer, and J. Lang. Voting on multiattribute domains with cyclic preferential dependencies. In *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (AAAI-2008)*, 2008. (p. 135)

L. Xia, V. Conitzer, and J. Lang. Aggregating preferences in multi-issue domains by using maximum likelihood estimators. In *9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, 2010. (p. 2)

L. Xia, V. Conitzer, and J. Lang. Strategic sequential voting in multi-issue domains and multiple-election paradoxes. In *Proceedings of the 12th ACM Conference on Electronic Commerce*, 2011. (p. 42)

Index

Samenvatting

Deze dissertatie presenteert het systeem van *binaire aggregatie met integriteitsvoorwaarden* en vergelijkt deze met bestaande systemen uit de theorie van collectieve besluitvorming die zijn ontwikkeld in de sociale keuzetheorie en de kunstmatige intelligentie. Daarbij verkent het theoretische resultaten die kunnen leiden tot de implementatie van dit systeem.

In ons systeem moet een verzameling van individuen een keuze maken uit een verzameling van binaire kwesties. Elk individu dient een ja/nee-keuze in voor elke kwestie en deze keuzes worden samengevoegd tot een collectieve keuze gebaseerd op een aggregatie-procedure. Individuele keuzes zijn afgegrensd door een rationaliteits-aanname die de verscheidenheid aan antwoorden definieert die als rationeel worden beschouwd. Wij representeren rationaliteits-aannames met eenvoudige propositionele formules die integriteitsvoorwaarden worden genoemd. Dit systeem kan worden gebruikt om een verscheidenheid aan situaties van collectieve besluitvorming te representeren, zoals samengestelde referenda en commissieverkiezingen alsook het probleem van aggregatie van individuele voorkeuren of oordelen.

Een concept dat centraal staat in de gehele dissertatie is dat van de *collectieve rationaliteit*: onderstellend dat elk individu voldoet aan een gegeven integriteitsvoorwaarde, zijn wij geïnteresseerd te achterhalen of het resultaat van een aggregatie-procedure nog steeds voldoet aan dezelfde integriteitsvoorwaarde.

We noemen een situatie waarin *niet* aan collectieve rationaliteit wordt voldaan een *paradox*. En we tonen aan dat de meeste klassieke paradoxen uit de literatuur over de sociale keuzetheorie kunnen worden beschouwd als voorbeelden van onze algemene definitie. We concentreren onze analyse op de Condorcet paradox, het discursieve dilemma, de Ostrogorski paradox en de samengestelde verkiezingen, en identificeren een gemeenschappelijke syntactische eigenschap in de integriteitsvoorwaarde die deze paradoxale situaties definieert.

We classificeren integriteitsvoorwaarden in syntactisch gedefinieerde talen en definieren $\mathcal{CR}[\mathcal{L}]$ als de verzameling van procedures die collectief rationeel zijn met betrekking tot elke integriteitsvoorwaarde in een bepaalde taal $\mathcal{L}$. We indiceren bijvoorbeeld met $\mathcal{CR}[cubes]$ de verzameling van aggregatie-procedures die collectief rationeel zijn met betrekking tot de integriteitsvoorwaarden die kunnen worden omschreven als conjuncties van literalen (i.e. cubes). Anderzijds worden verzamelingen van aggregatie-procedures normaal gesproken gedefinieerd in termen van axioma's, en indiceren we met $\mathcal{F}[\text{AX}]$ de verzameling van procedures die voldoen aan de axioma's AX. We onderzoeken de relatie tussen deze twee definities voor verscheidene fragmenten van de propositionele logische taal en voor verscheidene axioma's uit de literatuur van de sociale keuzetheorie. We bewijzen bijvoorbeeld dat de verzameling van collectief rationele procedures, met betrekking tot de cubes, gelijk is aan de verzameling van eensgezinde procedures, i.e., de aggregaten accepteren (wijzen af) een gegeven kwestie indien alle individuen overeenkomen tot het accepteren (afwijzen) van de kwestie: $\mathcal{CR}[\mathcal{L}] = \mathcal{F}[\text{Unanimity}]$.

De systemen van aggregatie-voorkeur en -oordeel zijn de belangrijkste systemen die zijn ontwikkeld voor het bestuderen van problemen gerelateerd aan de aggregatie van individuele uitdrukkingen. We verstrekken een inbedding van beide systemen in binaire aggregatie door het construeren van een geschikte integriteitsvoorwaarde en we bieden alternatieve bewijzen van een aantal van de klassieke resultaten in deze systemen door gebruikmaking van onze resultaten over collectieve rationaliteit. In het systeem van aggregatie-oordeel focussen we op het nieuw ontstane probleem van *veiligheid van de agenda*: hoe kunnen we met een verzameling van propositionele formules die het object van oordeel vormen , i.e., de agenda, garanderen dat het collectieve oordeel consistent zal zijn wanneer alle individuele oordelen dat zijn. Voor verscheidene verzamelingen van aggregatie-procedures gedefinieerd in axiomatische termen, verstrekken we noodzakelijke en toereikende condities voor een veilige agenda. En we laten zien dat het probleem van controleren van deze condities een zeer hardnekkig probleem is (Π_2^p-complete) voor alle overwogen verzamelingen.

Er zijn verscheidene voorbeelden van aggregatie-procedures die collectief rationeel zijn voor alle mogelijke integriteitsvoorwaarden. En we concluderen de dissertatie met het bestuderen van drie zulke procedures. We onderzoeken de computationele complexiteit van de twee klassieke problemen van de bepaling van de winnaar en de strategische manipulatie, i.e., we vergelijken de complexiteit van het berekenen van de winnaar van een verkiezing met het probleem van de bepaling of individuen stimulansen hebben om hun stem te veranderen ten gunste van hun eigen positie.

Deze dissertatie vormt een systematische studie naar het probleem van collectieve rationaliteit in binaire aggregatie. Dit vraagstuk staat centraal in de literatuur van de sociale keuzetheorie en heeft zich nuttig bewezen meer inzicht te verwerven in een variëteit aan toepassingen in de kunstmatige intelligentie.

Abstract

This dissertation presents the framework of *binary aggregation with integrity constraints*, positioning it with respect to existing frameworks for the study of collective decision making developed in Social Choice Theory and Artificial Intelligence, and exploring theoretical results which may pave the way for its implementation.

In our setting, a set of individuals need to make a choice over a set of binary issues. Each individual submits a yes/no choice for each of the issues, and these choices are then aggregated into a collective choice by means of an aggregation procedure. Individual choices are bound by a rationality assumption, specifying the range of answers that is considered to be rational. We represent rationality assumptions with formulas in a simple propositional language, calling them *integrity constraints*. This framework can be employed to model a variety of situations of collective decision making, such as multiple referenda, committee elections, as well as the problem of aggregating individual preferences or judgments.

A concept that is central to the whole dissertation is that of *collective rationality*: assuming that each individual satisfies a given integrity constraint, we are interested in finding out whether the output of a given aggregation procedure still satisfies the same integrity constraint.

We call a situation in which collective rationality is *not* satisfied a *paradox*, and we show that most of the classical paradoxes studied in the literature on Social Choice Theory can be seen as instances of our general definition. We focus our analysis on the Condorcet paradox, the discursive dilemma, the Ostrogorski paradox and the multiple election paradox, identifying a common syntactic property in the integrity constraints that define these paradoxical situations.

We classify integrity constraints into syntactically defined languages, and define $\mathcal{CR}[\mathcal{L}]$ as the class of collectively rational procedures with respect to all integrity constraints in a given language $\mathcal{L}$. For instance, we indicate with $\mathcal{CR}[cubes]$ the class of aggregation procedures that are collectively rational with respect to integrity constraints that can be expressed as conjunctions of literals (i.e., cubes). On the other hand, classes of aggregation procedures are usually defined in ax-

iomatic terms, and we indicate with $\mathcal{F}$[AX] the class of procedures satisfying a list of axiomatic properties AX. We investigate the relation between these two definitions for several natural fragments of the language of propositional logic, and for several axiomatic properties from the literature on Social Choice Theory. As an example, we prove that the class of collectively rational procedures with respect to cubes coincides with the class of unanimous procedures, i.e., those aggregators that accept (reject) a given issue if all individuals agree on accepting (rejecting) the issue: $\mathcal{CR}[cubes] = \mathcal{F}[\text{Unanimity}]$.

The frameworks of preference aggregation and judgment aggregation are the main settings developed in the literature to study problems related to the aggregation of individual expressions. We provide an embedding from each of the two frameworks into binary aggregation by devising a suitable integrity constraint, and we provide alternative proofs of some of the classical results in these settings by making use of our characterisation results of collective rationality. In the framework of judgment aggregation we focus on the novel problem of *safety of the agenda*: given a set of propositional formulas which constitute the objects of judgment, i.e., the agenda, how can we guarantee that the collective judgment will be consistent when all individual judgments are. For several classes of procedures defined in axiomatic terms, we provide necessary and sufficient conditions for an agenda to be safe, and we show that the problem of checking such conditions is a highly intractable problem (Π_2^p-complete) for all classes under consideration.

There are several examples of aggregation procedures that are collectively rational for every possible integrity constraint, and we conclude the dissertation by studying three such procedures. We investigate the computational complexity of the two classical problems of winner determination and strategic manipulation, i.e., we compare the complexity of computing the winner of an election with the problem of determining whether individuals have incentives to misrepresent their vote in order to favour their own position.

Overall, this dissertation constitutes a systematic study of the problem of collective rationality in binary aggregation, a problem that is central to the literature in Social Choice Theory and that proved useful to gain insight into a variety of applications in Artificial Intelligence.

Riassunto

Questa tesi introduce un modello formale per l'*aggregazione di questioni binarie con vincoli di integrità*, confrontandolo con i modelli di scelta collettiva sviluppati nella teoria della scelta sociale e in intelligenza artificiale, e ottenendo risultati teorici indirizzati ad una sua possibile implementazione.

La situazione di base nel nostro modello è composta da un insieme di individui che deve operare una scelta su una sequenza di *questioni* binarie. Ogni individuo esprime una scelta affermativa o negativa per ognuna delle questioni, e tali scelte sono in seguito aggregate in una scelta collettiva per mezzo di una procedura di aggregazione. Le scelte individuali sono vincolate da ipotesi di razionalità, atte a specificare l'insieme delle risposte considerate razionali. Nel nostro modello rappresentiamo le ipotesi di razionalità tramite formule espresse in un semplice linguaggio basato sulla logica proposizionale denomitate *vincoli di integrità*. Questo modello può essere utilizzato per rappresentare svariate situazioni di scelta collettiva, come referendum multipli, elezioni di comitati, nonché i classici problemi dell'aggregazione di preferenze e dell'aggregazione di giudizi.

Poniamo al centro dell'attenzione il concetto di *razionalità collettiva*: assumendo che ogni individuo soddisfi un dato vincolo di integrità, ci chiediamo se il risultato collettivo ottenuto tramite una data procedura di aggregazione continui a soddisfare lo stesso vincolo di integrità.

Definiamo *paradosso* ogni situazione in cui la razionalità collettiva *non* viene soddisfatta, e mostriamo come i classici paradossi studiati nell'ambito della teoria della scelta sociale possano per la maggior parte essere interpretati come casi particolari della nostra definizione di paradosso. Concentriamo la nostra analisi sul paradosso di Condorcet, sul dilemma discorsivo, sul paradosso di Ostrogorski e sul paradosso delle elezioni multiple, ed identifichiamo una proprietà sintattica comune a tutti i vincoli di integrità alla base di queste situazioni paradossali.

I vincoli di integrità possono essere suddivisi in linguaggi definiti in maniera sintattica, e definiamo $\mathcal{CR}[\mathcal{L}]$ la classe di procedure di aggregazione che sono collettivamente razionali rispetto ai vincoli di integrità nel linguaggio $\mathcal{L}$. Ad esempio,

$\mathcal{CR}[cubes]$ indica la classe di procedure collettivamente razionali rispetto ai quei vincoli di integrità che possono essere espressi attraverso congiunzioni di letterali, detti *cubes*. Le procedure di aggregazione sono solitamente classificate in termini assiomatici, e denotiamo con $\mathcal{F}[\text{AX}]$ la classe di procedure che soddisfano una data lista di assiomi AX. In questa tesi esaminiamo le relazioni che intercorrono tra queste due diverse definizioni di classi di procedure, variando da un lato il linguaggio in cui i vincoli di integrità vengono espressi, e dall'altro le possibili combinazioni di assiomi introdotte nella letteratura sulla teoria della scelta sociale. Per esempio, dimostriamo che la classe di procedure collettivamente razionali rispetto al linguaggio dei *cubes* coincide con la classe delle procedure unanimi, ossia quelle procedure che accettano (rifiutano) una questione quando tutti gli individui accettano (rifiutano) unanimemente tale questione: $\mathcal{CR}[cubes] = \mathcal{F}[\text{Unanimity}]$.

Nella letteratura sulla teoria della scelta sociale i principali modelli sviluppati per lo studio dell'aggregazione di espressioni individuali sono i modelli dell'aggregazione delle preferenze e dell'aggregazione dei giudizi. In questa tesi mostriamo come, sviluppando adeguati vincoli di integrità, entrambi questi modelli possono essere tradotti nel modello dell'aggregazione binaria e diamo dimostrazioni alternative di alcuni risultati classici facendo uso dei nostri risultati di caratterizzazione per procedure collettivamente razionali. Nel modello dell'aggregazione dei giudizi introduciamo il problema della *sicurezza dell'agenda*: dato un insieme di formule proposizionali che costituiscono l'oggetto del giudizio, i.e., l'agenda, ci chiediamo se sia possibile garantire che il giudizio collettivo sia coerente, sapendo che tutti i giudizi individuali lo sono. Per diverse classi di procedure di aggregazione definite in termini assiomatici troviamo condizioni necessarie e sufficienti per la sicurezza dell'agenda, e dimostriamo che il problema di controllare queste condizioni è altamente intrattabile (Π_2^p-completo) per tutte le classi di procedure prese in considerazione.

La tesi viene conclusa studiando tre procedure di aggregazione che sono collettivamente razionali rispetto ad ogni possibile vincolo di integrità. Studiamo la complessità computazionale di due problemi classici, la determinazione del vincitore e la manipolazione strategica, ossia compariamo quanto è complicato dal punto di vista computazionale determinare il vincitore di una data elezione rispetto al problema di determinare se sussistono incentivi a modificare il voto individuale per favorire la propria posizione a livello collettivo.

Nel complesso, questa tesi costituisce uno studio sistematico del problema della razionalità collettiva nell'aggregazione di questioni binarie, un problema che risulta essere centrale nella teoria della scelta sociale e che si è dimostrato di grande utilità per comprendere e sviluppare diverse applicazioni nel campo dell'intelligenza artificiale.

Titles in the ILLC Dissertation Series:

ILLC DS-2006-01: **Troy Lee**
Kolmogorov complexity and formula size lower bounds

ILLC DS-2006-02: **Nick Bezhanishvili**
Lattices of intermediate and cylindric modal logics

ILLC DS-2006-03: **Clemens Kupke**
Finitary coalgebraic logics

ILLC DS-2006-04: **Robert Špalek**
Quantum Algorithms, Lower Bounds, and Time-Space Tradeoffs

ILLC DS-2006-05: **Aline Honingh**
The Origin and Well-Formedness of Tonal Pitch Structures

ILLC DS-2006-06: **Merlijn Sevenster**
Branches of imperfect information: logic, games, and computation

ILLC DS-2006-07: **Marie Nilsenova**
Rises and Falls. Studies in the Semantics and Pragmatics of Intonation

ILLC DS-2006-08: **Darko Sarenac**
Products of Topological Modal Logics

ILLC DS-2007-01: **Rudi Cilibrasi**
Statistical Inference Through Data Compression

ILLC DS-2007-02: **Neta Spiro**
What contributes to the perception of musical phrases in western classical music?

ILLC DS-2007-03: **Darrin Hindsill**
It's a Process and an Event: Perspectives in Event Semantics

ILLC DS-2007-04: **Katrin Schulz**
Minimal Models in Semantics and Pragmatics: Free Choice, Exhaustivity, and Conditionals

ILLC DS-2007-05: **Yoav Seginer**
Learning Syntactic Structure

ILLC DS-2008-01: **Stephanie Wehner**
Cryptography in a Quantum World

ILLC DS-2008-02: **Fenrong Liu**
Changing for the Better: Preference Dynamics and Agent Diversity

ILLC DS-2008-03: **Olivier Roy**
Thinking before Acting: Intentions, Logic, Rational Choice

ILLC DS-2008-04: **Patrick Girard**
Modal Logic for Belief and Preference Change

ILLC DS-2008-05: **Erik Rietveld**
Unreflective Action: A Philosophical Contribution to Integrative Neuroscience

ILLC DS-2008-06: **Falk Unger**
Noise in Quantum and Classical Computation and Non-locality

ILLC DS-2008-07: **Steven de Rooij**
Minimum Description Length Model Selection: Problems and Extensions

ILLC DS-2008-08: **Fabrice Nauze**
Modality in Typological Perspective

ILLC DS-2008-09: **Floris Roelofsen**
Anaphora Resolved

ILLC DS-2008-10: **Marian Counihan**
Looking for logic in all the wrong places: an investigation of language, literacy and logic in reasoning

ILLC DS-2009-01: **Jakub Szymanik**
Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language

ILLC DS-2009-02: **Hartmut Fitz**
Neural Syntax

ILLC DS-2009-03: **Brian Thomas Semmes**
A Game for the Borel Functions

ILLC DS-2009-04: **Sara L. Uckelman**
Modalities in Medieval Logic

ILLC DS-2009-05: **Andreas Witzel**
Knowledge and Games: Theory and Implementation

ILLC DS-2009-06: **Chantal Bax**
Subjectivity after Wittgenstein. Wittgenstein's embodied and embedded subject and the debate about the death of man.

ILLC DS-2009-07: **Kata Balogh**
Theme with Variations. A Context-based Analysis of Focus

ILLC DS-2009-08: **Tomohiro Hoshi**
Epistemic Dynamics and Protocol Information

ILLC DS-2009-09: **Olivia Ladinig**
Temporal expectations and their violations

ILLC DS-2009-10: **Tikitu de Jager**
"Now that you mention it, I wonder...": Awareness, Attention, Assumption

ILLC DS-2009-11: **Michael Franke**
Signal to Act: Game Theory in Pragmatics

ILLC DS-2009-12: **Joel Uckelman**
More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains

ILLC DS-2009-13: **Stefan Bold**
Cardinals as Ultrapowers. A Canonical Measure Analysis under the Axiom of Determinacy.

ILLC DS-2010-01: **Reut Tsarfaty**
Relational-Realizational Parsing

ILLC DS-2010-02: **Jonathan Zvesper**
Playing with Information

ILLC DS-2010-03: **Cédric Dégremont**
The Temporal Mind. Observations on the logic of belief change in interactive systems

ILLC DS-2010-04: **Daisuke Ikegami**
Games in Set Theory and Logic

ILLC DS-2010-05: **Jarmo Kontinen**
Coherence and Complexity in Fragments of Dependence Logic

ILLC DS-2010-06: **Yanjing Wang**
Epistemic Modelling and Protocol Dynamics

ILLC DS-2010-07: **Marc Staudacher**
Use theories of meaning between conventions and social norms

ILLC DS-2010-08: **Amélie Gheerbrant**
Fixed-Point Logics on Trees

ILLC DS-2010-09: **Gaëlle Fontaine**
Modal Fixpoint Logic: Some Model Theoretic Questions

ILLC DS-2010-10: **Jacob Vosmaer**
Logic, Algebra and Topology. Investigations into canonical extensions, duality theory and point-free topology.

ILLC DS-2010-11: **Nina Gierasimczuk**
Knowing One's Limits. Logical Analysis of Inductive Inference

ILLC DS-2010-12: **Martin Mose Bentzen**
Stit, Iit, and Deontic Logic for Action Types

ILLC DS-2011-01: **Wouter M. Koolen**
Combining Strategies Efficiently: High-Quality Decisions from Conflicting Advice

ILLC DS-2011-02: **Fernando Raymundo Velazquez-Quesada**
Small steps in dynamics of information

ILLC DS-2011-03: **Marijn Koolen**
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

ILLC DS-2011-04: **Junte Zhang**
System Evaluation of Archival Description and Access

ILLC DS-2011-05: **Lauri Keskinen**
Characterizing All Models in Infinite Cardinalities

ILLC DS-2011-06: **Rianne Kaptein**
Effective Focused Retrieval by Exploiting Query Context and Document Structure

ILLC DS-2011-07: **Jop Briët**
Grothendieck Inequalities, Nonlocal Games and Optimization

ILLC DS-2011-08: **Stefan Minica**
Dynamic Logic of Questions

ILLC DS-2011-09: **Raul Andres Leal**
Modalities Through the Looking Glass: A study on coalgebraic modal logic and their applications

ILLC DS-2011-10: **Lena Kurzen**
Complexity in Interaction

ILLC DS-2011-11: **Gideon Borensztajn**
The neural basis of structure in language

ILLC DS-2012-01: **Federico Sangati**
Decomposing and Regenerating Syntactic Trees

ILLC DS-2012-02: **Markos Mylonakis**
Learning the Latent Structure of Translation

ILLC DS-2012-03: **Edgar José Andrade Lotero**
Models of Language: Towards a practice-based account of information in natural language

ILLC DS-2012-04: **Yurii Khomskii**
Regularity Properties and Definability in the Real Number Continuum: idealized forcing, polarized partitions, Hausdorff gaps and mad families in the projective hierarchy.

ILLC DS-2012-05: **David García Soriano**
Query-Efficient Computation in Property Testing and Learning Theory

ILLC DS-2012-06: **Dimitris Gakis**
Contextual Metaphilosophy - The Case of Wittgenstein

ILLC DS-2012-07: **Pietro Galliani**
The Dynamics of Imperfect Information